

Copyright  
by  
Carlos José Pérez Pardavila  
2025

The Dissertation Committee for Carlos José Pérez Pardavila  
certifies that this is the approved version of the following dissertation:

## **From Black Holes to Hitchin Systems: A Geometric Voyage**

### **Committee:**

Jacques Distler, Supervisor

Andreas Karch, Co-supervisor

Elena Caceres

Daniel Freed

Sonia Paban

# **From Black Holes to Hitchin Systems: A Geometric Voyage**

by

**Carlos José Pérez Pardavila**

## **Dissertation**

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

**Doctor of Philosophy**

**The University of Texas at Austin**

**August 2025**

# Acknowledgments

There are many people to whom I owe the deeply positive experience I've had over these years—individuals whose support, guidance, and presence made this academic journey possible. What follows is undoubtedly an incomplete list, but it is offered with immense gratitude.

During my time in Austin, I had the extraordinary privilege of learning physics and doing research alongside some of the leading physicists and mathematicians.

First and foremost, I would like to thank my advisor, Jacques Distler, whose dedication to science and the advancement of knowledge inspired me from the very beginning. His passion and rigorous approach were invaluable lessons in what it means to do research at the highest standard.

I am also deeply grateful to my coadvisor, Andreas Karch, for being an exceptional guide and collaborator. His work ethic and remarkable patience in answering my questions were instrumental—not only in the development of the papers that form this dissertation, but in my growth as a researcher.

I would also like to thank Elena Cáceres, not only for serving on my dissertation committee, but for being a supportive and loving friend. Her caring presence helped me—and many of my colleagues—manage the stress and uncertainty that so often accompany the PhD journey.

My thanks also go to my other committee members, Sonia Paban and Dan Freed. Without Sonia's steady support, I would have struggled to navigate the academic complexities of the PhD process. From Dan, I had the privilege of learning about geometry and mathematical physics in a way that only a truly gifted mathematician can offer—insightful, elegant, and profound.

Finally, I want to express my deep appreciation to Vadim Kaplunovsky, from whom I learned quantum field theory. His deep understanding and generous teach-



ing—offering insights you won’t find in any textbook—left a lasting impression on me.

In addition to the invaluable guidance of my professors, I greatly benefited from the many interactions with students, postdocs, and faculty—both on the 9th floor and throughout other departments. These conversations, collaborations, and shared moments of discovery enriched my understanding and made daily life in academia more rewarding. I would like to thank Behzat Ergun, Ali Shehper, Vikas Aragam, Niraj Desai, Hao Geng, Aswin Balasubramanian, Mario Martone, Lisa Randall, and Suvrat Raju for their insights and generosity.

I am also deeply grateful to my friends within the Weinberg Theory Group and beyond: Rafael Carrasco, Merna Youssef, Brian Kent, Sanjit Shashi, Alejandro Vilar López, Alejandro Franco, David Luján, Sergi Navarro, Marcos Riojas, Piero Visconte, Tyler Larginho, Mariana Escalona, Saba Baig, Ricky Ardila, and Yoel Villahermosa. Your friendship provided the emotional grounding and laughter that kept me going through the many ups and downs of graduate school.

A very special acknowledgment goes to Rodrigo Castillo and Nicolás Morales Durán, my brothers in arms throughout this journey. More than colleagues, you were constant companions in both the academic and personal sense—walking this path with me every step of the way.

Las siguientes palabras vienen en forma de dedicatoria, y por eso me tomo la libertad de escribirlas en español.

Quiero dedicar esta tesis, además de agradecer, a las personas sin las cuales este esfuerzo de casi siete años no hubiese sido posible. En primer lugar, a mi compañera de aventuras, la que me aguanta la locura: mi pareja, Irene Arias Quintero.

También quiero agradecer a mi familia, para quienes no siempre ha sido fácil tenerme tan lejos, pero que, aun así, me han dado la fuerza necesaria para enfrentar todos los desafíos y, a su manera, me han hecho sentir siempre cerca. A mis padres,

Carlos y Beatriz, por su amor y apoyo incondicional, por ser ejemplos de esfuerzo incansable, generosidad y entrega. Siempre han estado ahí, poniéndome por delante de todo. A mis hermanas, Valentina y Victoria, y a mis cuñados, Fernando y Fernando. Agradezco también a mis primos: Cecilia, Carolina, Ana Cristina, Javier, Rocío y Mateo, y a mis tías, Marité y Alba, por su cariño constante y por estar siempre presentes, incluso en la distancia.

Un recuerdo muy especial para mi abuela Argentina y mi tío Jorge, quienes me vieron comenzar esta etapa y, seguramente, me están viendo terminarla desde otro lugar. A mi abuela le debo mucho más de lo que estas líneas pueden expresar. Su valentía, fuerza y determinación son parte de la razón por la que hoy puedo celebrar este logro. A mi tío Jorge, por su cariño y su apoyo constante, también le guardo un lugar muy especial en mi corazón. Ambos siguen presentes en este logro de una forma profunda e inolvidable.

A mi familia española: María José Quintero y Santiago Arias, gracias por su cariño y apoyo.

A mis amigos en Uruguay, los Crispy, y a mi hermano de la vida, Agustín Scasso: gracias por hacerme sentir acompañado aunque la distancia sea grande. Y especialmente, a mi querida ahijada, María del Pilar, que aunque no me vio empezar este camino, sí podrá verme terminarlo.

A todos ustedes, les estoy eternamente agradecido por haber sido mi sostén en estos años.

# Abstract

## From Black Holes to Hitchin Systems: A Geometric Voyage

Carlos José Pérez Pardavila, PhD  
The University of Texas at Austin, 2025

SUPERVISORS: Jacques Distler, Andreas Karch

Geometry has played a pivotal role in deepening our understanding of physical phenomena. At the pinnacle of this interplay stands string theory. In this dissertation, we explore two distinct applications of string-theoretic ideas to fundamental problems in high-energy physics.

First, we examine recent progress on the black hole information paradox—a long-standing challenge in theoretical physics. Leveraging developments in the AdS/CFT correspondence and the concept of entanglement islands, we investigate how geometric surfaces in higher-dimensional spacetimes encode quantum information in lower-dimensional conformal field theories. We further analyze the limitations of the island prescription in gravitational theories with long-range effects, such as massive gravity, and propose consistency conditions for entanglement wedges in nonlocal contexts.

In the second part of the dissertation, we shift focus to the low-energy dynamics of superconformal field theories of class S. Exploiting the striking correspondence between Hitchin systems and Seiberg-Witten geometries, we construct these integrable systems locally and lay the groundwork for a family-theoretic construction of type-D class S theories.

# Table of Contents

|                                                                                                        |     |
|--------------------------------------------------------------------------------------------------------|-----|
| List of Tables . . . . .                                                                               | 10  |
| List of Figures . . . . .                                                                              | 11  |
| Chapter 1: Introduction . . . . .                                                                      | 23  |
| 1.1 Geometry in holography . . . . .                                                                   | 24  |
| 1.2 Geometry in quantum field theory . . . . .                                                         | 64  |
| Chapter 2: Information Transfer with a Gravitating Bath . . . . .                                      | 94  |
| 2.1 Introduction . . . . .                                                                             | 95  |
| 2.2 Entanglement Entropy of a Dynamical $\mathcal{R}$ . . . . .                                        | 102 |
| 2.3 Left/Right Entanglement at Finite Temperature . . . . .                                            | 121 |
| 2.4 The Critical Angle . . . . .                                                                       | 137 |
| 2.5 Conclusion . . . . .                                                                               | 145 |
| Appendix 2.A: Hartman-Maldacena Area in the Black String . . . . .                                     | 148 |
| Chapter 3: Inconsistency of Islands in Theories with Long-Range Gravity . .                            | 151 |
| 3.1 Introduction . . . . .                                                                             | 152 |
| 3.2 Asymptotic regions in entanglement wedges . . . . .                                                | 160 |
| 3.3 A puzzle with islands . . . . .                                                                    | 166 |
| 3.4 A resolution using massive gravity . . . . .                                                       | 173 |
| 3.5 Islands in decoupled systems . . . . .                                                             | 190 |
| 3.6 Discussion . . . . .                                                                               | 197 |
| Appendix 3.A: Loopholes in Possible Counterarguments . . . . .                                         | 200 |
| Chapter 4: Entanglement Phase Structure of a Holographic BCFT in a Black-<br>Hole Background . . . . . | 210 |
| 4.1 Introduction . . . . .                                                                             | 211 |
| 4.2 The Doubly Holographic Black-String Model . . . . .                                                | 216 |
| 4.3 Numerical Results for the Area Difference . . . . .                                                | 228 |
| 4.4 Discussion of Results . . . . .                                                                    | 245 |
| Appendix 4.A: Analytic Results for the Area Difference in the $\Gamma \rightarrow 1$ Limit             | 246 |
| Appendix 4.B: Validation of Numerics . . . . .                                                         | 248 |
| Chapter 5: Jackiw-Teitelboim Gravity from the Karch-Randall Braneworld .                               | 250 |
| 5.1 Introduction . . . . .                                                                             | 250 |
| 5.2 Karch-Randall Braneworld and Wedge Holography . . . . .                                            | 253 |
| 5.3 Realizing JT Gravity . . . . .                                                                     | 254 |
| 5.4 A Puzzle and Its Resolution . . . . .                                                              | 258 |
| 5.5 Conclusions . . . . .                                                                              | 260 |

|                                                                             |     |
|-----------------------------------------------------------------------------|-----|
| Chapter 6: Subregion Entropy for the Doubly-Holographic Global Black String | 261 |
| 6.1 Introduction                                                            | 261 |
| 6.2 Setup and Equations of Motion                                           | 267 |
| 6.3 Numerical Results                                                       | 279 |
| 6.4 Conclusion                                                              | 291 |
| Appendix 6.A: Useful Results about Islands and the Photon Sphere            | 295 |
| Appendix 6.B: Tiny Island Surfaces and Flat Page Curves                     | 295 |
| Chapter 7: Entropy of Radiation with Dynamical Gravity                      | 298 |
| 7.1 Introduction                                                            | 298 |
| 7.2 Double Holography and Dynamical Gravity                                 | 302 |
| 7.3 RT Structure                                                            | 312 |
| 7.4 Numerical Results                                                       | 317 |
| 7.5 Conclusions                                                             | 330 |
| Appendix 7.A: Supporting Plots                                              | 332 |
| Chapter 8: Constraining braneworlds with entanglement entropy               | 334 |
| 8.1 Introduction                                                            | 334 |
| 8.2 The holography of two-brane models                                      | 342 |
| 8.3 Constraining the DGP term                                               | 355 |
| 8.4 Discussion                                                              | 377 |
| Appendix 8.A: Effective Action with RS + DGP Gravity                        | 381 |
| Appendix 8.B: Dimension Dependence of the Positive-Entropy Constraint       | 386 |
| Chapter 9: The Hitchin Image in Type-D                                      | 392 |
| 9.1 Introduction                                                            | 392 |
| 9.2 The Local Hitchin Base in Type-D                                        | 396 |
| 9.3 A Conjecture                                                            | 407 |
| 9.4 Even and Very Even Type Constraints                                     | 408 |
| 9.5 Odd Type Constraints                                                    | 415 |
| 9.6 Global Considerations                                                   | 430 |
| Appendix 9.A: Illustrative Examples for Theorem 2                           | 436 |
| Appendix 9.B: The Equivariant Fundamental Group and Lusztig's Quotient      | 442 |
| Appendix 9.C: Hitchin Systems for 3-Punctured Spheres                       | 443 |
| Appendix 9.D: The Work of Kraft-Procesi on Special Pieces                   | 445 |
| Works Cited                                                                 | 450 |
| Vita                                                                        | 498 |

# List of Tables

|     |                                                                                                                                                         |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 9.1 | Examples illustrating the inclusions $\bar{A}_b \subseteq \bar{A}(\mathcal{O}) \subseteq A(\mathcal{O})$ for various partitions $\mathcal{O}$ . . . . . | 443 |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------|-----|

# List of Figures

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            |     |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 1.1 | A diagram illustrating the connections between different theories, tracing their descent from the six-dimensional $(2,0)$ theory down to zero-dimensional invariants. The elements within each row share the same spacetime dimension. Arrows indicate the compactifications and corresponding low-energy limits that relate these theories. The colour coding distinguishes different compactification manifolds: red for $S^1$ , green for a two-dimensional surface $C$ , blue for a three-dimensional manifold, and purple for a four-dimensional manifold. Inspired by a similar figure in Moore (2012). . . . .                                                                                                                                                                      | 69  |
| 2.1 | (a) Embedding of a KR braneworld with two subcritical branes in anti-de Sitter space. We refer to the left brane as the “physical” brane, and the right brane as the “bath” brane. $\mathcal{R}$ denotes the radiation region whose entanglement entropy we wish to calculate, and the dashed green line connected to the boundary of $\mathcal{R}$ is the candidate RT surface. $\mathcal{I}$ is the entanglement island on the brane corresponding to the RT surface in green. (b) Embedding of a KR braneworld with the same tensions in the black string geometry. The dashed black line is the black string horizon separating the exterior and interior regions. . . . .                                                                                                             | 98  |
| 2.2 | A constant- $t$ slice of global $\text{AdS}_{d+1}$ with two branes present. The defect is shown in red, and various candidate extremal surfaces (in order of decreasing area from left to right) are in green. For each surface, $\mathcal{R}$ and $\mathcal{I}$ are respectively the radiation region and island, which end orthogonally on both branes. The minimal, zero-area entanglement surface $r(\mu) = 0$ is shown on the right as a limit, cutting through the middle of the space. . . . .                                                                                                                                                                                                                                                                                      | 116 |
| 2.3 | The critical anchor on the brane, given here in four dimensions, is a monotonically decreasing function of the brane angle which vanishes at the critical angle, which in this case is approximately $\theta_c \approx .98687$ . The critical anchor tends toward the horizon distance ( $u_h = 1$ ) as $\theta \rightarrow 0$ and tends toward zero as $\theta \rightarrow \theta_c$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                            | 128 |
| 2.4 | Overview for potential island surfaces within the black string geometry, each of which satisfies the boundary condition for some corresponding brane placed below the critical angle. Each surface ends at its critical anchor, which is where the island begins. One loses the island surfaces as the angle decreases, and the island fills the brane at the critical angle. The area difference between the Hartman-Maldacena surface and the island surface vanishes slightly below the critical angle. While one would need many digits of precision to show this, the island surfaces appear to fill the <i>island commonwealth</i> , labelled as region I, without crossing one another. They never make it into the region II, which we call <i>independent territory</i> . . . . . | 130 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |     |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 2.5  | The $t = 0$ area difference (4.11) between the Hartman-Maldacena surface and the island surface, given as a function of brane angle is in orange, given for the black string at $d = 4$ . This should be compared to the behavior of the brane's induced Newton's constant $G_N^{-1} \sim 1/\theta^2$ in blue. As we approach the critical angle the area of the island surface approaches the area of Hartman-Maldacena in empty AdS. This means the difference becomes negative in the region near the critical angle, which for $d = 4$ is approximately $\theta_c \approx 0.98687$ . . . . .                                                                                                                                                                                                                                                                         | 133 |
| 2.6  | Page curves obtained for sample angles $\theta = \frac{\pi}{10}, \frac{2\pi}{10}, \frac{3\pi}{10}$ starting just below the critical angle $\theta_c \approx 0.98687$ . As the brane angle decreases, the Page time goes to infinity. As the brane approaches the critical angle $\theta_c$ , the Page time decreases. To obtain the area difference, we integrate numerically up to some small region at which we model the area of the RT surface with a series. . . . .                                                                                                                                                                                                                                                                                                                                                                                                | 134 |
| 2.7  | A phase diagram for $d=4$ showing the dominant minimal surface as a function of time $t$ and the larger brane-angle $\theta$ . The transition between the island surface and the tiny island surface always takes place across the critical angle. The transition between the Hartman-Maldacena surface and the island surface is across the blue curve, which intersects $t = 0$ at the Page angle. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                             | 137 |
| 2.8  | The critical angles in $\text{AdS}_{d+1}$ for various dimensions $d > 2$ , analytically continued to real values. The critical angle vanishes at precisely $d = 2$ . Both empty AdS and the black string feature the same critical angles; the orange curve was obtained using (2.59), and the black points were obtained numerically by solving the shooting problem for the black string. The reason for the equivalence is that black string island surfaces near the critical angle are located near the defect, where the metric is well approximated by empty AdS. . . . .                                                                                                                                                                                                                                                                                         | 140 |
| 2.9  | Possible island surfaces in empty $\text{AdS}_{d+1}$ , which exist only for a brane placed at the critical angle. To summarize, we have two distinct cases. If $\theta_1 < \theta_P$ and $\theta_2 < \theta_P$ , the left/right entanglement entropy for the AdS black string follows a Page curve, with the Page time determined by the stronger brane. Otherwise, we have no Page curve, and the left/right entanglement entropy is simply constant. These are relevant for the black string geometry, since island surfaces near the boundary are increasingly well approximated by empty AdS. The island surfaces completely fill the island commonwealth ( $\theta_c < \mu < \pi - \theta_c$ ), region I, without crossing one another. They also fail to make it into region II, the independent territory ( $\mu < \theta_c$ or $\mu > \pi - \theta_c$ ). . . . . | 141 |
| 2.10 | The minimum area versus time for the black string geometry with $u_h = 1$ and $d = 4$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 149 |



|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |     |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.1 | <i>A cartoon of a constant-time slice of a black hole with a brane embedded. <math>R</math> is the union of regions on two asymptotic boundaries and <math>\bar{R}</math> is its complementary region. The horizons in the bulk are marked by <math>H</math>. The separation between horizons is meant to convey that the Cauchy slice under examination is a late-time slice on which the wormhole is of a finite length. The dominant RT surface for the region <math>R</math> is shown in purple. The region on the brane marked <math>I</math> becomes the “island” in the lower-dimensional picture of Figure 3.2. In this figure, both the horizontal and the vertical directions are spatial.</i> | 156 |
| 3.2 | <i>A spacetime diagram of the same system of branes in a black hole in the <math>d</math> dimensional description. The entanglement wedge for the region <math>R</math> is now an “island”. In this figure, the horizontal direction is spatial and time runs along the vertical direction.</i>                                                                                                                                                                                                                                                                                                                                                                                                          | 157 |
| 3.3 | <i>A point <math>P</math> that is part of multiple entanglement wedges. The quasilocal bulk operator must be dressed to <math>R_1</math> on the left and to <math>R_2</math> on the right. The figure shows a time slice of global AdS.</i>                                                                                                                                                                                                                                                                                                                                                                                                                                                              | 167 |
| 3.4 | <i>An excitation at the point <math>P</math> inside the island (pink shaded disk) can be detected using a two-point correlator outside the island. The two-point correlator involves an integral of the asymptotic metric (indicated by the dashed line) and another operator obtained by taking the limit of point <math>P'</math> to <math>P_B</math> on the boundary of the gravitational region. The bath is not shown in this figure, which shows a time slice of AdS.</i>                                                                                                                                                                                                                          | 170 |
| 3.5 | <i>In the higher-dimensional setup, an operator at point <math>P</math> can be dressed to the boundary through the higher-dimension. The setup is the same as that of Figure 3.1 and this Figure illustrates how an operator at <math>P</math> can be dressed to the boundary while bypassing the entanglement wedge of <math>\bar{R}</math>. In this figure both the horizontal and vertical directions are spatial.</i>                                                                                                                                                                                                                                                                                | 175 |
| 3.6 | <i>A lower-dimensional picture of the setup of Figure 3.5. This is a space-time diagram like Figure 3.2. In the lower-dimensional description, an operator at point <math>P</math> in the island cannot be dressed to the boundary without affecting the region outside the island. So the island cannot constitute a consistent entanglement wedge in a theory where the Gauss law applies. The lower-dimensional description of Figure 3.5 involves massive gravity where the Gauss law does not hold. In this figure, the horizontal direction is spatial and time runs along the vertical direction.</i>                                                                                             | 176 |
| 3.7 | <i>An <math>\text{AdS}_d</math> brane (thick red line) in <math>\text{AdS}_{d+1}</math>. The conformal boundary is the thick black line. Lines of constant <math>\varrho</math> are red dashed lines and coordinates transverse to <math>\varrho</math> are constant along the blue dashed circles.</i>                                                                                                                                                                                                                                                                                                                                                                                                  | 186 |
| 3.8 | <i>One possibility for an island in a decoupled pair of systems. The island cannot affect the degrees of freedom beyond either the left horizon or the right horizon. The solid black lines are the left and right asymptotic AdS boundaries. The broken blue lines are singularities and the dashed lines are horizons. The dashed lines also demarcate the boundaries between regions marked, <math>L</math>, <math>R</math>, <math>I</math>, <math>J</math>. In this figure, the horizontal direction is spatial and time runs along the vertical direction.</i>                                                                                                                                      | 194 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 3.9  | <i>A single sided black hole formed from the collapse of some matter (shown in brown). The dashed line on the left is the origin of polar coordinates and the solid black line on the right is the asymptotic AdS boundary. The Figure shows a signal <math>S_1</math> that originates on the boundary and reaches deep inside the interior. In this figure, the horizontal direction is spatial and time runs along the vertical direction.</i>                                                                                                                                                                                                                                                                                                                                                                                                                                             | 195 |
| 3.10 | <i>In the eternal black hole, it is always possible to send signals to any point in the interior, even classically, provided these signals originate at an early enough time. The Figure shows a signal <math>S_1</math> that originates on the left boundary and another signal <math>S_2</math> that originates on the right boundary. Therefore, no part of the interior of a standard eternal black hole can constitute an island that is redundant with degrees of freedom from a decoupled system. In this figure, the horizontal direction is spatial and time runs along the vertical direction.</i>                                                                                                                                                                                                                                                                                 | 196 |
| 3.11 | <i>A possible attempt to produce islands in standard gravity. An island is first formed in a theory of massive gravity that involves a nongravitational bath. At time <math>t_D</math> the coupling to the bath is switched off. The change in boundary conditions is expected to make the graviton massless at all points in the causal future of the decoupling event (shown in red). But the island (shown in green) is unaffected by this event since every point in the island is spacelike to the decoupling event. So the graviton remains massive in the island.</i>                                                                                                                                                                                                                                                                                                                 | 208 |
| 4.1  | <i>This cartoon illustrates the doubly holographic black-string model, where the region behind the KR brane has been excised. There are three perspectives, with the <math>\text{AdS}_{d+1}</math> gravitating bulk illustrated at the top of the diagram. The <math>d</math>-dimensional BCFT and <math>d</math>-dimensional <math>\text{AdS}_d</math> brane are depicted in colors which correspond with the intermediate picture indicated at the bottom of the diagram. Notice the island region envelops the black-hole horizon on the brane, and the defect cannot be reached from the island in the intermediate picture without crossing through its complement. The two parameters in the theory are the anchor point <math>\Gamma</math> and brane angle <math>\theta_b</math>. Note that “left/right” has a different meaning than in our previous paper Geng et al. (2021b).</i> | 218 |
| 4.2  | <i>The bulk gravitational system we study in this paper is dual to a BCFT thermal state on a nongravitating black-hole background. We compute the entropy of the subregion <math>\mathcal{R}</math>, a connected interval that runs through the <i>bifurcation surface</i> (the horizon at <math>t = 0</math> separating the past and future horizons, represented by the point in the middle of the Penrose diagram above), using the holographic RT prescription. This bipartition is similar to that in a recent work Geng et al. (2021f) studying entanglement islands in de Sitter space.</i>                                                                                                                                                                                                                                                                                           | 219 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     |     |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.3 | This cartoon illustrates the relevant RT surfaces for different brane angles $\theta_{1,2}$ below the critical angle. Island surfaces end on the brane at right angles, and HM surfaces cross the black-string horizon and traverse the Einstein-Rosen bridge. Islands always attach to the brane inside the atoll, which begins at a point called the <i>critical anchor</i> and extends to the black-hole horizon. Decreasing the brane angle reduces the size of the atoll, which shrinks to the horizon as $\theta \rightarrow 0$ . Increasing the size of $\mathcal{R}$ , bounded by $\Gamma$ , pushes the anchors (in purple) toward the black-hole horizon; increasing the brane angle pushes the anchors toward the defect. The critical anchor, which is the point on the brane that determines the size of the largest island on the brane (which always corresponds to $\Gamma = 0$ ), coincides with the defect above the critical angle. See Figure 4.5 for numerical results. . . . . | 223 |
| 4.4 | These figures were included in our previous work Geng et al. (2021b), but we repeat them here for emphasis. The critical angle is a monotonically increasing function of the number of spatial dimensions $d$ and takes the same value for any geometry which is asymptotically $\text{AdS}_{d+1}$ . This is distinct from the behavior of the critical anchor defining the beginning of the atoll for the black string. The critical anchor instead <i>decreases</i> monotonically with brane angle and coincides with the defect at the critical angle (about 0.98687 for $d = 4$ ). It can be shown that the critical anchor also increases monotonically with the number of dimensions for the space. . . . .                                                                                                                                                                                                                                                                                   | 227 |
| 4.5 | Here we see the setup together with numerical data in the planar black string. The KR brane is green, the HM surfaces are red, the island surfaces are blue, and the black-string horizon is the semicircle in black. The area difference between the island surface and the HM surface for various anchor points $\Gamma$ is displayed in the set of second plots below. Sending $\Gamma \rightarrow 0$ recovers the finite area differences from our previous paper Geng et al. (2021b) for angles below the critical angle. When the brane lies above the critical angle, the area difference diverges to negative infinity. These divergences are explained physically in Section 4.3.3. . . . .                                                                                                                                                                                                                                                                                                | 229 |
| 4.6 | The area difference is displayed as a function of brane angle $\theta_b$ for many possible CFT anchor points $\Gamma$ . The area difference is clearly monotonically increasing in $\Gamma$ and monotonically decreasing in $\theta_b$ . The anchor points were determined numerically for many values of $\theta_{HM}$ using the shooting method with the boundary conditions in (7.12). The area difference diverges to negative infinity as $\Gamma \rightarrow 0$ for $\theta_b > \theta_c$ , but when $\Gamma \rightarrow 0$ for $\theta_b < \theta_c$ it converges to a finite value. . . . .                                                                                                                                                                                                                                                                                                                                                                                                 | 231 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |     |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.7  | The area difference is displayed as a function of CFT anchor point $\Gamma$ for many possible brane angles $\theta_b$ . The area difference is clearly monotonically increasing in $\Gamma$ and monotonically decreasing in $\theta_b$ . The islands become dominant when the area difference is negative, which is possible only above the Page angle. The Page angle Geng et al. (2021b) is very slightly less than the critical angle ( $\theta_c \approx 0.98687$ in $d = 4$ ). The area difference diverges to negative infinity as $\Gamma \rightarrow 0$ for $\theta_b > \theta_c$ , but when $\Gamma \rightarrow 0$ for $\theta_b < \theta_c$ it converges to a finite value. We explain in Section 4.3.3 that the area difference diverges to positive infinity as $\theta_b \rightarrow 0$ . . . . . | 232 |
| 4.8  | This cartoon illustrates how the area difference is affected by each limiting case for the boundary anchor point $\Gamma$ , with angles below the critical angle on the right, and angles above the critical angle on the left. Taking the difference between the surfaces essentially acts as a renormalization scheme. The area difference between the island surface and HM surface diverges when one of them shrinks to a point because the difference between the surfaces reaches farther into the boundary region. When the brane lies below the critical angle, the island surface cannot shrink to a point when $\Gamma \rightarrow 0$ because it anchors to the brane inside the atoll. . . . .                                                                                                      | 233 |
| 4.9  | Here we present the positive area difference as a function of $\theta_b$ and $\theta_{HM}$ . The parameter subspace for which the area difference vanishes is an approximately straight line. Negative brane tensions are most likely not physical but they have been included without much trouble. This figure is more complete than Figures 4.6 and 4.7, which simply correspond to slices along the $x$ and $y$ axes, with the main difference being that $\Gamma$ has been swapped out for $\theta_{HM}$ . Note the area difference increases monotonically with $\theta_{HM}$ and decreases monotonically with $\theta_b$ , and that the area difference vanishes along an approximately straight line. . . . .                                                                                          | 236 |
| 4.10 | The phase diagram depends on the anchor point $\Gamma$ and the brane angle $\theta_b$ . When the brane lies above the Page angle, the island surfaces are dominant when $\Gamma$ is sufficiently small; as the brane angle increases the island surfaces are dominant for increasingly large values of $\Gamma$ . For anchor points outside the boundary black hole's horizon ( $\Gamma < 1$ ), the island surfaces become dominant as we take $\theta_b \rightarrow \pi$ . . . . .                                                                                                                                                                                                                                                                                                                            | 239 |
| 4.11 | The phase diagram is approximately linear when the anchor point $\Gamma$ is exchanged for the HM angle $\theta_{HM}$ ; this can also be seen in Figure 4.9. The phase diagram is otherwise the same as in Figure 4.10, with $\theta_{HM} = \pi$ corresponding to $\Gamma = 1$ and $\theta_{HM} = \pi/2$ corresponding to $\Gamma = 0$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 240 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  |     |
|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 4.12 | When an RT surface anchors within the constant entropy belt on the bath, the island surfaces are dominant on the initial time slice and the entanglement entropy is constant in time. Decreasing the brane tension $(d-1)\cos\theta_b$ increases the width of the belt, indicated by the color-coded arrows on the boundary which correspond to the vertical axis in Figure 4.10. The belt first appears at the Page angle, slightly below the critical angle, and covers the entire region outside the boundary black hole horizon (at $\Gamma=1$ ) in the limit where $\theta_b\rightarrow 1$ . All angles are measured from the boundary, and the tensionless limit is achieved at $\theta_b=\pi/2$ . . . . . | 241 |
| 4.13 | Here we illustrate the <i>Russian doll rule</i> , which states that nested surfaces can be constructed by either decreasing the brane angle $\theta_b$ or increasing the anchor point $\Gamma$ . This always pushes the surfaces closer to the event horizon into the bulk, decreases the size of the island on the brane, and increases the area difference between the island surface and the HM surface on the initial time slice. The area difference increases without bound as the event horizon is approached by the island surface. . . . .                                                                                                                                                              | 242 |
| 4.14 | The island region can be reconstructed using boundary data in $\bar{\mathcal{R}}$ only if both sides are accessible and only until the Page time. According to the Russian doll rule, the size of the region which is lost at the Page time increases monotonically with the brane angle $\theta_b$ and decreases monotonically with $\Gamma$ . . . . .                                                                                                                                                                                                                                                                                                                                                          | 243 |
| 4.15 | Here we illustrate the doubly holographic black string setup for a small radiation region of width $s$ . This corresponds to the limit $\Gamma\rightarrow 1$ . The island region becomes infinitesimally small as well; this is consistent with the Russian doll rule. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                   | 247 |
| 4.16 | Area of the HM surface as a function of $\zeta_*$ with $u_h=1$ . The red curve is what we expect for the near-horizon region (small $\zeta_*$ ) from theory, namely $-0.030/\zeta_*$ . Close to the horizon, the data roughly matches the theory, but for some large enough $\zeta_*$ the finite temperature kicks in and we should not expect the data to match. It is not visible, but there is a small error in the “large” $\zeta_*$ region. . . . .                                                                                                                                                                                                                                                         | 249 |
| 4.17 | Area difference as a function of brane angle, for $\Gamma=0.32$ for angular cutoffs ranging from 0.0005 to 0.2 in steps of 0.0005 (with $u_h=1$ ). Each curve is shown in a different colour, but since they are very close it is hard to distinguish one from the other. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                | 249 |
| 5.1  | A schematic representation of the doubly holographic nature of the wedge. (a) is the bulk description, which is just gravity on the $(d+1)$ -dimensional wedge with a codimension-two corner in red. (b) is the intermediate description in which the two $d$ -dimensional theories share a transparent interface (red). (c) is the boundary description on the interface. . . . .                                                                                                                                                                                                                                                                                                                               | 252 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |     |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 5.2 | (a) A constant time slice of the empty $\text{AdS}_3$ bulk with one Karch-Randall brane (in black). The brane intersects the asymptotic boundary (in blue) at the two defects (red dots). The holographic dual is a $\text{BCFT}_2$ living on the blue interval. (b) The empty $\text{AdS}_3$ wedge consisting of two KR branes. The two branes intersect at the two defects (red and blue dots). The dual boundary description is hence a conformal quantum mechanics living on the two defects. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 255 |
| 5.3 | The shapes of the two branes are fluctuating (though more rigorously we have also to impose the orbifold projection). The $\text{AdS}_2$ cutoff is induced by the $\text{AdS}_3$ cutoff (the red curve). The shaded regions are excised. Now the regulated asymptotic boundary of the wedge are two intervals. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 258 |
| 5.4 | The empty $\text{AdS}_3$ wedge consisting of two KR branes. The bulk geometry is the $\text{AdS}_3$ global patch. The green curves are representatives of the degenerate RT surfaces calculating the entanglement entropy between the two defects (the blue dot and the red dot). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 259 |
| 6.1 | The diameter of the black string is determined by the order parameter in the theory, $u_h$ . The bulk theory is dual to a thermal $\text{BCFT}$ state on the surface of the sphere. When $u_h$ becomes larger than $u_{\text{HP}} = \frac{1}{2^{1/3}}$ we encounter the Hawking-Page transition and the string is no longer the global minimum of the free energy. The black string reaches minimum temperature at the spinodal point $u_s = \frac{\sqrt{3}}{2^{2/3}}$ , where the specific heat changes sign from positive to negative. Also see Figure 6.2. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                        | 265 |
| 6.2 | The temperature of the black hole scales non-monotonically with $u_h$ , but for large black holes, the relationship is approximately linear. Unlike in asymptotically flat space, where black holes always have negative specific heat, $\text{AdS}$ black holes have positive specific heat when they are large enough. The canonical ensemble in $\text{AdS}$ is sensible because a temperature can be defined in the dual theory. For $\beta$ above the Hawking-Page phase transition, where the radius of the black hole is the $\text{AdS}$ radius, the system will transition to a gas of thermal radiation. Such black holes cannot be in thermal equilibrium in the canonical ensemble. For $u_h$ past the spinodal point, black holes become hotter as they evaporate, so their specific heat is negative. Since $\text{AdS}$ black holes do not exist above the value of $\beta$ where the spinodal point occurs, they do not contribute at all to the canonical ensemble. . . . . | 273 |
| 6.3 | This figure shows the critical anchor on the brane, as a function of brane angle, as a percentage of the horizon radius, for various choices of order parameter $u_h$ . The black string's diameter, and correspondingly, the black hole's size on the KR brane, decreases (increases) as $u_h$ is increased (decreased). The Hawking-Page transition lies at $u_h = \frac{1}{2^{1/3}} \approx .8$ . We recover our result for a large black string, discussed in Geng et al. (2021b), by taking the limit $u_h \rightarrow 0$ . There is a "hole" on the atoll when the critical anchor plot is not 1-to-1, which occurs for any non-zero value of $u_h$ . As $u_h \rightarrow \infty$ , the black string's diameter shrinks to a point, and the plot approaches a square shape. There is only one critical anchor below the critical angle ( $\theta_c \approx .98687$ ), but there are two critical anchors between the critical angle and the saturation angle $\theta_s$ . . . . .      | 282 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |     |
|-----|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.4 | This cartoon illustrates typical RT surfaces, above the critical angle $\theta_c$ , for an eternal black string dual to a BCFT on the surface of a sphere. When the KR brane lies above the critical angle, RT surfaces can travel from the KR brane to the defect from either the upper or lower critical anchor. RT surfaces can reach the boundary only from the atoll, and the region between the critical anchors is always a hole in the atoll. Increasing the brane angle decreases the size of the hole; at the saturation angle, the upper and lower critical anchors annihilate at the saturation anchor. Note that similar observations apply to anchors within a finite distance of the defect on the non-gravitating bath. . .                                                                                                                                                                                                                                                                                                                                                                     | 283 |
| 6.5 | This figure shows the saturation angle for the brane on the left, and the saturation anchor for the brane on the right, both as a function of $u_h$ . For a given value of $u_h$ , the saturation angle is the largest brane angle with a hole on the atoll. As the brane angle increases, the hole shrinks to a point that we call the saturation anchor. Our numerics are consistent with Result 6.2, in the Appendix, which states that the saturation anchor always lies inside the photon sphere. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 285 |
| 6.6 | According to Figure 6.3, the critical anchor lies at a finite distance from the defect, which depends on the order parameter $u_h$ , at the critical angle $\theta_c$ . Then for $u_h \neq 0$ , which corresponds to non-planar black holes on each slice, the phase transition at the critical angle is discontinuous because tiny island surfaces dominate the entropy functional at and above $\theta_c$ . For a review of tiny islands, see the Appendix 6.B. By sending $u_h \rightarrow 0$ , one recovers the result for planar black holes in our previous note Geng et al. (2021b) – the critical anchor drops into the defect, and the phase transition becomes continuous. Note that, in contrast to Geng et al. (2021b), the atoll does not saturate the brane at the critical angle when $u_h \neq 0$ . . . . .                                                                                                                                                                                                                                                                                     | 288 |
| 6.7 | Here are the numerically computed area differences between the Hartman-Maldacena (HM) surface and the island RT surface for each value of $u_h$ below the critical angle $\theta_c \approx .98687$ . The bottom panel zooms in on the region where the area difference vanishes. We find numerically that the Page angle equals the critical angle at roughly $u_h \approx .62777$ . While the area differences <i>above</i> the critical angle are also displayed for completeness, the actual dominant surface for $\theta > \theta_c$ will be the tiny island surface, for which the area difference actually diverges to negative infinity – see the Appendix 6.B. By sending $u_h \rightarrow 0$ , one obtains our old result Geng et al. (2021b), given by the thick dark curve, where the island RT surface degenerates into the tiny island surface at the critical angle – so for $u_h = 0$ , the area difference diverges to negative infinity at $\theta_c$ . For ease of comparison with our previous note, the vertical axis has been rescaled by the square of the horizon radius $u_+$ . . . . . | 289 |

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             |     |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 6.8  | Here we present our phase diagram for the doubly-holographic global black string model. Since the thermal BCFT state lives on the surface of a sphere, there is a scale dependence in the model that depends on $u_h$ – see Figure 6.1 and Figure 6.2. Each point on this plot gives the brane angle for which the area difference vanishes, which we call the Page angle $\theta_p$ . The Page angle was found to increase with $u_h$ . The island surfaces are dominant when the area difference is negative. It was shown in previous work Geng et al. (2021b) that the tiny island phase is dominant above the critical angle because the area difference diverges to $-\infty$ . In contrast, finite-size islands dominate below the critical angle when the brane lies above the Page angle. When $u_h > u_h^{\text{crit}}$ , which is $\approx .62777$ , the island surfaces are dominant only when the brane lies above the critical angle. . . . . | 290 |
| 7.1  | Cartoon of one copy of the KR braneworld considered in this note. The region behind the Karch-Randall brane has been excised. Two copies of this braneworld are glued along the brane. The intersection of the brane with the RT surface is not necessarily at right angles and will be determined by equation (7.9). . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 304 |
| 7.2  | Bound on the parameters $\lambda_b$ and $u_b$ for which we can reach the bath. The region inside the red dashed curves shows the theoretical bound we derived above. The region inside the blue solid curves is the numerical region we obtained from solving the differential equation. For one $\lambda_b$ a critical anchor is marked by a green dot and the shooting points making up the atoll are marked with a green line. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 310 |
| 7.3  | Bound on the parameters $\lambda_b$ and $u_b$ for which we can reach the bath for several brane angles $\theta_b$ . The region inside curves of the same color represent the values of $\lambda_b$ and $u_b$ for which we can reach the bath. This region gets wider as we increase $\theta_b$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    | 311 |
| 7.4  | Critical anchors as a function of brane angle $\theta_b$ for several values of $\lambda_b > -1$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 314 |
| 7.5  | Critical anchors as a function of brane angle $\theta_b$ for several values of $\lambda_b \leq -1$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 315 |
| 7.6  | Plots of the shooting points on the brane $u_b$ vs the anchoring points on the bath $\Gamma$ for several values of $\theta$ for $\lambda = -1$ . The yellow lines show the $\Gamma_{\min}$ values. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 316 |
| 7.7  | Polar plot of the three RT-candidates described above for $\Gamma = 0.9$ . The value brane angle is $\theta_b = 1.2$ , for which $\Gamma_{\min} = 0.873845$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 318 |
| 7.8  | Area of the RT surface as a function of $\Gamma$ and $\lambda_b$ for angles between $\theta = 0.4$ and $\theta = 1$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 322 |
| 7.9  | Area of the RT surface as a function of $\Gamma$ and $\lambda_b$ for angles $\theta = 1.2$ and $\theta = 1.4$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 323 |
| 7.10 | Area of the RT surface as a function of $\Gamma$ and $\theta_b$ for $\lambda_b < 0$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 325 |
| 7.11 | Area of the RT surface as a function of $\Gamma$ and $\theta_b$ for $\lambda_b \geq 0$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 326 |



|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |     |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 7.12 | $\Gamma$ and $\theta_b$ for which $\Delta\mathcal{A} = 0$ for several values of $\lambda_b$ . The horizon is at $\Gamma = 1$ . For the first few curves the area is positive for a small range near the horizon, independently of the angle. Our previous results are reproduced by the black curve $\lambda_b = 0$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 327 |
| 7.13 | Page angles $\theta_{Page}$ as a function of $\lambda_b$ . At $\lambda_b = 0$ , we obtain $\theta_P \approx 0.972$ , matching our previous results. Furthermore the curve monotonically increases. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           | 329 |
| 7.14 | Plot of $\Gamma_{min}$ for $\lambda_b = -1$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 332 |
| 7.15 | Areas for RT-1, RT-2 and RT-3 for $\lambda_b = -1$ and $\theta_b = 1.35$ . The yellow line marks the position of $\Gamma_{min}$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 333 |
| 8.1  | A schematic representation of the two-brane setup of interest in this paper. (a) is the classical bulk configuration consisting of two KR branes (in red) embedded in an ambient $\text{AdS}_{d+1}$ spacetime. (b) is the brane-localized “intermediate picture” of two interacting $\text{AdS}_d$ universes coupled together along a $(d-1)$ -dimensional interface (the “defect”). Both pictures can be seen as dual to a $(d-1)$ -dimensional CFT on this defect.                                                                                                                                                                                                                                                                                                                 | 344 |
| 8.2  | Schematic representations of the Ryu–Takayanagi surfaces for a fixed-time CFT subregion $\mathcal{R}$ in (a) AdS with no branes and (b) AdS with two KR branes. In (a), the surface is anchored to the boundary of $\mathcal{R}$ (i.e. $\partial\gamma_{\mathcal{R}} = \partial\mathcal{R}$ ); this is a Dirichlet condition on $\gamma_{\mathcal{R}}$ . In (b), the CFT subregion is the $(d-1)$ -dimensional defect at fixed time, and $\gamma_{\mathcal{R}}$ is anchored to the branes via a Neumann condition. We use Neumann for physical reasons; a Dirichlet condition would break diffeomorphism invariance on the brane. . . . .                                                                                                                                            | 346 |
| 8.3  | On the left, a cartoon of causal wedge inclusion (CWI) for a CFT subregion $\mathcal{R}$ at fixed boundary time $t = 0$ . The inner (blue) wedge is the causal wedge $\text{CW}[\mathcal{R}]$ , while the outer (green) wedge is the entanglement wedge $\text{EW}[\mathcal{R}]$ . We can further consider the cross-section of this picture with the initial bulk Cauchy slice (right), in which $\gamma_{\mathcal{R}}$ (green, dashed) is the RT surface wrapping around $\text{EW}[\mathcal{R}]$ while $\xi_{\mathcal{R}}$ (blue, dotted) is the codimension-2 surface wrapping around $\text{CW}[\mathcal{R}]$ . CWI implies that the region on this Cauchy slice bounded by $\xi_{\mathcal{R}}$ must be within the analogous region bounded by $\gamma_{\mathcal{R}}$ . . . . . | 350 |
| 8.4  | A fixed- $t$ , fixed- $\vec{x}$ slice of a planar black string in AdS consisting of two planar KR branes (in red), each at a fixed angle $\theta_1, \theta_2 \in (0, \frac{\pi}{2})$ with respect to the conformal boundary of the underlying $\text{AdS}_{d+1}$ spacetime. The defect (the red point) is at $u = 0$ and the horizon (the dashed line) is at $u = u_h$ . The branes have constant extrinsic curvature, and so this geometry solves the equations of motion with the RS action (8.5)–(8.6).                                                                                                                                                                                                                                                                           | 354 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                |     |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 8.5 | A plot of the two-brane DGP parameter space $(\lambda_1, \lambda_2)$ for $d = 4$ . The upper-right (green) region $\{\lambda_1 \geq 0\} \cap \{\lambda_2 \geq 0\}$ cannot be excluded by our swampland criteria because the horizon is always the (positive-entropy) RT surface. The colored lines demarcate the boundaries of regions excluded by the positive-horizon-entropy requirement for fixed $\theta_1$ and $\theta_2$ , which parameterize the branes as per (8.31)–(8.32). We exclude points below these lines. Additionally, the constraints have a reflection symmetry seen by swapping the labeling of the branes ( $1 \leftrightarrow 2$ ). We see that thinner wedges yield more restrictive bounds on $(\lambda_1, \lambda_2)$ . The intersection of all $(\theta_1, \theta_2)$ -dependent regions describes the couplings for which the horizon entropy is negative for all brane angles. This is shown in black and lies entirely within the unphysical $G_{\text{eff}} < 0$ region, with latter being shaded grey. . . . . | 362 |
| 8.6 | The points in $(\lambda_1, \lambda_2)$ parameter space for which we find non-horizon extremal surfaces, for fixed angles $\theta_1$ and $\theta_2$ . The colors correspond to (8.70) (red), (8.71) (blue), and (8.72) (green). The thick dashed lines are the $(\theta_1, \theta_2)$ -dependent bounding lines (8.52) below which $S_h < 0$ , while the light dashed line is $\lambda_1 + \lambda_2 = 0$ . We have also shaded the region in which either effective Newton constant is negative. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 373 |
| 8.7 | A cartoon of a possible time-dependent entangling surface in the two-brane setup with DGP couplings, shown on the $t = 0$ slice of the maximally extended black-string geometry, in which the horizons are identified. This surface would violate causal wedge inclusion—we can see that the entanglement wedge (in green) does not include the full exterior of either side and thus never includes the causal wedge of either defect. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 377 |
| 8.8 | The $(\lambda_1, \lambda_2)$ parameter space with the $d$ -dependent contours demarcating the intersection (over brane angles) of all regions excluded by horizon-entropy positivity. We depict contours for a variety of (brane) dimensions $d$ . The intersection for a given $d$ is the region below the corresponding contour. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                     | 388 |
| 9.1 | A Newton polygon showing the non-zero coefficients in $\text{NP}([5, 3, 2^2])$ . We have chosen the conventions from Baraglia and Kamgarpour (2018) and taken the vertices to be at $(j, n - s_j)$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   | 403 |
| 9.2 | An example of the (a) move of Kraft-Procesi . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 447 |
| 9.3 | An example of the (g) move of Kraft-Procesi . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          | 447 |

# Chapter 1: Introduction

Physics and mathematics have always been deeply interconnected. Just as a language consists of vocabulary—the words themselves—and syntax—the rules for forming coherent sentences—nature speaks in the language of physics, with mathematical structures serving as its vocabulary. Perhaps nowhere is this relationship more profound than in geometry. Geometry has played a pivotal role in shaping modern physics, while physics, in turn, has provided deep insights into mathematical questions about geometry.

One of the earliest and most well-known examples of this interplay is Einstein’s theory of general relativity. The groundbreaking realization that gravitational forces could be understood as the curvature of spacetime marked a revolutionary shift in both physics and mathematics. This insight catalyzed major developments in differential geometry, demonstrating the first profound impact of geometric thinking on fundamental physics.

The next great leap in our understanding of fundamental interactions came in the mid-20th century with the development of quantum field theory (QFT). A crucial aspect of this framework is the concept of gauge fields, which provide a natural way to describe the fundamental forces of nature. Mathematically, gauge fields correspond to connections on principal fibre bundles, a geometric formalism that has become indispensable in theoretical physics. The geometrisation of gauge theory has been key to understanding what is often called “the most successful scientific theory ever”<sup>1</sup>.

Over the past half-century, quantum field theory and gravity have been integrated into a broader theoretical framework—string theory. String theory extensively relies on geometric ideas, from Calabi-Yau compactifications to dualities that map

---

<sup>1</sup>This is how David Tong refers to the Standard Model in Merali (2021).

seemingly different physical regimes to one another. This unification represents one of the most sophisticated applications of geometry in modern physics.

This historical trajectory—from general relativity to gauge theory and ultimately to string theory—illustrates how geometry has shaped our deepest theories of nature. The themes of this dissertation lie at the intersection of these developments, focusing on how geometric structures continue to influence our understanding of quantum fields and gravity. Here, we provide some historical context to how this progress was developed, in the words of the great Steven Weinberg Weinberg (2003): *“Learn something about the history of science, or at a minimum the history of your own branch of science.(...) As a scientist, you’re probably not going to get rich. Your friends and relatives probably won’t understand what you’re doing. And if you work in a field like elementary particle physics, you won’t even have the satisfaction of doing something that is immediately useful. But you can get great satisfaction by recognizing that your work in science is a part of history.”*

## 1.1 Geometry in holography

We begin this introduction by examining the pivotal role that geometry plays in our understanding of gravitation. One of the central unresolved questions in this context is the *black hole information paradox* Hawking (1975). This paradox challenges the reconciliation of quantum mechanics with general relativity, and its resolution is expected to shed light on the nature of quantum gravity. In this section, we explore various geometric approaches to the problem and discuss how different formulations of holography provide crucial insights into potential resolutions.

### 1.1.1 Black hole information paradox

The black hole information paradox has been a central problem in theoretical physics, challenging the fundamental principles of quantum mechanics and gravity. At its core, the paradox arises from an apparent tension between the semi-classical

description of black hole evaporation and the expectation that quantum evolution should be unitary. In simple terms, if black holes evaporate thermally as predicted by Hawking radiation, the final state of the evaporation process is a mixed thermal state, seemingly in contradiction with the unitary evolution of quantum mechanics.

One of the key challenges in reconciling general relativity with our understanding of the other fundamental interactions is that gravity is inherently geometric—it manifests as the curvature of spacetime rather than as a force mediated by quantum fields. Unlike in conventional quantum field theory, where fields are defined on a fixed, typically flat, background, gravity requires us to consider quantum fields propagating in a dynamically curved spacetime. In most experimental scenarios, this does not pose a problem, as we operate within a regime where gravitational effects are well-approximated by the “*solar-system physics*” limit Mathur (2009). However, the question of how quantum fields behave at Planck-scale wavelengths, where curvature effects become extreme, remains an open problem. For wavelengths significantly longer than the Planck scale, a semi-classical approximation provides a well-understood framework for describing quantum fields in curved spacetime.

Perhaps one of the reasons the black hole information paradox is such a crucial problem is that black holes are both quantum objects and sources of strong gravitational effects. This unique combination makes them the perfect lab subject for testing ideas in quantum gravity. Any consistent theory of quantum gravity must be able to resolve the paradox, making black holes a natural arena for probing the fundamental principles of spacetime and information.

To gain insight into the black hole information paradox, it is useful to examine a concrete formulation of the problem. For this, we will closely follow the approach outlined in Raju (2021). We begin by presenting a general result concerning the degrees of freedom in QFT across a null surface, which will serve as the foundation for our discussion of Hawking radiation and its implications for unitarity.

### 1.1.2 Quantum Fields in Curved Spacetime and Thermal Entanglement

To understand the behavior of quantum fields in curved spacetime, we begin by considering a local region around a point—taken to be the origin—in a  $(d + 1)$ -dimensional spacetime. In this region, we introduce a set of coordinates where two directions,  $u$  and  $v$ , are taken to be null, while the remaining  $(d - 1)$  coordinates,  $\xi^a$ , parametrise the transverse spatial directions. The local line element can then be written as

$$ds^2 = -dudv + \delta_{ab}d\xi^a d\xi^b + \dots \quad (1.1)$$

where the ellipsis denotes potential higher-order corrections that become relevant away from the origin. However, in the immediate vicinity of the origin, the metric smoothly reduces to this simple Minkowski-like form, providing a convenient setup for analyzing quantum field correlations across the null surface  $u = 0$ .

Now, consider a freely propagating quantum field  $\phi$  in the vicinity of the origin. We are interested in the correlations of states associated with the semiclassical metric and, for this purpose, we will extract the modes on each side of the null surface. For distances shorter than the curvature scale but still larger than the Planck length, the spacetime appears locally Minkowskian. In this regime, quantum field correlations exhibit a universal short-distance behavior:

$$\langle \psi | \phi(x_1) \phi(x_2) | \psi \rangle \sim \frac{1}{s_{12}^{d-1}} \quad (1.2)$$

where  $s_{12}$  denotes the geodesic distance between the points  $x_1$  and  $x_2$ . To leverage this result, we now extract a set of local modes on either side of the null surface  $u = 0$ . This is done by smearing the quantum fields  $\phi(x)$  over a small region around  $u = 0$ , defining the smeared modes as follows:

$$\begin{aligned} a &= \int \frac{\partial_u \phi(u, v = 0, \xi^a)}{\sqrt{\pi \omega_0 \text{Vol}}} (-u)^{-i\omega_0} T(-u) du d^{d-1} \xi^a, \\ \tilde{a} &= \int \frac{\partial_u \phi(u, v = -\epsilon, \xi^a)}{\sqrt{\pi \omega_0 \text{Vol}}} u^{i\omega_0} T(u) du d^{d-1} \xi^a. \end{aligned} \quad (1.3)$$

Here,  $T(u)$  is a smooth, normalised <sup>2</sup> function localised near  $u = 0$ . The oscillatory factors  $(-u)^{-i\omega_0}$  and  $u^{i\omega_0}$  ensure that we extract the right-moving modes propagating along the  $u$  direction on either side of  $u = 0$ . These modes are causally disconnected, separated by a small distance  $\epsilon$  in the  $v$  direction.

Using the universal short-distance behavior established above, we can now compute the two-point function of the smeared modes. Applying the correlation structure of the quantum field  $\phi(x)$  across the null surface, we obtain:

$$\langle \Psi | aa^\dagger | \Psi \rangle = \langle \Psi | \tilde{a} \tilde{a}^\dagger | \Psi \rangle = \frac{1}{1 - e^{-2\pi\omega_0}}. \quad (1.4)$$

This result reveals a striking feature: the expectation value of the number operator follows a thermal Bose-Einstein distribution with an effective temperature given by

$$T_{\text{eff}} = \frac{\omega_0}{2\pi}. \quad (1.5)$$

Physically, this means that the extracted modes near the null surface exhibit a thermal character, as if they were thermally populated at this temperature. This thermal behavior arises solely from the structure of quantum correlations across the null surface and is independent of any particular choice of background curvature, as long as the curvature scale remains much larger than the local wavelengths of interest. This observation will play a crucial role in understanding Hawking radiation and the black hole information paradox.

In addition to the occupation number, we can also compute the two-point correlations between the extracted modes on either side of the null surface. This yields the expectation value:

$$\langle \Psi | a \tilde{a} | \Psi \rangle = \frac{e^{-\pi\omega_0}}{1 - e^{-2\pi\omega_0}}. \quad (1.6)$$

This nonzero correlation demonstrates that the two sets of modes—separated by the null surface—are entangled. The strength of this entanglement is frequency-dependent: it decays exponentially with  $\omega_0$ , indicating that low-energy modes (small

---

<sup>2</sup>For the normalisation conditions see Papadodimas et al. (2020)

$\omega_0$ ) exhibit stronger entanglement, while high-energy modes (large  $\omega_0$ ) are only weakly correlated. This entanglement structure will play a key role in the discussion of Hawking radiation, as it is precisely this type of entanglement across a black hole horizon that leads to the thermal properties of the outgoing radiation.

Furthermore, each set of modes behaves as a quantum harmonic oscillator, satisfying the standard commutation relations:  $[a, a^\dagger] = [\tilde{a}, \tilde{a}^\dagger] = 1$  and  $[a, \tilde{a}] = [a, \tilde{a}^\dagger] = 0$ . These relations confirm that the extracted modes on either side of the null surface form independent quantum degrees of freedom.

So far, we have considered a local region where the transverse directions are taken to be flat. However, in the context of black holes, we are primarily interested in spacetimes with spherical symmetry. Importantly, the key results derived here—particularly equations (1.4) and (1.6)—remain valid in the presence of curvature, as long as the local spacetime near the horizon can still be approximated as a Rindler-like geometry. This robustness will be essential when applying these results to black hole horizons and the emergence of Hawking radiation.

### 1.1.3 Hawking Radiation and the Formation of a Mixed State

We now apply the universal behavior of quantum correlations across a null surface to understand the emergence of Hawking radiation. The key insight is that a black hole horizon functions as a null surface, and the quantum correlations across this horizon give rise to thermal radiation. This observation leads directly to the conclusion that black holes emit radiation with a thermal spectrum, ultimately raising profound questions about the fate of information during black hole evaporation.

Consider a black hole formed by the gravitational collapse of infalling matter. At sufficiently late times, the exterior geometry can be well approximated by the Schwarzschild black hole solution in  $(d + 1)$ -dimensional spacetime. The metric is given by

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega_{d-2}^2 \quad (1.7)$$



where the blackening function takes the form

$$f(r) = 1 - \frac{M}{r^{d-2}} \quad (1.8)$$

Here,  $M$  is a parameter related to the mass of the black hole<sup>3</sup>, and the event horizon is located at  $r = r_h$ , where the blackening function vanishes,  $f(r_h) = 0$ . This metric describes the static, spherically symmetric exterior of the black hole and serves as the starting point for analyzing the quantum effects near the horizon that lead to Hawking radiation.

To proceed, it is convenient to introduce the tortoise coordinate  $r_*$ , defined by

$$\frac{dr_*}{dr} = \frac{1}{f(r)} \quad (1.9)$$

This transformation ensures that the radial coordinate smoothly accounts for the behavior near the black hole horizon. In particular, as we approach the horizon from the outside ( $r \rightarrow r_h^+$ ) or from the inside ( $r \rightarrow r_h^-$ ), the tortoise coordinate diverges to negative infinity,

$$r_* \rightarrow -\infty \quad \text{as} \quad r \rightarrow r_h. \quad (1.10)$$

The use of  $r_*$  will allow us to express the wave equation in a form analogous to a scattering problem in flat space, facilitating the analysis of Hawking radiation.

We consider a minimally coupled scalar field  $\phi$  in the black hole background and focus on time scales where the back-reaction of the evaporation process remains negligible. Since the Schwarzschild metric is static and spherically symmetric, the field modes can be characterised by three quantum numbers: the frequency  $\omega$ , the angular momentum  $\ell$ , and a radial component. The scalar field can thus be expanded in terms of a mode decomposition outside the horizon:

$$\phi = \sum_{\ell} \int d\omega \left[ A_{\omega,\ell} f^{out}(\omega, \ell, r_*) + B_{\omega,\ell} f^{in}(\omega, \ell, r_*) \right] e^{-i\omega t} Y_{\ell}(\Omega) + \text{h.c.} \quad (1.11)$$

---

<sup>3</sup>In particular it is related to the mass  $m$  via:  $(d-1)M = 8\pi^{\frac{2-d}{2}} \Gamma(\frac{d}{2}) Gm$

where  $Y_\ell(\Omega)$  are the spherical harmonics, and we have decomposed the field into a basis of purely ingoing and outgoing radial solutions, denoted by  $f^{\text{in}}$  and  $f^{\text{out}}$ , respectively. Near the horizon, these take the asymptotic forms<sup>4</sup>:

$$\begin{aligned} f^{\text{in}} &\xrightarrow[r \rightarrow r_h^+]{} h_{\omega,\ell} e^{-i\omega r_*}, \\ f^{\text{out}} &\xrightarrow[r \rightarrow r_h^+]{} e^{i\omega r_*} + g_{\omega,\ell} e^{-i\omega r_*}. \end{aligned} \quad (1.12)$$

The coefficients  $A_{\omega,\ell}$  and  $B_{\omega,\ell}$  correspond to mode operators associated with the outgoing and ingoing wave solutions in the exterior region. A similar expansion can be performed for the field inside the black hole, which we now turn to.

$$\phi = \sum_\ell \int d\omega \left( \tilde{A}_{\omega,\ell} e^{i\omega t} Y_\ell^*(\Omega) + C_{\omega,\ell} e^{-i\omega t} Y_\ell(\Omega) \right) \tilde{f}_{\text{out}}(\omega, \ell, r_*) + \text{h.c.} \quad (1.13)$$

where we have introduced new operators  $\tilde{A}_{\omega,\ell}$  and  $C_{\omega,\ell}$  for the interior region, along with a new basis function. Near the horizon, the outgoing mode takes the form:

$$f^{\text{out}} \xrightarrow[r \rightarrow r_h^-]{} e^{-i\omega r_*} \quad (1.14)$$

Since the quantum field  $\phi$  is a single field defined throughout the entire spacetime, it must be continuous across the horizon. This continuity condition imposes a relation between the mode coefficients on the exterior and interior sides of the horizon:

$$C_{\omega,\ell} = A_{\omega,\ell} h_{\omega,\ell} + B_{\omega,\ell} g_{\omega,\ell}. \quad (1.15)$$

To further analyze the near-horizon behavior of the field modes, we introduce Kruskal coordinates, which are well-suited for describing the full black hole spacetime and smoothly covering the horizon. The transformation is given by:

$$u = -\frac{1}{\kappa} e^{\kappa(r_*-t)}, \quad v = \frac{1}{\kappa} e^{\kappa(r_*+t)}. \quad (1.16)$$

---

<sup>4</sup>Here  $h_{\omega,\ell}$  and  $g_{\omega,\ell}$  are the greybody factors, which describe the partial transmission and reflection of waves due to the black hole potential. While they are important for a detailed understanding of Hawking radiation, they will play no role in our discussion.

where  $\kappa$  is the *surface gravity* of the black hole. In these coordinates, the metric retains the form of equation (1.1), but now with spherical transverse coordinates. Importantly, the event horizon is located at  $u = 0$ , explicitly identifying it with the null surface we previously considered. This coordinate system provides a natural framework for analyzing quantum field modes near the horizon.

With this setup, we define the near-horizon modes as described earlier. To relate these to the global modes in the black hole spacetime, we introduce the identifications:

$$a = a_{\omega_0\kappa,\ell}, \quad \tilde{a} = \tilde{a}_{\omega_0\kappa,\ell}. \quad (1.17)$$

where the operators  $a_{\omega,\ell}$  and  $\tilde{a}_{\omega,\ell}$  correspond to smeared versions of the global mode operators  $A_{\omega,\ell}$  and  $\tilde{A}_{\omega,\ell}$ :

$$a_{\omega,\ell} = \int A_{\omega',\ell} q(\omega', \omega) d\omega', \quad \tilde{a}_{\omega,\ell} = \int \tilde{A}_{\omega',\ell} \tilde{q}(\omega', \omega) d\omega' \quad (1.18)$$

These expressions describe how the local field modes near the horizon relate to the global field modes in Schwarzschild spacetime. This mapping will allow us to extract the thermodynamic properties of the outgoing radiation.

Using the universal short-distance correlations derived earlier, we now compute the expectation value of the number operator for the outgoing modes. We find that it follows a thermal Bose-Einstein distribution, confirming the thermal nature of the radiation emitted by the black hole:

$$\langle \Psi | a_{\omega,\ell} a_{\omega,\ell}^\dagger | \Psi \rangle = \frac{1}{1 - e^{-\beta\omega}}. \quad (1.19)$$

where the temperature, also known as the *hawking temperature* is given by  $T_H = \frac{\kappa}{2\pi}$ . We remark that we have imposed no constraints on the ingoing modes  $B_{\omega\ell}$ . There are two common choices for these, one is having them in their vacuum state (*Unruh state*) and another is having them thermally populated (*Kruskal state*).

Therefore, using general arguments about the correlations of a quantum field across a null hypersurface in a curved spacetime, we have established that black holes

formed from collapse emit radiation, that the emitted pairs are entangled across the horizon, and that the outgoing modes exhibit a thermal spectrum. These observations lead naturally to the question of how the statistical properties of the radiation can be characterised more precisely.

To formalise this, we define the number operator for outgoing modes with angular momentum  $\ell$  and frequency  $\omega$  as

$$N_{\omega,\ell} = a_{\omega,\ell}^\dagger a_{\omega,\ell}. \quad (1.20)$$

For a black hole state  $|\Psi\rangle$ , the expectation value of any power of the number operator satisfies

$$\langle \Psi | (N_{\omega,\ell})^q | \Psi \rangle = \text{Tr} (\rho_{\omega,\ell} (N_{\omega,\ell})^q), \quad (1.21)$$

where the density matrix associated with each mode is given by

$$\rho_{\omega,\ell} = \frac{1}{1 - e^{-\beta\omega}} e^{-\beta\omega N_{\omega,\ell}}. \quad (1.22)$$

This result demonstrates that all statistical moments of the emitted radiation match those of a thermal state, reinforcing the interpretation of Hawking radiation as thermal. A natural question that arises is whether there exist correlations between different frequency modes that could account for the thermal nature of the outgoing radiation. However, as we will now show, no such correlations exist.

If a pure state undergoes gravitational collapse to form a black hole, the argument above demonstrates that the resulting black hole radiates thermally, leading to a final state described by a mixed thermal density matrix rather than a pure state. This apparent transition from a pure state to a mixed state violates unitary evolution, implying a fundamental loss of information.<sup>5</sup> Since pure states encode intricate

---

<sup>5</sup>However, more recent developments suggest that information about the black hole interior may be encoded in a highly state-dependent manner Papadodimas and Raju (2013a, 2014a), potentially providing a resolution to the information paradox without requiring firewalls. This framework suggests that the interior degrees of freedom are effectively encoded in the radiation subsystem, in a way consistent with unitary evolution.

quantum correlations, whereas thermal states lack such structure, the semiclassical description suggests that information about the initial state is irretrievably lost in the evaporation process. This conclusion challenges the foundational principles of quantum mechanics and gives rise to the black hole information paradox.

One might be tempted to accept this conclusion and assume that black holes inherently violate the principles of quantum mechanics in some unknown way. However, there are strong reasons to believe that black hole evolution remains unitary. In the following, we present two key arguments supporting this hypothesis.

#### 1.1.4 State Counting and Microstates in String Theory

To diagnose the issue raised by the black hole information paradox, we examine the entropy of the Hawking radiation. A fundamental principle of quantum mechanics dictates that the von Neumann entropy of a pure state is zero. If black hole evaporation were a unitary process, the final state—including both the emitted radiation and any remaining remnants—should also have zero entropy. However, Hawking’s calculation predicts that the radiation follows a thermal distribution, leading to a mixed-state density matrix with nonzero entropy.

If this conclusion were taken at face value, it would imply a fundamental violation of unitarity, suggesting that information is lost in the evaporation process. Resolving this contradiction requires either modifying Hawking’s semiclassical analysis or identifying a mechanism by which quantum correlations restore coherence in the final state.

Bekenstein originally conjectured Bekenstein (1973a), and Hawking later confirmed Bekenstein (1973a); Hawking (1976a), that black holes obey a thermodynamic law in which their entropy, known as the *Bekenstein-Hawking entropy*, is proportional to the horizon area:

$$S_{\text{BH}} = \frac{A}{4G_N}. \quad (1.23)$$

This result strongly suggests that black holes should be regarded as thermodynamic

objects, with their entropy reflecting an underlying count of microscopic degrees of freedom. However, the precise nature of these microstates remains a key question in the search for a quantum theory of gravity.

Microscopic state counting is, in general, a formidable challenge, requiring a complete theory of quantum gravity. String theory provided a breakthrough in this direction when Strominger and Vafa [Strominger and Vafa (1996a)] successfully performed such a counting for a specific class of black holes. They considered extremal black holes in five-dimensional string compactifications, particularly within type II string theory on  $K3 \times S^1$ . The low-energy effective action in this setup includes the Einstein metric  $g_{\mu\nu}$ , a dilaton  $\phi$ , and two gauge field strengths:  $\tilde{H}$ , associated with the NS-NS 3-form, and  $F$ , the RR 2-form field strength. The action takes the form:

$$S_{\text{eff}} = \frac{1}{16\pi} \int d^5x \sqrt{-g} \left( R - \frac{4}{3}(\nabla\phi)^2 - \frac{e^{-4\phi/3}}{4}H^2 - \frac{e^{2\phi/3}}{4}F^2 \right). \quad (1.24)$$

In this theory, black holes carry two charges,  $Q_H$  and  $Q_F$ , associated with the field strengths  $\tilde{H}$  and  $F$ , respectively. These charges play a crucial role in identifying the microscopic degrees of freedom that account for the black hole entropy.

The black hole solutions that preserve 1/4 of the  $\mathcal{N} = 4$  supersymmetry correspond to BPS states. These solutions satisfy a charge-dependent mass bound, ensuring their stability against decay into lower-mass configurations. The metric for an extremal charged black hole in this theory is given by

$$ds^2 = - \left( 1 - \frac{r_0^2}{r^2} \right)^2 dt^2 + \left( 1 - \frac{r_0^2}{r^2} \right)^{-2} dr^2 + r^2 d\Omega_3^2, \quad (1.25)$$

where the event horizon is located at  $r = r_0$ , with

$$r_0 = \left( \frac{8Q_H Q_F^2}{\pi^2} \right)^{1/6}. \quad (1.26)$$

From this, the horizon area is determined to be

$$A = 8\pi \sqrt{\frac{Q_H Q_F^2}{2}} \quad (1.27)$$

which leads to the Bekenstein-Hawking entropy:

$$S_{\text{BH}} = \frac{A}{4G_N} = 2\pi\sqrt{Q_H \frac{Q_F^2}{2}}. \quad (1.28)$$

Up to this point, the result is entirely macroscopic, derived from the classical general relativity description of black hole entropy. The key insight of Strominger and Vafa was that the same black hole can be understood microscopically within string theory as a bound state of D-branes. By counting the degeneracy of such bound states, they were able to reproduce the Bekenstein-Hawking entropy, providing a statistical interpretation of black hole entropy in terms of fundamental quantum degrees of freedom.<sup>6</sup>

Type IIB string theory has odd-dimensional D-branes Polchinski (2007a); Dai et al. (1989); Joh, which can wrap cycles of the internal space  $K3 \times S^1$ . The RR charge Polchinski (1995)  $Q_F$  is carried by D-branes wrapping  $K3 \times S^1$  Horowitz and Strominger (1996). The momentum charge  $Q_H$  is carried by left-moving excitations along  $S^1$ .

Therefore, the black hole corresponds to a bound state of  $D1$ -branes wrapping  $S^1$  (carrying  $D1$  charge  $Q_F$ ),  $D5$ -branes wrapping  $K3 \times S^1$  (carrying  $D5$  charge), and left-moving excitations along  $S^1$  (carrying momentum charge  $Q_H$ ). This bound state preserves precisely 1/4 of the supersymmetries.

The system can be effectively described by a  $(1+1)$ -dimensional theory on the  $D1$ - $D5$  worldvolume. This effective theory is a two-dimensional conformal field theory (CFT) with central charge

$$c = 6 \times \left( \frac{Q_F^2}{2} + 1 \right). \quad (1.29)$$

The number of states with left-moving momentum, which preserve 1/4 of the supersymmetries, is given by

$$n = Q_H \quad (1.30)$$

---

<sup>6</sup>For a broad review of black hole microstates, see Mathur (2005). For discussions of supergravity microstate geometries, see Bena and Warner (2008).

Using Cardy's formula for the degeneracy of left-moving excitations Cardy (1986),

$$d(n, c) \sim \exp\left(2\pi\sqrt{\frac{nc}{6}}\right) \quad (1.31)$$

the statistical entropy is

$$S_{stat} = \ln d(Q_F, Q_H) \sim 2\pi\sqrt{Q_H\left(\frac{1}{2}Q_F^2 + 1\right)} \quad (1.32)$$

For large values of  $Q_F$  and  $Q_H$ , the leading-order term precisely matches the Bekenstein-Hawking entropy, providing a microscopic justification for black hole entropy within string theory.

This precise agreement demonstrates that black hole entropy genuinely counts microstates in a manner consistent with statistical mechanics. Since the entropy derived from microstate counting matches the Bekenstein-Hawking entropy, black hole thermodynamics is not merely an analogy but a true statistical description. This strongly suggests that black holes possess a unitary quantum description in terms of their underlying microstates. Consequently, this result provides compelling motivation for resolving the black hole information paradox within the framework of string theory.

### 1.1.5 The Holographic Correspondence and the Black Hole Information Problem

Another strong piece of evidence for the unitarity of black hole evolution comes from the AdS/CFT correspondence. One can formulate the black hole information paradox in the context of an asymptotically Anti-de-Sitter (AdS) black hole. If the AdS/CFT duality holds, the black hole should have a description in terms of a dual conformal field theory, which is manifestly unitary. Since the CFT is a unitary quantum mechanical theory with a well-defined Hilbert space, any dual description, including black hole evaporation in AdS, must also be unitary. This reasoning offers a powerful argument that the apparent information loss in semiclassical gravity must be resolved in the full quantum gravitational theory.



Perhaps one of the earliest and often underappreciated pieces of evidence for gauge/gravity duality came from Gerard 't Hooft's seminal work in 1974 Hooft (1974). He observed that the large- $N$  expansion of  $SU(N)$  Yang-Mills theory exhibits a striking resemblance to perturbative string theory. By classifying Feynman diagrams according to the topology of the Riemann surfaces on which they can be drawn, he showed that planar diagrams dominate at leading order, scaling as  $N^2$ , while higher-genus contributions are suppressed by factors of  $1/N$ . This observation suggested that large- $N$  gauge theories could be reorganised as a sum over worldsheets of a dual string theory. Decades later, this perspective found a precise realization in the AdS/CFT correspondence. 't Hooft's insights laid the foundation for modern holography, where the strong coupling limit of certain gauge theories is conjectured to describe gravitational dynamics in higher-dimensional spacetimes.

During the next decade, efforts to formulate a string theory that directly describes quantum chromodynamics (QCD) were largely abandoned due to difficulties in reproducing its asymptotic freedom. Instead, theoretical physics shifted its focus toward superstring theory and supergravity, which provided a natural framework for unifying fundamental interactions, including gravity. In 1981, Polyakov Polyakov (1981) introduced the path integral quantization of string theory using Liouville field theory, offering a systematic approach to summing over random surfaces. His work reinforced the idea that gauge theories could have a dual string-theoretic description, as he argued that the confining flux tubes of gluons in QCD could be understood as forming string-like worldsheets. Despite these insights, a concrete realization of a string-theoretic formulation of QCD remained elusive. Instead, string theory research increasingly shifted toward higher-dimensional superstrings, the study of D-branes, supersymmetry, and potential gravitational duals of gauge theories.<sup>7</sup>

---

<sup>7</sup>Much of this progress is documented in the influential “Green Books” Green et al. (1988a,b), which were later succeeded by more modern treatments such as Polchinski (2007a,b); Becker et al. (2006).

It was not until 1993 that 't Hooft 't Hooft (1993) proposed a novel reinterpretation of the Bekenstein-Hawking entropy law for black holes. Ordinarily, one might expect the number of degrees of freedom—and consequently, the entropy—of a physical system to scale with its volume. However, the black hole entropy formula, which is proportional to the horizon area rather than the enclosed volume, suggests a fundamentally different organizing principle for quantum gravity. 't Hooft argued that this result "suggests that quantum gravity should be described entirely by a topological quantum field theory, in which all physical degrees of freedom can be projected onto the boundary." In the same work, he coined the term *the hologram*, laying the foundation for what is now known as the *holographic principle*. The following year, Susskind Susskind (1995) further developed this idea within the context of string theory, demonstrating that string interactions and longitudinal boosts naturally lead to a holographic description of physics.

#### 1.1.5.1 AdS/CFT

We now consider a modification of the Strominger-Vafa argument that leads to an important conjecture. For this discussion, we follow the review by Hubeny Hubeny (2015). Consider a stack of  $N$  coincident  $D3$ -branes in Type IIB string theory. This system admits two complementary descriptions.

At weak coupling, when  $g_s N \ll 1$ , the theory is best understood in terms of open strings ending on the  $D3$ -branes and closed strings propagating in the 10-dimensional flat background. In this regime, the system is described by a supersymmetric  $U(N)$  Yang-Mills theory living on the branes, while closed strings mediate bulk interactions.

At strong coupling, when  $g_s N \gg 1$ , the backreaction of the branes on the surrounding spacetime cannot be ignored. The system is now more naturally described in terms of a curved gravitational background, given by the extremal black 3-brane

solution:

$$ds^2 = f(r)^{-1/2} \eta_{\mu\nu} dx^\mu dx^\nu + f(r)^{1/2} (dr^2 + r^2 d\Omega_5^2), \quad (1.33)$$

where  $x^\mu$  are the four coordinates along the  $D3$ -branes, and the warp factor takes the form:

$$f(r) = 1 + \frac{4\pi g_s N \ell_s^4}{r^4}. \quad (1.34)$$

To establish a precise duality, we take a low-energy limit that isolates the relevant degrees of freedom. In this limit, high-energy bulk excitations are removed from the open string sector. The presence of  $D3$ -branes breaks half of the supersymmetry in Type IIB string theory<sup>8</sup>, leading to a decoupling of massive string modes. Consequently, the low-energy dynamics on the branes is governed by a four-dimensional  $SU(N)$  super Yang-Mills (SYM) theory with  $\mathcal{N} = 4$  supersymmetry.

To analyze the strongly coupled regime, we take the low-energy limit of the black brane geometry by focusing on its near-horizon region. In this limit, as  $r \rightarrow 0$ , the metric simplifies to

$$ds^2 = \frac{r^2}{\ell^2} \eta_{\mu\nu} dx^\mu dx^\nu + \frac{\ell^2}{r^2} dr^2 + \ell^2 d\Omega_5^2. \quad (1.35)$$

This describes an  $AdS_5 \times S^5$  spacetime, where both factors share the same characteristic radius, given by

$$\ell = (4\pi g_s N)^{1/4} \ell_s. \quad (1.36)$$

Since the gauge theory description remains valid at all values of the coupling, Maldacena conjectured Maldacena (1998a) that these two seemingly different descriptions—one in terms of a gravitational background and the other in terms of a gauge theory—are in fact equivalent, leading to the celebrated duality:

**Type IIB string theory on  $AdS_5 \times S^5$     $\iff$     $\mathcal{N} = 4$  SYM on  $\mathbb{R}^{1,3}$ .**

---

<sup>8</sup>Type IIB string theory has 32 supercharges, and the  $D3$ -branes preserve only 16 of them.

This correspondence implies that the dynamics of a strongly coupled gauge theory can be equivalently described as a weakly coupled gravitational theory in a higher-dimensional curved spacetime.

The key parameters on both sides of the duality satisfy the following relations:

$$4\pi g_s = g_{\text{YM}}^2, \quad \frac{\ell}{\ell_s} = (4\pi g_s N)^{1/4} \sim \lambda^{1/4}, \quad (1.37)$$

where  $\lambda = g_{\text{YM}}^2 N$  is the 't Hooft coupling. The duality is best understood in the classical gravity limit, which requires  $\lambda \gg 1$  and  $N \gg 1$ . In this regime, the AdS curvature remains small compared to the string length  $\ell_s$ , ensuring that higher-order string corrections are suppressed. Additionally, taking  $N \gg 1$  suppresses quantum gravity effects, allowing a semi-classical supergravity description on the AdS side.

In the classical gravity limit, it is useful to focus on the  $AdS_5$  factor, as the  $S^5$  primarily encodes the internal R-symmetry structure of the dual conformal field theory. Thus, classical gravity in  $AdS_5$  is dual to a strongly coupled CFT living on its boundary. The AdS/CFT dictionary establishes that every bulk field in  $AdS$  corresponds to an operator in the CFT, and the generating functional of correlation functions in the CFT is related to the partition function of the gravitational theory in AdS.

$$\left\langle \exp \int d^d x \phi_0(x) O(x) \right\rangle_{\text{CFT}} = Z_{\text{gravity}}[\phi_0] \approx e^{-I_{\text{SUGRA}}[\phi_0]}. \quad (1.38)$$

Here,  $O(x)$  is a local operator in the boundary CFT (located at  $z = 0$ ), sourced by a bulk field  $\phi$  in AdS. The function  $Z_{\text{gravity}}[\phi_0]$  represents the gravitational partition function with boundary condition  $\phi(z \rightarrow 0, x) \rightarrow \phi_0(x)$ . The term  $I_{\text{SUGRA}}[\phi_0]$  denotes the classical supergravity action, evaluated on-shell with the same boundary condition. In the strong coupling and large- $N$  limit, the classical action  $I_{\text{SUGRA}}$  dominates the gravitational path integral. This relation encodes the generating functional of CFT correlation functions in terms of bulk gravitational dynamics.

With the holographic duality established, we now return to the problem of black hole evaporation. If this process is analyzed from the string theory side of the

duality, there must be an equivalent description within the conformal field theory. Since the CFT is a manifestly unitary quantum theory, it follows that the black hole evaporation process should also be unitary. This provides strong evidence that information is not lost, but rather encoded in a subtle way in the final state of the evaporation process.

We can make this connection more precise by leveraging the computational power of the holographic duality. As previously discussed, a key diagnostic of the black hole information paradox is the behavior of the entropy of the radiation. In general, computing the entropy of subsystems in an arbitrary region of a conformal field theory is a challenging problem. However, as we will now explore, AdS/CFT provides powerful tools to address this question, offering a framework to systematically compute entanglement entropy in strongly coupled quantum systems.

#### 1.1.5.2 Ryu-Takayanagi formula

We recall that the entropy of a black hole is encoded in the area of its horizon, as given by the Bekenstein-Hawking formula. In the AdS/CFT correspondence, black holes in the bulk are dual to thermal states in the boundary CFT. Since bulk and boundary physics are identified through the duality, the entropy of the black hole should have a natural interpretation within the CFT. In particular, the entropy of the Hawking radiation must correspond to the entanglement entropy of a specific region in the CFT. Understanding this correspondence provides a crucial step toward resolving the black hole information paradox within a holographic framework.

Conversely, given a constant time slice, one can ask: for a given subregion  $\mathcal{A}$  in a  $d$ -dimensional conformal field theory, what is the corresponding geometric quantity in the  $(d + 1)$ -dimensional bulk that encodes the entropy of a state defined within this region? Ryu and Takayanagi (2006a,b) provided a striking answer to this question, proposing a generalisation of the Bekenstein-Hawking entropy

formula to entanglement entropy:

$$S_{\mathcal{A}} = \frac{\text{Area}(\gamma_{\mathcal{A}})}{4G_N^{(d+1)}} \quad (1.39)$$

where  $\gamma_{\mathcal{A}}$  is the minimal-area bulk surface that is anchored to the AdS boundary and homologous to  $\mathcal{A}$ <sup>9</sup>. Here,  $G_N^{(d+1)}$  is the  $(d+1)$ -dimensional Newton's constant. This formula provides a remarkable bridge between quantum entanglement and spacetime geometry, suggesting that the structure of entanglement in the boundary CFT is encoded in the bulk as extremal surfaces. This insight laid the foundation for the modern holographic approach to quantum information in gravity.

Originally, the Ryu-Takayanagi formula was tested in the context of the  $\text{AdS}_3/\text{CFT}_2$  correspondence. In this case, the central charge  $c$  of the boundary CFT is related to the AdS radius  $\ell$  via

$$c = \frac{3\ell}{2G} \quad (1.40)$$

where  $G$  is the Newton's constant of the  $\text{AdS}_3$  spacetime. The metric for global  $\text{AdS}_3$  is given by

$$ds^2 = \ell^2(-\cosh^2 \rho dt^2 + d\rho^2 + \sinh^2 \rho d\theta^2) \quad (1.41)$$

where  $\rho$  is the radial coordinate. Here, the dual CFT is  $(1+1)$ -dimensional and defined on a spatial circle of length  $R$ . On a constant time slice, minimal-area surfaces in the bulk reduce to geodesics. The entanglement entropy of an interval of length  $r$  in the boundary CFT, subtended between  $\theta = 0$  and  $\theta = \frac{2\pi r}{R}$ , is then given by

$$S_{\mathcal{A}} = \frac{c}{3} \log \left( e^{\rho_0} \sin \frac{\pi r}{R} \right) \quad (1.42)$$

This result precisely matches the independent CFT computation of entanglement entropy Calabrese and Cardy (2004); Holzhey et al. (1994), providing the first compelling evidence in support of the Ryu-Takayanagi formula. This agreement established a key link between holographic geometry and quantum entanglement, reinforcing the broader AdS/CFT framework.

---

<sup>9</sup>Homology in this context means that  $\gamma_{\mathcal{A}}$  shares the same boundary as  $\mathcal{A}$ , i.e.,  $\partial\mathcal{A} = \partial\gamma_{\mathcal{A}}$ , and can be continuously deformed into  $\mathcal{A}$ .

### 1.1.5.3 Replica Trick and Holographic Derivation

To better understand the entanglement entropy of a QFT region, it is useful to first consider its discrete counterpart, where the degrees of freedom reside on a spatial lattice with spacing  $\epsilon$ . The Hilbert space of the entire system is then given by a tensor product over individual lattice sites:

$$\mathcal{H} = \otimes_i \mathcal{H}_i \quad (1.43)$$

We want to compute how the degrees of freedom of a subregion  $\mathcal{A}$  are entangled with those in the complement  $\mathcal{A}^c$  in a pure state  $|\Psi\rangle \in \otimes_i \mathcal{H}_i$ . The standard way of doing this is by partitioning the Hilbert space over these regions

$$\otimes \mathcal{H}_i = \mathcal{H}_A \otimes \mathcal{H}_{A^c} \quad (1.44)$$

and constructing the reduced density matrix associated with the subregion  $\mathcal{A}$ . This is obtained by tracing out the degrees of freedom in  $\mathcal{A}^c$ :

$$\rho_A = \text{Tr}_{A^c} (|\Psi\rangle \langle\Psi|) \quad (1.45)$$

The entanglement entropy is defined as the von Neumann entropy of this reduced density matrix:

$$S_A = -\text{Tr}_A (\rho_A \log \rho_A) \quad (1.46)$$

In a finite discrete system, directly computing  $S_A$  can be challenging, particularly as the number of lattice sites increases. However, in principle, one can diagonalise  $\rho_A$  and compute its entropy explicitly. This formulation provides an intuitive foundation for understanding entanglement entropy in continuous quantum field theories.

One can define a more general class of entanglement measures known as the *Rényi entropies*, which provide a more detailed characterization of the reduced density matrix  $\rho_A$ . These entropies depend on a parameter  $q \in \mathbb{Z}$  and are defined as

$$S_A^{(q)} = \frac{1}{1-q} \log \text{Tr}_A (\rho_A^q) \quad (1.47)$$

By analytically continuing  $q$  to real values, one recovers the von Neumann entropy in the limit

$$S_A = \lim_{q \rightarrow 1} S_A^{(q)}. \quad (1.48)$$

Rényi entropies encode more than just the leading order entanglement, as they are sensitive to the full eigenvalue spectrum of  $\rho_A$ .

To extend these ideas to a continuum field theory, we define the reduced density matrix  $\rho_A$  on a Cauchy slice  $\Sigma_{t=0}$ . The field configurations naturally decompose into those supported on the region  $\mathcal{A}$  and its complement  $\mathcal{A}^c$ , so that a given configuration takes the form  $\Phi = \{\Phi^{\mathcal{A}}(x), \Phi^{\mathcal{A}^c}(x)\}$ . The reduced density matrix elements is then obtained by tracing over the degrees of freedom in  $\mathcal{A}^c$ , leading to a Euclidean path integral representation:

$$[\rho_A]_{ab} = \int \mathcal{D}\Phi e^{-S_E} \delta(\Phi^{\mathcal{A}}(t=0^-) - \Phi_a^{\mathcal{A}}) \delta(\Phi^{\mathcal{A}}(t=0^+) - \Phi_b^{\mathcal{A}}). \quad (1.49)$$

This expression describes a transition amplitude between field configurations  $\Phi_a^{\mathcal{A}}$  and  $\Phi_b^{\mathcal{A}}$  at  $t=0$ .

For states with nontrivial evolution, special care must be taken to ensure that the causal structure is respected. The standard way to handle this is via the Schwinger-Keldysh formalism, in which  $\Psi[\Phi_a]$  and  $\Psi[\Phi_b]$  correspond to the forward and backward time evolution branches.<sup>10</sup>

---

<sup>10</sup>More formally in a QFT we replace our wavefunction with a wavefunctional, which has a Euclidean path integral representation

$$\Psi[\Phi_0(\vec{x})] = \langle \Phi_0(\vec{x}) | \Psi | \Phi_0(\vec{x}) \rangle = \int_{t=-\infty}^{t=0, \Phi(t=0, \vec{x})=\Phi_0} \mathcal{D}\Phi e^{-S_E}$$

and similarly for  $\Psi^*$ . Thus the density matrix is given by the Euclidean path integral

$$\rho_A \sim \int \mathcal{D}\Phi^{\mathcal{A}^c}(t=0, \vec{x} \in \mathcal{A}^c) \langle \Phi^{\mathcal{A}^c} | \Psi | \Phi^{\mathcal{A}^c} \rangle \langle \Psi | \Phi^{\mathcal{A}^c} | \Psi | \Phi^{\mathcal{A}^c} \rangle$$

Specifying the boundary conditions at  $t=0^-$  to be  $\Phi_a^{\mathcal{A}}(\vec{x} \in \mathcal{A})$  and at  $t=0^+$  to be  $\Phi_b^{\mathcal{A}}(\vec{x} \in \mathcal{A})$ , we obtain the matrix elements:

$$[\rho_A]_{ab} = \langle \Phi_a^{\mathcal{A}} | \rho_A | \Phi_b^{\mathcal{A}} \rangle = \langle \Phi_a^{\mathcal{A}} | \rho_A | \Phi_b^{\mathcal{A}} \rangle \sim \int \mathcal{D}\Phi^{\mathcal{A}^c}(t=0, \vec{x} \in \mathcal{A}^c) (\langle \Phi_a^{\mathcal{A}} | \langle \Phi^{\mathcal{A}^c} | | \Psi \rangle \langle \Psi | \langle \Phi_b^{\mathcal{A}} | \langle \Phi^{\mathcal{A}^c} | )$$



To compute arbitrary powers of  $\rho_{\mathcal{A}}$ , we use the replica trick. The idea is to interpret the computation the computation above as cutting the path integral along  $\mathcal{A}$  at  $t = 0$  and forcing the fields on either side to take some specific values  $\Phi_{a,b}$ . Repeating this procedure a number of times allows us to compute  $\text{Tr} \rho_{\mathcal{A}}^q$ . To do this we introduce  $q$  copies (or replicas) of the field theory, each of which is prepared in the same state and sewn together in a cyclic manner. This corresponds to computing the path integral on an  $q$ -sheeted Riemann surface  $\mathcal{B}_q$ , where the fields are identified cyclically (like above) across the sheets at the cut along  $\mathcal{A}$ .

$$Z_q = \text{Tr}(\rho_{\mathcal{A}}^q) \equiv Z[\mathcal{B}_q] \quad (1.50)$$

Thus, after normalizing  $\rho_{\mathcal{A}}^q$ , the Rényi entropies take the form:

$$S_q(\mathcal{A}) = \frac{1}{1-q} \log(\text{Tr}(\rho_{\mathcal{A}}^q)) = \frac{1}{1-q} \log \left( \frac{Z_q[\mathcal{A}]}{(Z_1[\mathcal{A}])^q} \right) \equiv \frac{1}{1-q} \log \left( \frac{Z[\mathcal{B}_q]}{(Z[\mathcal{B}])^q} \right). \quad (1.51)$$

where the normalization factor is given by the partition function of the original (uncut) manifold

$$Z[\mathcal{B}] = \text{Tr} e^{-\beta H} = \int_{\mathcal{B}} \mathcal{D}\Phi e^{-S_E} \quad (1.52)$$

The task now is to analytically continue  $q$  to non-integer values and extract the entanglement entropy by taking the  $q \rightarrow 1$  limit. This continuation is highly nontrivial in the field theory. However, Lewkowycz and Maldacena (2013a) recognised that this problem is significantly easier to handle in the dual gravitational description.

In the holographic setting, we denote the bulk spacetime dual to  $\mathcal{B}_q$  as  $\mathcal{M}_q$ , which has  $\mathcal{B}_q$  as its conformal boundary. Applying the AdS/CFT duality from equation (1.38), we obtain:

$$Z[\mathcal{B}_q] = Z[\mathcal{M}_q] \approx e^{-I_{SUGRA}[\mathcal{M}_q]} \quad (1.53)$$

---

which after using the expression for  $\Psi[\Phi(\vec{x})]$  and integrating over the fields supported on  $\mathcal{A}^c$  becomes 1.49. For a detailed review on entanglement entropy in quantum field theory, see for example Nishioka (2018). For a discussion more closely related to the Ryu-Takayanagi formula and holography, see Rangamani and Takayanagi (2017a,b).

Substituting this into the definition of the Rényi entropy, we get:

$$S_q = \frac{1}{1-q} \log \left( \frac{Z[\mathcal{B}_q]}{(Z[\mathcal{B}])^q} \right) = \frac{1}{1-q} (\log(Z[\mathcal{M}_q]) - q \log(Z[\mathcal{M}])) \quad (1.54)$$

A key insight from Lewkowycz and Maldacena is that the bulk spacetime  $\mathcal{M}_q$  inherits the same  $\mathbb{Z}_q$  symmetry as  $\mathcal{B}_q$ , allowing us to construct the quotient space  $\hat{\mathcal{M}}_q = \mathcal{M}_q/\mathbb{Z}_q$ . This enables the orbifold relation

$$Z[\mathcal{M}_q] = qZ[\hat{\mathcal{M}}_q]. \quad (1.55)$$

Then using equations 1.53 and 1.55 in 1.54, we have that the Renyi entropy is

$$S_q = \frac{q}{q-1} (I[\hat{\mathcal{M}}_q] - I[\hat{\mathcal{M}}]) \quad (1.56)$$

This orbifold construction is straightforward as long as  $\mathbb{Z}_q$  acts smoothly on  $\mathcal{M}_q$ . However, when fixed points of the  $\mathbb{Z}_q$  action are present, singularities arise that significantly affect the computation of  $Z[\hat{\mathcal{M}}_q]$ . We now explore the implications of these fixed points in the holographic setup.

In the construction described above, one key element has yet to be examined in detail. After constructing the  $q$ -sheeted cover  $\mathcal{B}_q$ , we obtain a manifold containing a special region  $\mathcal{A}$ , whose entropy we aim to compute. The replica procedure is implemented by making a cut along  $\mathcal{A}$  and cyclically gluing the  $q$  sheets together. Specifically, the “top” side of the cut on one sheet is smoothly identified with the “bottom” side of the cut on the next sheet. This ensures continuity of field configurations across replica copies while introducing a nontrivial monodromy structure around the entangling surface. This monodromy plays a central role in the analytic continuation required to extract the entanglement entropy in the  $q \rightarrow 1$  limit.

A crucial aspect of this construction is the angular identification in the multi-sheeted geometry. If we traverse a closed path that does not intersect the cut, the total angular range covered remains  $2\pi$ , as in the original manifold. However, a path that crosses the cut transitions between different sheets of  $\mathcal{B}_q$ , such that a complete

loop around the entangling surface  $\partial\mathcal{A}$  traverses all  $q$  sheets before returning to its starting point. Consequently, the total angular range around  $\partial\mathcal{A}$  is not  $2\pi$  but  $2\pi q$ . This modified angular periodicity is a key feature of the replica geometry, as it encodes the presence of a conical singularity at the entangling surface in the  $q$ -sheeted space.

However, when constructing the quotient space  $\hat{\mathcal{B}}_q = \mathcal{B}_q/\mathbb{Z}_q$ , we identify points related by the  $\mathbb{Z}_q$  symmetry, effectively reducing the  $q$ -sheeted geometry to a single-sheeted one. This identification modifies the local structure near the entangling surface  $\partial\mathcal{A}$ . In the original  $q$ -sheeted manifold  $\mathcal{B}_q$ , a path encircling  $\partial\mathcal{A}$  completes a full  $2\pi q$  rotation before returning to its starting point. However, in the quotient space  $\hat{\mathcal{B}}_q$ , the periodicity is reduced: a full loop around  $\partial\mathcal{A}$  now covers only an angular range of  $2\pi/q$ . This implies that the quotient space is locally equivalent to a manifold with an angular deficit

$$\Delta\theta = 2\pi - \frac{2\pi}{q} = 2\pi \left(1 - \frac{1}{q}\right). \quad (1.57)$$

Thus,  $\hat{\mathcal{B}}_q$  exhibits a *conical singularity* at  $\partial\mathcal{A}$ , where the local curvature is concentrated. The presence of such a singularity is a crucial feature of the replica geometry, as it enables the computation of entanglement entropy through the variation of the gravitational action with respect to  $q$ . In the holographic context, this corresponds to evaluating the gravitational path integral on the backreacted bulk spacetime, which will play a key role in deriving the Ryu-Takayanagi prescription from the gravitational perspective.

In the bulk, the conical singularity at  $\partial\mathcal{A}$  manifests as a codimension-2 cosmic brane  $C_q$ , which serves as the fixed-point locus of the  $\mathbb{Z}_q$  quotient. This brane backreacts on the bulk geometry, modifying the local curvature in a manner consistent with the angular deficit introduced by the replica construction. Our goal is to compute the partition function  $Z[\hat{\mathcal{M}}_q]$  in the presence of this conical singularity. Since directly evaluating the full gravitational path integral is intractable, we instead approximate it by determining the dominant saddle-point geometry and computing its on-shell action  $I[\hat{\mathcal{M}}_q]$ . This allows us to extract entanglement entropy holographically through

variations of the action with respect to  $q$ , as we will now explore.

Near the locus  $C_q$ , we introduce normal polar coordinates  $(x, t)$  so that the local geometry can be expressed as<sup>11</sup>:

$$ds^2 = dx^2 + dt^2 + (\gamma_{ij} + 2K_{ij}^x x + 2K_{ij}^t t) dy^i dy^j + \dots \quad (1.58)$$

where  $\gamma_{ij}$  represents the induced metric on the brane,  $K_{ij}^a$  are extrinsic curvature components, and the dots indicate higher-order corrections. The presence of the conical singularity at  $C_q$  introduces curvature terms that scale as  $(q-1)\frac{K^a}{r}$ , where  $r = \sqrt{t^2 + x^2}$  is a radial coordinate. Since the original manifold  $\mathcal{M}_q$  is smooth before the quotient, these divergences must cancel. However, higher-order terms alone cannot accomplish this, necessitating the condition  $K^a = 0$ .

This constraint implies that in the limit  $q \rightarrow 1$ , the cosmic brane  $C_q$  coincides with a minimal-area surface  $\gamma_{\mathcal{A}}$  homologous to  $\mathcal{A}$ :

$$\lim_{q \rightarrow 1} C_q = \gamma_{\mathcal{A}}. \quad (1.59)$$

This result establishes a direct connection between the replica trick in gravity and the Ryu-Takayanagi prescription for holographic entanglement entropy. Specifically, the entanglement entropy of  $\mathcal{A}$  is governed by the area of the extremal surface  $\gamma_{\mathcal{A}}$  in the bulk, a principle that will be rigorously derived in the next steps.

To compute the entropy, we need to evaluate the gravitational action  $I[\hat{\mathcal{M}}_q]$  in the presence of the cosmic brane  $C_q$ . As noted by Dong Dong (2016), a more convenient approach is to compute the variation  $\partial_q I[\hat{\mathcal{M}}_q]$  rather than evaluating the full action explicitly. This approach leverages the fact that  $q$  controls the strength of the conical singularity, allowing us to extract the entropy via differentiation.

Conceptually,  $\partial_q$  can be understood as varying the bulk solution. In classical gravity, variations of the action take the standard form:

$$\delta I[\hat{\mathcal{M}}_q] = \text{EOM} + \text{Boundary terms} \quad (1.60)$$

---

<sup>11</sup>Here we follow Rangamani and Takayanagi (2017b)

where the first term enforces the bulk equations of motion (EOM), while the second term accounts for contributions from the boundaries of  $\hat{\mathcal{M}}_q$ . These include both the asymptotic AdS boundary  $\mathcal{B}_q$  and potential internal boundaries, which arise due to the presence of the cosmic brane  $C_q$ . Since the equations of motion vanish on-shell, only the boundary terms contribute to the entropy calculation, reducing the problem to an evaluation of the gravitational action at the conical singularity.

In our case, regularizing the cosmic brane  $C_q$  involves excising a small tubular neighborhood around the singular locus. This introduces a codimension-1 surface, denoted  $C_q(\epsilon)$ , which acts as a boundary of the regularised spacetime. To ensure a well-posed variational problem, it is necessary to include a Gibbons-Hawking-York (GHY) boundary term to account for contributions from this surface:

$$I_{\text{bdry}}[\hat{\mathcal{M}}_q] = \frac{1}{8\pi G_N^{d+1}} \int_{C_q(\epsilon)} K_\epsilon \quad (1.61)$$

where  $K_\epsilon$  is the trace of the extrinsic curvature associated with  $C_q(\epsilon)$ . Since the variation of the gravitational action primarily arises from boundary contributions, differentiating the action with respect to  $q$  yields:

$$\partial_q I[\hat{\mathcal{M}}_q] = -\partial_q I_{\text{bdry}}[\hat{\mathcal{M}}_q] \quad (1.62)$$

To proceed, we introduce local coordinates adapted to the neighborhood of  $C_q$ , in which the extrinsic curvature takes the form

$$K_\epsilon = \frac{1}{q\epsilon}. \quad (1.63)$$

Thus, differentiating the gravitational action with respect to  $q$  gives:

$$\partial_q I[\hat{\mathcal{M}}_q] = \frac{1}{8\pi G_N^{d+1}} \int_0^{2\pi q} d\tau \int \frac{1}{q^2 \epsilon} \epsilon dA_{C_q} = \frac{\text{Area}(C_q)}{4q^2 G_N^{d+1}}. \quad (1.64)$$

Using equation (1.56), we define the *modular entropy* as:

$$\partial_q \left( \frac{q-1}{q} S_q \right) = \partial_q I[\hat{\mathcal{M}}_q] \equiv S_{\text{mod}} \quad (1.65)$$

Taking the limit  $q \rightarrow 1$ <sup>12</sup>, we recover the Ryu-Takayanagi formula:

$$S_{\mathcal{A}} = \frac{\text{Area}(\gamma_{\mathcal{A}})}{4G_N^{d+1}}. \quad (1.68)$$

As mentioned above, the Ryu-Takayanagi formula applies under specific conditions. In particular, we have thus far assumed a static setup with trivial time evolution. A natural question is how this prescription generalises to dynamical settings where the entangling region evolves in time. We previously noted that the appropriate replica construction in quantum field theory requires the Schwinger-Keldysh formalism to properly account for time evolution. This raises the question of how the holographic dictionary should be extended in such cases. This problem was addressed by Hubeny, Rangamani, and Takayanagi (HRT) Hubeny et al. (2007a), who proposed a natural generalisation of the RT prescription for computing entanglement entropy in time-dependent states.

The derivation follows an analogous procedure to the static case, but with a crucial modification. Previously, the vanishing of the extrinsic curvature ensured that the surface  $\gamma_{\mathcal{A}}$  was minimal. In the time-dependent setting, however, this condition now selects a surface that is *extremal*, denoted  $\varepsilon_{\mathcal{A}}$ . The entanglement entropy formula remains unchanged, but the nature of the bulk surface is different. While  $\varepsilon_{\mathcal{A}}$  still satisfies the homology constraint, it is not necessarily confined to a single time slice—it can extend fully into the Lorentzian bulk. Moreover,  $\varepsilon_{\mathcal{A}}$  is characterised

---

<sup>12</sup>It is important to note that the modular entropy and the Rényi entropies converge to the von Neumann entropy in this limit. This is evident from expanding  $\rho^q = e^{q \log \rho} \approx \rho + (q-1)\rho \log \rho$  near  $q = 1$ , leading to:

$$\lim_{q \rightarrow 1} S^q = \lim_{q \rightarrow 1} \frac{1}{1-q} \log \text{Tr} [\rho + (q-1)\rho \log \rho] = -\text{Tr} [\rho \log \rho] = S_{\text{vN}}, \quad (1.66)$$

where we have used the density matrix normalization  $\text{Tr} \rho = 1$ . Similarly, for the modular entropy:

$$\lim_{q \rightarrow 1} S_{\text{mod}}^q = \lim_{q \rightarrow 1} \frac{q}{q-1} \log \text{Tr} [\rho^q \log \rho] = -\text{Tr} [\rho \log \rho] = S_{\text{vN}}. \quad (1.67)$$

by extremizing the area functional,  $\delta \text{Area}(\varepsilon_{\mathcal{A}}) = 0$ , and among all extremal surfaces satisfying this condition, we must select the one with minimal area.

#### 1.1.5.4 Quantum Corrections and Quantum Extremal Surfaces (QES)

The next refinement to the holographic prescription involves incorporating quantum effects. Thus far, our derivation has relied on a purely classical saddle point solution, capturing only the leading-order contributions from bulk geometry. To extend this framework to include the effects of quantum fields, Faulkner, Lewkowycz, and Maldacena Faulkner et al. (2013a) proposed a prescription that generalises the Ryu-Takayanagi formula by incorporating quantum corrections. Engelhardt and Wall Engelhardt and Wall (2015a) further developed this framework providing a refined prescription for computing holographic entanglement entropy beyond the classical limit which we see below.

The key idea follows from a well-established result in black hole thermodynamics. As first recognised by Bekenstein in 1973 Bekenstein (1973a), the entropy of a black hole should include not only the geometric contribution from the event horizon but also the entropy of the quantum fields outside it. This led to the definition of the *generalised entropy*:

$$S_{gen} = \frac{\text{Area of horizon}}{4\hbar G_N} + S_{outside} \quad (1.69)$$

where  $S_{outside}$  accounts for the von Neumann entropy of matter and gravitons outside the black hole.

Just as the Ryu-Takayanagi formula was inspired by early black hole entropy formulas, the inclusion of quantum corrections follows naturally from the concept of generalised entropy. The prescription for computing holographic entanglement entropy beyond the classical limit involves minimising the generalised entropy functional:

$$S \sim \min \left[ \frac{\text{Area}(\gamma)}{4G_N} + S_{out} \right] \quad (1.70)$$

where  $S_{out}$  captures the effects of quantum fields on either side of the extremal surface.

The precise formulation of this prescription is more subtle due to the fact that the surface in question is a codimension-2 object. This implies that it is localised in both a spatial direction and time. In the classical case, without quantum corrections, the prescription follows the HRT formulation described earlier, where the extremal surface is found by extremising the area functional. When quantum effects are included, this generalises to an extremisation over both spatial and temporal directions, requiring a careful analysis of how the entanglement entropy varies in a fully dynamical setting. As before, if multiple candidate surfaces satisfy the extremisation condition, the prescription is to select the one yielding the global minimum.

Since the fine-grained entropy remains constant over time, any covariant surface used to compute it should not introduce spurious time dependence. This follows from the fact that we are measuring the von Neumann entropy, which is a fundamental, time-invariant quantity in a closed quantum system. One natural way to select the appropriate codimension-2 surface is to first find the minimal-area surface within a fixed Cauchy slice and then extremise over all possible choices of Cauchy slices, selecting the one that maximises the result. This procedure is often referred to as the *min-max* construction and provides a systematic framework that can be readily generalised to include quantum corrections.

The final prescription for computing the quantum-corrected holographic entropy takes the form:

$$S = \min_{\varepsilon} \left\{ \text{ext}_{\varepsilon} \left[ \frac{\text{Area}(\varepsilon)}{4G_N} + S_{\text{semi-cl}}(\varepsilon) \right] \right\} \quad (1.71)$$

where now we use  $S_{\text{semi-cl}}$  to label the entropy contribution from quantum fields evaluated on the surface  $\varepsilon$ . The expression in brackets defines the *generalised entropy*, denoted by  $S_{\text{gen}}$ .

Whenever a black hole is present, the surface  $\varepsilon$  is not restricted to lie entirely outside the horizon; rather, it can extend into the black hole interior. This flexibility allows it to probe different interior regions that may share a common exterior. As



a result, the prescription correctly captures the fine-grained entropy of a black hole, incorporating both the contributions from its exterior and quantum effects within the interior.<sup>13</sup> This refined extremal surface, which generalises the classical HRT prescription to include quantum corrections, is referred to as a *quantum extremal surface* (*QES*) in the literature.<sup>14</sup>

#### 1.1.5.5 The Page Curve and the Black Hole Information Paradox

We have now systematically constructed the fine-grained entropy in a quantum field theory and explored how the AdS/CFT correspondence provides a powerful framework for computing this quantity in the gravitational dual. The introduction of quantum extremal surfaces extends the holographic dictionary beyond the classical regime, incorporating quantum effects in a controlled manner. With these tools in hand, we are now equipped to revisit the black hole evaporation process introduced earlier and assess whether the fine-grained entropy follows the expected unitary evolution.

In particular, we aim to examine whether the entropy of the radiation, computed using the quantum extremal surface prescription, exhibits the characteristic behavior known as the Page curve. If black hole evaporation is unitary, the entropy of the radiation should initially grow as entangled Hawking pairs are produced, but at late times, it must decrease as information is recovered, ultimately returning to zero when the black hole fully evaporates. This provides a crucial test for any proposed resolution to the black hole information paradox.

If we accept the conclusions drawn from string theory, it should be evident that

---

<sup>13</sup>Here we adopt the *central dogma* perspective on black hole interiors, which asserts that black holes should be treated analogously to ordinary quantum systems, such as a lump of coal. This viewpoint assumes that the entropy of a black hole is fully accounted for by its exterior degrees of freedom, much like any other thermal object. However, the appropriate way to think about black hole interiors remains an active and evolving area of research.

<sup>14</sup>This is perhaps one of the rare instances in physics where a naming convention is both intuitive and precisely descriptive: a QES is simply an extremal surface modified by quantum effects.

the evaporation process must be unitary. In terms of the entropy considerations we have been discussing, this implies that the initial state, being pure, has zero entropy. Likewise, if the final state after complete evaporation is also pure, the entanglement entropy of the remaining Hawking radiation must return to zero. However, the behavior of the entropy in the intermediate stages remains unclear. One might argue that in a gravitational system, the inability to factorise the Hilbert space could lead to an entropy evolution that remains constant throughout the process.<sup>15</sup>

Don Page (1993a) argued that if an observer collects the radiation emitted by the black hole, they are effectively “seeing” only one twin of each entangled pair produced across the horizon, as described earlier. This implies that the entanglement entropy of the radiation initially increases as the black hole radiates. While it is theoretically possible for the black hole to emit both twins of a pair at some point—thus purifying an earlier radiation mode—such an event is statistically unlikely.<sup>16</sup> Instead, a more reasonable assumption is that the black hole preferentially emits one twin of each entangled pair first. At a critical moment, known as the *Page time*, the black hole transitions to predominantly emitting the second twins of these pairs. This second phase of radiation purifies the previously emitted first twins, causing the entanglement entropy to decrease.

Thus, the entanglement entropy follows a characteristic evolution: an initial rise as first twins are emitted, reaching a peak at the Page time, followed by a decrease as the second twins are emitted and purification occurs. This behavior is known as the *Page curve* and serves as a crucial benchmark for any framework seeking to explain the fine-grained entropy of evaporating black holes. Consequently, a primary goal of computing the entanglement entropy of Hawking radiation is to recover the Page curve, thereby providing a concrete test of the unitarity of the evaporation process.

---

<sup>15</sup>Several proposed resolutions following this line of reasoning are comprehensively reviewed in Raju (2021).

<sup>16</sup>Here we adopt the terminology of “twins” to refer to each of the two particles in an entangled pair.

Now, in order to perform specific computations, we need a concrete model for an evaporating black hole. Nearly two decades ago, there was considerable interest in the string theory community in understanding how our four-dimensional universe could emerge from the ten-dimensional framework of superstring theory. Much of this research focused on compactifications, where the higher-dimensional theory is formulated on a product space of the form  $\mathbb{R}^{1,3} \times C$ , with  $C$  representing a compact six-dimensional internal manifold. By considering scenarios in which the compact dimensions shrink to scales below observational reach, one recovers an effective four-dimensional spacetime description.

This idea, called *compactifications*, is a vast subject that one can trace back to the work of Kaluza and Klein, although it gained a more concrete form after the *Phase 3* of string theory evolution (1984-1994), where the focus was set on string phenomenology Rickles (2014). Compact dimensions in string theory are directly accessible to string excitations, unlike Standard Model particles, for which very small dimensions remain effectively invisible. However, string theory provides a richer structure, particularly through T-duality, which ensures that even small compact dimensions remain physically meaningful.

### 1.1.6 Randall-Sundrum braneworlds

While compactifications provide a mechanism for hiding extra dimensions, they are not the only possible way to reconcile higher-dimensional physics with our observed four-dimensional universe. An alternative perspective arises in the context of braneworld scenarios, where our four-dimensional universe is embedded as a hypersurface (a brane) inside a higher-dimensional bulk. Unlike traditional compactifications, in which all fields propagate in the full higher-dimensional space, braneworld models allow certain degrees of freedom—such as Standard Model particles—to be confined to the brane while gravity extends into the bulk. This approach has deep implications for holography and black hole physics, as we will now explore.

Randall and Sundrum Randall and Sundrum (1999a,b) proposed a particularly influential braneworld model in which our universe is a *codimension-one brane* embedded in a five-dimensional bulk spacetime. Instead of compactifying the extra dimension in the usual sense, they showed that a warped geometry could localise gravity near the brane, effectively recovering four-dimensional gravity at long distances. This discovery provided an elegant resolution to the hierarchy problem and opened new avenues for understanding holography, black hole physics, and cosmology from a higher-dimensional perspective.

The key idea behind the Randall-Sundrum (RS) braneworld is to introduce a thin brane of matter into an  $AdS$  background governed by Einstein-Hilbert gravity in  $d+1$  dimensions. The action for this system consists of two terms: the standard bulk Einstein-Hilbert term with a negative cosmological constant, and an additional term localised on the brane that accounts for its intrinsic tension. This takes the form

$$S = \frac{1}{16\pi G} \int d^{d+1}x (R - 2\Lambda) + T \int d^d x \sqrt{\tilde{g}} \quad (1.72)$$

where  $\tilde{g}$  is the metric induced on the brane, and  $T$  represents the brane tension. The bulk cosmological constant is given by  $\Lambda = -\frac{d(d-1)}{2L^2}$ . For the present discussion, we focus on the case  $d = 5$ , meaning that the four-dimensional brane is embedded in a five-dimensional bulk. This construction forms the basis of the *Randall-Sundrum (RS) braneworld*.

The brane acts as a localised source, much like a delta function potential in quantum mechanics. Since the gravitational equations remain unchanged in the bulk, the solution away from the brane remains a pure  $AdS_5$  background. However, the presence of the brane introduces a discontinuity in the extrinsic curvature, which must be accounted for by the appropriate matching conditions. These are given by the Israel junction conditions:

$$K_{\mu\nu} - \tilde{g}_\mu K = 8\pi G T_{\mu\nu} = -4\pi G T \tilde{g}_{\mu\nu} \quad (1.73)$$

where  $K_{\mu\nu}$  is the extrinsic curvature of the brane and  $T_{\mu\nu}$  represents the stress-energy tensor localised on the brane. The second equality follows from assuming the brane tension  $T$  is the only source of energy-momentum.

To ensure a well-posed variational principle, we implicitly include a Gibbons-Hawking-York (GHY) boundary term, which accounts for the contribution of the brane as a boundary to the bulk spacetime. Physically, the junction conditions encapsulate how the embedding of the brane modifies the surrounding geometry, dictating how the gravitational field responds to its presence.

This setup provides a crucial departure from conventional compactification schemes: rather than simply wrapping extra dimensions into a small internal space, the brane-world scenario allows for an effectively four-dimensional gravitational dynamics to emerge from a higher-dimensional bulk. The extent to which gravity is localised near the brane, and how extra-dimensional effects modify its behavior, will play a central role in understanding holographic entanglement entropy and black hole physics in braneworld scenarios.

To ensure that the worldvolume of the brane exhibits maximal symmetry, we require it to take the form of a  $(3 + 1)$ -dimensional maximally symmetric spacetime, which could be  $AdS_4$ , Minkowski, or  $dS_4$ . A natural choice of coordinates for the bulk is to express the  $AdS_5$  metric in a warped slicing:

$$ds_{AdS_5}^2 = e^{2A(r)} ds_4^2 + dr^2 \quad (1.74)$$

where  $ds_4^2$  denotes the four-dimensional metric on the brane, and the coordinate  $r$  parameterises the extra dimension. The brane is located at  $r = 0$ , serving as a transition hypersurface between the two sides of the bulk.

Applying the junction condition from equation (1.73), we obtain a condition on the derivative of the warp factor  $A(r)$  across the brane:

$$A'(\epsilon) - A'(-\epsilon) = -\frac{2T}{T_c} A(0) \quad (1.75)$$

where the critical brane tension is defined as  $T_c = \frac{3}{4\pi GL}$ .

The solutions to the junction condition depend on the value of the brane's tension. If the tension satisfies  $T < T_c$ , the induced geometry on the brane is  $AdS_4$ . This scenario will be of particular relevance in the following chapters. A Randall-Sundrum braneworld with an  $AdS_4$  brane is referred to as a *Karch-Randall (KR) braneworld*. If instead  $T = T_c$ , the brane is flat, leading to a Minkowski brane, while for  $T > T_c$ , the induced geometry on the brane is  $dS_4$ .

To analyze the behavior of gravitational fluctuations around the brane, we consider perturbations of the bulk metric. This results in a Schrödinger-like equation governing the gravitational modes DeWolfe et al. (2000); Karch and Randall (2001a). The solutions to this equation give rise to a localised Kaluza-Klein tower of modes. In the case of a Minkowski brane, the lowest-lying mode is massless and can be interpreted as the four-dimensional graviton. However, when the brane is  $AdS_4$ , this mode acquires a positive mass Miemiec (2001a), which has important consequences for the effective gravitational theory on the brane. This feature will play a significant role in the discussions of later chapters.

The Randall-Sundrum braneworld can be naturally incorporated into the AdS/CFT framework, leveraging its computational advantages. We begin with a five-dimensional bulk spacetime in which a four-dimensional brane is embedded. The brane tension is tuned such that the induced geometry on the brane is  $AdS_4$ , corresponding to a subcritical brane.

A key feature of this setup is that it allows for the application of holography at two different levels. First, we can apply the standard AdS/CFT correspondence to the five-dimensional bulk, associating it with a dual four-dimensional conformal field theory. However, since the brane itself is  $AdS_4$ , we can apply holography again at the level of the brane. In this second iteration, the usual duality does not yield a standard CFT, but rather a version with a *boundary*—a *boundary conformal field theory (BCFT)* Takayanagi (2011a).

This iterative application of holography, where the bulk AdS/CFT correspondence is combined with an additional holographic duality on the brane, is known as *double holography*. This framework will play a crucial role in the upcoming discussion, providing a powerful tool for studying quantum gravity and black hole information in the presence of branes.

### 1.1.7 Entanglement Wedge and Islands: Quantum Information in Gravity

As first pointed out by Almheiri, Mahajan, Maldacena, and Zhao (AMMZ) Almheiri et al. (2020d), this model can be used to compute the entanglement entropy of Hawking radiation from an evaporating black hole. In this setup, the black hole is localised on the brane and emits radiation into a bath CFT that is coupled to the brane. Since the radiation is collected in a BCFT, the task is to compute the entanglement entropy of a subregion  $\mathcal{A}$  within this BCFT. The first step in this computation is to apply the QES formula using the first iteration of

$$S_{\mathcal{A}} = \min_{\varepsilon \sim \mathcal{A}} \left\{ \text{ext}_{\varepsilon} \left[ \frac{\text{Area}(\varepsilon)}{4G_N} + S_{\text{QFT}}(EW(\mathcal{A})) \right] \right\} \quad (1.76)$$

Here, the second term explicitly accounts for the quantum corrections arising from fields in the bulk region enclosed by  $\varepsilon$  and  $\mathcal{A}$ , a region referred to as the *entanglement wedge*. A key feature of the generalised entropy in this formulation is that it allows for the consideration of disconnected surfaces  $\varepsilon$ . In the classical RT/HRT prescription, such surfaces are typically suboptimal, as their combined area would be larger than a single connected surface. However, once quantum corrections are included, it is possible for a disconnected extremal surface to minimise the generalised entropy, making it the dominant contribution. This key insight plays a crucial role in reproducing the Page curve.

Therefore, we should look for a surface  $\varepsilon$  in  $AdS_4$ , possibly with a disconnected component, that minimises  $S_{\text{gen}}$ . However, directly computing this quantity is highly nontrivial, as it requires tracking quantum field contributions and performing a full

QFT computation on the brane. This is where double holography provides a crucial simplification. By applying a second iteration of holography, we can recast the quantum field contribution  $S_{\text{QFT}}(EW(\varepsilon))$  as a geometric quantity: the area of a minimal surface in the 5-dimensional  $AdS$  bulk. In other words, the entropy computation can be mapped onto a classical Ryu-Takayanagi formula in the  $5d$  bulk, making the problem significantly more tractable.

In low dimensions, the computation can be performed explicitly, as demonstrated by AMMZ, who found that the entanglement entropy of radiation follows the Page curve as expected Almheiri et al. (2019b, 2020a).<sup>17</sup> To obtain this behavior, one must include a disconnected component of the extremal surface, known as an *island*, which contributes to the entropy in the manner described above, effectively purifying it. This insight naturally leads to the *island proposal*:

$$S_{\mathcal{A}} = \min_{\varepsilon \sim \mathcal{A}} \text{ext}_{\varepsilon} S_{\text{gen}}(\mathcal{A} \cup \mathcal{J}) \quad (1.77)$$

where  $\mathcal{J}$  is the disconnected component representing the island, which is dynamically determined as part of the problem's resolution.

The computation in lower dimensions was soon extended to higher dimensions by Almheiri, Mahajan, and Santos, who analyzed the entanglement entropy of a planar black hole Almheiri et al. (2020f). Around the same time, Gen and Karch identified a technical issue with these models, despite their tractability Geng and Karch (2020a). Specifically, in the subcritical brane scenario, the lowest-lying graviton mode acquires a positive mass, implying that the effective four-dimensional gravity is not truly massless. As a result, although these models successfully reproduce the expected entropy behavior, they may not be physically accurate. Another critique concerns the artificial nature of the setup: the radiation is always collected in a region of space (the bath) where gravity is turned off, which may not fully capture the dynamics of a realistic evaporating black hole.

---

<sup>17</sup>A comprehensive review by some of the key contributors can be found in Almheiri et al. (2021).



Motivated by these issues, Geng et al. extended the computation to higher dimensions while incorporating a gravitating bath Geng et al. (2021a). This was achieved using a two-brane Karch-Randall (KR) braneworld, in which the setup was modified to include two black holes connected by a black string in the bulk. Interestingly, the entropy curve in this model remained flat, with the minimal-area surface always coinciding with the horizon. Although this evolution is unitary, it does not align with the conventional expectations of the central dogma. Moreover, in this scenario, islands are always absent ( $\mathcal{I} = \emptyset$ ), meaning they are not necessary to maintain unitarity.

Notably, this geometry features a non-gravitating region, known as the defect, which is attached to both the brane and the bath. If one computes the entropy of the “left” and “right” degrees of freedom of the defect, islands and nontrivial entropy evolution do appear. However, this is distinct from the entropy of the defect when paired with the thermofield double, to which the black string model is dual.

Following the higher-dimensional generalizations, the study of entanglement entropy in these setups was extended to global coordinates in order to investigate whether entanglement entropy could capture the Hawking-Page transition Karch et al. (2023a). While this attempt was ultimately unsuccessful, the global black string model revealed an intricate phase structure for entanglement entropy, highlighting new transitions that do not appear in the planar case.

Meanwhile, in lower dimensions, it was shown that one can recover Jackiw-Teitelboim (JT) gravity as the effective description on the brane Geng et al. (2022c). However, this setup presents a puzzling degeneracy of Ryu-Takayanagi (RT) surfaces, suggesting multiple extremal surfaces contribute equally to the entropy. This issue is resolved by accounting for fluctuations of the brane’s radion field, which lift the degeneracy and select a preferred surface. These results illustrate the subtle interplay between bulk geometry, brane dynamics, and entanglement entropy in braneworld models.

Islands, by construction, exhibit an unusual property: they provide one of the first examples of an entanglement wedge that is completely disconnected from the boundary. At first, this might not seem problematic, but in a theory of massless gravity, we expect gauge-invariant bulk operators to be dressed to the boundary, typically via a gravitational Wilson line. According to subregion duality, these dressed operators should be expressible in terms of elements of the algebra  $\mathcal{A}(\varepsilon)$  associated with the entanglement wedge of their dual region. However, in the case of islands, this poses a fundamental issue: operators localised within the island cannot be dressed to the boundary without crossing the complement entanglement wedge.

One way to address this in the Karch-Randall (KR) braneworld is through the introduction of massive gravity, where the explicit breaking of diffeomorphism invariance allows for local gauge-invariant observables, circumventing the need for asymptotic dressing. More generally, in theories of massive gravity, the absence of exact diffeomorphism symmetry enables the construction of local observables that do not require dressing to infinity. This observation was the basis of the argument by Geng et al., who proposed that in massless gravitational theories, islands inherently violate Gauss’s law—i.e., the Hamiltonian constraint—due to the obstruction in dressing operators to the boundary Geng et al. (2022a).

For the moment, one can set aside the issue of the massive graviton and instead analyze the phase structure of entanglement entropy for the black string in the case of a non-gravitating bath. This corresponds to returning to a one-brane KR braneworld. In contrast to the planar black hole studied by Almheiri, Mahajan, and Santos Almheiri et al. (2020f), where entanglement entropy was computed only for a few fixed values of the theory parameters due to the complexity of the problem, the black string setup allows for a more systematic classification. Surprisingly, in this case, it is computationally feasible to determine, for all values of the theory parameters, whether the entropy remains flat or transitions to a Page curve. The key parameter governing this transition is known as the *Page angle* Geng et al. (2022b), which effectively dictates when the entanglement entropy undergoes a phase change.

One of the key advantages of the two-brane model over the one-brane model is that the superposition of modes from the left and right branes can be fine-tuned to recover a massless graviton. This mechanism, known as *bi-gravity*, has been understood for several decades Kogan et al. (2001a). Interestingly, in this context, islands do not appear because the boundary conditions on each brane are too restrictive for minimal-area surfaces to form. Around the same time, I was studying the phase structure of entanglement entropy in the one-brane setup, incorporating higher-order corrections in the form of *Dvali-Gabadadze-Porrati (DGP)* terms Dvali et al. (2000a,b) on the brane Perez-Pardavila (2023c). Miao later realised that these corrections could be exploited to bypass the restrictive boundary conditions and generate islands in the two-brane setup Miao (2024, 2023a).

However, this proposal was quickly challenged by Geng et al., who demonstrated that these island surfaces must necessarily violate a fundamental holographic principle known as causal wedge inclusion (CWI) Wall (2014a); Geng et al. (2023b). The core idea is that an island defines an entanglement wedge, while the black hole horizon also defines an entanglement wedge. According to CWI, the entanglement wedge must always contain the causal wedge, meaning that information accessible within a causal domain should also be encoded within its corresponding entanglement wedge. However, almost by construction, these island surfaces lie outside the horizon, meaning that their entanglement wedge does not fully contain the associated causal wedge, explicitly violating CWI.<sup>18</sup> This fundamental inconsistency casts doubt on the viability of islands in this particular setup.

The status of these island surfaces, at the time of writing this dissertation, remains an open question, with ongoing debate in the community about whether the issues discussed here—summarized in the preceding paragraphs—pose fundamental problems for their proponents. While the resolution of these concerns is still uncertain,

---

<sup>18</sup>The term “outside” here may be somewhat ambiguous. Since the setup involves a thermofield double, the islands are included in the causal wedge of the corresponding copy of the TFD, leading to an explicit violation of CWI.

what is clear is that many of the key ideas, from the earliest formulations of the holographic correspondence to the present day, have emerged from viewing the black hole information paradox through a geometric lens.

## 1.2 Geometry in quantum field theory

The perspective that geometry plays a crucial role in understanding quantum gravity extends far beyond the context of black holes. Indeed, throughout modern theoretical physics, geometry has emerged as a powerful organising principle, offering deep insights into the structure of quantum field theories. From the study of anomalies and dualities to the classification of supersymmetric field theories, geometric methods provide a unifying framework that links diverse physical phenomena. In particular, the interplay between quantum field theory and geometry has been especially striking in the study of supersymmetric theories, where rich mathematical structures—such as moduli spaces, integrable systems, and holographic correspondences—naturally arise. This connection is exemplified by the six-dimensional  $(2,0)$  superconformal field theories, which serve as a higher-dimensional origin for many lower-dimensional gauge theories and encode profound links between string theory, M-theory, and mathematical physics.

### 1.2.1 $(2,0)$ theories: From string theory to class S

Six-dimensional theories with  $\mathcal{N} = (2,0)$  supersymmetry play a fundamental role in modern high-energy theoretical physics Strominger (1996); Witten (1996, 1995); Seiberg and Witten (1996a); Seiberg (1998); Aharony et al. (1998); Seiberg (1997b). As the highest-dimensional supersymmetric conformal field theories (SCFTs), they serve as parent theories for a vast landscape of lower-dimensional quantum field theories, encoding deep structural insights into gauge theory, supersymmetric field theory, conformal field theory, and quantum gravity. These theories also establish remarkable mathematical correspondences, such as the Alday-Gaiotto-Tachikawa

(AGT) duality Gaiotto (2012a); Alday et al. (2010), connections to the Geometric Langlands program Beilinson and Drinfeld; Witten (2008); Gaiotto and Witten (2024), applications in knot homology, and implications for  $q$ -deformations of algebraic structures Witten (2011). The influence of  $\mathcal{N} = (2, 0)$  theories extends far beyond conventional field theory, reaching into string theory, M-theory, and topological invariants of manifolds. While a comprehensive exploration of these interconnections is beyond the scope of this dissertation, our focus will be on a specific yet central aspect of this framework: the compactification of six-dimensional  $(2, 0)$  theories to four dimensions, leading to the rich and intricate structure of *Class S* theories.

