Towards the Black Hole Interior

Henry Wanjune Lin

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Juan Martín Maldacena

September 2022

Abstract

This thesis elucidates various aspects of the black hole interior in low-dimensional models of quantum gravity. In Chapter 2, we construct 3 symmetry generators that form a Lie algebra $\mathrm{SL}(2,\mathbb{R})$. These generators act on objects in the interior of black hole and either boost or translate them in spacetime, allowing us to explore the interior region. In the dual quantum mechanical description of the black hole, these generators are related to the notion of "operator size" in quantum mechanics, giving a symmetries-based perspective on quantum chaos and teleportation. Chapters 3 and 4 are devoted to exploring novel consequences of non-perturbative quantum effects associated with the black hole interior. In Chapter 3, we show how global symmetries are violated in the black hole interior. Using certain relative entropies defined on the Hawking radiation to diagnose the symmetry violation, one sees a large $O(1)$ violation after the Page time, whereas more conventional scattering amplitudes/correlation functions only receive exponentially suppressed violations. In Chapter 4, we quantify the extent to which the black hole interior is highly sensitive to the specific values of the couplings of the theory. We consider a setup where a black hole is in an entangled superposition of states with an auxiliary reference system. The reference records the specific values of the couplings associated to each of the black hole states. At late times, the entanglement wedge of the reference contains an island, signaling the failure of bulk reconstruction in the black hole interior without precise information about the couplings.

This thesis is based on three papers: [1] with Juan Maldacena and Ying Zhao, [2] with Yiming Chen, and [3] with Ahmed Almheiri.

"In the beginner's mind, there are many possibilities, in the expert's mind there are few."

— Shunryu Suzuki

Acknowledgements

It takes a village to raise a child, but it takes a spaceship full of misfits to raise a physicist.

To Lenny: you introduced me to the world of physics through your online lectures, yet somehow that wasn't generous enough. Thank you for sharing your intuitions, teaching me how to think imprecisely, and encouraging me to listen to my own scientific voice.

Thank you Nadir Jeevanjee, Jason Rhodes, Mike McDonald, and Adam Brown for your mentorship and guidance in my early days, without which I would not have found this path. Thanks to Daniel Harlow, Douglas Stanford, and Dan Roberts who patiently explained many things to me when I was getting started. I particularly remember Douglas's advice: "Instead of agonizing over whether to calculate, just calculate."

I am grateful to Igor Klebanov, Herman Verlinde, and Edward Witten for memorable discussions and support over the years, and to the NDSEG and Charles and Elizabeth Proctor fellowship for funding during my PhD.

Thank you Ahmed, Adam, Alex, Baur, Geoff, Phil, Ven, Yiming, and Ying for somehow creating genuine camaraderie at the IAS during the pandemic. To Makinde, Wayne, Himanshu, Kurt, Brad, and Megan, thanks for the sorely needed levity.

And to my collaborators — Ahmed Almeiri, Ying Zhao, and Yiming Chen, thank you for teaching me so much. Ahmed, I will miss leaving my coffee in your office.

And lastly, thank you Juan for sharing your many gifts with me. I will leave a thorough enumeration of Juan's gifts for future hagiographies; suffice it to say that it has been a great privilege to have an advisor who is always simultaneously the expert and the beginner.

To my parents, who encouraged me to question.

And to Cindy, for always answering my questions.

Contents

Chapter 1

Introduction

Viewed from faraway, a black hole is a rather mundane quantum mechanical system. At a gross level of approximation, a black hole is not so different from any other large N, strongly interacting quantum mechanical object, like a complicated atomic nucleus. An outside observer, call him Bob, can perform many experiments on the black hole and learn things like its energy levels, correlation functions, etc. A quantum theory of gravity should predict the outcomes of Bob's experiments, and indeed there is a growing body of work that shows how to compute quantities of interest to Bob in simple models of quantum gravity. (For recent progress on near-extremal black holes, see [4–9]) Unlike the case of a heavy nuclei, however, an adventurous observer, call her Alice, can jump into a black hole. Presumably the rules of quantum gravity should also dictate the results of Alice's experiments that she may conduct in the interior, but much less is known about these rules. It would be interesting to understand these rules not just for Alice's sake, but also for our own: after all, our universe is described by a cosmology with a cosmic event horizon that shares some similarities with the black hole horizon. So understanding what happens to Alice seems like a good warmup to understanding our own cosmology.

However part of the difficulty with the black hole interior is that it is not so clear *what* to compute, even putting aside *how* to do the computations. Indeed, in quantum gravity, simple observables, such as S-matrix elements in flat space or correlators in AdS/CFT, are anchored to asymptotic boundaries. These observables naively seem better suited to answering Bob's questions over Alice's. It is challenging to extract from boundary-anchored observables information about the interior of the black hole, which is shrouded from the boundary by a horizon. (For some attempts, see [10–12].)

In this thesis, we explore some aspects of the interior by asking questions motivated from quantum

information theory. Such questions involve more complicated quantities that go beyond simple correlators, but the advantage is that they probe the interior more directly. While we are quite far from a general understanding of the black hole interior, we hope to have found clues that will play a role in the eventual understanding.

This thesis is organized as follows:

In Chapter 2, we consider a pair of near-extremal black holes that form a two-sided wormhole, described by nearly-AdS_2 ($NAdS_2$) gravity. We construct three gauge-invariant operators in $NAdS_2$ which move bulk matter relative to the dynamical boundaries. In a two-sided system, these operators satisfy an $SL(2,\mathbb{R})$ algebra (up to non perturbative corrections). In a semiclassical limit, these generators act like $SL(2,\mathbb{R})$ transformations of the boundary time, or conformal symmetries of the two sided boundary theory. These can be used to define an operator-state mapping. A particular large N and low temperature limit of the SYK model has precisely the same structure, and this construction of the exact generators also applies. We also discuss approximate, but simpler, constructions of the generators in the SYK model. These are closely related to the size operator and are connected to the maximal chaos behavior captured by out of time order correlators.

Chapters 3 and 4 are devoted to exploring novel consequences of non-perturbative quantum effects associated with the black hole interior. These "replica wormholes" contributions are crucial for the unitarity of the Hawking process [13, 14]. Here we will see that that have other qualitative lessons in store for the black hole interior.

In Chapter 3, we reconsider the lore that exact global symmetries do not exist in theories that admit quantum black holes in light of replica wormholes. We propose a way to quantify the degree of global symmetry violation in the Hawking radiation of a black hole by using certain relative entropies. While the violations of global symmetry that we consider are non-perturbative effects, they nevertheless give $O(1)$ contributions to the relative entropy after the Page time. Furthermore, using island formulas, these relative entropies can be computed within semi-classical gravity, which we demonstrate with explicit examples. These formulas give a rather precise operational sense to the statement that a global charge thrown into an old black hole will be lost after a scrambling time. The relative entropies considered here may also be computed using a replica trick. At integer replica index, the global symmetry violating effects manifest themselves as charge flowing through the replica wormhole.

In Chapter 4, we continue to study non-perturbative effects in the black hole interior. From various points of view, these non-perturbative effects are believed to be highly theory-dependent. We sharpen this statement by considering a setup where the state of the black hole is in a superposition of

states corresponding to boundary theories with different couplings, entangled with a reference which keeps track of those couplings. The entanglement wedge of the reference can then be interpreted as the bulk region most sensitive to the values of the couplings. In simple bulk models, e.g., JT gravity + a matter BCFT, the QES formula implies that the reference contains the black hole interior at late times. We also analyze the Renyi-2 entropy of the reference, which can be viewed as a diagnostic of chaos via the Loschmidt echo. We find explicitly the replica wormhole that diagnoses the island and restores unitarity. Numerical and analytical evidence of these statements in the SYK model is presented. Similar considerations are expected to apply in higher dimensional AdS/CFT, for marginal and even irrelevant couplings.

Chapter 2

Symmetries Near the Horizon

Any black hole with finite temperature has a near horizon geometry that can be approximated by flat space. The boost symmetry of this flat space region corresponds to the full modular Hamiltonian of the outside region of the black hole, and it is an exact symmetry of the full wormhole geometry. The two translation symmetries of this flat space region are more mysterious. It is important to understand them because they can take matter into the black hole interior. In this chapter, we construct explicitly these symmetries for nearly AdS_2 gravity and also for the related SYK model.

Nearly AdS_2 gravity [15–18] captures the gravitational dynamics of near extremal black holes after a Kaluza-Klein reduction. The important gravitational mode is non-propagating and can be viewed as living at the boundaries of the nearly AdS_2 region. The action of these boundary modes is universal and can be written in terms of a Schwarzian action for a variable that can be viewed as a map from the boundary proper time to a time coordinate in a rigid AdS_2 spacetime. A similar mode appears in the description of Nearly CFT_1 ($NCFT_1$) quantum systems, such as the SYK model [19–22], which exhibit nearly conformally-invariant correlation functions at relatively low energies.

In these systems, there is an approximation where matter appears to move in a rigid AdS_2 background geometry displaying an $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ isometry group[1], henceforth denoted by $\mathrm{SL}(2,\mathbb{R})$ This approximation becomes better and better as the boundaries are further and further away. However, this does not obviously translate into a physical symmetry, since only the relative position between the boundaries and the bulk matter is physical. Nevertheless, we will find three $\mathrm{SL}(2,\mathbb{R})$generators that act on the full physical Hilbert space of the system. These generators obey an exact $\mathrm{SL}(2,\mathbb{R})$algebra, but they do not commute with the Hamiltonian. However, they have a relatively simple behavior

[1] In Euclidean signature, the isometry group is $\mathrm{PSL}(2,\mathbb{R})$. In this chapter, we are mostly concerned about the algebra and not the group.

under Hamiltonian evolution, which can be used to define "conserved" charges through a more subtle construction.

It is convenient to describe the $\text{NAdS}_2/\text{NCFT}_1$ system in terms of an extended Hilbert space with a gauge constraint. The extended Hilbert space factorizes into three pieces: two systems describing the two boundaries of the eternal black hole and a system describing the bulk matter fields. Each of them can be viewed in terms of particles moving on an exact AdS_2 spacetime [4, 6]. The physical Hilbert space is obtained by imposing an $\text{SL}(2)_g$ gauge constraint that sets to zero the overall $\text{SL}(2)_g$ charges of the three systems. This constraint imposes that only the relative position between the two boundaries, or between the boundary and the bulk matter, are physical. In this chapter we discuss physical $\text{SL}(2,\mathbb{R})$ generators which are invariant under the $\text{SL}(2)_g$ gauge symmetry. It is important not to confuse these two $\text{SL}(2,\mathbb{R})$ groups, the gauge one and the physical one. This chapter is about they physical one. Our construction will define these physical generators relative to boundary positions in such a way that they are invariant under the gauge symmetries. Due to the fact that they involve the boundary positions, they are not conserved under time evolution, since the boundary positions change in time. However, the dynamics of the boundary positions is integrable [4–6, 23–25], and one could use this fact to define "conserved" charges by simply "undoing" the boundary evolution.

Our discussion is exact in a scaling limit where we go to low temperatures, but we scale up the size of the black hole so that we keep fixed the coupling of the Schwarzian mode, or the quantum gravitational effects in AdS_2. This is a limit where the near extremal entropy $\Delta S = S - S_e$ is kept fixed[2]. In SYK variables this is the limit $N \to \infty$, $\beta J \to \infty$ with $N/(\beta J)$ fixed. We have not included finite βJ effects or finite N effects. The generators we define involve variables, such as the distance between the two boundaries, which are well defined in the gravity theory, in the scaling limit we define, but are not expected to make sense when non-perturbative effects are taken into account. In particular, they are not expected to make sense for finite N in the SYK model. This should not be surprising since unitary $\text{SL}(2,\mathbb{R})$ representations are infinite dimensional. However, we also relate the generators we defined to other operators which are well defined for finite N, but agree with the generators in the semiclassical limit. This allows us to identify operators in both a gravity theory and the SYK model which approximately obey an $\text{SL}(2,\mathbb{R})$ algebra and should be identified with the symmetries of AdS_2. These approximate symmetries behave as $SL(2)_u$ transformations of the physical boundary time of a pair of NCFT_1 s, and we give an approximate state-operator map

[2]For a four dimensional charged near extremal black hole this is $\Delta S \propto r_e^3 T/l_p^2$, where r_e is the extremal radius. We take ΔS fixed with $r_s \to \infty$, $T \to 0$.

that organizes the NCFT_1 Hilbert space into primaries and descendants, in analogy with higher dimensions.

These generators are connected to the operators that generate traversable wormholes [26]. These move matter from one side of the horizon to the other. In fact, the approximate expression for the global time translation operator is essentially the same as the coupled Hamiltonian in [27]. These approximate generators also make contact with another approach for describing bulk motion via the "size" operator in [28–30]. So the discussion in this chapter explains why such operators act like approximate $\text{SL}(2,\mathbb{R})$ isometries in the bulk. We also point out that the structure of the approximate generators is similar to the structure of the exact generators in the case of higher dimensional conformal field theories in Rindler space.

Outline. In section **two**, we review nearly AdS_2 gravity at low energies in the embedding space formalism. The Hilbert space of the system consists of two boundary modes plus arbitrary matter, with an overall $\text{SL}(2,\mathbb{R})$ gauge constraint. We briefly review how this structure also emerges in the SYK model.

In section **three**, we construct the generators which satisfy an exact $\text{SL}(2,\mathbb{R})$ algebra. Although the quadratic Casimir commutes with the usual Hamiltonians H_l or H_r, the individual generators do not. We nevertheless explain how to obtain conserved charges.

In section **four**, we consider the charges in the semi-classical limit. In this limit, the generators can be viewed as conformal symmetries of the boundary time. We show that the Hilbert space organizes into primaries and descendants and give a state-operator correspondence analogous to the higher dimensional versions.

In section **five**, we show how to use these charges to explore bulk physics. We comment on drama near the inner horizon. We also discuss applications of our construction to previous work. Our charges are closely related to the coupled Hamiltonian in [27].

In section **six**, we explain how these approximate charges can be realized in the SYK model. We also relate our charges to "size" in SYK [29–31]. We note that the generators written in terms of the microscopic variables has an analogous form in higher dimensional CFTs.

In section **seven**, we discuss some issues and draw conclusions.

In the **appendices**, we explain how to construct gauge-invariant SO(3) generators in a rather pedestrian system involving two non-relativistic particles on a sphere, plus some arbitrary matter, with an overall angular momentum gauge constraint. We also explain how to compute commutators in both the canonically quantized Schwarzian theory and its linearized cousin. We also discuss an alternative to the embedding space formalism which uses $\text{SL}(2,\mathbb{R})$ spinors instead of the vectors.

Finally, we comment on a modified eternal traversable wormhole where the oscillation frequency of the Schwarzian mode is very large.

Notation. In most of this chapter we work in units, where in the SYK language $\alpha_s N/\mathcal{J} = 1$, or in 4d near extremal charge black hole language, $\frac{r_e^3}{G_N} = 1$, where r_e is the extremal radius. Such factors can be restored by dimensional analysis.

2.1 Review

2.1.1 Review of the symmetries of AdS_2

In the embedding space formalism, AdS_2 is the universal cover of the surface defined by

$$Y.Y = \eta_{ab} Y^a Y^b = -(Y^{-1})^2 - (Y^0)^2 + (Y^1)^2 = -1. \tag{2.1}$$

From this definition, it is clear that AdS_2 has an $SO(2,1) \simeq SL(2,R)$ symmetry generated by

$$Q^a = \frac{1}{2}\epsilon^{abc} J_{bc} \ , \qquad J_{ab} = -iY_a \frac{\partial}{\partial Y^b} + iY_b \frac{\partial}{\partial Y^a}. \tag{2.2}$$

These generators satisfy the algebra

$$[Q^a, Q^b] = i\epsilon^{abc}\eta_{cd} Q^d. \tag{2.3}$$

To see how these generators act on AdS_2 more explicitly, we can solve the constraint 2.1 using global coordinates:

$$Y = (Y^{-1}, Y^0, Y^1) = \left(\frac{\cos T}{\sin\sigma}, \frac{\sin T}{\sin\sigma}, \frac{-1}{\tan\sigma}\right), \quad \sigma \in [0, \pi]. \tag{2.4}$$

Then, the Killing vectors

$$\begin{aligned} B &= Q^{-1} = J_{0,1} = i\left(-\cos T \cos\sigma \partial_T + \sin T \sin\sigma \partial_\sigma\right), \\ P &= Q^0 = -J_{-1,1} = -i\left(\sin T \cos\sigma \partial_T + \cos T \sin\sigma \partial_\sigma\right), \\ E &= Q^1 = J_{-1,0} = i\partial_T. \end{aligned} \tag{2.5}$$

Note that $P = -i[B, E]$. Near the bifurcation point $T = 0, \sigma = \pi/2$ these symmetries act as time translation/energy E, spatial translation/momentum P, and boost B. See figure 2.1. Of course, the

algebra is SL(2), not Poincare, so that $[E, P] = iB$.

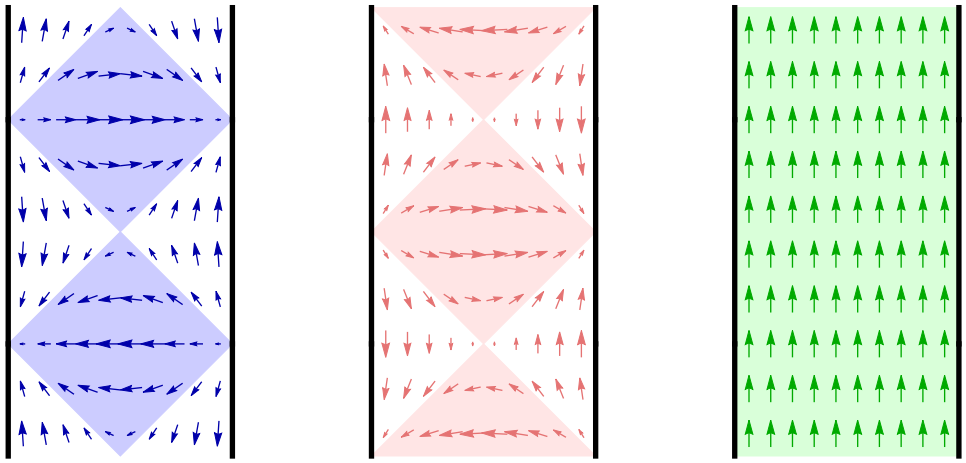


Figure 2.1: The three Killing vectors: boost B (blue), momentum P (pink), and global energy E (green) given in (2.5) in the coordinate system (2.6a). The shaded regions delineate different orbits of the symmetries.

We can choose coordinate systems [3] that simplify the action of these generators

$$\begin{aligned}
&\text{Rindler}: && ds^2 = -dt^2 \sinh^2\rho + d\rho^2, && \rho \in [-\infty, \infty], && B = i\partial_t \\
&\text{FLRW}: && ds^2 = -d\tau^2 + \sin^2\tau dx^2, && \tau \in [0, \pi], && P = -i\partial_x \\
&\text{Global}: && ds^2 = \frac{-dT^2 + d\sigma^2}{\sin^2\sigma}, && \sigma \in [0, \pi], && E = i\partial_T
\end{aligned} \tag{2.6a}$$

Notice that, given a vector W^a, we can assign a charge $Q_W = W_a Q^a$. This charge has the property that it leaves the bulk point $Y^a \propto W^a$ fixed[4]. This is basically just the familiar fact that a rotation about some axis fixes the axis.

Points at the boundary are naturally described in terms of projective coordinates, $\tilde{X}^a$ with the constraint $\tilde{X}.\tilde{X} = 0$ and the identification $\tilde{X}^a \sim \lambda \tilde{X}^a$. If we have a charge associated to the vector W^a, $Q_W = W.Q$, then this charge will leave invariant the boundary points that are light-like separated from W^a, $W.\tilde{X} = 0$.

A particle moving in AdS_2 can be described by a trajectory $Y^a(u)$ constrained to live on the surface $Y.Y = -1$. Since Y is a vector, $[Q^a, Y^b] = i\epsilon^{abc}Y_c$. For a standard massive particle, of mass

[3] The second coordinates describe a Friedmann-Lemaitre-Robertson-Walker (FLRW) cosmology.

[4] If the vector W^a is spacelike, then we will not have any fixed point in AdS_2. An example is the generator E in (2.5), see figure 2.1.

m, the charges are given by

$$Q^a = m\epsilon^a{}_{bc} Y^b \dot{Y}^c \ , \qquad Y.Y = -1 \ , \qquad \dot{Y}.\dot{Y} = -1 \tag{2.7}$$

If the particle is also charged under an electric field that is uniform in AdS_2, then the charges are

$$Q^a = m\epsilon^a{}_{bc} Y^b \dot{Y}^c - \mathfrak{q} Y^a \ , \qquad Y.Y = -1 \ , \qquad \dot{Y}.\dot{Y} = -1 \tag{2.8}$$

The charges Q^a are conserved, and the particle trajectories are given by $Q.Y = \mathfrak{q}$.

Alternatively we can say that if we have a particle moving in AdS_2, its Hilbert space has operators satsifying

$$\begin{aligned} &[Y^a, Y^b] = 0, \\ &[Q^a, Q^b] = i\epsilon^{abc}\eta_{cd} Q^d, \\ &[Q^a, Y^b] = i\epsilon^{abc}\eta_{cd} Y^d. \end{aligned} \tag{2.9}$$

This is the Poincare algebra $R^{2,1} \rtimes SL(2,R)$. The Casimirs of this algebra are

$$r^2 = Y.Y, \quad \mathfrak{q} = Y.Q. \tag{2.10}$$

The values of these Casimirs are inputs of the physical theory. For example, for a spin-less particle freely propagating in AdS_2 we have $r^2 = -1, \mathfrak{q} = 0$. From this point of view $\mathfrak{q}$ is the spin of the particle.

If we have quantum fields moving on AdS_2 the charges can be written in terms of the stress tensor and the associated Killing vector

$$Q_\zeta = \int_\Sigma n^\mu T_{\mu\nu} \zeta^\nu \tag{2.11}$$

where ζ^μ are each of the Killing vectors in (2.5). These charges are constant and independent of the spatial slice Σ used to evaluate them, if the fields obey appropriate reflecting conditions at the AdS_2 boundary.

2.1.2 Review of the nearly-AdS_2 gravity theory

We will be considering the JT theory coupled to matter as [32–35]

$$S = \phi_0 \left[\int R - 2 \int K \right] + \int \phi(R+2) - 2\phi_b \int K + S_m[g_{\mu\nu}, \chi] \tag{2.12}$$

where we have also indicated the boundary terms. The first term is topological and only contributes to the extremal entropy. We have also assumed that the matter couples to the metric but not to ϕ. We will also assume that the boundary is very far away so that matter effectively feels as if it was in exactly AdS_2 space. This is sometimes called the "Schwarzian" limit because in this case the boundary dynamics is governed by [16–18]

$$S_{\rm Sch}[t] = -\int du \, \{e^{t(u)}, u\} \ , \qquad \{f, u\} = \frac{f'''}{f'} - \frac{3}{2}\frac{f''^2}{f'^2} \ , \quad d\tau_p = 2\phi_b \, du \tag{2.13}$$

where u is a rescaled version of proper time τ_p, and t can be viewed as the Rindler time t in (2.6a) near the boundary. We can view the curve $t(u)$ as parametrizing the position of the boundary. The action (2.13) captures a gravitational degree of freedom that we can view as living on the boundary[5].

We will consider spacetimes describing a two sided eternal black hole, so that we have two boundaries and two variables t_r, t_l, each with the action (2.13). The dynamics of the full system (2.12) reduces to the dynamics of three decoupled systems connected only by an overall $\mathrm{SL}(2)_g$ constraint

$$S = S_{\rm Sch}[t_r] + S_{\rm Sch}[t_l] + S_m[g_{\mu\nu}, \chi] \tag{2.14}$$

These three decoupled systems are the following. First we have the matter which lives in exactly AdS_2 space and has $\mathrm{SL}(2)_g$ charges Q^a_m. Then we have the right and left boundaries. In this limit, they are not directly coupled to each other or to the matter. However, in NAdS_2 gravity, an overall $\mathrm{SL}(2)_g$ transformation is a redundancy of our description. Hence the physical Hilbert space is [16–18]

$$\mathcal{H}_{\rm Physical} = (\mathcal{H}_l \times \mathcal{H}_{matter} \times \mathcal{H}_r)/SL(2)_g \ , \qquad Q^a_l + Q^a_m + Q^a_r = 0 \tag{2.15}$$

where the charges Q^a are the $\mathrm{SL}(2)_g$ charges of each of the systems. Physically, this says that only the relative positions of the matter and the boundaries matter. As pointed out in [6], we can view it as "Mach's" principle, where the boundaries are the "distant stars". These are part of the usual

[5] This should not be confused with a possible holographically dual boundary quantum mechanical theory, which would describe the full system.

constraints of general relativity.

One can find explicit expressions for the $SL(2)_g$ charges of the right and left boundaries by using the Noether procedure on (2.13) [17]

$$\begin{aligned} Q_r^{-1} &= \frac{t_r'''}{t_r'^2} - \frac{t_r''^2}{t_r'^3} - t_r' \\ Q_r^{+} &= e^{t_r}\left[\frac{t_r'''}{t_r'^2} - \frac{t_r''^2}{t_r'^3} - \frac{t_r''}{t_r'}\right] \\ Q_r^{-} &= e^{-t_r}\left[-\frac{t_r'''}{t_r'^2} + \frac{t_r''^2}{t_r'^3} - \frac{t_r''}{t_r'}\right]. \end{aligned} \tag{2.16}$$

where $Q^{\pm} = Q^0 \pm Q^1$. The left side charges may also be obtained by analytic continuation $Q_l \to -Q_r$ with $t_l = -t_r + i\pi$, $u_l \to -u_r + i(\text{constant})$. We are defining u_l and t_l so that they go forwards in time in the thermofield double interpretation.

$$\begin{aligned} Q_l^{-1} &= -\frac{t_l'''}{t_l'^2} + \frac{t_l''^2}{t_l'^3} + t_l' \\ Q_l^{+} &= e^{-t_l}\left[\frac{t_l'''}{t_l'^2} - \frac{t_l''^2}{t_l'^3} + \frac{t_l''}{t_l'}\right] \\ Q_l^{-} &= e^{t_l}\left[-\frac{t_l'''}{t_l'^2} + \frac{t_l''^2}{t_l'^3} + \frac{t_l''}{t_l'}\right] \end{aligned} \tag{2.17}$$

One can check that $\frac{1}{2}Q_r.Q_r = \{e^{-t_r}, u\}$.

For our subsequent discussion it is convenient to write a nicer expression for the boundary position so that its $SL(2)_g$ transformations properties are more manifest. The dynamics of the boundary is closely related to the dynamics of a charged massive particle, or a particle with spin, in the limit that the mass and the charge (or spin) becomes both very large, while keeping the total $SL(2)_g$ charges Q^a finite [4, 6]. We have (2.9) with $\mathfrak{q} = 2\phi_b$. In this case, the coordinates Y^a become very large because we approach the boundary. So it is convenient to define rescaled coordinates X^a via

$$X_r^a = \frac{Y_r^a}{Y_r.Q_r} = \frac{Y_r^a}{\mathfrak{q}}\ , \qquad X_r.Q_r = 1\ , \qquad X_r^2 \to 0 \tag{2.18}$$

So, from the point of view of (2.9) we have $r^2 = 0$, $\mathfrak{q}_x = 1$ for the variable X_r^a. For X_l we get $\mathfrak{q}_x = -1$. We can also rescale proper time by the same factor so that now we obey $\dot{X}_r.\dot{X}_r = -1$. In terms of our previous variables these can be written as

$$\begin{aligned} &X_r = \left(X^{-1}, X^+, X^-\right) = \left(\frac{1}{t_r'}, \frac{e^{t_r}}{t_r'}, -\frac{e^{-t_r}}{t_r'}\right), \qquad X^{\pm} \equiv X^0 \pm X^1 \\ &X_l = \left(\frac{1}{t_l'}, -\frac{e^{-t_l}}{t_l'}, \frac{e^{t_l}}{t_l'}\right). \end{aligned} \tag{2.19}$$

We can check that $-X_l.Q_l = X_r.Q_r = 1$. In Appendix 2.8, we verify that in the canonically quantized Schwarzian theory, the above operators satisfy the Poincare algebra (2.9) with the appropriate Casimirs $r^2 = 0$, $\mathfrak{q}_x = \pm 1$. (In appendix 2.10 we give an alternative description in terms of spinors.) In addition, the Hamiltonian corresponding to the Schwarzian action (2.13) is

$$2H_r = -Q_r^2 = \ddot{X}_r^2. \tag{2.20}$$

Using (2.9) this gives us the quantum mechanical relation

$$\dot{X}_r^a = i[H_r, X_r^a] = -\frac{i}{2}[Q_r^2, X_r^a] = \frac{1}{2}\epsilon^a{}_{bc}(X_r^b Q_r^c + Q_r^b X_r^c). \tag{2.21}$$

Taking a second derivative we get the operator equations

$$Q_r^a = \ddot{X}_r^a - H_r X_r^a - X_r^a H_r \ , \qquad X_r.Q_r = 1 \ , \qquad \dot{X}_r^2 = -1, \tag{2.22}$$

$$Q_l^a = -\ddot{X}_l^a + H_l X_l^a + X_l^a H_l \ , \qquad X_l.Q_l = -1 \ , \qquad \dot{X}_l^2 = -1 \ , \tag{2.23}$$

where the difference in signs is due to the difference in sign of $\mathfrak{q}$ for the left boundary. We may also use the algebra to compute commutators between $X, \dot{X}, \ddot{X}$. For example,

$$[\dot{X}_r^a, X_r^b] = -iX_r^a X_r^b \tag{2.24}$$

Using these coordinates it is also possible to write the correlation functions of operators dual to matter fields in the bulk. If we have a massive field in the bulk giving rise to an operator of dimension Δ, then its left right correlator is given by

$$\langle O(u_l)O(u_r)\rangle \propto \frac{1}{(-2X_l(u_l).X_r(u_r))^\Delta} = \left(\frac{t_l'(u_l)t_r'(u_r)}{4\cosh^2(\frac{t_l(u_l)+t_r(u_r)}{2})}\right)^\Delta \tag{2.25}$$

where we used (2.19). We get a similar formula for correlators on the same side.

2.1.3 Review of SYK

The SYK model contains N Majorana fermions with random interactions affecting q fermions at a time, $q = 4, 6, \cdots$ [19–21]. In the large N limit, one can write down an effective action in terms of a bilocal field $G(u_1, u_2)$, which becomes equal to the average two point function once we impose the equations of motion. At low energies this action becomes nearly reparameterization-invariant,

except for a low action reparametrization mode (or soft mode), which has a Schwarzian action with an overall coefficient scaling as N/J, with J an energy scale setting the strength of the interactions of the original model.

In more detail, we start with a scaling solution

$$G_0(t_1, t_2) \propto |t_1 - t_2|^{-2\Delta}. \tag{2.26}$$

The soft mode corresponds to functions G obtained by a reparametrization of (2.26), $G = [f'(u_1)f'(u_2)]^{\Delta}G(f(u_1), f(u$
We can also generate new configurations by having fluctuations $\delta_\perp G(t_1, t_2)$ which lie in the directions orthogonal to the soft mode. For now the coordinates t_1 and t_2 are some coordinates that appear in the solution of the low energy equations of the SYK model and are defined by the form of the unperturbed solution (2.26). We introduce the soft mode by writing the full physical G as [21]

$$G(u_1, u_2) = [f'(u_1)f'(u_2)]^{\Delta}G_0(f(u_1), f(u_2)) + [f'(u_1)f'(u_2)]^{\Delta}\delta_\perp G(f(u_1), f(u_2)) \tag{2.27}$$

This can be viewed as parametrization the full space of functions G. Namely, we think of the integration variables as $f(u)$ and $\delta G_\perp(t_1, t_2)$. Inserting this expression into the SYK action, and taking a low energy limit, we find that the full (Euclidean) action becomes

$$S = S[G_0(t_1, t_2) + \delta_\perp G(t_1, t_2)] - \frac{N\alpha_S}{\mathcal{J}} \int du\{f, u\} \tag{2.28}$$

where we used the approximate reparametrization invariance of the low energy action. This means that the first term in (2.28) is independent of f. All the dependence on f is in the second term of (2.28) and it comes from a small violation of the reparametrization symmetry [21]. To evaluate the path integral, one should sum over different f and $\delta_\perp G$. An important point is that in this parametrization, we have an $\mathrm{SL}(2)_g$ symmetry

$$f \to \frac{af + b}{cf + d}\,, \qquad \delta_\perp G(t_1, t_2) \to \frac{1}{[(a - ct_1)(a - ct_2)]^{2\Delta}} \delta_\perp G\left(\frac{dt_1 - b}{-ct_1 + a}, \frac{dt_2 - b}{-ct_2 + a}\right) \tag{2.29}$$

The arguments of $\delta_\perp G$ are transforming in the inverse way than f so that the second term in (2.27) remains invariant. The first term in (2.27) also remains invariant under this transformation. Therefore (2.29) is a redundancy in our parametrization of the space of $G(u_1, u_2)$ (2.27) and we should demand that everything is invariant. Note that when we write the action as the sum of two terms such as in (2.28) (or three terms if we wrote the Lorentzian action for the thermofield double),

then the $\mathrm{SL}(2)_g$ symmetry will act in a non-trivial way on the variables of each term. In particular, the $\mathrm{SL}(2)_g$ action transforms $\delta_\perp G(t_1, t_2)$, as in (2.29). So, even though the two terms of the action (2.28) are decoupled, they become connected by the total $\mathrm{SL}(2)_g$ constraint.

The conclusion is that in the SYK model we have a structure which is similar to the one we had in nearly AdS_2 gravity. We have three separate systems connected by an overall gauge constraint. The Schwarzian parts are identical to what we had in gravity. But the analog of the matter action $S_m[g_{\mu\nu}, \chi]$ is the first term in (2.28). It is independent of the Schwarzian variables, but its variables transform nontrivially under $SL(2)_g$.

2.2 Exact generators

2.2.1 Construction of gauge invariant $\mathrm{SL}(2,\mathbb{R})$ generators

In NAdS_2 gravity, bulk matter "feels" as if it was moving in empty AdS_2. This suggests that we could define $SL(2)$ generators that move the matter. Naively these would be Q^a_m. However, these are not physical because they are not invariant under the $SL(2)$ gauge symmetry. Said slightly differently, once we go from quantum field theory on a fixed background to quantum gravity, we must gravitationally dress all observables. Since the metric of NAdS_2 is essentially rigid, the dressing should involve the boundary degrees of freedom.

For example, given two boundary positions X^a_l and X^a_r we can define the vector $W^a = \epsilon^{abc} X_{lb} X_{rc}$ and the generator

$$G^0 = \tilde{P} = \frac{\epsilon_{abc} Q^a_m X^b_l X^b_r}{X_l . X_r} \tag{2.30}$$

where we introduced two different notations for the generator. This generator leaves the boundary points X_l and X_r invariant. It is a translation in the bulk along the geodesic that joins these two boundary points, see figure 2.2. In addition, we have normalized it so that it generates translations by a "unit" proper distance in the bulk. In the case that X_l and X_r correspond to the points with $T = 0$ and $\sigma = 0, \pi$ in (2.6a) get we the generator P in (2.5). For general positions for X_l and X_r we get a linear combination of the generators in (2.5). A nice feature of (2.30) is that it is invariant under the $SL(2)$ gauge transformations. Another nice feature of (2.30) is the fact that it acts within the so called "Wheeler-de-Witt" patch, which is the set of points that are spacelike separated from both boundary points, X_l and X_r, see figure 2.2.

We can now wonder whether we can define two other generators in a similar way. Natural

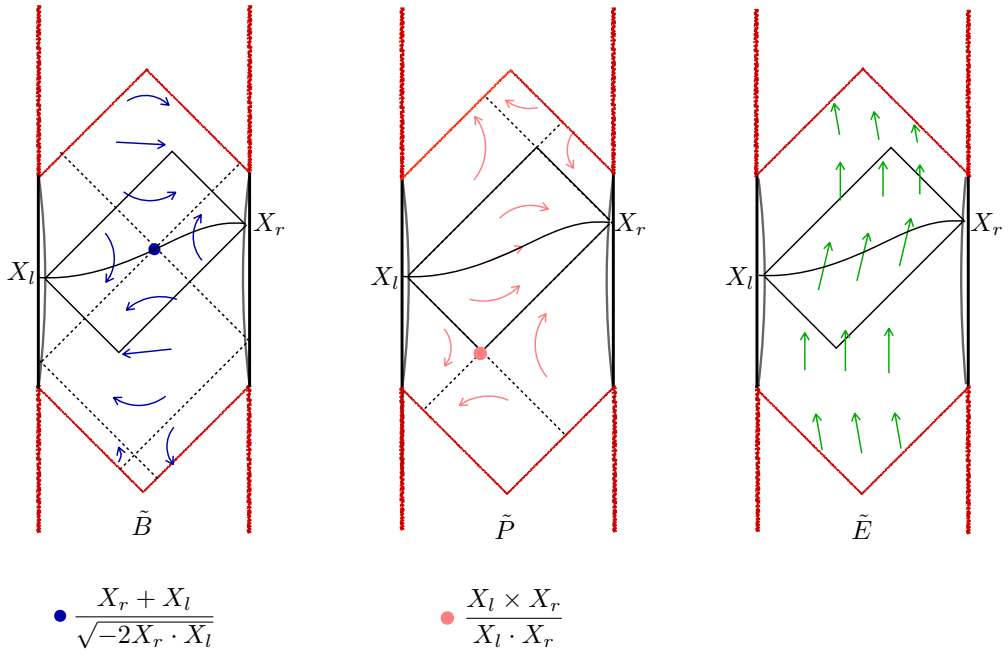


Figure 2.2: Geometrical action of the gauge invariant charges $\tilde{P}, \tilde{B}, \tilde{E}$. The points which are fixed by the symmetry generators are also indicated.

candidates are

$$G^1 + G^{-1} = \tilde{E} + \tilde{B} = -2\frac{Q^a_m X_{la}}{\sqrt{-2X_l.X_r}} , \qquad G^1 - G^{-1} = \tilde{E} - \tilde{B} = 2\frac{Q^a_m X_{ra}}{\sqrt{-2X_l.X_r}} \tag{2.31}$$

These are generators which leave one of the points fixed, (X_l for the first, and X_r for the second). They do not act within the Wheeler de Witt patch, and can map points inside to points outside, see figure 2.2. These generators have been defined so that they obey the same algebra as the generators in (2.5), but are defined relative to the two boundary positions. Notice that they involve a matter operator, Q^a_m and operators of the boundary systems $X^a_{l,r}$. Since they are gauge invariant, they map physical states to other physical states.

We can think of the generators $\tilde{B}$ as defined by the following procedure. Imagine that have have two points X_l and X_r that are very far away, but not yet at the boundary. Then we join them by a geodesic and determine their midpoint. Then $\tilde{B}$ is the boost around this midpoint. Then the third generator, $\tilde{E}$, results from commuting the previous ones and gives a generator that locally looks like a time translations around the midpoint, see figure 2.2. These are time translations locally orthogonal to the geodesic joining X_l and X_r.

Acting on a state with given boundary coordinates X_l and X_r, this state moves the matter around

leaving the boundary points fixed. The resulting time evolution of X_l and X_r can be changed by the action of these generators, but not their instantaneous positions.

We have found the action of a physical $SL(2)$ symmetry on the physical Hilbert space. In particular this means that the physical Hilbert space is infinite dimensional due to the matter degrees of freedom, and their descendants.

In this discussion, we have neglected the possibility of topology changes, such as the ones in [8], since we assumed that the topology is essentially a strip. Therefore we are assuming that ϕ_0 in (2.12) is very large so that topology changes are highly suppressed. It would be interesting to understand how other topologies change the picture; presumably it should be related to cutting off the algebra to a finite dimensional Hilbert space.

An alternative way to describe this same construction is to say that we have defined three vectors e^A_a, where A is an index running over the three vectors, and then we defined three gauge invariant generators

$$G^A = e^A_a Q^a_m. \tag{2.32}$$

The three vectors were the ones in (2.30) (2.31), i.e.,

$$\begin{aligned} e^0_a &= \frac{\epsilon_{abc} X^b_l X^c_r}{X_l.X_r} \ , \qquad e^{-1}_a = -\frac{1}{\sqrt{-2X_l.X_r}}(X_{ra} + X_{la}) \ , \\ e^1_a &= \frac{1}{\sqrt{-2X_l.X_r}}(X_{ra} - X_{la}). \end{aligned} \tag{2.33}$$

The G^A also obey the $SL(2)$ algebra, $[G^A, G^B] = i\epsilon^{ABC} G_C$, due to the properties of e^A_a and the commutation relations of Q^a_m. We can also write the matter Casimir

$$C \equiv G^A G^B \eta_{AB} = Q^a_m Q^b_m \eta_{ab} = \tilde{E}^2 - \tilde{B}^2 - \tilde{P}^2 = E^2_m - B^2_m - P^2_m \tag{2.34}$$

which is $\mathrm{SL}(2)_g$ gauge invariant and commutes with the Hamiltonian.

As a side comment, we may preserve the algebra by rescaling E, B, P by a factor depending on $X_l.X_r$ if we also rescale η by a compensating factor.

Writing the charges purely in terms of boundary quantities

We can use the fact that $Q^a_m = -(Q^a_r + Q^a_l)$, (2.15), together with (2.21), to write

$$G^0 = \tilde{P} \quad = \quad -\frac{\epsilon_{abc} X^b_l X^c_r (Q^a_r + Q^a_l)}{X_l.X_r} = \frac{\dot{X}_l.X_r - X_l \dot{X}_r}{X_l.X_r}$$

$$\tilde{P} = (\partial_{u_l} - \partial_{u_r}) \log[-2X_l.X_r] = (\partial_{u_l} - \partial_{u_r})\ell \tag{2.35}$$

$$\tilde{P} = \frac{\alpha_S N}{\mathcal{J}}(\partial_{u_l} - \partial_{u_r})\ell = \frac{r_e^3}{G_N}(\partial_{u_l} - \partial_{u_r})\ell \tag{2.36}$$

where we noted that $\log[-2X_l.X_r]$ is the regularized distance between the two boundaries, in units of the radius of AdS_2. By "regularized" we mean that we have subtracted an infinite additive constant to the actual proper distance[6]. In (2.36) we have restored the constants that we had set to one for the SYK case or the 4d near extremal charged black hole.

Notice that, due to (2.24), the numerator commutes with the denominator, even though each term in the numerator does not commute with the denominator. In this formula, (2.35), we see that the total momentum is expressed purely in terms of boundary quantities, or gravitational quantities[7]. In addition, this generator can be interpreted as the matter momentum in a frame set by the boundary positions.

Notice that (2.35) is a rather pleasing expression because it can be interpreted as saying that the momentum of matter is minus the momentum of the left plus right boundaries. Namely, if we choose a coordinate x along the geodesic connecting the two boundaries, then the distance is $\ell = x_r - x_l$ and the momentum is

$$\tilde{P} = (\partial_{u_l} - \partial_{u_r}) \left[x_r(u_r) - x_l(u_l)\right] = -(\dot{x}_r + \dot{x}_l). \tag{2.37}$$

This is saying that the matter momentum is minus the sum of the momenta of the boundary particles. We get the naive expression for the momentum of the boundary particles because the term involving $\mathfrak{q}$ in (2.8) drops out when we contract $\vec{Y}_l \times \vec{Y}_r$ with $\vec{Q}$. So we get the same result as for an ordinary massive particle.

Note that in writing (2.35) we assumed a particular form for the Hamiltonian that generates the u dependence. In particular, we have assumed that we have a decoupled evolution, by H_l and H_r in (2.20). In contrast, the expressions (2.30) (2.31) did not use the form of the Hamiltonian and are valid more generally (for example we could have a small coupling between the left and right sides). We can obtain expressions that are more generally valid by writing Q_l^a and Q_r^a in terms of the boundary positions and their conjugate momenta, see appendix 2.8 and (2.132).

We can also consider the expressions for the other generators. Again, we start from (2.31) and we express the matter charges in terms of the left and right charges, and use (2.22) (2.23) to obtain

[6] This infinite constant is independent of time and independent of the X_l or X_r variables.

[7] Note that we are talking about the boundary gravitational degrees of freedom, and not the holographically dual boundary quantum mechanical theory.

(ignoring operator ordering issues)

$$
\begin{aligned}
G^1 = \tilde{E} &= (-2X_l.X_r)^{1/2}\left(H_l + H_r + \frac{1}{X_l.X_r} - \frac{\ddot{X}_l.X_r + X_l.\ddot{X}_r}{2X_l.X_r}\right), \\
G^{-1} = \tilde{B} &= -(-2X_l.X_r)^{1/2}\left(H_l - H_r + \frac{X_l.\ddot{X}_r - \ddot{X}_l.X_r}{2X_l.X_r}\right).
\end{aligned}
\tag{2.38}
$$

We can also express the Casimir (2.34) in terms of purely boundary quantities. Of course, these expressions depend on *both* boundaries.

These observables can be expressed in terms of energies and distances between left and right sides. We will discuss how to measure distances in Section 2.6.1.

2.2.2 Defining "conserved" charges

The generators we have defined above act on the physical Hilbert space but they do not commute with the Hamiltonians of the system, H_l or H_r. Therefore we cannot call them conserved quantities. (Of course, the Casimir (2.34) is indeed conserved.) However, one feature of the gauge-non-invariant matter charges is that they are conserved $[H_{l,r}, Q^a_m] = 0$ in the unphysical Hilbert space.

The charges we defined depend on the left and right times through the boundary positions $X_l(u_l)$ and $X_r(u_r)$. Then the charges in (2.30) (2.31) depend on the two times $G^A(u_l, u_r)$. However, the dynamics of the left and right boundaries is solvable as a quantum mechanical theory. This means that the change in the charges follows a reasonably predictive pattern. In particular, we would obtain time independent expressions for the generators by solving the boundary dynamics so that we can work with $X_l(0)$ and $X_r(0)$. Therefore we can simply say that the "conserved" charges are simply $G^A(0_l, 0_r)$ where we have set both times to zero. Now, this looks like we are cheating since we can always define a conserved quantity by undoing the time evolution. However, in this case, the statement has non-trivial content because we *only* have to undo the evolution of the boundary mode, the Schwarzian degree of freedom. In particular, we are not undoing the evolution of matter, which could be a complicated self interacting theory. In addition, in the classical limit, we can undo the classical evolution of the boundary theory in a simple way. In principle, we can also express $G^a(0,0)$ in terms of the correlators at zero as in (2.35) (2.38).

Formally we can write down

$$
G^A(0,0) = e^{-i(H_l u_l + H_r u_r)} G^A(u_l, u_r) e^{i(H_l u_l + H_r u_r)} = \Lambda^A{}_B G^B(u_l, u_r) \tag{2.39}
$$

with

$$\Lambda^A{}_B = e^A_a(0,0)\left(e^{-1}(u_l,u_r)\right)^a{}_B = e^A_a(0,0)\eta^{ab}e^C_b(u_l,u_r)\eta_{CB}. \tag{2.40}$$

This expression for Λ involves the quantum operators $X_{l,r}$ evaluated at zero and also evaluated at u_l, u_r, so it is a rather complex expression in the quantum Schwarzian theory. Note that the operator $X_r(0)$ can be expressed explicitly in terms of operators at time $u_l,\ u_r$ by using the propagators for the Schwarzian theory [4, 6]. Unfortunately, the operators are not diagonal in the $X_r(u_r)$ basis, so it is hard to express them in terms of correlators at time u_r, and we will not attempt to do it here.

Let us mention that even the standard expressions for the matter charges (2.11) involve some explicit time dependent expressions, since (some of) the Killing vectors depend explicitly on time. In (2.11), this dependence is very simple. In our problem the time dependence is a bit more complicated, but in principle solvable.

One case where the dynamics can be solved simply is the classical limit, as we will see in (2.48).

2.3 Approximate expressions for the generators

2.3.1 The generators in the semiclassical limit

It is instructive to consider the above construction in the semiclassical limit. So we start with a two sided black hole solution with $\beta \ll 1$,

$$t_r = \tilde{u}\ , \qquad t_l = \tilde{u}\ , \qquad \tilde{u} \equiv \mathfrak{s}u\ , \qquad \mathfrak{s} \equiv \frac{2\pi}{\beta} = \frac{2\pi\alpha_S N}{\beta\mathcal{J}} = \frac{2\pi r_e^3}{\beta G_N} \tag{2.41}$$

where we have defined the coefficient of the Schwarzian action $\mathfrak{s}$, which has the interpretation of the $SL(2)$ spin of the state in the Schwarzian theory ($j = \frac{1}{2} + i\mathfrak{s}$). It is also related to the near extremal entropy, $S - S_0 = 2\pi\mathfrak{s}$. We have also restored the constants we had set to one for the case of SYK or 4d near extremal charged black holes.

It is common in these discussions to keep two parameters N and $\mathcal{J}$ as independent, and the fact that we have removed them completely might confuse some readers. Indeed they are independent parameters in a model such as SYK. However, all of our discussion is centering on the low energy regime and the coupling of the Schwarzian theory to the conformal sector. These parameters appear only in the overall coefficient of the Schwarzian action. We have rescaled the units of time u, so as to set this constant to one, see (2.13). This highlights the fact that there is only an overall lengthscale appearing the problem, and we have chosen units where this lengthscale is set to one. This is the

timescale at which the Schwarzian theory becomes strongly coupled. In these units the limit $\beta \ll 1$ corresponds to the semiclassical limit of the Schwarzian theory. We can restore the full dependence on $N/\mathcal{J}$ by restoring such constants by using dimensional analysis. For example, thinking of $\mathfrak{s}$ as the entropy we restore them as in (2.41). If we had β appearing in a dimensionless quantity, such as $e^{2\pi u/\beta}$, then no further change is needed, since the rescaling of u and β cancel out. The semiclassical limit corresponds to $\mathfrak{s} \gg 1$.

Inserting (2.41) into the right-left charges (2.16) (2.17) we find that $Q_r^a + Q_l^a = 0$ as expected. The only non-zero components of these charges are $Q_l^{-1} = -Q_r^{-1} = \mathfrak{s}$. We now add a relatively small amount of bulk matter Q_m. By small we mean that the changes in the boundary trajectories are small,

$$t_r = \tilde{u} + \epsilon_r(\tilde{u}) \ , \qquad t_l = \tilde{u} + \epsilon_l(\tilde{u}) \tag{2.42}$$

with $\epsilon_{r,l} \ll 1$. In this case we can expand the charges Q_r^a, Q_l^a in ϵ. In fact, it is convenient to expand the sum of the charges because this sum is then equated to $Q_m^a = -(Q_l^a + Q_r^a)$.

This gives

$$\begin{aligned}
Q_m^{-1} &\simeq \mathfrak{s}\left[\epsilon_r' - \epsilon_r''' - \epsilon_l' + \epsilon_l'''\right] \ , \qquad \text{where} \quad {}' \equiv \partial_{\tilde{u}} \\
Q_m^0 + Q_m^1 = Q_m^+ &\simeq \mathfrak{s}\left[e^{\tilde{u}_r}(\epsilon_r'' - \epsilon_r''') + e^{-\tilde{u}_l}(-\epsilon_l'' - \epsilon_l''')\right] = \mathfrak{s}\left[\epsilon_r''(0) - \epsilon_r'''(0) - \epsilon_l''(0) - \epsilon_l'''(0)\right] \\
Q_m^0 - Q_m^1 = Q_m^- &\simeq \mathfrak{s}\left[e^{-\tilde{u}_r}(\epsilon_r'' + \epsilon_r''') + e^{\tilde{u}_l}(-\epsilon_l'' + \epsilon_l''')\right] = \mathfrak{s}\left[\epsilon_r''(0) + \epsilon_r'''(0) - \epsilon_l''(0) + \epsilon_l'''(0)\right]
\end{aligned} \tag{2.43}$$

where the primes on $\epsilon_l(\tilde{u}_l)$ and $\epsilon_r(\tilde{u}_r)$ are derivatives with respect to $\tilde{u}_l$, $\tilde{u}_r$ respectively. These expressions are naively u dependent, but the equations of motion for ϵ make sure that they are u independent, and we have given the expressions for zero times. Namely, from the conservation of energy,

$$H_r = \frac{\mathfrak{s}^2}{2} + \mathfrak{s}^2(\epsilon_r' - \epsilon_r''') + \cdots \ , \qquad H_l = \frac{\mathfrak{s}^2}{2} + \mathfrak{s}^2(\epsilon_l' - \epsilon_l''') + \cdots \tag{2.44}$$

we get

$$\epsilon'''' - \epsilon'' = 0 \ , \qquad \epsilon = \epsilon_{r,l}(\tilde{u}_{l,r}) \tag{2.45}$$

which ensures that the right hand sides of (2.43) are all independent of time (as are the left hand sides).