### 1.2.1.1 Six dimensional $\mathcal{N} = (2, 0)$ theories

We begin by outlining the original construction of these six-dimensional  $(2, 0)$  theories<sup>19</sup>. These theories were first identified by Witten Witten (1995) in the context of Type IIB string theory compactified on a K3 surface, a construction known as *geometric engineering* Katz et al. (1997, 1998). In this framework, the K3 surface develops singularities, corresponding to degenerations in the compactification geometry. These singularities can be resolved by introducing massless degrees of freedom—specifically, tensionless strings—that emerge from  $D3$ -branes wrapping vanishing cycles. The classification of these singularities follows the well-known ADE scheme, resulting in an interacting<sup>20</sup> six-dimensional theory that shares the same ADE classification.

The six-dimensional  $(2, 0)$  superconformal field theories exhibit maximal superconformal symmetry, possessing 16 real supercharges Nahm (1978). Their full superconformal algebra is given by  $OSp(2, 6|2)$ , where the bosonic subgroup consists of the Lorentz symmetry  $Spin(5, 1)$ , which governs six-dimensional spacetime

---

<sup>19</sup>We refer to “theories”, but strictly speaking, there is a unique  $(2, 0)$  theory, which has no free parameters. This is not inconsistent with standard notions of quantum field theory, as its lack of a Lagrangian description does not preclude its existence.

<sup>20</sup>i.e., not free, where free” refers to a theory in which all equations of motion are linear.

transformations, and the R-symmetry group  $Sp(2)_R$ , which acts on the supercharges and scalar fields. This maximal supersymmetry makes  $(2, 0)$  theories highly constrained, inherently non-Lagrangian <sup>21</sup>, and deeply intertwined with string theory and M-theory.

The supersymmetry algebra of six-dimensional  $(2, 0)$  theories takes the form Distler and Hanany (1997)

$$\{Q_\alpha^a, Q_\beta^b\} = 2\omega^{ab}\gamma^\mu\alpha\beta P_\mu + \gamma^\mu\alpha\beta Z_\mu^{ab}, \quad (1.78)$$

where  $\omega^{ab}$  is the  $Sp(2)$ -invariant tensor, ensuring the R-symmetry structure of the theory, and  $Z_\mu^{ab}$  is a central charge transforming in the  $(\mathbf{6}, \mathbf{5})$  representation of  $\text{Spin}(5, 1) \times Sp(2)_R$ .

The field content of the six-dimensional  $(2, 0)$  theory consists of a single *tensor multiplet*, whose degrees of freedom are:

- A self-dual two-form gauge field  $B_{\mu\nu}$ , whose field strength  $H = dB$  satisfies the self-duality condition  $H = *H$ . This constraint prevents a conventional covariant Lagrangian formulation, reflecting the intrinsically non-Lagrangian nature of the theory.
- Five real scalar fields  $\Phi^I$ , transforming as a  $\mathbf{5}$  of the R-symmetry group  $Sp(2)_R$ . These scalars correspond to the transverse positions of the M5-brane in the eleven-dimensional M-theory spacetime.
- Four Weyl fermions, transforming in the  $\mathbf{4}$  of  $Sp(2)_R$ , which provide the fermionic degrees of freedom required by supersymmetry.

When expressed in terms of representations under the little group  $\text{Spin}(4) \cong SU(2) \times SU(2)$  of  $\text{Spin}(5, 1)$ , these fields decompose as:

$$(1, 3; \mathbf{1}) \oplus (1, 1; \mathbf{5}) \oplus (1, 2; \mathbf{4}), \quad (1.79)$$

---

<sup>21</sup>i.e., a theory that cannot be formulated as a functional map from the space of fields to  $\mathbb{C}$  via a conventional path integral. As we see here, all known constructions arise from string theory.

where  $(1, 3; \mathbf{1})$  represents the self-dual two-form field,  $(1, 1; \mathbf{5})$  corresponds to the scalar fields, and  $(1, 2; \mathbf{4})$  describes the fermionic superpartners.

Due to the self-duality of  $B_{\mu\nu}$  and the absence of a weakly coupled regime, the  $(2, 0)$  theory does not admit a standard Lagrangian formulation. Instead, its dynamics are best understood through indirect constructions, such as its realization in M-theory via M5-branes or through its compactifications to lower dimensions. In particular, its low-energy behavior is typically probed using dualities, such as its relation to five-dimensional super Yang-Mills theory upon compactification on a circle.

From an M-theory perspective, the  $(2, 0)$  theories naturally arise as the low-energy worldvolume theories of coincident M5-branes Strominger (1996). The simplest case corresponds to the  $A_{N-1}$  series, which emerges when  $N$  M5-branes are placed in flat eleven-dimensional Minkowski space. The resulting six-dimensional superconformal field theory possesses an  $A_{N-1}$  gauge algebra but, being inherently non-Lagrangian, lacks a conventional weakly coupled description. Nevertheless, its structure can still be analyzed through its moduli space of vacua, which organizes the possible vacuum configurations and characterizes the space of allowed deformations. This perspective provides a natural way to study its strongly coupled dynamics despite the absence of a Lagrangian formulation.

A crucial component of this moduli space is the *Coulomb branch*, which is parametrised by the vacuum expectation values (VEVs) of the five scalar fields in each tensor multiplet. These scalars encode the transverse positions of the M5-branes in the ambient eleven-dimensional spacetime. When all  $N$  M5-branes coincide, the theory enters a strongly interacting phase with an enhanced gauge symmetry of type  $A_{N-1}$  Seiberg and Witten (1996a). Conversely, when the branes are separated in the transverse space, the gauge symmetry is broken to its maximal abelian subgroup, and the low-energy description consists of  $N$  independent free  $(2, 0)$  theories, each described by a single tensor multiplet.

Since the overall center-of-mass motion of the stack of  $N$  M5-branes corre-

sponds to a free, decoupled mode, the effective number of interacting degrees of freedom is reduced. The remaining interacting sector defines a theory of *rank*  $(N - 1)$ , corresponding to the dimension of the Cartan subalgebra of the  $A_{N-1}$  gauge algebra. This rank determines the number of independent massless tensor multiplets in the low-energy description and parametrises the moduli space of the interacting theory. A more detailed perspective on this structure will emerge in our discussion of Seiberg-Witten theory and its role in describing low-energy dynamics in four-dimensional compactifications.

A similar construction applies to the  $D_N$  series, which arises when  $2N$  M5-branes are placed at the fixed locus of an M-theory orientifold singularity of the form  $\mathbb{R}^5 \times \mathbb{R}^5/\mathbb{Z}_2$ . Here, the  $\mathbb{Z}_2$  action reflects the five transverse coordinates to the M5-branes.

More generally, for a  $(2, 0)$  theory associated with a simply-laced Lie algebra  $\mathfrak{g}$ , the Coulomb branch is parametrised by the vacuum expectation values of the five scalar fields in the  $\text{rank}(\mathfrak{g})$  tensor multiplets. This can be written as

$$B = (\mathbb{R}^5)^{\text{rank}(\mathfrak{g})}/W(\mathfrak{g}), \tag{1.80}$$

where  $W(\mathfrak{g})$  is the Weyl group of the  $\mathfrak{g}$ . Alternatively, these parameters can be described in terms of the VEVs of chiral primary operators associated with the Casimir invariants of  $\mathfrak{g}$ . These chiral primary operators are protected by supersymmetry, and their scaling dimensions are determined by the exponents of the corresponding Lie algebra.

For the  $A_{N-1}$  series, these chiral primary operators have scaling dimensions:  $2, 3, 4, \dots, N$ . For the  $D_N$  series, they have scaling dimensions:  $2, 4, 6, \dots, 2N - 2, N$ , with the last operator known as the Pfaffian.

Despite these well-established constructions for the  $A$ - and  $D$ -series theories, the exceptional  $(2, 0)$  theories of type  $E_6, E_7$ , and  $E_8$  remain more mysterious. No explicit M5-brane construction in M-theory is known to realise these theories directly.



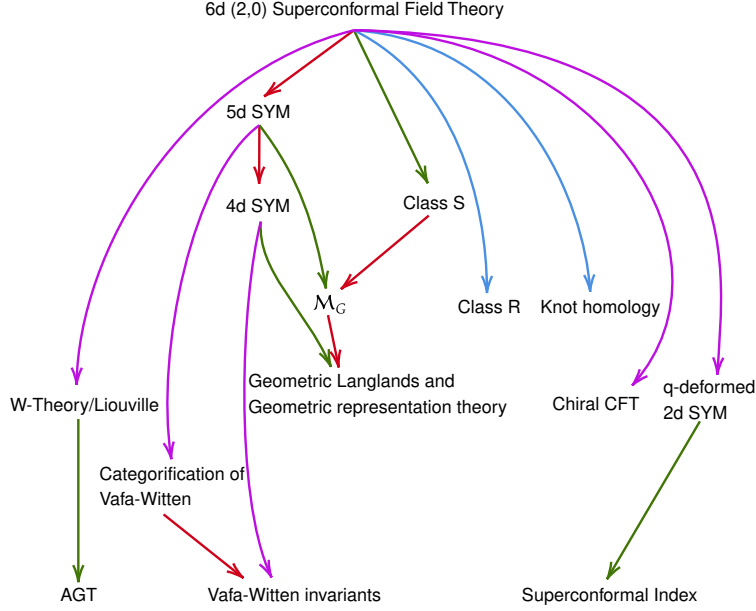


Figure 1.1: A diagram illustrating the connections between different theories, tracing their descent from the six-dimensional  $(2,0)$  theory down to zero-dimensional invariants. The elements within each row share the same spacetime dimension. Arrows indicate the compactifications and corresponding low-energy limits that relate these theories. The colour coding distinguishes different compactification manifolds: red for  $S^1$ , green for a two-dimensional surface  $C$ , blue for a three-dimensional manifold, and purple for a four-dimensional manifold. Inspired by a similar figure in Moore (2012).

Instead, their existence is inferred through indirect arguments, including the classification of six-dimensional superconformal algebras, anomaly cancellation conditions, and dualities in string theory Seiberg (1997a).

### 1.2.1.2 Class-S Theories

Having established the foundations of the six-dimensional  $\mathcal{N} = (2,0)$  theories, we now turn to their compactifications and the resulting four-dimensional physics. A particularly rich class of 4D  $\mathcal{N} = 2$  superconformal field theories arises from compactifying these 6D theories on a two-dimensional Riemann surface  $C$ , leading to what are

known as *Class S theories* Gaiotto et al. (2013a); Gaiotto (2012a). This construction provides a vast landscape of interacting SCFTs that play a central role in modern theoretical physics, deeply intertwining supersymmetric gauge theory, integrable systems, geometric representation theory, and string theory dualities.

The framework of Class S theories unifies and extends our understanding of  $4d$   $\mathcal{N} = 2$  physics. It provides a natural origin for many well-known gauge theories, offers a geometric interpretation of dualities, and connects directly to mathematical structures. Furthermore, this approach reveals how  $6d$  physics encodes information about  $4d$  gauge dynamics, bridging different areas of high-energy theory and mathematical physics.

Our goal in this section is to systematically explore the compactification process, the role of topological twisting, the emergence of  $4d$  gauge theories, and the classification of Class S theories through punctured Riemann surfaces and codimension-2 defects. We will begin by reviewing the essential mechanisms that allow us to preserve supersymmetry upon compactification.

In order to obtain a  $4d$  theory, we consider compactifying the  $6d$   $\mathcal{N} = (2, 0)$  theory on a two-dimensional manifold. A crucial aspect of this procedure is ensuring that some amount of supersymmetry is preserved in the resulting four-dimensional theory. If all supersymmetry were broken, the analysis of these theories would become significantly more complicated due to uncontrolled quantum corrections. Conversely, if we compactified in a way that preserved more than  $\mathcal{N} = 2$  supersymmetry, the theory would be overly constrained, often leading to trivial or well-understood results.

Thus, to obtain a nontrivial, yet sufficiently controlled, four-dimensional theory, we aim to preserve exactly  $\mathcal{N} = 2$  supersymmetry in the compactification process. This choice not only ensures a rich class of interacting theories but also allows for a direct connection to many known results in  $\mathcal{N} = 2$  supersymmetric gauge theory, such as Seiberg-Witten theory. The key to achieving this controlled compactification is the use of a *topological twist*, which we now discuss in detail.

Compactifying on a general two-dimensional manifold does not necessarily preserve the desired amount of supersymmetry. To retain  $\mathcal{N} = 2$  supersymmetry in four dimensions, we implement a *topological twist*. This procedure modifies the embedding of the holonomy group of the compactification manifold within the symmetry group of the original six-dimensional theory. In particular, it alters the way the Lorentz and R-symmetries act on the fields, ensuring that some supercharges remain globally well-defined after compactification.

The  $6d$   $(2, 0)$  theory has a bosonic symmetry group given by  $\text{Spin}(5, 1) \times Sp(2)_R$ . Upon compactification on a two-dimensional Riemann surface  $C$ , the symmetry is broken as follows:

$$\text{Spin}(5, 1) \times Sp(2)_R \rightarrow \text{Spin}(3, 1) \times U(1)_C \times SU(2) \times U(1)_R \quad (1.81)$$

Here,  $\text{Spin}(3, 1)$  is the four-dimensional Lorentz group,  $U(1)_C$  is the rotation symmetry of the Riemann surface, and  $SU(2) \times U(1)_R$  is a subgroup of the original  $Sp(2)_R$  R-symmetry.

To ensure that four-dimensional  $\mathcal{N} = 2$  supersymmetry is preserved, we introduce a partial twist, which redefines how the Lorentz group acts on the supercharges. Specifically, we modify the spin connection by shifting the charges under  $U(1)_C$  by an appropriate combination of the R-symmetry. This introduces a new  $U(1)$  symmetry:

$$U(1)_0 \subset U(1)_C \times U(1)_R. \quad (1.82)$$

This construction guarantees that some of the original six-dimensional supercharges transform as scalars under  $U(1)_0$ , allowing them to survive the compactification and yielding a well-defined  $\mathcal{N} = 2$  supersymmetric theory in four dimensions.

As a result, after compactification, the surviving supercharges transform under

$$\text{Spin}(3, 1) \times SU(2) \times U(1)_0, \quad (1.83)$$

which matches the structure of four-dimensional  $\mathcal{N} = 2$  supersymmetry.<sup>22</sup> This guarantees that exactly eight real supercharges survive in four dimensions, forming an  $\mathcal{N} = 2$  supersymmetric theory.<sup>23</sup>

More specifically, before performing the twist, but after compactifying on the Riemann surface, the supercharges transform under the symmetry group  $\text{Spin}(3, 1) \times U(1)_C \times SU(2) \times U(1)_R$  as

$$\left( ((2, 1)_{1/2} \oplus (1, 2)_{-1/2}) \otimes (2_{1/2} \oplus 2_{-1/2}) \right). \quad (1.84)$$

After implementing the twist, the relevant subgroup of the symmetry group changes, and the supercharges now transform under  $\text{Spin}(3, 1) \times SU(2) \times U(1)_0$  as

$$(2, 1; 2)_1 \oplus (2, 1; 2)_0 \oplus (1, 2; 2)_0 \oplus (1, 2; 2)_{-1}. \quad (1.85)$$

The two middle summands, which are uncharged under  $U(1)_0$ , correspond to the eight real supercharges that survive in four dimensions, forming the  $\mathcal{N} = 2$  supersymmetry algebra of the resulting  $4d$  theory.

We should mention an important exception to the above argument. If the Riemann surface is a torus, then, due to its flat geometry, it has trivial holonomy. As a result, no twisting is necessary, and the original supercharges remain unbroken. This leads to a four-dimensional theory with enhanced supersymmetry, preserving the full  $\mathcal{N} = 4$  structure. Consequently, compactifying the  $(2, 0)$  theory on a torus produces a  $\mathcal{N} = 4$  superconformal field theory in four dimensions.

The story so far produces a rather small class of intrinsically interacting four-dimensional  $\mathcal{N} = 2$  theories, known as the  $T_N$  theories. A key ingredient in Gaiotto's construction that significantly expands the landscape of Class S theories is the inclusion of *codimension-2 defects* Gukov and Witten (2006a, 2010); Chacaltana et al. (2013a). These defects extend along the four-dimensional spacetime and are localised

---

<sup>22</sup>Namely c.f. with  $\text{Spin}(3, 1) \times SU(2)_R \times U(1)_R$

<sup>23</sup>More explicitly, the preserved symmetry group is  $\text{Spin}(3, 1) \times SU(2)_R \times U(1)_R$ .

at specific points on the Riemann surface. These distinguished points, where the defects are placed, are referred to as *punctures* on  $C$ . The presence of punctures modifies the structure of the resulting  $4d$  theory, introducing additional global symmetries and allowing for a wider range of gauge theories to emerge from the compactification.

The defects introduced in the compactification process are *half-BPS*, meaning they preserve half of the original  $6d$  supersymmetry that remains unbroken after the twist Wess and Bagger (1992). While these defects do not break additional supersymmetry, they do induce a controlled breaking of the gauge symmetry. Let  $\mathfrak{g}$  denote the original ADE-type Lie algebra associated with the  $6d$   $(2, 0)$  theory. The symmetry breaking pattern induced by the defects can be understood in terms of the residual, unbroken global symmetry. This preserved symmetry is classified by a Levi subalgebra  $\mathfrak{l} \subset \mathfrak{g}$ , which encodes how the defect modifies the gauge symmetry locally.

The classification of Levi subalgebras is closely related to the structure of nilpotent orbits in the Lie algebra  $\mathfrak{g}$ . The seminal work of Bala and Carter Bala and Carter (1976a,b) established a correspondence between Levi subalgebras and nilpotent orbits, meaning that specifying a nilpotent orbit in  $\mathfrak{g}$  uniquely determines a Levi subalgebra and vice versa Collingwood and McGovern (1993). As a result, in the Class S construction, each puncture—where a defect is located—is labeled by a nilpotent orbit of  $\mathfrak{g}$ . This data dictates the boundary conditions in the compactification and determines key aspects of the resulting  $4d$  theory, including its flavor symmetry and moduli space.

Thus, the essential ingredients for defining a Class S theory are:

- A choice of a  $6d$   $(2, 0)$  theory of type  $G$ , where  $G$  is an ADE-type Lie group, namely the associated Lie algebra is  $\mathfrak{g} \in A_N, D_N, E_6, E_7, E_8$ .
- A choice of a genus- $g$  Riemann surface  $C_{g,n}$  with  $n$  punctures, serving as the compactification manifold.

- A decoration of each puncture by a nilpotent orbit of the Lie algebra  $\mathfrak{g}$ , which determines the local symmetry breaking at the defect.

### 1.2.2 Seiberg-Witten theory

The study of supersymmetric field theories is largely inspired by the appearance of string theory. We have laid out several instances of specific examples which are of great interest to understand the universe around us. In the mid-90s due to the efforts of Intriligator, Seiberg and Witten a lot of progress was made in understanding  $\mathcal{N} = 1$  and  $\mathcal{N} = 2$  theories Intriligator (1994); Intriligator and Seiberg (1995); Seiberg and Witten (1994a,b). Here we want to discuss in a very brief and compact way (which probably doesn't do justice to the great insights developed in these years) some of these advancements <sup>24</sup>.

#### 1.2.2.1 Four dimensional $\mathcal{N} = 2$ supersymmetric field theories

We begin by considering a generic  $\mathcal{N} = 2$  supersymmetric field theory. The most general  $\mathcal{N} = 2$  Lagrangian is formulated in terms of two fundamental building blocks: the  $\mathcal{N} = 2$  vector multiplet and the  $\mathcal{N} = 2$  hypermultiplet. Their field content is given by:

$$\mathcal{N} = 2 \text{ Vector Multiplet : } \begin{cases} \phi^A \\ \lambda_\alpha^A, \tilde{\lambda}_{\dot{\alpha}}^A \\ A_\mu \end{cases} \quad (1.86)$$

$$\mathcal{N} = 2 \text{ Hypermultiplet : } \begin{cases} \psi_\alpha^i \\ q^i, \tilde{q}^i \\ \tilde{\psi}_i^{\dot{\alpha}} \end{cases} \quad (1.87)$$

where  $\alpha$  and  $\dot{\alpha}$  are the spinor indices, and  $A$  labels the supersymmetry generators. The rows in each multiplet correspond to representations of the  $\mathfrak{su}(2)_R$  R-symmetry.

---

<sup>24</sup>There are many excellent reviews on this topic Tachikawa (2013); Martone (2020); Akhond et al. (2022)

Additionally, when viewed diagonally, these fields decompose into their respective  $\mathcal{N} = 1$  representations.

We can now construct the Lagrangian for an  $\mathcal{N} = 2$  gauge theory using these fields. The pure super Yang-Mills theory is formulated in terms of an  $\mathcal{N} = 1$  vector multiplet,  $V = (A_\mu, \lambda_\alpha)$ , and an  $\mathcal{N} = 1$  chiral multiplet,  $\Phi = (\phi, \tilde{\lambda}_\alpha)$ , where we have suppressed some indices for clarity. By introducing interactions between these components, we obtain the  $\mathcal{N} = 2$  vector multiplet Lagrangian:

$$\mathcal{L}_{\text{vector}} = -\text{Tr} \int d^2\vartheta \frac{i\tau}{8\pi} W^\alpha W_\alpha + \frac{\text{Im}\tau}{4\pi} \int d^4\vartheta \Phi^\dagger e^{2V} \Phi \quad (1.88)$$

Here,  $W_\alpha$  is the superfield strength of  $V$ , given by

$$W_\alpha = -\frac{1}{8} \overline{\mathcal{D}}^2 (e^{-2V} \mathcal{D}_\alpha e^{2V}), \quad (1.89)$$

where  $\mathcal{D}_\alpha$  denotes the superspace covariant derivatives. Additionally, we introduce the complexified gauge coupling:

$$\tau = \frac{\theta}{2\pi} + \frac{4\pi i}{g^2} \quad (1.90)$$

We can introduce matter fields by coupling the theory to an  $\mathcal{N} = 2$  hypermultiplet. This multiplet decomposes into two  $\mathcal{N} = 1$  chiral multiplets:  $Q^i = (\psi_\alpha^i, \tilde{q}^i)$  and  $\tilde{Q}^i = (\tilde{\psi}_\alpha^i, q^i)$ .<sup>25</sup> The interaction terms are given by:

$$\mathcal{L}_{\text{hyper}} = \int d^4\vartheta (Q_i^\dagger e^{2V} Q^i + \tilde{Q}_i^\dagger e^{-2V} \tilde{Q}^i) + \sqrt{2} \int d^2\vartheta (\tilde{Q}_i \Phi Q^i + h.c.) + \int d^2\vartheta \sqrt{2m} \tilde{Q}_i Q^i \quad (1.91)$$

Here, the first term represents the coupling of the matter fields to the vector multiplet, the second term describes the Yukawa coupling to the chiral multiplet  $\Phi$ , and the third term is a mass term.

Combining these ingredients, we obtain the most general form of an  $\mathcal{N} = 2$  Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{vector}} + \mathcal{L}_{\text{hyper}} \quad (1.92)$$

---

<sup>25</sup>This notation for the chiral multiplets is slightly unconventional, as it coincides with that used for supercharges. However, since it is standard in many references, we adopt it here as well.

### 1.2.3 Low energies

We now turn to an important aspect of these theories: their low-energy dynamics. The classical vacuum is determined by minimizing the potential, which takes the form

$$P = \frac{1}{2}D^2 + F^{i\dagger}F_i \quad (1.93)$$

where the D-term and F-term contributions are given by  $D^A = \phi_i^\dagger (T^A)^i_j \phi^j$  and  $F_i = \partial W / \partial \phi_i$ . Here,  $D^A$  is written in terms of the Lie algebra  $\mathfrak{g}$  generators  $T^A$ , and  $F_i$  is obtained from the superpotential  $W$ . The vacuum structure is determined by solving  $D^A = 0$  and  $F_i = 0$ , which we now analyze.

From this we can find the field configuration which constrains the allowed vacuum  $P = 0$ . We first consider the case where  $m = 0$ . It turns out there are two equally good choices for this configuration. Either we have  $\phi = 0$  with hypermultiplets  $q, \tilde{q}$  turned on, or  $\tilde{q} = q = 0$  with  $\phi$  turned on, but satisfying  $[\phi, \phi^\dagger] = 0$ . Geometrically this means that the moduli space has three components or *branches*:

- *Coulomb branch*:  $q = \tilde{q} = 0, \phi \neq 0$ . Here the field  $\phi$  is constrained by  $[\phi, \phi^\dagger] = 0$ . From an algebraic perspective, it means that  $\phi$  must lie in the Cartan subalgebra of  $\mathfrak{g}$ . When  $G = SU(N)$ , this means that  $\phi = \text{diag}(a_1, a_2, \dots, a_N)$ , where  $\text{Tr}\phi = 0$ . At a generic point of the Coulomb branch, these scalar VEVs break the  $G$ -symmetry to the Cartan subalgebra. For  $G = SU(N)$ , this is broken to  $U(1)^{N-1}$ . The only exception occurs at the origin of the Coulomb branch, where  $a_i = 0 \quad \forall i$ . Here the full  $G$ -symmetry is restored.
- *Higgs branch*:  $q = \tilde{q} \neq 0, \phi = 0$ . The low-energy theory is that of an  $\mathcal{N} = 2$  theory with free decoupled hypermultiplets. At generic points of the Higgs branch the gauge group is fully Higgsed. Since we are mainly interested in the Coulomb branch we will not discuss this branch in more detail here.
- *Mixed branch*: There are points where the two branches above intersect and these define another branch, called the mixed branch.



Since we will focus on the behaviour of the theory at generic points on the Coulomb branch, let us examine its structure in more detail. As mentioned earlier, at a generic point on the Coulomb branch, the gauge symmetry is broken to the Cartan subalgebra. The low-energy effective field theory is then described by an abelian  $U(1)^{N-1}$  theory, with the Lagrangian:

$$\mathcal{L}_{\text{eff}} = \frac{1}{8\pi} \int d^4\vartheta K(\bar{a}^i, a^j) - \frac{i}{8\pi} \int d^2\vartheta \tau^{ij}(a) W_i^\alpha W_{\alpha,j} + h.c. \quad (1.94)$$

where the relevant quantities are defined as follows<sup>26</sup>:

$$\begin{aligned} \tau_{ij} &= \frac{\partial^2 \mathcal{F}}{\partial a_i \partial a_j}, \\ a_D^i &= \frac{\partial \mathcal{F}}{\partial a_i}, \\ K &= i(\bar{a}_D^i a_i - \bar{a} i a_D^i), \end{aligned} \quad (1.95)$$

where  $\mathcal{F}$  is a locally holomorphic function known as the *prepotential*, and the pair  $(a^D, a)$  are the *special coordinates* of the Coulomb branch. In this formulation, the Coulomb branch naturally takes the structure of a (rigid) *special Kähler manifold*.<sup>27</sup>

What makes this structure *special Kähler* is that the metric is not arbitrary but derived from a holomorphic function, the prepotential  $\mathcal{F}$ . Specifically, the metric is given by:

$$g_{ij} = \text{Im} \left( \frac{\partial^2 \mathcal{F}}{\partial a^i \partial a^j} \right). \quad (1.97)$$

This constraint, which ensures that the entire geometry is encoded in a single holomorphic function, distinguishes special Kähler manifolds from general Kähler manifolds. For further details, see Freed (1999a). Because of this structure, describing the infrared (IR) behaviour of a  $\mathcal{N} = 2$  theory reduces to determining the prepotential  $\mathcal{F}(a)$ .

---

<sup>26</sup>We have included dual variables, denoted by the subscript  $D$

<sup>27</sup>A special Kähler manifold is equipped with a closed two-form, the Kähler form:

$$\Omega = 2ig_{ij} da^i \wedge da^j \quad (1.96)$$

where  $d\Omega = 0$ . This condition ensures the existence of a (non-unique) Kähler potential that determines the metric, as we explain below.

The key insight of Seiberg and Witten is that the geometric data of the Coulomb branch can be encoded in an algebraic curve and a one-form. Conversely, specifying this algebraic curve and one-form allows us to reconstruct the IR physics at generic points on the Coulomb branch. In this framework, the special coordinates are identified with the periods of the one-form, while the algebraic curve defines a fibration over the Coulomb branch.

So far we have used special coordinates  $(a^D, a)$  to describe the Coulomb branch. However, these coordinates are not gauge-invariant, so to parameterise this space, we may instead use gauge-invariant quantities. For instance, in the case of  $G = SU(N)$ , one can define

$$u_i = \text{Tr} \phi^{i+1} \tag{1.98}$$

where  $i = 1, \dots, \text{rank } \mathfrak{g}$ . As a concrete example, for  $SU(2)$ , there is a single special coordinate  $a$  and a corresponding gauge-invariant parameter  $u = \text{Tr} \phi = 2a^2$ . This immediately highlights an important feature: since  $a = \pm\sqrt{u/2}$ , it is not single-valued and therefore does not provide a globally well-defined coordinate on the Coulomb branch. More generally, as one moves around certain points in the Coulomb branch, the coordinate  $a$  undergoes a monodromy transformation, failing to return to its original value.

In this example, the monodromy around  $u = 0$  acts as  $a \rightarrow -a$ , and similarly affects the dual coordinate as  $a^D \rightarrow -a^D + 2a$ . In general, these transformations relate different but equivalent descriptions of the theory, connected by *electric-magnetic duality* (or *S-duality*). For  $SU(2)$ , the full duality group is  $SL(2, \mathbb{Z})$  <sup>28</sup>.

Interestingly, these dualities reflect the non-existence of a single, globally valid coordinate system and instead indicate the necessity of multiple equivalent descriptions. This structure manifests geometrically in the form of monodromies around

---

<sup>28</sup>It is important to note that this is a *duality*, not a symmetry. From a field-theoretic perspective, electric-magnetic duality acts on the complexified coupling constant  $\tau$ . In particular, the transformation  $a \rightarrow -a$  corresponds to shifting  $\tau$  as  $\tau \rightarrow \tau + 1$ , which is known as (the negative of) a *T transformation*.

singular points of the Seiberg-Witten curve. Consequently, the duality transformations are naturally encoded in the monodromy structure of the algebraic curve.

It should be sufficiently clear that we need to describe this curve. Thus, let us examine its encoding more closely. Rather than viewing the special coordinates  $a$  as mere functions of the gauge-invariant coordinates  $u$ , we can leverage the monodromies discussed in the preceding paragraph to adopt a more geometric perspective. For  $r = \text{rank}; \mathfrak{g}$ , we can interpret  $(a^D(u), a(u))$  as  $2r$ -dimensional holomorphic sections of an  $Sp(2r, \mathbb{Z})$ <sup>29</sup> vector bundle over the Coulomb branch. The fibres of this bundle are  $r$ -dimensional complex tori.

Thus, given the Coulomb branch of the theory, we can construct this specific fibration, encoding the monodromy structure. One of the key insights of Seiberg and Witten was to reverse-engineer this construction: rather than starting from the Coulomb branch and deriving the fibration, they proposed the fibration itself as the fundamental object, from which the IR physics at generic points on the Coulomb branch can be extracted.

The way to achieve this is by specifying an algebraic curve  $\Sigma = P(x, u) = 0$  along with a one-form  $\lambda = x, du$ . Given this pair, one can determine the special coordinates by integrating the one-form over a basis of the cycles on the torus:

$$a_i^D = \oint_{A^i} \lambda \quad , \quad a_i = \oint_{B^i} \lambda \quad (1.99)$$

where  $i = 1, \dots, r$ , and  $A_i, B_i$  form a basis for the homology group  $H_1(\text{CB}, \mathbb{Z}) = \mathbb{Z}^{2r}$ . From this point, the rest follows directly from equation 1.95. In particular, the prepotential can be obtained via

$$\mathcal{F}(a) = \int a^D da \quad (1.100)$$

This completely determines the IR physics of the theory. The algebraic curve  $\Sigma$  is known as the *Seiberg-Witten curve*, while the one-form  $\lambda$  is referred to as the *Seiberg-Witten differential*.

---

<sup>29</sup>We remind the reader that  $Sp(2r, \mathbb{Z}) = SL(2, \mathbb{Z})$  for  $r = 1$

For example, returning to the  $SU(2)$  case considered above, the Seiberg-Witten curve<sup>30</sup> is given by

$$y^2 = x^3 - 2x^2u + \Lambda^4x \quad (1.101)$$

with the Seiberg-Witten differential given by  $dx/y$ . The curve develops three singularities over the Coulomb branch, located at  $u = +\Lambda^2, -\Lambda^2, \infty$ . The point  $u = \infty$  corresponds to the weak coupling limit<sup>31</sup>, while the points  $u = +\Lambda^2, -\Lambda^2$  are strongly coupled limits<sup>32</sup>. These singularities give rise to three monodromy matrices,  $M_{-\Lambda^2}$ ,  $M_{\Lambda^2}$ , and  $M_\infty$ , which satisfy the relation

$$M_{\Lambda^2} = M_\infty M_{-\Lambda^2}. \quad (1.102)$$

On the other hand, the mass of a BPS particle<sup>33</sup> is given by

$$M = |ma + na_D| \quad (1.103)$$

where the monodromy associated with an electric-magnetic charge  $(m, n)$  is

$$\begin{pmatrix} 1 + mn & -m^2 \\ n^2 & 1 - mn \end{pmatrix} \quad (1.104)$$

By comparing with the known monodromies at  $u = \pm\Lambda^2$ , we can identify the massless states at these points. At  $u = +\Lambda^2$ , a monopole of charge  $(0, 1)$  becomes massless, while at  $u = -\Lambda^2$ , a dyon of charge  $(2, 1)$  becomes massless. Thus, in addition to extracting local information via the prepotential, we also obtain global information about the BPS spectrum, specifically the states that become massless in the IR<sup>34</sup>. As expected, any global information we can extract helps to clarify the duality structure of the theory at strong coupling.

---

<sup>30</sup>In its Weierstrass form

<sup>31</sup>Where  $g \rightarrow 0$ , and  $(a^D, a)$  are good variables for describing the effective field theory.

<sup>32</sup>Where  $g \rightarrow \infty$ , requiring an S-dual frame to describe the IR physics.

<sup>33</sup>These are states that saturate the BPS bound  $M \geq |Z|$ , where  $Z$  is the central charge.

<sup>34</sup>In cases where the explicit Seiberg-Witten curve is unknown, this spectrum can serve as a guiding principle for reconstructing it. This provides a route from global monodromy data to the local structure of the Coulomb branch, encapsulated in the prepotential.

Before proceeding to the next section, let us recast the Seiberg-Witten curve in a form that will facilitate future connections. We rewrite the curve in its branched cover form as

$$\Lambda^2 z + \frac{\Lambda^2}{z} = -x^2 - u \quad (1.105)$$

where the Seiberg-Witten differential takes the form  $\lambda = xdz/z$ . In this formulation, the Seiberg-Witten curve can be interpreted as the spectral curve of an integrable system:

$$\lambda^2 - \phi_2(z) = 0 \quad , \quad \text{where } \phi_2(z) = \left( \frac{\Lambda^2}{z} + u + \Lambda^2 z \right) \frac{dz^2}{z^2} \quad (1.106)$$

Here, we set  $\lambda = ydz$ , and the coordinates  $(y, z)$  naturally live in the cotangent bundle  $T^*C$ , where  $C$  is the *UV curve* parameterised by the coordinate  $z$ . This formulation makes it clear that the Seiberg-Witten curve takes the form of a spectral curve, expressed as the vanishing locus  $\det(\lambda - \Phi) = 0$ . The differentials  $\phi_i$  appearing in this formulation will play a significant role in what follows. Our next step is to explore which integrable system naturally gives rise to this spectral curve.

#### 1.2.4 Hitchin systems

Before we proceed, let us pause to recast what we have found here in a slightly more mathematical language <sup>35</sup>. The Seiberg-Witten fibration with fibres being Lagrangian tori is a *complex integrable system*. This particular geometric structure, is called the *Seiberg-Witten system* Gorsky et al. (1995); Martinec and Warner (1996); Donagi and Witten (1996a), but our presentation shows the following important fact:

**Any  $\mathcal{N} = 2$  theory naturally defines a complex integrable system.**

Loosely speaking, an integrable system consists of a holomorphic symplectic bundle together with a projection  $p : \mathcal{M} \rightarrow B$ , such that the fibres  $F_b$  are compact

---

<sup>35</sup>Here we follow Neitzke (2016)

complex tori. There is, of course, a dictionary between the two sides of the same coin that we introduced earlier. For example, on the mathematical side,  $H_1(\mathcal{M}, \mathbb{Z})$  corresponds to the lattice of electromagnetic charges in the physical description. The choice of an electric-magnetic splitting corresponds to selecting a symplectic basis for  $H_1(\mathcal{M}, \mathbb{Z})$ , while the automorphisms of  $\mathcal{M}$  are encoded in the EM duality group  $Sp(2r, \mathbb{Z})$ .

Now, continuing along this mathematical perspective, let us proceed with the explicit construction of the Hitchin system. As we will see, this system arises naturally as the Seiberg-Witten system for many theories. The reason for this ubiquity is that a large class of  $\mathcal{N} = 2$  theories belong to the so-called class S.

To make this relation explicit, we consider the compactification of the six-dimensional  $(2, 0)$  theory on the product space  $C \times S^1$ , which gives rise to a three-dimensional  $\mathcal{N} = 4$  supersymmetric theory. There are two key regimes to analyze, depending on the relative sizes of the compactification spaces:

$$|S^1| \ll |C|, \tag{1.107}$$

where the circle  $S^1$  is much smaller than the Riemann surface  $C$ , and

$$|S^1| \gg |C|, \tag{1.108}$$

where  $C$  is much smaller than  $S^1$ . Each limit provides a different perspective on the theory. However, thanks to the partial topological twist, BPS-protected quantities remain unchanged, allowing us to extract meaningful physical insights from either approach.

Let us first consider the small- $S^1$  limit, where we first reduce the  $(2, 0)$  theory on  $S^1$  to obtain five-dimensional  $\mathcal{N} = 2$  super Yang-Mills theory. The bosonic content consists of a gauge connection  $A$  and five adjoint-valued scalars  $\Phi^I$ ,  $I = 1, \dots, 5$ , transforming in the fundamental representation of  $SO(5)$ . These are the same scalars we encountered in the six-dimensional  $(2, 0)$  theory, where they originally transformed

as a  $\mathbf{5}$  under the  $Sp(2)_R$  R-symmetry group. In that higher-dimensional context, they described the transverse fluctuations of an M5-brane in eleven-dimensional M-theory. Upon dimensional reduction to five dimensions, they become components of the adjoint-valued fields in the  $\mathcal{N} = 2$  vector multiplet.

When we further compactify on  $C$ , the  $SO(5)$  symmetry is broken, leading to the definition of the complex adjoint-valued *Higgs field*

$$\Phi = \frac{1}{2}(\Phi^4 + i\Phi^5)dz, \quad (1.109)$$

which is a  $(1,0)$ -form on  $C$ . These scalar fields play a central role in the emergence of the Hitchin system, where  $\Phi^4$  and  $\Phi^5$  together form the Higgs field, while the remaining scalars contribute to the characterization of possible Higgs or mixed branches of the theory.

The low-energy dynamics, governed by the vacua invariant under the three-dimensional Poincaré symmetry and supersymmetry, are described by the *Hitchin equations*:

$$F_A + [\Phi, \Phi^\dagger] = 0, \quad (1.110)$$

$$d_A \Phi = 0 \quad , \quad d_A * \Phi = 0 \quad (1.111)$$

where  $d_A = d + A$ . The moduli space of solutions to these equations, modulo gauge equivalence, defines an integrable system, namely the *Hitchin integrable system*  $\mathcal{M}_G(C)$  Hitchin (1987). The base space of this integrable system is parameterised by the VEVs of  $\Phi$ , also called in the literature as its Casimir values. In other words this is a fibration from the moduli space of pairs  $(A, \Phi)$  <sup>36</sup> which solve Hitchin's equations  $(A, \Phi) \rightarrow \{\text{Casimir of } \Phi\}$ . Physically,  $\mathcal{M}_G(C)$  serves as the Coulomb branch of the three-dimensional theory, and the correlation functions of the compactified theory are naturally interpreted as functions on this moduli space.

---

<sup>36</sup>More generally this pair defines a principal  $G$ -bundle with connection  $A$  and field  $\Phi$  called a Higgs field. In the literature, the Hitchin system is sometimes said to be a fibration of the moduli space of Higgs bundles.

From a different perspective, we could first compactify on  $C$ , yielding a four-dimensional  $\mathcal{N} = 2$  theory with Coulomb branch  $B$ . In this picture,  $B$  is identified with the base space of the Hitchin fibration. The Hitchin base is generally given by Chacaltana and Distler (2010a, 2013a); Chacaltana et al. (2013a)

$$B = \bigoplus_k H^0(C, K_C^{\otimes k}), \quad (1.112)$$

where  $K_C$  is the canonical bundle of  $C$  and  $k$  depends on the Lie algebra of the gauge group. For example, as we mentioned before, for  $\mathfrak{sl}_N$ ,  $k = 2, \dots, N$  and  $2, 4, 6, \dots, 2N - 2, N$  for  $\mathfrak{so}_{2N}$ . The spectral data of this system is encoded in the spectral curve  $\Sigma \subset T^*C$  of the Hitchin system.

The Seiberg-Witten solution naturally emerges from this setup as the spectral curve of the associated Hitchin system. More generally, the spectral curve

$$\det(\lambda - \Phi) = 0 \quad (1.113)$$

describes deformations of the Seiberg-Witten curve across the Coulomb branch, encoding the full structure of the integrable system. However, at the conformal point, where the UV theory retains scale invariance, there is a direct identification between the Seiberg-Witten curve and the Hitchin spectral curve. Here,  $\lambda = xdz$  defines the Seiberg-Witten differential, and the eigenvalues of  $\Phi$  determine the positions of the sheets of  $\Sigma_u$  in the cotangent directions, with  $\Sigma_u$  varying over the Coulomb branch parameterised by  $u$ . Here we are interested in conformal field theories and we will therefore, without ambiguity, identify the Seiberg-Witten curve to the spectral curve of the Hitchin system.

Since pure  $SU(2)$  gauge theory belongs to Class-S, it admits a description in terms of a Hitchin system. However, it is not conformal due to its asymptotically free nature, which introduces a dynamically generated scale  $\Lambda$ . In this case, the Seiberg-Witten curve takes the form given in equation 1.106.



Again it is important to pause here to understand what we have done. When the theory is in addition conformal, the spectral curve coming from the Seiberg-Witten integrable is the same as that from the Hitchin system. More generally:

**for four-dimensional class-S theories the Coulomb branch is identified with the base space of a Hitchin system**

or equivalently

**for three-dimensional class-S theories compactified on a circle, the Seiberg-Witten integrable system is a Hitchin system**

#### 1.2.4.1 Adding punctures

Thus far, we have seen that the Hitchin system naturally encodes the low-energy dynamics of four-dimensional  $\mathcal{N} = 2$  theories in Class S via the identification of the Seiberg-Witten curve with the spectral curve of the Hitchin system. To extend this construction and incorporate more general four-dimensional theories, we now introduce codimension-two defects Chacaltana et al. (2013a) in the Hitchin system. It is important to remark that the inclusion of these punctures will modify the Hitchin system parallelly to the physical construction, but the facts stated above will remain true.

We mentioned that these defects are inserted at distinguished points or punctures on the Riemann surface. From the perspective of the  $3d$  we just described they represent singular boundary conditions for the Hitchin equations at specific points on the UV curve  $C$ . From the perspective of the Hitchin system, punctures introduce singularities in the Higgs field  $\Phi$  at prescribed locations on  $C$  while preserving sufficient structure to maintain supersymmetry.

We are particularly interested in defects that respect  $\mathcal{N} = 2$  superconformal symmetry. Such punctures must be independent of the choice of complex structure on

$C$  since, following Gaiotto Gaiotto (2012a), we want to identify the moduli space of complex structures of  $C$  with the moduli space of exactly marginal gauge couplings in the four-dimensional theory. Physically, this means that the nature of the punctures should not depend on how we tune the gauge couplings.

### 1.2.5 Superconformal Punctures and Hitchin's Equations

As we have seen, the Hitchin equations on  $C$  provide the BPS equations for the effective four-dimensional  $\mathcal{N} = 2$  theories obtained from compactification. Within this framework, punctures introduce singular boundary conditions for the fields of the Hitchin system at discrete points on  $C$ . More precisely, punctures correspond to *complex codimension-one* (i.e., zero-dimensional) defects in the Hitchin system.

Since the Hitchin system is constructed from a reduction of five-dimensional  $\mathcal{N} = 2$  super Yang-Mills, whose gauge group is inherited from the original six-dimensional  $(2,0)$  theory, we restrict our attention to Hitchin systems with simply-laced gauge groups  $G$ . This ensures consistency with the global structure of Class  $\mathcal{S}$  theories.

For punctures that preserve  $\mathcal{N} = 2$  superconformal symmetry, the Higgs field  $\Phi$  exhibits a singularity at the puncture location. These punctures are characterised by a leading-order term given by a simple pole in the expansion of  $\Phi$ . Locally, near a puncture at  $z = 0$ , we write:

$$\Phi(z) = \frac{X}{z} + \dots, \quad (1.114)$$

where the ellipsis denotes subleading terms, and  $X$  is a representative of a nilpotent or semisimple orbit in the Lie algebra  $\mathfrak{g}$  of the gauge group  $G$ . The origin of this element comes from the classification of punctures in terms of nilpotent orbits we mentioned before.

Although it will not be the main focus of this dissertation and is left for future work, we briefly mention that conformal invariance can be broken by mass deforming

the theory. This involves adding a semisimple element  $\mathfrak{g}$  to  $\Phi$ , which encodes these mass deformations. This topic is part of a larger framework of *sheets* and has been studied by Balasubramanian and Distler Balasubramanian and Distler (2021a). We will revisit this concept in some form in the following chapters.

Using the grading in equation 1.112, we can express the spectral curve in terms of the graded sections:

$$\Sigma = \{0 = \det(\Phi - \lambda \mathbb{I})\} = \begin{cases} 0 = w^N - \sum_{k=2}^N \phi_k w^{N-k}, & \mathfrak{g} = A_{N-1} \\ 0 = w^{2N} + \sum_{k=1}^{N-1} \phi_{2k} w^{2(N-k)} + \tilde{\phi}_N^2, & \mathfrak{g} = D_N \end{cases} \quad (1.115)$$

where  $\phi_k \in H^0(C, L_k)$  for  $L_k = K_C^{\otimes k}$ . If  $X$  belongs to a nilpotent orbit, then the singular behaviour of the differentials is

$$\phi_k(z) = \frac{u_k}{z^{\pi_k}} + \dots, \quad (1.116)$$

where  $1 \leq \pi_k \leq \text{rank}; \mathfrak{g}$ . The values of  $\pi_k$  are determined by the specific nilpotent element. We will refer to this as the *pole picture*.

Rather than using this picture, we will consider twisting the canonical bundle on  $C$  by a divisor consisting of the sum of the points  $D = \sum_i^n p_i$  of the graded sections Balasubramanian et al. (2022a), where  $n$  is the number of punctures. In this way, we can write  $\Phi \in H^0(C, K_C(D) \otimes \text{ad}(A))$  near the puncture as

$$\Phi(z) = X + z(\dots) \quad (1.117)$$

where  $X$  is our nilpotent representative and the term with  $(\dots)$  vanishes as we approach the point. We will call this the *zero picture*. The reason the zero picture is convenient is that the spectral curve we wrote down is compact, as opposed to the one with the pole picture.

Now, our graded sections, near the puncture, will be

$$\phi_k = t^{X_k}(c_k + z(\dots)) \quad (1.118)$$

where  $\chi_k = k - \pi_k$ , marking the zero structure, is determined by the nilpotent element we choose to decorate the puncture. To ensure this zero structure, we must make the global sections

$$\phi_k \in H^0(C, L_k) \quad \text{with} \quad L_k = K_C(D)^{\otimes k} \otimes \mathcal{O} \left( - \sum_{i=1}^n \chi_k^{(n)} p_n \right) \quad (1.119)$$

where now  $\chi_k^{(j)}$  is the zero determined by the nilpotent element at puncture  $j$  at degree  $k$ .

There is a straightforward algorithm to extract this information for the types  $A_N$  and  $D_N$ . To begin, we recall that nilpotent orbits of  $\mathfrak{g}$  are in one-to-one correspondence with partitions of the rank of  $\mathfrak{g}$ . For the  $D_N$ -series, we specifically consider *D-partitions*, which are partitions of  $2N$  where even parts appear with even multiplicities. Given a nilpotent orbit associated with a partition  $[P]$ , we define the partial sums as:

$$k_j = \sum_{i=1}^j P_i \quad (1.120)$$

and for each  $k$  such that  $k_{j-1} < k \leq k_j$ , we set  $\chi_j = j$ .

### 1.2.5.1 Families

Up to this point, we have constructed the spectral curve of a Hitchin system for a specific gauge group  $G$  and a genus  $g$  curve  $C_{g,n}$ , decorated by  $n$  punctures. This construction corresponds to the Seiberg-Witten curve for the associated class-S theory obtained by compactifying a  $(2, 0)$  superconformal field theory on  $C_{g,n}$ . However, these theories naturally form families, which allows us to extend the construction in a more general sense. Remarkably, this familial structure leads to a nontrivial global construction that we briefly summarise here. Since the Hitchin system—or equivalently, the Seiberg-Witten system—provides a global description of the IR physics, we refer to this new overarching construction as the *very global construction*.<sup>37</sup>

---

<sup>37</sup>In this section, we summarise the key points from Balasubramanian et al. (2022a).

The class-S theories we have described form a family that is parameterised by the marginal parameters. These parameters control the gauge couplings and the complex structure of the curve  $C_{g,n}$ , and thus the moduli space of this family corresponds to the Deligne-Mumford space  $\bar{M}_{g,n}$ , which represents the space of complex structures on  $C_{g,n}$ . This moduli space captures the entire spectrum of possible class-S theories, each associated with a distinct point in  $\bar{M}_{g,n}$ , and it plays a central role in understanding the structure and dynamics of the theory as we vary the complex structure of the curve.

Within this framework, different codimension-1 boundaries of  $\bar{M}_{g,n}$  correspond to different weakly coupled gauge theory descriptions of the same underlying SCFT. These boundaries represent various limiting configurations of the theory, where the gauge coupling becomes small, and the theory can be approximated by a weakly coupled gauge theory. These limits are the same as those studied by Seiberg and Witten, who investigated the dynamics of the Coulomb branch in supersymmetric gauge theories. Later, Argyres further explored these limits, specifically focusing on the interactions between different vacua and the corresponding gauge theory descriptions, as described in earlier works Argyres and Seiberg (2007); Argyres and Wittig (2008). These studies reveal how distinct gauge theory descriptions emerge from the same SCFT depending on the choice of parameters within the moduli space  $\bar{M}_{g,n}$ .

This identification, however, depends on a condition imposed on the space  $H^1(C_{g,n}, \mathcal{L}_k)$ , specifically that the dimension of this space vanishes for all values of  $k$ . The line bundles  $\mathcal{L}_k$  play a crucial role in characterising the very global properties of the construction, and we will define them in more detail below. In the literature, these conditions are often refined further into categories such as *good*, *bad*, and *ugly* (see Balasubramanian et al. (2022a)). For the purposes of this discussion, however, the OK condition will suffice.

Let  $C_{g,n}$  be the  $n$ -punctured genus- $g$  Riemann surface with complex structure  $\lambda \in \bar{M}_{g,n}$ . We have already seen that the  $\phi_k$  are sections of line bundles on  $C_{g,n}$ ,

specifically the bundles  $L_k$ . As we vary  $\lambda$ , i.e., as we move within the moduli space, these line bundles also change. This variation corresponds to the modification of our family representative in the SCFT. In other words, as we adjust the marginal parameter, the bundles  $L_k$  will vary, and they will fit together to form a nontrivial vector bundle  $\mathcal{L}_k$ . This important phenomenon was first studied by Balasubramanian, Distler, and Donagi in Balasubramanian et al. (2022a).

Therefore, we have the following nontrivial fibration:

$$\mathcal{L}_k \rightarrow \mathcal{C}_{g,n} \xrightarrow{\pi} \bar{M}_{g,n} \quad (1.121)$$

where  $\mathcal{C}_{g,n}$ <sup>38</sup> is a *universal curve* that fibers over  $\bar{M}_{g,n}$ . The map  $\pi$  is the projection for this fibration. To make the fitting more precise, we want to find

$$\phi_k \in H^0(\mathcal{C}_{g,n}, \mathcal{L}_k) = H^0(\bar{M}_{g,n}, \pi_* \mathcal{L}_k). \quad (1.122)$$

This requires us to determine  $\pi_* \mathcal{L}_k$ .

Before proceeding with an example, let us briefly recall what these sections provide. First, computing the direct images  $\pi_* \mathcal{L}_k$  corresponds to determining the nontrivial vector bundles that encode the very global construction we are describing. Second, computing the  $\phi_k$  enables us to construct the spectral curves associated with the Coulomb branches (or Hitchin bases) of the SCFTs. This gives us the following decomposition:

$$\mathcal{B} = \bigoplus_k \mathcal{B}_k \quad (1.123)$$

where  $\mathcal{B}_k = \pi_*(\mathcal{L}_k)$ .

Now, let's consider the case of the four-punctured sphere. In this scenario, we have  $\bar{\mathcal{M}}_{0,4} = \mathbb{CP}^1$ , and the universal curve is a complex surface called *del Pezzo surface of degree four*, denoted as  $\mathcal{C}_{0,4} = dP_4$ . This surface is equivalent to  $\mathbb{CP}^2$  blown up at four points—specifically, the four points where the punctures are located.

---

<sup>38</sup>Not to be confused with  $C_{g,n}$

Now, specifying the case of four full punctures, we have the following condition for the zero structure:

$$\chi_k^{(i)} = 1 \quad \forall k = 2, \dots, N, \quad a = 1, \dots, 4 \quad (1.124)$$

Since every vector bundle over  $\mathbb{CP}^1$  splits as a direct sum of line bundles, we can explicitly compute the direct image sheaves. For each value of  $k$ , we get the following:

$$\mathcal{B}_k = \begin{cases} 4\mathcal{O}(1) \oplus 4\mathcal{O}(2) \oplus \dots \oplus 4\mathcal{O}(l-1) \oplus \mathcal{O}(l) & k = 2l \\ 4\mathcal{O}(1) \oplus 4\mathcal{O}(2) \oplus \dots \oplus 3\mathcal{O}(l-1) \oplus \mathcal{O}(l) & k = 2l + 1 \end{cases}$$

where  $\mathcal{O}(i)$  denotes the tautological line bundle over  $\mathbb{CP}^1$  of degree  $i$ .

In general, for any puncture, the direct image  $\pi_*(\mathcal{L}_k)$  may fail to be locally free. This occurs when the section dimensions over the boundary points, specifically  $h^0(\mathcal{C}_{g,n,\lambda=0,1,\infty}, \mathcal{L}_k)$ , experience jumps. These jumps can violate the OK condition that we previously discussed. In such cases, it becomes necessary to perform an additional twist, as outlined in the work of Balasubramanian, Distler, and Donagi Balasubramanian et al. (2022a). This twist ensures that the bundle behaves in a manner consistent with the desired properties, thus preserving the structure we are constructing.

Therefore, for a generic value of  $\lambda$ , the fibre  $\pi^{-1}(\lambda)$  is smooth. However, as we approach the boundaries  $\lambda = 0, 1, \infty$ , the curve  $\mathcal{C}_{g,n}$  degenerates into a pair of lines. In this situation, the Hitchin system is defined over a *nodal curve*. While this system does not have a mathematically canonical definition, we can adopt a physics-inspired construction to handle this case. At these specific points, we have a clear understanding of the physics, as these are precisely the weak-coupling limits of the theory that Seiberg and Witten studied.

#### 1.2.5.2 $SU(N) + 2N(\square)$

To illustrate some of the concepts discussed, we now consider a specific example:  $SU(N) + 2N(\square)$ . This theory has a class-S realisation associated with the

4-punctured sphere, where the punctures are labelled by  $[1^N]$ ,  $[1^N]$ ,  $[N-1]$ , and  $[N-1]$ . Using the framework developed by Balasubramanian, Distler, and Donagi (BDD), we can compute the degrees of the corresponding sheaves, which are found to be 1 at each degree, i.e.,  $\mathcal{B}_k = \mathcal{O}(1)$  for all  $k$ .

As an example, when  $N = 4$ , we can explicitly write down the family of spectral curves (choosing a trivialisation away from  $\lambda = \infty$ ) as follows:

$$\det(w\mathbb{I} - \phi) = w^4 - x(y-z)[c_2w^2 + c_3(x-y)w - c_4(x-y)^2] = 0$$

where  $x$ ,  $y$ , and  $z$  are the pullbacks of homogeneous coordinates on  $\mathcal{C}_{0,4} \cong dP_4$ , and  $c_i$  are the Coulomb branch parameters.

For  $\lambda \neq 0$ , there is another family with a similar structure. Through the projection  $\pi$ , these curves depend on the parameter  $\lambda \in \bar{M}_{g,n} \cong \mathbb{CP}^1$ . Thus, the family of Hitchin bases is given by

$$\mathcal{B} = \overbrace{\mathcal{O}(1)}^{k=2} \oplus \overbrace{\mathcal{O}(1)}^{k=3} \oplus \overbrace{\mathcal{O}(1)}^{k=4} = 3\mathcal{O}(1) \quad (1.125)$$

Now, this theory has a weakly-coupled Lagrangian description near  $\lambda = 0, \infty$ , which was already studied by Seiberg and Witten. At this point, we can relate the Coulomb branch parameters between the two weak-coupling descriptions. The relation between them is governed by the transition functions of  $\mathcal{O}(1)$ . Specifically, the Coulomb branch parameters transform as  $c_k \rightarrow c'_k \lambda$ , where  $c'_k$  are the Coulomb branch parameters that serve as good coordinates away from  $\lambda = 0$ .

With this, we further deepen our understanding of class S theories by exploring their global structure. As we will explore shortly, the challenge when extending this very global construction to the  $D_N$  series lies in a new phenomenon that arises in the Hitchin image, even at the local level, near a puncture. In type  $D$ , in addition to different partitions—specifically D-partitions—we encounter the fact that not all Coulomb branch parameters are independent. There will be relations between the  $c_{2k}$  parameters, one of the most apparent being that  $c_{2N} = \tilde{c}^2$ , which arises from the



Pfaffian, as seen in equation 1.115. These relations were first discovered by CD and depend on the choice of nilpotent element. We will explore how this plays out in the following sections Balasubramanian et al. (2023).

## Chapter 2: Information Transfer with a Gravitating Bath

### Abstract

Late-time dominance of entanglement islands plays a critical role in addressing the information paradox for black holes in AdS coupled to an asymptotic non-gravitational bath. A natural question is how this observation can be extended to gravitational systems. To gain insight into this question, we explore how this story is modified within the context of Karch-Randall braneworlds when we allow the asymptotic bath to couple to dynamical gravity. We find that because of the inability to separate degrees of freedom by spatial location when defining the radiation region, the entanglement entropy of radiation emitted into the bath is a time-independent constant, consistent with recent work on black hole information in asymptotically flat space. If we instead consider an entanglement entropy between two sectors of a specific division of the Hilbert space, we then find non-trivial time-dependence, with the Page time a monotonically decreasing function of the brane angle—provided both branes are below a particular angle. However, the properties of the entropy depend discontinuously on this angle, which is the first example of such discontinuous behavior for an AdS brane in AdS space.

---

<sup>0</sup>This chapter is based on published work in collaboration with Geng et al. (2021a). All authors contributed to the final version of the manuscript. My contributions included performing numerical computations (together with Marcos Riojas), which were presented in the figures—specifically,

## 2.1 Introduction

An important and interesting advance in our understanding of the quantum mechanics of black holes is the set of recent calculations Penington (2019); Almheiri et al. (2019a) of the time evolution of the entanglement entropy between a holographic system that contains a black hole and an external bath. (See Almheiri et al. (2020c); Liu and Sonner (2020); Raju (2020) for recent reviews). These calculations have yielded the so-called *Page curve* for the time-dependence of the entanglement entropy, which is consistent with unitarity and general expectations from quantum information theory Page (1993b); Lubkin (1978). One crucial ingredient in this calculation is the appearance of *entanglement islands* Almheiri et al. (2019c,d, 2020b); Penington et al. (2019), which are seemingly disconnected parts in the holographic system that contribute to the entanglement entropy of a region of the bath.

Attempts have been made to generalize this calculation to higher-dimensional models as well Almheiri et al. (2019d); Geng and Karch (2020b); Krishnan (2020); Ling et al. (2020); Chen et al. (2020a,b). A common feature of most such models to date is that the external bath into which the black hole evaporates is non-gravitating. It was recently pointed out in Geng and Karch (2020b) that in all such models with a propagating graviton, this graviton is massive as a direct consequence of coupling the gravitational theory to the bath. Take away the coupling, then you take away the mass. To what extent this coupling to the bath, and hence the mass of the graviton, is crucial in these calculations remains to be seen.<sup>1</sup>

Using somewhat orthogonal methods, the time evolution of the entanglement

---

identifying RT surfaces and using them to compute entanglement entropy, tracking the growth of the Hartman-Maldacena surface, and generating the Page curves. Additionally, I contributed to the conceptual development through discussions.

<sup>1</sup>One recent paper that explores the construction of entanglement islands and Page curves for black hole radiation in a massless theory of gravity, and where the radiation region is also gravitating, is Krishnan (2020). It deserves to be mentioned that we can construct entanglement islands in a massless theory of gravity without any information paradoxes to start with Hartman et al. (2020a); Balasubramanian et al. (2020a,b).

entropy of black holes in asymptotically flat space was recently studied in Laddha et al. (2020). One of the claims of Laddha et al. (2020) is that for black holes in asymptotically flat space, all information about the exact quantum state of the full spacetime (including the black hole) is accessible from an infinitesimal neighborhood of the past boundary of future null infinity. If true, the entanglement between the black hole degrees of freedom and the radiation reaching null infinity is completely time-independent, and the entropy curve would be flat instead of first rising and later falling.

In order to compare and contrast this result with the results of Penington (2019); Almheiri et al. (2019a, 2020c), the authors of Laddha et al. (2020) noted that the constructions of Penington (2019); Almheiri et al. (2019a, 2020c) have a non-gravitating external bath, so local observables in the bath in these calculations are well-defined, as is always true in quantum field theory. Since the radiation region in this scenario is taken to be part of the non-gravitating bath, the Hilbert space can be factorized into a part that describes the radiation and its complement. In contrast, gravity in asymptotically flat space does not shut off near null infinity. So it was suggested by Laddha et al. (2020) that, in order to mimic this situation and see if the Page curve calculations persist in gravitating systems, one would need to at least weakly couple the external bath to gravity. However, even the presence of weak gravity can potentially change the answers to fine-grained quantum information questions as was recently emphasized in Chowdhury et al. (2020).

Most work on islands in higher dimensions Almheiri et al. (2019d); Geng and Karch (2020b); Krishnan (2020); Ling et al. (2020); Chen et al. (2020a,b); Bousso and Wildenhain (2020); Hernandez et al. (2020); Bhattacharya et al. (2020); Chandrasekaran et al. (2020); Chou et al. (2020), as well as some of the results in lower dimensions Almheiri et al. (2020e); Rozali et al. (2020); Bak et al. (2020); Cooper et al. (2020); Caceres et al. (2020); Kumar Basak et al. (2020a), have been obtained in the context of the *Randall-Sundrum (RS)* braneworlds Randall and Sundrum (1999d)

with subcritical tension branes Karch and Randall (2001b,c), also known as the *Karch-Randall (KR)* braneworlds. Compared to the generic RS setup, KR has several specific features. It allows not just two but in fact three different holographically equivalent descriptions. Furthermore, gravity is generically massive, with an almost massless graviton only emerging in the near flat limit.

It is worth noting that, like the original calculation of Page (1993b), previous calculations also assumed a non-gravitating bath whereas we seem to live in a world with gravitating degrees of freedom. So in this paper we consider another important question allowed by this type of set-up: how the entanglement entropy calculation would generalize to a gravitating bath.

As with any RS scenario, the KR setup has the nice property that it is straightforward to introduce weak gravity in the bath region: one simply includes a second KR brane. The resulting setup is sketched in Figure 2.1a. A critical property of this generalization is that in this double KR branes setup, in contrast to the single brane case discussed in Geng and Karch (2020b), the model contains a massless graviton since the warped extra dimension now has a finite volume. The light (but massive) graviton that appears in the limit where the original brane is near the boundary co-exists with the truly massless graviton—a phenomenon that was dubbed bi-gravity in Kogan et al. (2001b). The massless graviton is a superposition of the modes localized near the left brane and the right brane. In the limit that one brane, designated as the bath, is weakly gravitating (in the sense that it’s closer to the boundary), the massless graviton is mostly localized on this weakly gravitating “bath” brane but also has some small overlap with the strongly gravitating “physical” brane. Together with the arguments put forward in Laddha et al. (2020) this seems to be an ideal scenario to investigate the significance of entanglement islands for black hole radiation in a system with gravitating bath.

Based on these motivations and observations, we will study two distinct types of entanglement entropy for our system with a gravitating bath: one closer to pre-

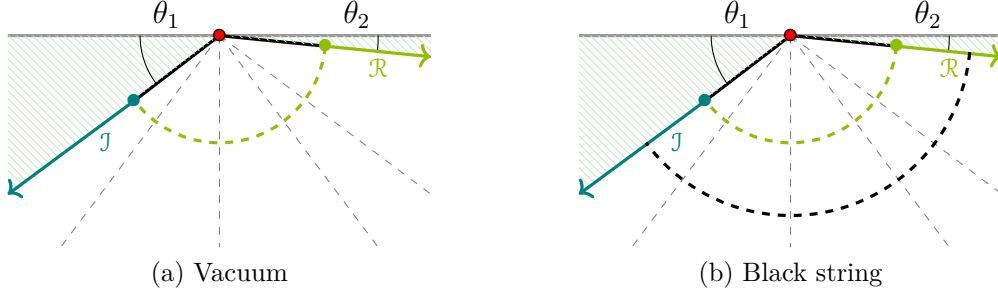


Figure 2.1: (a) Embedding of a KR braneworld with two subcritical branes in anti-de Sitter space. We refer to the left brane as the “physical” brane, and the right brane as the “bath” brane.  $\mathcal{R}$  denotes the radiation region whose entanglement entropy we wish to calculate, and the dashed green line connected to the boundary of  $\mathcal{R}$  is the candidate RT surface.  $\mathcal{J}$  is the entanglement island on the brane corresponding to the RT surface in green. (b) Embedding of a KR braneworld with the same tensions in the black string geometry. The dashed black line is the black string horizon separating the exterior and interior regions.

vious work on the entropy of the black hole radiation that we demonstrate has no time-dependent Page curve and another entropy that we define which follows an interesting Page curve even in this system. In both cases we consider a zero temperature (vacuum) setup as well as one with a thermal black string configuration Gregory and Laflamme (1993); Chamblin and Karch (2005). The latter is perhaps the simplest and most readily calculable way to induce a black hole on the brane.<sup>2</sup> We consider eternal,  $(d + 1)$ -dimensional geometries. Our numerical results primarily focus on  $d = 4$ .

Our first approach mimics earlier calculations in that we study the entanglement entropy of a radiation region on the bath brane. A key difference between a gravitating bath and a non-gravitating bath is that in the former, there is no natural, diffeomorphism-invariant way to separate local degrees of freedom into different regions, so the radiation region  $\mathcal{R}$  on the bath brane should be determined *dynamically*.

<sup>2</sup>Though both branes contain a black hole, we divide the setup by fiat into a physical and bath region.

We suggest a modest extension of previous dynamical principles to determine  $\mathcal{R}$  for KR two-brane setups by studying *quantum extremal surfaces (QES)* Engelhardt and Wall (2015b) via a doubly-holographic model, for which the location and the associated entropy of the QES in the  $d$ -dimensional theory is determined by a classical minimal (RT) surface in a  $(d + 1)$ -dimensional bulk.

Following this procedure, we will see that in the zero temperature case (global AdS),  $\mathcal{R}$  on the bath brane reduces to a point, and the entanglement entropy is exactly zero, which is consistent with the setup being dual to a pure state. However, for the black string geometry, we find that the RT surface is the horizon and  $\mathcal{R}$  is the region behind the horizon, and the entanglement entropy is given by the area of the horizon, *i.e.* the Bekenstein-Hawking entropy. Regardless, for both configurations we reach an agreement with arguments in Laddha et al. (2020), in the sense that, as the entanglement surface remains constant in time, the corresponding entropy is also constant.

In fact, Laddha et al. (2020) also suggested that while the Page curve may not describe the fine-grained entropy of the radiation when gravity is dynamical, there are *other questions* to which the Page curve might still be the answer. And, indeed, this is exactly what we find in this paper.

There is a second kind of question we can ask in our setup. Gravity switches off completely on the  $(d - 1)$ -dimensional boundary defect where both branes meet. (This point is marked as a red dot in both Figure 2.1a and Figure 2.1b). Therefore, it is again possible to factorize the Hilbert space on this boundary. This factorization allows us to perform an analysis that is analogous to those of Penington (2019); Almheiri et al. (2019a, 2020c). We can divide the boundary Hilbert space into different subsectors and ask about the entanglement entropy between them. We refer to the resulting quantity as *left/right entanglement entropy*. In the absence of entanglement islands this would correspond to entanglement between the left and right branes, although we explain below that this interpretation must be modified in the presence

of islands.

We will place our results in the context of both older Mollabashi et al. (2014) and more recent Akal et al. (2020); Miao (2020) progress on such systems, in which a  $(d - 1)$ -dimensional CFT is dual to a  $(d + 1)$ -dimensional gravitational space. In the more recent papers, this has been termed *wedge holography*. The left/right division alluded to above, refers to a well-defined *internal division* of the degrees of freedom of this  $(d - 1)$ -dimensional CFT into two subsystems. These subsystems are coupled but correspond to distinct factors of the Hilbert space, and so one can ask about the entanglement entropy between them and also the time-dependence of this entanglement entropy.

The left/right entanglement entropy is calculated in the  $(d + 1)$ -dimensional bulk as the smallest of two classes of candidate minimal surfaces anchored to the defect:

- the *Hartman-Maldacena surface* Hartman and Maldacena (2013), which stretches from the defect, where the two branes meet on the asymptotic conformal boundary, to its thermofield double partner, and
- the minimal surface starting from the defect and shooting towards one of the branes.

This second class of surfaces tells us that sometimes degrees of freedom on the right brane may be described by degrees of freedom from the left sector of the  $(d - 1)$ -dimensional CFT. This is why the left/right entanglement entropy should not be interpreted as the entanglement entropy between the left and the right brane.

We also show that in the black string geometry, the Hartman-Maldacena surface must cross an Einstein-Rosen (ER) bridge, so it grows in time, with the growth being linear at late times. Put another way, because the geometry has a horizon and



is a fast scrambler Maldacena and Susskind (2013), the growth of this surface is indefinite.<sup>3</sup> As the other candidate surface is constant in time, the Hartman-Maldacena surface should, at some point, be larger in area, allowing for the possibility of a phase transition for the entangling surface.

This is exactly the recent story of entanglement islands in the KR braneworld that gives a Page curve resolving a version of the information paradox for the eternal black hole Almheiri et al. (2019c). We study the left/right entanglement entropy in empty  $\text{AdS}_{d+1}$  and the  $\text{AdS}_{d+1}$  black string. Surprisingly, we find that the time-dependent Page curve for this left/right entropy appears if and only if the two branes are below a particular angle called the *Page angle*,  $\theta_P$ .

In our analysis, we find that above an angle slightly larger than the Page angle, called the *critical angle*,  $\theta_c$ , the island surfaces cease to exist altogether. The entropy is then governed by a limiting surface that we call the *tiny island surface* and describe in more detail in section 2.3.2.3. The value of  $\theta_c$ , as we will find in our  $d = 4$  numerics, is independent of the specifics of the bulk geometry *except* for the dimension because this value is controlled by the asymptotic region near the AdS boundary. We use this property to derive an analytical formula for  $\theta_c$  using empty  $\text{AdS}_{d+1}$ .

Our paper is organized as follows. In Section 2.2, we explore the entanglement entropy of dynamical radiation regions in empty  $\text{AdS}_{d+1}$  and the black string geometry with two branes. In Section 2.3, we study the left/right entanglement entropy of the black string, for which we may find a Page curve. In Section 2.4, we discuss the apparent universality of the critical angle  $\theta_c$  which controls the appearance of the Page curve. We compute  $\theta_c$  for general  $d$  using empty  $\text{AdS}_{d+1}$  and examine some of its properties. Finally, in Section 2.5, we conclude with additional discussions and remarks guiding future directions of interest. Appendix 2.A provides details of the area growth of the Hartman-Maldacena surface in the black string geometry.

---

<sup>3</sup>A similar statement would be true to other types of geometries with event horizons, such as de Sitter space which was studied in Geng (2020); Aalsma and Shiu (2020); Haque and Underwood (2020).

## 2.2 Entanglement Entropy of a Dynamical $\mathcal{R}$

In this section, we study the entanglement entropy of a radiation region  $\mathcal{R}$  on the bath brane which, as we will see in the following examples, can be redundant with degrees of freedom on the physical brane. We consider both zero temperature and finite temperature geometries.

In our zero temperature example, the bulk geometry consists of two  $\text{AdS}_d$  branes in empty  $\text{AdS}_{d+1}$ , both of which have positive tension. At finite temperature, we take the bulk geometry to be the  $\text{AdS}_{d+1}$  black string. In this case the induced geometries on both branes will be  $\text{AdS}_d$  planar black holes.

We argue that  $\mathcal{R}$  should be dynamically determined by the extremization of the entangling surface, as with previous quantum extremal surface prescriptions in the gravitating region, with important consequences for the resultant entanglement entropy. In Section 2.2.1, we encapsulate the motivation for the dynamical extremization principle to determine  $\mathcal{R}$ . In Section 2.2.2, we show that in the context of Karch-Randall braneworlds, the extremization procedure specifies that the entangling surface is perpendicular to both branes (Neumann boundary conditions). We emphasize that the entanglement entropy of  $\mathcal{R}$  is constant in time for both cases discussed in Sections 2.2.3 and 2.2.4, as anticipated by the analysis in Laddha et al. (2020).

### 2.2.1 Double Holography and Quantum Extremal Surfaces

In this subsection, we will review how KR braneworlds can be naturally understood as *doubly-holographic* and elaborate how holography helps us calculate entanglement entropy. This holographic dictionary encodes the von Neumann entropy of the black hole radiation in a classical minimal surface residing in a higher-dimensional gravity theory, with the boundary of this surface setting the location of the entanglement island. The ultimate purpose of this section is to propose an extension of this holographic dictionary from the usual one-brane to a two-brane setup. We will

see that because the radiation degrees of freedom whose entropy we calculate are also located on a gravitating region, the minimization of the bulk classical surface serves as a dynamical principle which decides not just the location of the island, but also the location of the radiation region.

Let us first briefly review the story of a single KR brane coupled to a non-gravitating bath and its three descriptions. We call such a system doubly-holographic because these three descriptions are related to each other by applying the standard AdS/CFT holography twice Karch and Randall (2001b,c); Almheiri et al. (2020e):

- (I) a  $d$ -dimensional CFT on a flat background with a  $(d - 1)$ -dimensional boundary (i.e. a BCFT $_d$  Cardy (2004); McAvity and Osborn (1995)), a description also emphasized in Takayanagi (2011b); Fujita et al. (2011),
- (II) a  $d$ -dimensional CFT with some characteristic UV cutoff coupled to gravity on an asymptotically AdS $_d$  space<sup>4</sup>  $\mathcal{M}_d$ , with the boundary of  $\mathcal{M}_d$  connected to another  $d$ -dimensional CFT on a half-Minkowski space via transparent boundary conditions,
- (III) Einstein gravity in an asymptotically AdS $_{d+1}$  space  $\mathcal{M}'_{d+1}$  containing a Karch-Randall brane Karch and Randall (2001b)  $\mathcal{M}_d$  as an end-of-the-world brane.

In keeping with previous calculations with a non-gravitating bath, we first consider the von Neumann entropy  $S(\mathcal{R})$  of a codimension-one (i.e. on a constant time slice) subregion  $\mathcal{R}$  defined in the description (I). In the description (II), using the quantum corrected *Ryu-Takayanagi formula* Ryu and Takayanagi (2006c); Lewkowycz and Maldacena (2013b); Faulkner et al. (2013b), this is given by the so-called *island rule* Penington (2019); Almheiri et al. (2020e) as,

$$S(\mathcal{R}) = \min_{\mathcal{J}} \text{ext } S_{\text{gen}}(\mathcal{R} \cup \mathcal{J}), \quad (2.1)$$

---

<sup>4</sup>For the minimal scenario, this is just AdS $_d$  itself.

where  $\mathcal{I}$  is a codimension-one region in  $\mathcal{M}_d$  disconnected from  $\mathcal{R}$  and thus called the *island*.  $S_{\text{gen}}$  denotes the *generalized entropy* functional used in the quantum extremal surface (QES) prescription of Engelhardt and Wall (2015b),

$$S_{\text{gen}}(\mathcal{R} \cup \mathcal{I}) = \frac{A(\partial\mathcal{I})}{4G_N} + S_{\text{matter}}(\mathcal{R} \cup \mathcal{I}), \quad (2.2)$$

and  $\partial\mathcal{I}$  is the QES. Here we want to emphasize that  $\mathcal{I}$  lives on the gravitating region and  $\mathcal{R}$  is on a non-gravitating region. Therefore, in (2.1) we minimize  $S_{\text{gen}}$  over  $\partial\mathcal{I}$  to obtain the von Neumann entropy, and this minimization in the gravitating region can be thought of as a quantum manifestation of the diffeomorphism invariance. The area term comes from using holography to promote the  $(d-1)$ -dimensional degrees of freedom in the BCFT $_d$  of description (I) to gravitational degrees of freedom on  $\mathcal{M}_d$ , yielding (II).

$G_N$  is the Newton's constant on  $\mathcal{M}_d$ . In the context of RS braneworlds, gravity can be thought of as entirely induced by matter loops after removing UV degrees of freedom. In this interpretation  $G_N^{-1}$  vanishes at tree level and is induced only by matter loops. Without a brane, this induced  $G_N^{-1}$  would be divergent, but since on the brane the CFT $_d$  has a cutoff, this induced contribution is finite. In practice, this means that the area term in the special case of induced gravity is included in  $S_{\text{matter}}$  as long as no explicit Ricci scalar term<sup>5</sup> is included in the brane action Emparan (2006); Chen et al. (2020b).

Using holography a second time helps us compute  $S_{\text{matter}}$ , which corresponds to the CFT degrees of freedom on the  $d$ -dimensional configuration of description (II). Specifically, the CFT matter in (II) is promoted to bulk gravity in (III), so to the leading semiclassical order  $S_{\text{matter}}$  can be computed using only the geometrical data in  $\mathcal{M}'_{d+1}$ . Since gravity on the brane is induced only by matter effects,  $S_{\text{gen}}$  itself is

---

<sup>5</sup>Almheiri et al. (2020e) and Chen et al. (2020b,a) included a *Dvali-Gabadadze-Porrati (DGP) gravity* Dvali et al. (2000c) term in the brane action, that is an explicit term proportional to the Ricci scalar of the induced metric. This is a higher curvature term in the brane's action. In this case, more care is needed beyond simply having induced RS gravity Chen et al. (2020b).

encoded by the geometry of  $\mathcal{M}'_{d+1}$  in accordance with the standard Ryu-Takayanagi prescription (assuming that  $\mathcal{R}$  is on a constant-time slice).

Therefore, we can write the resulting generalized entropy in terms of the area of a  $(d-1)$ -dimension surface  $\gamma$  cutting through the  $(d+1)$ -dimensional spacetime from  $\partial\mathcal{R}$  to  $\partial\mathcal{J}$ ,

$$S_{\text{gen}}(\mathcal{R} \cup \mathcal{J}) = \frac{A(\gamma)}{4G_{d+1}}, \quad (2.3)$$

where  $G_{d+1}$  is the Newton's constant of  $\mathcal{M}'_{d+1}$  in the description (III).

Combining this with (2.1), we get the doubly-holographic rule to compute  $S(\mathcal{R})$ —we minimize  $S_{\text{gen}}$  by varying  $\mathcal{J}$  and thus the position of  $\partial\mathcal{J}$ . This dynamical principle is essential because  $\mathcal{J}$  is defined in a gravitating region. However,  $\mathcal{R}$  is fixed beforehand, since the system described by (I) is non-gravitating and thus can be partitioned.

Now we may ask what happens when the bath is gravitating as well. We still have a doubly-holographic set-up, but involving the *wedge holography* discussed by Akal et al. (2020); Miao (2020):

- (I\*) a  $(d-1)$ -dimensional CFT on  $\mathcal{M}_{d-1}^{(0)}$ . Here the BCFT is replaced by  $(d-1)$ -dimensional surface since both branes have a lower-dimensional holographic dual.
- (II\*) two  $d$ -dimensional CFTs with characteristic UV cutoffs coupled to gravity on distinct asymptotically  $\text{AdS}_d$  spaces  $\mathcal{M}_d^L$  and  $\mathcal{M}_d^R$ , with these systems being connected via transparent boundary conditions at a defect  $\mathcal{M}_{d-1}^{(0)} = \partial\mathcal{M}_d^L = \partial\mathcal{M}_d^R$ ,
- (III\*) Einstein gravity in an asymptotically  $\text{AdS}_{d+1}$  space  $\mathcal{M}'_{d+1}$  containing two end-of-the-world Karch-Randall branes  $\mathcal{M}_d^{(L)}$  and  $\mathcal{M}_d^{(R)}$ , both of which are gravitational regions and both of which are attached at a defect and thus form a *wedge*.

However, as there is no fundamental distinction between the physical and the bath branes, now the quantum manifestation of the diffeomorphism invariance should be a straightforward generalization of the island rule to demand that instead of the single brane instructions in (2.1) and (2.2), we extremize  $S_{\text{gen}}$  over *both*  $\partial\mathcal{I}$  and  $\partial\mathcal{R}$ :

$$S = \min_{\mathcal{I}, \mathcal{R}}^{\text{ext}} S_{\text{gen}}(\mathcal{R} \cup \mathcal{I}), \quad (2.4)$$

$$S_{\text{gen}}(\mathcal{R} \cup \mathcal{I}) = S_{\text{matter}}(\mathcal{R} \cup \mathcal{I}). \quad (2.5)$$

As we mentioned above, in the context of RS braneworlds, where gravity is entirely induced, the generalized entropy is determined by matter effects alone so we don't include separate area terms. Now using holography, we know that  $S_{\text{matter}}(\mathcal{R} \cup \mathcal{I})$  can be calculated in terms of a classical RT surface  $\gamma$  connecting  $\partial\mathcal{R}$  to  $\partial\mathcal{I}$  in the higher-dimensional spacetime  $\mathcal{M}'_{d+1}$

$$S = \min_{\mathcal{I}, \mathcal{R}}^{\text{ext}} \frac{A(\gamma)}{4G_{d+1}}. \quad (2.6)$$

Our proposal (2.6) yields both an island  $\mathcal{I}$  and a radiation region  $\mathcal{R}$  whose boundaries are chosen to minimize the area of the surface  $\gamma$ . That we have no inputs at all is a result of the entire system being gravitating—we simply cannot choose a radiation region because this would involve partitioning gravitational degrees of freedom. The extremization over  $\gamma$  also includes extremizing over its endpoints, which is the equivalent of extremizing over  $\mathcal{R}$  and  $\mathcal{I}$ . This minimization over the boundary of  $\gamma$  eventually translates to boundary conditions of  $\gamma$  near the two branes. Since these boundary conditions are dynamically determined, we will call them dynamical boundary conditions.

We emphasize that there are several physical arguments one can make to support this claim about boundary conditions. First of all, with both branes gravitating, it seems unjustified to treat them asymmetrically as we would do by imposing a fixed radiation region as a boundary condition only on the physical brane. Second, in a theory with gravity, it is not meaningful to localize degrees of freedom in a region that

does not extend to asymptotic infinity. This is in accordance with the well-known result that there are no local gauge-invariant observables in quantum gravity DeWitt (1963); Kuchar (1991); Giddings et al. (2006). Furthermore, allowing for dynamical boundary conditions also appears to be what is required by the philosophy of quantum extremal surfaces Engelhardt and Wall (2015b); in a theory with dynamical gravity, the position of the surface across which we calculate the entanglement entropy of the matter fields<sup>6</sup> is not fixed a priori, but needs to be extremized over.

Our procedure is distinct from Krishnan et al. (2020) and also from Dong et al. (2020), where the authors attempted to define the entropy of a region with dynamical gravity using a “relational” procedure. In our context, the analog of this procedure would be to identify points in the radiation region by shooting geodesics from the defect and following them for a given proper distance along the brane. However, note that this procedure does not actually divide the algebra of operators on the brane into two commuting subalgebras. The geodesic that connects the operator in the radiation region to the defect can be thought of as a gravitational Wilson line that extends *outside* the radiation region. Since every operator that appears to be localized in the radiation region is actually attached to such a Wilson line, it does not commute with operators outside the region. The notion of a fixed radiation region may still be useful to define a more coarse-grained entropy Ghosh and Raju (2018), where one somehow sidesteps these gravitational effects, but not to compute a fine-grained entropy, as we are doing here. This issue was also recently discussed in Das et al. (2020), where it was emphasized that some observables may need to be systematically discarded from the algebra to define the entanglement-entropy of a fixed region in gravity.

---

<sup>6</sup>In our doubly-holographic model, recall that the  $(d + 1)$ th bulk dimension may be viewed as matter on the branes.

### 2.2.2 Extremization and Neumann Boundary Conditions

The geometries we wish to consider are depicted in Figure 2.1. The physical brane on the left is at an angle  $\theta_1$ , the bath brane on the right at  $\theta_2$ . We will generally be interested in scenarios in which  $\theta_2 \ll \theta_1$  so that the bath is weakly gravitating when compared to the physical brane, but most results will apply for general angles. Both branes are taken to be subcritical, so the geometry of the weakly gravitating bath is also asymptotically  $\text{AdS}_d$ , with a curvature radius that is much larger than that of the physical asymptotically  $\text{AdS}_d$  brane when the bath brane is at a small angle. When we take much smaller  $\theta_2$ , the truly massless graviton mostly overlaps with the bath brane, but also has an exponentially suppressed overlap with the physical brane. The radiation region  $\mathcal{R}$  is a subregion of the bath, indicated by a solid green line in Figure 2.1a.

In establishing some technical details and obtaining the boundary conditions determining  $\mathcal{R}$ , we first consider the vacuum solution. While islands are of particular importance when studying black holes on the physical brane (potentially in thermal equilibrium with the bath if the latter is kept at the same finite temperature as the black hole), it was recently shown in Chen et al. (2020b) that a lot can be learned already about the existence of islands from studying the zero temperature case.

In order to describe the system quantitatively, we need to commit to a coordinate system. There are two useful coordinate systems. One is the standard *Poincaré patch* coordinate system on  $\text{AdS}_{d+1}$  (setting the bulk curvature radius to 1),

$$ds^2 = \frac{1}{z^2}(-dt^2 + dz^2 + d\vec{x}^2 + dy^2). \quad (2.7)$$

Here  $y \in \mathbb{R}$  is the horizontal direction in Figure 2.1,  $z > 0$  is the vertical direction, and  $\vec{x} = (x^1, \dots, x^{d-2})$  represents the  $d - 2$  real transverse directions which are suppressed in the figure together with  $t \in \mathbb{R}$ . Here the boundary of  $\text{AdS}_{d+1}$  is located at  $z \rightarrow 0$ , and the brane ends on this boundary at  $y = 0$ .



The other coordinate system is the one adapted to the geometry of the subcritical KR brane,

$$ds^2 = \frac{1}{\sin^2 \mu} \left( \frac{-dt^2 + du^2 + d\vec{x}^2}{u^2} + d\mu^2 \right). \quad (2.8)$$

In Figure 2.1,  $u > 0$  and  $\mu \in (0, \pi)$  are radial and angular coordinates of a *spherical* coordinate system centered on the defect at which the two branes meet. In these coordinates, each subcritical brane is located at a constant  $\mu$ . For example, the two branes in Figure 2.1a are located at  $\mu = \theta_1$  and  $\mu = \pi - \theta_2$ . Several other lines of constant  $\mu$  are indicated as dashed lines. The geometry of a slice with constant  $\mu$  is itself  $\text{AdS}_d$ , with  $u$  being the radial Poincaré patch coordinate on each slice, *i.e.* the brane analog of  $z$ . The change of coordinates from  $y$ - $z$  to  $u$ - $\mu$  is given by,<sup>7</sup>

$$z = u \sin \mu, \quad y = -u \cos \mu. \quad (2.9)$$

Let us work out the minimal area surfaces connecting to a brane in this space-time. To find the equations of motion of the bulk minimal surface, it is easiest to work with the coordinates in (2.7). In this coordinate system, the area density per unit volume<sup>8</sup> in the transverse space labelled by  $\vec{x}$  for an embedding  $y(z)$  can be written as,

$$\mathcal{A} = \int \frac{dz}{z^{d-1}} \sqrt{1 + y'(z)^2}. \quad (2.10)$$

Since  $y(z)$  appears only in the “action” via its derivative, we can straightforwardly integrate the Euler-Lagrange equation to,

$$\frac{1}{z^{d-1}} \frac{y'(z)}{\sqrt{1 + y'(z)^2}} = \pm \frac{1}{z_*^{d-1}} \implies y'(z) = \pm \frac{z^{d-1}}{\sqrt{z_*^{2(d-1)} - z^{2(d-1)}}}. \quad (2.11)$$

Here, the integration constant  $z_* > 0$  denotes the depth of the *turnaround point* of the minimal surface. The minimal area surface reaches towards the boundary

---

<sup>7</sup>The sign here is chosen so that  $y$  is negative when  $\mu$  is between 0 and  $\pi/2$  and  $y$  is positive when  $\mu$  is between  $\pi/2$  and  $\pi$ .

<sup>8</sup>The transverse volume will play no role in our work. Strictly speaking all entropies we calculate are entropy densities per unit transverse volume.

(meaning to smaller  $z$ ) at both sides of the turnaround point. At  $z_*$  the two branches of the solution get glued together. Note that the *only* solution which does not have a turnaround point is  $y' = 0$ , the straight vertical line corresponding to  $z_* \rightarrow \infty$ . This is the only solution that passes through the Poincaré patch horizon. All other solutions will turn around and so reach either one of the two branes, hence giving rise to an island (J in Figure 2.1a).

The novel aspect of our calculation involves the boundary condition imposed on the bath brane. To determine the correct boundary conditions on the brane, it is easier to work in the  $u$ - $\mu$  coordinates. We consider the case where the actions for the branes contain only their tension, as is natural from the point of view of a low energy effective action. The tension term is leading in a derivative expansion. Thus in this case, there is no contribution to the entanglement entropy from the brane action, and as we discussed in Sec.2.2.1, the correct boundary conditions simply arise from minimizing the area functional over all possible locations of the surface along the boundary.

In this setup, the locations where the RT surface meets the branes are allowed to vary, and we minimize the area of the surface with respect to this location. Let us first review this analysis. In  $u$ - $\mu$  coordinates, our action for an embedding  $u(\mu)$  reads,

$$\mathcal{A} = \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + u'(\mu)^2}. \quad (2.12)$$

Note that the integration boundaries are important in specifying the boundary conditions. The action contains both  $u$  and  $u'$ , and so the equations of motion appear complicated, even though we know that the most general solution still must be given by (2.11) processed via the coordinate change (2.9).

When deriving the equations of motion we need to integrate by parts. Keeping

track of boundary terms, this means,

$$\begin{aligned}
0 = \delta\mathcal{A} &= \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{(u \sin \mu)^{d-1}} \frac{u'}{\sqrt{u^2 + u'^2}} \delta u' + \int_{\theta_1}^{\pi-\theta_2} \frac{\delta A}{\delta u} \delta u d\mu \\
&= \frac{\delta u}{(u \sin \mu)^{d-1}} \frac{u'}{\sqrt{u^2 + u'^2}} \Big|_{\theta_1}^{\pi-\theta_2} - \int_{\theta_1}^{\pi-\theta_2} d\mu \text{ (EOM)} \delta u.
\end{aligned} \tag{2.13}$$

The second term is the bulk equation of motion. While non-trivial in these coordinates, we know that the most general solution is given by (2.11) above. The first term is the boundary term obtained by integrating by parts. For it to vanish, we impose the following Neumann boundary conditions,

$$u'(\theta_1) = u'(\pi - \theta_2) = 0. \tag{2.14}$$

That is, the RT surface ends orthogonally to both branes. Note that this is the simplest type of boundary condition we may take which does not involve fixing a point on the brane and thus partitioning a gravitational system. Using the change of coordinates (2.9), we can readily see what this implies for a curve  $y(z)$  in the  $y$ - $z$  coordinate system.

$$\frac{dz}{d\mu} = u' \sin \mu + u \cos \mu, \quad \frac{dy}{d\mu} = -u' \cos \mu + u \sin \mu, \tag{2.15}$$

and hence,

$$\frac{dy}{dz} = \frac{dy/d\mu}{dz/d\mu} = \frac{u \sin \mu - u' \cos \mu}{u' \sin \mu + u \cos \mu}. \tag{2.16}$$

Using (2.14), this implies that on the brane

$$y' = \tan \mu|_{\mu=\theta_1, \pi-\theta_2} = -\frac{z}{y}. \tag{2.17}$$

This is how orthogonality reads in the  $y$ - $z$  coordinates. How could we have seen this boundary condition directly from the  $y$ - $z$  coordinates? The problem is that, in the action (2.10), varying the endpoint of the RT surfaces along the brane does not just change the variable  $y$  of the endpoint, but also the  $z$ -value that appears as a limit

of the integration region. To deal with this complication, it is easiest to parameterize the RT surface as  $(y(\tau), z(\tau))$ , where  $\tau$  runs from 0 to 1. With this, the  $y$ - $z$  action reads,

$$\mathcal{A} = \int_0^1 d\tau \frac{\sqrt{\dot{y}^2 + \dot{z}^2}}{z^{d-1}}, \quad (2.18)$$

where dots represent differentiation with respect to  $\tau$ .

This time, we see that the boundary term we pick up from integrating by parts in deriving the equations of motions demands,

$$\dot{y}(\delta y) + \dot{z}(\delta z)|_{\theta_1} = \dot{y}(\delta y) + \dot{z}(\delta z)|_{\pi-\theta_2} = 0. \quad (2.19)$$

To see what this means for  $y(z)$ , note that the boundary variation of  $y$  and  $z$  aren't independent. Let us focus on the brane at  $\mu = \theta_1$ ; the boundary condition at  $\mu = \pi - \theta_2$  is analogous. For the boundary of the RT surface to lie on the brane, whose embedding is given by the equation  $z = -y \tan \theta_1$ , we need to require,

$$\delta z = -(\delta y) \tan \theta_1. \quad (2.20)$$

This means that our boundary condition is,

$$\frac{dy}{dz} = \frac{\dot{y}}{\dot{z}} = -\frac{\delta z}{\delta y} = \tan \theta_1, \quad (2.21)$$

in perfect agreement with (2.17). We find a similar expression at the  $\mu = \pi - \theta_2$  brane.

We will also be studying the  $AdS_{d+1}$  *black string*,

$$ds^2 = \frac{1}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + d\vec{x}^2 + u^2 d\mu^2 \right], \quad h(u) = 1 - \frac{u^{d-1}}{u_h^{d-1}}. \quad (2.22)$$

These coordinates are the same as those of  $AdS_{d+1}$  in spherical coordinates, so  $u > 0$  and  $\theta_1 \leq \mu \leq \pi - \theta_2$  with the branes present. The action for an embedding  $u(\mu)$  which is analogous to (2.12) is,

$$\mathcal{A} = \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + \frac{u'(\mu)^2}{h(u)}}. \quad (2.23)$$

The equations of motion are more complicated because of the blackening factor  $h(u)$ , but the boundary conditions at the branes are still those of empty  $\text{AdS}_{d+1}$  (2.14).

For a non-gravitating bath, one demands Dirichlet boundary conditions, *i.e.* fixing a point  $y_0$  as the endpoint of the RT surface on the bath so that  $y_0$  defines the radiation region of interest. Fixing the endpoint would also satisfy the general boundary condition in (2.13). However, as we discussed in Section 2.2.1, for a gravitating bath, the correct prescription is to minimize over all possible endpoints of the RT surface on the bath brane. We saw in this section that this is translated into imposing the Neumann condition (2.14) on the bath brane. So instead of working with a fixed radiation region  $\mathcal{R}$ , we let the location of  $\partial\mathcal{R}$  adjust in order to minimize the entropy. Note that this doesn't contradict with the Dirichlet boundary condition if we continuously dial the bath brane to be a non-gravitating bath, because any extremal surface will automatically satisfy  $y' = 0$  as  $z \rightarrow 0$ , according to (2.11).

The procedure of imposing this boundary condition on the brane has an interesting implication. Immediately, we see that the trivial surface, *i.e.* the  $y' = 0$  surface such as that of Hartman and Maldacena Hartman and Maldacena (2013), is no longer allowed. As mentioned above, the only island-free RT surface, that is the only surface that does not turn around, is this trivial surface. But it can never be orthogonal to the bath brane if  $\theta_2 \neq 0$ . Thus, we have a unique RT surface whose area doesn't grow in time, and whose detailed form we will work out below, which always dominates. Hence, the entropy curve is flat, as in Laddha et al. (2020).

We conclude this subsection by linking our work to the analysis of Laddha et al. (2020) and the wedge holography of Akal et al. (2020); Miao (2020). In asymptotically Minkowski space, gravity is dynamical everywhere. As a result, the analysis of Laddha et al. (2020) concluded that the full information about (massless) excitations is available to any asymptotic observer on null infinity. Our setup is similar in that gravity is dynamical on both the physical and bath branes. Since both branes are  $\text{AdS}_d$ , joined at their common interface which we have been calling the defect

above, the role of the asymptotic observer is played by this defect. It is here that gravity shuts off; we are basically studying a universe whose spatial slices are compact, but have a codimension-two wedge on which gravity disappears, in accordance with Akal et al. (2020); Miao (2020). In direct analogy with Laddha et al. (2020), one would want to claim that for a geometry with a single asymptotic region, all information about excitations in the geometry is encoded on this defect, leading to constant entropy. For a geometry with two asymptotic regions like the eternal black string that we study in Section 2.2.3, the arguments of Laddha et al. (2020) would imply that all the information is encoded in the union of the two defects. This is, in fact, exactly what wedge holography would indicate: the CFT on the defect encodes the physics of the bulk geometry.

We would like to make two additional comments. First, both the arguments of Laddha et al. (2020) and of wedge holography would suggest that the fine-grained entropy of the defect as a whole should remain constant even when the bulk geometry is time-dependent. We do not check this claim in our paper since, in the black string geometry, both branes contain black holes that are in thermal equilibrium. But in principle, one could consider a dynamical solution where one starts with a black hole on the physical brane that evaporates into the bath brane. It would be interesting to extend our analysis to this case. Second, even if all the information is available near the defect, it is possible to still obtain a Page curve by dividing the defect itself into two parts as we describe in section 2.3. This is also consistent with the analysis of Laddha et al. (2020), which suggested that even when gravity is dynamical it may be possible to find a division of the Hilbert space that yields a Page curve, even though this does not correspond to a division of degrees of freedom into those “outside” and “inside” a black hole.

### 2.2.3 Empty AdS

We have two options for the bulk geometry; we may either use the equations of Section 2.2.2 and study the entanglement surfaces in a Poincaré patch of empty

$\text{AdS}_{d+1}$ , or we may use global  $\text{AdS}_{d+1}$ . We start with the former, comparing the result to the field theoretic answer. We then check the global case. Ultimately, all three results will match.

In the Poincaré patch, we need to find solutions of the form (2.11) obeying (2.14) on both branes. However, if the endpoints of the RT surface are allowed to vary, then nothing stops them from moving to larger  $u$ , with the surface expanding to contain the full space. Furthermore, we can see from (2.12) that the  $u^{-2}$  suppression of the metric leads to an area that goes to 0 as  $u \rightarrow \infty$ . Thus, we conclude that, in the Poincaré patch, both  $\mathcal{R}$  and  $\mathcal{J}$  for the true minimal, extremal surface are reduced to points at infinity, and  $S = 0$ .

Going to the true field theory dual allows us to make sense of this answer. As in the case of a non-gravitating bath, the two-brane KR setup has a triality of descriptions: the  $(d + 1)$ -dimensional classical bulk with two branes, a  $d$ -dimensional gravity theory coupled to a holographic CFT on the branes, and the true field theory dual. In the one-brane case, the purely field theoretic system is a BCFT. However, in the two-brane case, this non-gravitational description is a  $(d - 1)$ -dimensional CFT living on the defect, as discussed in Akal et al. (2020); Miao (2020). This is because the  $\mu$  direction is compactified on an interval, so, technically, the bulk theory is  $\text{AdS}_d$  times this interval, similarly to how a 10-dimensional string theory on  $\text{AdS}_5 \times S^5$  has a 4-dimensional dual—not a 9-dimensional dual.

In this CFT, we do not introduce any division of the degrees of freedom into subsets, so the entropy we are calculating is simply the von Neumann entropy of the entire CFT state. As the geometry is the vacuum, the CFT state is pure. Thus, the entropy should be zero. The consistency between this field theoretic answer and what we find in the bulk is another piece of evidence that allowing the endpoints of the RT surface to move so as to satisfy (2.14) is the correct course of action.

One may wonder if  $S = 0$  is also the correct answer when studying the solution in global  $\text{AdS}_{d+1}$ , rather than in the Poincaré patch we have been looking at so far.

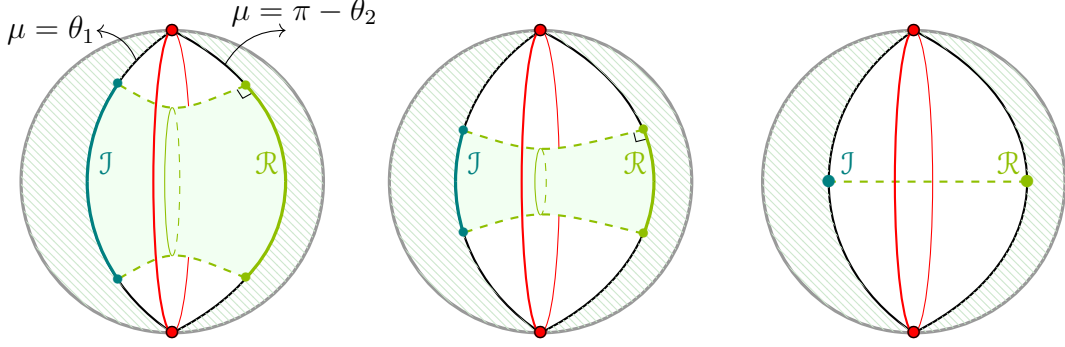


Figure 2.2: A constant- $t$  slice of global  $\text{AdS}_{d+1}$  with two branes present. The defect is shown in red, and various candidate extremal surfaces (in order of decreasing area from left to right) are in green. For each surface,  $\mathcal{R}$  and  $\mathcal{I}$  are respectively the radiation region and island, which end orthogonally on both branes. The minimal, zero-area entanglement surface  $r(\mu) = 0$  is shown on the right as a limit, cutting through the middle of the space.

This means that the field theory is now living on the sphere,<sup>9</sup> so there is now a scale in the problem: the curvature radius of the sphere. Global  $\text{AdS}_{d+1}$  is also the setting of Chen et al. (2020b) which established the existence of islands at zero temperature in modified braneworld setups. Our two-brane configuration is depicted in Figure 2.2.

It is quite straightforward to see that, with our prescription of imposing (2.14) on both branes,  $S = 0$  is still the global answer, even with a scale in the problem. In global  $\text{AdS}_{d+1}$ , the metric (2.8) is replaced by,

$$ds^2 = \frac{1}{\sin^2 \mu} \left( -\cosh^2 r dt^2 + dr^2 + \sinh^2 r d\Omega_{d-2}^2 + d\mu^2 \right), \quad (2.24)$$

where  $r$  goes from 0 to  $\infty$ . Instead of the  $d - 2$  planar coordinates  $\vec{x}$ , we now have the coordinates of a  $(d - 2)$ -sphere with volume element  $\sim d\Omega_{d-2}$ .

The branes are sitting at  $\mu = \theta_1, \pi - \theta_2$ , as before, with the boundary condition (2.14) implying  $r' = 0$  at both branes. The corresponding area density,

$$\mathcal{A} = \int_{\theta_1}^{\pi - \theta_2} d\mu \frac{(\sinh r)^{d-2}}{(\sin \mu)^{d-1}} \sqrt{1 + r'(\mu)^2}, \quad (2.25)$$

<sup>9</sup>Technically, the defect is homeomorphic to  $\mathbb{R} \times S^{d-2}$ , with the sphere parameterized by spatial coordinates. In Poincaré coordinates, this defect is homeomorphic to  $\mathbb{R}^{d-1}$ .



has a simple solution consistent with the  $r' = 0$  boundary conditions,

$$r(\mu) = 0. \quad (2.26)$$

with the corresponding entropy again yielding  $S = 0$ .

The shrinking of  $\mathcal{R}$  and  $\mathcal{J}$  is clearer in global coordinates than in the Poincaré patch, in which both regions run off to  $u \rightarrow \infty$ . As one can see in Figure 2.2,  $\mathcal{R}$  and  $\mathcal{J}$  are topologically  $d$ -dimensional balls. In the limiting surface capturing the entanglement entropy, these balls contract to points. The resulting generalized entropy of their union (2.5) is simply 0.

#### 2.2.4 Black String

To compare to the results of Penington (2019); Almheiri et al. (2019a, 2020c), we would like to study the system with a black hole on the physical brane. However, a uniform radiation bath in time-independent AdS space is not a solution to Einstein's equations, so we have a choice on how to proceed. We can keep the radiation uniform on the brane and allow for a time-dependent solution. Einstein's equations, which include the jump equations on the brane, force the brane to move in response to the radiation. The position would then be time-dependent and the induced worldvolume metric would be an FRW universe driven by the radiation. The exact solution for a single such brane was worked out long ago in Kraus (1999).

Alternatively, we can let the radiation in the bath region also collapse into a black hole, in which case we will have a black hole on both branes with matching Hawking temperatures. As long as both branes are above the Hawking-Page transition temperature, the system would be in stable equilibrium. This is exactly the scenario analyzed in Chamblin and Karch (2005). One bulk metric corresponding to this scenario is the  $\text{AdS}_{d+1}$  black string metric,

$$ds^2 = \frac{1}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + d\vec{x}^2 + u^2 d\mu^2 \right], \quad h(u) = 1 - \frac{u^{d-1}}{u_h^{d-1}}. \quad (2.27)$$

This has the same  $\mu$ -dependence as for empty  $\text{AdS}_{d+1}$  in (2.8), but this time with planar  $\text{AdS}_d$ -Schwarzschild black holes, instead of empty  $\text{AdS}_d$ , on each constant- $\mu$  slice. The corresponding spacetime is sketched in Figure 2.1b.

Such black string metrics can be afflicted by a Gregory-Laflamme instability Gregory and Laflamme (1993), but in the case of subcritical branes the instability is absent for large black holes, which includes the case of the planar black hole of interest to us. In fact, it was shown in Chamblin and Karch (2005) that the onset of the Gregory-Laflamme instability is the bulk manifestation of the Hawking-Page phase transition on the brane.

We now want to find the entangling surfaces. Since  $h(u)$  vanishes at  $u_h$ , the coordinate system in which we write the metric (2.27) is not well-suited to analyze the behavior of an RT surface near the horizon. We use the *tortoise-like coordinates*, instead; that is, we redefine the radial coordinate to be,

$$\frac{du}{\sqrt{h(u)}} = dr, \quad (2.28)$$

so that the  $t = 0$  slice of (2.27) now reads

$$ds_{t=0}^2 = \frac{1}{\sin^2 \mu} \left( \frac{dr^2 + d\vec{x}^2}{u^2} + d\mu^2 \right), \quad (2.29)$$

where it is understood that  $u$  is a function of  $r$ , obtained from integrating (2.28). The precise form of  $u(r)$  is not needed for this analysis, however.

Parameterizing  $r(\mu)$ , the corresponding action now reads,

$$\mathcal{A} = \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + r'(\mu)^2}. \quad (2.30)$$

Insisting that the boundary variation of  $\mathcal{A}$  vanishes implies the boundary condition  $r' = 0$  at both  $\mu = \theta_1$  and  $\mu = \pi - \theta_2$ . Furthermore, in the bulk, the two terms of the Euler-Lagrange equation read,

$$\frac{d}{d\mu} \left( \frac{\partial \mathcal{A}}{\partial r'} \right) = \frac{d}{d\mu} \left[ \frac{r'}{(u \sin \mu)^{d-1} \sqrt{u^2 + r'^2}} \right], \quad (2.31)$$

$$\frac{\partial \mathcal{A}}{\partial r} = \frac{\partial \mathcal{A}}{\partial u} \frac{du}{dr} = \frac{\partial \mathcal{A}}{\partial u} \sqrt{h(u)}. \quad (2.32)$$

While it is no longer possible to find the most general solution in closed form, we can readily write down one solution obeying the correct boundary conditions,<sup>10</sup>

$$u(\mu) = u_h. \quad (2.33)$$

This is the unique solution both obeying the Euler-Lagrange equations and satisfying the orthogonality boundary conditions at both branes, so the black string horizon itself is the single RT surface.

To see that the horizon is indeed the unique solution, we study the equations of motion to rule out other potential candidates. Away from the horizon, we do not need to use tortoise-like coordinates. If we parameterize our surface as  $u(\mu)$ , and then solve for the equations of motion for general bulk dimension  $d + 1$ , we obtain,

$$u'' = -(d-2)u h(u) + (d-1)u' \cot \mu \left( 1 - \frac{\tan \mu}{2} \frac{u'}{u h(u)} + \frac{u'^2}{u^2 h(u)} \right) - \left( \frac{d-5}{2} \right) \frac{u'^2}{u}. \quad (2.34)$$

If  $u'$  vanishes at some angle  $\mu = \mu_0$ ,

$$u''(\mu_0) = -(d-2)u(\mu_0) \left[ 1 - \frac{u(\mu_0)^{d-1}}{u_h^{d-1}} \right]. \quad (2.35)$$

In other words,  $u''$  is negative unless the surface lies precisely at the horizon. Since  $u' = 0$  at the left brane where the RT surface starts, this means that  $u' < 0$  just to the right of that brane. Now let's assume that there exists a second brane at which  $u'$  also vanishes. By the same argument,  $u'' < 0$  here as well. This second brane is supposed to be the endpoint of the RT surface, so we care about the value of  $u'$  to its left. Since  $u'' < 0$  and  $u' = 0$  at the brane,  $u' > 0$  just to the left of it.

Thus,  $u'$  must change sign between the two branes. There are two cases to consider—either  $u'(\mu)$  is continuous or the parameterization breaks down, corresponding to  $d\mu/du = 0$ . In the former, there exists an angle  $\mu_m$  between the two branes at

---

<sup>10</sup>(2.31) vanishes trivially by  $r' = 0$ , which is implied when  $du/d\mu = 0$ . Additionally, (2.32) vanishes because  $\partial A/\partial u$  is finite, but  $du/dr = \sqrt{h(u)} = 0$  at  $u = u_h$ .

which  $u(\mu_m) < u(\theta_1)$ ,  $u'(\mu_m) = 0$ , and  $u''(\mu_m) \geq 0$ . This third condition allows for  $u'$  to change sign, but it is in direct contradiction with (2.35). We deduce that any extremal surface besides the horizon cannot have a continuous derivative.

Thus, for  $u'$  to change sign, there must be an angle  $\mu_\infty \in (\theta_1, \pi - \theta_2)$  at which,

$$\lim_{\mu \rightarrow \mu_\infty^-} u'(\mu) = -\infty, \quad \lim_{\mu \rightarrow \mu_\infty^+} u'(\mu) = +\infty. \quad (2.36)$$

However, such a surface runs along a radial line of constant  $\mu$ , thus failing to make it “across” the defect and to the other brane. We have thus ruled out surfaces which are not the horizon (2.33).

The fact that we find our RT surface to be the horizon makes sense physically. Horizons are minimal area surfaces, and the horizon itself obeys the appropriate boundary conditions that we require. Also, the corresponding entropy satisfies,

$$S = \frac{A_{d+1}}{4G_{d+1}} = \frac{A_d}{4G_d}. \quad (2.37)$$

Here,  $A_{d+1}$  stands for the  $(d+1)$ -dimensional horizon area,  $G_{d+1}$  is the  $(d+1)$ -dimensional Newton constant, and  $G_d$  is the  $d$ -dimensional Newton constant after compactifying  $\mu$  on the interval. Note that  $G_d$  is not the same as  $G_{eff}$ , the Newton constant for the massive gravity theory on the physical brane.  $G_d$  is instead the Newton constant that governs the dynamics of the true zero mode, which is a superposition of bath and physical brane gravitons.  $G_d$  is given by the standard Kaluza-Klein reduction formula for compactification on an interval of finite volume,

$$\frac{1}{G_d} = \frac{1}{G_{d+1}} \int_{\theta_1}^{\pi - \theta_2} \frac{d\mu}{(\sin \mu)^{d-1}}. \quad (2.38)$$

The same factor of  $\int \frac{d\mu}{(\sin \mu)^{d-1}}$  also relates  $A_{d+1}$  to  $A_d$ , so that it is indeed straightforward to confirm that the  $(d+1)$ -dimensional area in  $(d+1)$ -dimensional Planck units is the same as the  $d$ -dimensional area in  $d$ -dimensional Planck units.

Lastly, from the  $(d-1)$ -dimensional CFT perspective, the answer is once again easy to understand. Here we are studying a mixed state that is given by the standard

thermal density matrix. The entropy is reproduced in the bulk by the horizon area of a static AdS-Schwarzschild black hole Maldacena (2003).

### 2.3 Left/Right Entanglement at Finite Temperature

We have seen that when gravity is dynamical for both the brane and bath, the conventional entanglement entropy of radiation leads to trivial results. In empty  $\text{AdS}_{d+1}$  bounded by KR branes, this entanglement entropy is always zero, whereas for a black string this entanglement entropy is always calculated by the horizon area. In neither case do we get a Page curve. In the first case, this is a consequence of empty AdS being in a pure state, whereas for the black string the finite temperature calculation captures the entanglement between the defect system and its thermofield double. This makes it appear that there is no evolution of the system with time. However that is not the case.

In this section we argue that there is time-dependence in some other form of entanglement entropy for finite temperature setups. Even with dynamical gravity on both KR branes, the defect is special in that it is non-gravitating. So we may divide its degrees of freedom into a “left” part and a “right” part. We call the entanglement entropy between these parts the *left/right entanglement entropy*.

We will see that, for a range of  $\theta_1$  and  $\theta_2$  parameters bounded from above by what we call the *Page angle*  $\theta_P$  this left/right entanglement entropy can yield a Page curve indicating the evolution of the structure of entanglement of a different type for our system. The physics of the Page angle is driven by the appearance of a *critical angle*  $\theta_c$ , beyond which some of the potential saddles contributing to the entanglement entropy cease to exist. We will see that this critical angle plays an important role in several different systems and seems to be the critical dividing point for at least two different types of RT surfaces. The apparently universal nature of this critical angle and its implications for the Page curve are not yet fully understood. They might however be a clue to the microscopic origin of the existence of non-trivial

saddle points that can lead to a Page curve and in any case might constrain the possibilities. We summarize what we find in Section 2.3.1 and present the details in those following.

### 2.3.1 Motivation and Summary

As discussed previously, holography dictates that the  $(d + 1)$ -dimensional classical bulk encodes the quantum effects of the  $d$ -dimensional description. Equivalently, the same entanglement entropy density can also be calculated in the  $(d - 1)$ -dimensional purely field theoretic dual. In computing the left/right entanglement entropy, we consider potential RT surfaces that start from the defect where the two branes meet. Broadly speaking, there are two classes of candidate surfaces: the Hartman-Maldacena surface Hartman and Maldacena (2013) connecting the defect to its thermofield double and the *island surfaces* connecting the defect to either brane. The appropriate quantum entangling surface on the brane is chosen via minimization.

In the exterior region of the black string geometry, the Hartman-Maldacena surface goes from the defect to the horizon. However, in the maximally-extended, two-sided black string geometry, there are two defects: the defect displayed in Figure 2.1b and its thermofield double. The Hartman-Maldacena surface connects these defects going through the bulk Einstein-Rosen bridge, and the portion of the surface located behind the horizon (which has zero area at  $t = 0$ ) increases indefinitely in time due to the growth of the Einstein-Rosen bridge, see Appendix 2.A. Naively, this would introduce an information paradox Hartman and Maldacena (2013).

Meanwhile, an island surface is defined as connecting the defect to a brane. As the name suggests, these surfaces feature islands, which appear as disconnected regions in the  $d$ -dimensional computation of the entanglement entropy. In the black string geometry, this class of surfaces saves us from the paradox mentioned above; the island surfaces are located strictly outside of the horizon, so they do *not* grow in time for the eternal black string geometry. The presence of these surfaces also

prevents us from strictly interpreting the left/right entanglement entropy as being the entanglement between the entire left and right brane. That they form islands indicates that some degrees of freedom on the left brane are redundant with degrees of freedom on the right brane. However, the division of the defect’s internal degrees of freedom into “left” and “right” is well-defined with or without islands. The lack of precise correspondence to the branes themselves is due to the interactions of the left and right degrees of freedom.

A brief summary of our results follows.

We will first demonstrate that in the black string geometry there exists a critical angle  $\theta_c$  for the brane location that demarcates regions of consistent RT surfaces. For a single brane with  $\theta > \theta_c$ , no RT surface obeying the correct boundary conditions (discussed in Section 2.2.2) and connecting the brane to the defect exists.

However, even though no local extrema exist when minimizing the area as a function of the intersection point of the brane with the surface at angles greater than the critical angle, the minimum value is obtained in an asymptotic limit. In the black string geometry, these surfaces become infinitesimal, but their area is not zero due to the UV divergence of the asymptotic  $\text{AdS}_{d+1}$  metric. We call such pseudo-island surfaces *tiny island surfaces*. In the  $\theta > \theta_c$  regime, we find that the area of the tiny island surface, while divergent, is infinitely smaller than that of the Hartman-Maldacena surface and so is, in fact, the dominant solution. The validity of these tiny island surfaces can be justified by slightly deforming the system in such a way that the tiny island surface actually has a finite negative area difference with the Hartman-Maldacena surface. We’ll provide such a justification both in the single-brane and the two-brane scenario.

On the other hand, for  $\theta < \theta_c$ , we find a genuine extremal surface ending orthogonally at a finite depth on the brane. We note that the asymptotic tiny island surface mentioned above also exists below the critical angle. However, for small angles, this surface has an area infinitely *larger* than that of Hartman-Maldacena. In fact,

we can alternatively define the critical angle as the value of  $\theta$  at which the leading divergence in the area difference between Hartman-Maldacena and the tiny island surface flips sign. As such,  $\theta_c$  only cares about the near boundary behavior of the metric and should not be sensitive to bulk geometrical properties. We will explicitly see this play out when comparing the black string to empty  $\text{AdS}_{d+1}$  (Section 2.4). So below the critical angle we do not need to consider the tiny island surfaces.

For angles much smaller than the critical angle, the true island surface has an area that is finitely larger than that of the Hartman-Maldacena surface at  $t = 0$ .<sup>11</sup> This means that the initial entangling surface is the Hartman-Maldacena surface, and there is a phase transition of the entangling surface after some time to the finite-depth island surface. However, at the critical angle, and in a small range of angles below it, the  $t = 0$  Hartman-Maldacena surface is already larger than the island surface, and so in this narrow range of angles the entropy is a time-independent constant and the entropy curve is trivial. We call the lower boundary of this region, with constant entropy curve dominated by the island surface, the Page angle.

With the two branes in Figure 2.1b, we compare the Hartman-Maldacena surface to all possible island surfaces. We find that a Page curve indeed exists for left/right entanglement entropy, but if and only if  $\theta_1$  and  $\theta_2$  are both below the Page angle. We emphasize that this Page angle, whose existence is driven by the critical angle, is a benchmark for us to get the Page curve.

The critical angle is perhaps the most intriguing aspect of our results and we have already seen that it plays several roles. It is the point where the tiny island surfaces become irrelevant and also where the leading divergence of the area of the finite-depth island surface below the critical angle equals that of the Hartman-Maldacena surface. In addition,  $\theta_c$  is the upper bound of the angles for which an

---

<sup>11</sup>We technically have two copies of the island surface (one in each of the maximally-extended black string's exterior regions). However, we may also compare the area of one copy to the area of Hartman-Maldacena in one of the exterior regions.



island surface can be found at all. This angle also played a role in the analysis of Chen et al. (2020a). We will further explore the generality and universality of the critical angle in a later section. Our main interest in the following section will be the physics of the black string geometry, representing a finite temperature state in the defect system.

### 2.3.2 Black String

Let us now turn to the black string. We will use the exactly known finite temperature solution to understand the evolution of left/right entanglement entropy at finite temperature. This entanglement entropy can be understood as originating from the left and right degrees of freedom on the boundary system, which consists of a  $(d - 1)$ -dimensional thermofield double state.

#### 2.3.2.1 Initial Time Hartman-Maldacena Surface

We choose to use polar Poincaré patch coordinates  $u$ - $\mu$ , where we are confronted with the metric and blackening factor  $h(u)$  given by (2.27). We start by analyzing the Hartman-Maldacena surface, which crosses the event horizon, passes through the Einstein-Rosen bridge, and anchors to the thermofield double defect on the other side of the Einstein-Rosen bridge. As expected for a surface passing through the Einstein-Rosen bridge, the area of this surface increases linearly following a short non-linear growth phase, a result of Hartman and Maldacena (2013) which we review in Appendix 2.A. Since the RT surface must be anchored to the defect, the Hartman-Maldacena surface must be anchored on both sides of the fully extended black string spacetime.

The general Hartman-Maldacena surface would be parameterized by a curve in the  $t$ - $u$  plane. However, for the purpose of understanding the Page curve, the first question is whether this surface contributes to the entropy at all. This is determined by the value of the area of this surface at  $t = 0$ , which is when it is the

smallest. In this subsection, we analyze this special  $t = 0$  surface, leaving the general time-dependent analysis for Appendix 2.A. Furthermore since we are interesting in numerically comparing the area of this surface to the area of other surfaces, we will specialize to  $d = 4$ .

Since the surface lies at a constant angle, the  $u(\mu)$  parameterization is inconvenient so we change coordinates to  $\mu(u)$ . For our initial  $t = 0$  slice we have the induced metric on the RT surface

$$d\tilde{s}^2|_{t=0} = \frac{1}{u^2 \sin^2 \mu} \left[ \left( \frac{1}{h(u)} + u^2 \mu'(u)^2 \right) du^2 + d\vec{x}^2 \right]. \quad (2.39)$$

For convenience we have set  $u_h = 1$  which gives  $h(u) = 1 - u^3$ . Using the induced metric, we get the area functional,

$$\mathcal{A} = \int \frac{du}{(u \sin \mu)^3} \sqrt{\frac{1}{h(u)} + u^2 \mu'(u)^2}, \quad (2.40)$$

from which we can obtain the equations of motion for the extremal surfaces:

$$\begin{aligned} \mu''(u) = 3 \cot[\mu(u)] & \left[ \frac{1}{u^2(u^3 - 1)} - \mu'(u)^2 \right] \\ & - \left[ \frac{2 + u^3}{2u(u^3 - 1)} \right] \mu'(u) - 2u(u^3 - 1)\mu'(u)^3. \end{aligned} \quad (2.41)$$

It is straightforward to see that there is a constant solution,  $\mu(u) = \pi/2$ , which corresponds to the Hartman-Maldacena surface at  $t = 0$ . The computation of the area is most easily done in Cartesian coordinates, where we can rewrite the area functional using  $\mu'(u) = 0$  and  $u \sin \mu = z$ . The area is divergent but can be regulated by introducing a small cutoff in the  $z$ -coordinate.

$$\mathcal{A}_{HM}(0) = \lim_{\epsilon \rightarrow 0} \left[ -\frac{1}{2\epsilon^2} + \int_{\epsilon}^1 \frac{dz}{z^3} \sqrt{\frac{1}{1 - z^3}} \right] = \frac{\sqrt{\pi} \Gamma\left(\frac{1}{3}\right)}{12 \Gamma\left(\frac{5}{6}\right)}. \quad (2.42)$$

This value will serve as a reference when we compute (and compare) the areas of other surfaces in the following subsections.

### 2.3.2.2 Island Surfaces

Our next task is to numerically determine those RT surfaces, referred to in this work as island surfaces, which start on the defect and end on one brane where they satisfy the boundary conditions (2.14). It is important to realize, as we said in the previous section, that we are forced to choose the defect as the point dividing off the radiation region.

These surfaces which start on the defect will end on the brane at some potentially arbitrary anchoring point  $u_i$ . The boundary conditions enforced by the extremization process, which require the surface trajectory to land orthogonally on the brane, are so restrictive that this point turns out to be determined completely by the strength of gravity on the brane. Each brane then enjoys its own unique anchoring point because the strength of gravity on the brane is uniquely determined by its angle. This point on the brane furthermore divides those surfaces which anchor to the brane and shoot across the defect from those which land on the same side; for this reason we will call this point the *critical anchor*.

For numerical convenience, instead of shooting from the defect to the brane, we enforce the boundary conditions on the brane when determining the solution to the differential equations and shoot to the defect. These conditions are,

$$u(\theta_1) = u_1, \quad u'(\theta_1) = 0, \quad (2.43)$$

$$u(\pi - \theta_2) = u_2, \quad u'(\pi - \theta_2) = 0, \quad (2.44)$$

where  $\theta_1$  and  $\theta_2$  are the angles for each brane and  $u_i$  ( $i = 1, 2$ ) gives the corresponding anchoring point for the RT surface.

Before discussing island surfaces in the full two-brane geometry, let us first discuss the case with a single brane, say the physical brane. That is, we first calculate the left/right entanglement entropy in the case of a non-gravitating bath. In this case, one could have fixed a generic radiation region in the bath region. The left/right

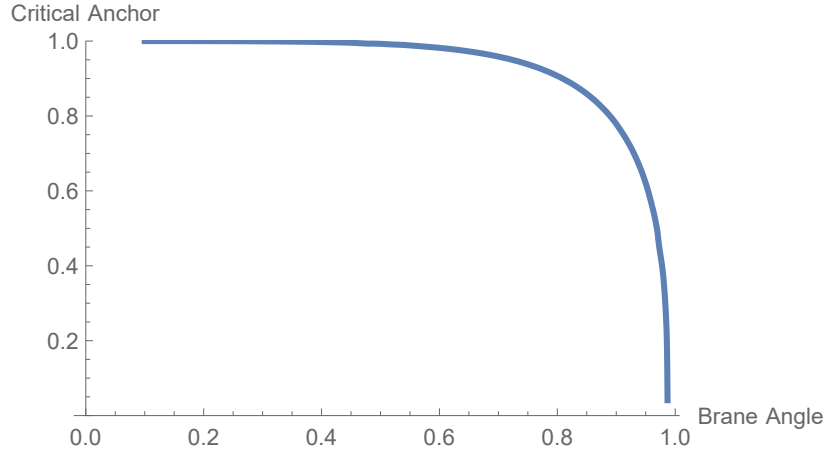


Figure 2.3: The critical anchor on the brane, given here in four dimensions, is a monotonically decreasing function of the brane angle which vanishes at the critical angle, which in this case is approximately  $\theta_c \approx .98687$ . The critical anchor tends toward the horizon distance ( $u_h = 1$ ) as  $\theta \rightarrow 0$  and tends toward zero as  $\theta \rightarrow \theta_c$ .

entanglement entropy here is just the limiting case that the radiation region is chosen to be the entire bath.

By imposing orthogonality, we ensure these surfaces are extremal, but they are not guaranteed to end on the defect for generic  $u_i$ , so we need to scan through all possible values of  $u_i$  to find the critical anchor which allows us to hit the defect. We numerically find those surfaces which hit the defect are only available for branes below a special value for the angle  $\theta_c$ ,

$$0 < \theta_c < \pi/2, \quad (2.45)$$

which we call the *critical angle*. This angle depends only on the dimension of the space (Figure 2.8), and the critical anchor is a monotonically decreasing function of the brane angle which vanishes at the critical angle (Figure 2.3).

The possible island surfaces for a selection of brane angles below the critical angle are illustrated in Figure 2.4. The critical anchor exhibits interesting limiting

behavior as a function of brane angle. Since we cannot shoot to the defect when above the critical angle, we cannot satisfy the boundary conditions. However, every extremal surface starting on the brane above the critical angle makes it over the defect. In fact, the critical anchor slides towards the defect as we approach  $\theta_c$  from below. Moreover, the area of the island surface approaches the area of the Hartman-Maldacena surface as one approaches the critical angle. Correspondingly the Page time gets smaller as we approach the critical angle (as illustrated in Figure 2.6, Section 2.3.2.4) and vanishes slightly below. Notice that island surfaces near the critical angle lie in the asymptotic region near the defect; this indicates that the value of the critical angle will be tightly constrained by the asymptotic geometry. Another interesting limit is  $\theta \rightarrow 0$  in which case the critical anchor tends toward the horizon distance  $u_h$ , the local strength of gravity on the brane vanishes, and the Page time diverges.

The limiting behavior of the critical anchor corresponds directly to the limiting behavior of the island surfaces—one loses the part of the island surfaces outside the horizon as the critical anchor approaches the horizon. In contrast, the island region fills the brane as the critical anchor approaches the defect when the brane hits the critical angle. Phrased in terms of the gravity on the brane, the Page time vanishes continuously as one increases the strength of gravity on the brane; during this process, an island forms and saturates the brane at the critical angle.

### 2.3.2.3 Above the Critical Angle and Tiny Island Surfaces

For a single brane, the finite island surfaces are lost when the brane lies above the critical angle. We want to show that in this case the area functional is dominated by infinitesimal tiny island surfaces, whose area difference with Hartman-Maldacena (4.11) is  $-\infty$ . The left/right entanglement will be a time-independent constant. As mentioned in Section 2.3.1, these are infinitesimal surfaces obtained as asymptotic limits near the defect that are subdominant when the brane is below the critical angle. However, above the critical angle, the tiny island surfaces are infinitely *smaller* than even the Hartman-Maldacena surface. Coupled with the  $\theta \rightarrow \theta_c$  behavior of Figure

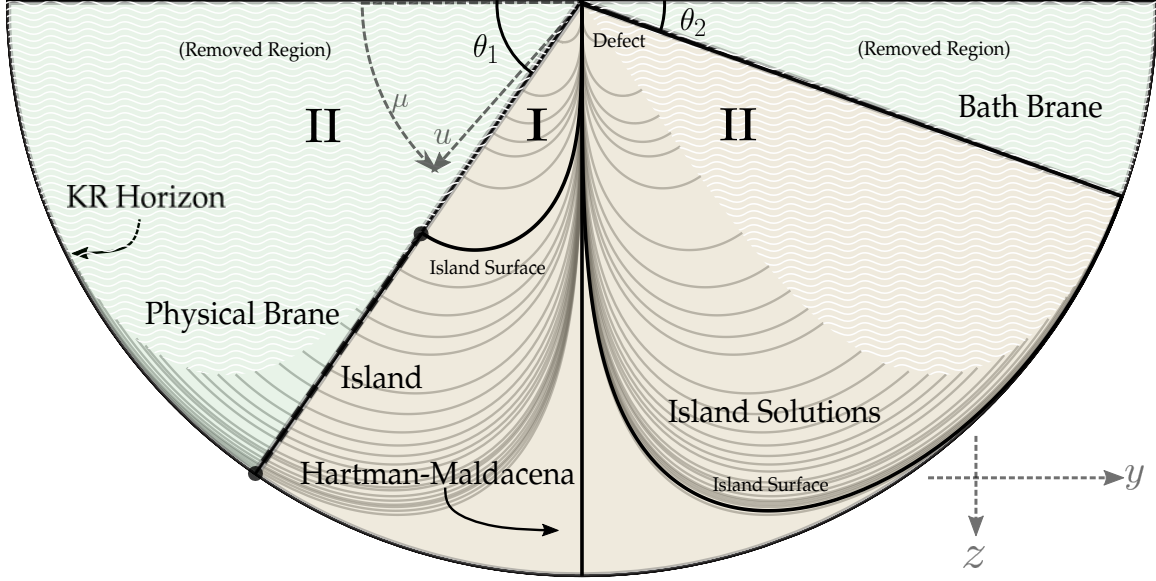


Figure 2.4: Overview for potential island surfaces within the black string geometry, each of which satisfies the boundary condition for some corresponding brane placed below the critical angle. Each surface ends at its critical anchor, which is where the island begins. One loses the island surfaces as the angle decreases, and the island fills the brane at the critical angle. The area difference between the Hartman-Maldacena surface and the island surface vanishes slightly below the critical angle. While one would need many digits of precision to show this, the island surfaces appear to fill the *island commonwealth*, labelled as region I, without crossing one another. They never make it into the region II, which we call *independent territory*.

2.3, the tiny island surface is the only alternative candidate to Hartman-Maldacena above the critical angle. In fact, we can even extract the critical angle from the behavior of the infinitesimal tiny island surfaces.

If we launch a surface from the defect with a classical trajectory and let it end at a finite  $u_i$ , but do not impose the Neumann boundary condition  $u' = 0$ , the area difference (4.11) becomes a function of  $u_i$ . We write this difference as  $\mathcal{A}(u_i)$ . The island surfaces for which  $u' = 0$  are genuine local minima of  $\mathcal{A}(u_i)$ .

One would expect that the properties of these minima depend on the full metric. However, the asymptotic form of  $\mathcal{A}(u_i)$  at small  $u_i$  is completely determined by the asymptotic form of the metric. In particular, the leading contribution to  $\mathcal{A}$  goes as

$$\mathcal{A} \sim \frac{I}{u_i^2}, \quad (2.46)$$

with  $I > 0$  below the critical angle and  $I < 0$  above the critical angle.<sup>12</sup> Although  $\mathcal{A}$  is a non-trivial function of  $u_i$ , its asymptotic form near  $u_i = 0$  determines the properties of its minima.

Only when  $I > 0$  can we possibly hope to find a single minimum for  $\mathcal{A}(u_i)$ . Thus, above the critical angle,  $\mathcal{A}$  cannot have a single minimum. Any single extremum would have to be a maximum. Thus the characterization of the potential extrema of  $\mathcal{A}$  has to change at the same critical angle in any background that is asymptotically AdS. In other words, the critical angle, defined this way, is *universal*.

The case of  $I < 0$  corresponds to the dominance of tiny island surfaces as we claimed above. The area difference functional is minimized by sending  $u_i$  to zero, where it diverges to negative infinity.

To properly define the tiny island surface in the single-brane case, one can make use of the freedom to change the radiation region when the bath is non-gravitating.

---

<sup>12</sup>At zero temperature, scale invariance indicates that  $\mathcal{A} = I/u_i^2$ , so island surfaces exist only at the critical angle, where  $I = 0$ . We will discuss and exploit this in more detail in Section 2.4.

Instead of taking the radiation region to be the entire bath, we can instead exclude a small strip of infinitesimal length  $\epsilon$  next to the defect from the radiation region. In this case, there exists an island surface whose critical anchor is of order  $\epsilon$  yielding a surface which has a large negative, yet finite, area difference with Hartman-Maldacena. In the  $\epsilon \rightarrow 0$  limit, one recovers the tiny island surface with its negative infinite area difference to Hartman-Maldacena as already noted in Geng and Karch (2020b) and consistent with general AdS/BCFT results Fujita et al. (2011).

### 2.3.2.4 Phase Transitions and the Page Curves

The Hartman-Maldacena and island surfaces are both available when working with a single brane below the critical angle. While both have UV-divergent areas, they have the same divergent terms and their area difference is finite with or without regularization.

$$\Delta\mathcal{A}(t) = \mathcal{A}_{IS} - \mathcal{A}_{HM}(t) < \infty. \quad (2.47)$$

$\mathcal{A}_{IS}$  is the area of the island surface, and  $\mathcal{A}_{HM}(t)$  is the area of the Hartman-Maldacena surface at time  $t$ .  $\mathcal{A}_{HM}(t)$  can be said to consist of two terms—one constant in  $t$  and another monotonically increasing for  $t > 0$ .

We display our numerical results for  $\Delta\mathcal{A}$  at  $t = 0$  in Figure 2.5. On this initial  $t = 0$  time slice the island surfaces are subdominant as long as they attach to a critical anchor on a brane satisfying  $\theta \ll \theta_c$ . This initial area difference vanishes close to but at a value smaller than the critical angle, the *Page angle*  $\theta_P$ . Increasing the angle beyond this point, the island surface dominates already at  $t = 0$ . Once we reach the critical angle the island saturates the brane. Increasing the angle even further, we lose the island surfaces and the tiny islands take over.

The Page time is determined by this computation of the finite area difference on the initial time slice. For branes above the Page angle, the island dominates from the very beginning and the entropy curve is flat. Even below the Page angle, the island surfaces will eventually dominate. While the island surfaces do not evolve



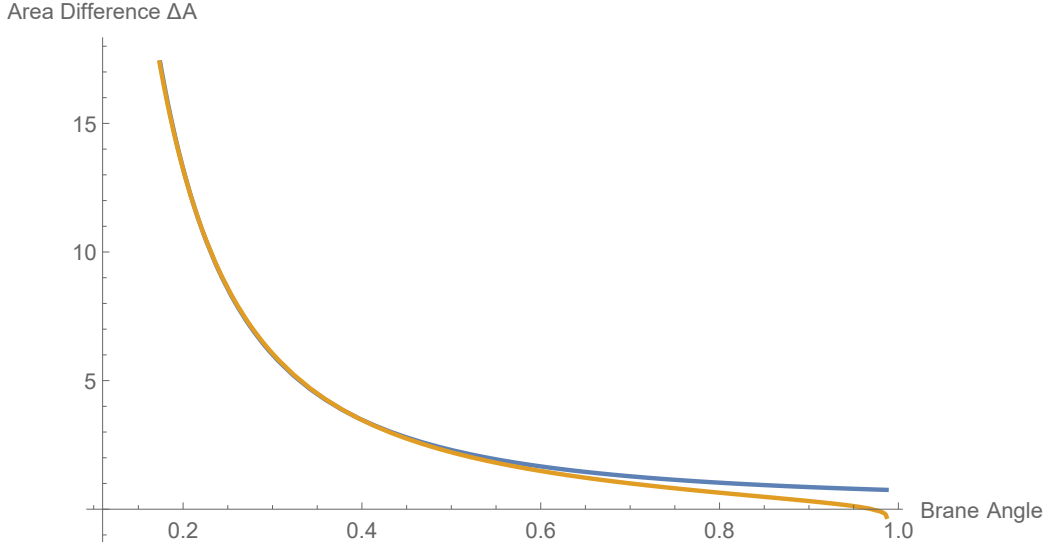


Figure 2.5: The  $t = 0$  area difference (4.11) between the Hartman-Maldacena surface and the island surface, given as a function of brane angle is in orange, given for the black string at  $d = 4$ . This should be compared to the behavior of the brane's induced Newton's constant  $G_N^{-1} \sim 1/\theta^2$  in blue. As we approach the critical angle the area of the island surface approaches the area of Hartman-Maldacena in empty AdS. This means the difference becomes negative in the region near the critical angle, which for  $d = 4$  is approximately  $\theta_c \approx 0.98687$ .

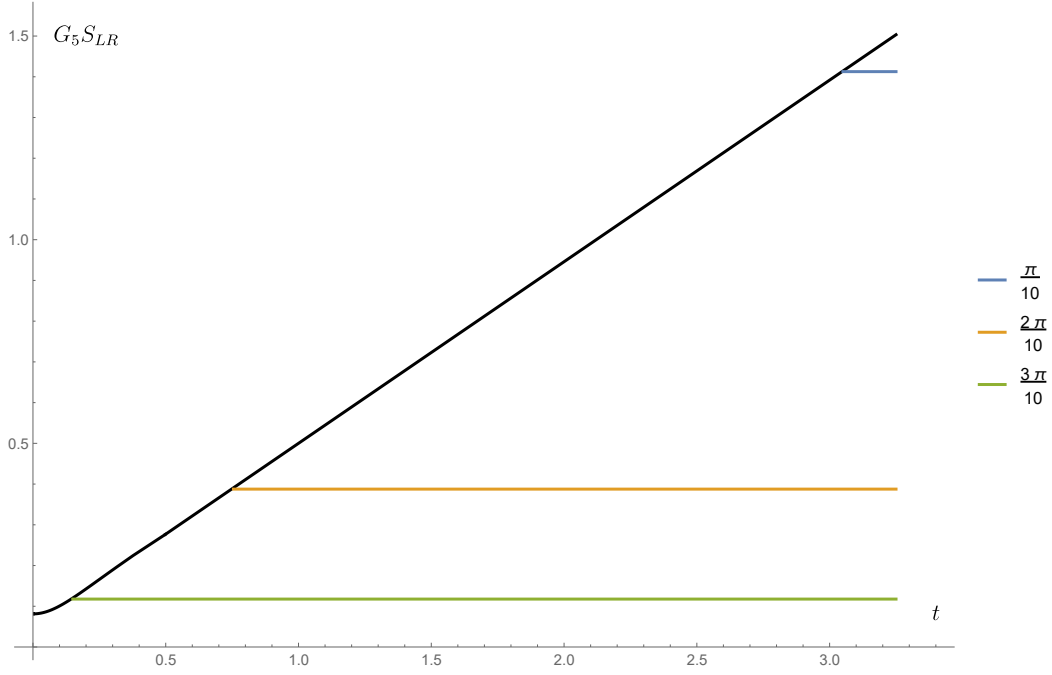


Figure 2.6: Page curves obtained for sample angles  $\theta = \frac{\pi}{10}, \frac{2\pi}{10}, \frac{3\pi}{10}$  starting just below the critical angle  $\theta_c \approx 0.98687$ . As the brane angle decreases, the Page time goes to infinity. As the brane approaches the critical angle  $\theta_c$ , the Page time decreases. To obtain the area difference, we integrate numerically up to some small region at which we model the area of the RT surface with a series.

with time even below the Page angle, the Hartman-Maldacena surface eventually grows linearly with time and depletes the area difference after some finite interval called the Page time. At this moment, the island surfaces become dominant, there is a phase transition, and the left/right entanglement entropy saturates. Because the area difference changes sign slightly below the critical angle, the Page time vanishes there and the islands take over immediately. The resulting single-brane Page curves are sketched for various angles in Figure 2.6. Note that while Figure 2.6 displays the entanglement entropy for  $t > 0$ , this quantity is symmetric about  $t = 0$  since  $\mathcal{A}_{HM}(-t) = \mathcal{A}_{HM}(t)$ .

As we emphasized, our numerics indicate that at the critical angle, the area difference between the island surface and the Hartman-Maldacena surface is actually

slightly negative. As a consequence, the area difference vanishes at the Page angle slightly below the critical angle. We can understand this finite value analytically. As noted in Section 2.3.2.3, at the critical angle the island surface lives entirely in the asymptotic region, and so its properties are equivalent to those of an island surface in empty AdS. This implies, in particular, that the area difference between the island surface and the empty AdS Hartman-Maldacena surface vanishes. Unlike the island surface, the latter actually explores the entire geometry, so we can see that,

$$\begin{aligned}
\Delta\mathcal{A}^{\theta_c} &= \mathcal{A}_{IS,BS}^{\theta_c} - \mathcal{A}_{HM,BS} \\
&= (\mathcal{A}_{IS,BS}^{\theta_c} - \mathcal{A}_{HM,AdS}) + (\mathcal{A}_{HM,AdS} - \mathcal{A}_{HM,BS}) \\
&= \mathcal{A}_{HM,AdS} - \mathcal{A}_{HM,BS}
\end{aligned} \tag{2.48}$$

Here all areas are calculated at  $t = 0$ . The BS and AdS subscripts indicate the black string and empty AdS geometries, respectively. The superscript  $\theta^c$  reminds us that we do this calculation at the critical angle. The first term on the second line vanishes by the fact that the area difference is 0 at the critical angle in empty AdS. Furthermore, the final result is exactly the negative of that calculated in (2.42),

$$\Delta\mathcal{A}^{\theta_c} = -(\mathcal{A}_{HM,BS} - \mathcal{A}_{HM,AdS}) = -\mathcal{A}_{HM}(0) = -\frac{\sqrt{\pi}\Gamma\left(\frac{1}{3}\right)}{12\Gamma\left(\frac{5}{6}\right)}. \tag{2.49}$$

Let us also note the interesting behavior of the area difference, which to a very good approximation is inversely proportional to the induced Newton's constant on the brane.<sup>13</sup> Such behavior would be consistent with the claim that gravity in a KR braneworld is purely induced by matter effects Emparan (2006).

These results apply as well when we include a weakly gravitating bath if both branes are below the Page angle. While the Hartman-Maldacena surface is unaffected, one of the island surfaces will dominate the other. This means the Page time is completely determined by the strongly gravitating brane, and the addition of the

---

<sup>13</sup>The induced Newton's constant refers to the gravitational coupling of the massive graviton localized to the brane—not the zero mode.

weakly gravitating bath does not change our result. If either brane is above the Page angle, it dominates the entanglement entropy already at  $t = 0$ , leading to a time-independent entropy curve. Last but not least, if either of the branes is above the critical angle we will once again be dominated by a tiny island surface and in this region we will also have a time-independent left/right entanglement.

Similar to what we did in the one brane case, we can ensure that the tiny island surface in the two-brane case is properly defined by thinking of it as a limiting case. This time we separate the branes at the conformal boundary by a small interval of infinitesimal size  $\epsilon$  instead of meeting at a defect. We then take an RT surface ending orthogonally on the brane at a very small  $u_i$  and divide the interval into two regions. This regulated situation sports a genuine RT surface whose area is less than that of the Hartman-Maldacena surface, with the difference becoming more and more negative as one shrinks the interval back to zero size.

To summarize, we have two distinct cases. If  $\theta_1 < \theta_P$  and  $\theta_2 < \theta_P$ , the left/right entanglement entropy for the AdS black string follows a Page curve, with the Page time determined by the stronger brane. Otherwise, we have no Page curve, and the left/right entanglement entropy is simply constant. Within this second case with at least one of the angles above  $\theta_P$ , there are two subcases. Either we still have  $\theta_1 < \theta_c$  and  $\theta_2 < \theta_c$  in which case the left/right entanglement surface is always given by an island surface with finite critical anchor, or one of the angles is above  $\theta_c$  and a corresponding tiny island surface dominates.

So there are two important angles. One is the Page angle which is relevant only for the black string. The second is the critical angle, which is the value at which the island surface disappears. This angle can also be defined as the angle at which the island saturates the brane. We can interpret these two angles in terms of two phase transitions. The first at the Page angle is a genuine phase transition from a regime with a time-dependent page curve to one without. The second is more like a spinoidal transition at which the island surfaces disappear. The tiny island surfaces exist at

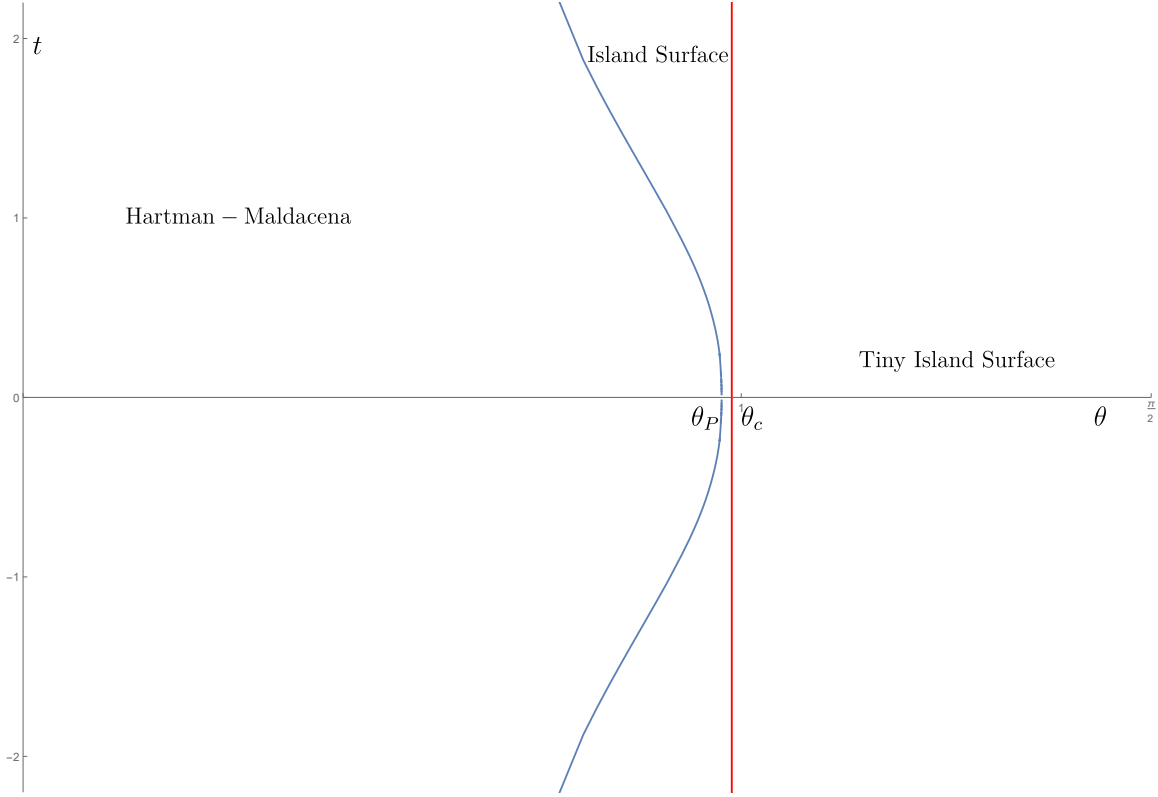


Figure 2.7: A phase diagram for  $d=4$  showing the dominant minimal surface as a function of time  $t$  and the larger brane-angle  $\theta$ . The transition between the island surface and the tiny island surface always takes place across the critical angle. The transition between the Hartman-Maldacena surface and the island surface is across the blue curve, which intersects  $t = 0$  at the Page angle.

all angles but are important only when their contribution to the entropy difference is negative infinity and hence the smallest area, which is the regime above the critical angle.

The full phase diagram of the system is sketched in Figure 2.7.

## 2.4 The Critical Angle

We have seen that the critical angle  $\theta_c$  can be used to characterize left/right entanglement surfaces for doubly-holographic models through our numerics for  $d = 4$ .

In particular, this parameter plays an important role in determining whether a given surface can access degrees of freedom on both sides of the wedge. We have encountered this feature in the black string geometry, but we have also argued that the existence of a critical angle should be determined solely by the asymptotic geometry, where the spacetime increasingly resembles empty AdS (Section 2.3.2.3). In this section, we first complement the numerical study for the black string with an analytic determination of the critical angle at zero temperature. In the second half of this section, we make some comments about the broader significance of this angle.

#### 2.4.1 The Critical Angle at Zero Temperature

We consider  $d > 2$  and study the left/right entanglement surfaces in empty  $\text{AdS}_{d+1}$  using  $y$ - $z$  coordinates. We have already found that any extremal surface  $y(z)$  satisfies (2.11),

$$y'(z) = \pm \frac{(z/z_*)^{d-1}}{\sqrt{1 - (z/z_*)^{2(d-1)}}}, \quad (2.50)$$

where  $z_*$  is the depth of the turnaround point. If we consider a trajectory starting at the defect and with  $z_* \rightarrow \infty$ , we obtain the empty  $\text{AdS}_{d+1}$  analog of the Hartman-Maldacena surface, for which,

$$y(z) = 0. \quad (2.51)$$

Alternatively, assuming that the turnaround is finite, one can eliminate the  $z_*$  parameter by a change in variables,

$$x = \frac{z}{z_*}, \quad \tilde{y} = \frac{y}{z_*}. \quad (2.52)$$

The solution then reads,

$$\tilde{y}'(x) = \pm \frac{x^{d-1}}{\sqrt{1 - x^{2(d-1)}}}. \quad (2.53)$$

This equation can be explicitly integrated to obtain the trajectories, which are clearly multi-valued functions  $\tilde{y}(x)$  with branch points at  $x = 1$ .

At this stage, we consider the trajectories which reach the KR brane located at  $\mu = \theta < \pi/2$ .<sup>14</sup> For such a brane, the island surface consists of a  $-$  (left-moving) branch anchored to the defect and a  $+$  (right-moving) branch starting orthogonally (Section 2.2.2) to the brane. Starting from the defect, we integrate to obtain an outgoing trajectory,

$$\tilde{y}_{\text{out}}(x) = -\frac{x^d}{d} {}_2F_1\left(\frac{1}{2}, \frac{d}{2(d-1)}; \frac{3d-2}{2(d-1)}; x^{2(d-1)}\right). \quad (2.54)$$

At the turnaround point  $x = 1$ , we may use an identity of the hypergeometric function to determine the value of this branch,

$$\tilde{y}_* = -\frac{\sqrt{\pi}}{d} \frac{\Gamma(\frac{3d-2}{2d-2})}{\Gamma(\frac{2d-1}{2d-2})}. \quad (2.55)$$

Now, we switch to the right-moving branch and integrate until we reach a point on the brane at  $x = x_b$  and  $\tilde{y} = \tilde{y}_b$ , which satisfy the relationship,

$$\tilde{y}_b = -x_b \cot \theta. \quad (2.56)$$

For the island surface to be continuous, the value of  $\tilde{y}_b$  must also be,

$$\begin{aligned} \tilde{y}_b &= \tilde{y}_* + \int_1^{x_b} dx \frac{x^{d-1}}{\sqrt{1-x^{2(d-1)}}} \\ &= \tilde{y}_* + \int_0^1 dx \left[ -\frac{x^{d-1}}{\sqrt{1-x^{2(d-1)}}} \right] - \int_0^{x_b} dx \left[ -\frac{x^{d-1}}{\sqrt{1-x^{2(d-1)}}} \right] \\ &= 2\tilde{y}_* - \tilde{y}_{\text{out}}(x_b). \end{aligned} \quad (2.57)$$

Next, we impose orthogonality at the brane. This reads as (2.17), which in  $\tilde{y}$ - $x$  coordinates is,

$$\tan \theta = \frac{x_b^{d-1}}{\sqrt{1-x_b^{2(d-1)}}} \iff \sin \theta = x_b^{d-1}. \quad (2.58)$$

---

<sup>14</sup>Considering instead  $\mu = \pi - \theta$ ,  $\theta < \pi/2$  instead will ultimately yield the same results, but the signs will be flipped on most intermediate steps.

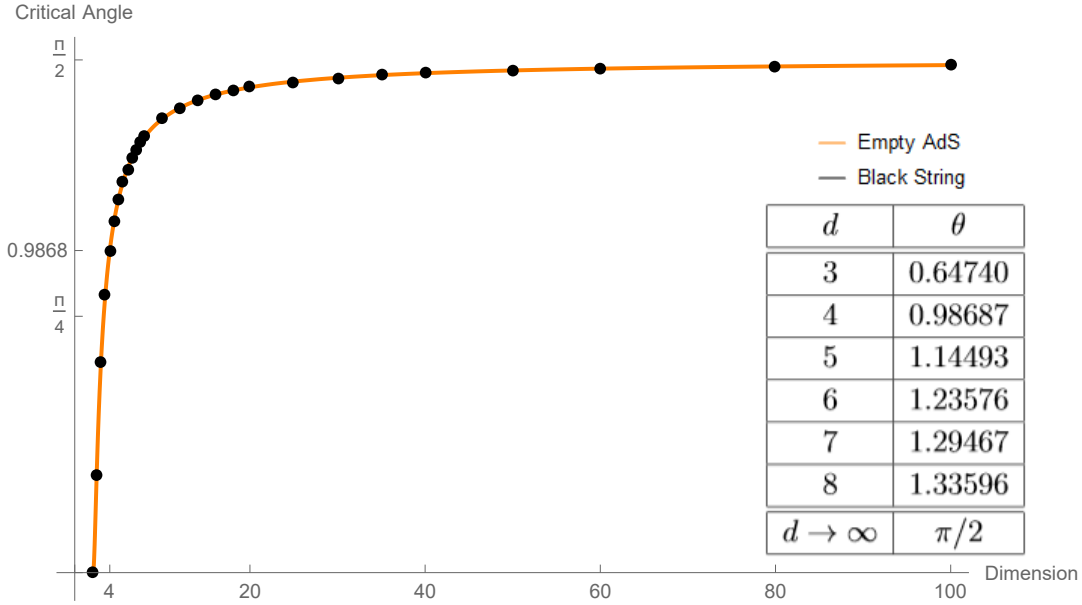


Figure 2.8: The critical angles in  $\text{AdS}_{d+1}$  for various dimensions  $d > 2$ , analytically continued to real values. The critical angle vanishes at precisely  $d = 2$ . Both empty AdS and the black string feature the same critical angles; the orange curve was obtained using (2.59), and the black points were obtained numerically by solving the shooting problem for the black string. The reason for the equivalence is that black string island surfaces near the critical angle are located near the defect, where the metric is well approximated by empty AdS.

Combining our results thus far yields a constraint on  $\theta$ .

$$\cot \theta = \frac{2\sqrt{\pi}}{d(\sin \theta)^{1/(d-1)}} \frac{\Gamma(\frac{3d-2}{2d-2})}{\Gamma(\frac{2d-1}{2d-2})} - \frac{\sin \theta}{d} {}_2F_1\left(\frac{1}{2}, \frac{d}{2(d-1)}; \frac{3d-2}{2(d-1)}; \sin^2 \theta\right). \quad (2.59)$$

This can be solved numerically, using standard root-finding methods, to compute the values presented in Figure 2.8 at various dimensions. Observe that the critical angle approaches  $\pi/2$  as  $d \rightarrow \infty$ . Additionally in Figure 2.8, the critical angles computed numerically for  $\text{AdS}_{d+1}$  black strings in various dimensions overlap with the plot of real solutions to (2.59). This emphasizes the universality discussed in Section 2.3.2.3.

By this method, we have obtained the critical angle analytically by studying island surfaces in empty  $\text{AdS}_{d+1}$ . In this case, we find that island surfaces exist *only*



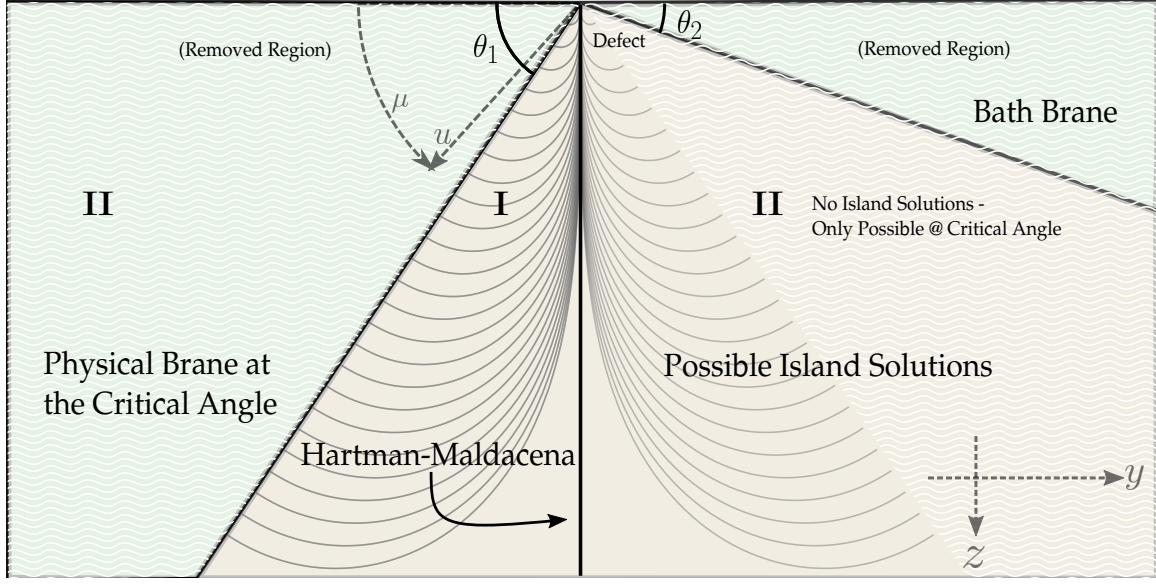


Figure 2.9: Possible island surfaces in empty  $\text{AdS}_{d+1}$ , which exist only for a brane placed at the critical angle. To summarize, we have two distinct cases. If  $\theta_1 < \theta_P$  and  $\theta_2 < \theta_P$ , the left/right entanglement entropy for the AdS black string follows a Page curve, with the Page time determined by the stronger brane. Otherwise, we have no Page curve, and the left/right entanglement entropy is simply constant.

These are relevant for the black string geometry, since island surfaces near the boundary are increasingly well approximated by empty AdS. The island surfaces completely fill the island commonwealth ( $\theta_c < \mu < \pi - \theta_c$ ), region I, without crossing one another. They also fail to make it into region II, the independent territory ( $\mu < \theta_c$  or  $\mu > \pi - \theta_c$ ).

at the critical angle, since the orthogonal boundary conditions can be satisfied only at that position for the brane. Candidates for such a surface are depicted in Figure 2.9.

As observed in Section 2.3.2.3 the leading divergence in the area difference between Hartman-Maldacena and the tiny island surface vanishes at the critical angle (Figure 2.5). We now confirm this to also be the case in empty  $\text{AdS}_{d+1}$ .

The area density of the surface (2.53) is given by  $\mathcal{A} = z_*^{2-d} I$ , where,

$$I = \lim_{\epsilon \rightarrow 0} \left[ -\frac{1}{(d-2)\epsilon^{d-2}} + \int \frac{dx}{x^{d-1}} \sqrt{1 + \tilde{y}'(x)^2} \right]. \quad (2.60)$$

To define the integral in (2.60), we introduce a small cutoff  $x = \epsilon > 0$ . The integral is performed starting at this cutoff and along the rest of an island surface again hitting a brane at  $\theta < \pi/2$ . In other words, we will need to integrate along both the outgoing trajectory and the returning trajectory. At the end, the cutoff is taken to zero.

Below we will calculate a regulated version of  $I$ , using a regularization scheme in which the area of the Hartman-Maldacena area vanishes. It is this regulated  $I$  that featured prominently in Section 2.3.2. We will see that the regulated  $I$  vanishes exactly at the critical angle. Furthermore, from the analytic expressions it is apparent that  $I$  is positive below and negative above the critical angle, as we already stated in Section 2.3.2.3.

To determine  $I$ , we first do the integral along the outgoing trajectory. Denoting this by  $I_{\text{out}}$ , we find that,

$$\begin{aligned} I_{\text{out}}(x) &= \lim_{\epsilon \rightarrow 0} \left[ -\frac{1}{(d-2)\epsilon^{d-2}} + \int_{\epsilon}^x \frac{dx}{x^{d-1}} \sqrt{1 + \tilde{y}'(x)^2} \right] \\ &= -\frac{1}{(d-2)x^{d-2}} {}_2F_1 \left( -\frac{d-2}{2(d-1)}, \frac{1}{2}; \frac{d}{2(d-1)}; x^{2(d-1)} \right). \end{aligned} \quad (2.61)$$

This integral is *already regulated* and yields a finite answer for a finite value of  $x$ . In particular, at the turnaround point,

$$I_{\text{out}}(1) = -\frac{\sqrt{\pi}}{d-2} \frac{\Gamma(\frac{d}{2d-2})}{\Gamma(\frac{1}{2d-2})} = \frac{1}{d-2} \tilde{y}_*, \quad (2.62)$$

where we have used  $\Gamma$ -function identities to obtain the second part of the equality above.

If we do the integral in (2.60) along the outgoing trajectory and the returning trajectory up to  $x_b$ , we obtain,

$$I(x_b) = 2I_{\text{out}}(1) - I_{\text{out}}(x_b). \quad (2.63)$$

In our regularization scheme, the area of the Hartman-Maldacena surface is just 0. So, the area difference vanishes if  $I(x_b) = 0$ , which is satisfied if,

$$2\tilde{y}_* = -\frac{1}{x_b^{d-2}} {}_2F_1\left(-\frac{d-2}{2(d-1)}, \frac{1}{2}; \frac{d}{2(d-1)}; x^{2(d-1)}\right). \quad (2.64)$$

We rewrite (2.57) and introduce the boundary condition (2.58) to alternatively express  $2\tilde{y}_*$  as,

$$2\tilde{y}_* = -\frac{\sqrt{1-x_b^{2(d-1)}}}{x_b^{d-2}} - \frac{x_b^d}{d} {}_2F_1\left(\frac{1}{2}, \frac{d}{2(d-1)}; \frac{3d-2}{2(d-1)}; x_b^{2(d-1)}\right). \quad (2.65)$$

We conclude that the area difference vanishes if the right-hand sides of (2.64) and (2.65) match. We prove that this is indeed the case by some hypergeometric identities. Writing  $z = x^{2(d-1)}$  and  $a = d/(2d-2)$ , we want,

$$\sqrt{1-z} + \left(\frac{2a-1}{2a}\right) z {}_2F_1\left(\frac{1}{2}, a; a+1; z\right) - {}_2F_1\left(a-1, \frac{1}{2}; a; z\right) = 0. \quad (2.66)$$

Once we recognize that the first two entries of  ${}_2F_1$  commute and  $\sqrt{1-z} = (1-z) {}_2F_1(a, 1/2; a; z)$ , (2.66) follows immediately using one of the so-called Gauss contiguous relations.<sup>15</sup> As we mentioned, requiring that the area difference vanishes yields the same critical angle as looking for the existence of solutions.

To summarize, we have two distinct cases. If  $\theta_1 < \theta_P$  and  $\theta_2 < \theta_P$ , the left/right entanglement entropy for the AdS black string follows a Page curve, with the Page time determined by the stronger brane. Otherwise, we have no Page curve, and the left/right entanglement entropy is simply constant.

## 2.4.2 Significance of the Critical Angle

The critical angle has appeared in several different contexts. It is a consequence of surfaces being dominated by the asymptotic region of AdS space so that

---

<sup>15</sup>See identity 15.5.16 in Lozier (2003).

the geometry away from the boundary plays little or no role. Here we see how this property manifests itself in the existence of islands in several different contexts; gravitating and non-gravitating bath as well as black string or empty AdS. Some of the points following might seem redundant but here we clarify the different role it plays in these distinct systems.

### **I Empty AdS with a gravitating bath.**

- (a) Island surfaces ending at the defect exist only at the critical angle.
- (b) The area difference between the island surface and the Hartman-Maldacena surface vanishes at the critical angle.
- (c) The island surface at the critical angle can be taken to end anywhere on the brane by scale invariance. In particular, there exists a choice of island surface for which the island saturates the brane.

### **II Empty AdS with a non-gravitating bath.**

- (a) Island surfaces that cross the defect and end in the bath exist only above the critical angle. There is an extremal surface connecting every point on any brane above the critical angle to the bath.
- (b) If one fixes one endpoint of the surface in the bath, and approaches the critical angle from above, the other endpoint of the surface on the brane runs off to the Poincaré horizon. This point was uncovered independently in Chen et al. (2020a).
- (c) It is only for the critical angle that surfaces starting at a finite anchor point on the physical brane can be found to hit the defect in empty  $\text{AdS}_{d+1}$ .

### **III Black string with a gravitating bath.**

- (a) For the black string, island surfaces ending at the defect exist only below the critical angle.

- (b) The critical angle is where the leading divergence in the area difference between the island surface and the Hartman-Maldacena surface vanishes. This divergence changes sign above the critical angle, which is the region where the tiny island surfaces dominate the entanglement entropy calculation.
- (c) The Page time vanishes slightly below the critical angle. This is because the island surface at the critical angle has the same area as the Hartman-Maldacena surface at zero temperature, which is slightly smaller than its value at finite temperature (2.42).

#### IV. Black string with a non-gravitating bath

- (a) Below the critical angle, only those surfaces that are anchored beyond a critical anchor on the brane reach the bath. The critical anchor is an asymptotically decreasing function of the angle which decreases from the horizon and vanishes at the critical angle.

The universality of this angle for a given dimension suggests a non-trivial phase transition occurring at the critical angle in many different aspects of KR brane-worlds. For now, we have focused on the implications of our result for the Page curve and leave further study concerning universality of the critical angle to future work, with some speculations in the concluding section.

## 2.5 Conclusion

In this work, we have analyzed the Page curve for black holes in AdS coupled to a gravitating bath using KR branes as a tool. Our main findings were that gravity in the bath leads to several major modifications of the Page curve story. While coupling to a non-gravitating bath via boundary conditions implies massive gravitons, extending beyond this analysis to include a gravitating bath means that we are actually studying a theory with a massless graviton. We have shown how gravity in the

bath completely changes the behavior of the Page curve for conventional entanglement entropy. If one asks the same questions as has been done in previous work with a non-gravitating bath, *i.e.* computing the entanglement entropy of some region in the bath identified as the radiation region, the entropy is simply constant. This is consistent with the findings of Laddha et al. (2020) for the entanglement entropy of black holes in asymptotic flat space.

However, we have identified a different quantity, what we call the left/right entanglement entropy, which still yields a Page curve for a fixed range of brane tensions. In the analysis of this left/right entanglement, a crucial role is played by a critical angle. As a function of dimension, this angle goes to zero for  $d = 2$  and to  $\frac{\pi}{2}$  for infinitely large dimension, and is approximately 0.98687 for  $d = 4$ . Well below the critical angle, we find a Page curve with initial entropy growth terminated by island formation. At an angle slightly below the critical angle, the Page curve disappears.

In all previous work on massive gravitons in KR brane worlds, the physics has always been found to depend smoothly on the angle and thus the graviton mass. Here there are two dominant effects as we increase the angle: the UV cutoff decreases and the graviton mass increases, with the latter due to quantum CFT effects. For both of these phenomena, the transition has been gradual. However by studying the entanglement entropy, we have found the first quantity (near the critical angle) which displays a sharp phase transition at a fixed value of the mass. Gravity appears to be very different when we cross the critical angle, and it will be very interesting to better understand this phase transition in future work.

We now briefly speculate on this point. One possible answer may be found by looking at the dual field theory description. Here, there are two competing quantities: the total Hilbert space dimension and the initial entanglement entropy at  $t = 0$ . If the Hilbert space dimension is small enough such that the initial entanglement entropy saturates its maximal value, then there is no room for the entanglement entropy to

grow. Hence, we would have a constant curve for all time. On the other hand, if the Hilbert space dimension is large enough such that the initial entanglement entropy is relatively small, then there is still room for it to grow. Therefore, due to the fast scrambling dynamics of holographic theories Maldacena and Susskind (2013); Shenker and Stanford (2014); Geng (2020), we get a Page curve in which the entanglement entropy initially grows, then saturates its maximal bound.

If we consider  $\text{AdS}_d$  gravity coupled to a CFT, both living on the KR branes, it is well known that the boundary conditions encoded by the angle lead to an angle-dependent UV cutoff and graviton mass Karch and Randall (2001b). At generic angles, the mass is of order the inverse  $\text{AdS}_d$  curvature radius on the brane, but at small angles specifically, the mass is parametrically lighter. Therefore the UV cutoff decreases whereas the mass increases as we go to larger angle, perhaps decreasing the size of the Hilbert space to below the required amount for a Page curve transition. It is natural to think that the tension of the brane encodes the Hilbert space dimension of the  $(d - 1)$ -dimensional conformal field theory.<sup>16</sup> Hence, since a Page curve requires a sufficiently large Hilbert space dimension, this implies the existence of a critical angle  $\theta_c$  as described above. To see the Page curve for left/right entanglement entropy in the two-brane configuration, both the left system and right systems need to be large enough. This might explain why we would require both  $\theta_1 < \theta_c$  and  $\theta_2 < \theta_c$ .

One crucial take-home lesson from our work is that the coupling to a non-gravitating bath really is an important ingredient in the recent Page curve calculations. The bath is not just a spectator—it influences the physics. Once we make the bath itself gravitating, although we recover a massless graviton, the original Page curve vanishes. Nevertheless, different quantities can be defined that still sometimes lead to a Page curve, even though its interpretation is unclear. This Page curve is not *the* Page curve of an evaporating black hole. Rather, it indicates the time-dependence

---

<sup>16</sup>A straightforward example is the tensionless probe brane  $\theta = \pi/2$  studied in Geng and Karch (2020b), where the  $(d - 1)$ -dimensional dual is empty.

of the entanglement entropy between two parts of a holographic system, even though the system as a whole is in thermal equilibrium.

## Appendix 2.A: Hartman-Maldacena Area in the Black String

We review the analysis of Hartman and Maldacena (2013) and generalize it slightly to suit our work. We are interested in Hartman-Maldacena minimal surfaces in the  $\text{AdS}_{d+1}$  black string geometry bounded by two branes. The branes are  $d$ -dimensional, and the metric is given by (2.27).

Let us consider a surface that travels at a constant value of  $\mu = \frac{\pi}{2}$  from  $t = t_{\text{diff}}$  at the original asymptotic boundary to the same value of  $t$  at the asymptotic boundary in the thermofield double. This surface can be parameterized by a function  $u(t)$ . Defining,

$$L \equiv \frac{1}{u^{d-1}} \sqrt{-h(u) + \frac{\dot{u}^2}{h(u)}}, \quad (2.67)$$

the area that we need to compute is given by the extremum of the action,

$$\mathcal{A} = \frac{1}{(\sin \mu)^{d-1}} \int dt L. \quad (2.68)$$

Since the Lagrangian is not explicitly dependent on  $t$ , we obtain a quantity that is conserved along any solution of the equations of motion. We call this  $C$ .

$$C = \dot{u} \frac{\partial L}{\partial \dot{u}} - L = \frac{h(u)u^{1-d}}{\sqrt{\frac{\dot{u}^2}{h(u)} - h(u)}}. \quad (2.69)$$

We need to be careful about the *sign* of  $\dot{u}$ , as it changes across the horizon. Solving the equation above tells us,

$$\dot{u} = \begin{cases} +\frac{h(u)}{C} \sqrt{C^2 + u^{2(1-d)}h(u)} & \text{if } u < u_h, \\ -\frac{h(u)}{C} \sqrt{C^2 + u^{2(1-d)}h(u)} & \text{if } u > u_h. \end{cases} \quad (2.70)$$



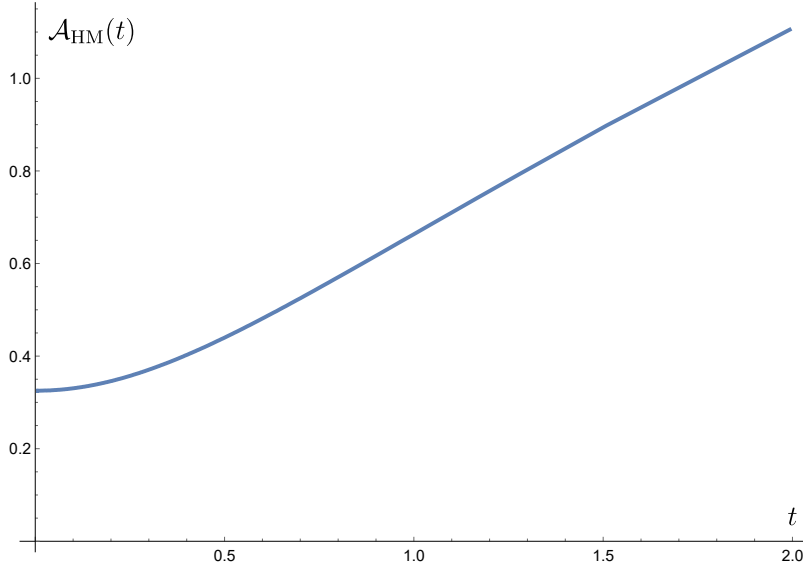


Figure 2.10: The minimum area versus time for the black string geometry with  $u_h = 1$  and  $d = 4$ .

We can use the symmetry of the system to integrate the area up to the point where  $\dot{u} = 0$ . This happens at a value  $u_s$  that solves,

$$C^2 = -h(u_s)u_s^{2(1-d)}. \quad (2.71)$$

Note that this point is inside the horizon and so  $h(u_s) < 0$  at this point. The minimal trajectory  $u(t)$  is symmetric about this point. Recall that the  $t$  variable we have displayed here is the *Schwarzschild time*. We are looking for a surface that reaches the thermofield double CFT boundary at the same value of the *CFT time*, which is related to the Schwarzschild time by a minus sign.

We cannot choose the constant  $C$  arbitrarily. For example, denote the maximum value of the right-hand side of (2.71) for  $u \in [u_h, \infty)$  by  $C_{\max}^2$ . If we take  $C > C_{\max}$ , then (2.71) ceases to have a solution.

We can also fix  $\sin \mu = 1$ . The minimal area, which is denoted by  $\mathcal{A}_{\text{HM}}$  in

section 2.3.2.4, is then just given by,

$$\mathcal{A}_{HM}(t_{\text{diff}}) = 2 \lim_{\epsilon \rightarrow 0} \left[ \frac{-1}{(d-2)\epsilon^{d-2}} + \int_{\epsilon}^{u_s} \frac{du}{|\dot{u}|} \frac{1}{u^{d-1}} \sqrt{-h(u) + \frac{\dot{u}^2}{h(u)}} \right], \quad (2.72)$$

where we have integrated up to  $u_s$  and multiplied by 2 using the symmetry of the minimal surface. We have also removed a universal divergent piece near the boundary. In this integral,  $\dot{u}$  is substituted using (2.70). The time-difference between  $u = 0$  and  $u = u_s$  is given by,

$$t_{\text{diff}} = \lim_{\delta \rightarrow 0} \left( \int_0^{u_h - \delta} \frac{du}{\dot{u}} + \int_{u_h + \delta}^{u_s} \frac{du}{\dot{u}} \right). \quad (2.73)$$

We can increase the value of  $t_{\text{diff}}$  by increasing the value of  $C$ . As we take  $C \rightarrow C_{\text{max}}$ , both the minimal area and the time-difference start to increase in an unbounded manner. Asymptotically, we find that  $\mathcal{A}_{HM}(t_{\text{diff}})$  varies linearly with  $t_{\text{diff}}$ , as shown in Figure 2.10 for  $d = 4$ .

This area can be interpreted as one possible contribution to the left/right entanglement entropy. However, the unbounded linear growth above potentially leads to a paradox since the entanglement entropy cannot grow without limit. The resolution to this paradox—as explained in section 2.3.2.4—is that whenever these surfaces contribute to the entanglement entropy at  $t_{\text{diff}} = 0$ , at late times this growth is cut off when the area exceeds the area of the island surface of section 2.3.2.2.

# Chapter 3: Inconsistency of Islands in Theories with Long-Range Gravity

## Abstract

In ordinary gravitational theories, any local bulk operator in an entanglement wedge is accompanied by a long-range gravitational dressing that extends to the asymptotic part of the wedge. Islands are the only known examples of entanglement wedges that are disconnected from the asymptotic region of spacetime. In this paper, we show that the lack of an asymptotic region in islands creates a potential puzzle that involves the gravitational Gauss law, independently of whether or not there is a non-gravitational bath. In a theory with long-range gravity, the energy of an excitation localized to the island can be detected from outside the island, in contradiction with the principle that operators in an entanglement wedge should commute with operators from its complement. In several known examples, we show that this tension is resolved because islands appear in conjunction with a massive graviton. We also derive some additional consistency conditions that must be obeyed by islands in decoupled systems. Our arguments suggest that islands might not constitute consistent entanglement wedges in standard theories of massless gravity where the Gauss law applies.

---

<sup>0</sup>This chapter is based on published work in collaboration with Geng et al. (2022a). All authors contributed to the final version of the manuscript. My contributions included bringing key ideas to the discussions that shaped the conceptual development of the paper. I also contributed to technical

### 3.1 Introduction

Recent literature Penington (2019); Almheiri et al. (2020e, 2019d, 2020b, 2019a); Penington et al. (2019) has presented significant progress in understanding the evaporation of AdS black holes coupled to an auxiliary non-gravitational bath. In these settings, the fine-grained entropy of a part of the bath, called the radiation region, can be computed through an elegant “island rule.” It is further believed that operators that are localized within the island can be reconstructed from operators in the radiation region in the same sense that, in standard AdS/CFT, operators from part of a boundary of AdS can be used to reconstruct operators in the corresponding entanglement wedge. Using these techniques, the entropy of the radiation region has been found to follow a Page curve.

In spacetime dimensions larger than two, precise computations of the Page curve have been performed using a doubly-holographic setup where the AdS black hole and non-gravitational bath are realized through a Karch-Randall brane Karch and Randall (2001b,c) embedded in a higher-dimensional AdS spacetime Almheiri et al. (2019d). (See Rozali et al. (2020); Liu and Vardhan (2021a); Sully et al. (2021); Emparan et al. (2020); Liu and Vardhan (2021b); Ling et al. (2020); Kumar Basak et al. (2020b); Caceres et al. (2021a,b); Deng et al. (2021); Karlsson (2021); Miao (2021); Bachas and Papadopoulos (2021); May and Wakeham (2021); Kawabata et al. (2021a); Bhattacharya et al. (2021b); Kim and Nam (2021); Aalsma and Sybesma (2021); Neuenfeld (2021b); Geng et al. (2021e); Balasubramanian et al. (2021b); Uhlemann (2021b); Neuenfeld (2021a); Kawabata et al. (2021b); Chu et al. (2021); Kruthoff et al. (2021); Akal et al. (2021b); Kumar Basak et al. (2021); Lu and Lin (2021); Omiya and Wei (2021); Manu et al. (2021) for recent related work.) In

---

computations, particularly working with Hao Geng on performing the 3+1 split of the quadratic gravity action and deriving the Hamiltonian constraint. We explored different approaches to obtain the constraint from the quadratic action, investigating its structure and implications for the Gauss law in the presence of a mass term. While we made progress in this direction, the final derivation was ultimately completed by Suvrat Raju after we encountered difficulties in the computation.

this setting, it was pointed out in Geng and Karch (2020b) that the lower-dimensional graviton is always massive.<sup>1</sup> This is a manifestation of a more general phenomenon: when a gravitational theory in AdS is coupled to a non-gravitational bath, the graviton in AdS picks up a mass. Nevertheless, it is sometimes believed that the non-gravitational bath and the massive graviton that appear in such models are merely technicalities, and that the general lessons regarding islands and the Page curve should be applicable to other physical systems including realistic black holes in asymptotically flat space Krishnan (2020); Almheiri et al. (2020c).

However, in previous work Geng et al. (2021b), we pointed out that the non-gravitational bath is not just a spectator but an important participant in the physics. When gravity is dynamical in the bath, as it should be in realistic models of black holes, it was found that the fine-grained entropy of radiation was constant, consistent with the previously obtained results Laddha et al. (2020) that the Page curve of radiation is trivial for black holes in asymptotically flat spaces.<sup>2</sup>

In this paper, we return to the system with a non-gravitational bath and present an argument that suggests the mass of the graviton plays a significant physical role in allowing islands to constitute an entanglement wedge.

The crux of our argument is very simple. In ordinary gravitational theories, there are no local gauge-invariant operators. We note that if one is studying a non-gravitational observable where gravity would be a small perturbation, it is possible to define approximately local observables by choosing gauge that would suffice for such a measurement. However, when studying processes where both gravitational and quantum effects are important, there is no procedure for defining a local observable, perturbatively or otherwise.

---

<sup>1</sup>Specifically, there is a tower of gravitational KK modes whose lightest graviton is massive.

<sup>2</sup>Even in the presence of dynamical gravity, the Page curve may be the answer to appropriate non-gravitational questions Geng et al. (2021b). Also see Ghosh and Krishnan (2021); Geng et al. (2021f); Langhoff and Nomura (2020) and section 4.2 of Laddha et al. (2020) for a discussion of whether coarse-graining the entropy of the radiation in the presence of dynamical gravity may lead to a nontrivial Page curve.

We now restrict our attention to theories with gravity. As we review below, in standard gravitational theories with massless gravitons, every “localized” operator must be dressed in some way to the asymptotic boundary. This dressing is sometimes referred to as a gravitational Wilson line, which terminates at the asymptotic boundary. In ordinary examples of entanglement wedges in AdS/CFT Maldacena (1998b); Gubser et al. (1998); Witten (1998a), as we review in section 3.2, the connected components of the wedge contain a piece of the asymptotic boundary. So one can meaningfully localize operators from such a wedge by dressing them to this part of the boundary. These operators commute with operators from the complement of the wedge when the latter are dressed to the complementary asymptotic region.

However, an island represents a unique type of entanglement wedge that is entirely surrounded by its complement and where the wedge itself does not extend to the asymptotic boundary. When a region is surrounded by its complement, and it is the complement that extends to the asymptotic boundary, it was argued in Laddha et al. (2020); Chowdhury et al. (2020) based on a careful analysis of the gravitational constraints that the state of the region could be completely determined through observations in its complement. A review of this result, termed the “principle of holography of information,” can be found in Raju (2022a). It is apparent that the picture of islands is already in tension with this principle. Nevertheless, in this paper, we will *not* need to invoke the full power of the principle of holography of information. We will demonstrate that simple physical principles suffice to generate the following puzzle for islands.

Consider a simple unitary operator that adds a localized excitation to the island. Since the excitation is confined to the island, which has finite extent, it must have some nonzero energy by the Heisenberg uncertainty principle. In theories with massless gravitons, the energy can be measured from the falloff of the asymptotic metric using the Gauss law. This would imply that any unitary operator that creates an excitation in the island must fail to commute with the metric in the complement

of the island. This is inconsistent with the idea that the algebra of an entanglement wedge should be closed and commute with the algebra of its complement.

This puzzle becomes even more acute if the island under consideration is described by operators from a radiation region in a non-gravitational system, the bath, and its complement is described by operators from the complement of the radiation region. In this setting, operators that act on the island must commute with operators that act on its complement since, in the non-gravitational theory, operators in the radiation region and its complement are spacelike to each other and so commute by microcausality.

In this paper, we argue that one way to address this puzzle is with massive gravity. In fact, the mass of the graviton appears naturally when gravity in AdS is coupled to an external bath Aharony et al. (2003). The reason is simply that, since the stress-tensor on the boundary of AdS is no longer conserved, it picks up an anomalous dimension. This corresponds to a nonzero mass for the graviton in AdS.

This can be studied in the Karch-Randall scenario that is commonly used to study islands in  $d > 2$  dimensions. Here, as mentioned above, an  $\text{AdS}_d$  brane is embedded in an  $\text{AdS}_{d+1}$  black hole spacetime. The radiation region  $R$  is just a part of the non-gravitational boundary of this  $\text{AdS}_{d+1}$ . The entropy of  $R$  can be computed using the standard RT/HRT prescription Ryu and Takayanagi (2006c,d); Hubeny et al. (2007b) and is given by the area of a bulk minimal surface. In some cases, this surface may end on the brane as shown in Figure 3.1.

The entanglement wedge of the radiation region  $R$  in Figure 3.1 is a conventional entanglement wedge. This entanglement wedge takes on the form of an island if one uses the dual  $d$ -dimensional description to obtain a black hole in  $\text{AdS}_d$  coupled to a non-gravitational bath as shown in Figure 3.2. A striking aspect of this  $d$  dimensional description with a non-gravitational bath is that although there is a localized graviton that “locally” generates a lower-dimensional gravitational theory, this lower-dimensional graviton is massive Karch and Randall (2001b); Porrati (2003,

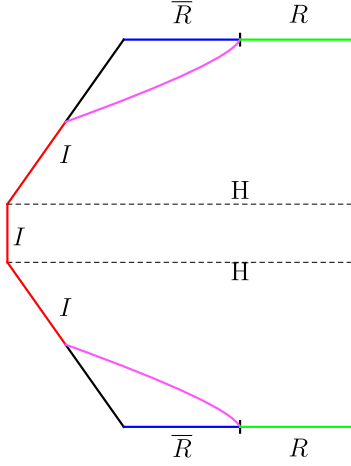


Figure 3.1: A cartoon of a constant-time slice of a black hole with a brane embedded.  $R$  is the union of regions on two asymptotic boundaries and  $\bar{R}$  is its complementary region. The horizons in the bulk are marked by  $H$ . The separation between horizons is meant to convey that the Cauchy slice under examination is a late-time slice on which the wormhole is of a finite length. The dominant  $RT$  surface for the region  $R$  is shown in purple. The region on the brane marked  $I$  becomes the “island” in the lower-dimensional picture of Figure 3.2. In this figure, both the horizontal and the vertical directions are spatial.

2002b); Porrati and Starinets (2002). The constraints in massive gravity are significantly weaker than in massless gravity and do *not* disallow localized excitations, thereby resolving the puzzle presented above.

This is why both pictures of Figure 3.1 and Figure 3.2 are consistent. There is no puzzle in Figure 3.1 because the entanglement wedge extends to the asymptotic region and there is no “island” in the entanglement wedge, i.e. the entanglement wedge has no connected component that is separated from the boundary. There is an “island” in Figure 3.2, but the graviton is massive.

This discussion suggests that the paradigm of islands is inapplicable to standard theories of gravity with massless gravitons.



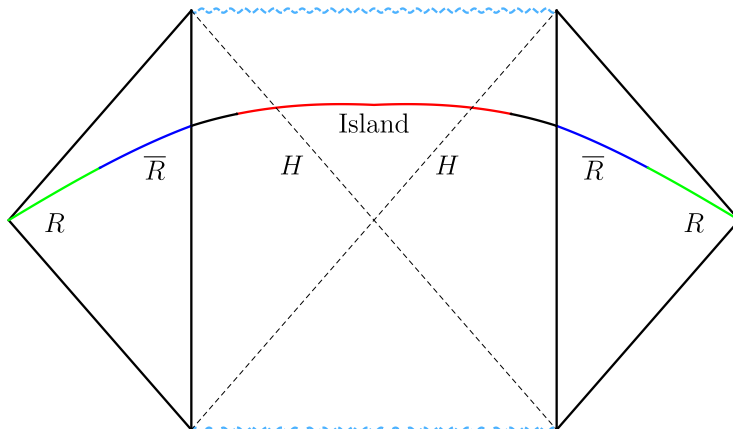


Figure 3.2: A spacetime diagram of the same system of branes in a black hole in the  $d$  dimensional description. The entanglement wedge for the region  $R$  is now an “island”. In this figure, the horizontal direction is spatial and time runs along the vertical direction.

### 3.1.1 Definitions and clarifications

In this paper we will use the phrase “island” strictly in accordance with the following definition.

*Definition.* An island is an entanglement wedge in the gravitating spacetime that does not extend to the asymptotic boundary of the gravitating spacetime.

Accordingly, when “islands” are studied by embedding a brane in a higher-dimensional theory, we use the term “island” only in the lower dimensional dual description. We do *not* use the term “island” to describe the full higher-dimensional entanglement wedge since that does extend to the boundary of the higher-dimensional spacetime. We urge the reader to keep this definition in mind for the rest of this paper, particularly since the term “island” is used more loosely in other parts of the literature.

We also clarify that we use the phrase “entanglement wedge” according to its original definition Headrick et al. (2014) so that it is a region of the gravitating space-

time that can be reconstructed from some part of the non-gravitational spacetime. Some papers in the literature use the phrase “entanglement wedge” to include a part of the non-gravitational spacetime as well, but we will not use this convention.

**Gauss law.** When we refer to the *Gauss law* in this paper, we are referring to the relationship between the total energy on a Cauchy slice and the integral of an appropriate component of the asymptotic metric. This relationship follows from the Hamiltonian constraint in standard theories of gravity. We emphasize that we will use the Gauss law not just as a relationship between expectation values but also inside quantum correlation functions. For clarity, we will always refer to the Hamiltonian constraint as a constraint equation to distinguish it from the Gauss law, as defined above. Note that even in massive gravity, states must obey local constraint equations, which we review below. However these constraints do not lead to a Gauss law.

### **Bulk reconstruction**

In this paper, we adopt the perspective that a consistent entanglement wedge is one where it is possible to reconstruct approximately local bulk operators that, in the limit  $\ell_{\text{pl}} \rightarrow 0$ , reduce to standard quantum-field operators. From its beginning Hamilton et al. (2006) the bulk reconstruction program has sought to reconstruct such operators. So, our perspective aligns with the standard perspective on subregion duality Faulkner and Lewkowycz (2017).

We note that unless one can understand local physics in the bulk from the boundary, it does not make sense to state that a boundary subregion is dual to a bulk subregion. Moreover, the idea that an entanglement wedge can be demarcated by a precisely defined quantum extremal surface presupposes that one can localize bulk operators in the wedge to sub-AdS scales. Finally we note that our perspective is consistent with every known example of subregion duality that has been studied in the literature.

In appendix 3.A we relax the criterion of strict locality. For the reasons outlined above, we do not consider proposals where the only operators that can be reconstructed in the island are infinitely delocalized or spread out over a parametrically large spacetime region. We do however examine (subject to the above restriction) the reconstruction of multi-local operators in the island and show that the proposals suggested so far are also subject to our puzzle.

### **Localization of quantum information in quantum gravity**

The arguments in this paper do *not* contradict the idea that quantum information is localized very differently in theories of gravity than it is in quantum field theories. The conclusion of Laddha et al. (2020), which is consistent with the results of this paper, can be interpreted as the claim that information about a black hole microstate is always available outside for sufficiently detailed measurements in a standard theory of long-range gravity. This could not possibly be true in a local quantum field theory, where spacelike-separated operators commute.

However, islands that lead to a Page curve involve a “halfway” description. The island picture suggests that while operators in what is called the “radiation region” can be used to reconstruct the island, they *cannot* be used to reconstruct the complement of the island which is described by a commuting set of operators. In this halfway description, the unusual localization of information in gravity is important because the radiation region describes degrees of freedom that are in a region that is spacelike separated from it. Nevertheless, the Hilbert space still effectively factorizes as in a local quantum field theory. The puzzles that we present below suggest that this halfway description can be valid only under certain conditions.

Our focus in this paper is on higher-dimensional theories of gravity in which the notion of a “graviton” makes sense. We briefly comment on two-dimensional models in section 3.6. We caution the reader that additional subtleties might arise in two-dimensional theories, and the analysis of section 3.3 and 3.4 is not directly applicable to models of islands in two dimensions.

This paper is organized as follows. In section 3.2, we review conventional entanglement wedges in AdS/CFT and emphasize the importance of the asymptotic region. In section 3.3, we explore the puzzle sketched above, which involves the tension between the Gauss law and the appearance of entanglement wedges that are disconnected from the boundary of the gravitating space time. In section 3.4, we examine the form of the constraints in massive gravity and show how the puzzle is avoided in this setting. We also study islands in doubly holographic settings and show how a dimensional reduction of the higher-dimensional constraints leads to the constraints of a lower-dimensional *massive* theory. In section 3.5, we discuss some additional constraints that are important when islands emerge in decoupled pairs of systems.

### 3.2 Asymptotic regions in entanglement wedges

In this section, we recount some simple properties of entanglement wedges and the algebra of operators associated with them. Here, we focus on conventional entanglement wedges in AdS/CFT; we will turn to islands in later sections. Our objective is to emphasize the significance of the fact that conventional entanglement wedges always have an asymptotic region.

In a holographic theory, the entropy of a boundary region  $R$  (see, for instance, Figure 3.3) is given by Ryu and Takayanagi (2006c,d); Hubeny et al. (2007b); Engelhardt and Wall (2015b); Faulkner et al. (2013b),

$$S(R) = \min \left[ \text{ext} \left( \frac{A(X)}{4G} + S_{\text{bulk}}(E) \right) \right]. \quad (3.1)$$

Here  $X$  is a surface that is homologous to  $R$ ,  $E$  is the region bounded by  $X$  and  $R$  and  $S_{\text{bulk}}(E)$  is the bulk entropy of the region  $E$  computed while *ignoring* gravitational interactions. The homology constraint Headrick et al. (2014); Headrick (2014) on  $X$  states that  $E$  should have *no other* boundaries except for  $X$  and  $R$ . The bulk causal diamond constructed on the region  $E$  is called the entanglement wedge

of  $R$ . Since we will be interested in bulk and boundary regions on a single Cauchy slice, we simply refer to  $E$  as the entanglement wedge and to its complement by  $\overline{E}$ .

The “subregion duality” proposal Jafferis et al. (2016); Dong et al. (2016a) states that boundary operators in  $R$  are dual to bulk operators in  $E$ . This can be made precise as follows. The boundary theory is non-gravitational and so one can associate an *algebra* of operators  $\mathcal{A}(R)$  with the region  $R$ . This algebra comprises all boundary operators that are localized within  $R$  and so, by construction, it is closed under products, linear combinations, and Hermitian conjugation Haag (1992). One can similarly associate an algebra with the complementary region,  $\overline{R}$ , and we denote this algebra by  $\overline{\mathcal{A}(R)}$ .<sup>3</sup> Operators in  $\mathcal{A}(R)$  and  $\overline{\mathcal{A}(R)}$  commute,

$$[A_1, A_2] = 0, \quad \forall A_1 \in \mathcal{A}(R), A_2 \in \overline{\mathcal{A}(R)}, \quad (3.2)$$

since every point in  $R$  is separated by a spacelike interval from every point in  $\overline{R}$ . So equation (3.2) follows from microcausality in the boundary theory. The subregion duality proposal is then that it is possible to find a representation of bulk operators associated with  $E$  within  $\mathcal{A}(R)$  and a representation of bulk operators associated with  $\overline{E}$  within  $\overline{\mathcal{A}(R)}$ .

The subregion duality proposal involves a subtle point that is sometimes glossed over. This aspect of the proposal can be illustrated by considering the case where the bulk theory has a gauge symmetry and charged matter. We caution the reader in advance that there are also important differences between gauge theories and gravitational theories that we will mention below, and so the gauge-theory discussion is provided only as a simplified warm-up.

When there is a gauge symmetry in the bulk, we expect to have a corresponding global symmetry in the boundary theory. We denote the generator of this global symmetry by  $Q$  and we expect that it is given by the integral of a local boundary

---

<sup>3</sup>When the boundary theory is a gauge theory, these algebras have a center Casini et al. (2014); Ghosh et al. (2015). However, this issue will be unimportant for the discussion in this paper.

current  $J$ .

$$Q = \int_R J + \int_{\bar{R}} J. \quad (3.3)$$

Consider a point in the entanglement wedge which we denote by  $P$ , and consider the operator that probes the charged matter field at the point  $P$  which we denote by  $\phi(P)$ . By itself,  $\phi(P)$  is clearly not gauge-invariant. One way to make it gauge-invariant is to attach a Wilson line  $W(P, P_B)$  to this operator that extends from  $P$  in the bulk to another point  $P_B$  on the asymptotic boundary. We now have a gauge-invariant operator that, nevertheless, transforms nontrivially under the global charge.

$$[Q, W(P, P_B)\phi(P)] = W(P, P_B)\phi(P), \quad (3.4)$$

where we have normalized the charge to unity for simplicity. Note that although the operator transforms under the global charge, it is an allowed operator in the theory because it is invariant under “small” gauge transformations that die off at the asymptotic boundary.

There is no unique choice of the path  $P$  to  $P_B$  and not even of the boundary point  $P_B$  itself. But if we want to represent the operator  $W(P, P_B)\phi(P)$  as an element of  $\mathcal{A}(R)$ , then the subregion duality proposal (3.2) implies that it is necessary to choose  $P_B \in R$  and also ensure that the path between  $P$  to  $P_B$  lies entirely within  $E$ . With this choice we have,

$$\left[ \int_R J, W(P, P_B)\phi(P) \right] = W(P, P_B)\phi(P); \quad \left[ \int_{\bar{R}} J, W(P, P_B)\phi(P) \right] = 0, \quad (3.5)$$

in accordance with (3.2).

There are other options for making the operator  $\phi(P)$  gauge-invariant. For example, one can simply fix the gauge. But the gauge-fixed operator must still obey (3.5) for it to be an element of  $\mathcal{A}(R)$ . This leads to the following simple but robust conclusion.

*Observation.* A charged operator in an entanglement wedge can be represented as an operator in the dual boundary region only when it is dressed to that boundary region.

We now turn to the analogous phenomenon in gravity. The difference between gauge theories and gravity is as follows. Gauge theories contain both positive and negative charges. Consequently, gauge theories contain an infinite number of exactly local gauge-invariant operators. An example of such an operator is a small Wilson loop that is entirely localized within a region. Similarly, in the example above, it is possible to construct a gauge-invariant localized operator entirely within an entanglement wedge by considering another operator of the opposite charge  $\phi^*(\tilde{P})$  and connecting the two with a Wilson line:  $\phi(P)W(P, \tilde{P})\phi^*(\tilde{P})$ . But in gravity, there are no “negative charges,” so the gravitational dressing must extend to infinity and cannot terminate in the bulk.

A second way to understand the same physical fact is as follows. If an operator could be localized to a finite region, it would have nonzero energy just by the Heisenberg uncertainty principle. But since the energy can be measured near infinity in ordinary theories of gravity by the Gauss law, this operator cannot commute with the metric near infinity. This is a sign of the fact that even what may appear to be a “local operator” in gravity is secretly delocalized and must extend to the asymptotic boundary DeWitt (1963); Kuchar (1991); Giddings et al. (2006); Jacobson (2013); Jacobson and Nguyen (2019).

In the context of subregion duality, if we want bulk operators in  $E$  to be dual to operators in  $R$  then they must be dressed to  $R$ . At leading order, this dressing can be described as follows. To make the operator  $\phi(P)$  invariant under the gauge transformations of the theory, which comprise small diffeomorphisms—those that vanish near the asymptotic boundary—the position of the point  $P$  is specified by relating it to a part of the asymptotic boundary. In the semiclassical approximation this can be done by specifying the point  $P$  to be the endpoint of a geodesic that starts at some point in the region  $R$  and has a certain renormalized proper length. This picture is not very precise when fluctuations of the metric are themselves important. But, at an intuitive level, this relational prescription can be thought of as the analogue of a Wilson line that must be attached to charged local operators in gauge theories.

Say that the operator  $\phi(P)$  has been specified in a diffeomorphism-invariant manner as described above. Then the analogue of (3.4) in gravity is that this operator transforms nontrivially under the boundary Hamiltonian.

$$[H, \phi(P)] = -i \frac{\partial \phi(P)}{\partial t}. \quad (3.6)$$

In this equation, the coordinate  $t$  involves an extension of the boundary time coordinate into the bulk. By dressing the operator  $\phi(P)$  in different ways, it is possible to choose different  $t$  coordinates in the bulk and so the commutator of the boundary Hamiltonian with the bulk operator depends on the dressing. Note that the reason this is analogous to (3.4) is that if we Fourier transform,

$$\phi(P) = \int_{-\infty}^{\infty} \phi_{\omega} e^{i\omega t} d\omega, \quad (3.7)$$

then (3.6) tells us that the boundary Hamiltonian measures the energy of the Fourier components of  $\phi(P)$ ,

$$[H, \phi_{\omega}] = \omega \phi_{\omega}. \quad (3.8)$$

The boundary Hamiltonian can also be written as the integral of a local current density that is the sum of a term in  $R$  and another in  $\bar{R}$ .

$$H = \int_R T_{00} + \int_{\bar{R}} T_{00}. \quad (3.9)$$

Thus if we want the operator in the entanglement wedge  $E$  to have a representation in the boundary region  $R$ , then the gravitational dressing must be chosen so that,

$$\left[ \int_R T_{00}, \phi(P) \right] = -i \frac{\partial \phi(P)}{\partial t}; \quad \left[ \int_{\bar{R}} T_{00}, \phi(P) \right] = 0. \quad (3.10)$$

The extrapolate dictionary Banks et al. (1998) tells us that the boundary Hamiltonian is itself obtained as the limit of the bulk metric fluctuation in a certain gauge. Let  $g_{\mu\nu}^{\text{AdS}}$  denote the background AdS metric with radial coordinate  $r$ . If the bulk metric is expanded as  $g_{\mu\nu}^{\text{AdS}} + h_{\mu\nu}$ , where  $h_{\mu\nu}$  is the deviation from AdS, then



upon choosing Fefferman-Graham gauge near the boundary  $r \rightarrow \infty$ , i.e.  $h_{r\mu} = 0$ , the extrapolate dictionary reads de Haro et al. (2001),

$$T_{00} = \frac{d}{16\pi G} \lim_{r \rightarrow \infty} r^{d-2} h_{00}. \quad (3.11)$$

We will provide a covariant version of this formula below. For now, we just note that the integral of equation (3.11) on the boundary of AdS provides the definition of the energy of the bulk state in a theory of gravity. Therefore (3.11) is just a manifestation of the Gauss law in the bulk since it tells us that the integral of the boundary metric fluctuation measures the energy of the state in the bulk.

The choice of dressing that ensures that equation (3.10) holds also ensures that the operator  $\phi(P)$  commutes with the metric fluctuation near the boundary in the region  $\overline{E}$ . This is consistent with the idea that operators in an entanglement wedge should commute with operators in its complement.

We can summarize this discussion in terms of the following observation.

*Observation.* Any bulk operator in an ordinary gravitational theory must be dressed to asymptotic infinity to make it invariant under small diffeomorphisms. The asymptotic part of an entanglement wedge provides a base that can be used to define relational observables in the bulk of the wedge.

We should clarify that, for the purpose of this discussion, we need only the relationship (3.11) and the commutator (3.6) to hold within low-point correlators, or the “code subspace” Almheiri et al. (2015). The gravitational dressing is unimportant if we consider the limit where  $G \rightarrow 0$ , but it is already significant at leading nontrivial order in the gravitational constant. In fact, this nonzero commutator of bulk operators with the boundary Hamiltonian was emphasized when the concept of a code subspace was first introduced in the literature by considering small fluctuations about black hole microstates and termed the “little Hilbert space” in Papadodimas and Raju (2014b). In particular, even if (3.11) and (3.6) receive corrections at higher orders

in the gravitational constant or nonperturbative corrections, such corrections are not relevant for the discussion in this paper.

Second, this discussion has interesting consequences when we consider points that belong to two different entanglement wedges. In Figure 3.3 we show a point  $P$  that belongs to the entanglement wedge of the region  $R_1$  and *also* to the entanglement wedge of the region  $R_2$ . When one implements the subregion duality proposal for region  $R_1$ , one picks an operator  $\phi_1(P)$  that satisfies,

$$\left[ \int_{R_1} T_{00}, \phi_1(P) \right] = -i \frac{\partial \phi_1(P)}{\partial t}; \quad \left[ \int_{R_1} T_{00}, \phi_1(P) \right] = 0. \quad (3.12)$$

When one implements the subregion duality proposal for region  $R_2$  one must pick an operator,  $\phi_2(P)$ , that satisfies

$$\left[ \int_{R_2} T_{00}, \phi_2(P) \right] = -i \frac{\partial \phi_2(P)}{\partial t}; \quad \left[ \int_{R_2} T_{00}, \phi_2(P) \right] = 0. \quad (3.13)$$

Note that some operators may satisfy both equations (3.12) and (3.13) but the point  $P$  belongs to an infinite number of entanglement wedges and it is not possible to find a single operator that can serve as the bulk dual in all entanglement wedges. So it is important that one has some freedom in how to dress the bulk operator, and this freedom can be used to “move around” the commutator of the bulk operator with the asymptotic metric and the boundary stress tensor in order to ensure consistency with the subregion duality proposal. Nevertheless, whatever choice one makes for the dressing, there is always a nonzero commutator with the boundary Hamiltonian that is the integral of a component of the asymptotic metric on the entire boundary. In the example above this can be seen from the fact that both equations (3.12) and (3.13) lead to equation (3.6).

### 3.3 A puzzle with islands

We now turn to islands and describe our puzzle. Islands can be understood as follows. We consider a  $\text{CFT}_{\tilde{d}+1}$  propagating in a gravitating  $\text{AdS}_{\tilde{d}+1}$  geometry. We

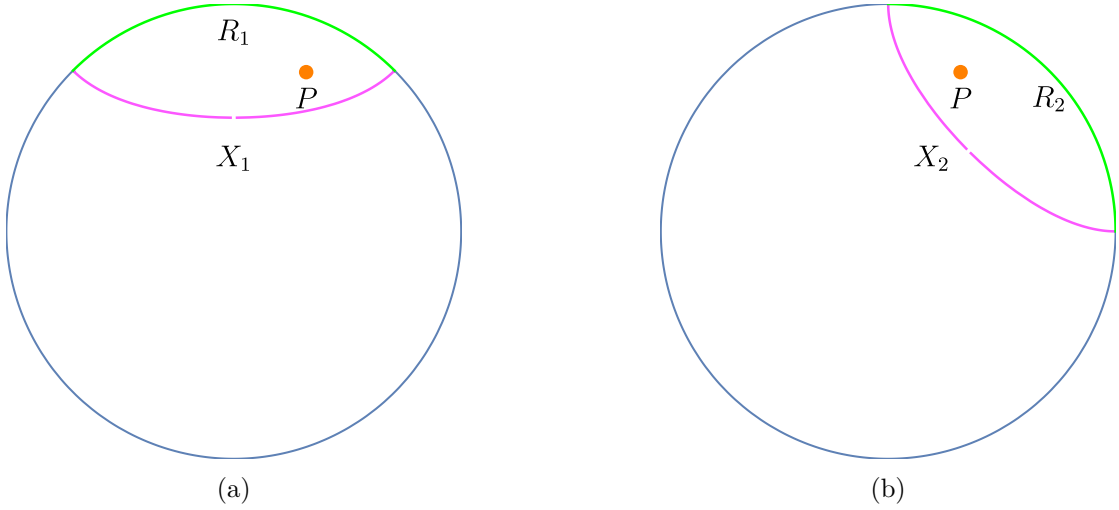


Figure 3.3: A point  $P$  that is part of multiple entanglement wedges. The quasilocal bulk operator must be dressed to  $R_1$  on the left and to  $R_2$  on the right. The figure shows a time slice of global  $AdS$ .

couple this gravitational system to another system where the  $CFT_{\tilde{d}+1}$  is supported in a non-gravitational spacetime. The coupling is designed to lead to “transparent boundary conditions” so that excitations in the CFT can propagate freely from the gravitating to the non-gravitational system.<sup>4</sup> The island rule Penington (2019); Almheiri et al. (2019c) then provides a method of deriving the entanglement entropy of a region,  $R$ , in the non-gravitational system. The island rule is that this entropy is obtained by extremizing,

$$S(R) = \min \left[ \text{ext} \left( \frac{A(\partial I)}{4G} + S_{\text{bulk}}(I \cup R) \right) \right], \quad (3.14)$$

where  $I$  is a part of the gravitating system. The natural extension of the subregion duality proposal suggests that operators in  $R$  can describe the physics of  $I$ .

The formula (8.3) has been carefully derived in JT gravity using a replica trick Almheiri et al. (2020b). In other settings, the formula has been justified but again when  $R$  is in the non-gravitational region Bousso and Wildenhain (2020). In

---

<sup>4</sup>We use  $\tilde{d}$  in this section since in section 3.4 we will study islands that are realized on an  $AdS_d$  brane embedded in  $AdS_{d+1}$ , in which case  $\tilde{d} = d - 1$ .

some parts of the literature (8.3) is directly applied even when  $R$  is in a region with dynamical gravity. It has already been pointed out in Geng et al. (2021b); Laddha et al. (2020) that since the entanglement entropy is a fine-grained quantity, even the presence of weak gravity can alter its magnitude. Therefore (8.3) is not directly applicable to settings where gravity is dynamical everywhere. The puzzle that we describe below adds additional evidence for this claim.

The puzzle arises from the following simple observation.

*Observation.* Islands are the only known example of entanglement wedges that do not extend to the asymptotic boundary of the gravitating spacetime. They are disconnected from the region  $R$  in the sense that not even a spatial geodesic from  $I$  can reach  $R$  without passing through the complement  $\bar{I}$  of the island.

Here we are using  $\bar{I}$  to denote the complement of the island in the gravitating system to be consistent with the notation used above.

In light of the discussion of section 3.2, this leads to a puzzle because in an ordinary theory of gravity, there is no way to dress operators in  $I$  without making reference to operators in  $\bar{I}$ . This is an obstacle to making operators in  $I$  invariant under diffeomorphisms. This puzzle can be made sharp as follows.

Let  $\phi(P)$  be a Hermitian scalar field operator that probes physics at a point  $P \in I$ . Consider the unitary operator  $U = e^{i\lambda\phi(P)}$  where  $\lambda$  is a small parameter introduced for convenience below. Then since  $\phi(P)$  is described by an operator in  $R$ , the unitary operator  $U$  should leave the expectation value of all operators in the algebra  $\overline{\mathcal{A}(R)}$  unchanged. This can be seen as follows. Let  $|\Psi\rangle$  denote the state of the entire system, including the bath. Let  $A_{\bar{R}} \in \overline{\mathcal{A}(R)}$ . We then expect to have,

$$\langle\Psi|U^\dagger A_{\bar{R}} U|\Psi\rangle = \langle\Psi|A_{\bar{R}} U^\dagger U|\Psi\rangle = \langle\Psi|A_{\bar{R}}|\Psi\rangle, \quad (3.15)$$

where we have used the commutator (3.2) and the unitarity of  $U$ . Since operators in  $\bar{I}$  are dual to operators in  $\bar{R}$ , this means that the action of  $U$  should also leave the expectation value of all operators in  $\bar{I}$  unchanged.

But, in an ordinary theory of gravity, the Gauss law tells us that the asymptotic metric near the boundary of AdS (which is in  $\bar{I}$ ) measures the energy of the bulk, which includes  $I$ . Let  $A_{\bar{I}}$  be another simple operator that acts on the complement of the island. Then, using equation (3.11) and the Gauss law we find that,

$$\begin{aligned}
\lim_{r \rightarrow \infty} r^{\tilde{d}-2} \int_{(\partial \text{AdS})} \frac{\partial}{\partial \lambda} \langle \Psi | U A_{\bar{I}} h_{00} U^\dagger | \Psi \rangle \Big|_{\lambda=0} &= \lim_{r \rightarrow \infty} r^{\tilde{d}-2} \int_{\partial(\text{AdS})} \frac{\partial}{\partial \lambda} \langle \Psi | A_{\bar{I}} U h_{00} U^\dagger | \Psi \rangle \Big|_{\lambda=0} \\
&= \lim_{r \rightarrow \infty} r^{\tilde{d}-2} \int_{\partial(\text{AdS})} \langle \Psi | A_{\bar{I}} [\phi(P), h_{00}] | \Psi \rangle \\
&= \frac{-16\pi G}{\tilde{d}} \langle \Psi | A_{\bar{I}} \frac{\partial \phi(P)}{\partial t} | \Psi \rangle,
\end{aligned} \tag{3.16}$$

which is different from (3.15).

Physically, equation (3.16) has a simple interpretation. The unitary  $U$  inserts a small excitation in the region  $I$ . The metric at infinity should be able to measure the energy of this excitation. Note that (3.16) involves an insertion of the gravitational constant and so this commutator appears at leading nontrivial order in perturbation theory.

Observe that if one takes the nongravitational limit, the right hand side of (3.16) vanishes. This is why it is possible to discuss local measurements when gravity can be neglected.

**Black holes.** We would like to make a few important comments about islands in the presence of black holes. First note that (3.16) is not just about the *expectation value* of the energy. When there is a black hole in the bulk, one might attempt to consider operators that somehow “extract” energy from inside the black hole and “insert” it in some part of the island that is outside the black hole. However, even such an operator would have to change the distribution of energy in the island, so it would not commute with the metric near infinity. This nonzero commutator can be detected by the insertion of an appropriate operator  $A_{\bar{I}}$  in the correlator (3.16).

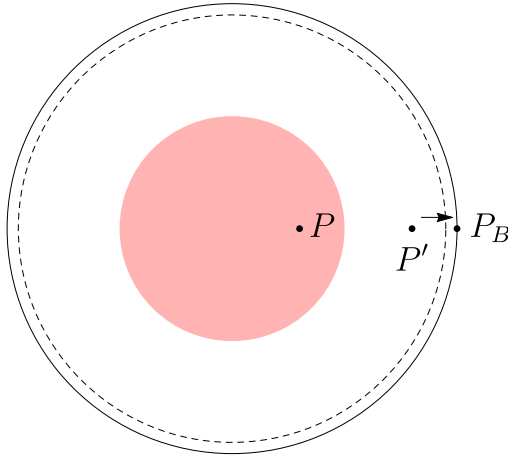


Figure 3.4: An excitation at the point  $P$  inside the island (pink shaded disk) can be detected using a two-point correlator outside the island. The two-point correlator involves an integral of the asymptotic metric (indicated by the dashed line) and another operator obtained by taking the limit of point  $P'$  to  $P_B$  on the boundary of the gravitational region. The bath is not shown in this figure, which shows a time slice of AdS.

Let us consider a more explicit example. Let the state  $|\Psi\rangle$  correspond to a black hole that has equilibrated with a bath at the same temperature. Assume that the field  $\phi$  describes a scalar excitation of mass  $\mu$  and let  $\Delta = \frac{\tilde{d}}{2} + \sqrt{\mu^2 + \frac{\tilde{d}^2}{4}}$ . Then an explicit choice of  $A_{\bar{I}}$  that leads to a nonzero value for the correlator in equation (3.16) is simply,

$$A_{\bar{I}} = \lim_{P' \rightarrow P_B} r^{\Delta} \frac{\partial \phi(P')}{\partial t}, \quad (3.17)$$

where  $P'$  is a point in  $\bar{I}$  with radial coordinate  $r$  that is taken to a point on the boundary of the gravitational region  $P_B$  and scaled up to yield a finite operator. (See Figure 3.4.) In the absence of the unitary operator in equation (3.16) we find that,

$$\lim_{r \rightarrow \infty} r^{\tilde{d}-2} \int_{\partial \text{AdS}} \langle \Psi | A_{\bar{I}} h_{00} | \Psi \rangle = 0. \quad (3.18)$$

This equation follows because in an equilibrium state, there is no preferred direction of time and so correlators of the Hamiltonian and a time-derivative of a bulk scalar field vanish; even the fluctuations of such a correlator are exponentially

small. On the other hand, in the presence of the unitary operator we see that the final result in equation (3.16) becomes a two-point *Wightman function* of the time-derivative of the field—with one insertion inside the island and another insertion near the boundary. This Wightman function does not vanish even when one point is inside the horizon and the other is outside the horizon.

We emphasize that our objective here is not to completely reconstruct the state  $|\Psi\rangle$ , which would require exponential precision. It is only to show that *some* simple excitations, including those produced by unitary operators involving bulk fields, can be detected from outside the island. These simple excitations should be in the “little Hilbert space” or “code subspace” built about the state  $|\Psi\rangle$ .

Moreover, since we are considering only simple excitations it is safe to evaluate the correlator in (3.16) in the black-hole background. The simple correlators that appear in (3.16) do not receive significant contributions from nonperturbatively suppressed branches of the wavefunction, which might be important for the computation of very high-point correlators. In particular, the perturbative calculation leading to (3.16) cannot be invalidated by considering exotic configurations where  $\phi(P)$  can be dressed to the non-gravitational bath using wormholes that bypass the complement of the island.<sup>5</sup>

All of the comments above hold if  $|\Psi\rangle$  corresponds to a state of an eternal black hole coupled to a bath at the same temperature. The only subtlety is that since the gravitational part of the geometry has two asymptotic regions the boundary of AdS in (3.16) should be interpreted as the union of the two asymptotic boundaries.

One might wonder if the integrated asymptotic metric is the only operator that causes a problem in interpreting the island as an entanglement wedge. This is not the case. Even if one attempts to “discard” the asymptotic metric from the algebra

---

<sup>5</sup>In the next section, we will see that when the island and the bath are realized in a higher-dimensional doubly-holographic setting,  $\phi(P)$  can be dressed to the non-gravitational bath through the higher dimension. But, in this setting, the  $\text{AdS}_{\tilde{d}+1}$  theory of gravity is massive.

of operators in  $\bar{I}$  at leading order, it will reappear in the algebra at subleading order El-Showk and Papadodimas (2012). This is simply due to the fact that the OPE of boundary operators produces the boundary stress tensor. In the bulk, this means that the algebra of other asymptotic operators produces the asymptotic metric. So, for consistency with (3.16), the operator  $U$  must fail to commute with other operators in  $\bar{I}$  although such commutators may appear only at subleading orders in perturbation theory.

For the reader who would like additional details, we note that the detection of simple excitations about black holes has been studied previously in the literature. In Papadodimas and Raju (2014b) (see section 5) a class of excitations of black holes leading to “near-equilibrium” states was studied and it was shown how they could be detected. Excitations of the island in the region outside the black hole horizon correspond to the “near-equilibrium” states of Papadodimas and Raju (2014b). In Papadodimas and Raju (2016) (see section 8.1) and also Papadodimas (2017), a more general class of excitations of the black-hole interior was studied and it was shown how correlators of the Hamiltonian and other operators could be used to detect them as well.

We have therefore arrived at the following puzzle.

*Puzzle.* The Gauss law suggests that the action of any operator in the island must be accompanied by a disturbance in the metric outside the island. This is in contradiction with the idea that operators in the island are described by operators in  $R$  and commute with operators in the complement of the island that are described by operators in  $\bar{R}$ .

We remind the reader that we use the phrases “island” and “Gauss law” in the precise sense described in section 3.1.1.

Note that our puzzle pertains to whether the islands can constitute consistent entanglement wedges. If the island rule (8.3) is used merely as a trick to compute the entropy without a concomitant claim that the island is the entanglement wedge of the radiation region, then our puzzle would not apply. However, we do not consider



this possibility further since the proof of subregion duality follows directly from the entropy formula Jafferis et al. (2016); Dong et al. (2016a), and so the computation of the entropy and the determination of the entanglement wedge cannot usually be separated.

We also note that the mere existence of solutions to (8.3) in standard theories of long-range gravity Ghosh and Krishnan (2021); Krishnan (2020); Hollowood et al. (2021b); Wang et al. (2021a,b) cannot be used to conclude that such theories must exhibit a Page curve or have islands as entanglement wedges. The physical interpretation of such solutions is unclear since the island rule has not been justified in standard theories of gravity. In particular, the Gauss-law puzzle above implies that even if an island is obtained as a geometric solution to a minimization problem in a theory of long-range gravity, it does not constitute a consistent entanglement wedge.

We have formulated a puzzle above using the gravitational Gauss law. A similar puzzle can be formulated using the Gauss law in gauge theories. The action of a charged operator in the island must be accompanied by a change in the gauge field outside the island. The puzzle in gauge theories is less acute than it is in gravity since gauge theories contain local gauge-invariant operators and there is no obstacle to localizing such operators in the island. In standard theories of gravity, as we have already explained, there are no local gauge-invariant operators.

### 3.4 A resolution using massive gravity

In this section we describe how the puzzle of section 3.3 can be resolved in the setting of massive gravity. We will argue that the puzzle of section 3.3 does not appear in this scenario.

Because there are few well-understood examples of massive gravity and because it has naturally occurred in the context of entropy calculations in higher dimensions, our starting point is the Karch-Randall setup that has been used to study islands in higher dimensions that we reviewed earlier. Here one embeds a  $d$ -dimensional AdS

brane in a  $(d + 1)$ -dimensional AdS bulk. The boundary dual to this geometry is believed to be a  $\text{BCFT}_d$ , a  $\text{CFT}_d$  on a space with a boundary and with conformal boundary conditions for CFT fields as one approaches this boundary Takayanagi (2011b). In addition, the boundary of this half-space can support additional degrees of freedom and is sometimes referred to as a “defect” Gaiotto and Witten (2009). One can consider a thermofield double state of two such BCFTs which is then dual to an eternal black hole with two asymptotic boundaries and a brane that runs from one boundary to the other. (See Figure 3.5.)

It is believed that the correct bulk generalization Almheiri et al. (2019c) of the holographic entanglement entropy prescription is obtained through the following generalization of the homology constraint: the entropy of a region  $R$  on the boundary is given by formula (3.1), where one is allowed to consider all those surfaces  $X$  so that  $E$  has no boundaries except for  $R$ ,  $X$ , and *a possible portion of the brane*. In particular, if one takes  $R$  to be the union of a region on one asymptotic boundary with a similar region on the other asymptotic boundary, then at late enough times, one finds a phase transition for the surface  $X$  between what is called a Hartman-Maldacena surface Hartman and Maldacena (2013), which runs from  $\partial R$  on one boundary to  $\partial R$  on the other boundary, and a second surface that runs from  $\partial R$  to the brane as shown in Figure 3.5.<sup>6</sup>

For this entanglement wedge, it is clear that the puzzle of section 3.3 does not arise. If we consider an operator that acts near the brane as shown in Figure 3.5, then this operator can be dressed to the asymptotic boundary in region  $R$  entirely within the entanglement wedge  $E(R)$  without ever entering its complement. We can ensure that the operator commutes with all operators on  $\overline{R}$  but not that it commutes with operators in  $R$ . On the other hand, as Figure 3.5 shows, the entanglement wedge does not contain an “island” (in the sense of subsection 3.1.1). This entanglement

---

<sup>6</sup>There are technically two copies of this second surface, with each residing in a respective exterior patch of the black hole.



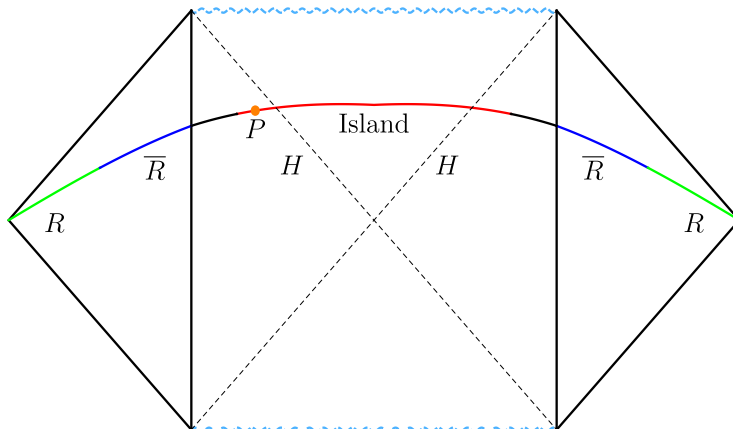


Figure 3.6: *A lower-dimensional picture of the setup of Figure 3.5. This is a spacetime diagram like Figure 3.2. In the lower-dimensional description, an operator at point  $P$  in the island cannot be dressed to the boundary without affecting the region outside the island. So the island cannot constitute a consistent entanglement wedge in a theory where the Gauss law applies. The lower-dimensional description of Figure 3.5 involves massive gravity where the Gauss law does not hold. In this figure, the horizontal direction is spatial and time runs along the vertical direction.*

Since the higher-dimensional picture does not involve any violation of the Gauss law, the lower-dimensional picture must also be consistent. A resolution to the apparent Gauss-law puzzle *must* therefore lie in the details of the dimensional reduction.

A notable aspect of the dimensionally-reduced picture is that the lower-dimensional theory of gravity is always *massive*. This can be understood from the perspective of the gravitational theory on the brane Porrati (2003, 2002b); Porrati and Starinets (2002).

But another simple way to understand the origin of the mass is from the boundary. In the absence of the bath, the theory of gravity on the AdS brane is dual to a conformal field theory with a conserved stress tensor. The coupling to the bath leads to the nonconservation of the stress tensor on the boundary. This allows the stress tensor to pick up an anomalous dimension Aharony et al. (2006) which, in the

bulk, corresponds to a massive graviton. This interpretation is important because it would generalize to any such theory coupled to a bath such as the one analyzed in Penington (2019). The coupling should always generate a mass for a propagating graviton. We will show that this mass will always be present in consistent scenarios that include an island. Note that even if, as is sometimes suggested, the coupling to the bath is turned off after evaporation, islands always appear concomitantly with a mass for the graviton.

As we further elaborate below, the mass of the graviton resolves the apparent Gauss-law puzzle for the simple reason that the constraints of massive gravity cannot be integrated to obtain a Gauss law of the form that exists in the massless theory. Consequently, in a theory of massive gravity, even when a region is surrounded by its complement, it *is possible* to modify the state of the region without modifying the state of its complement. This is an example of how massive gravity can have *qualitatively different* properties from massless gravity.

In flat space, the gravitational force law changes discontinuously as the mass of the graviton goes to zero, and this is known as the vDVZ discontinuity van Dam and Veltman (1970); Zakharov (1970). In AdS, there is no such discontinuity in the gravitational force law. However, there is still a qualitative difference in the way massless and massive theories of gravity *store quantum information*. It is this difference that allows islands to exist in theories of massive gravity but not in theories of massless gravity where the Gauss law applies.

The rest of this section is divided into two parts. First, we provide a simple explanation of our main idea in subsection 3.4.1, which illustrates the difference between massive and massless gravity in flat space. This discussion avoids technical details but is sufficient to understand the point which we wish to emphasize.

Then, in subsection 3.4.2, we explore the form of the gravitational constraints in the higher-dimensional  $\text{AdS}_{d+1}$  theory, focusing on the so-called Hamiltonian constraint. We linearize the Hamiltonian constraint and perform a Kaluza-Klein reduc-

tion of the constraint in a braneworld geometry. We show that these constraints of the  $(d + 1)$ -dimensional bulk theory go over to the constraints of  $d$ -dimensional massive gravity in subsection 3.4.3. We show in subsection 3.4.3 that the mass spectrum obtained in this manner is precisely the known mass spectrum of Karch-Randall braneworlds.

### 3.4.1 Massless vs. massive constraints in flat-space linearized gravity

In this section, we illustrate the difference between massless and massive gravity in a simple setting: linearized gravity in flat space.

Consider a theory of gravity coupled to matter in flat space. We are interested in a state that is close to the vacuum with some matter energy density  $\rho$ . Since we would like to examine the constraints of the theory, we focus on a *spatial slice* at a specific instant of time. We take the spatial components of the metric on this slice to be  $\delta_{ij} + h_{ij}$ . Note that  $i, j$  do not run over the time coordinate. Then this metric fluctuation and the energy density are not entirely uncorrelated. In standard massless gravity, at the linearized level, they must obey the constraint,

$$-\partial_j \partial_j h_{ii} + \partial_j \partial_i h_{ij} = 16\pi G \rho, \quad (3.19)$$

This constraint follows from the linearized  $TT$  component of the Einstein equation, but it can also be derived from a standard canonical analysis in the Hamiltonian formalism.

Integrating this constraint over a volume  $V$  leads to,

$$\frac{1}{16\pi G} \int_B d^{d-1}x \, n_j (\partial_i h_{ij} - \partial_j h_{ii}) = \int_V d^d x \, \rho. \quad (3.20)$$

Here the left integral is performed over the boundary  $B$  of the region  $V$ , and  $n_j$  is the unit normal vector of the boundary. On the right hand side, we have a bulk integral over the entire region  $V$ . This is just the standard Gauss law, used in the sense of section 3.1.1, which relates the total energy to the asymptotic metric.

We note an aspect of equation (3.19) that will be relevant below. It is convenient to decompose the metric perturbation, following ADM Arnowitt et al. (2008); Kuchar (1970), into a “longitudinal” (L) component, a “transverse traceless” (TT) component and what we call a “T” component,

$$h_{ij} = {}_{ij} + \partial_{(i} \epsilon_{j)}, \quad (3.21)$$

where,

$${}_{ij} = \partial_{(i} \epsilon_{j)}, \quad (3.22)$$

for some vector field  $\epsilon_j$  and,

$$\partial_{ii} = \partial_{ij} = 0; \quad {}_{ii} = 0. \quad (3.23)$$

For an explicit decomposition of any metric perturbation into (3.21) see the nice discussion in Kuchar (1991).

Both  ${}_{ij}$ —which corresponds to spatial diffeomorphisms of the slice—and  ${}_{ij}$ —which parameterizes the dynamical graviton—drop out of equation (3.19). Thus, equation (3.19) constrains only the T component of the metric perturbation. We will return to the relevance of this observation when we study the constraints in massive gravity.

We would like to make a few comments.

1. The equation above was derived in the linearized approximation. However, when  $V$  is taken to be an entire Cauchy slice, the term on the left hand side of equation (3.20) turns into the famous ADM Hamiltonian Arnowitt et al. (2008). So when  $B$  is the large- $r$  region of the Cauchy slice, the left hand side of equation (3.20) is the *definition* of the energy, even in the full theory of general relativity.
2. Upon the insertion of an excitation in the bulk, which changes the integral of  $\rho$ , the constraint *forces* a concomitant change in the metric at infinity. So even if the excitation appears to insert energy only in a bounded region, the metric around that region all the way up to infinity must be changed.

3. In the classical theory, it is possible to “move” energy from one spot to another in the bulk while keeping the metric unchanged outside a large ball. This is guaranteed in the classical theory by the Birkhoff theorem and its generalization by Corvino and Schoen Corvino and Schoen (2006). Such a possibility does *not* exist in the quantum theory. In the quantum theory, the boundary Hamiltonian not only knows about the expectation value of the energy but also about *moments* of the energy and of *correlators* of the energy with other observables. This information is enough to ensure that it is impossible to change even the distribution of the energy in the bulk region without affecting at least some correlator of the boundary metric with other boundary degrees of freedom. This was termed the principle of “holography of information” in Raju (2022a).
4. The constraints of the theory hold on a single Cauchy slice, and so they hold even if the spacetime has a horizon, since the horizon arises from the global causal properties of the spacetime. In particular, the constraints are important even in the presence of black holes. The insertion of an excitation in the interior of the black hole still changes the asymptotic Hamiltonian. As mentioned above, it can be shown that simple excitations in the interior can be detected by correlators of the asymptotic Hamiltonian and other operators as described in Papadodimas (2017). (See also section 8 of Papadodimas and Raju (2016).)

We now turn to massive gravity. Even in massive gravity, we find that if one studies a state close to the Minkowski vacuum then the matter energy-density and the metric perturbation are related by the following constraint Hinterbichler (2012):

$$-\partial_j \partial_j h_{ii} + \partial_j \partial_i h_{ij} + m^2 h_{ii} = 16\pi G \rho. \quad (3.24)$$

We note an important difference between equation (3.24) and (3.19). The left hand side of (3.24) is *not* a gradient. Consequently the integral of the energy-density over a volume cannot be expressed in terms of the integral of the boundary metric and its derivatives. This is why the energy of the state is *not* just given by a boundary



term and there is no analogue of the Gauss law (as defined in subsection 3.1.1) in theories of massive gravity.

We emphasize that equation (3.24) should not be thought of simply as a “screened” version of equation (3.19). This is because equation (3.24) now involves  $_{ij}$  from the decomposition (3.21) and is not just a constraint on  $_{ij}$ . The constraint relates the longitudinal mode, which is now an additional degree of freedom, to the other modes of the graviton.

Therefore, at least at this linearized level and when the graviton has a mass, it *is* possible to insert energy at a location while keeping the metric far away unchanged. We simply use the  $m^2 h_{ii}$  term to compensate for the change in  $\rho$ . So, it is possible for energy to “appear” in the middle of a bounded region, i.e. for  $\rho$  to change, without an alteration of the metric at the boundary of that region. Some more discussion of this effect from a phenomenological perspective can be found in Goldhaber and Nieto (2010).

We are not aware of any detailed analysis of quantum wavefunctionals that satisfy the constraint (3.24) or any analysis that carefully accounts for the effects of possible nonlinearities. However, since the Hamiltonian is not a boundary term in massive gravity we do not see any *a priori* obstruction to preparing “split states” in massive gravity: states that differ within a bounded region but are identical outside that region.

This suggests that the puzzle of section 3.3 is removed when the bulk graviton has a mass. The island is surrounded by its complement. But if the bulk theory of gravity is massive, it is possible for degrees of freedom in the island to correspond to degrees of freedom in a disconnected “radiation” region. An insertion of energy, performed via the action of a unitary operator in the radiation region, does not need to modify the metric in the complement of the island.

We now turn to a more detailed investigation of the constraints when the massive graviton is realized through dimensional reduction of a higher-dimensional

gravitational theory.

### 3.4.2 Gravitational constraints for AdS spacetimes with branes

The diffeomorphism-invariance of gravity leads to a set of constraints that must be obeyed by valid wavefunctionals even when we go beyond the linearized approximation. These constraints are commonly divided into what are called the “momentum constraints” and the “Hamiltonian constraint” DeWitt (1967). We review these constraints below in the context of asymptotically AdS spacetimes that support a brane, and we explain how they directly lead to a Gauss law in the higher-dimensional spacetime. We then dimensionally reduce these constraints on the brane and show how the mass of the graviton appears in the dimensionally-reduced constraint. This supports the idea introduced above that the mass is key to the consistency of islands on the brane. The analysis in this section has some overlap with the analysis of Chowdhury et al. (2021) where the reader will find further details.

The analysis of gravitational constraints is commonly performed after a  $d + 1$  split of the metric, so that the line element takes the form,

$$ds^2 = -N^2 dt^2 + \gamma_{ij}(dx^i + N^i dt)(dx^j + N^j dt). \quad (3.25)$$

Here,  $N$  is called the “lapse function,”  $N^i$  is called the “shift vector,” and  $\gamma_{ij}$  is the metric on spatial  $d$ -dimensional slices. In addition, the theory might contain matter fields that we simply denote by  $\phi_{\text{matter}}$  since they will play a limited role in the analysis.

The nontrivial constraint equation in gravity is the so-called Hamiltonian constraint which tells us that any valid state of the theory must satisfy

$$\mathcal{H} = \frac{1}{2\sqrt{\gamma}} \left( \gamma_{ik}\gamma_{jl} + \gamma_{il}\gamma_{jk} - \frac{2}{d-1}\gamma_{ij}\gamma_{kl} \right) \Pi^{ij}\Pi^{kl} - \sqrt{\gamma}R + 2\sqrt{\gamma}\Lambda + 16\pi G\mathcal{H}_{\text{matter}} = 0. \quad (3.26)$$

Here  $\mathcal{H}_{\text{matter}}$  is the Hamiltonian density of the matter sector.  $R$  is the  $d$ -dimensional Ricci scalar and we have included a possible cosmological constant  $\Lambda$ .  $\Pi^{ij}$  is the

momentum conjugate to the metric. In the quantum theory,  $\Pi^{ij}$  is represented as  $-i\frac{\partial}{\partial\gamma_{ij}}$ , but in the classical limit the conjugate momentum is related to the extrinsic curvature of the spatial  $d$ -dimensional slices  $K_{ij}$  via,

$$\Pi^{ij} = -\sqrt{\gamma} \left( K^{ij} - \gamma^{ij} \gamma_{kl} K^{kl} \right). \quad (3.27)$$

The constraint (3.26) must be obeyed by any valid state and in a standard theory of long-range gravity it implies a Gauss law. To illustrate the physics, we show how this Gauss law emerges when the constraint is linearized for a family of simple states. We assume that we are considering a vacuum background solution to (3.26) with,

$$\gamma_{ij} = g_{ij}; \quad \Pi_{ij} = 0; \quad \mathcal{H}_{\text{matter}} = 0; \quad N^i = 0. \quad [\text{Background solution}] \quad (3.28)$$

We will use  $N$  to denote the value of the lapse in this background. Although equation (3.28) describes a simple background, the final expression we will obtain for the energy, as measured from the boundary, will be more general.

We are interested in a nearby solution with a nonzero matter energy-density  $\delta\rho$ . To see how the metric must respond to this energy density, we expand the metric about this background as,

$$\gamma_{ij} = g_{ij} + h_{ij}, \quad (3.29)$$

where  $g_{ij}$  is the background metric and  $h_{ij}$  is the metric fluctuation. After some algebra we find that the linear term in  $h_{ij}$  is of the form

$$\sqrt{\gamma}(R - 2\Lambda) = \sqrt{g} \left[ -h^{ij} R_{ij}^{(g)} + \frac{1}{2} h_i^i R^{(g)} - h_i^i \Lambda + \nabla_i \nabla_j h^{ij} - \nabla^j \nabla_j h_i^i \right] + \dots \quad (3.30)$$

where  $\dots$  indicates higher-order terms,  $\nabla_i$  is the covariant derivative with respect to  $g_{ij}$ ,  $R_{ij}^{(g)}$  and  $R^{(g)}$  are the background Ricci tensor and scalar, and all indices are raised using  $g^{ij}$ .

This may not immediately seem like a total derivative. However, for the background solution (3.28), we have the identity,

$$G_{ij}^{(g)} = R_{ij}^{(g)} - \frac{1}{2}g_{ij}R^{(g)} - \frac{1}{N}(\nabla_i\nabla_j - g_{ij}\nabla^2)N. \quad (3.31)$$

The vacuum Einstein equations are,

$$G_{ij}^{(g)} + \Lambda g_{ij} = 0, \quad (3.32)$$

so therefore we can write (3.30) as,

$$\begin{aligned} \sqrt{\gamma}(R - 2\Lambda) &= \frac{-\sqrt{g}}{N}(\nabla^i\nabla^j N - g^{ij}\nabla^2 N)h_{ij} + \sqrt{g}(\nabla_i\nabla_j h^{ij} - \nabla^j\nabla_j h_i^i) \\ &= \frac{\sqrt{g}}{N}[\nabla_i(N\nabla_j h^{ij}) - \nabla_i(h^{ij}\nabla_j N) - \nabla_i(N\nabla^i h_j^j) + \nabla_i(h_j^j\nabla^i N)]. \end{aligned} \quad (3.33)$$

In the last line above we have also relabeled some dummy indices. The constraint can now be written as,

$$\frac{1}{16\pi G}\sqrt{g}\nabla_i J^i = N\sqrt{g}\delta\rho, \quad (3.34)$$

Here we have defined

$$J^i \equiv N\nabla_j h^{ij} - h^{ij}\nabla_j N - N\nabla^i h_j^j + h_j^j\nabla^i N. \quad (3.35)$$

We now integrate both sides of (3.34) on the entire Cauchy slice. We denote the boundary of the slice by  $S_\infty$  and the unit normal vector to the boundary by  $n_i$ . The left hand side of (3.34) then integrates to a quantity that we denote by  $E$ ,

$$E \equiv \frac{1}{16\pi G} \int_{S_\infty} d^{d-1}x \sqrt{g} n_i J^i, \quad (3.36)$$

and (3.34) reduces to,

$$E = \int d^d x N \sqrt{g} \delta\rho. \quad (3.37)$$

This is the generalization of the Gauss law (3.20) to curved space, which tells us that an insertion of energy density at a point in the bulk *must* be accompanied by a change of the boundary metric in order to satisfy the constraints.

The expression above was derived in the presence of some approximations. In the full theory, we need to account for higher-order terms in the expression (3.26) and also account for the energy of transverse-traceless gravitons themselves. Nevertheless, even in the full theory of general relativity, the boundary integral (3.36) provides the correct definition of the energy. It can be seen that this coincides with the expression of Hawking and Horowitz (1996) and Balasubramanian and Kraus (1999) for energy in asymptotically anti-de Sitter spacetimes and can also be shown to be equivalent to the expression that follows (3.11). We refer the reader to Chowdhury et al. (2021) for details.

Equation (3.36) also provides the correct expression for the energy in the quantum theory. In the quantum theory, denoting the matter fields collectively by  $\phi_{\text{matter}}$ , states are represented by wavefunctionals  $\Psi[\gamma_{ij}, \phi_{\text{matter}}]$ . Valid states must satisfy (3.26) in the sense that,

$$\mathcal{H}\Psi[\gamma_{ij}, \phi_{\text{matter}}] = 0. \quad (3.38)$$

where  $\Pi^{ij} = -i\frac{\partial}{\partial\gamma_{ij}}$ . The value of the energy for any valid state is then again given by the expression (3.36).

### 3.4.3 Dimensional reduction of constraints in warped geometries

We now specialize the constraints above to the geometry with branes. We would like to show that the local constraint (3.34) goes over, after dimensional reduction, to the local constraint of a theory of massive gravity. The mass spectrum of the lower-dimensional graviton has been studied in geometries without black holes and so, for simplicity, we consider such geometries here. The results that we derive below will have a clear generalization.

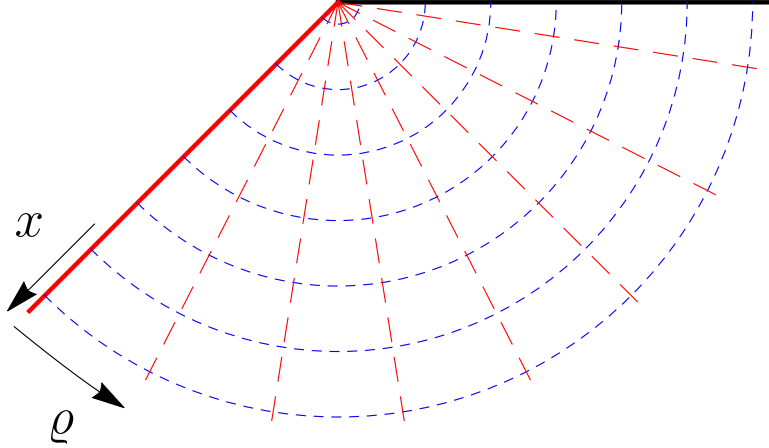


Figure 3.7: An  $AdS_d$  brane (thick red line) in  $AdS_{d+1}$ . The conformal boundary is the thick black line. Lines of constant  $\varrho$  are red dashed lines and coordinates transverse to  $\varrho$  are constant along the blue dashed circles.

We consider a family of “warped geometries” in  $d + 1$  dimensions that can be written in the form,

$$ds^2 = d\varrho^2 + e^{2A(\varrho)} \left( -\bar{N}^2 dt^2 + \bar{\gamma}_{\bar{i}\bar{j}} dx^{\bar{i}} dx^{\bar{j}} \right). \quad (3.39)$$

Our notation for the coordinates is as follows. We consider a brane placed at  $\varrho = 0$ . Our objective is to dimensionally reduce along the  $\varrho$ -direction. We use  $\bar{i}, \bar{j}$  to run over the spatial directions excluding  $\varrho$ . We use  $\bar{a}, \bar{b}$  to run over the spacetime directions excluding  $\varrho$ . Here we are adopting the notation of Aharony et al. (2003) and  $\varrho$  here should be equated with  $r$  in Aharony et al. (2003). (See Figure 3.7.)

We will focus on warped geometries where  $\bar{N}$  is independent of  $\varrho$  and consider a background solution of the vacuum Einstein equations where  $\bar{\gamma}_{\bar{i}\bar{j}} = \bar{g}_{\bar{i}\bar{j}}$ . We then look for nearby solutions where,

$$\bar{\gamma}_{\bar{i}\bar{j}} = \bar{g}_{\bar{i}\bar{j}} + \bar{h}_{\bar{i}\bar{j}}. \quad (3.40)$$

Note that the perturbation displayed in (3.40) is not the most general perturbation of the warped geometry. In general the higher-dimensional graviton reduces,

upon dimensional reduction in the  $\varrho$ -coordinate, to a scalar, a vector and a lower-dimensional graviton. We focus only on the lower-dimensional graviton and neglect the scalar and vector modes.

In terms of the perturbation defined in (3.29) we have,

$$h_{\bar{i}\bar{j}} = e^{2A(\varrho)} \bar{h}_{\bar{i}\bar{j}}. \quad (3.41)$$

We denote the trace of this perturbation by  $\bar{h} \equiv h^i_i = \bar{g}^{\bar{i}\bar{j}} \bar{h}_{\bar{i}\bar{j}}$ , where  $\bar{g}^{\bar{i}\bar{j}}$  is the inverse of  $\bar{g}_{\bar{i}\bar{j}}$ . Note that the warp factor drops out of this trace.

Some algebra leads to the following metric-compatible connection coefficients:

$$\begin{aligned} \Gamma_{\bar{a}\bar{b}}^\varrho &= -e^{2A(\varrho)} A'(\varrho) \bar{g}_{\bar{a}\bar{b}}, \\ \Gamma_{\varrho\bar{b}}^{\bar{a}} &= \Gamma_{\bar{b}\varrho}^{\bar{a}} = A'(\varrho) \delta_b^{\bar{a}}, \\ \Gamma_{\bar{b}\bar{c}}^{\bar{a}} &= \bar{\Gamma}_{\bar{b}\bar{c}}^{\bar{a}}. \end{aligned} \quad (3.42)$$

where  $\bar{\Gamma}$  refers to the connection coefficients of the metric  $\bar{g}_{\bar{a}\bar{b}}$ . Therefore, for any spatial vector field  $J^i$  we have,

$$\nabla_i J^i = \partial_\varrho J^\varrho + A'(\varrho)(d-1)J^\varrho + \bar{\nabla}_{\bar{i}} J^{\bar{i}}, \quad (3.43)$$

where  $\bar{\nabla}_{\bar{a}}$  is the  $\bar{g}_{\bar{a}\bar{b}}$ -compatible covariant derivative. Then from the definition (3.35), a straightforward calculation shows that,

$$J^\varrho = -\bar{N} e^{A(\varrho)} \partial_\varrho \bar{h}, \quad (3.44)$$

where we have used the fact that  $\bar{N}$  is independent of  $\varrho$ . Hence we have,

$$\nabla_i J^i = -e^{-(d-1)A(\varrho)} \bar{N} \partial_\varrho \left( e^{dA(\varrho)} \partial_\varrho \bar{h} \right) + \bar{\nabla}_{\bar{i}} J^{\bar{i}}. \quad (3.45)$$

The natural modes  $\phi_n$  in the  $\varrho$ -direction are those that satisfy

$$-e^{-(d-1)A(\varrho)} \partial_\varrho \left( e^{dA(\varrho)} \partial_\varrho \phi_n \right) = e^{-A(\varrho)} m_n^2 \phi_n, \quad (3.46)$$

where  $m_n$  gives the spectrum of the masses of the KK descendants of the graviton. This is the same eigenvalue equation that was found to govern the spectrum of massive

transverse traceless graviton fluctuations in this system and it is well known that the eigenvalues  $m_n^2$  are all nonzero Karch and Randall (2001b). Here we use these same eigenvalues in the constraint equations for the “T” component of the metric fluctuation.

It is natural to expand,

$$\bar{h} = \sum \bar{h}_n \phi_n. \quad (3.47)$$

In the equation above, it is understood that  $h_n$  varies only along the brane directions and the  $\varrho$ -dependence is completely captured by  $\phi_n$ . We will use the same convention for other decompositions below.

It is natural to also define a lower-dimensional current,

$$\bar{J}^i = e^{A(\varrho)} J^i, \quad (3.48)$$

and lower-dimensional energy density,

$$\bar{\rho} = e^{2A(\varrho)} \rho. \quad (3.49)$$

The powers of the warp-factor that appear above can be obtained by carefully keeping track of the relative factors that appear if one uses the lower-dimensional metric  $\bar{g}_{\bar{i}\bar{j}}$  to define a lower-dimensional Hamiltonian density and the lower-dimensional perturbation  $\bar{h}_{\bar{i}\bar{j}}$  in (3.35) to define a lower-dimensional current.

By also expanding the divergence of this lower-dimensional current and energy-density in terms of the modes (3.46),

$$\bar{J}^i = \sum \bar{J}_n^i \phi_n; \quad \bar{\rho} = \sum \bar{\rho}_n \phi_n, \quad (3.50)$$

we find from (3.34) the constraint equation,

$$m_n^2 \bar{N} \bar{h}_n + \bar{\nabla}_{\bar{i}} \bar{J}_n^{\bar{i}} = 16\pi G \bar{N} \bar{\rho}_n. \quad (3.51)$$

We again see that this equation allows for a nonzero matter stress-energy that cannot necessarily be measured at the boundary of the brane because of the  $m_n^2$  term.



Using this term, it is possible to find a solution to the constraint (3.51) where the value of  $\bar{\rho}_n$  is not equal to the divergence of  $\bar{J}_n^i$  and the difference is made up by the  $m_n^2$  term.

The underlying physics can be understood as follows. In the higher-dimensional setting, it is clear that the constraint equation can be satisfied by shooting the gravitational field lines “off of” the brane and allowing them to end at the asymptotic boundary. In the lower-dimensional theory, this effect is captured by  $m_n^2$  term. This is because the  $m_n^2$  term arises directly from the variation of the metric in the directions that point “off” the brane. Therefore, the nonzero  $m_n^2$  term is a signal in the lower-dimensional effective field theory that the gravitational field lines can escape off the brane in the higher-dimensional description.

As we have already mentioned, it is known through direct calculation that  $m_n^2$  is nonzero for AdS branes embedded in AdS spaces. But even if this had not been known, our analysis above could have been used to deduce this fact.

We would like to make a few comments.

1. In a previous paper Geng et al. (2021b), we studied the system with two branes, where a specific linear combination of the localized gravitons is massless. One of the results of Geng et al. (2021b) was that if one studies minimal surfaces with endpoints on both branes then the only such surface is the horizon. In the higher-dimensional description, this implies that the entanglement wedge of the defect, where the branes meet, is the entire exterior of the black hole. Therefore, even in the dimensionally-reduced description, this entanglement wedge extends up to the asymptotic boundary and is not an island. In Geng et al. (2021b), we also studied islands on a single brane that are relevant for the entanglement between internal degrees of freedom on the defect. But if one studies islands on one brane then the effective description in terms of the localized-gravity theory on that brane has access only to a linear combination of a massive and massless graviton. Correspondingly, the asymptotic form of the metric on a single brane

does not give us access to the total energy on that brane. Therefore the results of Geng et al. (2021b) are entirely consistent with the reasoning advanced in this paper.

2. It is sometimes proposed that an island with a massless graviton can be studied by starting with an AdS black hole that partially evaporates into a non-gravitational bath allowing for a massive graviton, after which one “switches off” the coupling between the bath and the AdS space. However, a closer analysis of the causal structure of this process reveals that the graviton is always massive in the island. This is explained in appendix 3.A.5.
3. It is also not possible to obtain operators that commute with the asymptotic Hamiltonian by dressing them to an “end-of-the-world” brane, since such a brane must interact with the ambient metric. Therefore, in theories with long-range gravity, the brane must itself be dressed to the asymptotic boundary. The asymptotic boundary Hamiltonian then measures the energy of the brane and that of additional excitations. So even if one dresses an excitation to the brane, the excitation still transforms under time-translations generated by the asymptotic Hamiltonian.

### 3.5 Islands in decoupled systems

In the sections above, we described a puzzle and then showed how the puzzle could be resolved if the gravitational theory that supports islands is massive. When a standard theory of gravity in AdS is coupled to a bath, the mass of the graviton arises as a consequence of the coupling.

Islands have also been studied when a gravitating system and another system are entangled but not coupled Penington et al. (2019); Hartman et al. (2020a); Balasubramanian et al. (2020a); Anderson et al. (2020); Fallows and Ross (2021); Miyata and Ugajin (2021a). In higher dimensions, islands in decoupled systems are also sub-

ject to the puzzle described in section 3.3. In the absence of a coupling, the graviton does not pick up a dynamically generated mass, and so the resolution outlined in section 3.4 does not apply to decoupled systems. Therefore, in higher dimensions, unless the theory of gravity is massive to start with, our analysis above implies that islands in decoupled systems would suffer from an inconsistency with the Gauss law.

However, most of the studies of decoupled islands have been performed only in  $1+1$ -dimensions and so the analysis of section 3.3 and section 3.4 does not directly apply to these studies. Although it will be of interest to see if the considerations above apply to this lower-dimensional case, in this section, we focus on another elementary consistency condition that must be obeyed by all islands in decoupled systems, including islands in  $1+1$ -dimensional theories. We will explain how this simple separate consistency condition—though not necessarily ruling out islands—is sufficient to indicate that islands in decoupled systems cannot be used to model realistic evaporating black holes.

### 3.5.1 Consistency condition

Consider two decoupled systems described by Hilbert spaces,  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively. The joint system is described by the Hilbert space,

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2. \quad (3.52)$$

That the systems are *decoupled* means, by definition, that the Hamiltonian in the joint Hilbert space  $H$  is just a sum of the Hamiltonians  $H_1$  and  $H_2$ ,

$$H = H_1 \otimes 1 + 1 \otimes H_2, \quad (\text{decoupled systems}) \quad (3.53)$$

We now derive a few consequences of the elementary equations (3.53) and (3.52). First note that given any density matrix  $\rho$  that describes the state of the joint system, we can obtain a density matrix for system 1 using

$$\rho_1 = \text{Tr}_2(\rho). \quad (3.54)$$

Now consider the action of an *arbitrary unitary operator*  $U_2$  that acts on  $\mathcal{H}_2$ , i.e. it acts as the identity on  $\mathcal{H}_1$ . This modifies the joint density matrix as

$$\rho \rightarrow U_2 \rho U_2^\dagger. \quad (3.55)$$

But under such a transformation of state, the density matrix of the first system is unchanged,

$$\text{Tr}_2(U_2 \rho U_2^\dagger) = \text{Tr}_2(\rho) = \rho_1. \quad (3.56)$$

Consequently, the expectation value of *all* observables in system 1 are unchanged by the action of a unitary in system 2.

Equation (3.56) holds whenever the Hilbert space factorizes. But in the case of decoupled systems, as a consequence of (3.53), even the time-evolution operator in system 2 commutes with all operators in system 1. Therefore (3.56) must hold even if we evolve  $U_2$  forward or backward by an *arbitrary* amount of time  $t$ .

$$\text{Tr}_2 \left[ e^{-i(1 \otimes H_2)t} U_2 e^{i(1 \otimes H_2)t} \rho e^{-i(1 \otimes H_2)t} U_2^\dagger e^{i(1 \otimes H_2)t} \right] = \text{Tr}_2(\rho) = \rho_1. \quad (3.57)$$

Now consider a setup where we have two decoupled systems, and where a region  $R$  in the second system is believed to also describe the degrees of freedom in an island  $I$ . The argument above tells us that no possible unitary operator acting on  $R$ , at any point in time, should have the ability to affect any degrees of freedom from the first system. Conversely no possible unitary acting on the first system at any point of time should affect the island. This leads us to the following elementary consistency condition.

*Observation.* When islands are redundant with degrees of freedom from one part of a decoupled pair of systems, it should be impossible for the islands to either send signals to or receive signals from the degrees of freedom described by the other part of the decoupled pair of system.

### 3.5.2 Implications of the consistency condition

We now explain how the consistency condition derived above rules out the possibility of modeling the interior of physical black holes using islands in decoupled systems.

An example of a geometry that obeys the consistency condition above is provided by Figure 3.8, which was obtained in Balasubramanian et al. (2020a).<sup>7</sup> Note that in Figure 3.8 it is insufficient for the island to be behind the future horizon to satisfy the consistency condition. Starting from the left asymptotic boundary, one has to cross *two* horizons to reach the island region, and the same is true if one starts from the right asymptotic boundary.

Referring to Figure 3.8, the two asymptotic boundaries completely describe the physics in region  $L$  and region  $R$ . The island is dual to a decoupled system that completely describes the physics in region  $I$ . The regions marked  $J$  are described by a combination of the decoupled system and the system that lives on the asymptotic boundaries. This Penrose diagram meets our consistency condition since every point in region  $L$  and region  $R$  is separated by a spacelike interval from every point in region  $I$ . Note that signals sent from region  $L$  (or region  $R$ ) and region  $I$  can meet in the regions marked  $J$ . This does not contradict the consistency condition but reflects the familiar feature, also present in the duality between an eternal black hole and the thermofield doubled state Israel (1976); Maldacena (2003), that signals from decoupled systems can meet inside a wormhole.

This example illustrates that any spacetime geometry that meets the above consistency condition must have a qualitatively different causal structure from that of a black hole formed from collapse. For contrast, a Penrose diagram of a single-

---

<sup>7</sup>The setup of Hartman et al. (2020a) also satisfies our consistency condition although this is more subtle to see from the Penrose diagrams of Hartman et al. (2020a). In Figure 13 of Hartman et al. (2020a) for instance, the independent degrees of freedom of the gravitational system should be associated only with the “tip” of the Penrose diagram at spatial infinity, and one can neither send nor receive signals from spatial infinity. We thank Tom Hartman for explaining this point to us.

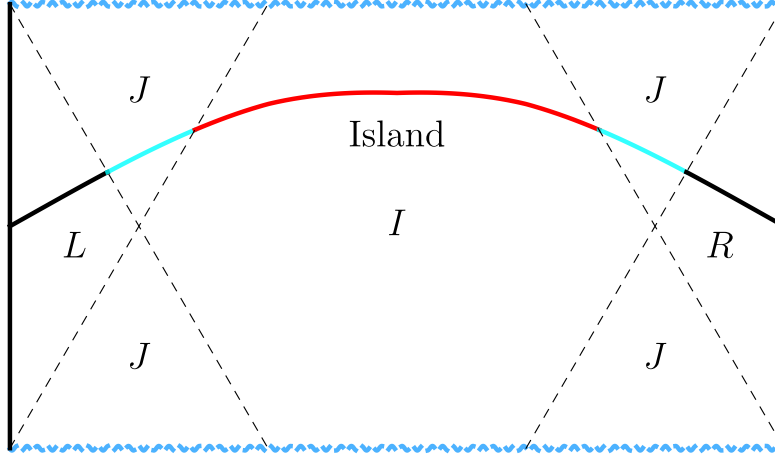


Figure 3.8: *One possibility for an island in a decoupled pair of systems. The island cannot affect the degrees of freedom beyond either the left horizon or the right horizon. The solid black lines are the left and right asymptotic AdS boundaries. The broken blue lines are singularities and the dashed lines are horizons. The dashed lines also demarcate the boundaries between regions marked,  $L$ ,  $R$ ,  $I$ ,  $J$ . In this figure, the horizontal direction is spatial and time runs along the vertical direction.*

sided AdS black hole is shown in Figure 3.9. It is clear that, even classically, it is possible to send signals to *every* point in the interior of the black hole from the exterior, provided the signal is sent early enough. If any part of the interior had been part of an island from a decoupled system, this would not have been possible by our consistency condition. Therefore no part of the interior of such a black hole can constitute an island that is described by degrees of freedom from a decoupled system.