Inserting (2.41) into (2.19) and then computing the generators G^A to zeroth order in ϵ we find

$$G^A(0,0) \simeq Q_m^a \ , \qquad A = a. \tag{2.46}$$

This equation is valid in the gauge we used to write the background solution (2.41). In general we can also compute the generators G^A at more general times. Using the classical evolution to evolve the vectors e_a^A we can express them as a linear combination of the ones in (2.46), see (2.40),

$$G^A(0,0) \simeq \Lambda^A_{\ B} G^B(u_l, u_r) \ , \tag{2.47}$$

with

$$\begin{aligned} \Lambda \ &= \ \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh\gamma & \sinh\gamma \\ 0 & \sinh\gamma & \cosh\gamma \end{pmatrix} \cdot \begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \ , \\ &\text{where} \qquad \sin\alpha = \tanh\frac{\tilde{u}_l + \tilde{u}_r}{2} \ , \qquad \gamma = \frac{\tilde{u}_l - \tilde{u}_r}{2}. \end{aligned} \tag{2.48}$$

This is reflecting the fact that when we pick arbitrary left and right times, in the classical limit, the generators G^A have been rotated relative to the ones at zero times, see figure 2.2. In this case, the time dependence of the generators is simple and can be extracted to define the time independent generators $G^A(0,0)$. This is the classical version of the formula (2.40).

It is also interesting to note that we can also obtain (2.43) by evaluating $G^A(u_l, u_r)$ using correlators, as in (2.35) and (2.38). In detail, what we have in mind is the following. Let us consider (2.35), for example. We express $X_l.X_r$ in terms of the boundary times as in (2.25). We then expand the times as in (2.42), to obtain

$$\tilde{P}(u_l, u_r) = (\partial_{u_l} - \partial_{u_r}) \log[-2X_l.X_r] = \mathfrak{s}\left[\epsilon_r'' - \epsilon_l'' + \frac{1}{2}(\epsilon_l' - \epsilon_r') \tanh\frac{\tilde{u}_l + \tilde{u}_r}{2}\right]. \tag{2.49}$$

In a similar way we get can get the other generators. Then applying the inverse of the matrix in (2.48) we get the generators $G^A(0,0)$ which are the ones in (2.46), with the expression (2.43).

2.3.2 The semiclassical limit and $\mathrm{SL}(2,\mathbb{R})$ symmetries of the physical boundary time

In this semiclassical limit, we can think of the $G^A(0,0)$ as generating a symmetry that acts as ordinary reparametrizations of $\tilde{u}$, generated by the infinitesimal transformations

$$\mathrm{SL}(2)_u : \qquad \tilde{u} \to \tilde{u} + \alpha_{-1} + \alpha_+ e^{\tilde{u}} + \alpha_- e^{-\tilde{u}} \ , \qquad \tilde{u} = \mathfrak{s} u \ , \qquad \mathfrak{s} = \frac{2\pi}{\beta} \tag{2.50}$$

In order to make this manifest we will analyze how the charges $G^A(0,0)$ act on states created by the insertion of operators in Euclidean time. In fact, we will discuss a state/operator map for the nearly-CFT_1 that is dual to this gravity theory.

More precisely, we view the state at $u_l = u_r = 0$ as created by Euclidean time evolution over a time $\Delta u_e = \beta/2$ (or $\delta\tilde{u}_e = \pi$). This Euclidean evolution generates the empty wormhole. We can then create excitations by acting by operators during the euclidean evolution period. To simplify the notation we will denote by φ the rescaled Euclidean time $\varphi \equiv \frac{2\pi}{\beta} u_e$, so that $\varphi \sim \varphi + 2\pi$.

We can now consider a general two point function between an operator inserted on the top half of the circle and one on the bottom half

$$\langle O(\varphi_t) O(\varphi_b)\rangle = \langle O(\varphi_t)|O(\varphi_b)\rangle \propto \frac{\mathfrak{s}^{2\Delta}}{\left[\sin\frac{\varphi_t-\varphi_b}{2}\right]^{2\Delta}} \tag{2.51}$$

We can view this as the overlap of two states. One is a state that is obtained by doing the path integral over the bottom half and the other is the one obtained by doing the path integral over the top half. This defines an operator/state map. See figure 2.3.

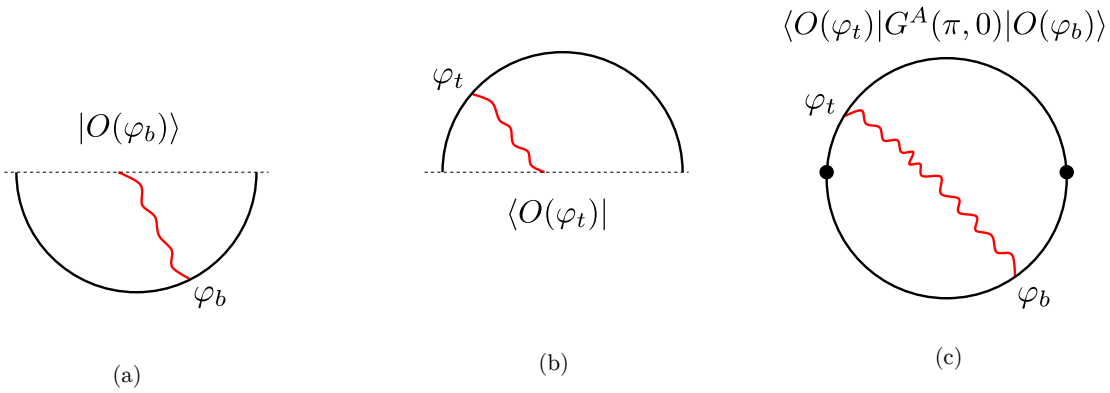


Figure 2.3: (a) By performing euclidean evolution over time $\beta/2$ and inserting an operator at some point during the euclidean evolution we create a state. This defines a map between operators and states. These are states of a wormhole or states living in the Hilbert space of two copies of the dual boundary quantum system. (b) The same for the bra. (c) We can take the inner product and add the action of a charge $G^A(\pi,0)$, represented by the black dots. In the semiclassical regime, these charges act as $SL(2)$ generators on the states or the operators.

We will now act with the charges $G^A(0,0)$, or in this case $G^A(\pi,0)$ and will demonstrate that they act as expected on the states created by these operator insertions, in other words

$$\langle O(\varphi_t)|G^A(\pi,0)|O(\varphi_b)\rangle = \left[\zeta^A(\varphi_b)\partial_{\varphi_b} + \Delta(\partial_{\varphi_b}\zeta^A(\varphi_b))\right]\langle O(\varphi_t)|O(\varphi_b)\rangle \tag{2.52}$$

where ζ^A is a linear combination of the vectors generating the infinitesimal $SL(2)$ reparametrizations (2.50), see (2.53). This is physically saying that the action of $G^A(\pi, 0)$ is acting with an infinitesimal reparametrization on the bottom part, see figure 2.3(c). We can equally view it as acting with (minus) the reparametrization on the top part, since acting with the reparametrization both on the top and bottom leaves the correlator (2.51) invariant.

We will demonstrate (2.52) as follows. First we write down the three generators and their associated vectors.

$$\begin{aligned} \tilde{B} &\simeq \mathfrak{s}\left[\epsilon'(0) + \epsilon'''(0) - \epsilon'(\pi) - \epsilon'''(\pi)\right] , & \zeta^{\tilde{B}} &= 1 \\ \tilde{P} &\simeq \mathfrak{s}\left[\epsilon''(0) + \epsilon''(\pi)\right] , & \zeta^{\tilde{P}} &= -\sin\varphi \\ \tilde{E} &\simeq \mathfrak{s}\left[\epsilon'''(0) + \epsilon'''(\pi)\right] , & \zeta^{\tilde{E}} &= \cos\varphi \end{aligned} \tag{2.53}$$

These are the expressions appropriate for Euclidean time[8]. We can picture the geometric action of these generators as in figure 2.4.

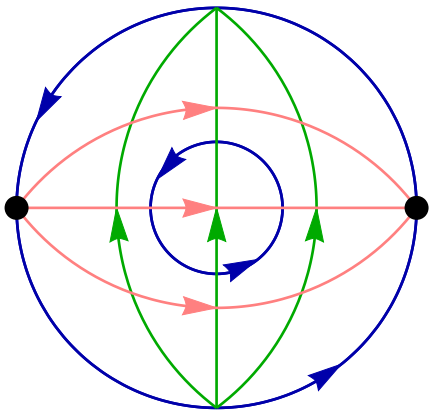


Figure 2.4: Geometric action of the generators in Euclidean AdS_2 . Here the charges are inserted at the black points, at $\varphi = 0$ and $\varphi = \pi$. Blue lines follow the boost Killing vectors $\tilde{B}$; pink lines, the momentum $\tilde{P}$; and green lines, the global energy vectors $\tilde{E}$.

Notice that in this classical limit $\tilde{B}$ is proportional to $H_r - H_l$, see (2.44) (2.43) and it generates shifts in φ. Now, in order to evaluate the left hand side of (2.52) we will use the first order expression for G^A in (2.53). We then also expand the correlators to first order in ϵ. In other words, we write

[8] Relative to (2.43) we have flipped the signs of ϵ_l and of u_l. This arises due to a different definition of the left time. In addition, when we go from Lorentzian to Euclidean time we need to say that $\epsilon \to i\epsilon$, $u \to iu$. We also removed an extra i in $\tilde{P}$.

the correlators as in (2.25) and expand the times as in (2.42) to obtain

$$\langle O(\varphi_t)O(\varphi_b)\rangle \to \langle O(\varphi_t)O(\varphi_b)\rangle \left\{1+\Delta\left[\epsilon_t'+\epsilon_b'-\frac{(\epsilon_t-\epsilon_b)}{\tan\frac{\varphi_t-\varphi_b}{2}}\right]\right\} \tag{2.54}$$

where the subindices of ϵ indicate where they are evaluated, $\epsilon_t=\epsilon(\varphi_t)$, etc. Then the computation of (2.52) boils down to a computation in the linearized Schwarzian theory with action

$$S_E=\mathfrak{s}\int d\varphi\frac{1}{2}(\epsilon''^2-\epsilon'^2)+\cdots \tag{2.55}$$

We see that the classical limit is indeed large $\mathfrak{s}$. The propagator associated to this action is [17]

$$\langle\epsilon(\varphi)\epsilon(0)\rangle=\frac{1}{\mathfrak{s}}\left[G(|\varphi|)+a+b\cos\varphi\right]\ ,\qquad G(\varphi)\equiv-\frac{(\varphi-\pi)^2}{4\pi}+\frac{(\varphi-\pi)}{2\pi}\sin\varphi \tag{2.56}$$

where a and b are constants that drop out when we compute $SL(2)$ gauge invariant quantities, such as the ones we are computing. So we can set them to zero. For example, to compute an insertion of $\tilde{P}$ we need to compute the correlator

$$\frac{\langle O(\varphi_t)|\tilde{P}|O(\varphi_b)\rangle}{\langle O(\varphi_t)|O(\varphi_b)\rangle}=\Delta\mathfrak{s}\left\langle\left[\epsilon_t'+\epsilon_b'-\frac{(\epsilon_t-\epsilon_b)}{\tan\frac{\varphi_t-\varphi_b}{2}}\right][\epsilon''(0)+\epsilon''(\pi)]\right\rangle \tag{2.57}$$

Using the propagator (2.56) we find that this is equal to the expression we need to generate the right hand side of (2.52). In other words, it is

$$\frac{\langle O(\varphi_t)|\tilde{P}|O(\varphi_b)\rangle}{\langle O(\varphi_t)|O(\varphi_b)\rangle}=\Delta\left[\frac{1}{\tan\frac{\varphi_t-\varphi_b}{2}}\zeta^{\tilde{P}}(\varphi_b)+\left(\zeta^{\tilde{P}}(\varphi_b)\right)'\right]\ ,\qquad \zeta^{\tilde{P}}=-\sin\varphi \tag{2.58}$$

with $\zeta^{\bar{P}}=-\sin\varphi$, as in (2.53). The diagrams we need to compute can be seen in figure 2.5(a). For $\tilde{E}$ and $\tilde{B}$ we also get results consistent with (2.52), (2.53).

In computing the matrix elements of G^A, we ignored 1-loop corrections to the 2-pt function. This is justified because such corrections actually cancel, since the zero-th order term in G^A is actually zero. We also ignored the 1-loop correction to G^A itself. If we use the exact charges, this too must vanish because G^A exactly annihilates a state without matter. However, if we used approximate expressions (as we will discuss in Section 2.3.3) for G^A, one should in principle subtract off these contributions order by order in perturbation theory, see equation (2.69).

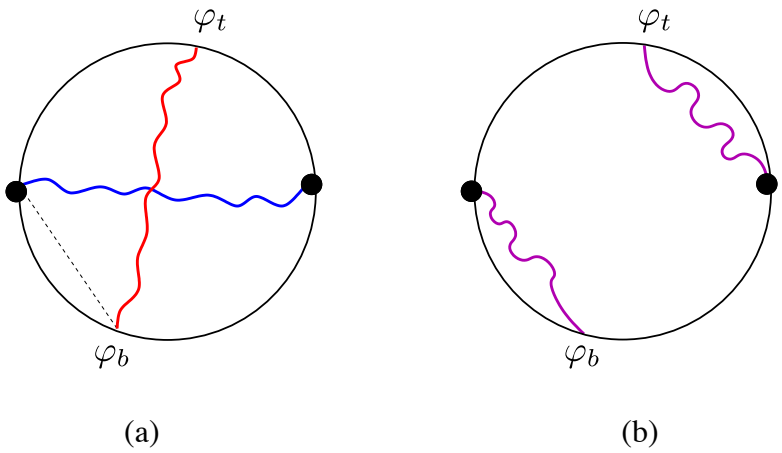


Figure 2.5: (a) We consider the action of the charges. We have matter fields propagating from the bottom to the top indicated in red. These cause some backreaction on the boundary positions. These are summarized by the coupling to ϵ at the insertion points of operators. The definition of the charges involves computing a distance, which implicitly, or more explicitly (in the generators $\hat{G}^A$ in (2.63), (2.69),(2.71)), involve the propagation of other matter fields. The interesting terms come from correlators between these ϵ insertions. We only have two point functions of ϵ, only one of which is indicated in the diagram by a doted line. Other diagrams contain a dotted line between black points and $\varphi_{t,b}$. (b) In some specific models we might get contractions between the fields in the definition of the charges and the insertions. We want to suppress this type of diagrams. They are indeed suppressed relative to those in (a) in the SYK model.

As a more specific example we can consider the expectation values of all three generators on a state created by inserting the operator in Euclidean time at $\varphi_b < 0$, see figure (2.6):

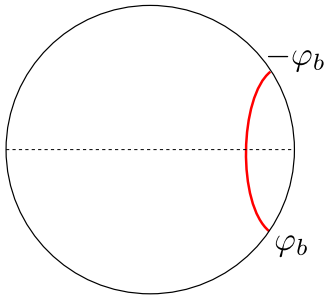


Figure 2.6: Inserting an operator in Euclidean time creates a particle at rest.

We simply evaluate expressions like (2.58) setting $\varphi_t = -\varphi_b$ and we obtain

$$\langle \tilde{P} \rangle = 0 \ , \qquad \langle \tilde{B} \rangle = \frac{\Delta}{-\tan \varphi_b} \ , \qquad \langle \tilde{E} \rangle = \frac{\Delta}{-\sin \varphi_b} \tag{2.59}$$

We should think of a state which contains a particle at rest on the initial slice. At small $|\varphi_b|$, the particle is located at a propert distance of the order $-\log(-\varphi_b)$ from the horizon, and the redshift difference between the horizon and its position if of order $1/(-\varphi_b)$. See figure 2.6.

The conclusion of this discussion is that around these classical states, the exact generators $G^A(0,0)$ are acting as $SL(2)$ generators transforming the boundary time.

The primary states and their descendants defined by the state-operator correspondence are eigenstates of $\tilde{E}$ in the semiclassical approximation, but this is not expected to be exact in the Schwarzian theory. Presumably the exact eigenstates of $\tilde{E}$ could be obtained by smearing the primary in some suitable fashion.

In previous sections we have seen that G^A maps physical states to physical states. We have seen here that this map changes states as we expect from $SL(2)_u$ symmetries of the boundary time u (2.50). In other words, the generators G^A that are always well defined, become the $SL(2)_u$ generators of a NCFT_1 in this limit. This correspondence is not expected to hold away from the semiclassical limit. In fact, the boundary dynamics is not invariant under $SL(2)_u$. But this is a good approximate symmetry in this classical limit. Note that the semiclassical limit is really hardwired in our description of the symmetry itself, since the action of the approximate symmetry depends explicitly on β (2.50). This represents a state dependence of the symmetry action. And it is reflected in the dependence of the generators on β (2.53) (and will be more explicitly seen below). Furthermore, in section 2.3.2, we will see that, as we insert matter at early lorentzian times, this semiclassical picture also breaks down.

Finally, we would like to caution the reader that this physical $SL(2)_u$ should not be confused with the exact gauge symmetry $SL(2)_g$ that has previously been discussed (and which we review in Section 2.3). The $SL(2)_u$ generators act on physical states $|O(u)\rangle$ and give new physical states, e.g., $\tilde{B}\,|O(\phi_b)\rangle \approx |\partial O(\phi_b)\rangle$.

Inserting matter at early lorentzian times

In the previous section we have discussed that inserting operators in euclidean time gives us states at $u_l = u_r = 0$ that contain bulk excitations, and we explained how to read off the $SL(2)$ charges of these states in the semiclassical limit.

Of course, these formulas also work in Lorentzian signature. More specifically, imagine that we start with the thermofield double state at early times, say $u_l = 0,\ u_r \ll 0$, which we can obtain by evolving the TFD state backwards in time on one of the sides. We then insert an operator at time $u_{r0} < 0$, and evolve up to $u_r = 0$. See figure 2.7(a). We will need to slightly smear it in order to create a relatively low energy state that can be described within the conformal regime. This will also create a matter state inside the wormhole. We can find the $SL(2)$ transformation properties of this state by acting with the charges, and we will obtain the expected action, as indicated in (2.52). It is interesting that now some of the "conformal Killing vectors" have an exponential depedence on time,

$$\zeta^{\tilde{B}} = 1 \ , \qquad \zeta^{\tilde{P}} = \sinh \frac{2\pi u}{\beta} \ , \qquad \zeta^{\tilde{E}} = \cosh \frac{2\pi u}{\beta} \tag{2.60}$$

This implies that if we insert the same operator O, earlier and earlier in time, we will get exponentially growing values for its energy and its momentum, from (2.52). At least this is true as long as these semiclassical expressions hold. It turns out that for early times, times larger than the scrambling time, $u_{\rm scr} = \frac{\beta}{2\pi} \log \mathfrak{s}$, it becomes important to take into account the backreaction beyond the leading order in ϵ.

Inserting matter beyond the scrambling time and corrections to the semiclassical limit

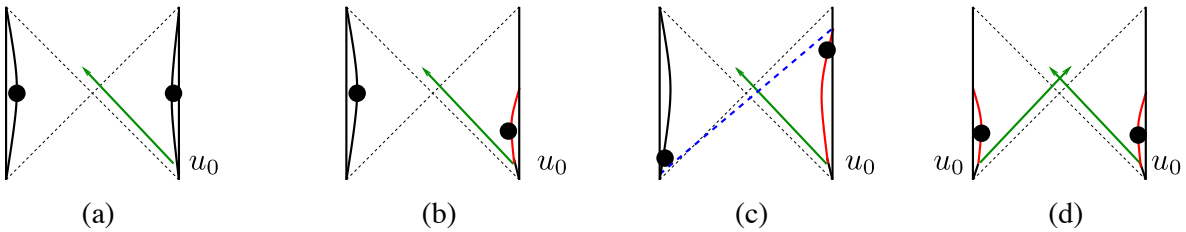


Figure 2.7: (a) We insert matter at some early Lorentzian time $u_0 < 0$ smaller than the scrambling time so that the backreaction on the boundary trajectory is small. (b) Same insertion but beyond the scrambling time. There is a large change in the boundary trajectory. (c) The same as (b) but after an overall boost (a gauge transformation) that explains better why there is a maximum momentum. Here we have kept the point u_0 fixed in the figure so that increasing it corresponds to moving the left and right dots (the times where we evaluate the charges) down and up. The point is that the boost angle between the green line (the matter we inserted) and the blue dotted line is finite. The blue dotted line represents that highest boost angle for a geodesic joining the two black dots at late times. (d) We insert matter on both sides. In this case, one can show that the energy continues to increase as we take u_0 beyond the scrambling time.

We have seen in the previous subsection that if we insert some mode of energy ω at some early time u_0, then its charges evaluated at $u_l = u_r = 0$ grow exponentially as $\beta\omega e^{-\tilde{u}_0}$ ($\tilde{u}_0 = \mathfrak{s} u_0$ is

negative). Then, even if $\mathfrak{s} \gg 1$, there can be a time when the simple small ϵ approximation breaks down. The expansion parameter is really $\beta\omega e^{-\tilde{u}_0}/\mathfrak{s}$, and the small ϵ approximation breaks down when this is of order one. This is the so called scrambling time [36]. The picture is that, by this time, the excitation has an order one commutator with any other simple excitation. Now, our basic expressions for the generators are exact and can be evaluated beyond the scrambling time. In this section we sketch the results for the exact generators (2.30) (2.31) when we go beyond the scrambling time. We will work in the large $\mathfrak{s}$ approximation, and for simplicity we will further assume that $\omega/\mathfrak{s} \ll 1$ and $|u_0| \ll 1$ but we will work exactly in

$$\alpha \equiv \frac{\beta\omega}{\mathfrak{s}} e^{-\tilde{u}_0} \ , \qquad \tilde{u}_0 = \mathfrak{s} u_0 = \frac{2\pi u_0}{\beta} \tag{2.61}$$

In this regime, we can find the correction to the classical trajectory and compute the generators, see figure 2.7(b). We find that the generators are equal to, see appendix 2.10.1,

$$\tilde{P} \sim \tilde{E} \sim 2\mathfrak{s}\frac{\hat{\alpha}}{1+\hat{\alpha}} \ , \qquad \tilde{B} \sim 0 \ , \qquad \hat{\alpha} \propto \frac{\omega\beta}{\mathfrak{s}} e^{-\tilde{u}_0} \tag{2.62}$$

where $\hat{\alpha}$ is equal to α in (2.61), up to a numerical constant. It is also worth noting that $X_l.X_r \propto (1+\hat{\alpha})^2$. This implies that the physical distance, which is the logarithm of this quantity increases linearly with u_0, as $\ell \sim 2|\tilde{u}_0| \sim 4\pi|u_0/\beta|$. The standard semiclassical expression discussed in section 2.3.2 amounts to expanding (2.62) to first order in $\hat{\alpha}$. Interestingly we find that the generators saturate at an amount of order $\mathfrak{s}$ which is independent of the energy of the particle we have sent in. This might seem surprising, since it naively looks like we are inserting a higher and higher energy state as we take $u_0 \to -\infty$. However, this insertion is moving the dynamic right boundary and is changing the notion of momentum. Notice that the bulk Casimir is zero in this limit, see figure 2.7(b,c). The saturation of (2.62) will be related to the decay of out of time order correlators in section 2.3.4.

We could consider a different experiment where we send matter from both sides, see figure 2.7(c). In this case, $\tilde{P} \sim \tilde{B} \sim 0$, but $\tilde{E}$ continues to increase exponentially.

This computation illustrates how the exact generators (2.30) and (2.31) can be defined and used, beyond the scrambling time.

2.3.3 Other semiclassical expressions for the generators

We have seen that we can get approximate expressions for the $\mathrm{SL}(2,\mathbb{R})$ generators. These approximate expressions relied purely on the small ϵ expansion of the boundary trajectories around a given thermofield double state. Here we want to relate these expressions to correlators in the boundary theory. Of course, we have already given an expression of the exact generators in terms of the distances that are probed by boundary correlators, (2.35) (2.38). Here we want to provide simple expressions that give the same answer in the semiclassical limit.

We have already mentioned one of them. Namely, the boost generator can be approximately given in terms of the difference of Hamiltonians

$$\tilde{B} \simeq \hat{B} = \frac{\beta}{2\pi}(H_r - H_l) \tag{2.63}$$

This is also the modular Hamiltonian that arises when we split the system into left and right sides. Note that $\hat{B}$ is an exact symmetry of the thermofield double state.

The chapter [27] discussed a coupled system whose Hamiltonian could be viewed as the global time translation, $\tilde{E}$, in AdS_2. This Hamiltonian was defined as

$$H_{\text{coupled}} = H_r + H_l - \tilde{\eta}\sum_j O_l^j O_r^j \tag{2.64}$$

$$\sim H_r + H_l - \eta\left(\frac{t_l' t_r'}{\cosh^2\frac{t_l+t_r}{2}}\right)^{\Delta}, \qquad \eta = \tilde{\eta} N 2^{-2\Delta} \tag{2.65}$$

where we have indicated the approximate expression in the Schwarzian theory in the approximation that the effect of the boundary coupling on the bulk matter is very small. We normalized the operators so that they go like $\langle O_r(u_1)O_r(r_2)\rangle \sim |u_{12}|^{-2\Delta}$ at short distances. Then the main effect of the coupling is on the Schwarzian variables [27]. We expand around a solution of the form

$$t_l = t_r = \mathfrak{s} u \tag{2.66}$$

The solution that minimizes the energy (and obeys all necessary equations of the two Schwarzian theories) is such that

$$\mathfrak{s}^{2-2\Delta} = \Delta\eta \ , \tag{2.67}$$

Since the semiclassical limit involves $\mathfrak{s} \gg 1$, we need that $\eta \gg 1$. This can be achieved by having a large number of operators in (2.64). In other words, we take small $\tilde{\eta}$, but large N, so that η in (2.64)

is large. In the construction of [27] this equation, (2.67), was viewed as determining $\mathfrak{s}$, or β, in terms of η. The ground state of the system is close to the thermofield double at inverse temperature

$$\beta = \frac{2\pi}{\mathfrak{s}} \tag{2.68}$$

This is not the physical temperature of the coupled system, it is rather the effective temperature of the density matrix of each side on its own. Finally, the normalized global time translation symmetry is then

$$\tilde{E} \simeq \hat{E} \equiv \frac{1}{\mathfrak{s}} \left[H_{\text{coupled}} - \langle H_{\text{coupled}} \rangle_0 \right] \tag{2.69}$$

$$\hat{E} \simeq -\mathfrak{s}(\epsilon_l''' + \epsilon_r''') \tag{2.70}$$

where the last expression agrees with (2.43), as expected. Here $\langle \cdots \rangle_0$ indicates the expectation value in the ground state of the coupled system.

For the purposes of this chapter, we can simply view the TFD state at a given inverse temperature, β, as given. And we then write (2.69), solving for η in terms of $\mathfrak{s} = \frac{2\pi}{\beta}$ via (2.67), and construct $\hat{E}$ as in (2.69). The advantage of this procedure is that it gives an approximate expression for $\tilde{E}$ that is relatively simple, we only need to couple the two sides.

Finally, we can get a simple expression for $\tilde{P}$ by taking the commutator of (2.63) and (2.69), to obtain

$$\tilde{P} \simeq \hat{P} \equiv -i[\hat{B}, \hat{E}] = \frac{-i}{\mathfrak{s}} \left[H_r - H_l, -\tilde{\eta} \sum_j O_l^j O_r^j \right] = \frac{\tilde{\eta}}{\mathfrak{s}^2} \sum_j (O_l^j \dot{O}_r^j - \dot{O}_l^j O_r^j) \tag{2.71}$$

$$\simeq \frac{1}{\mathfrak{s}^2} (\partial_{u_r} - \partial_{u_l}) \eta \left(\frac{\dot{t}_l \dot{t}_r}{\cosh^2 \frac{t_l + t_r}{2}} \right)^{\Delta}$$

$$\hat{P} \simeq \mathfrak{s}(\epsilon_r'' - \epsilon_l'') \tag{2.72}$$

where we have used (2.67).

Notice that the generators $\hat{B}$, $\hat{E}$, $\hat{P}$ are completely well defined if the system has a quantum mechanical dual. For example, they are well defined in the SYK model. However they do not obey an exact $\text{SL}(2, \mathbb{R})$ algebra. In addition, their definition depends on β (via $\mathfrak{s}$). This means that they behave as $\text{SL}(2, \mathbb{R})$ generators only for states close to the thermofield double state with that inverse temperature. The fact that they obey the right algebra for such states comes from their connection to the matter charges in (2.43). Notice that the thermofield double state, or empty wormhole, really

comes in a two parameter family, parametrized by the temperature and a relative time shift between the two sides, see [37,38]. Again, these generators act as desired only for a particular synchronization of the two times. This is implicit in the above formulas when we write left-right correlators "at the same time".

2.3.4 Order from chaos

We can wonder what happens if we take the generators we defined, which are defined in terms of correlators at $u_l = u_r = 0$ and we "evolve" them with the boost Hamiltonian. We then get, in Lorentzian time,

$$\begin{aligned} e^{i\tilde{u}\hat{B}}\hat{E}e^{-i\tilde{u}\hat{B}} &= \frac{\beta}{2\pi}\left[H_r + H_l - \tilde{\eta}\sum_j O_l^j(-u)O_r^j(u) - \langle\cdots\rangle_{\rm TFD}\right] \simeq \cosh\tilde{u}\hat{E} - \sinh\tilde{u}\hat{P} \\ e^{i\tilde{u}\hat{B}}\hat{P}e^{-i\tilde{u}\hat{B}} &= \partial_{\tilde{u}}\left(\frac{\beta\tilde{\eta}}{2\pi}\sum_j O_l^j(-u)O_r^j(u)\right) \simeq -\sinh\tilde{u}\hat{E} + \cosh\tilde{u}\hat{P} \end{aligned} \tag{2.73}$$

where $\langle\cdots\rangle_{TFD}$ indicates the expectation value of the previous three terms in the TFD state. The first equality is what we get from the explicit definition of the hatted generators. The second equality is expected to hold for states that are close to the thermofield double, and it holds to the extent that we can approximate the hatted generators by the matter ones in the semiclassical limit, see (2.43) and to the extent that the hatted operators obey an approximate SL(2) algebra.

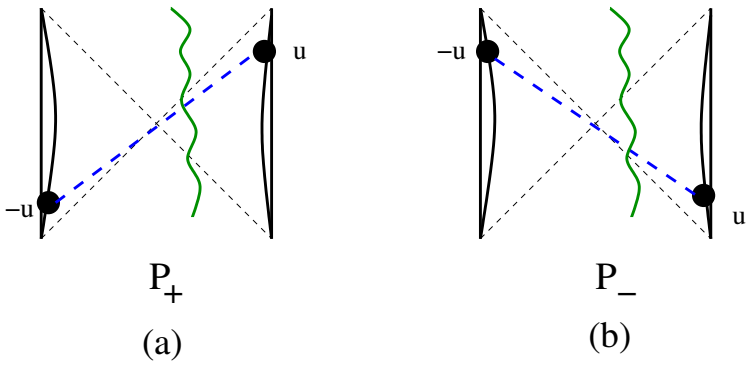


Figure 2.8: (a) In order to measure P_+ we can consider a correlator in this configuration with large u, but smaller than the scrambling time. (b) To measure P_- we consider instead a correlator at these times.

We can think of (2.73) as an approximate expression for the approximate symmetries at zero time in terms of operators at other times.

Notice that in (2.73) we have exponentially growing terms in the right hand side as $\tilde{u} \to \infty$. Such

terms can *only* come from the term involving $O_l^j(-u)O_r^j(u)$, which indeed can lead to exponential growth. The reason is the following. The expectation values of these operators on a state created by acting with operators on the thermofield double is an out of time order correlator. This is an analytic continuation to Lorentzian time of a configuration of operators as in figure 2.5(a). In (2.73) we are computing the *difference* between this out of time order correlator and the disconnected correlator contained in the thermofield double expectation value $\langle O_l^j(-u)O_r^j(u)\rangle_{\rm TFD}$. The latter is time independent due to the boost symmetry of the empty wormhole or thermofield double. On the other hand the out of time order correlator decays as u increases [39, 40]. This initial decay is given by an exponentially growing deviation from the disconnected diagram [39,40]. Since we have a difference between the two correlators in (2.73), we only pick up the correction that is exponentially growing in time.

We can concentrate on these growing terms and write a simple expression for $P_\pm$ at time equal to zero in terms of correlators at other times

$$\begin{aligned} -P_+ &= \frac{\hat{E}-\hat{P}}{2} = -\lim_{u\to+\text{large}} e^{-\frac{2\pi u}{\beta}}\frac{\beta\tilde{\eta}}{2\pi}\sum_j\left[O_l^j(-u)O_r^j(u) - \langle O_l^j(-u)O_r^j(u)\rangle_{\rm TFD}\right] \\ -P_- &= \frac{\hat{E}+\hat{P}}{2} = -\lim_{u\to-\text{large}} e^{\frac{2\pi u}{\beta}}\frac{\beta\tilde{\eta}}{2\pi}\sum_j\left[O_l^j(-u)O_r^j(u) - \langle O_l^j(-u)O_r^j(u)\rangle_{\rm TFD}\right] \end{aligned} \tag{2.74}$$

where $\tilde{\eta}$ is fixed by (2.65) and (2.67). The explicit exponential prefactors are decreasing in the corresponding limits and extract the growing pieces of the correlator corrections. In these equations when we say "large" we mean a large time but smaller than the scrambling time. In other words, a time obeying

$$1 \ll \tilde{u} \ll \log \mathfrak{s}. \tag{2.75}$$

Therefore, these formulas make sense only in the semiclassical limit, where $\mathfrak{s} \gg 1$.

The growing nature of the left-right correlators in the presence of matter is related to chaos [39, 41]. It was found that this growth is related to gravitational shockwaves which inducing null shifts of the bulk matter. Here we are inverting the logic and using these growing pieces to *define* the action of the generators. In this sense we are getting a symmetry (order) from chaos.

Alternatively, it was shown in [26] (see also [42]) that the two sided correlators induce null displacements of the matter propagating inside, when there is a large relative boost between the two. This is related to the phenomenon of quantum teleportation. Here we are using this phenomenon to talk about the symmetries.

In fact, the present discussion suggests that we will be able to use this growth for any non-zero

temperature black hole, not just near extremal ones. The only difference will be in the value of the commutator between $[P_+, P_-]$. In our case this gives B. But for a generic black hole we expect that this should be zero, because the symmetry near any horizon is just Poincare. In fact, even in our case, if we consider excitations that are very close to the horizon, they will have large values of P_+ and P_- so that is natural to rescale the generators, making them smaller. This in turn will also rescale the commutator. On the other hand, near any black hole horizon the boost generator has a natural universal normalization which is that of the "modular" Hamiltonian (conjugate to standard Rindler time).

Finally, we should remark that by looking at (2.73), which was derived from algebraic and symmetry considerations, we can deduce that the expectation value of $\sum_j O_l^j(-u)O_r^j(u)$, in a perturbed thermofield double state, should contain a term growing exponentially with maximal Lyapunov exponent in order to match the right hand side of (2.73). So we can view this as an algebraic derivation of the maximal chaos behavior. Of course, this is essentially the same as the original gravitational derivation using shock waves [39, 40] after we use the particular features of nearly AdS_2 gravity.

2.3.5 Generators for the one-sided case

All of the charges we have been discussing use two-sided operators. To what extent can we define physical matter charges if we only have access to one side? Clearly it is impossible to determine the matter charges for a general state with only one-sided observables. However, it is plausible that we could detect the matter charges on restricted states, where matter is inserted only from one side.

Let us be more concrete. The generators in section 2.2.1 were of the form $G^A = -e^A_a(Q^a_l + Q^a_r)$. Below we will consider choices of e^A that only depend on the right side. Nevertheless, the presence of Q^a_l seems troublesome. Now imagine that we start with a state with no matter, e.g., as $u \to \infty$, the matter charge vanishes $Q^a_m = 0$. On such states $Q^a_l(-\infty) = -Q^a_r(-\infty)$. If we furthermore assume that the left boundary evolves with the standard Hamiltonian H_l with no matter insertions, then Q^a_l is conserved for all times, so we may write

$$Q^a_l = -Q^a_r(-\infty). \tag{2.76}$$

We may also write this as

$$\Delta G^A_{os} = -e^A_a \Delta Q^a_r, \qquad \Delta Q^a_r = Q^a_{\mathrm{r,\,after}} - Q^a_{\mathrm{r,\,before}}. \tag{2.77}$$

ΔQ_r^a measures the change in the right gauge charges before and after the matter insertion. The subscript "os" means one sided. Note that we are really defining the *change* in the generators, not the generators themselves.

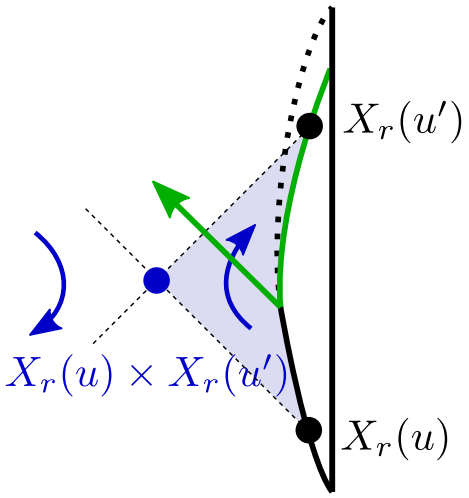


Figure 2.9: The one-sided generator. An unknown amount of matter is inserted at some time between u and u'. Our task is to detect it using right-sided observables. The blue point in the center is proportional to $X_r(u) \times X_r(u')$. The generator corresponding to this point fixes the causal wedge (shaded in light blue).

We now consider various choices of e^A. A relatively natural one is to use X_r^a at two different times $X_r(u), X_r(u')$. For the rest of this subsection, all quantities will be on the right side, so we will drop subscripts.

$$e_a^A = (e_a^{-1}, e_a^+, e_a^-) = \left(\frac{(X_r \times X_r')_a}{X_r . X_r'}, \sqrt{2} \frac{X_{r\,a}(u)}{\sqrt{X_r . X_r'}}, \sqrt{2} \frac{X_{r\,a}(u')}{\sqrt{X_r . X_r'}} \right) \tag{2.78}$$

The generators associated to such vectors have a nice geometric interpretation in terms of causal wedges. Namely, if $u < u'$ then we shoot a future directed light ray from u and past directed light ray from u'. Then the "causal wedge" is defined to be what is enclosed by these light rays, see figure 2.9. The first vector in (2.78) gives a generator that performs a boost around the intersection of the light rays, see figure 2.9. This boost generator maps points in the causal wedge to points in the causal wedge. In the QFT approximation to the bulk physics, ignoring gravity, one can view it as the modular Hamiltonian of the causal wedge[9].

This construction is also closely related to the recent work of [43]. There, they associate to two boundary times u and u' the point $X_r(u) \times X_r(u')$. They consider a bulk operator at such a

[9] In the full gravity theory the causal wedge should be only an approximate notion that arises when we restrict to simple operators in the boundary theory.

point. In other words, they gravitationally dress a bulk matter fields with the gravitational operators $X_r(u) \times X_r(u')$ in order to define a diffeomorphism invariant operator. Here we are dressing the matter charge with the same gravitational operators to obtain G_{os}^{-1}.

One problem with the vectors in (2.78) is that they fail to commute in the quantum theory. To quantify the extent of this problem, let us compute the commutator of the last two vectors in the semiclassical limit. We can do this by writing the general classical solution

$$X^a = -\frac{Q^a}{2H} + \left(X^a(0) + \frac{Q}{2H}\right)\cosh\left(\sqrt{2H}u\right) + \frac{\dot{X}^a(0)}{\sqrt{2H}}\sinh\left(\sqrt{2H}u\right) \tag{2.79}$$

We can then compute the Poisson bracket $\{X^a(u), X^b(0)\}_{PB}$. There will be many terms; the important point is that

$$\{X^a(u), X^b(0)\}_{PB} = M^{ab}e^{\sqrt{2H}u} + \cdots, \tag{2.80}$$

where M^{ab} is a function of $X(0), \dot{X}(0), H$, but independent of u. The important point is that there is an exponentially growing contribution to the commutator. On a thermal state, $\lambda = \sqrt{2H} = 2\pi/\beta$; this is exactly the maximal Lyapunov exponent. It is somewhat ironic that the maximal chaos, which we said was useful for constructing the two sided generators, is also what prevents us from defining good 1-sided charges.

This motivates us to consider $u \to u'$ so that the commutator is smaller. This is equivalent to choosing

$$e^A_a = \left(\dot{X}_a, \frac{\ddot{X}_a}{\sqrt{2H}}, \frac{(\dot{X}\times\ddot{X})_a}{\sqrt{2H}}\right). \tag{2.81}$$

Since $\dot{X}.\dot{X} = -1$, the first two vectors are automatically orthogonal $\ddot{X}.\dot{X} = 0$. Note, however, that there will be a non-trivial commutator between the different components of e^A. For example, the commutator of the first two components will get a contribution from

$$[\ddot{X}^a, \dot{X}^b] = i\epsilon^{abc}\dot{X}^c + i\{H, X^aX^b\} - i\{\ddot{X}^a, X^b\}. \tag{2.82}$$

Here $\{A, B\} = AB + BA$. If we contract this expression with the matter charges $\Delta Q^a_m \Delta Q^b_m$ the last two terms can become large. In particular, it grows exponentially as a function of the boundary time when the matter was inserted. So even if we use vectors at the same time, chaos will lead to

the breakdown of the one-sided algebra if we wait too long after perturbing the right side.

Finally, let us turn to the momentum discussed in [29]. They consider the momentum of a particle thrown in from one side. "momentum" here means the variable conjugate to distance. But distance from what? It is most natural to take the distance from the bifurcating surface on the left side, since the bifurcating surface on the right changes when matter is inserted. The left bifurcating surface sits at a point $Y^a \propto Q_l^a$, which corresponds to a definition of momentum

$$\tilde{P}_{\rm os} = Q_m . \frac{Q_l \times X_r}{Q_l . X_r} = \frac{\dot{X}_r . Q_l}{X_r . Q_l} = \frac{d}{du_r} \log(-X_r . Q_l). \tag{2.83}$$

From the last line, it is clear that $\tilde{P}_{\rm os}$ is proportional to the velocity of the right boundary particle relative to the entangling surface. Note, as above, we may replace $Q_l = -Q_r(-\infty)$ to arrive at a purely one-sided quantity.

If we consider semiclassical states with a particle thrown in at some time $u_0 < 0$ from the right, the one-sided momentum approximates the two-sided momentum $\tilde{P}_{\rm os} \approx \tilde{P}$ as long as $|u_r|$ is less than the scrambling time. This is because the geodesic connecting X_r and X_l approximately intercepts the bifurcating surface Q_l.

2.4 Exploring the bulk

2.4.1 Evolving with the charges

In section 2.3.3 we pointed out that some generators, such a $\tilde{E}$, can be approximated, in the semi-classical limit, by a simple coupled Hamiltonian (2.64) (2.69). It is natural to ask whether it is possible to systematically correct this coupled Hamiltonian so that it gives the exact generator $\tilde{E}$.

One simple way to think about this is to declare that the full Hamiltonian of the coupled system *is* simply

$$\tilde{H}_{\rm coupled} = \tilde{E} \tag{2.84}$$

This is not the same as (2.69), hence the tilde. This seems a legitimate Hamiltonian from the point of view of the gravity theory[10].

This Hamiltonian has a number of differences with (2.64) or $\hat{E}_c$ in (2.69). A simple difference is that we do not need to subtract the ground state energy as in (2.69). In fact, by construcution, $\tilde{E}$ annihilates the thermofield double state. A more important difference is that $\hat{E}$ in (2.69) depends

[10] Non-perturbative corrections that can render the distance between the two boundaries ill defined, and therefore (2.84) ill defined. We ignore such corrections here.

on β (through the temperature dependence on (2.67) (2.68). This means that $\hat{E}$ is close to the generator $\tilde{E}$ only for states that are close enough to the TFD states with inverse temperature β. In contrast, the Hamiltonian (2.84) is β independent. TFD states with any temperature and any relative synchronization between left and right times are ground states of (2.84). In (2.84) only states with nontrivial bulk matter have non-zero energy (under the Hamiltonian $\tilde{H}_{\text{coupled}}$).

In the presentation of the charges in (2.31), we see that $\tilde{E}$ moves the matter along the global time translation symmetry generator but leaves the boundaries at the same positions. Up to an $SL(2)_g$ gauge symmetry this is the same as moving the boundaries and keeping the bulk matter fixed. This is close to what we mean by the time evolution of the bulk observer. The picture is very similar to the one for the evolution with (2.64), [27], where the two physical boundaries move vertically in global AdS_2. The difference is that these physical boundaries can have any location here, while under (2.64) they had a preferred location.

In order to explore the relation between H_{coupled} and $\tilde{H}_{\text{coupled}}$ a bit further, it is useful to write the expression for $\tilde{E}$ in terms of the global time T_l and T_r. In particular, if we act by physical symmetries and choose a special gauge we can classically restrict[11] to symmetric configurations $T_l(u) = T_r(u)$. We find from (2.38)

$$\tilde{E} = 2\left(-T' + \frac{T''^2}{T'^3} - \frac{T'''}{T'^2}\right) = -2e^{-\varphi}(\varphi'' + e^{2\varphi}), \qquad \tilde{P} = \tilde{B} = 0 \tag{2.85}$$

with $\varphi \equiv \log T'$. This is also the matter energy in our gauge, and this the same as the gauge constraint $E_m = -Q_l^1(T_l) - Q_r^1(T_r) = \tilde{E}$. To derive this formulas we can write $X_r^A = (\cos T_r, \sin T_r, 1)/T_r'$ and a similar expression for T_l.

We can interpret the full expression for $\tilde{E}$ in (2.38) as follows. The prefactor

$$(-2X_l.X_r)^{1/2} = 2e^{-\varphi} \tag{2.86}$$

simply gives a redshift factor of order β when acting on states near the thermofield double. The first three terms in parentheses of (2.38) is precisely the Hamiltonian in [27] with $\Delta = 1$ and $\eta = 1/2$

$$H_{\text{coupled}} = -2\varphi'' + (\varphi')^2 - e^{2\varphi} - \eta e^{2\Delta\varphi}\ , \qquad \eta = \frac{1}{2}\ , \qquad \Delta = 1 \tag{2.87}$$

[11] By acting on a general state with the physical symmetries, we can set $\tilde{P} = \tilde{B} = 0$ at $u_l = u_r = 0$. We then use the SL(2)$_g$ gauge symmetry to set $t_l(0) = t_r(0) = 0$ and $t_l'(0) = t_r'(0)$. So the matter and physical charges align $e_a^A = \delta_a^A$, at $u_l = u_r = 0$. Now the two equations $\tilde{P}(0,0) = 0$, $\tilde{B}(0,0) = 0$ set $t_l''(0) = t_r''(0)$ and $t_l'''(0) = t_r'''(0)$. So at $u_l = u_r = 0$ the coordinates and momenta are equal on both sides. Then using the classical equations of motion, the coordinates and momenta are the same for all times. Hence $t_l(u) = t_r(u)$. Note that this argument could be rerun with the coupled Hamiltonian with almost no modification.

The last terms in (2.38) give a similar expression which combines to

$$\tilde{E} = (-2X_l.X_r)^{\frac{1}{2}}(-\varphi'' - e^{2\varphi}) = 2e^{-\varphi}(\frac{1}{2}e^{\varphi}E_m) = E_m \tag{2.88}$$

where we used (2.85) viewed as a gauge constraint. We see that the dynamical boundary variables have disappeared. So with this Hamiltonian, the boundary has no dynamics. This expression, (2.88), looks misleadingly simple because it was written in a special gauge. In order to act with $\tilde{E}$, we do need to know the boundary positions. We need to know the relative synchronization of the two times, for example, and to extract this information we need to measure some left-right correlators. This is a common feature in gauge theories, where an expression in a fixed gauge might look local (here $\tilde{E}$ appears to involve only one factor in the Hilbert space), but the full gauge invariant operator is not local.

Before proceeding, let us mention a subtlety in the above discussion. In the above, we wrote expressions which involved derivatives with respect to u. These expressions implicitly assumed that u-translation was generated by $H_l + H_r$. However, when we imagine evolving with $\tilde{E}$ or with the coupled Hamiltonian H_{coupled}, our expressions will be modified. A better approach is to write this in terms of coordinates and momenta. This is developed in Appendix 2.8. There, we give explicit formulas for the charges and $X^a_{l,r}$ in terms of the coordinates T, T' and their conjugate momenta p_1, p_2 (see equation (2.132) and (2.126)). Using these formulas, we can express $\tilde{E}$ in terms of coordinates and momenta via the relation (2.31) or via (2.33).

If we denote by p^r_1, p^r_2 the momenta conjugate to the global times $T_r(u)$ and $T'_r(u)$, the statement is that when coordinates and momenta for different sides are always equal (up to minus signs), then we get the simple expression

$$T'\tilde{E} = -T'(p^l_1 + p^r_1). \tag{2.89}$$

Then evolving by $T'\tilde{E}$ gives the solution $T = -T'u$.

In [27], the low energy spectrum of the coupled wormhole was approximately a tensor product of the bulk matter $e^{\varphi}E_m$ and the boundary Schwarzian degrees of freedom whose breathing mode is an anharmonic oscillator. While the bulk matter Hamiltonian $e^{\varphi}E_m$ organizes into SL(2) multiplets, the boundary degrees of freedom does not, since for one thing the oscillator's frequency differs from the AdS_2 frequency. Viewed as a Hamiltonian $\tilde{E}$ solves the above problem by subtracting off the kinetic terms and then flattening out the potential energy of the oscillator. Finally, there is an

overall factor which removes the redshift factor so that the Hamiltonian is exactly the bulk energy. The result is an energy spectrum of $\tilde{E}$ that is independent of the Schwarzian modes. One might also imagine an opposite strategy of achieving an approximate $\mathrm{SL}(2,\mathbb{R})$ spectrum where instead of removing the energy of the boundary degrees of freedom, one makes the frequency of oscillation so large that the boundary modes are essentially frozen, and we have an effective description that only involves the $\mathrm{SL}(2,\mathbb{R})$ matter. A preliminary exploration of this idea is given in Appendix 2.12.

While $\tilde{H}_{\text{coupled}}$ makes sense in JT-gravity plus matter, one can question whether we can really construct it from a more microscopic theory, such as a full boundary quantum mechanical theory. This is of course a question about all G^A generators. In the next section, we discuss a particular large N scaling limit of SYK, where these generators make sense. On the other hand, in a boundary quantum mechanical theory with a finite Hilbert space, we should not be able to construct the generators G^A (since they generate an infinite number of states). The difficulty lies in measuring distance, as we will discuss in section 2.6.1.

2.4.2 Exploring behind the horizon or moving the horizon

One of our motivations was to understand better how matter moves in the bulk and how that is represented in the boundary theory. The generators we constructed allow us to move matter in the bulk relative to the boundaries, so they allow us to explore the bulk. One would like to be able to explore the region behind the horizon. Indeed these charges allow us to move matter within the Wheeler-de-Witt patch, see figure 2.2.