Figure 3.9 shows a large single-sided black hole, but the same conclusion holds for an evaporating black hole.

For completeness, we mention that the Penrose diagram in Figure 3.8 is also qualitatively different from that of the standard eternal black hole, which is shown in Figure 3.10. Here, it is always possible to send a signal to any point in the interior from either the left or the right asymptotic boundary. The consistency condition above then implies that no part of the interior of the eternal black hole of Figure 3.10

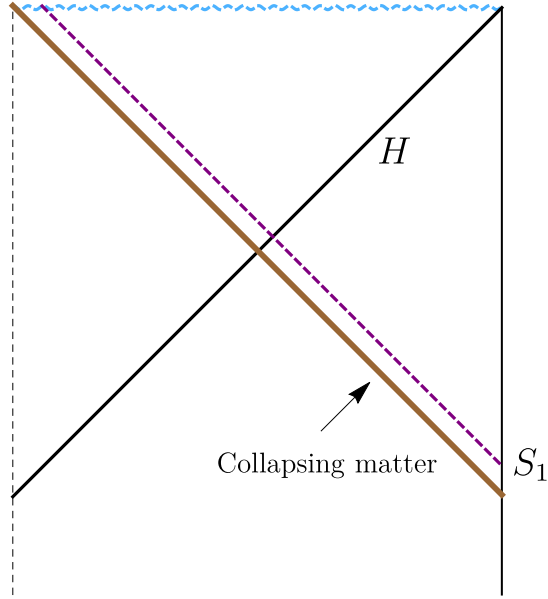


Figure 3.9: A single sided black hole formed from the collapse of some matter (shown in brown). The dashed line on the left is the origin of polar coordinates and the solid black line on the right is the asymptotic AdS boundary. The Figure shows a signal  $S_1$  that originates on the boundary and reaches deep inside the interior. In this figure, the horizontal direction is spatial and time runs along the vertical direction.

can constitute an island corresponding to degrees of freedom from a system that is decoupled from both asymptotic boundaries.

We emphasize that although the Penrose diagram in Figure 3.8 satisfies the consistency condition described in this section, this picture of an island is not exempt from the analysis described in section 3.3. In higher dimensions, Figure 3.8 would be consistent for massive gravity but would violate the Gauss law in ordinary massless gravity since the energy of excitations in the island could be measured by combining observations in the left and right asymptotic regions. The study in Balasubramanian et al. (2020a), where this diagram was obtained, was performed in the  $1 + 1$  dimensional context of JT gravity, where we have yet to determine if some analogue of our puzzle exists.

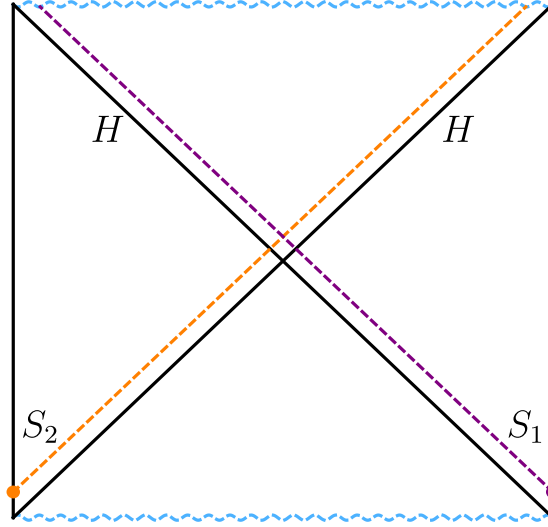


Figure 3.10: *In the eternal black hole, it is always possible to send signals to any point in the interior, even classically, provided these signals originate at an early enough time. The Figure shows a signal  $S_1$  that originates on the left boundary and another signal  $S_2$  that originates on the right boundary. Therefore, no part of the interior of a standard eternal black hole can constitute an island that is redundant with degrees of freedom from a decoupled system. In this figure, the horizontal direction is spatial and time runs along the vertical direction.*

Islands in the presence of a decoupled bath were also studied in Penington et al. (2019). The theory used in Penington et al. (2019) is defined via a precise rule for Euclidean path integrals. We are unable to determine if the Euclidean saddle points corresponding to islands found in Penington et al. (2019) satisfy the consistency condition outlined here, since the Lorentzian action that gives rise to the rules for the Euclidean path integral above is not known. It would be interesting to obtain the Lorentzian theory corresponding to the Euclidean rules of Penington et al. (2019). This would allow us to check if the consistency condition outlined in this section holds for this theory and would more readily allow an extension of the puzzle of section 3.3 to this model.



### 3.6 Discussion

In this paper, we have described how models of black hole evaporation that involve “islands” lead to puzzles in theories with long-range gravity where the Gauss law applies. In quantum mechanics, the Gauss law is stronger than it is in classical mechanics. A localized operator in the bulk must have nonzero energy, and so it must fail to commute with the Hamiltonian which, in a theory of gravity, is a boundary term. This nonzero commutator can be thought of as arising because any localized operator must be “dressed” to the asymptotic boundary. Such a dressing is sometimes called a gravitational Wilson line. (See Donnelly and Giddings (2016b); Giddings (2019) and references there.)

Since there are no negative charges in gravity, gravitational Wilson lines end only on asymptotic boundaries. In conventional island-free entanglement wedges in AdS/CFT, such asymptotic regions are always part of the wedge. Therefore operators in the wedge can form a self-contained algebra, since the entire operator, including its dressing, can be localized within the wedge.

Islands would be entanglement wedges that do not extend to the asymptotic boundary and are surrounded by their complements. Consequently, in an ordinary theory of gravity in which the Gauss law applies, it is not possible to localize an operator and its dressing entirely to an island. This means that, in an ordinary gravitational theory, an island cannot constitute an entanglement wedge.

We showed how this puzzle was resolved in a concrete setting. When islands are realized by embedding a brane in a higher-dimensional anti-de Sitter spacetime, the lower-dimensional theory of gravity obtained on the brane has a massive graviton. Such theories do not have a Gauss law, and so the puzzle does not arise at all.

An extension of our analysis to  $1 + 1$  dimensional theories of gravity would allow contact with the significant literature where islands have been studied using JT gravity Almheiri et al. (2020b); Penington et al. (2019) and other  $1 + 1$  dimensional models Hartman et al. (2020b); Akal et al. (2021a); Gautason et al. (2020). One

obstacle to extending our analysis is purely technical: the analysis of gravitational constraints reviewed in Raju (2022a) which suggests that gravity stores information differently from local quantum field theories is valid only for spacetime dimension larger than two. Moreover, the mechanism whereby imposing transparent boundary conditions on a theory of gravity coupled to matter in AdS leads to a mass for the graviton is understood only for higher-dimensional theories Porrati (2003, 2002b); Porrati and Starinets (2002).

Nevertheless, we would like to mention some results in the existing literature that suggest that it might be possible to generalize our puzzle. Even in  $1 + 1$  dimensional dilaton-gravity models with a negative cosmological constant and appropriate boundary conditions, it is possible to define conserved charges Grumiller et al. (2017); Grumiller and McNees (2007); Ruzziconi and Zwickel (2021) that are related to the constraints of the theory. Moreover Harlow and Jafferis (2020), extending the techniques of Harlow (2016); Guica and Jafferis (2017), analyzed the significance of the bulk constraints in pure JT gravity without matter and showed that these constraints prevent the factorization of the Hilbert space on the boundary, which is a puzzle that is somewhat similar in flavor to the puzzle that we have described in this paper. Some further discussion of these issues can be found in Gao et al. (2021); Harlow and Shaghoulian (2021). One aspect of the higher-dimensional analysis that clearly does carry over even to  $1 + 1$  dimensions is the fact that coupling to an external bath will allow energy to leak out of the system and thereby ruin any conservation laws that were present before the coupling to the bath. It is an interesting open problem to build on these papers and see whether and how our puzzle extends to  $1 + 1$  dimensions.

Similarly, we also point out that we have addressed only the theories of massive gravity that emerge when AdS branes are embedded in AdS spaces. Our analysis does not necessarily indicate whether other theories of massive gravity can support islands.

Finally we contrast the picture of how quantum information is localized in theories of gravity that support islands with the corresponding picture in theories of

long-range gravity. Even in theories with long-range gravity, degrees of freedom that appear to be localized to a region may be equated with a scrambled version of degrees of freedom in another region. This physical effect follows from a careful analysis of gravitational constraints as reviewed in Raju (2022a). It is also important for black hole evaporation and for understanding the interior of black holes in AdS/CFT as emphasized in Papadodimas and Raju (2013b, 2014b). However, in theories with long-range gravity, degrees of freedom near the asymptotic boundary capture *all* the degrees of freedom on a bulk Cauchy slice and not only the degrees of freedom in an island. Therefore, in the presence of long-range gravity, information about the black hole is always available outside, regardless of the stage of black-hole evaporation Laddha et al. (2020).

On the other hand, in models of black hole evaporation that involve islands, information emerges gradually during black hole evaporation according to the Page curve. This agrees with intuition obtained from non-gravitational systems, where such a Page curve is expected on very general grounds. The key reason that the entropy of the radiation first increases and then decreases in such models is that the radiation region describes the island but does not describe its complement. But not only do such models necessarily involve non-gravitational baths; they seem to be consistent only in theories without a Gauss law.

We emphasize that our puzzle is not simply a question of the examples presented in the literature. Theories with long-range gravity allow the energy to be measured at the asymptotic boundary, which is in contradiction with the notion of an island as a connected component of an entanglement wedge, disconnected from the boundary. This is why all consistent constructions of islands so far in gravitational theories with more than two dimensions involve massive gravity.

## Appendix 3.A: Loopholes in Possible Counterarguments

In the main text we have shown that islands cannot support approximately local operators that commute with operators in its complement.

In this appendix, we will examine some proposals that attempt to sidestep our puzzle. We will conclude, as the puzzle in this paper would lead us to expect, that none of these proposals lead to operators that are viable for entanglement wedge reconstruction in a standard theory of gravity.

### 3.A.1: Products of Modes in Frequency Space

It is not difficult to construct operators that commute with the Hamiltonian to an excellent approximation. As explained in the text, the difficulty is in localizing such operators. This follows from the uncertainty principle; an operator that is localized to a finite region must have nonzero energy and an operator that has zero energy must extend to infinity.

This can be illustrated using the following example.<sup>8</sup> As in section 3.3, consider a scalar field  $\phi$  propagating in a spherically symmetric black hole in an asymptotically global  $\text{AdS}_{\tilde{d}+1}$  spacetime. In this subsection, we write the field as  $\phi(t, r, \Omega)$ , thereby explicitly specifying its position in time, the radial coordinate and the  $S^{\tilde{d}-1}$ . To leading order in  $G$ , outside the black hole horizon, the field can be expanded in terms of modes,

$$\phi(t, r, \Omega) = \sum_{\ell} \int d\omega a_{\omega, \ell} e^{-i\omega t} \psi_{\omega, \ell}(r) Y_{\ell}(\Omega) + \text{h.c.}, \quad (3.58)$$

where  $Y_{\ell}(\Omega)$  is a spherical harmonic corresponding to the angular momentum quantum numbers  $\ell$  and  $\psi_{\omega, \ell}$  are radial mode functions that are discussed in greater detail in Papadodimas and Raju (2013b). With respect to the boundary Hamiltonian  $H$  that appeared in section 3.3 as the integral of the boundary metric we have,

$$[H, a_{\omega, \ell}] = -\omega a_{\omega, \ell}; \quad [H, a_{\omega, \ell}^{\dagger}] = \omega a_{\omega, \ell}^{\dagger}; \quad [a_{\omega, \ell}, a_{\omega', \ell'}^{\dagger}] = \delta(\omega - \omega') \delta_{\ell \ell'}. \quad (3.59)$$

---

<sup>8</sup>We thank an anonymous referee for drawing our attention to this proposal.

Now consider the following operator,

$$X = a_{\omega_1,0}^\dagger a_{\omega_2,0} a_{\omega_3,0}, \quad (3.60)$$

where  $\omega_1 = \omega_2 + \omega_3$  and where the 0 in the subscript indicates that we are focusing on the  $s$ -wave sector. It is easy to see from (3.59) that

$$[H, X] = 0. \quad (3.61)$$

The trilinear operator  $X$  above is just an example and the discussion below easily generalizes to any polynomial in the modes that commutes with the Hamiltonian.

However, the creation and annihilation operators that enter inside  $X$  are not local operators. As a result, whereas  $X$  commutes with the Hamiltonian it does not commute with other operators outside the island. We will now show, through an explicit computation, that the nonzero commutator between  $X$  and operators outside the island can be seen at  $O(1)$  and does not require us to even keep track of gravitational effects.

Consider the limit of the field operator near the boundary at a time  $t_1$ . We can take  $t_1$  to be sufficiently small so that the boundary point is spacelike to every point in the island. Note that such points always exist since, by our definition in section 3.1.1, the island does not extend to the boundary. Using the mode expansion and commutation relations above we find that (to  $O(G^0)$ ),

$$\lim_{r \rightarrow \infty} r^\Delta [X, \phi(r, t_1, \Omega)] = -C_{\omega_1} e^{-i\omega_1 t_1} a_{\omega_2,0} a_{\omega_3,0} + C_{\omega_2} e^{i\omega_2 t_1} a_{\omega_1,0}^\dagger a_{\omega_3,0} + C_{\omega_3} e^{i\omega_3 t_1} a_{\omega_1,0}^\dagger a_{\omega_2,0}, \quad (3.62)$$

where we have normalized the spherical harmonics through  $Y_0(\Omega) = 1$  and  $C_\omega$  is a real number defined by

$$C_\omega = \lim_{r \rightarrow \infty} r^\Delta \psi_{\omega,0}(r). \quad (3.63)$$

The precise value of  $C_\omega$  must be computed numerically in general  $\tilde{d}$  but will not be required below.

The nonzero commutator (3.62) can be easily detected inside a correlation function by inserting two more field operators, at points  $(r, t_2, \Omega_2)$  and  $(r, t_3, \Omega_3)$  that we again choose to be spacelike to the island, and using the correlators

$$\langle \Psi | a_{\omega, \ell} a_{\omega', \ell'}^\dagger | \Psi \rangle = G_\omega^+ \delta(\omega - \omega') \delta_{\ell, \ell'}; \quad \langle \Psi | a_{\omega, \ell}^\dagger a_{\omega', \ell'} | \Psi \rangle = G_\omega^- \delta(\omega - \omega') \delta_{\ell, \ell'}, \quad (3.64)$$

where

$$G_\omega^+ = \frac{1}{1 - e^{-\beta\omega}}; \quad G_\omega^- = G_\omega^+ - 1 = \frac{e^{-\beta\omega}}{1 - e^{-\beta\omega}}, \quad (3.65)$$

and  $\beta$  is the inverse temperature of the black hole. We find that

$$\begin{aligned} \lim_{r \rightarrow \infty} r^{3\Delta} \langle \Psi | \phi(r, t_2, \Omega_1) \phi(r, t_3, \Omega_2) [X, \phi(r, t_1, \Omega)] | \Psi \rangle &= C_{\omega_1} C_{\omega_2} C_{\omega_3} \\ &\times \left[ -e^{-i\omega_1 t_1} G_{\omega_2}^- G_{\omega_3}^- \left( e^{i\omega_2 t_2 + i\omega_3 t_3} + e^{i\omega_3 t_2 + \omega_2 t_3} \right) + e^{i\omega_2 t_1} G_{\omega_1}^+ G_{\omega_3}^- \left( e^{-i\omega_1 t_2 + \omega_3 t_3} + e^{-i\omega_1 t_3 + \omega_3 t_2} \right) \right. \\ &\left. + e^{i\omega_3 t_1} G_{\omega_1}^+ G_{\omega_2}^- \left( e^{-i\omega_1 t_2 + \omega_2 t_3} + e^{-i\omega_1 t_3 + \omega_2 t_2} \right) \right]. \end{aligned} \quad (3.66)$$

Note that once we use  $\omega_1 = \omega_2 + \omega_3$  we see that the result depends only on the time-differences  $(t_2 - t_1)$ ,  $(t_3 - t_1)$  and  $(t_3 - t_2)$ .

We can present this result in a manner that is identical to section 3.3 by showing how a unitary operator containing  $X$  changes correlators outside the island.  $X$  is not Hermitian, but we can construct a unitary operator via  $U_X = e^{i\lambda(X+X^\dagger)}$ . We then find that

$$\begin{aligned} \lim_{r \rightarrow \infty} r^{3\Delta} \frac{\partial}{\partial \lambda} \langle \Psi | \phi(r, t_1, \Omega_1) \phi(r, t_2, \Omega_2) U_X \phi(r, t, \Omega) U_X^\dagger | \Psi \rangle \Big|_{\lambda=0} \\ = \lim_{r \rightarrow \infty} i r^{3\Delta} \langle \Psi | \phi(r, t_1, \Omega_1) \phi(r, t_2, \Omega_2) [X + X^\dagger, \phi(r, t_1, \Omega)] | \Psi \rangle \neq 0. \end{aligned} \quad (3.67)$$

The exact expression on the right hand side above can be read off from (3.66) but we do not write it since it is not illuminating.

We pause to note another subtlety with the unitary  $U_X$ . One might have thought that the unitary changes the occupation number of modes with frequency  $\omega_1, \omega_2$  and  $\omega_3$ . However, a simple check shows that the correlators (3.64) are unchanged in the state  $U_X | \Psi \rangle$  for all frequencies including these three frequencies. More

precisely,

$$\langle \Psi | U_X^\dagger a_\omega a_\omega^\dagger U_X | \Psi \rangle = \langle \Psi | a_\omega a_\omega^\dagger | \Psi \rangle, \quad \text{and} \quad \langle \Psi | U_X^\dagger a_\omega^\dagger a_\omega U_X | \Psi \rangle = \langle \Psi | a_\omega^\dagger a_\omega | \Psi \rangle, \quad \forall \omega \quad (3.68)$$

This is a general property of unitary operators that commute with the Hamiltonian. They do not “create excitations” in typical states because correlators in typical states are thermal correlators, and so if  $A$  is any operator,

$$\langle \Psi | U_X^\dagger A U_X | \Psi \rangle = \frac{1}{Z(\beta)} \text{Tr}(e^{-\beta H} U_X^\dagger A U_X) = \frac{1}{Z(\beta)} \text{Tr}(U_X e^{-\beta H} U_X^\dagger A) = \langle \Psi | A | \Psi \rangle, \quad (3.69)$$

where we have used the cyclicity of the trace and the commutator of  $U_X$  and the Hamiltonian, and  $Z(\beta)$  is the partition function.

So  $U_X |\Psi\rangle$  is not an excited state at all and can be thought of as just another black hole microstate. This observation is not in contradiction with (3.67) since the unitary operator there is inserted in the middle of other operators.

To summarize, our calculation implies that  $X$  cannot be thought of, in any sense, as an operator localized to the island. Moreover, the action of a unitary made up of  $X$  does not even create an excitation. So its existence does not resolve the puzzle presented in the text and, moreover, is irrelevant for the discussion of entanglement wedge reconstruction.

The discussion above is quite general and can be generalized to all operators that can be formed from the operators displayed in (3.58). We note that by using state-dependent operators it is possible to construct operators that cannot be detected by means of (3.66) Papadodimas and Raju (2016) but even these operators are infinitely delocalized and are not relevant for entanglement wedge reconstruction.

### 3.A.2: Swapping Excitations

Another proposal is to only study operators that swap one excitation for another. More specifically, one starts not with an empty black hole but rather with

an excited black hole state of the form  $U|\Psi\rangle$  where  $U$  is defined in section 3.3. Now imagine that one has another field with precisely the same mass, which we denote by  $\tilde{\phi}$ . One might then consider the operator that *swaps* an excitation made up of  $\phi$  for an excitation made up of  $\tilde{\phi}$ . More precisely we study the unitary operator  $V = \tilde{U}U^\dagger$  where  $\tilde{U} = e^{i\lambda\tilde{\phi}(P)}$  in the notation of section 3.3.

We note that to leading order in  $G$ ,

$$\lim_{r \rightarrow \infty} r^{\tilde{d}-2+\Delta} \int_{(\partial\text{AdS})} \frac{\partial}{\partial\lambda} \langle \Psi | \tilde{U} \frac{\partial\phi(P')}{\partial t} h_{00} \tilde{U}^\dagger | \Psi \rangle \Big|_{\lambda=0} = 0. \quad (3.70)$$

But the insertion sandwiched between the unitaries above is precisely the same insertion that appeared in section 3.3 where we found a nonzero value. Therefore the insertion of the boundary Hamiltonian and a matter field can distinguish between the state with an excitation of  $\phi$  and the state with an excitation of  $\tilde{\phi}$ . Therefore the operator  $V$  does not commute with the product of the boundary Hamiltonian and a matter field at infinity.

This should not be puzzling since the correlator we are measuring keeps track of more than just the energy; the insertion of  $\phi$  breaks the symmetry between the two states. Indeed, one can check that for a different correlator,

$$\lim_{r \rightarrow \infty} r^{\tilde{d}-2+\Delta} \int_{(\partial\text{AdS})} \frac{\partial}{\partial\lambda} \langle \Psi | \tilde{U} \frac{\partial\tilde{\phi}(P')}{\partial t} h_{00} \tilde{U}^\dagger | \Psi \rangle \Big|_{\lambda=0} \neq 0, \quad (3.71)$$

and the nonzero value is precisely the value that was explored in section 3.3.

### 3.A.3: Background Fields as Coordinates

In the presence of suitable background fields, it is possible to define perturbatively gauge-invariant local operators. More precisely, imagine that in  $\tilde{d}+1$  spacetime dimensions we have  $\tilde{d}+1$  background scalar fields  $X^1 \dots X^{\tilde{d}+1}$  and a classical background where a point in spacetime can be *uniquely specified* by specifying the values of the fields. Let  $\phi$  be another propagating field. It is then possible to define the operator

$$O(Z^i) = \int \phi(x) F(Z^1 - X^1(x), \dots, Z^{\tilde{d}+1} - X^{\tilde{d}+1}(x)) dx. \quad (3.72)$$



where  $F$  is a sharply peaked function that has support only when all its arguments are close to zero. Such an operator commutes with boundary operators within perturbation theory (although not nonperturbatively) and approximates a local operator. For more details we refer the reader to Marolf (2015) and references there.

However, such a construction is not possible for a black hole at late times. By the classical no-hair theorem, at late times and in the absence of asymptotic sources, the black hole solution does not support classical fields that take on distinct values at each point in spacetime. On the other hand, if we consider asymptotic sources that stay on permanently and stabilize a background of classical scalar hair, we would break diffeomorphism invariance and therefore give the graviton a mass as discussed in Blake et al. (2014).

Indeed, it is the emergent time-translational symmetry of the black hole solution that allows us to phrase our puzzle within perturbation theory. For general time-dependent backgrounds, it might be nonperturbatively difficult to detect the presence of a local excitation from infinity due to the existence of operators like (3.72). But because the late-time geometry of the black hole is so simple, local excitations can also be detected within perturbation theory as in empty AdS.

As opposed to the eternal black hole, the time-translational symmetry is not exact for the evaporating black hole in that the evaporation of the black hole itself defines a “clock.” One could attempt to use this clock to define operators along the lines of (3.72) that commute with the boundary Hamiltonian within perturbation theory. However since the evaporation happens over a very long time scale, this clock allows one to only define a class of very diffuse operators. These operators do not act at a well-defined time and are smeared out over an  $O(1)$  fraction of the evaporation time scale. Such operators are very different from the approximately local operators that one seeks in standard entanglement-wedge reconstruction. However, it would be interesting to explore the graviton mass and such operators more precisely and their significance for the island proposal for evaporating black holes.

### 3.A.4: Mundane Locality

In this paper, we have pointed out that gravitational effects pose an obstruction to defining operators that are confined to the island. The reader might wonder how this analysis is consistent with everyday experience: although we live in a world where gravity is presumably quantized, it is perfectly sensible to discuss local experiments without having to worry about the effect of these experiments at a distant location.

This has to do with the weakness of gravity as we now explain. As we have mentioned above, even in the presence of long-range gravity it is easy to obtain approximately local gauge-invariant operators. One option is simply to fix gauge. Such gauge-fixed operators depend on the choice of gauge and obey a nonlocal algebra Donnelly and Giddings (2016a). But these nonlocal commutators are suppressed by a factor of  $\left(\frac{E}{M_{\text{pl}}}\right)^{\tilde{d}-2}$  where  $E$  is a characteristic energy scale. (This is also the factor that appears on the right hand side of (3.16), where  $E$  depends on the precise operators inserted in that correlator.) For the energy scales probed in ordinary experiments, this factor is well below the threshold of experimental accuracy, and so the unusual localization of information in gravity is of no practical consequence.

On the other hand, gravitational effects are a key aspect of entanglement-wedge reconstruction. It is due to such effects that the entanglement wedge can be larger than the causal wedge, as happens when islands appear. Therefore such effects cannot be ignored in the study of islands.

### 3.A.5: Decoupling the Bath

Yet another proposal for reconstructing operators in the island in a standard theory of gravity is to start with a theory of massive gravity coupled to a nongravitational bath, wait until an island forms in the bulk, and then turn off the coupling. One might naively expect that the graviton becomes massless in the bulk while operators in the island are still redundant with operators in the radiation region. Therefore a naive analysis might suggest that this procedure can be used to obtain islands in a

standard theory of gravity.

We now show that this does not happen. Even in this scenario, the island forms and exists only in the presence of a massive graviton.

In simple models, where the coupling is given by the product of a light operator on the boundary of AdS with an operator in the nongravitational bath, the coupling can be turned on and off by means of a simple change of boundary conditions at the boundary of AdS. One expects the effect of this change in boundary conditions to propagate causally in the bulk.

Physically, one can think of a “shock wave” that propagates inwards and changes the graviton from being massive to being massless in its wake. The graviton therefore becomes massless only at points in the bulk that are in the *causal future* of the decoupling event on the boundary. This is shown in Figure 3.11: the coupling is turned off in a coordinated manner at the time  $t_D$  on the boundary and the graviton becomes massless in the red “triangular” regions on the top left and top right corner of the Penrose diagram.

We see that the entanglement wedge corresponding to the island is a causal diamond in the bulk with the property that every point in the diamond is space-like to the decoupling event. The island is depicted as a green diamond in Figure 3.11. Therefore the island never learns of the change in boundary conditions and the graviton remains massive everywhere in the island.

In the black-hole geometry, if the bath is decoupled at late times the graviton remains massive on at least some part of every spacelike slice in the bulk. This is a consequence of the fact that the black-hole spacetime contains nontrivial causal patches. In contrast if a bath is coupled and then decoupled in empty AdS then, after a “transitional period” in the bulk, the graviton becomes massless everywhere. However, even in empty AdS the island is always confined to the region where the graviton is massive.

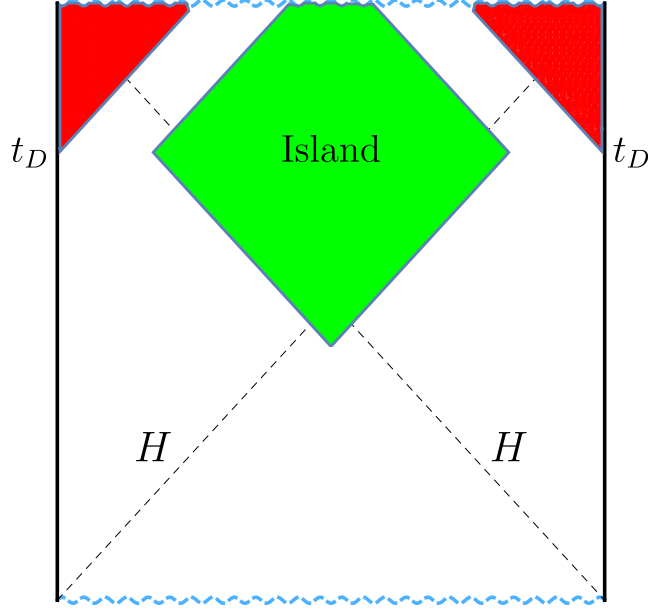


Figure 3.11: A possible attempt to produce islands in standard gravity. An island is first formed in a theory of massive gravity that involves a nongravitational bath. At time  $t_D$  the coupling to the bath is switched off. The change in boundary conditions is expected to make the graviton massless at all points in the causal future of the decoupling event (shown in red). But the island (shown in green) is unaffected by this event since every point in the island is spacelike to the decoupling event. So the graviton remains massive in the island.

We note that were the graviton in the vicinity of the island to indeed become massless,<sup>9</sup> the longitudinal mode essential to the consistency with Gauss' law would disappear when the mass turns off, leading to a violation of Gauss' law even if the mass had been present at an intermediate stage.

---

<sup>9</sup>We note that this possibility is the more complicated option in that it would require that the action of turning on and off the coupling is achieved by means of a complicated operator of AdS that can, in principle, distort the geometry deep in the bulk in a noncausal fashion.

# Chapter 4: Entanglement Phase Structure of a Holographic BCFT in a Black-Hole Background

## Abstract

We compute holographic entanglement entropy for subregions of a BCFT thermal state living on a nongravitating black hole background. The system we consider is doubly holographic and dual to an eternal black string with an embedded Karch-Randall brane that is parameterized by its angle. Entanglement islands are conventionally expected to emerge at late times to preserve unitarity at finite temperature, but recent calculations at zero temperature have shown such islands do not exist when the brane lies below a critical angle. When working at finite temperature in the context of a black string, we find that islands exist even when the brane lies below the critical angle. We note that although these islands exist when they are needed to preserve unitarity, they are restricted to a finite connected region on the brane which we call the atoll. Depending on two parameters—the size of the subregion and the brane angle—the entanglement entropy either remains constant in time or follows a Page curve. We discuss this rich phase structure in the context of bulk reconstruction.

---

<sup>0</sup>This chapter is based on published work in collaboration with Geng et al. (2022b). All authors contributed to the final version of the manuscript. My contributions included actively participating in discussions and bringing key ideas that shaped the project. I played a central role in conceptualizing the study of the entanglement phase structure of the BCFT in a black-hole background. Additionally, I worked with Andreas Karch to determine the appropriate boundary conditions for the Hartman-

## 4.1 Introduction

The *Ryu-Takayanagi (RT) prescription* Ryu and Takayanagi (2006c,d); Hubeny et al. (2007b), and its generalization to incorporate quantum effects Faulkner et al. (2013b); Lewkowycz and Maldacena (2013b); Engelhardt and Wall (2015b), have proven to be powerful tools in the AdS/CFT correspondence Maldacena (1998b); Gubser et al. (1998); Witten (1998a). These techniques yield remarkably simple formulas for the entropy of subregions of strongly coupled conformal field theories with bulk gravitating AdS duals, and they have since been extended to compute entropy in conformal field theories on spaces with boundaries (BCFTs).

Some strongly coupled BCFTs admit multiple dual descriptions and are called “doubly holographic” Karch and Randall (2001c); Takayanagi (2011b); Fujita et al. (2011). The dual descriptions comprise  $\text{AdS}_{d+1}$  gravity containing a subcritical end-of-the-world *Karch-Randall (KR) brane* Randall and Sundrum (1999d); Karch and Randall (2001b) and an “intermediate” description where brane-localized  $\text{AdS}_d$  gravity coupled to a CFT is connected to a nongravitating bath with the same CFT Almheiri et al. (2019a); Penington (2019); Almheiri et al. (2020e, 2019d,c). The  $(d - 1)$ -dimensional boundary of the BCFT is sometimes called a *defect*.

In this paper, we will study a black-string state in the bulk theory illustrated in Figure 4.1. This is dual to an  $\text{AdS}_d$  black hole coupled to a bath but it is also described by a CFT in a thermal state that itself resides on a nongravitating eternal black hole background. The braneworld black-hole horizon is where the black-string horizon connects to the brane and the horizon also extends to the boundary theory.

Since the BCFT description of this system is entirely nongravitational we expect that, in a generic state, the entanglement entropy of subregions in the dual

---

Maldacena surfaces. I performed entanglement entropy computations, alongside Marcos Riojas, and mapped out the entanglement phase structure, identifying the different regimes of behaviour. Furthermore, I compared the near-horizon entanglement entropy with analytical results, validating our numerical computations against known limits.

BCFT will be governed by a universal expression depending on the size of the subregion and its complement Page (1993b); Lubkin (1978). However, the original state of the BCFT at  $t = 0$  is not generic; this is reflected in correlations between operators in the left and right wedges of the background black hole. The subregions that we study have support in both of these wedges, and so their initial entropy is sometimes lower than the generic value. Time evolution leads to the loss of correlations between the two sides, and this causes the entanglement entropy of these subregions to approach that of a typical state, leading to a “Page curve” for the evolution of this entropy with time.

From the  $(d + 1)$ -dimensional bulk perspective, the Page curve is expected to emerge classically because, while the early time entropy is controlled by an extremal *Hartman-Maldacena (HM) surface* Hartman and Maldacena (2013) which crosses the black-string horizon and traverses the growing Einstein-Rosen bridge, the late-time entropy should be controlled by an extremal *island surface* which connects the brane to the conformal boundary and is constant in time Almheiri et al. (2019c). Such island surfaces are standard RT surfaces ending on the KR brane, but in the  $d$ -dimensional braneworld these are dual to *quantum extremal surfaces (QES)* Engelhardt and Wall (2015b); Almheiri et al. (2020e) which are disconnected from the  $\text{AdS}_d$  boundary and enclose (semiclassical) “islands” on the brane.<sup>1</sup>

---

<sup>1</sup>Entanglement islands have been studied extensively in the literature Ling et al. (2020); Kumar Basak et al. (2020b); Emparan et al. (2020); Caceres et al. (2021a,b); Deng et al. (2021); Krishnan (2020); Balasubramanian et al. (2020a,b); Manu et al. (2021); Karlsson (2021); Wang et al. (2021b); Miao (2021); Bachas and Papadopoulos (2021); May and Wakeham (2021); Kawabata et al. (2021a); Bhattacharya et al. (2021b); Anderson et al. (2020); Miyata and Ugajin (2021a); Kim and Nam (2021); Hollowood et al. (2021b); Wang et al. (2021a); Aalsma and Sybesma (2021); Ghosh and Krishnan (2021); Neuenfeld (2021b); Geng et al. (2021e); Balasubramanian et al. (2021b); Uhlemann (2021b); Neuenfeld (2021a); Kawabata et al. (2021b); Chu et al. (2021); Kruthoff et al. (2021); Akal et al. (2021b); Kumar Basak et al. (2021); Lu and Lin (2021); Omiya and Wei (2021); Ahn et al. (2021); Balasubramanian et al. (2021a); Li et al. (2021); Kames-King et al. (2021); Sun (2021); Hollowood et al. (2021c); Miyaji (2021); Bhattacharya et al. (2021a); Goswami et al. (2021); Chu and Miao (2021); Aref’eva and Volovich (2021); Shaghoulain (2021); García-García and Godet (2021); Buoninfante et al. (2021); Yu and Ge (2021); Nam (2021); He et al. (2021); Langhoff et al. (2021); Ageev (2021); Pedraza et al. (2021); Iizuka et al. (2021); Miyata and Ugajin (2021b); Gaberdiel et al. (2021); Uhlemann (2021a); Collier et al. (2021); Hollowood et al. (2021a); Emparan et al.



The story clearly breaks down if the bulk island surfaces do not exist in the first place,<sup>2</sup> and recent computations have shown Chen et al. (2020a); Geng et al. (2021b) that this may happen in the empty  $\text{AdS}_{d+1}$  bulk for  $d > 2$ . More precisely, there is a *critical angle* for the brane that depends on  $d$ , and island solutions do not exist when the brane lies below that angle. This is not necessarily a problem for empty  $\text{AdS}_{d+1}$  because there will be no Einstein-Rosen bridge, but island solutions are needed to obtain the Page curve when there is a black hole in the bulk. This concern is not unique to the black string model. The critical angle plays an important role in every braneworld model which is asymptotically AdS Geng et al. (2021b), and there seems to be no guarantee that island solutions exist in all configurations, even when they are needed to preserve unitarity.

To address this issue, this paper considers a tractable geometric configuration in which islands and their properties can be studied numerically. First a KR brane is embedded in an eternal  $\text{AdS}_{d+1}$  planar black string. The symmetries of the background allow the brane to be inserted at an arbitrary angle using the Israel junction conditions Israel (1966) that uniquely associate the brane tension with a given brane angle. Note that the planar black string is stable in that it does not have the usual Gregory-Laflamme instability Gregory and Laflamme (1993). Furthermore the horizon extends all the way to the conformal boundary and this is why the dual BCFT state itself resides on a nongravitating black hole background Marolf et al. (2014).

Although the bulk computation of the entropy is simpler than an ab initio computation in the strongly coupled BCFT Rozali et al. (2020); Sully et al. (2021), it still often requires the numerical solution of partial differential equations Almheiri et al. (2019d); Uhlemann (2021b). In this paper we show how the entropy of subregions for this BCFT state can be computed in the bulk using simple numerical

---

(2021); Omid (2021); Bhattacharya et al. (2021c), and useful reviews are Almheiri et al. (2020c); Raju (2022a,b); Liu and Sonner (2020); Nomura (2021); Kibe et al. (2021).

<sup>2</sup>Concretely, “nonexistence” of island surfaces means that the area functional of a codimension-2 bulk surface ending on the brane does not have any extrema. This is qualitatively different from the island surfaces existing but being subleading in the entanglement entropy.

techniques. Our main result is a presentation of the rich phase structure for this system, in which island solutions exist for all values of brane angle and finite subregion size, even below the critical angle. Instead of failing to exist, the semiclassical islands for fixed  $\theta_b$  are restricted to a region of the brane which we call the *atoll*, which we illustrate in Figure 4.3.

To determine which surface is dominant (i.e. minimizes the area functional) on the initial time slice, one simply computes the area difference between the HM and island surfaces for the same parameters.<sup>3</sup> We find that subregion entropy depends on both subregion size and brane angle. Depending on the values of these parameters, the initial dominant bulk surface is either an HM surface or the union of an island surface and its thermofield partner.

We obtained some simple rules for the phase structure. One can tune the size of the island by varying either radiation region size or the angle of the brane. Shrinking the radiation region size or decreasing the brane angle decreases the island size (except for the special case when the radiation region is the entire bath and the island covers the whole brane). In Figure 4.13 we illustrate the Russian doll rule, which states that nested surfaces can be constructed by either decreasing the brane angle  $\theta_b$  or increasing the anchor point  $\Gamma$ . We show how this “Russian doll rule” can be used to qualitatively compare the area difference for different parameter choices. This rule shows that the Page time diverges as the parameters are varied to probe regions near the black-string horizon in the bulk. We comment on this rule in the context of entanglement wedge reconstruction.

Another interesting feature of this model is the “constant entropy belt,” which is a connected interval on the boundary defined as the values of  $\Gamma$  for which island surfaces dominate on the initial time slice. When the angle of the brane  $\theta_b$  is increased

---

<sup>3</sup>The areas are technically UV-divergent, but the area difference is manifestly UV-finite since all extremal surfaces anchored to the conformal boundary exhibit the same divergence structure Graham and Witten (1999).

from 0 to  $\pi$ , the belt first appears slightly below the critical angle (at what we call the *Page angle*) and grows monotonically until it covers the entire boundary as  $\theta_b \rightarrow \pi$ . We find an interesting formula describing this region, which we elaborate on in Section 4.3.4.

**Relationship with previous work.** Before we proceed to our analysis, we pause to clarify how the results of this paper should be interpreted in light of our previous results on gravitating baths and the relationship between massive gravity and islands Geng et al. (2021b,c).

In Geng et al. (2021b), we showed that the physics of an  $\text{AdS}_d$  black hole coupled to a bath is qualitatively different when the bath is *gravitating* as opposed to nongravitating. Specifically, for a gravitating bath the Page curve disappears, even if gravity in the bath is parameterically weak (i.e. the ratio of  $G_N$  between the “bath brane” and the “system brane” is small). This is consistent with the finding that the full fine-grained entropy at  $\mathcal{I}^+$  does not follow a Page curve for black holes in asymptotically flat space Laddha et al. (2020); Raju (2022a). The nongravitating bath also induces a mass for the graviton in the region where gravity is dynamical Aharony et al. (2006) as a result of a bulk one-loop effect Porrati (2003). In Geng et al. (2021c) we showed that this mass causes a breakdown of the Gauss law that is essential for islands to constitute consistent entanglement wedges; in theories of gravity where the Gauss law applies, islands are inconsistent even in perturbation theory.

So we caution the reader that, in contrast to some proposals in the literature Krishnan et al. (2020); Almheiri et al. (2020c), we do not necessarily view the non-gravitating bath that appears in this paper as a model for the region far from a black hole in asymptotically flat space. Moreover, the graviton that is localized on the brane is massive. Therefore the Page curve that we discuss below should not be thought of as the Page curve for Hawking radiation in a standard theory of gravity. Our results

have a clean interpretation in terms of the entanglement entropy of subregions in the nongravitating BCFT—which is a problem of independent interest and one for which the Page curve is relevant—and we use the bulk theory of gravity simply as a tool to perform this computation.

**Overview of the Paper.** We first discuss the geometrical aspects of our model in Section 4.2, starting with a review of double holography and then delving into the types of extremal surfaces that appear in the black string with a KR brane present. We then discuss the numerics in Section 4.3, presenting our results for the phase structure (summarized above) in detail. We conclude with some discussion of our results and potential future directions in Section 4.4.

## 4.2 The Doubly Holographic Black-String Model

This section describes general features of the doubly holographic black-string model, illustrated in Figure 4.1. The first two subsections are mostly qualitative and we refer the reader to Geng et al. (2021b) for additional technical details. Readers who are already familiar with the literature may want to skip the overview of doubly holographic models in Section 4.2.1. The main point is to review the doubly holographic setup and point out that transparent boundary conditions make the graviton on the KR brane massive.

Section 4.2.2 provides an overview of the doubly holographic black-string model, illustrated in Figure 4.1, and determines the area functional for the codimension-2 surfaces. Section 4.2.3 explains how boundary conditions distinguish the extremal surfaces into two distinct classes. These are the *Hartman-Maldacena (HM) surfaces* Hartman and Maldacena (2013), which avoid the branes and traverse the Einstein-Rosen bridge, and *island surfaces*, which connect the boundary to the brane and avoid the Einstein-Rosen bridge. Due to the growth of the Einstein-Rosen bridge, the former grow with boundary time and the latter do not. The interplay between

these two classes of surfaces ensures the growth of the entanglement entropy will be consistent with unitarity, provided the island surfaces exist. In Section 4.2.4 we review recent work performed independently by our group Geng et al. (2021b) and others Chen et al. (2020a) that revealed that island solutions sometimes *fail to exist* when the boundary has more than two spatial dimensions. This potentially leads to a serious puzzle in any braneworld model.

#### 4.2.1 Review of Doubly Holographic KR Models

While KR models have recently become more visible, they have been well understood for some time as having three equivalent descriptions Karch and Randall (2001b,c); Almheiri et al. (2020e); Takayanagi (2011b); Fujita et al. (2011):

- **Boundary:**  $d$ -dimensional CFT with a  $(d - 1)$ -dimensional defect (BCFT $_d$  Cardy (2004); McAvity and Osborn (1995)).
- **Bulk:** Einstein gravity in an asymptotically AdS $_{d+1}$  spacetime containing an AdS $_d$  Karch-Randall (KR) brane Karch and Randall (2001b).
- **Intermediate:**  $d$ -dimensional CFT coupled to “localized” gravity on the AdS $_d$  brane Karch and Randall (2001b), with transparent boundary conditions between the brane at infinity and a nongravitating CFT $_d$  (the *bath*) on half space.

The transparent boundary conditions are responsible for an important and sometimes overlooked result, which is that the graviton on the KR brane’s gravity theory picks up a mass Porrati (2002b); Porrati and Starinets (2002); Porrati (2003) from energy exchange. In the dual “boundary” description, the transparent boundary conditions leads to the nonconservation of the stress tensor at the BCFT defect, which causes the stress tensor to pick up an anomalous dimension corresponding to a mass for the graviton Aharony et al. (2006). The graviton’s mass is roughly quadratic in brane angle for small angles (the “near-critical” regime) Miemiec (2001b); Schwartz (2001)

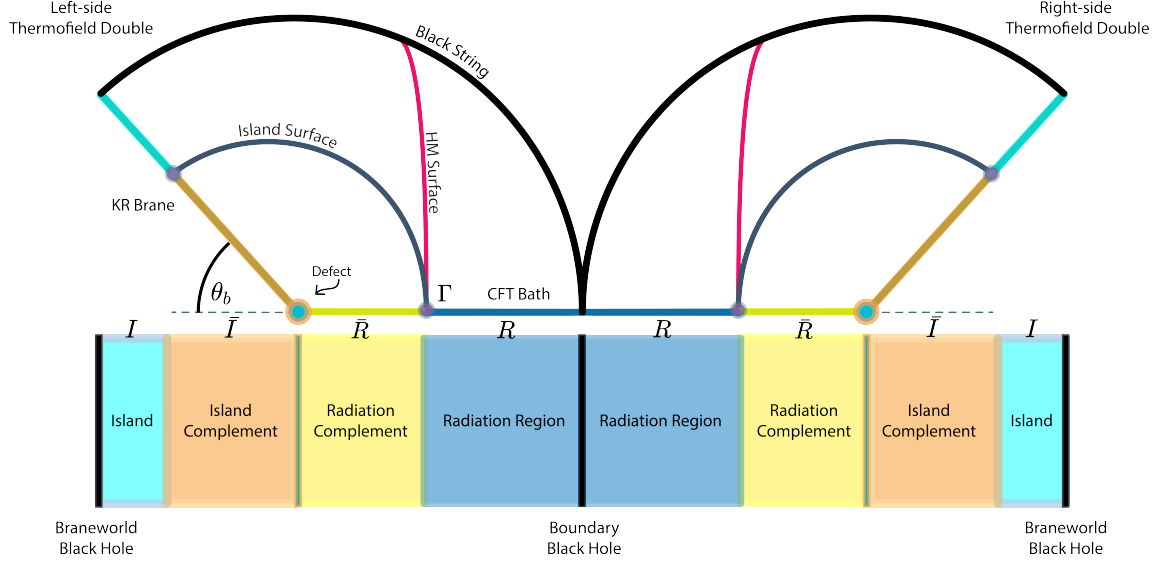


Figure 4.1: This cartoon illustrates the doubly holographic black-string model, where the region behind the KR brane has been excised. There are three perspectives, with the  $\text{AdS}_{d+1}$  gravitating bulk illustrated at the top of the diagram. The  $d$ -dimensional BCFT and  $d$ -dimensional  $\text{AdS}_d$  brane are depicted in colors which correspond with the intermediate picture indicated at the bottom of the diagram. Notice the island region envelops the black-hole horizon on the brane, and the defect cannot be reached from the island in the intermediate picture without crossing through its complement. The two parameters in the theory are the anchor point  $\Gamma$  and brane angle  $\theta_b$ . Note that “left/right” has a different meaning than in our previous paper Geng et al. (2021b).

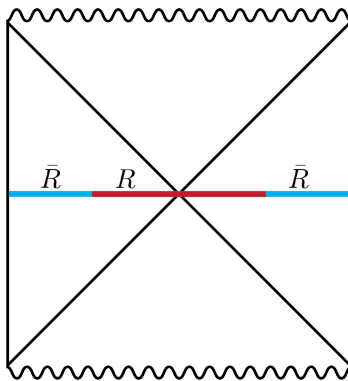


Figure 4.2: The bulk gravitational system we study in this paper is dual to a BCFT thermal state on a nongravitating black-hole background. We compute the entropy of the subregion  $\mathcal{R}$ , a connected interval that runs through the *bifurcation surface* (the horizon at  $t = 0$  separating the past and future horizons, represented by the point in the middle of the Penrose diagram above), using the holographic RT prescription. This bipartition is similar to that in a recent work Geng et al. (2021f) studying entanglement islands in de Sitter space.

and comparable to the  $\text{AdS}_{d+1}$  length scale when the angles are not parametrically small.

As discussed in Geng and Karch (2020b), all reliable calculations of the semi-classical Page curve in more than  $1 + 1$  spacetime dimensions ( $d > 2$ ) have been performed in systems with massive gravitons. It was further explained in Geng et al. (2021c) that this mass plays an essential role in enabling islands to constitute a consistent entanglement wedge. Consistent with these statements, the setup in which we study islands in this work features massive gravity on the brane.

#### 4.2.2 The Black-String Model

This section provides an overview of the doubly holographic black string illustrated in Figure 4.1. Starting with a  $(d + 1)$ -dimensional black string, we embed a  $d$ -dimensional KR brane Randall and Sundrum (1999d); Karch and Randall (2001b). By the AdS/BCFT correspondence Takayanagi (2011b); Fujita et al. (2011), the setup is dual to a thermal BCFT state on a black hole background whose degrees

of freedom Affleck and Ludwig (1991); Azeyanagi et al. (2008) are counted by the brane tension. The intermediate picture is a gravitating black hole coupled to a nongravitating black-hole bath.<sup>4</sup>

The planar black string (with the  $\text{AdS}_{d+1}$  radius set to 1) is described by the metric,

$$\begin{aligned} ds^2 &= \frac{1}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + d\vec{x}^2 + u^2 d\mu^2 \right], \\ h(u) &= 1 - \frac{u^{d-1}}{u_h^{d-1}}, \end{aligned} \tag{4.1}$$

where  $t \in \mathbb{R}$ ,  $u > 0$ ,  $0 < \mu < \pi$ , and  $\vec{x} \in \mathbb{R}^{d-2}$ . The blackening factor  $h(u)$  gives a black hole on each constant  $\mu$  slice—one such slice will be the end-of-the-world KR brane.

We briefly mention that one could also consider the global version, in which case the metric would be,

$$\begin{aligned} ds^2 &= \frac{1}{u^2 \sin^2 \mu} \left[ -h_G(u) dt^2 + \frac{du^2}{h_G(u)} + d\Omega_{d-2}^2 + u^2 d\mu^2 \right], \\ h_G(u) &= 1 + u^2 - \frac{u^{d-1}}{u_h^{d-1}}. \end{aligned} \tag{4.2}$$

In the global metric, instead of the transverse coordinates comprising  $\mathbb{R}^{d-2}$ , they form a  $(d-2)$ -sphere with a line element proportional to  $d\Omega_{d-2}$ . The dual boundary state is a finite temperature CFT state on a sphere. However, the global black string features a Gregory-Laflamme instability Gregory and Laflamme (1993). This thermodynamic instability was identified by Chamblin and Karch (2005) with a Hawking-Page transition on the brane controlled by a dimensionless combination of the horizon temperature and the boundary sphere’s radius. In this paper we will consider *only* the planar black string, for which the Hawking-Page transition is absent.

---

<sup>4</sup>One can also start with empty  $\text{AdS}_{d+1}$  instead of the black string Chen et al. (2020b). This would produce an analogous picture lacking any black holes.



The brane is introduced by adding the standard Randall-Sundrum (RS) term Randall and Sundrum (1999d) to the Einstein action, and then restricting to subcritical brane tensions. The brane is parametrized as  $\mu = \theta_b$ . The model dual to BCFT is then obtained by treating this as an *end-of-the-world brane* and restricting the  $\mu$  coordinate to  $\mu \in [\theta_b, \pi)$ .<sup>5</sup>

Our main objective is to use the RT prescription to calculate the entanglement entropy between a BCFT subregion  $\mathcal{R}$  and its complement  $\bar{\mathcal{R}}$  induced by a generic bipartition of the initial time ( $t = 0$ ) slice, as depicted in Figure 4.2. From here the prescription is to determine the minimal extremal surfaces  $\Sigma$  that are either homologous to  $\mathcal{R}$  or terminate on the end-of-the-world brane Fujita et al. (2011); Almheiri et al. (2019c), with the latter leading to islands on the brane Penington (2019); Almheiri et al. (2019a). First we consider a generic interval on the brane  $\mathcal{J}$ —a “candidate” island.<sup>6</sup> We then determine extremal surfaces  $\Sigma$  that satisfy the homology constraint,

$$\partial\Sigma = \partial\mathcal{R} \cup \partial\mathcal{J}. \quad (4.3)$$

The next step is to apply a variational principle to the *area density functional*  $\mathcal{A}$ ,

$$\mathcal{A} = \int_{\theta_b}^{\pi} \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + \frac{u'(\mu)^2}{h(u)}}, \quad (4.4)$$

to find the  $\Sigma$  and corresponding  $\mathcal{J}$  for which the area is minimized. Note that  $u_h$  can be brought to the front of the action as an overall prefactor  $u_h^{2-d}$  in planar coordinates. For simplicity we choose  $u_h = 1$  for all of our numerical calculations.

The action (7.3) is extracted from the constant- $t$  slices of the line element (7.1) as the area density of a surface  $u = u(\mu)$  with respect to the transverse space  $\vec{x}$ .<sup>7</sup> This determines the area of the surface up to a factor, proportional to the volume of the suppressed dimensions, that does not affect the phase structure.

---

<sup>5</sup>The tension of the brane is known to be  $(d-1)\cos\theta_b$  Takayanagi (2011b); Fujita et al. (2011).

<sup>6</sup> $\Sigma$  simply being homologous to  $\mathcal{R}$  corresponds to  $\mathcal{J} = \emptyset$ , i.e. the empty set.

<sup>7</sup>For the rest of this paper we will use area density and area interchangeably.

The Euler-Lagrange equation for the action (7.3), which we also call the “equation of motion,” is an ordinary differential equation because of the parameterization  $u = u(\mu)$ . Furthermore, we ensure that the boundary terms in the variation of  $\mathcal{A}$  vanish by imposing boundary conditions on  $\Sigma$ . As discussed in Geng et al. (2021b), we impose a *Dirichlet condition* at the conformal boundary and a *Neumann condition*—requiring that the first derivative  $u'$  vanishes—at the brane. The latter tells us that such surfaces must anchor onto the brane at right angles. Solving the Euler-Lagrange equation for the second derivative  $u''$  gives,

$$u'' = -(d-2)u h(u) + (d-1)u' \cot \mu \left( 1 - \frac{\tan \mu}{2} \frac{u'}{u h(u)} + \frac{u'^2}{u^2 h(u)} \right) - \left( \frac{d-5}{2} \right) \frac{u'^2}{u}. \quad (4.5)$$

The nonlinear first term leads directly to nontrivial physics below the critical angle that we explore in Section 4.2.4. Figure 4.4 reviews how the critical angle depends on the number of dimensions for the space. Note that the critical angle vanishes for  $d = 2$ —the first term in (7.2.3) vanishes and the trajectories of the RT surfaces become circular geodesics that exist for all brane angles.

While we have so far presented the equations for general dimension, we work exclusively with  $d = 4$  in the remaining sections of this paper.

### 4.2.3 Different Classes of RT Surfaces

We now take a closer look at the Dirichlet and Neumann boundary conditions that play an important role in the construction of the RT surfaces. The relevant surfaces are illustrated in Figure 4.3. The two parameters of interest are  $\Gamma$ , which is the endpoint of the relevant subregion  $\mathcal{R}$ ,<sup>8</sup> and  $\theta_b$ , the angle between the brane and the excised boundary region. There are two classes of surfaces because, while all surfaces

---

<sup>8</sup>Technically, there are two endpoints, with one on each side. This subtlety is particularly important for the surfaces explored in Section 4.2.3.1, but for now it suffices to say that we assume that  $\mathcal{R}$  is symmetric with respect to the boundary black hole’s bifurcation surface.

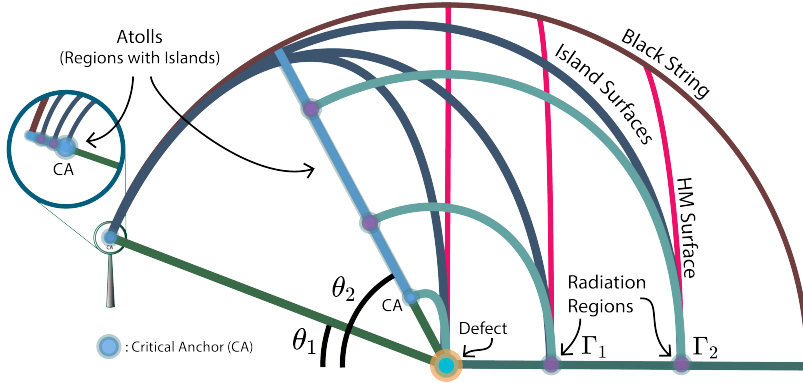


Figure 4.3: This cartoon illustrates the relevant RT surfaces for different brane angles  $\theta_{1,2}$  below the critical angle. Island surfaces end on the brane at right angles, and HM surfaces cross the black-string horizon and traverse the Einstein-Rosen bridge. Islands always attach to the brane inside the atoll, which begins at a point called the *critical anchor* and extends to the black-hole horizon. Decreasing the brane angle reduces the size of the atoll, which shrinks to the horizon as  $\theta \rightarrow 0$ . Increasing the size of  $\mathcal{R}$ , bounded by  $\Gamma$ , pushes the anchors (in purple) toward the black-hole horizon; increasing the brane angle pushes the anchors toward the defect. The critical anchor, which is the point on the brane that determines the size of the largest island on the brane (which always corresponds to  $\Gamma = 0$ ), coincides with the defect above the critical angle. See Figure 4.5 for numerical results.

obey Dirichlet boundary conditions at  $\Gamma$ , only some surfaces connect to the brane where the boundary condition must be Neumann by diffeomorphism-invariance.<sup>9</sup>

The Hartman-Maldacena (HM) surfaces avoid the branes as they plunge from the boundary into the black-string horizon, pass through the Einstein-Rosen bridge to the other side, and anchor to the boundary of the thermofield double space. As the Einstein-Rosen bridge grows, the HM surfaces are stretched out behind the horizon, and their areas grow Hartman and Maldacena (2013).

The counterparts to the HM surfaces are the “island surfaces” that connect the brane to the conformal boundary. Since these avoid the Einstein-Rosen bridge, their areas are constant in time. Near the conformal boundary they are roughly the same

<sup>9</sup>It is mathematically possible to impose a Dirichlet boundary condition on the brane Ghosh and Krishnan (2021); the Neumann condition is motivated by the presence of localized gravity.

as the HM surfaces at  $t = 0$ , since they also obey the Dirichlet condition there by virtue of being extremal. However, the island surfaces satisfy a Neumann boundary condition at the brane.

The interplay between these two surfaces makes it possible for the entanglement entropy to obey unitarity. Essentially, island surfaces prevent the entropy from growing forever, since they are constant in time and always dominate (i.e. are minimal) at sufficiently late times. Specifically, even if an island surface is subdominant on the initial time slice to an HM surface in the entanglement entropy, the latter will eventually outgrow the former.

This phase transition of surfaces takes place at the *Page time*. Its value depends on parameters that are clearly connected to the physics of the BCFT state; these are  $\theta_b$ , which determines the number of degrees of freedom on the defect Takayanagi (2011b); Fujita et al. (2011), and  $\Gamma$ , which determines the size of  $\mathcal{R}$ . The Page time is roughly proportional to the area difference between the RT surfaces at  $t = 0$ .

#### 4.2.3.1 Hartman-Maldacena (HM) Boundary Conditions

Here we look at the Hartman-Maldacena surfaces in more detail. An alternative  $\mu = \mu(u)$  parameterization for the area functional is sometimes useful:

$$\mathcal{A} = \int \frac{du}{(u \sin \mu)^3} \sqrt{\frac{1}{h(u)} + u^2 \mu'(u)^2} . \quad (4.6)$$

Since the HM surface traverses the Einstein-Rosen bridge, we must specify the location of  $\mathcal{R}$  on *both* sides of the geometry. Calling the endpoint of  $\mathcal{R}$  and its thermofield double partner  $\Gamma$  and  $\tilde{\Gamma}$  respectively, for simplicity we consider the symmetric case  $\Gamma = \tilde{\Gamma}$ . In our previous paper Geng et al. (2021b), in which we computed “left/right” entanglement entropy with a *gravitating* bath, we could consider only the case where  $\Gamma = 0$ . With the nongravitating bath, the choice of  $\Gamma$  is no longer restricted.

Choosing the symmetric case leads to a constraint for the HM surfaces on the

initial time slice—there will be a “kink” unless the surfaces cross the black-string horizon orthogonally in tortoise-like coordinates:

$$dr = \frac{du}{u\sqrt{h}}, \quad \mu'(r=0) = 0. \quad (4.7)$$

This can be translated into our  $\mu(u)$  parameterization by expanding around the horizon:

$$\mu(r) = \sum_{n=0}^{\infty} \mu_n r^n \sim \sum_{n=0}^{\infty} \mu_n (u_h - u)^{n/2}. \quad (4.8)$$

This expansion has to be done in half-integer powers of  $(u_h - u)$  because  $r \sim \sqrt{u_h - u}$ .

The condition that  $\mu'(r=0) = 0$  in the tortoise-like coordinates implies the absence of a term linear in  $r$ . Thus the  $\sqrt{u_h - u}$  term in the expansion in the  $u$  coordinate vanishes, so  $\mu(u)$  becomes a regular power series. Imposing the equation of motion in terms of  $\mu(r)$  now fixes  $\mu'(u_h)$  in terms of  $\mu(u_h)$ . So even though it naively looks like we imposed only one boundary condition at  $u_h$ , regularity as implied by the symmetric ansatz tells us that the solution is completely fixed by simply specifying  $\mu(u_h)$ .

To ensure that we construct an HM surface with a particular value of  $\Gamma = u/u_h$ , we need to solve the shooting problem with boundary conditions:

$$\mu(u_h \Gamma) = \pi, \quad \mu(u_h) = \theta_{HM}, \quad \mu'(u_h) = \frac{2 \cot \theta_{HM}}{u_h}, \quad (4.9)$$

where  $\Gamma$  is the boundary of  $\mathcal{R}$  relative to  $u_h$ ,  $\theta_{HM}$  is the angle  $\mu$  where the HM surface intersects the horizon, and the prime indicates a derivative with respect to the radial coordinate  $u$ . The Dirichlet condition  $\mu'(u_h \Gamma) = \infty$  at the conformal boundary is enforced automatically by the equations of motion. If one prefers the  $u(\mu)$  parameterization, one can obtain the same results by inverting the boundary conditions instead.<sup>10</sup>

---

<sup>10</sup>This involves the reciprocals of the derivatives in (7.12). This reparameterization from  $\mu(u)$  to  $u(\mu)$  is consistent because as we proved in Geng et al. (2021b) that any solution  $u(\mu)$  of (7.2.3) is a monotonic function.

#### 4.2.3.2 Island Surface Boundary Conditions

The island surfaces avoid the Einstein-Rosen bridge and terminate on the brane  $\mu = \theta_b$ . On the initial time slice they are distinguished from the Hartman-Maldacena surfaces by their resulting brane boundary conditions, which are Neumann and thus require the island surfaces to connect to the brane orthogonally. So, we need to solve the shooting problem for:

$$u(\pi) = u_h \Gamma, \quad u(\theta_b) = u_b, \quad u'(\theta_b) = 0, \quad (4.10)$$

where  $u_b$  indicates the location on the brane where the surface terminates. The condition  $u'(\pi) = 0$  is enforced automatically by the equations of motion.

Since these surfaces do not pass through the black-string horizon, they do not traverse the Einstein-Rosen bridge and their areas do not change in time. The generic late-time dominance of island surfaces over HM surfaces leads to the saturation of entanglement entropy at the Page time, *provided the island surfaces actually exist*.<sup>11</sup> But past work Chen et al. (2020a); Geng et al. (2021b) has shown these solutions *do not exist* for empty AdS below the critical angle, as we now review.

#### 4.2.4 A Potential Puzzle for Unitarity in Higher Dimensions

Recent work performed independently by our group Geng et al. (2021b), and others Chen et al. (2020a), found that island solutions *cease to exist* below a critical angle in empty  $\text{AdS}_{d+1}$ . The angle of the brane, which we reiterate encodes the number of degrees of freedom at the defect and uniquely determines the brane tension in the bulk, plays an important role in determining whether (and where) islands can exist on the brane. This potentially leads to a serious problem for unitarity in all thermal doubly holographic states—without an island surface in the bulk the HM surface would grow without bound.

---

<sup>11</sup>We define the Page time as  $t = 0$  when the island dominates at the beginning.

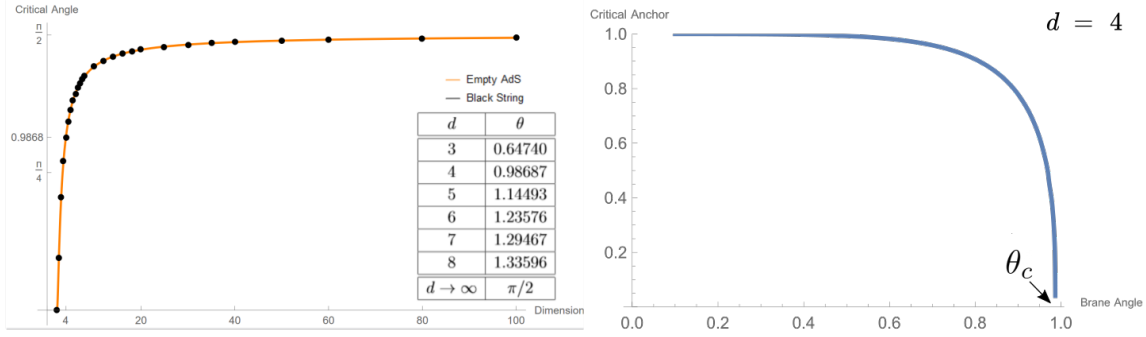


Figure 4.4: These figures were included in our previous work Geng et al. (2021b), but we repeat them here for emphasis. The critical angle is a monotonically increasing function of the number of spatial dimensions  $d$  and takes the same value for any geometry which is asymptotically  $\text{AdS}_{d+1}$ . This is distinct from the behavior of the critical anchor defining the beginning of the atoll for the black string. The critical anchor instead *decreases* monotonically with brane angle and coincides with the defect at the critical angle (about 0.98687 for  $d = 4$ ). It can be shown that the critical anchor also increases monotonically with the number of dimensions for the space.

We note that the previously observed failure of islands to exist occurs when they are not needed to preserve unitarity. Empty  $\text{AdS}_{d+1}$  with a KR brane is the prime example—RT surfaces are forbidden from traveling between the boundary and any brane below the critical angle, which means the region containing islands on the brane (the atoll) does not exist in this regime. However, since there is no Einstein-Rosen bridge there is no potential paradox involving a monotonically growing entanglement entropy and therefore there is no need for a phase transition between entanglement surfaces in such vacuum states. The nonexistence of island surfaces would, however, lead to a breakdown of unitarity in *thermal* states, such as the black-string geometry considered in this paper.

The preservation of unitarity is thus more subtle for the black-string model because it features an Einstein-Rosen bridge. We find that while island surfaces still exist below the critical angle, they connect only to a finite region of the brane—the atoll. These islands are sufficient to prevent the entanglement entropy from growing without bound, as we showed in a previous paper Geng et al. (2021b) in the context

of "left/right" entanglement, in which we are forced to  $\Gamma = 0$ . There we found a critical anchor, which is the boundary of the atoll we find in the current context. In this paper, we find a more complex phase structure associated with the freedom to choose the size of the radiation region on the boundary.

Our results show that for each choice of brane angle below the critical angle, the largest possible island is formed when the island surface begins at the critical anchor and terminates on the defect. This suffices to determine the atolls (see Figure 4.3). The critical anchor coincides with the defect above the critical angle, so the atoll completely covers the brane in that case.

### 4.3 Numerical Results for the Area Difference

Here we present our numerical results for the initial area difference between the competing island and HM surfaces as a function of the two parameters  $\Gamma$  and  $\theta_b$ . The intricacy of our results compared to those found with a gravitating bath Geng et al. (2021b) arises from the freedom to vary  $\Gamma$  on the nongravitating bath. The phase structure is determined by the sign of the UV-finite area difference,

$$\Delta\mathcal{A}(t) = \mathcal{A}_{IS} - \mathcal{A}_{HM}(t). \quad (4.11)$$

The sign is important because it tells us which type of surface is dominant in the entanglement entropy (i.e. minimal in area). The HM surface is dominant when  $\Delta\mathcal{A}(t) > 0$ , and the island surface is dominant when  $\Delta\mathcal{A}(t) < 0$ . To determine whether or not we have a Page curve at all, it is sufficient to look at the  $t = 0$  slice, since the HM surface is at its smallest in  $t$  when it crosses the bifurcation surface of the black string.

Through the boundary conditions (7.12)–(4.10), the initial area difference  $\Delta\mathcal{A}(0)$  depends on just two free parameters, the anchor point on the bath  $\Gamma$  and the angle of the brane  $\theta_b$ . We mentioned before that  $u_h$  scales out of the action, so we are free to choose  $u_h = 1$  for our numerical results. We find that  $\Delta\mathcal{A}(0)$  is finite



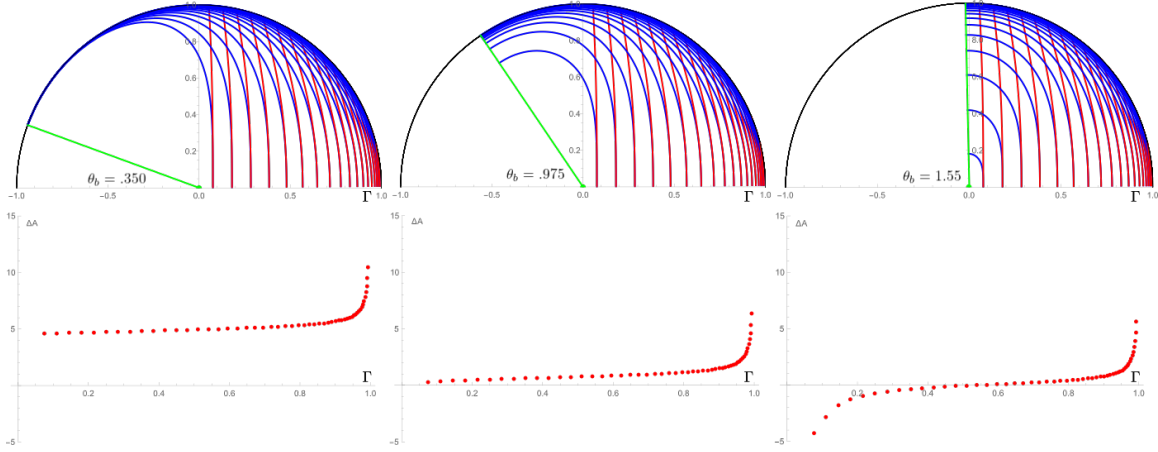


Figure 4.5: Here we see the setup together with numerical data in the planar black string. The KR brane is green, the HM surfaces are red, the island surfaces are blue, and the black-string horizon is the semicircle in black. The area difference between the island surface and the HM surface for various anchor points  $\Gamma$  is displayed in the set of second plots below. Sending  $\Gamma \rightarrow 0$  recovers the finite area differences from our previous paper Geng et al. (2021b) for angles below the critical angle. When the brane lies above the critical angle, the area difference diverges to negative infinity. These divergences are explained physically in Section 4.3.3.

when the parameters lie within the following open intervals,

$$\Gamma \in (0, 1), \quad \theta_b \in (0, \pi). \quad (4.12)$$

See Figure 4.5 for samples of our numerical data at brane angles below, at, and above the critical angle. The question of what happens when one parameter is fixed while the other varies is explored in Figures 4.6 and 4.7.  $\Delta\mathcal{A}(0)$  as a function of both  $\Gamma$  and  $\theta_b$  is plotted in 3D in Figure 4.9.

#### 4.3.1 The Difference between the Page Angle and Critical Angle

There are two special angles that play an important role in this section—the *Page angle*  $\theta_P$  and the *critical angle*  $\theta_c$  Geng et al. (2021b). For clarity, we briefly review these before presenting our numerical results.

The Page angle is the value of  $\theta_b$  at which  $\Delta\mathcal{A}(0)$  (for  $\Gamma = 0$ ) vanishes and the entanglement entropy becomes constant, and the critical angle is the slightly larger

value of  $\theta_b$  below which extremal island surfaces do not exist in empty  $\text{AdS}_{d+1}$ . Note however that the critical angle still plays a role in other geometries so long as the asymptotic region is still empty  $\text{AdS}_{d+1}$ . In particular, when the brane is below the critical angle, island surfaces approach the horizon but do not probe the asymptotic region. The islands on the braneworld are restricted to the atoll discussed in Section 4.2.4, which is only a finite portion of the brane below the critical angle, whereas above the critical angle the atoll covers the entire brane.

In empty AdS, the Page angle and the critical angle are equal, but in other geometries they are generally slightly different. The area difference vanishes slightly below the critical angle in the black string geometry because the area of the HM surface is bigger than it would be in empty AdS, but the island solution for  $\Gamma = 0$  is unchanged at the critical angle.

### 4.3.2 Numerical Results for Fixed Parameters

Here we take a closer look at the initial area differences in our numerical results, which were computed for thousands of combinations of the anchor point  $\Gamma$  and the brane angle  $\theta_b$ . We depict a representative portion of the full  $(\Gamma, \theta_b)$  parameter space.

The  $t = 0$  area differences<sup>12</sup> ( $\Delta\mathcal{A} = \mathcal{A}_{HM} - \mathcal{A}_{island}$ ) for fixed  $\Gamma$  as a function of  $\theta_b$  are illustrated numerically in Figure 4.6. The area difference decreases monotonically with the value of  $\theta_b$ , and from comparing the different curves on constant- $\theta_b$  slices it can be seen that the area difference increases monotonically with anchor point  $\Gamma$ . The area difference is negative—and thus the island surface is initially dominant—only when the brane lies above the Page angle.<sup>13</sup> Above the critical angle the area difference diverges at the edges of the intervals  $\Gamma \in (0, 1)$  and  $\theta_b \in (0, \pi)$ , while below the critical angle the area difference is finite as  $\Gamma \rightarrow 0$ . For all other values of  $\Gamma$  and  $\theta_b$  the area difference is finite.

---

<sup>12</sup>By area differences, we always mean at  $t = 0$ .

<sup>13</sup>Recall that the brane angle is measured from the excised boundary region, see Figure 4.1.

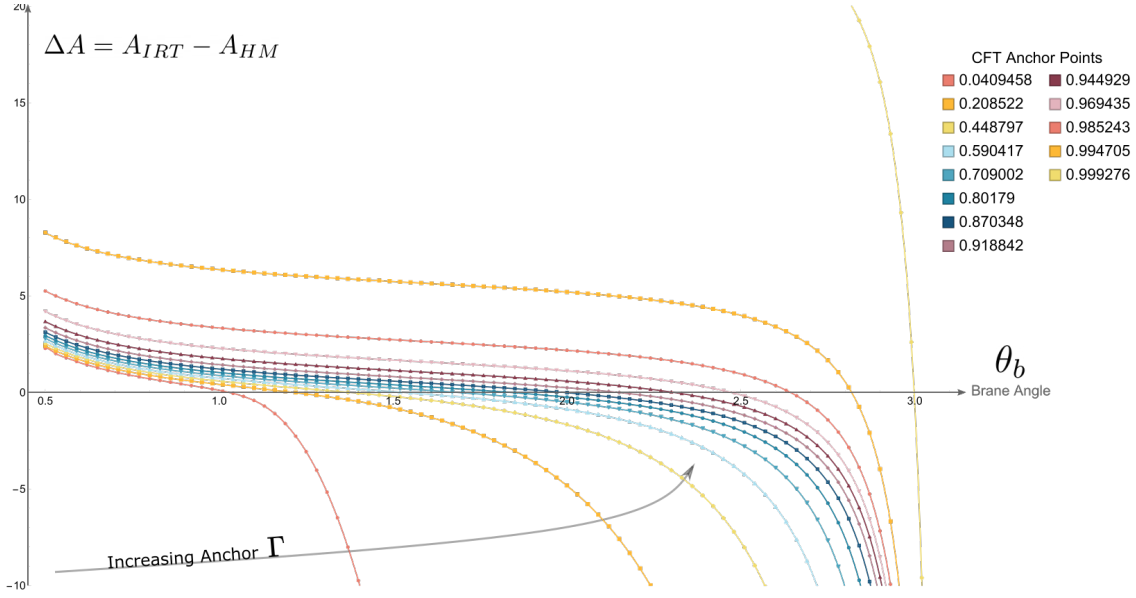


Figure 4.6: The area difference is displayed as a function of brane angle  $\theta_b$  for many possible CFT anchor points  $\Gamma$ . The area difference is clearly monotonically increasing in  $\Gamma$  and monotonically decreasing in  $\theta_b$ . The anchor points were determined numerically for many values of  $\theta_{HM}$  using the shooting method with the boundary conditions in (7.12). The area difference diverges to negative infinity as  $\Gamma \rightarrow 0$  for  $\theta_b > \theta_c$ , but when  $\Gamma \rightarrow 0$  for  $\theta_b < \theta_c$  it converges to a finite value.

Similar considerations hold for fixed brane angles  $\theta_b$  as a function of anchor point  $\Gamma$ , illustrated numerically in Figure 4.7. Again we see that the area difference increases monotonically with the value of  $\Gamma$ . From comparing the different curves on constant- $\Gamma$  slices, one can see the area difference is also monotonically decreasing in brane angle  $\theta_b$ . Again the area difference is negative only when the brane lies above the Page angle.

Note that numerical results are included in the figures only above  $\theta_b = 0.5$ . This is because the computation time for the most stable<sup>14</sup> version of the shooting method becomes very large as  $\theta_b$  becomes small. As a check, the calculation was also performed using a slightly different method for smaller brane angles. While this is not

<sup>14</sup>In order to validate our numerics, we have repeated the calculations of the area difference for a large range of different cutoffs. As we show in Appendix 4.B, our results are very stable.

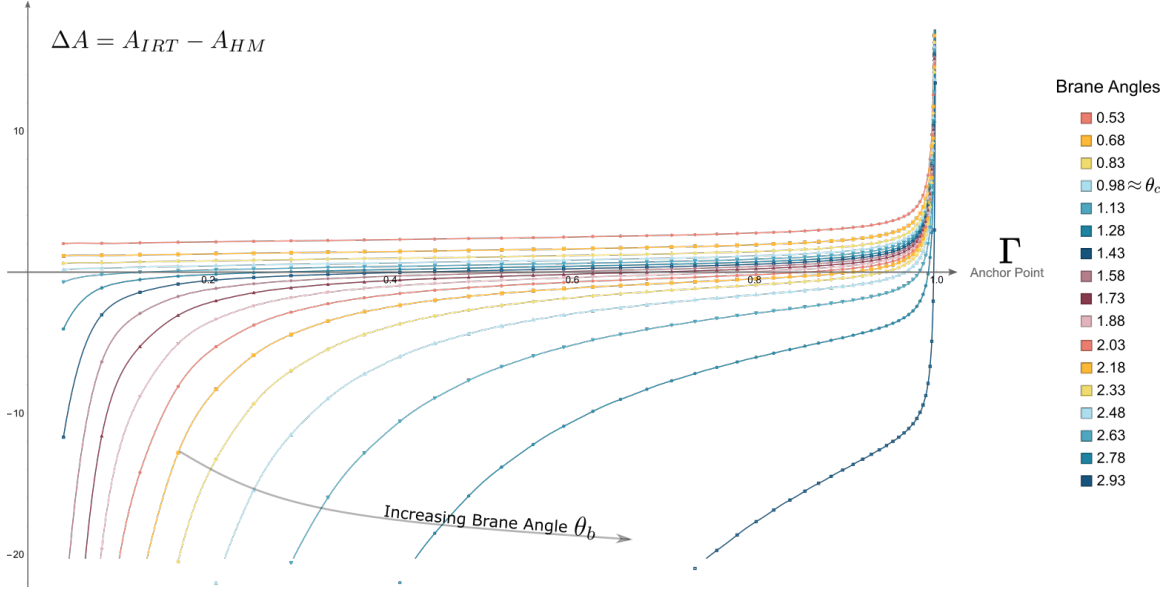


Figure 4.7: The area difference is displayed as a function of CFT anchor point  $\Gamma$  for many possible brane angles  $\theta_b$ . The area difference is clearly monotonically increasing in  $\Gamma$  and monotonically decreasing in  $\theta_b$ . The islands become dominant when the area difference is negative, which is possible only above the Page angle. The Page angle Geng et al. (2021b) is very slightly less than the critical angle ( $\theta_c \approx 0.98687$  in  $d = 4$ ). The area difference diverges to negative infinity as  $\Gamma \rightarrow 0$  for  $\theta_b > \theta_c$ , but when  $\Gamma \rightarrow 0$  for  $\theta_b < \theta_c$  it converges to a finite value. We explain in Section 4.3.3 that the area difference diverges to positive infinity as  $\theta_b \rightarrow 0$ .

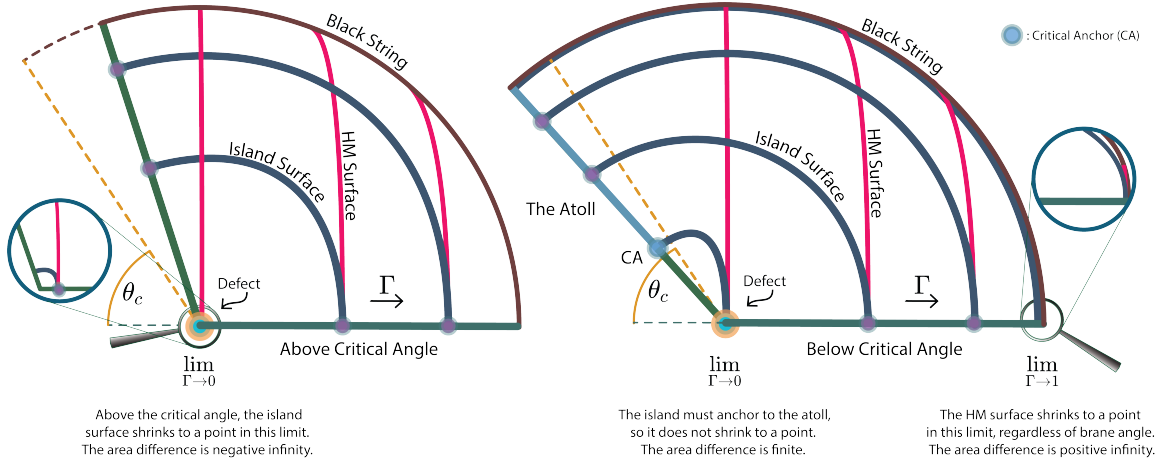


Figure 4.8: This cartoon illustrates how the area difference is affected by each limiting case for the boundary anchor point  $\Gamma$ , with angles below the critical angle on the right, and angles above the critical angle on the left. Taking the difference between the surfaces essentially acts as a renormalization scheme. The area difference between the island surface and HM surface diverges when one of them shrinks to a point because the difference between the surfaces reaches farther into the boundary region. When the brane lies below the critical angle, the island surface cannot shrink to a point when  $\Gamma \rightarrow 0$  because it anchors to the brane inside the atoll.

visible in the figures, we found that the area difference diverges to positive infinity as  $\theta_b \rightarrow 0$ . These numerical results are found to match our earlier work Geng et al. (2021b) in the limit where  $\Gamma \rightarrow 0$ .

### 4.3.3 The Geometric Origin of the Divergences

It is simple to understand the diverging area difference for parameters at the edges of the intervals  $\Gamma \in (0, 1)$  and  $\theta_b \in (0, \pi)$  geometrically. We reiterate that the surfaces themselves have been illustrated in Figure 4.5, the area differences for fixed values of  $\Gamma$  and  $\theta_b$  have been given in Figures 4.6 and 4.7. The positive area differences are presented in 3D in Figure 4.9.

#### 4.3.3.1 Positive Divergence

First we consider each case where the area difference diverges to positive infinity, meaning the island surface has infinitely bigger area than the HM surface. This happens when  $\theta_b \rightarrow 0$  for each fixed  $\Gamma$  and when  $\Gamma \rightarrow 1$  for each fixed  $\theta_b$ . The former can be understood by counting how many times each RT surface hits the conformal boundary; in a single exterior patch of the black string, the island surface touches the boundary twice while the HM surface touches it only once. The latter case where  $\Gamma \rightarrow 1$  is illustrated in Figure 4.8. Since the HM surface shrinks to a point on the boundary, it only serves to “cut off the end” of the island surface, so the area difference diverges to positive infinity.

The divergence can be understood as follows. As we approach the limiting value  $\Gamma = 1$ , we are effectively computing the entropy of a small strip on the boundary (see Figure 4.15 in Appendix 4.A). In this limit, the curvature of the boundary black hole is not important so the HM surface reduces to the familiar RT surface for a small strip in empty AdS. It is well known that the renormalized entanglement entropy of a small strip—and, correspondingly, the renormalized area of the HM surface—tends to  $-\infty$  in this limit.<sup>15</sup> Therefore the area difference between the area of the island surface and the HM diverges.

#### 4.3.3.2 Negative Divergence

Infinite *negative* area differences do not lead to eternally-increasing entropy, since island surfaces start out dominant on the initial time-slice for *every* negative area difference. The HM surface starts out infinitely larger than the island surface, so the entropy is simply constant and lacks a phase transition. But we can still consider the cases leading to such divergences.

The area difference diverges to negative infinity whenever the island surface

---

<sup>15</sup>See the second term in (4.15) in Appendix 4.A; this term is the renormalized entanglement entropy

shrinks to a point, as illustrated in Figure 4.8; this takes place when  $\Gamma \rightarrow 0$  for each fixed  $\theta_b$  above the critical angle  $\theta_c$ . Note that these are the “tiny” island surfaces found in Geng et al. (2021b).

There are also some nonphysical limits which give diverging negative area differences, for example in the limit where  $\theta_b \rightarrow \pi$ , but note the brane tensions are negative when  $\theta_b > \pi/2$ . For example, the area difference also diverges to negative infinity when  $\theta_b \rightarrow \pi$  for a fixed  $\Gamma$ , with the caveat that the associated HM surface takes a nonphysical path *through* the brane and into the excised region.<sup>16</sup> Fortunately, it turns out that the HM surfaces never travel through the brane when the area difference is positive,<sup>17</sup> and the critical angle *never* takes a value greater than  $\pi/2$  for any dimension.

We have seen what happens when  $\theta_b \rightarrow \pi$  and when  $\Gamma \rightarrow 1$  separately, but it is interesting to note what happens if both limits are applied simultaneously. This is a sensible procedure because you can keep  $\theta_{HM} > \theta_b$  while taking the limit. It turns out that the area difference vanishes, which can be seen geometrically because the boundary conditions for the HM surface and island surface become degenerate. This can be seen in (7.12); when the parameterization is switched over to  $u(\mu)$ , the boundary condition for the HM surface becomes,

$$u'(\mu)|_{u=u_h} = \frac{u_h}{2 \cot \theta_{HM}}. \quad (4.13)$$

Recalling that  $\theta_{HM}$  is the polar angle  $\mu$  where the HM surface crosses the black-string horizon, and considering the geometry of the planar black string, we see that  $\theta_{HM} \rightarrow \pi$  is the same limit as  $\Gamma \rightarrow 1$ . When this limit is taken, the boundary conditions in (4.10) and (4.13) match, the surfaces are the same, and the area difference vanishes.

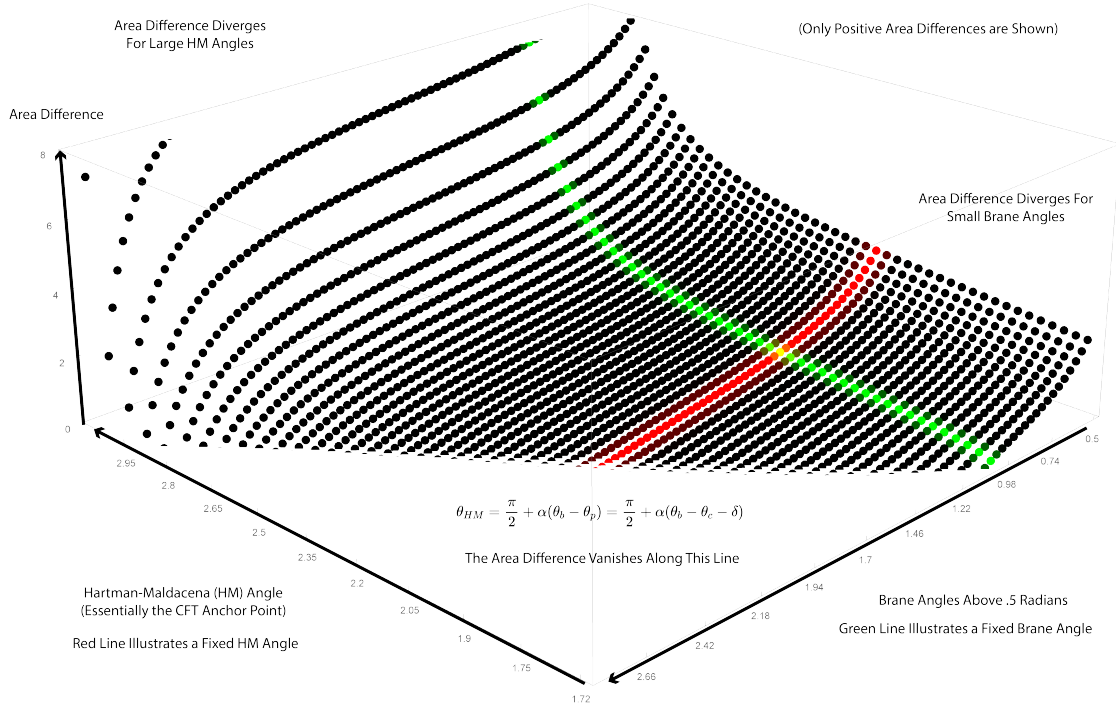


Figure 4.9: Here we present the positive area difference as a function of  $\theta_b$  and  $\theta_{HM}$ . The parameter subspace for which the area difference vanishes is an approximately straight line. Negative brane tensions are most likely not physical but they have been included without much trouble. This figure is more complete than Figures 4.6 and 4.7, which simply correspond to slices along the  $x$  and  $y$  axes, with the main difference being that  $\Gamma$  has been swapped out for  $\theta_{HM}$ . Note the area difference increases monotonically with  $\theta_{HM}$  and decreases monotonically with  $\theta_b$ , and that the area difference vanishes along an approximately straight line.



#### 4.3.4 Symmetry in Brane Angle and HM Angle

We observe a striking symmetry between Figure 4.6 and Figure 4.7 which can be made more apparent by replacing  $\Gamma$  with the HM angle  $\theta_{HM}$ , which we recall is the polar angle where the HM surface strikes the event horizon. This redefinition is not as arbitrary as it may seem at first glance. Since the boundary condition is given in polar coordinates, the coordinate transformation is natural from the perspective of solving the shooting problem in (7.12).

From looking at Figure 4.9, one can see the area difference is roughly symmetric in brane angle and HM angle; this can be made more apparent by including  $\theta_b < 0.5$  but at the cost of some accuracy. Only positive tension branes are considered on the left, but negative tension branes have been included for completeness on the right.

The symmetry can be used to write an approximately linear constraint on the region of parameter space for which the area difference is positive, with  $\Delta\mathcal{A}(0) = 0$  at equality,

$$\theta_{HM} \geq \frac{\pi}{2} + \alpha(\theta_b - \theta_c - \delta) \iff \Delta\mathcal{A}(0) \geq 0, \quad (4.14)$$

where  $\theta_c$  is the critical angle,  $\theta_b$  is the brane angle,  $\delta \equiv \theta_P - \theta_c$  ( $\delta \approx -0.02$  in  $d = 4$ ), and  $\alpha \approx 0.75$  in  $d = 4$ . The value of  $\alpha$  generically depends on  $d$ ; it equals unity for  $d = 2$  and vanishes as  $d \rightarrow \infty$ .

Choosing  $\Gamma = 0$  recovers the result of Geng et al. (2021b), since the area difference vanishes at the Page angle  $\theta_P = \theta_c + \delta$  for  $\Gamma = 0$ . Furthermore, the area difference vanishes when  $\theta_{HM} = \theta_b = \pi$ ; in that case it can be seen that the surfaces are the same, since the HM and island surface boundary conditions become degenerate in that limit.

---

<sup>16</sup>Such HM surfaces do not satisfy the homology constraint in any case in the bulk spacetime cutoff by the KR brane.

<sup>17</sup>This was also noted in Geng et al. (2021e) in a lower dimensional setup together with a precise field theory calculation.

It can be seen in this equation that, as we mentioned before, the area difference is always positive when  $\theta_{HM} \geq \theta_b$ . This serves as a useful consistency check, since the HM surface never travels through the brane for positive area differences, and the area difference vanishes when  $\theta_{HM} = \theta_b = \pi$ .

### 4.3.5 Phase Structure and the Constant Entropy Belt

Here we explain our main result, the phase structure for the black-string model with a nongravitating bath as a function of the anchor point  $\Gamma$  and the brane angle  $\theta_b$ . Figure 4.10 illustrates the phase transition that occurs for a given  $\theta_b > \theta_P$  when  $\Gamma$ , which separates  $\bar{\mathcal{R}}$  from  $\mathcal{R}$  in Figure 4.2, becomes large enough for the entropy to be described by a time-dependent Page curve. When discussing Figures 4.6 and 4.7 we pointed out that the area difference is monotonically decreasing in  $\theta_b$  and monotonically increasing in  $\Gamma$ . This behavior manifests in the phase diagram.

It can be seen in Figure 4.9 and Figure 4.10 that for a given value of  $\Gamma$  the initial area difference is negative (thus the island surface is dominant) if  $\theta_b$  is large enough. This is precisely the condition (4.14). In such cases the entanglement entropy does not grow with time, since the island surface does not traverse the Einstein-Rosen bridge. In the limit  $\Gamma \rightarrow 0$ , or equivalently  $\theta_{HM} \rightarrow \pi/2$ , the area difference vanishes at the Page angle Geng et al. (2021b).

The above properties make it possible to nearly summarize the phase structure in terms of what we call the *constant entropy belt*, illustrated in Figure 4.12. The definition is that the entropy curve is constant, for a given  $\theta_b$ , when  $\Gamma$  lies within the constant entropy belt. In all such cases the area difference is negative, which means the island surface will be dominant on the initial time slice, and therefore the entanglement entropy will be constant in time.

The belt does not exist when  $\theta_b$  is below the Page angle because the area difference is positive for all  $\Gamma$ . The belt first forms at the Page angle, only very slightly below the critical angle, so there is a small parametric range  $\theta_b \in (\theta_P, \theta_c)$  for

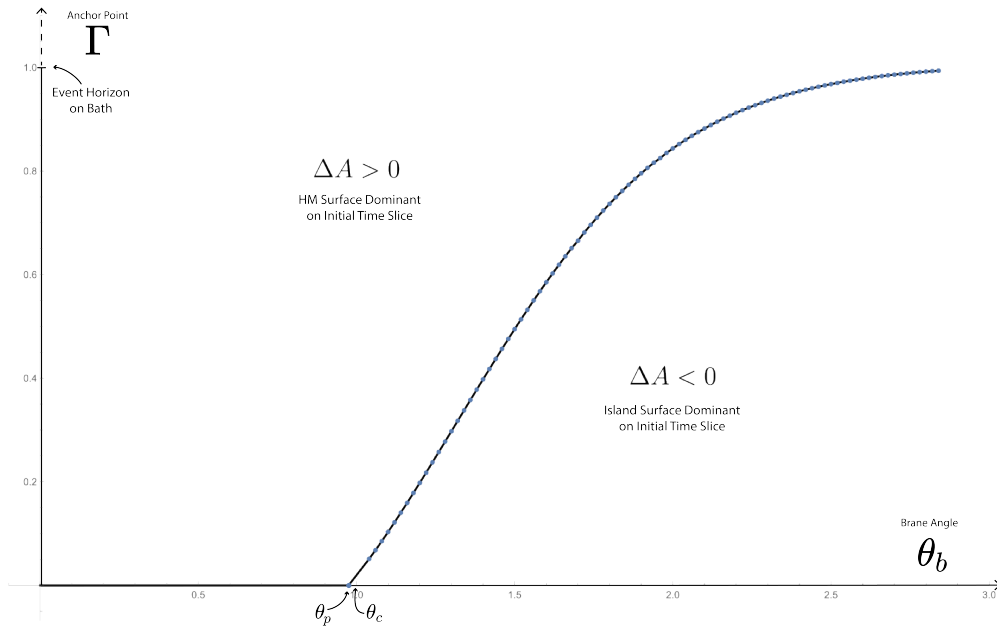


Figure 4.10: The phase diagram depends on the anchor point  $\Gamma$  and the brane angle  $\theta_b$ . When the brane lies above the Page angle, the island surfaces are dominant when  $\Gamma$  is sufficiently small; as the brane angle increases the island surfaces are dominant for increasingly large values of  $\Gamma$ . For anchor points outside the boundary black hole's horizon ( $\Gamma < 1$ ), the island surfaces become dominant as we take  $\theta_b \rightarrow \pi$ .

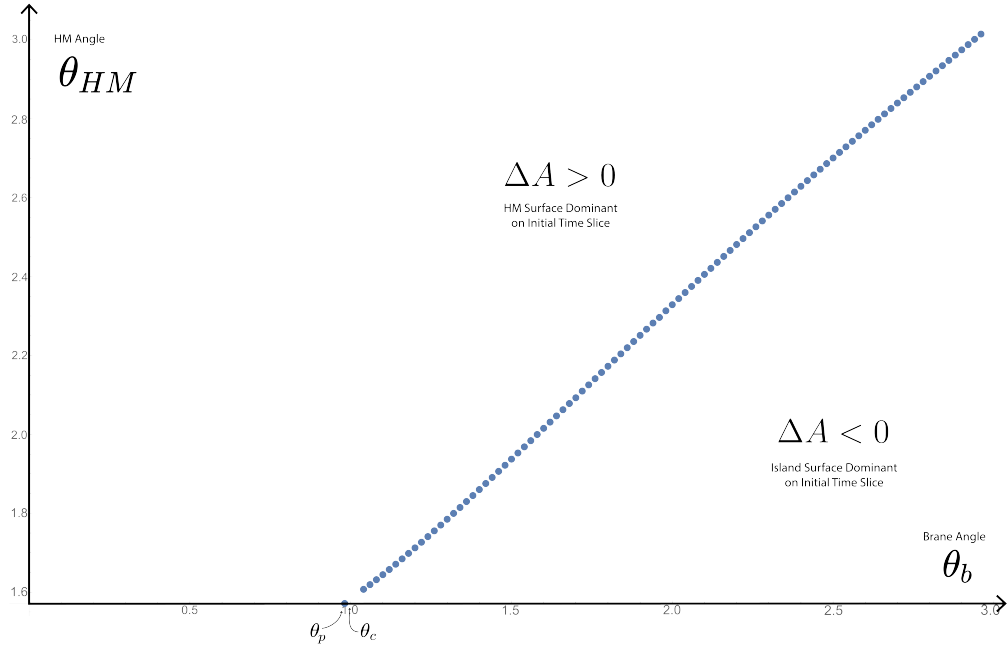


Figure 4.11: The phase diagram is approximately linear when the anchor point  $\Gamma$  is exchanged for the HM angle  $\theta_{HM}$ ; this can also be seen in Figure 4.9. The phase diagram is otherwise the same as in Figure 4.10, with  $\theta_{HM} = \pi$  corresponding to  $\Gamma = 1$  and  $\theta_{HM} = \pi/2$  corresponding to  $\Gamma = 0$ .

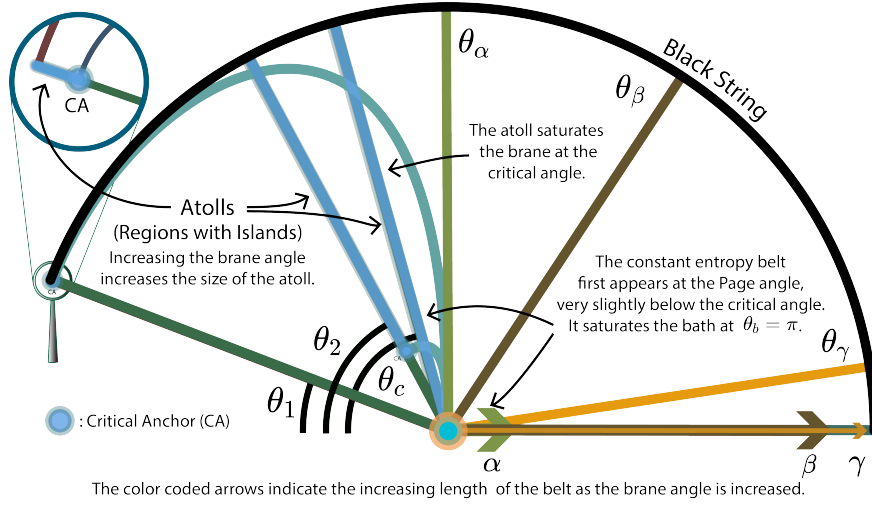


Figure 4.12: When an RT surface anchors within the constant entropy belt on the bath, the island surfaces are dominant on the initial time slice and the entanglement entropy is constant in time. Decreasing the brane tension  $(d-1)\cos\theta_b$  increases the width of the belt, indicated by the color-coded arrows on the boundary which correspond to the vertical axis in Figure 4.10. The belt first appears at the Page angle, slightly below the critical angle, and covers the entire region outside the boundary black hole horizon (at  $\Gamma = 1$ ) in the limit where  $\theta_b \rightarrow 1$ . All angles are measured from the boundary, and the tensionless limit is achieved at  $\theta_b = \pi/2$ .

which we have both a constant entropy belt and a finite critical anchor and hence a nontrivial atoll. Whereas the entropy belt (on the boundary) always includes the defect, the atoll (on the brane) extends almost to the defect at the Page angle at which the constant entropy belt first appears. Since the area difference decreases monotonically with brane angle, increasing  $\theta_b$  increases the range of  $\Gamma$  values for which the entropy is constant, so the belt grows with angle and is always connected. The belt completely covers the region on the boundary outside the boundary black hole's horizon, which is located at  $\Gamma = 1$ , in the limit where  $\theta_b \rightarrow \pi$ .

Figure 4.13 illustrates one more noteworthy property of the phase structure; the initial area difference is geometrically related to the size of the island on the brane. Decreasing  $\theta_b$  or increasing  $\Gamma$  pushes the island surfaces deeper into the bulk and reduces the size of the island. The aforementioned monotonicity in those parameters

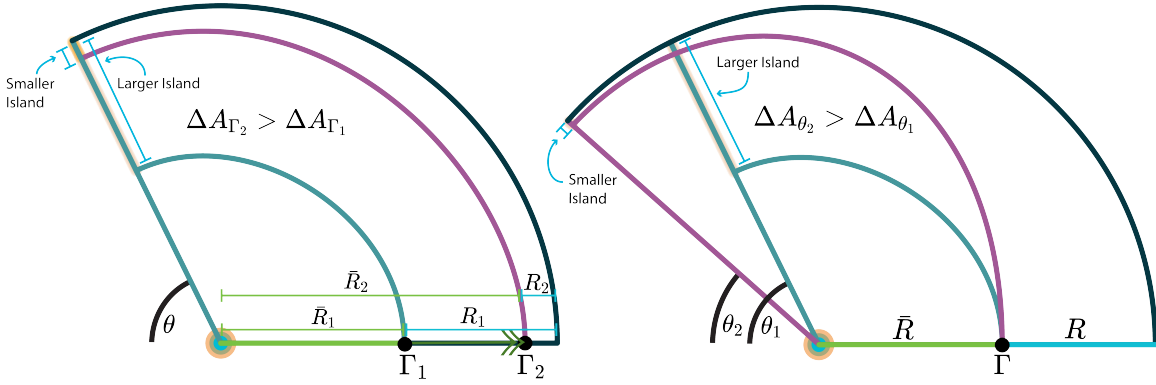


Figure 4.13: Here we illustrate the *Russian doll rule*, which states that nested surfaces can be constructed by either decreasing the brane angle  $\theta_b$  or increasing the anchor point  $\Gamma$ . This always pushes the surfaces closer to the event horizon into the bulk, decreases the size of the island on the brane, and increases the area difference between the island surface and the HM surface on the initial time slice. The area difference increases without bound as the event horizon is approached by the island surface.

leads to what we call the *Russian doll rule*, which states that the area difference always increases when  $\Gamma$  is increased, which always pushes the island surfaces closer to the horizon in the bulk. This area difference diverges to positive infinity as the black-string horizon is approached by the island surfaces. We expect that the Page time also diverges in this limit.

#### 4.3.6 Consequences for Entanglement Wedge Reconstruction

Armed with many copies of this doubly holographic model, experimenters on the boundary could go about reconstructing the bulk region by making many measurements of observables on the boundary region Jafferis et al. (2016); Faulkner and Lewkowycz (2017); Almheiri et al. (2019c). We suppose that we have two sets of experimenters who can access specific boundary degrees of freedom outside the horizon. For simplicity we will call these groups Cixin and Bindu.

The first group (Cixin) has access to boundary degrees of freedom on both sides of the Einstein-Rosen bridge shown in Figure 4.2, namely the open interval  $(0, \Gamma)$  and its thermofield double partner  $(0, \tilde{\Gamma})$ . The other group (Bindu) has access

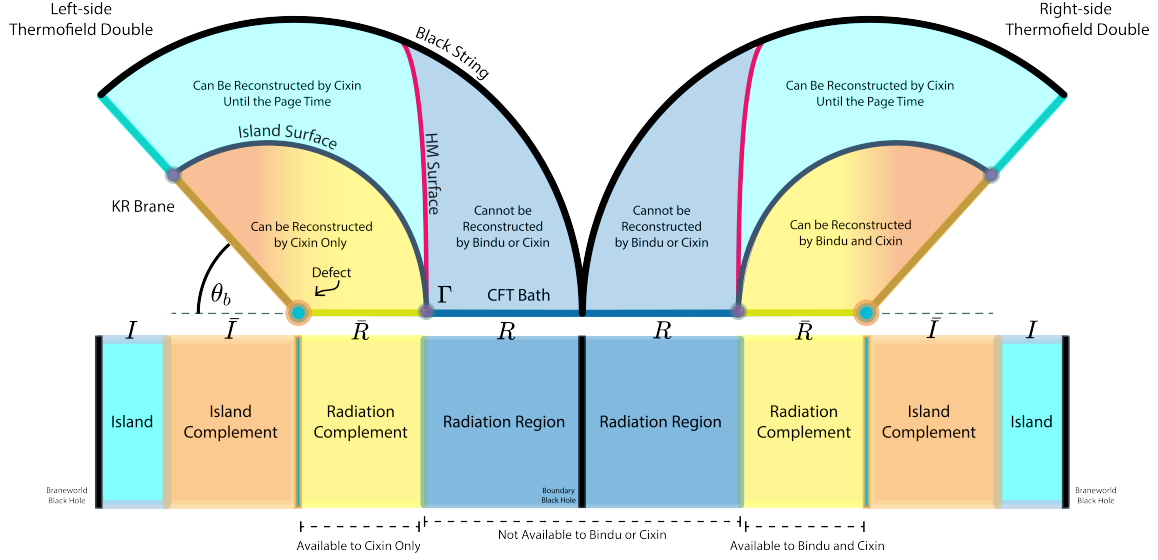


Figure 4.14: The island region can be reconstructed using boundary data in  $\bar{\mathcal{R}}$  only if both sides are accessible and only until the Page time. According to the Russian doll rule, the size of the region which is lost at the Page time increases monotonically with the brane angle  $\theta_b$  and decreases monotonically with  $\Gamma$ .

to only one side, which may as well be the right side, since  $\Gamma = \tilde{\Gamma}$  in our model.<sup>18</sup> The punchline is that Cixin can use correlations between the right and left sides of the geometry to reconstruct a larger region of the *right* side of the dual bulk geometry than Bindu, but only up until the Page time.

Neither group can reconstruct the entanglement wedge for  $\mathcal{R}$  on the boundary. Cixin can always reconstruct the entanglement wedge of the complement  $\bar{\mathcal{R}}$  that we will call the *entanglement co-wedge* of  $\mathcal{R}$ . This means she can never reconstruct the island region on the brane. Cixin can reconstruct the more encompassing co-wedge bounded by the HM surfaces. Since Bindu has access to only one side, her only option is to reconstruct the co-wedge bounded by the island surfaces on the right side of the bulk. After the Page time, which will be zero when  $\Gamma$  lies inside the constant entropy

<sup>18</sup>We could have instead considered a class of observers who have access to one side of the radiation region  $(\Gamma, 1)$ . To find the entanglement wedge for this region we would have to construct a separate extremal surface that is confined to one side of the horizon and has endpoints at  $\Gamma$  and 1.

belt, the island surfaces are dominant and the island region cannot be reconstructed. However Cixin and Bindu can regain access to the island region by increasing their value of  $\Gamma$ , and both have access to the same information in the limit where  $\Gamma \rightarrow 1$ .

There are some special considerations for higher-dimensional ( $d > 2$ ) finite temperature models. The most immediate implications concern the critical angle. For the black string, when the brane is below the critical angle, the island  $\mathcal{I}$  cannot extend beyond the critical anchor on the brane. In other words, the islands are confined to the atoll. This means the complementary region on the brane,  $\bar{\mathcal{I}}$ , can always be reconstructed when the brane lies below the critical angle. This information must be “stored” on the defect itself because the co-wedge will always contain  $\bar{\mathcal{I}}$  in the limit where  $\Gamma \rightarrow 0$ . This is not true above the critical angle, since the co-wedge then vanishes in the limit where  $\Gamma \rightarrow 0$ .<sup>19</sup> We suspect similar considerations will hold in other finite temperature models.

It is interesting to consider the limit as  $\Gamma \rightarrow 1$ , where the area difference diverges to infinity according to the Russian doll rule. We need  $\Gamma \rightarrow 1$  to reconstruct the event horizon at late times, and the Page time goes to infinity in that limit. This could indicate that information stored near the black string’s event horizon is difficult to access but also robust against the loss of correlation between the two sides induced by time evolution. Suppose for example that Cixin chooses increasingly large values of  $\Gamma$  and reconstructs regions of the entanglement co-wedge which are increasingly close to the event horizon. The Russian doll rule states that these surfaces will be nested and that the area difference will increase without bound as  $\Gamma \rightarrow 1$ . Since these larger area differences will prolong the Page time, Cixin loses access to the island regions at increasingly late times. Meanwhile the island surface coincides with the black-string horizon as  $\Gamma \rightarrow 1$  as the area difference diverges to positive infinity. Therefore Cixin will need to access increasingly large regions of the boundary to reconstruct regions

---

<sup>19</sup>In other words, using the language in our previous paper Geng et al. (2021b), we have a tiny island surface in the limit as  $\Gamma \rightarrow 0$  above the critical angle.



near the event horizon at late times, but the reconstruction can always be performed, provided  $\Gamma$  is large enough.

## 4.4 Discussion of Results

We have presented our computation of subregion entanglement entropy for the doubly holographic black-string model. The system is dual to a strongly coupled BCFT living on a black-hole background at finite temperature *and* to an  $\text{AdS}_d$  black hole coupled to a nongravitating thermal bath. In empty AdS, which corresponds to the zero-temperature limit of this system, it is known that the bulk geometry does not support island solutions for  $d > 2$  when the brane is below the critical angle.

If this feature were to persist in the presence of any finite-temperature horizon, it would pose a serious problem for unitarity. Therefore an important consistency criterion for any braneworld model is that islands must exist even below the critical angle in the presence of a finite-temperature horizon. In this study we have shown that this is precisely what happens in the black-string model: islands exist at finite temperature, albeit with limited extent bounded by a critical anchor. The islands are restricted to the atoll. These solutions are consistent with empty AdS because the critical anchor stays away from the defect below the critical angle.

We have summarized the phase structure for the area differences (Figure 4.10) with some simple geometric rules. For a given brane angle  $\theta_b$ , the island surfaces are dominant on the initial time slice only when the anchoring point in the bath  $\Gamma$  lies within the constant entropy belt. In other words, the initial area difference in such cases is negative. This belt starts at the defect, grows monotonically with brane angle, and extends to the horizon on the bath in the limit where  $\theta_b \rightarrow \pi$ . The constant entropy belt forms at almost the same brane angle where the atoll first saturates the brane.

The signed area difference, which is monotonically decreasing in  $\theta_b$  (Figure 4.6) and monotonically increasing in  $\Gamma$  (Figure 4.7), features some additional geometric

structure which is captured by the Russian doll rule. Nested island surfaces can be constructed by either decreasing  $\theta_b$  or increasing  $\Gamma$ ; this pushes surfaces closer to the black-string horizon, decreases the size of the island on the brane, and increases the area difference between the island surface and the HM surface. This has interesting implications for entanglement wedge reconstruction—as described in Section 4.3.6, an experimenter would need to access increasingly large regions of the boundary to reconstruct bulk regions increasingly near the black-string horizon at late times.

It would be interesting to determine the phase structure for other doubly holographic  $\text{AdS}_d$  braneworlds with  $d > 2$ . There are some results which we believe will be quite general even if the geometry is changed, with the critical angle playing an important role. The atoll should include a region near the horizon on the brane, because the horizon in the bulk is an extremal surface, and above the critical angle the atoll should include a region near the defect, but the atoll need not be connected. We suspect area differences will remain monotonic in both subregion size and brane angle; reducing the brane angle increases the number of degrees of freedom on the defect, and since the defect encodes the braneworld black hole, this should increase the Page time. The reduced Page time correlates with the reduced number of degrees of freedom on the defect. We hope that such general considerations will yield insights into the dynamics of BCFTs with holographic duals and help to uncover the physics underlying the critical angle, the Page angle and other aspects of the remarkable phase structure that we have found.

## **Appendix 4.A: Analytic Results for the Area Difference in the $\Gamma \rightarrow 1$ Limit**

In the bulk of this paper we have considered RT surfaces anchored at a given point  $\Gamma$  in the nongravitating black hole bath, and an symmetric point  $\tilde{\Gamma} = \Gamma$  in the second asymptotic region of the same black hole. In the limit that  $\Gamma \rightarrow 1$  these two anchor points approach the bifurcation surface from both sides. Furthermore, the

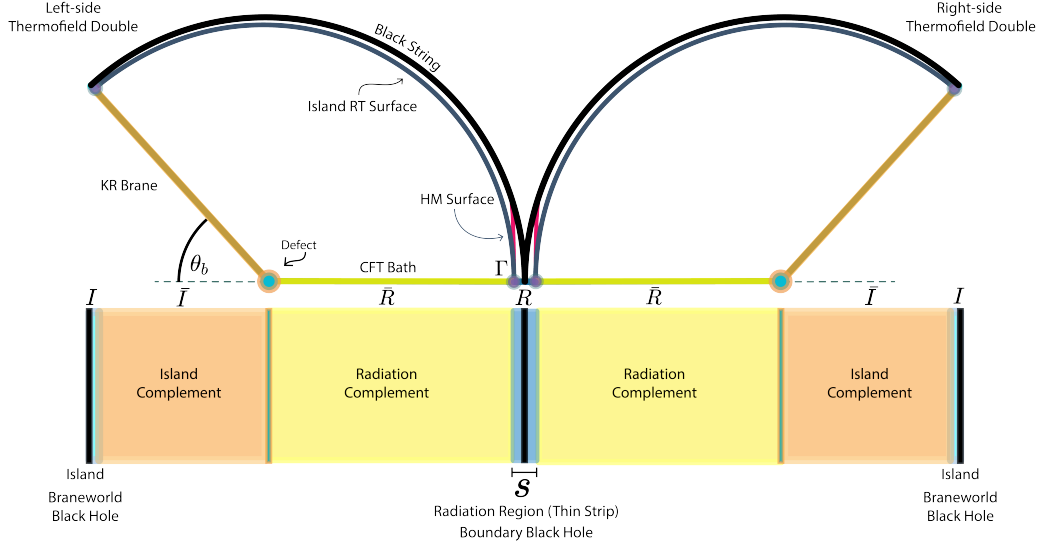


Figure 4.15: Here we illustrate the doubly holographic black string setup for a small radiation region of width  $s$ . This corresponds to the limit  $\Gamma \rightarrow 1$ . The island region becomes infinitesimally small as well; this is consistent with the Russian doll rule.

spacetime near the black horizon becomes essentially flat. So the initial HM surface in this limit is really calculating the entanglement entropy of a thin strip in empty AdS, illustrated in Figure 4.15, for which analytic results have been known for a long time Ryu and Takayanagi (2006d).

The entanglement entropy of a strip of length  $s$  in a nongravitating  $\text{CFT}_d$  with dual empty  $\text{AdS}_{d+1}$  is given by Ryu and Takayanagi (2006d),

$$S = \frac{1}{4G_N^{(d+1)}} \left[ \frac{2}{d-2} \left( \frac{L}{\epsilon} \right)^{d-2} - \frac{2^{d-1} \pi^{\frac{d-1}{2}}}{d-2} \left( \frac{\Gamma\left(\frac{d}{2(d-1)}\right)}{\Gamma\left(\frac{1}{2(d-1)}\right)} \right)^{d-1} \left( \frac{L}{s} \right)^{d-2} \right], \quad (4.15)$$

where the first term is the divergent piece, which is subtracted upon renormalization.  $L^{d-2}$  is the volume factor of to the orthogonal directions which were generally omitted in this note. Since the near-horizon geometry is flat, in order to compare with the above expression we need to transform our coordinate  $u = u_h - \zeta$ , so that,

$$\frac{du^2}{h(u)} = \frac{du^2}{1 - \frac{u^{d-1}}{u_h^{d-1}}} \approx \frac{d\zeta^2}{\frac{d-1}{u_h} \zeta}. \quad (4.16)$$

Upon transforming the coordinate, we take the near-horizon limit. In these new coordinates, we define  $\zeta = \zeta_*$  to correspond to  $u = u_h \Gamma$ .

Now we define the flat coordinate such that  $dy = d\zeta / \sqrt{\zeta \frac{d-1}{u_h}}$ . The length of the near-horizon belt in the flat setup is  $s = 2 \int_0^{y_*} dy = 2y_*$ , where the factor of 2 comes from the fact that our geometry is replicated on the other side of the horizon. In this coordinate, we have  $y_* = 2\sqrt{\zeta_* \frac{u_h}{d-1}}$ . Now, substituting this into (4.15), dividing by the transverse volume factor  $L^{d-2}$ , and renormalizing the expression we obtain,

$$2\mathcal{A}_{HM} = -\frac{2^{d-1}\pi^{\frac{d-1}{2}}}{d-2} \left( \frac{\Gamma\left(\frac{d}{2(d-1)}\right)}{\Gamma\left(\frac{1}{2(d-1)}\right)} \right)^{d-1} \left( \frac{1}{4} \sqrt{\frac{d-1}{\zeta_* u_h}} \right)^{d-2}, \quad (4.17)$$

where we have multiplied the entropy by  $4G_N^{(d+1)}$  to obtain the corresponding area. However, note that this is the area of the *full* initial HM surface spanning both exterior regions of the bulk black string; to appropriately match with the conventions of this paper, we use  $\mathcal{A}$  to mean the area of the HM surface restricted to a single exterior region of the bulk.

In the limit where  $\Gamma \rightarrow 1$ , we take the above expression to  $\zeta_* \rightarrow 0$ . After substituting  $d = 4$  and  $u_h = 1$ , we find that,

$$\mathcal{A}_{HM} = -\frac{3\pi^{3/2}\Gamma\left(\frac{2}{3}\right)^3}{8\Gamma\left(\frac{1}{6}\right)^3} \frac{1}{\zeta_*} \approx -\frac{0.030}{\zeta_*}. \quad (4.18)$$

We plot our numerical results for the Hartman-Maldacena area close to  $\zeta_* = 0$  in Figure 4.16 and demonstrate that they agree well with the analytic form (4.18), validating our method.

## Appendix 4.B: Validation of Numerics

One may wonder whether the results we are obtaining are indeed physical. To show that this is the case, we compute the area difference as a function of the brane angle for different cutoffs. We see in Figure 4.17 that our area differences do not vary substantially for a large range of cutoff choices.

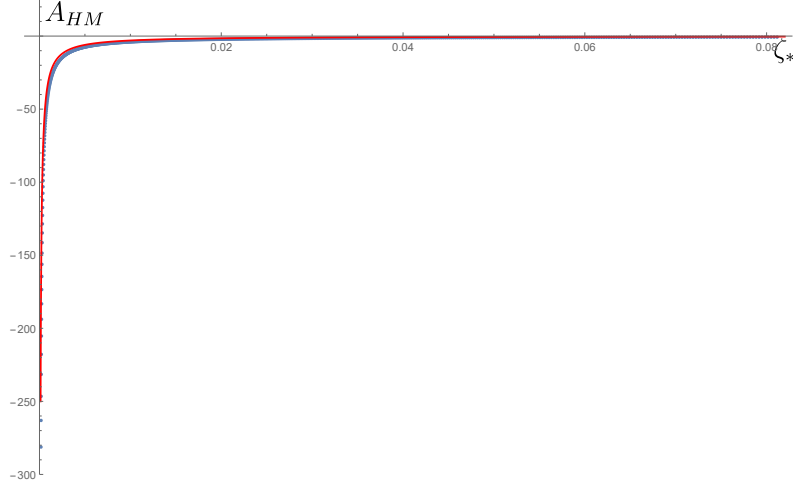


Figure 4.16: Area of the HM surface as a function of  $\zeta_*$  with  $u_h = 1$ . The red curve is what we expect for the near-horizon region (small  $\zeta_*$ ) from theory, namely  $-0.030/\zeta_*$ . Close to the horizon, the data roughly matches the theory, but for some large enough  $\zeta_*$  the finite temperature kicks in and we should not expect the data to match. It is not visible, but there is a small error in the “large”  $\zeta_*$  region.

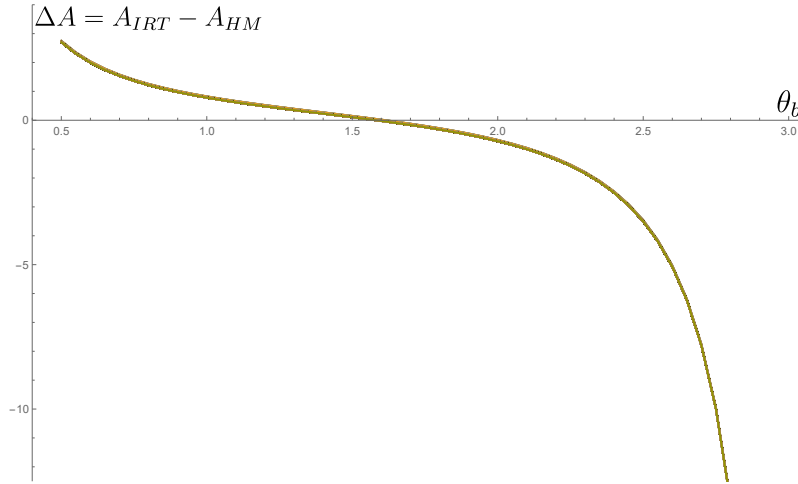


Figure 4.17: Area difference as a function of brane angle, for  $\Gamma = 0.32$  for angular cutoffs ranging from 0.0005 to 0.2 in steps of 0.0005 (with  $u_h = 1$ ). Each curve is shown in a different colour, but since they are very close it is hard to distinguish one from the other.

# Chapter 5: Jackiw-Teitelboim Gravity from the Karch-Randall Braneworld

## Abstract

In this letter, we show that Jackiw-Teitelboim (JT) gravity can be naturally realized in the Karch-Randall braneworld. Notably the role of the dilaton in JT gravity is played by the radion in a suitably orbifolded version of the setup. In the classical entanglement entropy calculation, there is an apparent degeneracy of Ryu-Takayanagi surfaces. We demonstrate how quantum fluctuations of the radion/dilaton resolves this would-be classical puzzle regarding entanglement wedge reconstruction.

## 5.1 Introduction

Recent utilization of the doubly holographic Karch-Randall braneworld Karch and Randall (2001b,c) has yielded substantial insights into quantum gravity. These include analytical models that have been used to produce Page curves describing black hole radiation in higher-dimensional (i.e in  $d > 2$  spacetime dimensions) setups Almheiri et al. (2019d); Geng and Karch (2020b)<sup>1</sup>, demonstrating a connection between the existence of entanglement islands and massive gravity theories Geng and Karch (2020b); Geng et al. (2021b,c), providing simple holographic models to study quantum field theory in black hole backgrounds Geng et al. (2021d, 2022f), and moti-

---

<sup>0</sup>This chapter is based on published work in collaboration with Geng et al. (2022c). All authors contributed to the final version of the manuscript. The main ideas of this work emerged through discussions, in which I actively participated, contributing to the conceptual development of the project.

<sup>1</sup>This is motivated by the earlier work Penington (2019); Almheiri et al. (2019a, 2020e, 2019c).

vating a novel holographic setup—wedge holography Akal et al. (2020); Miao (2020); Geng et al. (2021b,c).

Wedge holography is a simple canonical deformation of the original Karch-Randall braneworld which manifests a codimension-two holographic duality Akal et al. (2020); Miao (2020); Geng et al. (2021b,c). This setup starts with two  $\text{AdS}_d$  Karch-Randall branes in an ambient  $\text{AdS}_{d+1}$  with the two branes intersecting each other at the conformal boundary of the ambient spacetime. These branes are treated as “end-of-the-world” (EOW) branes, meaning that the part of the ambient spacetime behind them is excised. Thus, the leftover bulk spacetime is a wedge (see Fig. 5.1a). In analogy with the original Karch-Randall braneworld, the wedge system is *doubly holographic* in that there are three equivalent descriptions related to each other by the AdS/CFT correspondence Maldacena (1998b); Witten (1998a); Gubser et al. (1998) (see Fig. 5.1):

- **The bulk description:** Einstein-Hilbert gravity in the wedge.
- **The intermediate description:** Two  $\text{CFT}_d$ ’s coupled with gravity living on separate  $\text{AdS}_d$ ’s and glued to each other along their asymptotic boundaries by imposing transparent boundary conditions.
- **The boundary description:** A  $\text{CFT}_{d-1}$  living on the common asymptotic boundary of the two branes (which is the corner of the wedge in Fig. 5.1a).

In this letter, we study wedge holography when the bulk is  $\text{AdS}_3$  with two  $\text{AdS}_2$  branes. As opposed to the original wedge holography setup in higher dimensions where the brane fluctuations can be neglected, we will consider fluctuating branes with orbifold projection so that there is only one fluctuating scalar degree of freedom. This scalar is the *radion* in the language of the Randall-Sundrum I setup Randall and Sundrum (1999c). We will derive an action for the radion which is coupled with the 2d localized braneworld gravity. This can be thought of as the low-energy effective theory

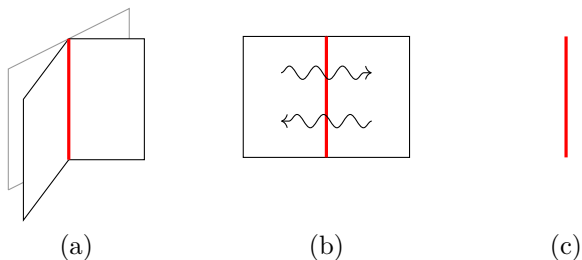


Figure 5.1: A schematic representation of the doubly holographic nature of the wedge. (a) is the bulk description, which is just gravity on the  $(d + 1)$ -dimensional wedge with a codimension-two corner in red. (b) is the intermediate description in which the two  $d$ -dimensional theories share a transparent interface (red). (c) is the boundary description on the interface.

of the intermediate description. We will see that this theory is precisely the Jackiw-Teitelboim (JT) gravity Teitelboim (1983); Jackiw (1985); Fukuyama and Kamimura (1985), in which the dilaton field is the radion.

The importance of our study is that it ensures the consistency of wedge holography in  $\text{AdS}_3$ . This is because when we ignore the brane fluctuations and study the entanglement entropy in  $\text{AdS}_3$  wedge holography using the Ryu-Takayanagi (RT) formula, we find an infinite number of degenerate RT surfaces. This renders the entanglement wedge, which is an essential element in holography Dong et al. (2016a), as undefined. Taking into account the brane fluctuations, this degeneracy is lifted, as we show explicitly using our effective JT description.

We note that JT gravity Almheiri and Polchinski (2015); Jensen (2016); Maldacena et al. (2016); Engelsöy et al. (2016); Sachdev (2019) has also played an important role in the recent development of low-dimensional ( $d = 2$ ) quantum gravity Penington (2019); Almheiri et al. (2019a); Saad et al. (2019). This includes providing an analytically controllable model to compute the Page curve of evaporating black holes Penington (2019); Almheiri et al. (2019a) and a precise duality between Euclidean JT gravity to a random matrix ensemble Saad et al. (2019). Our realization of JT gravity in the Karch-Randall braneworld further argues for its significance in our study of quantum gravity Geng (2022).



## 5.2 Karch-Randall Braneworld and Wedge Holography

In this section, we briefly review the Karch-Randall braneworld and wedge holography, focusing on the geometric pictures for later holographic study. The original theory of Karch and Randall is described by the following action

$$S = - \frac{1}{16\pi G_{d+1}} \int d^{d+1}x \sqrt{-g_{\text{bulk}}} (R[g_{\text{bulk}}] - 2\Lambda) - \frac{1}{8\pi G_{d+1}} \int d^d x \sqrt{-h} (K - T), \quad (5.1)$$

where the first term is the standard Einstein-Hilbert action describing the gravitational in the ambient  $\text{AdS}_{d+1}$  with a cosmological constant  $\Lambda = -\frac{d(d-1)}{2}$ ,<sup>2</sup> the second term describes the embedding of the brane in the ambient  $\text{AdS}_{d+1}$ ,  $h_{ab}$  is the induced metric on the brane,  $K$  is the trace of the extrinsic curvature of the brane and  $T$  is the tension of the brane and satisfies  $T \leq (d-1)$  as a Karch-Randall brane.<sup>3</sup> The bulk metric fluctuations satisfy the Neumann boundary condition near the brane

$$\nabla_n \delta g_{\text{bulk } \mu\nu} |_{\text{near brane}} = 0, \quad (5.2)$$

where  $n$  denotes the normal direction of the brane. With this boundary condition taken into account, setting the variation of the action Equ. (5.1) to zero we obtain two equations of motion—the usual Einstein's field equation which determines the bulk geometry and the brane-localized equation of motion describing the embedding of the brane in the bulk. The latter is given by Fujita et al. (2011)

$$K_{ab} = (K - T)h_{ab}. \quad (5.3)$$

This system of equations can be solved by

$$ds_{\text{bulk}}^2 = dr^2 + \cosh^2(r) ds_{\text{AdS}_d}^2, \quad (5.4)$$

---

<sup>2</sup>We have set the AdS curvature length to one.

<sup>3</sup>The critical value  $T = d - 1$  corresponds to the Randall-Sundrum brane Randall and Sundrum (1999c,d).

where  $r \in (-\infty, \infty)$ , with the brane given by  $r = r_1 = \text{const.}$  and  $T = (d-1)|\tanh r_1|$  (we only consider positive tension branes).<sup>4</sup> The bulk metric is just empty  $\text{AdS}_{d+1}$  written as a foliation into  $\text{AdS}_d$  slices. This is particularly useful in the study of the Karch-Randall braneworld as here the brane geometry  $r = \text{const.}$  is clearly  $\text{AdS}_d$ . In this letter, we focus on the case  $d = 2$ , in which case the geometry can be easily visualized as that on the left of Fig. 5.2a.

In the wedge holography setup Akal et al. (2020); Miao (2020); Geng et al. (2021b,c), we add another (positive tension) Karch-Randall brane. This system is described by the following action

$$S = -\frac{1}{16\pi G_3} \int d^3x \sqrt{-g_{\text{bulk}}} (R[g_{\text{bulk}}] - 2\Lambda) - \frac{1}{8\pi G_3} \int_{i=1,2} d^d x \sqrt{-h_i} (K_i - T_i), \quad (5.5)$$

where we impose the Neumann boundary conditions Equ. (5.2) for the bulk metric fluctuation near both branes. The equations of motion still consist of two parts—the bulk Einstein’s equation and the two brane-localized equations, with each one the same as Equ. (5.3) but with their own tensions  $T_1$  and  $T_2$ . Hence, a solution is given by the same bulk metric Equ. (5.4) with  $d = 2$  where  $r \in (r_1, r_2)$ ,  $T_1 = \tanh(-r_1)$ ,  $T_2 = \tanh r_2$ ,  $r_1 < 0$  and  $r_2 > 0$ . This geometry is visualized in Fig. 5.2b.

### 5.3 Realizing JT Gravity

In the original wedge holography papers Akal et al. (2020); Miao (2020); Geng et al. (2021b,c) the branes are treated as rigid. This assumption is rooted in the original Karch-Randall setup where we just have one brane and we can take the brane to be sitting at an orbifold fixed point to model an EOW brane in analog to O-planes in string theory Polchinski (2007c); Johnson (2005). However, reminding

---

<sup>4</sup>If the brane is given by  $r = r_1 < 0$  understood as an EOW brane with positive tension the domain of the bulk is  $r \in (r_1, \infty)$  and if the brane is given by  $r = r_2 > 0$  understood as an EOW brane with positive tension the bulk domain is  $r \in (-\infty, r_2)$  Karch and Randall (2020).

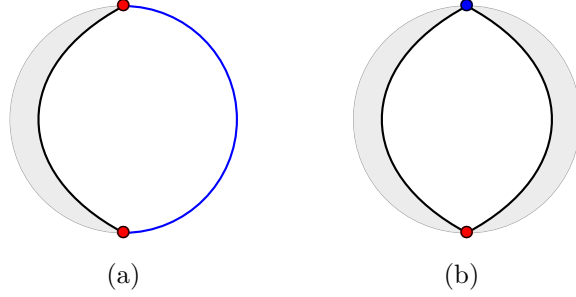


Figure 5.2: (a) A constant time slice of the empty  $\text{AdS}_3$  bulk with one Karch-Randall brane (in black). The brane intersects the asymptotic boundary (in blue) at the two defects (red dots). The holographic dual is a  $\text{BCFT}_2$  living on the blue interval. (b) The empty  $\text{AdS}_3$  wedge consisting of two KR branes. The two branes intersect at the two defects (red and blue dots). The dual boundary description is hence a conformal quantum mechanics living on the two defects.

ourselves of the Randall-Sundrum I (RS1) setup Randall and Sundrum (1999c) with two RS branes, the relative distance between the branes is a dynamical variable called the radion, which has important phenomenological implications Arkani-Hamed et al. (2001b). Hence in the case of wedge holography, we should not neglect the possibility of brane fluctuations.

In this section, we consider small brane fluctuations (relative to the bulk  $\text{AdS}_3$  curvature scale) and derive the effective action for the coupling between these small fluctuations and the brane-localized graviton mode by doing dimensional reduction. While brane fluctuations can be neglected in higher dimensions, we will see that this small contribution plays an important role when the bulk is  $\text{AdS}_3$ . The detailed derivation and a more careful analysis will be presented in Geng (2022).

To be precise, we will consider two branes with tensions  $T_1$  and  $T_2$  sitting at  $r = r_1 + \delta\phi_1(x)$  and  $r = r_2 + \delta\phi_2(x)$  slices of the following bulk geometry

$$ds^2 = dr^2 + \cosh^2(r)g_{ab}(x)dx^a dx^b, \quad (5.6)$$

with the relationships  $T_1 = -\tanh r_1 > 0$  and  $T_2 = \tanh r_2 > 0$  satisfied and  $\delta\phi_1, \delta\phi_2 \ll 1$  (the AdS length scale has been set to one). The wedge runs from

$r = r_1 + \delta\phi_1(x)$  to  $r = r_2 + \delta\phi_2(x)$ , and the metric ansatz Equ. (5.6) describes the lowest 2d graviton mode (the higher modes would depend on the coordinate  $r$  as  $g_{ab}(r, x)$ ).

To obtain the low-energy effective action, we have to plug the above ansatz of the metric and brane locations into Equ. (5.5) and keep terms only up to order  $\mathcal{O}(\delta\phi^2)$ . This yields the following action<sup>5</sup>

$$S_{\text{eff}} = S_0 - \frac{1}{16\pi G_3} \int d^2x \sqrt{-g} \phi(x) \left[ R[g] + 2 \right] + S_{\text{dilaton}}, \quad (5.7)$$

where  $\phi(x) = \delta\phi_2(x) - \delta\phi_1(x)$  is the radion mode,  $S_0$  is a purely topological term

$$S_0 = -\frac{r_2 - r_1}{16\pi G_3} \int d^2x \sqrt{-g} R[g], \quad (5.8)$$

and  $S_{\text{dilaton}}$  is the kinetic term for the brane fluctuations

$$S_{\text{dilaton}} = -\frac{1}{8\pi G_3} \int d^2x \sqrt{-g} \left[ \frac{T_2}{2} \nabla_\mu \delta\phi_2 \nabla^\mu \delta\phi_2 + T_2 (\delta\phi_2)^2 + \frac{T_1}{2} \nabla_\mu \delta\phi_1 \nabla^\mu \delta\phi_1 + T_1 (\delta\phi_1)^2 \right]. \quad (5.9)$$

In analogy to the Randall-Sundrum scenario, to isolate the dynamics of the radion we have to orbifold our setup. We can project out either  $\delta\phi_1(x)$  or  $\delta\phi_2(x)$  by a canonical orbifold prescription: start with an array of an infinite number of paired wedges, mod out the translation symmetry on the array to a single pair of wedges with  $\mathbb{Z}_2$  symmetry, and then mod out *this*  $\mathbb{Z}_2$  symmetry (see Geng (2022) for details) to get a single wedge with only one fluctuating brane. The possibility of projecting out either  $\delta\phi_1(x)$  or  $\delta\phi_2(x)$  are gauge-equivalent. To see this gauge equivalence, without loss of generality we can project out  $\delta\phi_1(x)$  and redefine  $g_{ab}(x)$  such that the final action is independent of  $r_1$  and  $r_2$ . Specifically, by setting

$$\delta\phi_1(x) = 0, \quad g_{ab}(x) \rightarrow e^{-T_2 \delta\phi_2(x)} g_{ab}(x), \quad (5.10)$$

---

<sup>5</sup>For details of the derivation, we refer the readers to Geng (2022).

the total action becomes

$$S_{\text{eff}} = S_0 - \frac{1}{16\pi G_3} \int d^2x \sqrt{-g} \phi(x) \left[ R[g] + 2 \right], \quad (5.11)$$

where now the dilaton field is  $\phi(x) = \delta\phi_2(x)$ . This is precisely Jackiw-Teitelboim gravity Jackiw (1985); Teitelboim (1983) for  $\text{AdS}_2$ .

However, there is no nontrivial profile for the dilaton on the  $\text{AdS}_2$  background without a cutoff Maldacena et al. (1999). This cutoff breaks the asymptotic conformal symmetry and allows the theory to have finite energy excitations Maldacena et al. (2016). Hence we fix a small cutoff  $\epsilon$  and fix the boundary metric for the nearly- $\text{AdS}_2$  and the dilaton field to be

$$g_{ab}(x)|_{\text{bdy}} dx^a dx^b = -\frac{du^2}{\epsilon^2}, \quad \phi(x)|_{\text{bdy}} = \phi_b = \frac{\phi_r}{\epsilon}. \quad (5.12)$$

One subtlety of our derivation is that we have to take  $\phi_b \sim \epsilon$  (or  $\phi_r \sim \epsilon^2$ ) to ensure that the brane fluctuation is always small, i.e.  $\delta\phi_i(x) \ll 1$ . We emphasize that this scaling is special to our setup. For example, it does not appear in dimensional reduction for the near-horizon region of a four-dimensional near-extremal black holes Nayak et al. (2018); Yang (2019). In that case  $\phi_b$  is determined by the linear-in-temperature coefficient of the near-extremal black hole entropy Sachdev (2019). In our case, without a UV completion of the underlying higher-dimensional theory,  $\phi_b$  is a free parameter.<sup>6</sup> In our case, the UV completion is a conformal quantum mechanics system Geng (2022), which differs from the UV completion into a black hole.

Translating to the  $\text{AdS}_3$  language Equ. (5.6), we have to take a cutoff of the  $\text{AdS}_3$  wedge and fix the metric there to be (see Fig. 5.3 for an indication)

$$ds_{\text{bdy}}^2 = dr^2 - \cosh^2(r) \frac{du^2}{\epsilon^2}. \quad (5.13)$$

---

<sup>6</sup>We thank Subir Sachdev for useful discussions regarding this interesting point.

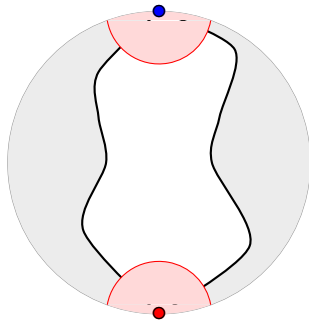


Figure 5.3: The shapes of the two branes are fluctuating (though more rigorously we have also to impose the orbifold projection). The  $\text{AdS}_2$  cutoff is induced by the  $\text{AdS}_3$  cutoff (the red curve). The shaded regions are excised. Now the regulated asymptotic boundary of the wedge are two intervals.

Now that we have identified this cutoff and boundary conditions for the metric in the  $\text{AdS}_3$  bulk, we have to add the corresponding boundary term to the total action Equ. (5.5). The boundary term is the usual Gibbons-Hawking term

$$S_{\text{bdy}} = S_{\text{GH}} = -\frac{1}{8\pi G_3} \int d^2 x_{\text{bdy}} \sqrt{-h^{\text{bdy}}} K^{(3)}, \quad (5.14)$$

which can be shown to be

$$S_{\text{bdy}} = -\frac{2(r_2 - r_1)}{16\pi G_3} \int_{\partial\mathcal{M}} dx \sqrt{-h} K - \frac{2\phi_b}{16\pi G_3} \int_{\partial\mathcal{M}} dx \sqrt{-h} K, \quad (5.15)$$

where  $K$  is the trace of the extrinsic curvature of the cutoff boundary of the  $\text{AdS}_2$ . With this term added to Equ. (5.11), we get the full-fledged JT gravity action as that in Harlow and Jafferis (2020). For more details we refer the readers to Geng (2022).

## 5.4 A Puzzle and Its Resolution

A canonical quantity to study in  $\text{AdS}_3$  wedge holography is the entanglement entropy between the two asymptotic defects which support the dual quantum mechanics (see Fig. 5.2b). This can be calculated using the Ryu-Takayanagi (RT) formula Ryu and Takayanagi (2006c,d) by studying the RT surfaces that connect the two branes Akal et al. (2020); Geng et al. (2021b). Choosing global  $\text{AdS}_2$  for those  $\text{AdS}_2$

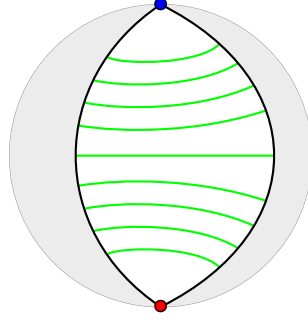


Figure 5.4: The empty  $\text{AdS}_3$  wedge consisting of two KR branes. The bulk geometry is the  $\text{AdS}_3$  global patch. The green curves are representatives of the degenerate RT surfaces calculating the entanglement entropy between the two defects (the blue dot and the red dot).

slices, the bulk metric Equ. (5.4) can be written as

$$ds^2 = dr^2 + \cosh^2(r) \left[ -\cosh^2(\eta) d\tau^2 + d\eta^2 \right], \quad (5.16)$$

if we ignore the brane fluctuations  $\delta\phi_i(x)$ . The RT surface can be parametrized as  $\eta = \eta(r)$ ,  $\tau = \text{const.}$  which leads to the area functional

$$A = \int_{r_1}^{r_2} dr \sqrt{1 + \cosh^2(r) \eta'(r)^2}. \quad (5.17)$$

It is easy to show that, with the boundary term carefully taken into account, the variation of the area functional vanishes if and only if  $\eta'(r) = 0$ . Therefore, any constant  $\eta$  surface is an RT surface and they all have equal area  $A = r_2 - r_1$  (see Fig. 5.4 for an illustration). According to this result, there would be no definite RT surface for the entanglement entropy between the two defects and as a corollary there would be no definite entanglement wedge for either of the defects. This would be problematic because it would preclude a holographic interpretation from the point of view of bulk (or more precisely, entanglement wedge Jafferis et al. (2016)) reconstruction.

Fortunately, this puzzling result occurs only when the branes are treated as rigid. We expect brane fluctuations to lift the degeneracy of the RT surfaces. Because the brane fluctuation is in general time-dependent, we should use the HRT

formula Hubeny et al. (2007b) to compute the entanglement entropy from the wedge. However, this would be complicated as the result is heavily dependent on properly treating the boundary condition near the brane.

This challenge of computing entanglement entropy with brane fluctuations is readily met if we instead do it using our effective braneworld theory Equ. (5.11). In that case we are just computing the entanglement entropy between the two cutoff boundaries of the nearly-AdS<sub>2</sub> space. By solving the equation of motion of the dilaton and 2d metric, we have that

$$ds_{2d}^2 = -\cosh^2(\eta)d\tau^2 + d\eta^2, \quad (5.18)$$

$$\phi(x) = \phi_r \cosh \eta \cos \tau, \quad (5.19)$$

where  $\phi_r$  is a constant. Thus, using the maxmin prescription Wall (2014b) of the HRT formula we have

$$S_{EE} = \max_{\tau} \min_{\eta} \left( \frac{\phi_0 + \phi(x)}{4G_3} \right) = \frac{\phi_0 + \phi_r}{4G_3}, \quad (5.20)$$

where the minimization has a solution  $\eta = 0$ ,  $\tau = 0$ . As a result, from the bulk AdS<sub>3</sub> perspective Equ. (5.16), the RT surface is just the  $\eta = 0$  surface which sits precisely in the middle of the bulk (i.e. the middle green surface in Fig. 5.4). Thus, the entanglement wedge of either defect is the half bulk region next to it, which is consistent with the symmetry between the two defects.

## 5.5 Conclusions

In this letter, we studied wedge holography when the bulk is three-dimensional. We found that it is necessary to consider brane fluctuations, identified as the radion mode in the braneworld language Randall and Sundrum (1999c,d), for the apparent consistency of the holographic setup. We found that the low-energy dynamics of the brane fluctuations with orbifold projection is precisely Jackiw-Teitelboim gravity.



# Chapter 6: Subregion Entropy for the Doubly-Holographic Global Black String

## Abstract

We study the growth of entanglement entropy in a doubly holographic model of gravity for a spherical AdS black hole. Compared to previous work, which was limited to the case of planar black holes, this introduces an extra scale to the problem. This allows us to analyze the interplay between the reorganization of entanglement entropy due to island formation and the onset of the Hawking-Page phase transition and to find the appearance of a new critical black hole radius unrelated to the thermodynamics. We also find that the geometry of the Ryu-Takayanagi surface capturing the physics of islands exhibits drastically different behavior than in the planar case.

## 6.1 Introduction

The AdS/CFT correspondence Maldacena (1998b); Gubser et al. (1998); Witten (1998a), which is the most concrete realization of the holographic principle to date, is a conjectured duality between two seemingly different physical theories. One of its most remarkable aspects is that, since the correspondence relates a higher-dimensional theory of gravity in anti-de Sitter space (AdS) to a dual conformal field theory (CFT) on its boundary, it provides a consistent framework where the unitarity

---

<sup>0</sup>This chapter is based on published work in collaboration with Karch et al. (2023c). All authors contributed to the final version of the manuscript. My contributions, together with Merna Youssef, included performing the entanglement entropy computations and mapping the phase structure of the model. Additionally, I participated in discussions that helped develop key ideas in the project.

of a black hole’s evaporation process can be studied.

The crucial ingredient in recent calculations, which have yielded a unitary Page curve for the time-evolution of the entanglement entropy, has been entanglement islands Penington (2019); Almheiri et al. (2019a) – regions of spacetime which, while seemingly disconnected from the holographic system, arise when the boundary region is coupled to a system with additional gravitational degrees of freedom. They are responsible for the late-time contribution to the entropy of a subregion on the boundary theory, and their location on a slice of  $\text{AdS}_{d+1}$  is determined by a quantum extremal surface (QES) Engelhardt and Wall (2015b) whose location, if perturbed, won’t affect the measurements of entanglement entropy.

However, determining the location of these islands is highly nontrivial as one needs to have quantitative control of the quantum entanglement across the QES. But by introducing a boundary to the conformal field theory, one obtains a doubly holographic model where the entanglement entropy can be computed using the classical Ryu-Takayanagi (RT) prescription in one dimension higher Ryu and Takayanagi (2006c,d); Hubeny et al. (2007b); Faulkner et al. (2013b); Lewkowycz and Maldacena (2013b); Chen et al. (2020b,a).<sup>1</sup> In such models the boundary of the CFT, which we will refer to as the defect, is dual to a Karch-Randall (KR) brane where RT surfaces can connect, thus defining a quantum extremal surface and forming an island on the brane. More precisely, following previous work Geng et al. (2021d), the island is defined as the region extending from the RT surface’s anchor point on the brane to the other side of the thermofield double, and the family of anchor points on the brane to which RT surfaces can connect is defined as the atoll. RT surfaces that connect the brane to the boundary are called island surfaces, and when they connect the brane to the defect, the location where it attaches to the brane is known as a critical anchor.

---

<sup>1</sup>The main advantage of this approach is that analytic control of the intermediate picture is not needed to determine the location of the quantum extremal surface, though such computations could presumably be done in principle.

One way to realize this model is by starting with a Karch-Randall (KR) brane on a  $d$ -dimensional slice of  $\text{AdS}_{d+1}$ . Through the AdS/CFT correspondence, this model has three equivalent descriptions Karch and Randall (2001b); Randall and Sundrum (1999d); Karch and Randall (2001c); Takayanagi (2011b); Fujita et al. (2011) which we will refer to as the boundary, intermediate, and bulk pictures. The boundary picture is a unitary  $d$ -dimensional CFT ending on a  $d-1$  dimensional defect, and the bulk picture is an  $\text{AdS}_{d+1}$  spacetime with an embedded  $\text{AdS}_d$  KR brane. We emphasize that the bulk picture is particularly notable because we can use this description to compute entanglement entropy using classical RT surfaces. The location of the RT surface in the bulk description determines the location of the QES in the intermediate picture. Here we have a  $d$ -dimensional CFT on the KR brane ending on the defect, with semi-transparent boundary conditions, allowing energy transfer between the KR brane and a non-gravitating bath housing a  $d$ -dimensional CFT. The intermediate description is useful because, once a Schwarzschild black hole is introduced on each  $\text{AdS}_{d+1}$  spacetime slice, we can use it to interpret the areas of the bulk RT surfaces – which determine the entanglement entropy of the radiation region in the non-gravitating bath – as the entanglement entropy of the radiation escaping the black hole on the brane.

We pause briefly to describe the coupling of the conformal field theory to a semi-transparent boundary, which makes sense from a physical perspective. There is a negative cosmological constant on the brane which necessitates the introduction of boundary conditions for the quantum fields. The conventional approach is to choose reflecting boundary conditions, and when this choice is made, AdS black holes essentially cannot evaporate on their own because the Hawking radiation they send to the boundary is reflected. One way to define an evaporation process for them is to couple the system to a heat bath Rocha – in this model, from the intermediate perspective, this is achieved through the transparent boundary conditions at the defect that separates the degrees of freedom on the brane from the degrees of freedom on the boundary CFT.

The main difference between the setup we consider here, and the one previously explored in Geng et al. (2021b,d), is that we are interested in spherical Schwarzschild black holes instead of planar ones. This means we use AdS in global coordinates. To understand the significance of this, let us first recall some facts about standard AdS/CFT at finite temperature without the brane. For global AdS, the topology of the boundary is  $S^{d-1} \times S^1$ , and that's where the CFT is located. If we instead place the CFT on flat background  $R^{d-1}$ , as has been done in previous work on double holographic islands Almheiri et al. (2020e, 2019d); Geng and Karch (2020b); Geng et al. (2021b,d,e), this differs from a sphere  $S^{d-1}$  in that the latter introduces an extra phase transition and associated instability. For a CFT on  $R^{d-1}$ , the temperature  $T$  is the only scale, and therefore there is no notion of a hot or cold black hole; physics is equivalent at different temperatures. In this case, the boundary topology makes the AdS Schwarzschild black hole the only candidate solution in the bulk to express the thermal CFT on the boundary. As for the CFT on a sphere, we now have the radius of the sphere  $R$  and temperature  $T$  composing a dimensionless number  $RT$  which can form nontrivial physics at different temperatures Witten (1998b). This is to say the CFT at finite  $T$  on a sphere introduces two candidate solutions: the thermal AdS (empty AdS) and the Schwarzschild AdS black hole. It is the well-known Hawking-Page phase transition, above which the black hole is no longer the preferred phase, which signals that there exists a point at which a phase transition from a black hole to a thermal gas of gravitons occurs.

We employ the black string solution, a well-known solution to vacuum gravity in  $\text{AdS}_5$  or higher dimensional spacetimes, to realize this important phase transition of a finite temperature CFT on the sphere within double holography. The scale in the theory is represented by the parameter  $u_h$ , which determines the size of the Schwarzschild black hole on each spacetime slice – see Figure 6.1. From the bulk perspective, introducing the black string in global  $\text{AdS}_{d+1}$  is equivalent to placing spherical Schwarzschild black holes on each  $\text{AdS}_{d+1}$  spacetime slice, including on the boundary on which the BCFT lives. From the intermediate perspective, this amounts

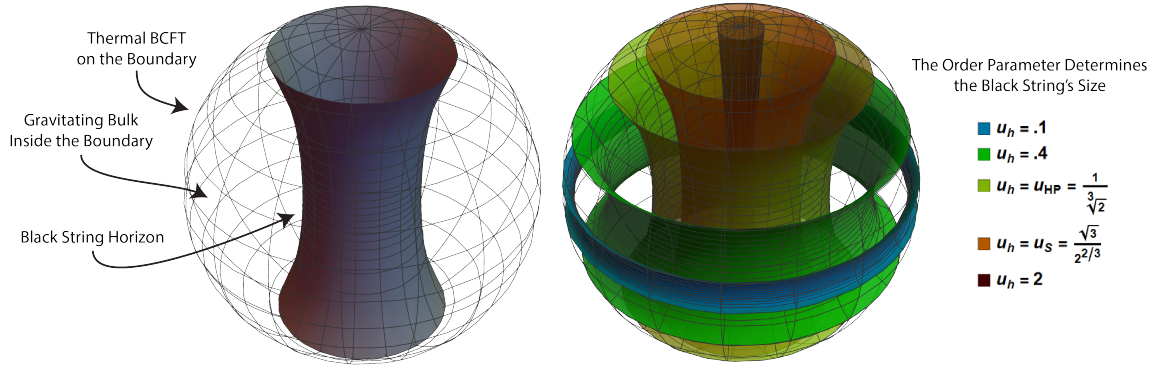


Figure 6.1: The diameter of the black string is determined by the order parameter in the theory,  $u_h$ . The bulk theory is dual to a thermal BCFT state on the surface of the sphere. When  $u_h$  becomes larger than  $u_{\text{HP}} = \frac{1}{2^{1/3}}$  we encounter the Hawking-Page transition and the string is no longer the global minimum of the free energy. The black string reaches minimum temperature at the spinodal point  $u_s = \frac{\sqrt{3}}{2^{2/3}}$ , where the specific heat changes sign from positive to negative. Also see Figure 6.2.

to placing equal size black holes on the gravitating brane and the non-gravitating bath, which allows them to be in equilibrium while still exchanging Hawking quanta. From the boundary perspective, this is just a theory at finite temperature. In our case, the finite temperature in the non-gravitating bath is maintained by having a black hole in the bath as well. While somewhat unnatural from the field theory point of view, this situation serves just as well as a heat bath. It greatly simplifies the problem due to the underlying simplicity of the black string. Furthermore, since this is a static geometry, the RT surfaces can be constructed on a single time slice.

It can be seen in Figure 6.1 that the black string has the topology of a cylinder. Such black string solutions are known to have an instability, corresponding to the rippling of the horizon, which is known as the Gregory Laflamme instability Gregory and Laflamme (1993). This instability is first encountered for the global AdS black string when the black holes on each slice dip below a certain size. Early on it was believed that this instability is dual to the Hawking-Page phase transition of the braneworld black hole Chamblin and Karch (2005); Gregory et al. (2008) based on early numerical evidence Hirayama and Kang (2001). More recent numerical evidence

Marolf and Santos (2019) and analytic arguments in a large number of dimensions Emparan et al. (2022) have shown that the GL instability, in fact, only sets in for black holes smaller than the radius at the Hawking Page transition. This is in fact consistent with the Gubser Mitra conjecture Gubser and Mitra (2001), which states that mechanical instabilities in the higher dimensional bulk should be dual to local thermodynamic instabilities in the dual CFT. The Hawking Page transition being first order does not correspond to a local instability.

Our objective in this note is to examine whether the Hawking-Page transition encountered in the global AdS black string model at a certain temperature, set by the order parameter  $u_h$ , plays an important role in this doubly holographic model which places a BCFT on the surface of a sphere. We are interested in the interplay of islands with the Hawking-Page phase transition. In particular, we are interested in obtaining a connection between the Hawking-Page transition and a general phase transition that occurs for higher-dimensional RT surfaces at the critical angle Geng et al. (2021b). The area difference between the Hartman-Maldacena (HM) surface, which encodes the increasing entanglement entropy of the black hole on the brane, and the island surface is known to diverge to negative infinity at the critical angle when there is a black hole on the brane. However, note that the area difference vanishes there in empty AdS. But we are mainly interested in the fact that island RT surfaces do not exist at all below the critical angle in empty AdS, and that for any asymptotically AdS geometry, the entanglement entropy above the critical angle is governed by tiny islands: global minima that live in the asymptotic region near the defect.

In this work, where we examine the RT phase structure for the global black string, we obtain new scale-dependent behavior realized at the brane’s critical angle ( $\theta_c$ ). We find there exists a hole in the atoll when the brane lies above  $\theta_c$  – namely, there is a region on the brane where RT surfaces cannot anchor. From the intermediate perspective, this leads directly to a discontinuous phase transition for quantum extremal surfaces at  $\theta_c$ . For fixed brane angles above the critical angle, the size of this

region grows monotonically as the black hole decreases in size and vanishes at what we call the saturation angle,  $\theta_s$ . We also find that RT surfaces can connect to the boundary points both above and below the hole, which in some cases leads to multiple candidate RT surfaces that need to be compared. This gives multiple candidate RT surfaces that travel from the brane to the defect, and we have found that the critical anchors for these surfaces lie exactly at the boundaries of the hole – see Figure 6.3 and Figure 6.4. The hole’s origin lies in the non-monotonic behavior of  $h(u)$ , and some surprising connections to the photon sphere for the black hole on the brane will soon be outlined in a separate note. Note that, while the hole exists for any non-zero  $u_h$ , the size of the hole starts to become large on the order of the Hawking-Page phase transition. Finally, we remark that the area difference between the HM and island surfaces vanishes at  $\theta_c$  for a special value of  $u_h$  which we call  $u_h^{\text{crit}}$  – in the language of Geng et al. (2021b), the Page angle equals the critical angle for this value of the order parameter. It can be seen in Figure 6.8 that, even though it is invisible to purely thermodynamic arguments, its value nonetheless plays an important role in the phase structure of this model.

## 6.2 Setup and Equations of Motion

### 6.2.1 The Doubly-Holographic Global Black String

This section provides an overview of the doubly-holographic black string model which we are using to model the entanglement entropy of an evaporating black hole on the KR brane. It is important to realize that there is a major difference between this setup and the one we explored in our previous note Geng et al. (2021d) – in brief, the area density for global  $\text{AdS}_d$  depends on the scale set by the size of the black hole. Once we place a spherical  $\text{AdS}_d$  black hole on each  $\text{AdS}_d$  spacetime slice, we will obtain the global  $\text{AdS}_{d+1}$  black string.

We begin by writing down the metric for an  $\text{AdS}_d$  black hole in a convenient coordinate system. Conventionally setting the  $\text{AdS}_d$  curvature radius to 1, one can

write spherically symmetric metrics as:

$$ds^2 = \frac{1}{u^2} \left[ -h(u)dt^2 + \frac{du^2}{h(u)} + d\Omega_{d-2}^2 \right]. \quad (6.1)$$

To obtain a spherical Schwarzschild black hole, which we will soon place on each constant angle  $\mu$  slice, we choose the following form for  $h(u)$ :

$$h(u) = 1 + u^2 - \frac{u^{d-1}}{u_h^{d-1}}. \quad (6.2)$$

Decreasing the value of  $u_h$  always increases the size of the black hole. The coordinates we are using differ from the more conventional ones – the difference is that we made the coordinate substitution:

$$r \rightarrow \frac{1}{u}. \quad (6.3)$$

and defined:

$$h(u) \equiv u^2 f\left(\frac{1}{u}\right), \quad (6.4)$$

where  $f(1/u) = f(r)$  is the standard blackening function:

$$f(r) = 1 + r^2 - \frac{\omega_{d-1}M}{r^{d-3}}. \quad (6.5)$$

Where  $\omega$  is a constant that is introduced so that  $M$  is the mass of the black hole. By comparing these two coordinate systems, we can see that the parameter  $u_h$  is given by:

$$u_h = \left( \frac{1}{\omega_{d-1}M} \right)^{\frac{1}{d-1}} \quad (6.6)$$

The main advantage of using  $u$  coordinates, instead of  $r$  coordinates, is that the boundary of AdS lies at  $u = 0$  instead of at  $r = \infty$ ; similarly,  $u = \infty$  corresponds to  $r = 0$  in standard coordinates. From a numerical perspective, these coordinates are



much easier to work with than the standard  $f(r)$  coordinates, and they also helped us gain insight into the equations of motion.

In order to make our setup doubly holographic, we take the standard approach of employing the black string by embedding a  $d$ -dimensional KR brane in  $\text{AdS}_{d+1}$  Randall and Sundrum (1999d); Karch and Randall (2001b). This can be done by adding the standard Randall-Sundrum (RS) term Randall and Sundrum (1999d) to the Einstein action while restricting to subcritical brane tensions. The simplest approach is to treat this as an *end-of-the-world brane* by orbifolding the original 2-sided RS geometry. For the black string geometry the brane is embedded in we use:

$$ds^2 = \frac{1}{u^2 \sin^2 \mu} \left[ -h(u)dt^2 + \frac{du^2}{h(u)} + d\Omega_{d-2}^2 + u^2 d\mu^2 \right]. \quad (6.7)$$

In contrast to Poincare-patch AdS, where the transverse coordinates are Cartesian, the transverse coordinates form a sphere  $S_{d-2}$  with a line element proportional to  $d\Omega_{d-2}$ . Due to the presence of the brane, the angular coordinate must be larger than the angle of the brane location as we are taking a positive subcritical brane tension. This removes part of the universe with the angle variable being limited to:

$$\mu \in [\theta_b, \pi). \quad (6.8)$$

The introduction of the KR brane induces an AdS radius on the brane,  $b$ , which differs from the bulk value,  $l$ , according to the following equation:

$$b = \frac{l}{\sin \mu}. \quad (6.9)$$

### 6.2.2 The Doubly-Holographic System and Instabilities

Here we discuss the doubly-holographic system which we are studying, beginning with a review of the black string and some well-known instabilities: the Hawking-Page phase transition and the spinodal point.

### 6.2.2.1 The Location of the Black String

The location of the Schwarzschild black string in the bulk, which has been plotted numerically in Figure 6.1, can be determined by determining where the blackening function  $\frac{h(u)}{u^2}$  vanishes, which is equivalent to solving for the roots of the *non-monotonic* function  $h(u)$ :

$$h(u) = 1 + u^2(\mu) - \frac{u^{d-1}}{u_h^{d-1}} = 0 \quad (6.10)$$

In our case, where  $d = 4$ , we have:

$$u_+ = \frac{1}{3} \left( u_h^3 + \left( \frac{2}{\chi} \right)^{\frac{1}{3}} u_h^5 + \left( \frac{2}{\chi} \right)^{-\frac{1}{3}} u_h \right), \quad (6.11)$$

with  $\chi(u_h) \equiv 2u_h^6 + 3(9 + \sqrt{81 + 12u_h^6})$  and  $u_+ \in \mathbb{R}^+$ , introduces a scale to the theory with  $u_h$  acting as an order parameter. The advantage to focusing on  $d = 4$  is that the system has enough dimensions to have interesting dependence on scale. There are also some less interesting cases where  $d = 2$  or  $d = 3$ , where the horizon is given by:

$$R_2 = \frac{1}{2u_h} - \frac{\sqrt{1 - 4u_h^2}}{2u_h} \text{ if } u_h < \frac{1}{2} \quad (6.12)$$

$$R_3 = \frac{u_h}{\sqrt{1 - u_h^2}} \text{ if } u_h < 1 \quad (6.13)$$

These systems lack interesting dependence on scale for different reasons – the  $d = 2$  system is scale invariant, and the  $d = 3$  system is the BTZ black hole on the brane. Note that, for a general number of dimensions, the location of the horizon can be straightforwardly obtained using standard root-finding methods.

### 6.2.2.2 The Hawking-Page Phase Transition

Before discussing the full braneworld geometry, let us focus on  $\text{AdS}_d$  gravity and its  $\text{CFT}_{d-1}$  dual. The standard lore is that conformal field theories, on a manifold

$M$  of dimension  $d - 1$ , can be studied holographically by summing over the contributions of different gravitational bulk theories,  $B$ , of dimension  $d$ , which have  $M$  as their boundary. Our bulk theory  $B$  is susceptible to an interesting phase transition, known as the Hawking-Page phase transition Hawking and Page (1983), because the thermal state on the boundary,  $M$ , of the bulk lives on a sphere.

When the thermal boundary state lives on a sphere, there are two known bulk theories,  $B$ , which need to be compared Witten (1998b) – the first is empty AdS with a spherical boundary theory, also called thermal AdS,

$$ds^2 = \left( \frac{r^2}{b^2} + 1 \right) dt^2 + \frac{dr^2}{\left( \frac{r^2}{b^2} + 1 \right)} + r^2 d\Omega^2, \quad (6.14)$$

and the second is a standard AdS Schwarzschild black hole,

$$ds^2 = \left( \frac{r^2}{b^2} + 1 - \frac{\omega_{d-1} M}{r^{d-3}} \right) dt^2 + \frac{dr^2}{\left( \frac{r^2}{b^2} + 1 - \frac{\omega_{d-1} M}{r^{d-3}} \right)} + r^2 d\Omega^2. \quad (6.15)$$

Here  $\omega_{d-1}$  is a constant we introduce so that  $M$  is the mass of the black hole, but its precise value does not play an important role in our discussion.

The difference between the infinite actions associated to these geometries,  $I$ , was first computed for three spatial dimensions,  $d - 1 = 3$ , by Hawking and Page Hawking and Page (1983). It is also computed by Witten Witten (1998b) for a general number of spatial dimensions. The action turns out to be proportional to the volume and is given by:

$$I = \frac{d-1}{8\pi G_N} \lim_{R \rightarrow \infty} (V_2(R) - V_1(R)) = \frac{\text{Vol}(\mathbf{S}^{d-2}) (b^2 r_+^{d-2} - r_+^d)}{4G_N ((d-1)r_+^2 + (d-3)b^2)}, \quad (6.16)$$

where  $r_+$  is the radius of the black hole in radial coordinates and  $b$  is the  $AdS_d$  curvature radius. These radial coordinates,  $r$ , are related to our coordinates,  $u$ , by the coordinate transformation<sup>2</sup>:

---

<sup>2</sup>This differs from Equation 6.3 because the  $\sin(\mu)$  term arises when the braneworld Schwarzschild black hole is placed into a warped geometry.

$$r = \frac{1}{u \sin \mu} \quad (6.17)$$

The Hawking-Page transition occurs when  $I$  changes sign – in other words, an AdS black hole is unstable when its radius, measured from the origin *on* the KR braneworld in  $r$  coordinates, is less than the brane AdS radius:

$$r_+ < b \implies \text{The AdS Black Hole on the Brane is Unstable} \quad (6.18)$$

The next step is to convert to the coordinates used in our note. Taking  $u_+$  to be the location of the horizon in our coordinates, Equation 6.17 gives:

$$u_+ \sin \mu = \frac{1}{r_+} > \frac{1}{b} = \frac{\sin \mu}{l} \quad (6.19)$$

$$u_+ > \frac{1}{l} \implies \text{The AdS Black Hole on the Brane is Unstable} \quad (6.20)$$

Here we choose a bulk AdS radius  $l = 1$ , so the Hawking-Page transition occurs when  $u_+ = 1$ . From Equation (6.11), it can be seen immediately that the Hawking-Page phase transition occurs at  $u_h = 2^{-\frac{1}{3}}$ . Note that the change in the AdS radius on the brane was canceled by the  $\sin(\mu)$  term from the warped geometry.

### 6.2.2.3 The Spinodal Point

To facilitate our review of the spinodal transition, we derive a well-known and generically useful result that relates the temperature of a black hole to the derivative of its blackening function at the horizon. Following the standard procedure Witten (1998b), we perform a Taylor expansion in the near horizon region:

$$ds^2 = \frac{dr^2}{f(r)} + f(r)dt^2. \quad (6.21)$$

When working in Euclidean signature, the metric encounters a conical singularity at the horizon which poses a problem for the periodic time coordinate. This can

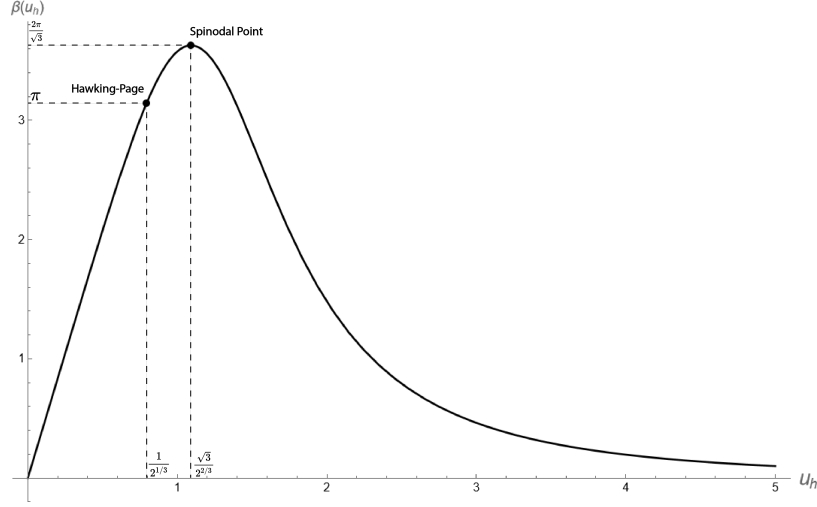


Figure 6.2: The temperature of the black hole scales non-monotonically with  $u_h$ , but for large black holes, the relationship is approximately linear. Unlike in asymptotically flat space, where black holes always have negative specific heat, AdS black holes have positive specific heat when they are large enough. The canonical ensemble in AdS is sensible because a temperature can be defined in the dual theory. For  $\beta$  above the Hawking-Page phase transition, where the radius of the black hole is the AdS radius, the system will transition to a gas of thermal radiation. Such black holes cannot be in thermal equilibrium in the canonical ensemble. For  $u_h$  past the spinodal point, black holes become hotter as they evaporate, so their specific heat is negative. Since AdS black holes do not exist above the value of  $\beta$  where the spinodal point occurs, they do not contribute at all to the canonical ensemble.

be repaired by carefully tuning the circle's circumference and amounts to fixing an inverse temperature,  $\beta$ , which makes the time-circle smooth at and outside the horizon. Start by noting that the metric near the horizon, written in Euclidean signature, is asymptotically flat with the periodic  $\phi$  playing the role of a time coordinate:

$$ds^2 = dr^2 + r^2 d\phi^2 \quad (6.22)$$

Defining a new coordinate that is measured from the horizon,  $\xi$ , the Taylor expansion for the metric near the horizon yields:

$$f(\xi) = f(r_+) + f'(r_+)(\xi - r_+) + \mathcal{O}(\xi^2) \quad (6.23)$$

One can match coefficients in the region near the horizon while pairing the periodic coordinates – that is, one should match the time coordinate,  $t$ , with the  $\phi$  coordinate. Solving the resulting system of equations gives an expression for the differential coordinate along the circle,  $d\phi$ :

$$\frac{d\xi^2}{f'(r_+)\xi} + f'(r_+)\xi dt^2 = d\rho^2 + \rho^2 d\phi^2 \quad (6.24)$$

$$d\phi = \frac{f'(r_+)}{2} dt \quad (6.25)$$

When working in Euclidean signature, the system's time coordinate is periodic in the system's inverse temperature,  $\beta$ . Requiring that the time coordinate makes it "all the way around" the circle at the horizon yields the known expression for the inverse temperature of the system:

$$2\pi = \frac{f'(r_+)}{2}\beta \implies \beta = \frac{4\pi}{f'(r)|_{r_+}} \quad (6.26)$$

The blackening function for the spherical AdS Schwarzschild black hole, given by Equation (6.15), then gives an expression for its temperature:

$$\beta = \frac{4\pi}{f'(r)|_{r_+}} = \frac{4\pi r_+ b^2}{(d-1)r_+^2 + (d-3)b^2} = \frac{4\pi u_+ b^2}{(d-1) + (d-3)b^2 u_+^2} \quad (6.27)$$

The local minimum occurs when the derivative with respect to  $r_+$  vanishes:

$$r_+ = b\sqrt{\frac{d-3}{d-1}} \implies r_+ = \frac{b}{\sqrt{3}} \iff u_+ = \sqrt{3} \quad (6.28)$$

As in the previous section, the relationship 6.9 between the AdS radius on the brane,  $b$ , and the AdS radius in the bulk,  $l$ , cancels the  $\sin(\mu)$  term that arises due to the warped geometry. It can be seen in Figure 6.2 that, at least for large black holes,  $u_h$  is roughly proportional to the inverse temperature. This can be shown analytically. In the limit of large  $M$ , we can use an equation from Witten (1998b), together with Equation 6.6, to show that the temperature of the black hole is approximately:

$$\beta = \frac{4\pi}{f'(r)|_{r_+}} \approx \frac{4\pi b^2}{(d-1)(\omega_{d-1}b^2)^{\frac{1}{d-1}}(M)^{\frac{1}{d-1}}} = u_h \left( \frac{4\pi}{d-1} \right) b^{\frac{2(d-2)}{d-1}}. \quad (6.29)$$

#### 6.2.2.4 The Gregory Laflamme Instability

From the perspective of the intermediate picture, where the conformal boundary and the KR brane are connected through the defect, the black string is perceived as two eternal AdS Schwarzschild black holes. These correspond to where the black string crosses the gravitating KR-brane and the non-gravitating conformal boundary, and one black hole lives in each region. Since the brane is gravitating, the AdS-Schwarzschild black hole on the brane will be susceptible to the Hawking-Page phase transition.

Beyond a certain value of  $u_h$ , the cylindrical topology of the black string is unstable to small perturbations which lead to the formation of a series of black holes along the extra dimension Gregory and Laflamme (1993). This is the well-known Gregory Laflamme instability for the black string. Recent results on this instability Marolf and Santos (2019); Emparan et al. (2022) indicate that the Gregory Laflamme instability in fact happens at black hole radii even smaller than the spinodal point. This is rather surprising as one would have thought that the local thermodynamic

instability that sets on at the spinodal should be represented by exactly such an instability of the black string against small fluctuations. It would be interesting to understand the interplay between Gregory Laflamme and the well-known thermodynamic transitions of the black hole on the brane in more detail. In this work we will not consider Gregory Laflamme any further and will focus on the interplay between Hawking Page, the spinodal point, and the formation of islands.

### 6.2.3 Enforcing The Ryu-Takayanagi Prescription

Now that we have constructed our doubly-holographic model, our objective is to determine the time-dependent entanglement entropy between the two black holes in the intermediate picture by computing the areas of classical RT surfaces anchored to the defect in the bulk picture. To do this we will need to follow the RT prescription.

The RT prescription, which requires us to determine the minimal extremal surfaces  $\Sigma$  that are homologous to  $\mathcal{R}$  Fujita et al. (2011); Almheiri et al. (2019c), leads directly to the formation of islands when these surfaces terminate on the end-of-the-world brane Penington (2019); Almheiri et al. (2019a). First we consider a generic interval on the brane  $\mathcal{J}$ —a “candidate” island. We then determine extremal surfaces  $\Sigma$  that satisfy the homology constraint,

$$\partial\Sigma = \partial\mathcal{R} \cup \partial\mathcal{J}. \quad (6.30)$$

The next step is to apply a variational principle to the *area density functional*  $\mathcal{A}$ , which can be obtained by computing the square root of the determinant of the constant-time slices of the metric. This determines the area of the surface up to a factor, proportional to the volume of the suppressed dimensions, that does not affect the phase structure:

$$\mathcal{A} = \int \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + \frac{u'(\mu)^2}{h(u)}}, \quad (6.31)$$

where  $h(u)$  is given by:



$$h(u) = 1 + u^2(\mu) - \frac{u^{d-1}}{u_h^{d-1}}. \quad (6.32)$$

The action parameterized as  $\mu(u)$  can be obtained using the same procedure:

$$\mathcal{A} = \int \frac{du}{(u \sin \mu)^{d-1}} \sqrt{u^2 \mu'(u)^2 + \frac{1}{h(u)}} \quad (6.33)$$

Now we vary the area functional to determine the  $\Sigma$  and corresponding  $\mathcal{J}$  for which the area is minimized. Note that, in contrast to our previous work, the size of the black hole plays an important role in this note. While we can bring the  $u_h$  outside the integral for the action as an overall prefactor  $u_h^{2-d}$  when the black string is large, we cannot make this simplification for a scale-dependent black string in global AdS.<sup>3</sup> In our case, the size of the black hole – which now depends on  $u_h$  – introduces a scale to the theory which determines the width of the black string and plays a major role in the phase structure.

The Euler-Lagrange equations for the action (6.31) will be ordinary differential equations because of the parameterizations  $u = u(\mu)$  and  $\mu = \mu(u)$ . As discussed in Geng et al. (2021b), the boundary terms in the variation of  $\mathcal{A}$  vanish by imposing boundary conditions on  $\Sigma$ . Following the same argument, we impose a *Dirichlet condition* at the conformal boundary and a *Neumann condition* – requiring that the first derivative  $u'$  vanishes – at the brane. The Neumann boundary condition demands that RT surfaces anchor to the brane at right angles. Depending on which coordinate system –  $u(\mu)$  or  $\mu(u)$  – we choose to use, we have:

$$u'(\theta_b) = 0 \quad (6.34)$$

$$\mu'(u_b) = \infty \quad (6.35)$$

---

<sup>3</sup>For this reason, for ease of comparison with our previous note we have rescaled the axis by  $u_+^2$  in Figure 6.7.

It's possible to gain some physical intuition into these boundary conditions. The first states that the RT surface is struck at a right angle at the brane, since increasing the angle does not change the radial distance. The second statement follows from taking the reciprocal.

### 6.2.3.1 The Equations of Motion

Since we wrote the action in terms of two different parametrizations, we can obtain the corresponding equation of motion for each of them separately. Varying the action (6.31) and solving the resulting Euler-Lagrange equation for the second derivative  $u''(\mu)$  yields:

$$u''(\mu) = -(d-2)h(u)u(\mu) + u'(\mu) \left( (d-1) \cot \mu - (d-3) \frac{u'(\mu)}{u(\mu)} \right) + \left( \frac{u'(\mu)}{u(\mu)} \right)^2 \left( \frac{u(\mu)^2 h'(u) + 2(d-1) \cot(\mu) u'(\mu)}{2h(u)} \right). \quad (6.36)$$

We can similarly vary the action (6.33) and solve the resulting Euler-Lagrange equations for the second derivative  $\mu''(u)$ , which yields:

$$\mu''(u) = (d-2)h(u)u\mu'(u)^3 + \left( \frac{\mu'(u)}{u} \right) ((d-3) - (d-1)u \cot(\mu(u))\mu'(u)) - \frac{2(d-1) \cot(\mu(u)) + u^2 h'(u)\mu'(u)}{2u^2 h(u)} \quad (6.37)$$

These equations of motion are useful in different situations. The first set of equations is typically more useful when solving the shooting problem from the brane, since the RT surface will lie at a right angle to the radial direction, which sets  $u'(\mu) = 0$ . The second set is typically more useful when shooting from the defect, since the RT surface will point along the radial direction, which sets  $\mu'(u) = 0$ .

For example, when shooting from the brane we have  $u'(\mu) = 0$ , which reduces our  $u(\mu)$  equations of motion to:

$$u''(\mu) = -(d-2)h(u)u(\mu) \quad (6.38)$$

It follows immediately that RT surfaces that shoot from the horizon of the black hole, where  $h(u) = 0$ , will remain on the horizon until they reach the boundary. Indeed, in such cases we have  $u''(\mu) = 0$  and  $u'(\mu) = 0$ , so  $u(\mu)$  lies at a stationary point.

Similarly, when shooting from the defect we have  $\mu'(u) = 0$ , which reduces our  $u(\mu)$  equations of motion to:

$$\mu''(u) = -\frac{(d-1)\cot(\mu(u))}{u^2h(u)} \quad (6.39)$$

Under similar reasoning, this equation is stationary when  $\mu = \pi/2$ . Such solutions correspond to the Hartman-Maldacena surface Hartman and Maldacena (2013) which travels from the defect, crosses the black hole horizon, and travels through the Einstein-Rosen bridge to the other side of the thermofield double.

### 6.3 Numerical Results

Here we present our main results. Our objective is to understand how the scale-dependent location of the horizon, given by Equation 6.11, affects the phase structure, especially at the critical angle. We find that introducing a scale induces a discontinuous phase transition in the entanglement wedge at the critical angle for the brane. This is in contrast to Geng et al. (2021b), where the scale-less theory had a continuous phase transition at the critical angle. We will explain how this new behavior comes about due to a "hole" in the *atoll*, which is the region on the brane where RT surfaces can shoot from the brane to the bath.

The standard approach is followed throughout. By using our equations of motion, given by Equations 6.36 and 6.37 respectively, together with the appropriate boundary conditions, given by Equations 6.34 and 6.35, we identify the candidate

extremal surfaces which compete to dominate the entropy functional, given by Equations 6.31 and 6.33. As in our previous work Geng et al. (2021b), we are interested in the boundary conditions satisfying the homology constraint in Equation 6.30, namely,  $\Gamma = 0$ , which anchors the RT surfaces to the defect and determines which regions can, in principle, be reconstructed using information localized at the defect.

### 6.3.1 Numerical Approach: The Shooting Method

The RT surfaces were obtained using the following procedure, which amounts to using the shooting method at the defect:

1. Using the equations of motion – in our case, Equation 6.33 – expand  $\mu(u)$  in series in the asymptotic region near the defect. Solutions can then be obtained near the defect using an asymptotic expansion.
2. The asymptotic expansion cannot satisfy the boundary conditions, given by Equation 6.35, which require  $\mu'(u)$  to diverge at some finite value of  $u$ . One must define a suitable cutoff region for the asymptotics and switch to numerics when the asymptotic solution crosses the boundary of that region. Since the equations of motion are second order, the values of  $\mu(u)$  and  $\mu'(u)$  at the cutoff region will suffice to perform the numerics.
3. For some given parameters, the first derivative  $\mu'(u)$  is found, numerically, to diverge to positive infinity at some angle  $\theta_b$ . Therefore the boundary conditions, given by Equation 6.35, are obeyed at this angle. While such solutions will not necessarily be unique, each corresponds to an RT surface anchored to a KR brane at an angle  $\theta_b$ .

This method should be widely applicable to doubly-holographic models where RT surfaces anchored to the defect need to be calculated.

### 6.3.2 The Series Expansion

The series expansion near the defect for  $\mu(u)$ , which is needed to perform the shooting method described in the previous section, can be written as:

$$\mu(u) = \mu(0) + \alpha_1 \mu'(0) + \alpha_2 \mu''(0) + \alpha_3 \mu'''(0) + \dots \quad (6.40)$$

In our coordinate system the RT surface intersects the defect at a right angle at  $\mu = \frac{\pi}{2}$ , so we have  $\mu(0) = \pi/2$  and  $\mu'(0) = 0$ . It is straightforward to show that  $\mu''(0) = 0$  by using the equations of motion for  $\mu(u)$  in (6.39), and so:

$$\mu(u) = \frac{\pi}{2} + \alpha_3 \mu'''(0) + \dots \quad (6.41)$$

This equation gives  $\mu(u)$  in the asymptotic region as a function of just one parameter  $\alpha_3$ , with each of its values potentially yielding a solution satisfying the boundary conditions for a brane at some angle  $\theta_b$ . Result 1, see Appendix 6.A, points out that the RT surfaces anchored to the defect will not necessarily be unique, and in fact, we have obtained multiple solutions for each brane angle above the critical angle.

### 6.3.3 Multiple Critical Anchors, and The Hole in the Atoll

It was shown in previous work Geng et al. (2021b,d) that RT surfaces are restricted to parts of the brane in higher dimensional models. The region where candidate RT surfaces can anchor to the KR brane, thereby forming islands on the brane when they dominate the entropy functional, has been called the *atoll*. The point where defect-anchored RT surfaces form an island by anchoring to the brane is called a *critical anchor*. We can, for example, use these definitions to restate a relevant result from an earlier paper – when there is a *planar* black hole on each constant  $\mu$  slice, as in Geng et al. (2021b,d), the *atoll* is the region beyond the critical anchor on the brane. Note that, above the *critical angle*, tiny island surfaces dominate the entropy functional. Please see Figure 6.6 and Appendix 6.B for more details.

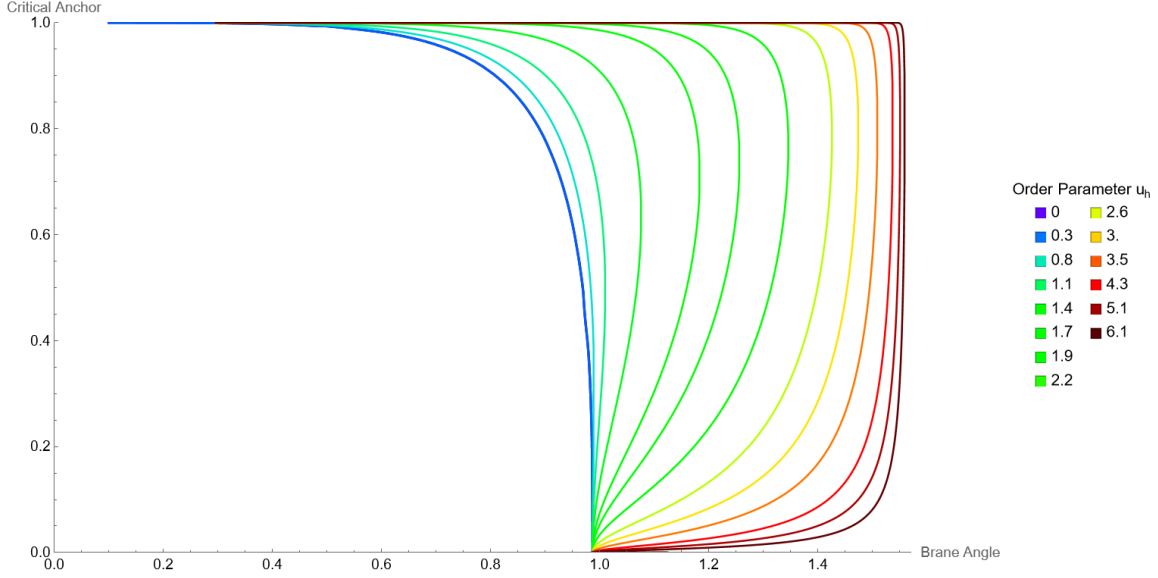


Figure 6.3: This figure shows the critical anchor on the brane, as a function of brane angle, as a percentage of the horizon radius, for various choices of order parameter  $u_h$ . The black string’s diameter, and correspondingly, the black hole’s size on the KR brane, decreases (increases) as  $u_h$  is increased (decreased). The Hawking-Page transition lies at  $u_h = \frac{1}{2^{1/3}} \approx .8$ . We recover our result for a large black string, discussed in Geng et al. (2021b), by taking the limit  $u_h \rightarrow 0$ . There is a “hole” on the atoll when the critical anchor plot is not 1-to-1, which occurs for any non-zero value of  $u_h$ . As  $u_h \rightarrow \infty$ , the black string’s diameter shrinks to a point, and the plot approaches a square shape. There is only one critical anchor below the critical angle ( $\theta_c \approx .98687$ ), but there are two critical anchors between the critical angle and the saturation angle  $\theta_s$ .

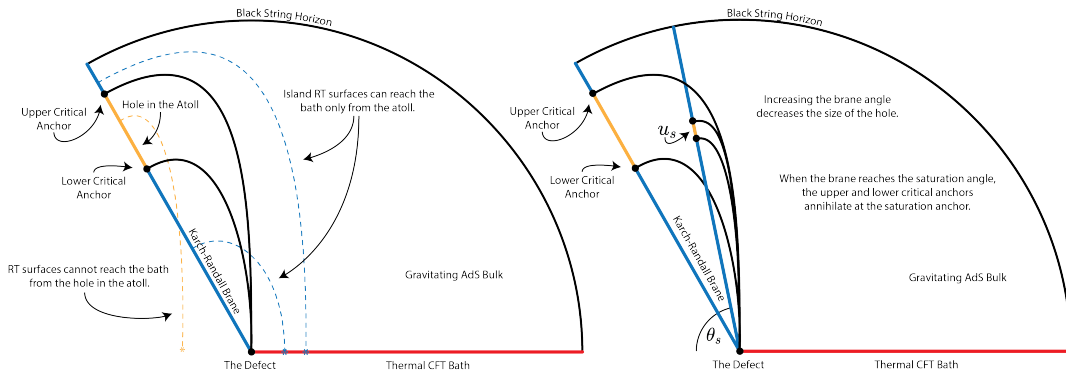


Figure 6.4: This cartoon illustrates typical RT surfaces, above the critical angle  $\theta_c$ , for an eternal black string dual to a BCFT on the surface of a sphere. When the KR brane lies above the critical angle, RT surfaces can travel from the KR brane to the defect from either the upper or lower critical anchor. RT surfaces can reach the boundary only from the atoll, and the region between the critical anchors is always a hole in the atoll. Increasing the brane angle decreases the size of the hole; at the saturation angle, the upper and lower critical anchors annihilate at the saturation anchor. Note that similar observations apply to anchors within a finite distance of the defect on the non-gravitating bath.

Here we generalize our previous results by extending them to *spherical Schwarzschild* black holes on each slice, with a horizon distance set by  $u_h$ . The scale set by  $u_h$  leads to interesting new behavior – note, however, that we can obtain our old results by taking the limit where  $u_h \rightarrow 0$ . We have found that – above the critical angle – there are *two* critical anchors which border a region on the brane which we call the “hole” in the atoll. The critical anchors as a function of  $u_h$  and the angle of the brane have been computed and can be seen in Figure 6.3. Those critical anchors are determined as a percentage of the horizon distance. The hole defines a region on the brane from which RT surfaces cannot reach the bath – please see Figure 6.4 for an illustrative example.

The extent of the hole on the atoll depends on the value of  $u_h$ , according to Equation (6.11), with larger values of  $u_h$  leading to smaller Schwarzschild black holes and larger holes in the atoll. Its size can be seen by inspecting Figure 6.3 – for any non-zero value of  $u_h$ , the critical anchor plot stops being 1-to-1 above the critical angle,  $\theta_c$ , and increasing the value of  $u_h$  causes the plot to become not 1-to-1 within an increasingly wide strip of  $\theta_b$  values. For those brane angles, there is more than one critical anchor, and the hole is defined as the region between the two critical anchors.

It can be seen in Figure 6.3 and Figure 6.4 that, for any fixed value of  $u_h$ , increasing the brane angle decreases the size of the hole up until what we call the *saturation angle*  $\theta_s$ . The hole does not exist beyond that point because it shrinks to a point on the brane which we call the *saturation anchor*  $u_s$ . Beyond the saturation angle, the atoll saturates the brane, and an RT surface can reach the bath from any point on the brane – i.e., there is no hole in the atoll beyond the saturation angle. The saturation angles and saturation anchors are given in Figure 6.5. We observe that in the limit where  $u_h \rightarrow \infty$ , the saturation angle approaches  $\pi/2$ , and the saturation anchor approaches  $\approx 81\%$  of the horizon distance. This generalizes our result for the planar black string Geng et al. (2021b,d), where the atoll first saturates the brane at the critical angle  $\theta_c$  as we approach it from below and continues to saturate it above



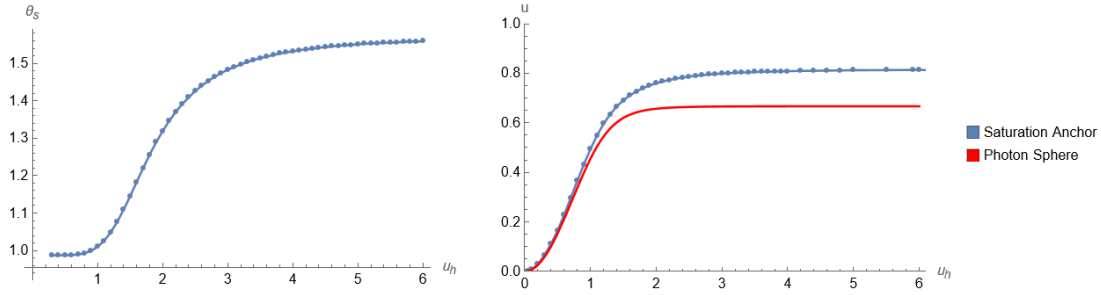


Figure 6.5: This figure shows the saturation angle for the brane on the left, and the saturation anchor for the brane on the right, both as a function of  $u_h$ . For a given value of  $u_h$ , the saturation angle is the largest brane angle with a hole on the atoll. As the brane angle increases, the hole shrinks to a point that we call the saturation anchor. Our numerics are consistent with Result 6.2, in the Appendix, which states that the saturation anchor always lies inside the photon sphere.

– for the global black string, the atoll saturates the brane at the saturation angle  $\theta_s > \theta_c$ .

### 6.3.4 Area Differences

We have used the equations of motion to determine which extremal surfaces can reach the defect, located at  $\Gamma = 0$ , and we have identified a "hole" in the atoll on the brane. There is a second question concerning these surfaces that is also interesting – which bulk region is holographically dual to the defect, in the sense that we can use boundary information located on the defect to reconstruct its contents? This is equivalent to asking which extremal surface, for a given brane angle  $\theta_b$ , and KR braneworld Schwarzschild black hole size (set by  $u_h$ ), dominates the entropy functional on a given time-slice.

The answer to this question depends critically on the tiny island surfaces, which always (never) dominate the area functional for  $\Gamma = 0$  above (below) the critical angle  $\theta_c$ . More details are available in Geng et al. (2021b), and we also review the tiny islands in Appendix 6.B. To answer this question below the critical angle, where the tiny islands are never dominant, we have computed the area differences between

the "candidate" RT surfaces: the Hartman-Maldacena surfaces, which plunge from the defect through the horizon, and the island RT surfaces, which travel from the defect to the brane. While these areas are formally infinite, their area difference is finite and given by:

$$\Delta A(t) = A_{IS} - A_{HM}(t), \quad (6.42)$$

for a fixed time slice, which we take to be  $t = 0$ . Note that the Hartman-Maldacena surface is dominant (subdominant) when  $\Delta A$  is positive (negative). In other words, when the area difference is positive,  $A_{IS} > A_{HM}$  the Hartman-Maldacena surface is dominant on the initial time-slice and we get an initially increasing (non-trivial) Page curve. Similarly, when the area difference is negative,  $A_{HM} > A_{IS}$  the island surface is dominant on the initial time-slice and we get a constant (trivial) Page curve.

This area difference depends on the time coordinate,  $t$ , because the Hartman-Maldacena (HM) surface traverses the Einstein-Rosen bridge – hence, its area increases roughly linearly at late times, and its area will eventually exceed that of the island RT surface. So when the island RT surface does not start out dominant on the initial time slice, it will dominate after the area of the Hartman-Maldacena surface has grown by an amount equal to the area difference on the initial time slice. This happens at a coordinate time which we conventionally call the Page time, denoted  $t_p$ .

We are mainly interested in whether the islands dominate on the initial time slice. To determine which RT surface is dominant below the critical angle, it will suffice to compute  $A(0)$ , beginning with the HM surface. Since the anchor point on the bath lies on the defect (in other words,  $\Gamma = 0$ ) the HM surface drops straight down, on the  $\mu = \pi/2$  slice, and crosses the black string horizon at  $u = u_+$ . Since  $\mu'(u) = 0$  along its trajectory, we can use 6.33 to show that its area outside the horizon is given by:

$$A_{HM}(0) = \int_0^{u_+} \frac{1}{u^3} \sqrt{\frac{1}{h(u)}}. \quad (6.43)$$

This equation has closed-form solutions, which can be obtained using Mathematica but are cumbersome to write down. Meanwhile, the areas of the island RT surfaces must be obtained by solving the equations of motion (6.37) using the shooting method, as described in Section 3.1, and evaluating the action (6.33) numerically. The results we obtained by carrying out this procedure can be seen Figure 6.7. The next step, below, is to use these results to determine the phase structure.

### 6.3.5 Phase Structure and the Critical Angle

One of the key features of the phase structure for the black string is a discontinuous phase transition in the entanglement wedge at the critical angle,  $\theta_c$ , for the brane. It was established in Geng et al. (2021b) that, since tiny island solutions always dominate the entropy functional above the critical angle, the island phase will dominate on the initial time slice above the critical angle. This leads to interesting new behavior for the black string – when the angle for the brane,  $\theta_b$ , crosses the critical angle,  $\theta_c$ , the entanglement wedge shrinks discontinuously from finite to infinitesimal size – see Figure 6.6 for an illustration. More details about the tiny islands can be found in Geng et al. (2021b) and in the Appendix 6.B.

The area differences are also interesting and can be seen in Figure 6.7, where our numerical data is presented as a function of brane angle,  $\theta_b$ , for various  $u_h$  values.<sup>4</sup> Some general features from the planar model are preserved in the spherical scale-dependent case. For any fixed value of black hole size, determined by  $u_h$  and Equation (6.11), the area difference decreases monotonically with the brane angle. Hence, as before, island surfaces trend toward dominance as the brane angle increases. On the other hand, the dependence on the scale,  $u_h$ , is new. We find that, for any fixed value of the brane angle,  $\theta_b$ , the area difference increases monotonically with  $u_h$ .<sup>5</sup> According to (6.11), and Figure 6.1, increasing the value of  $u_h$  decreases the size of

---

<sup>4</sup>The numerics become more challenging as  $u_h$  increases.

<sup>5</sup>Since  $u_+$  increases with  $u_h$ , the rescaling of our axis by  $u_+^2$  does not affect this result.

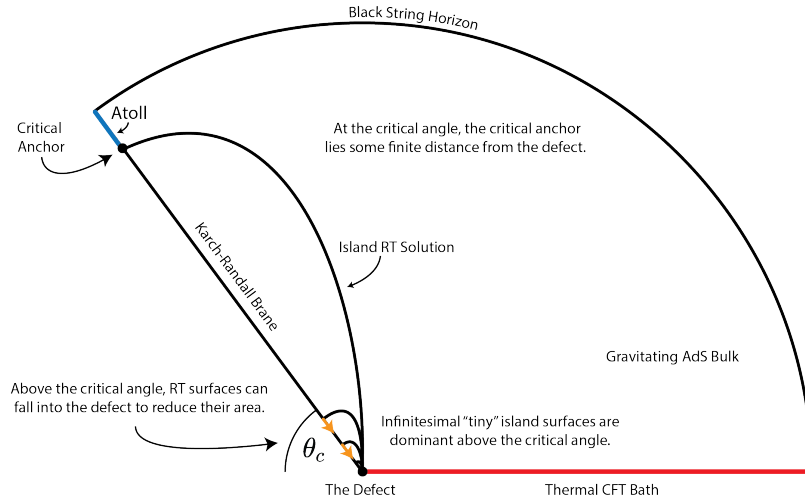


Figure 6.6: According to Figure 6.3, the critical anchor lies at a finite distance from the defect, which depends on the order parameter  $u_h$ , at the critical angle  $\theta_c$ . Then for  $u_h \neq 0$ , which corresponds to non-planar black holes on each slice, the phase transition at the critical angle is discontinuous because tiny island surfaces dominate the entropy functional at and above  $\theta_c$ . For a review of tiny islands, see the Appendix 6.B. By sending  $u_h \rightarrow 0$ , one recovers the result for planar black holes in our previous note Geng et al. (2021b) – the critical anchor drops into the defect, and the phase transition becomes continuous. Note that, in contrast to Geng et al. (2021b), the atoll does not saturate the brane at the critical angle when  $u_h \neq 0$ .

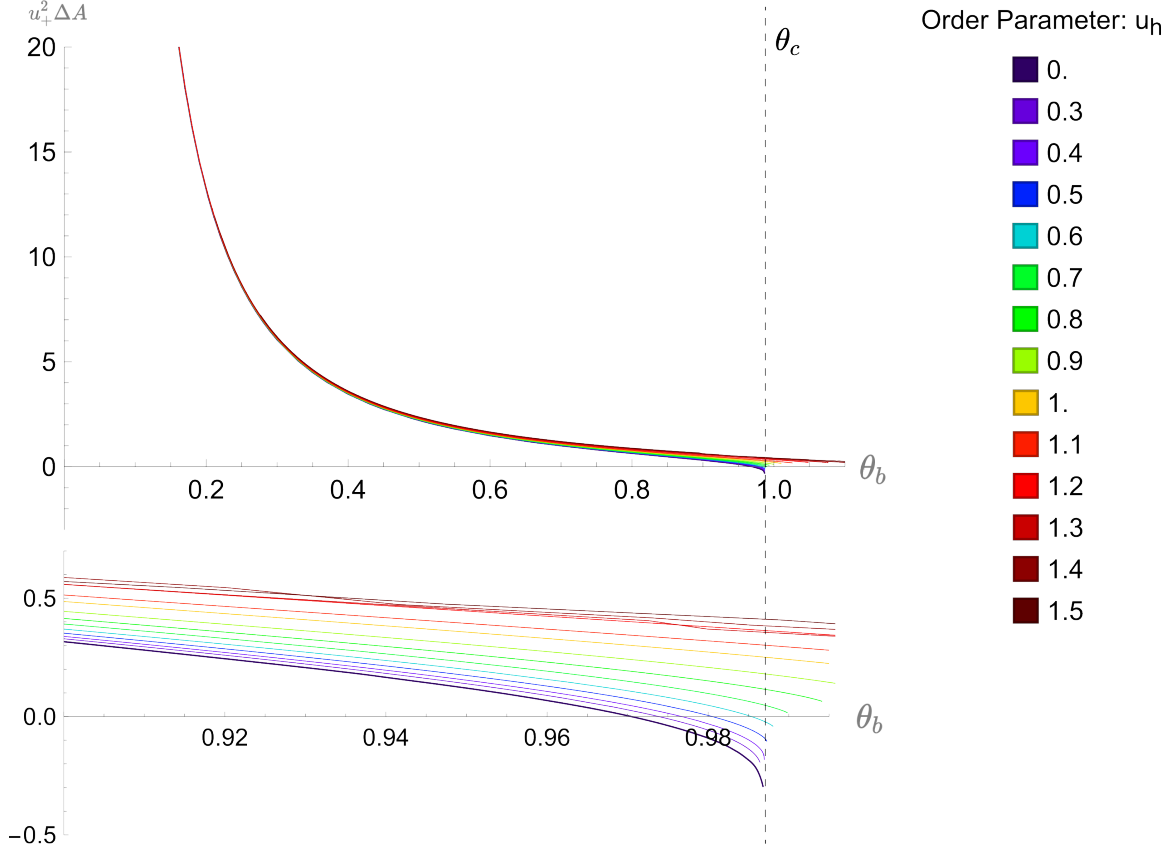


Figure 6.7: Here are the numerically computed area differences between the Hartman-Maldacena (HM) surface and the island RT surface for each value of  $u_h$  below the critical angle  $\theta_c \approx .98687$ . The bottom panel zooms in on the region where the area difference vanishes. We find numerically that the Page angle equals the critical angle at roughly  $u_h \approx .62777$ . While the area differences *above* the critical angle are also displayed for completeness, the actual dominant surface for  $\theta > \theta_c$  will be the tiny island surface, for which the area difference actually diverges to negative infinity – see the Appendix 6.B. By sending  $u_h \rightarrow 0$ , one obtains our old result Geng et al. (2021b), given by the thick dark curve, where the island RT surface degenerates into the tiny island surface at the critical angle – so for  $u_h = 0$ , the area difference diverges to negative infinity at  $\theta_c$ . For ease of comparison with our previous note, the vertical axis has been rescaled by the square of the horizon radius  $u_+$ .

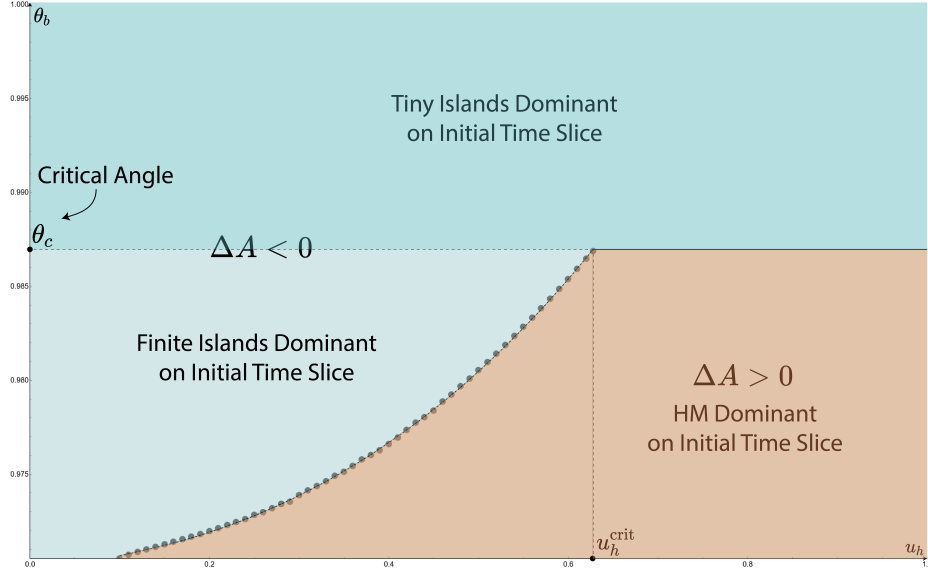


Figure 6.8: Here we present our phase diagram for the doubly-holographic global black string model. Since the thermal BCFT state lives on the surface of a sphere, there is a scale dependence in the model that depends on  $u_h$  – see Figure 6.1 and Figure 6.2. Each point on this plot gives the brane angle for which the area difference vanishes, which we call the Page angle  $\theta_p$ . The Page angle was found to increase with  $u_h$ . The island surfaces are dominant when the area difference is negative. It was shown in previous work Geng et al. (2021b) that the tiny island phase is dominant above the critical angle because the area difference diverges to  $-\infty$ . In contrast, finite-size islands dominate below the critical angle when the brane lies above the Page angle. When  $u_h > u_h^{\text{crit}}$ , which is  $\approx .62777$ , the island surfaces are dominant only when the brane lies above the critical angle.

the black hole on the KR brane and narrows the diameter of the black string. Hence, islands tend to lose dominance as the size of the black string decreases.

The phase diagram for the eternal black string in  $d = 4$  is presented in Figure 6.8. The value of  $\theta_b$  where the area difference vanishes is interesting since the island phase becomes dominant when the sign of the area difference changes. In previous work, we have called this value the Page angle  $\theta_p$ . Below the critical angle,  $\theta_c$ , there is genuine competition for dominance between the HM and island surfaces because the tiny islands are not dominant. For each value of  $u_h$  below  $u_h^{\text{crit}}$ , which will be explained in a moment, we obtain a value of  $\theta_b$  where the area difference vanishes

below the critical angle. We find that the Page angle equals the critical angle at  $u_h = u_h^{crit}$ , where  $u_h^{crit} \approx 0.62777$  – beyond this point, the island phase is dominant precisely when defect-anchored RT surfaces lie above the critical angle. Beyond this critical value for  $u_h$ , which amounts to fine-tuning the radius of the Schwarzschild black hole on the brane, the entropy curve is constant if and only if the brane lies above the critical angle.

Some additional connections to previous work are noted here. We have studied the case for  $u_h = 0$ , for which the diameter of the black string is *much* larger than the AdS radius, before Geng et al. (2021b). In that case,  $\theta_p$  was slightly less than  $\theta_c$  – it was also found that, for empty AdS,  $\theta_c$  was exactly equal to  $\theta_p$  due to the scale invariance in the model. One of the authors has explored another situation where the Page and critical angles are equivalent under certain conditions after an Einstein-Hilbert term is used to introduce DGP gravity to the brane Perez-Pardavila (2023a).

## 6.4 Conclusion

We have detailed the entanglement phase structure for a doubly-holographic black string formed from spherical AdS black holes placed on each spacetime slice. In such BCFT models, the boundary of the CFT, which we refer to as the defect throughout this note, is dual to a Karch-Randall (KR) brane where RT surfaces can connect. When they do so, this defines a quantum extremal surface (QES) and forms an island on the brane. The black holes on the brane and boundary are located where the black string crosses their respective regions, and we have computed the entanglement entropy between them by placing the anchor point for the bulk RT surfaces on the defect. It should be noted that since the RT surfaces in our model are homologous to the defect, our results apply equally well to the case where the bath is gravitating.

Spherical AdS black holes differ from planar ones because they are susceptible

to the Hawking-Page (HP) phase transition. This is connected to the fact that their dual thermal BCFT state lives on the surface of a sphere. Unlike in asymptotically flat space, where black holes have negative specific heat, AdS black holes below the HP phase transition can exist in stable thermal equilibrium with a heat bath. In contrast, when the inverse temperature,  $\beta$ , lies above the HP phase transition the system will transition to a gas of thermal radiation. This important phase transition is related to the observation that the blackening function  $f(r)$  is not monotonic and the temperature achieves a minimum value at the spinodal point. Some intriguing connections will be explored more generally in an upcoming note Riojas.

We have shown that the phase structure, presented in Figure 6.8, for the entanglement entropy between two spherical AdS black holes exhibits drastically different behavior than the planar case explored in Geng et al. (2021b,d). There is a direct connection to the diameter of the black string set by  $u_h$ , which also sets the system's temperature and acts as an order parameter in our model. Recall that when the island surface is dominant on the initial time slice, the entanglement entropy does not increase and there is a constant Page curve. For a fixed brane angle, as  $u_h$  decreases, the area difference decreases. Thus islands become increasingly dominant as we decrease  $u_h$ . When the value of  $u_h$  lies above a critical value, which we call  $u_h^{\text{crit}}$ , entanglement islands are always subdominant unless the KR brane lies above the critical angle  $\theta_c$  identified in Geng et al. (2021b). It is pointed out in Riojas that the Hawking-Page transition, the spinodal point, and the critical transition at  $u_h^{\text{crit}}$  arise because of the non-monotonicity of  $h(u)$ ; also, the photon sphere lies at the unstable maximum value of  $h(u)$  and only exists when  $h(u)$  is monotonic. Nonetheless, we find in this note that these phase transitions take different numerical values. While the quantum information reorganizes itself at a special black hole radius associated with  $u_h^{\text{crit}}$ , this radius does not appear to correspond directly to the thermodynamics of black hole phase transitions.

One of our main findings is that the diameter of the black string in the bulk picture influences the existence of candidate RT surfaces. This is due to new scale-



dependent behavior realized at the brane’s critical angle. When the angle of the brane is greater than  $\theta_c$ , island RT surfaces are forbidden from anchoring to a region of the KR brane which we call the hole in the atoll. It is interesting to note that, from the intermediate perspective, the QES is forbidden from occupying the region inside the hole – this seems to indicate that quantum corrections to the QES are large within the hole. Since the photon sphere always probes the region inside the hole Riojas, it would be interesting to see if it plays a direct role. Our results are illustrated in Figures 6.3 and Figure 6.4. There are potentially up to two candidate RT surfaces that need to be considered for any choice of boundary anchor, with the extremal surfaces anchoring either above or below the hole. At the critical angle, the critical anchors – defined in Geng et al. (2021b) as the anchor points on the brane for extremal surfaces which connect the brane and the defect – lie precisely at the boundaries of the hole. For fixed  $u_h$ , we find that the hole shrinks with increasing brane angle until it vanishes at what we call the saturation angle  $\theta_s$ . We also found that the critical anchors depend on the scale set by the diameter of the black string, and that the hole’s existence induces a discontinuous change in the entanglement wedge at the critical angle.

Several interesting connections exist between this work and our previous note Geng et al. (2021b,d). In particular, we once again find that the area difference between the HM and island surfaces diverges to negative infinity at the critical angle – the main difference is that the change in the size of the entanglement wedge is discontinuous instead of continuous. Note that, when we take the limit as  $u_h \rightarrow 0$ , we obtain the same results which were detailed in Geng et al. (2021b,d). In particular, the hole in the atoll vanishes, the entanglement wedge again vanishes continuously at the critical angle, and the uniqueness of (finite) candidate island surfaces is restored. The restoration of the diverging area difference at the critical angle in this limit can be seen in Figure 6.7. We also noted in Geng et al. (2021d) that the atoll need not be connected, and we have realized that in this particular doubly-holographic model.

It would be interesting to analyze how the two competing surfaces, which appear for general  $\Gamma$  on a non-gravitating bath when a scale is introduced, change the phase structure in Geng et al. (2021d). Since tiny island solutions are unavailable when the RT surface anchors to a finite point on the boundary, there will be competition from the candidate RT surfaces anchoring above and below the hole on the atoll. We anticipate that, for fixed brane angles and anchor points on the boundary, there will be a phase transition between them at some value of  $u_h$ . We also expect that each boundary anchor will have its own saturation angle  $\theta_s$ . It might also be interesting to compute the Page time for various combinations of parameters in this model.

As before, we conclude our discussion by noting that several general considerations should apply to a wide variety of braneworld models, even if the geometry is changed. First we note that, for doubly-holographic models with  $d > 2$ , there will be a critical phase transition at the critical angle  $\theta_c$ . The existence of the hole on the atoll, which appears above the critical angle, is consistent with our predictions for the general structure of the atoll that were laid out in Geng et al. (2021d) – the atoll should contain the region near the horizon on the brane, since the horizon is an extremal surface. It should also contain the region near the defect above the critical angle, since island surfaces exist in that region in empty AdS. Nonetheless, as we find here, it need not be connected. Our results are also consistent with our observation that increasing the brane angle should decrease the area difference between the HM and island surface. Physically one expects that the number of degrees of freedom on the defect decreases with angle, and the defect encodes the braneworld black hole; hence, increasing the brane angle should decrease the Page time. Finally, in this note, for the relatively small  $u_h$  that we have studied, we have observed that the area difference increases with  $u_h$ . This suggests that smaller spherical AdS black holes take longer to saturate their entanglement entropy, at least in this system. It is possible that the area difference turns around as the black hole shrinks with  $u_h$  and heats up. It would be interesting to study this observation in more detail.

## Appendix 6.A: Useful Results about Islands and the Photon Sphere

Here we quote some results from another paper, in preparation by one of the authors, which were relevant to the analysis of this note Riojas.

**Theorem 6.1.** *For spherically symmetric blackening functions, there will be exactly one solution which travels from the brane to the defect unless the black hole on the brane has a photon sphere. In such cases where there are multiple solutions, at most one solution will be outside the photon sphere.*

**Theorem 6.2.** *The atoll is defined as the region on the brane where RT surfaces can anchor, forming an island. In some cases there is a "hole" in the atoll, with islands being possible on both sides of the hole. The hole exists only when there is a photon sphere; when it exists, part of the hole will always lie within the interior of the photon sphere.*

## Appendix 6.B: Tiny Island Surfaces and Flat Page Curves

A crucial role in the phase diagram is played by what we refer to as tiny island surfaces Geng et al. (2021b). These are tiny surfaces that connect the defect to the brane in the asymptotic region near the defect, and their dominance above the critical angle leads directly to the second-order phase transition noted in Section 3. The main point of this subsection is that tiny island surfaces are legitimate *global* minima that dominate the area functional above the critical angle. Nonetheless, while tiny islands have many features in common with RT surfaces, we will review some important distinctions here. Note that the Page curve is flat when tiny islands are dominant, and furthermore, that these comments hold quite generally in higher dimensional AdS Geng et al. (2021b).

Unlike tiny island surfaces, RT surfaces are local minima of the area functional. As such, their area functional vanishes under infinitesimal fluctuations of the RT

surface. This is true both in the bulk and on the boundary – indeed, the vanishing of the variation of the area with respect to the location of the boundary point is a physical requirement that gave rise to the condition (6.34). And since tiny islands do not satisfy condition (6.34), one might well wonder if these saddles are legitimate.

But one must not lose sight of the fact that perfectly legitimate saddles of the area functional can be obtained under finite variations, so long as one obtains a *global* minima for the area functional. Indeed, when the RT endpoints on the brane (the brane anchors) are allowed to fluctuate by a finite amount, the true global minima of the area functional can also occur at the boundaries of the range of the allowed anchors. And this is what tiny island surfaces are – when the brane anchor approaches the defect under a finite variation, the corresponding RT surface degenerates into an infinitesimal surface localized very near the defect.

It was found in Geng et al. (2021b) that the regulated areas of these surfaces are always formally infinite, despite their tiny nature, but the infinity can be either positive or negative. Since these areas are *positive* infinity below the critical angle, tiny islands are *never* dominant in that regime. But it was also found that the areas are *negative* infinity above the critical angle, which means they are *always* dominant in that regime. The physical interpretation of this is as follows: when one remembers to include finite variations above the critical angle, one realizes that the would-be RT surface can simply slide toward the defect to decrease its area without bound. See Figure 6.6. As its endpoint slides down the brane, the area difference between the island surface and the HM surface blows up to negative infinity – in other words, the HM surface becomes infinitely larger – and one obtains a tiny island that dominates the area functional.

It may seem somewhat unsatisfactory to find that the area difference between these surfaces diverges, but this apparent infinity can be resolved using a limiting procedure. Consider a finite endpoint for the RT surface on the boundary, the boundary anchor, by sliding it slightly off the defect – i.e.,  $\Gamma = \epsilon$  for some small  $\epsilon$ . One finds that

the apparently infinite area difference is a large but finite one Geng et al. (2021b). To obtain the tiny island surfaces, one takes the limit as the endpoint on the boundary approaches the defect, i.e.  $\epsilon \rightarrow 0$ . The area difference diverges to negative infinity and one realizes that, for  $\Gamma = 0$ , the tiny islands are indeed the dominant RT surfaces above the critical angle. In such cases, the tiny islands are dominant on the initial time slice and one obtains a flat Page curve.

# Chapter 7: Entropy of Radiation with Dynamical Gravity

## Abstract

We compute the subregion entanglement entropy for a doubly holographic black string model. This system consists of a non-gravitating bath and a gravitating brane, where we incorporate dynamic gravity by adding a DGP term. This opens up a new parameter directly extending previous work and raises an important question about unitarity. In this note we analyse which theories in this big parameter space, will have unitary entropy evolution, in particular, we will distinguish which of those will follow a Page curve.

## 7.1 Introduction

Over the last few years, several new results have addressed and attempted to resolve a long-standing paradox about information and black holes. One important tool utilised in these recent developments is the Ryu-Takayanagi prescription Ryu and Takayanagi (2006c,d); Hubeny et al. (2007b); Faulkner et al. (2013b); Lewkowycz and Maldacena (2013b); Engelhardt and Wall (2015b). In essence, it exploits the AdS/CFT correspondence Maldacena (1998b); Gubser et al. (1998); Witten (1998a) enabling to perform complicated entropy computations in the strongly conformal field theory by identifying them with area computations on the  $AdS$  bulk.

Boundary conformal field theories or BCFTs are  $d$ -dimensional strongly cou-

---

<sup>0</sup>This chapter is based on published work Perez-Pardavila (2023c). Notably, this work laid the foundation for a subsequent paper that incorporates DGP corrections for the brane. Many of the key equations and theoretical developments presented there were first derived in this work.

pled conformal field theories on manifolds with boundary Cardy (2004). The boundary, a  $d - 1$  manifold, is usually called the defect. Using holography once, we arrive at the so-called “intermediate picture”, in which the boundary is dual to  $AdS_d$  gravity Karch and Randall (2001c); Takayanagi (2011b); Fujita et al. (2011) localised in a Karch-Randall (KR) Karch and Randall (2001b); Randall and Sundrum (1999d) brane coupled to a non-gravitating bath where our CFT lives in Penington (2019); Almheiri et al. (2019a, 2020e, 2019d,c). A second application of holography yields the full picture in which the brane is an end-of-the-world brane embedded in an  $AdS_{d+1}$  space. The CFT now lives in the boundary of this bulk. These double descriptions are usually called “doubly holographic” and throughout this note we will adhere to this convention.

In the past we have used this doubly holographic model to study the phase structure of the entropy in a subregion of a non-gravitating bath Geng et al. (2022d). This was itself an extension to the case where the bath was gravitating, forming a wedge-holographic model Geng et al. (2021b). In this note, however, we will keep the bath non-gravitating as in the previous study. In other words, we will consider a black string state in an  $AdS_{d+1}$ , in which we embed a KR brane. The advantage of this model is that it allows us to place a black hole in the background of the CFT, because the bulk theory is dual to a CFT thermal state in a non-gravitating eternal black hole background. On the other hand, the black string horizon is connected to the horizon of the black hole in the brane and the bath.

The main motivation behind this note is to continue analysing the time evolution of the entropy of a subregion of the bath in relation to the Page curve. The Page curve describes the unitary evolution one should expect entropy to follow according to our understanding of quantum mechanics Lubkin (1978); Page (1993b). In particular, we have previously observed the emergence of a Page curve from the bulk perspective by comparing two competing minimal area surfaces or *RT surfaces* at a classical level. The early time entropy is generally controlled by a horizon penetrating surface, namely the *HM surface* Hartman and Maldacena (2013) which under some

circumstances is then replaced by a lower area surface called the *island surface* whose area is constant Almheiri et al. (2019c). When this happens at some finite time, we say that the evolution of the entropy follows a Page curve. The latter type of surfaces anchor on the brane, enclosing the so-called *islands* on it. The whole surface is the union of the island surface and its partner on the thermofield double side. These have been previously studied in Ling et al. (2020); Kumar Basak et al. (2020b); Emparan et al. (2020); Caceres et al. (2021a,b); Deng et al. (2021); Krishnan (2020); Balasubramanian et al. (2020a,b); Manu et al. (2021); Karlsson (2021); Wang et al. (2021b); Miao (2021); Bachas and Papadopoulos (2021); May and Wakeham (2021); Kawabata et al. (2021a); Bhattacharya et al. (2021b); Anderson et al. (2020); Miyata and Ugajin (2021a); Kim and Nam (2021); Hollowood et al. (2021b); Wang et al. (2021a); Aalsma and Sybesma (2021); Ghosh and Krishnan (2021); Neuenfeld (2021b); Geng et al. (2021e); Balasubramanian et al. (2021b); Uhlemann (2021b); Neuenfeld (2021a); Kawabata et al. (2021b); Chu et al. (2021); Kruthoff et al. (2021); Akal et al. (2021b); Kumar Basak et al. (2021); Lu and Lin (2021); Omiya and Wei (2021); Ahn et al. (2021); Balasubramanian et al. (2021a); Li et al. (2021); Kames-King et al. (2021); Sun (2021); Hollowood et al. (2021c); Miyaji (2021); Bhattacharya et al. (2021a); Goswami et al. (2021); Chu and Miao (2021); Aref'eva and Volovich (2021); Shaghoulian (2021); García-García and Godet (2021); Buoninfante et al. (2021); Yu and Ge (2021); Nam (2021); He et al. (2021); Langhoff et al. (2021); Ageev (2021); Pedraza et al. (2021); Iizuka et al. (2021); Miyata and Ugajin (2021b); Gaberdiel et al. (2021); Uhlemann (2021a); Collier et al. (2021); Hollowood et al. (2021a); Emparan et al. (2021); Omidi (2021); Bhattacharya et al. (2021c); Karch et al. (2023c), and reviewed in Almheiri et al. (2020c); Raju (2022a,b); Liu and Sonner (2020); Nomura (2021); Kibe et al. (2021)

It is clear that whenever these island surfaces exist and depending on their relative area to the HM surface, we can either have constant for all times or initially rising Page curves. However, when such surfaces don't exist this story breaks and we lose any hope of describing evolution as unitary. This is for example the case of



empty  $AdS_{d+1}$  where the existence of such surfaces depends on a single parameter called the *critical angle*.

This problem with unitarity is resolved if we include a black string in the bulk. As explained above, in this case we found that an island surface always exists Geng et al. (2022d). This existence is independent of the two parameters in the theory, which are the anchoring point in the CFT, namely  $\Gamma$  and the brane angle  $\theta_b$ . Therefore, one could classify the two types of entropy curves (i.e. constant-for-all-times or initially rising Page curve) in a phase diagram by studying the area difference between the island surface and the HM surface. This phase diagram is driven by the Page angle, which is defined as the angle for which the island surface and the HM surface have the same area. The critical angle still appears in this picture, since below the critical angle, island surfaces exist only above some anchoring point in the brane, called the critical anchor. Similar entropy phase structures have also been studied analytically in lower dimensions Geng et al. (2022g, 2021f).

Following these results, one may wonder if they are still true in the case where one includes dynamical gravity on the brane. In lower dimensions this was achieved by including JT gravity on the brane Almheiri et al. (2020e). In higher dimensions we can mimic this by including an Einstein-Hilbert term in the brane action similar to Dvali-Gabadadze-Porrati (DGP) gravity in an  $AdS$  background; this was already done for empty  $AdS$  Chen et al. (2020b), for topological black holes Chen et al. (2020a) and for the wedge black string model Miao (2023c, 2022b); Liu et al. (2022). This opens a new parameter that our theory depends on, extending our previous result. This begs the question of whether the parameter space in this family of theories is limited. These limitations may arise in the same way as we observed in empty  $AdS$ , where the lack of RT-surfaces meant that entropy couldn't possibly evolve unitarily for some brane angles. Therefore, our first task will be to classify the structure of RT surfaces and compute their areas. In doing this we will not only address the unitarity problem, but also check when we have a Page curve.

In particular, Miao (2023c, 2022b) assert that for some part of parameter space, the wedge holographic model does reproduce the Page curve, while maintaining the massless graviton. Although we do not study this model here, we refer the reader to future work Geng et al. (2023a).

Before we summarise these results let us observe that in all our previous studies Geng et al. (2022d,e, 2021b,c), the natural choice of boundary conditions on the brane were Neumann boundary conditions. Without a DGP term these conditions translate to right-angle anchoring of the RT surface on the brane. An important consequence of the boundary conditions in the two-brane scenario is that the only surface which satisfied them on both branes was the horizon. This leads to the trivial result that the entropy between the defect and its thermofield double does not evolve with time (for empty  $AdS$ ). Although we do not study the two-gravitating-brane picture in this note, we observe that relaxing the Neumann boundary conditions can enable new RT surfaces for this case which were not previously studied. In this former paper Geng et al. (2021b), as well as in the case of a non-gravitating bath, time-dependence came from those surfaces anchoring in the non-gravitating region, which is the case we are going to study in this note.

**Overview of the Paper.** In section 7.2 we present the setup of the system studied here. We remind the reader about double holography and the one-brane black string model. We then proceed to introduce dynamical gravity on the brane at the action level and derive the corresponding equations of motion to find the RT surfaces in section 7.3. In section 7.4 we compute and present the corresponding areas and the phase structure. The note concludes in section 7.5 with some remarks.

## 7.2 Double Holography and Dynamical Gravity

In this section we give a brief overview of the holographic model and the black string setup we will use in this paper, depicted in figure 7.1. Since the setup is mostly

the same as we have been using in the past few papers Geng et al. (2021b, 2022d), we refer the unfamiliar reader to those papers. The new ingredients in this study will be covered in the latter parts of this section, so the familiar reader may want to skip to that part.

### 7.2.1 Doubly holographic model and the black string

The main setup we want to study in this note is the KR braneworld model. This model has a good description in terms of three equivalent pictures Karch and Randall (2001b,c); Almheiri et al. (2020e); Takayanagi (2011b); Fujita et al. (2011):

- **Boundary:**  $d$ -dimensional CFT with a  $(d - 1)$ -dimensional defect (BCFT $_d$  Cardy (2004); McAvity and Osborn (1995)).
- **Bulk:** Einstein gravity in an asymptotically AdS $_{d+1}$  spacetime containing an AdS $_d$  Karch-Randall (KR) brane Karch and Randall (2001b). This KR brane also has dynamical gravity.
- **Intermediate:**  $d$ -dimensional CFT coupled to “dynamical” gravity on the AdS $_d$  brane Karch and Randall (2001b), with transparent boundary conditions between the brane at infinity and a nongravitating CFT $_d$  (the *bath*) on half space.

One of the most important consequences of the transparent boundary conditions between the brane and the non-gravitating bath is that the graviton on the brane acquires a mass Porrati (2002b); Porrati and Starinets (2002); Porrati (2003). This can be seen from the boundary description, since these boundary conditions lead to the non-conservation of the stress tensor. This in turn picks up an anomaly dimension, which translates to the mass of the graviton Aharony et al. (2006)<sup>1</sup>. The

---

<sup>1</sup>It is a standard fact in *AdS/CFT* that the anomalous dimension of operators in the boundary appears as a mass term in the *AdS* equations.

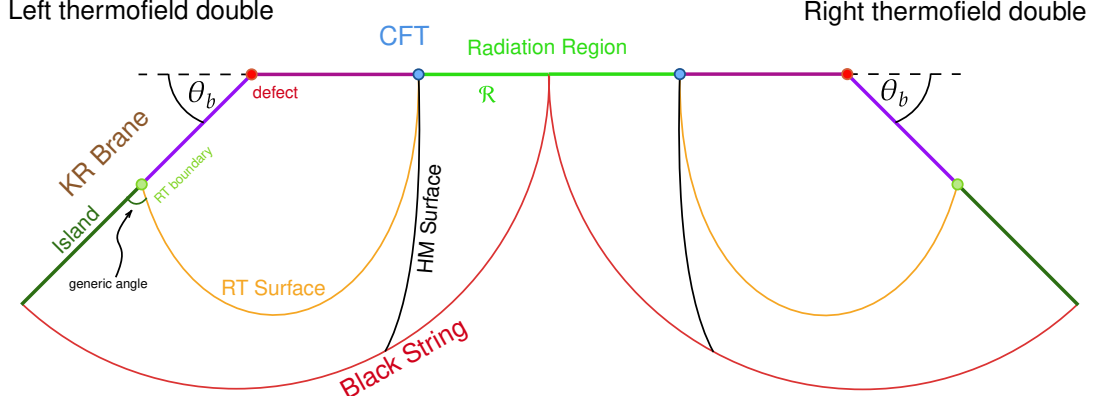


Figure 7.1: Cartoon of one copy of the KR braneworld considered in this note. The region behind the Karch-Randall brane has been excised. Two copies of this braneworld are glued along the brane. The intersection of the brane with the RT surface is not necessarily at right angles and will be determined by equation (7.9).

mass of the graviton can be related to the brane angle (i.e. it is roughly quadratic in the brane angle) for small angles Miemiec (2001b); Schwartz (2001) and to the  $AdS_{d+1}$  length scale for large angles. This means that most computations of the Page curve in  $d > 2$  at a semiclassical level have been performed in models with a massive graviton Geng and Karch (2020b). Since we are using the same model, this note and all computations and results presented here have the same feature. We remark that in Miao (2023c, 2022b) the authors claim they're able to reproduce the Page curve evolution in massless gravity.

Now that we have seen the general idea of coupling a gravitating region to a non-gravitating bath and the three equivalent descriptions that holography provides, it is time to include the black string. From the bulk perspective, we start with a  $(d+1)$ -dimensional black string environment in which we embed a  $d$ -dimensional KR brane. Setting the  $AdS_{d+1}$  radius to 1 we can describe the bulk geometry with the metric,

$$ds^2 = \frac{1}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + d\vec{x}^2 + u^2 d\mu^2 \right], \quad (7.1)$$

$$h(u) = 1 - \frac{u^{d-1}}{u_h^{d-1}},$$

where  $t \in \mathbb{R}$ ,  $u > 0$ ,  $0 < \mu < \pi$ , and  $\vec{x} \in \mathbb{R}^{d-2}$ . These coordinates slice the  $AdS_{d+1}$  space into  $AdS_d$  subspaces, one for each constant  $\mu$ . The black string induces a black hole in each slice, one of which is the KR brane, i.e. the  $\mu = \theta_b$  slice. In particular, by the  $AdS/CFT$  correspondence, this setup is a gravitating black-hole coupled to a non-gravitating black-hole background in the bath.

Typically, the black string features a thermodynamic instability, called the Gregory-Laflamme instability. In global  $AdS$  this instability is controlled by a combination of two of the parameters in the theory, the  $AdS$  radius and the horizon distance  $u_h$ . Considering large enough black strings (compared to the horizon distance) we avoid this instability. For the case at hand, by considering the Poincaré patch we automatically avoid this, i.e. we only consider large black holes. Small black holes will be considered in Karch et al. (2023c).

We are interested in finding the entanglement entropy between a subregion  $\mathcal{R}$  (the radiation region) in the bath and its complement  $\bar{\mathcal{R}}$  at a fixed time slice (which we choose to be  $t = 0$ ). We will use the usual RT prescription in which we first find all extremal surfaces  $\Sigma$  satisfying the homology constraint

$$\partial\Sigma = \partial\mathcal{R} \cup \partial\mathcal{I}. \quad (7.2)$$

The second boundary term in the homology constraint, written in terms of the island  $\mathcal{I}$ , allows for surfaces to cross the brane into the other KR braneworld.

### 7.2.2 Area functional

Now we wish to explicitly find the surfaces satisfying the constraint above. As we will see shortly, and at the level of finding the RTs, the DGP term will affect the boundary condition so we will start by not writing it explicitly. The final step will be to extremise the set of these surfaces and pick the minimal area one.

Looking at the metric, (7.1), we can exploit the translational symmetry along  $x_i$  to parameterise the potential RT surfaces as  $u(\mu)$ . Substituting this into our metric,

for a constant  $t = 0$  slice, we can write an area density functional

$$\mathcal{A} = \int_{\theta_b}^{\pi} \frac{d\mu}{(u \sin \mu)^{d-1}} \sqrt{u^2 + \frac{u'(\mu)^2}{h(u)}}, \quad (7.3)$$

We observe that  $u_h$  can be written as an overall factor of  $u_h^{2-d}$  in the integral above, meaning that it scales out of the problem. Thus, we can without loss of generality set  $u_h = 1$ . Furthermore, we remark that we will be, mostly, omitting the word “density” throughout this note

Our objective now is to extremise this functional and if the set of such extremising surfaces contains more than one element, to pick the minimal area one. This is done by solving the Euler-Lagrange equation (i.e. the equation of motion) and imposing appropriate boundary conditions. This equation can be written explicitly as

$$u'' = -(d-2)u h(u) + (d-1)u' \cot \mu \left( 1 - \frac{\tan \mu}{2} \frac{u'}{u h(u)} + \frac{u'^2}{u^2 h(u)} \right) - \left( \frac{d-5}{2} \right) \frac{u'^2}{u}. \quad (7.4)$$

Until now, this is apparently the same as we have done in the previous study, so let's see what changes now.

### 7.2.3 DGP term

The area of the RT surface to compute entanglement entropy arises from evaluating the Ricci scalar on a solution with a conical defect, required by the replica trick Lewkowycz and Maldacena (2013b). Therefore, following Chen et al. (2020b), by introducing a Ricci scalar (via an Einstein-Hilbert term) on the brane, namely induced by the bulk gravity, we can get an additional contribution to the entanglement entropy, called the DGP term. The full action on the brane will thus be,

$$S = -T \int d^d x \sqrt{-\tilde{g}} + \frac{1}{16\pi G_b} \int d^d x \sqrt{-\tilde{g}} \tilde{R} \quad (7.5)$$

where the  $\tilde{g}$  is the induced metric on the brane,  $\tilde{R}$  is the corresponding Ricci scalar,  $T$  is the brane's tension <sup>2</sup> and  $G_b$  is the Newton's constant on the brane which parameterises the strength of gravity there (i.e. the DGP term). With this, the entanglement entropy computed from the area of the RT surface is

$$S_{EE} = \frac{1}{2} \left( \frac{2}{4G} \mathcal{A} + \frac{1}{4G_b} \mathcal{A}_b \right) = \frac{1}{8G} (2\mathcal{A} + \lambda_b \mathcal{A}_b) \quad (7.6)$$

where  $\mathcal{A}$  is calculated in (7.3) and  $\lambda_b = \frac{G}{G_b}$  <sup>3 4</sup>. There is also a factor of 2, which comes from the fact that we have two copies of spacetime in our KR braneworld. The factor of  $\frac{1}{2}$  comes from the  $\mathbb{Z}_2$  orbifolding, although this is a conventional choice. The latter is the independent parameter that we are going to be free to tune and it is the main new ingredient in this note. The term  $\mathcal{A}_b$  is the area density of the point where the RT surface intersects the brane, it is thus given by

$$\mathcal{A}_b = \frac{1}{(u \sin \mu)^{d-2}} \Big|_{u=u_b, \mu=\theta_b} \quad (7.7)$$

where  $u_b$  is the  $u$ -coordinate where the brane intersects the RT surface. This is the last piece that we need in order to derive the correct boundary conditions for the problem at hand, so let's proceed to this now.

The equation of motion (7.4) will not be affected by the boundary term  $\mathcal{A}_b$ . In this way, abbreviating the equation of motion by *EOM* and extremising

$$\begin{aligned} 0 \equiv & 2\delta\mathcal{A} + \lambda_b\delta\mathcal{A}_b = - \int_{\theta_b}^{\pi} d\mu (\delta u) (EOM) \\ & + \frac{2\delta u}{(u \sin \mu)^{d-1}} \frac{u'}{h(u) \sqrt{u^2 + \frac{(u')^2}{h(u)}}} \Big|_{\theta_b}^{\pi} + \frac{(2-d)\lambda_b\delta u}{u^{(d-1)}(\sin \mu)^{d-2}} \Big|_{u=u_b, \mu=\theta_b} \end{aligned} \quad (7.8)$$

---

<sup>2</sup>This is related to the brane angle via  $T = \frac{(d-1)}{4\pi G} [\cos \theta_b + \frac{\lambda_b}{2} (d-2) \sin^2 \theta_b]$ . The author would like to thank Hao Geng for highlighting this important point.

<sup>3</sup>We will occasionally omit the subindex and write  $\lambda$  instead of  $\lambda_b$  to simplify notation.

<sup>4</sup> $G$  is the bulk's Newton's constant.

For Neumann boundary conditions on the brane we must set the second line to zero, evaluated at the boundary on the brane, namely

$$u' = \pm(2-d) \frac{u \lambda_b h(u) \sin \mu}{\sqrt{4 - (d-2)^2 h(u) \sin^2 \mu \lambda^2}} \Big|_{u=u_b, \mu=\theta_b} \quad (7.9)$$

If we further assume  $d > 2$ , and since  $0 < \mu < \pi/2$  then the sign of  $u'$  is fixed to be the same as the sign of  $-\lambda_b$ , by requiring that it indeed solves the Neumann boundary conditions. We observe that the square root in the denominator of (7.9) means that  $\lambda_b$  can only take values in  $(-\frac{2}{(d-2) \sin \theta_b \sqrt{h(u_b)}}, \frac{2}{(d-2) \sin \theta_b \sqrt{h(u_b)}})$ , for which  $u'$  can take any value, for fixed  $\theta_b$  and  $u_b$ .

For the rest of this note we will fix  $d = 4$ <sup>5</sup>. Following the discussion in Chen et al. (2020b) we can also write the induced Newton's constant on the brane as

$$\frac{1}{G_{eff}} = \frac{1}{G}(1 + \lambda_b) \quad (7.10)$$

This means that at the special value of  $\lambda_b = -1$  gravity on the brane becomes infinitely strong and we should be careful when analysing the results. Furthermore, although we will analyse all values of  $\lambda_b$ , if  $\lambda_b \leq -1$ , the effective Newton's constant  $G_{eff}$  is negative and so the theory is non-physical.

As we can see from the equation of motion, solving it analytically is quite difficult so we are going to follow the same approach as in Geng et al. (2021b) and solve it numerically, using a shooting method.

We will now proceed to present some results and therefore (in addition to fixing  $d = 4$ ) we fix  $u_h = 1$ . The latter will not affect any result, since this is an overall scale in the functional which doesn't affect the equations of motion.

---

<sup>5</sup>This means, when comparing with Chen et al. (2020b), that the induced Newton's constant  $G_{RS} = (d-2)G/2$  is the same as the bulk Newton's constant  $G$  and hence their  $\lambda_b$  is the same as our  $\lambda_b$ . This is not true in general dimensions (or general radius curvature).



### 7.2.3.1 The $\lambda$ -parameter

Geometrically we can think of the boundary condition given by  $u'$ , as the angle at which the RT surface intersects the brane. For instance, when  $u'(\theta_b) = 0$  we recover our old right-angle condition, for example if  $\lambda_b = 0$ . However for other values of  $\lambda_b$  the situation is more complicated since, by looking at (7.9), we also have to consider  $u_b$  and  $h(u_b)$ .

For example, if we fix the shooting point to  $u_b = 0$ , then  $u' = 0$  independently of the value of  $\lambda_b$ . We observe that this will not always be possible because the critical anchor  $u_{crit}$  also plays a role in this story. We remind the reader that the *critical anchor* Geng et al. (2021b) is the lowest point on the brane from which it is not possible to reach the bath, by shooting from it. In other words, RT candidates are defined in the region of the brane  $(u_{crit}, u_h)$ , which receives the name of *atoll*. For  $\lambda_b = 0$  the critical anchor is zero for a brane angle  $\theta_b \geq \theta_c$ , where  $\theta_c$  is the critical angle. For general  $\lambda_b$  the critical anchor is non-zero even above the critical angle and hence in general we are not always able to shoot from  $u_b = 0$ .

On the other hand if we fix  $u_b = u_h$  then equation (7.9) implies that  $u'(\theta_b) = 0$ , for any value of  $\lambda_b$ . This means that if we shoot to the bath from the horizon, we will stay at the horizon and we will always reach the bath. If the surface penetrates the horizon and does not reach the bath on the TFD side from which it was launched, its area will represent increasing entropy over time and will never result in a Page curve. Therefore, for the parameter combinations where this occurs, the entropy evolution in our theory will not be unitary and we will reject this theory. In other words, in addition to the reality condition we imposed a priori on  $\lambda_b$ , we will further impose that the corresponding RT surfaces reach the bath.

Doing this numerically gives us a critical anchor and an *atoll* for each angle and  $\lambda_b$ . Surprisingly the bounds imposed by this are lower than the theoretical reality bounds, as seen in figure 7.2.

This figure shows how for fixed values of  $\lambda_b$  we get a critical anchor and a

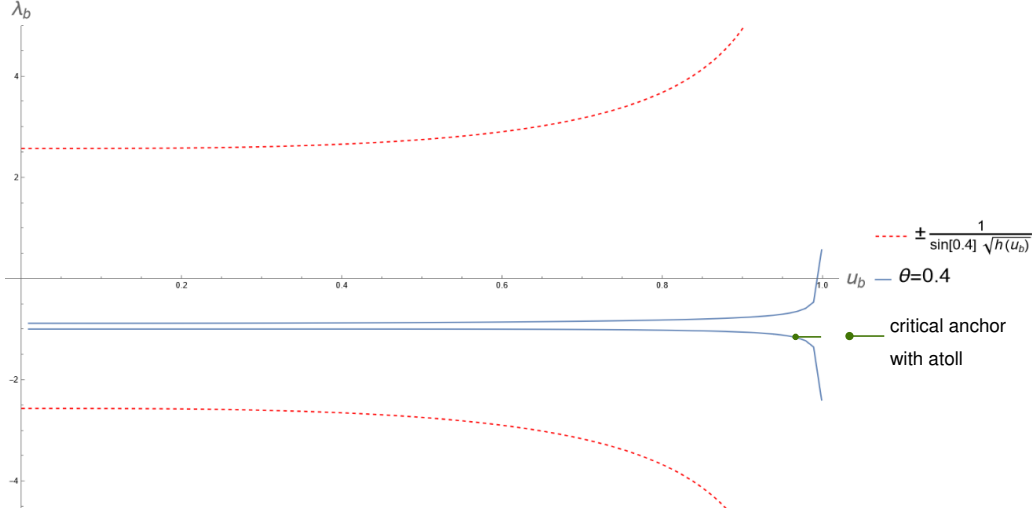


Figure 7.2: Bound on the parameters  $\lambda_b$  and  $u_b$  for which we can reach the bath. The region inside the red dashed curves shows the theoretical bound we derived above. The region inside the blue solid curves is the numerical region we obtained from solving the differential equation. For one  $\lambda_b$  a critical anchor is marked by a green dot and the shooting points making up the atoll are marked with a green line.

corresponding atoll which changes as we change  $\lambda_b$ . This generalises our old story of critical anchors, which is now both  $\lambda_b$  and  $\theta_b$  dependent (see figure 7.3). Furthermore, we observe that there is a range of  $\lambda_b$  values, for which shooting anywhere on the brane, we can reach the bath. In other words, there is a range where the critical anchor shrinks to zero. As an example, for the case shown in figure 7.2, this is  $\lambda_b \in (-0.884, -0.999)$

By performing the computation for multiple values of the brane angle, we can see how the range of values gets wider as the  $\theta_b$  increases, shown in figure 7.3. We also notice that as the brane angle increases, the lower bound curves decrease, while the upper bound curves increase. For small values of the anchoring point  $u$ , when  $|\lambda_b| < 1$  this increase (or decrease) is not substantial and so all curves accumulate near  $\lambda_b = \pm 1$ . These are somewhat special values, at  $\lambda_b = -1$  the effective Newton's constant on the brane becomes infinite. On the other hand, at  $\lambda_b = 1$ , the value

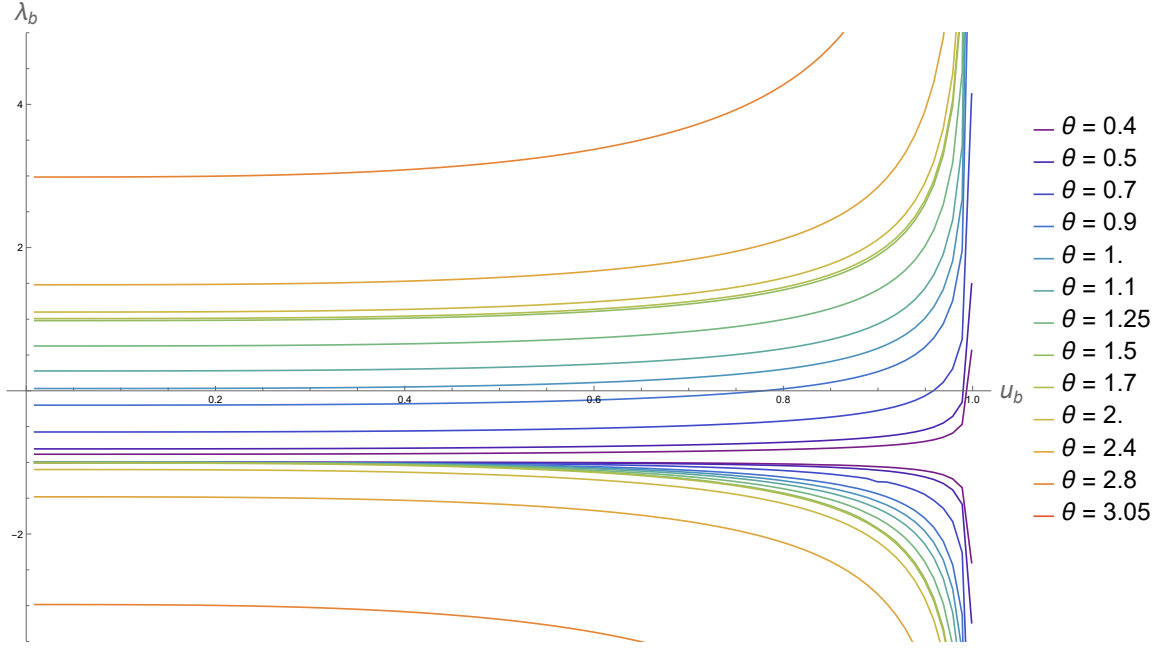


Figure 7.3: Bound on the parameters  $\lambda_b$  and  $u_b$  for which we can reach the bath for several brane angles  $\theta_b$ . The region inside curves of the same color represent the values of  $\lambda_b$  and  $u_b$  for which we can reach the bath. This region gets wider as we increase  $\theta_b$ .

of the brane and the bulk Newton's constants match numerically,  $G_b = G$ <sup>6</sup>. In particular, for the angles considered here, we will see that  $\lambda_b = -1$  is also special when we study the RT structure. For  $\theta_b = \theta_c \approx 0.98687$  the lower bound curve of the range crosses entirely the  $\lambda_b = 0$  axis. This is consistent with previous results in Geng et al. (2021b) since for angles above the critical angle the atoll should cover the whole brane. Additionally, for angles below the critical angle, the critical anchor monotonically decreases as the angle increases for any fixed  $\lambda_b \geq -0.884$  and in particular for  $\lambda_b = 0$ .

---

<sup>6</sup>We remark that these do not have the same dimensions and hence this equality is just a numerical one.

## 7.3 RT Structure

In this section we are going to explore the different types of Ryu-Takayanagi (RT) surfaces we can have for a fixed set of parameters, that is for a fixed anchoring point in the bath  $\Gamma$ , a fixed value for  $\lambda_b$  and a fixed brane angle  $\theta_b$ . We begin by describing the Hartman-Maldacena (HM) surface, which penetrates the Einstein-Rosen bridge and extends into the thermofield-double side. The other surfaces are going to be different versions of island surfaces, which anchor on the brane and cross into the second copy of spacetime in our KR braneworld.

### 7.3.1 HM surface

To study the Hartman-Maldacena surface we are going to change the parametrisation of our surface, such that  $\mu = \mu(u)$ . In this way the area functional becomes:

$$\mathcal{A} = \int \frac{du}{(u \sin \mu)^3} \sqrt{\frac{1}{h(u)} + u^2 \mu'(u)^2} . \quad (7.11)$$

where, since we know the surface doesn't intersect the brane, the DGP term doesn't affect the solutions to the Euler-Lagrange equations of this functional. To determine the boundary conditions at the horizon we can, in general, consider the Hartman-Maldacena surface anchoring at  $\Gamma$  on one side of the thermofield double and  $\tilde{\Gamma}$  at the other side. For simplicity and following Geng et al. (2022d), we will choose the symmetric case  $\Gamma = \tilde{\Gamma}$ . This imposes Dirichlet boundary conditions on the bath as with any other RT surface.

The other boundary is at the horizon, so we need to determine what the conditions are here. In order to do so we are going to assume this surface is smooth across the Einstein-Rosen bridge, in other words, and since we assumed the HM anchors at the same distance from the defect on both sides of the TFD, that it has no “kinks”. In adapted (tortoise-like) coordinates,  $dr = \frac{du}{u\sqrt{h}}$ , this can be mathematically written as:  $\mu'(r = 0) = 0$

Following Geng et al. (2022d), we can either expand the solution to the equation of motion in this coordinates, or translate to our old parametrisation  $\mu(u)$ . If we do the latter, we note that the expansion will contain half integer powers of  $(u_h - u)$ <sup>7</sup>. After doing this, we uniquely fix the solution by choosing the  $\Gamma$  and the angle at which the HM surface crosses the horizon  $\theta_{HM}$ . This is equivalent (using our power expansion) as choosing the boundary conditions:

$$\mu(u_h \Gamma) = \pi, \quad \mu(u_h) = \theta_{HM}, \quad \mu'(u_h) = \frac{2 \cot \theta_{HM}}{u_h}, \quad (7.12)$$

### 7.3.2 Island surfaces

In this section we are going to analyse the RT surfaces giving rise to islands, i.e. surfaces which anchor on the bath and the brane. In the doubly holographic models, islands are generally considered to be the regions on the brane going into the horizon, depicted in figure 7.1. In previous work we found that, when there was no DGP term, the existence of RT surfaces was driven by a specific angle value, which we called the critical angle  $\theta_c$ . For angles  $\theta < \theta_c$ , there is a specific value of anchoring point, the critical anchor, on the brane below which we can never reach the bath Geng et al. (2021b).

Something similar will occur in our setup; for a fixed angle and a fixed DGP coupling  $\lambda_b$ , there is a finite range of anchoring points from which we can reach the bath. In other words, we generalise the critical anchor slightly, where now this depends on these two parameters. We already noted this in figure 7.3, where we gave the minimum range of values for which we can reach the bath by shooting from anywhere in the brane. As we increase or decrease  $\lambda_b$ , we can reach the whole bath only by shooting from a section of the brane, namely the atoll. We can extract the critical anchor information and plot them as a function of  $\theta_b$  for each  $\lambda_b$ , like we show in figure 7.4 for  $\lambda_b > -1$  and figure 7.5 for  $\lambda_b < -1$ .

---

<sup>7</sup>This is because  $r \sim \sqrt{u_h - u}$

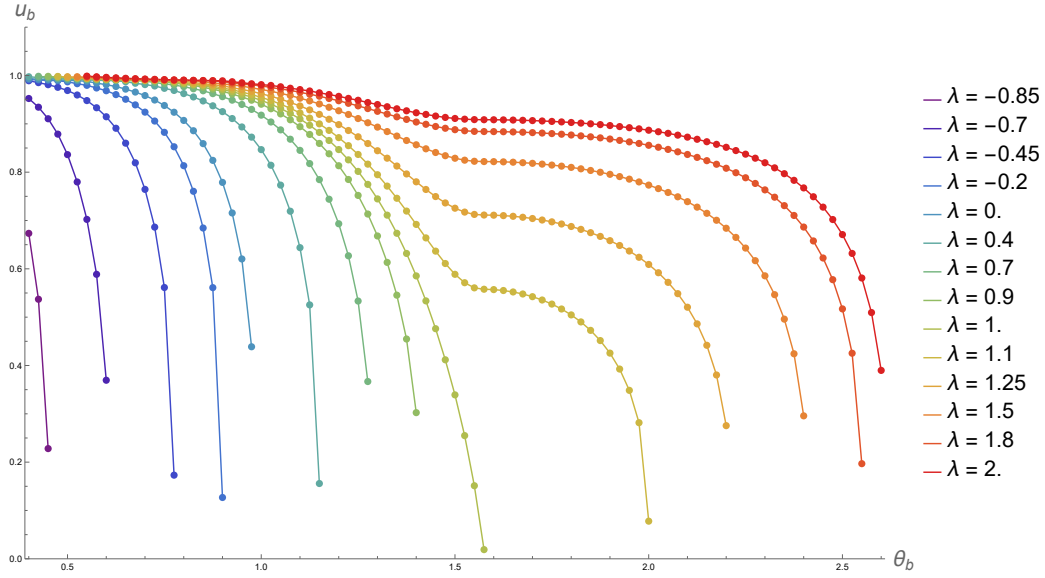


Figure 7.4: Critical anchors as a function of brane angle  $\theta_b$  for several values of  $\lambda_b > -1$ .

As we can see in figure 7.4 the critical anchors are monotonically decreasing with increasing brane angle or decreasing  $\lambda_b$ . There are two main behaviours we observe in this figure. When  $1 < \lambda_b < 2$  the curves don't decrease with decreasing slope, but show a small bump. The  $\lambda_b = 1$  value is somewhat significant because it is where the Newton's constants on the brane and the bulk numerically match.

When  $-1 < \lambda_b < 1$ , we see that the critical anchors, all follow similar curves to our previous results, i.e.  $\lambda_b = 0$  Geng et al. (2021b). In fact, if we think of the critical angle as the angle at which the critical anchor shrinks to the defect  $u_b = 0$ , then by following the curves in figure 7.4 to their intersection with the  $\theta_b$ -axis, we can also define a generalised critical angle. As we can see this decreases with increasing  $\lambda_b$ . Of course, one may think of the critical angle as determined in empty  $AdS$  and hence there is no possible generalisation in this case. In the next section we are going to take this latter interpretation.

When  $\lambda_b \leq -1$ , the behaviour of the critical anchors in relation to  $\lambda_b$  is reversed. As seen in figure 7.5, decreasing  $\lambda_b$  increases the critical anchor, while the

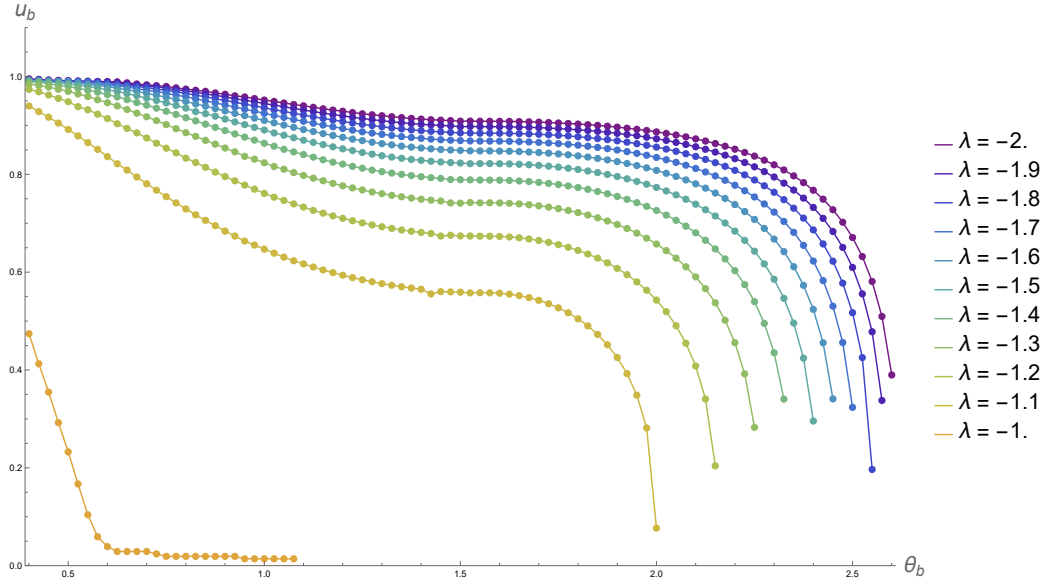


Figure 7.5: Critical anchors as a function of brane angle  $\theta_b$  for several values of  $\lambda_b \leq -1$ .

behaviour in relation to the brane angle remains unchanged. In this case, we also notice that the critical anchors seem to have a similar behaviour as when  $1 < \lambda_b < 2$  with a slope that appears to be small (near zero) for some angles. This produces generalised critical angles above  $\pi/2$  for  $|\lambda_b| > 1$ . When  $\lambda_b = -1$ , i.e. when  $G_{eff}$  becomes infinite making the theory on the brane non-physical, the critical anchors approach  $u_b = 0$  asymptotically.

Another way of putting the above observations is that the size of the atoll increases when we increase the brane angle or when the value of  $\lambda_b$  approaches  $-1$ , from above in figure 7.4 or from below in figure 7.5. If we view the degrees of freedom on the brane as redundant with those on the bath, in the limit when the effective Newton's constant is infinite, i.e.  $G_{eff} \rightarrow \infty$  information on the brane gets delocalised. On the other hand when this ratio has a large absolute value information on the brane is localised near the horizon, i.e. the resulting size of the island is small and is located near the horizon.