Actually, it is also important to understand the sense in which we can move matter. When we act with the generator $\tilde{E}$, for example, we are either moving matter or moving the boundaries (these are two equivalent descriptions). Let us take the point of view that we leave the matter in the bulk as it is but we move the boundaries forwards in time along the vertical direction in the Penrose diagram (i.e. by performing a global time translation $T \to T+$constant (2.6a)). But, if after doing this, we let the boundaries evolve with decoupled Hamiltonians, then we would find now the horizon at a new position, see figure 2.10(c). So, we can say that the $\tilde{E}$ generator, allows us to explore the region that would have been behind the horizon if we had done nothing. By the very act of evolving with $\tilde{E}$, we have brought it out of the horizon (see [26]). The horizon is a teleological object, and in the quantum theory, it is related to the limitation on the types of experiments we can do. For example, it depends on whether we allow a coupling between the two boundaries. We see an important point: a black hole is not just a "state", but a state with some evolution law. Only after specifying the

evolution process can we can say that the black hole is "black" (it has a horizon). More specifically, if we start with the usual two boundary wormhole and we do not allow any information exchange between the two boundaries, then we have two black holes. But if we allow information exchange and also allow operators that couple the two boundaries, then we could have an eternal traversable wormhole, as in [27]. In both of these cases the "state" at $u_l = u_r = 0$ is the same, or very similar.

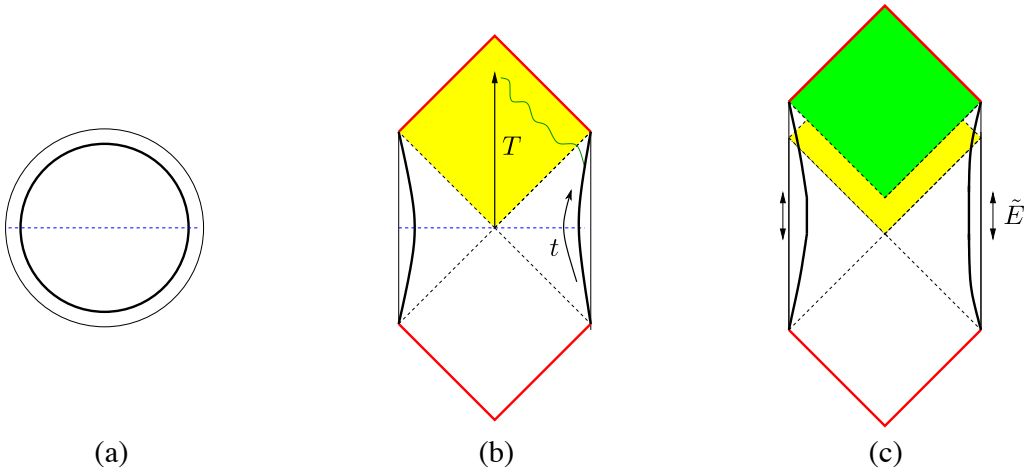


Figure 2.10: (a) The standard Euclidean AdS_2 picture and its continuation to Lorentzian signature (b). This looks similar to the vacuum decays to AdS studied by Coleman and de Luccia [44]. They have shown that the addition of irrelvant operators at the domain wall, which is the boundary for us, leads to a divergence in the bulk at the red line. A bulk observer moving along the central black arrow gets to see a lot of the boundary in a very small proper time. This is the usual blue-shift near the inner horizon. In (c) we evolve for some time using the generator $\tilde{E}$. This allows us to explore some of the region that was behind the horizon, but the new horizon moves up. The yellow region was behind the horizon and now it is outside.

2.4.3 The inner horizon or Cauchy horizon

From the point of view of the matter in the bulk, the evolution is given by E_m and it would seem at first sight that we could continue such evolution "forever". However, our present discussion does not allow us to move past the inner horizon or Cauchy horizon. The reason is the following. We assumed that the matter fields have standard boundary conditions at the AdS_2 boundary. These are implied by the boundary conditions at the physical boundaries (curved black lines in figure 2.2). However, beyond the region where the physical boundaries extend, we have no guarantee that the matter boundary conditions are the same as when we had a physical boundary. For this reason we cannot extend the bulk evolution beyond the dotted red lines in figure (2.2). Note that this is also the boundary of the Wheeler de Witt patch when we move both the left and right times to the far future. Of course, it is an important problem to figure out what happens beyond that region!

It has been argued by Penrose that the inner horizon would be generically singular. (Though it has been demonstrated that classically the singularity is not too bad [45] and could be traversed, in some cases.) Here we can connect this expectation to the related discussion of vacuum decay into AdS that was studied by Coleman and de Luccia [44], see figure 2.10(a,b). In the thin wall approximation, the action of the bubble has a surface term and a boundary term, which reproduces JT gravity (2.12), [4, 6]. In our case, only the "true vacuum" part is present, not the false vacuum. Coleman and de Luccia have argued that in such situations there will be a singularity at the inner horizon. The reason is the following: imagine that we have a scalar field ϕ in the bulk and that there is some source for it on the boundary. Then, even if the field ϕ is massive, and thus corresponds to an irrelevant perturbation [46], it is expected to have a non-zero expectation value at the usual horizon. This is translation invariant in the FLRW patch (the yellow patch in figure 2.10(b)). Then the FLRW evolution will generically make it singular along the red line. (For a free bulk field one can avoid the singularity if the corresponding dimension Δ is an integer). In the case of the SYK model, we have other operators turned on when we are at finite $\beta\mathcal{J}$. For a near extremal 4d charged black hole, the fact that the boundary conditions for the fields allow a leakage into the flat space region implies that we have some double trace operators turned on. This could imply that operators such as $\phi\phi$ would get divergent expectation values. This is a quantum effect. It is also suppressed in the scaling limit we took, but it seems important if $\beta\mathcal{J}$ is large and finite or the black hole throat has finite length.

Just to put in some formulas into this discussion we can consider a scalar field ϕ and imagine that there is some source on the boundary. We can then use the bulk-to-boundary propagator to compute

$$\langle\phi\rangle \propto \int_{\mathcal{C}} du \frac{1}{[-Y.X(u)]^{\Delta}} \propto \int_{\mathcal{C}} du \frac{1}{[\cos T + \sin T \sinh t(u)]^{\Delta}} \tag{2.90}$$

where we took a point at the center of bulk with $Y^a = (\cos T, \sin T, 0)$ and a point at the boundary with $X \propto (1, \sinh t, \cosh t)$, see figure 2.10(b). The inner horizon corresponds to $T = \pi$, and we see that there the $\sin T$ factor becomes zero and there could be a divergence from the integral over large real times. To analyze this properly we need to specify the contour of integration $\mathcal{C}$ which goes over the Keldysh contour appropriate for this problem. The regions that contribute are those where Y and $X(u)$ are timelike separated, where the $i\epsilon$ prescription for the forwards and backwards parts Keldysh contour are different and do not cancel if Δ is not an integer. Formula (2.90) applies also for deformations by products of bulk field operators ("double trace"), where Δ in (2.90) is the total dimension of the operator.

Notice that this singularity can be moved by evolving the system for some time using the generator $\tilde{E}$, see figure 2.10(c).

Note that the yellow region in figure 2.10(b) looks like a two dimensional Friedman-Lemaitre-Robertson-Walker two dimensional cosmology. With the perturbations we discussed, it seems to develop a bulk singularity. A proper boundary understand of this region from the boundary theory would give us a toy model for a 2d FRLW cosmology. This is just the two dimensional version of a general connection between nearly conformal theories on dS_{D-1} and negative cosmological constant FLRW cosmologies with $D-1$ hyperbolic slices [46].

2.4.4 Moving operators into the bulk

When we study gravitational systems with a boundary, one sometimes wants to express operators in the interior in terms of operators closer to the boundary. For example, if we had a bulk field $\phi(\tau, x)$ defined in the bulk of AdS_2, we want to express the operator deep inside in terms of an operator closer to the boundary. One way to do it is via the HKLL construction [47] which involves solving the bulk wave equation and expressing the field at a point in the bulk as an integral of the field near the boundary over a range of times. Here we will provide an alternative construction.

We have constructed an operator $\tilde{P}$ which performs translations in the bulk, so we could use it to translate an operator deep in the interior to an operator closer to the boundary. Roughly we want

$$\phi(\tau, x) = e^{i(x-x_{\rm bdy})\tilde{P}}\phi(\tau, x_{\rm bdy})e^{-i(x-x_{\rm bdy})\tilde{P}} \tag{2.91}$$

However, this expression is not good enough and we need to clarify some subtleties before writing a better expression.

First, recall that in the construction of $\tilde{P}$ we assumed that the actual UV boundaries where infinitely far away. Therefore the point $x_{\rm bdy}$ should still be far away from the boundary, but it could sit at a relatively large value of the coordinates x in (2.6a), a value that is large but fixed when we take the UV regulator to zero. In other words, we want $x_{\rm bdy}$ to be larger than any value of x of other operators in the bulk that we want to consider, but finite in the limit that we send the boundaries far away. A similar assumption goes into the standard HKLL [47] construction if one wants to use simple AdS wavefunctions.

A second issue is that the boundaries are dynamical objects and the coordinate points x are not physical by themselves. When we construct the operator relatively close to the boundary we only determine its position relative to the boundary, we will call this $\ell_{\rm bdy}$, or $x_{\rm bdy} = \ell_{\rm bdy} + x_r$, where

x_r is the position of the right boundary. Such an operator can be constructed in various ways, see e.g. [43]. For an operator with a large value of $\ell_{\rm bdy}$ we expect that quantum fluctuations of the Schwarzian variables are small and the construction will be fairly accurate.

With all these caveats, we can now construct a better expression for a bulk operator at some distance ℓ from the UV boundary as

$$\Phi(\tau,\ell) = e^{i(\ell-\ell_{\rm bdy})\tilde{P}}\Phi(\tau,\ell_{\rm bdy})e^{-i(\ell-\ell_{\rm bdy})\tilde{P}} \tag{2.92}$$

$$\Phi(\tau,\ell_{bdy}) = \int dx_l dx_r \phi(\tau,\ell_{\rm bdy}+x_r)|x_l,x_r\rangle\langle x_l,x_r| \tag{2.93}$$

The first expression translates the operator from a point closer to the boundary to points deeper into the bulk. The second line expresses the operator close to the boundary in a gauge invariant fashion. This operator involves an operator ϕ acting on the matter Hilbert space. It also involves projection operators onto definite coordinate values for the boundary particles. Thus, it acts on the full Hilbert space of the theory (2.15). The momentum $\mathrm{SL}(2)_g$ gauge generator acts by shifting all x coordinates in (2.93) by a constant, which can be absorbed by a shift of integration variables. We have used a capital Φ to express the final dressed operators.

Finally, a more precise description of the bulk operator would also involve the time boundary variables and is

$$\Phi(u_l,u_r,\ell) = \int d^2X_l d^2X_l\, \phi[Y^a(X_l,X_r;\ell)]|X_l,X_r\rangle\langle X_l,X_r| \tag{2.94}$$

where Y^a is a point determined as follows. First we find the geodesic going between X_l^a to X_r^a . Then we determine its midpoint. Then we move by a distance ℓ to the right along that geodesic to determine Y^a.

One can check that $[\Phi(\tau,\ell),\Phi(\tau,\ell')] = 0$ for $\ell \neq \ell'$. This should be compared to the HKLL construction where the construction has to be modified order by order to ensure this commutativity [48, 49]. The construction discussed here includes the full gravitational dressing to all orders in the G_N expansion[12]. Furthermore, all matter self interactions have been taken into account. The prescription is somewhat similar to the prescription of "shooting a geodesic orthogonal to the boundary" [51, 52]. One issue is that, in AdS, all spacelike geodesics are orthogonal to the boundary[13]. Here we choose a precise geodesic by selecting two boundary points, one on the left and one on the right.

Note that the whole discussion in this subsection is about the bulk theory, not the holographic

[12] Due to topology changing corrections, of order e^{-S_0}, this prescription is not well defined non-perturbatively, see [50].

[13] One could still select a unique geodesic by choosing more than one point along the boundary, we discuss this in section (2.3.5). In our case, this can only work in the semiclassical limit.

boundary theory. Of course, it is convenient for holography to have the bulk operators written in terms of operators near the boundary.

2.5 Connection to SYK and other systems

2.5.1 Generators in SYK

Here we discuss these generators in the context of SYK. If we consider the system at temperature β the effective coupling is βJ. It is convenient to consider the limit

$$\beta J \to \infty \ , \qquad N \to \infty \ , \qquad \frac{N}{\beta J} = \text{fixed} \tag{2.95}$$

where the Schwarzian action becomes exact [25]. It is important to remark that, in this limit, we can consider quantum mechanical effects in the Schwarzian action. Also, since $S_0 \propto N \to \infty$, other topologies do not contribute.

The constructions we discussed above for the charges can be discussed in this model. Whenever we got an expression involving $X_l.X_r$ we could replace it by a fermion operators via

$$\frac{1}{(-2X_l.X_r)^{\Delta}} \propto \frac{i}{N}\sum_j \psi_l^j \psi_r^j \tag{2.96}$$

Our expressions for the charges involved functions of $X_l.X_r$. We can then consider functions of these correlators, such as the logarithm or other powers. In the limit (2.95) these functions are well defined for the low energy states under consideration. Namely, one can be worried that the operator in the right hand side of (2.96) has zero or negative eigenvalues. However, in the large N limit (2.95), we do not access such eigenvalues from low energy states. For such low energy states, and in the limit (2.95), the operator in the right hand side (2.96) is a *positive* operator. Therefore we can raise it to arbitrary powers (positive and negative) and we can also take its logarithm in order to construct the exact generators (2.35) (2.38).

Now, one can say that in the infinite N limit, (2.95), we have an infinite number of fermions anyway so it is not at all surprising that we can find some exact $SL(2)$ algebra. What is interesting is that the quantum effects of the boundary mode are still finite in this limit, (2.95). In particular, the scrambling time for excitations of thermal energies is still finite, in this limit. And also the dynamics of the boundary mode is not conformal invariant. The non-trivial statement we are making is that, despite these facts, we still have exact $\mathrm{SL}(2,\mathbb{R})$generators.

Most of our discussion used the language of nearly-AdS_2 gravity. However, the SYK model displays the same structure, as reviewed in section (2.1.3). So everything we did in this chapter also holds for the SYK model. Note that we made an important assumption when we discussed the JT gravity theory, right after (2.12). We said that the boundary was very far away. This scaling limit of the JT theory, where $\phi_b \to \infty$, is essentially the same as (2.95) in SYK. Unfortunately, the G, Σ action does not give us a simple Hilbert space description, other than (2.15). In particular, one would like to understand how this emerges from the fermions. Or how the Hilbert spaces in (2.15) are embedded into the full Hilbert space of the fermions. The analysis of aspects of these symmetries directly in terms of the femions was undertaken in [30, 53]. The results in this chapter provide a "target" for such discussions.

Of course, an important question is how this structure is broken at finite N. We will not discuss that in this chapter. However, we will now make the following simple observation. The action of the generator $\hat{E}_c$ (or $\hat{P}_c$) on a state created by a fermion was discussed near (2.52). Once we use the expression (2.69) (2.63) for the $\hat{E}_c$ generator, then the computation boils down to a fermion four point function computation of the general form

$$\mathfrak{s} \frac{\langle \psi^i(\pi)\psi^j(u_t)\psi^i(0)\psi^j(u_b)\rangle}{\langle \psi^i(\pi)\psi^i(0)\rangle\langle \psi^j(u_t)\psi^j(u_b)\rangle} \ , \qquad \mathfrak{s} \propto \frac{N}{\beta J} \tag{2.97}$$

and there is no sum over i or j. What we want is that this ratio of correlators is a certain particular order one function of the $\tilde{u}_i$ variables. Now, the general structure of the four point function is

$$\frac{\langle \psi^i(\pi)\psi^j(u_t)\psi^i(0)\psi^j(u_b)\rangle}{\langle \psi^i(\pi)\psi^i(0)\rangle\langle \psi^j(u_t)\psi^j(u_b)\rangle} = 1 + \frac{\beta J}{N}\mathcal{F}_{\text{enhanced}} + \frac{1}{N}\mathcal{F}_{\text{finite}} \tag{2.98}$$

where $\mathcal{F}_{\text{enhanced}}$ comes from the Schwarzian mode. The 1 is subtracted with the expectation value in (2.69). Due to the factor of of $\mathfrak{s}$ in (2.97) we pick up *only* the part involving $\mathcal{F}_{\text{enhanced}}$ and the $\mathcal{F}_{\text{finite}}$ part drops out. In fact this last term contains the conformal invariant contribution that features an infinite sequence of composite operators, etc. Such terms depend on something which we can call "bulk interactions" [54]. Such terms drop out in the limit (2.95). If had not taken that limit, then we see that there are some specific $1/\mathfrak{s}$ corrections that would change the action of $\hat{E}$ relative to that expected for an $SL(2,\mathbb{R})$generator. We will not discuss here whether this can be fixed up, or to what extent. But it is of course an interesting problem!

2.5.2 Relation to "size"

It is worth noting that in SYK language the three generators in (2.63), (2.69), (2.71) can be expressed as

$$\hat{B} = \frac{\beta}{2\pi}(H_r - H_l) \tag{2.99}$$

$$\hat{E} = \frac{\beta}{2\pi}\left[H_r + H_l + i\mu\sum_j \psi_l^j\psi_r^j - \langle H_r + H_l + i\mu\sum_j \psi_l^j\psi_r^j\rangle_{\rm TFD}\right] \tag{2.100}$$

$$\begin{aligned}\hat{P} &= -i[\hat{B},\hat{E}] = -i\frac{\beta^2}{4\pi^2}\left\{i\mu\psi_l^i[H_r,\psi_r^i] - i\mu[H_l,\psi_l^i]\psi_r^i\right\} \\ &= -i\mu\frac{\beta^2}{4\pi^2}(\psi_l^i\dot{\psi}_r^i - \dot{\psi}_l^i\psi_r^i)\end{aligned} \tag{2.101}$$

$$\text{with} \quad \frac{\mu}{\mathcal{J}} = \frac{4\alpha_S}{\Delta c_\Delta}\left(\frac{\pi}{\beta\mathcal{J}}\right)^{2-2\Delta} = \frac{\alpha_s}{\Delta}\left(\frac{2\pi}{\beta\mathcal{J}}\right)^2\frac{1}{G(\frac{\beta}{2})} \tag{2.102}$$

where $G\left(\frac{\beta}{2}\right) = c_\Delta\left(\frac{\pi}{\beta\mathcal{J}}\right)^{2\Delta}$.

On the other hand, in previous investigations, the concept of the "size" of an operator was of interest [28–31]. Roughly speaking, in SYK, the size of an operator counts the number of fermions that will be affected by applying this operator. One way to characterize this is to consider its commutator with all the fermion operators. [14] For an operator normalized so that $2^{-\frac{N}{2}}\operatorname{tr} O^\dagger O = 1$, the operator size is given by [31]:

$$S_\infty(\mathcal{O}) = \frac{1}{2}\sum_{i=1}^{N} 2^{-\frac{N}{2}}\operatorname{tr}([\mathcal{O},\psi^i]^\dagger[\mathcal{O},\psi^i]). \tag{2.103}$$

For an operator $\psi(u_r) = \sum_k\sum_{i_1<\ldots<i_k} c_{i_1\ldots i_k}\left(2^{\frac{k}{2}}\psi_{i_1}\ldots\psi_{i_k}\right)$, this expression[15] gives

$$S_\infty = \sum_k \sum_{i_1<\ldots<i_k} k|\sqrt{2}c_{i_1\ldots i_k}|^2.$$

This can be written as the expectation value of some "size" operator defined as

$$\hat{S} = \sum_j i\psi_l^j\psi_r^j + \frac{N}{2} \tag{2.104}$$

$$S_\infty(\mathcal{O}) = \langle I|\,\mathcal{O}_r^\dagger\hat{S}\mathcal{O}_r\,|I\rangle\,, \tag{2.105}$$

[14] In the following expressions we assume the operator $\mathcal{O}$ is bosonic. When the operator we are interested in is fermionic, we can multiply it by a fermionic operator from some external system that anticommutes with all ψ^i's.

[15] There is a factor of $\sqrt{2}$ due to the normalization $(\psi_i)^2 = \frac{1}{2}$.

where $|I\rangle$ is the infinite temperature thermofield double state.[16] Note that part of the global energy operator $\hat{E}$ in (2.100) appears here.

The above notion of size only depends on the operator. A generalization of this notion of size which depends on both the operator and the temperature $1/\beta$ of the system is defined in [30], see also [29]:

$$S_\beta(\mathcal{O}) = \delta_\beta^{-1}\left(S_\infty(\mathcal{O}\rho^{\frac{1}{2}}) - S_\infty(\rho^{\frac{1}{2}})\right). \tag{2.106}$$

Here ρ is the thermal density matrix and δ_β is some normalization factor[17] which fixes $S_\beta(\psi^1) = 1$. This overall renormalization also implies $S_\beta(\psi^1(u))$ saturates at the value $\frac{N}{2}$. In the limit that $\beta \to \infty$, we recover (2.103). We may also write this in terms of the size operator $\hat{S}$:

$$S_\beta(\mathcal{O}) = \delta_\beta^{-1}\left(\frac{\langle \mathsf{TFD}|\mathcal{O}_r^\dagger \hat{S}\mathcal{O}_r|\mathsf{TFD}\rangle}{\langle TFD|\mathcal{O}_r^\dagger\mathcal{O}_r|\mathsf{TFD}\rangle} - \langle \mathsf{TFD}|\hat{S}|\mathsf{TFD}\rangle\right) \tag{2.107}$$

where the expectation value is taken on thermofield double at temperature $1/\beta$.

Here we see that size is related to global energy. Let $\mathcal{U}_r$ be a unitary operator acting from the right side. Then, once the coefficients are adjusted and the factors of H_l and H_r are added as in (2.100), (2.99), (2.102), the size of $\mathcal{U}_r$ is

$$S_\beta(\mathcal{U}_r) = \frac{\Delta}{2\alpha_s}\frac{\beta\mathcal{J}}{2\pi}\langle \mathsf{TFD}|\mathcal{U}_r^\dagger(\hat{E}-\hat{B})\mathcal{U}_r|\mathsf{TFD}\rangle. \tag{2.108}$$

The fact that $\mathcal{U}_r$ is unitary is useful because it ensures that the term involving H_l drops out from (2.108).

There are some interesting consequences of this expression. First, the $\hat{B}$ term is not important if one is only interested in the time dependence of size as we move a given operator insertion to earlier and earlier times (which was the focus of [29, 30]), since in that case the energies H_r and H_l are conserved. Let $\mathcal{U}_r(u_r) = e^{iH_r u_r}\mathcal{U}_r e^{-iH_r u_r}$. It was found that its size grows as $|u_r|$ gets large. This growth is the growth of the corrections to the usual out of time ordered correlators if we expand out (2.107).

Using (2.108), we have

$$\frac{\partial S_\beta(\mathcal{U}(u_r)))}{\partial u_r} = -i\frac{\Delta}{2\alpha_s}\frac{\beta\mathcal{J}}{2\pi}\langle \mathcal{U}_r(u_r)^\dagger\left[\frac{2\pi}{\beta}\hat{B},\hat{E}\right]\mathcal{U}_r(u_r)\rangle = \frac{\Delta}{2\alpha_s}\mathcal{J}\langle \mathcal{U}_r(u_r)^\dagger\hat{P}\mathcal{U}_r(u_r)\rangle. \tag{2.109}$$

[16] $|I\rangle$ satisfies $(\psi_L^j + i\psi_R^j)\,|I\rangle = 0$. As a consequence $\langle I|\,\hat{S}\,|I\rangle = 0$.

[17] $\delta_\beta = 2G(\frac{\beta}{2}) = 2c_\Delta\left(\frac{\pi}{\beta\mathcal{J}}\right)^{2\Delta}$, see [30].

We see that the momentum is related to the time derivative of size. This is also explored in [55].

The size of such an operator at large q was computed in SYK in [30]. To make such a perturbation at $u_r = 0$, we insert an operator of dimension Δ to the Euclidean circle at $\varphi_b = \frac{2\pi}{\beta}(-\frac{1}{2\mathcal{J}} + iu_r)$, $\varphi_t = \frac{2\pi}{\beta}(\frac{1}{2\mathcal{J}} + iu_r)$. We give a small real part to φ to smear the operator by the amount $\sim \frac{1}{\mathcal{J}}$.[18] Using (2.58) (and similar formulas for the other two generators obtained through (2.52)) we can get the charges of such an excitation.

$$\begin{aligned}\left(\hat{B}, \hat{P}, \hat{E}\right) &= \left(\Delta \frac{1}{\tan\frac{\varphi_t-\varphi_b}{2}}, -i\Delta \frac{\sin\frac{\varphi_t+\varphi_b}{2}}{\sin\frac{\varphi_t-\varphi_b}{2}}, \Delta \frac{\cos\frac{\varphi_t+\varphi_b}{2}}{\sin\frac{\varphi_t-\varphi_b}{2}}\right) \\ &\sim \Delta \frac{2\beta\mathcal{J}}{2\pi}\left(1, \sinh\left(\frac{2\pi}{\beta}u_r\right), \cosh\left(\frac{2\pi}{\beta}u_r\right)\right)\end{aligned} \tag{2.110}$$

where we assumed $|\varphi_t - \varphi_b| \ll 1$. Using (2.108) we find its size

$$S_\beta(\psi(u_r)) = \frac{\Delta}{2\alpha_s}\frac{\beta\mathcal{J}}{2\pi}(\hat{E} - \hat{B}) = 2\frac{\Delta^2}{\alpha_s}\left(\frac{\beta\mathcal{J}}{2\pi}\right)^2 \sinh^2\left(\frac{\pi}{\beta}u_r\right) \tag{2.111}$$

This agrees with the explicit fermion counting calculation in [30] at large q.

In the above example, the excitation has boost energy $2\Delta\mathcal{J}$. It starts from near the boundary and falls toward the horizon. A lower energy excitation can be made by creating a particle at rest at a finite proper distance to the horizon ρ_m. We denote an operator which creates such an excitation by $\tilde{\psi}(\rho_m)$. We can compute the size of such an operator using (2.108). To create such an excitation, one can consider insertions at some finite angle $\varphi_t = -\varphi_b$ on the Euclidean circle, see figure 2.6. Its distance to the horizon is $\rho_m = -\log\left(\tan\left(\frac{\varphi_t}{2}\right)\right)$. We write its charges (2.59) in terms of ρ_m:

$$\left(\hat{B}, \hat{P}, \hat{E}\right) = (\Delta \sinh\rho_m, 0, \Delta\cosh\rho_m)\,. \tag{2.112}$$

The size of this single-fermion perturbation[19] is

$$S_\beta(\rho_m) = \frac{\Delta^2}{2\alpha_s}\frac{\beta\mathcal{J}}{2\pi}e^{-\rho_m}. \tag{2.113}$$

When the excitation is near the boundary, $\rho_m \sim \log(\beta\mathcal{J})$ and the size is of order 1. As we move the excitation deeper and deeper into the throat its size increases exponentially and reaches $\sim \beta\mathcal{J}$

[18] The amount of smearing needed to go from UV to the conformal regime is $\sim \frac{1}{\mathcal{J}}$. We determined the $\frac{1}{2}$ factor in the real part of φ_t and φ_b, by matching the energy to the energy of the exact solution at large q [22]. The energy can be computed by taking time derivative of the two-point function.

[19] In this formula, as well as in (2.111), one factor of Δ originates from the dimension of the operators used to couple the two sides in our expressions for the global energy, whereas the other factor of Δ is the dimension of the operator inserted. So the size of a more general operator O of dimension Δ_O would be $S_\beta(O) \propto \Delta\Delta_O$.

before it enters the near-Rindler region.

In fact, one can directly see this exponential growth from the commutation relation $[P, E-B] = i(E-B)$.[20] We consider the operators $\tilde{\psi}(\rho) = e^{-i\hat{P}(\rho-\rho_b)}\psi e^{i\hat{P}(\rho-\rho_b)}$.

$$\begin{aligned}\frac{d}{d\rho}S_\beta\left(\tilde{\psi}(\rho)\right) &= \frac{\Delta^2}{2\alpha_s}\frac{\beta\mathcal{J}}{2\pi}2\,\langle \mathsf{TFD}|\,\tilde{\psi}(\rho)[i\hat{P}, \hat{E}-\hat{B}]\tilde{\psi}(\rho)\,|\mathsf{TFD}\rangle \\ &= -S_\beta(\tilde{\psi}(\rho))\end{aligned}$$

Part of the message of this chapter is that the symmetry structure in (2.14) (2.15) is giving us the correct "target" generators that should be reproduced by the microscopic analysis. Moreover, to the extent that we have derived the Schwarzian theory from SYK, we have also derived the existence of these generators from the SYK model.

In [55] Susskind discusses the following relation between the momentum of an infalling particle and the complexity of a precursor preparing this perturbation

$$P \sim \frac{d}{du_r}\mathcal{C}\ . \tag{2.114}$$

The complexity-volume conjecture relates the length of the wormhole with the state complexity [56]. Using this we see that (2.114) follows from (2.36).

2.5.3 Analogy to the Rindler (and AdS-Rindler) decomposition of a higher dimensional field theory

In section 2.3.3 we discussed the construction of approximate symmetry generators. One generator, $\hat{B}$ in (2.63), is the sum of an operator defined purely on the left and one purely on the right. This generator acts within the entanglement wedges of each of the two sides. On the other hand, the operator (2.69) is the sum of pieces that act on each side separately plus a term containing a product of left and right operators. This might seem exotic for a symmetry generator.

Here, we will point out that this type of structure is actually what we get in an ordinary quantum field theory when we use Rindler coordinates. The analogy is most clear if think of a quantum field theory on a spatial sphere. We then divide the sphere in two halves. We then automatically have a generator like $\hat{B}$. This is just the full modular Hamiltonian of the hemisphere (or a solid ball). The full state of the quantum field theory is exactly the thermofield double for the modular energy. In

[20] Here we only consider small excitations on the throat so to a good approximation we can assume the $SL(2)$ algebra holds. See section 2.3.3.

the bulk, each entanglement wedge looks like the outside of a hyperbolic black hole [57].

Now we can consider a generator such as the global time translation Killing vector of the full system. This generator is exactly related to an integral of the form

$$E = \int_{\text{Sphere}} T_{00} = E_l + E_r + \sum_i \phi_l^i \phi_r^i \ , \qquad E_l = \int_{l-\text{Hemisphere}} T_{00} \tag{2.115}$$

where E_r is the integral over the right hemisphere and $\phi_l \phi_r$ is related to the stress tensor on exactly the boundary between the two hemispheres. In order to make it more manifest, we could regularize the theory using a lattice so that there is no lattice point precisely at the boundary. However, there is a term in the full global Hamiltonian which acts on the sites that are to the left and the right of the boundary and these terms have the structure of an operator on the left and one on the right side . The sum over i is over all fields of the boundary theory[21].

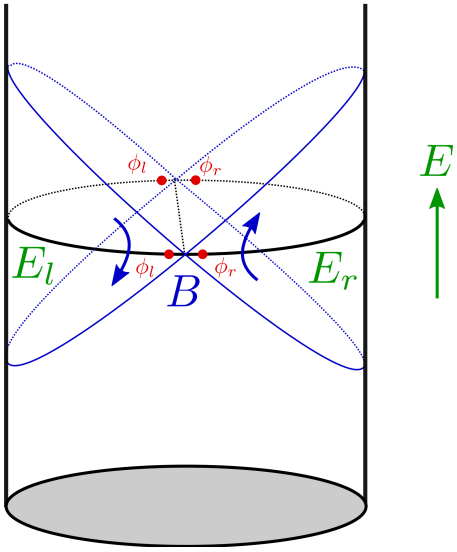


Figure 2.11: We show a two dimensional boundary theory on a cylinder. We divide the cylinder in two parts, a left and right part. The full modular Hamiltonian corresponds the the boost like generator B, which also acts like a boost in the bulk near horizon region. The global time translation generator E is the sum of three parts. E_l and E_r are completely contained in each region. The third term involves a product of operators in the two regions, depicted here in red.

So we see that the structure of $\hat{E}$ in (2.69) is very reminiscent of what we have in an ordinary QFT when we split it into two parts and we consider the generator E. Instead of $H_r + H_l$ (which here would be $B_l + B_r$), we have two other operators that are defined purely on the left and the right in (2.115).

[21] If we had a gauge theory we can do the splitting by introducing extra boundary charges (in an entangled state) and we have the same structure [58].

The fact that there are terms that act on both the left and the right is crucial to generate the appropriate isometry which can be used to transfer information between the left and right system. In fact, this is the teleportation operator introduced in [26]. Note also that in (2.115) we naturally have many operators coupling the left and the right because we have many fields in a CFT that is dual to a weakly coupled gravity system. In fact, we can view a particle moving in the gravitational bulk from one side to the other as a particular example of the traversable wormhole discussion in [26], but applied to the hyperbolic black holes.

However, one important difference between the higher dimensional case and the two dimensional gravity theory is that in higher dimensions we can define an operator E, as in (2.115), that is part of an *exact* $SO(2, D-1)$ conformal algebra. In our case, we only had an approximate expression.

Notice that in the CFT problem, even if all we cared about was the theory on $R \times H_{d-1}$, or the theory on the hemisphere, then when we consider the thermofield double of that system we generate an entangled state which, in the bulk, could be extended beyond the union of the two entanglement wedges. This region can be easily explored by evolving the system with the energy generator E (2.115).

2.6 Discussion

2.6.1 Measuring distance

In the context of a theory of gravity plus matter the distance between the two boundaries look like a reasonable observable (at least if we ignore the effects of topology change, see [38]).

On the other hand, from the point of view of the holographic boundary theory (the quantum mechanical theory with a finite number of degrees of freedom), this distance is not an obvious operator. We will not attempt to give it a precise meaning in the boundary theory here. It was proposed in [56, 59–61] that this distance is related to "complexity".

It will suffice to note that we can define it by considering correlation functions of operators. Namely, one can consider correlation functions of operators across the two boundaries, such as

$$\langle O_l O_r \rangle = \frac{1}{(-X_l.X_r)^{\Delta}} \tag{2.116}$$

This is a correlation function on the vacuum and by measuring it, we can indirectly infer the distance [62]. However, if we define the distance in terms of this operator, then we could change it by changing the state of this field. For example, we can act with an operator that creates a highly correlated

state between left and right values of the field. We can avoid this by imagining that we have many fields, then we can take the average of the correlators of many fields. In that case, changing just one field is not enough to change the distance. But it could be that, if we changed all N of the fields, then we could generate a new geometry. For this to work, the number N has to be comparable to S_0, the extremal entropy[22]. In the SYK we indeed have N fields and an $S_0 \propto N$, so that this method makes sense. This simple method of measuring distances works in the limit (2.95). In Appendix 2.13, we explore how to measure distance in the opposite regime: a bulk theory with only one free scalar field.

As a qualitative side comment we could say that by increasing the correlations for just one field, we are introducing a small microscopic wormhole, which could become macroscopic when we correlate all N fields.

A second problem with the definition of distance in terms of correlators is that if we have bulk matter, we can have interactions with matter. If we have such interactions, the correlators could change in the presence of bulk matter and our definition of distance will change. In the particular case of SYK model, these self interactions go like $1/N$, and are suppressed in the limit (2.95). Therefore, in this strict limit this definition makes sense.

In more general NAdS_2 theories, it is more challenging problem to define this distance, and we will not attempt to do it here. By more general theories we mean, for example, ones with a small number of matter fields where these fields are self interacting.

A potentially promising framework for understanding the distance between the two boundaries is the following[23]. In many body physics, there are some properties that are not defined for single particles but emerge when we have many particles. An example is the phase of a superconductor (by a "superconductor" here we just mean a system of interacting fermions with a U(1) global symmetry that is spontaneously broken). This phase is good classical variable in the large N limit, but it is not well defined for a small number of fermions. In our case, we have a rather similar variable which is the relative time shift between the two sides of the thermofield double. This relative time shift behaves as a classical variable and it is indeed one of the classical variables of an empty wormhole [37, 38]. It can be measured by considering correlation functions, in the same way that the phase of a superconductor can be measured looking at fermion two point functions. One subtle point for our discussion is that in our case this is both spontaneously and explicitly broken. In other words, if we evolve by $H_l + H_r$ we increase the value of this time shift. This is analogous to adding a term

[22] Ideas for measuring distance were mentioned in [62].
[23] See a related discussion in [63].

to the superconductor Hamiltonian that is proportional to the U(1) charge; the phase would then move linearly with time. We think that we can call this a "time superfluid" in the sense that the overall time translation symmetry is broken by the wormhole or thermofield double state. In fact, we can think about any state that has a classical time dependence as a "time superfluid" in the sense that the time translation symmetry is broken. The wormhole seems special because this time translation is also related to other symmetries of the problem, the approximate SL(2) symmetries of the quantum theory. So another very close to flat direction is the temperature. Namely, states with different temperatures are also related. They are connected by acting with an overall dilation the system. This affects the wormhole by making it longer. In fact, in the wormhole both the time and space directions are related, so the state should be more fittingly called a "space-time superfluid". We hope to expand on these remarks in a future publication.

2.6.2 Conclusions and open questions

In this chapter we have studied the symmetries near the horizon of a black hole. The main reason to study them is that these allow us to move into the region behind the horizon, in the sense discussed in section 2.4.2. So a thorough understanding of these symmetries is crucial for understanding how the interior region emerges in the full quantum theory. We have considered near extremal black holes because in this case we have a connection to the SYK model. We considered a scaling limit where the temperature becomes very small and the extremal entropy very large, but the near extremal entropy remains fixed. This physically means that we are keeping the scrambling time fixed in units of the temperature and quantum gravity effects are important, but calculable. But topology changing effects are negligible. We have defined three SL(2) generators (2.30) (2.31). They act on both the matter degrees of freedom and the boundary graviton degrees of freedom. If we have a state with fixed values of the boundary positions, the generators do not move the boundaries but move the matter relative to the boundaries as in figure 2.2. These generators act on physical states of the theory grouping them into SL(2) representations. In particular, they imply that the number of states for the wormhole is infinite. This is not a contradiction because we are working in a limit where the extremal entropy is infinite. The non-trivial aspect is that we are keeping the scrambling time finite. So we are making the non-trivial statement that the generators are well defined even after the scrambling time. These generators do not commute with the Hamiltonian, so we cannot call them "symmetries", though from the point of view of a bulk observer made out of matter they act pretty much like symmetries. The total Hamiltonian changes the generators because it changes

the boundary positions. Fortunately the evolution of the boundary positions is a solvable quantum problem so that, in principle, we can undo it. This allows us to define time independent (conserved) charges as in (2.39).

We have also expressed the generators in terms of the distance between the left and right boundaries and their time derivatives. In this way we obtain expressions that depend purely on the boundary and its dynamics. (Of course, this is related to matter via the constraints.) These formulas also allow us to express the generators in terms of correlation functions of operators, after making some extra assumptions. These assumptions are necessary to relate correlation functions to distances between the boundaries, even in the presence of matter. We have assumed that we have a number of fields scaling like N and that their interactions with the matter we are probing goes as $1/N$. These conditions are met in the SYK model.

In the semiclassical limit, where $\Delta S = \mathfrak{s} \gg 1$, we can further simplify the expressions for the generators and write them in terms of products of operators on the two boundaries. These are similar to the operators that appear in the traversable wormhole discussion [26]. This is not a coincidence, it is because these operators generate bulk matter displacements across the horizon. These displacements also play a role in the growth of chaos or in out of time order correlators. We have also shown that these operators approximately act like SL(2) transformations on the boundary time u, so that they can be viewed as conformal[24] transformations of the boundary theory. This also allows us to define an operator-state mapping. It relates operators inserted during Euclidean evolution by $\beta/2$ to states on the wormhole. These approximate symmetry generators depend explicitly on the temperature of the background wormhole we are expanding around.

We should emphasize that our discussion is valid both for a nearly-AdS_2 gravity theory with a large number of fields or the SYK model. In both cases, we can repeat all the steps of this construction. And the discussion is valid in the large N limit with finite $N/(\beta J)$. This means that scrambling effects are included, but not effects due to the finite distance to the boundary (finite βJ), or topology changing finite N effects. As an example, we have computed the growth of the charges before and beyond the scrambling time in section 2.3.2 and found that beyond the scrambling time the momentum or energy of an excitation does not continue to grow. This saturation is related to the fact that out of time order correlators decay to zero. In fact, when we evaluate the charges on a state that is created by inserting operators during euclidean evolution, we are computing the same type of correlation functions that appear in out of time order correlators. These out of time order correlators have some pieces that display an exponential growth related to chaos. It is also

[24]We are talking about three generators, and not the infinite dimensional group of reparametrizations.

possible to define the approximate generators in such a way that they depend only on these growing pieces. Of course, this is related to the well known fact that such pieces are related to shock wave scattering, which in this two dimensional context, generate simple bulk displacements [39, 40]. Here we are inverting the logic and saying that we can use these growing correlators to define symmetry generators. It is tempting to say that in any system with maximal chaos we can define translation generators that, together with the boost generator obey an approximately Poincare algebra near the horizon. For near extremal black holes they obey an SL(2) algebra, which is a bit more constraining. The difference is just whether $[E, P]$ is zero or $[E, P] \propto B$. Even in the SL(2) case, if we consider excitations very close to the horizon, at distances smaller than the radius of AdS_2, we find that E and P become relatively large, so that after a rescaling we obtain a Poincare looking algebra. It would be nice to understand the conditions under which this structure emerges in a general maximally chaotic large N system.

The approximate generators are related to the "size" operator that has been studied recently [28–31]. This connection to a symmetry algebra explains many of its features and helps to clarify its relation to bulk dynamics in AdS_2. A particularly simple relation is that the time derivative of size is the same as the bulk momentum. This is explored further in [55]. In [30] a direct SYK analysis at large q and large N showed how size evolved up to the scrambling time. Presumably, similar methods could explain why it saturates after the scrambling time. In other words, one would like to have a better microscopic picture for these generators in the SYK model, one which does not go through the usual route of the G, Σ action but rather constructs them more manifestly in terms of the UV operators of the model, as in [30].

There are some straightforward looking generalizations of this discussion, for example one could consider a supersymmetric system and a super-Schwarzian.

It would be interesting to understand how this story is modified when the length of the throat is not infinite (or βJ is finite) and eventually covering the case of a generic finite temperature black hole. We know that the chaos correlators are also given in terms of simple displacements in this case too. So we expect that they should also be useful for constructing the local 2d Poincare symmetry near any horizon (the two dimensions are time and the radial direction). We have also recalled the general expectation that the inner horizon would have some kind of singularity. It would be interesting to understand it better, and it is likely that these finite throat length corrections are relevant.

2.7 Appendix: SO(3) analogy

Consider a non-relativistic particle in Euclidean space. We have the position operators and the angular momenta:

$$[J_i, J_j] = i\varepsilon_{ijk}J_k, \qquad [J_i, X_j] = i\varepsilon_{ijk}X_k, \qquad [X_i, X_j] = 0, \qquad i,\, j,\, k = 1,\, 2,\, 3. \tag{2.117}$$

This is the Euclidean algebra $E(3)$. (Note that what is normally the momentum generator, P^i, is here the position operator). This algebra has two Casimirs,

$$X^2 = r^2, \qquad J.X = \lambda. \tag{2.118}$$

Our toy model consists of two non-relativistic particles, which we dub "left" and "right," each constrained to live on the surface of the sphere $r^2 = 1$. We assume for simplicity that our non-relativistic particles do not carry intrinsic spin, so $\lambda = 0$. In addition, there is some "matter" which can carry angular momentum. This matter could for instance be some spinning particle, which lives at the center of the sphere. (Later we will see that the matter has to have integer spin.) Finally, we demand that the overall system has vanishing angular momentum $J_l + J_r + J_m = 0$; e.g., the overall SO(3) symmetry is gauged. Thus the Hilbert space of our system may be denoted

$$\mathcal{H} = (\mathcal{H}_l \otimes \mathcal{H}_r \otimes \mathcal{H}_m)/SO(3). \tag{2.119}$$

This also means that all physical operators O must commute with the total angular momentum, $[J_l^a + J_r^a + J_m^a, O] = 0$.

2.7.1 Exact SO(3) algebra from two copies

Armed with this Hilbert space, we may consider the physical operators

$$X_l.J_r, \quad X_r.J_l, \quad (X_l \times X_r).\,(J_l + J_r)\,. \tag{2.120}$$

Note that we can $\lambda = 0$ and the gauge constraint to rewrite these as

$$G^1 = -X_l.J_m, \quad G^2 \sim -X_r.J_m, \quad G^3 \sim -\,(X_l \times X_r)\,.J_m. \tag{2.121}$$

This suggests that the 3 operators form an SO(3) algebra. (More precisely, we should replace X_r with the Gram-Schmidt linear combination of X_l and X_r that is orthonormal to X_l.) We can check this by computing their commutators directly in the 2-body Hilbert space without the gauge constraint. To do so, it is worth introducing a little bit of notation. Let V and W transform be vector operators under SO(3) such that

$$\begin{aligned} &[V_i, V_j] = [W_i, W_j] = [V_i, W_j] = 0 \\ &V.V = W.W = 1, \\ &V.W = 0. \end{aligned} \tag{2.122}$$

Then defining $e^1 = V, e^2 = W, e^3 = V \times W$, we may write $G^A = -e_i^A J_i$ where $J = J_l + J_r$. Here the capital indices are physical, and the lower indices are the gauge index.

$$\begin{aligned} [G^A, G^B] &= [e_i^A J_i, e_j^B J_j] \\ &= i\epsilon_{ijk}\left(e_i^A e_k^B J_i + e_j^B e_i^A J_k + e_j^B e_k^A J_i\right) \\ &= -i(e^A \times e^B).J \\ &= i\epsilon^{ABC} G_C, \end{aligned} \tag{2.123}$$

where in the last line, we used $e^A \times e^B = \epsilon^{ABC} e_C$. Note that in order to construct this algebra, it was crucial that we had other vector operators that were not just the angular momentum operators. The left and right Hilbert spaces were representations of the Euclidean algebra, and not just the SO(3) algebra, so we are necessarily considering an infinite dimensional Hilbert space.

It is also interesting to construct the Casimir operator

$$C = G^A G^A = e_i^A J_i e_k^A J_k = \delta_{ik} J_i J_k = J^2. \tag{2.124}$$

So we see that the Casimir of the gauge charges is equal to the Casimir of these physical charges.

Note that evolving with one of the G^A leads the X_l and X_r vectors to precess about some axis. Alternatively, if we replace $J_l + J_r$ with $-J_m$, we can view the positions of the particles as fixed under time evolution, but the matter rotates. The invariant statement is that the generators move the left and right particles with respect to the matter.

2.7.2 Action on states/operators

We would now like to use the physical algebra to organize states and operators in this theory. Let us imagine that the matter sector contains a vector operator W^a which is not J^a. For example, if the matter is actually another particle, W^a could be its position vector Y^a. We may form the gauge invariant operator $W^A = W^a e_a^A$. Then $[W^A, G^B] = i\epsilon^{ABC} W_C$. Furthermore, we can use W^A to construct tensor operators. So we can generate operators which transform in any integer spin representation of the physical $SO(3)$.

Now consider a state $|0\rangle$ that is in a singlet under the gauge SO(3) and the physical SO(3)

$$|0\rangle = \sum_{m,j} \psi(j)(-1)^m |j,m\rangle_l \otimes |j,-m\rangle_r \otimes |0,0\rangle_{\text{matter}} . \tag{2.125}$$

Here $\psi(j)$ can be any normalizable function. For example, the thermofield double corresponds to $\psi \sim e^{-\beta j^2/4}$. We may generate states which transform under integer spin by repeatedly applying W^A, e.g., $W^A W^B W^C |0\rangle$. We can then organize these into irreps $|j,m\rangle$ of the physical algebra.

One interesting question is: what matter states are allowed in this gauge theory? Since there are no half-integer states in the two-particle Hilbert space, we conclude that the matter is restricted to be integer spin. For example, the matter cannot be a spin 1/2 qubit.

2.8 Appendix: Canonical quantization of Schwarzian theory

The purpose of this section is to show that the most naive quantization of the Schwarzian theory explicitly realizes the commutation relations needed in our construction of the charges. We will quantize using global time coordinates $T_r(u)$, although one could also use Rindler coordinates related by $\tan(T_r/2) = \tanh(t_r/2)$. **Friendly warning**: when using global time, we adopt different embedding space conventions from the rest of the paper. Defining $\mathcal{X}^0 = X^1$ and $\mathcal{X}^\pm = X^{-1} \pm iX^0$, we may write

$$\begin{aligned} X_r = (\mathcal{X}_r^0, \mathcal{X}_r^+, \mathcal{X}_r^-) &= \left(\frac{1}{T_r'}, -\frac{e^{iT_r}}{T_r'}, -\frac{e^{-iT_r}}{T_r'} \right), \\ X_l &= \left(-\frac{1}{T_l'}, -\frac{e^{iT_l}}{T_l'}, -\frac{e^{-iT_l}}{T_l'} \right). \end{aligned} \tag{2.126}$$

The metric in these coordinates is $X.Y = \mathcal{X}^0\mathcal{Y}^0 - \frac{1}{2}(\mathcal{X}^+\mathcal{Y}^- + \mathcal{X}^-\mathcal{Y}^+)$. The advantage of this convention is that the components of X are simple when written in terms of global time, and furthermore (2.9) becomes $[\mathcal{Q}_m, \mathcal{Q}_n] = (m-n)\mathcal{Q}_{m+n}$, where m, n take values from $\{0, +, -\}$. This

convention is perhaps more familiar from $d \geq 2$ CFT, where L_0 generates dilation, or global time translation along the cylinder.

Let's proceed with the canonical quantization. We start with the Lagrangian on one side, setting $T = T_r(u_r)$:

$$L = -\{\tan T/2, u\} = \frac{1}{2}\frac{T''^2}{T'^2} - \frac{1}{2}T'^2 - \left(\frac{T''}{T'}\right)'. \tag{2.127}$$

Dropping the overall derivative, we search for a 4-dimensional phase space with 2 canonical coordinates T, T'. The canonical momentum are $p_i = \sum_{k=i}^{2}(-\partial_u)^{k-i}\partial L/\partial T^{(k)}$:

$$p_1 = -T' + \frac{T''^2}{T'^3} - \frac{T'''}{T'^2}, \quad p_2 = \frac{T''}{T'^2}. \tag{2.128}$$

Note that p_1 is just the global T gauge charge $p_1 = \mathcal{Q}^0$. Then the Hamiltonian is

$$H = \sum p\dot{q} - L = -\frac{1}{2}T'^2 + \frac{3T''^2}{2T'^2} - \frac{T'''}{T'} = -\{\tan T/2, u\}. \tag{2.129}$$

We then impose canonical commutation relations

$$[T, T'] = 0, \quad [T, p_1] = i, \quad [T', p_2] = i. \tag{2.130}$$

We would like to check that the gauge charges satisfy an $\mathrm{SL}(2,\mathbb{R})$algebra. The other two charges are

$$\mathcal{Q}^{\pm} = e^{\pm iT}\left(\pm i\frac{T''}{T'} + \frac{T''^2}{T'^3} - \frac{T'''}{T'^2}\right) \tag{2.131}$$

Now to compute these commutators, we first rewrite these charges in terms of coordinates and momenta by eliminating T''', T'' in favor of p_1, p_2.

$$\mathcal{Q}^{\pm} = e^{\pm iT}\left(T'(1 \pm ip_2) + p_1\right) \ , \qquad \text{and} \qquad \mathcal{Q}^0 = p_1. \tag{2.132}$$

In writing these expressions, we have chosen an operator ordering where all the momenta are to the right of the coordinates. Note that these charges generate the expected infinitesimal transformations

on the times and satisfy the $\mathrm{SL}(2)_g$ algebra:

$$\begin{aligned} &i[\mathcal{Q}^0, T] = 1, && i[\mathcal{Q}^0, T'] = 0, \\ &i[\mathcal{Q}^\pm, T] = e^{\pm iT}, && i[\mathcal{Q}^\pm, T'] = \pm ie^{\pm iT}T', \\ &[\mathcal{Q}^0, \mathcal{Q}^\pm] = \pm\mathcal{Q}^\pm, && [\mathcal{Q}^+, \mathcal{Q}^-] = -2\mathcal{Q}^0. \end{aligned} \tag{2.133}$$

With these gauge charges, H is simply the Casimir

$$H = -\frac{1}{2}Q.Q. \tag{2.134}$$

One can check that, e.g., the equation of motion correctly links the two coordinates $i[H, T] = T'$.