Furthermore we can study in more detail how the anchoring point on the brane

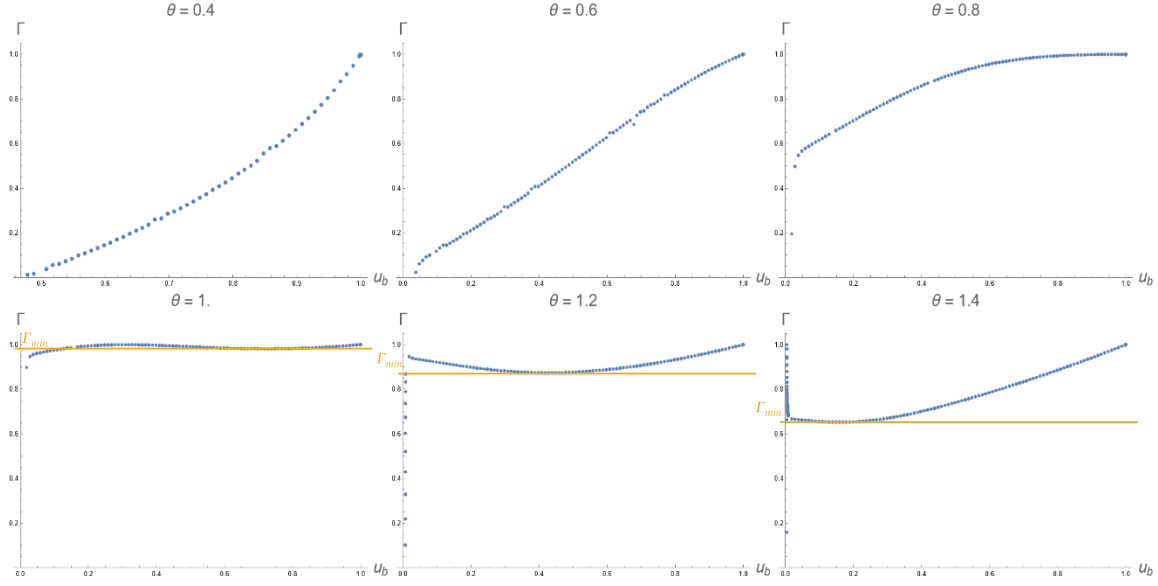


Figure 7.6: Plots of the shooting points on the brane  $u_b$  vs the anchoring points on the bath  $\Gamma$  for several values of  $\theta$  for  $\lambda = -1$ . The yellow lines show the  $\Gamma_{min}$  values.

affects the anchoring point on the bath. When doing so we discover an interesting behaviour at  $\lambda_b \leq -1$ , which we depict in figure 7.6. We distinguish two behaviours. When  $\theta_b \leq 0.825$  there is a one-to-one correspondence between shooting points on the brane, above some critical anchor, and an anchoring point  $\Gamma$  on the bath. This is similar to the situation we observed for other values of  $\lambda_b$ . However when we increase the angle past this value, the critical anchor shrinks to the defect and there is a region close to the horizon, for which a single point in the bath may have up to three corresponding RT-candidates. This region is characterised by some  $\Gamma_{min}$  above which these three surfaces exist and below which we have a unique RT candidate. This unique RT candidate is the one which has a smallest shooting point  $u_b$  and which we call *RT-1*. The next RT-candidate, ordering them by shooting point is what we call *RT-2*. Finally the RT-candidate which is shot farthest from the defect receives the name of *RT-3*. This situation is depicted in figure 7.7.

Moreover, the value of  $\Gamma_{min}$  monotonically decreases with increasing brane



angle  $\theta_b$ <sup>8</sup>. Associated to this value there is a shooting point  $u_{min}$ , such that the point  $(u_{min}, \Gamma_{min})$  is in the curves shown in figure 7.6. These  $u_{min}$  points also decrease with increasing  $\theta_b$ . The important thing for the case at hand, is that following the RT prescription, when there is a competition between RT-candidates we need to choose the minimal area one. Thus, if there is some transition in the RT surface between the RT-candidates listed, this occurs for decreasing values of  $\Gamma$ . This is in fact what will happen as we will see shortly, once we compute the areas of the RT-candidates.

When  $\lambda_b = -1.1$ , the same three competing surfaces and the same behaviour is found for angles  $\theta_b \geq 1.225$ . When  $\lambda_b = -1.2$  this behaviour is found for angles  $\theta_b \geq 1.4$  and when  $\lambda_b = -1.3$  it is found for  $\theta_b \geq 1.525$ . This poses a technical difficulty when computing areas for  $\lambda_b \leq -1$  and angles above  $\pi/2$ . Furthermore, as we will see in the next section, the areas generally decrease with  $\lambda_b$  for any given brane angle. Therefore, with the aim of characterising when the areas vanish and producing a phase diagram, we will decide to compute the areas up to a brane angle of  $\pi/2$ .

## 7.4 Numerical Results

Now that we have given a detailed account of what the possible RT-candidates are we can proceed to compute their areas. As we mentioned above we will have two competing RT surfaces: the island surface and the HM surface. The former has a constant area in time while the latter has an area which grows in time, because of the growing Einstein-Rosen bridge. In this section we compute the area differences between the island surfaces and the Hartman-Maldacena surface

$$\Delta A(t) = A_{IS} - A_{HM}(t) \tag{7.13}$$

which is automatically finite in the UV. We perform the computation at  $t = 0$ , since we know what the time-evolution of each term. We use the results from our previous

---

<sup>8</sup>see figure 7.14 in appendix 7.A.

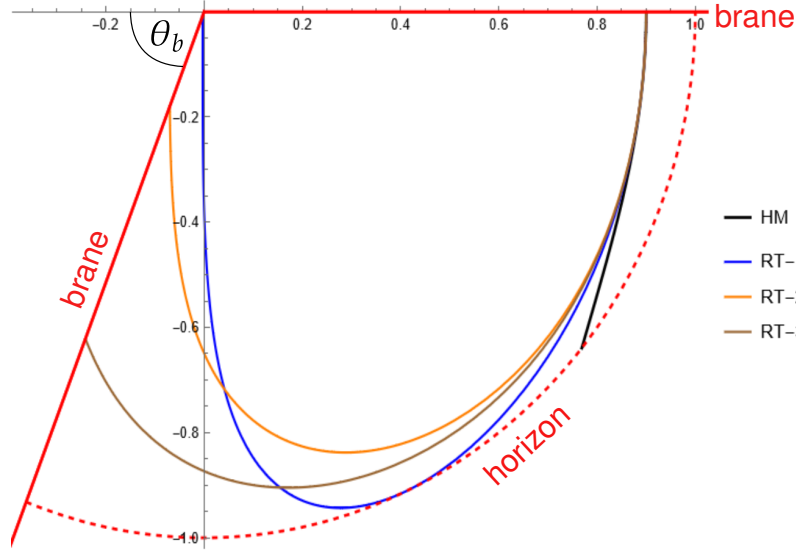


Figure 7.7: Polar plot of the three RT-candidates described above for  $\Gamma = 0.9$ . The value brane angle is  $\theta_b = 1.2$ , for which  $\Gamma_{min} = 0.873845$ .

section about the existence of RT surfaces to find the areas<sup>9</sup>. Finally, from this data we obtain the entropy phase structure for the present case.

The situation where  $\lambda \leq -1$  brings a new challenge, since for certain angles we have several competing surfaces, so we will also need to be careful here to compute the smaller area surface.

#### 7.4.1 Area difference for fixed parameters

Before proceeding we make a small philosophical distinction about the parameters. Although there is no a priori hierarchy between the three parameters and we are free to choose,  $\lambda_b$ ,  $\theta_b$  and  $\Gamma$ , we would like to think about the first two as in some sense fixed before  $\Gamma$ . What we mean is that the brane angle and the gravity on the brane are fixed in the particular universe we study. However, the location on the bath where we measure the entropy depends on the location of the observer. In

<sup>9</sup>Formally we are computing area differences given by (7.13), but in the text we will loosely refer to “areas”.

other words, an observer in a particular universe can always move further from the defect to choose another region on the bath to measure the entropy. That observer cannot change the brane angle nor the strength of the gravity on the brane unless that observer changes universes. Although we'll vary all of the parameters in our computations, we are going to do so by fixing  $\lambda_b$  and  $\theta_b$  and plotting the areas as a function of  $\Gamma$ .

Methodologically, we have decided to compute the areas for angles  $\theta \in (0.4, \pi/2)$  (in steps of  $\theta = 0.025$ )<sup>10</sup>. The parameter  $\lambda_b$  will run in the range  $(-2, 2)$ , in steps of  $\lambda_b = 0.05$ . It's worth noting that this range includes values that correspond to theories that are considered non-physical. However, for the sake of completeness, we will still consider these values in our analysis. As we noted above, the particular value of  $\lambda_b$  may limit the range of shooting points on the brane, i.e. the size of the atoll, but in terms of the position on the bath  $\Gamma$ , we will still cover the whole bath, so  $\Gamma \in (0, 1)$ .

#### 7.4.1.1 $\lambda = -1$

We now present some of the technical difficulties, in particular for the  $\lambda = -1$  case, where the situation is a bit more intricate. We have noticed above that for a fixed  $\Gamma$  we can have up to three competing RT surfaces. We have ordered these according to the shooting point on the brane: RT-1, the one closer to the defect, RT-2, the one in the middle and RT-3, the one closer to the horizon (see figure 7.7). In figure 7.6 we can see that when the angle is  $\theta \leq 0.825$ , there is a unique surface for each shooting point on the brane and hence there is no competition. Furthermore, when  $\Gamma < \Gamma_{min}$ , there is also a unique surface, so again, this is the smallest area surface. However, when we increase the angle above 0.85 two surfaces, RT-2 and

---

<sup>10</sup>We remark that the tensionless branes are generally not at  $\theta_b = \pi/2$ , given that the tension vanishes at some angle depending on  $\lambda_b$ , which can be obtained from the relation between brane tension and angle quoted above. The choice of this range of angles is based on practical reasons, namely that the region of interest for this note lies in this range, as we will see from the phase diagram.

RT-3 appear for  $\Gamma > \Gamma_{min}$ . Let us analyse each of the areas independently to explain this situation.

For RT-1, the area is positive near  $\Gamma = 0$  and diverges as  $\Gamma \rightarrow 1$ . The shooting point for RT-1 is located on the brane, close to the *AdS* boundary, and it is therefore not regulated. This might lead one to expect a positive divergence for this area, but this is compensated by subtracting a similar DGP contribution, in (7.6). When  $\Gamma = 1$  the area blows up again, because here the HM has large negative renormalised area. This behaviour was noted in Ryu and Takayanagi (2006d). This competition between the non-boundary area and the DGP term, produces a nearly constant area for every angle, for this surface.

For RT-2 and RT-3, we see a similar behaviour, reaching a smallest area at the point where we shoot directly to  $\Gamma_{min}$ , since this shooting point is closer to the defect. More precisely, RT-2 and RT-3 have monotonically increasing area for  $\Gamma > \Gamma_{min}$ . However, in this range of  $\Gamma$ s, RT-3 has slightly smaller area, only equating RT-2 precisely at  $\Gamma_{min}$ , where these surface are degenerate<sup>11</sup>. This rules out RT-2 as a competing surface. Nevertheless, these surfaces have decreasing area with increasing angle and hence, immediately after appearing they will dominate over RT-1. Both surfaces have the same divergence when  $\Gamma \rightarrow 1$ , which we observed in RT-1 and is what we would expect for any of these.

Collecting these results, we have the following behaviour in the area for a fixed angle as a function of  $\Gamma$ . Below  $\theta_b = 0.85$ , the area (determined by the area of RT-1) is slightly positive close to the defect and then increases with  $\Gamma$ . It diverges again for  $\Gamma$  close to the horizon because of the Hartman-Maldacena surface. When the angle is at or above 0.85 this behaviour is maintained at  $\Gamma < \Gamma_{min}$ , being RT-1 the dominant surface. At  $\Gamma_{min}$ , there is a transition in which RT-3 becomes dominant, with a smaller area and keeps dominating all the way up to the stage where  $\Gamma$  reaches the horizon and the area diverges. This produces a jump from a larger area RT-1,

---

<sup>11</sup>See figure 7.15 in appendix 7.A which illustrates this explanation.

to a smaller area, RT-3 which looks discontinuous. Although here we won't resolve this discontinuity, we suspect that sampling more points in the bath can show a steep but smooth transition between these surfaces. This would produce two points for which  $\Delta\mathcal{A} = 0$ . However the conservative approach we'll take here is to think of this jump as discontinuous. This is similar to the discontinuity observed in Geng et al. (2021b); Karch et al. (2023c), when the tiny islands take over. As we can see in figure 7.6, increasing the angle means that  $\Gamma_{min}$  is decreased and appears for lower shooting points  $u_b$ . In other words RT-2 and RT-3 appear "sooner" and thus this jump also occurs sooner.

#### 7.4.1.2 Area difference as a function of $\lambda_b$

Here we look closer at the numerical results, depicting only a representative subset of the areas computed. These show the main features we observed in our computations. In particular, we can plot the area difference  $\Delta\mathcal{A}$  for  $t = 0$  as a function of  $\Gamma$  for several  $\lambda_b$ -values. This produces a 2-D surface for each angle. Below, in figures 7.8 and 7.9 we just show a few of these surfaces.

As we can see in figures 7.8 and 7.9, the area increases monotonically with  $\lambda_b$  until we reach  $\lambda_b = -1$ . Here we observe that the areas near the defect,  $\Gamma = 0$ , are positive, as we described above and which will be more evident when we present these areas for fixed  $\lambda_b$ s. In figure 7.8, for  $\lambda_b$  between  $-1$  and  $0$  we see how the near defect area becomes increasingly negative as we increase the angle. This behaviour is exacerbated until we reach  $\theta_b = 1$ , for which the negative areas also appear closer to  $\lambda_b = 0$ . In figure 7.9, this trend is maintained even for positive  $\lambda_b$ s. To explain this we recall that for  $\lambda_b = 0$  we recover our results in Geng et al. (2021b). Here we observed this divergence, which was caused by infinitesimal surfaces which wrapped the defect, namely the tiny islands. For non-zero values of  $\lambda_b$ , we can now have a non-zero critical anchor, even above the critical angle. Hence we don't obtain these tiny islands as limits of any surface, namely we cannot continuously shoot to the defect. However, the critical anchor gets smaller as  $\lambda_b \rightarrow -1$  from below (see figure

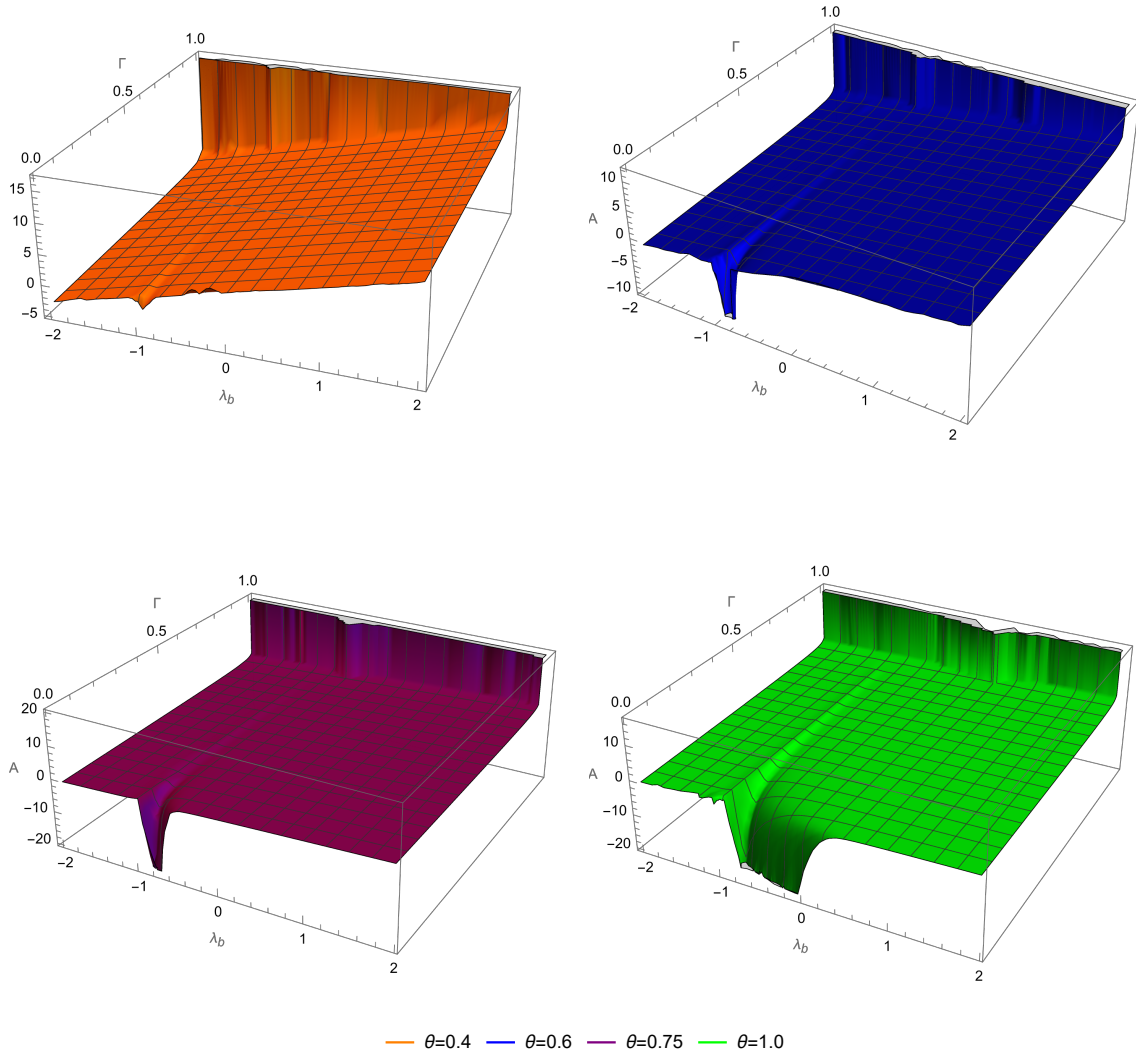


Figure 7.8: Area of the RT surface as a function of  $\Gamma$  and  $\lambda_b$  for angles between  $\theta = 0.4$  and  $\theta = 1$ .

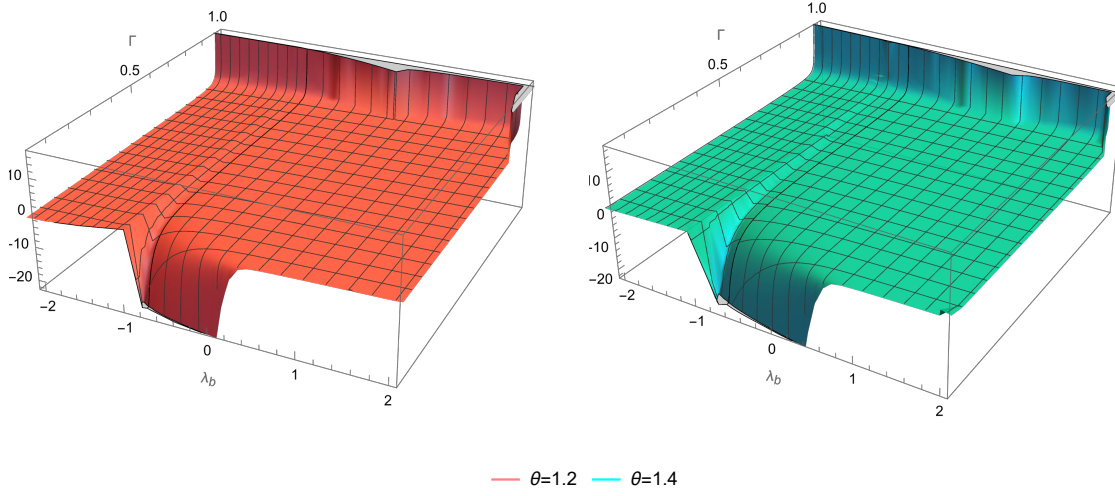


Figure 7.9: Area of the RT surface as a function of  $\Gamma$  and  $\lambda_b$  for angles  $\theta = 1.2$  and  $\theta = 1.4$ .

7.5) and so, when we have large negative  $\lambda_b$  ( $\lambda_b \leq -1$ ) the RT-surface has a finite large shooting point  $u_b$  and doesn't diverge to negative infinity. As we increase  $\lambda_b$  the critical anchor sinks into the defect and the tiny islands take over. For small negative values of  $\lambda_b$ , we are subtracting a large negative area from an already negative term (tiny islands) by the DGP contribution as given in equation (7.7). Since the range of  $\lambda_b$ s for which the critical anchor shrinks into the defect increases with angle, this produces an increasingly pronounced dip in figures 7.8 and 7.9 near  $\Gamma = 0$ . Something similar occurs for larger values of  $\lambda_b$ , where again the appearance of a critical anchor truncates the existence of tiny islands, preventing the area difference to diverge to negative infinity.

This behaviour suggests that any definition of the critical angle that depends on the existence of tiny islands, or as mentioned before, on the critical anchor sinking to the defect, will also be affected by the inclusion of the DGP term. However, our focus is on comparing our parameters to the critical angle of empty *AdS* without a DGP term, and thus we will not define the critical angle in this manner.

Furthermore, away from the  $\Gamma = 0$  defect the areas are monotonically increasing with one exception. Although it is less noticeable in the other angles, in figure 7.9, for  $\theta_b = 1.4$  and  $\lambda_b = -1$ , the area monotonically increases until it reaches the  $\Gamma_{min}$  we described before. At this point we have a discontinuity corresponding to the transition between RT-1 and RT-3. Away from the  $\lambda_b = -1$  value the area is finite and monotonically increases with  $\Gamma$  for every angle. In all cases, the area diverges to infinity as  $\Gamma \rightarrow 1$ . Above we mentioned that this divergence was caused by the fact that the HM surface has a renormalised area which diverges to  $-\infty$ . In other words, the explanation of this divergence comes from the fact that in (7.13) we are subtracting a  $-\infty$  term. Another way to see this divergence is using the fact that when  $\Gamma$  gets close to the horizon, the corresponding HM shrinks to a point. This means that it “cuts off” less of the RT surface which still picks up a divergence.

We recall that that a negative area difference  $\Delta\mathcal{A}$ , means that the island surface is smaller than the HM surface, which generally grows with time. Therefore, we can summarise these observation by: an observer sitting close to the defect for  $\lambda_b \geq -1$ , will (depending on the angle) measure a constant entropy curve. Furthermore, because of the competing surfaces at larger angles, this is also true for an observer sitting closer to the horizon. We remind the reader that exactly at  $\lambda_b = -1$ , the effective Newton’s constant on the brane,  $G_{eff}$ , becomes infinite and hence the theory is considered non-physical. As we increase  $\lambda_b$  the area differences will become positive and the observer will be able to measure a non-constant entropy curve.

#### 7.4.1.3 Area difference as a function of $\theta_b$

To get some more intuition and to show some of the things we described above, we can similarly fix the  $\lambda_b$  value and plot  $\Delta\mathcal{A}$  vs  $\Gamma$  for several angles. In this way each 2D surface corresponds to one  $\lambda_b$ -value. We can see these surfaces in figures 7.10 and 7.11

In figure 7.10 we show some of the surfaces obtained for  $\lambda_b \leq 0$ . When  $\lambda_b$  is



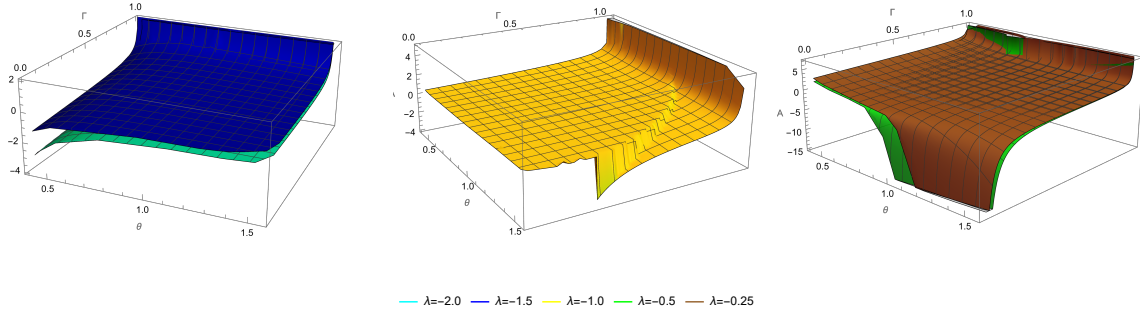


Figure 7.10: Area of the RT surface as a function of  $\Gamma$  and  $\theta_b$  for  $\lambda_b < 0$ .

sufficiently large and negative (panel on the left of the figure) we see that the area is finite independently of the angle <sup>12</sup>. This is what we observed above and it is because there is a non-zero critical anchor which, even when  $\Gamma = 0$ , keeps our surface anchored away from the defect. The area monotonically increases with the brane angle  $\theta_b$  and  $\lambda_b$ . As before, the area is divergent when  $\Gamma \rightarrow 1$ .

In the centre figure, the areas are approximately constant with brane angle, but increasing with  $\Gamma$  for most of parameter space. However for larger angles we observe the transition between RT-1 and RT-3. The latter surface has monotonically increasing area with  $\Gamma$ , which is also appreciated in this panel. Furthermore, we observe how this transition occurs at decreasing  $\Gamma_{min}$  as we increase the brane angle. We remind the reader that thus far, these theories have  $G_{eff} < 0$  and hence can be considered as non-physical.

On the right panel of figure 7.10, the areas still increase with  $\Gamma$ , but the trend with respect to  $\theta_b$  is reversed, namely now we have decreasing areas with brane angle. This behaviour with respect to the two parameters will be maintained for all  $\lambda_b > -1$  in figure 7.11. The other new ingredient is the tiny island effect which appears when  $\Gamma \rightarrow 0$  and  $\theta > \theta_{crit}$ . The dependence of the critical anchor on  $\lambda_b$  can clearly be

---

<sup>12</sup>We expect the areas to diverge as the brane angle approaches  $\theta_b \rightarrow 0$ , from it becoming the *AdS* boundary, but this is outside the range of values computed in this note

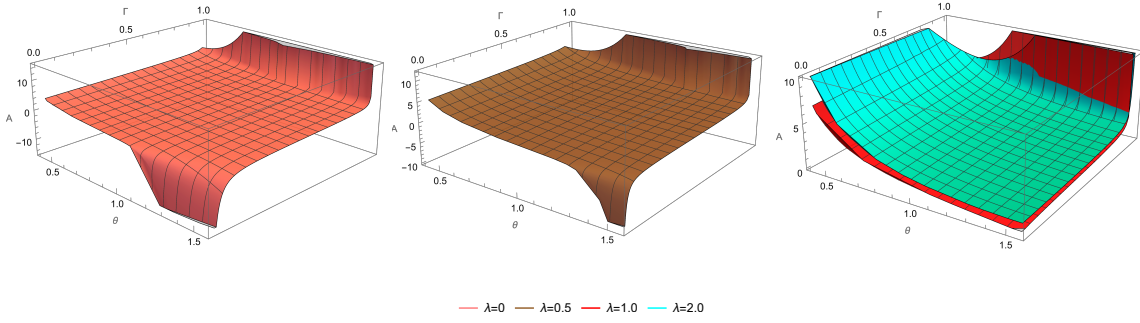


Figure 7.11: Area of the RT surface as a function of  $\Gamma$  and  $\theta_b$  for  $\lambda_b \geq 0$ .

seen by the fact that the areas diverge faster to  $-\infty$  for angles larger than  $\theta_c$  as  $\lambda_b$  approaches zero from below. The leftmost panel of figure 7.11 shows the case where the DGP coupling is turned off,  $\lambda_b = 0$  and this matches the results in Geng et al. (2021b).

Comparing these figures we see how the tiny islands are somewhat tamed by the DGP coupling. However as we can see by comparing the first and third panel of 7.10 the DGP term has a greater effect on the smaller angles, reversing the monotonic behaviour of the area as a function of brane angle. This is evident from equation (7.7) which is proportional to  $1/\sin \theta_b^2$  and since the anchors are roughly fixed in a small range, as we would expect for small angles. We will comment on the signs of the areas below.

In general, the areas for any  $\lambda_b$  and  $\theta_b$  can be positive or negative so we can ask the question about when the areas are precisely zero, this is what we proceed to do now.

#### 7.4.2 Phase structure

We have presented a detailed account for the area differences as a function of the different parameters in our theory;  $\Gamma$ ,  $\theta_b$  and  $\lambda_b$ . Now that we have observed the different trends we can analyse the entropy phase structure given in this parameter

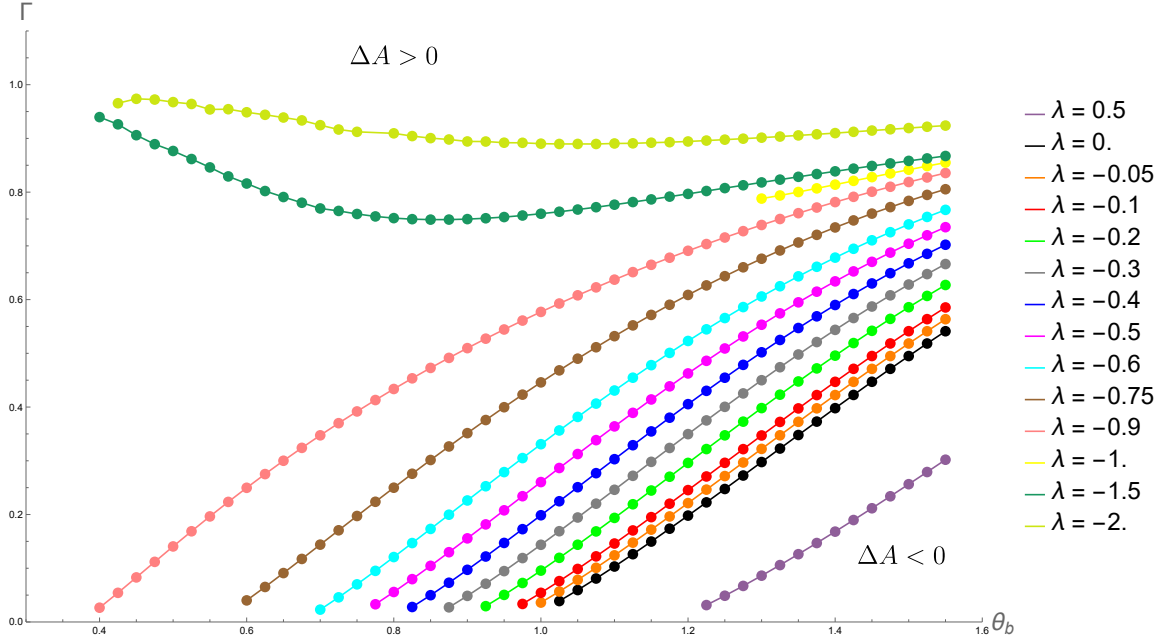


Figure 7.12:  $\Gamma$  and  $\theta_b$  for which  $\Delta\mathcal{A} = 0$  for several values of  $\lambda_b$ . The horizon is at  $\Gamma = 1$ . For the first few curves the area is positive for a small range near the horizon, independently of the angle. Our previous results are reproduced by the black curve  $\lambda_b = 0$ .

space. We recall that the sign of  $\Delta\mathcal{A}$  will determine what surface dominates at  $t = 0$ . Since the HM has a linear area evolution with time, a positive  $\Delta\mathcal{A}$  means that the entropy follows a Page curve evolution. However, when  $\Delta\mathcal{A}$  is negative the constant area island surface is the RT-surface whose area is proportional to the entropy and hence the entropy is just constant. Therefore, the question of when  $\Delta\mathcal{A}$  vanishes is a question about when we have an entropy evolution following a Page curve. We remark that both situations satisfy unitarity, since we already discarded the non-unitary theories, that is how we limited the parameter space in the first place.

Here we present the main result in this note which is the entropy phase structures diagrams. These are curves showing the values of  $\Gamma$  and  $\theta_b$  for which  $\Delta\mathcal{A} = 0$  for fixed  $\lambda_b$  values.

From the plot in figure 7.12 we observe several different behaviours. When  $\lambda_b$

is between  $-2$  and  $-1$  we observe that the curves in the phase diagram never intersect the  $\theta_b$ -axis. This means that the area never vanishes at the defect (for this range of angles), or in other words there is no proper definition of the Page angle, the angle for which the vanishing area surface has  $\Gamma = 0$ . In fact, for these few first curves, the region where  $\Delta\mathcal{A} > 0$  is a smaller region of parameter space which is also close to the horizon. Since  $\Delta\mathcal{A} < 0$  means that there is a flat entropy curve (because the island surface starts dominant at  $t = 0$ ), this means that for most of the parameter space there is no Page curve. In particular when  $\lambda_b = -2$  or  $\lambda_b = -1.5$ , we see the tendency of the curve to drop, but it is still mostly close to the horizon. Again we remark that these theories (when  $\lambda_b \leq -1$ ) are unphysical from the start. However in the limit where  $G_{eff} \rightarrow \infty$ , or in other words when  $\lambda_b \rightarrow 1$ , we see that the competing surfaces described above only allows for vanishing areas in a small region of angles. Therefore, the corresponding  $\lambda_b = -1$  curve (in yellow) is shorter than the other curves. In this case we can imagine a vertical line from the left end of this curve delimiting the  $\Delta\mathcal{A} > 0$  region from the  $\Delta\mathcal{A} < 0$  region.

When  $\lambda_b > -1$ , the curves intersect the  $\theta_b$ -axis and roughly half of parameter space will give us a page curve,  $\Delta\mathcal{A} > 0$  and half will have a constant entropy evolution. As we increase  $\lambda_b$  the constant entropy region gets smaller. In fact when  $\lambda_b > 1$ , we didn't observe negative areas and hence there is no vanishing area curves, although we suspect this will not be the case if we look at larger angles. This can also be seen in figure 7.11. This means that the corresponding curves (which are not shown in the figure), would just be flat lines sitting on the  $\theta_b$ -axis.

We can also reformulate these observations in terms of what, in previous works, we called the constant entropy belt. This is the region of the bath containing anchoring points  $\Gamma$ , for which the area is  $\Delta\mathcal{A} < 0$ . In the past, Geng et al. (2022d), this belt had a size which was monotonic with angle and this is what we observe here for  $\lambda_b > -1$ . In particular we observe this for what would reproduce our previous results  $\lambda_b = 0$ . In this way we remark that this figure, directly extends figure 10 in Geng

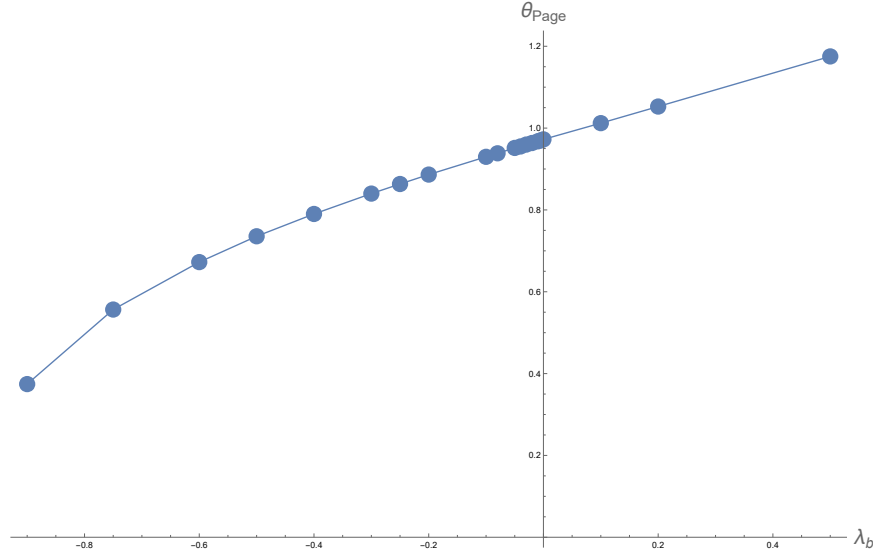


Figure 7.13: Page angles  $\theta_{Page}$  as a function of  $\lambda_b$ . At  $\lambda_b = 0$ , we obtain  $\theta_P \approx 0.972$ , matching our previous results. Furthermore the curve monotonically increases.

et al. (2022d). However, when  $\lambda_b = -2$  or  $\lambda_b = -1.5$ , this monotonicity is broken. We see the belt first decreases and then increases in both cases.

Finally, since some of the curves appear to intersect the  $\theta_b$ -axis, we can extract, by extrapolating these curves, a Page angle for each value of  $\lambda_b$ . We now proceed to do this extrapolation. This extrapolation will carry some intrinsic error, but to show the robustness of our method, we observe that for  $\lambda_b = 0$  the Page angle we obtain is 0.972. This is only marginally different from the value we had in our previous study Geng et al. (2022d) of  $\theta_P \approx 0.975$ .

### 7.4.3 Page angles

Extracting the Page angle, namely the angle for which the area of the curve shooting to the defect vanishes, for each value of  $\lambda_b$  in figure 7.12 we obtain the figure 7.13.

As we noted before, we can check our method by comparing the  $\theta_P$  we obtain now compared to what we obtained before Geng et al. (2022d) and we observe that

this is roughly the same, namely approximately 0.972.

We can see that the Page angle monotonically increases with the value of  $\lambda_b$ . Therefore, there is another interesting quantity we can extract from this plot and this is the value of the DGP coupling for which the Page angle is the critical angle. We recall that in empty  $AdS$  the Page angle is the same as the critical angle. The reason why this is not true for the black string, is because the HM has a non-zero renormalised area and hence subtracts a small number from the island surface area at the critical angle. Thus, the area at the critical angle is slightly negative and hence the value of  $\theta_b$  for which the area vanishes lies slightly below  $\theta_c$ . In the case at hand, by fine tuning the strength of the gravity in the brane, namely the DGP coupling parameter,  $\lambda_b$ , we can lift the effect of the black string and make the Page angle match the critical angle again. A similar thing has been done in Karch et al. (2023c), where the fine tuned parameter is the black hole size. This happens for approximately  $\lambda_b \approx 0.0364$ .

## 7.5 Conclusions

We have extended the study in Geng et al. (2022d) analysing the subregion entropy for a doubly holographic black string model. The system is dual to a  $BCFT_d$  with a black-hole background and to a non-gravitating bath coupled to an  $AdS_d$  brane in which dynamical gravity can be turned on. This dynamical gravity is parameterised by a coupling  $\lambda_b$  which measures the strength of gravity on the brane. In this way we are testing the parameter space of the black string model with respect to the anchoring point on the bath, the  $AdS_d$  brane angle and the strength of the DGP term,  $(\Gamma, \theta_b, \lambda_b)$ .

In empty  $AdS$  we previously observed that island surfaces did not exist below the critical angle. This issue was resolved at finite temperature, namely for the  $\lambda_b = 0$  system giving us a rich phase structure. However this phase structure could be modified by the inclusion of dynamical gravity. Therefore our aim here was to

limit the parameter space for well-behaved theories supporting islands and to study the subregion entropy phases when the strength of gravity on the brane was modified. In particular, throughout this note we focused on unitary evolution of entropy and the existence of a Page curve.

In order to do this, we have first proceeded to study the RT structure of the system, namely the existence and number of competing RT-candidates. For most of the parameter space there was a single island RT-candidate for a shooting point on the brane. This shooting point was limited to only a fraction of the brane, characterising the size of the island. This extends our critical anchor story, since now, changing the  $\lambda_b$  parameter allows for non-zero critical anchors even above the critical angle. For the particular case of  $\lambda_b = -1$  we observed 3 competing RT-candidates and steep phase transition between two of these. This phase transition was captured by a special value of anchoring point on the bath  $\Gamma_{min}$ . All of this shows a richer RT structure which we needed to take into account when computing their areas.

With this in mind we computed the areas and presented them as 2D surfaces. Some of the behaviour they exhibited matched our previous work but some of it was new and needed to be explained. In particular, as mentioned, we highlight the case of  $\lambda = -1$  and how the tiny island effects were lifted by large absolute value strength of DGP coupling. All of this is summarised in figure 7.12, which constitutes the main result in this note. Here we distinguish the parts of parameter space which give us a Page curve from those that give us an eternally constant entropy. In any case it seems both situations present a unitary evolution and hence means that the theories in the parameter space are stable against unitarity tests.

We can use this information to obtain the value of  $\lambda_b$  for which the critical angle matches the Page angle, which we obtained to be  $\lambda_b \approx 0.0364$ . In other words, one can fine tune the strength of gravity on the brane so that the effect of the black string on the renormalised area of HM is lifted and we obtain what we had in the case of empty *AdS*. The difference is that now, unlike in empty *AdS* we do have island

surfaces below the critical angle.

In recent work it was discussed that the DGP terms may solve the issue of massive gravity in island models Miao (2023c, 2022b). In this note we have seen how the parameter space is confined in order for islands to exist. Something similar might be expected for wedge holographic models where we can also impose other constraints. This is going to be analysed in more detail in future work Geng et al. (2023a).

This study has focused on the entropy phase structure and RT surfaces for angles  $\theta_b < \pi/2$ . However, the relationship between brane tension and angle indicates that for larger  $\lambda_b$ , there may be a range of subcritical brane angles that are still physical above  $\pi/2$ . It would be interesting to explore this further for larger angles and values of DGP coupling, and determine how the competition between three RT surfaces is resolved and if any phase transitions occur.

## Appendix 7.A: Supporting Plots

Here we show some of the plots which were used in part of the computations in the main text. They illustrate some of the descriptions and claims above.

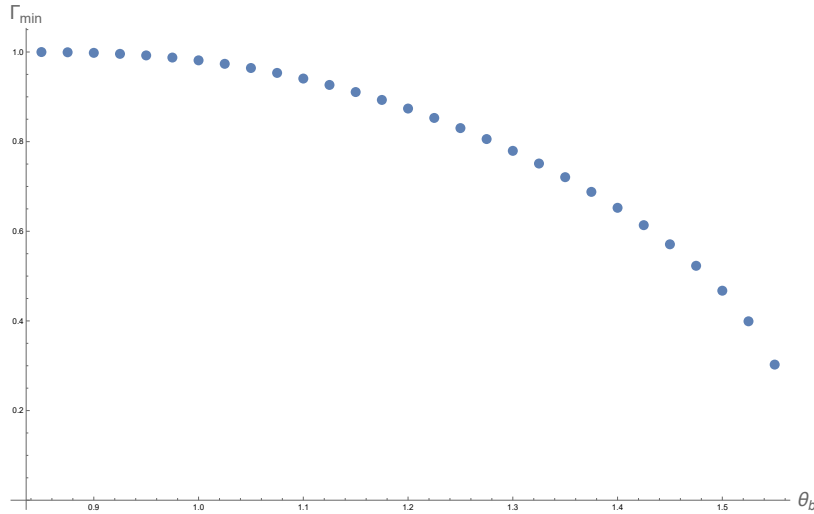


Figure 7.14: Plot of  $\Gamma_{\min}$  for  $\lambda_b = -1$ .



Figure 7.14 shows what we claimed for  $\lambda_b = -1$ . This is that  $\Gamma_{min}$  decreases as we increase the brane angle.

Figure 7.15 shows the areas for two of the competing surfaces when  $\lambda_b = -1$  and  $\theta_b = 1.35$ . The surfaces RT-2 and RT-3 become degenerate at  $\Gamma_{min}$ , while the surfaces RT-1 and RT-2 become degenerate at  $\Gamma = 1$ . Above this value we have three competing RT surfaces: RT-1, RT-2 and RT-3. We can see how RT-2 and RT-3 have increasing area as  $\Gamma \rightarrow 1$ , but overall, RT-3 has a slightly smaller area than RT-2

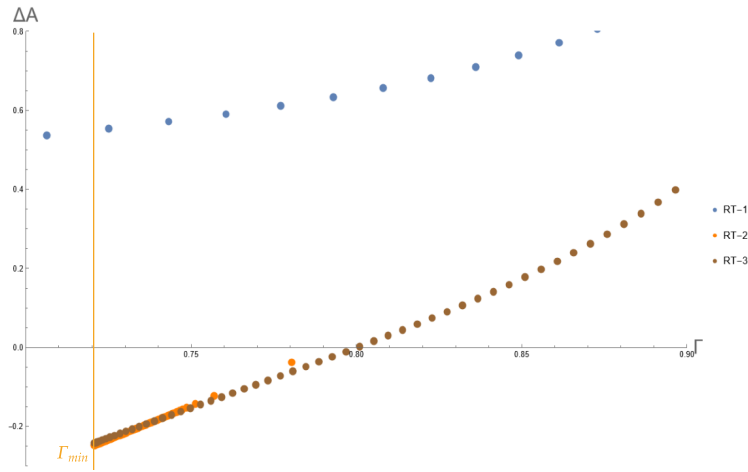


Figure 7.15: Areas for RT-1, RT-2 and RT-3 for  $\lambda_b = -1$  and  $\theta_b = 1.35$ . The yellow line marks the position of  $\Gamma_{min}$ .

# Chapter 8: Constraining braneworlds with entanglement entropy

## Abstract

We propose swampland criteria for braneworlds viewed as effective field theories of defects coupled to semiclassical gravity. We do this by exploiting their holographic interpretation. We focus on general features of entanglement entropies and their holographic calculations. Entropies have to be positive. Furthermore, causality imposes certain constraints on the surfaces that are used holographically to compute them, most notably a property known as causal wedge inclusion. As a test case, we explicitly constrain the Dvali–Gabadadze–Porrati term as a second-order-in-derivatives correction to the Randall–Sundrum action. We conclude by discussing the implications of these criteria for the question on whether entanglement islands in theories with massless gravitons are possible in Karch–Randall braneworlds.

## 8.1 Introduction

Braneworld models of the universe as being embedded in a higher-dimensional spacetime have led to a plethora of insights into particle physics and quantum gravity

---

<sup>0</sup>This chapter is based on published work in collaboration with Geng et al. (2023c). All authors contributed to the final version of the manuscript. The theoretical framework and equations were derived in my earlier solo work, forming the foundation of this project. Additionally, I participated in discussions that helped shape key ideas, particularly in collaboration with Sanjit Shashi. I also performed numerical computations (together with Merna Youssef) to support our analysis.

Agrawal et al. (2022). The original scenario of Randall and Sundrum Randall and Sundrum (1999c) considers  $(d+1)$ -dimensional<sup>1</sup> “bulk” Einstein gravity with Newton constant  $G_{d+1}$ , cosmological constant  $\Lambda = -\frac{d(d-1)}{2L^2}$  (with  $L$  a length scale), and two  $d$ -dimensional branes, both with action proportional to their worldvolumes. The Randall-Sundrum (RS) brane actions take the schematic form

$$I_{\text{RS}} = -T \int_{\text{brane}} \sqrt{-\tilde{g}}. \quad (8.1)$$

$T$  is a *tension* parameter for the brane with induced metric  $\tilde{g}$ . Furthermore, one may instead consider only one brane Randall and Sundrum (1999d), which results in an infinite anti-de Sitter (AdS) bulk. While the action (8.1) does not directly arise in string theory, it can be thought of as a low-energy effective action for any brane-like object.<sup>2</sup> For a brane whose only degree of freedom is its geometric embedding, the action (8.1) is the only zero-derivative term consistent with symmetries, in particular reparameterization invariance.

Among its many phenomenological aspects (see for example Arkani-Hamed et al. (2000); Goldberger and Wise (1999); Kim and Kim (2000); Csaki et al. (2000); Arkani-Hamed et al. (2001a)), the braneworld program provides a framework for studying quantum features of gravity on the branes Gubser (2001); Verlinde (2000). Furthermore, if the branes themselves have AdS geometry—in which case we call them Karch–Randall (KR) branes Karch and Randall (2001b,c)—then we can apply the AdS/CFT correspondence Maldacena (1998b) to establish a “*doubly holographic*” principle under which the braneworld has an additional third description as a non-gravitating, strongly-coupled conformal field theory. Thus, we can apply the usual tools of holography relating the classical bulk to the CFT in order to study semiclassical gravity on the branes indirectly while circumventing several, more complicated

---

<sup>1</sup>Technically, Randall and Sundrum (1999c) takes  $d = 4$  to reflect our universe, but we keep  $d$  general for now, only requiring  $d > 2$ .

<sup>2</sup>By “brane-like object,” we mean any  $(d+1)$ -dimensional source of energy density in the bulk whose size in one spatial direction is much smaller than the AdS scale  $L$  and in other directions is of  $O(L)$ . For such an object, we can use the low-energy effective action in its derivative expansion to describe its dynamics.

calculations. As discussed below, one recent application of double holography has been to tackle the black-hole information problem beyond 2d gravity (as in Almheiri et al. (2019d)).

We can go beyond the RS term. In principle, the branes are coarse descriptions of warped compactifications in string theory, since, as mentioned above, the RS Lagrangian (8.1) is only the leading-order possible term in a derivative expansion Georgi et al. (2001); Sundrum (2005).<sup>3</sup> The action could have higher order corrections that go beyond the RS term to include higher-derivative interactions. One such example is the Dvali-Gabadadze-Porrati (DGP) Dvali et al. (2000c); Chen et al. (2020b) term

$$I_{\text{DGP}} = \frac{1}{16\pi G_{\text{b}}} \int_{\text{brane}} \sqrt{-\tilde{g}} \tilde{R}, \quad (8.2)$$

where  $\tilde{R}$  is the Ricci scalar computed from  $\tilde{g}$  and  $G_{\text{b}}$  is a new coupling. This is one of several two-derivative terms one can include as a correction to the RS action, following the standard logic of effective field theory (EFT) Weinberg (1979)—we expect “effective terms” in the Lagrangian that are higher order in the derivatives and come from the (assumed) UV completion. Furthermore, features of the UV completion in turn put constraints on what the coupling coefficients of these terms may be. Any choice of couplings in the EFT inconsistent with the existence of some UV completion describes a theory in the *swampland* Vafa (2005); Arkani-Hamed et al. (2007); Ooguri and Vafa (2007).

Holography is intrinsic to the underlying string (or UV-complete) theory. Nonetheless, it is often assumed to persist to semiclassical regimes of quantum gravity Aharony et al. (2000). However, if this is true then we may use the tools of holography to put swampland constraints on EFTs of gravity (cf. Akers et al. (2016);

---

<sup>3</sup>We should be clear that not all holographic boundary conditions are described by brane-like objects, as discussed by Reeves et al. (2021); Belin et al. (2021). For example, the Janus solutions to type IIB supergravity Bak et al. (2003); Clark et al. (2005); Bak et al. (2007); Bachas et al. (2023); Baig and Shashi (2023) describe a class of smooth geometries dual to conformal defect systems. These situations are simply not described by what we referred to as a brane-like object above and so are not a part of our analysis.

Geng et al. (2019); Cáceres et al. (2019); Folkestad (2022); Lee et al. (2022)). This observation has been used to constrain two-dimensional branes equipped with RS and Jackiw–Teitelboim terms Lee et al. (2022). In the same spirit, we apply this idea to the analogous brane-localized couplings [(8.1) and (8.2)] in higher ( $d > 2$ ) dimensions by specifically focusing on aspects of “entanglement structures” (particularly Ryu–Takayanagi surfaces Ryu and Takayanagi (2006c); Hubeny et al. (2007b)) constructed in the bulk spacetime and from some subregion of the dual CFT. Our goal is to put physical limitations on doubly holographic braneworlds using two criteria proposed in Geng (2023) which are based on the following considerations.

Entanglement structures are built from minimal, extremal codimension-2 bulk surfaces whose areas—including possible codimension-3 boundary terms—compute von Neumann entropies of CFT subregions. On quantum-theoretic grounds, such entropies must be positive. However, certain ranges of brane-localized couplings induce negative entropies through the codimension-3 boundary terms. Such couplings are in the swampland.

Additionally, we can look to the relationship between entanglement structures and causal structures. Both entanglement and causality can be used to devise proposals for a bulk region associated with the degrees of freedom on a CFT subregion  $\mathcal{R}$  Bousso et al. (2012); Hubeny and Rangamani (2012); Czech et al. (2012); Hubeny et al. (2013); Headrick et al. (2014). Fundamental causality conditions in the field theory imply that the resulting bulk “causal wedge” associated with  $\mathcal{R}$  must always be included in the analogous “entanglement wedge” of  $\mathcal{R}$ —a property called *causal wedge inclusion (CWI)*. And so, we employ CWI as a swampland criterion on brane-localized couplings in braneworld theories.

We should say that using CWI as a swampland condition (but without branes) is not new. It has also been used to constrain Gauss–Bonnet gravity Cáceres et al. (2019) and explored in Einstein cubic gravity Cáceres et al. (2021). However, causality in these “higher-derivative” theories need not be the same as in Einstein gravity,

in that the fastest causal modes may not be lightlike but rather superluminal Izumi (2014); Reall et al. (2014), so causal structures often need to be reformulated. Fortunately, this is not a problem if we only have higher-derivative terms localized to the brane Gao and Wald (2000); Omiya and Wei (2021).

### 8.1.1 Relevance to black-hole information

In recent years, doubly holographic braneworld models like those studied in this paper have served as tools by which to probe the problem of black-hole information in  $d > 2$  dimensions Hawking (1976b); Almheiri et al. (2019d). Indeed, one of our motivations for constraining such models is to rein in their application, so that we are not led to artificial conclusions by unphysical setups.

Let us review the recent progress of black-hole information in AdS. The basic setup of Penington (2019); Almheiri et al. (2019a) is to couple a gravitating black hole to a non-gravitating thermal bath. We work in a semiclassical regime with quantum fields turned on but not backreacting onto the metric. Hawking radiation is collected in a subregion  $\mathcal{R}$  on some Cauchy slice of the bath. As we want to include quantum effects, the entropy of this Hawking radiation  $S[\mathcal{R}]$  is computed by the *quantum extremal surface* prescription Engelhardt and Wall (2015b), which yields the “island rule” Almheiri et al. (2020c):

$$S[\mathcal{R}] = \min_{\mathcal{J}} \text{ext} \left( S_{\text{scl}}[\mathcal{R} \cup \mathcal{J}] + \frac{A[\partial\mathcal{J}]}{4G_{\text{N}}} \right). \quad (8.3)$$

$\mathcal{J}$  is a bulk region that is disconnected from  $\mathcal{R}$ , and so we call it an “island.”  $S_{\text{scl}}[\mathcal{R} \cup \mathcal{J}]$  is the entropy of matter fields in  $\mathcal{R} \cup \mathcal{J}$ , and  $A[\partial\mathcal{J}]$  is the area of the boundary of the island. We need to find  $\mathcal{J}$  such that the expression in parentheses, which is a “generalized entropy,” is minimized. For both evaporating Penington (2019); Almheiri et al. (2019a) and eternal Almheiri et al. (2019c) 2d black holes,  $S[\mathcal{R}]$  was found to follow the usual time-dependent Page curve Page (1993c)—thereby preventing an information paradox—thanks to the emergence of an island (i.e. a transition from  $\mathcal{J} = \emptyset$  to  $\mathcal{J} \neq \emptyset$ ).

The higher-dimensional setup of Almheiri et al. (2019d) relies on treating the matter as holographic so that  $S_{\text{scl}}$  can be computed using bulk geometry Almheiri et al. (2020e). This is accomplished by embedding a single  $d$ -dimensional KR brane in a  $(d + 1)$ -dimensional AdS black geometry. However, Geng and Karch (2020b) shortly after claimed that the existence of a time-dependent Page curve and an island may be due to the semiclassical theory on the brane being one of *massive* gravity, a feature long attributed to the presence of a non-gravitating bath Porrati (2002a); Porrati and Starinets (2002).

This aspect of the one-brane setup motivated the study of a two-brane setup embedded in bulk AdS spacetime Geng et al. (2021b). Here, one of the branes acts as a gravitating bath, and so the semiclassical theory on the branes maintains a massless graviton in its spectrum Kogan et al. (2001b). To be concrete, we introduce the following terminology for the three different perspectives of the (two-KR-brane) braneworld afforded by double holography:

- (I) the bulk, which is classical  $(d + 1)$ -dimensional Einstein gravity on a “wedge” Bousso and Wildenhain (2020); Akal et al. (2020) with  $d$ -dimensional KR branes (i.e. with AdS geometry);
- (II) the intermediate picture, which is the effective field theory describing semiclassical  $d$ -dimensional (and notably, massless Kogan et al. (2001b)) gravity on AdS; and
- (III) the defect system, i.e. the  $(d - 1)$ -dimensional CFT on the tip of the bulk wedge.

Any version of the black hole information problem must be posed in the intermediate picture, in which we must explicitly account for quantum corrections. However, if there is a known bulk picture, then we can map the problem to an easier one of classical geometry in the bulk Almheiri et al. (2020e). This is the power of double holography, as we will see in Section 8.2.

In Geng et al. (2021b), we studied two  $d > 2$  dimensional, single-sided AdS-Schwarzschild black holes in the intermediate picture coupled to one another at infinity. However, unlike with a non-gravitating bath, the radiation region  $\mathcal{R}$  (assumed to not be anchored to the defect) must be dynamically determined to protect diffeomorphism invariance, i.e. we minimize over both  $\mathcal{J}$  and  $\mathcal{R}$  in the island rule (8.3) to compute the entropy  $S$ .

With that in mind, one can consider several scenarios of how to divide the system in order to define an entanglement entropy. The most suitable for our purposes is to study in the fully quantum defect system (III) the von Neumann entropy of the defect in the thermofield double state (TFD). That is, we trace out the degrees of freedom of the double to obtain a density matrix for the CFT on the defect and calculate the associated entropy. In prescriptions (II) and (I) we are respectively looking either for a quantum extremal surface on the brane or, equivalently, a classical RT surface in the bulk that separates the two defects of the thermofield double state from each other.

In Geng et al. (2021b), we solved this particular problem using a bulk geometry with two  $d$ -dimensional KR branes [with the RS action (8.1)] embedded in a  $(d + 1)$ -dimensional AdS black string. We found that the entropy is computed by the bulk black string’s horizon area, thus following a “trivial” (time-independent) Page curve rather than a “nontrivial” (time-dependent) one. The Page curve’s triviality in theories of massless gravity was earlier suggested by Geng and Karch (2020b); Laddha et al. (2020) and asserted in Raju (2022a); Geng et al. (2021c). Furthermore unlike in the case with a non-gravitating bath, no entanglement island (in the sense of a disconnected bulk region) is formed.

Recent work Miao (2022a, 2023b); Li and Miao (2023) has claimed that one may get nontrivial Page curves and islands even with a gravitating bath (and thus, massless gravity) by turning on brane-localized couplings—particularly DGP terms (8.2)—treated as higher-derivative corrections to RS terms (8.1). Such a result re-



quires that the coupling  $G_b$  in (8.2) be negative on at least one of the branes, but this alone is not necessarily an unphysical assumption for AdS branes. However, in such theories, one may get entropies that are smaller than Bekenstein–Hawking or even time-dependent. Such answers violate *bulk* causal wedge inclusion, putting such scenarios in the swampland. In other words, the radiation entropy for UV-consistent couplings must be the bulk Bekenstein–Hawking value at all times.

### 8.1.2 Outline

In Section 8.2, we will set the stage by reviewing the holography of the two-brane setup, in particular discussing how the branes modify the usual entanglement entropy prescriptions Ryu and Takayanagi (2006c); Hubeny et al. (2007b). We will also justify both the positivity of entropy computed by the codimension-2 entanglement surfaces and the rationale for CWI.

In Section 8.3, we will scrutinize RS + DGP gravity [(8.1) + (8.2)] through both analytic and numerical methods. Instead of using  $T$  and  $G_b$  as our brane parameters, we find it more convenient to use the Euclidean angle of the AdS brane with the conformal boundary (the “brane angle”  $\theta$ ) and a particular dimensionless combination of  $G_b$  with the bulk Newton constant and curvature radius (the “DGP coupling”  $\lambda$ ). The formulas that switch between these pairs of parameters are written in Section 8.3.1. The parameter space for two-brane systems is four-dimensional, so we explore how our swampland criteria exclude ranges of the DGP couplings for particular brane angles. See Figure 8.5 for the analytic constraints from entropy positivity and Figure 8.6 for the numerical points excluded by CWI.

In Section 8.4, we discuss the consequences of our exclusion criteria on the physics of two-brane braneworld models in the intermediate picture. In particular, we describe the implications for black-hole information in higher-dimensional ( $d > 2$ ) KR branes with massless gravity. Essentially, our criteria—particularly CWI—mandate that UV-consistent two-brane setups furnish a trivial semiclassical Page curve for the

entropy of Hawking radiation on the brane, in accordance with the ideas of Laddha et al. (2020); Geng et al. (2021b,c), with the entropy always given by the horizon area. Thus, braneworld theories with massless gravity on the brane and whose entanglement structures yield nontrivial Page curves Miao (2022a, 2023b); Li and Miao (2023) contradict CWI and should be considered in the swampland.

## 8.2 The holography of two-brane models

We first review the holographic interpretations of the two-KR-brane setups with just RS terms so as to lay the groundwork for applying the tools of holography to more exotic braneworld constructions.

We start by recalling the setup in question. It consists a bulk spacetime  $\mathcal{M}$  with two codimension-1 boundaries  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ . The action (including Gibbons–Hawking terms) is

$$I_{\text{RS1}} = \frac{1}{16\pi G_{d+1}} \int_{\mathcal{M}} \sqrt{-g} \left[ R + \frac{d(d-1)}{L^2} \right] + \sum_{i=1}^2 \frac{1}{8\pi G_{d+1}} \int_{\mathcal{Q}_i} \sqrt{-\tilde{g}_i} (K_i - 8\pi G_{d+1} T_i), \quad (8.4)$$

where  $\tilde{g}_i$  is the metric of  $\mathcal{Q}_i$  and  $K_i$  is the trace of the extrinsic curvature  $K_{i,\mu\nu} = \nabla_\mu n_{i,\nu}$  of  $\mathcal{Q}_i$  (with  $n_{i,\nu}$  being the unit normal and the  $\mu, \nu$  indices projected onto  $\tilde{g}_i$ ). For now, we have suppressed the coordinate dependence to simplify the notation. Additionally,  $i = 1, 2$  labels the two branes while Greek letters are spacetime indices. The equations of motion of this action are

$$G_{\mu\nu} = \frac{d(d-1)}{2L^2} g_{\mu\nu}, \quad (8.5)$$

$$K_{i,\mu\nu} = (K_i - 8\pi G_{d+1} T_i) \tilde{g}_{i,\mu\nu}. \quad (8.6)$$

(8.5) implies that the bulk geometry is locally  $\text{AdS}_{d+1}$ . (8.6) contracted with  $\tilde{g}_i$  yields

$$K_i = \frac{8\pi G_{d+1} d}{d-1} T_i. \quad (8.7)$$

In other words, solutions to (8.6) are constant-curvature boundaries in  $\text{AdS}_{d+1}$ .

We are interested in branes that have timelike boundary and thus extend eternally in the bulk. This is only the case when the tensions are “subcritical” Karch and Randall (2001b, 2020) or

$$|T_i| < \frac{d-1}{8\pi G_{d+1}L}. \quad (8.8)$$

For such tensions, the resulting brane has  $\text{AdS}_d$  geometry Karch and Randall (2001b) and is thus deemed a Karch–Randall (KR) brane. Critical tensions saturating this bound yield Minkowski spacetime instead, and so the near-critical limit is the regime in which the brane’s cosmological constant is small. This is often the limit of interest because it is the regime where the lowest-mass Kaluza–Klein (KK) mode of linearized fluctuations of the brane “localizes,” yielding a bound graviton on the brane coupled to parametrically heavier CFT modes Karch and Randall (2001b). However, we will generally allow any subcritical tension and assume that there is some holographic description of the  $(d+1)$ -dimensional bulk theory as  $d$ -dimensional semiclassical gravity on the brane<sup>4</sup> and coupled to CFT fields Porrati (2002b). We call this lower-dimensional theory the “intermediate picture” due to its role in double holography Almheiri et al. (2020e).

KR branes can be engineered by foliating  $\text{AdS}_{d+1}$  into  $\text{AdS}_d$  slices (cf. Aharony et al. (2003)). The individual foliates in such a slicing reach the conformal boundary. Taking just one of these slices as a KR brane yields a setup in which the intermediate picture describes a gravitating universe on  $\text{AdS}_d$  coupled at infinity to a non-gravitating CFT. When we instead construct a setup using two KR branes, they intersect along a  $(d-1)$ -dimensional surface (called the *defect*) at the conformal boundary. In the intermediate picture, this defect is viewed as an interface between the two gravitating  $\text{AdS}_d$  universes, and the full system has a massless graviton.<sup>5</sup> See Figure 8.1 for schematic representations of these dual setups.

---

<sup>4</sup>In this interpretation of the universe on the brane, the lowest-mass KK mode is still the graviton, but it has finite mass and is no longer parametrically lighter than the higher KK modes constituting the CFT.

<sup>5</sup>Technically, the massless graviton exists in tandem with massive modes. This is called “bigravity” Kogan et al. (2001b).

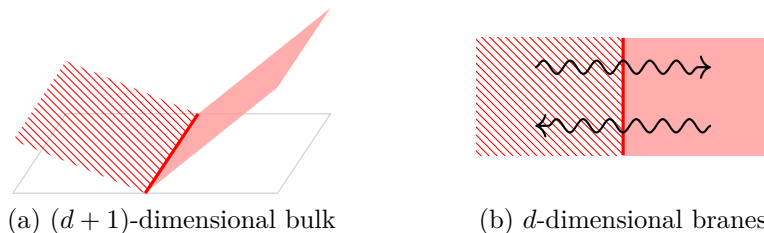


Figure 8.1: A schematic representation of the two-brane setup of interest in this paper. (a) is the classical bulk configuration consisting of two KR branes (in red) embedded in an ambient  $\text{AdS}_{d+1}$  spacetime. (b) is the brane-localized “intermediate picture” of two interacting  $\text{AdS}_d$  universes coupled together along a  $(d - 1)$ -dimensional interface (the “defect”). Both pictures can be seen as dual to a  $(d - 1)$ -dimensional CFT on this defect.

So far, we have discussed the bulk perspective and the intermediate picture. However, the AdS/CFT correspondence Maldacena (1998b) suggests that there is a third description—a conformal field theory on the defect (since it is the boundary of the bulk spacetime) Karch and Randall (2001c). The correspondence between the bulk and defect systems has more recently been dubbed “wedge holography” Bousso and Wildenhain (2020); Akal et al. (2020)<sup>6</sup> and falls under the broader umbrella of “double holography” equating the  $(d + 1)$ -dimensional bulk,  $d$ -dimensional intermediate, and  $(d - 1)$ -dimensional defect systems. The upshot is that physical quantities of the  $(d - 1)$ -dimensional CFT are dual to various geometric data in the  $(d + 1)$ -dimensional bulk theory, and the dictionary translating between these two pictures can be used to further understand semiclassical physics in the intermediate picture without *explicitly* accounting for its quantum effects.

---

<sup>6</sup>It behooves us to mention that this logic also goes through for the configuration consisting of just one brane. There, the dual field theory is a BCFT Cardy (1984, 2004). The relationship between this BCFT and the bulk theory has been dubbed the “AdS/BCFT correspondence” Takayanagi (2011b); Fujita et al. (2011).

### 8.2.1 Entanglement entropy

The main entry of the holographic dictionary of relevance to our paper is the classical prescription for computing entanglement entropies of CFT subsystems. In AdS/CFT with no branes, the formula is given by the Ryu–Takayanagi (RT) prescription Ryu and Takayanagi (2006c) (or its covariant extension of Hubeny–Rangamani–Takayanagi Hubeny et al. (2007b)); for a subregion  $\mathcal{R}$  of the boundary CFT, the entanglement entropy to leading order in the limit  $G_{d+1} \ll L^{d-1}$  is

$$S[\mathcal{R}] = \frac{1}{4G_{d+1}} \min_{\gamma_{\mathcal{R}} \sim \mathcal{R}} \text{ext } A[\gamma_{\mathcal{R}}]. \quad (8.9)$$

Here,  $\gamma_{\mathcal{R}}$  is a  $(d-1)$ -dimensional surface that is homologous to the CFT subregion  $\mathcal{R}$ , where “homology” means that there exists some codimension-1 bulk region  $\Sigma$  such that

$$\partial\Sigma = \gamma_{\mathcal{R}} \cup \mathcal{R}, \quad \partial\gamma_{\mathcal{R}} = \partial\mathcal{R}, \quad (8.10)$$

and  $A[\gamma_{\mathcal{R}}]$  is the area functional for such surfaces. Essentially, this formula equates entanglement entropy with the area of the smallest extremal surface “anchored” to the CFT subregion  $\mathcal{R}$  (Figure 8.2a). However, note that this does not include contributions from quantum fields in the bulk.

We can ask how this prescription is modified with two KR branes. The proposal Almheiri et al. (2019d); Bousso and Wildenhain (2020); Akal et al. (2020); Geng et al. (2021b) is to equate the von Neumann entropy of the defect at some fixed time with a minimal extremal surface in the bulk via a modification of (8.9). Essentially, the homology condition (8.10) is modified such that the surface is anchored to the branes, rather than to the conformal boundary (Figure 8.2b). This generically introduces boundary terms in the entropy functional, and the extremization procedure also requires us to choose boundary conditions (dynamical or fixed) for the intersection of the surface with the branes. While a fixed boundary condition like Dirichlet is technically a mathematically consistent choice Ghosh and Krishnan (2021), imposing it amounts to partitioning degrees of freedom on the brane by hand, a procedure

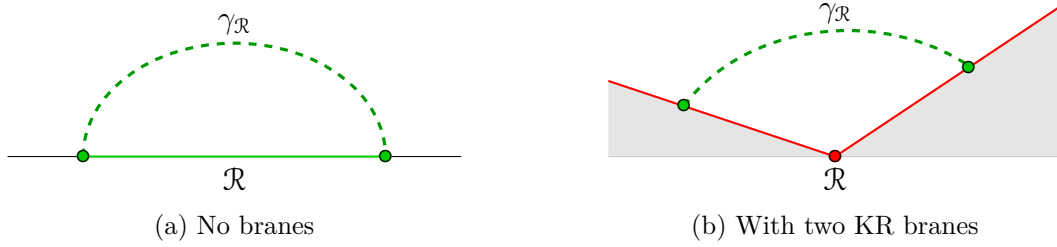


Figure 8.2: Schematic representations of the Ryu–Takayanagi surfaces for a fixed-time CFT subregion  $\mathcal{R}$  in (a) AdS with no branes and (b) AdS with two KR branes. In (a), the surface is anchored to the boundary of  $\mathcal{R}$  (i.e.  $\partial\gamma_{\mathcal{R}} = \partial\mathcal{R}$ ); this is a Dirichlet condition on  $\gamma_{\mathcal{R}}$ . In (b), the CFT subregion is the  $(d-1)$ -dimensional defect at fixed time, and  $\gamma_{\mathcal{R}}$  is anchored to the branes via a Neumann condition. We use Neumann for physical reasons; a Dirichlet condition would break diffeomorphism invariance on the brane.

that we assert is physically incompatible with diffeomorphism invariance Geng et al. (2021b). And so, we impose dynamical (e.g. Neumann) conditions on the RT surface as in Takayanagi (2011b).<sup>7</sup>

In the intermediate picture, this same entropy is interpreted as a generalized entropy that accounts not just for a leading-order semiclassical entropy, but also for quantum fields on the brane. Typically, the generalized entropy is calculated by a quantum extremal surface (QES) Engelhardt and Wall (2015b) rather than an RT surface, and the QES may even be the boundary of disconnected pieces of the entanglement wedge called “entanglement islands” Penington (2019); Almheiri et al. (2019a). By the power of double holography, the generalized entropy in the intermediate picture can be computed by the area of a classical RT surface in the  $(d+1)$ -dimensional bulk Almheiri et al. (2020e). This is the key feature of double holography that we will exploit in this paper.

In other words, the semiclassical entropy on the brane, which might normally

---

<sup>7</sup>We should be clear that there is no proof in the sense of Lewkowycz–Maldacena Lewkowycz and Maldacena (2013b); Dong et al. (2016b) for this prescription. However, there is evidence in the literature for this approach, e.g. in Chen et al. (2020b); Anous et al. (2022).

be computed by employing the quantum extremal surface prescription, is encoded by a classical surface in the bulk. As classical surfaces are rather tractable to compute, double holography has allowed the use of one-brane setups to study the entropy of quantum fields in the presence of a gravitating black hole coupled to a non-gravitating thermal bath. Indeed, this has been the main approach to study higher-dimensional ( $d > 2$ ) semiclassical gravity, starting with Almheiri et al. (2019d) and extended by Geng and Karch (2020b); Chen et al. (2020b,a) and many others as part of the entanglement island program.

While we will briefly discuss black-hole information and islands, our primary focus is on how the bulk entropy prescription can be used to constrain brane-localized couplings added by hand to the theory. In other words, we are assuming the holographic dictionary to be a fundamental aspect of UV physics that is capable of providing swampland criteria on effective field theories of gravity (cf. Akers et al. (2016); Geng et al. (2019); Cáceres et al. (2019); Folkestad (2022)).

We now briefly justify the two main swampland criteria rooted in entanglement. We will later explain how we concretely use them in Section 8.2.2.

**Positivity of entropy** One immediate swampland criterion is the positivity of entropy. Generally, microcanonical entropy is a logarithmic count of the microstates  $\Omega$  in a system,

$$S \sim \log \Omega, \quad \Omega \geq 1. \quad (8.11)$$

Entanglement entropy (i.e. the von Neumann entropy of a reduced density matrix) can also be written as (8.11), albeit with  $\Omega$  being real rather than an integer. By employing a Schmidt decomposition of the initial state, we can identify  $\Omega$  with a function of Schmidt coefficients that is necessarily greater than 1,<sup>8</sup> and this crucially lets us insist that entanglement entropy be positive.

---

<sup>8</sup>To be concrete, suppose that we have a normalized state  $|\Psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ . For simplicity, if we take these Hilbert spaces to be finite-dimensional, we can define a finite  $n \equiv \min(\dim \mathcal{H}_A, \dim \mathcal{H}_B)$ . We then Schmidt-decompose  $|\Psi\rangle = \sum_{i=1}^n \alpha_i |a_i\rangle \otimes |b_i\rangle$ , where  $\{|a_i\rangle\} \subset \mathcal{H}_A$  and  $\{|b_i\rangle\} \subset \mathcal{H}_B$  are

However, the entanglement entropy of a subregion formed by splitting a connected interval is typically UV-divergent in QFT Witten (2018). This is reflected in the RT prescription with no branes present through the fact that RT surfaces reach the conformal boundary, thus formally being infinite in area. Nonetheless, these “absolute” entropies are still positive. While we can subtract away the infinity through some renormalization scheme and get a negative answer, such a result does not contradict the positivity of the “bare” entropy.

But in two-brane configurations, we no longer have this positive divergence in the area. This is because we are not splitting any CFT regions. Furthermore, the full entropy functional picks up “boundary-term” contributions from the couplings on the branes, and so it is *a priori* possible to have a functional with negative values. However, a negative-entropy UV-finite RT surface contradicts (8.11), and so couplings giving such a result would be in the swampland.

**The basis of CWI** Another swampland condition coming from AdS/CFT is causal wedge inclusion (CWI)—the property that the entanglement wedge always contains the causal wedge. Let us discuss why such a criterion is fundamental.

First, we recall the motivations and definitions of the causal and entanglement wedges. Both ideas came out of attempts Bousso et al. (2012); Hubeny and Rangamani (2012); Czech et al. (2012); Hubeny et al. (2013); Headrick et al. (2014) to answer to the question, “What is the holographic dual of the reduced density matrix  $\rho_{\mathcal{R}}$  on the CFT subregion  $\mathcal{R}$ ?” Recall that this reduced density matrix is defined by tracing out complementary degrees of freedom  $\mathcal{R}^c$  and encodes the degrees of freedom of  $\mathcal{R}$ :

$$\rho_{\mathcal{R}} \equiv \text{Tr}_{\mathcal{R}^c}(\rho), \quad S[\mathcal{R}] \equiv -\text{Tr}_{\mathcal{R}}(\rho_{\mathcal{R}} \log \rho_{\mathcal{R}}). \quad (8.12)$$

---

orthonormal sets and each  $\alpha_i$  lies between 0 and 1 (and  $\sum_{i=1}^n \alpha_i^2 = 1$ ). Pushing through the von Neumann entropy formula yields  $\Omega \equiv \prod_{i=1}^n c_i^{-2c_i^2}$ , which is easily seen to be  $\geq 1$ .



Starting with  $\mathcal{R}$ , we define its future domain of dependence  $D_+[\mathcal{R}]$  as the points  $p$  for which all past-directed causal curves starting at  $p$  intersect  $\mathcal{R}$ . We analogously define the past domain of dependence  $D_-[\mathcal{R}]$  as the points from which all future-directed causal curves intersect  $\mathcal{R}$ . The union  $D[\mathcal{R}] = D_+[\mathcal{R}] \cup D_-[\mathcal{R}]$  is the full domain of dependence and is shaped like a diamond.

The causal wedge  $\text{CW}[\mathcal{R}]$  is defined in terms of the domain of dependence in the boundary system  $D[\mathcal{R}]$ . Basically, we take all bulk causal curves which start in  $D_-[\mathcal{R}]$  and end in  $D_+[\mathcal{R}]$  Hubeny and Rangamani (2012). If  $\mathcal{R}$  is a fixed-time subregion (say at  $t = 0$ ), this comes with a codimension-2 surface  $\xi_{\mathcal{R}}$  that is anchored to  $\mathcal{R}$  and bounds the  $t = 0$  cross-section of the causal wedge.

Meanwhile, the entanglement wedge  $\text{EW}[\mathcal{R}]$  is defined in terms of the RT surface  $\gamma_{\mathcal{R}}$  of  $\mathcal{R}$ . Suppose that  $\mathcal{R}$  is spacelike. We then take the bulk codimension-1 “homology surface” bounded by  $\mathcal{R}$  and  $\gamma_{\mathcal{R}}$  within the Cauchy slice containing  $\mathcal{R}$  [i.e.  $\Sigma$  in (8.10)] and define  $\text{EW}[\mathcal{R}]$  as the domain of dependence of this homology surface Headrick et al. (2014). Just like  $\text{CW}[\mathcal{R}]$ , the entanglement wedge asymptotes to  $D[\mathcal{R}]$ .<sup>9</sup> However,  $\text{EW}[\mathcal{R}]$  is constructed from entanglement structure and is generally distinct from  $\text{CW}[\mathcal{R}]$ . In particular, for  $\mathcal{R}$  at  $t = 0$ , the corresponding RT surface  $\gamma_{\mathcal{R}}$  need not be the same as  $\xi_{\mathcal{R}}$ .

In fact, one can argue that the entanglement wedge is generally supposed to reach further into the bulk than the causal wedge Headrick et al. (2014). This is precisely the statement of CWI and is depicted in Figure 8.3. Furthermore, for a boundary subregion  $\mathcal{R}$  at  $t = 0$ , CWI implies a condition on the codimension-2 surfaces  $\xi_{\mathcal{R}}$  and  $\gamma_{\mathcal{R}}$  residing on the bulk  $t = 0$  Cauchy slice—the region bounded by  $\gamma_{\mathcal{R}}$  must contain the region bounded by  $\xi_{\mathcal{R}}$ . Physically, CWI states that the RT surface of  $\mathcal{R}$  is generally causally disconnected from the domain of dependence  $D[\mathcal{R}]$ ,

---

<sup>9</sup>This is because the bulk and boundary causal structures are compatible with one another (cf. Gao and Wald (2000); Omiya and Wei (2021)), and so null rays confined to the boundary can also be seen as bulk null rays “at infinity.”

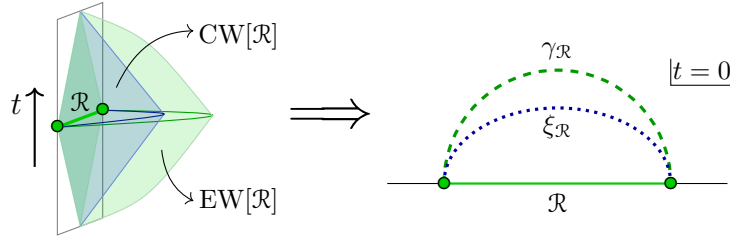


Figure 8.3: On the left, a cartoon of causal wedge inclusion (CWI) for a CFT subregion  $\mathcal{R}$  at fixed boundary time  $t = 0$ . The inner (blue) wedge is the causal wedge  $CW[\mathcal{R}]$ , while the outer (green) wedge is the entanglement wedge  $EW[\mathcal{R}]$ . We can further consider the cross-section of this picture with the initial bulk Cauchy slice (right), in which  $\gamma_{\mathcal{R}}$  (green, dashed) is the RT surface wrapping around  $EW[\mathcal{R}]$  while  $\xi_{\mathcal{R}}$  (blue, dotted) is the codimension-2 surface wrapping around  $CW[\mathcal{R}]$ . CWI implies that the region on this Cauchy slice bounded by  $\xi_{\mathcal{R}}$  must be within the analogous region bounded by  $\gamma_{\mathcal{R}}$ .

at best only being accessible via null signals, and the argument relies on causality in the boundary theory.

Let us give the intuition behind this argument; for a more thorough proof, see Headrick et al. (2014). Again, take  $\mathcal{R}$  to be a CFT subregion on the  $t = 0$  Cauchy slice. Suppose that the RT surface  $\gamma_{\mathcal{R}}$  is located between  $\xi_{\mathcal{R}}$  and  $\mathcal{R}$  on this slice. This implies that the RT surface can be reached causally by timelike signals sent from the domain of dependence  $D[\mathcal{R}]$ .

Such timelike signals can be realized in the field theory as small perturbations of the Hamiltonian which evolves the reduced state on  $\mathcal{R}$ . The perturbed state is

$$\rho_{\mathcal{R}'} = U \rho_{\mathcal{R}} U^\dagger, \quad (8.13)$$

where  $U$  is some unitary with support on just  $D[\mathcal{R}]$ .

If  $\gamma_{\mathcal{R}}$  is in the causal wedge, there exists a perturbation which corresponds to a deformation of the bulk metric in the vicinity of  $\gamma_{\mathcal{R}}$ . Thus, applying this perturbation would alter the area of  $\gamma_{\mathcal{R}}$  and change the entropy. However, on the CFT side, we can use the invariance of the partial trace under change of basis (since the Hilbert spaces on  $\mathcal{R}$  and  $\mathcal{R}'$  are isomorphic) to equate the  $n$ th Rényi entropy of the perturbed state

to that of the original state:

$$S_n[\rho_{\mathcal{R}'}] = -\frac{1}{n-1} \log \text{Tr}_{\mathcal{R}'}[(\rho_{\mathcal{R}'})^n] = -\frac{1}{n-1} \log \text{Tr}_{\mathcal{R}}[(\rho_{\mathcal{R}})^n] = S_n[\rho_{\mathcal{R}}]. \quad (8.14)$$

Recall that the entanglement entropy is the  $n \rightarrow 1$  limit of the Rényi entropy. So, the entanglement entropy should *not* change under the perturbation. This is a contradiction and we thus conclude that  $\gamma_{\mathcal{R}}$  cannot be strictly within the causal wedge of  $\mathcal{R}$ , after all.

Note that the arguments of Headrick et al. (2014) and the discussion above are set in the realm of AdS/CFT without KR branes. However, we expect CWI to also hold with branes. This is because all that is necessary for CWI is the notion of a causal wedge and an entanglement wedge. Just having a working proposal for both is enough. More specifically, the causal wedge in a two-brane system can again be constructed from causal curves. Meanwhile, we consider the entanglement wedge to be constructed from the surface found via the modified entanglement prescription outlined in the previous exposition of Ryu–Takayanagi.

### 8.2.2 Black spacetimes as a testing ground

As one application of entropy positivity and CWI as swampland constraints on higher-derivative corrections to the RS action, we examine two-KR-brane systems embedded in eternal two-sided “black” spacetimes in the bulk. All of our analysis will take place on one side of the initial-time slice. In principle, we may consider other two-brane setups in other geometries, and doing so may yield additional nontrivial constraints on brane couplings. We focus on black spacetimes because they have particularly useful entanglement and causal features.

**Positivity of horizon entropy** The original motivation of the RT prescription was the Bekenstein–Hawking identification of a black-hole horizon’s area with the black hole’s entropy Bekenstein (1973b); Hawking (1976b), which was also found in string

theory Strominger and Vafa (1996b). Indeed, RT is meant to be a generalization of the Bekenstein–Hawking formula. We recover the latter when we consider a CFT in a thermal state, which is realized holographically by a one-sided black hole. When  $\mathcal{R}$  is taken to be the full boundary, then the associated RT surface is the horizon.

In two-brane setups embedded in black holes, the entropy functional is generically modified by boundary terms coming from brane-localized couplings. Nonetheless, the horizon is still expected to be an extremal surface. Given this, the quantity  $A_h$  (defined as the sum of the “bulk” horizon area with its boundary terms at the branes) must then at least upper-bound the minimum of the entropy functional  $S_{\min}$ , which represents entropy in the semiclassical limit and is thus positive. Thus,

$$S_h \equiv \frac{A_h}{4G_N} > S_{\min} > 0. \quad (8.15)$$

A negative value of  $S_h$  (induced by sufficiently negative boundary terms in the functional) contradicts the positivity of entropy. However, we might mathematically get such a result from the brane-localized couplings, but such theories live in the swampland.

**CWI in black spacetimes** Consider a one side of a two-sided eternal black geometry at some initial Cauchy slice  $t = 0$ . We take  $\mathcal{R}$  to be the boundary of the chosen side at this slice. In this case,  $\mathcal{D}[\mathcal{R}]$  is the full boundary, with the earliest and latest points respectively being past and future timelike infinity. The bulk causal wedge  $\text{CW}[\mathcal{R}]$  can then be constructed by shooting null rays from these points. Upon doing so, we see that the  $t = 0$  slice of  $\text{CW}[\mathcal{R}]$  is precisely the exterior region outside of the horizon,<sup>10</sup> and so the  $t = 0$  bounding surface  $\xi_{\mathcal{R}}$  of the causal wedge is itself the horizon.

The point is that applying the CWI criterion in black geometries simply becomes a matter of explicitly computing the RT surface  $\gamma_{\mathcal{R}}$  with  $\mathcal{R}$  being the boundary

---

<sup>10</sup>In fact, the causal wedge is the full exterior on the same side of the horizon as  $\mathcal{R}$ . One can see this by drawing the Penrose diagram and shooting null rays from past and future timelike infinity.

system at  $t = 0$ . For a general braneworld with Einstein gravity in the bulk,  $\gamma_{\mathcal{R}}$  will be either the horizon (in accordance with the Bekenstein–Hawking formula Bekenstein (1973b); Hawking (1976b)) or some other surface. The former is consistent with CWI, so the only point of concern is the latter case.

Fortunately, non-horizon surfaces can easily be seen to contradict CWI. In the maximally extended spacetime at  $t = 0$ , there is no interior geometry. Thus, the only conceivable non-horizon candidates for  $\gamma_{\mathcal{R}}$  are those that intersect the horizon or those that reside completely in the exterior. In either case, there would be a part of the  $t = 0$  slice of the exterior that is not a part of the entanglement wedge, which then implies  $\text{CW}[\mathcal{R}] \not\subset \text{EW}[\mathcal{R}]$ .

So, to summarize how to use CWI to potentially rule out a theory, we start by taking a two-KR-brane setup embedded in a black spacetime. We then search for  $t = 0$  non-horizon extremal surfaces anchored to the branes. If we find one, we can compute its entropy  $S_{\text{ext}}$  and compare it against that of the horizon  $S_{\text{h}}$ . *A priori* we have the following possibilities:

$$S_{\text{ext}} > S_{\text{h}} \implies \text{no violation of CWI}, \quad (8.16)$$

$$S_{\text{h}} > S_{\text{ext}} \implies \text{theory violates CWI}. \quad (8.17)$$

A theory with a non-horizon extremal surface for which the latter holds is in the swampland.

**Black string with RS terms** A simple test is to check that both criteria are satisfied in two-KR-brane configurations solving the equations (8.5)–(8.6). We do not expect braneworld theories with only RS terms to reside in the swampland, since the bulk action is just Einstein gravity while the brane action has only the terms corresponding to zeroth and first derivatives of the metric.<sup>11</sup> One black solution in

---

<sup>11</sup>The zeroth-derivative term is the RS term and has one free parameter—the tension. The first-derivative term is the Gibbons–Hawking term proportional to extrinsic curvature, and its coefficient

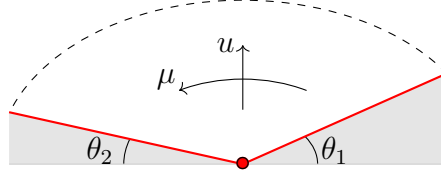


Figure 8.4: A fixed- $t$ , fixed- $\vec{x}$  slice of a planar black string in AdS consisting of two planar KR branes (in red), each at a fixed angle  $\theta_1, \theta_2 \in (0, \frac{\pi}{2})$  with respect to the conformal boundary of the underlying  $\text{AdS}_{d+1}$  spacetime. The defect (the red point) is at  $u = 0$  and the horizon (the dashed line) is at  $u = u_h$ . The branes have constant extrinsic curvature, and so this geometry solves the equations of motion with the RS action (8.5)–(8.6).

which we can test this claim is the following  $(d + 1)$ -dimensional planar AdS black-string geometry:

$$ds^2 = \frac{L^2}{u^2 \sin^2 \mu} \left[ -h(u)dt^2 + \frac{du^2}{h(u)} + u^2 d\mu^2 + d\vec{x}^2 \right], \quad h(u) = 1 - \frac{u^{d-1}}{u_h^{d-1}}. \quad (8.18)$$

Here,  $t \in \mathbb{R}$ ,  $u > 0$ , and  $\mu \in (0, \pi)$  are respectively the time, radial, and angular coordinates while  $\vec{x} \in \mathbb{R}^{d-2}$  represents the  $d - 2$  remaining transverse, planar directions. The defect is located at  $u = 0$ , and  $u = u_h$  is the horizon. The equation of motion (8.6) is solved by “planar KR branes” parameterized as

$$\mu = \theta_1, \quad \mu = \pi - \theta_2, \quad (8.19)$$

for constants  $\theta_1, \theta_2 \in (0, \pi)$  with  $\theta_1 + \theta_2 < \pi$ , and taking  $\mu \in (\theta_1, \pi - \theta_2)$ . See Figure 8.4 for a visual depiction. Not only do these hypersurfaces each have constant extrinsic curvature, but the induced geometry of each is a planar AdS-Schwarzschild black hole. Both the tensions  $T_1, T_2$  and the induced curvature radii  $\ell_1, \ell_2$  of these branes are well-defined functions of the angles:

$$T_i = \frac{(d-1)}{8\pi G_{d+1} L} \cos \theta_i, \quad \ell_i = L \csc \theta_i. \quad (8.20)$$

---

is fixed (relative to the bulk Einstein action’s coupling) by requiring that the variational principle is well-defined.

In Geng et al. (2021b), we had explored the bulk entanglement surfaces in this two-KR-brane configuration. With just RS terms, the entropy functional has no boundary terms (see Chen et al. (2020b)), so we had found that the RT surface computing the entropy of the defect at  $t = 0$  is always the horizon. Then, the entropy is computed by just the bulk area of the horizon:

$$S_{\text{h}}|_{\text{RS}} = \frac{L^{d-1}}{4G_{d+1}u_{\text{h}}^{d-2}} \int_{\mathbb{R}^{d-2}} d\vec{x} \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1}\mu}. \quad (8.21)$$

This is always positive. Furthermore, because the horizon is the RT surface, the entanglement wedge is precisely the causal wedge. Thus, there is no violation of CWI.

### 8.3 Constraining the DGP term

We can apply the criteria for entanglement surfaces in Section 8.2 to KR braneworlds beyond those with just RS terms. We do so in an explicit simple model, finding that violations are indeed mathematically possible and can be used to rule out particular brane-localized couplings on physical grounds.

The simplest higher-derivative term we can add to the RS action is one that is linear in the brane’s intrinsic Ricci curvature. Following the conventions of Chen et al. (2020b); Perez-Pardavila (2023b), we write the full action as

$$I = \frac{1}{16\pi G_{d+1}} \int_{\mathcal{M}} \sqrt{-g} \left[ R + \frac{d(d-1)}{L^2} \right] + \sum_{i=1}^2 \frac{1}{8\pi G_{d+1}} \int_{\mathcal{Q}_i} \sqrt{-\tilde{g}_i} \left[ K_i - 8\pi G_{d+1} T_i + \frac{G_{d+1}}{2G_{\text{b},i}} \tilde{R}_i \right], \quad (8.22)$$

which has an additional term—the DGP term Dvali et al. (2000c)—for each brane. The DGP Lagrangian is proportional to the Ricci scalar  $\tilde{R}_i$  computed from the metric  $\tilde{g}_i$ . The  $G_{\text{b},i}$  couplings are new parameters each controlling the strength of this term on the associated brane.

We should mention that the original DGP model describes flat branes living in a flat bulk Dvali et al. (2000c). In the original model, there is no RS term and no bulk

cosmological constant, and the gravitational coupling flows from  $G_b$  at high energies to  $G_{d+1}$  at low energies. However, we are mainly interested with using the machinery of AdS/CFT to put theories in the swampland. To learn anything about flat branes in flat space, we may take flat-space limits of our AdS constraints. The punchline is that the constraints from entanglement lose all power under flat-space limits, and that the only constraint we are left with is that  $G_b > 0$  in order to not have a wrong-sign Euclidean action with maxima instead of minima.<sup>12</sup> We will discuss these points in Appendix 8.A (specifically, 8.A.2).

In the current section, we first discuss the parameter space of two-KR-brane setups with RS + DGP terms, as well as how the entropy functional gets modified by DGP terms. We then compute the equations for extremal surfaces in the presence of these branes. What follows are our main results, i.e. the constraints put on this class of braneworld models, which we obtain via both an analytic approach and a numerical approach. Analytically, we will find that violations of CWI require at least one DGP coupling to be negative. We find further analytic lower-bounds on DGP couplings by requiring positivity of the horizon entropy. Last but not least we numerically find theories violating CWI leading to even more refined bounds.

### 8.3.1 Parametrizing KR branes with RS + DGP terms

Here we consider the DGP term as a higher-derivative correction to the RS term on a KR brane as in Chen et al. (2020b), so we are studying AdS branes in an AdS bulk. The brane-embedding equations are modified relative to (8.6) by adding terms that depend on Ricci curvature:

$$K_{i,\mu\nu} = (K_i - 8\pi G_{d+1} T_i) \tilde{g}_{i,\mu\nu} - \frac{G_{d+1}}{G_{b,i}} \left[ \tilde{R}_{i,\mu\nu} - \frac{1}{2} \tilde{R}_i \tilde{g}_{i,\mu\nu} \right]. \quad (8.23)$$

---

<sup>12</sup>This is a very general constraint that we will also impose in AdS, but note that it is not rooted in entanglement structure. Only minima of the action dominate the semiclassical gravitational path integral.



However, the solutions turn out to be geometrically the same as without the DGP term, but with the tension parameter being redefined.<sup>13</sup> In terms of the angular parameterization (8.19), the brane tension of a KR brane with RS + DGP terms (derived in Appendix 8.A) is

$$T_i = \frac{d-1}{8\pi G_{d+1}L} \cos \theta_i - \frac{(d-1)(d-2)}{16\pi G_{b,i}L^2} \sin^2 \theta_i. \quad (8.24)$$

We mention that the DGP term is one of three second-derivative terms we can add to the brane action. We can also add an action  $\int \sqrt{-\tilde{g}_i}(a_i K_i^2 + b_i K_i^{\mu\nu} K_{i,\mu\nu})$  whose terms are products two first-derivative factors. Interestingly, Lee et al. (2022) finds that the consistency of the bulk variational principle implies  $a_i K_i \tilde{g}_i^{\mu\nu} + b_i K_i^{\mu\nu} = 0$ , and contracting this against  $K_{i,\mu\nu}$  reveals that the corresponding action vanishes on-shell. So, these terms do not alter the embedding equations (8.23). However, it is not immediate if or how these terms modify the entropy functional; see Section 8.4 for details. For simplicity, we only add a DGP term.

It is convenient to define a dimensionless parameter explicitly describing the relative strength of the DGP term against the RS term. We do so in terms of each brane’s “effective” gravitational coupling, which is calculated in detail in Appendix 8.A by integrating out the bulk extra dimension and isolating the  $d$ -dimensional Einstein–Hilbert term in the resulting effective action Chen et al. (2020b).<sup>14</sup> From that analysis, this coupling (the “effective Newton constant”) on an brane with no DGP term is just

$$\frac{1}{G_{\text{RS}}} = \frac{L}{(d-2)G_{d+1}}. \quad (8.25)$$

---

<sup>13</sup>Another approach taken by Chen et al. (2020b) to studying DGP is to shift the tension by some amount proportional to the Ricci curvature of the branes. This cancels the explicit contribution of Ricci terms to the brane stress tensors and yields the equations (8.6).

<sup>14</sup>Note that Chen et al. (2020b) takes two copies of a bulk geometry with a brane glued together via Israel junction conditions, whereas we consider one bulk geometry with an “end-of-the-world” brane. This changes the KR brane’s effective Newton constant by a factor of 2 and also alters the effective cosmological constant.

If we apply that same procedure with a DGP term with coupling  $G_b$  present, however, then we get

$$\frac{1}{G_{\text{eff}}}\Big|_{\text{RS+DGP}} = \frac{L}{(d-2)G_{d+1}} + \frac{1}{G_b} = \frac{1}{G_{\text{RS}}} \left(1 + \frac{G_{\text{RS}}}{G_b}\right). \quad (8.26)$$

So, in (8.22) we define the dimensionless parameters

$$\lambda_1 = \frac{G_{\text{RS}}}{G_{b,1}}, \quad \lambda_2 = \frac{G_{\text{RS}}}{G_{b,2}}, \quad (8.27)$$

as stand-ins for the couplings. In terms of  $\theta_i$  and  $\lambda_i$ , the tension of brane  $i$  is

$$T_i = \frac{d-1}{8\pi G_{d+1}L} \left( \cos \theta_i - \frac{\lambda_i}{2} \sin^2 \theta_i \right). \quad (8.28)$$

We can and will trade the physical parameters  $(T_i, G_{b,i})$  for  $(\theta_i, \lambda_i)$  in our analysis. The latter will turn out to be mathematically more convenient.

The limit  $\lambda_1, \lambda_2 \rightarrow 0$  recovers the RS action (8.4). Meanwhile taking either  $\lambda < -1$  yields a negative effective Newton constant. This leads to the “wrong sign” in the Euclidean actions, in that classical saddles would be maxima rather than minima.<sup>15</sup> Such saddles are actually suppressed in the gravitational path integral and thus ill-defined as semiclassical configurations, and so the region

$$\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}, \quad (8.29)$$

is considered pathological for these branes.

While the bulk classical solutions to (8.22) are essentially the same as for without DGP terms, there is a key change in applying the RT prescription that we must address. In the entropy functional, we pick up contributions from the areas of the intersection points between  $\gamma_{\mathcal{R}}$  and the branes, in accordance with the semiclassical

---

<sup>15</sup>As mentioned above, this constraint—that the brane’s  $G_{\text{eff}}$  coupling must be positive—applies to flat-space DGP constructions, such as the original one of Dvali, Gabadadze, and Porrati Dvali et al. (2000c).

island rule Almheiri et al. (2020e); Chen et al. (2020b). Specifically, the entropy functional (derived by Chen et al. (2020b)) is

$$\begin{aligned} S[\mathcal{R}] &= \min_{\gamma_{\mathcal{R}} \sim \mathcal{R}} \text{ext} \left( \frac{A[\gamma_{\mathcal{R}}]}{4G_{d+1}} + \frac{\tilde{A}[\gamma_{\mathcal{R}} \cap \mathcal{Q}_1]}{4G_{b,1}} + \frac{\tilde{A}[\gamma_{\mathcal{R}} \cap \mathcal{Q}_2]}{4G_{b,2}} \right) \\ &= \frac{1}{4G_{d+1}} \min_{\gamma_{\mathcal{R}} \sim \mathcal{R}} \text{ext} \left[ A[\gamma_{\mathcal{R}}] + \frac{L}{d-2} \left( \lambda_1 \tilde{A}[\gamma_{\mathcal{R}} \cap \mathcal{Q}_1] + \lambda_2 \tilde{A}[\gamma_{\mathcal{R}} \cap \mathcal{Q}_2] \right) \right], \end{aligned} \quad (8.30)$$

where  $A$  is a  $(d-1)$ -dimensional “bulk” area functional while  $\tilde{A}$  is a  $(d-2)$ -dimensional “boundary” area functional. In the RS limit  $\lambda_1, \lambda_2 \rightarrow 0$ , we recover (8.9). The inclusion of boundary terms in the extremization does not alter the bulk equation for extremal curves  $\gamma_{\mathcal{R}}$ , but it does affect boundary conditions.

We reiterate that we apply both the positivity of horizon entropy and causal wedge inclusion (CWI) as swampland criteria on entanglement entropy. With that in mind, our goal is to use the DGP-modified entropy functional in the black-string solution (8.18) (Figure 8.4) to constrain the space of DGP couplings  $(\lambda_1, \lambda_2)$ . Technically the two-KR-brane RS + DGP models have four free parameters which include not just these two couplings but also the brane angles  $(\theta_1, \theta_2)$ . Thus we find  $\theta$ -dependent constraints on  $\lambda_1$  and  $\lambda_2$ . This will allow us to explore how these constraints depend on the brane angles.

We should mention that our approach of using entanglement here builds on some previous work of Chen et al. (2020b), which finds DGP couplings yielding negative  $G_{\text{eff}}$  (i.e.  $\lambda < -1$ ) to be unphysical through using the entropy functional rather than the wrong-sign-action argument. Using a class of extremal surfaces in AdS described as “bubbles,” Chen et al. (2020b) argues that the entropy functional supplemented by DGP terms (8.30) has no global minimum when  $G_{\text{eff}} < 0$ . Thus, the  $\lambda_1, \lambda_2 < -1$  regimes of the DGP couplings are deemed pathological. Notably, the constraints we find are independent from the considerations of Chen et al. (2020b).

### 8.3.2 Extremal curves in the RS + DGP black string

We again consider the  $(d+1)$ -dimensional black string in the bulk with planar branes given by (8.18)–(8.19). For convenience, we fix  $L = 1$ . To reiterate, the metric and branes are

$$ds^2 = \frac{1}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + u^2 d\mu^2 + d\vec{x}^2 \right], \quad h(u) = 1 - \frac{u^{d-1}}{u_h^{d-1}}, \quad (8.31)$$

$$\mathcal{Q}_1 : \mu = \theta_1, \quad \mathcal{Q}_2 : \mu = \pi - \theta_2. \quad (8.32)$$

We take  $\mathcal{R}$  to be the defect system at  $t = 0$ . The time-independent extremal curves may then be parameterized as

$$u = u(\mu), \quad \mu \in (\theta_1, \pi - \theta_2). \quad (8.33)$$

From the metric (8.31), we have that the bulk area functional is

$$\begin{aligned} A[u(\mu)] &= V_{d-2} \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{[u(\mu) \sin \mu]^{d-1}} \sqrt{\frac{u'(\mu)^2}{h[u(\mu)]} + u(\mu)^2} \\ &\equiv V_{d-2} \int_{\theta_1}^{\pi-\theta_2} d\mu \mathcal{L}_s[\mu, u(\mu), u'(\mu)]. \end{aligned} \quad (8.34)$$

where  $V_{d-2} = \int d\vec{x}$  and we denote the integrand as a Lagrangian  $\mathcal{L}_s$ . Furthermore, the boundary area functional for the intersection between  $u(\mu)$  and  $\mu = \mu_0$  is

$$\tilde{A}[u(\mu_0)] = \frac{V_{d-2}}{[u(\mu_0) \sin \mu_0]^{d-2}}. \quad (8.35)$$

And so, the entropy functional is

$$S[\mathcal{R}] = \frac{V_{d-2}}{4G_{d+1}} \min_{\gamma_{\mathcal{R}} \sim \mathcal{R}} \text{ext} \left( \int_{\theta_1}^{\pi-\theta_2} d\mu \mathcal{L}_s[\mu, u(\mu), u'(\mu)] + \frac{1}{d-2} \sum_{i=1}^2 \frac{\lambda_i}{(u_i \sin \theta_i)^{d-2}} \right), \quad (8.36)$$

where we have defined  $u_1 \equiv u(\theta_1)$  and  $u_2 \equiv u(\pi - \theta_2)$  to make the notation more compact. Note that the DGP couplings  $\lambda_1$  and  $\lambda_2$  only explicitly appear in the boundary terms, although they will implicitly affect the extremal surfaces through boundary conditions.

Now, we extremize the functional in (8.36). Taking the variational derivative and integrating by parts yields (up to the fixed factor of  $\frac{V_{d-2}}{4G_{d+1}}$ )

$$\begin{aligned} \delta S \sim & \int_{\theta_1}^{\pi-\theta_2} d\mu \left[ \frac{\partial \mathcal{L}_s}{\partial u} - \frac{d}{d\mu} \left( \frac{\partial \mathcal{L}_s}{\partial u'} \right) \right] \delta u \\ & - \left[ \frac{\partial \mathcal{L}_s}{\partial u'} \Big|_{\mu=\theta_1} + \frac{\lambda_1 \sin \theta_1}{(u_1 \sin \theta_1)^{d-1}} \right] \delta u_1 + \left[ \frac{\partial \mathcal{L}_s}{\partial u'} \Big|_{\mu=\pi-\theta_2} - \frac{\lambda_2 \sin \theta_2}{(u_2 \sin \theta_2)^{d-1}} \right] \delta u_2, \end{aligned} \quad (8.37)$$

where  $\delta u_1$  and  $\delta u_2$  are first-order variations of the surface's endpoints. Setting the first term to 0 yields the Euler–Lagrange equation for  $\mathcal{L}_s$ , which we solve to write a second-order ODE for  $u(\mu)$ :

$$u'' = -(d-2)u h(u) + (d-1)u' \cot \mu \left[ 1 - \frac{\tan \mu}{2h(u)} \frac{u'}{u} + \frac{1}{h(u)} \frac{u'^2}{u^2} \right] - \left( \frac{d-5}{2} \right) \frac{u'^2}{u}. \quad (8.38)$$

The second and third terms are the boundary conditions on  $\mathcal{Q}_1$  and  $\mathcal{Q}_2$ , respectively. Taking the dynamical boundary conditions [i.e.  $\delta u_1 \neq 0$  and  $\delta u_2 \neq 0$  in (8.37)] yields

$$\frac{u'_1}{\lambda_1} = -(\sin \theta_1)h(u_1)\sqrt{\frac{u_1'^2}{h(u_1)} + u_1^2} \leq 0, \quad \frac{u'_2}{\lambda_2} = (\sin \theta_2)h(u_2)\sqrt{\frac{u_2'^2}{h(u_2)} + u_2^2} \geq 0, \quad (8.39)$$

where we have defined  $u'_1 \equiv u'(\theta_1)$  and  $u'_2 \equiv u'(\pi - \theta_2)$ . The inequalities follow from

$$\theta_i \in (0, \pi) \implies \sin \theta_i > 0, \quad u_i \in (0, u_h) \implies h(u_i) > 0, \quad u'_i \in \mathbb{R}, \quad (i = 1, 2), \quad (8.40)$$

and fix the signs of  $\frac{u'_1}{\lambda_1}$  and  $\frac{u'_2}{\lambda_2}$ . With this in mind, we rearrange these expressions to write

$$u_1'^2 = \frac{(\lambda_1 \sin \theta_1)^2 h(u_1)^2 u_1^2}{1 - (\lambda_1 \sin \theta_1)^2 h(u_1)}, \quad u_2'^2 = \frac{(\lambda_2 \sin \theta_2)^2 h(u_2)^2 u_2^2}{1 - (\lambda_2 \sin \theta_2)^2 h(u_2)}, \quad (8.41)$$

and solve for  $u'_1$  and  $u'_2$ :

$$u'_1 = -\frac{(\lambda_1 \sin \theta_1)h(u_1)u_1}{\sqrt{1 - (\lambda_1 \sin \theta_1)^2 h(u_1)}}, \quad u'_2 = \frac{(\lambda_2 \sin \theta_2)h(u_2)u_2}{\sqrt{1 - (\lambda_2 \sin \theta_2)^2 h(u_2)}}. \quad (8.42)$$

We immediately observe that the horizon  $u(\mu) = u_h$  is still a solution even with the DGP boundary conditions above. This means that the horizon is always a valid candidate for the RT surface. We will use this fact in the following discussion.

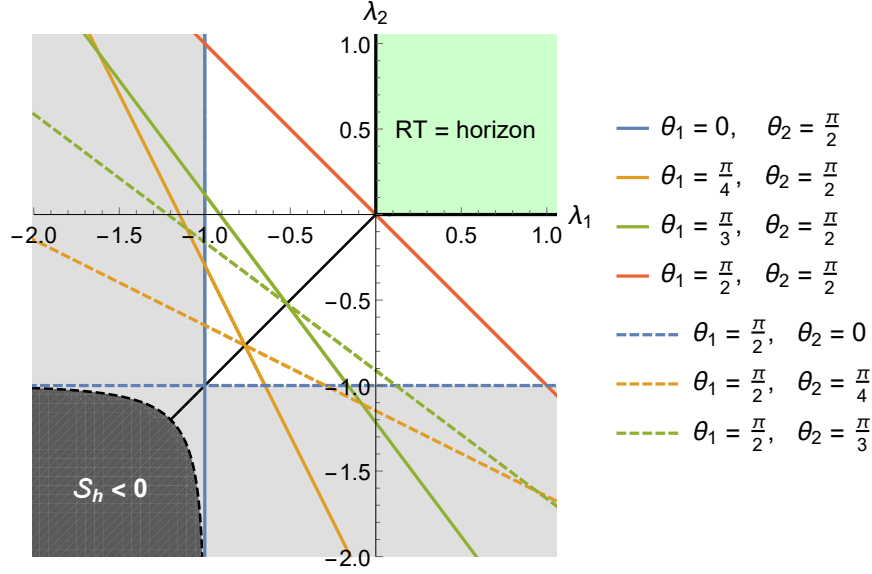


Figure 8.5: A plot of the two-brane DGP parameter space  $(\lambda_1, \lambda_2)$  for  $d = 4$ . The upper-right (green) region  $\{\lambda_1 \geq 0\} \cap \{\lambda_2 \geq 0\}$  cannot be excluded by our swampland criteria because the horizon is always the (positive-entropy) RT surface. The colored lines demarcate the boundaries of regions excluded by the positive-horizon-entropy requirement for fixed  $\theta_1$  and  $\theta_2$ , which parameterize the branes as per (8.31)–(8.32). We exclude points below these lines. Additionally, the constraints have a reflection symmetry seen by swapping the labeling of the branes ( $1 \leftrightarrow 2$ ). We see that thinner wedges yield more restrictive bounds on  $(\lambda_1, \lambda_2)$ . The intersection of all  $(\theta_1, \theta_2)$ -dependent regions describes the couplings for which the horizon entropy is negative for all brane angles. This is shown in black and lies entirely within the unphysical  $G_{\text{eff}} < 0$  region, with latter being shaded grey.

### 8.3.3 Analytic criteria for DGP couplings

We now take an analytic approach to our exploration of the parameter space of DGP couplings  $(\lambda_1, \lambda_2)$  by employing the RT prescription. By noting that the horizon of the black string is always at least a candidate surface, we can make the following assertions:

- (1) Theories in which the RT prescription selects the horizon and computes a positive entropy cannot be excluded.
- (2) The horizon entropy upper bounds the RT entropy.

In this section, we will use these statements to explore the DGP parameter space without numerically constructing extremal surfaces. The results are summarized in Figure 8.5. Furthermore, we reiterate that the region  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  is already deemed unphysical by both our wrong-sign-action argument and an entropy argument by Chen et al. (2020b).

**Entanglement does not constrain  $\{\lambda_1 \geq 0\} \cap \{\lambda_2 \geq 0\}$**  We first consider the case of nonnegative DGP couplings on both branes. In this case, we find that our swampland criteria do not apply. Specifically, the entropy functional (8.36) is always positive, and we can and will also show that the RT surface is always the horizon, which is consistent with CWI. To do the latter, we study locally smooth extremal surfaces assumed to not be the horizon, finding that they are always UV-divergent and, thus, always subdominant to the (UV-finite) horizon in the functional integral that computes the entropy. In other words, we find that non-horizon candidates (i.e. finite extremal surfaces) for the RT prescription require either DGP coupling being negative.

Suppose that we have an extremal surface that is anchored to brane 1 at a point  $u = u_1 < u_h$  and brane 2 at a point  $u = u_2 < u_h$ . Then, from the boundary

conditions (8.42), we have that

$$u'_1 < 0, \quad u'_2 > 0. \quad (8.43)$$

Pictorially in Figure 8.4, such an extremal surface bends “away” from the horizon at both branes. Thus,  $u'(\mu)$  must change sign at some  $\mu = \mu_0$  somewhere in the bulk, i.e. there exists  $\mu_0 \in (\theta_1, \pi - \theta_2)$  for which

$$u'(\mu_0 - \epsilon) < 0, \quad u'(\mu_0 + \epsilon) > 0, \quad (8.44)$$

for all sufficiently small  $\epsilon > 0$ . Since the metric is non-singular, the solutions to the Euler–Lagrange equation do not accommodate finite discontinuities in  $u'$ . We also observe that  $u'(\mu_0)$  cannot vanish; if it does, then the equations of motion (8.38) give

$$u''(\mu_0) = -(d-2)u(\mu_0)h[u(\mu_0)], \quad (8.45)$$

where  $u(\mu_0) \in (0, u_1)$  and thus  $h[u(\mu_0)] > 0$ . It then follows that  $u''(\mu_0) < 0$ , but this is inconsistent with (8.44).

The only remaining possibility is that  $u'$  has an infinite discontinuity at  $\mu = \mu_0$ , i.e. that taking  $\epsilon \rightarrow 0$  in (8.44) respectively yield  $-\infty$  and  $+\infty$ . The resulting surface would not be globally smooth, but it would consist of two locally smooth components, where each reaches the defect at  $u = 0$ . Nonetheless, we can discard such a surface because it has a UV divergence in its entropy and thus is larger than the (UV-finite) horizon entropy.

We emphasize that the argument for this statement goes through if either DGP coupling is zero. Indeed, we had already proven the statement for  $\lambda_1 = \lambda_2 = 0$  in Geng et al. (2021b). For just one positive and one zero DGP coupling, we can set  $\lambda_1 = 0$  without loss of generality. In this case, we must exploit the fact that  $u''(\theta_1) = 0$  by the equation of motion (8.38). Thus, for some small  $\epsilon \ll 1$ ,  $u'(\theta_1 + \epsilon) < 0$  and, because we still have  $u'_2 > 0$ , there exists a  $\mu = \mu_0$  where  $u'(\mu)$  changes sign. However, we reiterate that such a sign change cannot happen without an infinite discontinuity. This completes the proof.



To summarize, for nonnegative DGP couplings  $\lambda_1, \lambda_2 \geq 0$ , the RT surface is always the horizon because the only other extant extremal surfaces are UV-divergent, so we never find a violation of CWI in this part of parameter space. Our search for CWI-violating theories in the swampland must involve at least one negative DGP coupling.

**Positivity of entropy constrains  $\{\lambda_1 < 0\} \cup \{\lambda_2 < 0\}$**  For a  $(d+1)$ -dimensional AdS black string with two branes, the entropy *density*, which we define as the functional (8.36) in units of the prefactor  $\frac{V_{d-2}}{4G_{d+1}}$ , is computed to be

$$S_h = \frac{1}{u_h^{d-2}} \left[ \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu} + \frac{1}{d-2} \left( \frac{\lambda_1}{\sin^{d-2} \theta_1} + \frac{\lambda_2}{\sin^{d-2} \theta_2} \right) \right]. \quad (8.46)$$

It is evident that if either DGP coupling is sufficiently negative, then the horizon entropy is negative. While this is mathematically possible, we argue that this is unphysical.

We briefly reiterate the main points regarding horizon-entropy positivity discussed in Section 8.2. Recall that the RT prescription computes a von Neumann entropy, and such an entropy must be positive. In a typical AdS calculation with no branes, applying the RT prescription to a CFT subregion gives a UV-divergent result, since the surface reaches the conformal boundary. In the field theory, the divergence comes about because we have split the degrees of freedom. We often subtract this infinity via some ultraviolet renormalization procedure. The renormalized entropy can be negative (depending on the scheme), but this does not violate the original positivity requirement.

Meanwhile in a two-brane setup, the RT prescription gives UV-finite answers because we never split a connected CFT region. Another way to say this is that the algebra of observables on a defect system is type I, rather than type III. However, such an entropy computed from the gravity side may concernedly negative, which does not happen when there are no branes.

With that said, recall that the black-string horizon is extremal for all DGP couplings, since it solves (8.38) and (8.41). Thus, the horizon entropy upper-bounds the entropy  $S$  from RT and

$$\mathcal{S}_h < 0 \implies S \leq \frac{V_{d-2}}{4G_{d+1}} \mathcal{S}_h < 0. \quad (8.47)$$

In other words, a negative horizon entropy implies a negative RT entropy.

Now we rule out parts of  $(\lambda_1, \lambda_2)$  parameter space in which either DGP coupling is negative enough to induce a negative horizon entropy. First, consider fixed brane angles  $\theta_1$  and  $\theta_2$ . The excluded region is then determined by a linear constraint:

$$\begin{aligned} \mathcal{S}_h &= \frac{1}{u_h^{d-2}} \left[ \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu} + \frac{1}{d-2} \left( \frac{\lambda_1}{\sin^{d-2} \theta_1} + \frac{\lambda_2}{\sin^{d-2} \theta_2} \right) \right] < 0 \\ \iff \frac{\lambda_1}{\sin^{d-2} \theta_1} + \frac{\lambda_2}{\sin^{d-2} \theta_2} &< -(d-2) \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu}. \end{aligned} \quad (8.48)$$

The “bounding line” of this region depends on the values of  $\theta_1$  and  $\theta_2$ . Furthermore, this constraint depends on the dimension parameter  $d$ . As such, for concreteness we set  $d = 4$ , discussing general  $d$  in Appendix 8.B. For  $d = 4$ , the horizon’s entropy density is

$$\begin{aligned} \mathcal{S}_h &= \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{u_h^2 \sin^3 \mu} + \frac{\lambda_1}{2(u_h \sin \theta_1)^2} + \frac{\lambda_2}{2(u_h \sin \theta_2)^2} \\ &= \frac{1}{2u_h^2} \sum_{i=1}^2 \left[ \csc \theta_i \cot \theta_i - \log \left( \tan \frac{\theta_i}{2} \right) + \lambda_i \csc^2 \theta_i \right]. \end{aligned} \quad (8.49)$$

We can then write the unphysical region as

$$\lambda_2 < - \left( \frac{\sin \theta_2}{\sin \theta_1} \right)^2 \lambda_1 - \sin^2 \theta_2 \sum_{i=1}^2 \left[ \csc \theta_i \cot \theta_i - \log \left( \tan \frac{\theta_i}{2} \right) \right]. \quad (8.50)$$

Let us now explore more  $\theta$ -independent bounds. For convenience, we define notation for the fixed- $(\theta_1, \theta_2)$  unphysical regions in  $(\lambda_1, \lambda_2)$  space and the corresponding bounding lines:

$$\mathcal{B}(\theta_1, \theta_2) \equiv \{ \lambda_2 < L_{\mathcal{B}}(\lambda_1, \theta_1, \theta_2) \}, \quad (8.51)$$

$$L_{\mathcal{B}}(\lambda_1; \theta_1, \theta_2) \equiv - \left( \frac{\sin \theta_2}{\sin \theta_1} \right)^2 \lambda_1 - \sin^2 \theta_2 \sum_{i=1}^2 \left[ \csc \theta_i \cot \theta_i - \log \left( \tan \frac{\theta_i}{2} \right) \right]. \quad (8.52)$$

We can construct two  $\theta$ -independent regions: the *intersection*  $\cap \mathcal{B}$  and the *union*  $\cup \mathcal{B}$  of  $\mathcal{B}(\theta_1, \theta_2)$  over all valid choices of brane angles. We do so in that order.

The intersection describes the part of parameter space in which the horizon entropy is always negative regardless of the brane angles. Thus, the intersection is in the swampland. To find this region, it is helpful to define the contribution of an individual brane to the entropy:

$$\mathcal{S}_h^{(i)} \equiv \frac{1}{2} \left[ \csc \theta_i \cot \theta_i - \log \left( \tan \frac{\theta_i}{2} \right) + \lambda_i \csc^2 \theta_i \right]. \quad (8.53)$$

$\mathcal{S}_h^{(i)}$  has a positive singularity as  $\theta_i \rightarrow 0^+$  for  $\lambda_i \geq -1$ . So, if either DGP coupling is at or above  $-1$ , then we can get a positive horizon entropy by taking the corresponding angle to be sufficiently small. This means that

$$\cap \mathcal{B} \subset \{\lambda_1 < -1\} \cap \{\lambda_2 < -1\}. \quad (8.54)$$

Meanwhile, for  $\lambda_i < -1$  the sign of the singularity flips, and  $\mathcal{S}_h^{(i)}$  also has a negative singularity as  $\theta_i \rightarrow \pi^-$ . Thus, the profile of  $\mathcal{S}_h^{(i)}$  develops a maximum in  $\theta_i$  at

$$\cos \theta_{i*} = -\frac{1}{\lambda_i}. \quad (8.55)$$

We can explicitly compute the value of this maximum:

$$\mathcal{S}_{i*} \equiv \mathcal{S}_h^{(i)} \Big|_{\theta_i=\theta_{i*}} = \frac{\lambda_i}{2} - \frac{1}{4} \log \left( \frac{\lambda_i + 1}{\lambda_i - 1} \right). \quad (8.56)$$

Up to a positive factor, the total horizon entropy at a point in  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  is upper-bounded by the sum of these maxima, and so

$$\mathcal{S}_{1*} + \mathcal{S}_{2*} < 0 \iff \mathcal{S}_h < 0, \quad \forall (\theta_1, \theta_2). \quad (8.57)$$

Hence, the couplings  $(\lambda_1, \lambda_2)$  for which  $\mathcal{S}_{1*} + \mathcal{S}_{2*} < 0$  are precisely those in  $\cap \mathcal{B}$ . We find the boundary of this region by setting the sum of maxima to 0, which is equivalent to

$$\frac{\lambda_1 + \lambda_2}{2} - \frac{1}{4} \log \left[ \left( \frac{\lambda_1 + 1}{\lambda_1 - 1} \right) \left( \frac{\lambda_2 + 1}{\lambda_2 - 1} \right) \right] = 0. \quad (8.58)$$

This produces a curve in  $(\lambda_1, \lambda_2)$  space bounding the region of couplings for which horizon entropy is always negative, and thus we write

$$\bigcap \mathcal{B} = \left\{ \frac{\lambda_1 + \lambda_2}{2} - \frac{1}{4} \log \left[ \left( \frac{\lambda_1 + 1}{\lambda_1 - 1} \right) \left( \frac{\lambda_2 + 1}{\lambda_2 - 1} \right) \right] < 0 \right\}. \quad (8.59)$$

The takeaway is that combinations of two *very* negative DGP couplings should be seen as be highly pathological, particularly since this regime is also ruled out by other arguments (namely our wrong-sign-action argument and the RT bubbles of Chen et al. (2020b)).

Next, the union only tells us where at least one set of brane angles furnishes a negative horizon entropy. Put another way, the complement of the union consists of the DGP parameters that are unaffected by the positivity of horizon entropy, and knowing the union helps visualize which part of the parameter space might conceivably be subjected to swampland constraints. With that in mind, we now compute the union. We argue that<sup>16</sup>

$$\bigcup \mathcal{B} = \{\lambda_1 < -1\} \cup \{\lambda_2 < -1\} \cup \{\lambda_1 + \lambda_2 < 0\}. \quad (8.60)$$

Notably, (8.60) contains positive-negative combinations of DGP couplings, including ones which yield  $G_{\text{eff}} > 0$  on both branes.

To prove (8.60), we first check that  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\} \cup \{\lambda_1 + \lambda_2 < 0\} \subset \bigcup \mathcal{B}$ . To do so, we compute (8.51) in the following three limits:

$$\mathcal{B} \left( \frac{\pi}{2}, \frac{\pi}{2} \right) = \{\lambda_1 + \lambda_2 < 0\}, \quad (8.61)$$

$$\mathcal{B} \left( 0, \frac{\pi}{2} \right) = \{\lambda_1 < -1\}, \quad (8.62)$$

$$\mathcal{B} \left( \frac{\pi}{2}, 0 \right) = \{\lambda_2 < -1\}. \quad (8.63)$$

Now, we want to show that points that are not in the union of (8.61)–(8.63) also never accommodate negative horizon entropies. Since all of the fixed- $(\theta_1, \theta_2)$  bounds

---

<sup>16</sup>This turns out to be the union of all fixed- $(\theta_1, \theta_2)$  constraints in any number of spacetime dimensions. In contrast, the intersection depends on  $d$ . We show this in Appendix 8.B.

are linear in DGP couplings, we just need to confirm three statements for the bounding lines  $L_{\mathcal{B}}$  (8.52):

- (i) any bounding line must have negative slope;
- (ii) any bounding line must intersect  $\lambda_1 = 0$  at some  $\lambda_2 < 0$ ; and,
- (iii) any bounding line which intersects  $\lambda_1 + \lambda_2 = 0$  once does so in  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$ .

These conditions ensure that any point contained in some excluded region  $\mathcal{B}(\theta_1, \theta_2)$  will be in the union of (8.61)–(8.63), thus proving (8.60). (i) is immediate from (8.52):

$$\frac{\partial L_{\mathcal{B}}}{\partial \lambda_1} = - \left( \frac{\sin \theta_2}{\sin \theta_1} \right)^2 < 0, \quad (8.64)$$

To check the other two statements, it helps to fix  $\theta_2 = \theta$  and set  $\theta_1 = \pi - \theta - \delta$ , where  $\delta \in (0, \pi - \theta)$ . We then write

$$L_{\mathcal{B}}(0, \pi - \theta - \delta, \theta) = -\sin^2 \theta F(\theta, \delta), \quad (8.65)$$

$$F(\theta, \delta) \equiv \csc \theta \cot \theta - \csc(\theta + \delta) \cot(\theta + \delta) - \log \left[ \tan \left( \frac{\theta}{2} \right) \cot \left( \frac{\theta + \delta}{2} \right) \right]. \quad (8.66)$$

The derivative of  $F$  with respect to  $\delta$  is  $2 \csc^3(\theta + \delta)$ , which is positive for all valid brane angles. Furthermore,  $F(\theta, 0) = 0$ , so we infer that  $F > 0$ . Thus, (8.65) is negative, confirming (ii).

To check (iii) directly,<sup>17</sup> we start by denoting the intersection between a bounding line  $L_{\mathcal{B}}$  and  $\lambda_1 + \lambda_2 = 0$  by  $(\lambda_1^*, \lambda_2^*)$ . We then use this intersection constraint to solve for  $\lambda_1^*$  as a function of  $\theta$  and  $\delta$ :

$$-\lambda_1^* = L_{\mathcal{B}}(\lambda_1^*, \pi - \theta - \delta, \theta) \implies \lambda_1^*(\theta, \delta) = \frac{F(\theta, \delta)}{\csc^2 \theta - \csc^2(\theta + \delta)}. \quad (8.67)$$

---

<sup>17</sup>We use a more indirect method in Appendix 8.B which can be applied in any number of dimensions.

Over the full domain of  $\delta$ , this function has the endpoint values

$$\lambda_1^*(\theta, 0) = \sec \theta, \quad \lambda_1^*(\theta, \pi - \theta) = -1, \quad (8.68)$$

We also observe that the derivative of  $\lambda_1^*$  with respect to  $\delta$  is positive for  $\theta \in (0, \pi)$  and  $\delta \in (0, \pi - \theta)$ . So now if we assume  $\theta > \frac{\pi}{2}$ , then we have enough to deduce that  $\lambda_1^* < -1$ . However, if  $\theta < \frac{\pi}{2}$ , then we note that there is a pole at  $\delta = \pi - 2\theta$ . Nonetheless, the derivative is still positive, so it turns out that

$$\delta \in (0, \pi - 2\theta) \implies \lambda_1^* > 1, \quad \delta \in (\pi - 2\theta, \pi - \theta) \implies \lambda_1^* < -1. \quad (8.69)$$

The points for which  $\lambda_1^* < -1$  are in  $\{\lambda_1 < -1\}$ . Meanwhile, since  $\lambda_2^* = -\lambda_1^*$ , the points for which  $\lambda_1^* > 1$  are in  $\{\lambda_2 < -1\}$ . Thus, (iii) is true.

With (i)–(iii) in hand, we can conclude that any bounding line  $L_{\mathcal{B}}(\lambda_1; \theta_1, \theta_2)$  must be contained within the union of (8.61)–(8.63), and so (8.60) is true. We reiterate that  $\bigcup \mathcal{B}$  includes DGP parameters for which the effective Newton constant on each brane is positive, unlike  $\bigcap \mathcal{B}$ .

### 8.3.4 Numerical violations of causal wedge inclusion

From the equation of motion (8.38) and boundary conditions (8.42), we can numerically solve for the (static) extremal surfaces for particular choices of brane angles  $(\theta_1, \theta_2)$  and DGP parameters  $(\lambda_1, \lambda_2)$ . We may then check whether or not there exists an extremal surface with total entropy strictly smaller than the horizon. This would be a violation of causal wedge inclusion (CWI) and put that combination of parameters in the swampland.

We first obtain a plethora of numerical extremal surfaces. Specifically, we shoot from brane 1 with fixed angle  $\theta_1$  and over a range of initial parameters  $u_1$  and couplings  $\lambda_1$ . Note that the value of  $u'_1$  is set by (8.42). We then filter the list of these solutions down to those consistent with a particular value of  $\theta_2$  at the second brane. For these remaining points, we compute the entropy density  $\mathcal{S}_{\text{ext}}$  numerically

and the entropy density of the horizon  $\mathcal{S}_h$  analytically (8.49). Then, we repeat the above process by shooting from brane 2. In the end, we have a large list of points  $(\lambda_1, \lambda_2)$  on a fixed- $(\theta_1, \theta_2)$  plane for which there is a non-horizon extremal surface, along with the associated entropies.

With that list in hand, we delete any points which seem pathological, such as those with large (and thus potentially numerically unstable) areas. Guided by both the positivity of horizon entropy  $\mathcal{S}_h > 0$  and possible cases (8.16)–(8.17) diagnosing whether or not we have a violation of CWI, we sort the points  $(\lambda_1, \lambda_2)$  in this list into three subsets:

$$(\lambda_1, \lambda_2) \text{ for which } \mathcal{S}_{\text{ext}} > \mathcal{S}_h > 0, \quad (8.70)$$

$$(\lambda_1, \lambda_2) \text{ for which } \mathcal{S}_h < 0, \quad (8.71)$$

$$(\lambda_1, \lambda_2) \text{ for which } \mathcal{S}_h > \mathcal{S}_{\text{ext}} \text{ and } \mathcal{S}_h > 0. \quad (8.72)$$

(8.70) consists of points that cannot be ruled out by the entropic swampland conditions considered in this paper, but we emphasize that these points may still appear in the pathological region with at least one negative effective Newton constant  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$ . Meanwhile, the regions (8.71) and (8.72) are unphysical on the fixed- $(\theta_1, \theta_2)$  plane. Specifically, the points in (8.71) must be contained within the analytically excluded region  $\mathcal{B}(\theta_1, \theta_2)$  defined in (8.51). Points in (8.72) are not in the analytically excluded region since they furnish positive horizon entropy, but they still violate CWI.<sup>18</sup>

Note that a pair of DGP couplings in satisfying (8.71) or (8.72) for some combination of brane angles need not always be in the swampland. In fact, a pair  $(\lambda_1, \lambda_2)$  is only generically be in the swampland (based on our criteria) if it is in the union of (8.71) with (8.72) and taken over *all*  $(\theta_1, \theta_2)$ . The excluded points presented

---

<sup>18</sup>*A priori* this set may also include more strongly pathological points for which  $\mathcal{S}_{\text{ext}} < 0$ . Not only would such points violate CWI, but they would also have a negative entropy as determined by the RT prescription.

below (Figure 8.6) should be interpreted with a specific combination of brane angles in mind.