Now consider the embedding space vectors in (2.126). We see that the X's are only a function of the coordinates and not the momenta. So we immediately conclude that all components of X commute amongst themselves. We can also check that $[Q^a, X^b] = i\epsilon^{abc}\eta_{cd}X^d$ transforms like a vector. This verifies that the Schwarzian realizes the Poincare algebra postulated in (2.9).

One advantage of using the Hamiltonian formalism is that expressions for operators written in terms of T, T' and p_1, p_2 are fairly agnostic about what time-evolution rule is being used. For example, we may evolve by $H_l + H_r$ or the coupled Hamiltonian $H_l + H_r + \eta(X_l.X_r)^{-\Delta}$. This is not the case when we write expressions in the Lagrangian formalism because the expressions for Q^a change (for example, they depend on η in (2.65)).

As another application of the Hamiltonian formalism, we may check that the approximate construction of the charges discussed in section 2.3.3 does not lead to a closed algebra. In other words, the failure of the algebra happens not just at the microscopic level (in terms of the fermions) but even in the Schwarzian limit. This is not too surprising because as discussed in [27] and in section 2.4.1, the spectrum of the eternal traverable wormhole (even in the Schwarzian limit) contains a tensor factor which is not $\mathrm{SL}(2, \mathbb{R})$invariant.

2.9 Appendix: Evaluating commutators perturbatively

One might wonder whether we can check the commutator computation using standard perturbation theory. In this section, we show how this may be done. We focus on checking the commutation relations between the gauge charges; the commutation relations between the physical charges G^A can be checked quite analogously.

We start by reviewing the linearized Schwarzian action [17]. We use the same variables as in Section 2.3.1. We write the Schwarzian as

$$I = \frac{\mathfrak{s}}{2}\int d\tilde{u}\left(\epsilon''^2 + \epsilon'^2\right) = \frac{\mathfrak{s}}{2}\int d\tilde{u}\left(-r'^2 - r^2 + q'^2\right). \tag{2.135}$$

The first line is the same expression as in (2.55), except now in Lorentzian signature. We have rewritten this higher derivative action by introducing the fields q and the ghost field r such that $q = -\epsilon'' + \epsilon$, $r = \epsilon''$. From this action we can read off the commutation relations

$$[q, q'] = \frac{i}{\mathfrak{s}}, \quad [r, r'] = -\frac{i}{\mathfrak{s}}. \tag{2.136}$$

Now as in (2.43), we expand Q in powers of ϵ, and then substitute for q, r. To quadratic order, we find

$$\begin{aligned} Q^{-1} &\simeq \mathfrak{s}\left(-1 - q' - (r^2 + r'q' + r'^2)\right) \\ Q^{+} &\simeq \mathfrak{s}\left(r - r' + q(r - r') + r(q' - 2r') + 2r'(q' + r') + 2r^2\right) \\ Q^{-} &\simeq \mathfrak{s}\left(-r - r' + q(r + r') + r(q' + 2r') + 2r'(q' + r') + 2r^2\right). \end{aligned} \tag{2.137}$$

Then we can verify that $[Q^{-1}, Q^{\pm}] = \pm iQ^{\pm}$ and $[Q^{+}, Q^{-}] = 2iQ^{-1}$ in line with (2.9). To reproduce the linear terms in $Q^{\pm}$ on the RHS, it is important to expand to quadratic order the charges in the commutator.

Note that in order to check the algebra at higher orders, one should not only expand Q to higher powers but also modify the commutation relations due to higher order terms in the action. This quickly gets cumbersome; hence the purpose of the previous section.

2.10 Appendix: Spinor description of the boundary variables

We have described the boundary motion in terms of a vector X^a, with $X^2 = 0$. Alternatively, we can use a two component spinor λ_α and construct X^a as

$$X^a = (\sigma^a)^{\alpha\beta}\lambda_\alpha\lambda_\beta, \tag{2.138}$$

where σ^a are Pauli matrices with one index raised, $(\sigma^a)^{\alpha\beta} = \epsilon^{\alpha\gamma}(\sigma^a)_\gamma{}^\beta$, and $\sigma^{-1} = (\sigma^{-1})_\gamma{}^\beta = \hat{\sigma}_3$, $\sigma^0 = -\hat{\sigma}_1$ and $\sigma^1 = i\hat{\sigma}_2$ where $\hat{\sigma}_i$ are the three usual Pauli matrices. The inner product between

two spinors is antisymmetric $\langle \lambda, \chi \rangle = \epsilon^{\alpha\beta}\lambda_\alpha \chi_\beta$. This implies that we automatically obey $X^2 = 0$, since we cannot build any non zero $SL(2)$ invariant from a single λ. The evolution equation is then

$$\dot{\lambda} = \mathcal{Q}\lambda \ , \qquad \ddot{\lambda} = (\mathcal{Q})^2\lambda \propto E\lambda \tag{2.139}$$

where $\mathcal{Q} = Q_a\sigma^a$. We could normalize the spinor by a condition of the form $\langle \dot{\lambda}, \lambda \rangle$ =constant so that $\dot{X}^2 = -1$, this also ensures $Q.X = 1$. This gives a simple expression for the evolution. Also, in terms of left and right spinors then the three generators have expressions proportional to

$$\tilde{P} \propto \frac{\langle \lambda_l, \mathcal{Q}_m \lambda_r \rangle}{\langle \lambda_l, \lambda_r \rangle} \ , \qquad \tilde{E} + \tilde{B} \propto \frac{\langle \lambda_l, \mathcal{Q}_m \lambda_l \rangle}{\langle \lambda_l, \lambda_r \rangle} \ , \qquad \tilde{E} - \tilde{B} \propto \frac{\langle \lambda_r, \mathcal{Q}_m \lambda_r \rangle}{\langle \lambda_l, \lambda_r \rangle} \tag{2.140}$$

2.10.1 Computation of the charges beyond the scrambling time

As an application of this formalism, we will consider the setup in section (2.3.2). We consider a background solution where $\mathcal{Q}_r = \sigma^3/2$ (we have set $\beta = 2\pi$) and

$$\lambda_r(0) = (i, i)/\sqrt{2} \ , \qquad \lambda_l(0) = (1, -1)/\sqrt{2} \tag{2.141}$$

Then, it is easy to find the time evolution, $\lambda_r(t) = e^{u\mathcal{Q}_r}\lambda_r(0) = i(e^{u/2}, e^{-u/2})/\sqrt{2}$, for example. We want to insert a perturbation. If we inserted the perturbation at zero time, we would expect it to change Q_r. The charge will be both rotated and rescaled. The rescaling is negligible in the limit $\omega \to 0$, see (2.61). So the main effect is a small infinitesimal rotation

$$\mathcal{Q}_r \to \mathcal{Q}'_r = L\mathcal{Q}_r L^{-1} \ , \qquad L \sim 1 + \vec{\gamma}\vec{\sigma} \tag{2.142}$$

where γ is a vector of size of order ω. If we insert this same excitation at some early time, then we should conjugate it by the time evolution,

$$L(u_0) = e^{u_0\mathcal{Q}_r} L e^{-u_0\mathcal{Q}_r} \sim 1 + 2\hat{\alpha}\sigma_- \ , \qquad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{2.143}$$

where $\hat{\alpha} \propto \omega e^{-u_0}$ and have kept only the terms that are finite in the limit $\omega \to 0$ and $u_0 \to -\infty$ with ωe^{-u_0} fixed. This implies that

$$\mathcal{Q}_m = -\delta\mathcal{Q}_r = \mathcal{Q}_r - L(u_0)\mathcal{Q}_r L(u_0)^{-1} = -2\hat{\alpha}[\sigma_-, \mathcal{Q}_r] \tag{2.144}$$

Now the new value of λ_r at the origin is just given by

$$\begin{aligned}\lambda_r'(0) &= e^{-u_0 \mathcal{Q}_r'} e^{u_0 \mathcal{Q}_r} \lambda_r(0) = L(u_0) e^{-u_0 \mathcal{Q}_r} L(u_0)^{-1} e^{u_0 \mathcal{Q}_r} \lambda_r(0) = \\ &= L(u_0)\lambda_r = (1 + 2\hat{\alpha}\sigma_-)\lambda_r(0) \end{aligned} \tag{2.145}$$

$\lambda_l(0)$ continues to be as in (2.141). Inserting (2.144), (2.141), (2.145) into (2.140) we get (2.62). We also find that

$$\langle \lambda_l, \lambda_r' \rangle \propto (1 + \hat{\alpha}) \Longrightarrow (X_l.X_r') \propto (1 + \hat{\alpha})^2 \tag{2.146}$$

2.11 Appendix: OTOC correlators and the expectation values of charges

Section 6.2 of [17] considered an OTOC correlator of four operators of dimension Δ. It is also possible to do the same computation for the case that two operators have dimension Δ_1 and the other two have dimension Δ_2. Then the same method gives the answer

$$\frac{\langle V_1 W_3 V_2 W_4 \rangle}{\langle V_1 V_2 \rangle \langle W_3 W_4 \rangle} = \frac{U(2\Delta_1, 1 + 2\Delta_1 - 2\Delta_2, \frac{1}{z})}{z^{2\Delta_1}} = \frac{1}{z^{2\Delta_1}\Gamma(2\Delta_1)} \int_0^\infty dt e^{-t/z} t^{2\Delta_1 - 1} (1+t)^{-2\Delta_2} \tag{2.147}$$

Here, $z = \mathfrak{s} e^{\tilde{u}}/8$. Despite appearances this expression is symmetric under $\Delta_1 \leftrightarrow \Delta_2$. We now want to relate this to the computation in section (4.2.2). To evaluate the charges it is necessary to evaluate the distance.

We can extract the distance from (2.147) by taking the limit $\Delta_2 \to 0$, since

$$\frac{1}{(X_3.X_4)^{\Delta_2}} \sim 1 - \Delta_2 \log(X_3.X_4) = 1 - \Delta_2 \ell \tag{2.148}$$

where ℓ is the distance. By going to higher orders in Δ_2 we could get higher moments for the distance. Expanding (2.147) to first order in Δ_2,

$$\langle \ell \rangle = 2 \frac{1}{z^{2\Delta_1}\Gamma(2\Delta_1)} \int_0^\infty dt e^{-t/z} t^{2\Delta_1 - 1} \log(1 + t). \tag{2.149}$$

This has the following expansion for small and large z

$$\langle \ell \rangle \sim 4\Delta_1 z \ , \quad z \ll 1 \ , \qquad \langle \ell \rangle \sim 2 \log z \ , \qquad z \gg 1 \tag{2.150}$$

(2.149) is the general answer. But we are interested in the classical limit, which corresponds to the case where $C \gg 1$. We are also interested in computing this for a state that has definite energy ω.

From classical computations, we expected the distance to have the form

$$\langle \ell \rangle = 2 \log[1 + 2\Delta_1 z] \tag{2.151}$$

But this answer is supposed to hold for states with definite energy $\tilde{E}$. One way to get the energy to be definite is to imagine that Δ_1 is large, then we could imagine creating states with definite energy, of order Δ_1 times the appropriate redshift factors, which would lie inside z. In this case we can do the integral (2.149) by saddle point. If we ignore the logarithm, the saddle point and the integral around it gives the gamma function. Again, ignoring the log, the saddle point is at

$$\frac{1}{z} = \frac{2\Delta_1}{t_s} \to t_s = 2\Delta_1 z \tag{2.152}$$

Then we see that we get the expected answer (2.151) by inserting this value of t_s into the log term, $\log(1 + t)$, in the integral (2.149).

It is also clear that, to the extent that we can use this saddle point evaluation of this second term, in (2.147), we will get that the distance behaves classically, with the value (2.151). To check whether the saddle point is valid, we can look at the second derivative of the exponent at the saddle which gives

$$\partial_t^2(-t/z + 2\Delta_1 \log t) = -\frac{2\Delta_1}{t_s^2} \tag{2.153}$$

So we see that the spread around the saddle is $\delta t \sim t_s/\sqrt{\Delta_1}$ which is always smaller than t_s, which justifies the saddle point approximation. The conclusion is that the large Δ_1 limit of the OTOC computation in [17] reproduces the expression for the distance used in previous sections.

2.12 Appendix: Stiffer traversable wormholes

In this section, we consider other alternative to the global energy generator $\tilde{E}$.

In the Maldacena-Qi wormhole, the spectrum at low energies includes both SL$(2, \mathbb{R})$excitations and excitations of the boundary. Both the energies are $\sim T'$. We would like to modify the eternal wormhole so that the boundary excitations become very heavy. Then we can see the SL$(2, \mathbb{R})$spectrum just by going to low energies.

In the Maldacena-Qi eternal wormhole, one can determine the frequency of oscillation of the

boundary graviton by writing down an effective action for the variable $\phi = \log T'$ in the Schwarzian approximation. Assuming no bulk matter, the gauge constraints simplify the action so that ϕ becomes the coordinate of a non-relativistic particle in a potential $V(\phi)$:

$$V(\phi) = e^{2\phi} - \sum_{\Delta} \eta_\Delta e^{2\Delta\phi}. \tag{2.154}$$

The first term is from the Schwarzian, the second term is from the Maldacena-Qi interaction. We are considering a slight generalization of their interaction that could arise, if there are matter fields with different dimensions in the bulk, or by imposing more complicated boundary conditions on the matter fields. For the SYK model, we could generate this interaction by not just adding a term $\sim i\eta \left(\sum_i \psi_l^i \psi_r^i\right)$, but also adding an interaction $\sim \frac{\rho}{N}\left(\sum_i \psi_l^i \psi_r^i\right)^2$. We want to know if the frequency can be made large by a judicious choice of η_Δ. Note that the validity of this effective potential requires $\eta_\Delta << 1$ and $\phi_{\min} \ll 0$. Let us first consider a minimal extension:

$$V = e^{2\phi} - \eta e^{2\Delta\phi} + \rho e^{2\tilde{\Delta}\phi}. \tag{2.155}$$

To simplify the algebra a little, we start with $\tilde{\Delta} = 2\Delta$. Now let us ignore the first term (from the Schwarzian) and see if we can find a solution where this is a valid approximation. We have

$$\phi_m = \frac{1}{2\Delta}\log(\eta/2\rho), \qquad T' = \left(\frac{\eta}{2\rho}\right)^{\frac{1}{2\Delta}}, \qquad \left(\frac{\omega}{T'}\right)^2 = 2^{\frac{1}{\Delta}}\Delta^2\rho\left(\frac{\rho}{\eta}\right)^{\frac{1}{\Delta}-2}. \tag{2.156}$$

For some fixed value $0 < \Delta < 1/2$, taking $\eta/\rho \ll 1$ while keeping ρ small and fixed, we can make $\phi_m \ll 0$ and make the frequency very large. In order for it to be a good approximation to drop the first term, we need

$$2^{-1/\Delta}\left(\frac{\eta}{\rho}\right)^{1/\Delta} \ll \frac{\eta^2}{2\rho} \tag{2.157}$$

Thus in the limit where $\eta/\rho \ll 1$ and ρ is small and held fixed, this condition will be satisfied for $0 < \Delta < 1/2$. In the SYK model, $\tilde{\Delta} = 2/q$ and $\Delta = 1/q$.

Another question we can ask is: how many bound N_b states exist for this spectrum? We can

compute this in the WKB approximation, with $p = 2N\dot{\phi} = 2N\sqrt{V}$.

$$\begin{aligned} N_b &= \frac{1}{2\pi}\oint p\, dq = \frac{4N}{2\pi}\int_{-\infty}^{\phi_m} d\phi\, \sqrt{\eta e^{2\Delta\phi} - \rho e^{4\Delta\phi}} \\ &= \frac{2N}{\pi}\left(\frac{\eta^2}{2\rho}\right)^{1/2}\int_{-\infty}^{0} d\phi \sqrt{e^{2\Delta\phi} - \frac{1}{2}e^{4\Delta\phi}} \\ &= \left(\frac{N\rho^{1/2}}{\pi}\right)\left(\frac{\eta}{\rho}\right)\frac{2+\pi}{4\Delta} \end{aligned} \tag{2.158}$$

So we see that the number of states is large in the classical limit, although it is suppressed now by $\eta/\rho^{1/2}$. So we need this quantity to be small but not too small if we want the classical analysis to be correct.

Now if we add some matter in the bulk with energy E_{bulk}, then should set the gauge charges of the Schwarzian mode $Q_0 = E_{\text{bulk}}$. This changes the effective potential by an amount

$$V \to V + E_{\text{bulk}} e^{\phi}/N. \tag{2.159}$$

Notice that for $\Delta < 1/4$, there is a significant difference between our model and the Malda+Qi wormhole. In particular, suppose we add N particles of energy Δ. In the regime of interest, neither T' nor the frequency will change appreciably! This is completely different from the Maldacena-Qi wormhole, where the e^{ϕ} term will typically dominate over the $e^{2\phi}$ term in the potential and change both T' and ω.

It is interesting to compute the ground state energy in the presence of matter q_0: is the non-linear term in q_0 suppressed? The first order correction is $T' E_{\text{bulk}}/N$. The second order correction is obtained by correcting $\delta\phi = -T'_* E/(NV'')$. Defining the frequency $\omega^2 = V''_*/T'^2$:

$$N\delta V = ET'_* - \frac{1}{2N}\frac{E^2}{\omega^2} + O\left(\frac{1}{N^2}\right). \tag{2.160}$$

This comes from both the contribution of the original potential and the matter piece. We see that when ω is large, the correction is suppressed.

It would be interesting to analyze this model further, e.g., explore its thermodynamics, study it at higher temperatures in large-q SYK, etc. We leave this to future work.

2.13 Appendix: Measuring distance using a single free field

We suspect that the distance could be measured even if we have a small number of fields in the bulk. As a simple example, consider a single *free* bosonic field in the bulk with dimension Δ. Then we may measure the distance using considering the operator $d\Delta \sim -\log \phi_l \phi_r$. This operator can be defined by the replica trick, e.g., we consider higher-pt correlation functions and then continue to $m = 0$.

Now we would like to see how well this operator tracks the distance when there are ϕ particles around. We can consider the states defined previously in Section 2.3.2,

$$\langle \phi(u_t)| \, \phi_l^m \phi_r^m \, |\phi(u_b)\rangle = m! G_{lr}^m G_{tb} + 2m^2(m-1)! G_{lr}^{m-1} \left(G_{lt}G_{rb} + G_{rt}G_{lb}\right). \tag{2.161}$$

Here the first term comes from the "good" diagrams in Figure 2.12; the "bad" diagrams ones are contractions where one of the ϕ_l contract with the matter created from ϕ_b or ϕ_t. The factor of 2 comes from the choice of connecting one of the ϕ_l to ϕ_t or to ϕ_b. The operators ϕ^m should be regulated, for instance by point-splitting. Analytically continuing,

$$\begin{aligned} d\Delta &\sim -\lim_{m\to 0} \frac{1}{G_{tb}} \langle \phi(u_t)| \left(\frac{(\phi_l \phi_r)^m - 1}{m} \right) |\phi(u_b)\rangle \\ &= -\log\left(G_{lr}\right) + \gamma - 2\frac{G_{lt}G_{rb} + G_{lb}G_{rt}}{G_{lr}G_{tb}}. \end{aligned} \tag{2.162}$$

In the limit of large distances (e.g., when we push the bulk fields towards the boundary) or when Δ is large, the leading contribution is the first term, which is precisely the distance.

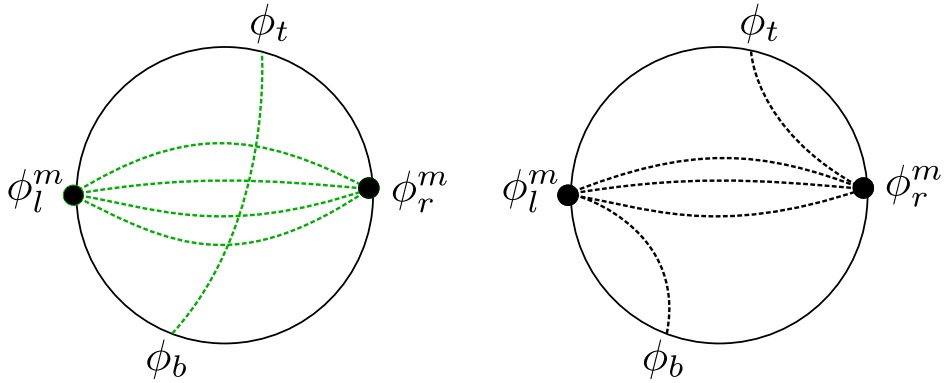


Figure 2.12: We consider the diagrams contributing to the correlator $\langle \phi(u_t)| \, \phi_l^m \phi_r^m \, |\phi(u_b)\rangle$ in (2.161). In green is the desired contribution which mimics the action of the charges in Figure 2.5(a). The "bad" diagrams (displayed in black) are morally similar to the ones in Figure 2.5(b).

As a generalization of this calculation[25], we may consider inserting not just a single particle but several particles, e.g., $\phi_b \to \phi_b^k$, $\phi_t \to \phi_t^k$. The analog of (2.162) will involve a sum over the number of bonds broken between the left and right. This will generate terms that depend on k. The k-dependence of the "bad" terms means that there is likely a bound on how many particles we can insert before the operator defined above no longer tracks the distance. This is not too surprising, for example, if we considered some coherent state with a large number of quanta, we expect that the expectation value of $\log \phi_l \phi_r$ is determined by the classical field values, which does not have to track the distance.

The preliminary conclusion is that at least for small number of quanta, the operator $-\log(\phi_l \phi_r)$ seems to be a good proxy for the distance between the two sides. We leave to future work a more thorough understanding of the limitations of this distance operator, the effects of interactions, etc.

[25]We thank Wayne Zhao for discussions about the combinatorics.

Chapter 3

Global Symmetry Violation in the Black Hole Interior

3.1 Introduction

An old conjecture states that in any theory of quantum gravity with black holes, there are no exact global symmetries [64–69]. A semiclassical treatment of black hole evaporation [70] suggests that the Hawking radiation has a thermal spectrum, independent of whatever the global charge of the matter that formed the black hole. After the black hole evaporates completely, any initial global charge is lost (up to fluctuations in the thermal ensemble). One might hope that the charge can be stored in some remnant of the black hole, but such remnants would violate entropy bounds [67] among other problems [71]. These arguments are most clear at the final stage of the evaporation, where the black hole is not big enough to store the information about the infalling charge, while they are less clear when the black hole is still large, say at the Page time of the black hole evaporation. Said differently, an experimentalist who wishes to check whether symmetry violation has occurred must gather almost all of the Hawking radiation from the black hole. If she is unable to capture the high energy modes at the endpoint of evaporation, or if she is unable to access certain fields (e.g. dark matter fields), she may not be able to conclusively convince herself that charge conservation is violated in nature.

Since we expect that the global symmetry violation happens throughout the evaporating process, it is therefore natural to ask that whether there are some quantities that can be computed with only some portion of the Hawking radiation, that are able to show explicitly the global symmetry has

been violated. (There is a strong analogy here with the Page curve [72]: unitarity implies that the entropy of *all* the Hawking radiation from a fully evaporated black hole is zero; Page famously showed that unitarity also has implications for a large subset of the radiation.) One obvious candidate is to compute correlators which would be zero if the global symmetry is exact. However, if there are no explicit symmetry violating processes in the semiclassical physics, these correlators are expected to have non-perturbatively small values $\mathcal{O}(e^{-S_{\rm BH}})$, where $S_{\rm BH}$ is the Bekenstein-Hawking entropy of the black hole. Computing these quantities will require non-perturbative knowledge beyond Hawking's description. To be clear, we are not claiming that the approximate global symmetry, such as the $B-L$ symmetry in nature, cannot be violated within the semiclassical physics. What we are imagining is the worst case scenario, namely even if the symmetry is preserved semiclassically, non-perturbative gravitational effect will eventually violate it.

Although understanding how to compute such exponentially small violations is a worthy goal for the future, the point of this chapter is to point out there exist other quantities which quantify the symmetry violation in the Hawking radiation, that can be computed just with the semiclassical knowledge, and give $\mathcal{O}(1)$ results. Our discussion is motivated by the recent progress in deriving the unitary Page curve [72] for an evaporating black hole within the semiclassical description [73, 74]. In those developments, the key tool is the gravitational fine grained entropy formula, or the quantum extremal surface (QES) prescription [75–78]. When applying the formula to compute the von Neumann entropy of the Hawking radiation after the Page time, one finds an island in the black hole interior that belongs to the entanglement wedge of the Hawking radiation [79]. In this chapter, we consider the relative entropies between different density matrices of the Hawking radiation R. More precisely, we consider the exact density matrix of the Hawking radiation $\rho_1 = \rho_{\rm exact}(R)$, and the density matrix $\rho_2 = U_R(g)\rho_{\rm exact}(R)U_R^\dagger(g)$ after applying a global symmetry transformation $U_R(g), g \in G$ to the radiation. If we form the black hole from neutral (singlet) matter, then if the global symmetry were exact, we would expect $\rho_2 = \rho_1$ since a reduced density matrix of the system should be invariant under the symmetry transformation. Thus a non-zero relative entropy $S(\rho_2|\rho_1) > 0$ is a way to quantify the degree to which the global symmetry is violated. Similar to the von Neumann entropy, the relative entropy can be computed with only the knowledge of the semiclassical physics [80], and we find it to be an $\mathcal{O}(1)$ result whenever an island arises. We demonstrate the calculation explicitly using simple examples. The relative entropy that we compute quantifies the amount of global symmetry violation in the Hawking radiation, and we discuss how it depends on various parameters of the problem. As we will see in §3.2, the relative entropy is zero before the Page time, but becomes non-zero at the Page time and grows as more radiation comes

out.

The island in the QES prescription is closely related to the replica wormhole geometry in the replica trick for deriving the entropy using the gravitational path integral [13, 14, 81]. It is thus natural to ask what roles they play in showing that the global symmetry is violated. In §3.3 we discuss the replica versions of the relative entropy, and point out that they are non-zero exactly due to charge flowing through the replica wormhole. This resonates with a long history of works anticipating that Euclidean wormholes violate global symmetries [65, 82]. In previous arguments, it was never completely clear whether wormholes should be included or not in a particular calculation; here, unitarity demands them to be included. We contrast the global symmetry case with a gauge symmetry, which is allowed because charge cannot flow through the wormhole due to the gauge constraint (Gauss's law). An exact global symmetry of a holographic theory must therefore be realized as a gauge symmetry in gravity.

We also consider in §3.2.3 a setup where the bulk relative entropy may be enhanced by creating pairs of particles with opposite charges in the bath, and waiting a scrambling time for the particles to fall into the island. This is related to the Hayden-Preskill [83] experiment, as we elaborate on in §3.4 along with other discussions and conclusions.

It is worth noting that a recent paper by Harlow and Shaghoulian [84] also connected recent developments in the black hole information problem with the violation of global symmetries, using arguments along the lines of [68, 69]. Our result is based on similar assumptions as theirs, but in some sense we take one step further by quantifying and computing the amount that an approximate global symmetry is violated by non-perturbative effects.

Note added: as we were finishing the manuscript on which this chapter is based, we learned that [85] are considering similar issues in an upcoming work.

3.2 Quantifying global symmetry violation with relative entropy

3.2.1 General argument

We start by reviewing the QES prescription and the relative entropy formula. In the context of an evaporating black hole, the QES prescription as applied to the exact density matrix $\rho_{\text{exact}}(R)$ of the

Hawking radiation states

$$S\left(\rho_{\text{exact}}(R)\right) = \text{Min}\left\{\text{Ext}_I\left[\frac{A(\partial I)}{4G_N} + S(\rho_{\text{semi}}(R\cup I))\right]\right\}, \tag{3.1}$$

where on the right hand side, one extremizes over the generalized entropy function over possible islands in the region where gravity is dynamical, and choose the minimal one if there are multiple extremums. Note that the right hand side of (3.1) is calculable within the semiclassical approximation. The region $R\cup I$ on the right hand side is called the entanglement wedge [86–88] of the region R. As first recognized in the context of the AdS/CFT correspondence [80], the validity of (3.1) among a family of "nearby" states implies an equality between the relative entropies in the exact description and the semiclassical description:

$$S\left(\rho_{\text{exact}}(R)|\sigma_{\text{exact}}(R)\right) = S\left(\rho_{\text{semi}}(R\cup I)|\sigma_{\text{semi}}(R\cup I)\right), \tag{3.2}$$

where the relative entropy $S(\rho|\sigma)$ is defined as

$$S(\rho|\sigma) \equiv \text{tr}\left[\rho(\log\rho - \log\sigma)\right], \tag{3.3}$$

which is a non-negative quantity quantifying how different ρ and σ are. The relative entropy formula has further subleading corrections, but in this chapter we will only be focusing on the leading order result as in (3.2).

Our argument will give non-trivial results in cases where there is an island in the entanglement wedge of the radiation. The phenomena of islands has been proposed for a variety of scenarios involving different dimensions and also different signs of the cosmological constant (see [89–104] for some works in this direction and more in a recent review [105]). Our argument is general, but we will use the toy examples in [89] as a concrete set-up for our discussion. More specifically, we will consider the case with a zero temperature black hole in the two-dimensional Jackiw-Teitelboim (JT) gravity [15, 32, 33] coupled to a flat space bath. There is a two-dimensional CFT with central charge $c \gg 1$ propagating both in the JT gravity region and in the flat space region. This gravity description is semiclassical, namely we will be neglecting nonperturbative effects that are suppressed by e^{-S_0}, where S_0 is the extremal entropy of the black hole. We will also assume that the model has a dual nonperturbative description as a quantum mechanical dot coupled to a half infinite line where the CFT lives. The two descriptions of the system are illustrated in Figure 3.1 (a). For the

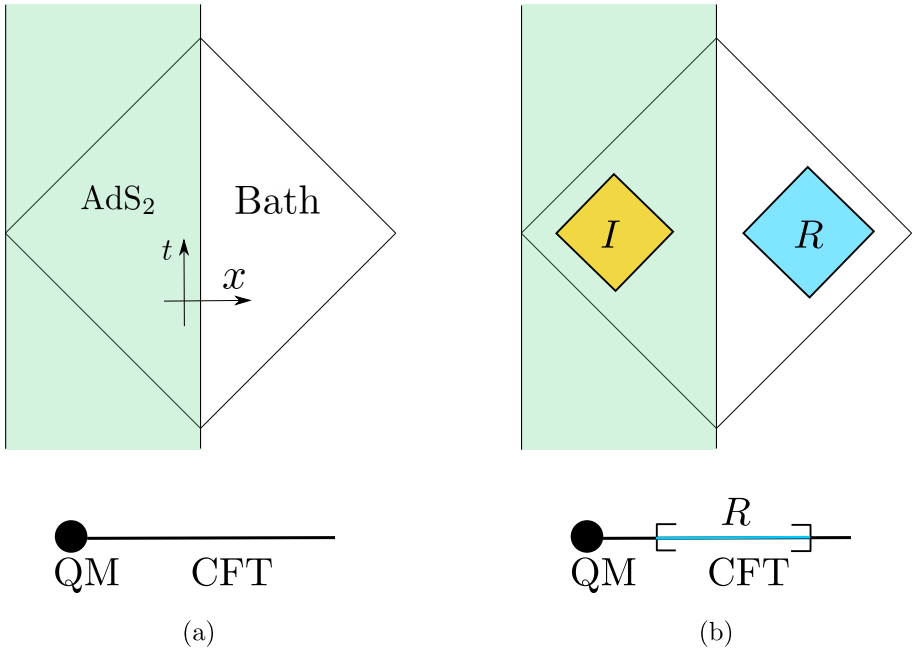


Figure 3.1: (a) A zero temperature black hole in JT gravity coupled to bath. It has a nonperturbative description in terms of a boundary quantum mechanical system coupled to a CFT in half infinite space. (b) When we compute the entropy of a region R in the bath that is large enough, there will be an island I in the gravity region.

ground state of the system, the metric and the dilaton profile in the AdS_2 region are

$$ds^2 = \frac{-dt^2 + dx^2}{x^2}, \quad x < 0, \quad \phi = S_0 - \frac{\phi_r}{x}, \tag{3.4}$$

where ϕ_r defines a length scale of the problem. The metric in the flat space region is given by

$$ds^2 = -dt^2 + dx^2, \quad x > 0, \tag{3.5}$$

and the conformal fields are in the Minkowski vacuum with respect to the global time coordinate t. For a region R in the flat space region that is large enough, applying the QES prescription to compute the entropy (and the location of the entanglement wedge), one finds that the entanglement wedge of R includes an island I in the JT gravity region (see Figure 3.1 (b)).

We consider situations in which there is a global symmetry in the semiclassical gravity description, with a symmetry group G that can generally be discrete or continuous, and possibly non-Abelian. We will assume that the symmetry is *not* gauged in the gravity description, leaving a discussion of gauge symmetries to §3.3.3. In terms of the exact "boundary" description, the flat space CFT has an exact global symmetry without coupling to the boundary quantum mechanical system, and the

coupling breaks this symmetry explicitly. However, the couplings that break the symmetry should be nonperturbatively small so that the global symmetry is still realized in the gravity region in the semiclassical description.

To simplify the discussion, we will first consider the case where the entire system is in a state which is invariant under the global symmetry transformation in the semiclassical description. For example, we could consider the ground state of the model in Figure 3.1 (a), for which the semiclassical state has the global symmetry.[1] If the state of the entire system is invariant under a symmetry transformation $U(g)$, then when we restrict to a spatial subregion R, the reduced density matrix $\rho(R)$ should also be invariant under the symmetry transformation restricted in R, which we denote by $U_R(g)$. (We will mostly leave the dependence on the group element $g \in G$ implicit from now on.) The reason is that the symmetry transformation U of a global symmetry has the split property [68]; morally, this means it can be written as a product of the symmetry transformation in R and its compliment $\bar{R}$, $U = U_R U_{\bar{R}}$.[2] It follows that

$$U_R \rho_R U_R^\dagger = t_{r\,\bar{R}} \left(U_R \rho U_R^\dagger \right) = t_{r\,\bar{R}} \left(U_{\bar{R}} U_R \rho U_R^\dagger U_{\bar{R}}^\dagger \right) = t_{r\,\bar{R}} \left(U \rho U^\dagger \right) = \rho_R. \tag{3.6}$$

Applied to our case, since the semiclassical state has the symmetry, we have

$$U_R \rho_{\text{semi}}(R) U_R^\dagger = \rho_{\text{semi}}(R), \tag{3.7}$$

or equivalently

$$S\left(U_R \rho_{\text{semi}}(R) U_R^\dagger \middle| \rho_{\text{semi}}(R) \right) = 0. \tag{3.8}$$

This motivates us to compare the two density matrices $\rho_{\text{exact}}(R)$ and $U_R^\dagger \rho_{\text{exact}}(R) U_R$, by computing the relative entropy between them, i.e. $S\left(U_R^\dagger \rho_{\text{exact}}(R) U_R \middle| \rho_{\text{exact}}(R) \right)$. Interestingly, to compute this quantity, we do not need to have a complete knowledge about $\rho_{\text{exact}}(R)$. We can compute it just from the semiclassical description, by applying the relative entropy formula (3.2):[3]

$$S\left(U_R \rho_{\text{exact}}(R) U_R^\dagger \middle| \rho_{\text{exact}}(R) \right) = S\left(U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger \middle| \rho_{\text{semi}}(R \cup I) \right). \tag{3.9}$$

We should stress that the transformation U_R on the left hand side is a completely well defined

[1] We will be assuming that the semiclassical global symmetry is not spontaneously broken throughout the chapter.

[2] In a continuum quantum field theory, there is an edge term U_{edge} that depends on how we regulate the boundary between R and $\bar{R}$. Since it is not essential to our discussion, we will keep it implicit.

[3] The relative entropy is valid either by considering a transformation U_R that is close to identity in the case of continuous symmetry, or by considering a global symmetry that only acts on a small subset of fields (note that we have $c \gg 1$).

operator in the nonperturbative description. For a continuous symmetry with Noether current J_a^μ, it is simply given by $e^{i\epsilon_a \int_R *J_a}$. An important assumption in the formula however is that when we pass from the exact description to the semiclassical description in (3.9), the transformation only acts in the region R but not the island, and still acts as the same symmetry transformation in the semiclassical theory.[4] One should be cautious about this, because the entanglement wedge of R includes both the region itself and the island, and in principle a general operation in R in the exact description can act both in the region R and the island I. Indeed, as advocated in [106], a modular flow using the exact modular Hamiltonian of R corresponds to a modular flow on $R \cup I$ with the semiclassical modular Hamiltonian. However, in contrast with the exact modular Hamiltonian, which is expected to be non-local and encodes the fine grained information of the state, the operator U_R factorizes into local pieces and does not contain information about the state. In other words, it is a "simple" operation. A more careful argument is the following. We can divide the region R into smaller pieces $\{R_i\}$, and the operator U_R factorizes as

$$U_R = \prod_i U_{R_i}. \tag{3.10}$$

If we divide R such that R_i are small enough, the entanglement wedge of each R_i will only contain the region itself. Thus each U_{R_i} should only act within R_i without any ambiguity, so the product should only act in the union of R_i, which is R. This is part of the idea that has been used in [68, 69, 84] to argue that there cannot be an exact global symmetry. In §3.3, we will also justify the formula (3.9) from the replica computation point of view.

After justifying the formula (3.9), we now turn to its implications. The density matrix $\rho_{\text{semi}}(R \cup I)$ is invariant under the semiclassical global symmetry transformation on the union of R and I, but it is *not* invariant under U_R alone. This is because the fields in R are entangled with the fields in I and so $\rho_{\text{semi}}(R \cup I) \neq \rho_{\text{semi}}(R) \otimes \rho_{\text{semi}}(I)$; otherwise it would be invariant under U_R because of (3.7). For this reason, for a fixed U_R, one expects the result to be larger if the region R and the island have larger mutual information. To summarize, we expect that the bulk formula (3.9) gives a nonzero relative entropy when there is an island:

$$S\left(U_R \rho_{\text{exact}}(R) U_R^\dagger \middle| \rho_{\text{exact}}(R)\right) = S\left(U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger \middle| \rho_{\text{semi}}(R \cup I)\right) = \mathcal{O}(1), \tag{3.11}$$

where $\mathcal{O}(1)$ denotes that this quantity is not suppressed by e^{-S_0}. As a side note, assuming that the

[4] We thank Ahmed Almheiri and Edgar Shaghoulian for asking about this assumption.

global semiclassical state is invariant under the symmetry, $\rho_{\text{semi}}(R \cup I)$ is invariant under $U_R U_I$, so we can also compute the relative entropy via

$$S\left(U_R \rho_{\text{exact}}(R) U_R^\dagger \middle| \rho_{\text{exact}}(R)\right) = S\left(U_I^\dagger \rho_{\text{semi}}(R \cup I) U_I \middle| \rho_{\text{semi}}(R \cup I)\right). \tag{3.12}$$

The quantity $S\left(U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger \middle| \rho_{\text{semi}}(R \cup I)\right)$ is computable solely using the semiclassical description. Since the von Neumann entropies of $U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger$ and $\rho_{\text{semi}}(R \cup I)$ are the same as they only differ by an unitary, the relative entropy can be expressed as the difference of the expectation value of the modular Hamiltonian $K_{\text{semi}}(R \cup I) \equiv -\log \rho_{\text{semi}}(R \cup I)$:

$$S\left(U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger \middle| \rho_{\text{semi}}(R \cup I)\right) = \Delta \langle K_{\text{semi}}(R \cup I) \rangle, \tag{3.13}$$

where we take the expectation value of the modular Hamiltonian in the state after acting with U_R, and subtract that in the original state. In simple cases, this computation can be carried out explicitly, as we will demonstrate in §3.2.2.

Before we move on to concrete examples, let us offer some qualitative comments. For the relative entropy to be nonzero, we need the region R to be large enough so that R claims the island. If we divide R into small pieces, then when we only look at the relative entropies for the small spatial pieces, we will not see such an effect.

Similar statement holds if we divide the fields in R into small subsets. For example, imagine the CFT is N copies of free fermions ($N \gg 1$), which has an $U(N)$ global symmetry. We could consider a symmetry transformation U_R that only acts among a small subset of the fermions, say M fermions with $M \sim \mathcal{O}(1)$. If we only look at the reduced density matrix of the M fermions while tracing out the rest, since it does not claim the island, we would conclude that it is invariant under the symmetry transformation. However, the density matrix of all the N fermions will not be invariant by (3.11). This suggests that the main contribution to the order one result in (3.11) comes from correlations between the M fermions and the rest of the fermions. For this to be true, there must be sufficient correlation between the M fermions and the rest of the fields at the non-perturbative level, despite the fact that they are decoupled and uncorrelated in the semiclassical description.[5] This correlation can also be quantified using the gravitational fine grained entropy formula, and we comment more on this in Appendix 3.5.

[5]Here we are neglecting the universal coupling through the boundary graviton in the semiclassical description. In the current model, this can be justified in the limit of $c \gg 1$ and S_0/c not too large.

3.2.2 Example: $U(1)$ global symmetry of Dirac fermions

As a concrete, calculable example, we consider a setup where the CFT is a tensor product of a massless Dirac fermion and some other CFT with central charge $c-1$ which we will not specify. For simplicity, we will assume that in the semiclassical description, this fermion field does not couple to the rest of the CFT directly, and we also neglect the Schwarzian fluctuations. In the semiclassical description, the model has a $U(1)$ global symmetry which rotates the fermion as $\psi \to e^{i\theta}\psi$. The time component of the conserved current corresponding to this symmetry is the fermion number density:

$$J_0 \equiv \psi_+^\dagger \psi_+ + \psi_-^\dagger \psi_-, \tag{3.14}$$

and the transformation we apply in R is

$$U_R \equiv \exp\left(i\alpha Q_R\right) = \exp\left(i\alpha \int_R dx\, J_0\right). \tag{3.15}$$

As we explained in (3.13), the relative entropy is given by the change of the expectation value of the semiclassical modular Hamiltonian $K_{\text{semi}}(R \cup I)$:

$$S\left(U_R \rho_{\text{exact}}(R) U_R^\dagger \middle| \rho_{\text{exact}}(R)\right) = \langle 0| U_R^\dagger K_{\text{semi}}(R\cup I) U_R |0\rangle - \langle 0| K_{\text{semi}}(R\cup I)|0\rangle, \tag{3.16}$$

where the $|0\rangle$ state is the vacuum state of the fermion in the semiclassical description. Since the fermion is massless, we can neglect the warp factor inside the AdS region. In other words, up to local terms at the boundary of the island that depends on the warp factor, the modular Hamiltonian $K_{\text{semi}}(R\cup I)$ is just the modular Hamiltonian for two disjoint intervals in flat space, whose form was given explicitly in [107] which we review in appendix. 3.6. We will not need the detailed form of the modular Hamiltonian here to understand the calculation, but only need its basic structure. For the vacuum state, the modular Hamiltonian is quadratic in the fermion operators. It can be separated into two pieces, the local piece K_{loc} which couples operators within R or I, and the non-local piece K_{noloc} that couples a fermion operator in R to another in I. It is not hard to see that

$$U_R^\dagger K_{\text{loc}} U_R = K_{\text{loc}}, \tag{3.17}$$

as K_{loc} can be expressed as an integral of the local energy density operator $T(x)$, which is invariant under U_R (the piece in the island is invariant trivially because it is spacelike separated from U_R).

On the other hand, K_{noloc} transforms non-trivially under U_R. K_{noloc} has the form

$$K_{\text{noloc},\pm} = i\int_I dx \int_R dy\, K(x,y)(\psi^\dagger_{\pm,I}(x)\psi_{\pm,R}(y) - \psi^\dagger_{\pm,R}(y)\psi_{\pm,I}(x)), \tag{3.18}$$

where the kernel $K(x,y)$ is given explicitly in Appendix 3.6. The symmetries act as

$$U_R^\dagger \psi^\dagger_{\pm,I}(x)\psi_{\pm,R}(y)U_R = e^{i\alpha}\psi^\dagger_{\pm,I}(x)\psi_{\pm,R}(y), \quad U_R^\dagger \psi^\dagger_{\pm,R}(y)\psi_{\pm,I}(x)U_R = e^{-i\alpha}\psi^\dagger_{\pm,R}(y)\psi_{\pm,I}(x). \tag{3.19}$$

By noting that $\langle 0|\, \psi^\dagger_{\pm,I}(x)\psi_{\pm,R}(y)\,|0\rangle = -\langle 0|\, \psi^\dagger_{\pm,R}(y)\psi_{\pm,I}(x)\,|0\rangle$, we obtain

$$S\left(U_R\rho_{\text{exact}}(R)U_R^\dagger\middle|\rho_{\text{exact}}(R)\right) = (\cos\alpha - 1)\,\langle 0|\, K_{\text{noloc}}\,|0\rangle\,. \tag{3.20}$$

The expression is periodic in 2π since the transformation with $\alpha/2\pi = n \in Z$ acts trivially. The expectation value of K_{noloc} is computed explicitly in Appendix 3.6. In Poincar coordinates x, if we parametrize the regions as $I = (a_1, b_1), R = (a_2, b_2)$, then

$$S\left(U_R\rho_{\text{exact}}(R)U_R^\dagger\middle|\rho_{\text{exact}}(R)\right) = \sin^2\frac{\alpha}{2}\left(1 + \frac{(2\eta-1)\arctan\sqrt{\frac{\eta}{1-\eta}}}{\sqrt{\eta(1-\eta)}}\right), \quad \eta \equiv \frac{(b_1-a_1)(b_2-a_2)}{(a_2-a_1)(b_2-b_1)}. \tag{3.21}$$

We see that (3.21) diverges as $1/\sqrt{1-\eta}$ when $\eta \to 1$, where the island and the bath region get closer. This is consistent with the intuition that the result should be larger if there is more mutual information between R and I.[6]

As discussed in [89], if we choose the bath region

$$a_2 \ll \frac{\phi_r}{c} \ll b_2, \quad \log\left(\frac{cb_2}{\phi_r}\right) > \frac{12S_0}{c} + \mathcal{O}(1), \tag{3.22}$$

then the island will be at

$$a_1 \approx -b_2, \quad b_1 \approx -\frac{6\phi_r}{c}. \tag{3.23}$$

Plugging (3.22) and (3.23) into (3.21), we find

$$S\left(U_R\rho_{\text{exact}}(R)U_R^\dagger\middle|\rho_{\text{exact}}(R)\right) \approx \frac{\pi}{4}\sin^2\frac{\alpha}{2}\sqrt{\frac{cb_2}{3\phi_r}}. \tag{3.24}$$

We observe that the relative entropy grows with the total central charge c, even though we are

[6]However, it is interesting to note that it grows faster than the mutual information as $\eta \to 1$, since the mutual information only grows logarithmically.

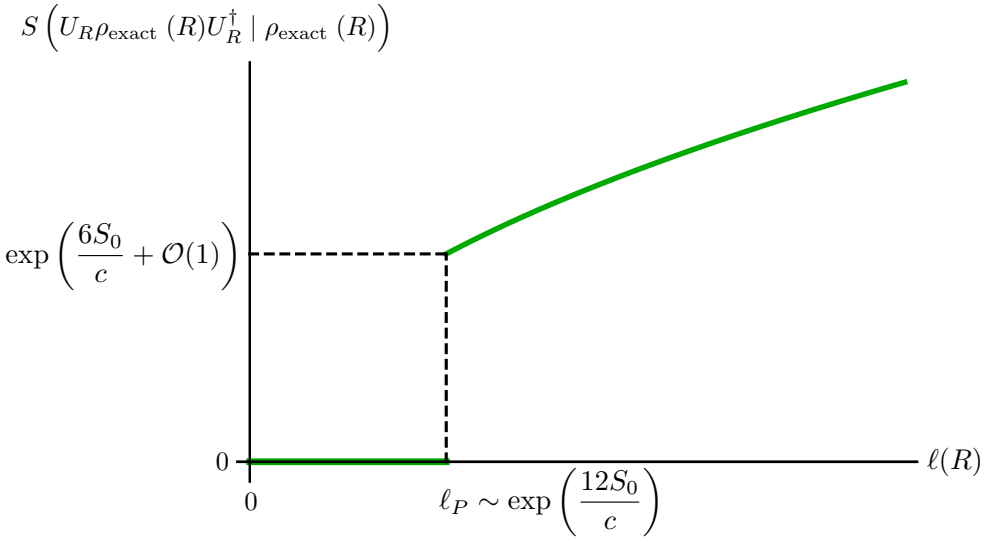


Figure 3.2: The relative entropy in the free fermion model, as a function of the size of the region R at fixed non-zero α. The horizontal axis is $\ell(R)$ the length of R in units of ϕ_r/c ($\ell(R) \equiv c(b_2 - a_2)/\phi_r \approx cb_2/\phi_r$). In the semi-classical approximation, there is a discontinuity in the relative entropy at the "Page length" $\ell_P(R) \sim e^{12S_0/c}$ when an island appears in the gravity region, see (3.24). It might seem that the relative entropy is large because it has S_0 in the exponential, but recall that in this model the ratio S_0/c is kept finite and not very large [89], so the relative entropy is $\mathcal{O}(1)$.

considering the symmetry that only involves one fermion. This confirms our previous expectation that the symmetry violation is contained in the correlation between the fermion field and other fields. We stress that $\rho_{\text{exact}}(R)$ is the density matrix for all the fields, not just the one fermion field. Note that ϕ_r/c in (3.24) is roughly the time it takes for a particle to reach the island if it is released from the boundary. We will give a physical explanation of why this "island time" is relevant in §3.2.3. We sketch the result (3.24) in Figure 3.2. There is an analogue of the "Page time" (perhaps more aptly called the Page length) ℓ_P in this problem that is the minimal size of R to claim the island. Below the minimal size, the relative entropy is zero, while above that size, the relative entropy becomes finite and grows with the size. At finite S_0, we expect to have a smooth curve instead of a sharp transition, see [108–110].

3.2.3 Increasing the relative entropy with charged particles

So far, we've only considered the ground state of the model in Figure 3.1. It is natural to ask whether we can choose different bulk semi-classical states which will have a larger relative entropy when there is an island. To motivate the discussion, let us first look at a rough description of the semi-classical vacuum state. There are three regions, the island I, the radiation R, and the region in between — call it S for sea. The global state is charge neutral but locally there can be vacuum

fluctuations of the charge density. We will focus on two possible forms of the fluctuation, and look at the density matrices they lead to. There are of course contributions from other types of terms and cross terms, but our goal here is just to draw some basic intuition. The first case corresponds to virtual loops that run between the sea and the radiation, as represented by ℓ_2 in Figure 3.3 (a):

$$|\text{vac}\rangle_{\text{bulk}} \sim \sum_q |0\rangle_I |-q\rangle_S |q\rangle_R , \tag{3.25}$$

where $Q_i |q\rangle_i = q_i |q\rangle_i$, $i = I, S, R$. The density matrix that it leads to is

$$\rho_{\text{semi}}(R \cup I) \sim \sum_q (|0\rangle_I \langle 0|_I \otimes |q\rangle_R \langle q|_R) . \tag{3.26}$$

It is invariant under U_R, and thus will not contribute to the relative entropy. The second case involves loops that run between the island and the radiation (loop ℓ_1 in Figure 3.3 (a)):

$$|\text{vac}\rangle_{\text{bulk}} \sim \sum_q |-q\rangle_I |0\rangle_S |q\rangle_R , \tag{3.27}$$

and it leads to

$$\rho_{\text{semi}}(R \cup I) \sim \sum_{q,\tilde{q}} (|-q\rangle_I \langle -\tilde{q}|_I \otimes |q\rangle_R \langle \tilde{q}|_R) . \tag{3.28}$$

We see that the semiclassical density matrix is not invariant under U_R, and will lead to a non-trivial dependence on α in the relative entropy via (3.2). These loops are also the virtual processes that build up the entanglement between the island and the radiation, so we see that the relative entropy is closely related to entanglement, or mutual information. See Appendix 3.5 for some further discussion.

The above toy model is supposed to depict the quantum fluctuations in the vacuum. However, it also suggests a way to increase the bulk relative entropy by making on shell particles with correlated charges. To neglect the fluctuations that already exist in the vacuum state, we could consider a matter theory which consists of a large c CFT + a massive "probe" theory with an $U(1)$ symmetry (for example, a single massive free fermion field). By "probe" field, we mean that when we can neglect the massive theory in determining the location of the island. When we compute the relative entropy with $U(1)$ symmetry generators acting on the massive theory using (3.13), we expect it to be suppressed by $e^{-m\ell}$ where ℓ is the proper distance between the island and the bath region R, and

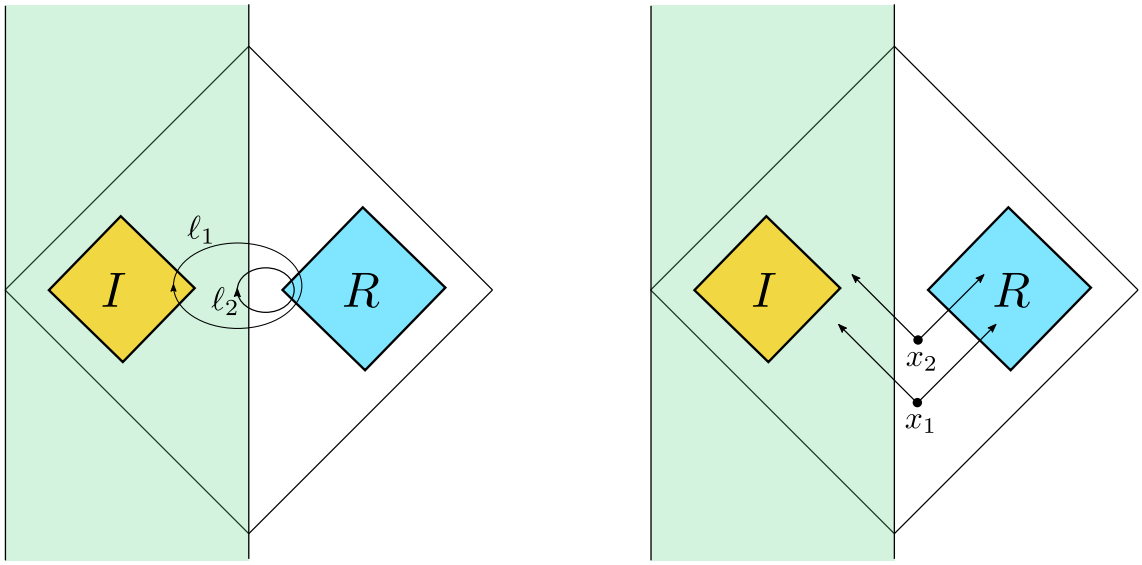


Figure 3.3: (a) We consider two virtual processes ℓ_1 and ℓ_2 where charged particles run in loops. Only the larger loop ℓ_1 contributes to the bulk relative entropy (as well as the entanglement between I and R). (b) We imagine the decay of a neutral particle into 2 fermions at x_1 and x_2. Only the decay at x_1 contributes significantly to the bulk relative entropy. Note also that in an evaporating black hole, a morally similar picture would also predict that the relative entropy increases after the Page time because each Hawking mode will be entangled with a mode of opposite charge behind the horizon.

is thus small for large mass. So if we create many pairs of particles, the leading contribution to the relative entropy will come from the real particles instead of vacuum fluctuations. In other words, we can interpret the charge fluctuations in (3.25) and (3.27) as coming from real particles.

Imagine that we start with a massive neutral particle at rest in the bath region. It then decays into two massive free fermions, with charge $+q$ and $-q$. These fermions propagate in opposite directions: one towards the black hole, and the other towards null infinity. However, we do not know which charge went where. It is interesting to consider the dependence of the relative entropy on the location x of the fermion. If the particle which falls into the black hole is not in the island, then we have a state similar to (3.25), and it will not increase the relative entropy significantly. However, if we drop in the particle and then wait some time t_{island}, the particle will appear in the island, see Figure 3.3. In cases where the island is near the horizon, for example in the evaporating black hole, or in the 2-sided setup at late times, we will need to wait about a scrambling time for a perturbation applied near the boundary of the gravity region.[7] This will lead to a nonzero relative

[7] In the 1-sided zero-temperature case, the scrambling time is infinite but the $t_{\text{island}} \sim \phi_r/c$.

entropy, by considerations nearly identical to those in (3.27) and (3.28).

We would like to interpret the nonzero relative entropy in this case as saying that the charge is lost from the outside observer point of view. If the symmetry was exact, we should find a zero relative entropy at all time. The fact that we have a nonzero relative entropy after creating the particle pair simply means that the charge that is outside the region R is not conserved. From our discussion, we see that this non-conservation happens at the time scale which is the scrambling time. We will return to this point in the discussion.

3.2.4 Evaporating black hole

In the discussion of §3.2.1 and 3.2.2, we mainly considered time independent cases where a black hole is in thermal equilibrium with the environment. However, our discussion also applies to general time dependent cases, such as evaporating black holes. Consider a theory that has a global symmetry in the semiclassical description, and we start from a pure state that has zero global charge and form a black hole. If the symmetry is not violated nonperturbatively, then if we look at the Hawking radiation emitted from the black hole, its density matrix should always be block diagonal in the charge basis, and thus is invariant under the symmetry transformation acting on the Hawking radiation. However, after the Page time, since the entanglement wedge of the Hawking radiation contains an island in the black hole interior, we can similarly compute relative entropies and see that the global symmetry is violated from a nonzero result. We also expect the relative entropy to grow as one collects more radiation from the black hole.

Of course, it is well known how to see that global symmetry is violated by looking at the final stage of evaporation as we reviewed in the introduction [67]. Here we are relating the violations of global symmetry to the appearance of an "island." An advantage of the argument presented here is that it provides a quantitative way to see global symmetry violation just *after the Page time*, and does not rely on the physics close to the final stage of the black hole evaporation.

Notice that if we gather the entire radiation of a completely evaporated black hole, it will be in a pure state. The pure state will be a superposition of states with different charge. Hence we expect that the relative entropy after the black hole has completely evaporated to formally diverge. On the other hand, the Rényi versions of the relative entropy that we study in §3.3 will still be well defined and give sensible answers. A related point is discussed in Appendix 3.7, where we explore global symmetry violations in a different model of JT gravity.

Instead of starting with a state that is annihlated by the charge, one could also imagine forming

the black hole from a general initial state ρ_0 that is not invariant under the global symmetry U that acts on the entire system.[8] In this case, one could still see that global symmetry is violated as follows. If the symmetry were exact, then when we look at any subsystem R at later time t, we should always find

$$\begin{aligned} S\left(U_R\rho_R(t)U_R^\dagger\middle|\rho_R(t)\right) &\leq S\left(Ue^{iHt}\rho_0 e^{-iHt}U^\dagger\middle|e^{iHt}\rho_0 e^{-iHt}\right) \\ &= S(U\rho_0 U^\dagger|\rho_0), \end{aligned} \tag{3.29}$$

where on the first line we are using the monotonicity of relative entropy under restricting to a subsystem R, and we used the assumption that U is an exact symmetry to get to the second line. However, if we use the gravitational formula to compute $S\left(U_R\rho_R(t)U_R^\dagger\middle|\rho_R(t)\right)$, we expect to get an increasing function with t after the Page time,[9] so it cannot be bounded by a constant $S(U\rho_0U^\dagger|\rho_0)$ as in (3.29), which necessarily means that the global symmetry is violated.

3.3 Global symmetry violation and replica wormholes

3.3.1 Charge flowing through the replica wormhole

As we've seen in §3.2, when there is an island in the entanglement wedge of R, the global symmetry violation is reflected in the relative entropy. In the computation of the von Neumann entropy using the replica trick, the island arises from the replica wormholes [13, 14]. So it is natural to consider the replica version of the relative entropy, and ask what the role of replica wormholes are in these quantities. Similar to the von Neumann entropy, the relative entropy can be computed via a replica trick [111] by

$$S(\rho|\sigma) = \lim_{n\to 1}\frac{1}{1-n}t_r\left[\rho(\sigma^{n-1}-\rho^{n-1})\right]. \tag{3.30}$$

To be concrete, we will take the two-sided black hole setup in [89] as an example, whose replica wormhole solution has been constructed explicitly in certain limits in [14]. The goal of this section is not to study the replica wormhole geometries in detail, but to understand what features of the replica wormholes are related to charge violation. We will be considering the thermofield double state $|\text{TFD}\rangle$ of the system, and study the density matrix of a region R which contains part of both

[8] Here we are assuming that before the black hole forms, all the matter are far apart and the gravity effect is weak enough such that the operator U is well defined.

[9] We expect the curve to be qualitatively similar to Figure 3.2, if we replace the size of the interval by time t. The reason is that there will be more and more correlated charge fluctuations between I and the Hawking radiation R (this is similar in spirit to Figure 3.3). It would be interesting to understand the calculation for evaporating black holes in more detail.

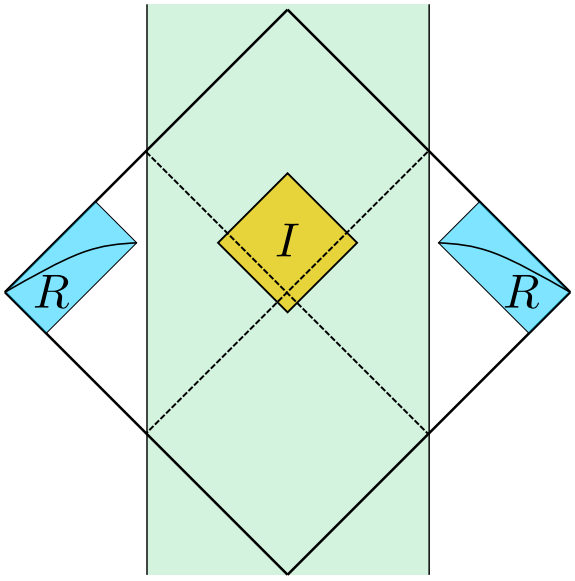


Figure 3.4: A two-sided black hole in JT gravity coupled to two bath regions [14]. When we compute the entropy of a large region R in the bath, depending on the parameters, we can have an island I in the gravity region. At sufficiently late times $t > 0$ there will always be an island.

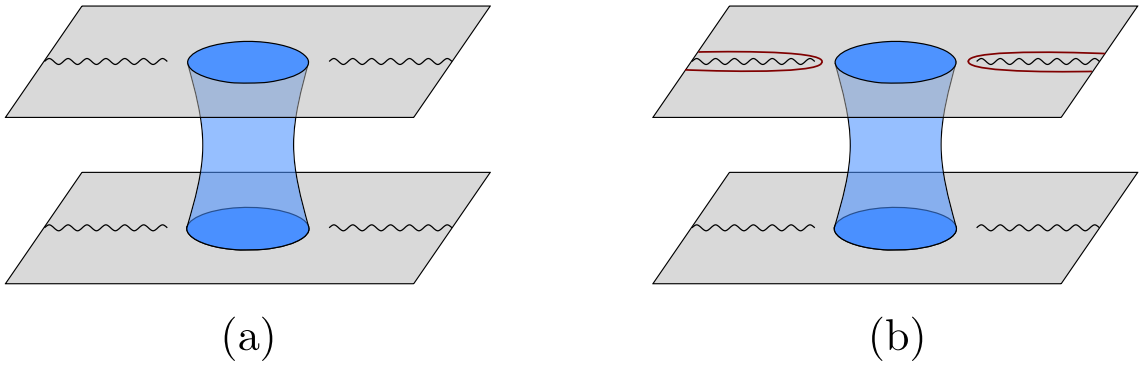


Figure 3.5: (a) The replica wormhole geometry $\mathcal{M}_2$ for computing $Z_2(0)$. (b) For $Z_2(\alpha)$, we have the symmetry transformation operators inserted around the cuts in one of the replicas.

the left bath and the right bath (see Figure 3.4). Depending on the parameters, we can have an island I that is slightly outside the horizon already at $t = 0$; our arguments in §3.2 would then apply.

Here we would like to consider the $n = 2$ replica:

$$\frac{Z_2(\alpha)}{Z_1^2} = \mathrm{tr}\left[U_R \rho_{\mathrm{exact}}(R) U_R^\dagger \rho_{\mathrm{exact}}(R)\right], \quad U_R = e^{i\alpha Q_R}, \tag{3.31}$$

where we take the symmetry to be an $U(1)$ global symmetry just for simplicity. We also divided by Z_1^2 where Z_1 is given by the path integral on a single replica, to get the expression for a normalized density matrix. If the symmetry were exact, then we should have $Z_2(\alpha) = Z_2(0)$, for the reasons

discussed in §3.2.1. As explained in [14], $Z_2(0)$ can be computed via a gravitational path integral with two replicas, and in the case with an island, the path integral is dominated by a replica wormhole geometry (see Figure 3.5 (a)), which we denote by $\mathcal{M}_2$. The only difference in the computation of $Z_2(\alpha)$ is that we have the extra insertions of U_R and $U_R^\dagger$, which are represented by the red lines in figure 3.5 (b).

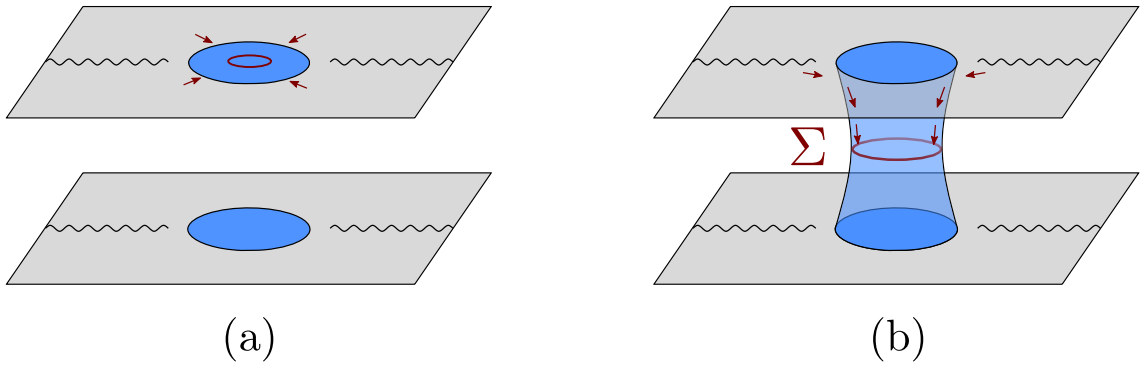


Figure 3.6: (a) Without the replica wormhole, the topological operator can be shrunk to a point. (b) With the replica wormhole, we can deform the topological operator to surround the throat of the wormhole. Therefore the dependence on the parameter α diagnoses the charge flowing through the wormhole.

Importantly, since we are computing (3.31) in the semiclassical description that has the global symmetry, the red lines in Figure 3.5 (b) represent topological operators.[10] This means we can deform the curves arbitrarily as long as they do not cross any charged operators (in which case they would pick up a phase). Since we are considering the thermofield double state, which does not contain any insertions, we can deform the topological operator freely. If there were no replica wormholes (see Figure 3.6 (a)), we could then deform and shrink the topological operator to a point, which tells us

$$\frac{Z_2(\alpha)}{Z_2(0)} = 1, \quad \text{(no replica wormhole)}. \tag{3.32}$$

On the contrary, with the replica wormhole, the topological operator cannot be shrunk to a point, but it can be deformed to the throat Σ of the wormhole, see Figure 3.6 (b). So we have

$$\frac{Z_2(\alpha)}{Z_2(0)} = \langle e^{i\alpha Q_{\text{wormhole}}} \rangle_{\mathcal{M}_2}, \quad \text{(with replica wormhole)}, \tag{3.33}$$

where $Q_{\text{wormhole}} = \int_\Sigma *J$ and J^μ is the Noether current. This is the finite replica analogue of equation (3.12). In general, we expect that $\langle e^{i\alpha Q_{\text{wormhole}}} \rangle_{\mathcal{M}_2} < 1$, simply because the value of Q_{wormhole} can

[10] The topological operator is well-defined for other symmetries as well, including non-Abelian cases and discrete symmetries.

fluctuate in the path integral. Thus we would have

$$Z_2(\alpha) < Z_2(0), \tag{3.34}$$

which means that the density matrix $\rho_{\text{exact}}(R)$ is not invariant under the symmetry transformation, as we concluded by studying the relative entropy.

We can use the intuition from a weakly-coupled theory to understand this effect slightly better. If we consider a small (meaning contractible) loop of virtual particles near the throat of the geometry, the net charge flow will be zero, since on a spatial slice Σ, we will have one particle and one anti-particle, see Figure 3.7 (a). However, the virtual loop could be non-contractible due to the cuts in the bath region, see Figure 3.7 (b).

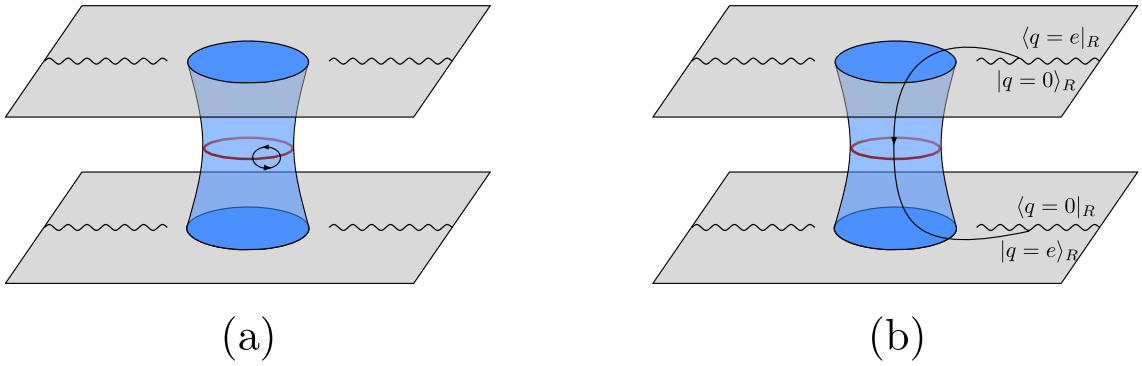


Figure 3.7: Two virtual processes that contribute to $Z_2(\alpha)$. (a) The smaller, contractible loop does not give a contribution that depends on α, since on the red circle there is both a particle and an anti-particle. (b) The larger loop is non-contractible, contributes to the off-diagonal part $\text{tr}\,[\rho_{e,0}\rho_{e,0}]$ (see §3.3.2), and therefore to the α dependence. This should be compared with Figure 3.3.

3.3.2 Off-diagonal contributions to the density matrix

The fact that $U_R\rho_{\text{exact}}(R)U_R^\dagger$ and $\rho_{\text{exact}}(R)$ are different indicates that $\rho_{\text{exact}}(R)$ must have some off-diagonal elements in the charge basis. We can try to "open up" the quantity in (3.31) and look at the contribution to the off-diagonal elements more explicitly:

$$\text{tr}\left[e^{i\alpha Q_R}\rho_{\text{exact}}(R)e^{-i\alpha Q_R}\rho_{\text{exact}}(R)\right] = \sum_q \text{tr}\left[\rho_{q,q}^2\right] + e^{i\alpha(q-q')}\sum_{q\neq q'}\text{tr}\left[\rho_{q,q'}\rho_{q',q}\right], \tag{3.35}$$

where

$$\rho_{q,q'} \equiv P_q\rho_{\text{exact}}(R)P_{q'} \tag{3.36}$$

and $P_q \equiv \sum_i |q,i\rangle \langle q,i|$ is the projector into the subspace with $Q_R = q$, with i running over the states in the same-charge subspace. Note that the projection only acts on the fields that transform under the symmetry, which we take to be a small subset of the fields. Since we are tracing out all the other fields, we expect a similar replica wormhole solution for each term in (3.35) independently. We see from (3.35) that all the dependence on α comes from the off-diagonal terms in the density matrix. In this section, we give a more direct explanation of the appearance of off-diagonal terms, and how they arise from Euclidean wormholes in gravity.

One way to understand the off-diagonal term $\mathrm{tr}\,[\rho_{q,q'}\rho_{q',q}]$ is as follows. We can rewrite the quantity as

$$\mathrm{tr}\,[\rho_{q,q'}\rho_{q',q}] = \sum_{i,j} |\,\langle \mathrm{TFD}|q,i|\mathrm{TFD}|q,i\rangle\,\langle q',j|\mathrm{TFD}|q',j|\mathrm{TFD}\rangle\,|^2, \tag{3.37}$$

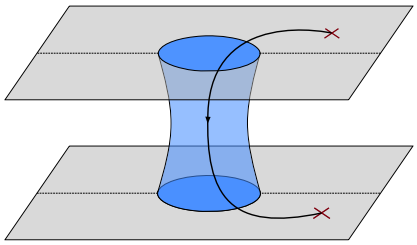
and in the path integral, it corresponds to putting boundary conditions for the fields along the branch cuts such that they have definite charges. (This should not be taken too literally, as such a state might be a very high energy.) In the semiclassical approximation, we can think of the boundary conditions as places where the worldline of a charged particle can be created or destroyed. Let's take $\mathrm{tr}\,[\rho_{e,0}\rho_{0,e}]$ as a simple example, where e is the unit charge of a charged particle. Due to the boundary condition, we will have a worldline starting from the bra part of the cut in the first replica, and the only place where it can land on is the ket part of the cut in the second replica. The only possibility for the contribution to be nonzero is that the charge can go through the replica wormhole, as shown in Figure 3.7 (b).

Despite that we have $\mathrm{tr}\,[\rho_{q,q'}\rho_{q',q}] \neq 0$, if we compute $\mathrm{tr}\,[\rho_{q,q'}]$ with $q \neq q'$ using gravity, we will find a zero answer, up to nonperturbative corrections. This is because we only have one replica, and there is nowhere the charged particle can go to. This effect can be viewed as a special case of a more general phenomena involving Euclidean wormholes. Consider a state of the effective quantum field theory in the bath obtained by inserting some charged operators in the thermofield double state $|\psi_q\rangle = O_{q_1}O_{q_2}\cdots O_{q_n}|\mathrm{TFD}\rangle$. Here $O_q = O_q(\tau,x)$ is some local operator with charge q evaluated at a general Euclidean or Lorentzian time. Ignoring the effects of higher topologies, this state will be orthogonal to the state $|\psi_{\tilde{q}}\rangle = O_{\tilde{q}_1}O_{\tilde{q}_2}\cdots O_{\tilde{q}_m}|\mathrm{TFD}\rangle$ unless the total charges are equal $q = \sum_i q_i = \tilde{q} = \sum_i \tilde{q}_i$, namely $\langle\psi_q|\psi_{\tilde{q}}|\psi_q|\psi_{\tilde{q}}\rangle = 0$ up to exponentially small errors. However, wormhole geometries will lead to a small squared-overlaps when $q \neq \tilde{q}$ [112]:

$$|\,\langle\psi_q|\psi_{\tilde{q}}|\psi_q|\psi_{\tilde{q}}\rangle\,|^2 = \delta_{q-\tilde{q}}\,(\mathrm{disconnected}) + e^{-S_0}(\mathrm{connected}) \tag{3.38}$$

The connected contribution comes from an Euclidean wormhole, where the charged particles go through the wormhole. For example, the square of a 1-pt function $\langle O_q \rangle$ with charge q is given by

$$|\langle \text{TFD}| O_q |\text{TFD}\rangle|^2 \quad = \quad \text{(3.39)}$$



Note however that (3.39) might require an ensemble average interpretation, since the left hand side is the square of $\langle \text{TFD}| O_q |\text{TFD}\rangle$, while the right hand side does not factorize manifestly. In a single unitary system, the proper interpretation of (3.39) might be that the wormhole computes the average of $|\langle \text{TFD}| O_q |\text{TFD}\rangle|^2$ among a suitable family of operators with charge q. We refer the readers to [7–9, 112–119] for more discussions on these issues about Euclidean wormholes. We should stress that the quantity $\text{tr}\,[\rho_{q,q'}\rho_{q',q}]$ does not factorize in the first place, since each charge subspace still contains a huge number of states that we sum over. So the quantity we compute does not suffer from factorization problem immediately. Due to the summation over a large number of intermediate states, we have a replica wormhole geometry which is a saddle point solution that dominates the path integral, while in general the wormhole in (3.39) will be an off-shell configuration (see however an on-shell construction in [112]).

If we imagine that the wormhole region is far away from the operator insertions, we can summarize the effects of the wormhole using an effective picture. If we think of the gravity region as a complicated boundary condition for the bath fields, this boundary condition will break explicitly the global symmetry. If the gravity theory has a holographic description as we've assumed in §3.2, then this boundary condition would be realized by the coupling between the quantum mechanical system and the bath CFT. More generally, we expect the same picture to hold even when the gravity theory does not have a holographic description. Far away from this boundary condition, the circle is hard to distinguish from a point, so we can replace the boundary condition by a sum of operator insertions, see Figure 3.8. We can imagine that there is a small coefficient for some operator $c_q O_q$ where c_q has mean zero in some appropriate averaged sense. Then ignoring these operators is equivalent to only considering disconnected geometries. When we include connected geometries, we are taking into account the non-zero variance of these coefficients $|c_q|^2 \sim e^{-S_0}$. Of course, this is just the old story about "α-parameters" and Euclidean wormholes [64, 116, 120, 121], what we did in this section is making the relation to replica wormholes more explicit.

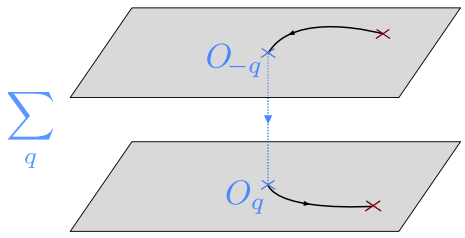


Figure 3.8: We can reproduce the effects of the wormhole by an effective picture where we sum over correlated operators. This is also what happens from the boundary point of view, where the gravity region is replaced by a boundary condition which completely breaks the global symmetry. Here we are imagining that the operator insertions in the bath are far away from the gravity region. In some theories with a disorder average, the dashed blue line may represent a contraction of couplings.

From this perspective, it is clear that a Euclidean wormhole with charge $\pm q$ propagating through the throat is giving rise to a state with charge $\pm q$ different than the "naive" charge of the state (e.g. the charge of the state we would have assigned in the effective description of the quantum fields on the trivial topology).

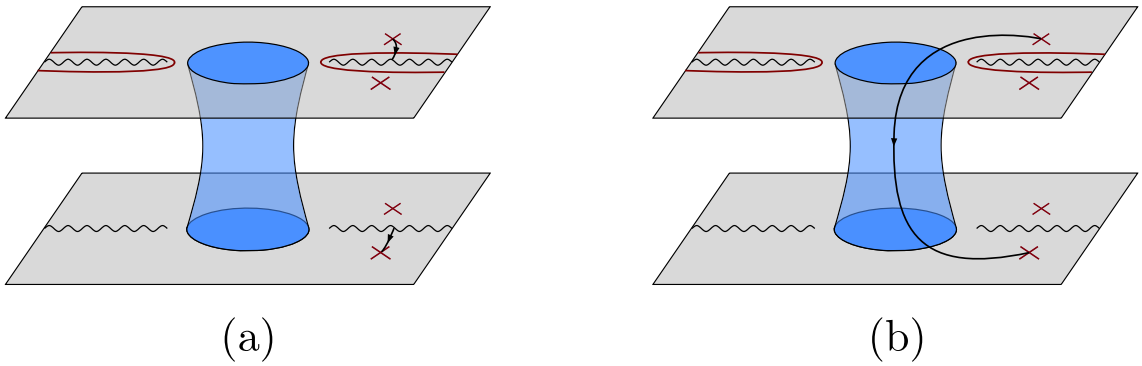


Figure 3.9: We consider a global state which is not the vacuum but contains some operator insertions in the bath region. (a) The diagonal contribution to the density matrix comes from particles propagating through the cuts. (b) Off-diagonal contributions come from world-lines with non-zero net charge propagating through the wormhole.

Now let us return to the density matrix of the fields in a subregion R. We will slightly generalize the previous discussion by considering a state $|\psi_q\rangle$ which is obtained by inserting some charged operators in the Euclidean evolution. Let us denote elements of a charge eigenbasis by a composite

index r such that $|r\rangle = |q_r, i\rangle$, where i runs over states with the same charge q_r. Then,

$$\begin{aligned}
\rho_{\text{exact}}(R) &= t_{r\,\bar{R}}\,|\psi_q\rangle\,\langle\psi_q|\,, \\
Z_2(\alpha)/Z_1^2 &= \sum_{r,s} \langle r|\, e^{-i\alpha Q}\rho_{\text{exact}}(R)e^{i\alpha Q}\,|s\rangle\,\langle s|\,\rho_{\text{exact}}(R)\,|r\rangle \\
&= \sum_{r,s} e^{i\alpha(q_s-q_r)}\,\langle\psi_q|\,O^R_{s,r}\,|\psi_q\rangle\,\langle\psi_q|\,O^R_{r,s}\,|\psi_q\rangle \\
&= \sum_{r,s} e^{i\alpha(q_s-q_r)}|\,\langle\psi_q|\psi_{\tilde{q}}|\psi_q|\psi_{\tilde{q}}\rangle\,|^2 \\
|\psi_{\tilde{q}}\rangle &= O^R_{r,s}\,|\psi_q\rangle\,, \quad \tilde{q} = q + q_r - q_s.
\end{aligned} \tag{3.40}$$

where $O^R_{r,s} \equiv |r\rangle\,\langle s|$ is an operator with charge $q_r - q_s$. Neglecting gravity, we would have expected to only receive a contribution when $q_r = q_s$. However, from (3.38) we see that each off-diagonal contribution to the wormhole corresponds to diagrams where particles carrying total charge $q_r - q_s$ propagate through the wormhole, see Figure 3.9.

As a side remark, let us comment on the relation between scattering in the black hole background. We have seen that the same effect which gives the non-zero relative entropy is also responsible for non-zero values of correlation functions which would naively vanish from symmetry considerations. These correlation functions in the bath are closely related to scattering in the black hole background. For example, a 4-pt function in the bath region in the appropriate out-of-time-order configuration could be interpreted as the traversable wormhole signal [26] that violates global symmetry.

3.3.3 Comments on gauge symmetries

While exact global symmetries are not allowed in quantum gravity, gauge symmetries of course can exist. Here we explain the difference between the two from the replica wormhole point of view. To be concrete, we will consider the same set up as in (3.3.1), where the CFT in the flat space region has an $U(1)$ global symmetry, while the difference is that the symmetry is gauged in the gravity region. By the same consideration as in (3.3.1), we would still have

$$\frac{Z_2(\alpha)}{Z_2(0)} = \langle e^{i\alpha Q_{\text{wormhole}}}\rangle_{\mathcal{M}_2}, \quad \text{(with replica wormhole)}. \tag{3.41}$$

However, in the path integral, there will be an integral over the spatial zero mode of the gauge field $\int dA_0 \exp\left(iA_0 \int_\Sigma j^0\right) = \delta(Q_{\text{wormhole}})$ which sets the charge propagating through the wormhole to

zero. Thus we have

$$Z_2(\alpha) = Z_2(0) \tag{3.42}$$

for arbitrary α as enforced by the gauge constraint. Similarly we have $Z_n(\alpha) = Z_n(0)$ for higher replicas as well. It then follows that as $n \to 1$ the relative entropy must vanish. This shows that the symmetry is exact in the boundary description as expected.[11]

We can rephrase this argument slightly differently as follows. Consider $(d+1)$-dimensional electromagnetism on a compact space with no boundary $\Sigma\times$ time. In such a theory, it is clear that the overall charge on Σ must vanish. A net charge would source electric field lines, which have nowhere to go. Integrating Gauss's law $\nabla \cdot E = \rho$ on a spatial manifold Σ with no boundary, we get the constraint that $Q = 0$. If we think of the direction along the Euclidean wormhole as time, this makes it clear that no net charge can propagate through the wormhole. This is true even when the gauge coupling goes to zero. Note that if we view gravity as a gauge theory, the analogous equation says that the ADM Hamiltonian of a closed universe vanishes. If we perform a Kaluza-Klein reduction of some of the compact dimensions along the wormhole, the momenta in the compact dimensions will become gauge charges. So we will get a constraint that the total momentum in each compact dimension vanishes.

3.4 Discussion

It has been suggested for a long time that Euclidean wormholes lead to symmetry violation since charged particles may propagate through the wormhole. Normally, this is an exponentially small effect. The main point of this chapter is showing that when we compute certain relative entropies, the tiny effects add up and give an $\mathcal{O}(1)$ result. This is morally similar to what happens in the Page curve after the Page time, where the tiny correlations in the Hawking radiation bring down the entropy. The novelty here is that instead of computing the entropy, we are computing a relative entropy which involves symmetry generators acting in the bath region, so these quantities tell us about the symmetry violation. Importantly, this effect can be computed with just knowledge of the semi-classical description. Our results are based on the recent paradigm of islands and replica wormholes. The same logic which demands that replica wormholes be included in the computation of the Page curve also demands that they should be included in computations of the relative entropy. Of course, there are unresolved puzzles about the inclusion of Euclidean wormholes in gravitational

[11] It would be interesting to understand how to see the vanishing of the relative entropy directly from the relative entropy formula (3.9), without going through the replica argument.

path integrals [7–9, 112–116, 118, 119]. In some cases, it is known that Euclidean wormholes are computing a disorder average over theories. In these cases, it is possible that the ensemble may have a symmetry group G that each individual theory does not. (This means that the probability distribution over theories is invariant under the action of G on theory-space). It is tempting to say that when there is an ensemble interpretation, if we have a matter theory with global symmetry group G coupled to gravity, the group G should be re-interpreted as the symmetry of the ensemble or the third-quantized theory. See Appendix 3.7.3 for some extra comments.

The relative entropy $S(U_R\rho_{\text{exact}}(R)U_R^\dagger|\rho_{\text{exact}}(R))$ that we considered involves only the density matrix in the bath region. Although it is not a direct observable, it might be able to provide constraints or bounds on physical observables, which requires further investigation. If one can create (or simulate holographically) black holes in the lab, then the relative entropy and its Rényi versions are in principle measurable for observers who have multiple copies of the system. In principle, measuring the say $n = 2$ version of the relative entropy would not be hard (in the sense of computational complexity) if one had access to $n = 2$ copies of the system. We do not need to measure each component of the density matrix independently.

We stress that it is surprising that a quantity that characterizes the nonperturbative violation of global symmetry can be computed in the semiclassical description. This suggests that in some sense, the way gravity violates global symmetry is far from generic. By this we mean the following: if we take some quantum field theory and put a random boundary condition that contains tiny terms that breaks any global symmetries, we would not expect the exact relative entropy to satisfy any formula given by the effective theory where we have erased the tiny terms. Of course, the surprise we are referring to is similar to the surprise that semiclassical gravity knows about the Page curve.

We've focused on idealized situations where there is no global symmetry violation within the semiclassical description. However, the argument we had can be easily generalized to situations where the symmetry has already been violated semiclassically. (For example, there could be Planck-suppressed operators in the Lagrangian which break explicitly any global symmetries.) In such theories, the relative entropy of the semiclassical density matrix will already be non-zero, i.e. $S\left(U_R\rho_{\text{semi}}(R)U_R^\dagger\middle|\rho_{\text{semi}}(R)\right) > 0$. However, if we consider the relative entropy of the exact density matrix, as long as the island still exists, we will have

$$\begin{aligned} S\left(U_R\rho_{\text{exact}}(R)U_R^\dagger\middle|\rho_{\text{exact}}(R)\right) &= S\left(U_R\rho_{\text{semi}}(R\cup I)U_R^\dagger\middle|\rho_{\text{semi}}(R\cup I)\right) \\ &> S\left(U_R\rho_{\text{semi}}(R)U_R^\dagger\middle|\rho_{\text{semi}}(R)\right), \end{aligned} \tag{3.43}$$

where we used the monotonicity of relative entropy from the first line to the second line. The difference between the exact answer and the semiclassical answer will still be order one due to the inclusion of the island. Thus for an evaporating black hole, we will still find a sudden increase of the relative entropy at the Page time, which is a clear signature of the global symmetry violation from nonperturbative effects.

We have emphasized that the symmetry violating effects we are considering are related to Euclidean wormholes. On the other hand, it has long been known that the Hawking process violates symmetries; these previous arguments only involved the naive disconnected geometries. In order to emphasize what our new arguments buy, let us return to the setup discussed in §3.2.3. The matter theory in the bulk is a large c CFT + a probe massive field with a $U(1)$ symmetry and particles of mass m. Let us consider a large black hole which has a Hawking temperature $T \ll m$ that is older than the Page time. Imagine that it formed from neutral matter. Now consider the same experiment in §3.2.3. If a few charged particles fall into the black hole, we might wonder whether the experimentalist which remains outside the hole can see charge violation. Naively to check the argument of [67], she would have to wait an extremely long time for the black hole to evaporate. In particular, she would have to wait for the black hole to become small enough that $T \gg m$, at which point it could start to radiate the $U(1)$ charge.

However, our proposal is that even after a scrambling time, the clever experimentalist can in fact report a violation of charge conservation by measuring off-diagonal components of the density matrix, or if she can control multiple copies of the system, a suitable Rényi relative entropy.

Note that our result is similar in spirit to the result of Hayden and Preskill [83]. Naively, one would have guessed that to recover the information from a diary thrown into the hole, one would have to wait until almost the end of the evaporation process. But [83] showed that one only needs to wait for a scrambling time if the black hole is older than the Page time. Here we are in some sense refining their result and showing that the information in the Hawking radiation recovered after the scrambling time is sufficient to demonstrate that charge conservation is violated.[12]

The connection between the appearance of islands and violation of global symmetries implies that we can see clear signatures of symmetry violation even in situations where charge conservation considerations like those in [67] do not immediately apply. Besides the Hayden-Preskill-like experiment discussed above, an example is the extremal (zero temperature) black hole. Another case

[12]Despite the analogy, we should not confuse the ability to recover the charge of the particle using entanglement reconstruction with the effect that we are talking about here. The violation of charge conservation is more directly related to the fact that naively orthogonal black hole states are not orthogonal in the exact description, this is particularly clear in the model of Appendix 3.7.

is the finite temperature black hole in equilibrium with a bath that never evaporates. We expect our results to generalize to the cosmological setups of [96, 97, 101] where the old arguments would not establish a clear signature of symmetry violation. Of course, the old arguments only involve measuring simple observables like the total charge of the Hawking radiation. In our work, we are discussing the relative entropy which is a much more complicated quantity.

It would be interesting to see if our arguments can be generalized to address questions about gauge charges like the weak gravity conjecture, completeness conjecture, etc. [66, 122].

3.5 Appendix: Nonperturbative correlation between subsets of fields in the Hawking radiation

As we mentioned in §3.2, the island formula implies that there exists substantial correlation between different fields in the Hawking radiation. In this appendix, we illustrate this point in more details. Instead of dividing the Hawking radiation into spatial subsystems, we can also divide the Hawking radiation in a fixed region into subsets of fields. In other words, we are dividing the Hilbert space of some subregion into the tensor product of some smaller Hilbert spaces. This division is at least well defined when the Hawking radiation is in an asymptotic region, where we can neglect the dynamical gravity effects.

Let's consider a region R that is big enough such that we have islands. Now we separate the fields in R into a large portion (labeled by L) and a small portion (labeled by S). We take the number of degrees of freedom in S to be much smaller than L. In the example of §3.2.2, S will be the massless fermion field under the $U(1)$ global symmetry, while L will be the rest of the fields.

We could compute the entropies of the three density matrices $\rho_{\text{exact}}(R_L \cup R_S)$, $\rho_{\text{exact}}(R_L)$ and $\rho_{\text{exact}}(R_S)$ using the QES prescription. For $\rho_{\text{exact}}(R_L \cup R_S)$ and $\rho_{\text{exact}}(R_L)$, we will find an island in the entanglement wedge:

$$S(\rho_{\text{exact}}(R_L \cup R_S)) = \frac{A(\partial I)}{4G_N} + S(\rho_{\text{semi}}(R_L \cup R_S \cup I_L \cup I_S)), \tag{3.44}$$

$$S(\rho_{\text{exact}}(R_L)) = \frac{A(\partial I)}{4G_N} + S(\rho_{\text{semi}}(R_L \cup I_L \cup I_S)), \tag{3.45}$$

where we neglected the change in the position of the island due to the addition of the field S, since it will only give subleading corrections. Importantly, in (3.45), one gets contributions from both I_L and I_S in the island. This is because the island should be for all the fields, as it follows from the

replica wormhole geometry which all the fields live on. (In other words, the emergent twist operators act on all fields.) For $\rho_{\text{exact}}(R_S)$, since its semi-classical entropy is small, there will not be an island:

$$S(\rho_{\text{exact}}(R_S)) = S(\rho_{\text{semi}}(R_S)). \tag{3.46}$$

Combining (3.44) to (3.46), one finds

$$I_{\text{exact}}(R_S : R_L) = I_{\text{semi}}(R_S : R_L \cup I_L \cup I_S). \tag{3.47}$$

In particular, by the monotonicity of mutual information, one has

$$I_{\text{exact}}(R_S : R_L) > I_{\text{semi}}(R_S : R_L), \tag{3.48}$$

and the difference will be an order one amount. The difference comes from nonperturbative effects that are not present in the semiclassical description. Notice that in the vacuum state, the semiclassical mutual information $I_{\text{semi}}(R_S : R_L)$ will vanish in the limit that S and L are decoupled in the semiclassical theory. On the other hand, the right hand side of (3.47) reduces to $I_{\text{semi}}(R_S : I_S)$, which is nonzero and finite even when the coupling between S and L is negligible in the semiclassical description.

3.6 Appendix: Details on the relative entropy for massless free fermion in 2D

As derived in [107,123], for a two dimensional free massless fermion in the vacuum state, the modular Hamiltonian of the union of $I = (a_1, b_1)$ and $R = (a_2, b_2)$ takes the following form:

$$\begin{aligned} K &= K_{\text{loc}} + K_{\text{noloc}}, \\ K_{\text{loc}} &= 2\pi \int_{I\cup R} \mathrm{d}x\, \omega'(x)^{-1} T(x), \\ K_{\text{noloc}} &= i2\pi \int_{I\cup R} \mathrm{d}x\, \psi^\dagger(x) \frac{(b_1-a_1)(a_2-b_1)(b_2-a_1)(b_2-a_2)}{\omega'(x)^2 (x-a_1)(x-a_2)(x-b_1)(x-b_2)} \\ &\quad \frac{1}{x(a_1+a_2-b_1-b_2)+(b_1b_2-a_1a_2)} \psi(\bar{x}), \end{aligned} \tag{3.49}$$

where $T(x) = \frac{1}{2}\left[i\partial_x\psi^\dagger(x)\psi(x) - \psi^\dagger(x)i\partial_x\psi(x)\right]$ is the energy density operator, and $\omega(x), \bar{x}$ are given by

$$\begin{aligned}\omega'(x) &= \frac{1}{x-a_1} + \frac{1}{x-a_2} - \frac{1}{x-b_1} - \frac{1}{x-b_2} \\ \bar{x} &= \frac{a_1a_2\,(x-b_1-b_2) - b_1b_2\,(x-a_1-a_2)}{x\,(a_1+a_2-b_1-b_2) + (b_1b_2-a_1a_2)}\end{aligned}. \tag{3.50}$$

The formulas are the same for ψ_+ and ψ_-, so we've omitted the chirality of the fermion. As discussed in (3.20), we are interested in computing the expectation value $\langle 0|\, K_{\text{noloc}}\, |0\rangle$. Using

$$\langle 0|\, \psi^\dagger(x)\psi(y)\, |0\rangle = \frac{i}{2\pi}\frac{1}{x-y}, \tag{3.51}$$

when $x \neq y$, we can compute the integral in (3.49), and find

$$\langle 0|\, K_{\text{noloc}}\, |0\rangle = -\frac{1}{4}\left(1 + \frac{(2\eta-1)\arctan\sqrt{\frac{\eta}{1-\eta}}}{\sqrt{\eta(1-\eta)}}\right), \tag{3.52}$$

where the cross ratio η is defined as

$$\eta \equiv \frac{(b_1-a_1)(b_2-a_2)}{(a_2-a_1)(b_2-b_1)}. \tag{3.53}$$

When computing the integral, it is convenient to set $\{a_1, b_1, a_2, b_2\} = \{0, \eta, 1, \infty\}$. (3.52) is the contribution from one of the chiral modes, so to get the total contribution, we multiply it by two.

3.7 Appendix: Global symmetry in JT gravity + EoW branes

In this section, we explore issues related to global symmetry violations in the "West coast" model [13], namely JT gravity with end-of-the-world (EoW) branes. We refer the reader to [13] for definitions and conventions. Some similar comments could be made in the simplified model of [116].

3.7.1 coarse-grained entropy

A natural question is whether one can define a coarse-grained density matrix $\tilde{\rho}_R$ of the radiation, whose entropy does not follow the Page curve but instead is given by the naive, non-minimal quantum extremal surface. In the setup of [13], we want the coarse-grained entropy to be $\approx \log k$, even when k is bigger than e^{S_0}. Here we point out that if we have a large symmetry group, we can use

group-averaging to define a coarse-grained density matrix:

$$\tilde{\rho}_R = \int dg\, U(g)\rho_R U(g^{-1}). \tag{3.54}$$

Note that the group-averaging is a completely positive trace preserving (CPTP) map. It takes any density matrix into a density matrix which is invariant under the symmetry group. The CPTP property ensures that

$$S(\tilde{\rho}_R) \geq S(\rho_R). \tag{3.55}$$

For a simple model where we can explicitly compute $S(\tilde{\rho}_R)$, consider the west coast model [13]. From the bulk (semi-classical) perspective, the model is defined by JT gravity + end-of-the-world (EoW) branes. The EoW branes carry an index i such that

$$\langle i|j\rangle_{\text{bulk}} = \delta_{ij}. \tag{3.56}$$

This equation should not be confused with a statement about the actual boundary states that correspond to these branes, which are not exactly orthonormal. The content of this equation is simply to prescribe rules for evaluating the bulk path integral. The point we would like to make here is that the bulk theory at the semi-classical level has a $U(k)$ global symmetry, where the EoW branes transform in the fundamental.

Now let us consider the replica wormhole computation of $S(\tilde{\rho}_R)$. We will start by computing the Rényi entropies. Let us consider the 2-replica computation of $t_r(\tilde{\rho}_R^2)$. Since the overall state of the black hole and the radiation is an entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{k}} \sum_{i=1}^{k} |\psi_i\rangle_{\text{B}} |i\rangle_{\text{R}}\,, \tag{3.57}$$

acting with the symmetry generator on the radiation $|i\rangle \to U|i\rangle_{\text{R}} = U_{ji}|j\rangle_{\text{R}}$ is equivalent to acting on the black hole states $|\psi_i\rangle_{\text{B}} \to \sum_j U_{ji}|\psi_j\rangle_{\text{B}}$. Clearly if we consider the disconnected saddle, the symmetry generators act trivially. But for the connected saddle, we actually get the constraint $i = j$.

So whereas

$$\begin{aligned}\mathrm{Tr}\left(\rho_{\mathrm{R}}^2\right) &= \frac{kZ_1^2 + k^2 Z_2}{\left(kZ_1\right)^2} = \frac{1}{k} + \frac{Z_2}{Z_1^2}, \\ \mathrm{Tr}\left(\tilde{\rho}_{\mathrm{R}}^2\right) &= \frac{kZ_1^2 + kZ_2}{\left(kZ_1\right)^2} = \frac{1}{k}\left(1 + \frac{Z_2}{Z_1^2}\right).\end{aligned} \tag{3.58}$$

Notice that the connected saddle never dominates, no matter how big k is.

In fact, from the boundary point of view we can compute the exact density matrix after group averaging:

$$\begin{aligned}\rho_{\mathrm{R}} &= \frac{1}{k}\sum_{i,j=1}^{k} |j\rangle \langle i|_{\mathrm{R}} \langle\psi_i|\psi_j\rangle_{\mathrm{B}}\,. \\ \tilde{\rho}_{\mathrm{R}} &= \frac{1}{k}\sum_{i,j=1}^{k} \int dU\, U\,|j\rangle \langle i|_{\mathrm{R}}\, U^\dagger \langle\psi_i|\psi_j\rangle_{\mathrm{B}} \\ &= \frac{1}{k}\sum_{i} |i\rangle \langle i|_R \langle\psi_i|\psi_i\rangle \\ &\approx \frac{1}{k}\sum_{i} |i\rangle \langle i|_R\,.\end{aligned} \tag{3.59}$$

Note that this is almost a completely mixed density matrix. (It is not perfectly mixed because the norms of the states $\langle\psi_i|\psi_i\rangle$ have small fluctuations. However, to leading order in e^{-S_0}, this effect is negligible.) Notice that at large k, this gives a concrete meaning to a sub-dominant QES.

3.7.2 Relative entropy

Let us start with computing the relative purity $t_r U\rho U^\dagger\rho$ in the West coast model. Up to a normalization,

$$\begin{aligned}t_r U\rho U^\dagger\rho = U_{ij}\rho_{jk}U^\dagger_{kl}\rho_{li} &= \\ &= U_{ij}U^\dagger_{kl}\langle\psi_j|\psi_k\rangle\langle\psi_l|\psi_i\rangle \\ &= \sum_{i,k}\delta_{ik}Z_1^2 + U_{ii}U^\dagger_{kk}Z_2\end{aligned} \tag{3.60}$$

Notice that the unitary does not affect the disconnected contribution. Furthermore, when we set the group element to the identity $U = I$, we recover the result in [13].

Now we can also use the QES prescription to compute the relative entropy. At small k, there is

no island and the QES prescription gives 0. Since the semi-classical density matrix is a pure state:

$$\rho_{\text{semi}} = \sum_{i,j} (|\psi_i\rangle \langle\psi_j| \otimes |i\rangle \langle j|) \tag{3.61}$$

the relative entropy would be infinite if $U \neq 1$:

$$S(U_R \rho_{\text{semi}}(R \cup I) U_R^\dagger | \rho_{\text{semi}}(R \cup I)) = \infty, \quad U \neq 1. \tag{3.62}$$

In the exact description, the density matrix of the radiation is not pure. We therefore expect a finite answer, so the relative entropy formula is not immediately applicable. It would be interesting to apply the planar resummation techniques of [13] to compute the finite answer.

3.7.3 Global symmetry in the ensemble

Here we would like to make some distinctions about the various kinds of symmetries in JT gravity + EoW branes.

1. The simplest symmetry is the semi-classical symmetry of the theory. This means that if we only include the trivial, disconnected topologies in the path integral, we will have a matter theory which has an exact global $U(k)$ symmetry. In particular, the overlaps between different brane states $\langle\psi_i|\psi_j\rangle = \delta_{ij}$ are exactly invariant under $|\psi_i\rangle \to \sum_j U_{ij} |\psi_j\rangle$.

2. If we include higher topologies in the path integral, it has been shown that the path integral is dual to an ensemble of theories. The ensemble is a probability distribution over the Hamiltonian (a random matrix) and a set of pure states (k random vectors). If we consider one instance of the disorder average, the resulting theory will generically not have a $U(k)$ global symmetry. For example, the inner products between different pure states will be slightly different.

3. Nevertheless, the ensemble has a $U(k)$ symmetry. Note that unlike in point 1, the $U(k)$ symmetry acts in *theory space*. It is not the symmetry group of an individual theory. Specifying each theory means specifying the values of the Hamiltonian and the pure state. The statement about the symmetry of the ensemble is that if we take such one instance of the ensemble t_1, and then act with $g \in U(k)$ on the random vectors to obtain a new theory $t_2 = g(t_1)$, the probability $P(t_2) = P(t_1)$. If we think of the ensemble as a third-quantization [116], we might say that $U(k)$ is a symmetry in the third-quantized theory.

Chapter 4

The Entanglement Wedge of Unknown Couplings

The black hole interior is a mysterious region of spacetime where non-perturbative quantum gravity effects are sometimes important. Despite the importance of such non-perturbative effects, work from the past few decades supports the idea that "plain" gravity (e.g. a sum over metrics and possibly a few light fields) knows a lot about fine-grained quantum information. The paradigmatic examples of this include the geometrization of von Neumann entropy in holographic systems, including its suitable generalization to generate the Page curve of Hawking radiation of an evaporating black hole [73, 74]. Such features were thought to require a UV complete theory of gravity, such as string theory, and obtaining them from gravity came as a pleasant surprise.