The primary fact that want to demonstrate is that this last set of points overlaps with the part of the  $(\lambda_1, \lambda_2)$  parameter space that is not already excluded by either the positivity of horizon entropy or the positivity of the brane’s effective Newton constant, i.e. that (8.72) includes points for which [recalling (8.51)–(8.52)]

$$\lambda_2 > L_{\mathcal{B}}(\lambda_1; \theta_2, \theta_2) \quad \text{and} \quad \lambda_1, \lambda_2 > -1. \quad (8.73)$$

We can be more ambitious and ask for a complete understanding of the DGP couplings that allow for non-horizon extremal surfaces, as well as how such a set partitions into subsets (8.70)–(8.72). While our approach does this to some extent, we emphasize that our numerics provide only evidence for such information, rather than proof. However, the numerics are still sufficient to exclude a region of parameter space within (8.73).

**Inspecting the found points** For various combinations of fixed  $(\theta_1, \theta_2)$ , we present the resulting sets of points in Figure 8.6. We observe points in (8.70) (red), (8.71) (blue), and (8.72) (green). By definition, the blue points are below the bounding line (8.52) along which  $\mathcal{S}_h = 0$ , while the red and green points are above this bounding line. Based solely on these numerics, we make several observations.

First, note that the blue points are in the region  $\{\lambda_1 < -1\} \cap \{\lambda_2 < -1\}$ , i.e. they appear at couplings for which both effective Newton constants are negative. This suggests that any couplings which allow negative horizon entropy and a non-horizon extremal surface are deeply pathological, since they violate multiple swampland criteria. Furthermore, these points are completely disconnected from the others.

The red and green points appear to form a single connected region above the bounding line  $L_{\mathcal{B}}$  (8.52). Most of the red points, which we reiterate correspond to



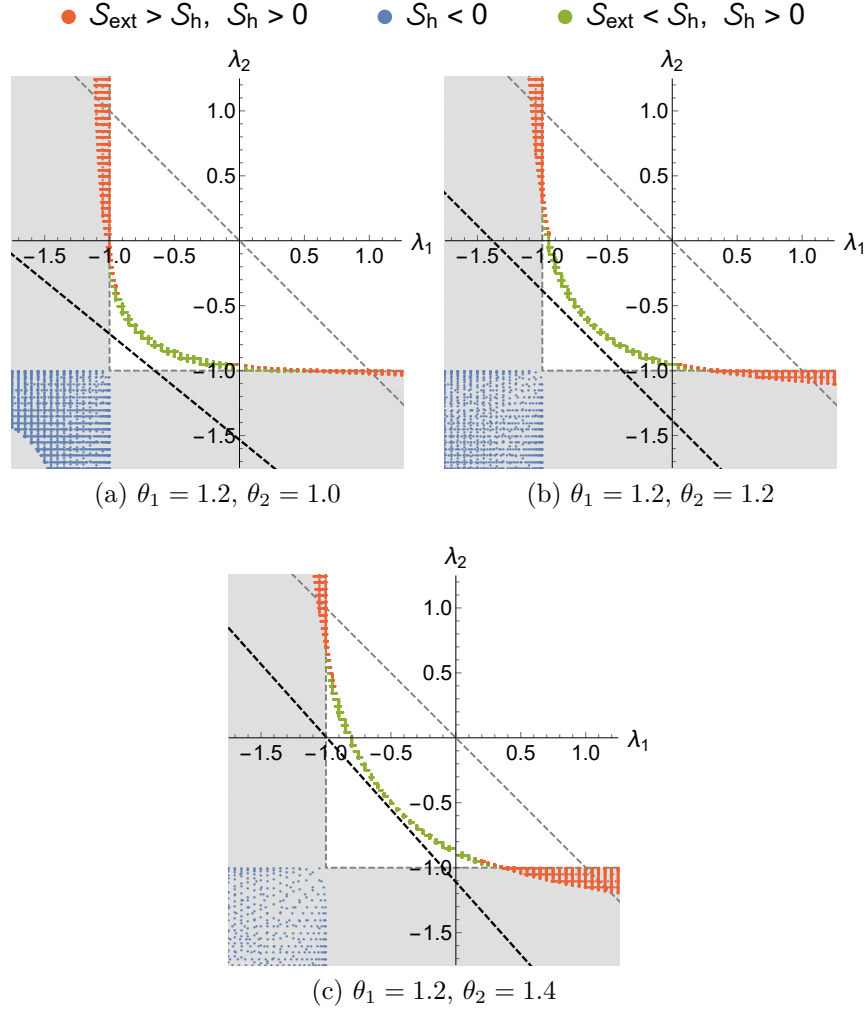


Figure 8.6: The points in  $(\lambda_1, \lambda_2)$  parameter space for which we find non-horizon extremal surfaces, for fixed angles  $\theta_1$  and  $\theta_2$ . The colors correspond to (8.70) (red), (8.71) (blue), and (8.72) (green). The thick dashed lines are the  $(\theta_1, \theta_2)$ -dependent bounding lines (8.52) below which  $S_h < 0$ , while the light dashed line is  $\lambda_1 + \lambda_2 = 0$ . We have also shaded the region in which either effective Newton constant is negative.

couplings that do not violate CWI, are pathological anyway because they furnish at least one negative effective Newton constant, i.e.  $\lambda_1 < -1$  or  $\lambda_2 < -1$ . Interestingly, some are not sick, and these points are near but notably not arbitrarily close to the  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  region. However, we caution that this part of the parameter space near the boundary of  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  is subject to higher numerical errors and such points may be artifacts of those errors. We will elaborate on this point later.

Meanwhile, the green points designating CWI-violating couplings appear within the region  $\{\lambda_1 > -1\} \cap \{\lambda_2 > -1\} \cap \{\lambda_1 + \lambda_2 < 0\}$ .<sup>19</sup> So for a given  $\theta_1$  and  $\theta_2$ , CWI excludes points not already ruled out by the positivity of horizon entropy and effective Newton constant. Thus, CWI provides another nontrivial swampland criteria for braneworlds.

We should say that these numerics do not suggest that CWI is a *stronger* condition than the positivity of horizon entropy. This would only be a true statement if the set of points ruled out by CWI contains the region  $\mathcal{B}(\theta_1, \theta_2)$  (8.51) in which  $\mathcal{S}_h < 0$ , but our numerics only find non-horizon extremal surfaces within a subset of  $\mathcal{B}(\theta_1, \theta_2)$ .

How does this story depend on the combination of angles? While we have sampled and shown only a few combinations as shown in Figure 8.6, we can make some conjectures based on the observed patterns. For one thing, the green and red points are above the bounding line  $L_{\mathcal{B}}$  but seem to follow the line as the angles are dialed. Numerically the CWI-excluded points are above the bounding line by a gap of order 0.1 at most in the sampled angles in Figure 8.6. For combinations of brane angles that yield a “wider” wedge, i.e. smaller  $\theta_1 + \theta_2$ , this line resides within the region  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$ , with the limit  $\theta_1, \theta_2 \rightarrow 0$  completely assuring positivity of entropy for all  $(\lambda_1, \lambda_2)$ . So if the CWI-excluded points must move with

---

<sup>19</sup>Some green points are hidden below the red points, but even those green points are within this region.

the bounding line, then in the limit  $\theta_1, \theta_2 \rightarrow 0$  we would expect CWI to become redundant with the positivity of the effective Newton constants.

It would be worthwhile to search for CWI-violating surfaces in the regime of small brane angles. However, taking both  $\theta_1, \theta_2 \rightarrow 0$  gives rise to instabilities in our numerics, so we are unable to confidently rule out points in this regime of parameter space. A more refined approach to searching for the CWI-violating surfaces is necessary. Alternatively it may be possible to organize the contributions of higher-derivative terms to the entropy into an ordered series in this limit, which would allow for analytic statements.

For “thinner” wedges, the bounding line intersects the region with positive effective Newton constants, i.e. increasing  $\theta_1 + \theta_2$  also pushes  $L_{\mathcal{B}}$  up and to the right in the  $(\lambda_1, \lambda_2)$  parameter space. This also pushes the red and green points up, and such combinations of brane angles is when CWI distinguishes itself from the other swampland conditions.

**Possible expansion of excluded zone** The key takeaway is that CWI certainly excludes DGP couplings that otherwise seem okay. The green points in Figure 8.6 are certainly sick. However, there are various nontrivial features of the plots. In particular, the switch between the green region and red region seems somewhat arbitrary, and there is also a gap between the red/green region and the bounding line  $L_{\mathcal{B}}$ . This latter feature in particular suggests that CWI does not actually lower-bound the DGP couplings.

To be more precise, on the fixed  $(\theta_1, \theta_2)$ -slices shown in Figure 8.6, the excluded zone is both not convex and consists of two disconnected pieces. But we reiterate that the disconnectedness is not expected to be a feature of the excluded manifold in the full four-dimensional parameter space. This is because of our expectation that in the  $\theta_1, \theta_2 \rightarrow 0$  limit, both CWI and positivity of entropy become subsumed by the condition that the effective Newton constants are positive, which by itself excludes a

connected region.

We caution that the numerics could be subject to some slight instability originating from the shooting procedure, particularly when dealing with surfaces which are numerically close to the horizon. Such surfaces appear in our numerics near the boundary of the  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  region. Furthermore, in our approach we only completely specify one of the DGP couplings, whereas we read off the other coupling from a list of solutions. Thus, some of the green or red points are only really specified up to some error, and if the difference between  $\mathcal{S}_h$  (computed analytically) and  $\mathcal{S}_{\text{ext}}$  (computed numerically) is within the margin of error of the shooting then the switch between red points and green points is also ambiguous. However, we emphasize that our numerics are relatively stable against changes in precision, and so we believe these errors to be insignificant for most couplings—particularly for green points that are further from the red region in Figure 8.6.

It should be possible to further improve on our numerics. We emphasize that our primary goal was to find points excluded only by CWI. We have accomplished that much by considering a range of angles in which the numerics are under control. However, if one wants to better understand how the excluded zone is embedded in parameter space [for example, the shape of the excluded points in  $(\lambda_1, \lambda_2, \theta_1)$  keeping  $\theta_2$  fixed], then it would be necessary to use more refined approaches.

We also do not search for potential extremal surfaces that go into the black-string interior (Figure 8.7) and thus correspond to a time dependence in the entanglement entropy for the one-sided defect. Such surfaces require that  $\lambda_1 < 0$  and  $\lambda_2 < 0$  due to their shape, but they are harder to construct because they are not  $\mathbb{Z}_2$  symmetric with respect to the maximally extended geometry.<sup>20</sup> Nonetheless, these would also violate causal wedge inclusion and could in principle be used to rule out theories, pos-

---

<sup>20</sup>One may consider surfaces that are anchored to the branes, dive into the horizon, and satisfy this  $\mathbb{Z}_2$  symmetry as in Miao (2022a). However, those surfaces contribute to the entropy functional of the *purified* defect system. This entropy is 0, and the corresponding bulk RT surface “shrinks” onto the brane’s horizon Geng (2023).

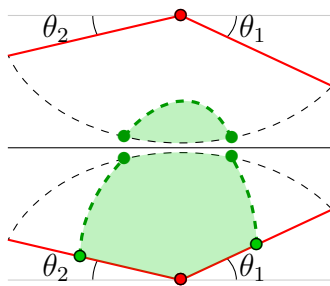


Figure 8.7: A cartoon of a possible time-dependent entangling surface in the two-brane setup with DGP couplings, shown on the  $t = 0$  slice of the maximally extended black-string geometry, in which the horizons are identified. This surface would violate causal wedge inclusion—we can see that the entanglement wedge (in green) does not include the full exterior of either side and thus never includes the causal wedge of either defect.

sibly filling the gap between the CWI-excluded and positive-entropy-excluded regions observed in Figure 8.6.

However, it is possible that such surfaces are not even mathematically feasible. In AdS black geometries without branes, entanglement surfaces akin to Figure 8.7 and anchored to the conformal boundary are not allowed because horizons in those settings are “extremal surface barriers” Engelhardt and Wall (2014). So, the only way to get time-dependent surfaces without branes is to compute the entropy of a boundary region supported on both sides of the maximally extended geometry (cf. Hartman and Maldacena (2013)). However, it is not clear that such a mathematical obstruction exists for surfaces ending on KR branes. If it does, then it would be a mathematical statement—rather than just a physical one—that the entanglement entropy of the one-sided defect geometry can only follow a trivial (time-independent) Page curve.

## 8.4 Discussion

To summarize, we have explored how entanglement structure constrains effective field theories of gravity on branes in AdS. Our core philosophy is that the

AdS/CFT dictionary, being a feature of UV physics, gives rise to swampland criteria that may be implemented in semiclassical holographic models. As a case study, we have used this idea to put nontrivial constraints—both analytic and numerical—on the RS + DGP braneworld model of gravity.

We should emphasize that the end-of-the-world branes we consider in this paper are of a particular class. They are “constant-angle” branes (8.32) embedded in a bulk AdS space (8.31). In principle, one may consider other foliations of the bulk into KR branes. This would yield more configurations, and in principle holography could be used to put different constraints on higher-derivative brane couplings. We leave this to future explorations.

Another notable caveat is that we only consider one particular two-derivative correction to the brane action—the DGP term. Even at the level of two derivatives in the effective theory, we can write additional independent terms proportional to squares of extrinsic curvature. Schematically, this correction takes the form

$$I_{K^2} = \int_{\text{brane}} (aK^2 + bK^{\mu\nu}K_{\mu\nu}). \quad (8.74)$$

So in truth, general two-brane setups with actions truncated at two derivatives occupy an eight-dimensional parameter space parameterized as  $(\theta_1, \theta_2, \lambda_1, \lambda_2, a_1, a_2, b_1, b_2)$ . Fortunately as found by Lee et al. (2022), (8.74) does not affect the embedding equations (8.23), so the shapes of the branes are only determined by the RS tensions and DGP couplings. However, (8.74) may alter the  $(d - 2)$ -dimensional area terms in the entropy functional (8.30). To see how, one would need to apply the recipe of Chen et al. (2020b) [specifically their equation (A.7)] to (8.74).

Our goal in this paper was merely to show the possibility that a swath of effective brane-localized theories can be excluded on general physical principles. We do this by restricting to a slice of theory space where the terms (8.74) (and all higher couplings) are dialed to zero. We leave exploration of other higher-derivative couplings to future work.

**Trivial massless Page curves** We briefly comment on the significance of these results to the recent discourse on the black hole information paradox. The typical ( $d > 2$ ) higher-dimensional realization of Page curves as describing the evolution of semiclassical black hole information relies on the construction with one-KR-brane embedded in an ambient AdS geometry with a nontrivial blackening factor. The first such construction was numerical Almheiri et al. (2019d), and a simple analytical model was provided by Geng and Karch (2020b) shortly after. Moreover, Geng and Karch (2020b) claimed that the existence of a nontrivial Page curve may be due to the semiclassical theory on the brane being a *massive* gravity theory, a feature that has long been ascribed to the presence of a non-gravitating bath glued on at asymptotic infinity Porrati (2002a); Porrati and Starinets (2002).

This aspect of the one-brane setup motivated our study of a two-brane setup embedded in a bulk black string geometry in Geng et al. (2021b). Here, one of the branes acts as a gravitating bath, and so the semiclassical theory on the branes maintains a massless graviton in its spectrum Kogan et al. (2001b). We had found the entropy of radiation collected in the gravitating bath to be eternally computed by the bulk horizon area. Thus, the entropy of radiation was found to be constant in time, in agreement with the claims of Laddha et al. (2020).

Recent work Miao (2022a, 2023b); Li and Miao (2023) has proposed that our prior result is changed by adding DGP terms to the two-brane setup. The basic claim is that for at least some combinations of DGP couplings (with at least one being negative) the bulk horizon no longer computes the minimum of the entropy functional, allowing for time-dependent RT surfaces. Furthermore, one may also get static RT surfaces that are not the horizon. For example in making this case, Miao (2022a) explicitly finds the latter type of surface for a particular combination

of parameters, which in our conventions are<sup>21</sup>

$$\theta_1 \approx 1.090, \quad \theta_2 \approx 0.585, \quad \lambda_1 = 0, \quad b \lambda_2 \approx -0.984. \quad (8.75)$$

However, the DGP couplings allowing for *any* such non-horizon RT surfaces in the first place are in conflict with causal wedge inclusion, since the causal wedge of one of the defect systems includes the entire corresponding side’s exterior region. That is also true of time-dependent surfaces (see Figure 8.7).<sup>22</sup> Thus, semiclassical massless-gravity theories furnishing nontrivial time-dependent Page curves or any static non-horizon bulk RT surfaces are in the swampland, and the lessons of Miao (2022a) are unphysical. See Geng (2023) for more specific discussion of this point.

In fact, our numerics loosely indicates that most (but conceivably not all) DGP couplings for which there even *exists* a UV-finite extremal surface apart from the horizon are sick, either because of a violation of causal wedge inclusion or some other physical pathology such as a negative effective Newton constant.

In conclusion, we have shown how swampland constraints can restrict the viability of AdS gravity theories with a holographic interpretation. Indeed, holography is fundamentally a feature of UV physics, and so care must be applied in assuming that it holds semiclassically to avoid misleading conclusions. Our work here focuses on entanglement structure in braneworld models, with an eye towards their recent application to the black hole information problem. The takeaway is that the existence of a holographic dictionary is an important limitation on physically consistent semiclassical models of gravity.

---

<sup>21</sup>As a consistency check, we have checked our entropies associated with these values against those of Miao (2022a). The analytic horizon entropies are precisely the same. Numerically, shooting from either brane yields  $\mathcal{S}_{\text{ext}} \approx 0.838$ , which is consistent with the value  $\mathcal{S}_{\text{ext}} \approx 0.842$  of Miao (2022a) up to small  $O(0.1\%)$  error.

<sup>22</sup>As a side note, explicitly constructing time-dependent surfaces may allow one to further expand the excluded zone presented in Figure 8.6. There may also exist brane angles for which time-dependent surfaces are the *only* way to violate CWI, but searching for such points is beyond the methods in the current paper.



## Appendix 8.A: Effective Action with RS + DGP Gravity

We briefly review the derivation of the semiclassical gravitational action on the brane with both an RS and a DGP term. We discuss this in the case of a semi-infinite  $(d + 1)$ -dimensional bulk geometry  $\mathcal{M}$  with one  $d$ -dimensional brane  $\mathcal{Q}$  (as in Chen et al. (2020b); Geng et al. (2021d); Karch et al. (2023b); Perez-Pardavila (2023b)), but the basic results carry over to the two-brane configuration. The starting point is the bulk action with brane-localized terms (leaving Gibbons–Hawking terms implicit),

$$I = \underbrace{\frac{1}{16\pi G_{d+1}} \int_{\mathcal{M}} \sqrt{-g} \left[ R + \frac{d(d-1)}{L^2} \right]}_{\text{Einstein–Hilbert}} + \underbrace{\int_{\mathcal{Q}} \sqrt{-\tilde{g}} \left( -T + \frac{1}{16\pi G_b} \tilde{R} \right)}_{\text{RS + DGP}}, \quad (8.76)$$

where  $g$  is the bulk metric,  $\tilde{g}$  is the induced brane metric, and tildes denote curvature invariants of the latter. We implicitly work with a bulk metric of the form

$$ds^2 = \frac{L^2}{u^2 \sin^2 \mu} \left[ -h(u) dt^2 + \frac{du^2}{h(u)} + d\vec{x}^2 + u^2 d\mu^2 \right], \quad (8.77)$$

where  $u > 0$ ,  $(t, \vec{x}) \in \mathbb{R}^{d-1}$ , and  $\mu \in (\theta, \pi)$ .  $\mu = \theta$  designates the position of the brane, and  $\theta$  is geometrically the angle between this brane and the conformal boundary.  $h(u)$  is some analytic function for which this metric satisfies Einstein’s equations (8.5).

We then treat the brane like a cutoff surface and integrate the Einstein–Hilbert term over the bulk radial direction. Chen et al. (2020b) computes the resulting dimensionally-reduced action to be of the form

$$I_{\text{diver}} = \frac{1}{16\pi G_{d+1}} \int \sqrt{-\tilde{g}} \left[ \frac{2(d-1)}{L} + \frac{L}{d-2} \tilde{R} + O(\tilde{R}^2) \right]. \quad (8.78)$$

There is an infinite tower of higher-derivative corrections, but in the semiclassical regime we are simply concerned with the Einstein–Hilbert terms on the brane consisting of  $\tilde{R}^0$  and  $\tilde{R}^1$  terms. Because we are dealing with just one copy of the bulk instead of two (cf. Chen et al. (2020b)), the effective action includes only one copy of  $I_{\text{diver}}$ :

$$I_{\text{eff}} = I_{\text{diver}} + \int_{\mathcal{Q}} \sqrt{-\tilde{g}} \left( -T + \frac{1}{16\pi G_b} \tilde{R} \right). \quad (8.79)$$

Plugging in (8.78) then yields

$$I_{\text{eff}} = \int_{\Omega} \sqrt{-\tilde{g}} \left[ \left( \frac{d-1}{8\pi G_{d+1} L} - T \right) \tilde{R}^0 + \left( \frac{L}{16\pi(d-2)G_{d+1}} + \frac{1}{16\pi G_b} \right) \tilde{R} + O(\tilde{R}^2) \right]. \quad (8.80)$$

We then write the action in the form

$$I_{\text{eff}} = \frac{1}{16\pi G_{\text{eff}}} \int_{\Omega} \sqrt{-\tilde{g}} \left[ \frac{(d-1)(d-2)}{\ell_{\text{eff}}^2} + \tilde{R} + O(\tilde{R}^2) \right], \quad (8.81)$$

where the  $G_{\text{eff}}$  is the effective Newton constant (i.e. the coupling of the effective Einstein–Hilbert terms) and  $\ell_{\text{eff}}$  is the effective curvature scale,<sup>23</sup> each of which are respectively identified as

$$\frac{1}{G_{\text{eff}}} = \frac{L}{(d-2)G_{d+1}} + \frac{1}{G_b}, \quad \frac{1}{\ell_{\text{eff}}^2} = \frac{2}{L^2} \left( 1 - \frac{8\pi G_{d+1} L}{d-1} T \right) \left[ \frac{(d-2)G_{d+1}}{G_b L} + 1 \right]^{-1}. \quad (8.82)$$

### Appendix 8.A.1: The $\lambda$ and $\theta$ Parameters

Throughout this paper, we study a particular class of “planar” KR branes (8.32) with  $\text{AdS}_d$  geometry embedded in a bulk  $\text{AdS}_{d+1}$  spacetime (8.31). We find it more useful to parameterize the embeddings of our branes not by  $(T, G_b)$  but by alternate parameters  $(\theta, \lambda)$ .  $\theta$  measures the angle between the brane and the conformal boundary while  $\lambda$  captures the relative contributions between the RS and DGP terms to the effective Newton constant. We now define the mapping between these different sets of parameters.

Taking  $\frac{1}{G_b} \rightarrow 0$  in (8.76) to remove the DGP term, we are simply left with the RS term. The effective RS Newton constant and curvature scale are then identified as

$$\frac{1}{G_{\text{RS}}} = \frac{L}{(d-2)G_{d+1}}, \quad \frac{1}{\ell_{\text{RS}}^2} = \frac{2}{L^2} \left( 1 - \frac{8\pi G_{d+1} L}{d-1} T \right). \quad (8.83)$$

---

<sup>23</sup>We emphasize that this effective curvature scale is not the same as the induced curvature radius of the brane read off the induced metric, e.g. (8.20).

The RS Newton constant partly contributes to the effective Newton constant of the brane with RS + DGP terms (8.82). This motivates us to define a dimensionless parameter,

$$\lambda \equiv \frac{G_{\text{RS}}}{G_{\text{b}}}, \quad (8.84)$$

describing the relative strength of these two terms. The effective Newton constant is then

$$\frac{1}{G_{\text{eff}}} = \frac{1}{G_{\text{RS}}}(1 + \lambda). \quad (8.85)$$

In particular, we have a negative effective Newton constant if  $\lambda < -1$ .

Next, we write the dynamical boundary condition of the brane with a DGP term. In RS, recall that it is (8.6). More generically, the boundary condition takes the form

$$K_{\mu\nu} - K\tilde{g}_{\mu\nu} = T_{\mu\nu}, \quad (8.86)$$

where  $T_{\mu\nu}$  is the brane's stress tensor (with a factor of  $8\pi G_{d+1}$  included in the definition),

$$T_{\mu\nu} = -\frac{16\pi G_{d+1}}{\sqrt{-\tilde{g}}} \frac{\delta I_Q}{\delta \tilde{g}^{\mu\nu}}. \quad (8.87)$$

For the brane action in (8.76), this is simply

$$\begin{aligned} T_{\mu\nu} &= -8\pi G_{d+1} T \tilde{g}_{\mu\nu} - \frac{G_{d+1}}{G_{\text{b}}} \left( \tilde{R}_{\mu\nu} - \frac{1}{2} \tilde{R} \tilde{g}_{\mu\nu} \right) \\ &= -8\pi G_{d+1} T \tilde{g}_{\mu\nu} - \frac{\lambda L}{d-2} \left( \tilde{R}_{\mu\nu} - \frac{1}{2} \tilde{R} \tilde{g}_{\mu\nu} \right). \end{aligned} \quad (8.88)$$

We have used the relationship between  $G_{\text{RS}}$  and  $G_{d+1}$  to write the right-hand side in terms of  $\lambda$ . We can then solve for the tension by contracting (8.86) with the induced metric tensor. This yields

$$8\pi G_{d+1} T = \frac{d-1}{d} K + \frac{\lambda L}{2d} \tilde{R}. \quad (8.89)$$

Now by computing the curvature invariants associated with  $\mu = \theta$  in (8.77),

$$K = \frac{d}{L} \cos \theta, \quad \tilde{R} = -\frac{d(d-1)}{L^2} \sin^2 \theta. \quad (8.90)$$

we can at last write the tension as a function of  $\lambda$  and  $\theta$ :

$$T = \frac{d-1}{8\pi G_{d+1}L} \left( \cos \theta - \frac{\lambda}{2} \sin^2 \theta \right). \quad (8.91)$$

Equipped with (8.84) and (8.91), we may use the parameters  $(\theta, \lambda)$  to fix the explicit couplings  $(T, G_b)$ . We thus parameterize the effective theories in terms of the former.

### Appendix 8.A.2: Flat-Space Limits

The DGP term is often used to construct models of flat branes in flat space Dvali et al. (2000c), with the RS tension being a bare vacuum energy Clifton et al. (2012). However, our paper is about swampland criteria emerging from AdS/CFT, and so we are considering the DGP term in the context of AdS branes in AdS space as in Chen et al. (2020b). Nonetheless, to connect to braneworld phenomenology, we can ask what happens to our constraints by taking one of two flat limits.

Roughly speaking, the first limit is to take the bulk curvature radius  $L$  to be large (making the bulk cosmological constant vanish) and the brane tension to be 0. The second is a flat limit of just the effective theory on the brane in which we take the effective length scale (8.82) to be large. However, we find that the constraints from entanglement lose all power in either case.

Let us be more specific. The first flat-space limit we can consider is one in which we take

$$L \rightarrow \infty, \quad T \rightarrow 0, \quad G_{d+1}, G_b \text{-fixed}. \quad (8.92)$$

In other words, we are eliminating the cosmological constant terms in (8.76). This yields the original DGP model Dvali et al. (2000c) with flat branes in a flat bulk. Instead of (8.80), the resulting semiclassical theory on the brane is

$$I_{\text{eff}} \xrightarrow{L \rightarrow \infty, T \rightarrow 0} \frac{1}{16\pi G_b} \int \sqrt{-\tilde{g}} \left[ \tilde{R} + O(\tilde{R}^2) \right], \quad (8.93)$$

i.e. we have that  $G_{\text{eff}} = G_b$  and  $\ell_{\text{eff}} \rightarrow \infty$ . Note that in this regime, the requirement that  $G_{\text{eff}} > 0$  turns into the constraint  $G_b > 0$ .

(8.92) is perfectly compatible with the semiclassical regime of the Ryu–Takayanagi prescription  $G_{d+1} \ll L^{d-1}$ . However, by keeping  $G_{d+1}$  and  $G_b$  finite while sending  $L \rightarrow \infty$ , we are essentially imposing the limit

$$\lambda = \frac{G_{\text{RS}}}{G_b} = \frac{(d-2)G_{d+1}}{LG_b} \rightarrow 0. \quad (8.94)$$

Our entanglement-based swampland criteria, which require at least one negative DGP coupling in the two-brane setup, are no longer applicable.

Another possibility is to consider a “flat-brane” limit in which the effective cosmological constant  $\sim \frac{1}{\ell_{\text{eff}}^2}$  vanishes. This can be done by taking the following limit of the tension:

$$T \rightarrow \frac{d-1}{8\pi G_{d+1} L}. \quad (8.95)$$

For a given  $\lambda$ , we can rephrase this in terms of a limit on the brane angle  $\theta$ . From (8.91), we get (8.95) by taking one of two values for the brane angle:

$$\theta \rightarrow 0^+ \quad \text{or} \quad \theta \rightarrow \text{Cos}^{-1} \left( 1 + \frac{2}{\lambda} \right). \quad (8.96)$$

For the first brane angle, we are essentially taking the planar brane to the conformal boundary. In the case of two planar branes forming a wedge, this makes the wedge wider. As seen in the main text however, the excluded points would then be pushed into the  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  region, in which  $G_{\text{eff}} < 0$  on at least one of the branes. Meanwhile, the second brane angle only makes sense when we assume  $\lambda < -1$ . So for this flat limit, we see that the swampland constraints coming from entanglement are simply subsumed by the requirement that  $G_{\text{eff}} > 0$  and thus lose all power.

However, we should emphasize that we are working with a particular slicing whereby the brane is a constant- $\mu$  hypersurface in (8.77). It is conceivable that another configuration of branes could yield well-defined bounds in flat-space limits. We leave this to future work.

## Appendix 8.B: Dimension Dependence of the Positive-Entropy Constraint

In general, the entropy density of the  $(d+1)$ -dimensional black string's horizon with two RS + DGP branes is given by (8.46), which we recall here:

$$\mathcal{S}_h = \frac{1}{u_h^{d-2}} \left[ \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu} + \frac{1}{d-2} \left( \frac{\lambda_1}{\sin^{d-2} \theta_1} + \frac{\lambda_2}{\sin^{d-2} \theta_2} \right) \right]. \quad (8.97)$$

$\theta_1$  and  $\theta_2$  are the angles of the branes while  $\lambda_1$  and  $\lambda_2$  are their respective DGP couplings. Furthermore,  $d$  is the number of brane dimensions. Physically, (8.97) must never be negative, but mathematically there exist ranges of  $\lambda_1$  and  $\lambda_2$  for which  $\mathcal{S}_h < 0$ . For fixed brane angles  $(\theta_1, \theta_2)$ , we denote this part of the  $(\lambda_1, \lambda_2)$  parameter space by

$$\mathcal{B}_d(\theta_1, \theta_2) \equiv \left\{ \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu} + \frac{1}{d-2} \left( \frac{\lambda_1}{\sin^{d-2} \theta_1} + \frac{\lambda_2}{\sin^{d-2} \theta_2} \right) < 0 \right\} \quad (8.98)$$

and assert that it resides in the swampland.

In the main text (Section 8.3.3), we use this condition to rule out a section of the DGP parameter space for  $d = 4$ . For fixed angles, the constrained region is linear, and we also compute both the *intersection*  $\cap \mathcal{B}_d$  and the *union*  $\cup \mathcal{B}_d$  of these excluded regions over all choice of  $(\theta_1, \theta_2)$ . The intersection is the region of DGP parameter space in which we get a negative horizon entropy for any choice of brane angles, while the union is that in which we find a negative horizon entropy for at least one choice of brane angles. However, the details may depend on  $d$ . Our goal in this appendix is to describe this dependence. Roughly speaking, the intersection becomes smaller as  $d$  is tuned larger, but the union is  $d$ -independent.

**Intersection of excluded regions** First, we observe that the horizon-entropy density can be rewritten as a sum over the two branes:

$$\mathcal{S}_h = \frac{1}{u_h^{d-2}} \sum_{i=1}^2 \mathcal{S}_h^{(i)}(d), \quad \mathcal{S}_h^{(i)}(d) \equiv \int_{\theta_i}^{\frac{\pi}{2}} \frac{d\mu}{\sin^{d-1} \mu} + \frac{\lambda_i}{d-2} \csc^{d-2} \theta_i. \quad (8.99)$$

As in the  $d = 4$  case,  $\mathcal{S}_h^{(i)}$  generally<sup>24</sup> diverges as  $\theta_i \rightarrow 0^+$  or  $\theta_i \rightarrow \pi^-$ . The signs of these divergences depend on  $\lambda_i$ , and the divergence at  $\theta_i = 0$  is particularly important for determining the intersection. If there is a positive divergence there, then we may force the entire sum to be positive by taking  $\theta_i \ll 1$ , .

We can find the signs of these divergences by examining the derivative of  $S_h^{(i)}$  with respect to  $\theta_i$ ,

$$\frac{\partial \mathcal{S}_h^{(i)}}{\partial \theta_i} = -\csc^{d-1} \theta_i (1 + \lambda_i \cos \theta_i). \quad (8.100)$$

For  $|\lambda_i| < 1$ , this quantity is negative, and so  $S_h^{(i)}$  monotonically decreases. Thus the divergence as  $\theta_i \rightarrow 0^+$  must be positive, while the divergence as  $\theta_i \rightarrow \pi^-$  is negative. For  $|\lambda_i| > 1$ , however, the derivative has a zero at  $\theta_i = \theta_{i*}$ , where

$$\cos \theta_{i*} = -\frac{1}{\lambda_i}. \quad (8.101)$$

Furthermore, we observe that  $\lambda_i > 1$  implies that this is a minimum, whereas  $\lambda_i < -1$  implies it is a maximum. So, both divergences are positive for  $\lambda_i > 1$ , whereas they are negative for  $\lambda_i < -1$ .

Taken together, these observations imply that if either  $\lambda_1 > -1$  or  $\lambda_2 > -1$ , then we may take one of the angles to be very small, inducing a positive blow-up in the total entropy. And so,

$$\bigcap \mathcal{B}_d \subset \{\lambda_1 < -1\} \cap \{\lambda_2 < -1\}. \quad (8.102)$$

We had found this in  $d = 4$ , but we have just shown that (8.102) is true in any number of dimensions. Recalling that  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$  is already deemed pathological both because of our wrong-sign-action argument in the main text and the RT bubble argument of Chen et al. (2020b), (8.102) means that the DGP couplings for which we get negative horizon entropy for all combinations of brane angles are already sick.

---

<sup>24</sup>The exception is for  $d = 3$ , for which the setting  $\lambda_i = 1$  cancels the  $\theta_i \rightarrow \pi^-$  divergence and taking  $\lambda_i = -1$  cancels the  $\theta_i \rightarrow 0^+$ . These are edge cases, so we ignore them.

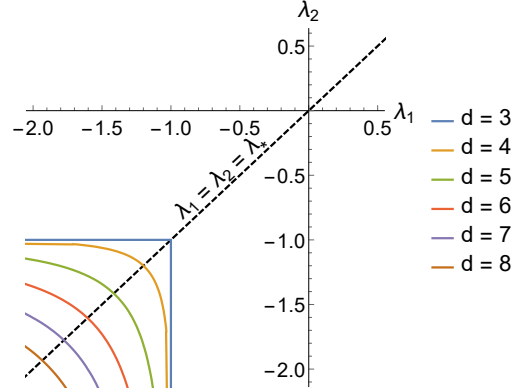


Figure 8.8: The  $(\lambda_1, \lambda_2)$  parameter space with the  $d$ -dependent contours demarcating the intersection (over brane angles) of all regions excluded by horizon-entropy positivity. We depict contours for a variety of (brane) dimensions  $d$ . The intersection for a given  $d$  is the region below the corresponding contour.

To compute the intersection more precisely, we must specify  $d$ . Upon doing so, we compute the maximum value of  $\mathcal{S}_h^{(i)}(d)$  as a function of  $\lambda_i$ . For example, these maxima for the first three values  $d = 3, 4, 5$  are

$$\mathcal{S}_h^{(i)}(3)\big|_{\theta_i=\theta_{i*}} = -\sqrt{\lambda_i^2 - 1}, \quad (8.103)$$

$$\mathcal{S}_h^{(i)}(4)\big|_{\theta_i=\theta_{i*}} = \frac{\lambda_i}{2} - \frac{1}{4} \log \left( \frac{\lambda_i + 1}{\lambda_i - 1} \right), \quad (8.104)$$

$$\mathcal{S}_h^{(i)}(5)\big|_{\theta_i=\theta_{i*}} = \frac{1}{3} \frac{2 - \lambda_i^2}{\sqrt{\lambda_i^2 - 1}}. \quad (8.105)$$

The point is that a pair of couplings  $(\lambda_1, \lambda_2)$  is in  $\cap \mathcal{B}_d$  precisely when the sum of the maxima for brane 1 ( $i = 1$ ) and brane 2 ( $i = 2$ ) is negative. We can find this region by finding the contour in  $(\lambda_1, \lambda_2)$  parameter space along which the sum of maxima are 0, then by noting that the horizon entropy becomes negative if either DGP coupling decreases. In other words, the intersection of excluded regions is everything to the lower-left of the zero contour. These contours for various dimensions are shown in Figure 8.8.



**Union of excluded regions** We can also ask what part of the  $(\lambda_1, \lambda_2)$  parameter space allows for *at least* one combination of brane angles furnishing a negative entropy, so as to get a sense of how much of the parameter space could potentially be impacted by the swampland constraints.<sup>25</sup> This turns out to be

$$\bigcup \mathcal{B}_d = \{\lambda_1 < -1\} \cup \{\lambda_2 < -1\} \cup \{\lambda_1 + \lambda_2 < 0\}, \quad (8.106)$$

regardless of the number of dimensions. The argument is similar to that of the main text. First, we note that

$$\mathcal{S}_h^{(i)}(d) = \frac{\sqrt{\pi}\Gamma\left(1 - \frac{d}{2}\right)}{2\Gamma\left(\frac{3}{2} - \frac{d}{2}\right)} - \frac{1}{2}B\left(\sin^2 \theta_i; 1 - \frac{d}{2}, \frac{1}{2}\right) + \frac{\lambda_i}{d-2} \csc^{d-2} \theta_i, \quad (8.107)$$

where  $B$  is the incomplete beta function and we employ a dimensional regularization scheme<sup>26</sup> to write a finite entropy. We then take the following extremal combinations of brane angles:

$$\theta_1 \rightarrow \frac{\pi}{2}, \quad \theta_2 \rightarrow \frac{\pi}{2} \implies \mathcal{S}_h^{(1)}(d) + \mathcal{S}_h^{(2)}(d) \sim \frac{1}{d-2}(\lambda_1 + \lambda_2), \quad (8.108)$$

$$\theta_1 \rightarrow 0, \quad \theta_2 \rightarrow \frac{\pi}{2} \implies \mathcal{S}_h^{(1)}(d) + \mathcal{S}_h^{(2)}(d) \sim \frac{(1 + \lambda_1)}{d-2} \csc^{d-2} \theta_1, \quad (8.109)$$

$$\theta_1 \rightarrow \frac{\pi}{2}, \quad \theta_2 \rightarrow 0 \implies \mathcal{S}_h^{(1)}(d) + \mathcal{S}_h^{(2)}(d) \sim \frac{(1 + \lambda_2)}{d-2} \csc^{d-2} \theta_1. \quad (8.110)$$

For these respective combinations of brane angles, requiring positive horizon entropy excludes the region  $\{\lambda_1 < -1\}$ ,  $\{\lambda_2 < -1\}$ , and  $\{\lambda_1 + \lambda_2 < 0\}$ , and so we at least have

$$\bigcup \mathcal{B}_d \subset \{\lambda_1 < -1\} \cup \{\lambda_2 < -1\} \cup \{\lambda_1 + \lambda_2 < 0\}. \quad (8.111)$$

Now, we argue that any point outside of the set on the right-hand side cannot furnish negative horizon entropy for any combination of angles. To do so, we write the

---

<sup>25</sup>To be more specific, the complement of the union is the part of the parameter space in which we will never get a negative horizon entropy and thus have no theories ruled out by the positivity requirement.

<sup>26</sup>The dimensional regulator cancels the even- $d$  divergence of the beta function, allowing us to ignore it.

general- $d$  bounding line in slope-intercept form:

$$\lambda_2 = L_{\mathcal{B},d}(\lambda_1; \theta_1, \theta_2) \equiv - \left( \frac{\sin \theta_2}{\sin \theta_1} \right)^{d-2} \lambda_1 - (d-2) \sin^{d-2} \theta_2 \int_{\theta_1}^{\pi-\theta_2} \frac{d\mu}{\sin^{d-1} \mu}. \quad (8.112)$$

As the constraint for fixed  $(\theta_1, \theta_2)$  is linear, it is sufficient to show the following three statements, which we had shown for  $d = 4$  in the main text:

- (i) any bounding line must have negative slope;
- (ii) any bounding line must intersect  $\lambda_1 = 0$  at some  $\lambda_2 < 0$ ; and,
- (iii) any bounding line which intersects  $\lambda_1 + \lambda_2 = 0$  once does so in  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\}$ .

(i) is immediately true; the slope read from (8.112) is always negative. (ii) is true because the integral  $\int \csc^{d-1} \mu$  is of a positive function over a positive interval ( $\theta_1 < \pi - \theta_2$ ), and  $\sin \theta_2$  is also positive.

(iii) is difficult to prove directly in general  $d$ , but we can prove it indirectly by using the linearity of the bound for fixed  $(\theta_1, \theta_2)$  in conjunction with (i) and (ii). Denote the intersection of a bounding line  $L_{\mathcal{B},d}$  with  $\lambda_1 + \lambda_2 = 0$  by  $(\lambda_1^*, \lambda_2^*)$ . Then, for any  $L_{\mathcal{B},d}$  with a negative slope and negative  $\lambda_2$ -intercept,  $\lambda_1^* \in (-1, 0)$  implies that  $L_{\mathcal{B},d}$  intersects  $\lambda_1 = -1$  at some  $\lambda_2 > 1$ , whereas  $\lambda_1^* \in (0, 1)$  implies that  $L_{\mathcal{B},d}$  intersects  $\lambda_2 = -1$  at some  $\lambda_1 > 1$ .

So, we now show that the bounding line evaluated at  $\lambda_1 = -1$  must have a value below 1. This argument along with the symmetry of the problem under interchanging  $1 \leftrightarrow 2$  is sufficient to also assert that bounding lines intersect  $\lambda_2 = -1$  only at  $\lambda_1 < 1$ . Fixing  $\theta_2 = \theta$  and setting  $\theta_1 = \pi - \theta - \delta$  where  $\delta \in (0, \pi - \theta)$ , we evaluate

$$L_{\mathcal{B},d}(-1; \pi - \theta - \delta, \theta) = \frac{\sin^{d-2} \theta}{\sin^{d-2}(\theta + \delta)} - (d-2) \sin^{d-2} \theta \int_{\pi-\theta-\delta}^{\pi-\theta} \frac{d\mu}{\sin^{d-1} \mu}. \quad (8.113)$$

At  $\delta = 0$ , this evaluates to 1. Furthermore, its derivative is

$$\left. \frac{\partial L_{\mathcal{B},d}}{\partial \delta} \right|_{\lambda_1=-1} = -(d-2) \frac{\sin^{d-2} \theta}{\sin^{d-1}(\theta + \delta)} [1 + \cos(\theta + \delta)] < 0. \quad (8.114)$$

The punchline is that a bounding line  $L_{\mathcal{B},d}$  cannot actually intersect  $\lambda_1 = -1$  at a point where  $\lambda_2 > 1$ . By essentially the same reasoning, this line cannot intersect  $\lambda_2 = -1$  at a point where  $\lambda_1 > 1$ . Lastly, we note that  $\lambda_1^* \neq 0$  unless the bounding line is itself  $\lambda_1 + \lambda_2 = 0$ . Ergo,  $\lambda_1^* \notin (-1, 1)$ , and so (iii) is true.

With (i)–(iii) proven, we conclude that any excluded region  $\mathcal{B}_d(\theta_1, \theta_2)$  is contained within the union  $\{\lambda_1 < -1\} \cup \{\lambda_2 < -1\} \cup \{\lambda_1 + \lambda_2 < 0\}$ . Thus, (8.106) is true for any number of spacetime dimensions.

# Chapter 9: The Hitchin Image in Type-D

## Abstract

Motivated by their appearance as Coulomb branch geometries of Class S theories, we study the image of the local Hitchin map in tame Hitchin systems of type-D with residue in a special nilpotent orbit  $\mathcal{O}_H$ . We describe two important features which distinguish it from the type A case studied in Balasubramanian et al. (2022b). The first feature, which we term *even type constraints*, arises iff the partition label  $[\mathcal{O}_H]$  has even parts. Our Hitchin image is non-singular and thus different from the one studied by Baraglia and Kamgarpour. We argue that our Hitchin image always globalizes to being the Hitchin base of an integrable system. The second feature, which we term *odd type constraints*, is related to a particular finite group  $\overline{A}_b(\mathcal{O}_H)$  being non-trivial. This finite group parametrizes the choices for the local Hitchin base. Additionally, we also show that the finite group  $\overline{A}_b(\mathcal{O}_H)$  encodes the size of the dual special piece.

## 9.1 Introduction

Complex integrable systems are a powerful tool in high energy physics. The Coulomb branch of a 4D  $\mathcal{N} = 2$  supersymmetric quantum field theory is a special-Kähler manifold, which arises as the base of such an integrable system. Seiberg and

---

<sup>0</sup>This chapter is based on published work in collaboration with Balasubramanian et al. (2023). All authors contributed to the final version of the manuscript. My contributions included participating in group discussions that helped shape the project, writing portions of the manuscript, and performing computations for several of the examples presented in the paper.

Witten’s breakthrough Seiberg and Witten (1994c,d, 1996b) amounted to interpreting the low energy effective action on the Coulomb branch of certain  $\mathcal{N} = 2$  supersymmetric gauge theories in terms of a complex integrable system, and then solving it explicitly in terms of theta functions. The connection with complex integrable systems was generalized in Donagi and Witten (1996b) to a wider class of  $\mathcal{N} = 2$  theories, while Freed (1999b) gave a precise mathematical description of special Kähler manifolds.

A far-reaching generalization Gaiotto (2012b); Gaiotto et al. (2013b) involves class-S theories. These start their lives as 6-dimensional (2,0) super conformal field theories depending on a simply laced Lie algebra of type ADE. This can then be compactified on a Riemann surface  $C$  with punctures, to give a family of 4-dimensional theories. The resulting theories are described in terms of a meromorphic Hitchin system Markman (1994); Bottacin (1995). The total space parametrizes meromorphic Higgs bundles of type ADE on  $C$ , where the Higgs fields have first order poles at the punctures, with nilpotent residues in specified conjugacy classes.

The *superconformal* theories of class-S (where all masses are tuned to zero) are governed by particular symplectic leaves of the Hitchin integrable system. In turn, if  $C$  has genus- $g$  and  $n$  punctures, the corresponding superconformal theories come in a *family* of complex dimension  $3g - 3 + n$ , parameterized by the complex structure of the punctured curve  $C$ . We are thus naturally led to consider *families* of Hitchin systems, fibered over the moduli space of complex structures,  $\mathcal{M}_{g,n}$ . From the physics, it is natural to extend this to the boundary, where  $C$  degenerates. In Balasubramanian et al. (2022b), we studied this for type-A, and showed how the Hitchin bases fit together to form a holomorphic vector bundle over  $\overline{\mathcal{M}}_{g,n}$ . For  $\overline{\mathcal{M}}_{0,4} \simeq \mathbb{CP}^1$ , this vector bundle can be calculated very explicitly, and its splitting type<sup>1</sup> constitutes a new global property of the family of SCFTs.

---

<sup>1</sup>Every holomorphic vector bundle,  $E \rightarrow \mathbb{CP}^1$  splits as a direct sum of line bundles, and the degrees of those line bundles (the “splitting type”) completely characterize  $E$ .

Extending that construction to types  $D$  and  $E$  is the subject of the present series of papers. In this first installment, we will study the *local* behaviour of the Higgs field near one of the punctures, where it is assumed to have a simple pole with nilpotent residue. *Therefore, and unless otherwise instructed, we'll always refer to the local Hitchin base.*

In type-A, the *local Hitchin base* is parameterized by the Laurent tails of the coefficients of powers of  $\lambda$  in  $(\phi) = \det(\phi(t) - \lambda I)$ . The new ingredient in type-D is that the *leading* Laurent coefficients are not all independent; rather, they obey certain polynomial *constraints*. Those constraints were first computed in Chacaltana and Distler (2013b). Later, Baraglia and Kamgarpour Baraglia and Kamgarpour (2018) found a general expression for the constraints when the residue of  $\phi$  lies in a *Richardson* nilpotent orbit<sup>2</sup>. Part of the purpose of the present paper is to extend that to the (physically-relevant) case of an arbitrary *special* nilpotent orbit (see Def 9.1). Outside of Cartan type A, Richardson nilpotent orbits form a proper subset of the set of special nilpotent orbits.

Moreover, there are a few crucial differences between the constructions of Chacaltana-Distler (CD) and Baraglia-Kamgarpour (BK). They construct different Hitchin bases. CD's base is birational to (a finite cover of) BK's base. This will have a dramatic effect when we go to construct global examples, as we briefly do in Section 9.6.

In Section 9.2, we outline the local Hitchin setup. We review Spaltenstein's results on the factorization of  $(\phi)$  in Section 9.4.1. The factorization that Spaltenstein found is at the root of what we call the “even type” constraints of Baraglia-Kamgarpour Baraglia and Kamgarpour (2018) and Chacaltana-Distler Chacaltana and Distler (2013b) (who called them “ $c$ -type constraints”). Those two sets of authors take different approaches to implementing these constraints and the difference

---

<sup>2</sup>A Richardson nilpotent orbit is one that is induced from the zero orbit inside some Levi subalgebra (Collingwood and McGovern, 1993, pp. 107).

is summarized in Theorem 2 of Section 9.4.2.

In addition to the even type constraints, Spaltenstein’s factorization results also lead to a set of what we call “odd type” constraints (Chacaltana-Distler called these “ $a$ -type constraints”). We explain this in Section 9.5, where we relate the existence of odd type constraints to the presence of a nontrivial dual special piece (see Def 9.9 ), dual to the Hitchin nilpotent orbit. Each odd type constraint is associated with passing to a double-cover (setting  $c_{2k} = a_k^2$  in the notation below) of the naive Hitchin base. More precisely, we define a subgroup of Lusztig’s canonical quotient,  $\overline{A}_b \subset \overline{A}(\mathcal{O}_H)$  such that the nilpotent orbits in the dual special piece are in 1-1 correspondence with subgroups  $C \subset \overline{A}_b$ . If (as is the case in type-D)  $\overline{A}_b = (\mathbb{Z}_2)^s$ , then there are  $2^s$  choices for the Hitchin base, which has a factor  $\mathbb{C}^s/C$ . These choices are labeled by the choice of a subgroup  $C \subset \overline{A}_b$  or equivalently by a choice of nilpotent orbit in the dual special piece.

Both part (iii) of Theorem 2 and the existence of the odd-type constraints discussed in Section 9.5 rely on a non-obvious feature of the coefficients that appear in Spaltenstein’s factorization of  $(\phi)$ . A-priori, the coefficients are algebraic (in particular, not necessarily single-valued) functions on the Lie algebra. Certain of those coefficients, however, turn out to be polynomial functions on the Lie algebra. The result is actually independent of the details of Spaltenstein’s factorization and can be summarized in Conjecture 9.3. As discussed in Section 9.3, we have a lot of evidence for, but no proof of, the conjecture<sup>3</sup>.

In Section 9.5.5, we explain how the present work is compatible with the existence of mass deformation (or Poisson deformations) of the integrable system associated to the superconformal field theory. As explained in Balasubramanian and Distler (2021b), these deformations, in their local incarnation, correspond to *sheets* in a Lie algebra. In Section 9.5.5, we outline a conjecture that relates the local Hitchin image and Losev’s theory of birational sheets Losev (2022), which is refinement of the

---

<sup>3</sup>**Note added in proof:** Conjecture 9.3 has since been proven in Distler et al. (2024).

usual theory of sheets. A special case of this conjecture reduces to a duality between the smallest nilpotent orbit in a special piece and a birational sheet attached to a special nilpotent orbit. Such a duality was first noticed by J. Adams Adams (see our Remark 9.5.5 ) .

Finally, in Section 9.6, we give some global examples which illustrate the difference between the Hitchin bases constructed à la Chacaltana-Distler versus Baraglia-Kamgarpour and the different Hitchin bases that result from different choices of odd type constraints.

## 9.2 The Local Hitchin Base in Type-D

We study the local Hitchin base for tame Higgs bundles of type  $D_n$  with special nilpotent residues for the Higgs field. In this paper, we work in the framework of meromorphic Higgs bundles Markman (1994); Bottacin (1995).

Our primary motivation is to extend the work in Balasubramanian et al. (2022b) where we studied Hitchin base(s) of Cartan type A as a *family* over the moduli space of curves. In order to pursue a similar study in type-D, it turns out we need to understand some new phenomena that occur already in the study of the local Hitchin base. One can, in fact, study the Hitchin base at the following three levels

1. The purely local Hitchin base/Hitchin image (for a punctured disc).
2. The global Hitchin base (for a global curve  $C_{g,n}$  ).
3. The family of Hitchin bases over the moduli of curves  $\overline{\mathcal{M}}_{g,n}$  (“very global”).

In what follows, we confine ourselves to questions that can be addressed at the local level. The stories are, however, tightly linked and at a certain point, we will actually have to address the question of which features in the local story globalize. We will separately consider the new questions arising at the global and very global levels.



At the purely local level, when compared to type A, we find that describing the image of the Hitchin map in type-D is more subtle. Let  $\mathcal{B}^{\text{naive}}$  be the naive Hitchin base computed as though we were in type A (by this, we mean that we follow the Young tableaux based algorithm in Balasubramanian et al. (2022b) and just ignore the base contributions in odd degrees except for the Pfaffian in the case of  $D_N$ ,  $N$  odd). We say the Hitchin base *has constraints* iff the actual local Hitchin base  $\mathcal{B}$  is not the same as  $\mathcal{B}^{\text{naive}}$ .

Recall that nilpotent orbits in  $D_n$  are classified by D-type partitions of  $2n$  (Collingwood and McGovern, 1993, pp. 69-70). D-type partitions are those partitions in which even parts<sup>4</sup> occur with even multiplicity. We also follow the usual convention that the parts  $p_j$  of a partition  $[p_j]$  are always written in non-increasing order.

The constraints we encounter come in two types

1. **Even-type constraints** : These occur *if and only if* the Hitchin partition has even parts.
2. **Odd-type constraints** : These occur *if and only if* the  $2j^{\text{th}}$  and  $(2j + 1)^{\text{st}}$  parts of the Hitchin partition are both odd ( $p_{2j} = 2r + 1$ ,  $p_{2j+1} = 2s + 1$ ) with  $p_{2j} > p_{2j+1}$ .

When these constraints occur, the description of the Hitchin base requires more work. The naïve Hitchin base  $B^{\text{naive}}$  is a vector space. We let  $c_i$  denote the coordinates on it. The actual image,  $B^{BK}$ , is the subvariety of  $B^{\text{naive}}$  cut out by the constraints, which are certain polynomials in the  $c_i$ . In the case of Richardson nilpotent orbits, this was studied by Baraglia and Kamgarpour in Baraglia and Kamgarpour (2018). They find instances where their local Hitchin base  $B^{BK}$  is singular. On the other hand, Chacaltana and Distler in Chacaltana and Distler (2013b) note that one can always “solve” these constraints: they exhibit a new vector space  $B^{CD}$  and a surjective

---

<sup>4</sup>i.e. parts or elements of the partition which are even

morphism  $B^{CD} \rightarrow B^{BK}$ . We will exhibit some examples where this map is a birational morphism, and others where it has degree  $> 1$ . We let  $a_j$  denote the coordinates on  $B^{CD}$ ; the  $c_i$  are then polynomial functions of the  $a_j$ . From a physical perspective,  $B^{CD}$  is more natural, as it agrees with the expected properties of the Coulomb branch.

We restrict to special nilpotent residues for the Higgs field for a few reasons. First, these are the orbits that have appeared so far in the description of the Coulomb branches of superconformal field theories of class-S. Second, we will find that aspects of the story about even type constraints are easier to understand for special nilpotent residues. Even type constraints can be thought of as an extension of what was called *fingerprints of surface operators* in Gukov and Witten (2006b).

In the GIT framework for constructing the moduli of stable Higgs bundles, an important notion is that of *strongly parabolic Higgs bundles*. They correspond to the cases where the residue of the Higgs field is nilpotent and compatible with the parabolic structure at the puncture. From our point of view, the formalism of strongly parabolic Higgs bundles is somewhat limiting as it leads only to residues that live in Richardson nilpotent orbits (as in Baraglia and Kamgarpour (2018)). For our purposes, we needed to allow all special nilpotent orbits as possible residues. For any simple Lie algebra not of type- $A$ , the set of Richardson nilpotent orbits form a proper subset of the set of special nilpotent orbits. It would be interesting to know what replacement for the notion of strongly parabolic Higgs bundles would lead us exactly to the set of all special nilpotent orbits.

We note that some of the underlying mathematics (having to do with Spaltenstein's work on the Kazhdan-Lusztig map Spaltenstein (1988, 1990b) ) appearing in our study of the structure of the local Hitchin image had also appeared in a recent study of an analog of the Hitchin fiber in the local setting Yun (2021). It is interesting to ask if a direct connection can be found between the features we find in the local Hitchin image and the geometry of the fibers in Yun (2021).

While this paper is confined to the case of type-D Hitchin systems, it is natural

to wonder if similar features exist for other Cartan types. The connection to class-S theories extends most directly to all Hitchin systems corresponding to simply laced Lie algebras. So, the next natural set of examples to study would be Hitchin systems of type  $E_6, E_7, E_8$ . These have been studied and the explicit structure of constraints has been worked out for any special nilpotent orbit  $\mathcal{O}$  in  $E_6$  Chacaltana et al. (2015c) and  $E_7$  Chacaltana et al. (2018a). In the case of  $E_8$ , partial results on the constraints have been worked out in unpublished work by the authors of Chacaltana et al. (2018b).

As in type-D, the pattern of constraints is closely tied to the size of the dual special piece which, in some cases, is smaller than the size of Lusztig's quotient  $\overline{A}(\mathcal{O})$ . In the type-D case, we show in Theorem 3 that a smaller group  $\overline{A}_b(\mathcal{O}) \subset \overline{A}(\mathcal{O})$  captures the size of the dual special piece. Using the case-by-case computations in Chacaltana et al. (2015c, 2018a,b, 2016), we can define such a smaller group  $\overline{A}_b(\mathcal{O})$  for each special nilpotent orbit in  $E_6, E_7, E_8$ . We believe it should also be possible to give a case-free definition of the  $\overline{A}_b$  group in a way that extends naturally to exceptional cases and reduces to the one given in Def 9.7 for the type-D cases.

For the non-simply laced cases, the connection to class-S theories is less direct. We do, however, note some possible directions. On the one hand, we have the local results which relate twisted defects to nilpotent orbits in non-simply laced Lie algebras Chacaltana et al. (2015a,b,d). On the other hand, we have certain global results in the mathematical literature which relate Hitchin systems (without punctures) of type  $B, C, F, G$  to Hitchin systems of type  $A, D, E$  Beck et al. (2022). Extending the global results to the punctured case using the local results on twisted defects might be a way to study corresponding questions for the non-simply laced cases.

### 9.2.1 The local setup

Let  $\mathcal{O} = C[[t]]$  be the ring of formal power series and  $\mathcal{F} = C((t))$  be the field of formal Laurent series. The Higgs field in the neighbourhood of a puncture has the following form

$$\phi' = \frac{\mathbf{n}}{t} + \text{reg}$$

where  $\mathbf{n} \in \mathfrak{g}$  is a nilpotent element living in the (co-adjoint) orbit  $\mathcal{O}_H$  (ie. the Hitchin nilpotent orbit). Note that the above condition is equivalent to saying that  $t\phi'$  is of the form  $\mathbf{n} + m\mathfrak{g}$  where  $m$  is the unique maximal ideal in  $\mathcal{O}$ . This version will be helpful in relating to Kazhdan and Lusztig (1988); Spaltenstein (1988, 1990a).

Nilpotent orbits in the classical Lie algebras are labeled by partitions. For  $\mathfrak{so}(2N)$ , we have a D-partition  $[p]$ , a partition of  $2N$  such that every even part occurs with even multiplicity<sup>5</sup>. In what follows we will restrict to  $\mathbf{n}$  in a *special* nilpotent orbit.

**Definition 9.1.** (Special orbits in type-D) A special nilpotent orbit in type-D is a nilpotent orbit  $\mathcal{O}_H$  whose partition label  $[p]$  is a D-partition (all even parts occur with even multiplicity) such that there is an even number (possibly zero) of odd parts between any even parts and an even number (possibly zero) of odd parts at the beginning of the partition. Equivalently (Collingwood and McGovern, 1993, pp. 69-70),  $[p]^T$  is a C-partition<sup>6</sup>.

A Richardson orbit in type-D may be characterized as follows (Kempken (1983) 4.3):

- $[p]$  is special.
- Let  $r = \min_i \{i | p_i \text{ even}\} - 1$ . Then, for all  $j \geq \frac{1}{2}r + 1$ , if  $p_{2j} = 2l + 1$  and  $p_{2j+1} = 2m + 1$ , then  $l > m$ .

Evidently, all even orbits (orbits whose *weighted Dynkin diagram* consists of 0's and 2's) are Richardson.

---

<sup>5</sup>When all of the parts of  $[p]$  are even, then there are actually *two* nilpotent orbits corresponding to that partition. See (Collingwood and McGovern, 1993, pp. 69-70).

<sup>6</sup>A C-partition is a partition of  $2N$  such that every odd part occurs with even multiplicity.

Now, consider the map

$$\mu : \mathfrak{g}(\mathcal{F}) \rightarrow \mathbb{A}_{\mathcal{F}}^N$$

that is defined as follows

$$\mu : x \rightarrow (Q_1(x), Q_2(x), \dots, Q_N(x))$$

where  $Q_i$  are some chosen generators for the ring of invariant polynomials of  $\mathfrak{g}$ . We define the *local Hitchin image* associated to the nilpotent orbit  $\mathcal{O}_H$  to be the image  $\mu(\mathfrak{n} + m\mathfrak{g})$ .

**Zero Orders and Pole Orders:** A Hitchin system is an algebraic integrable system associated with the moduli space of Higgs bundles. In its simplest form, a Higgs bundle consists of a pair  $(V, \Phi)$ , where  $V$  is a principal  $\mathfrak{J}$ -bundle over a Riemann surface  $C$ , and the Higgs field  $\Phi$  is a section of  $\text{ad}(V) \otimes K$ , with  $K$  being the canonical bundle of  $C$ . The moduli space of such Higgs bundles is endowed with a natural Poisson structure, and the Hitchin map sends a Higgs bundle to the collection of coefficients of the characteristic polynomial of  $\Phi$ .

We have a choice to make as to where the Higgs field lives. If we choose  $\phi' \in H^0(\text{ad}(V) \otimes K)$ , tame Higgs bundles on the punctured disc correspond to those where the Higgs field has a simple pole ( $\phi' = \frac{a}{t} + \text{reg}$ ) at the puncture and the invariant polynomials built out of the Higgs field have poles of various orders. We denote these pole orders by  $\pi_k$  where  $k$  runs over the degrees of the invariant polynomials  $Q_i$ ,  $i \in \{1, 2, \dots, N\}$ .

On the other hand, if we choose  $\phi \in H^0(\text{ad}(V) \otimes K(D))$ , tame Higgs bundles correspond to those where, in a local coordinate  $t$  around the puncture, the Higgs field can be expressed as  $\phi \in \mathfrak{j} \otimes \mathbb{C}[[t]]$ , where  $\mathfrak{j}$  is the Lie algebra associated to  $V$ . In this formulation  $\phi$  has a nonzero constant term, and the invariant polynomials built

out of the Higgs field have zeros of order  $\chi_k$  (in the notation in Balasubramanian et al. (2022b), where  $k$  is the degree of the corresponding  $Q_i$ ). The relationship between the two definitions for the Higgs fields is simply

$$\phi = t\phi'$$

From this, it follows that the zero orders and pole orders are related by :

$$\chi_k = k - \pi_k$$

### 9.2.2 Newton polygons

It is sometimes convenient to draw Newton polygons to keep track of the non-zero coefficients in the Laurent tails of the coefficients of  $\lambda$  in  $(\phi - \lambda I)$ .

Let  $[p_j]$  be the Hitchin partition and let  $[s_j]$  be the partial sums built out of the Hitchin partition where  $s_j = \sum_{i=1}^j p_i$ . Let  $\sum_j p_j = n$ . We follow the conventions of Baraglia and Kamgarpour (2018) and define a Newton polygon  $NP([p_j])$  associated to a partition  $[p_j]$  to be the polygon bounded by line segments joining the points  $(j, n - s_j)$ .<sup>7</sup> In Fig 9.1, we give an example of a Newton polygon for  $n = 12$  and Hitchin partition  $[5, 3, 2^2]$ .

We denote a generic polynomial of degree  $n$  with coefficients in  $\mathcal{O}$  to be :

$$\mathcal{P}_n = \lambda^n + \sum_{\beta} \sum_{\alpha} \rho_{\alpha, \beta} t^{\alpha} \lambda^{\beta}$$

We say that a polynomial  $\mathcal{P}_n$  corresponds to a Newton polygon  $NP$  if the coefficients  $\rho_{\alpha, \beta}$  are non-zero only if  $(\alpha, \beta)$  lie in the region defined by the polygon (see shaded region in example above). Note that, by definition, we always have  $\rho_{(0, n)} = 1$ .

---

<sup>7</sup>We note here that Spaltenstein uses a different convention in Spaltenstein (1988) and uses a Newton polygon bounded by line segments joining the points  $(n - s_j, j)$ .

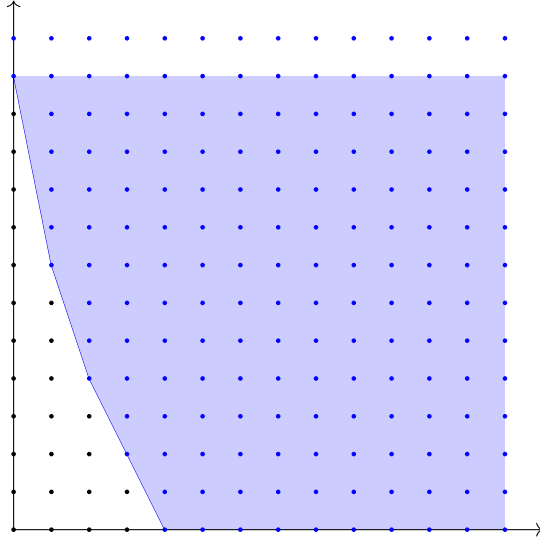


Figure 9.1: A Newton polygon showing the non-zero coefficients in  $NP([5, 3, 2^2])$ . We have chosen the conventions from Baraglia and Kamgarpour (2018) and taken the vertices to be at  $(j, n - s_j)$

We denote by  $NP([\mu_i])$  a Newton polygon determined by a partition  $[\mu_i]$ . Now, let  $(\phi)$  in type  $D_N$  be :

$$(\phi) = \lambda^{2N} + u_2 \lambda^{2N-2} + \cdots + u_{2N-2} \lambda^2 + u_N^2 \quad (9.1)$$

with coefficients  $u_{2j}, u_N \in \mathcal{O}[[t]]$ , the ring of formal power series. Let  $j \in 1, 2, \dots, N-1$ . We will denote the leading terms, as  $t \rightarrow 0$ , of  $u_{2j}[[t]], u_N[[t]]$  as  $c_{2j}t^{\chi_{2j}}, \tilde{c}t^{\chi_N}$  respectively. The leading terms have zero orders  $\chi_{2j}, \chi_N$  and  $c_{2j}, \tilde{c}$  are the coefficients. A basic result that determines the zero orders is the following :

*Lemma 1.* The Newton polygon corresponding to  $(\phi)$  is  $NP([p_j])$  where  $[p_j]$  is the Hitchin partition.

*Proof.* The Lemma follows directly from Cor 23 in Baraglia and Kamgarpour (2018).  $\square$

The space of  $u_{2j}[[t]], u_N[[t]]$  determines what we called  $\mathcal{B}^{\text{naive}}$ . In type  $D$ , there

are some additional constraints obeyed by the leading terms  $c_{2j}, \tilde{c}$ . These additional constraints are the main subject matter of this paper. In subsequent sections, we will turn to studying these constraints from a local point of view.

### 9.2.3 Spaltenstein's factorization

Our analysis will heavily rely on results of Spaltenstein on factorization patterns of characteristic polynomials and the Kazhdan-Lusztig map. In particular, we need Theorem B and Proposition 6.4 from Spaltenstein (1988). We specialize the theorem to the case of Higgs fields with residues in special orbits.

*Observation 1.* (Spaltenstein Spaltenstein (1988)) For any special nilpotent orbit  $\mathcal{O}_H$  in  $\mathfrak{so}_{2N}$ , let  $(\phi)$  be the characteristic polynomial built out of the Higgs field  $\phi = t\phi'$  with local behaviour  $\phi' = \mathfrak{n}/t + (\text{reg})$  and  $\mathfrak{n}$  being an element in the orbit  $\mathcal{O}_H$ . Then  $(\phi)$  has a factorization

$$(\phi) = \prod_i P_{\alpha,i}(\lambda) P_{\alpha,i}(-\lambda) \prod_j P_{\beta,j}(\lambda^2) \quad (9.2)$$

into irreducible polynomial factors labeled by a bi-partition  $(\alpha, \beta)$  of  $N$ .

First, the bi-partition  $(\alpha, \beta)$  is determined from the special partition  $[p]$  in the following way

- When  $p_{2j-1} = p_{2j}$ , then  $\alpha$  has one part equal to  $p_{2j-1}$ .
- When  $p_{2j-1} = 2r + 1$  and  $p_{2j} = 2s + 1$ , with  $r > s$ , then  $\beta$  has two parts:  $r$  and  $s + 1$ .

Then, the irreducible factors are determined from the parts of  $(\alpha, \beta)$  according to :

- Corresponding to every part  $\alpha_i$  of  $[\alpha]$ , there is a pair of irreducible factors:  $P_{\alpha,i}(\lambda), P_{\alpha,i}(-\lambda)$ , of  $(\phi)$  each of degree  $\alpha_i$  where

$$P_{\alpha,i}(\lambda) := (\lambda^{\alpha_i} + te_1^{(i)}[[t]]\lambda^{\alpha_i-1} + te_2^{(i)}[[t]]\lambda^{\alpha_i-2} + \cdots + te_{\alpha_i}^{(i)}[[t]]).$$



- Corresponding to every part  $\beta_i$  of  $\beta$ , there is a single irreducible factor, *i.e.* a polynomial  $P_{\beta,i}(\lambda^2)$ , of degree  $\beta_i$ .

$$P_{\beta,i}(\lambda^2) := (\lambda^{2\beta_i} + t f_2^{(i)}[[t]] \lambda^{2\beta_i-2} + t f_4^{(i)}[[t]] \lambda^{2\beta_i-4} + \cdots + t f_{2\beta_i}^{(i)}[[t]]).$$

The coefficients  $e_n[[t]]$  and  $f_n[[t]]$  are formal power series in  $t$ . The coefficients in those power series are, in general, algebraic (*i.e.*, not necessarily single-valued) functions on the Lie algebra. Of course,  $(\phi)$  is single-valued. In fact, its coefficients (as a double power series in  $\lambda$  and  $t$ ) are polynomials on the Lie algebra. Thus, whatever monodromies are experienced by the coefficient functions  $e_n[[t]]$  and  $f_n[[t]]$  must be such as to exchange the polynomial factors in (9.2). The condition that the partition be special implies that no  $\alpha_i$  can equal  $2\beta_j$ , so no  $P_\alpha(\lambda)$  can have the same degree as a  $P_\beta(\lambda^2)$ . In particular, monodromy permutes the  $P_\alpha$ 's and  $e$ 's among themselves, and similarly for the  $P_\beta$ 's and  $f$ 's, with no mixing.

The parts of  $[\beta]$  can only occur with multiplicity 2 or 1. For  $[p]$  special, the multiplicity is 2 iff  $p_{2j-1} = p_{2j} + 2 = 2r + 1$ , in which case  $\beta_i = \beta_{i+1} = r$ , for some  $i$ . In this case, it is the product  $P_{\beta,i}(\lambda^2)P_{\beta,i+1}(\lambda^2)$  that is carried into itself under monodromy. Hence (in particular)  $f_{2r}^{(i)} f_{2r}^{(i+1)}$  is a rational function on the Lie algebra.

When  $p_{2j-1} = p_{2j} + 2l = 2r + 1$  (for some  $l > 1$ ), then  $P_{\beta,i}(\lambda^2)$  and  $P_{\beta,i+1}(\lambda^2)$  are separately monodromy-invariant and hence  $f_{2r}^{(i)}[[t]]$  and  $f_{2(r-l+1)}^{(i+1)}[[t]]$  are separately rational functions on the Lie algebra.

If the parts of  $[\alpha]$  are all distinct, then the monodromy can *at worst* exchange  $P_{\alpha,i}(\lambda) \leftrightarrow P_{\alpha,i}(-\lambda)$ . In other words, it could act on the  $e_n[[t]]$  as  $e_n[[t]] \rightarrow (-1)^n e_n[[t]]$ . So the (coefficients of powers of  $t$  in)  $e_{2n}[[t]]$  are rational functions on the Lie algebra, whereas the coefficients in  $e_{2n+1}[[t]]$  can have the form of a rational function times the square-root of a polynomial.

When the parts of  $[\alpha]$  appear with multiplicity greater than 1, then more complicated monodromy patterns can appear. Let  $[p] = [\dots, (2r)^{2s}, \dots]$ , and let the

corresponding parts of  $[\alpha]$  be  $\alpha_i = \alpha_{i+1} = \dots = \alpha_{i+s-1} = 2r$ . Then Spaltenstein's factorization contains a product of the form

$$\prod_{j=0}^{s-1} P_{\alpha, i+j}(\lambda) P_{\alpha, i+j}(-\lambda)$$

of polynomials, each of which has degree  $2r$ . These can have monodromies among themselves, but the product

$$\prod_{j=0}^{s-1} e_{2r}^{(i+j)}[[t]]$$

is monodromy-invariant, and hence a rational function on the Lie algebra. To see this, simply note that  $Q_{2r}(\lambda) = \prod_{j=0}^{s-1} P_{\alpha_{i-j}}(\lambda)$  can only be exchanged with  $Q_{2r}(-\lambda)$  under monodromy and that  $\prod_{j=0}^{s-1} e_{2r}^{(i+j)}[[t]]$  is the coefficient of  $\lambda^0$  in both  $Q_{2r}(\lambda)$  and  $Q_{2r}(-\lambda)$ .

Similarly, if  $[p] = [\dots, (2r+1)^{4s}, \dots]$ , then Spaltenstein's factorization contains a product of the form

$$\prod_{j=0}^{2s-1} P_{\alpha, i+j}(\lambda) P_{\alpha, i+j}(-\lambda)$$

of polynomials of degree  $2r+1$ . Again, these polynomials can have monodromies among themselves, but the product

$$\prod_{j=0}^{2s-1} e_{2r+1}^{(i+j)}[[t]]$$

is monodromy-invariant, and hence a rational function on the Lie algebra.

**Summary:** The following products of coefficients are rational functions on the Lie algebra:

- $f_{2r}^{(2i-1)}[[t]]$  and  $f_{2(r-l+1)}^{(2i)}[[t]]$  when  $p_{2j-1} = p_{2j} + 2l = 2r + 1$  (for some  $l > 1$ )
- $f_{2r}^{(2i-1)} f_{2r}^{(2i)}$  when  $p_{2j-1} = p_{2j} + 2 = 2r + 1$ .
- $e_{2n}[[t]]$  if the parts of  $[\alpha]$  are all distinct.
- $\prod_{j=0}^{s-1} e_{2r}^{(i+j)}[[t]]$  if  $[p] = [\dots, (2r)^{2s}, \dots]$ .

- $\prod_{j=0}^{2s-1} e_{2r+1}^{(i+j)}[[t]]$  if  $[p] = [\dots, (2r+1)^{4s}, \dots]$ .

The conjecture of the next section is, in essence, the assertion that certain monodromy patterns do not, in fact, occur and that certain additional coefficients (or products of coefficients) in the  $e_n^{(i)}[[t]]$  and  $f_{2n}^{(i)}[[t]]$  are rational functions on the Lie algebra, even though the factorization (9.2) would allow them to have nontrivial monodromies.

### 9.3 A Conjecture

Given a special D-partition,  $[p]$  and a nilpotent,  $X$ , in the corresponding nilpotent orbit, let us write the characteristic polynomial  $(\phi) = \det(X + tg - \lambda I)$ ,  $g \in \mathfrak{so}(2N)$  as

$$(\phi) = \lambda^{2N} + \sum_{k=1}^N t^{\chi_{2k}} (c_{2k} + O(t)) \lambda^{2(N-k)} \quad (9.3)$$

Here, if  $[p]$  has parts  $p_i$  then

$$\chi_{2k} = j, \quad \text{where} \quad \sum_{i=1}^{j-1} p_i < 2k \leq \sum_{i=1}^j p_i$$

A-priori, the  $c_{2k}$  are polynomials on the Lie algebra, of degree  $\chi_{2k}$ . In certain cases, a stronger result obtains. Let  $[p]$  be a special D-partition. For every part  $p_{2i}$  such that  $p_{2i} > p_{2i+1}$ , let  $2k_i = \sum_{j=1}^{2i} p_j$ . Then  $c_{2k_i} = (a_{k_i})^2$  with  $a_{k_i}$  a polynomial on the Lie algebra. An obvious special case is when  $p_{2i}$  is the last part of  $[p]$ , *i.e.* when  $k = N$ . Then  $c_{2N}$  is the square of the Pfaffian. More generally, we have:

**Proposition 9.1.** *Conjecture 9.3 holds in the following cases:*

- If  $p_1, p_2, \dots, p_{2i}$  are all even then the Conjecture holds for  $c_{2k_i}$ .*
- If all of the parts  $p_j$ , for  $j > 2i$  are even, then the Conjecture hold for  $c_{2k_i}$ .*
- If all odd parts of  $[p]$  occur with multiplicity a multiple of 4, then Conjecture 9.3 holds, with  $a_{k_i} = \prod_{j=1}^i e_{p_{2j}}^{(j)}$ .*

*Proof.* Part (a) follows from the proof of part (i) of Theorem 2 below. Part (b) follows from the proof of part (iii) of Theorem 2. For Part (c), the individual  $e_{p_{2j}}^{(j)}$  are not necessarily rational functions on the Lie algebra. Nevertheless, for each even part of  $p$  (occurring with multiplicity  $2s$ ) and for each odd part (occurring with multiplicity  $4s$ ), we showed at the end of Section 9.2.3 that the *product* of the corresponding  $e_{p_{2j}}^{(j)}$  is rational. Taking the product of these rational functions, we obtain a rational function whose *square* is a polynomial on the Lie algebra. Hence it, itself, is a polynomial on the Lie algebra.  $\square$

We have not proven the Conjecture for general  $[p]$  with odd parts whose multiplicity is  $4s + 2$  or  $2s + 1$ . Nevertheless, Chacaltana-Distler (Chacaltana and Distler (2013b) and unpublished) checked the Conjecture explicitly for all nilpotent orbits in  $\mathfrak{so}(2N)$  for  $N \leq 7$ . We have spot-checked a large number of cases for higher  $N$ .

**Note added in proof:** Conjecture 9.3 has since been proven in Distler et al. (2024) for an arbitrary special D-partition.

## 9.4 Even and Very Even Type Constraints

### 9.4.1 Work of Baraglia-Kamgarpour

Now, we apply Spaltenstein's result (Theorem 1) to study the Hitchin image in the presence of even parts in the Hitchin partition. This discussion will closely follow the work of Baraglia and Kamgarpour (2018). Later, we will emphasize how the approach of Chacaltana and Distler (2013b) differs from that of Baraglia and Kamgarpour (2018). Let  $[p]$  be the Hitchin partition. For  $[p]$  a special orbit in type  $D$ , Theorem 1 implies that

$$(\phi) = \mathcal{P}^{odd} \prod_{i|\alpha_i \text{ even}} P_{\alpha,i}(\lambda) P_{\alpha,i}(-\lambda)$$

where  $\mathcal{P}^{odd}$  contains all of the factors in (9.2) corresponding to the odd parts of  $[p]$ . An even part  $[p] = [\dots, (2r)^{2s}, \dots]$  corresponds to  $[\alpha] = [\dots, (2r)^s, \dots]$ , and the product is over the even parts of  $[\alpha]$ .

Now, following Baraglia and Kamgarpour (2018), for each *distinct* even part  $r_d$  of  $[p]$ , occurring with multiplicity  $2e_d$ , let us *define* a polynomial  $Q$  that is built out of the leading terms corresponding to the points on the line segments of slope  $-r_d$ , in the Newton polygon  $NP([r_d])$ . We will have  $2e_d + 1$  points on this line segment. Let their co-ordinates be  $\alpha_s, \beta_s$  for  $s = 0, \dots, 2e_d$ .

$$Q_d(u) = \sum_{s=0}^{2e_d} \rho_{\alpha_s, \beta_s} u^{2e_d-s}$$

Note that when the first part in the Hitchin partition is even, the first term in the polynomial is just  $u^{2e_d}$  since  $\rho_{(0,N)} = 1$  by definition.

For special nilpotents,  $p_{2i-1}, p_{2i}$  have the same parity. In other words, even and odd parts occur in pairs. As derived in the proof of Prop 34 in Baraglia and Kamgarpour (2018), this implies that the polynomial  $Q_d$  is a perfect square. That is, we have an equation of the form

$$\sum_{s=0}^{2e_d} \rho_{\alpha_s, \beta_s} u^{2e_d-s} = \left( \sum_{n=0}^{e_d} \hat{a}_n u^{e_d-n} \right)^2 \quad (9.4)$$

Observe that the coefficients on the LHS involve just the leading terms in the  $t$  expansion of the coefficients in  $(\phi)$ . So, we could rewrite the above equation using notation introduced at the end of Section 9.2.2.

$$\sum_{s=0}^{2e_d} c_{2k+sr_d} u^{2e_d-s} = \left( \sum_{n=0}^{e_d} a_{k+nr_d} u^{e_d-n} \right)^2 \quad (9.5)$$

where  $2k$  is the following partial sum of the Hitchin partition  $[p_j]$

$$2k = \sum_{j \mid p_j > r_d} p_j \quad (9.6)$$

and  $\hat{a}_n(t) = a_{k+nr_d} t^{\frac{1}{2}\chi_{2k+2nr_d}} + \dots$  for some  $t$ -independent  $a_{k+nr_d}$ .

Since there are fewer independent coefficients on the polynomial on the RHS, this implies relations among the  $c_{2k+sr_d}$  on the LHS (called c-type constraints in CD).

We will henceforth use the notation  $c_i$  to refer to these leading coefficients where the  $i$  subscript encodes the  $\mathbb{C}^\star$  weight of the coefficient in the graded decomposition of the local Hitchin image.

We can use (9.5) to *solve* for as many of the  $a_n$ s as we can. Here is where CD and BK part company. BK eliminate the remaining (unsolved-for)  $a_n$ s, whereas CD treat them as independent gauge-invariant polynomials in the coefficients in the Higgs field, hence as generators of the local Hitchin base.

#### 9.4.2 Relationship between Chacaltana-Distler and Baraglia-Kamgarpour

We summarise the relationship between CD's description of the local Hitchin image and BK's description in the following theorem.

**Definition 9.2.** Given a Hitchin partition  $[p]$ , decompose it into blocks, with one block for

- each odd part (counted with multiplicity)
- each string of consecutive even parts

For both BK and CD, the image of the local Hitchin maps is a Cartesian product of factors, one factor for each block in the above decomposition.

$$S_{BK} = \prod_i S_{BK}^{(i)}, \quad S_{CD} = \prod_i S_{CD}^{(i)} \quad (9.7)$$

The  $S^{(i)}$  will be referred to as the *factors* in the local Hitchin image.

Consider the factor  $S^{(i)}$  corresponding to a particular string of consecutive even parts in  $[p]$ .

**Definition 9.3.** (BK's Hitchin image) BK describe  $S_{BK}^{(i)}$  as the space of  $\rho_{\alpha_s, \beta_s}$  subject to the constraints arising from equation (9.5). It describes a (possibly singular) subspace of co-dimension  $e_j$  in  $\mathbb{C}^{2e_j+1}$ .

**Definition 9.4.** (CD's Hitchin image) CD's description,  $S_{CD}^{(i)}$ , amounts to taking the space of  $a_n$  as the local Hitchin image. It is manifestly a vector space of dimension  $e_j + 1$ .

The map  $\gamma : S_{CD} \rightarrow S_{BK}$  respects the decomposition (9.7). That is, it consists of a collection of maps  $\gamma^{(i)} : S_{CD}^{(i)} \rightarrow S_{BK}^{(i)}$ .

*Observation 2.* Let  $\gamma : S_{CD} \rightarrow S_{BK}$  be the map between the two descriptions of the local Hitchin image and let  $\gamma^{(i)} : S_{CD}^{(i)} \rightarrow S_{BK}^{(i)}$  be the component corresponding to a particular string of consecutive even parts in  $[p]$ .

- i) If the block occurs at the beginning of the partition and consists of either a single even part (with even multiplicity) or is a very-even partition of the form  $[(2s)^{2l}, (2r)^2]$ , then the map  $\gamma^{(1)}$  is an isomorphism. Both  $S_{CD}^{(1)}$  and  $S_{BK}^{(1)}$  are smooth (in fact, vector spaces). For any other string of even parts at the beginning of the partition,  $S_{BK}^{(1)}$  is singular and  $\gamma^{(1)}$  is the normalization map. In particular,  $S_{CD}^{(1)}$  and  $S_{BK}^{(1)}$  are birational.
- ii) When the block occurs at the end of the partition (but does not extend to the beginning of the partition), then  $S_{BK}^{(i)}$  is singular and  $\gamma^{(i)}$  is the normalization map. In particular,  $S_{CD}^{(i)}$  and  $S_{BK}^{(i)}$  are birational.
- iii) When the block occurs in the middle of the partition (i.e. not at the beginning or the end) of the partition, then Conjecture 9.3 implies that  $\gamma^{(i)}$  is generically  $2 : 1$ .  $S_{CD}^{(i)}$  is a vector space,  $r : S_{CD}^{(i)} \rightarrow S_{CD}^{(i)}$  is reflection through the origin and  $S_{CD}^{(i)}/r$  is birational to  $S_{BK}^{(i)}$ .

Before presenting the proof, we refer the reader to the examples in Appendix 9.A, which may facilitate a better understanding of the proof.

*Proof of Theorem 2 (i).* For part (i) of the Theorem, consider the case where the string of consecutive even parts occurs at the beginning of the partition  $[p]$ . Say the

first part (with multiplicity) is  $(2s)^{2l}$ . The associated constraint equation reads

$$u^{2l} + c_{2s}u^{2l-1} + c_{4s}u^{2l-2} + \cdots + c_{4l} = (u^l + a_{2s}u^{l-1} + a_{4s}u^{l-2} + \cdots + a_{2l})^2$$

The equation for  $a_{2s}$  is linear:  $a_{2s} = \frac{1}{2}c_{2s}$ . Plugging its solution into the equation for  $a_{4s}$ , then becomes linear and can be solved. Proceeding in the same fashion, we solve for  $a_{2s}, \dots, a_{2l}$  which, in turn, allows us to solve for  $c_{2s(l+1)}, \dots, c_{4l}$  as polynomials in the  $c_{2s}, \dots, c_{2l}$ . If this is the only even part in the block, then we are done.

If this is a very-even partition of the form  $[(2s)^{2l}, (2r)^2]$ , then the second constraint equation

$$a_{2l}^2 u^2 + c_{4l+2r}u + \tilde{c}^2 = (a_{2l}u + a_{2l+2r})^2$$

which is solved by setting  $a_{2l+2r} = \pm \tilde{c}$  and  $c_{4l+2r} = \pm 2\tilde{c}a_{2l}$  (recall that we have already solved for  $a_{2l}$ ). The choice of sign corresponds to the two nilpotent orbits associated to this very-even partition.

In either of these two cases, we have solved for the  $a$ 's in terms of the  $c$ 's and  $S_{\text{CD}}^{(1)} \simeq S_{\text{BK}}^{(1)}$ . In any other case of a string of even parts at the beginning of the partition, we can continue to solve for the  $a$ 's, as we have done. However, we only obtain them as rational functions of the  $c$ 's. For concreteness, suppose  $[p] = [(2s)^{2l}, (2r)^{2m}, \dots]$ . We have already solved for  $a_{2l}$  as a polynomial in the  $c$ 's. The constraint associated to the second distinct even part in this block is

$$a_{2l}^2 u^{2m} + c_{4l+2r}u^{2m-1} + \cdots + c_{4l+4mr} = (a_{2l}u^m + a_{2l+2r}u^{m-1} + \cdots + a_{2l+2mr})^2$$

This is a linear equation for  $a_{2l+2r}$ , which is solved by setting  $a_{2l+2r} = \frac{c_{4l+2r}}{2a_{2l}}$ . Plugging this in, the equation for  $a_{2l+4r}$  becomes linear, ... In the end, we obtain expressions for  $a_{2l+4r}, \dots, a_{2l+2mr}$  which are rational functions of the  $c$ 's with some powers of  $a_{2l}$  as the denominator. In particular, we have that  $a_{2l+2mr}^2 = c_{4l+4mr}$  which is a polynomial on the Lie algebra. Since  $a_{2l+2mr}$  is a rational function on the Lie algebra and its square is a polynomial, we conclude that  $a_{2l+2mr}$  is a polynomial on the Lie algebra. For  $m = 1$ , we are done. For  $m = 2$ , the remaining constraint equation reads

$$c_{4l+4r} = 2a_{2l}a_{2l+4r} + a_{2l+2r}^2$$



Since the other terms are polynomial, we conclude that  $a_{2ls+2r}^2$  is polynomial and hence, since  $a_{2ls+2r}$  is rational, that  $a_{2ls+2r}$  is polynomial.

Things become progressively more complicated as we go to higher  $m$ . For  $m = 3$ , we have that  $a_{2ls}$  and  $a_{2ls+6r}$  are polynomials on the Lie algebra. For  $a_{2ls} \neq 0$ , we can solve for the rest of the  $a$ 's:

$$a_{2ls+2r} = \frac{1}{2a_{2ls}} c_{4ls+2r} \quad (9.8a)$$

$$a_{2ls+4r} = \frac{1}{8a_{2ls}^3} (4c_{4ls+4r}a_{2ls}^2 - c_{4ls+2r}^2) \quad (9.8b)$$

$$a_{2ls+6r} = \frac{1}{32a_{2ls}^5} (16c_{4ls+6r}a_{2ls}^4 - 4c_{4ls+2r}c_{4ls+4r}a_{2ls}^2 + c_{4ls+2r}^3) \quad (9.8c)$$

Since  $a_{2ls+6r}$  is a polynomial, (9.8c) can be rewritten as

$$\begin{aligned} 32a_{2ls}^5 a_{2ls+6r} &= 16c_{4ls+6r}a_{2ls}^4 - 4c_{4ls+2r}c_{4ls+4r}a_{2ls}^2 + c_{4ls+2r}^3 \\ &= 16c_{4ls+6r}a_{2ls}^4 - 8a_{2ls}^3 a_{2ls+4r} \end{aligned}$$

or

$$a_{2ls+4r} = 2c_{4ls+6r}a_{2ls} - 4a_{2ls}^2 a_{2ls+6r}$$

which proves that  $a_{2ls+4r}$  is a polynomial. Plugging (9.8a) into (9.8b), we have

$$a_{2ls+2r}^2 = c_{4ls+4r} - 2a_{2ls}a_{2ls+4r}$$

We just proved that the RHS is a polynomial. Since  $a_{2ls+2r}$  is rational, and its square is a polynomial, it too is a polynomial.

The procedure generalizes to higher  $m$ , though it becomes increasingly tedious to implement. We first solve for the  $a$ 's as rational functions of the  $c$ 's and  $a_{2ls}$  with a power of  $a_{2ls}$  as the denominator. Noting that  $a_{2ls+2mr}$  is a polynomial, we rearrange the expression for it to solve for  $a_{2ls+2(m-1)r}$ , which we see is a polynomial. Plugging that knowledge into the expression for  $a_{2ls+2(m-1)r}$ , we learn that  $a_{2ls+2(m-2)r}$  is a polynomial, etc. Proceeding in this fashion, we conclude that  $a_{2ls+2(m-2)r}, \dots, a_{2ls+2r}$  are also polynomials on the Lie algebra.

Away from some codimension-1 locus where  $a_{2ls} = 0$ , the map from the space of  $a$ 's to the space of  $c$ 's is invertible. On the locus  $a_{2ls} = 0$ , we have a set of coupled quadratic equations whose number of solutions is finite. If there are more even parts in this block, we continue as before. The  $a$ 's in the next block are rational functions of the  $c$ 's and  $a_{2ls+2mr}$ , with denominators that are powers of  $a_{2ls+2mr}$ .

We conclude that, away from a locus of codimension-1, the map  $\gamma^{(1)} : S_{\text{CD}}^{(1)} \rightarrow S_{\text{BK}}^{(1)}$  is 1-1. On that locus, it is finite-to-1. Over all of  $S_{\text{BK}}^{(1)}$ ,  $\gamma^{(1)}$  is birational and finite – hence a normalization.  $\square$

*Proof of Theorem 2 (ii).* More or less the same story holds for part (ii): a string of even parts at the end of the partition. Consider a partition of the form  $[p] = [\dots, (2s)^{2l}]$ . The constraint associated to this last even part is

$$c_{2N-4sl}u^{2l} + c_{2N-2s(2l-1)}u^{2l-1} + \dots + c_{2N-2s}u + \tilde{c}^2 = (a_{N-2sl}u^l + a_{N-2s(l-1)}u^{l-1} + \dots + a_{N-2s}u + a_N)^2$$

We can take  $a_N = \tilde{c}$ . Then we find  $a_{N-2s} = \frac{c_{2N-2s}}{2\tilde{c}}$ , and so on. The  $a$ 's are rational functions of the  $c$ 's, with denominators which are powers of  $\tilde{c}$ . The last equation then says that  $c_{2N-4sl} = a_{N-2sl}^2$ . Since  $c_{2N-4sl}$  is a polynomial on the Lie algebra,  $a_{N-2sl}$  must be denominator-free (as a rational function on the Lie algebra). Hence it (and the other  $a$ 's that we solved for) are also polynomials on the Lie algebra. The same considerations as before lead us to conclude that, for this block,  $\gamma^{(i)}$  is birational and finite – hence a normalization.  $\square$

*Proof of Theorem 2 (iii).* Finally, when the string of even parts occurs in the middle, the map  $\{a_j\} \rightarrow S_{\text{BK}}^{(i)}$  is generically 2-1, since if  $\{a_j\}$  is a solution, then so is  $\{-a_j\}$ . A-priori, now the  $a_j$  are only algebraic; for instance, the first one  $a_k = (c_{2k})^{1/2}$ . Here, we need to invoke Conjecture 9.3 which says that  $a_k$  is, in fact, a polynomial. We then proceed as before. In the same spirit as the previous cases, the map from the quotient  $S_{\text{CD}}^{(i)}/r \rightarrow S_{\text{BK}}^{(i)}$  is evidently birational and finite.  $\square$

We want to emphasize that the above theorem is purely about the local Hitchin image. We explore some of the subtleties that arise in studying the corresponding questions in the global setting in 9.6.1.

## 9.5 Odd Type Constraints

In this section, we turn our focus to the constraints associated with consecutive odd parts in the partition  $[p]$ , referred to as “odd-type constraints”. These constraints are defined in relation to the special Hitchin orbits and play a role in determining the structure of the Hitchin base. We will explore the mathematical formulation of these constraints and their implications, particularly in the context of gauge transformations and choices for local invariants.

The constraints of Section 9.4 had to do with strings of consecutive even parts in  $[p]$ . We now turn to the constraints associated to consecutive odd parts.

**Definition 9.5.** (Odd-type constraint) Let  $2k = \sum_{i=1}^{2j} p_i$ , for some  $j$ . We say that  $c_{2k}$ , for  $k \leq N$ , has an *odd-type constraint* iff  $p_{2j}$  and  $p_{2j+1}$  are odd and  $c_{2k}$  can be written as a perfect square of a polynomial function,  $a_k$ , on  $\mathfrak{g}$ .

**Definition 9.6.** (Marked pair) A **marked pair** is a pair of successive parts,  $p_{2j}, p_{2j+1}$ , where  $p_{2j}$  and  $p_{2j+1}$  are both odd and  $p_{2j} > p_{2j+1}$ .

It follows from Conjecture 9.3 that for every marked pair in the special Hitchin orbit  $\mathcal{O}_H$ , there is an odd type constraint  $c_{2k} = (a_k)^2$ . When such a constraint occurs, we have a choice in the (local) ring of invariant polynomials, we can either continue to use  $c_{2k}$  as the generator for the local ring of invariants or replace  $c_{2k}$  by  $a_k$ .

This situation is somewhat similar to what happens with  $c_{2N}$  for any Hitchin orbit with one key difference.  $c_{2N}$  is always a perfect square, whose square root is the Pfaffian  $\tilde{c}_N$ . The transformation that flips the sign of  $\tilde{c}_N$  is an element of  $O(2N)$  but not an element of  $SO(2N)$ . So, when we are studying the  $SO(2N)$  Hitchin system, we have no choice but to choose  $\tilde{c}_N$  as the generator for the ring of invariants.

In the case of odd type constraints, we have  $c_{2k} = a_k^2$  for some polynomial function  $a_k$  and  $kN$ . The subgroup of  $SO(2N)$  which preserves our chosen nilpotent is disconnected. The gauge transformations in that subgroup which flip the sign of  $a_k$  are not in the identity component. So, one has an **additional choice** (beyond deciding the group for the Hitchin system) of whether we choose to quotient, locally, by such transformations. Quotienting by such transformations corresponds to choosing  $c_{2k}$  as the local invariant while not quotienting corresponds to choosing  $a_k$  as the local invariant.

Depending on the choice we make, the corresponding local Hitchin image, which is always a graded vector space, will either have a contribution in degree  $2k$  or degree  $k$ . More generally, when  $l$  odd-type constraints occur (each for a different  $c_{2k}$ ) for given Hitchin nilpotent  $\mathcal{O}_H$ , we have  $2^l$  choices for the local Hitchin image. We say that the Hitchin image is spanned by the  $a_k$ 's when we chose  $a_k$ s as the local invariants and similarly for the  $c_k$ s.

### 9.5.1 Sommers-Achar subgroups

In this subsection, we introduce the notion of Sommers-Achar subgroups, which are subgroups of the group  $\overline{A}_b(\mathcal{O}_H)$ . These subgroups arise naturally when considering the geometry of the Hitchin base and are of particular interest due to their connection to the structure of the Hitchin image. We will define the group  $\overline{A}_b(\mathcal{O}_H)$  and explain the role of Sommers-Achar subgroups in the classification of Hitchin images.

**Definition 9.7.** (the group  $\overline{A}_b$ ) We define a finite group  $\overline{A}_b(\mathcal{O}_H)$  associated to every special Hitchin orbit  $\mathcal{O}_H$ .  $\overline{A}_b = (\mathbb{Z}_2)^d$ , where  $d$  is the number of marked pairs in  $[\mathcal{O}_H]$ . If we label the marked pair  $[\dots, (2r+1)_{2j}, (2s+1)_{2j+1}, \dots]$  by the positive integer  $j$  (the subscript denotes the position within the partition), then the generators of  $\overline{A}_b$  are labeled by the ordered set

$$S = \{j_1, j_2, \dots, j_d\} \tag{9.9}$$

with  $j_1 < j_2 < \cdots < j_d$ .

$\overline{A}_b(\mathcal{O})$  is a subgroup of Lusztig's quotient,  $\overline{A}(\mathcal{O})$  which, in turn, is a subgroup of the equivariant fundamental group  $A(\mathcal{O})$  (=component group of the centralizer) of the orbit  $\mathcal{O}$ . For the reader's convenience, we recall the definitions of these other groups in Appendix 9.B. The subscript 'b' in  $\overline{A}_b$  is meant to signify the fact that it is this subgroup of  $\overline{A}$  that is visible in the geometry of the Hitchin base. If Conjecture 9.3 holds, then we can take  $B_0$  to be the Hitchin image parametrized by the  $a_{k_i}$ . The  $i^{\text{th}}$  generator of  $\overline{A}_b(\mathcal{O})$  acts on  $B_0$  as  $a_{k_i} \rightarrow -a_{k_i}$ . In particular, the Hitchin image parametrized by the  $c_{2k}$ s is  $B_0/\overline{A}_b$ .

However, there are intermediate quotients that are of physical interest.

**Definition 9.8.** A *Sommers-Achar* subgroup  $C_i \subset \overline{A}_b(\mathcal{O}_H)$  is a subgroup whose generators are labeled by a subset  $S_i \subset S$ . Corresponding to the pair  $(\mathcal{O}_H, C_i)$ , we define the Hitchin image to be  $B_0/C_i$ .

The above definition is a sharper version of the one given in Chacaltana et al. (2013b). The terminology is motivated by the duality map due to Sommers Sommers (2001) and Achar-Sommers Achar and Sommers (2002) which, in turn, is a refinement of the duality maps of Lusztig-Spaltenstein Spaltenstein (1982) and Barbasch-Vogan Barbasch and Vogan Jr (1985). We should note here that the subgroups  $C_i$  arise in Sommers (2001), Achar and Sommers (2002) in a slightly different context. Their relevance to the geometry of the Hitchin image is one of the crucial insights from Chacaltana et al. (2013b).

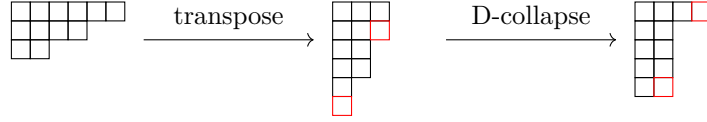
### 9.5.2 Special pieces

Here we introduce a duality operation on nilpotent orbits, which will help define the concept of special pieces and, consequently, non-special degenerations.

Recall that if  $[\mathcal{O}]$  is the partition label corresponding to a nilpotent orbit (not necessarily special), then duality acts by  $d : \mathcal{O} \rightarrow \mathcal{O}^\vee$  where  $[\mathcal{O}^\vee] := D([\mathcal{O}]^T)$ . Here,

$[\dots]^T$  denotes the transpose operation on partitions and  $D(\dots)$  is the D-collapse operation. For any  $\mathcal{O}$ ,  $\mathcal{O}^\vee$  is always a special nilpotent orbit, so  $d$  is neither one-one nor onto. Nevertheless,  $d^3 = d$ .

Here is an example of the duality operation. Take  $[\mathcal{O}] = [3^2, 2^2, 1^2]$ . We have  $[\mathcal{O}]^T = [6, 4, 2]$  and  $D([\mathcal{O}]^T) = [5^2, 1^2]$ . We depict the steps using Young diagrams below (where the column sizes encode the parts of a partition) :



**Note:** We will use  $D(\dots)$  to denote both the operation of D-collapse itself and the resulting special partition. It should be clear from the context which we mean.

For our given special nilpotent,  $O_H$ , we can ask about its pre-image under  $d$ .

**Definition 9.9.** (Special piece) A special piece consists of those nilpotent orbits in the pre-image of the duality map  $d$  which map to the same special nilpotent orbit  $O_H$  under  $d$ . Further,  $O_{N_0} := d(O_H)$  is the unique special nilpotent in the special piece.

In type  $D$ , it is known (see Theorem 4 below) that the size of special pieces are  $2^s$  for some natural number  $s$ . Furthermore,

**Definition 9.10.** (Non-special degeneration) We define a *non-special degeneration* to be a Kraft-Procesi small degeneration (see Appendix 9.D) that starts with a Nahm orbit  $\mathcal{O}$  and results in a non-special orbit  $\mathcal{O}'$  which is in the same special piece as  $\mathcal{O}$ . On the level of the corresponding partitions, a non-special degeneration is

$$\begin{aligned} & [\dots, (2r+1)_{2j}, (2r)_{2j+1}, \dots, (2r)_{2j+2m}, (2r-1)_{2j+2m+1}, \dots] \\ & \rightarrow [\dots, (2r)_{2j}, (2r)_{2j}, \dots, (2r)_{2j+2m}, (2r)_{2j+2m+1}, \dots]. \end{aligned} \tag{9.10}$$

We label it by the positive integer  $j$  (again, subscripts denote the location within the partition).  $m = 0$  corresponds to a degeneration of type (a);  $m > 0$  is a degeneration of type (g).

Starting with the special orbit  $O_{N_0} = d(O_H)$ , there are  $s$  independent non-special degenerations<sup>8</sup> that we can do. Let us label them by the set

$$S' = \{j_1, j_2, \dots, j_s\} \quad (9.11)$$

with  $j_1 < j_2 < \dots < j_s$ . Every orbit in the special piece is obtained by starting with the special orbit  $O_{N_0}$  and performing some set of such non-special degenerations. That is to say that the orbits  $O_{N_i}$  in the special piece are labeled by subsets  $S'_i \subset S'$ .

### 9.5.3 Special pieces and Hitchin images

Now, we are ready to prove one of the main theorems of this paper, regarding the number of marked pairs and non-special degenerations of a nilpotent orbit. As will become clear from this result, the elements in these sets will, in a certain sense, label choices of local Hitchin images. To achieve this, we will first prove a series of three lemmas.

Associated to each special Hitchin nilpotent,  $O_H$ , we have defined two ordered sets: the set  $S$  of marked pairs and the set  $S'$  of non-special degenerations in the dual special piece.

*Observation 3.* Given a special Hitchin orbit  $\mathcal{O}_H$  and the associated ordered sets

- $S$  of marked pairs in  $[O_H]$  and
- $S'$  of independent non-special degenerations of  $\mathcal{O}_{N_0} = d(\mathcal{O}_H)$ ,

we have:

- a) The ordered sets  $S$  and  $S'$  have the same number of elements.
- b) Let  $\sigma : S' \rightarrow S$  be the order-reversing bijection between them.  $\sigma$  induces a bijection between the set of nilpotent orbits  $\mathcal{O}_{N_i}$  in the dual special piece and

---

<sup>8</sup>see Appendix 9.D for more details on independent non-special degenerations.

the set of pairs  $(O_H, C_i)$ . By Conjecture 9.3, we can identify the latter with the set of Hitchin images  $B_0/C_i$ .

*Proof.* Once we have proven part (a), part (b) follows trivially from the definitions we have given. The  $\mathcal{O}_{N_i}$  are labeled by subsets  $S'_i \subset S'$  and the Hitchin images  $B_0/C_i$  are labeled by subsets  $S_i \subset S$ . So let us prove part (a).

Let  $d$  be the number of marked pairs (labelling the generators of  $\overline{A}_b(\mathcal{O}_H)$ ) and let  $s$  be the number of independent non-special degenerations of the dual special Nahm orbit  $\mathcal{O}_{N_0} = d(\mathcal{O}_H)$ . In other words  $|S| = d$  and  $|S'| = s$ . We proceed in two steps:

- We show  $s \geq d$  starting with a special Hitchin orbit with  $\overline{A}_b \neq 1$  and studying its dual special Nahm orbit (Lemma 2 below) .
- We show  $d \geq s$  using Kraft-Procesi's description of the size of the special piece in terms of number of *non-special degenerations* of the dual special Nahm orbit and then deducing the consequences for the special Hitchin orbit (Lemma 3 below).

First, we will need some combinatorial facts about how special nilpotent orbits in type-D behave under duality.

**Definition 9.11.** We define the notion of a **wall** for D-collapse operations. We say that the operation  $D(\lambda) = \sigma$  for some partition  $\lambda$  has a **wall** if there is a splitting of  $\lambda, \sigma$  of the form  $\lambda = \lambda_1 \mid \lambda_2$  ,  $\sigma = \sigma_1 \mid \sigma_2$  such that each part of  $\lambda_2$  is smaller than the smallest part of  $\lambda_1$  and

$$D(\lambda_i) = \sigma_i.$$

Here, the operation  $\lambda_1 \mid \lambda_2$  is defined to yield a partition  $\lambda$  that is formed by taking together the parts of  $\lambda_1$  and  $\lambda_2$ .



Let  $[q]$  be a partition and  $[\tilde{q}] = [q]^T$  be its transpose. We will prove that corresponding to every distinct even part in  $[q]$ , there is a wall in the operation  $D([\tilde{q}])$ .

*Lemma 2.* (The wall lemma)

1. Let  $\mathcal{O}$  be a special nilpotent orbit. Then, the operation  $D([\mathcal{O}]^T) = [\mathcal{O}^\vee]$  had  $l$  walls where  $l$  is the number of distinct even parts in  $[\mathcal{O}]$ .
2. The wall corresponding to an even part  $r_d^{e_d}$  occurs between the corresponding two parts  $\tilde{q}_j, \tilde{q}_{j+1}$  in  $[\mathcal{O}]^T$  that obey  $\tilde{q}_j - \tilde{q}_{j+1} = e_d$ . The wall divides  $[\mathcal{O}]^T$  into two pieces  $\lambda_1, \lambda_2$  such that  $\lambda_1 \mid \lambda_2 = [\mathcal{O}]^T$  and the dual special orbit  $d(\mathcal{O})$  is given by  $D(\lambda_1) \mid D(\lambda_2)$ . When the even part occurs at the beginning of  $[\mathcal{O}]$  we take  $\tilde{q}_{j+1} = 0$  and the wall in  $[\mathcal{O}]^T$  is between the last part and 0.

*Proof.* Let us also assume that  $[\mathcal{O}]$  is not very even. When  $[\mathcal{O}]$  is very even, D-collapse acts trivially on  $[\mathcal{O}]^T$  and the statement #1 of Lemma 1 is tautological.

Let us take the case where  $[\mathcal{O}]$  has a single distinct even part  $t$  occurring somewhere in the middle with multiplicity  $e$  for some even number  $e$ . So, the partition label  $[\mathcal{O}]$  is of the form

$$[\mathcal{O}] = [\dots, t^e, \dots]$$

and the transpose will be of the form

$$[\mathcal{O}]^T = [\tilde{q}_1, \dots, \tilde{q}_j, \tilde{q}_{j+1}, \dots]$$

where  $\tilde{q}_j$  and  $\tilde{q}_{j+1}$  are even numbers (because  $\mathcal{O}$  is special) that obey  $\tilde{q}_j - \tilde{q}_{j+1} = e$ . Now, since  $t$  is the only distinct even part in  $[\mathcal{O}]$ , it follows that the parts on both sides of  $t^e$  in  $[\mathcal{O}]$  will be odd parts. This implies that both  $\tilde{q}_j$  and  $\tilde{q}_{j+1}$  occur with odd multiplicity. Furthermore, the last part of  $[\mathcal{O}]$  will be an odd part and this implies

that the first part of  $[\mathcal{O}]^T$  will be an even part with odd multiplicity. In other words, the transpose partition actually has the form

$$[\mathcal{O}]^T = [\tilde{q}_1, \dots, \tilde{q}_j^{2r+1}, \tilde{q}_{j+1}^{2s+1}, \dots]$$

with  $\tilde{q}_1, \tilde{q}_j, \tilde{q}_{j+1}$  even. These are the only even parts that occur with odd multiplicity in  $[\mathcal{O}]^T$ . So, it is clear how D-collapse will act on  $[\mathcal{O}]^T$ .

$$D([\mathcal{O}]^T) = [\tilde{q}_1 - 1, \dots, \tilde{q}_j + 1, \tilde{q}_j^{2r}, \tilde{q}_{j+1}^{2s}, \tilde{q}_{j+1} - 1, \dots, 1]$$

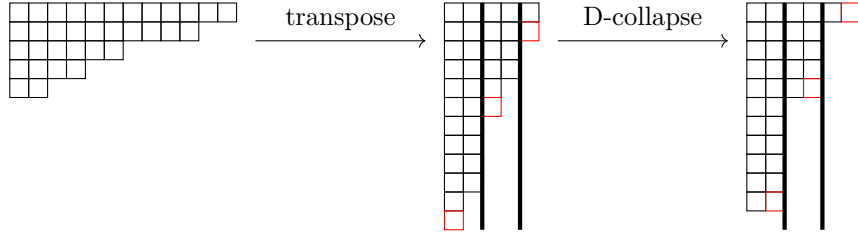
So, it follows that there is a wall between  $\tilde{q}_j$  and  $\tilde{q}_{j+1}$ .

The proof for multiple distinct even parts (occurring in the middle) proceeds in a similar way. The crucial point is that in the transpose, the corresponding  $\tilde{q}_j$  is always an even part that can not be a source of a D-collapse move. And the part  $\tilde{q}_{j+1}$  can not be a recipient.

When an even part  $t$  occurs as the last part of  $[\mathcal{O}]$ , the partition  $[\mathcal{O}]^T$  starts with  $[\tilde{s}_1^t, \dots]$  where  $\tilde{s}_1$  is an even number equal to the total number of parts in  $[\mathcal{O}]$ . Under D-collapse, the parts  $\tilde{s}_1^t$  will be unchanged. So, it follows that the D-collapse operation  $D([\mathcal{O}]^T)$  has a wall after the last occurrence of  $\tilde{s}_1$ .

When an even part  $t$  occurs as the first part of  $[\mathcal{O}]$  and with multiplicity  $e$ , the partition  $[\mathcal{O}]^T$  will end with  $[\dots, e]$  and there will be an odd number of even parts with odd multiplicities that are smaller than  $e$ . So, in the D-collapse operation, the part  $e$  will be turned into  $e + 1$ . It follows that there is a wall between  $e$  and 0.  $\square$

**Example** In terms of Young diagrams, the presence of a wall implies that no box is moved across wall during the D-collapse operation. Here is an example with  $[\mathcal{O}] = [5^2, 4^2, 3^2, 2^4, 1^2]$ .



Corresponding to the two distinct even parts in  $[\mathcal{O}]$ , we have two wall in the transpose  $[\mathcal{O}]^T = [12, 10|6, 4|2]$ . It is clear that no box moves across the walls during the subsequent D-collapse step.

We'll now prove  $s \geq d$ .

*Lemma 3.* Corresponding to every marked pair in a special nilpotent orbit, there is an independent *non-special degeneration* for the dual special Nahm orbit, i.e.  $s \geq d$ .

*Proof.* The statement is quite obvious when the special nilpotent orbit is very even. They have  $d = 0$  and the dual special piece is always trivial. So, we will not consider these below.

### Very odd orbits

We start instead with a *very odd* orbit with  $\bar{A}_b = (\mathbb{Z}_2)^d$ . This means that it has  $d$  marked pairs. Take the first marked pair from the right. Let it be of the form  $[\dots, p_{2j}, p_{2j+1}, \dots]$  where  $p_{2j}$  and  $p_{2j+1}$  are odd numbers that differ by some even number  $2k$  ie  $p_{2j} - p_{2j+1} = 2k$  for some  $k \neq 0$ . Now, the  $[\mathcal{O}]^T$  will have the following form :

$$[\mathcal{O}]^T = [(2m)^{2n+1}, \{\dots\}, (2j)^{2k}, \dots] \quad (9.12)$$

where  $2n + 1$  is the smallest part of  $[\mathcal{O}]$  and  $2m$  is the total number of parts in  $[\mathcal{O}]$  and the parts in  $\{\dots\}$  are all odd parts with even multiplicities. Now, D-collapse acts in the following way :

- If  $d = 0$ , we have  $j = 0$ . And D-collapse will yield  $[(2m)^{2n}, \dots, 1]$
- If  $d = 1$ , D-collapse yields  $[(2m)^{2n}, \dots, 2j + 1, (2j)^{2k-2}, 2j - 1, \dots, 1]$

It is clear that the part  $2j + 1$  in  $D([\mathcal{O}]^T)$  will have even parity. By Kraft-Procesi, the dual Nahm special orbit now admits a single small degeneration. If  $k = 1$ , it is a small degeneration of type (a) and if  $k > 1$ , it is a small degeneration of type (g) and the result of the deformation would yield a non-special orbit. We call (Def 9.10) such small degenerations *non-special degenerations*.

Now, when  $d > 1$ , we proceed in a similar way. We handle one marked pair in  $[\mathcal{O}]$  at a time. Each of these lead to a single non-special degeneration in  $D([\mathcal{O}]^T)$ . So, it is clear that  $D([\mathcal{O}]^T)$  will have at least  $d$  non-special degenerations.

### Orbits with even parts

Now, we consider special nilpotent orbits that have even parts. The story is similar except that the even parts in  $[\mathcal{O}]$  create walls in the D-collapse of  $[\mathcal{O}]^T$ .

Within the odd chunks  $u_i$  that are between the even parts, the argument proceeds as in the case of very odd orbits.

As a specific case, let us say there is one distinct even part in  $[\mathcal{O}]$ . Let us also assume that this even part is not the first or the last part. Corresponding to this, there is a wall in  $D([\mathcal{O}]^T)$  and there is a splitting of  $[\mathcal{O}]^T$  into two pieces  $\lambda_1, \lambda_2$  such that  $[\mathcal{O}]^T = \lambda_1 \mid \lambda_2$  and each part of  $\lambda_2$  is smaller than the smallest part of  $\lambda_1$ .

The first part of  $\lambda_i$  will necessarily be an even part occurring with odd multiplicity. In fact, the parts of  $\lambda_i$  will again be as in (9.12) above. So, the D-collapse of  $\lambda_i$  proceeds as in the previous section. Any marked pairs within  $\lambda_i$  will contribute the existence of an independent non-special degeneration in  $[\mathcal{O}^\vee]$ .

So, we get a single non-special degeneration in  $[\mathcal{O}^\vee]$  corresponding to every marked pair in  $[\mathcal{O}]$ .

So, we have shown that Lemma 3 is true, i.e.  $s \geq d$ .  $\square$

*Lemma 4.* Corresponding to every non-special degeneration in a special Nahm orbit, there is a marked pair in the dual Hitchin orbit, i.e.  $d \geq s$ .

*Proof.* Recalling Def 9.10, a non-special degeneration takes the form

$$[\dots, (2r+1)_{2j}, (2r)^{2m}, (2r-1)_{2j+2m+1}, \dots] \rightarrow [\dots, (2r)_{2j}, (2r)^{2m}, (2r)_{2j+2m+1}, \dots]. \quad (9.13)$$

The transpose of this statement is

$$[\dots, (2j+2m)_{2r}, (2j)_{2r+1}, \dots] \rightarrow [\dots, (2j+2m+1)_{2r}, (2j-1)_{2r+1}, \dots] \quad (9.14)$$

We now need to distinguish two cases:

$m > 0$ : a (g)-type non-special degeneration

$m = 0$ : an (a)-type non-special degeneration

When  $m > 0$ , by the wall lemma, the D-collapse of the first partition in (9.14) is

$$[\dots, (2j+2m+1)_{2r}, (2j-1)_{2r+1}, \dots],$$

so in both the source and the target we have the marked pair

$$[\dots, (2j+2m+1)_{2r}, (2j-1)_{2r+1}, \dots] \quad (9.15)$$

This, of course, is as it should be: by definition the dual of the two Nahm partitions in (9.13) is the *same* Hitchin partition.

When  $m = 0$ , we no longer have the wall lemma to show that

$$D([\dots, (2j)_{2r}, (2j)_{2r+1}, \dots]) = [\dots, (2j+1)_{2r}, (2j-1)_{2r+1}, \dots] \quad (9.16)$$

Nevertheless,

$$D([\dots, (2j+1)_{2r}, (2j-1)_{2r+1}, \dots]) = [\dots, (2j+1)_{2r}, (2j-1)_{2r+1}, \dots]$$

and since the two Nahm partitions in (9.13) map under duality to the same Hitchin partition, (9.16) continues to hold for the transposes.

Thus each non-special degeneration (9.13) of  $\mathcal{O}_N$  leads to a marked pair (9.15) in  $\mathcal{O}_H$ .  $\square$

Now consider  $\mathcal{O}$  a special Hitchin orbit in type-D with  $\bar{A}_b(\mathcal{O}) = (\mathbb{Z}_2)^d$ . By definition (Def 9.7) there are precisely  $d$  marked pairs. This proves part (a) of Theorem 3. Part (b) follows immediately.  $\square$

The 4d  $\mathcal{N} = 2$  SCFTs of class-S are labeled, in part, by the collection of *Nahm* nilpotent orbits,  $\mathcal{O}_N$ . The correspondence between Nahm nilpotent orbits  $\mathcal{O}_{N_i}$  in a special piece and *pairs*  $(\mathcal{O}_H, C_i)$  consisting of the dual Hitchin nilpotent  $\mathcal{O}_H = d(\mathcal{O}_{N_i})$  and a ‘‘Sommers-Achar’’ subgroup  $C_i$  was proposed in Chacaltana et al. (2013b). Their proposal was based on the observation that the 4d Coulomb branches of the  $\mathcal{N} = 2$  SCFTs have the form  $B_0/C_i$ . Theorem 3 offers a refinement of that proposal, by giving an *a priori* count of the number of orbits in the special piece along with a precise identification of the group  $\bar{A}_b(\mathcal{O})$  and its Sommers-Achar subgroups  $C_i$ .

#### 9.5.4 Independence of even and odd type constraints

We have devoted two separate sections to study the two different phenomena that are new when we study the Hitchin map in type-D. We have presented these as separate discussions without allowing for the possibility that they may interact/interfere with each other. It turns out this never happens.

Say we have a marked pair  $(p_{2j}, p_{2j+1})$  in the sense of Def 9.6. Let  $2k = \sum_{i=1}^{2j} p_i$ . There is an odd-type constraint for the term  $c_{2k}$ . Since  $p_{2j+1}$  is also odd, this guarantees that  $c_{2k}$  is not part of any system of even type constraints (of the kind we encountered in Section 9.4). So, even when odd and even type constraints occur in the local Hitchin image for the same special orbit  $\mathcal{O}$ , they necessarily involve

mutually distinct subsets of the  $c_{2k}$ . In this sense, the even and odd type constraints are independent of each other.

### 9.5.5 Odd type constraints, mass deformations and the Poisson integrable system

Throughout this paper, we have confined our study to the geometry of the Coulomb branch underlying a *conformal*  $\mathcal{N} = 2$  theory arising from a tame Hitchin system. However, there is a class of  $\mathcal{N} = 2$  preserving relevant deformations called *mass deformations* of the Coulomb branch geometry. Here, we briefly discuss the effect of these deformations on the geometry of the Hitchin system and their close connection with the phenomenon of odd-type constraints.

A key notion is that of a sheet, which is a connected component of the union of all orbits of a fixed dimension in the Lie algebra. Each such sheet contains a unique nilpotent orbit, sometimes referred to as the boundary of the sheet. However, outside of type A, a given nilpotent orbit can be the boundary of several sheets. This Lie-algebraic observation underlies the multiple Hitchin systems that we have encountered above.

In terms of the Hitchin system, the passage from a conformal theory to a mass deformed theory corresponds to allowing the residue of the Higgs field take generic values in a sheet in the Lie algebra Balasubramanian and Distler (2021b) whose boundary is the nilpotent orbit appearing in the description of the conformal limit. When we allow the mass deformation parameters to vary, they act as additional Casimirs (in the sense of Markman (1994)) and we get a Poisson integrable system.

Outside of type A, it is possible for a given nilpotent orbit to occur at the boundary of more than one sheet. Identifying the allowed mass deformations (i.e., identifying the Poisson integrable system in which the Hitchin system of our conformal theory is a symplectic leaf) is tantamount to a choice of one of those sheets. The procedure to identify the exact sheet corresponding to a local mass deformation was

outlined in Balasubramanian and Distler (2021b)<sup>9</sup>. For a given Hitchin nilpotent orbit  $\mathcal{O}_H$  we have the following pieces of data which are in 1-1 correspondence

1. a ‘‘Sommers-Achar’’ subgroup  $C \subset \overline{A}_b(\mathcal{O}_H)$ . For  $C = 1$ , the local Hitchin image has a factor of  $\mathbb{C}^s$ , on which  $\overline{A}_b(\mathcal{O}_H)$  acts; for the other choices of  $C$ , that factor in the Hitchin image is replaced by  $\mathbb{C}^s/C$ .
2. a ‘‘Nahm nilpotent orbit’’,  $\mathcal{O}_N$  in the dual special piece of  $\mathcal{O}_H$  (i.e., a nilpotent orbit such that  $d(\mathcal{O}_N) = \mathcal{O}_H$ )
3. a ‘‘special’’ sheet whose boundary is  $\mathcal{O}_H$ .

The correspondence between (1) and (2) was given in Theorem 3, where we related the set of Nahm nilpotent orbits in the dual special piece to ‘‘Achar-Sommers’’ subgroups  $C \subset \overline{A}_b(\mathcal{O}_H)$ . The correspondence between (2) and (3) is that the Bala-Carter Levi of the Nahm nilpotent orbit  $\mathcal{O}_N$  is the sheet Levi<sup>10</sup> of the special sheet.

The upshot is that the choice of which odd-type constraints to impose in the conformal case amounts to a choice of what the allowed mass deformations are (i.e., into which Poisson integrable system our conformal theory embeds as a symplectic leaf).

**Birational sheets:** In recent work, Losev has developed a theory of *birational sheets* Losev (2022), which is a refinement of the usual theory of sheets. We expect a close compatibility between Theorem 3 and Losev’s theory of birational sheets. In particular, we expect that the stabilizer of the Sommers-Achar subgroup  $C$  (as a subgroup in the component group  $A(\mathcal{O})$ ) is related to the degree of the generalized

---

<sup>9</sup>Interestingly, it was found that only *special sheets* (a notion defined in Balasubramanian and Distler (2021b)) arise as local mass deformations in Class S theories.

<sup>10</sup>The sheets in a complex Lie algebra  $\mathfrak{g}$  are labeled by pairs  $(\mathfrak{l}, \mathcal{O}_{\mathfrak{l}})$ , where  $\mathfrak{l}$  is a Levi subalgebra (which we call the ‘‘sheet Levi’’) and  $\mathcal{O}_{\mathfrak{l}}$  is a rigid nilpotent orbit in  $\mathfrak{l}$ . The (unique) nilpotent orbit  $\mathcal{O}$  in  $\mathfrak{g}$  which is the boundary of the sheet  $(\mathfrak{l}, \mathcal{O}_{\mathfrak{l}})$  is given by orbit induction :  $\mathcal{O} = \text{Ind}_{\mathfrak{l}}^{\mathfrak{g}}(\mathcal{O}_{\mathfrak{l}})$ . (See Appendix A.2 of Balasubramanian and Distler (2021b) for a review.)



Springer map(s) associated to a particular sheet. To make things more precise, we formulate the following conjecture.

Let  $\mathcal{O}$  be the special nilpotent orbit occurring in the conformal limit and let  $C \subset \bar{A}_b(\mathcal{O}) \subset A(\mathcal{O})$  be the Sommers-Achar subgroup associated to one of the Hitchin images corresponding to  $\mathcal{O}$  (according to Theorem 3). Denote by  $[A : C]$  the index of  $C$  as a subgroup of the component group  $A(\mathcal{O})$ . Now, let  $\mu$  be the generalized Springer map associated to the sheet corresponding to the mass deformed defect (identified as in Balasubramanian and Distler (2021b)). Our conjecture is that  $\deg(\mu) = [A : C]$ .

The case of a birational sheet corresponds to the special case where  $\mu = 1$ . In this case, the conjecture reduces to the claim that  $C = A(\mathcal{O})$  is a necessary and sufficient condition for a mass deformation to correspond to a birational sheet. If we have a non-trivial special piece on the Nahm side,  $C = A(\mathcal{O})$  can possibly occur only for the defect whose Nahm label is the smallest orbit in a special piece.<sup>11</sup> In other words, when multiple sheets share a nilpotent orbit as their boundary, the conjecture says that the birational sheet (if it exists) corresponds to the one whose sheet Levi is the Bala-Carter Levi of the smallest orbit in the dual special piece.

Although this paper is confined to discussing the case of type D Hitchin systems, we expect our Conjecture 9.5.5 to hold in the  $E_6, E_7, E_8$  cases as well.

In discussions with E. Sommers and J. Adams, we learnt of earlier work of J. Adams Adams where a duality between the smallest non-special orbit and the existence of birational Springer maps was first noticed from the point of view of the theory of unitary representations of non-compact groups. We believe that our conjecture above can offer a different perspective on this duality.

---

<sup>11</sup>It is a fact about the geometry of the special pieces that there is a unique such orbit.

## 9.6 Global Considerations

In this section, we discuss the global implications of the local constraints derived earlier. The total space of the global Hitchin integrable systems we consider are the moduli spaces of tame meromorphic Higgs bundles as formulated in Markman (1994); Bottacin (1995) for  $GL_n$  and generalized to arbitrary group  $G$  in Dalakov (2016).

### 9.6.1 Even type constraints

When the even part occurs in middle or at the end, we find a difference between the Hitchin bases in the CD approach and the BK approach. In a global context, these lead to *two distinct* Hitchin systems.

To every punctured Riemann surface, we associate a tame meromorphic Hitchin system of type  $\mathfrak{g}$  where each puncture is “decorated” by a nilpotent orbit in the Lie algebra  $\mathfrak{g}$ . Associated to this datum, there is a class-S theory. This is a 4D  $\mathcal{N} = 2$  superconformal field theory whose Coulomb-branch geometry is governed by the meromorphic Hitchin system. The Coulomb branch of this theory is  $B_{CD}$  together with its natural special Kähler structure.

On the other hand, we know that  $B_{BK}$  (when different from  $B_{CD}$ ) is the base of a *different* Hitchin system which may (or may not) arise from a 4D  $\mathcal{N} = 2$  theory.