Gravity, however, doesn't know everything. It appears to know of the underlying random unitary dynamics, but it fails to pin down a particular realization of those dynamics. In particular, while gravity is able to reproduce arbitrary moments of the signal drawn from an ensemble of random dynamics, it fails to capture the large fluctuations that come with a given realization. This is a pretty big miss since the size of those fluctuations is of order the signal.

A related point is that whereas the gravity calculations seem reliable for some universal quantities, it seems to know only statistical properties about the non-perturbative effects that are highly theory-dependent. For example, the spectral form factor, which at large times probes the detailed energy spectrum of the black hole, is a highly erratic function that depends sensitively on the couplings [7, 114]. Similarly, the black hole S-matrix that governs the formation and evaporation of a black hole is suspected to be an erratic and possibly pseudo-random matrix [124]. Presumably

a precise computation of such quantities from the bulk point of view will involve strings, branes, half-wormholes [125], etc, and the answers would depend on the particular string vacua.

The dependence of these non-perturbative effects on the couplings of the theory suggests that the interior of the black hole is in some sense highly theory dependent. The goal of our work will be to make this more precise in the context of bulk reconstruction. We will consider models which admit an ensemble of boundary theories parameterized by a set of boundary couplings, and analyze the sensitivity of bulk reconstruction on the level of precision in specifying those couplings. This means we will look for instances where the reconstruction fails, and map out which bulk regions are most sensitive to this. Hence, those bulk regions require exquisite knowledge of the couplings to reconstruct.

Note added: as we were finishing this work, we became aware of work by Qi, Shangnan, and Yang [126]. We have arranged to coordinate our preprints. See also [127].

The first point we make precise is the notion of "knowing the couplings." Suppose we have a system whose Hamiltonian depends on a set of parameters $\mathbf{J} = \{\lambda_i\}$. For example, in SYK it is natural to choose $\mathbf{J} = \{J_{ijkl}\}$ to be the set of random couplings. However, in general the couplings could be non-random, for example we could consider $\mathcal{N} = 4$ and take $\mathbf{J} = \{g_{YM}^2\}$. We denote the state of the system prepared with those couplings as

$$|\psi; \mathbf{J}\rangle_{\mathsf{sys}} \,. \tag{4.1}$$

As a concrete example, we could take $|\psi; \mathbf{J}\rangle_{\mathsf{sys}} = |\beta + i2T; \mathbf{J}\rangle_{\mathsf{sys}}$ to be the thermofield double state associated to the Hamiltonian $H(\mathbf{J})$ at some temperature β and some Lorentzian time T. Let's imagine that these couplings are drawn form some distribution $P(\mathbf{J})$. We can keep track of this by using a standard method in quantum information theory of entangling the system with a reference system labelled by the couplings

$$|\Psi\rangle_{\mathsf{sys} \cup \mathsf{journal}} = \sum_{\mathbf{J}} \sqrt{P(\mathbf{J})}\, |\psi; \mathbf{J}\rangle_{\mathsf{sys}} |\mathbf{J}\rangle_{\mathsf{journal}}. \tag{4.2}$$

We call the reference system journal since in the SYK context it records the J's. Tracing over the journal, we get a density matrix

$$\rho_{\mathsf{sys}} = \sum_{\mathbf{J}} P(\mathbf{J}) \rho_\psi(\mathbf{J}) = \langle \rho_\psi \rangle_{\mathbf{J}} \,, \quad \rho_\psi(\mathbf{J}) = |\psi; \mathbf{J}\rangle\langle\psi; \mathbf{J}|_{\mathsf{sys}} \,. \tag{4.3}$$

This state represents the situation where we have no information about the couplings. This ignorance is captured by the non-zero von Neumann entropy between sys and journal. If we think of $\mathbf{J}$ as parameters in a disorder average, we may use condensed matter jargon and refer to the von-Neumann entropy as the the *annealed* entropy $S\left(\langle\rho_\psi\rangle_\lambda\right)$. In the above discussion we started with a pure state density matrix ρ_ψ but after averaging over $\mathbf{J}$ we get a mixed state. As a slight generalization, we could consider any subsytem of sys, and get a similar formula, where ρ_ψ is replaced by some partial trace of $|\psi;\mathbf{J}\rangle\langle\psi;\mathbf{J}|_{\mathsf{sys}}$.

The main question we'd like to answer is the following: How much of the bulk can be reconstructed given the density matrix of the system after tracing out the reference? Or in other words, how much of the bulk is contained within the entanglement wedge of sys? In a theory with a semi-classical holographic dual, the answer is given by the quantum extremal surface (QES) formula, which states that the boundary of the entanglement wedge is given by the QES responsible for the von Neumann entropy $S(\mathsf{sys})$. An alternative version of this question is: how much of the bulk is contained within the entanglement wedge of the reference? Since the $S(\mathsf{sys}) = S(\mathsf{journal})$ for a pure state, the entanglement wedge is simply the complement of the entanglement wedge of sys.

Let us make some preliminary comments on what we mean by bulk reconstruction. First notice that semi-classically, the entropy $S(\mathsf{sys}) > 0$. The bulk matter is not pure, since it is entangled with journal. This semi-classical entropy can "pollute" the bulk and lead to problems with reconstructing any operator using traditional methods like HKLL [47]. Any matter in the bulk will generically interact (at least weakly) with the fields that are sourced by the couplings we turn on at the boundary of AdS; this will cause problems with reconstruction methods such as HKLL. We will not be discussing such semiclassical problems in this chapter. Instead, we will ask for bulk reconstruction in the modern sense of entanglement wedge reconstruction [80, 128]. The failure of our ability to do reconstruction will be a non-perturbative effect, signaled by replica wormholes and the appearance of an island.

In the above discussion, we entangled sys to journal in such a way that the global state is pure. However, one could also consider a setup where instead of entangling sys to a journal, we instead classically correlate sys to a "pointer" system ptr:

$$\rho_{\mathsf{sys}\cup\mathsf{ptr}} = \sum_{\mathbf{J}} P(\mathbf{J})\rho_\psi(\mathbf{J}) \otimes |\mathbf{J}\rangle\langle\mathbf{J}|_{\mathsf{ptr}}\,. \tag{4.4}$$

This mixed state would be the appropriate description of a setup where Alice flips some coins and uses the outcome to decide which couplings to prepare the system in. In fact, following the standard

discussion of decoherence/measurement theory, this state can be purified by adding an auxilary system env. Often one adopts the interpretation that the system ptr is a pointer or measurement device, and env is the environment. The purification of this state is simply (4.2), where $\mathsf{journal} = \mathsf{ptr} \cup \mathsf{env}$ and $|\mathbf{J}\rangle_{\mathsf{journal}} = |\mathbf{J}\rangle_{\mathsf{ptr}} |\mathbf{J}\rangle_{\mathsf{env}}$. Tracing over env gives (4.4). Alternatively, we can go the other direction: we start from the global pure state of (4.2) and perform a complete measurement in the λ basis:

$$|\Psi\rangle\langle\Psi|_{\mathsf{sys}\cup\mathsf{journal}} \to \sum_{\mathbf{J}} \Pi_{\mathbf{J}}^{\mathsf{journal}} |\Psi\rangle\langle\Psi|_{\mathsf{sys}\cup\mathsf{journal}} \Pi_{\mathbf{J}}^{\mathsf{journal}}. \tag{4.5}$$

Then relabelling $\mathsf{journal} \to \mathsf{ptr}$, we obtain (4.4). Thus including the ptr system makes precise the idea of having classical (e.g. decohered) knowledge of the couplings.

To ask about bulk reconstruction given perfect classical knowledge of the couplings is to ask for the entanglement wedge of $\mathsf{sys}\cup\mathsf{ptr}$. Following the QES rules, we should compute the von Neumann entropy of $\mathsf{sys}\cup\mathsf{ptr}$:

$$S(\mathsf{sys}\cup\mathsf{ptr}) = \sum_{\mathbf{J}} P(\mathbf{J}) S(\rho_\psi(\mathbf{J})) - \sum_{\mathbf{J}} P(\mathbf{J}) \log P(\mathbf{J}). \tag{4.6}$$

The first term is the von Neumann entropy, averaged over the couplings. If $\rho_{\mathbf{J}}$ is pure, this vanishes. But this formula also applies to an initially mixed state, or a subsystem, e.g., we could take sys to be the left side of the thermofield double. Then the first term would be the disorder-averaged thermal entropy, known as the *quenched* entropy, $\langle S(\rho_\psi)\rangle_\lambda$.

The second term in (4.6) is the Shannon entropy of the couplings $S(\mathsf{ptr})$. It is an entirely classical entropy and does not play a role in determining the QES. To see this more explicitly, note that (4.6) is a formula for the exact entropy, but in a holographic system a similar formula would hold for the semi-classical matter entropy of any subset A of the bulk:

$$S_{\text{matter}}(A\cup\mathsf{ptr}) = \sum_{\mathbf{J}} P(\mathbf{J}) S(\rho_{\mathbf{J}}^A) - \sum_{\mathbf{J}} P(\mathbf{J}) \log P(\mathbf{J}), \quad \rho_{\mathbf{J}} = \Pi_{\mathbf{J}} \rho_A^{\text{semi}} \Pi_{\mathbf{J}}. \tag{4.7}$$

The second term is independent of the choice of A, so it will never contribute to the derivative of the matter entropy (relevant for the QES). The conclusion is that the entanglement wedge of $\mathsf{sys}\cup\mathsf{ptr}$ can be diagnosed by computing $S(\mathsf{sys}|\mathsf{ptr}) = S(\mathsf{sys}\cup\mathsf{ptr}) - S(\mathsf{ptr})$, which is simply the quenched

entropy. In this context, we have a QES formula for the conditional entropy:

$$S(\mathsf{sys}|\mathsf{ptr}) = \min \mathrm{ext}_I \left[\mathrm{Area}(\partial I)/(4G) + S_{\mathsf{matter}}(I|\mathsf{ptr})\right]. \tag{4.8}$$

To summarize, if we are interested in the entanglement wedge of sys with no knowledge of the couplings whatsoever, we should compute the annealed entropy $S(\langle\rho_\psi\rangle_{\mathbf{J}})$. If we are interested in bulk reconstruction when we have perfect classical knowledge of the couplings $\mathsf{sys}\cup\mathsf{ptr}$, we should compute the quenched entropy $\langle S(\rho_\psi)\rangle_{\mathbf{J}}$.

Note that whether we choose to entangle sys with a reference or classically correlate sys to a pointer makes no difference for the density matrix ρ_{sys}. A crucial difference is for the complement of sys: the density matrix of ρ_{ptr} is diagonal in the λ basis whereas ρ_{journal} has off-diagonal elements. Furthermore, the entanglement wedge of journal can contain an island (as we will show), whereas a quick argument [129] rules out any possible island for the entanglement wedge of ptr: for the pointer system, the matter conditional entropy is positive $S(A|\mathsf{ptr}) = S(A\cup\mathsf{ptr}) - S(\mathsf{ptr}) \geq 0$ for any bulk region A, so adding any the bulk region A can only increase the entropy. Therefore, any island QES must be non-minimal.

4.0.1 Both known and unknown couplings

More generally, we can imagine that all the couplings in journal can be divided into known and unknown. We write $\mathbf{J} = \{\kappa, \mu\}$ with κ the known parameter(s) and μ the unknown, $\mathsf{journal} = \mathsf{known}\cup\mathsf{unknown}$. The global state is a density matrix on $\mathsf{sys}\cup\mathsf{journal}$:

$$\rho = \sum_{\kappa,\mu,\mu'} \sqrt{P(\mu,\kappa)P(\mu',\kappa)}\, |\psi;\kappa,\mu\rangle\langle\psi;\kappa,\mu'|_{\mathsf{sys}} \otimes |\kappa\rangle\langle\kappa|_{\mathsf{known}} \otimes |\mu\rangle\langle\mu'|_{\mathsf{unknown}} \tag{4.9}$$

Here we are treating the known couplings as "classical" pointer states ptr, whereas the unknown parameters are entangled. The reduced density matrices

$$\rho_{\mathsf{journal}} = \sum_{\mu,\mu',\kappa} |\mu\rangle\langle\mu'|_{\mathsf{unknown}}\, |\kappa\rangle\langle\kappa|_{\mathsf{known}} \sqrt{P(\mu,\kappa)\, P(\mu',\kappa)}\, \langle\psi;\mu,\kappa|\psi;\mu',\kappa\rangle_{\mathsf{sys}} \tag{4.10}$$

$$\rho_{\mathsf{known}} = \sum_{\kappa} |\kappa\rangle\langle\kappa|_{\mathsf{known}}\, P(\kappa), \quad P(\kappa) = \sum_{\mu} P(\mu,\kappa) \tag{4.11}$$

We see that the density matrix of known is completely classical as before. Once we include "known" couplings, the global state is not pure. Therefore, $S(\mathsf{journal}) \neq S(\mathsf{sys})$. Nevertheless, a close analog

is given by the conditional entropies:

$$\boxed{S(\mathsf{sys}|\mathsf{known}) = S(\mathsf{unknown}|\mathsf{known}).} \tag{4.12}$$

This follows from (4.10). As we have already argued, the entanglement wedge of $\mathsf{sys} \cup \mathsf{known}$ may be diagnosed by computing the conditional entropy $S(\mathsf{sys} \cup \mathsf{known})$, and similarly for the entanglement wedge of $\mathsf{known} \cup \mathsf{unknown}$. The equality of the above conditional entropies essentially shows that the entanglement wedges are in fact complementary even though the global state is not pure.

The fact that we are including the known couplings when computing the entanglement wedges reflects the fact that the known couplings are completely classical, so there is no obstruction to cloning. The experimentalist simply publishes in a journal the values of the couplings in which she prepared her system.

4.1 The entanglement wedge with uncertain couplings

In this section we analyze the candidate entanglement wedges of the reference keeping track of the unknown couplings. This will be done in a bulk model of JT gravity [15–18] coupled to general conformal matter:

$$-I[g] = -S_0\chi + \int_{\Sigma_2} \frac{\phi}{4\pi}(R+2) + \frac{\phi_b}{4\pi}\int_{\partial\Sigma_2} 2K + \log Z_{\mathsf{CFT}}[g], \tag{4.13}$$

where χ is the Euler characteristic. We will specialize to a concrete BCFT when needed.

The couplings in this model will be the choice of CFT boundary conditions along the AdS boundary, which we label $|J\rangle$ and interpret as arising from a holographic boundary Hamiltonian H_J. We require the state $|J\rangle$ to be a conformally invariant Cardy state. Here J could be a discrete variable (e.g., if the CFT is a minimal model) or continuous. We will study the state given by entangling the thermofield double to the $\mathsf{journal}$ as in (4.2):

$$|\Psi\rangle_{\mathsf{sys}\cup\mathsf{journal}} = \sum_J \sqrt{P(J)}|\beta + 2iT, J\rangle_{\mathsf{sys}}|J\rangle_{\mathsf{journal}} \tag{4.14}$$

where the time-evolved thermofield double $|\beta + 2iT, J\rangle_{\mathsf{sys}} = e^{-iH_L T}|\beta, J\rangle_{\mathsf{sys}}$. We will frequently drop the $2iT$ and think of β as a complex number. The reduced density matrix

$$\rho_{\mathsf{journal}} = \sum_{J,J'} \sqrt{P(J)P(J')}\langle\beta, J'|\beta, J\rangle\ |J\rangle\langle J'|_{\mathsf{journal}}. \tag{4.15}$$

To compute the entanglement wedge of the journal, we will use the QES formula [77, 78]:

$$S(\mathsf{journal}) = \min\left\{\mathrm{ext}_I\left[\sum_{\partial I}\phi(\partial I) + S_m(I\cup\mathsf{journal})\right]\right\}. \tag{4.16}$$

Here we are instructed to compute the generalized entropy of extremal islands, and then pick the smallest extremized entropy. Note that for marginal deformations of the boundary theory, we expect the dilaton to be independent of the boundary state J. Hence the non-trivial calculation is just the matter entropy in (4.16). We will compute this in some special cases in this section.

Before getting into the weeds, let us clarify somewhat the interpretation of the calculation. We are thinking of the above theory as an effective theory of the bulk. It is not UV complete due to divergences when wormholes get narrow [8]. We do not know what the precise boundary dual of

this bulk model is (or even if it exists). The simplest possibility is that the boundary dual is a theory with a small number of couplings. For example, one might be able to embed such a setup in a traditional higher dimensional example of AdS/CFT by considering near extremal black holes in AdS. In this case, we assume that all of the couplings besides J are fixed, and the state of the combined system is given by (4.14).

However, it is also possible that the gravity description only arises after a disorder average over other random couplings, like in pure JT gravity [8] or in SYK [22]. In this case, the above (4.14) and (4.15) are not quite right; instead, we assume that the additional couplings are "known" while the J couplings are unknown, and use (4.9). The density matrix of $\mathsf{sys} \cup \mathsf{unknown}$ would be

$$\rho_{\mathsf{sys}\cup\mathsf{unknown}} = \sum_{\kappa,J,J'} P(\kappa)\sqrt{P(J)P(J')}\,|\beta;\kappa,J\rangle\langle\beta;\kappa,J'|_{\mathsf{sys}} \otimes |J\rangle\langle J'|_{\mathsf{unknown}} \tag{4.17}$$

Then following (4.7) we would interpret the QES computation as giving the conditional entropy $S(\mathsf{unknown}|\mathsf{known})$:

$$S(\mathsf{unknown}|\mathsf{known}) = \min\left\{\mathrm{ext}_I\left[\sum_{\partial I}\phi(\partial I) + \langle S_{\mathrm{m}}(I\cup\mathsf{unknown})\rangle\right]\right\}. \tag{4.18}$$

Here we have used $S_{\mathrm{matter}}\left(I\cup\mathsf{unknown}|\mathsf{known}\right) = \langle S_{\mathrm{matter}}\left(I\cup\mathsf{unknown}\right)\rangle$, where $\langle\cdots\rangle$ is a disorder average with respect to the known couplings, see (4.7) and (4.8).

4.1.1 Inconsistency of the trivial surface

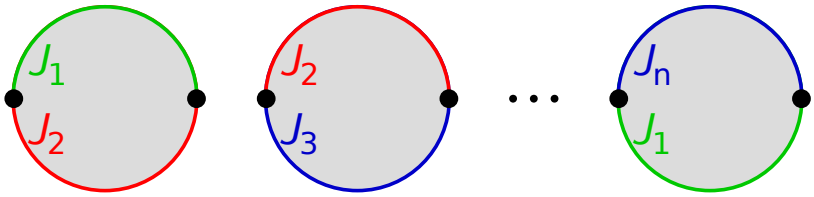
Here we analyze the contribution from the trivial surface, namely where the entanglement wedge of the reference doesn't include any part of the gravitational system. The entire bulk is encoded on the boundary. We will find that this contribution to the entanglement between the boundary and the reference leads to an ever growing entropy, producing to a Hawking-like information paradox. This signals that at late times, the couplings should contain an island, to prevent the entropy from growing.

To compute the entropy of the trivial surface, all we need is the bulk matter entropy. The most straightforward way of getting this is by computing the entropy of the density matrix (4.15) using the replica trick while freezing the gravitational saddle to be the product of Euclidean cigars on the n copies. The n-th Renyi entropy[1] of the journal is therefore given by the path integral with the

[1] In the standard quantum information literature, the Renyi's usually refer to $S_n = \frac{1}{1-n}\log R_n$; here we will refer to both S_n and R_n as Renyi entropy.

boundary conditions $Z_n = t_r \rho^n_{\text{unknown}}$

$$Z_n = \quad = \int \prod_{i=1}^{n} P(J_i)\, dJ_i \prod_{i=1}^{n} \langle \beta, J_i | \beta, J_{i+1} \rangle, \tag{4.19}$$



where $J_{n+1} = J_1$.

In the above drawing, we are supposed to sum over all the indices, weighted by the probability distribution $P(J_i)$. The overlaps can be represented as the BCFT disk partition function in the presence of boundary changing operators.

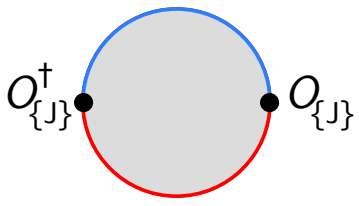
$$\langle \beta, J_i | \beta, J_{i+1} \rangle = \left\langle O_{J_i, J_{i+1}}(0) O_{J_{i+1}, J_i}(\tau) \right\rangle_{\text{disk}}, \quad \tau = \beta/2. \tag{4.20}$$

where $O_{J_i, J_{i+1}}(\tau)$ changes the boundary condition by $J_{i+1} - J_i$ as τ is crossed along the path integral contour. They satisfy $\left(O_{J_i, J_{i+1}}(\tau)\right)^\dagger = O_{J_{i+1}, J_i}(-\tau)$, and behave as primary operators with a dimension that depends on the boundary conditions $\Delta[J_{i+1}, J_i]$, and their two point function is

$$\left\langle O_{J_i, J_{i+1}}(0) O_{J_{i+1}, J_i}(\tau) \right\rangle_{\text{disk}} = \left[\frac{\pi \epsilon}{\beta \sin(\pi \tau / \beta)} \right]^{2\Delta[J_{i+1}, J_i]} \tag{4.21}$$

Here the ϵ comes from the Weyl factor evaluated on the boundary. Instead of consisting a product of n disks, we may equivalently consider the quotient picture where we instead consider the n-fold tensor product of the CFT on a single disk, with only 2 boundary condition changing operators:

$$Z_n = \tag{4.22}$$



In this picture, the black dot represents a composite boundary condition changing operator $O_{\{J\}} = O_{J_1, J_2} \otimes O_{J_2, J_3} \otimes \cdots \otimes O_{J_n, J_1}$ which shifts the boundary condition $\{J_1, J_2 \cdots J_n\} \to \{J_2, J_3 \cdots J_1\}$. This quotient picture is a bit overkill for the no-island computation, but we are introducing it now since it will be crucial when the QES is non-trivial.

Since we are interested in the time evolution of the entropy, we need to consider the Renyi computation in the time evolved state $|\beta + 2iT, J\rangle = e^{-iH_J^L T}|\beta, J\rangle$. We can achieve this by setting $\tau = \beta/2 + iT$. The overlaps become

$$\langle \beta + 2iT, J_i | \beta + 2iT, J_{i+1}\rangle = \left[\frac{\pi\epsilon}{\beta\cosh(\pi T/\beta)}\right]^{2\Delta[J_{i+1},J_i]} \tag{4.23}$$

Notice that all off-diagonal matrix elements are decaying to zero at large T; only the diagonal terms where $\Delta = 0$, e.g., $J_i = J_{i+1}$ do not decay. This implies that at late times the matter is close to maximally mixed.

Putting this back into the Renyi entropy, we get

$$Z_n = \int \prod_{i=1}^{n} dJ_i\, p(J_i) \ \left[\frac{\pi\epsilon}{\beta\cosh(\pi T/\beta)}\right]^{\sum_{i=1}^{n} 2\Delta[J_{i+1},J_i]} \tag{4.24}$$

So the only dynamical input we need from the particular CFT to evaluate these Renyi entropies is the boundary dimensions $\Delta[J_{i+1}, J_i]$. In Appendix 4.6 we consider a generic BCFT with 2 boundary states. Here we will consider a non-compact boson, with action

$$S = \frac{1}{2\pi\alpha'}\int d^2z \partial X \bar{\partial} X \tag{4.25}$$

The boundary conditions we consider are simply Dirichlet conditions on the free field, labeled by X. The boundary condition changing operator[2] changes the value of the field from X_1 to X_2. Its dimension is equal to the energy on the theory on the strip, which is given by

$$\Delta_b = \frac{1}{4\pi\alpha'}\int_0^\pi (\partial_\sigma X)^2 d\sigma = \frac{1}{\alpha'}\left(\frac{X_1 - X_2}{2\pi}\right)^2 \tag{4.26}$$

This is familiar from string theory. The mass of a bosonic open string of level $N = 1$ which stretches between two D-branes is $M^2 = (X_1 - X_2)^2/(2\pi\alpha')^2$. We will take a Gaussian measure over the boundary conditions $p(X) = \frac{m}{\sqrt{2\pi}} e^{-m^2X^2/2}$. In terms of the boundary dual, we are considering an ensemble of boundary Hamiltonians parameterized by a marginal coupling J that changes the Dirichlet condition of the bulk field X; the ensemble for this boundary coupling is Gaussian distributed.

[2] A quick way to see that changing Dirichlet conditions behaves as a local boundary primary is to use T-duality, which relates this setup to the insertion of a vertex operator with momentum k on the boundary (with standard Neumann condtions.) We thank Juan Maldacena for pointing this out.

Then Equation (4.24) becomes

$$\begin{aligned} Z_n &= \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \prod_{i=1}^{n} dx_i \exp\left[-\frac{1}{2}\left(x_i^2 + a(x_i - x_{i+1})^2\right)\right] \\ &= (\det M_n)^{-1/2}, \quad a = \frac{1}{2\pi^2\alpha' m^2} \log\left[\frac{\beta}{\epsilon\pi}\cosh\frac{\pi T}{\beta}\right] \end{aligned} \tag{4.27}$$

Here M_n is an $n \times n$ matrix; to take $n = 5$ for example:

$$M_5 = \begin{pmatrix} 2a+1 & -a & 0 & 0 & -a \\ -a & 2a+1 & -a & 0 & 0 \\ 0 & -a & 2a+1 & -a & 0 \\ 0 & 0 & -a & 2a+1 & -a \\ -a & 0 & 0 & -a & 2a+1 \end{pmatrix} \tag{4.28}$$

Here the determinant is a polynomial in a:

$$\det M_n = \sum_{k=0}^{n-1} \binom{2n-k}{k} \frac{n}{2n-k} a^k \tag{4.29}$$

This sum can be analytically continued to complex n:

$$\det M_n = \left(d_+^2\right)^n + \left(d_-^2\right)^n - 2a^n, \quad d_\pm = \frac{1 \pm \sqrt{4a+1}}{2} \tag{4.30}$$

One can show that for integer n the Taylor series of this expression in a reproduces the polynomial (4.29). Furthermore, this function satisfies Carlson's theorem, since it grows exponentially in n on the real axis, but only oscillates along the imaginary axis. Then differentiating and taking the $n \to 1$ limit of the normalized Renyi entropy we get

$$\boxed{S_m = -\,\partial_n Z_n|_{n=1} = \frac{\log(a)}{2} + \sqrt{4a+1}\coth^{-1}\left(\sqrt{4a+1}\right)} \tag{4.31}$$

At late times, $a \gg 1$ so[3]

$$S_m \approx \frac{1}{2}\log a + 2 \sim \frac{1}{2}\log\left(\frac{T}{\beta\alpha' m^2}\right). \tag{4.32}$$

In general, the entropy monotonically increases. This growth results from states with different

[3]Note that we could derive the large time expansion by simply taking $\det M \approx n^2 a^{n-1}$.

boundary conditions becoming more orthogonal under time evolution. If we had c independent free bosons, each with uncorrelated boundary conditions, the entropy would simply be c times the above answer.

This entropy eventually competes with the thermal entropy of the wormhole, which at low temperatures is $S = 2S_0 + O(1/\beta)$, and produces a unitarity problem when

$$u/\beta \sim 4\pi\alpha' m^2 e^{4S_0/c-4}. \tag{4.33}$$

This signals that the trivial surface eventually becomes subdominant in the full non-perturbative calculation of the entropy of the unknown couplings entropy, or equivalently of the system.[4] It suggests that some part of the bulk might become inaccessible to the boundary system, falling out of it's entanglement wedge as a result of the uncertainty in the couplings.

It is also interesting to note the dependence of this time scale on the uncertainty in the couplings. For that, we use the Shannon entropy of the dimensionless variable $(\alpha')^{1/2}X$, which for a Gaussian distribution is

$$S_{\rm Sh}(X) = \frac{c}{2}\left(1 + \log\frac{2\pi}{\alpha' m^2}\right) \tag{4.34}$$

Recall that m is the inverse width of the distribution, and the increase of $S_{\rm Sh}(X)$ as m decreases is consistent with there being more uncertainty. The time scale is

$$u/\beta \sim e^{\frac{2}{c}[2S_0 - S_{\rm Sh}(X)]}. \tag{4.35}$$

Thus, the transition time is sooner for higher uncertainty. This can be used to make the transition time less than exponential in the entropy, and seemingly as small as desired by considering a wide enough distribution for the couplings.

4.1.2 Matter entropy for semi-infinite interval

Before addressing the matter entropy of an island, let us consider the matter entropy of a semi-infinite interval (a "peninsula"). This may be viewed as a warmup to the island calculation, which we will see reduces to the peninsula calculation at late times. It is also the relevant computation for the entanglement wedge of the left side of the boundary system, tracing out the right side and the

[4]Note that whether we choose to regulate the problem by compactifying the boson $X \sim X + 2\pi R$ or by changing the measure so that $X^2 \sim 1/m$, the net result is quite similar. See Appendix 4.5.

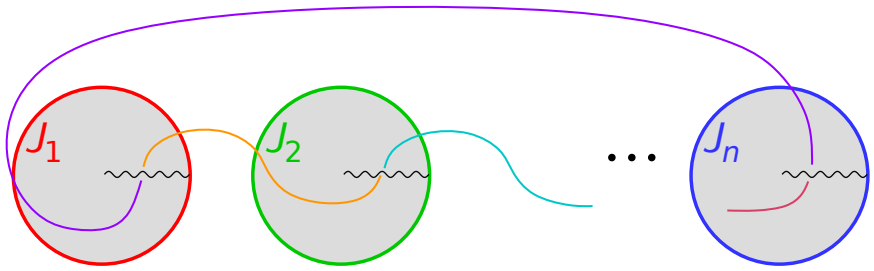
journal. If the journal is empty, the QES lives at the bifurcate horizon; with a journal, we expect the QES to shift to the left.

To compute the matter entropy, we will follow the same strategy as above. First note that the left density matrix

$$\rho_L = \sum_J P(J) \frac{e^{-\beta H_J}}{Z_J(\beta)} \tag{4.36}$$

For a free boson, the disk partition function $Z_J(\beta)$ is independent of J, so we can ignore the denominator (it will just contribute to an overall normalization of the density matrix). More generally, for a BCFT on a disk one can we write $Z_J(\beta) = g(J)Z(\beta)$, where $\log g(J)$ is the "boundary entropy" which is the boundary analog of the central charge [130].[5] Therefore we can absorb $g(J)$ into the definition of $P(J)$. Then the computation of the unnormalized density matrix $t_r \rho_L^n$ is given by

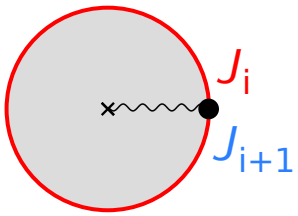
Z_n =



(4.37)

This is a path integral on a cone with boundary total length βn. Each segment of length β has some (generically different) boundary conditions J_i. So there are n boundary condition changing operators on the disk. We can quotient the picture to obtain:

Z_n =



(4.38)

In the quotient picture, the CFT is again the n-fold tensor product of the seed theory, now with only a single boundary condition changing operator $O_{\{J\}}$ and a twist operator σ. We will compute the matter entropy as a function of the position of the twist operator. The virtue of the quotient picture is that whereas before we had an n-pt function on the disk, now we only have a bulk-to-boundary

[5] $Z(\beta)$ is generated from the conformal anomaly, which only depends on the central charge.

2-pt function, which is fixed by conformal symmetry:

$$\langle \sigma(z,\bar{z}) O_b(y) \rangle_{\mathsf{disk}} \propto \frac{1}{(|1-|z|^2)^{2h-h_b}|1-z\bar{y}|^{2h_b}} \tag{4.39}$$

We can obtain this by mapping the disk to the upper half plane, and then using the doubling trick to relate the correlator to a chiral three-pt function on the plane. This is the 2-pt function on a flat disk. We are interested in the Poincaré disk $ds^2 = 4dzd\bar{z}/(1-|z|^2)^2$ with a circular boundary of circumference β/ϵ at $(1-|z|^2)/2 = 2\pi\epsilon/\beta$, so a Weyl transformation leaves us with

$$\left\langle \sigma(z,\bar{z}) O_{\{J\}}(\theta') \right\rangle_{\mathsf{AdS}} = c_n(\{J\}) \left(\frac{2\pi\epsilon}{\beta} \frac{(1-|z|^2)}{|1-z\bar{y}|^2} \right)^{h_b}. \tag{4.40}$$

Note that this expression is invariant under rotational (boost) symmetry $z \to ze^{i\theta}, y \to ye^{i\theta}$. In our problem, the boundary condition changing operator is on the left side at $\bar{y} = -1$ and z is real. The dependence on X and comes in via both h_B and the BOE coefficient $c_n(\{J\})$. In principle one can evaluate this coefficient, which is related to an n-pt function of boundary vertex operators on a disk:

$$c_n(\{J\}) = \prod_{i<j}^{n} |w_i - w_j|^{-2\left(\frac{X_i - X_j}{2\pi}\right)^2/\alpha'} \tag{4.41}$$

where $w_k = e^{2\pi i(k-1)/n}$. Together $h_b \propto -2\left(\frac{X_i - X_j}{2\pi}\right)^2/\alpha'$, (4.40) reduces to "just" a Gaussian integral. Nevertheless, the $c_n(\{J\})$ factors lead to a sufficiently complicated determinant that analytically continuing the answer in n is not easy. However, defining

$$\tilde{a} = \frac{1}{4\pi^2\alpha' m^2} \log\left(\frac{\beta}{2\pi\epsilon} \frac{1+z}{1-z} \right) = \frac{d_{\mathsf{twist}}}{2\pi^2\alpha' m^2}, \tag{4.42}$$

where d_{twist} is the distance from the boundary to the twist operator.2 we expect that when $\tilde{a} \gg 1$, the integral will be dominated by the second factor in (4.40), which is much more sharply peaked in X^2. The entropy with then be of the form (4.27) where $a \to \tilde{a}$.

The entropy in the limit of large $\tilde{a}$ is

$$S_m \approx \frac{1}{2} \ln \tilde{a} + ... \tag{4.43}$$

At fixed $\alpha' m^2$, this answer is expected[6] to be valid when $z + 1 \gg \epsilon^{1/2}$, e.g., as long as the twist

[6]The only possible complication is that the $n \to 1$ limit does not commute with $\tilde{a} \to \infty$

operator is far from the boundary cutoff.

To find the QES, we need the profile of the dilaton, which in these coordinates is

$$\phi = \frac{2\pi\phi_r}{\beta}\frac{1+|z|^2}{1-|z|^2} \tag{4.44}$$

In general, there will be a QES to the left of the horizon as long as $\phi_r/\epsilon \gg 1$, e.g., as long as ϕ_r is fixed in the $\epsilon \to 0$ limit. This is shown in figure 4.1. Balancing out the derivatives of the ϕ and S_m places the QES at

$$z_{\mathsf{QES}} = -\frac{\beta}{2\pi\phi_r \ln \beta/(2\pi\epsilon)} \tag{4.45}$$

This is just outside the left horizon. To leading order in $1/\tilde{a}$ we find no dependence of the location of the QES on the degree of uncertainty of the couplings m^2.

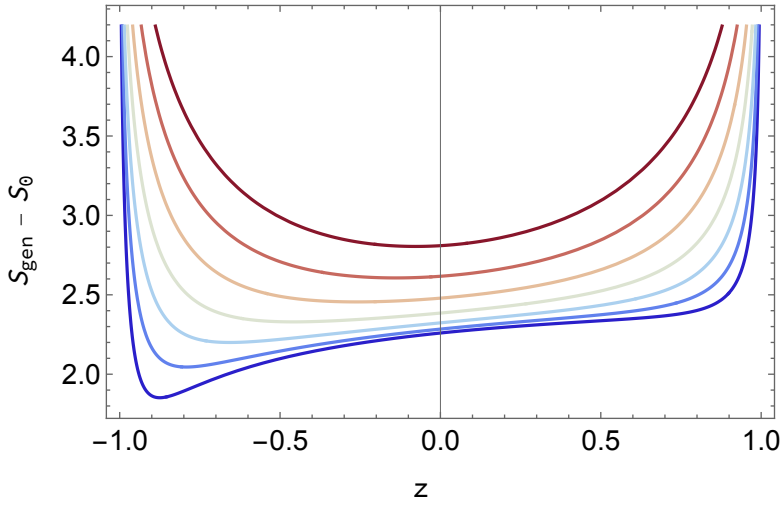
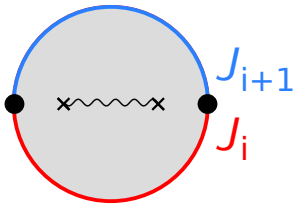


Figure 4.1: Generalized entropy $S_m + \phi - \phi_0$ as a function of the real coordinate z. Here S_m is the matter entropy of an interval $[-1+\epsilon^{1/2}, z]$. For large values of ϕ_h, there is a QES near the horizon. As ϕ_r gets smaller, the QES shifts closer to the left side.

4.1.3 Entropy of the island

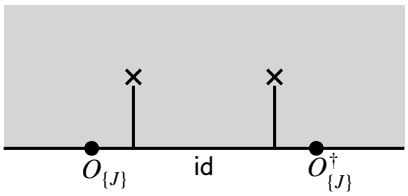
Let us finally consider the matter entropy of an island $S_m(I \cup \mathsf{journal})$. Again consider the n-th Renyi entropy of the semiclassical theory. Like in the case of the trivial surface, there are 2 boundary

condition changing operators; now, there are also 2 twist operators in the bulk:

$$Z_n \quad = \quad$$ 

(4.46)

One can ask what happens when we bring one of the twist operators very close to a boundary operator. This is a bulk-boundary OPE (sometimes referred to as a BOE) limit. More explicitly, we can expand both twist operators in terms of boundary primaries and their descendants. Then we are left with a boundary 4-pt function. In the limit that two of the points are close, only the boundary identity will contribute:

$$\langle O_{\{J\}}(y_1)\sigma(z_1,\bar{z}_1)\sigma(z_2,\bar{z}_2)O^\dagger_{\{J\}}(y_2)\rangle \quad = \quad$$ 

(4.47)

Let us consider the entropy of a finite interval with both endpoints near the horizon at very late times. This is a bulk-boundary OPE limit: the distance to the horizon is fixed whereas the length of the wormhole is growing linearly with time. So we expect the above four point function to factorize:

$$\langle O_{\{J\}}(y_1)\sigma(z_1,\bar{z}_1)\rangle\langle\sigma(z_2,\bar{z}_2)O^\dagger_{\{J\}}(y_2)\rangle \tag{4.48}$$

The Renyi entropy, however, doesn't factorize, but is instead given given by correlated sum of the above product:

$$\int\prod_{i=1}^{n} dJ_i\, p(J_i)\ \langle O_{\{J\}}(y_1)\sigma(z_1,\bar{z}_1)\rangle\langle\sigma(z_2,\bar{z}_2)O^\dagger_{\{J\}}(y_2)\rangle \tag{4.49}$$

In the frame where we evolve in time symmetrically on the left and right, the problem has a symmetry that interchanges the left and right sides. This allows us to consider instead the quantity

$$\int\prod_{i=1}^{n} dJ_i\, p(J_i)\ \langle\sigma(z_2,\bar{z}_2)O^\dagger_{\{J\}}(y_2)\rangle^2 \tag{4.50}$$

This is identical to the single interval case, but with $\tilde{a} \to 2\tilde{a}$. This minor modification does not change the location of the QES. See Figure 4.14.

This case demonstrates the strong sensitivity of the black hole interior to the values of the couplings. The Page-like transition gives a limit to the allowed uncertainty in the couplings, as measured by the entropy between the journal and the boundary, after which the entanglement wedge snaps and the interior falls outside the entanglement wedge of the boundary.

4.2 Renyi Entropies in gravity and SYK

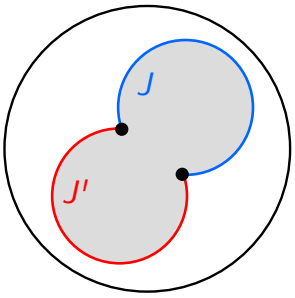
In the previous section, we discussed a model where the reconstructability absent precise knowledge of the couplings could be directly probed by finding the QES of the unknown couplings. Not all models enjoy this level solvability. However, it is sometimes easier to compute the Renyi entropy in these more general models. Thankfully, there are two signatures of the presence of a QES from the Renyi entropy. The first is a unitarity paradox where the semi-classical saddle produces a Renyi entropy too small to be consistent with the dimensionality of the Hilbert space of the system. The second is signature, and what fixes the first, is the presence of Replica wormholes. In solvable models, the QES prescription falls out of the $n \to 1$ limit of the n-th Renyi entropy computation.

We dedicate this section to computing the $n = 2$ Renyi entropy of SYK. This will be done in various regimes, including the low temperature Schwarzian regime, at large q, and also numerically via exact diagonalization of the SYK Hamiltonian. We will again find signatures of the failure of reconstruction due to insufficient knowledge of the couplings. We will also consider the case where the unknown couplings correspond to an irrelevant deformation of the system.

4.2.1 Disk contribution

As an intermediate step for the Renyi computation, we will compute elements of the journal density matrix. In a holographic theory, this is determined by a compution like

$$\langle J|\,\rho\,|J'\rangle \quad \propto \quad$$



(4.51)

where we fill in some gravity solution with the appropriate boundary conditions that correspond to evolution by $H + \chi(u)O_\Delta$. Here $\chi(u)$ takes the value J for $0 < u < \tau$, and the value J' for the other part of the circle $\tau < u < \beta$. By taking $\tau = \beta/2 + iT$ one can study the matrix elements as a function of real time T. If we assume that the field dual to O_Δ is a free field in AdS_2, we can

integrate out the field to get an effective matter action

$$-I_m = D\int du_1\, du_2 \left[\frac{t'(u_1)t'(u_2)}{2\sin^2\left(\frac{t(u_1)-t(u_2)}{2}\right)}\right]^{\Delta} \chi_r(u)\chi_r(u') \tag{4.52}$$

where $D = \frac{\left(\Delta-\frac{1}{2}\right)\Gamma(\Delta)}{\sqrt{\pi}\Gamma\left(\Delta-\frac{1}{2}\right)}$. In computing the overlap, we should divide by the norms of the states $\sqrt{\langle J|J\rangle\,\langle J'|J'\rangle}$; but these are not time dependent and do not play an important conceptual role. We will now specialize to the case of a marginal deformation $\Delta = 1$, where the computation simplifies significantly. For one thing, we may shift the $J \to J + a$, $J' \to J' + a$ without changing the value of the overlap, so without loss of generality we may set $J' = 0$. In Appendix 4.8 we evaluate this integral for the special case of a marginal deformation $\Delta = 1$, giving an answer that is remarkably bi-local in u_1 and u_2:

$$e^{-I_m} = \left[\frac{\epsilon^2 t'(u_1)t'(u_2)}{\sin^2\left(\frac{t(u)-t(0)}{2}\right)}\right]^{\delta}, \quad \delta = \frac{(J-J')^2}{2\pi} \tag{4.53}$$

where ϵ is a UV regulator. This precisely agrees with the dimension of the boundary condition changing operator we found in (4.26) if we set $\alpha' = 1/(2\pi)$ so that the scalar is canonically normalized. Note that the final form is consistent with a picture where we have inserted a bulk "domain wall" of mass $\sim \delta$ that separates the J and J' vacua. The domain wall is where the gradient of the bulk field is appreciable; it is not a thin wall.

This matter action is valid for an off-shell $t(u)$. In other words, we can include the gravitational backreaction by integrating over $t(u)$ with the Schwarzian action appropriate for the disk. This is the same action we would have gotten from inserting two local operators of NCFT$_1$ dimension δ at $u = 0$ and u. For a free scalar field in the bulk, these expressions combined with the results for the Schwarzian n-pt functions [5,6,9,131] give us the exact disk contribution, summing over all quantum fluctuations of the boundary mode.

The fact that we have an exact quantum expression for the Renyi entropy gives us confidence that there is really an information paradox if we just focus on the disk. Although in the classical approximation, the disk contribution decays exponentially, without the quantum expressions, we would not be confident that the disk decays to a value smaller than what is required by unitarity $\sim e^{-S_0}$. Indeed, the quantum modifications show that the exponential decay is replaced by a power law decay $\sim T^{-3}$, with the exponent independent of δ.

Here we would also like to comment that the above result also applies to SYK in the Schwarzian

limit. We will use the following conventions for SYK:

$$H_{\mathsf{SYK}} = i^{q/2} \sum_{1\leq i_1 \ldots \leq i_a \leq N} J_{i_1 \ldots i_q} \psi_{i_1} \ldots \psi_{i_q}, \quad \left\langle J^2_{i_1 \ldots i_q} \right\rangle = \frac{NJ^2}{q\binom{N}{q}} = \frac{N\mathcal{J}^2}{2q^2\binom{N}{q}}, \quad \{\psi_i, \psi_j\} = 2\delta_{ij} \tag{4.54}$$

To mimic the above expressions, we can imagine turning on a deformation by another SYK Hamiltonian,

$$H = H_{\mathsf{SYK}}(J_1, q_1) + \chi(u) H_{\mathsf{SYK}}(J_2, q_2) \tag{4.55}$$

Here J_1, J_2 are distributed like in (4.54). We can introduce a G, Σ action by following the usual steps, only integrating out both J_1 and J_2 to obtain an action $I_0 + I_m$, where

$$\begin{aligned} -I_0/N &= \log \mathrm{Pf}\left(\partial_t - \Sigma\right) - \frac{1}{2}\int d\tau_1 d\tau_2 \left[\Sigma\left(\tau_1, \tau_2\right) G\left(\tau_1, \tau_2\right) - \frac{J^2}{q} G\left(\tau_1, \tau_2\right)^q\right] \\ &\approx \frac{\alpha_S}{\mathcal{J}} \int du \left\{\tan\frac{\pi t(u)}{\beta}, u\right\} \\ -I_m/N &= \frac{J^2}{2q'} \int du_1\, du_2\, \left[G(u_1, u_2)\right]^{q'} \chi(u_1)\chi(u_2) \\ &\approx \int du_1\, du_2 \left[\frac{t'(u_1)t'(u_2)}{2\sin^2\left(\frac{t(u_1)-t(u_2)}{2}\right)}\right]^{q'/q} \chi_r(u_1)\chi_r(u_2), \quad \chi_r = \chi(u) b^{q'/2} \end{aligned} \tag{4.56}$$

where $J^2 b^q \pi = \left(\frac{1}{2} - \frac{1}{q}\right)\tan(\pi/q)$. Here I_0 comes from J_1 and I_m comes from the J_2 couplings. The relative normalization of the two terms is controlled by the magnitude of χ. Now in the low temperature, large N limit $I_0[G, \Sigma] \to \mathsf{Sch}[t(u)]$ is replaced by the Schwarzian action, where we integrate over $t(u)$ instead of G, Σ. Similarly, the action I_m can be re-written in terms of the Schwarzian mode. Note that in writing these expressions, we also assume that χ is small so that we can simply integrate over the near-zero mode. This is a similar approximation to what is discussed in [27].

We expect that the disk partition function at early times to be self-averaging both in J_1 and J_2. Therefore, this computation has multiple interpretations. The first interpretation is that we draw some particular choice of J_2 and view the term $H_{\mathsf{SYK}}(J_2)$ as a single operator that is deforming the original theory. The journal records not the values of J_2 but merely a single coupling χ_0 which is the overall normalization of the deformation. Then the above action would govern the matrix element of the journal density matrix, e.g., an overlap between TFD's with different values of χ_0. Since the disk answer (for early times) is self-averaging, we can average over J_1, J_2 in this computation and

derive the above effective action in terms of G, Σ.

A second physically different setup is when χ_0 is fixed, and the journal instead records the values of all $\binom{N}{q'}$ couplings J_2. Then we think of J_1 as "known" or fixed. We will see in the next section that (4.56) will still be relevant for computing Renyi entropies, or the overlaps averaged over J_2.

4.2.2 Renyi-2 wormhole

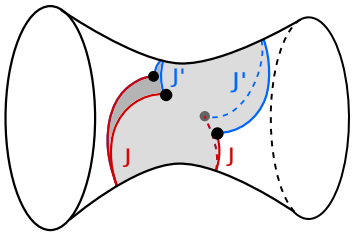
The Renyi-2 entropy of the journal is given by

$$t_r \rho^2_{\text{journal}} = \int p(J)p(J') |\langle J| \rho |J'\rangle|^2. \tag{4.57}$$

As an intermediate step, we should compute the squares of density matrix elements. If we plug in our answer for the density matrix elements (4.52), there will be a unitarity paradox at late times. In particular, $t_r \rho^2_{\text{journal}}$ will decay to zero, which is impossible since it is bounded by 2^{-N}.

The resolution to this paradox is that the squares of the density matrix elements have an additional contribution given by wormholes. In a theory with other "known" couplings, this is an acceptable resolution, but such wormholes would violate factorization in a theory with fixed couplings. To obtain the norm of the overlaps, we start with a Euclidean computation real $\tau, \bar{\tau}$ and then analytically continuing $\tau = \beta/2 + iT$, $\bar{\tau} = \beta/2 - iT$:

$$|\langle \beta + 2iT, J|\beta + 2iT, J'\rangle|^2 = t_r\left(e^{-\bar{\tau}H_J} e^{-\tau H_{J'}}\right) t_r\left(e^{-\tau H_J} e^{-\bar{\tau} H_{J'}}\right)$$
$$\supset$$



(4.58)

In addition to the disk topology which we have already discussed, we have drawn a wormhole contribution. We can think of this wormhole as being supported by the bulk "domain walls" that separate the J and J' region. Such a wormhole is closely related to the one described by Douglas Stanford in Appendix B of [132].

Let us outline the steps to obtaining the wormhole, leaving a more thorough discussion to Appendix 4.8. We will start with the double trumpet geometry, with a cutout parameterized by two boundary times $T_L(u_L)$ and $T_R(u_R)$. First, we turn on a source χ_L on the left side from times in $(0, \tau)$ and a source χ_R which is on from times in (τ, β). This leads to a matter action that has LL

correlators, RR correlators, as well as LR cross terms.

$$-I_m^{\text{wormhole}} = D\int du_1\, du_2 \Bigg\{ \left[\frac{T_L'(u_1)T_R'(u_2)}{\cosh^2\left(\frac{T_L-T_R}{2}\right)}\right]^{\Delta} \chi_L(u_1)\chi_R(u_2) + \left[\frac{T_L'(u_1)T_L'(u_2)}{\sinh^2\left(\frac{T_L(u_1)-T_L(u_2)}{2}\right)}\right]^{\Delta} \chi_L(u_1)\chi_L(u_1) + \left[\frac{T_R'(u_1)T_R'(u_2)}{\sinh^2\left(\frac{T_R(u_1)-T_R(u_2)}{2}\right)}\right]^{\Delta} \chi_R(u_1)\chi_R(u_2)\Bigg\} \tag{4.59}$$

In principle, one should evaluate this integral for off-shell T_L, T_R, and then find the classical solution. For $\Delta = 1$ the integral can be performed but the answer is a bit complicated. However, large Lorentzian times T, a sensible ansatz is that the distance between the quench sites on opposite sides of the wormhole and u_2 and u_3. In the above picture (4.58), we are saying that the domain walls that cross the wormhole on the "front" and "back" sides have fixed length at large T. With this ansatz, one can show that the above matter integral becomes the insertion of two domain wall operators:

$$e^{-I_m} = \mathcal{C}^{LR}(u_1,u_4)\mathcal{C}^{LR}(u_2,u_3), \quad \mathcal{C} = \left[\frac{T_L'(u_1)T_R'(u_2)}{\cosh^2\left(\frac{T_L-T_R}{2}\right)}\right]^{\delta}. \tag{4.60}$$

Since the overall effect of the matter is local, the on-shell solution must again be semi-circular arcs, with junctions at the quench sites. Therefore the wormhole can be obtained by starting with two copies of the disk geometry in (4.51), one for the "front side" of the wormhole and the other for the "back side." Then one cuts both of these disks and pastes them together along two geodesics so that there is no bulk discontinuity. The key equation we will need is the effective energy of the solution, or equivalently the circumference β_E of the pieces of the disk that we use to make the solution. This is the same energy on all pieces of the solution away from the quench sites. It is determined by imposing $\mathrm{SL}(2,\mathbb{R})$ charge conservation at the quench sites, see 4.8. The result is that for any T,

$$\tan\left(\frac{\pi\beta}{2\beta_E}\right) = \frac{\delta\beta_E}{2\pi}. \tag{4.61}$$

For small values of $\delta\beta$, we get $\beta_E^2 = \pi^2\beta/\delta$. Several aspects of the wormhole geometry are discussed in 4.8; here we will just check that the wormhole resolves the unitarity problem. To do so in a semi-classical approximation, we need to evaluate the on-shell value of

$$\exp\left(-I^{\text{wormhole}}\right)\mathcal{C}^{LR}(u_1,u_4)\mathcal{C}^{LR}(u_2,u_3). \tag{4.62}$$

Let us now evaluate I^{wormhole} in the limit of small $\delta\beta$ and show that this resolves the potential unitarity paradox at large T. There are two Schwarzian boundaries in the computation, both of which have the same action. The left boundary consists of two arcs:

$$\begin{aligned} -I_{\mathsf{Sch}} &= 2\int du_L \{\tan\frac{\pi t(u)}{\beta}, t\} = 2\left[\tau+\bar{\tau}\right]\left(\frac{\pi}{\beta_E}\right)^2 \\ &= 2\pi^2\beta/\beta_E^2. \end{aligned} \tag{4.63}$$

The correlators are essentially on opposite sides of an effective thermofield double β_E, so they take thermal values:

$$Z_{\text{wormhole}} \sim \exp\left(-4\pi^2\beta/\beta_E^2\right)\beta_E^{-4\delta}. \tag{4.64}$$

We see that the classical action approaches a value independent of T at late times, which resolves the potential unitarity paradox for the Renyi-2 entropy. More generally, from the results in Appendix 4.8, we can see that the solution at general iT will give an action independent of T as both the distance across the wormhole and the action, including corner contributions, will be time-independent.

4.2.3 Numerics for finite N SYK

In this section, we resort to the numerical analysis of the $n = 2$ Renyi entropy of the journal by exact diagonalization of the SYK Hamiltonian at finite N. We will present evidence of the wormholes. The "known" couplings in all of our setups will be in a standard SYK Hamiltonian with $\binom{N}{q}$ couplings. The "unknown" couplings will be additional unknown couplings.

We argued around (4.12) that the relevant entropy that diagnoses the entanglement wedge given the "known" couplings is the relative entropy $S(\mathsf{unknown}|\mathsf{known})$. The "conditional" Renyi entropy whose $n \to 1$ limit is equal to this conditional entropy is given by

$$\tilde{R}_n = \sum_{\kappa} P(\kappa) \sum_{\mu_1,\cdots,\mu_n} \prod_{i=1}^{n} P(\mu_i) \left\langle \psi;\mu_i,\kappa|\psi;\mu_{i+1},\kappa\right\rangle_{\mathsf{sys}}. \tag{4.65}$$

$$S(\mathsf{unknown}|\mathsf{known}) = -\partial_n \tilde{R}_n\big|_{n=1}. \tag{4.66}$$

An efficient numerical way to estimate the above quantity is to Monte Carlo sample the quantity $\prod_{i=1}^{n}\left\langle \psi;\mu_i,\kappa|\psi;\mu_{i+1},\kappa\right\rangle_{\mathsf{sys}}$ from the distribution $P(\mu_1)\cdots P(\mu_n)P(\kappa)$. The conditional Renyi $\tilde{R}_n$ has the appealing feature that it is linear in $P(\kappa)$, as opposed to the standard Renyi entropy of the

full journal, which comes with a $P^n(\kappa)$ factor.

We will now employ this method in studying $\tilde{R}_2$ for various "unknown" deformations of the SYK model. The first case we will consider is the $q = 4$ SYK, where the couplings are known with only limited precision. To model this, we decompose the usual SYK Hamiltonian into two terms:

$$H = (K_{ijkl} + J_{ijkl})\,\psi_i\psi_j\psi_k\psi_l, \tag{4.67}$$

Here K are the known couplings, and J are the unknown couplings (the ones stored in the journal). We can think of this setup as modeling a situation where the usual SYK couplings are known but with some Gaussian errors. This is similar to (4.55), except that there we were viewing the single coupling as an overall normalization of the second term; here, there are $\binom{N}{4}$ couplings in the journal. The conditional $n = 2$ Renyi entropy is given by

$$\tilde{R}_2 = \int dJ_0\, p(J_0) \int dJ dJ'\, p(J)p(J') \,|\, \langle \beta/2 + iT; J_0, J|\beta/2 + iT; J_0, J'\rangle \,|^2. \tag{4.68}$$

We would like to test whether wormholes contribute to the above calculation. To do so, we consider the following "disk" approximation to $\tilde{R}_2$:

$$\tilde{R}_2^{\mathsf{disk}} = \left| \int dJ_0 dJ dJ'\, p(J_0)p(J)p(J')\ \langle \beta/2 + iT; J_0, J|\beta/2 + iT; J_0, J'\rangle \right|^2. \tag{4.69}$$

In this approximation, we have averaged first, before squaring. In the large N analysis, only disconnected solutions can contribute to the above quantity. If $\tilde{R}_2$ is significantly larger than $\tilde{R}_2^{\mathsf{disk}}$, we interpret this as evidence that a wormhole is dominating over the disconnected solution.

To Monte Carlo sample these integrals, we draw two independent SYK Hamiltonians H_1 and H_2 from Gaussian distributions. Then $H_\pm = (\cos\theta)H_1 \pm (\sin\theta)H_2$ are also normalized Hamiltonians such that $\langle t_r H_\pm^2 \rangle = J^2$, $\langle t_r H_+ H_- \rangle = \lambda^2 J^2$, where $\lambda^2 = \cos(2\theta)$. The results are displayed in Figure 4.2. The disk approximation is a good one at small JT but at large values, we see that it significantly underestimates $\tilde{R}_2$. In fact, at very late times, the underestimate is so bad that even without computing $\tilde{R}_2$ exactly, we can rule it out on the grounds that it would violate unitarity.

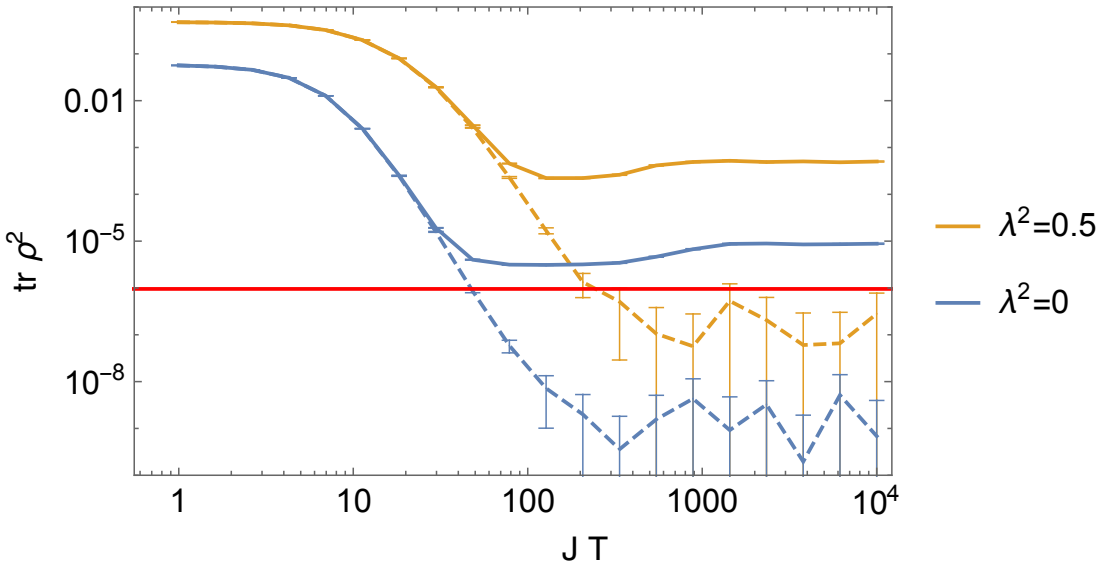


Figure 4.2: The conditional Renyi $\tilde{R}_2$ as a function of Lorentzian time evolution JT, obtained by numerically diagonalizing the SYK Hamiltonian for $N = 20$, $\beta J = 16$ and $n_{\text{trials}} = 3000$. 1σ error bars are displayed. The solid curve is the full answer while the dashed curve is the "disk" approximation $\tilde{R}_2^{\text{disk}}$. We see that the disk approximation is good for early times, but decays too rapidly at late times. Even without computing the full answer (in orange), the unitarity bound $t_r \rho^2 \geq 2^{-N}$ would rule out the disk answer at large times.

We also consider deforming a standard $q = 4$ SYK Hamiltonian with a *single* unknown coupling χ:

$$H = (K_{ijkl})\, \psi_i \psi_j \psi_k \psi_l + i^{q_2/2}\, (\psi_1 \cdots \psi_{q_2})\, \chi. \tag{4.70}$$

Here K_{ijkl} are drawn from a Gaussian as usual with a variance set by (4.54); we also choose χ to be normally distributed, with a variance $\langle \chi^2 \rangle = \epsilon^2 J^2$. This may be thought of as a variant of the problem in (4.55). The results are shown in figure 4.3. We see clear evidence of a wormhole, even when $q_2 > 4$. In fact, at the moderate values of N that we computed, there does not seem to be a large qualitative difference between $q_2 = 4$ and $q_2 > 4$. This is perhaps surprising since in the conformal regime, $q_2 > 4$ would correspond to an *irrelevant* coupling whereas $q_2 = 4$ would be marginal. To further test this interpretation, we compute the G_{LL} correlators numerically and see that they qualitatively agree with the predictions given in the subsequent section on the large q wormhole, see Figure 4.4 for the case of many marginal couplings and 4.5 for the case of a single

irrelevant coupling. We also computed G_{LR} numerically and saw that it is nearly zero at early times $Jt \sim 1$ but becomes appreciable at times that correspond to the wormhole/disk transition. More specifically, in Figures 4.4 and 4.5 we compare

$$\begin{aligned} |G_{\mathsf{LL}}^{\mathsf{disk}}| &= \frac{1}{N'} \sum_i \frac{\overline{|\langle \beta + 2iT, J| \, \psi_i(\beta/2)\psi_i(t) \, |\beta + 2iT, J'\rangle|}}{\overline{|\, \langle \beta + 2iT, J|\beta + 2iT, J'\rangle \,|}}, \\ |G_{\mathsf{LL}}^{\mathsf{WH}}| &= \frac{\overline{\langle \beta + 2iT, J| \, \psi_i(\beta/2)\psi_i(t) \, |\beta + 2iT, J'\rangle \, \langle \beta + 2iT, J'|\beta + 2iT, J\rangle}}{\overline{|\, \langle \beta + 2iT, J'|\beta + 2iT, J\rangle \,|^2}}, \end{aligned} \tag{4.71}$$

where in the above expressions the overbar indicates averaging with respect to J and J'. For the single coupling case, the fermion index i is summed over all $N' = N - q_2$ fermions except those appearing in the deformation, whereas in the case of many marginal couplings, the index i is summed over all $N' = N$ fermions. Of course our limited numerics cannot conclusively test the conformal regime ($N \to \infty$, $\beta\mathcal{J} \gg 1$) where the term "irrelevant" has a sharp meaning. Nevertheless, we conjecture that in the conformal/JT regime, there exist wormhole solutions for irrelevant couplings; we hope to report progress in this direction in a future publication.

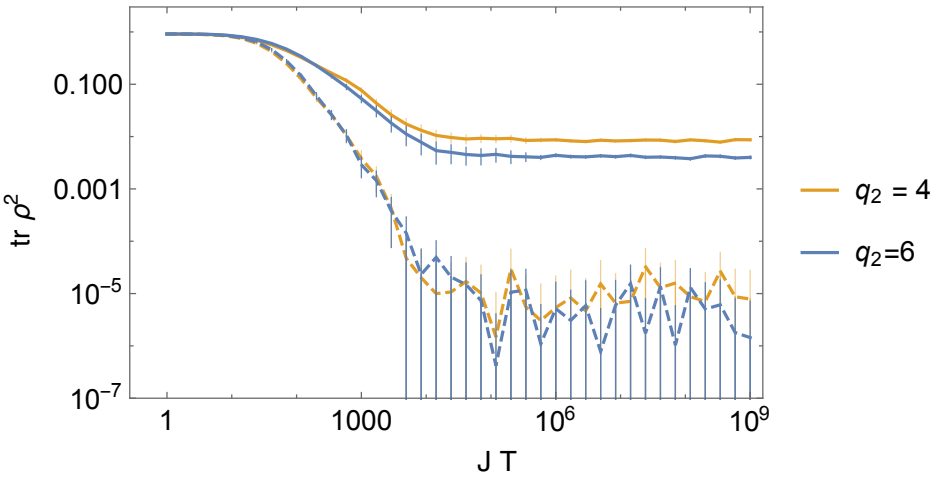


Figure 4.3: $N = 20$, $n_{\text{trials}} = 1000$, $\beta J = 16, \epsilon = 0.1$. We show $q_2 = 4, 6$ deformations and 1σ error bars. Note that $q_2 = 6$ is an irrelevant deformation, but the curves seem qualitatively similar to the marginal case.

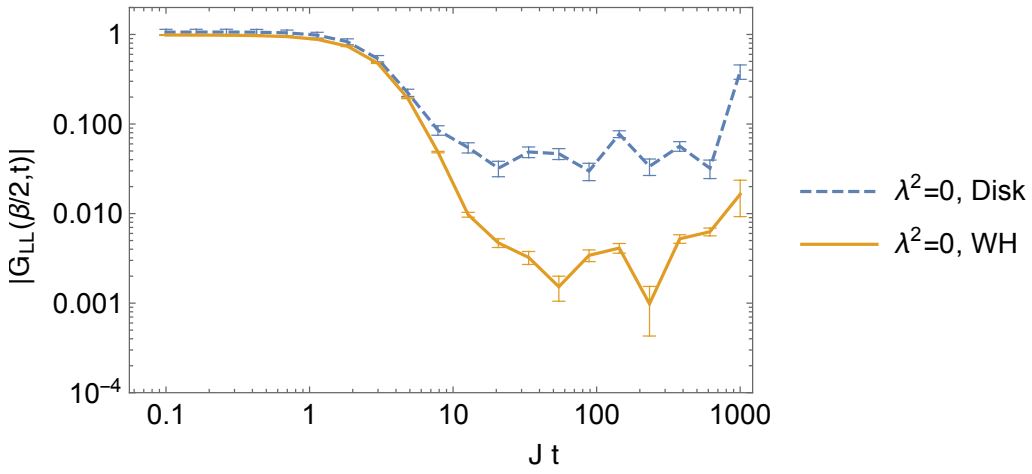


Figure 4.4: The 2-pt correlator $|G_{LL}|$ as a function of Lorentzian time for the $\lambda^2 = 0$ disk and wormhole for the case of many marginal couplings. We take $N = 20$, $n_{\text{trials}} = 600$, $\beta\mathcal{J} = 1$, $JT = 10^3$. Notice that as $t \to T$ we reach the second quench site. For the disk solution, a prediction from the large q analysis is that the correlator becomes large on the disk, whereas it remains small on the wormhole. This seems to be in rough agreement with the modest $N, q = 4$ numerics. In computing the correlator, we divide by the average norm $\langle \beta + 2iT, J | \beta + 2iT, J' \rangle$ for the disk and $| \langle \beta + 2iT, J | \beta + 2iT, J' \rangle |^2$ for the wormhole.

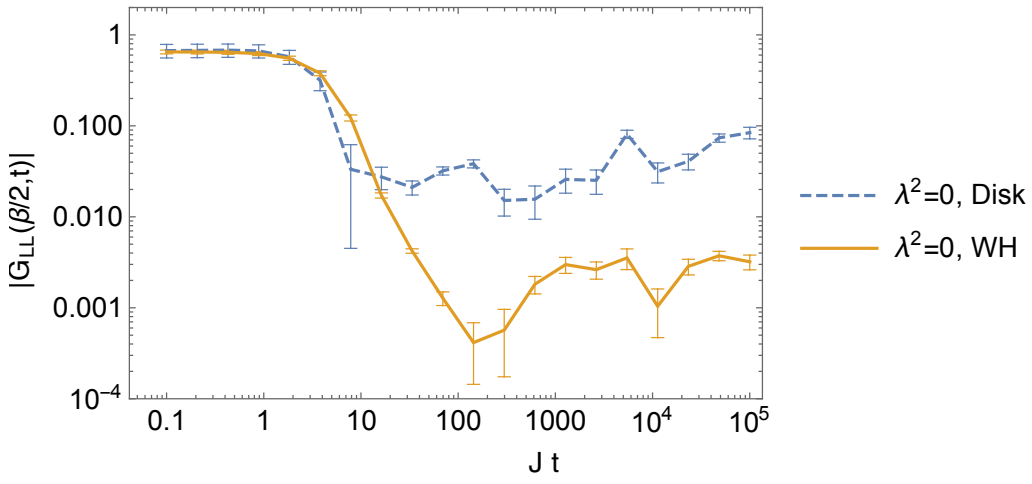


Figure 4.5: The 2-pt correlator $|G_{LL}|$ as a function of Lorentzian time, where a single irrelevant $q = 6$ parameter is varied, as in Figure 4.3. We take $N = 20$, $n_{\text{trials}} = 2000$, $\beta\mathcal{J} = 1$, $JT = 10^5$. The correlators are qualitatively similar to the case with many marginal parameters in Figure 4.4.

4.2.4 Large q SYK

In this section, we will consider the disk contribution to the Renyi-2 entropy in the $N \to \infty$ large q SYK model. We will show that the disk contribution decays to zero at large Lorentzian times, in conflict with unitarity. This suggests that there should be a wormhole that dominates at late times. We show that the wormhole in the low temperature limit is closely related to the Schwarzian wormhole described in the gravity section.

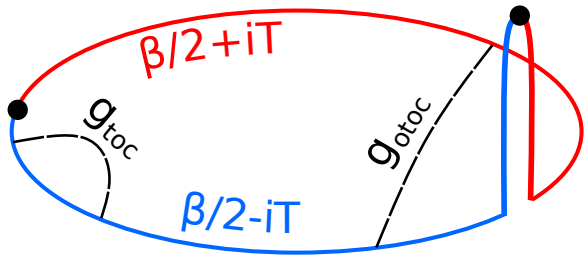


Figure 4.6: Contour relevant for matrix element of journal. This can be obtained by considering a red segment of length τ and a blue segment of length $\beta - \tau$, and then setting $\tau = \beta/2 + iT$.

We will consider the disk solution for the thermal circle of length β, with quench sites at times 0 and τ. The couplings are constant in the segments $(0, \tau)$ and (τ, β) and are partially correlated between these two segments. Let us assume the correlation between the couplings on the two sides is $\lambda^2 = e^{-\mu}$. When $\lambda = 1$ there is no change in the coupling at the quench sites; when $\lambda = 0$ the two sides are uncorrelated. The large q equations of motion are

$$\begin{aligned} \partial_1 \partial_2 g_{\mathsf{toc}} + 2\mathcal{J}^2 e^{g_{\mathsf{toc}}} = 0, \quad t_1, t_2 \text{ on same side} \\ \partial_1 \partial_2 g_{\mathsf{otoc}} + 2e^{-\hat{\mu}} \mathcal{J}^2 e^{g_{\mathsf{otoc}}} = 0, \quad t_1, t_2 \text{ on opposite sides} \end{aligned} \tag{4.72}$$

We will sometimes use the subscripts $(\mathsf{toc}, \mathsf{otoc})$ to emphasize when an equation applies when both times are on the same side or on opposite sides of the quench. It is convenient to define

$$\begin{aligned} \Omega_\tau &= \Theta(\tau - t_1)\Theta(t_2 - \tau) + \Theta(\tau - t_2)\Theta(t_1 - \tau) \\ \Omega_{\mathsf{toc}} &= 0, \quad \Omega_{\mathsf{otoc}} = 1. \end{aligned} \tag{4.73}$$

Then, we can summarize (4.72) as

$$\partial_1 \partial_2 g + 2\mathcal{J}^2 \exp\left(g - \mu \Omega_\tau\right) = 0. \tag{4.74}$$

In addition to the usual UV boundary condition $g(\tau,\tau)=0$, we impose continuity at the quench site:

$$\begin{aligned} g(t,-\epsilon) &= g(t,+\epsilon) \\ g(t,\tau-\epsilon) &= g(t,\tau+\epsilon). \end{aligned} \tag{4.75}$$

These boundary conditions follows from the fact that for very short times ϵ, we can neglect the interactions from $\mathcal{J}$. A useful trick is the following. Consider the field redefinition

$$g = \hat{g} + \mu\Omega. \tag{4.76}$$

Then the equations of motion and the boundary conditions are

$$\begin{aligned} \partial_1\partial_2\hat{g} + 2\mathcal{J}^2 e^{\hat{g}} &= 0, \\ \hat{g}_{\mathsf{toc}} &\to \hat{g}_{\mathsf{otoc}} - \mu. \end{aligned} \tag{4.77}$$

By using the variable $\hat{g}$ the equations of motion is uniform over the entire circle, at the price of a discontinuous boundary condition. This discontinuous boundary condition is considered in [30, 133, 134]. For general τ, see Streicher [133]. In our conventions,

$$e^{g_{\mathsf{toc}}} = \left(\frac{\alpha_1}{\mathcal{J}\sin\left[\alpha_1 t_{12}+\gamma_1\right]}\right)^2, \quad 0<t_1<t_2<\tau \tag{4.78}$$

$$e^{g_{\mathsf{toc}}} = \left(\frac{\alpha_2}{\mathcal{J}\sin\left[\alpha_2 t_{12}+\gamma_2\right]}\right)^2, \quad t_3<t_1<\bar{\tau} \tag{4.79}$$

$$e^{g_{\mathsf{otoc}}} = \left(\frac{\alpha_1\alpha_2\mathcal{J}^{-2}}{\lambda^2\sin\left[\alpha_1(t_1-\tau)\right]\sin\left[\alpha_2(t_2-\bar{\tau})\right]-\sin\left[\alpha_1(t_1-\tau)+\gamma_1\right]\sin\left[\alpha_2(t_2-\bar{\tau})-\gamma_2\right]}\right)^2. \tag{4.80}$$

We also have the UV boundary conditions:

$$\alpha_i = \mathcal{J}\sin\gamma_i \tag{4.81}$$

$$\sin\left(\frac{\alpha_1\tau\pm\alpha_2\bar{\tau}}{2}+\gamma_1\pm\gamma_2\right) = \lambda^2\sin\left(\frac{\alpha_1\tau\pm\alpha_2\bar{\tau}}{2}\right) \tag{4.82}$$

Here we are imagining solving the above equations for real $\tau,\bar{\tau}$ and then analytically continuing and setting $\tau=\beta/2+iT, \bar{\tau}=\beta/2-iT$. The physical branch of the above equations is the one that is continuously connected to the $T=0,\lambda=1$ solution. In figure 4.8 we solve for γ numerically at $T=0$ for varying $\beta\mathcal{J}$. We also solve for γ as a function of T for fixed $\beta\mathcal{J}$ in figure 4.9.

It is interesting to study these constraints for small $\gamma \approx \alpha/\mathcal{J}$:

$$-\frac{\gamma_1 \pm \gamma_2}{1-\lambda^2} \approx \tan\left(\frac{\alpha_1\tau \pm \alpha_2\bar{\tau}}{2}\right). \tag{4.83}$$

This can be compared with a classical Schwarzian analysis for the insertion of a heavy operator of dimension δ on the disk, which gives (see [135] equation 2.9):

$$-\frac{k_1 \pm k_2}{\Delta} = \tan\left(\frac{k_1\tau \pm k_2\bar{\tau}}{4C}\right). \tag{4.84}$$

This means that we may identify $k_i/2C = \alpha_i$ and $(1-\lambda^2)2C\mathcal{J} = \delta$. Here C is the coefficient of the Schwarzian action, which for large q SYK is $C \sim N/\mathcal{J}q^2$. In the Schwarzian theory, $k_i = 2\pi C/(\beta_E)_i$, where β_E is the length of the circle that is pieced together to form the disk solution.

The left hand side of the above equation is of order $(\beta\Delta)^{-1} \sim (\beta\mathcal{J}(1-\lambda^2))^{-1}$. So for large $\beta\Delta$, we have $\tan\left[\frac{1}{2}\left(\alpha_1\tau \pm \alpha_2\bar{\tau}\right)\right] = 0$. The correct branch is to take $\alpha_1\tau + \alpha_2\bar{\tau} = 2\pi$ and $\alpha_1\tau - \alpha_2\bar{\tau} = 0$, which yields

$$\alpha_1 = \frac{\pi}{\tau}, \quad \alpha_2 = \frac{\pi}{\bar{\tau}}. \tag{4.85}$$

This has a pleasing geometric interpretation for real $\tau, \bar{\tau}$. As we have noted, $\alpha_i = \pi/(\beta_E)_i$ where β_E is the effective inverse temperature (e.g. correlators on the same side of the quench agree with thermal correlators at a temperature β_E^{-1}.) The solution tells us to set $(\beta_E)_1 = \tau$ and $(\beta_E) = \bar{\tau}$. This means the geometry has the form of two circles which are completely "pinched." Indeed, if we plug in this solution, we can compute the correlator (4.78) at $t_1 = 0$ and $t_2 = \tau$. We see that $e^g = 1$, its maximal value. From the Schwarzian point of view, the domain wall is so massive that it pinches the disk together. This is depicted in Figure 4.7.

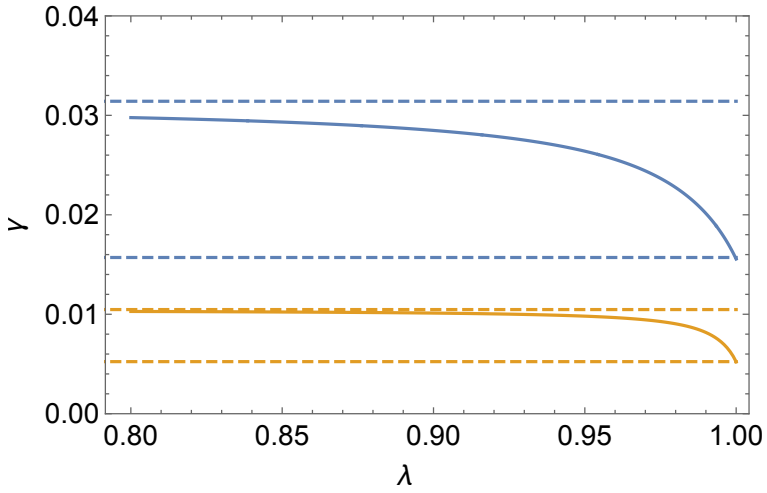


Figure 4.8: Here we show γ as a function of λ for $\beta\mathcal{J} = 200$ and $\beta\mathcal{J} = 800$ in blue and orange respectively. We also show in dashed lines the upper and lower bounds on γ obtained analytically (in a $1/\beta\mathcal{J}$ expansion.) Note that for larger $\beta\mathcal{J}$, γ quickly obtains its maximal value $2\pi/\beta$ away from $\lambda = 1$.

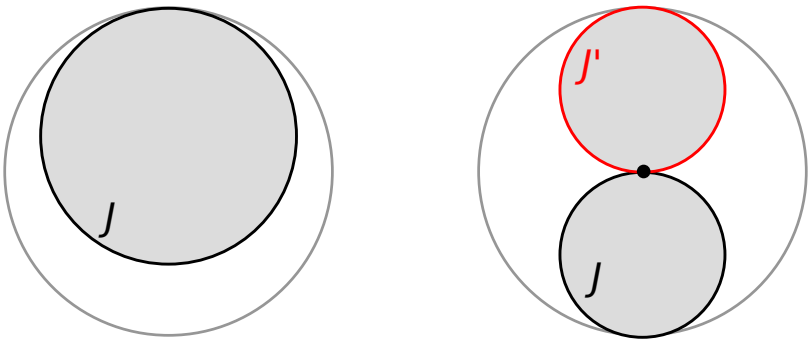


Figure 4.7: Left: the ordinary extremal or zero temperature solution. Since the boundary particle intercepts the asymptotic boundary, the total proper time is infinite. Right: if we take $\beta = \infty$ while holding the mass of the brane constant, the solution pinches. This means that the correlator between the quench sites is getting large. In the large q theory, we showed that the correlator in fact becomes maximal $e^g = 1$.

This should be contrasted with the case when $\Delta\beta \ll 1$ in which case $\alpha_i = 2\pi/\beta$. So for $\tau = \bar{\tau}$ large but general λ, we expect that $\pi/(\beta\mathcal{J}) \leq \gamma \leq 2\pi/(\beta\mathcal{J})$. These bounds are shown in figure 4.8.

We can use our results for $\gamma_1 = \gamma_2^*$ to compute the 2-pt correlator $e^{g(\beta,\tau)}$ as a function of τ, see figure 4.10. This will be needed for the evaluation of the on-shell action for the Renyi-2 entropy in the next subsection.

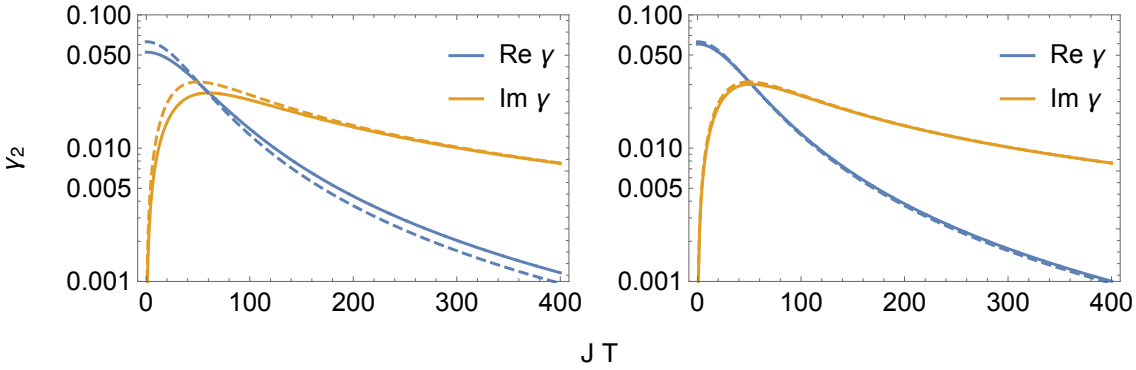


Figure 4.9: $\gamma_2 = \gamma_1^*$ as a function of $\mathcal{J}T$ for $\beta\mathcal{J} = 100$. On the left, $\lambda = 0.9$, whereas on the right $\lambda = 0$. Solid lines are obtained by numerically solving equation (4.82). The dashed lines are the "pinching" approximation derived in (4.85). At large $\beta\mathcal{J}$, pinching is a reasonable approximation even when $1 - \lambda^2$ is not small (left).

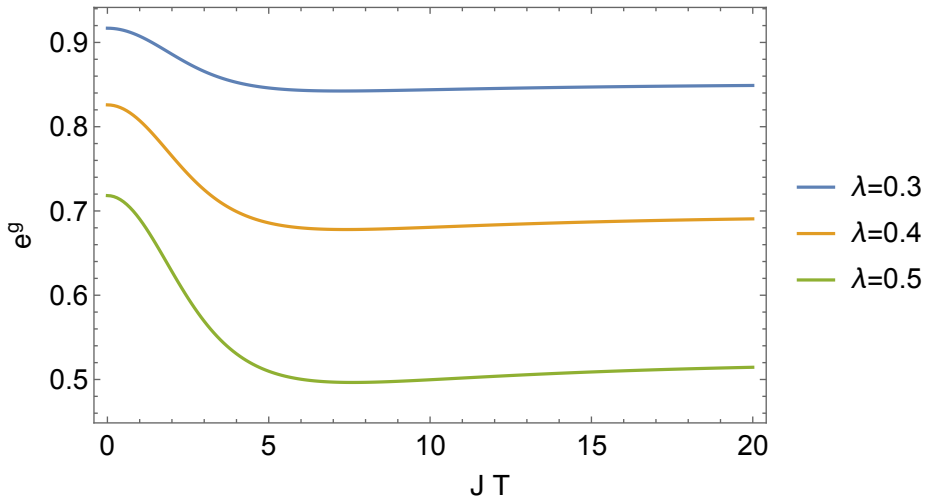


Figure 4.10: Here we show the disk contribution to the 2-pt correlator $e^{g(t_1,t_2)}$ evaluated at the quench sites ($t_1 = 0$, $t_2 = \beta/2 + iT$) for $\lambda = 0.3, 0.4, 0.5$, and $\beta\mathcal{J} = 5$. Note that at small values of λ the correlator stays large, which indicates that the distance between the quench sites is staying relatively small, even at large Lorentzian times. This is qualitatively similar to the Schwarzian behavior where the geometry gets pinched by the brane.

Evaluation of the action

Now we would like to evaluate the on-shell action of the disk. We start with the large-q Liouville action [114]:

$$S = \frac{N}{8q^2}\int dt_1 dt_2\ \partial_1 g \partial_2 g - 4\mathcal{J}^2 e^{-\mu\Omega} e^{g} \tag{4.86}$$

Then in terms of $\hat{g}$ (4.76):

$$S = \frac{N}{8q^2}\int dt_1 dt_2\ \partial_1 \hat{g} \partial_2 \hat{g} + \mu^2 \partial_1\Omega\partial_2\Omega - \mu\partial_1\hat{g}\partial_2\Omega - \mu\partial_1\Omega\partial_2\hat{g} - 4\mathcal{J}^2 e^{\hat{g}} \tag{4.87}$$

Let us consider $\partial S/\partial\mu$. Since we are evaluating the action on-shell, only the explicit μ dependence contributes:

$$\begin{aligned} -\partial S/\partial\mu &= \frac{N}{8q^2}\int \partial_1\hat{g}\partial_2\Omega + \partial_2\hat{g}\partial_1\Omega \\ &= -\frac{N}{8q^2}\int dt_1 \partial_1\hat{g}(t_1,\tau)\mathrm{sgn}(t_1-\tau) + \int dt_2 \partial_2\hat{g}(\tau,t_2)\mathrm{sgn}(t_2-\tau) \\ &= \frac{N}{2q^2}\left[\hat{g}(\tau,\tau) - \hat{g}(\beta,\tau)\right] \end{aligned} \tag{4.88}$$

So to compute the action, we may simply integrate the correlator as a function of λ

$$I(\tau,\bar{\tau}) = -\frac{N}{q^2}\int_{\tilde{\lambda}=\lambda}^{1} 2\tilde{\lambda}^3 d\tilde{\lambda}\log\left[\frac{\alpha_1(\tilde{\lambda})}{\sin\Big(\alpha_1(\tilde{\lambda})\tau + \gamma_1(\tilde{\lambda})\Big)}\right] \tag{4.89}$$

In this expression, we have subtracted the action at $\lambda = 1$, which is required when we normalize the thermofield double. This expression can be computed numerically. We display the result in Figure 4.11. We see that the answer grows at late times T. This suggests that the disk contribution alone will lead to a violation of unitarity at late times. (It does not prove that there is a violation since in principle quantum corrections could in principle stop the growth, but based on both the quantum Schwarzian computations and the finite N numerics, this seems unlikely.)

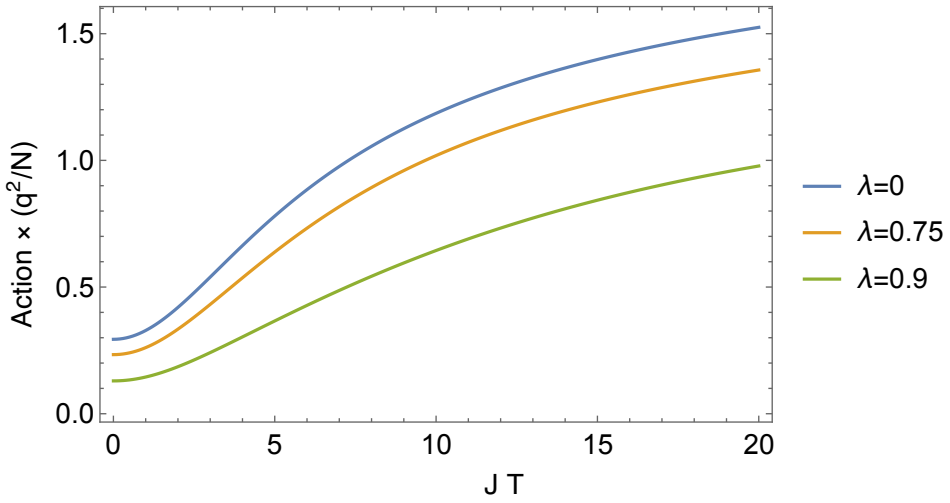


Figure 4.11: The large N on-shell action of the disk as a function of $\mathcal{J}T$ for $\beta\mathcal{J} = 5$ for various values of λ. The growth of the action implies that the typical overlap $\langle \mathsf{TFD}, J|\mathsf{TFD}, J'\rangle \sim e^{-I}$ is shrinking as a function of time. Smaller values of λ lead to more decorrelated states, which agrees with the larger values of the action. The growth at late times suggests that there is a unitarity problem if we only include the disk solution.

4.2.5 Wormhole in large q SYK

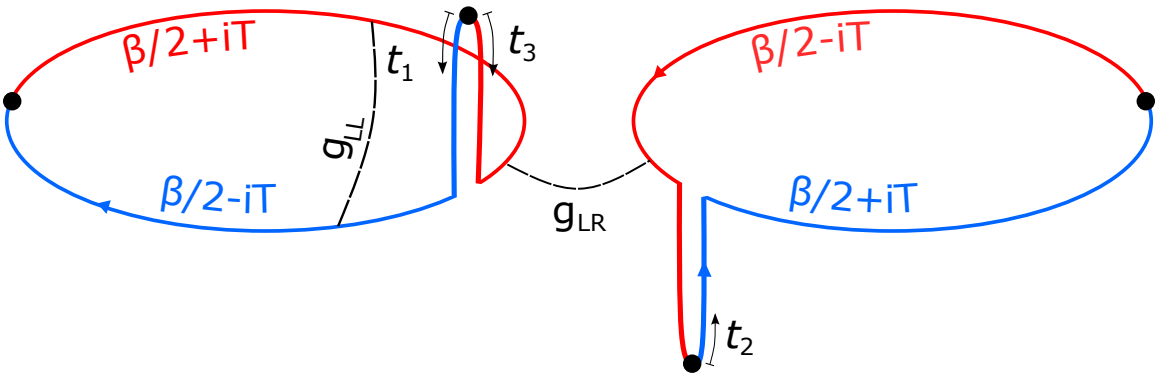


Figure 4.12: Contour relevant for the Renyi-2 wormhole, with our conventions labeled. The black arrows start at the points where $t_i = 0$ and they point in the direction of positive t_i. The correlator between points t_1 and t_2 behaves as if they are on two sides of a thermofield double.

Here we present a preliminary exploration of the wormhole in large q SYK, leaving a more thorough analysis for the future. When we introduce a second side, the large N variables include a g_{LR}. Remarkably in the standard large q approximation, the g_{LL} and g_{LR} variables are uncoupled in the

equations of motion:

$$\partial_1\partial_2 g_{\mathsf{LL}} = -2\mathcal{J}^2 \exp\left(g_{\mathsf{LL}} - \mu\Omega_{\mathsf{LL}}\right), \quad \partial_1\partial_2 g_{\mathsf{LR}} = 2\mathcal{J}^2 \exp\left(g_{\mathsf{LR}} - \mu\Omega_{\mathsf{LR}}\right). \tag{4.90}$$

Furthermore, the equations of motion just differ by a sign, which can be accounted for by changing the direction of time. The basic idea is that using the same tricks as above, we will have a g_{LR} that obeys a Liouville like equation. So we expect a solution that is very similar to the Schwarzian wormhole, except that we "build" the solution from the large q disk solution. More explicitly, consider an ansatz:

$$\begin{aligned}
\exp\left(g^{\mathsf{LR}}_{\mathsf{toc}}(u_1,u_2)\right) &= \left[\frac{\alpha}{\mathcal{J}\sin\left[\alpha\left(u_1-\tilde{u}_2\right)+\gamma\right]}\right]^2, \quad \tilde{u}_2 = u_2 - u_b \\
\exp\left(g^{\mathsf{LR}}_{\mathsf{otoc}}(u_3,u_2)\right) &= \frac{\alpha^2}{\mathcal{J}^2}\left[\sin\left(\alpha\left(u_3-\tilde{u}_2\right)+\gamma\right) - \frac{\mathcal{J}\left(1-\lambda^2\right)}{\alpha}\sin\left(\alpha u_3\right)\sin\left(\alpha\tilde{u}_2\right)\right]^{-2} \\
\exp\left(g^{\mathsf{LL}}_{\mathsf{otoc}}(u_1,u_3)\right) &= \frac{\alpha^2}{\mathcal{J}^2}\left[\sin\left(\alpha\left(u_1-u_3\right)+\gamma\right) - \frac{\mathcal{J}\left(1-\lambda^2\right)}{\alpha}\sin\left(\alpha u_1\right)\sin\left(\alpha u_3\right)\right]^{-2} \\
\exp\left(g^{\mathsf{LL}}_{\mathsf{toc}}(u_1,u_1')\right) &= \left[\frac{\alpha}{\mathcal{J}\sin\left[\alpha\left(u_1-u_1'\right)+\gamma\right]}\right]^2.
\end{aligned} \tag{4.91}$$

We are adopting a convention where $u_i = 0$ corresponds to a quench site, and the second quench site on the left (right) side is at $u_1 = \tau$ ($u_2 = \tau$). Furthermore, u_1, u_3 runs clockwise on the LHS, whereas u_2 runs counterclockwise on the RHS. So for example, the solution $g^{\mathsf{LR}}_{\mathsf{toc}}$ is valid when $u_1 u_2 < 0$ before u_1 or u_2 cross another quench site. This accounts for the sign difference in the equations of motion of (4.90). To analytically continue, we set $u_i = it_i$. This gives a sign convention t_i illustrated in Figure 4.12. One can confirm that the equations of motion are satisfied if $\alpha = \sin\gamma$. The remaining task is to determine α and the time shift u_b.

As a warm-up, it is useful to reconsider the Schwarzian wormhole from a slightly different perspective. There, too, the equations of motion are locally the same as for the disk solution. The main idea is that we can build the wormhole by considering the disk solution with some auxiliary temperature β_{aux}. The disk solution in the Schwarzian limit is obtained by taking two disks of circumference β_E and joining them together at the quench sites. The wormhole is constructed by cutting along the diameter of each partial disk and then gluing to another copy of the solution, see figure 4.16 and Appendix 4.8.1. Pasting together the two solutions gives a new solution that is a topological cylinder.

Now before cutting, the total length of the doubled disk solution is $2\beta_{\text{aux}}$. When we cut along

the diameter of the partial disks and join the two copies, we remove a portion of the boundary that has length $4 \times (\beta_E/2)$. Then requiring that the total length of the boundary is 2β (a factor of β for each side) gives

$$2\beta_{\text{aux}} - 2\beta_E = 2\beta \tag{4.92}$$

Now for the disk solution, we have from (4.84)

$$-\tan\left(\frac{\beta_{\text{aux}}\pi}{2\beta_E}\right) = \frac{4\pi C}{\delta\beta_E}. \tag{4.93}$$

Eliminating β_{aux} reproduces (4.61), as desired:

$$\tan\left(\frac{\pi\beta}{2\beta_E}\right) = \frac{\delta\beta_E}{4\pi C} \approx \frac{(1-\lambda^2)\beta_E\mathcal{J}}{2\pi}. \tag{4.94}$$

The last equality is based on the approximate relation derived in the previous subsection, around (4.83).

Now we can follow the same strategy to find β_E as a function of β in the large q solution. The key point is that although we do not have a clear geometric picture like in the Schwarzian case, nevertheless the toc correlator is exactly thermal at an effective temperature $1/\beta_E$, so the procedure is entirely analogous. In particular, just as we cut $\beta_E/2$ of the boundary particle on the partial disk, we remove $\beta_E/2$ of the toc correlator, and paste it to a second copy. If we enforce equation (4.92), this pasting procedure gives a smooth correlator. Indeed, the special thing about removing exactly $\beta_E/2$ of the solution is that the toc 2-pt function is minimized at this time. So when when we continue the 2-pt function g_{LR} onto the "back side" of the solution, there will be no discontinuity in any of its derivatives.

Now while (4.92) holds for both the Schwarzian and the large q theory, the relation between β_E and β will be modified. For the toc correlator with any α, we can define an effective inverse temperature β_E:

$$\frac{\alpha}{\mathcal{J}} = \cos\left(\frac{\alpha\beta_E}{2}\right). \tag{4.95}$$

In general, β_E differs from β_{aux} due to the added energy from the insertion of the operator. Together

with the relation between β_{aux} and α derived in the previous subsection

$$\sin\left(\alpha\beta_{\text{aux}}/2 + 2\gamma\right) = \lambda^2 \sin\left(\alpha\beta_{\text{aux}}/2\right), \alpha/\mathcal{J} = \sin\alpha \tag{4.96}$$

and (4.92), we can eliminate β_{aux} and obtain β_E as a function of β for the wormhole solution. We show the results for in 4.13.

This solution for $\alpha(\beta_E)$ can be substituted into (4.91) along with

$$u_b = \beta_{\text{aux}}/2 = (\beta + \beta_E)/2. \tag{4.97}$$

The last requirement just says that in the auxiliary disk solution, the second quench site appears after a boundary time $\beta_{\text{aux}}/2$. We can check that with the above requirements, all the toc correlators obey the correct boundary conditions, e.g. $g_{\mathsf{LR}}(0,0) = g_{\mathsf{LR}}(\tau,\bar{\tau}), g_{\mathsf{LR}}(0,\bar{\tau}) = g_{\mathsf{LR}}(\tau,0)$.

We expect that the wormhole solution discussed above to be an approximate solution to the large N equations of motion when the Lorentzian time T is large. To borrow the Schwarzian language, the reason why we need this condition is that we have essentially ignored "windings." In the Schwarzian description, there are multiple geodesics that connect two points; in general, a correlator will receive contributions from all of these geodesics unless the wormhole throat b is very long (which happens at large T, see Appendix 4.8). To state the issue in the large q formalism, note that we have described the solution in patches (4.91), but actually the patches overlap. While the toc solutions are valid whenever the two times are on the same side of a quench, the otoc solutions are supposed to be valid "near" the quench site at $u_1 = u_3 = 0$. When the u_i are near the opposite quench site, we should use a similar ansatz:

$$\begin{aligned}
\exp\left(g^{\mathsf{LL}}_{\mathsf{otoc}}(u_1,u_3)\right) &= \frac{\alpha^2}{\mathcal{J}^2}\left[\sin\left(\alpha\left(u_1-u_3\right)+\gamma\right) - \frac{\mathcal{J}\left(1-\lambda^2\right)}{\alpha}\sin\left(\alpha(u_1-\tau)\right)\sin\left(\alpha(u_3-\tau)\right)\right]^{-2} \\
\exp\left(g^{\mathsf{LR}}_{\mathsf{otoc}}(u_3,u_2)\right) &= \frac{\alpha^2}{\mathcal{J}^2}\left[\sin\left(\alpha\left(u_3-\tilde{u}_2\right)+\gamma\right) - \frac{\mathcal{J}\left(1-\lambda^2\right)}{\alpha}\sin\left(\alpha(u_3-\tau)\right)\sin\left(\alpha(\tilde{u}_2-\tau)\right)\right]^{-2}
\end{aligned} \tag{4.98}$$

We believe that the full solution to the large N equations of motion is close to a sum of the two ansatzs the large αT limit, with a sign appropriate for the periodicity/anti-periodicity of the correlators. Summing the two ansatz only makes sense if the correlators are small in the overlap region. For fixed

$\beta\mathcal{J}, \lambda < 1$ this is satisfied if $\alpha T \gg 1$, since the correlators decay exponentially in the overlapping region $e^g \sim \alpha^2 e^{-2\alpha T}$. A similar issue arises in finding the double cone in SYK [7] and also for the finite temperature wormhole in the coupled SYK model see also [27]. To check this more carefully, one would need to use a different large q approximation, discussed in [27] that is valid when the correlators G_{LL} are small (and therefore $|g_{\mathsf{LL}}/q| \gg 1$.) It would also be interesting to solve the finite q G, Σ equations numerically.

A regime where the solution simplifies is when $\beta\mathcal{J} \gg 1$ while keeping $\lambda < 1$ fixed. The wormhole correlators become

$$\begin{aligned} \exp\left(g^{\mathsf{LL}}_{\mathsf{otoc}}(u_1, u_3)\right) &= \frac{1}{\left(-(1-\lambda^2)\mathcal{J}^2 u_1 u_3 + \mathcal{J} u_{13} + 1\right)^2} \\ \exp\left(g^{\mathsf{LR}}_{\mathsf{toc}}(u_1, u_2)\right) &= \frac{1}{\left(\mathcal{J}(u_1 - u_2) + 1\right)^2} \\ \exp\left(g^{\mathsf{LR}}_{\mathsf{otoc}}(u_3, u_2)\right) &= \frac{1}{\left(-(1-\lambda^2)\mathcal{J}^2 u_3 u_2 + \mathcal{J} u_{32} + 1\right)^2} \end{aligned} \tag{4.99}$$

Notice that the correlator across the wormhole is maximal at $u_1 = u_2 = 0$! This means that the wormhole is basically what is depicted in the right side of Figure 4.7. Since both circles intersect the asymptotic boundary, we can interpret the figure as either having a single boundary, or two boundaries that are infinitely long, where the two boundaries are separated at asymptotic infinity. This wormhole solution can be contrasted with the disk solution in the same limit. At large $(1-\lambda)^2\beta \gg 1$, the wormhole relation is $\beta_E \approx \beta$. On the other hand, the disk in this limit gives $\beta_E \approx \beta/2$. This means that for the disk, the 1-sided correlator $G_{\mathsf{LL}}(0, \beta/2 + iT) \approx 1$ is large whereas $G_{\mathsf{LR}} = 0$. But for the wormhole $G_{\mathsf{LL}}(0, \beta/2 + iT)$ is small, but the two-sided correlator $G_{\mathsf{LR}}(0, 0) \approx 1$. Such qualitative behavior of the correlators seems to be roughly in agreement with the $q = 4, N = 20$ numerics, see Figure 4.4.

Note that our setup has various exact discrete global symmetries, which implies that there are multiple wormhole solutions. These symmetries are the same as the ones for the "ramp" in SYK [7], so we will be brief.First note that $(-1)^F_L$ and $(-1)^F_R$ generate a $\mathbb{Z}_2 \times \mathbb{Z}_2$ global symmetry, that leaves G_{LL} and G_{RR} invariant but takes $G_{\mathsf{LR}} \to -G_{\mathsf{LR}}$ when only one $(-1)^F$ is applied. Thus any wormhole solution $G_{LR} \neq 0$ spontaneously breaks one of the two $\mathbb{Z}_2$ symmetries; further, the solutions come in pairs. For the large q solutions, we should write $G_{\mathsf{LR}} = \pm e^{g_{\mathsf{LR}}/q}$ When q is a multiple of 4, there is the additional symmetry $G_{\mathsf{LR}}(t, t') \to iG_{\mathsf{LR}}(t, -t')$, $\quad G_{\mathsf{RL}}(t, t') \to iG_{\mathsf{RL}}(-t, t')$, $\quad G_{\mathsf{RR}}(t, t') \to -G_{\mathsf{RR}}(-t, -t')$ that should be treated similarly.

Another setup in which one can study the wormhole of the journal analytically is Brownian SYK.