#### 9.6.1.1 Example with even part at the end

Consider the Hitchin system for  $\mathfrak{g} = \mathfrak{so}(10)$  on the three-punctured sphere with Hitchin partitions  $[3^2, 2^2], [3^3, 1], [5, 3, 1^2]$ . Before imposing the constraint(s), the naive Hitchin base has dimension

$$b_k = (0, 0, 1, 1; 0)$$

for  $k = 2, 4, 6, 8; 5$ . The corresponding spectral curve (in the conventions of Appendix 9.C) is

$$\Sigma = \left\{ 0 = w^{10} + xy(x-y) \left( w^4 xy(x-y) c_6 + w^2 x^2 y^2 (x-y) c_8 \right) \right\}$$

Since  $\tilde{c} = 0$ , this curve is reducible, with a non-reduced  $w^2 = 0$  and an irreducible octic

$$\Sigma = \left\{ 0 = w^2 \left[ w^8 + x^2 y^2 (x-y)^2 (w^2 c_6 + xy c_8) \right] \right\}$$

From the  $2^2$  at the end of  $[3^2, 2^2]$ , we have constraints which take the form

$$c_6 = a_3^2$$

$$c_8 = 2a_5 a_3$$

$$\tilde{c} = a_5$$

B&K would eliminate the  $a_k$  and write these as the singular surface in  $\mathbb{C}^3$ ,

$$c_8^2 = 4c_6 \tilde{c}^2$$

Globally, we have  $\tilde{c} = 0$ . So this just amounts to saying that  $c_8 = 0$ , yielding the spectral curve

$$\Sigma_{\text{BK}} = \left\{ 0 = w^4 \left[ w^6 + c_6 x^2 y^2 (x-y)^2 \right] \right\}$$

fibered over the complex  $c_6$ -plane. C&D would *solve* the constraints, yielding  $c_6 = a_3^2$ ,  $c_8 = 0$ . The spectral curve is then

$$\Sigma_{\text{CD}} = \left\{ 0 = w^4 \left( w^3 + ia_3 xy(x-y) \right) \left( w^3 - ia_3 xy(x-y) \right) \right\}$$

fibered over the complex  $a_3$ -plane. The sextic curve has become reducible, with the two components being cubics. The map from C&D's Hitchin base to B&K's is just the double-cover map  $a_3 \rightarrow a_3^2 \equiv c_6$ .

The involution  $\iota : w \rightarrow -w$  exchanges the two cubics. So the Prym of the sextic is the Jacobian of one of the cubics. Indeed, from the work of Argyres et al Argyres et al. (2016) who exhaustively classified all the possible 1-dimensional Hitchin

bases, we know that the Hitchin system in this case is an elliptic fibration over the  $a_3$ -plane with an isolated  $IV^*$  singularity at the origin<sup>12</sup>. The entire family of cubic curves has a  $\mathbb{Z}_6$  symmetry under which  $(x, y, w, a_3) \rightarrow (y, x, e^{2\pi i/6}w, a_3)$ . Under this  $\mathbb{Z}_6$ , the Seiberg-Witten differential,

$$\lambda_{\text{SW}} = \frac{w(xdy - ydx)}{xy(x - y)} \quad (9.17)$$

transforms by the same 6<sup>th</sup>-root of unity. The family has another  $\mathbb{Z}_2$  symmetry  $(x, y, w, a_3) \rightarrow (y, x, w, -a_3)$ , under which the Seiberg-Witten differential is invariant<sup>13</sup>. The quotient by this  $\mathbb{Z}_2$  is — after suitably resolving the singularities of the central fiber — an elliptic fibration over the  $c_6 = a_3^2$  plane with an isolated  $II^*$  singularity at the origin. Away from the origin, the fibers are again cubic curves with a  $\mathbb{Z}_6$  symmetry under  $(x, y, w, c_6) \rightarrow (y, x, e^{2\pi i/6}w, c_6)$ . The monodromy around the origin in the  $a_3$ -plane has order-3; the monodromy around the origin in the  $c_6$ -plane has order-6, as once around the origin in the  $a_3$ -plane is twice around the origin in the  $c_6$ -plane. The  $IV^*$  monodromy is the square of the  $II^*$  monodromy.

The upshot is that CD's prescription yields the  $IV^*$  elliptic fibration, which is the Hitchin system governing the Coulomb branch of the  $E_6$  Minahan-Nemeschansky theory, whereas BK's prescription yields the  $II^*$  elliptic fibration, which is the Hitchin system governing the  $E_8$  Minahan-Nemeschansky theory. Both of them are perfectly fine 4D  $\mathcal{N} = 2$  theories; but it is the former that arises in the class-S construction from the puncture data given.

Now consider replacing  $[5, 3, 1^2]$  in the above example with  $[5^2]$ . Before imposing the constraints, the spectral curve is

$$\Sigma_{\text{BK}} = \left\{ 0 = w^{10} + x^2 y^2 (x - y)^2 \left( w^4 c_6 + w^2 x y c_8 + x^2 y^2 \tilde{c}^2 \right) \right\} \quad (9.18)$$

---

<sup>12</sup>This is the  $A_2$  Hitchin system on the 3-punctured sphere with regular nilpotent orbits at each of the punctures.

<sup>13</sup>Since we are describing an *elliptic* fibration, we should specify a choice of section ( $x = y$ ) which is also invariant under the  $\mathbb{Z}_2$ .

Now,  $\tilde{c}$  no longer vanishes and so the B&K's spectral curve is fibered over the genuinely singular base

$$B_{\text{BK}} = \{c_8^2 = 4c_6\tilde{c}^2\} \subset \mathbb{C}^3 \quad (9.19)$$

C&D solve the constraints and their Hitchin base is the affine space spanned by  $a_3$  and  $a_5 = \tilde{c}$ . Their spectral curve

$$\Sigma_{\text{CD}} = \left\{ 0 = \left( w^5 + ixy(x-y)(a_3w^2 + \tilde{c}xy) \right) \left( w^5 - ixy(x-y)(a_3w^2 + \tilde{c}xy) \right) \right\} \rightarrow B_{\text{CD}}$$

Again  $\Sigma_{\text{CD}}$  is reducible, with two quintic components, the involution  $\iota$  exchanges the two quintics and the Prym of  $\Sigma_{\text{CD}}$  is the Jacobian of one of the quintics. This rank-2 Hitchin system is easily recognized to be the one associated to the class-S theory of type  $A_4$  on the 3-punctured sphere with nilpotent orbits  $[3, 2]$ ,  $[3, 2]$ ,  $[5]$  at the punctures (this is called the  $R_{2,5}$  SCFT in Chacaltana and Distler (2010b)<sup>14</sup>).

$B_{\text{CD}}$  is the normalization of  $B_{\text{BK}}$ . The map  $f : B_{\text{CD}} \rightarrow B_{\text{BK}}$  is 1-1 everywhere away from the locus  $\tilde{c} = 0$ , where it is 2-1. We know there is a Hitchin fibration over  $B_{\text{CD}}$  and this Hitchin system is the one which governs the Coulomb branch geometry of the class-S theory of type  $D_5$ , with the above puncture data. What is unclear is whether there is a Hitchin system with base (9.19) and spectral curve (9.18) and, if there is, whether it is the Hitchin system associated to some 4D  $\mathcal{N} = 2$  SCFT.

#### 9.6.1.2 Example with even part in the middle

Consider the Hitchin system for  $G = SO(12)$  on a thrice punctured sphere with Hitchin partitions  $[3^2, 2^2, 1^2][3^4][5^2, 1^2]$ . The spectral curve takes the form

$$\Sigma_{\text{BK}} = \left\{ 0 = w^2[w^{10} + x^2y^2(x-y)^2(c_6w^4 + c_8xyw^2 + c_{10}x^2y^2)] \right\}$$

There is a single even-type constraint associated to the  $2^2$  in  $[3^2, 2^2, 1^2]$ , which yields

$$c_6u^2 + c_8u + c_{10} = (a_3u + a_5)^2$$

---

<sup>14</sup>More precisely, it is the  $R_{2,5}$  SCFT with two free hypermultiplets.

BK's Hitchin base is the singular quadric

$$B_{\text{BK}} = \{c_8^2 = 2c_6c_{10}\} \subset \mathbb{C}^3$$

and we have the spectral curve (9.6.1.2) fibered over this base. Of course, this singular quadric is just isomorphic to  $\mathbb{C}^2/\mathbb{Z}_2$ .

CD's Hitchin base is  $B_{\text{CD}} = \mathbb{C}^2$ , with coordinates  $(a_3, a_5)$ . In those coordinates, the spectral curve is (further) reducible

$$\Sigma_{\text{CD}} = \left\{ 0 = w^2[w^5 + ixy(x-y)(a_3w^2 + a_5xy)][w^5 - ixy(x-y)(a_3w^2 + a_5xy)] \right\}$$

Ignoring the trivial non-reduced  $w = 0$  component, we see that this is once again the Hitchin system for the  $R_{2,5}$  theory.

As we said, when the even part occurs in the middle of the partition, the map  $B_{\text{CD}} \rightarrow B_{\text{BK}}$  is generically 2:1. Here,  $B_{\text{BK}}$  is the orbifold  $B_{\text{BK}} = B_{\text{CD}}/\mathbb{Z}_2$ , where  $\mathbb{Z}_2 : (a_3, a_5) \rightarrow (-a_3, -a_5)$

### 9.6.2 Odd type constraints

When *odd-type* constraints occur, we want to assert that there is a *choice*. We can choose to take  $c_{2k}$  as a generator in the (local) ring of invariants or take  $a_k$  as the generator. As discussed in Section 9.5.5, this corresponds to a choice of what the allowed mass deformations are (or, equivalently, how to embed our symplectic Hitchin system in a Poisson integrable system).

This choice leads to distinct Hitchin systems, as we will illustrate with some examples below. As in the previous subsection, our conventions are those of Appendix 9.C.

Consider the 3-punctured sphere in the  $D_5$  theory, where the nilpotent orbits at  $p_1, p_2, p_3$  are given by the Hitchin partitions  $[2^4, 1^2], [5, 3, 1^2], [9, 1]$ . The spectral curve is a homogeneous decic; Because  $\tilde{\phi} = 0$ , the curve is reducible, with one non-reduced component  $w^2 = 0$ , and the other component is a homogeneous octic. Before imposing the constraints, the space of coefficients in the octic is 4-dimensional.

$$0 = w^2 \left[ w^8 + x^2 y (x - y) \left( c_4 w^4 + w^2 x y c_6 + x^2 y (x c_8 - y c'_8) \right) \right]$$

The constraints associated to  $[2^4, 1^2]$  impose

$$\begin{aligned} c_6 &= \frac{1}{2} c_2 (c_4 - \frac{1}{4} c_2^2) \\ c'_8 &= \frac{1}{4} (c_4 - \frac{1}{4} c_2^2)^2 \end{aligned} \tag{9.20}$$

But, since  $c_2 = 0$ , this yields  $c_6 = 0$  and  $c'_8 = \frac{1}{4} c_4^2$ . So the spectral curve is

$$0 = w^2 \left[ w^8 + x^2 y (x - y) \left( c_4 w^4 + x^2 y (x c_8 - \frac{1}{4} y c_4^2) \right) \right] \tag{9.21}$$

The odd-type constraint associated to  $[5, 3, 1^2]$  is  $c_8 = a_4^2$ .

- If we impose that constraint, we obtain the Hitchin system  $\mathcal{E}_7 \times \mathcal{E}_7$ , where  $\mathcal{E}_7 \rightarrow \mathbb{C}$  is the elliptic fibration with an isolated  $III^*$  singularity at the origin. Setting

$$c_4 = (b_1 + b_2), \quad a_4 = \frac{1}{2} (b_1 - b_2)$$

the Hitchin bases for the two copies of  $\mathcal{E}_7$  are parameterized by  $b_1, b_2$  respectively.

- If we don't impose the constraint, we obtain the Hitchin system  $\text{Hilb}(2, \mathcal{E}_7)$ , the Hilbert Scheme of two points on  $\mathcal{E}_7$ .

As another (simpler) example, consider the 3-punctured sphere in the  $D_4$  theory with Hitchin partitions  $[3^2, 1^2], [3^2, 1^2], [7, 1]$ . Here, the spectral curve is

$$0 = w^2 \left[ w^6 + x^2 y^2 (x - y) (x c_6 - y c'_6) \right] \tag{9.22}$$

The two  $[3^2, 1^2]$  partitions have odd-type constraints associated to them

$$c_6 = (a_1 + a_2)^2, \quad c'_6 = (a_1 - a_2)^2 \tag{9.23}$$

- If we impose both constraints, then Ergun et al. (2021) showed by explicit computation that the Hitchin system we obtain is  $\mathcal{E}_6 \times \mathcal{E}_6$ , where  $\mathcal{E}_6 \rightarrow \mathbb{C}$  is the elliptic fibration with an isolated  $IV^*$  singularity at the origin. The Hitchin bases of the two  $\mathcal{E}_6$ s are parametrized by  $a_1$  and  $a_2$  respectively.
- If we impose just one of the constraints, we obtain the Hitchin system  $\text{Hilb}(2, \mathcal{E}_6)$ .
- If we impose neither constraint in (9.23), we obtain a Hitchin system whose succinct description eludes us at present.

## Appendix 9.A: Illustrative Examples for Theorem 2

The following examples illustrate the statements and proof of **Theorem 2** and help the reader follow its key steps.

1. Consider the partition  $[4^2, 1^4]$ . The string of even parts is at the beginning of the partition. Consequently,  $Q_1(u)$  has a leading term with coefficient 1.

$$u^2 + c_4u + c_8 = (a_0u + a_4)^2$$

This can be solved by setting  $a_0 = 1$ ,  $a_4 = c_4/2$ . Then we have a *constraint*,  $c_8 = \frac{1}{4}c_4^2$ . Eliminating  $c_8$ , CD and BK agree:  $S_{\text{CD}} = S_{\text{BK}}$

2. Consider the partition  $[2^6]$ . This is a very-even partition, so it consists of a single block of even parts. Since it starts with an even part,  $Q_1(u)$  has a leading term with coefficient 1.

$$u^6 + c_2u^5 + c_4u^4 + c_6u^3 + c_8u^2 + c_{10}u + \tilde{c}^2 = (a_0u^3 + a_2u^2 + a_4u + a_6)^2$$

We set  $a_0 = 1$  and solve for  $a_2 = c_2/2$ . The subsequent terms yield



$$a_2^2 + 2a_0a_4 = c_4$$

$$2(a_0a_6 + a_2a_4) = c_6$$

$$2a_2a_6 + a_4^2 = c_8$$

$$a_6 = \pm \tilde{c}$$

which can be successively solved for  $a_{4,6}$ :

$$a_4 = \frac{1}{2} \left( c_4 - \frac{1}{4}c_2^2 \right)$$

$$a_6 = \frac{1}{2}c_6 - \frac{1}{4}c_2(c_4 - \frac{1}{4}c_2^2)$$

Plugging back in yields the constraints

$$\begin{aligned} c_6 \mp 2\tilde{c} &= \frac{1}{2}c_2(c_4 - \frac{1}{4}c_2^2) \\ c_8 &= \frac{1}{4} \left( c_4 - \frac{1}{4}c_2^2 \right)^2 \pm c_2\tilde{c} \\ c_{10} &= \pm \tilde{c}(c_4 - \frac{1}{4}c_2^2) \end{aligned} \tag{9.24}$$

The two solutions for  $a_6$ , which yield the two choices of sign in (9.24), correspond to the two nilpotent orbits associated to this very-even partition. Again, CD and BK agree.

3. Consider the partition  $[4^2, 2^2]$ . Again, the partition is very-even, so there is a single block of even parts. However, there are two distinct even parts, so we get two perfect squares. Again, the leading part is even, so  $Q_1(u)$  has leading coefficient 1.

$$u^2 + c_4u + c_8 = (a_0u + a_4)^2$$

$$c_8u^2 + c_{10}u + \tilde{c}^2 = (a_4u + a_6)^2$$

Here, we set  $a_0 = 1$  and  $a_4 = c_4/2$ , which yields the constraint  $c_8 = \frac{1}{4}c_4^2$ . Plugging that into the second equation, we have  $a_4 = c_4/2$  and  $a_6 = \pm \tilde{c}$ . This yields the constraint  $c_{10} = \pm \tilde{c}c_4$ . We thus have two solutions, corresponding to the two nilpotent orbits determined by this very-even partition.

4. Consider the very-even partition  $[4^2, 2^4]$ . We again get two perfect squares

$$\begin{aligned} u^2 + c_4 u + c_8 &= (u + a_4)^2 \\ c_8 u^4 + c_{10} u^3 + c_{12} u^2 + c_{14} u + \tilde{c}^2 &= (a_4 u^2 + a_6 u + a_8)^2 \end{aligned}$$

Both BK and CD can solve for  $a_4 = \frac{1}{2}c_4$  and thence for  $c_8$

$$c_8 = \frac{1}{4}c_4^2$$

From the second equation, we have two possible solutions for  $a_8$ ,  $a_8 = \pm\tilde{c}$ , corresponding to the two choice of nilpotent orbit determined by this very-even partition. The remaining constraint equations read

$$\begin{aligned} c_{10} &= c_4 a_6 \\ c_{12} &= a_6^2 \pm c_4 \tilde{c} \\ c_{14} &= \pm 2\tilde{c} a_6 \end{aligned} \tag{9.25}$$

CD take the local Hitchin base to be the vector space spanned<sup>15</sup> by  $c_2, c_4, c_6, a_6, \tilde{c}$ . BK eliminate  $a_6$  from (9.25) and write the local Hitchin base as the singular affine variety

$$S_{\text{BK}} = \left\{ \begin{array}{l} c_{14}^2 - 4c_{12}\tilde{c}^2 \pm 4c_4\tilde{c}^3 = 0 \\ c_{10}c_{14} \mp 2c_4c_{12}\tilde{c} + 2c_4^2\tilde{c}^2 = 0 \\ c_{10}^2 - c_4^2c_{12} \pm c_4^3\tilde{c} = 0 \\ c_4c_{14} \mp 2c_{10}\tilde{c} = 0 \end{array} \right\} \subset \mathbb{C}^7$$

Away from the singular locus

$$\{c_4 = c_{10} = c_{14}^2 - 4c_{12}\tilde{c}^2 = 0\}$$

the map  $\gamma : S_{\text{CD}} \rightarrow S_{\text{BK}}$  is 1-1. On the singular locus, it is generically 2-1 (except on the line  $\{c_4 = c_{10} = c_{12} = c_{14} = 0\}$  where it is 1-1), hence a normalization.

---

<sup>15</sup> $c_2, c_6$  don't participate in the constraints, so CD and BK agree that they form a vector subspace of the local Hitchin base.

5. Consider the partition  $[5, 3, 2^2]$ . This partition has the block of even parts at the end. Hence CD and BK disagree. We have

$$c_8 u^2 + c_{10} u + \tilde{c}^2 = (a_4 u + a_6)^2$$

We can solve for  $a_6 = \tilde{c}$ . But the remaining constraints then read

$$c_8 = a_4^2, \quad c_{10} = 2a_4 \tilde{c} \quad (9.26)$$

CD say that  $a_4$  and  $\tilde{c}$  parametrize the Hitchin image (instead of  $c_8, c_{10}$  and  $\tilde{c}$ ). BK eliminate  $a_4$  and say this factor in the Hitchin image is parametrized by  $c_8, c_{10}$  and  $\tilde{c}$ , with the constraint  $c_{10}^2 = 4c_8 \tilde{c}^2$ .

The two spaces,  $\mathcal{S}_{\text{CD}}$  and  $\mathcal{S}_{\text{BK}}$  are birational. The map  $\mathcal{S}_{\text{CD}} \rightarrow \mathcal{S}_{\text{BK}}$  is 1:1 away from the locus  $\{\tilde{c} = 0, c_8 \neq 0\}$ , where it is 2:1. Again,  $\mathcal{S}_{\text{CD}}$  is the normalization of  $\mathcal{S}_{\text{BK}}$

6. Consider the partition  $[5^2, 4^2, 2^4, 1^2]$ . We again get two perfect squares

$$\begin{aligned} c_{10} u^2 + c_{14} u + c_{18} &= (a_5 u + a_9)^2 \\ c_{18} u^4 + c_{20} u^3 + c_{22} u^2 + c_{24} u + c_{26} &= (a_9 u^2 + a_{11} u + a_{13})^2 \end{aligned}$$

In this case equating coefficients in the polynomial equations yields:

$$\begin{aligned} c_{10} &= a_5^2 \\ c_{14} &= 2a_5 a_9 \\ c_{18} &= a_9^2 \\ c_{20} &= 2a_9 a_{11} \\ c_{22} &= a_{11}^2 + 2a_9 a_{13} \\ c_{24} &= 2a_{11} a_{13} \\ c_{26} &= a_{13}^2 \end{aligned}$$

Here BK would eliminate  $a_5, a_9, a_{11}$  and  $a_{13}$  and write this factor in the local Hitchin base,  $S^{(i)}$ , as the singular affine variety which is the intersection of a

quadric and seven cubics in  $\mathbb{C}^9$  (spanned by  $c_{10}, c_{12}, \dots, c_{26}$ ).

$$S = \left\{ \begin{array}{l} c_{14}^2 - 4c_{10}c_{18} = 0 \\ c_{20}^3 - 4c_{18}c_{20}c_{22} + 8c_{18}^2c_{24} = 0 \\ c_{20}^2c_{22} - 4c_{18}c_{22}^2 + 2c_{18}c_{20}c_{24} + 16c_{18}^2c_{26} = 0 \\ c_{20}^2c_{24} - 4c_{18}c_{22}c_{24} + 8c_{18}c_{20}c_{26} = 0 \\ c_{18}c_{24}^2 - c_{20}^2c_{26} = 0 \\ c_{20}c_{24}^2 - 4c_{20}c_{22}c_{26} + 8c_{18}c_{24}c_{26} = 0 \\ c_{22}c_{24}^2 - 4c_{22}^2c_{26} + 2c_{20}c_{24}c_{26} + 16c_{18}c_{26}^2 = 0 \\ c_{24}^3 - 4c_{22}c_{24}c_{26} + 8c_{20}c_{26}^2 = 0 \end{array} \right\} \subset \mathbb{C}^9 \quad (9.27)$$

while CD would take  $S^{(i)}$  to be the vector space  $\mathbb{C}^6$  spanned by  $a_5, a_9, a_{11}, a_{13}, c_{12}, c_{16}$ .

The map  $\gamma^{(i)} : S_{\text{CD}}^{(i)} \rightarrow S_{\text{BK}}^{(i)}$ , is 2-1 away from the singular locus

$$\{c_{14} = c_{18} = c_{20} = c_{24}^2 - 4c_{22}c_{26} = 0\}$$

of (9.27), where it is finite-1. If  $r : \{a_5, a_9, a_{11}, a_{13}\} \mapsto \{-a_5, -a_9, -a_{11}, -a_{13}\}$ , then  $S_{\text{CD}}^{(i)}/r$  is birational to  $S_{\text{BK}}^{(i)}$ .

7. Finally, let us consider a partition with three distinct blocks of even parts, one at the beginning, one in the middle and one at the end:  $[6^2, 5^2, 4^4, 3^2, 2^2]$  in  $D_{24}$ . We have

$$\begin{aligned} u^2 + c_6u + c_{12} &= (a_0u + a_6)^2 \\ c_{22}u^4 + c_{26}u^3 + c_{30}u^2 + c_{34}u + c_{38} &= (a_{11}u^2 + a_{15}u + a_{19})^2 \\ c_{44}u^2 + c_{46}u + \tilde{c}^2 &= (a_{22}u + a_{24})^2 \end{aligned}$$

The sign ambiguity among the  $a$ 's in the first equation can be resolved by setting  $a_0 = 1$ . Then we have  $2a_6 = c_6$ ,  $a_6^2 = c_{12}$ . Or

$$c_{12} = \frac{1}{4}c_6^2.$$

On this, CD and BK agree.

Similarly, among  $a_{22}, a_{24}$ , the sign ambiguity can be fixed by setting  $a_{24} = \tilde{c}$ .

Then we have  $2a_{22}\tilde{c} = c_{46}$ ,  $a_{22}^2 = c_{44}$ .

BK would eliminate  $a_{22}$  from these equations and write the singular hypersurface in  $\mathbb{C}^3$  (spanned by  $c_{44}, c_{46}, \tilde{c}$ )

$$4c_{44}\tilde{c}^2 = c_{46}^2$$

CD would construct the vector space,  $\mathbb{C}^2$ , spanned by  $a_{22}$  and  $\tilde{c}$ . The map to BK's space is given by

$$c_{44} = a_{22}^2$$

$$c_{46} = 2a_{22}\tilde{c}$$

$$\tilde{c} = \tilde{c}$$

This map is 1-1 away from the locus  $\{\tilde{c} = 0, c_{44} \neq 0\}$ . On that locus, it is 2-1.

Finally, let us turn to the block of even parts in the middle. Both CD and BK agree on the vector space factor, spanned by  $c_{24}, c_{28}, c_{32}, c_{36}$ . The constraints on the remaining  $c$ 's yield

$$c_{22} = a_{11}^2$$

$$c_{26} = 2a_{11}a_{15}$$

$$c_{30} = 2a_{11}a_{19} + a_{15}^2$$

$$c_{34} = 2a_{15}a_{19}$$

$$c_{38} = a_{19}^2$$

BK eliminate  $a_{11}, a_{15}, a_{19}$  and write  $S \subset \mathbb{C}^5$  as the singular intersection of 7 cubics:

$$S = \left\{ \begin{array}{l} c_{22}^3 - 4c_{26}c_{22}c_{30} + 8c_{26}^2c_{34} = 0 \\ c_{22}^2c_{30} - 4c_{26}c_{30}^2 + 2c_{26}c_{22}c_{34} + 16c_{26}^2c_{38} = 0 \\ c_{22}^2c_{34} - 4c_{26}c_{30}c_{34} + 8c_{26}c_{22}c_{38} = 0 \\ c_{26}c_{34}^2 - c_{22}^2c_{38} = 0 \\ c_{22}c_{34}^2 - 4c_{22}c_{30}c_{38} + 8c_{26}c_{34}c_{38} = 0 \\ c_{30}c_{34}^2 - 4c_{30}^2c_{38} + 2c_{22}c_{34}c_{38} + 16c_{26}c_{38}^2 = 0 \\ c_{34}^3 - 4c_{30}c_{34}c_{38} + 8c_{22}c_{38}^2 = 0 \end{array} \right\} \subset \mathbb{C}^5$$

CD take  $S$  to be the vector space  $\mathbb{C}^3$ , spanned by  $a_{11}, a_{15}, a_{19}$ .

## Appendix 9.B: The Equivariant Fundamental Group and Lusztig's Quotient

For the benefit of readers, we recall the formulas that describe the equivariant  $G_{ad}$ -fundamental group/component group  $A(\mathcal{O})$  and Lusztig's quotient  $\overline{A}(\mathcal{O})$ .

- $A(\mathcal{O})$  in type D (see (Collingwood and McGovern, 1993, pp. 92) ) : Let  $a$  be the number of distinct odd parts in the partition label  $[\mathcal{O}]$ . Then,  $A(\mathcal{O}) = (\mathbb{Z}_2)^m$  where  $m = \max(a - 1, 0)$  if all odd parts have even multiplicity and  $m = \max(a - 2, 0)$  otherwise.
- $\overline{A}(\mathcal{O})$  in type D (originally defined in Lusztig (1984), we use the summary provided in Sommers (2001) and Achar and Sommers (2002)) : Let  $[\mathcal{O}] = [\lambda_1^{a_1}, \lambda_2^{a_2}, \dots]$  and let  $\nu_i = \sum_{j=1}^i a_j$ . Now, identify the odd parts in  $[\mathcal{O}]$  for which  $\nu_i$  is odd. These are sometimes called *marked parts*. Form a string of integers  $\mu$  using the marked parts. Let  $m$  be the number of independent, non-empty, even cardinality subsets of  $\mu$ . We have that  $\overline{A}(\mathcal{O}) = (\mathbb{Z}_2)^m$ .

Note that the notions of a marked pair in Def 9.6 and that of a *marked part* are related but different. Every marked pair has an underlying marked part but not

every marked part is part of a marked pair. This will play an important role in what follows.

From the prescription to compute Lusztig's quotient and the definition of  $\bar{A}_b(\mathcal{O})$  in Def 9.7, it is clear that  $\bar{A}_b \subset \bar{A}(\mathcal{O})$ . More generally, we have

$$\bar{A}_b \subseteq \bar{A}(\mathcal{O}) \subseteq A(\mathcal{O})$$

Here is a table with some illustrative examples :

Table 9.1: Examples illustrating the inclusions  $\bar{A}_b \subseteq \bar{A}(\mathcal{O}) \subseteq A(\mathcal{O})$  for various partitions  $\mathcal{O}$ .

| $\mathcal{O}$                  | $\bar{A}_b$        | $\bar{A}$          | $A$                |
|--------------------------------|--------------------|--------------------|--------------------|
| $[3^2, 1^2]$                   | $\mathbb{Z}_2$     | $\mathbb{Z}_2$     | $\mathbb{Z}_2$     |
| $[7, 5, 3, 1]$                 | $\mathbb{Z}_2$     | $\mathbb{Z}_2$     | $(\mathbb{Z}_2)^2$ |
| $[3^2, 2^2, 1^2]$              | 1                  | $\mathbb{Z}_2$     | $\mathbb{Z}_2$     |
| $[11, 9, 7, 5^2, 3^2, 1]$      | $\mathbb{Z}_2$     | $\mathbb{Z}_2$     | $(\mathbb{Z}_2)^4$ |
| $[11, 9, 8^2, 7, 5^2, 3^2, 1]$ | 1                  | $\mathbb{Z}_2$     | $(\mathbb{Z}_2)^4$ |
| $[11, 9, 7, 5, 4^2, 3^2, 1^2]$ | $(\mathbb{Z}_2)^2$ | $(\mathbb{Z}_2)^4$ | $(\mathbb{Z}_2)^5$ |

## Appendix 9.C: Hitchin Systems for 3-Punctured Spheres

Here we set out our conventions for meromorphic Hitchin systems Markman (1994); Bottacin (1995); Dalakov (2016) on a punctured curve  $C$ . In particular, we will give a very concrete model for type- $D_N$  on a 3-punctured sphere.

Let  $D = \sum_i p_i$  be the divisor corresponding to the punctures. A  $D_N$  meromorphic Higgs bundle is a pair  $(E, \Phi)$ , composed of a  $D_N$ -bundle  $E$  on  $C$  and a holomorphic section  $\Phi$  of  $ad(E) \otimes L$ , where  $L = K_C(D)$ , called the Higgs field. We fix  $\Phi(p_i) \in O_i$ , where the  $O_i$  are a collection of nilpotent orbits in  $\mathfrak{so}(2N)$ .

The moduli space of Higgs bundles on  $C$ ,  $\mathcal{M}_{Higgs}$ , fibers over the Hitchin base. In type- $A_{N-1}$  Balasubramanian et al. (2022b), the Hitchin base is the graded vector space  $\bigoplus_{k=2}^N H^0\left(C, L^{\otimes k}(-\sum_i \chi_i^{(k)} p_i)\right)$ . The  $\chi_i^{(k)}$  are positive integers which depend on

the choice of nilpotent orbits,  $O_i$ . For the regular nilpotent orbit  $\chi^{(k)} = 1$ , whereas for the minimal nilpotent orbit  $\chi^{(k)} = k - 1$ .

More concretely, the sections  $\phi_k \in H^0\left(C, L^{\otimes k}(-\sum_i \chi_i^{(k)} p_i)\right)$  are symmetric polynomials in the Higgs field  $\Phi$

$$\phi_k = s_k(\Phi) = \frac{1}{k} \text{Tr}(\Phi^k) + \dots$$

In type- $D_N$ , the  $\phi_k$  vanish for odd  $k$  and  $\phi_{2N} = \tilde{\phi}^2$ , where

$$\tilde{\phi} = \text{Pfaff}(\Phi)$$

As sections of  $L^{\otimes k}$ , the  $\phi_k$  vanish to order  $\chi_i^{(k)}$  at  $p_i$ . In type- $D_N$  (in contrast to type- $A$ ), the leading coefficients of the  $\phi_k$  at  $p_i$  obey local constraints that are the subject of this paper. The Hitchin base, in type- $D_N$  is the space of solutions to those constraints.

The spectral curve  $\Sigma \rightarrow C$  is a hypersurface in the total space of  $L$

$$\Sigma = \left\{ 0 = w^{2N} + \sum_{k=1}^{N-1} w^{2(N-k)} \phi_k + \tilde{\phi}^2 \right\} \quad (9.28)$$

where  $w$  is the fiber coordinate on  $L$ . Every such  $D_N$  spectral curve admits an involution  $\iota : \Sigma \rightarrow \Sigma$ , under which  $w$  and the Seiberg-Witten differential are odd

$$\iota^*(w) = -w, \quad \iota^*(\lambda_{\text{SW}}) = -\lambda_{\text{SW}}$$

Let  $\pi : \Sigma \rightarrow \Sigma/\iota$  be the projection; the fibers of the Hitchin map are the Prym variety  $\text{Prym}(\pi)$ .

For the 3-punctured sphere, we can be even more concrete. Let  $x, y$  be homogeneous coordinates on  $C = \mathbb{CP}^1$ , such that the three punctures are located at  $p_1 = \{x = 0\}$ ,  $p_2 = \{y = 0\}$  and  $p_3 = \{x = y\}$ , respectively. The line bundle<sup>16</sup>

---

<sup>16</sup>If we take the homogeneous coordinates on  $\mathbb{CP}^2$  to be  $(x, y, w)$ , then the total space of  $L$  can be thought of as  $\mathbb{CP}^2 \setminus (0, 0, 1)$ .



$L = \mathcal{O}(1) \rightarrow \mathbb{CP}^1$ . The  $\phi_k(x, y)$  are then homogeneous polynomials of degree  $2k$ . Choosing three arbitrary Hitchin partitions, one for each puncture, the spectral curve takes the form of a homogeneous polynomial of degree  $2N$  in the variables  $x, y, w$

$$\Sigma = \left\{ 0 = w^{2N} + \sum_{k=1}^{N-1} \phi_{2k}(x, y) w^{2(N-k)} + \tilde{\phi}(x, y)^2 \right\} \quad (9.29)$$

Overall, the polynomial in (9.29) is homogeneous, of degree  $2N$  in the variables  $(x, y, w)$ .

For three regular nilpotents, the  $\phi_{2k}$  vanish at least linearly at each of the points and we can write

$$\phi_{2k} = xy(x - y)f_{2k}(x, y), \quad \tilde{\phi} = xy(x - y)\tilde{f}(x, y)$$

where the  $f_{2k}$  are homogeneous polynomials of degree  $2k - 3$  and  $\tilde{f}$  is a homogeneous polynomial of degree  $N - 3$ . The Seiberg-Witten differential is a meromorphic 1-form on the total space of  $L$ . In the case at hand, it has a simple expression in homogeneous coordinates

$$\lambda_{SW} = \frac{w(xdy - ydx)}{xy(x - y)}$$

and is manifestly odd under  $\iota$ . There is a second  $\mathbb{C}^*$  action, under which  $w$  scales with weight-1 and the *coefficients* in the polynomials  $f_{2k}(x, y)$  scale with weight- $2k$  (and the coefficients in  $\tilde{f}(x, y)$  scale with weight- $N$ ). As a consequence,  $\lambda_{SW}$  scales with weight-1. Since the coefficients in the polynomials parameterize the Hitchin base, the latter is graded by this  $\mathbb{C}^*$  action and the constraints are homogeneous.

## Appendix 9.D: The Work of Kraft-Procesi on Special Pieces

In this appendix we will briefly review the main result of Kraft and Procesi (1989) and some of the tools used in the proof of this result.

To begin, we recall that nilpotent orbits in  $D_N$  can be designated by  $D$ -partitions of  $2N$ , or alternatively by a Young diagram with columns of heights specified by the partition. The only requirement for these partitions is that even integers

must appear with even multiplicity, and in cases where all integers are even, there are actually two orbits linked to that partition. Another method of representing these nilpotent orbits involves using partitions of  $2N$ , where the corresponding Young diagrams have rows of length specified by those partitions (Hitchin). Among these orbits, there is a subset referred to as **special** which are those in the range of the Spaltenstein map  $d$ <sup>17</sup>. An alternative definition is that a partition  $[\mathcal{O}]$  is special if the transpose  $[\mathcal{O}]^T$  is a C-partition (Collingwood and McGovern, 1993, pp. 69-70)<sup>18</sup>. An easy way to check this is to check that there is an even number of odd parts between even parts at the beginning or ending of the partition. For example  $[3^2, 2^2, 1^2]$  is special, while  $[3, 2^2, 1]$  is not.

The Spaltenstein map,  $d$  takes Nahm partitions and maps them to Hitchin partitions and vice-versa. However, in type  $D$ , this map is not a one-to-one correspondence. When two Nahm partitions map to the same Hitchin partition, they are considered to be part of the same **special piece**. In other words, a special piece is defined as the set of nilpotent orbits that map to the same special nilpotent orbit under the action of  $d^2$ .

There is a partial-ordering on the nilpotent orbits in  $\mathfrak{g}$ , given by orbit closure. A *minimal degeneration*  $O \rightarrow O'$  is one where  $O'$  lies in the closure of  $O$  and there are no intervening *between*  $O$  and  $O'$  in the partial ordering<sup>19</sup>. These minimal degenerations correspond to *moves*, in which we rearrange the boxes in the corresponding Young tableaux. These moves were classified (a)–(g) in Kraft and Procesi (1989). Moves (b)–(f) involve moving two boxes. Moves (a) and (g) are distinguished by the fact that they involve moving only one box.

- **Type a:** For this move, the  $j^{\text{th}}$  column of the partition has odd height  $p_j = 2r+1$  and is followed by followed by the  $(j+1)^{\text{st}}$  column of height  $p_{j+1} = 2r-1$ . The

---

<sup>17</sup>Given by transposing and D-collapsing the Young tableau

<sup>18</sup>A partition of  $2N$  with the property that odd integers occur with even multiplicity

<sup>19</sup>More explicitly, there's no third orbit  $O''$  such that  $O''$  lies in the closure of  $O$  and  $O'$  lies in the closure of  $O''$ .

move consists of moving one box from the  $j^{\text{th}}$  column to the  $(j+1)^{\text{st}}$  column. After the move both columns have height  $p_j = p_{j+1} = 2r$ .

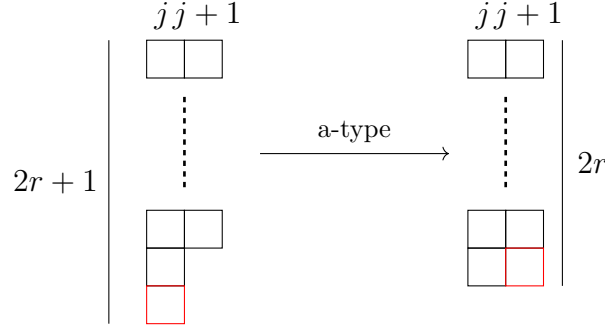


Figure 9.2: An example of the (a) move of Kraft-Procesi

- **Type g:** For this move,  $p_j = 2r + 1$ ,  $p_{j+1} = p_{j+2} = \cdots = p_{j+2m} = 2r$  and  $p_{j+2m+1} = 2r - 1$ . The move consists of moving one box from the  $j^{\text{th}}$  column to the  $(j+2m+1)^{\text{st}}$  column. After the move,  $p_j = p_{j+1} = \cdots = p_{j+2m+1} = 2r$ .

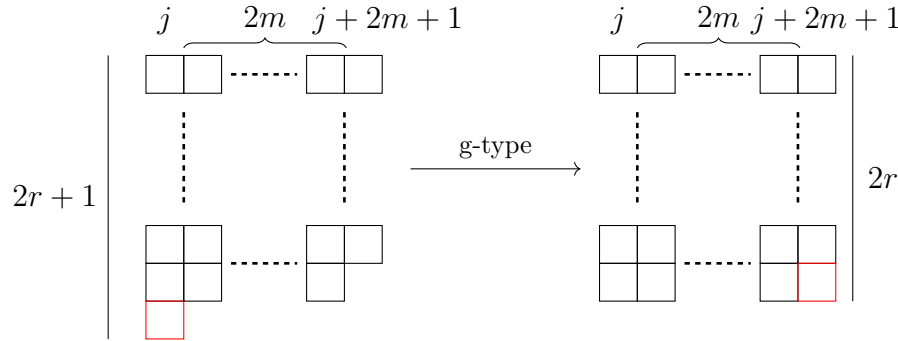


Figure 9.3: An example of the (g) move of Kraft-Procesi

These moves are what Kraft-Procesi call *small degenerations*. Two moves are considered *independent* if they can be composed. In other words, if making a move makes it impossible for us to perform a second move, those are considered to be dependent.

To understand what it means for two degenerations to be independent, consider example the  $\mathfrak{so}(24)$  partition  $\eta = [7^2, 5, 3, 1^2]$ . A priori, we can perform three

different **a**-moves, which we'll differentiate by using a subindex:

$$\eta \xrightarrow{a_1} \sigma_1 = [7, 6^2, 3, 1^2]$$

$$\eta \xrightarrow{a_2} \sigma_2 = [7^2, 4^2, 1^2]$$

$$\eta \xrightarrow{a_3} \sigma_3 = [7^2, 5, 2^2, 1]$$

Looking at  $\sigma_2$ , we see that we cannot perform another  $a_1$ -move, hence  $a_1$  and  $a_2$  are not independent. On the other hand, moves  $a_1$  and  $a_3$  can be composed in the following way (the order does not matter)

$$\eta \xrightarrow{a_1 a_3} \sigma_{13} = [7, 6^2, 2^2, 1]$$

and hence constitute independent degenerations.

Furthermore, we found it convenient to introduce the term “non-special degeneration” (Def 9.10) to refer to those small degenerations which lead to non-special orbits. It is only these small degenerations that play an important role in the structure of the special piece.

Since the terminology around various degenerations can be confusing, we recall them here in one place :

- *Minimal degeneration* (as in Kraft-Procesi Kraft and Procesi (1982)) : They are labeled from type (a) to (g).
- *Small degeneration* (as in Kraft and Procesi (1982)) : They are minimal degenerations of type (a) or (g)
- *Non-special degeneration* (implicit in Kraft and Procesi (1982), we use it explicitly) : These are small degenerations of a special Nahm orbit  $\mathcal{O}_s$  that result in a non-special orbit  $\mathcal{O}_{ns}$ . By Spaltenstein (cited in Kraft and Procesi (1982)), we know that  $\mathcal{O}_{ns}$  is in the same special piece as  $\mathcal{O}_s$ .

To see the need for the third definition above, consider the special nilpotent orbit  $\eta = [7^3, 5^2, 3^3, 1^4]$  in  $\mathfrak{so}(44)$  which has three independent small degenerations corresponding to the following operations :

$$\eta \xrightarrow{a_1} \sigma_1 = [7^2, 6^2, 5^2, 3^3, 1^4]$$

$$\eta \xrightarrow{a_2} \sigma_2 = [7^3, 5, 4^2, 3^2, 1^4]$$

$$\eta \xrightarrow{a_3} \sigma_3 = [7^3, 5^2, 3^2, 2^2, 1^2]$$

Out of this, only the third one  $\sigma_3$  is a non-special orbit and hence only  $a_3$  constitutes a non-special degeneration in the sense of Def 9.10.

With this terminology, we can state Theorem 6.2 of Kraft-Procesi Kraft and Procesi (1989) in the following way:

*Observation 4.* (Kraft-Procesi) If a special orbit  $\mathcal{O}_s$  admits  $s$  *independent* non-special degenerations, then the size of the special piece containing  $\mathcal{O}_s$  is  $2^s$ .

## Works Cited

*D-Branes*. Cambridge Monographs on Mathematical Physics. Cambridge University Press.

Lars Aalsma and Gary Shiu. Chaos and complementarity in de Sitter space. *JHEP*, 05:152, 2020. doi: 10.1007/JHEP05(2020)152.

Lars Aalsma and Watse Sybesma. The Price of Curiosity: Information Recovery in de Sitter Space. *JHEP*, 05:291, 2021. doi: 10.1007/JHEP05(2021)291.

Pramod N. Achar and Eric N. Sommers. Local systems on nilpotent orbits and weighted Dynkin diagrams. *Represent. Theory*, 6:190–201, 2002. doi: 10.1090/S1088-4165-02-00174-7.

Jeff Adams. Unpublished work.

Ian Affleck and Andreas W. W. Ludwig. Universal noninteger ‘ground state degeneracy’ in critical quantum systems. *Phys. Rev. Lett.*, 67:161–164, 1991. doi: 10.1103/PhysRevLett.67.161.

Dmitry S. Ageev. Shaping contours of entanglement islands in BCFT. 7 2021.

Prateek Agrawal, Cari Cesarotti, Andreas Karch, Rashmish K. Mishra, Lisa Randall, and Raman Sundrum. Warped Compactifications in Particle Physics, Cosmology and Quantum Gravity. In *2022 Snowmass Summer Study*, 3 2022.

Ofer Aharony, Micha Berkooz, and Nathan Seiberg. Light cone description of (2,0) superconformal theories in six-dimensions. *Adv. Theor. Math. Phys.*, 2: 119–153, 1998. doi: 10.4310/ATMP.1998.v2.n1.a5.

Ofer Aharony, Steven S. Gubser, Juan Martin Maldacena, Hirosi Ooguri, and Yaron Oz. Large N field theories, string theory and gravity. *Phys. Rept.*, 323: 183–386, 2000. doi: 10.1016/S0370-1573(99)00083-6.

Ofer Aharony, Oliver DeWolfe, Daniel Z. Freedman, and Andreas Karch. Defect conformal field theory and locally localized gravity. *JHEP*, 07:030, 2003. doi: 10.1088/1126-6708/2003/07/030.

Ofer Aharony, Adam B. Clark, and Andreas Karch. The CFT/AdS correspondence, massive gravitons and a connectivity index conjecture. *Phys. Rev. D*, 74:086006, 2006. doi: 10.1103/PhysRevD.74.086006.

Byoungjoon Ahn, Sang-Eon Bak, Hyun-Sik Jeong, Keun-Young Kim, and Ya-Wen Sun. Islands in charged linear dilaton black holes. 7 2021.

Ibrahim Akal, Yuya Kusuki, Tadashi Takayanagi, and Zixia Wei. Codimension two holography for wedges. 7 2020.

Ibrahim Akal, Yuya Kusuki, Noburo Shiba, Tadashi Takayanagi, and Zixia Wei. Entanglement Entropy in a Holographic Moving Mirror and the Page Curve. *Phys. Rev. Lett.*, 126(6):061604, 2021a. doi: 10.1103/PhysRevLett.126.061604.

Ibrahim Akal, Yuya Kusuki, Noburo Shiba, Tadashi Takayanagi, and Zixia Wei. Holographic moving mirrors. *Class. Quant. Grav.*, 38(22):224001, 2021b. doi: 10.1088/1361-6382/ac2c1b.

Chris Akers, Jason Koeller, Stefan Leichenauer, and Adam Levine. Geometric Constraints from Subregion Duality Beyond the Classical Regime. 10 2016.

Mohammad Akhond, Guillermo Arias-Tamargo, Alessandro Mininno, Hao-Yu Sun, Zhengdi Sun, Yifan Wang, and Fengjun Xu. The hitchhiker’s guide to 4d  $\mathcal{N} = 2$  superconformal field theories. *SciPost Phys. Lect. Notes*, 64:1, 2022. doi: 10.21468/SciPostPhysLectNotes.64.

Luis F. Alday, Davide Gaiotto, and Yuji Tachikawa. Liouville Correlation Functions from Four-dimensional Gauge Theories. *Lett. Math. Phys.*, 91: 167–197, 2010. doi: 10.1007/s11005-010-0369-5.

Ahmed Almheiri and Joseph Polchinski. Models of  $\text{AdS}_2$  backreaction and holography. *JHEP*, 11:014, 2015. doi: 10.1007/JHEP11(2015)014.

Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk Locality and Quantum Error Correction in  $\text{AdS}/\text{CFT}$ . *JHEP*, 04:163, 2015. doi: 10.1007/JHEP04(2015)163.

Ahmed Almheiri, Netta Engelhardt, Donald Marolf, and Henry Maxfield. The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole. *JHEP*, 12:063, 2019a. doi: 10.1007/JHEP12(2019)063.

Ahmed Almheiri, Raghu Mahajan, and Juan Maldacena. Islands outside the horizon. 10 2019b.

Ahmed Almheiri, Raghu Mahajan, and Juan Maldacena. Islands outside the horizon. 10 2019c.

Ahmed Almheiri, Raghu Mahajan, and Jorge E. Santos. Entanglement islands in higher dimensions. 11 2019d.

Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. Replica Wormholes and the Entropy of Hawking Radiation. *JHEP*, 05:013, 2020a. doi: 10.1007/JHEP05(2020)013.

Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. Replica Wormholes and the Entropy of Hawking Radiation. *JHEP*, 05:013, 2020b. doi: 10.1007/JHEP05(2020)013.

Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. 6 2020c.

Ahmed Almheiri, Raghu Mahajan, Juan Maldacena, and Ying Zhao. The Page curve of Hawking radiation from semiclassical geometry. *JHEP*, 03:149, 2020d. doi: 10.1007/JHEP03(2020)149.



Ahmed Almheiri, Raghu Mahajan, Juan Maldacena, and Ying Zhao. The Page curve of Hawking radiation from semiclassical geometry. *JHEP*, 03:149, 2020e. doi: 10.1007/JHEP03(2020)149.

Ahmed Almheiri, Raghu Mahajan, and Jorge E. Santos. Entanglement islands in higher dimensions. *SciPost Phys.*, 9(1):001, 2020f. doi: 10.21468/SciPostPhys.9.1.001.

Ahmed Almheiri, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini. The entropy of Hawking radiation. *Rev. Mod. Phys.*, 93(3):035002, 2021. doi: 10.1103/RevModPhys.93.035002.

Louise Anderson, Onkar Parrikar, and Ronak M. Soni. Islands with gravitating baths: towards  $ER = EPR$ . *JHEP*, 21:226, 2020. doi: 10.1007/JHEP10(2021)226.

Tarek Anous, Marco Meineri, Pietro Pelliconi, and Julian Sonner. Sailing past the End of the World and discovering the Island. *SciPost Phys.*, 13(3):075, 2022. doi: 10.21468/SciPostPhys.13.3.075.

Irina Aref'eva and Igor Volovich. A Note on Islands in Schwarzschild Black Holes. 10 2021.

Philip C. Argyres and Nathan Seiberg. S-duality in  $N=2$  supersymmetric gauge theories. *JHEP*, 12:088, 2007. doi: 10.1088/1126-6708/2007/12/088.

Philip C. Argyres and John R. Wittig. Infinite coupling duals of  $N=2$  gauge theories and new rank 1 superconformal field theories. *JHEP*, 01:074, 2008. doi: 10.1088/1126-6708/2008/01/074.

Philip C. Argyres, Matteo Lotito, Yongchao Lü, and Mario Martone. Expanding the landscape of  $N = 2$  rank 1 SCFTs. *JHEP*, 05:088, 2016. doi: 10.1007/JHEP05(2016)088.

Nima Arkani-Hamed, Savas Dimopoulos, G. R. Dvali, and Nemanja Kaloper. Infinitely large new dimensions. *Phys. Rev. Lett.*, 84:586–589, 2000. doi: 10.1103/PhysRevLett.84.586.

Nima Arkani-Hamed, Massimo Porrati, and Lisa Randall. Holography and phenomenology. *JHEP*, 08:017, 2001a. doi: 10.1088/1126-6708/2001/08/017.

Nima Arkani-Hamed, Massimo Porrati, and Lisa Randall. Holography and phenomenology. *JHEP*, 08:017, 2001b. doi: 10.1088/1126-6708/2001/08/017.

Nima Arkani-Hamed, Lubos Motl, Alberto Nicolis, and Cumrun Vafa. The String landscape, black holes and gravity as the weakest force. *JHEP*, 06:060, 2007. doi: 10.1088/1126-6708/2007/06/060.

Richard L. Arnowitt, Stanley Deser, and Charles W. Misner. The Dynamics of general relativity. *Gen. Rel. Grav.*, 40:1997–2027, 2008. doi: 10.1007/s10714-008-0661-1.

Tatsuo Azeyanagi, Andreas Karch, Tadashi Takayanagi, and Ethan G. Thompson. Holographic calculation of boundary entropy. *JHEP*, 03:054, 2008. doi: 10.1088/1126-6708/2008/03/054.

Constantin Bachas and Vassilis Papadopoulos. Phases of Holographic Interfaces. *JHEP*, 04:262, 2021. doi: 10.1007/JHEP04(2021)262.

Constantin Bachas, Stefano Baiguera, Shira Chapman, Giuseppe Policastro, and Tal Schwartzman. Energy Transport for Thick Holographic Branes. *Phys. Rev. Lett.*, 131(2):021601, 2023. doi: 10.1103/PhysRevLett.131.021601.

Saba Asif Baig and Sanjit Shashi. Transport across interfaces in symmetric orbifolds. 1 2023.

Dongsu Bak, Michael Gutperle, and Shinji Hirano. A Dilatonic deformation of  $\text{AdS}(5)$  and its field theory dual. *JHEP*, 05:072, 2003. doi: 10.1088/1126-6708/2003/05/072.

Dongsu Bak, Michael Gutperle, and Shinji Hirano. Three dimensional Janus and time-dependent black holes. *JHEP*, 0702:068, 2007. doi: 10.1088/1126-6708/2007/02/068.

Dongsu Bak, Chanju Kim, Sang-Heon Yi, and Junggi Yoon. Unitarity of Entanglement and Islands in Two-Sided Janus Black Holes. 6 2020.

P. Bala and R. W. Carter. Classes of unipotent elements in simple algebraic groups. i. *Mathematical Proceedings of the Cambridge Philosophical Society*, 79:401 – 425, 1976a. URL <https://api.semanticscholar.org/CorpusID:251096051>.

P. Bala and R. W. Carter. Classes of unipotent elements in simple algebraic groups. ii. *Mathematical Proceedings of the Cambridge Philosophical Society*, 80(1):1–18, 1976b. doi: 10.1017/S0305004100052610.

Aswin Balasubramanian and Jacques Distler. Masses, sheets and rigid SCFTs. *Adv. Theor. Math. Phys.*, 25(8):1953–2054, 2021a. doi: 10.4310/ATMP.2021.v25.n8.a1.

Aswin Balasubramanian and Jacques Distler. Masses, sheets and rigid SCFTs. *Adv. Theor. Math. Phys.*, 25(8):1953–2054, 2021b. doi: 10.4310/ATMP.2021.v25.n8.a1.

Aswin Balasubramanian, Jacques Distler, and Ron Donagi. Families of Hitchin systems and  $N = 2$  theories. *Adv. Theor. Math. Phys.*, 26(6):1585–1667, 2022a. doi: 10.4310/ATMP.2022.v26.n6.a2.

Aswin Balasubramanian, Jacques Distler, and Ron Donagi. Families of Hitchin systems and  $\mathcal{N} = 2$  theories. *Adv. Theor. Math. Phys.*, 26(6):1585–1667, 2022b. doi: 10.4310/ATMP.2022.v26.n6.a2.

Aswin Balasubramanian, Jacques Distler, Ron Donagi, and Carlos Perez-Pardavila. The Hitchin Image in Type-D. 10 2023.

Vijay Balasubramanian and Per Kraus. A Stress tensor for Anti-de Sitter gravity. *Commun. Math. Phys.*, 208:413–428, 1999. doi: 10.1007/s002200050764.

Vijay Balasubramanian, Arjun Kar, and Tomonori Ugajin. Entanglement between two disjoint universes. 8 2020a.

Vijay Balasubramanian, Arjun Kar, and Tomonori Ugajin. Islands in de Sitter space. 8 2020b.

Vijay Balasubramanian, Ben Craps, Mikhail Khramtsov, and Edgar Shaghoulian. Submerging islands through thermalization. *JHEP*, 10:048, 2021a. doi: 10.1007/JHEP10(2021)048.

Vijay Balasubramanian, Arjun Kar, and Tomonori Ugajin. Entanglement between two gravitating universes. 4 2021b.

Tom Banks, Michael R. Douglas, Gary T. Horowitz, and Emil J. Martinec. AdS dynamics from conformal field theory. 8 1998.

David Baraglia and Masoud Kamgarpour. On the image of the parabolic Hitchin map. *Q. J. Math.*, 69(2):681–708, 2018. ISSN 0033-5606. doi: 10.1093/qmath/hax055.

Dan Barbasch and David A Vogan Jr. Unipotent representations of complex semisimple groups. *Annals of Mathematics*, pages 41–110, 1985. doi: 10.2307/1971193. URL <https://www.jstor.org/stable/1971193>.

Florian Beck, Ron Donagi, and Katrin Wendland. Folding of Hitchin systems and crepant resolutions. *International Mathematics Research Notices*, 2022 (11):8370–8419, 2022. doi: 10.1093/imrn/rnaa375.

K. Becker, M. Becker, and J. H. Schwarz. *String theory and M-theory: A modern introduction*. Cambridge University Press, 12 2006. ISBN 978-0-511-25486-4, 978-0-521-86069-7, 978-0-511-81608-6. doi: 10.1017/CBO9780511816086.

A. Beilinson and Y. Drinfeld. Quantization of Hitchin’s Integrable System and Hecke Eigensheaves.

Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev. D*, 7:2333–2346, Apr 1973a. doi: 10.1103/PhysRevD.7.2333. URL <https://link.aps.org/doi/10.1103/PhysRevD.7.2333>.

Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev.*, D7:2333–2346, 1973b. doi: 10.1103/PhysRevD.7.2333.

Alexandre Belin, Shovon Biswas, and James Sully. The Spectrum of Boundary States in Symmetric Orbifolds. 10 2021.

Iosif Bena and Nicholas P. Warner. Black holes, black rings and their microstates. *Lect. Notes Phys.*, 755:1–92, 2008. doi: 10.1007/978-3-540-79523-0\_1.

Aranya Bhattacharya, Anindya Chanda, Sabyasachi Maulik, Christian Northe, and Shibaji Roy. Topological shadows and complexity of islands in multiboundary wormholes. 10 2020.

Aranya Bhattacharya, Arpan Bhattacharyya, Pratik Nandy, and Ayan K. Patra. Partial islands and subregion complexity in geometric secret-sharing model. 9 2021a.

Aranya Bhattacharya, Arpan Bhattacharyya, Pratik Nandy, and Ayan K. Patra. Islands and complexity of eternal black hole and radiation subsystems for a doubly holographic model. *JHEP*, 05:135, 2021b. doi: 10.1007/JHEP05(2021)135.

Aranya Bhattacharya, Arpan Bhattacharyya, Pratik Nandy, and Ayan K. Patra. Bath deformations, islands and holographic complexity. 2021c.

Mike Blake, David Tong, and David Vegh. Holographic Lattices Give the Graviton an Effective Mass. *Phys. Rev. Lett.*, 112(7):071602, 2014. doi: 10.1103/PhysRevLett.112.071602.

Francesco Bottacin. Symplectic geometry on moduli spaces of stable pairs. *Annales Scientifiques de l'École Normale Supérieure*, 28(4):391–433, 1995. doi: 10.24033/asens.1719. URL [http://www.numdam.org/item/ASENS\\_1995\\_4\\_28\\_4\\_391\\_0/](http://www.numdam.org/item/ASENS_1995_4_28_4_391_0/).

Raphael Bousso and Elizabeth Wildenhain. Gravity/ensemble duality. *Phys. Rev. D*, 102(6):066005, 2020. doi: 10.1103/PhysRevD.102.066005.

Raphael Bousso, Stefan Leichenauer, and Vladimir Rosenhaus. Light-sheets and AdS/CFT. *Phys. Rev. D*, 86:046009, 2012. doi: 10.1103/PhysRevD.86.046009.

Luca Buoninfante, Francesco Di Filippo, and Shinji Mukohyama. On the assumptions leading to the information loss paradox. 7 2021.

Elena Cáceres, Anderson Seigo Misobuchi, and Juan F. Pedraza. Constraining higher order gravities with subregion duality. *JHEP*, 11:175, 2019. doi: 10.1007/JHEP11(2019)175.

Elena Caceres, Arnab Kundu, Ayan K. Patra, and Sanjit Shashi. Warped Information and Entanglement Islands in AdS/WCFT. 12 2020.

Elena Caceres, Arnab Kundu, Ayan K. Patra, and Sanjit Shashi. Warped Information and Entanglement Islands in AdS/WCFT. *JHEP*, 07:004, 2021a. doi: 10.1007/JHEP07(2021)004.

Elena Caceres, Arnab Kundu, Ayan K. Patra, and Sanjit Shashi. Page Curves and Bath Deformations. 6 2021b.

Elena Cáceres, Rodrigo Castillo Vásquez, and Alejandro Vilar López. Entanglement entropy in cubic gravitational theories. *JHEP*, 05:186, 2021. doi: 10.1007/JHEP05(2021)186.

Pasquale Calabrese and John L. Cardy. Entanglement entropy and quantum field theory. *J. Stat. Mech.*, 0406:P06002, 2004. doi: 10.1088/1742-5468/2004/06/P06002.

John L. Cardy. Conformal Invariance and Surface Critical Behavior. *Nucl. Phys. B*, 240:514–532, 1984. doi: 10.1016/0550-3213(84)90241-4.

John L. Cardy. Operator content of two-dimensional conformally invariant theories. *Nuclear Physics B*, 270:186–204, 1986. ISSN 0550-3213. doi: [https://doi.org/10.1016/0550-3213\(86\)90552-3](https://doi.org/10.1016/0550-3213(86)90552-3). URL <https://www.sciencedirect.com/science/article/pii/0550321386905523>.

John L. Cardy. Boundary conformal field theory. 11 2004.

Horacio Casini, Marina Huerta, and Jose Alejandro Rosabal. Remarks on entanglement entropy for gauge fields. *Phys. Rev. D*, 89(8):085012, 2014. doi: 10.1103/PhysRevD.89.085012.

Oscar Chacaltana and Jacques Distler. Tinkertoys for Gaiotto Duality. *JHEP*, 11:099, 2010a. doi: 10.1007/JHEP11(2010)099.

Oscar Chacaltana and Jacques Distler. Tinkertoys for Gaiotto duality. *JHEP*, 11:099, 2010b. doi: 10.1007/JHEP11(2010)099.

Oscar Chacaltana and Jacques Distler. Tinkertoys for the  $D_N$  series. *JHEP*, 02:110, 2013a. doi: 10.1007/JHEP02(2013)110.

Oscar Chacaltana and Jacques Distler. Tinkertoys for the  $D_N$  series. *JHEP*, 02:110, 2013b. doi: 10.1007/JHEP02(2013)110.

Oscar Chacaltana, Jacques Distler, and Yuji Tachikawa. Nilpotent orbits and codimension-two defects of 6d  $\mathcal{N}=(2,0)$  theories. *Int. J. Mod. Phys. A*, 28:1340006, 2013a. doi: 10.1142/S0217751X1340006X.

Oscar Chacaltana, Jacques Distler, and Yuji Tachikawa. Nilpotent orbits and codimension-two defects of 6d  $\mathcal{N}=(2,0)$  theories. *Int. J. Mod. Phys. A*, 28:1340006, 2013b. doi: 10.1142/S0217751X1340006X.

Oscar Chacaltana, Jacques Distler, and Yuji Tachikawa. Gaiotto duality for the twisted  $A_{2N-1}$  series. *JHEP*, 05:075, 2015a. doi: 10.1007/JHEP05(2015)075.

Oscar Chacaltana, Jacques Distler, and Anderson Trimm. Tinkertoys for the twisted D-series. *JHEP*, 04:173, 2015b. doi: 10.1007/JHEP04(2015)173.

Oscar Chacaltana, Jacques Distler, and Anderson Trimm. Tinkertoys for the  $E_6$  theory. *JHEP*, 09:007, 2015c. doi: 10.1007/JHEP09(2015)007.

Oscar Chacaltana, Jacques Distler, and Anderson Trimm. Tinkertoys for the twisted  $E_6$  theory. 1 2015d.

Oscar Chacaltana, Jacques Distler, and Anderson Trimm. Tinkertoys for the  $\mathbb{Z}_3$ -twisted  $D_4$  theory. 1 2016.

Oscar Chacaltana, Jacques Distler, Anderson Trimm, and Yinan Zhu. Tinkertoys for the  $E_7$  theory. *JHEP*, 05:031, 2018a. doi: 10.1007/JHEP05(2018)031.

Oscar Chacaltana, Jacques Distler, Anderson Trimm, and Yinan Zhu. Tinkertoys for the  $E_8$  theory. 2 2018b.



A. Chamblin and A. Karch. Hawking and Page on the brane. *Phys. Rev. D*, 72:066011, 2005. doi: 10.1103/PhysRevD.72.066011.

Venkatesa Chandrasekaran, Masamichi Miyaji, and Pratik Rath. Including contributions from entanglement islands to the reflected entropy. *Phys. Rev. D*, 102(8):086009, 2020. doi: 10.1103/PhysRevD.102.086009.

Hong Zhe Chen, Robert C. Myers, Dominik Neuenfeld, Ignacio A. Reyes, and Joshua Sandor. Quantum Extremal Islands Made Easy, Part II: Black Holes on the Brane. 9 2020a.

Hong Zhe Chen, Robert C. Myers, Dominik Neuenfeld, Ignacio A. Reyes, and Joshua Sandor. Quantum Extremal Islands Made Easy, Part I: Entanglement on the Brane. 6 2020b.

Chia-Jui Chou, Bo-Han Lin, Bin Wang, and Yi Yang. Entanglement Entropy Inequalities in BCFT by Holography. 11 2020.

Chandramouli Chowdhury, Olga Papadoulaki, and Suvrat Raju. A physical protocol for observers near the boundary to obtain bulk information in quantum gravity. 8 2020.

Chandramouli Chowdhury, Victor Godet, Olga Papadoulaki, and Suvrat Raju. Holography from the Wheeler-DeWitt equation. 7 2021.

Chong-Sun Chu and Rong-Xin Miao. Conformal Boundary Condition and Perturbation Spectrum in AdS/BCFT. 10 2021.

Jinwei Chu, Feiyu Deng, and Yang Zhou. Page curve from defect extremal surface and island in higher dimensions. *JHEP*, 10:149, 2021. doi: 10.1007/JHEP10(2021)149.

A. B. Clark, D. Z. Freedman, A. Karch, and M. Schnabl. Dual of the Janus solution: An interface conformal field theory. *Phys. Rev. D*, 71:066003, 2005. doi: 10.1103/PhysRevD.71.066003.

Timothy Clifton, Pedro G. Ferreira, Antonio Padilla, and Constantinos Skordis. Modified Gravity and Cosmology. *Phys. Rept.*, 513:1–189, 2012. doi: 10.1016/j.physrep.2012.01.001.

Scott Collier, Dalimil Mazac, and Yifan Wang. Bootstrapping Boundaries and Branes. 12 2021.

David H. Collingwood and William M. McGovern. *Nilpotent Orbits in Semisimple Lie Algebras*. Van Nostrand, 1993.

Sean Cooper, Dominik Neuenfeld, Moshe Rozali, and David Wakeham. Brane dynamics from the first law of entanglement. *JHEP*, 03:023, 2020. doi: 10.1007/JHEP03(2020)023.

Justin Corvino and Richard M. Schoen. On the asymptotics for the vacuum Einstein constraint equations. *J. Diff. Geom.*, 73(2):185–217, 2006.

Csaba Csaki, Michael Graesser, Lisa Randall, and John Terning. Cosmology of brane models with radion stabilization. *Phys. Rev. D*, 62:045015, 2000. doi: 10.1103/PhysRevD.62.045015.

Bartłomiej Czech, Joanna L. Karczmarek, Fernando Nogueira, and Mark Van Raamsdonk. The Gravity Dual of a Density Matrix. *Class. Quant. Grav.*, 29:155009, 2012. doi: 10.1088/0264-9381/29/15/155009.

Jin Dai, R. G. Leigh, and Joseph Polchinski. New Connections Between String Theories. *Mod. Phys. Lett. A*, 4:2073–2083, 1989. doi: 10.1142/S0217732389002331.

Peter Dalakov. Meromorphic Higgs bundles and related geometries. *J. Geom. Phys.*, 109:44–67, 2016. ISSN 0393-0440,1879-1662. doi: 10.1016/j.geomphys.2016.01.005.

Sumit R. Das, Anurag Kaushal, Sinong Liu, Gautam Mandal, and Sandip P. Trivedi. Gauge Invariant Target Space Entanglement in D-Brane Holography. 11 2020.

Sebastian de Haro, Sergey N. Solodukhin, and Kostas Skenderis. Holographic reconstruction of space-time and renormalization in the AdS / CFT correspondence. *Commun. Math. Phys.*, 217:595–622, 2001. doi: 10.1007/s002200100381.

Feiyu Deng, Jinwei Chu, and Yang Zhou. Defect extremal surface as the holographic counterpart of Island formula. *JHEP*, 03:008, 2021. doi: 10.1007/JHEP03(2021)008.

Bryce S DeWitt. The quantization of geometry. In Louis Witten, editor, *Gravitation: an introduction to current research*. John Wiley & Sons, 1963.

Bryce S. DeWitt. Quantum Theory of Gravity. 1. The Canonical Theory. *Phys. Rev.*, 160:1113–1148, 1967. doi: 10.1103/PhysRev.160.1113.

O. DeWolfe, D. Z. Freedman, S. S. Gubser, and A. Karch. Modeling the fifth-dimension with scalars and gravity. *Phys. Rev. D*, 62:046008, 2000. doi: 10.1103/PhysRevD.62.046008.

Jacques Distler and Amihay Hanany. (0,2) Noncritical strings in six-dimensions. *Nucl. Phys. B*, 490:75–90, 1997. doi: 10.1016/S0550-3213(97)00046-1.

Jacques Distler, Nathan Donagi, and Ron Donagi. On generalized Pfaffians. 9 2024.

Ron Donagi and Edward Witten. Supersymmetric Yang-Mills theory and integrable systems. *Nucl. Phys. B*, 460:299–334, 1996a. doi: 10.1016/0550-3213(95)00609-5.

Ron Donagi and Edward Witten. Supersymmetric Yang-Mills theory and integrable systems. *Nucl. Phys.*, B460:299–334, 1996b. doi: 10.1016/0550-3213(95)00609-5.

Xi Dong. The Gravity Dual of Renyi Entropy. *Nature Commun.*, 7:12472, 2016. doi: 10.1038/ncomms12472.

Xi Dong, Daniel Harlow, and Aron C. Wall. Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality. *Phys. Rev. Lett.*, 117(2):021601, 2016a. doi: 10.1103/PhysRevLett.117.021601.

Xi Dong, Aitor Lewkowycz, and Mukund Rangamani. Deriving covariant holographic entanglement. *JHEP*, 11:028, 2016b. doi: 10.1007/JHEP11(2016)028.

Xi Dong, Xiao-Liang Qi, Zhou Shangnan, and Zhenbin Yang. Effective entropy of quantum fields coupled with gravity. *JHEP*, 10:052, 2020. doi: 10.1007/JHEP10(2020)052.

William Donnelly and Steven B. Giddings. Diffeomorphism-invariant observables and their nonlocal algebra. *Phys. Rev. D*, 93(2):024030, 2016a. doi: 10.1103/PhysRevD.93.024030. [Erratum: Phys.Rev.D 94, 029903 (2016)].

William Donnelly and Steven B. Giddings. Observables, gravitational dressing, and obstructions to locality and subsystems. *Phys. Rev. D*, 94(10):104038, 2016b. doi: 10.1103/PhysRevD.94.104038.

G. R. Dvali, G. Gabadadze, and M. Porrati. Metastable gravitons and infinite volume extra dimensions. *Phys. Lett. B*, 484:112–118, 2000a. doi: 10.1016/S0370-2693(00)00631-6.

G. R. Dvali, Gregory Gabadadze, and Massimo Porrati. 4-D gravity on a brane in 5-D Minkowski space. *Phys. Lett. B*, 485:208–214, 2000b. doi: 10.1016/S0370-2693(00)00669-9.

G.R. Dvali, Gregory Gabadadze, and Massimo Porrati. 4-D gravity on a brane in 5-D Minkowski space. *Phys. Lett. B*, 485:208–214, 2000c. doi: 10.1016/S0370-2693(00)00669-9.

Sheer El-Showk and Kyriakos Papadodimas. Emergent Spacetime and Holographic CFTs. *JHEP*, 10:106, 2012. doi: 10.1007/JHEP10(2012)106.

Roberto Emparan. Black hole entropy as entanglement entropy: A Holographic derivation. *JHEP*, 06:012, 2006. doi: 10.1088/1126-6708/2006/06/012.

Roberto Emparan, Antonia Micol Frassino, and Benson Way. Quantum BTZ black hole. *JHEP*, 11:137, 2020. doi: 10.1007/JHEP11(2020)137.

Roberto Emparan, Antonia Micol Frassino, Martin Sasieta, and Marija Tomašević. Holographic complexity of quantum black holes. 2021.

Roberto Emparan, David Licht, Ryotaku Suzuki, Marija Tomašević, and Benson Way. Black tsunamis and naked singularities in AdS. *JHEP*, 02:090, 2022. doi: 10.1007/JHEP02(2022)090.

Netta Engelhardt and Aron C. Wall. Extremal Surface Barriers. *JHEP*, 03:068, 2014. doi: 10.1007/JHEP03(2014)068.

Netta Engelhardt and Aron C. Wall. Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime. *JHEP*, 01:073, 2015a. doi: 10.1007/JHEP01(2015)073.

Netta Engelhardt and Aron C. Wall. Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime. *JHEP*, 01:073, 2015b. doi: 10.1007/JHEP01(2015)073.

Julius Engelsöy, Thomas G. Mertens, and Herman Verlinde. An investigation of  $\text{AdS}_2$  backreaction and holography. *JHEP*, 07:139, 2016. doi: 10.1007/JHEP07(2016)139.

Behzat Ergun, Qianyu Hao, Andrew Neitzke, and Fei Yan. Factorized class S theories and surface defects. *JHEP*, 12:041, 2021. doi: 10.1007/JHEP12(2021)041.

Seamus Fallows and Simon F. Ross. Islands and mixed states in closed universes. *JHEP*, 07:022, 2021. doi: 10.1007/JHEP07(2021)022.

Thomas Faulkner and Aitor Lewkowycz. Bulk locality from modular flow. *JHEP*, 07:151, 2017. doi: 10.1007/JHEP07(2017)151.

Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *JHEP*, 11:074, 2013a. doi: 10.1007/JHEP11(2013)074.

Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *JHEP*, 11:074, 2013b. doi: 10.1007/JHEP11(2013)074.

Asmund Folkestad. The Penrose Inequality as a Constraint on the Low Energy Limit of Quantum Gravity. 8 2022.

Daniel S. Freed. Special Kahler manifolds. *Commun. Math. Phys.*, 203:31–52, 1999a. doi: 10.1007/s002200050604.

Daniel S. Freed. Special Kähler manifolds. *Commun. Math. Phys.*, 203:31–52, 1999b. doi: 10.1007/s002200050604.

Mitsutoshi Fujita, Tadashi Takayanagi, and Erik Tonni. Aspects of  $\text{AdS}/\text{BCFT}$ . *JHEP*, 11:043, 2011. doi: 10.1007/JHEP11(2011)043.

Takeshi Fukuyama and Kiyoshi Kamimura. Gauge Theory of Two-dimensional Gravity. *Phys. Lett. B*, 160:259–262, 1985. doi: 10.1016/0370-2693(85)91322-X.

Matthias R. Gaberdiel, Bob Knighton, and Jakub Vošmera. D-branes in  $\text{AdS}_3 \times \text{S}^3 \times \mathbb{T}^4$  at  $k = 1$  and their holographic duals. 10 2021.

Davide Gaiotto. N=2 dualities. *JHEP*, 08:034, 2012a. doi: 10.1007/JHEP08(2012)034.

Davide Gaiotto.  $\mathcal{N} = 2$  dualities. *JHEP*, 08:034, 2012b. doi: 10.1007/JHEP08(2012)034.

Davide Gaiotto and Edward Witten. Supersymmetric Boundary Conditions in N=4 Super Yang-Mills Theory. *J. Statist. Phys.*, 135:789–855, 2009. doi: 10.1007/s10955-009-9687-3.

Davide Gaiotto and Edward Witten. Gauge Theory and the Analytic Form of the Geometric Langlands Program. *Annales Henri Poincare*, 25(1):557–671, 2024. doi: 10.1007/s00023-022-01225-6.

Davide Gaiotto, Gregory W. Moore, and Andrew Neitzke. Wall-crossing, Hitchin systems, and the WKB approximation. *Adv. Math.*, 234:239–403, 2013a. doi: 10.1016/j.aim.2012.09.027.

Davide Gaiotto, Gregory W. Moore, and Andrew Neitzke. Wall-crossing, Hitchin systems, and the WKB approximation. *Adv. Math.*, 234:239–403, 2013b. doi: 10.1016/j.aim.2012.09.027.

Ping Gao, Daniel L. Jafferis, and David K. Kolchmeyer. An effective matrix model for dynamical end of the world branes in Jackiw-Teitelboim gravity. 4 2021.

Sijie Gao and Robert M. Wald. Theorems on gravitational time delay and related issues. *Class. Quant. Grav.*, 17:4999–5008, 2000. doi: 10.1088/0264-9381/17/24/305.

Antonio M. García-García and Victor Godet. Half-wormholes in nearly  $\text{AdS}_2$  holography. 7 2021.

Friðrik Freyr Gautason, Lukas Schneiderbauer, Watse Sybesma, and Lárus Thorlacius. Page Curve for an Evaporating Black Hole. *JHEP*, 05:091, 2020. doi: 10.1007/JHEP05(2020)091.

Hao Geng. Non-local Entanglement and Fast Scrambling in De-Sitter Holography. 5 2020.

Hao Geng. Aspects of  $\text{AdS}_2$  Quantum Gravity and the Karch-Randall Braneworld. 6 2022.

Hao Geng. Revisiting Recent Progress in the Karch-Randall Braneworld. 6 2023.

Hao Geng and Andreas Karch. Massive islands. *JHEP*, 09:121, 2020a. doi: 10.1007/JHEP09(2020)121.

Hao Geng and Andreas Karch. Massive Islands. 6 2020b.

Hao Geng, Sebastian Grieninger, and Andreas Karch. Entropy, Entanglement and Swampland Bounds in  $\text{DS/dS}$ . *JHEP*, 06:105, 2019. doi: 10.1007/JHEP06(2019)105.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Information Transfer with a Gravitating Bath. *SciPost Phys.*, 10(5):103, 2021a. doi: 10.21468/SciPostPhys.10.5.103.



Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Information Transfer with a Gravitating Bath. *SciPost Phys.*, 10(5):103, 2021b. doi: 10.21468/SciPostPhys.10.5.103.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Inconsistency of Islands in Theories with Long-Range Gravity. 7 2021c.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Entanglement Phase Structure of a Holographic BCFT in a Black Hole Background. 12 2021d.

Hao Geng, Severin Lüster, Rashmish K. Mishra, and David Wakeham. Holographic BCFTs and Communicating Black Holes. *jhep*, 08:003, 2021e. doi: 10.1007/JHEP08(2021)003.

Hao Geng, Yasunori Nomura, and Hao-Yu Sun. Information paradox and its resolution in de Sitter holography. *Phys. Rev. D*, 103(12):126004, 2021f. doi: 10.1103/PhysRevD.103.126004.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Inconsistency of islands in theories with long-range gravity. *JHEP*, 01:182, 2022a. doi: 10.1007/JHEP01(2022)182.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Entanglement phase structure of a holographic BCFT in a black hole background. *JHEP*, 05:153, 2022b. doi: 10.1007/JHEP05(2022)153.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Jackiw-Teitelboim Gravity from the Karch-Randall Braneworld. *Phys. Rev. Lett.*, 129(23):231601, 2022c. doi: 10.1103/PhysRevLett.129.231601.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Entanglement phase structure of a holographic BCFT in a black hole background. *JHEP*, 05:153, 2022d. doi: 10.1007/JHEP05(2022)153.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, and Sanjit Shashi. Jackiw-teitelboim gravity from the karch-randall braneworld. *Physical Review Letters*, 129(23), nov 2022e. doi: 10.1103/physrevlett.129.231601. URL <https://doi.org/10.1103%2Fphysrevlett.129.231601>.

Hao Geng, Lisa Randall, and Erik Swanson. BCFT in a Black Hole Background: An Analytical Holographic Model. 9 2022f.

Hao Geng, Lisa Randall, and Erik Swanson. BCFT in a black hole background: an analytical holographic model. *Journal of High Energy Physics*, 2022(12), dec 2022g. doi: 10.1007/jhep12(2022)056. URL <https://doi.org/10.1007%2Fjhep12%282022%29056>.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Suvrat Raju, Lisa Randall, Marcos Riojas, Sanjit Shashi, and Merna Youssef. , *in preparation*, 2023a.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Lisa Randall, Marcos Riojas, Sanjit Shashi, and Merna Youssef. Constraining braneworlds with entanglement entropy. *SciPost Phys.*, 15(5):199, 2023b. doi: 10.21468/SciPostPhys.15.5.199.

Hao Geng, Andreas Karch, Carlos Perez-Pardavila, Lisa Randall, Marcos Riojas, Sanjit Shashi, and Merna Youssef. Constraining braneworlds with entanglement entropy. 6 2023c.

Howard Georgi, Aaron K. Grant, and Girma Hailu. Brane couplings from bulk loops. *Phys. Lett. B*, 506:207–214, 2001. doi: 10.1016/S0370-2693(01)00408-7.

Kausik Ghosh and Chethan Krishnan. Dirichlet baths and the not-so-fine-grained Page curve. *JHEP*, 08:119, 2021. doi: 10.1007/JHEP08(2021)119.

Sudip Ghosh and Suvrat Raju. Quantum information measures for restricted sets of observables. *Phys. Rev. D*, 98(4):046005, 2018. doi: 10.1103/PhysRevD.98.046005.

Sudip Ghosh, Ronak M Soni, and Sandip P. Trivedi. On The Entanglement Entropy For Gauge Theories. *JHEP*, 09:069, 2015. doi: 10.1007/JHEP09(2015)069.

Steven B. Giddings. Gravitational dressing, soft charges, and perturbative gravitational splitting. *Phys. Rev. D*, 100(12):126001, 2019. doi: 10.1103/PhysRevD.100.126001.

Steven B. Giddings, Donald Marolf, and James B. Hartle. Observables in effective gravity. *Phys. Rev. D*, 74:064018, 2006. doi: 10.1103/PhysRevD.74.064018.

Walter D. Goldberger and Mark B. Wise. Modulus stabilization with bulk fields. *Phys. Rev. Lett.*, 83:4922–4925, 1999. doi: 10.1103/PhysRevLett.83.4922.

Alfred Scharff Goldhaber and Michael Martin Nieto. Photon and Graviton Mass Limits. *Rev. Mod. Phys.*, 82:939–979, 2010. doi: 10.1103/RevModPhys.82.939.

A. Gorsky, I. Krichever, A. Marshakov, A. Mironov, and A. Morozov. Integrability and Seiberg-Witten exact solution. *Phys. Lett. B*, 355:466–474, 1995. doi: 10.1016/0370-2693(95)00723-X.

Kaberi Goswami, K. Narayan, and Hitesh K. Saini. Cosmologies, singularities and quantum extremal surfaces. 11 2021.

C. Robin Graham and Edward Witten. Conformal anomaly of submanifold observables in AdS / CFT correspondence. *Nucl. Phys. B*, 546:52–64, 1999. doi: 10.1016/S0550-3213(99)00055-3.

Michael B. Green, J. H. Schwarz, and Edward Witten. *Superstring Theory. Vol. 1: Introduction*. Cambridge Monographs on Mathematical Physics. 7 1988a. ISBN 978-0-521-35752-4.

Michael B. Green, J. H. Schwarz, and Edward Witten. *Superstring Theory. Vol. 2: Loop amplitudes, anomalies and phenomenology*. 7 1988b. ISBN 978-0-521-35753-1.

R. Gregory and R. Laflamme. Black strings and p-branes are unstable. *Phys. Rev. Lett.*, 70:2837–2840, 1993. doi: 10.1103/PhysRevLett.70.2837.

Ruth Gregory, Simon F. Ross, and Robin Zegers. Classical and quantum gravity of brane black holes. *JHEP*, 09:029, 2008. doi: 10.1088/1126-6708/2008/09/029.

Daniel Grumiller and Robert McNees. Thermodynamics of black holes in two (and higher) dimensions. *JHEP*, 04:074, 2007. doi: 10.1088/1126-6708/2007/04/074.

Daniel Grumiller, Robert McNees, Jakob Salzer, Carlos Valcárcel, and Dmitri Vassilevich. Menagerie of AdS<sub>2</sub> boundary conditions. *JHEP*, 10:203, 2017. doi: 10.1007/JHEP10(2017)203.

S. S. Gubser, Igor R. Klebanov, and Alexander M. Polyakov. Gauge theory correlators from noncritical string theory. *Phys. Lett. B*, 428:105–114, 1998. doi: 10.1016/S0370-2693(98)00377-3.

Steven S. Gubser. AdS / CFT and gravity. *Phys. Rev. D*, 63:084017, 2001. doi: 10.1103/PhysRevD.63.084017.

Steven S. Gubser and Indrajit Mitra. The Evolution of unstable black holes in anti-de Sitter space. *JHEP*, 08:018, 2001. doi: 10.1088/1126-6708/2001/08/018.

Monica Guica and Daniel L. Jafferis. On the construction of charged operators inside an eternal black hole. *SciPost Phys.*, 3(2):016, 2017. doi: 10.21468/SciPostPhys.3.2.016.

Sergei Gukov and Edward Witten. Gauge Theory, Ramification, And The Geometric Langlands Program. 12 2006a.

Sergei Gukov and Edward Witten. Gauge theory, ramification, and the geometric Langlands program. 12 2006b.

Sergei Gukov and Edward Witten. Rigid Surface Operators. *Adv. Theor. Math. Phys.*, 14(1):87–178, 2010. doi: 10.4310/ATMP.2010.v14.n1.a3.

R. Haag. *Local quantum physics: Fields, particles, algebras*. Springer, 1992.

Alex Hamilton, Daniel N. Kabat, Gilad Lifschytz, and David A. Lowe. Local bulk operators in AdS/CFT: A Boundary view of horizons and locality. *Phys. Rev. D*, 73:086003, 2006. doi: 10.1103/PhysRevD.73.086003.

S. Shajidul Haque and Bret Underwood. The Squeezed OTOC and Cosmology. 10 2020.

Daniel Harlow. Wormholes, Emergent Gauge Fields, and the Weak Gravity Conjecture. *JHEP*, 01:122, 2016. doi: 10.1007/JHEP01(2016)122.

Daniel Harlow and Daniel Jafferis. The Factorization Problem in Jackiw-Teitelboim Gravity. *JHEP*, 02:177, 2020. doi: 10.1007/JHEP02(2020)177.

Daniel Harlow and Edgar Shaghoulian. Global symmetry, Euclidean gravity, and the black hole information problem. *JHEP*, 04:175, 2021. doi: 10.1007/JHEP04(2021)175.

Thomas Hartman and Juan Maldacena. Time Evolution of Entanglement Entropy from Black Hole Interiors. *JHEP*, 05:014, 2013. doi: 10.1007/JHEP05(2013)014.

Thomas Hartman, Yikun Jiang, and Edgar Shaghoulian. Islands in cosmology. *JHEP*, 11:111, 2020a. doi: 10.1007/JHEP11(2020)111.

Thomas Hartman, Edgar Shaghoulian, and Andrew Strominger. Islands in Asymptotically Flat 2D Gravity. *JHEP*, 07:022, 2020b. doi: 10.1007/JHEP07(2020)022.

S. W. Hawking. Particle creation by black holes. *Commun. Math. Phys.*, 43(3):199–220, Aug 1975. ISSN 1432-0916. doi: 10.1007/BF02345020. URL <https://doi.org/10.1007/BF02345020>.

S. W. Hawking. Black holes and thermodynamics. *Phys. Rev. D*, 13:191–197, Jan 1976a. doi: 10.1103/PhysRevD.13.191. URL <https://link.aps.org/doi/10.1103/PhysRevD.13.191>.

S. W. Hawking. Black Holes and Thermodynamics. *Phys. Rev.*, D13:191–197, 1976b. doi: 10.1103/PhysRevD.13.191.

S. W. Hawking and Gary T. Horowitz. The Gravitational Hamiltonian, action, entropy and surface terms. *Class. Quant. Grav.*, 13:1487–1498, 1996. doi: 10.1088/0264-9381/13/6/017.

S. W. Hawking and Don N. Page. Thermodynamics of Black Holes in anti-De Sitter Space. *Commun. Math. Phys.*, 87:577, 1983. doi: 10.1007/BF01208266.

Song He, Yuan Sun, Long Zhao, and Yu-Xuan Zhang. The universality of islands outside the horizon. 10 2021.

Matthew Headrick. General properties of holographic entanglement entropy. *JHEP*, 03:085, 2014. doi: 10.1007/JHEP03(2014)085.

Matthew Headrick, Veronika E. Hubeny, Albion Lawrence, and Mukund Rangamani. Causality & holographic entanglement entropy. *JHEP*, 12:162, 2014. doi: 10.1007/JHEP12(2014)162.

Juan Hernandez, Robert C. Myers, and Shan-Ming Ruan. Quantum Extremal Islands Made Easy, PartIII: Complexity on the Brane. 10 2020.

Kurt Hinterbichler. Theoretical Aspects of Massive Gravity. *Rev. Mod. Phys.*, 84:671–710, 2012. doi: 10.1103/RevModPhys.84.671.

Takayuki Hirayama and Gungwon Kang. Stable black strings in anti-de Sitter space. *Phys. Rev. D*, 64:064010, 2001. doi: 10.1103/PhysRevD.64.064010.

Nigel J. Hitchin. The Selfduality equations on a Riemann surface. *Proc. Lond. Math. Soc.*, 55:59–131, 1987. doi: 10.1112/plms/s3-55.1.59.

Timothy J. Hollowood, S. Prem Kumar, Andrea Legramandi, and Neil Talwar. Grey-body Factors, Irreversibility and Multiple Island Saddles. 11 2021a.

Timothy J. Hollowood, S. Prem Kumar, Andrea Legramandi, and Neil Talwar. Islands in the stream of Hawking radiation. *JHEP*, 11:067, 2021b. doi: 10.1007/JHEP11(2021)067.

Timothy J. Hollowood, S. Prem Kumar, Andrea Legramandi, and Neil Talwar. Ephemeral Islands, Plunging Quantum Extremal Surfaces and BCFT channels. 9 2021c.

Christoph Holzhey, Finn Larsen, and Frank Wilczek. Geometric and renormalized entropy in conformal field theory. *Nucl. Phys. B*, 424:443–467, 1994. doi: 10.1016/0550-3213(94)90402-2.

G't Hooft. A planar diagram theory for strong interactions. *Nuclear Physics B*, 72(3):461–473, 1974. ISSN 0550-3213. doi: [https://doi.org/10.1016/0550-3213\(74\)90154-0](https://doi.org/10.1016/0550-3213(74)90154-0). URL <https://www.sciencedirect.com/science/article/pii/0550321374901540>.

Gary T. Horowitz and Andrew Strominger. Counting states of near-extremal black holes. *Physical Review Letters*, 77(12):2368–2371, September 1996. ISSN 1079-7114. doi: 10.1103/physrevlett.77.2368. URL <http://dx.doi.org/10.1103/PhysRevLett.77.2368>.

Veronika E. Hubeny. The AdS/CFT Correspondence. *Class. Quant. Grav.*, 32(12):124010, 2015. doi: 10.1088/0264-9381/32/12/124010.

Veronika E. Hubeny and Mukund Rangamani. Causal Holographic Information. *JHEP*, 06:114, 2012. doi: 10.1007/JHEP06(2012)114.

Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A Covariant holographic entanglement entropy proposal. *JHEP*, 07:062, 2007a. doi: 10.1088/1126-6708/2007/07/062.

Veronika E. Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A Covariant holographic entanglement entropy proposal. *JHEP*, 07:062, 2007b. doi: 10.1088/1126-6708/2007/07/062.

Veronika E. Hubeny, Mukund Rangamani, and Erik Tonni. Global properties of causal wedges in asymptotically AdS spacetimes. *JHEP*, 10:059, 2013. doi: 10.1007/JHEP10(2013)059.

Norihiro Iizuka, Akihiro Miyata, and Tomonori Ugajin. A comment on a fine-grained description of evaporating black holes with baby universes. 11 2021.



Kenneth A. Intriligator. 'Integrating in' and exact superpotentials in 4-d. *Phys. Lett. B*, 336:409–414, 1994. doi: 10.1016/0370-2693(94)90552-5.

Kenneth A. Intriligator and N. Seiberg. Duality, monopoles, dyons, confinement and oblique confinement in supersymmetric  $SO(N(c))$  gauge theories. *Nucl. Phys. B*, 444:125–160, 1995. doi: 10.1016/0550-3213(95)00159-P.

W. Israel. Singular hypersurfaces and thin shells in general relativity. *Nuovo Cim. B*, 44S10:1, 1966. doi: 10.1007/BF02710419. [Erratum: *Nuovo Cim.B* 48, 463 (1967)].

W. Israel. Thermo field dynamics of black holes. *Phys. Lett. A*, 57:107–110, 1976. doi: 10.1016/0375-9601(76)90178-X.

Keisuke Izumi. Causal Structures in Gauss-Bonnet gravity. *Phys. Rev. D*, 90(4):044037, 2014. doi: 10.1103/PhysRevD.90.044037.

R. Jackiw. Lower Dimensional Gravity. *Nucl. Phys. B*, 252:343–356, 1985. doi: 10.1016/0550-3213(85)90448-1.

Ted Jacobson. Boundary unitarity and the black hole information paradox. *Int. J. Mod. Phys. D*, 22:1342002, 2013. doi: 10.1142/S0218271813420029.

Ted Jacobson and Phuc Nguyen. Diffeomorphism invariance and the black hole information paradox. *Phys. Rev. D*, 100(4):046002, 2019. doi: 10.1103/PhysRevD.100.046002.

Daniel L. Jafferis, Aitor Lewkowycz, Juan Maldacena, and S. Josephine Suh. Relative entropy equals bulk relative entropy. *JHEP*, 06:004, 2016. doi: 10.1007/JHEP06(2016)004.

Kristan Jensen. Chaos in  $AdS_2$  Holography. *Phys. Rev. Lett.*, 117(11):111601, 2016. doi: 10.1103/PhysRevLett.117.111601.

Clifford V. Johnson. *D-branes*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2005. ISBN 9780511057694, 9780521030052, 9780521809122, 9780511606540. doi: 10.1017/CBO9780511606540. URL <http://books.cambridge.org/0521809126.htm>.

Joshua Kames-King, Evita Verheijden, and Erik Verlinde. No Page Curves for the de Sitter Horizon. 8 2021.

Andreas Karch and Lisa Randall. Locally localized gravity. *JHEP*, 05:008, 2001a. doi: 10.1088/1126-6708/2001/05/008.

Andreas Karch and Lisa Randall. Locally localized gravity. *JHEP*, 05:008, 2001b. doi: 10.1088/1126-6708/2001/05/008.

Andreas Karch and Lisa Randall. Open and closed string interpretation of SUSY CFT's on branes with boundaries. *JHEP*, 06:063, 2001c. doi: 10.1088/1126-6708/2001/06/063.

Andreas Karch and Lisa Randall. Geometries with mismatched branes. 6 2020.

Andreas Karch, Carlos Perez-Pardavila, Marcos Riojas, and Merna Youssef. Subregion entropy for the doubly-holographic global black string. *JHEP*, 05:195, 2023a. doi: 10.1007/JHEP05(2023)195.

Andreas Karch, Carlos Perez-Pardavila, Marcos Riojas, and Merna Youssef. Subregion Entropy for the Doubly Holographic Global Black String. 3 2023b.

Andreas Karch, Carlos Perez-Pardavila, Marcos Riojas, and Merna Youssef. Subregion Entropy for the Doubly Holographic Global Black String. 3 2023c.

Anna Karlsson. Concerns about the replica wormhole derivation of the island conjecture. 1 2021.

S. Katz, P. Mayr, and C. Vafa. Mirror symmetry and exact solution of 4-D N=2 gauge theories: 1. *Adv. Theor. Math. Phys.*, 1:53–114, 1998. doi: 10.4310/ATMP.1997.v1.n1.a2.

Sheldon H. Katz, Albrecht Klemm, and Cumrun Vafa. Geometric engineering of quantum field theories. *Nucl. Phys. B*, 497:173–195, 1997. doi: 10.1016/S0550-3213(97)00282-4.

Kohki Kawabata, Tatsuma Nishioka, Yoshitaka Okuyama, and Kento Watanabe. Probing Hawking radiation through capacity of entanglement. *JHEP*, 05:062, 2021a. doi: 10.1007/JHEP05(2021)062.

Kohki Kawabata, Tatsuma Nishioka, Yoshitaka Okuyama, and Kento Watanabe. Replica wormholes and capacity of entanglement. *JHEP*, 10:227, 2021b. doi: 10.1007/JHEP10(2021)227.

David Kazhdan and George Lusztig. Fixed point varieties on affine flag manifolds. *Israel Journal of Mathematics*, 62:129–168, 1988. doi: 10.1007/BF02787119.

Gisela Kempken. Induced conjugacy classes in classical Lie algebras. *Abh. Math. Sem. Univ. Hamburg*, 53:53–83, 1983. ISSN 0025-5858,1865-8784. doi: 10.1007/BF02941310.

Tanay Kibe, Prabha Mandayam, and Ayan Mukhopadhyay. Holographic space-time, black holes and quantum error correcting codes: A review. 10 2021.

Hang Bae Kim and Hyung Do Kim. Inflation and gauge hierarchy in Randall-Sundrum compactification. *Phys. Rev. D*, 61:064003, 2000. doi: 10.1103/PhysRevD.61.064003.

Wontae Kim and Mungon Nam. Entanglement entropy of asymptotically flat non-extremal and extremal black holes with an island. *Eur. Phys. J. C*, 81(10):869, 2021. doi: 10.1140/epjc/s10052-021-09680-x.

Ian I. Kogan, Stavros Mouslopoulos, and Antonios Papazoglou. A New bigravity model with exclusively positive branes. *Phys. Lett. B*, 501:140–149, 2001a. doi: 10.1016/S0370-2693(01)00096-X.

Ian I. Kogan, Stavros Mouslopoulos, and Antonios Papazoglou. A New bigravity model with exclusively positive branes. *Phys. Lett. B*, 501:140–149, 2001b. doi: 10.1016/S0370-2693(01)00096-X.

Hanspeter Kraft and Claudio Procesi. On the geometry of conjugacy classes in classical groups. *Comment. Math. Helv.*, 57(4):539–602, 1982. ISSN 0010-2571. doi: 10.1007/BF02565876.

Hanspeter Kraft and Claudio Procesi. A special decomposition of the nilpotent cone of a classical Lie algebra. *Astérisque*, 173(174):271–279, 1989. URL [http://www.numdam.org/item/AST\\_1989\\_\\_173-174\\_\\_271\\_0/](http://www.numdam.org/item/AST_1989__173-174__271_0/).

Per Kraus. Dynamics of anti-de Sitter domain walls. *JHEP*, 12:011, 1999. doi: 10.1088/1126-6708/1999/12/011.

Chethan Krishnan. Critical Islands. 7 2020.

Chethan Krishnan, Vaishnavi Patil, and Jude Pereira. Page Curve and the Information Paradox in Flat Space. 5 2020.

Jorrit Kruthoff, Raghu Mahajan, and Chitraang Murdia. Free fermion entanglement with a semitransparent interface: the effect of graybody factors on entanglement islands. *SciPost Phys.*, 11:063, 2021. doi: 10.21468/SciPostPhys.11.3.063.

K. Kuchar. Ground state functional of the linearized gravitational field. *J. Math. Phys.*, 11:3322–3334, 1970. doi: 10.1063/1.1665133.

Karel Kuchar. The problem of time in canonical quantization. In *Conceptual Problems of Quantum Gravity*. Birkhauser, 1991.

Jaydeep Kumar Basak, Debarshi Basu, Vinay Malvimat, Himanshu Parihar, and Gautam Sengupta. Islands for Entanglement Negativity. 12 2020a.

Jaydeep Kumar Basak, Debarshi Basu, Vinay Malvimat, Himanshu Parihar, and Gautam Sengupta. Islands for Entanglement Negativity. 12 2020b.

Jaydeep Kumar Basak, Debarshi Basu, Vinay Malvimat, Himanshu Parihar, and Gautam Sengupta. Page Curve for Entanglement Negativity through Geometric Evaporation. 6 2021.

Alok Laddha, Siddharth G. Prabhu, Suvrat Raju, and Pushkal Shrivastava. The Holographic Nature of Null Infinity. 2 2020.

Kevin Langhoff and Yasunori Nomura. Ensemble from Coarse Graining: Reconstructing the Interior of an Evaporating Black Hole. *Phys. Rev. D*, 102(8):086021, 2020. doi: 10.1103/PhysRevD.102.086021.

Kevin Langhoff, Chitraang Murdia, and Yasunori Nomura. Multiverse in an inverted island. *Phys. Rev. D*, 104(8):086007, 2021. doi: 10.1103/PhysRevD.104.086007.

Ji Hoon Lee, Dominik Neuenfeld, and Ashish Shukla. Bounds on gravitational brane couplings and tomography in  $\text{AdS}_3$  black hole microstates. *JHEP*, 10:139, 2022. doi: 10.1007/JHEP10(2022)139.

Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *JHEP*, 08:090, 2013a. doi: 10.1007/JHEP08(2013)090.

Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *JHEP*, 08:090, 2013b. doi: 10.1007/JHEP08(2013)090.

Dongqi Li and Rong-Xin Miao. Massless Entanglement Islands in Cone Holography. 3 2023.

Tianyi Li, Ma-Ke Yuan, and Yang Zhou. Defect Extremal Surface for Reflected Entropy. 8 2021.

Yi Ling, Yuxuan Liu, and Zhuo-Yu Xian. Island in Charged Black Holes. 9 2020.

Hong Liu and Julian Sonner. Quantum many-body physics from a gravitational lens. *Nature Rev. Phys.*, 2(11):615–633, 2020. doi: 10.1038/s42254-020-0225-1.

Hong Liu and Shreya Vardhan. A dynamical mechanism for the Page curve from quantum chaos. *JHEP*, 03:088, 2021a. doi: 10.1007/JHEP03(2021)088.

Hong Liu and Shreya Vardhan. Entanglement entropies of equilibrated pure states in quantum many-body systems and gravity. *P. R. X. Quantum.*, 2:010344, 2021b. doi: 10.1103/PRXQuantum.2.010344.

Yuxuan Liu, Zhuo-Yu Xian, Cheng Peng, and Yi Ling. Black holes entangled by radiation. *Journal of High Energy Physics*, 2022(9), sep 2022. doi: 10.1007/jhep09(2022)179. URL <https://doi.org/10.1007%2Fjhep09%282022%29179>.

Ivan Losev. Deformations of symplectic singularities and orbit method for semisimple Lie algebras. *Selecta Mathematica (New Series)*, 28(2):Paper no. 30, 52pp, 2022. doi: 10.1007/s00029-021-00754-y.

Daniel W Lozier. Nist digital library of mathematical functions. *Annals of Mathematics and Artificial Intelligence*, 38(1-3):105–119, 2003.

Yizhou Lu and Jiong Lin. Islands in Kaluza-Klein black holes. 6 2021.

Elihu Lubkin. Entropy of an n-system from its correlation with ak-reservoir. *Journal of Mathematical Physics*, 19(5):1028–1031, 1978.

George Lusztig. *Characters of reductive groups over a finite field*, volume 107 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1984. ISBN 0-691-08350-9; 0-691-08351-7. doi: 10.1515/9781400881772.

Juan Maldacena and Leonard Susskind. Cool horizons for entangled black holes. *Fortsch. Phys.*, 61:781–811, 2013. doi: 10.1002/prop.201300020.

Juan Maldacena, Douglas Stanford, and Zhenbin Yang. Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space. *PTEP*, 2016 (12):12C104, 2016. doi: 10.1093/ptep/ptw124.

Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Adv. Theor. Math. Phys.*, 2:231–252, 1998a. doi: 10.4310/ATMP.1998.v2.n2.a1.

Juan Martin Maldacena. The Large N limit of superconformal field theories and supergravity. *Adv. Theor. Math. Phys.*, 2:231–252, 1998b. doi: 10.1023/A:1026654312961.

Juan Martin Maldacena. Eternal black holes in anti-de Sitter. *JHEP*, 04:021, 2003. doi: 10.1088/1126-6708/2003/04/021.

Juan Martin Maldacena, Jeremy Michelson, and Andrew Strominger. Anti-de Sitter fragmentation. *JHEP*, 02:011, 1999. doi: 10.1088/1126-6708/1999/02/011.

A. Manu, K. Narayan, and Partha Paul. Cosmological singularities, entanglement and quantum extremal surfaces. *JHEP*, 04:200, 2021. doi: 10.1007/JHEP04(2021)200.

Eyal Markman. Spectral curves and integrable systems. *Compositio Mathematica*, 93(3):255–290, 1994. URL [http://www.numdam.org/item/CM\\_1994\\_\\_93\\_3\\_255\\_0/](http://www.numdam.org/item/CM_1994__93_3_255_0/).

Donald Marolf. Comments on Microcausality, Chaos, and Gravitational Observables. *Class. Quant. Grav.*, 32(24):245003, 2015. doi: 10.1088/0264-9381/32/24/245003.

Donald Marolf and Jorge E. Santos. Phases of Holographic Hawking Radiation on spatially compact spacetimes. *JHEP*, 10:250, 2019. doi: 10.1007/JHEP10(2019)250.

Donald Marolf, Mukund Rangamani, and Toby Wiseman. Holographic thermal field theory on curved spacetimes. *Class. Quant. Grav.*, 31:063001, 2014. doi: 10.1088/0264-9381/31/6/063001.

Emil J. Martinec and Nicholas P. Warner. Integrable systems and supersymmetric gauge theory. *Nucl. Phys. B*, 459:97–112, 1996. doi: 10.1016/0550-3213(95)00588-9.

Mario Martone. The constraining power of Coulomb Branch Geometry: lectures on Seiberg-Witten theory. In *Young Researchers Integrability School and Workshop 2020: A modern primer for superconformal field theories*, 6 2020.

Samir D. Mathur. The Fuzzball proposal for black holes: An Elementary review. *Fortsch. Phys.*, 53:793–827, 2005. doi: 10.1002/prop.200410203.

Samir D Mathur. The information paradox: a pedagogical introduction. *Classical and Quantum Gravity*, 26(22):224001, October 2009. ISSN 1361-6382. doi: 10.1088/0264-9381/26/22/224001. URL <http://dx.doi.org/10.1088/0264-9381/26/22/224001>.

Alex May and David Wakeham. Quantum tasks require islands on the brane. *Class. Quant. Grav.*, 38(14):144001, 2021. doi: 10.1088/1361-6382/ac025d.

D.M. McAvity and H. Osborn. Conformal field theories near a boundary in general dimensions. *Nucl. Phys. B*, 455:522–576, 1995. doi: 10.1016/0550-3213(95)00476-9.



Zeeya Merali. A video tour of the standard model, 2021. URL <https://www.quantamagazine.org/a-video-tour-of-the-standard-model-20210716/>. Quanta Magazine, Published July 16, 2021.

Rong-Xin Miao. An Exact Construction of Codimension two Holography. 9 2020.

Rong-Xin Miao. Codimension-n holography for cones. *Phys. Rev. D*, 104(8): 086031, 2021. doi: 10.1103/PhysRevD.104.086031.

Rong-Xin Miao. Massless Entanglement Island in Wedge Holography. 12 2022a.

Rong-Xin Miao. Massless Entanglement Island in Wedge Holography. 12 2022b.

Rong-Xin Miao. Entanglement island and Page curve in wedge holography. *JHEP*, 03:214, 2023a. doi: 10.1007/JHEP03(2023)214.

Rong-Xin Miao. Entanglement island and Page curve in wedge holography. *JHEP*, 03:214, 2023b. doi: 10.1007/JHEP03(2023)214.

Rong-Xin Miao. Entanglement Island and Page Curve in Wedge Holography. 1 2023c.

Rong-Xin Miao. Entanglement island versus massless gravity. *Eur. Phys. J. C*, 84(2):123, 2024. doi: 10.1140/epjc/s10052-024-12485-3.

Andre Miemiec. A Power law for the lowest eigenvalue in localized massive gravity. *Fortsch. Phys.*, 49:747–755, 2001a. doi: 10.1002/1521-3978(200107)49:7<747::AID-PROP747>3.0.CO;2-T.

Andre Miemiec. A Power law for the lowest eigenvalue in localized massive gravity. *Fortsch. Phys.*, 49:747–755, 2001b. doi: 10.1002/1521-3978(200107)49:7<747::AID-PROP747>3.0.CO;2-T.

Masamichi Miyaji. Island for Gravitationally Prepared State and Pseudo Entanglement Wedge. 9 2021.

Akihiro Miyata and Tomonori Ugajin. Evaporation of black holes in flat space entangled with an auxiliary universe. 3 2021a.

Akihiro Miyata and Tomonori Ugajin. Entanglement between two evaporating black holes. 11 2021b.

Ali Mollabashi, Noburo Shiba, and Tadashi Takayanagi. Entanglement between Two Interacting CFTs and Generalized Holographic Entanglement Entropy. *JHEP*, 04:185, 2014. doi: 10.1007/JHEP04(2014)185.

Gregory W. Moore. Lecture notes for felix klein lectures. 2012. URL <https://api.semanticscholar.org/CorpusID:34231309>.

W. Nahm. Supersymmetries and Their Representations. *Nucl. Phys. B*, 135: 149, 1978. doi: 10.1201/9781482268737-2.

Cao H. Nam. Entanglement entropy and Page curve of black holes with island in massive gravity. 8 2021.

Pranjal Nayak, Ashish Shukla, Ronak M. Soni, Sandip P. Trivedi, and V. Vishal. On the Dynamics of Near-Extremal Black Holes. *JHEP*, 09:048, 2018. doi: 10.1007/JHEP09(2018)048.

Andrew Neitzke. *Hitchin systems in  $N=2$  field theory.*, pages 53–77. 2016. doi: 10.1007/978-3-319-18769-3\_3.

Dominik Neuenfeld. Double Holography as a Model for Black Hole Complementarity. 5 2021a.

Dominik Neuenfeld. The Dictionary for Double Holography and Graviton Masses in  $d$  Dimensions. 4 2021b.

Tatsuma Nishioka. Entanglement entropy: holography and renormalization group. *Rev. Mod. Phys.*, 90(3):035007, 2018. doi: 10.1103/RevModPhys.90.035007.

Yasunori Nomura. From the Black Hole Conundrum to the Structure of Quantum Gravity. *Mod. Phys. Lett. A*, 36(08):2130007, 2021. doi: 10.1142/S021773232130007X.

Farzad Omidi. Entropy of hawking radiation for two-sided hyperscaling violating black branes. 2021.

Hidetoshi Omiya and Zixia Wei. Causal Structures and Nonlocality in Double Holography. 7 2021.

Hiroshi Ooguri and Cumrun Vafa. On the Geometry of the String Landscape and the Swampland. *Nucl. Phys. B*, 766:21–33, 2007. doi: 10.1016/j.nuclphysb.2006.10.033.

Don N. Page. Information in black hole radiation. *Phys. Rev. Lett.*, 71:3743–3746, 1993a. doi: 10.1103/PhysRevLett.71.3743.

Don N. Page. Average entropy of a subsystem. *Phys. Rev. Lett.*, 71:1291–1294, 1993b. doi: 10.1103/PhysRevLett.71.1291.

Don N. Page. Information in black hole radiation. *Phys. Rev. Lett.*, 71:3743–3746, 1993c. doi: 10.1103/PhysRevLett.71.3743.

Kyriakos Papadodimas. A class of non-equilibrium states and the black hole interior. 8 2017.

Kyriakos Papadodimas and Suvrat Raju. An Infalling Observer in AdS/CFT. *JHEP*, 10:212, 2013a. doi: 10.1007/JHEP10(2013)212.

Kyriakos Papadodimas and Suvrat Raju. An Infalling Observer in AdS/CFT. *JHEP*, 10:212, 2013b. doi: 10.1007/JHEP10(2013)212.

Kyriakos Papadodimas and Suvrat Raju. State-Dependent Bulk-Boundary Maps and Black Hole Complementarity. *Phys. Rev. D*, 89(8):086010, 2014a. doi: 10.1103/PhysRevD.89.086010.

Kyriakos Papadodimas and Suvrat Raju. State-Dependent Bulk-Boundary Maps and Black Hole Complementarity. *Phys. Rev. D*, 89(8):086010, 2014b. doi: 10.1103/PhysRevD.89.086010.

Kyriakos Papadodimas and Suvrat Raju. Remarks on the necessity and implications of state-dependence in the black hole interior. *Phys. Rev. D*, 93(8):084049, 2016. doi: 10.1103/PhysRevD.93.084049.

Kyriakos Papadodimas, Suvrat Raju, and Pushkal Shrivastava. A simple quantum test for smooth horizons. *JHEP*, 12:003, 2020. doi: 10.1007/JHEP12(2020)003.

Juan F. Pedraza, Andrew Svesko, Watse Sybesma, and Manus R. Visser. Semi-classical thermodynamics of quantum extremal surfaces in Jackiw-Teitelboim gravity. *JHEP*, 12:134, 2021. doi: 10.1007/JHEP12(2021)134.

Geoff Penington, Stephen H. Shenker, Douglas Stanford, and Zhenbin Yang. Replica wormholes and the black hole interior. 11 2019.

Geoffrey Penington. Entanglement Wedge Reconstruction and the Information Paradox. 5 2019.

Carlos Perez-Pardavila. Entropy of Radiation with Dynamical Gravity. 2 2023a.

Carlos Perez-Pardavila. Entropy of Radiation with Dynamical Gravity. 2 2023b.

Carlos Perez-Pardavila. Entropy of radiation with dynamical gravity. *JHEP*, 05:038, 2023c. doi: 10.1007/JHEP05(2023)038.

J. Polchinski. *String theory. Vol. 1: An introduction to the bosonic string.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 12 2007a. ISBN 978-0-511-25227-3, 978-0-521-67227-6, 978-0-521-63303-1. doi: 10.1017/CBO9780511816079.

J. Polchinski. *String theory. Vol. 2: Superstring theory and beyond.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 12 2007b. ISBN 978-0-511-25228-0, 978-0-521-63304-8, 978-0-521-67228-3. doi: 10.1017/CBO9780511618123.

J. Polchinski. *String theory. Vol. 1.* Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2007c. ISBN 9780511252273, 9780521672276, 9780521633031. doi: 10.1017/CBO9780511816079.

Joseph Polchinski. Dirichlet branes and ramond-ramond charges. *Phys. Rev. Lett.*, 75:4724–4727, Dec 1995. doi: 10.1103/PhysRevLett.75.4724. URL <https://link.aps.org/doi/10.1103/PhysRevLett.75.4724>.

A.M. Polyakov. Quantum geometry of bosonic strings. *Physics Letters B*, 103(3):207–210, 1981. ISSN 0370-2693. doi: [https://doi.org/10.1016/0370-2693\(81\)90743-7](https://doi.org/10.1016/0370-2693(81)90743-7). URL <https://www.sciencedirect.com/science/article/pii/0370269381907437>.

M. Porrati. Higgs phenomenon for 4-D gravity in anti-de Sitter space. *JHEP*, 04:058, 2002a. doi: 10.1088/1126-6708/2002/04/058.

M. Porrati. Higgs phenomenon for the graviton in ADS space. *Mod. Phys. Lett. A*, 18:1793–1802, 2003. doi: 10.1142/S0217732303011745.

Massimo Porrati. Mass and gauge invariance 4. Holography for the Karch-Randall model. *Phys. Rev. D*, 65:044015, 2002b. doi: 10.1103/PhysRevD.65.044015.

Massimo Porrati and Andrei Starinets. On the graviton selfenergy in AdS(4). *Phys. Lett. B*, 532:48–54, 2002. doi: 10.1016/S0370-2693(02)01490-9.

Suvrat Raju. Lessons from the Information Paradox. 12 2020.

Suvrat Raju. Lessons from the information paradox, 2021. URL <https://arxiv.org/abs/2012.05770>.

Suvrat Raju. Lessons from the information paradox. *Phys. Rept.*, 943:2187, 2022a. doi: 10.1016/j.physrep.2021.10.001.

Suvrat Raju. Failure of the split property in gravity and the information paradox. *Class. Quant. Grav.*, 39(6):064002, 2022b. doi: 10.1088/1361-6382/ac482b.

Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999a. doi: 10.1103/PhysRevLett.83.3370.

Lisa Randall and Raman Sundrum. An Alternative to compactification. *Phys. Rev. Lett.*, 83:4690–4693, 1999b. doi: 10.1103/PhysRevLett.83.4690.

Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.*, 83:3370–3373, 1999c. doi: 10.1103/PhysRevLett.83.3370.

Lisa Randall and Raman Sundrum. An Alternative to compactification. *Phys. Rev. Lett.*, 83:4690–4693, 1999d. doi: 10.1103/PhysRevLett.83.4690.

Mukund Rangamani and Tadashi Takayanagi. *Holographic Entanglement Entropy*, volume 931. Springer, 2017a. doi: 10.1007/978-3-319-52573-0.

Mukund Rangamani and Tadashi Takayanagi. *Holographic Entanglement Entropy*, volume 931 of *Lecture Notes in Physics*. Springer Cham, 1 edition,

May 2017b. ISBN 978-3-319-52571-6. doi: 10.1007/978-3-319-52573-0. URL <https://doi.org/10.1007/978-3-319-52573-0>.

Harvey Reall, Norihiro Tanahashi, and Benson Way. Causality and Hyperbolicity of Lovelock Theories. *Class. Quant. Grav.*, 31:205005, 2014. doi: 10.1088/0264-9381/31/20/205005.

Wyatt Reeves, Moshe Rozali, Petar Simidzija, James Sully, Christopher Waddell, and David Wakeham. Looking for (and not finding) a bulk brane. *JHEP*, 12:002, 2021. doi: 10.1007/JHEP12(2021)002.

Dean Rickles. *A Brief History of String Theory. From Dual Models to M-Theory*. The Frontiers Collection. Springer, Heidelberg, Germany, 2014. ISBN 978-3-642-45127-0, 978-3-662-50183-2, 978-3-642-45128-7. doi: 10.1007/978-3-642-45128-7.

Marcos Riojas. The photon sphere in double-holography. *in preparation*.

Jorge V. Rocha. Evaporation of large black holes in AdS: coupling to the evaporon. URL <https://arxiv.org/pdf/0804.0055.pdf>.

Moshe Rozali, James Sully, Mark Van Raamsdonk, Christopher Waddell, and David Wakeham. Information radiation in BCFT models of black holes. *JHEP*, 05:004, 2020. doi: 10.1007/JHEP05(2020)004.

Romain Ruzziconi and Céline Zwikel. Conservation and Integrability in Lower-Dimensional Gravity. *JHEP*, 04:034, 2021. doi: 10.1007/JHEP04(2021)034.

Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from AdS/CFT. *Phys. Rev. Lett.*, 96:181602, 2006a. doi: 10.1103/PhysRevLett.96.181602.

Shinsei Ryu and Tadashi Takayanagi. Aspects of Holographic Entanglement Entropy. *JHEP*, 08:045, 2006b. doi: 10.1088/1126-6708/2006/08/045.

Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from AdS/CFT. *Phys. Rev. Lett.*, 96:181602, 2006c. doi: 10.1103/PhysRevLett.96.181602.

Shinsei Ryu and Tadashi Takayanagi. Aspects of Holographic Entanglement Entropy. *JHEP*, 08:045, 2006d. doi: 10.1088/1126-6708/2006/08/045.

Phil Saad, Stephen H. Shenker, and Douglas Stanford. JT gravity as a matrix integral. 3 2019.

Subir Sachdev. Universal low temperature theory of charged black holes with  $\text{AdS}_2$  horizons. *J. Math. Phys.*, 60(5):052303, 2019. doi: 10.1063/1.5092726.

Matthew D. Schwartz. The Emergence of localized gravity. *Phys. Lett. B*, 502:223–228, 2001. doi: 10.1016/S0370-2693(01)00152-6.

N. Seiberg and Edward Witten. Monopoles, duality and chiral symmetry breaking in  $\mathcal{N}=2$  supersymmetric QCD. *Nucl. Phys. B*, 431:484–550, 1994a. doi: 10.1016/0550-3213(94)90214-3.

N. Seiberg and Edward Witten. Electric - magnetic duality, monopole condensation, and confinement in  $\mathcal{N}=2$  supersymmetric Yang-Mills theory. *Nucl. Phys. B*, 426:19–52, 1994b. doi: 10.1016/0550-3213(94)90124-4. [Erratum: Nucl.Phys.B 430, 485–486 (1994)].

N. Seiberg and Edward Witten. Monopoles, duality and chiral symmetry breaking in  $\mathcal{N}=2$  supersymmetric QCD. *Nucl. Phys.*, B431:484–550, 1994c. doi: 10.1016/0550-3213(94)90214-3.

N. Seiberg and Edward Witten. Electric-magnetic duality, monopole condensation, and confinement in  $\mathcal{N} = 2$  supersymmetric Yang-Mills theory. *Nucl. Phys.*, B426:19–52, 1994d. doi: 10.1016/0550-3213(94)90124-4. [Erratum: *Nucl. Phys.* **B430** (1994) 485].



N. Seiberg and Edward Witten. Comments on string dynamics in six-dimensions. *Nucl. Phys. B*, 471:121–134, 1996a. doi: 10.1016/0550-3213(96)00189-7.

Nathan Seiberg. Nontrivial fixed points of the renormalization group in six-dimensions. *Phys. Lett. B*, 390:169–171, 1997a. doi: 10.1016/S0370-2693(96)01424-4.

Nathan Seiberg. New theories in six-dimensions and matrix description of M theory on  $T^{*5}$  and  $T^{*5} / Z(2)$ . *Phys. Lett. B*, 408:98–104, 1997b. doi: 10.1016/S0370-2693(97)00805-8.

Nathan Seiberg. Notes on theories with 16 supercharges. *Nucl. Phys. B Proc. Suppl.*, 67:158–171, 1998. doi: 10.1016/S0920-5632(98)00128-5.

Nathan Seiberg and Edward Witten. Gauge dynamics and compactification to three-dimensions. In *The Mathematical Beauty of Physics: A memorial volume for Claude Itzykson. Proceedings, Conference, Saclay, France, June 5-7, 1996*, pages 333–366, 1996b.

Edgar Shaghoulian. The central dogma and cosmological horizons. 10 2021.

Stephen H. Shenker and Douglas Stanford. Black holes and the butterfly effect. *JHEP*, 03:067, 2014. doi: 10.1007/JHEP03(2014)067.

Eric Sommers. Lusztig’s canonical quotient and generalized duality. *J. Algebra*, 243(2):790–812, 2001. ISSN 0021-8693. doi: 10.1006/jabr.2001.8868.

N Spaltenstein. Polynomials over local fields, nilpotent orbits and conjugacy classes in Weyl groups. *Astérisque*, 168:191–217, 1988. URL [http://www.numdam.org/item/AST\\_1988\\_\\_168\\_\\_191\\_0/](http://www.numdam.org/item/AST_1988__168__191_0/).

N. Spaltenstein. A note on the Kazhdan-Lusztig map for even orthogonal Lie algebras. *Archiv der Mathematik*, 55(5):431–437, 1990a. ISSN 0003-889X. doi: 10.1007/BF01190263.

N Spaltenstein. On the Kazhdan-Lusztig map for exceptional Lie algebras. *Advances in Mathematics*, 83(1):48–74, 1990b. doi: 10.1007/BF01444622.

Nicolas Spaltenstein. *Classes unipotentes et sous-groupes de Borel*, volume 946. Springer, 1982.

Andrew Strominger. Open p-branes. *Phys. Lett. B*, 383:44–47, 1996. doi: 10.1201/9781482268737-13.

Andrew Strominger and Cumrun Vafa. Microscopic origin of the bekenstein-hawking entropy. *Physics Letters B*, 379(1–4):99–104, June 1996a. ISSN 0370-2693. doi: 10.1016/0370-2693(96)00345-0. URL [http://dx.doi.org/10.1016/0370-2693\(96\)00345-0](http://dx.doi.org/10.1016/0370-2693(96)00345-0).

Andrew Strominger and Cumrun Vafa. Microscopic origin of the Bekenstein-Hawking entropy. *Phys. Lett.*, B379:99–104, 1996b. doi: 10.1016/0370-2693(96)00345-0.

James Sully, Mark Van Raamsdonk, and David Wakeham. BCFT entanglement entropy at large central charge and the black hole interior. *JHEP*, 03:167, 2021. doi: 10.1007/JHEP03(2021)167.

Po-Chun Sun. Entanglement Islands from Holographic Thermalization of Rotating Charged Black Hole. 8 2021.

Raman Sundrum. Tasi 2004 lectures: To the fifth dimension and back. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics in  $D \geq 4$* , pages 585–630, 8 2005.

Leonard Susskind. The World as a hologram. *J. Math. Phys.*, 36:6377–6396, 1995. doi: 10.1063/1.531249.

Gerard 't Hooft. Dimensional reduction in quantum gravity. *arXiv: General Relativity and Quantum Cosmology*, 1993. URL <https://api.semanticscholar.org/CorpusID:119379556>.

Yuji Tachikawa. *N=2 supersymmetric dynamics for pedestrians*. 12 2013. doi: 10.1007/978-3-319-08822-8.

Tadashi Takayanagi. Holographic Dual of BCFT. *Phys. Rev. Lett.*, 107: 101602, 2011a. doi: 10.1103/PhysRevLett.107.101602.

Tadashi Takayanagi. Holographic Dual of BCFT. *Phys. Rev. Lett.*, 107: 101602, 2011b. doi: 10.1103/PhysRevLett.107.101602.

C. Teitelboim. Gravitation and Hamiltonian Structure in Two Space-Time Dimensions. *Phys. Lett. B*, 126:41–45, 1983. doi: 10.1016/0370-2693(83)90012-6.

Christoph F. Uhlemann. Information transfer with a twist. 11 2021a.

Christoph F. Uhlemann. Islands and Page curves in 4d from Type IIB. *JHEP*, 08:104, 2021b. doi: 10.1007/JHEP08(2021)104.

Cumrun Vafa. The String landscape and the swampland. 2005.

H. van Dam and M. J. G. Veltman. Massive and massless Yang-Mills and gravitational fields. *Nucl. Phys. B*, 22:397–411, 1970. doi: 10.1016/0550-3213(70)90416-5.

Herman L. Verlinde. Holography and compactification. *Nucl. Phys. B*, 580: 264–274, 2000. doi: 10.1016/S0550-3213(00)00224-8.

Aron C. Wall. Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy. *Class. Quant. Grav.*, 31(22):225007, 2014a. doi: 10.1088/0264-9381/31/22/225007.

Aron C. Wall. Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy. *Class. Quant. Grav.*, 31(22):225007, 2014b. doi: 10.1088/0264-9381/31/22/225007.

Xuanhua Wang, Ran Li, and Jin Wang. Page curves for a family of exactly solvable evaporating black holes. *Phys. Rev. D*, 103(12):126026, 2021a. doi: 10.1103/PhysRevD.103.126026.

Xuanhua Wang, Ran Li, and Jin Wang. Islands and Page curves of Reissner-Nordström black holes. *JHEP*, 04:103, 2021b. doi: 10.1007/JHEP04(2021)103.

Steven Weinberg. Phenomenological Lagrangians. *Physica A*, 96(1-2):327–340, 1979. doi: 10.1016/0378-4371(79)90223-1.

Steven Weinberg. Four golden lessons. *Nature*, 426(6965):389–389, 2003. ISSN 1476-4687. doi: 10.1038/426389a. URL <https://doi.org/10.1038/426389a>.

J. Wess and J. Bagger. *Supersymmetry and supergravity*. Princeton University Press, Princeton, NJ, USA, 1992. ISBN 978-0-691-02530-8.

Edward Witten. Some comments on string dynamics. In *STRINGS 95: Future Perspectives in String Theory*, pages 501–523, 7 1995.

Edward Witten. Five-branes and M-theory on an orbifold. *Nucl. Phys. B*, 463:383–397, 1996. doi: 10.1201/9781482268737-19.

Edward Witten. Anti-de Sitter space and holography. *Adv. Theor. Math. Phys.*, 2:253–291, 1998a. doi: 10.4310/ATMP.1998.v2.n2.a2.

Edward Witten. Anti-de Sitter space, thermal phase transition, and confinement in gauge theories. *Adv. Theor. Math. Phys.*, 2:505–532, 1998b. doi: 10.4310/ATMP.1998.v2.n3.a3.

Edward Witten. Mirror Symmetry, Hitchin’s Equations, And Langlands Duality. 2 2008. doi: 10.1093/acprof:oso/9780199534920.003.0007.

Edward Witten. Fivebranes and Knots. 1 2011.

Edward Witten. APS Medal for Exceptional Achievement in Research: Invited article on entanglement properties of quantum field theory. *Rev. Mod. Phys.*, 90(4):045003, 2018. doi: 10.1103/RevModPhys.90.045003.

Zhenbin Yang. The Quantum Gravity Dynamics of Near Extremal Black Holes. *JHEP*, 05:205, 2019. doi: 10.1007/JHEP05(2019)205.

Ming-Hui Yu and Xian-Hui Ge. Page Curves and Islands in Charged Dilaton Black Holes. 7 2021.

Zhiwei Yun. Minimal reduction type and the Kazhdan-Lusztig map. *Indag. Math. (N.S.)*, 32(6):1240–1274, 2021. ISSN 0019-3577. doi: 10.1016/j.indag.2021.06.007.

V. I. Zakharov. Linearized gravitation theory and the graviton mass. *JETP Lett.*, 12:312, 1970.

# Vita

Carlos José Pérez Pardavila was born in Montevideo, Uruguay, to Carlos Pérez Busto and Beatriz Pardavila Arcas. He is the grandchild of Galician immigrants José Pérez Rodríguez, Amalia Busto Lopez, Manuel Pardavila Iglesias, and Argentina Arcas Sanchez. He received a Master in Physics from the University of St Andrews and a Master of Advanced Studies in Mathematics from the University of Cambridge. In 2018, he moved to Austin, Texas, to pursue a Doctor of Philosophy in Physics at The University of Texas at Austin. His doctoral research was conducted under the supervision of Jacques Distler and co-supervised by Andreas Karch.

Address: [cjp3247@utexas.edu](mailto:cjp3247@utexas.edu)

This dissertation was typeset with  $\text{\LaTeX}^\dagger$  by the author.

---

<sup>†</sup> $\text{\LaTeX}$  is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's  $\text{\TeX}$  Program.