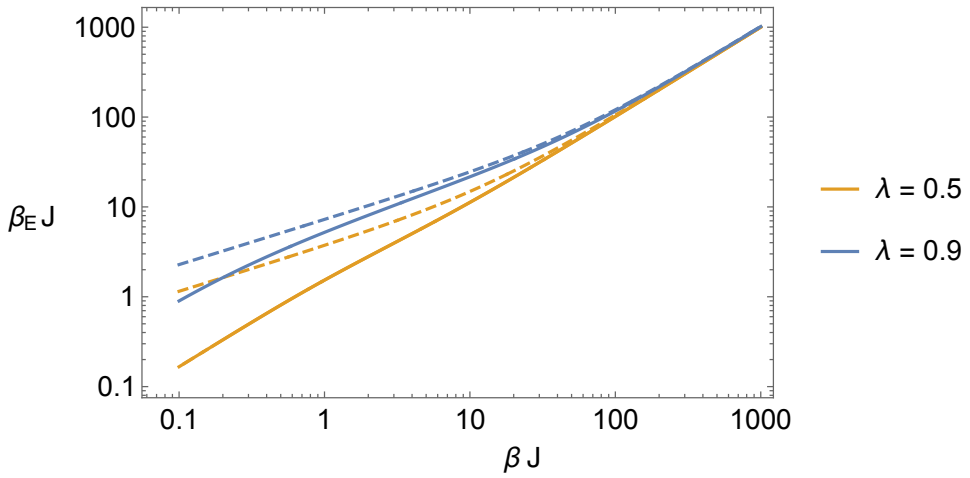


Figure 4.13: The effective inverse temperature β_E as a function of β in units of $\mathcal{J}$. We show the results for $\lambda = 0.9$ and $\lambda = 0.5$. The dashed lines are the Schwarzian predictions given by (4.94). At very large $\beta\mathcal{J}$, the growth is approximately linear $\beta_E \approx \beta$.

There, the journal contains a whole function's worth of couplings $\{J_{ijkl}(t)\}$. See Appendix 4.9.

4.3 Reconstruction with erroneous knowledge of the couplings

The goal of this section is to illustrate how bulk reconstruction using the boundary system is affected by erroneous or incomplete knowledge of the couplings. The discussion will be mostly qualitative. We will keep things simple and use the Petz Lite protocol of [136].

In the simplest setting of bulk reconstruction, one considers a "code subspace" of the boundary Hilbert space that share the same bulk background geometry but differ in the state of the matter. In our case, we consider a code subspace composed of perturbations of the time evolved thermofield double state $|\beta + iT; J\rangle$, where J labels the couplings of the Hamiltonian H_J used to prepare the state. We use $|i; J\rangle$ to denote the different matter states spanning the code subspace. Note that we are not yet averaging over couplings.

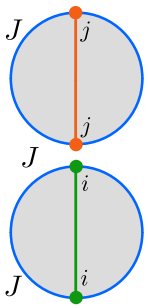
Implementing successful bulk reconstruction we will take to mean that there's an operator that can transition between any two states of the code subspace; for i, j, k, l in some orthonormal basis, there exists an operator $\mathcal{O}_{ji}$ such that

$$\langle k; J|\mathcal{O}_{ji}|l; J\rangle \approx \delta_{kj}\delta_{li}. \tag{4.100}$$

We use the approximate signs $\approx$ to indicate that we are ignoring non-perturbatively small effects from non-factorization wormholes, and also to account for the approximation of using Petz lite. This condition is easily satisfied by constructing the "global" J operator,

$$\mathcal{O}^J_{ji} \equiv |j; J\rangle\langle i; J|, \tag{4.101}$$

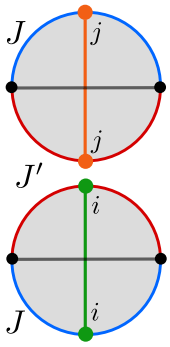
constructed out of the basis states of the code states. This gives the square of the norm of the states:

$\langle k; J|\mathcal{O}^J_{ji}|l; J\rangle =$  $\times\, \delta_{jk}\delta_{li}$ (4.102)

The setting we want to consider is where we don't know the value of the coupling J in the

Hamiltonian that's used in preparing the code subspace, and we are tasked with finding an operator that transitions between two specified code words i and j. This is risky because using a wrong value of the coupling, say J' instead of J, leads to an eventual breakdown of the reconstruction:

$$\langle j; J|\mathcal{O}_{ji}^{J'}|i; J\rangle = \langle j; J|j; J'\rangle\langle i; J'|i; J\rangle = \quad (4.103)$$



For $J \neq J'$, the right hand side decays in time due to the backreaction of the shocks created by the sudden change in the boundary conditions. We showed that in the case of JT + free boson BCFT model of section 4.2, that this decays exponentially

$$\langle j; J|\mathcal{O}_{ji}^{J'}|i; J\rangle \approx \left[\frac{\pi\epsilon}{\beta\cosh(\pi T/\beta)}\right]^{\frac{(J-J')^2}{2\alpha'\pi^2}} \quad (4.104)$$

Note that we are ignoring effects from computing the overlap in excited states from the TFD, hence the approximate sign. We are also ignoring gravitational backreaction and quantum Schwarzian effects. The reconstruction therefore fails at late times, since the states prepared with different couplings become more orthogonal under time evolution. This is the same effect that resulted in a unitarity problem, which hinted at the presence of an island.

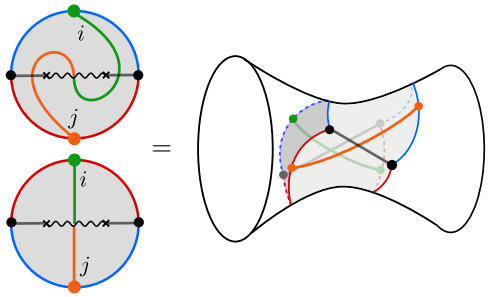
If we are given the knowledge of the distribution of the original couplings, then we can do better. A natural procedure is to average the couplings in the operator over this distribution. We can estimate the failure on average by also averaging over the couplings of the system:

$$\int dJdJ'P(J)P(J')\langle j; J|\mathcal{O}_{ji}^{J'}|i; J\rangle \approx \frac{1}{1+\frac{2}{\pi^2\alpha'm^2}\ln\left[\frac{\beta}{\pi\epsilon}\cosh(\pi T/\beta)\right]} \approx \frac{\beta\pi\alpha'm^2}{2T} \quad (4.105)$$

At first sight it appears that the reconstruction fails, even on average, albeit more slowly than if we don't average. However, we note that this decay only depends on the statistical properties of the couplings, and also weakly on i, j (not shown here). This means we can improve our reconstruction by simply scaling the operator by a time dependent factor $C_{ij,m}(T)$. We will take this to mean that the reconstruction is successful, at least on average.

However, there's another, more severe way in which this operator can fail: the sum over couplings enhances the contribution of a wormhole that connects the bra/ket of the operator/state its ket/bra. This in particular means that "flipped" matrix elements which should vanish actually get a contribution:

$$\int dJdJ' P(J)P(J')\langle i;J|\mathcal{O}^{J'}_{ji}|j;J\rangle \approx$$

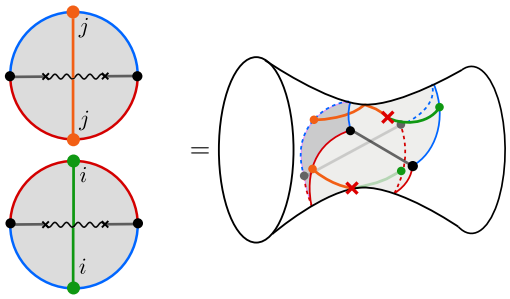


(4.106)

Averaging over the same value of the boundary coupling is indicated by like colors. At early times, this is a small contribution that only mildly violates our condition (4.100), but at large times, when the disk is highly suppressed, it is comparable to the "good" matrix elements $\langle j;J|\mathcal{O}^{J'}_{ji}|i;J\rangle$ that come from the disk (which are decaying). So, if we rescale the operator to set the "good" elements ≈ 1, we will also rescale these "flipped" elements $\langle j;J|\mathcal{O}^{J'}_{ji}|i;J\rangle \sim 1$. Thus we see that this reconstruction attempt fails badly at late times.

Furthermore, even when the disk contribution to the "good" matrix elements is small, there is not a substantial contribution from the wormhole since the bulk states in the throat are orthogonal, and winding contributions decay exponentially in $b \sim T$. One can see graphically that the overlap is zero by considering

$$\int dJdJ' P(J)P(J')\langle j;J|\mathcal{O}^{J'}_{ji}|i;J\rangle \approx$$

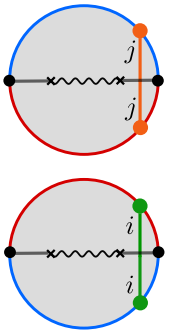


(4.107)

The red crosses above indicate the orthogonality of the bulk states. A similar use for wormholes was first pointed out in [136]. However, failure is not guaranteed because the particles can pair up

outside the branch cut:

$$\int dJdJ' P(J)P(J')\langle j;J|\mathcal{O}_{ji}^{J'}|i;J\rangle \approx$$



(4.108)

For such operators outside the horizon, the wormhole and the disk contribution give similar answers, so there are no significant contributions to the "flipped" elements and rescaling the operator with time should succeed.

The right interpretation of this is that the entanglement wedge of the boundary develops a blind spot in the bulk; the reconstruction fails because the particle falls into the island. This island is the entanglement wedge of reference purifying the system, namely what we've been calling the journal in the previous sections. With this interpretation, the averaged operator used in (4.105) is nothing but the Petz lite operator, obtained by tracing out the journal,

$$O_{ji}^{\text{Petz Lite}} = \text{Tr}_{\text{journal}}\left[\int dJ\sqrt{P(J)P(J')}|j;J\rangle\langle i;J'|_{\text{Sys}} \otimes |J\rangle\langle J'|_{\text{journal}}\right] \tag{4.109}$$

$$= \int dJP(J)\,|j;J\rangle\langle i;J|_{\text{Sys}} \tag{4.110}$$

We can adapt the above discussion to the case of having partial knowledge by defining a reconstruction controlled by the result of measuring the couplings. This is just a controlled Petz Lite,

$$\mathcal{O}_{ji} = \sum_{\kappa} \underbrace{\sum_{\mu} P(\mu,\kappa)\,|j;\kappa,\mu\rangle\langle i;\kappa,\mu|_{\text{sys}}}_{O_{ji}^{\text{Petz Lite}}(\kappa)} \otimes |\kappa\rangle\langle\kappa|_{\text{known}} \tag{4.111}$$

We note that this can be generalized to the case of full Petz with the replacement

$$O_{ji}^{\text{Petz Lite}}(\kappa) \to \left[\Pi_{\text{Sys}}^{\text{code}}(\kappa)\right]^{-\frac{1}{2}} O_{ji}^{\text{Petz Lite}}(\kappa) \left[\Pi_{\text{Sys}}^{\text{code}}(\kappa)\right]^{-\frac{1}{2}} \tag{4.112}$$

where

$$\Pi^{\mathsf{code}}_{\mathsf{Sys}}(\kappa) \equiv \mathrm{Tr}_{\mu,\kappa'}\left[\Pi_{\kappa'}\Pi^{\mathsf{code}}\Pi_{\kappa'}\right] \tag{4.113}$$

is the trace over the couplings of the projector onto the code subspace when projected onto a given value of the known couplings.

It is interesting to speculate how the failure of reconstruction due to islands manifests itself in a theory with fixed couplings. Given that it arises from the emergence of an island and wormholes (as a result of averaging) it suggests that something like half wormholes [125] are important for reconstructing the black hole interior. This suggests that we need to go beyond semi-classical gravity to address the firewall question [137, 138].

Besides the Petz map, one can also try to use simpler methods of bulk reconstruction. One option is to use the Maldacena-Qi Hamiltonian [27] which directly couples the left and right sides[7]. In [27], they showed that the wormhole does not require perfect correlation between couplings. This suggests that at early times, as long as the couplings are not too uncertain, evolution with the coupled Hamiltonian will still lead to an eternal traversable wormhole, which implies that the entanglement wedge of the two sides includes all of AdS. At late times, the wormhole is very long, so the correlations between the two sides is not strong enough to prevent the wormhole from growing. It would be interesting to understand this more quantitatively.

[7] We thank Juan Maldacena for suggesting this direction.

4.4 Discussion

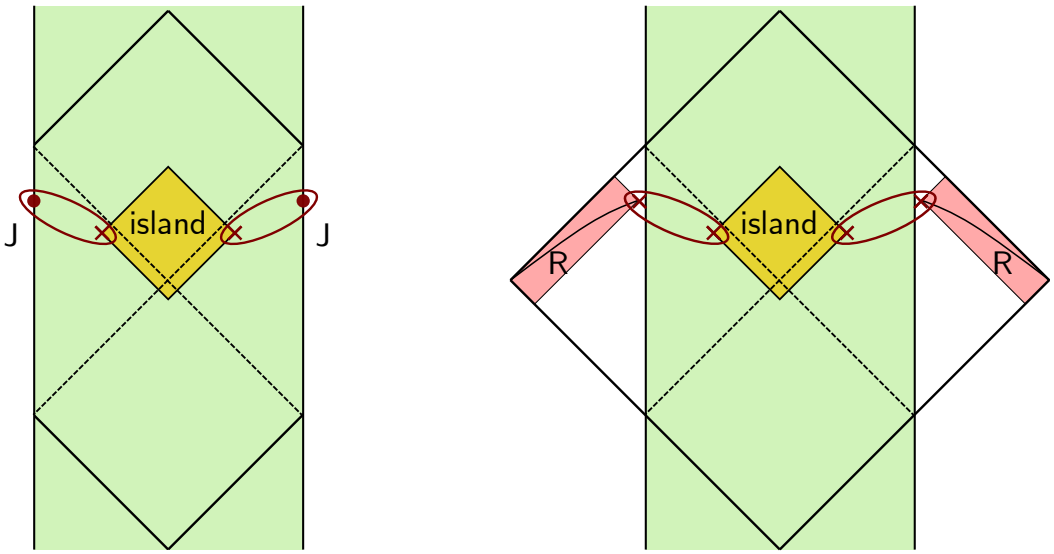


Figure 4.14: Island of the journal versus the island in the East coast setup. At late times, the separation between the twists from the island are nearly null separated from the boundary twists. In this OPE limit, one can approximate the entropy as the sum of two semi-infinite intervals discussed in section 4.1.2.

4.4.1 Bathing in the unknown

In the introduction of this chapter we motivated the entanglement entropy between the system and the journal as a measure of the uncertainty in the couplings, and we studied its growth under controlled time evolution and its effects on the entanglement wedge of the system.

Here we briefly motivate a different (but physically identical) picture of thinking of the journal as a bath with which the black hole system interacts. This makes the problem more analogous to the East Coast model [139], where the black hole interacts with a bath. Just as an example, we can consider the SYK system. The Hamiltonian [8] on the combined SYK and journal systems is

$$H = \sum_{i<j<k<l} \hat{J}_{ijkl}\psi_i\psi_j\psi_k\psi_l + \hat{J}^2_{ijkl}\frac{N^3}{6J^2} \tag{4.114}$$

Here we are using the undergraduate quantum mechanics "hat" notation to emphasize that

[8] To reproduce a probability distribution over the couplings that is independent of β, we will generally need to include an explicit β dependence in the Hamiltonian.

$\hat{J}$ is a quantum mechanical operator. The interaction is quite simple in a sense because there is no conjugate momentum to $\hat{J}$ in the Hamiltonian; however, it is sufficient to transfer quantum information to the journal. Then the formation of the journal is morally similar to the island of the radiation, see Figure 4.14.

4.4.2 Evaporating BH + journal

It would be interesting to study how the state of the Hawking radiation of an evaporating black hole is affected by the uncertainty in the couplings. A concrete case would be to consider the "East coast" model where we have JT gravity + CFT, joined to a flat space region where the CFT continues. In addition to CFT_1 in the bulk, we could also have a second CFT_2 that lives only in the black hole region. In other words, we have a boundary condition J for CFT_2 that prevents any transmission into the flat space region. Furthermore, we assume that in the semi-classical picture there is no interaction between CFT_1 and CFT_2.

We get information about the state of the radiation by looking for the entanglement wedge of the radiation. In this setup there are two systems, the radiation and the journal, vying for ownership of the island. We can guess at the winner by comparing the rate of entropy growth of the two systems. The first to $2S_{BH}$ wins. We found in section 4.1.1 that the entropy of the journal grows only logarithmically, while it was shown in [89] that the entropy of the radiation grows linearly time. Hence we believe that the radiation wins the race, at least at first.

An open question is whether ownership of the island is ever transferred to the journal. If it does get transferred, then one interpretation is that large uncertainty in the couplings reduces the system to the semi-classical state; this would support the recent ideas connecting premature ensemble averaging to Hawking's calculation [90, 116, 119]. In the case that the island doesn't get transferred, then a possible interpretation is that the radiation alone purifies the journal. This translates physically to the statement that the couplings can be determined by the measurements on the Hawking radiation.

4.4.3 Spectral form factor

One kind of coupling that always exists is the overall normalization of the Hamiltonian λ. This case is interesting because we can make contact with the spectral form factor. If the overall coupling is uncertain, time evolution with $H_\lambda = \lambda H$ and inverse time evolution with $H_{\lambda'} = \lambda' H$ will not cancel

by an amount $T(\lambda - \lambda')$. The Renyi-2 entropy of the journal is therefore non-trivial:

$$\begin{aligned} t_r \rho^2 &= \int p(\lambda) p(\lambda') \left| t_r \left(e^{i(\lambda-\lambda')HT - \frac{1}{2}(\lambda+\lambda')\beta H} \right) \right|^2 \\ &= \int p(\lambda) p(\lambda') \left| Z\left(\frac{1}{2}(\lambda+\lambda')\beta + i(\lambda-\lambda')T \right) \right|^2 \end{aligned} \tag{4.115}$$

Here we see that the Renyi-2 is essentially the time-averaged spectral form factor. The specific kind of time-averaging depends on $p(\lambda)$. The fact that the spectral form factor cannot decay to zero is just a consequence of the unitarity bound on this Renyi entropy. Hence the "journal" perspective conceptually unifies Maldacena's information paradox [140] with the information paradox of evaporating black holes. Given that the plateau of the spectral form factor is not explained by a single wormhole but doubly non-perturbative effects [7, 8], one could wonder if such effects play a role in the entropy of the journal.

4.4.4 Chaos and the Loschmidt Echo

The lack of our ability to do bulk reconstruction even when there is a small uncertainty in couplings is closely related to quantum chaos. In holography, one usually diagnoses chaos using the otoc, e.g, we create a perturbation at some time in the past $W(-T)$, and watch it grow. The perturbation is localized at some time $-T$.

An alternative diagnostic of chaos is the Loschmidt echo. Both diagonistics involve (backwards) time evolution $\exp(iH_1T)$, but in the Loschmidt echo, one changes the Hamiltonian going forward $\exp(iH_2T)$ by a small amount $H_2 = H_1 + \epsilon W$. So whereas the otoc is a localized pertubation in time, the Loschmidt echo is completely de-localized. The simplest diagnostic of chaos is the decay of the inner product

$$\langle \psi' | \exp(-iH_2T) \exp(iH_1T) | \psi \rangle \tag{4.116}$$

Clearly if $H_2 = H_1$ this product does not decay, but in general it should decay and then exhibit some erratic oscillations of order e^{-S}. What we showed was that the disk contribution leads to a decay, but wormholes are needed to explain the erratic oscillations. We showed this for $|\psi\rangle = |\beta, J_1\rangle$, $|\psi'\rangle = |\beta, J_2\rangle$ in SYK and in JT gravity, but we believe that the lessons should be fairly general.

One can also study the size of the Loschmidt operator $\exp(-iH_2T)\exp(iH_1T)$. At infinite temperature, this is given by sandwiching the size operator $\sim \psi_L \psi_R$ with the state $\exp(-iH_2T)\exp(iH_1T)|0\rangle$. To define a finite temperature version of size, we would need to specify whether the thermal ensemble

should correspond to H_1 or H_2, but for small perturbations this is a minor detail. For large q SYK, it follows from the disk solution at $\lambda \approx 1$ in 4.2.4 that the size grows exponentially with Loschmidt Lyapunov exponent

$$\lambda_L = 2\alpha, \tag{4.117}$$

e.g, the same Lyapunov exponent as the thermofield double. At low temperatures, the maximal chaos exponent follows from the bulk picture and the relation between size and the symmetry generators [1]. Indeed, in the bulk, a change in the couplings at time $-T$ inserts a matter shockwave; the gravitational backreaction is responsible for the exponential growth in size. But we have calculated the exponent at finite temperature in the large q theory, where SYK is not simply described by the Schwarzian mode.

An interesting question is whether one can argue for a bound on the Loschmidt Lypaunov exponent $\lambda_L \leq 2\pi/\beta$ along the lines of [141].

4.4.5 Singularity?

Our work shows that the black hole interior is quite sensitive to the precise values of the couplings of the theory. Even ignorance about the irrelevant couplings appears to be enough to prevent the interior from being reconstructed at late times. This is surprising if we believe that irrelevant operators have small effects in the bulk. On the other hand, irrelevant operators can have a large effect near the black hole singularity. In JT gravity, one can show that the profile of a free scalar field with mass determined by Δ will diverge near the inner horizon except for integer values of Δ [1]. It is tempting to speculate that the island associated to the journal that forms at late times is trying to censor us from accessing the region near the singularity. Adopting the spirit of Penrose's cosmic censorship conjecture [142], if we do not have access to the UV couplings of the boundary theory, it seems reasonable that we cannot reconstruct the region near the black hole singularity that is sensitive to those couplings. This obviously deserves more study.

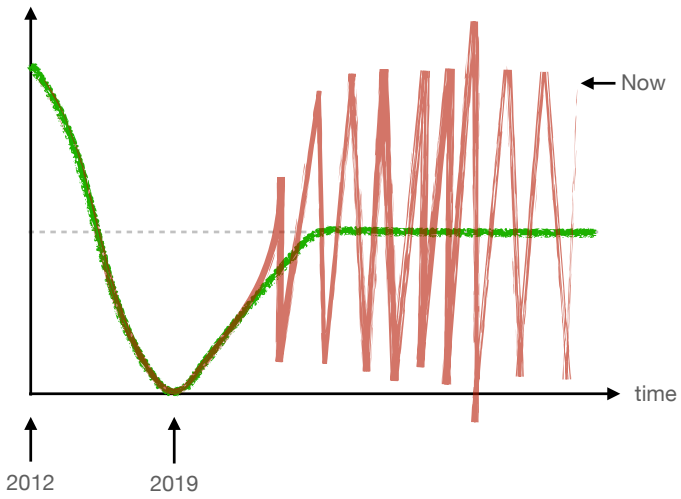
4.4.6 Janus's Journal

It would be interesting to consider the journal entangled to a "conventional" holographic theory which has no disorder average. We could consider a situation the journal records the value of λ in $\mathcal{N} = 4$ SYM. An off diagonal element of the journal density matrix would involve half of the

thermal cylinder with one value of the gauge coupling g and the other half with a different value g'. For the vacuum state, this problem is rather well-studied: one considers the Euclidean path integral on a sphere with gauge coupling g on half of the sphere and g' on the other half [143–146]. The problem at finite temperature would involve a path integral on a torus, with top and bottom halves of the torus differing [147]. It would be interesting to find wormholes in such a setup (or in other holographic setups with well-defined string duals), as this would pose a sharp factorization problem in $\mathcal{N} = 4$. More generally, one can consider a holographic CFT_1 joined to a CFT_2 at some interface, see [148–150] for some recent discussions. Our chapter suggests that the square of such quantites could have wormholes.

4.4.7 Conclusion

We are usually taught that gravity is a universal force, e.g., that spacetime is sourced by energy-momentum. What we seem to be learning here is that the spacetime in the interior of the black hole is in some sense highly non-universal. This seems related to the following equation:

$P(\text{firewall}) =$ 

(4.118)

What gravity doesn't know, neither do we...

4.5 Appendix: Compact free boson

In section 4.1, we considered the matter entropy of a non-compact boson, with Gaussian measure over the Dirichlet boundary conditions. Another case we can consider is the compact free boson $X \equiv X + 2\pi r$ with the same action as in (4.25). The boundary conditions we consider are again Dirichlet conditions on the free field, labeled by X. (For a review of boundary states in the compact

boson, see [151].) The boundary condition changing operator has dimension

$$\Delta_b = \frac{1}{\alpha'}\left(\frac{X_1 - X_2 + 2\pi r w}{2\pi}\right)^2, \quad w \in \mathbb{Z}. \tag{4.119}$$

In the compact case, there are actually infinitely many boundary condition changing operators that change $|X_1\rangle$ to $|X_2\rangle$; in string theory this is the fact that an open string ending on 2 D-branes can wind around the compact dimension w times. Then (4.24) becomes

$$Z_n = \frac{1}{(2\pi r)^n}\int \prod_{i=1}^{n} dX_i e^{-A(X_i - X_{i+1})^2/2} \tag{4.120}$$

We can evaluate this integral la Fadeev-Popov by first fixing $X_1 = 0$, performing $n-1$ integrals, and then integrating over X_1. This gives

$$Z_n = \frac{1}{(2\pi r)^{n-1}}\sqrt{\frac{(2\pi)^{n-1}}{\det \mathbf{M}}} = \sqrt{\frac{1}{n(2\pi A r^2)^{n-1}}}. \tag{4.121}$$

In writing the above equations, we implicitly summed over windings. This seems to be a choice that in principle might depend on UV regulator. The replica trick

$$S = -\,\partial_n\left(\frac{\log Z_n}{n}\right)\Big|_{n=1} = \frac{1}{2}\log 2\pi A r^2 - c'(1) \tag{4.122}$$

At large T, $Ar^2 \approx \frac{Tr^2}{2\pi\beta\alpha'}$. $S \sim \frac{1}{2}c\log(T/\beta)(r^2/\alpha')$ for c free bosons. When this entropy exceeds the entropy of the quantum dots $S \sim 2S_0$ there is a unitarity paradox. This happens at

$$T \sim \beta e^{2S_0/c}\alpha'/r^2. \tag{4.123}$$

4.6 Appendix: von Neumann entropy for 2 boundary states

Consider a general BCFT with two boundary states, each with equal probability. We can label the boundary conditions $\sigma = 1$ or $\sigma = -1$. Then Δ_i will be some value δ if the boundary conditions change, or else it will be 0. We can write this as $\Delta_i = \frac{\delta}{2}(\sigma_i\sigma_{i+1} - 1)$. Now we are instructed to

compute

$$Z_n = c_n \sum_{\sigma_i} \exp\left(-\frac{J}{2}\sum_{i=1}^{n}\sigma_i\sigma_{i+1} - 1\right) \tag{4.124}$$

This is an Ising model on a periodic lattice with n sites. Using the standard transfer matrix,

$$Z_n(J) = \left(1 - e^{-J}\right)^n + \left(1 + e^{-J}\right)^n, \tag{4.125}$$

$$S_m = (1 - \partial_n) \log Z_n = -\frac{1}{2}\log\left(1 - e^{-2J}\right) - e^{-J}\coth^{-1}\left(e^{J}\right) + \log(2) \tag{4.126}$$

This expression is real when $e^J > 1$. Notice that $J \to \infty$, we get $S_m = \log 2$. So if we have a very long interval, the matter entropy just reflects the uncertainty in the boundary condition. (The entropy of the Journal is saturated.) In writing this expression, we assumed that $c_1 = 1$ and that $\partial_n c|_{n=1} = 0$. The first condition is easy to justify but the second one is a bit mysterious.

More generally, we could choose the boundary conditions σ_i to appear with different probabilities. Then we would give an Ising model in a magnetic field.

4.7 Appendix: Zero temperature island

A case where we do not need to take the OPE limit is the zero-temperature or extremal limit. The geometry in this limit becomes the Poincaré plane:

$$ds^2 = (dx^2 + dy^2)/y^2, \quad y \in (\epsilon, \infty) \tag{4.127}$$

The Schwarzian boundary is located at $y = \epsilon$ and the Poincaré horizon is at $y \to \infty$. The BCFT bulk-boundary 2-pt function in these coordinates is

$$\left\langle O_{(h,0)}(y) O_{(0,h)}(\bar{y}) O_{(h_b,h_b)}(x')\right\rangle \sim \frac{y^{\Delta}\epsilon^{\Delta_\partial}}{(2z)^{\Delta-\Delta_\partial}(y^2+x^2)^{\Delta_\partial}} \tag{4.128}$$

Here we have included the warp factor $\Omega \sim 1/y$ for AdS. Notice that when the boundary dimension vanishes $\Delta_\partial = 0$, this correlator is independent of y, as required by symmetry. For our application, we take the bulk operator to be a twist and the boundary operator to be a boundary condition changing operator. When the separation $x = 0$, we get $c_n(\epsilon/y)^{\Delta_\partial}$.

For the case of 2 boundary states, we have $e^J = ye^{\delta}$. So the limit $J \to \infty$ corresponds to an

island that is very close to the Poincaré horizon. For a free boson,

$$\begin{aligned} S_m &= \frac{\log(\tilde{a})}{2} + \sqrt{4\tilde{a}+1}\coth^{-1}\left(\sqrt{4\tilde{a}+1}\right) + 1 - 2c_1' \\ \tilde{a} &\sim \log(y/\epsilon). \end{aligned} \tag{4.129}$$

4.8 Appendix: Renyi-2 entropy for a marginal journal

In this subsection we consider the problem of evaluating the integral (4.52) for a marginal deformation $\Delta = 1$:

$$-I_m = \frac{\chi_0^2}{2\pi}\int_0^{u_*} du_1 \int_0^{u_*} du_2 \frac{t'(u_1)t'(u_2)}{2\sin^2\left(\frac{t(u_1)-t(u_2)}{2}\right)} \tag{4.130}$$

To evaluate this integral, it is convenient to work in Poincare coordinates $f(u) = \tan t(u)/2$ instead of Rindler coordinates $t(u)$:

$$-\tilde{I}_m = \int du_1\, du_2 \left[\frac{f'(u_1)f'(u_2)}{\left(f(u_1)-f(u_2)\right)^2}\right] \tag{4.131}$$

One subtlety about this integral is that we need to regulate the divergence coming from $t_1 \to t_2$. A simple regulator is to restrict the integration domain to the region $|u_1 - u_2| > \epsilon$:

$$\begin{aligned} -I &= 2\int_0^{u_*-\epsilon} du_2 \int_{u_2+\epsilon} du_1 \left[\frac{f'(u_1)f'(u_2)}{(f(u_1)-f(u_2))^2}\right] \\ &= 2\int_0^{u_*-\epsilon} f'(u_2)du_2 \left[\frac{1}{f_2 - f(u_*)} - \frac{1}{f_2 - f(u_2+\epsilon)}\right] \\ &\approx 2\int_0^{u_*-\epsilon} du_2 \left[\frac{f'(u_2)}{f_2 - f(u_*)} + \epsilon^{-1} - \frac{f''(u_2)}{2t'(u_2)}\right] \\ &\approx 2\left[\log(f(u_*-\epsilon) - f(u_*)) - \log(f(0) - f(u_*)) - \frac{1}{2}\log(f'(u_*)) + \frac{1}{2}\log(f'(0)) + u_*\epsilon^{-1}\right] \\ &\approx 2\left[\frac{1}{2}\log\big(\epsilon^2 f'(u_*)f'(0)\big) - \log(f(0) - f(u_*)) + u_*\epsilon^{-1}\right] \\ &= (2u_*/\epsilon) + \log\left[\frac{\epsilon^2 f'(u_*)f'(0)}{\left(f(u_*)-f(0)\right)^2}\right] \end{aligned} \tag{4.132}$$

Mapping back to Rindler coordinates,

$$e^{-I} = e^{-\tilde{I}\delta} = e^{2u_*/\epsilon}\left[\frac{\epsilon^2 t'(u_*)t'(0)}{2\sin^2\left(\frac{t(u_*)-t(0)}{2}\right)}\right]^{\delta}, \quad \delta = \chi_0^2/(2\pi) \tag{4.133}$$

In SYK, we expect $\epsilon \sim 1/\mathcal{J}$, since the thermal two-pt function goes like $(\beta\mathcal{J})^{-\Delta}$. In computing the overlap between states, we should normalize the states, e.g., $\frac{|\langle\chi_0|0\rangle|^2}{\langle 0|0\rangle\langle\chi_0|\chi_0\rangle}$. The denominator $\langle\chi_0|\chi_0\rangle$ gives a similar factor, except that $u_* \to 2u_*$ and $\beta = 2u_*$ since we turn on the source over the entire thermal circle (both the bra and the ket). Using the same regulator as above, the denominator gives

$$\langle\chi_0|\chi_0\rangle \sim \left[e^{4u_*\delta/\epsilon}\frac{\epsilon^2 t'(u_*)t'(0)}{(\epsilon t'(0))^2}\right]^\delta \sim e^{4u_*\delta/\epsilon}, \tag{4.134}$$

So the effect of the denominator is to remove the $1/\epsilon$ divergence to the matter action. We conclude that the normalized overlap is

$$\frac{\langle\chi_0|0\rangle}{\sqrt{\langle 0|0\rangle\,\langle\chi_0|\chi_0\rangle}} = \frac{1}{\sqrt{Z(\beta)Z(2u_*)}}\int\frac{Dt}{\mathrm{SL}(2,\mathbb{R})}\,e^{-\mathsf{Sch}(t)}\left[\frac{\epsilon^2 t'(u_*)t'(0)}{2\sin^2\left(\frac{t(u_*)-t(0)}{2}\right)}\right]^\delta, \tag{4.135}$$

where $Z(\beta)$ is the Schwarzian thermal partition function, see [25]. For the SYK model, we expect that the UV regulator $\epsilon \sim 1/\mathcal{J}$ since the thermal two-pt function goes like $(\beta\mathcal{J})^{-\Delta}$. For JT gravity at finite cutoff, ϵ should be related to the JT cutoff $(ds/du)^2 = 1/\epsilon_{JT}^2$. Here we are imagining evaluating (4.133) with the classical thermal solution $t(u)$. However, our entire derivation also applies for off-shell $t(u)$. Hence a more accurate answer would be obtained by integrating (4.135) over all paths $t(u)$ with the Schwarzian action (4.135). Such an integral can be performed exactly, see e.g. [5, 6, 9, 131].

As a warmup for the wormhole, we can consider the thermal circle and turn on the source in two intervals, $[u_1, u_2]$ and $[u_3, u_4]$. This will clearly give a 2-pt function $\langle O(u_1)O(u_2)\rangle\,\langle O(u_3)O(u_4)\rangle$ when u and u' are both in the same interval. But we will also get a contribution when u, u' are in opposite intervals. There will be no divergences from this region of integration. We get

$$\begin{aligned}
-\tilde{I}_m &\supset 2\int_{u_1}^{u_2} du\int_{u_3}^{u_4} du'\left[\frac{f'(u)f'(u')}{(f(u)-f(u'))^2}\right] \\
&= 2\log\left(\left[\frac{f(u_2)-f(u_4)}{f(u_2)-f(u_3)}\right]\left[\frac{f(u_1)-f(u_3)}{f(u_1)-f(u_4)}\right]\right) \\
e^{-I} &\sim \left(\frac{f(u_2)-f(u_4)}{f'(u_2)f'(u_4)}\frac{f'(u_2)f'(u_3)}{f(u_2)-f(u_3)}\frac{f(u_1)-f(u_3)}{f'(u_1)f'(u_3)}\frac{f'(u_1)f'(u_4)}{f(u_1)-f(u_4)}\right)^\delta
\end{aligned} \tag{4.136}$$

Including the contribution from the two intervals, we get

$$e^{-I_m} \sim \mathcal{C}(u_1, u_2)\mathcal{C}(u_3, u_4) \left(\frac{\mathcal{C}(u_2, u_3)\mathcal{C}(u_1, u_4)}{\mathcal{C}(u_2, u_4)\mathcal{C}(u_1, u_3)} \right) \tag{4.137}$$

Here $\mathcal{C}(u_i, u_j)$ denotes the 2-pt function of a boundary primary of dimension δ:

$$\mathcal{C}(u_i, u_j) = \left[\frac{f'(u_i) f'(u_j)}{(f(u_i) - f(u_j))^2} \right]^{\delta} \tag{4.138}$$

Let us consider the OPE limit $u_2 \to u_3$ (on a circle, this is also equivalent to $u_1 \to u_4$). Then we get the simplification

$$e^{-I_m} \sim \mathcal{C}(u_2, u_3)\mathcal{C}(u_1, u_4) \tag{4.139}$$

Similarly in the OPE limit for the opposite channel $u_1 \to u_2$ or $u_3 \to u_4$:

$$e^{-I_m} \sim \mathcal{C}(u_1, u_2)\mathcal{C}(u_3, u_4) \tag{4.140}$$

So we see that in the OPE limit, the effect of the couplings is to insert a product of conformal 2-pt functions. We can think of this as inserting a pair of "domain walls" in the bulk. In the extreme limits, the domain walls will prefer to link up to their nearest partner.

More precisely, note that $\mathcal{C}_{12} \propto \exp(-\delta d_{12})$ where d is the renormalized geodesic distance in the bulk between u_1 and u_2. So when $d_{12} \ll d_{13}, d_{14}$ we get the first OPE limit, and similarly for the other limits.

Finally, note that one can generalize this computation easily to $n_I > 2$ intervals. In particular, since the matter action is bilocal in times, we will need to just consider each pair of intervals:

$$e^{-I_m} \sim \prod_{A=1}^{n_I} \mathcal{C}(u_{A1}, u_{A2}) \times \prod_{B<C} \left(\frac{\mathcal{C}(u_{B2}, u_{C1})\mathcal{C}(u_{B1}, u_{C2})}{\mathcal{C}(u_{B2}, u_{C2})\mathcal{C}(u_{B1}, u_{C1})} \right). \tag{4.141}$$

In this expression, A, B, C index the intervals; the endpoints of the interval are denoted by $A = [A1, A2]$.

4.8.1 Appendix: Renyi-2 wormhole

First consider the double trumpet geometry:

$$ds^2 = d\rho^2 + \left(\frac{b}{2\pi}\right)^2 \cosh^2\rho\, d\theta^2, \quad \rho \in (-\infty, \infty), \quad \theta \in [0, 2\pi] \tag{4.142}$$

This is conformally related to a cylinder of length π and radius b:

$$ds^2 = \frac{d\sigma^2 + \left(\frac{b}{2\pi}\right)^2 d\theta^2}{\sin^2\sigma}, \quad \sigma \in (0, \pi), \quad \theta \in [0, 2\pi] \tag{4.143}$$

To specify a cutout shape, one needs to specify some $\theta_L(u)$ and $\theta_R(u)$. We are interested in a setup in which the left circle has boundary conditions $\chi(u_L)$ and the right circle has boundary conditions $\chi(u_R)$. More specifically, $\chi(u_L)$ is constant and non-zero in some time interval $[0, u]$ and similarly for the right hand side. Now we can view the double trumpet (4.142) as a quotient of AdS_2 by a global time translation e^{ibE}. If we consider the universal cover of the double trumpet, we get the picture:

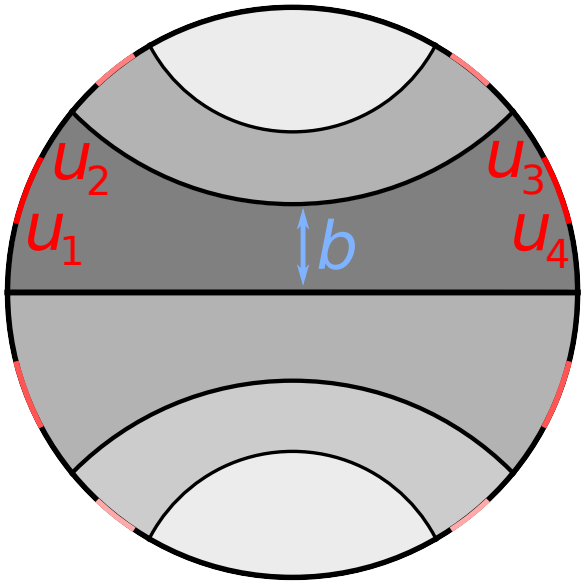


Figure 4.15: Here we display the fundamental domain of the double trumpet geometry (dark gray), as well as some of its images (lighter gray). We also show the parts of the boundary where a source is turned on (red) and its images (lighter pink).

We can evaluate the matter partition function by going to the universal cover. What started

out as a 2-interval computation on the cylinder becomes a computation involving infinitely many intervals. However, we can still apply (4.141). However, we claim that in the large Lorentzian time limit, we are in the OPE regime where (4.141) will just give

$$e^{-I_m} \sim \mathcal{C}^{LR}(u_1, u_4)\mathcal{C}^{LR}(u_2, u_3). \tag{4.144}$$

We will assume that this approximation holds; calculate the backreaction, and then check that our approximation is self-consistent.

To understand why this is the case, consider the expression for the matter action in these coordinates:

$$\begin{aligned} -I_m^{\mathsf{wormhole}} = &\frac{1}{2\pi}\int du_L\, du_R \left[\frac{\theta_L'(u_L)\theta_R'(u_R)}{2\cosh^2\left(\frac{b}{2\pi}\frac{\theta_L(u_L)-\theta_R(u_R)}{2}\right)}\right]\chi_L(u_L)\chi_R(u_R) \\ &+\frac{1}{2\pi}\int du\, du' \left[\frac{\theta_L'(u)\theta_L'(u')}{2\sinh^2\left(\frac{b}{2\pi}\frac{\theta_L(u)-\theta_L(u')}{2}\right)}\right]\chi_L(u)\chi_L(u') \\ &+\frac{1}{2\pi}\int du\, du' \left[\frac{\theta_R'(u)\theta_R'(u')}{2\sinh^2\left(\frac{b}{2\pi}\frac{\theta_R(u)-\theta_R(u')}{2}\right)}\right]\chi_R(u)\chi_R(u') \end{aligned} \tag{4.145}$$

The above formula is slightly imprecise since θ is a periodic variable. In principle we should sum over all windings $\theta \to \theta + 2\pi n$. However, we see that if b is very large, only a single winding will dominate.

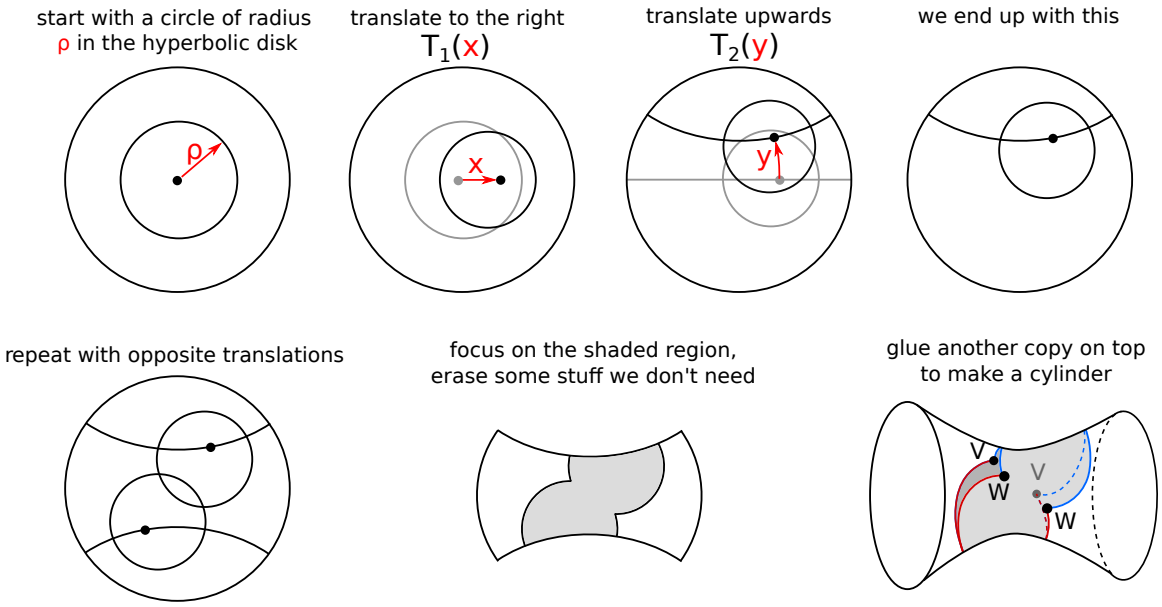


Figure 4.16: (Adapted from [132] with permission). This gives the cut-and-paste construction of the wormhole [132] that contributes to the Renyi-2 entropy. In Stanford's context, the wormhole arises from inserting 2 pairs of operators; in our context, these operators are boundary condition changing operators that arise from turning on a marginal source. At large Lorentzian times, we end up with essentially the same wormhole. If we already know the 2-operator disk solution, a "shortcut" procedure is to start with the solution on the bottom left, with some auxiliary inverse temperature β_{aux}. When we "erase some stuff we don't need," we are left with a boundary of length $\beta_{\text{aux}} - \beta_E$. After gluing to a second copy, we get the constraint $\beta = \beta_{\text{aux}} - \beta_E$.

Here ρ is the radius of the circle; the circumference of the circle defines an effective inverse temperature β_E,

$$2\pi \sinh \rho = \beta_E/\epsilon \tag{4.146}$$

where ϵ is the JT regulator. The relationship between the parameters $\{\beta_E, x, y\}$ and the physical

inputs $\{\beta, \tau, \delta\}$ is given in [132], equation (B.14-16):

$$\tan\left(\frac{\pi\beta}{2\beta_E}\right) = \frac{\delta\beta_E}{2\pi} \tag{4.147}$$

$$\sinh^2(y) = \frac{\left(\frac{\beta_E\delta}{2\pi}\right)^2}{1+\tan^2\left(\frac{\pi i t}{\beta_E}\right)} \tag{4.148}$$

$$\cosh^2(x) = \frac{\left(1+\left(\frac{\beta_E\delta}{2\pi}\right)^2\right)\left(1+\tan^2\left(\frac{\pi i\tau}{\beta_E}\right)\right)}{1+\left(\frac{\beta_E\delta}{2\pi}\right)^2+\tan^2\left(\frac{\pi i\tau}{\beta_E}\right)}. \tag{4.149}$$

In the small $\delta\beta$ limit, $\beta_E^2 = \pi^2\beta/\delta$. Now we can check that our approximation is self-consistent. At large τ,

$$b = 4y \approx \frac{4\pi}{\beta_E}\left(\tau - 2\tau_*\right), \quad \tau_* = \frac{\beta_E}{2\pi}\log\left(\frac{2\pi}{\beta_E\delta}\right). \tag{4.150}$$

Here τ_* can be interpreted as the scrambling time $\tau_* \sim \beta_E/2\pi \log\frac{S-S_0}{\Delta S}$. So at large times the wormhole is large and we expect the approximation to be valid.

Another regime we can explore is the extremal limit $\beta \to \infty, \tau = 0$. We also need to decide how δ scales in this limit. The simplest case is δ held fixed. Then $\pi\beta/(2\beta_E) \approx \pi/2$ so in this limit $\beta_E \approx \beta$. Furthermore, $y \sim \log\beta$, so at least classically the wormhole is getting large.

A useful parameter is the amount τ_a of the β_E circles that overlap, which governs the distance between W_L and W_R, or equivalently V_L and V_R. Consider the case $\tau = 0$ for simplicity. The disks of circumference β_E that are used to build the solutions are always cut along a geodesic that passes through the center of the disk, such that there is no discontinuinty in the bulk geometry. So requiring that the total boundary length is 2β gives

$$\tau_a = (\beta_E - \beta)/2. \tag{4.151}$$

$$\langle V_L V_R\rangle \sim \langle W_L W_R\rangle \sim \left(\frac{\pi}{\beta_E \sin\left(\pi\tau_a/\beta_E\right)}\right)^{2\delta} \tag{4.152}$$

To evaluate the action more carefully, we can follow [135].

Furthermore, the distances

$$\begin{aligned} e^{D_{12}} &\approx \left(\frac{\beta_E}{2\pi\epsilon}\right)^2 \sinh^2\left(\frac{\pi}{\beta_E}\left(-i\tau_a + t_{12}\right)\right) \\ e^{D_{13}} &\approx \left(\frac{\beta_E}{2\pi\epsilon}\right)^2 \left[i\sinh\left(\frac{\pi}{\beta_E}t_{13}\right) + \frac{\delta\beta_E}{\pi}\sinh\left(\frac{\pi}{\beta_E}t_1\right)\sinh\left(\frac{\pi}{\beta_E}t_3\right)\right]^2 \\ e^{D_{23}} &\approx \left(\frac{\pi}{\beta_E}\right)\left[\sinh\left(\frac{\pi}{\beta_E}\tau_{13}\right) + m\frac{\beta_E}{\pi}\sin\left(\frac{\pi}{\beta_E}(\beta-\tau_2)\right)\sin\left(\frac{\pi}{\beta_E}\tau_3\right)\right]^2 \end{aligned} \tag{4.153}$$

An interesting limit of these formula is when $\beta \to \infty$ with δ held fixed. Then $\tau_a \to 0$ and $\beta_E \to \infty$.

$$\begin{aligned} e^{D_{12}} &\approx (t_{12}/2\epsilon)^2 \\ e^{D_{13}} &\approx \left(i\frac{t_{13}}{2\epsilon} + \frac{\delta}{2\epsilon}t_1 t_3\right)^2 \end{aligned} \tag{4.154}$$

This agrees with the large q wormhole in the appropriate limit, see (4.99).

4.9 Appendix: Brownian SYK

in this section, we compute the Renyi-2 entropy of the journal, where the journal records the entire history of couplings $\{J_{ijkl}(t)\}$ for Brownian SYK. So the Hilbert space of the journal is that of a $\binom{N}{q}$ "fields" and not just that of a collection of point particles.

4.9.1 Decorrelated spectral form factor

Let us start by considering a simple generalization of the spectral form factor as a warmup. Usually in Brownian SYK, we write $\left\langle |Z(T)|^2 \right\rangle_J = \left\langle t_r U_J(T) t_r U_J(T)^{-1} \right\rangle_J$ where $\langle\cdots\rangle_J$ means to disorder average over J with the appropriate Gaussian measure. Here we consider the quantity $\left\langle t_r U_J(T) U_{J'}(T)^{-1} \right\rangle_{J,J'}$ where we disorder average over J, J' with a Gaussian measure that correlates J and J' by an amount $r = \langle JJ' \rangle$.

$$\begin{aligned} \left\langle t_r U_J(T) U_{J'}(T)^{-1} \right\rangle \approx & \int \mathcal{D}G\mathcal{D}\Sigma \exp\left\{-\frac{N}{2}\int_0^T dt \left[\frac{2J}{q}\left(\frac{1}{2^q} - i^q G(t)^q\right) + \Sigma(t)G(t)\right]\right\} \\ & \times \int \mathcal{D}\psi_a^{(L)}\mathcal{D}\psi_a^{(R)} \exp\left\{-\frac{1}{2}\int_0^T dt \left[\psi_a^{(A)}\partial_t\psi_a^{(A)} - \psi_a^{(L)}(t)\psi_a^{(R)}(t)\Sigma(t)\right]\right\} \end{aligned} \tag{4.155}$$

Here we are thinking of the L system as one of the $U(T)$'s and the R system as the other $U(-T)$. THe second line computes the trace over an $L \cup R$ of $2N$ fermions.

If we now consider time-independent solutions, we get a simple action

$$\exp\left\{N\left[\log\left(2\cos\frac{T\Sigma_{\mathsf{LR}}}{4}\right)-\frac{JT}{q2^q}+ri^q\frac{JT}{q}G_{\mathsf{LR}}^q-\frac{T}{2}\Sigma_{\mathsf{LR}}G_{\mathsf{LR}}\right]\right\} \tag{4.156}$$

Let's set $J=1$ from now on. We get the equations of motion

$$G=-(1/2)\tan(\Sigma T/4),\quad i^qG^{q-1}r=\Sigma/2 \tag{4.157}$$

This gives

$$\begin{aligned}\Sigma/2=ri^q\left(-1/2\right)^{q-1}\tan\left(\Sigma T/4\right)^{q-1},ri\sigma 2^{-q}=\Sigma\\ \sigma=\tanh^{q-1}\left(\sigma rT2^{-q}\right)\end{aligned} \tag{4.158}$$

There can be at most $q+1$ solutions to the above equation. Let us consider the case $q=2$ for simplicity. We get three solutions when $rT/4>1$. We get only 1 solution, $G=0$ when $rT/4<1$. The solution $G=0$ gives the exponential decay.

$$\langle Z_1(T)Z_2(-T)\rangle_{\text{dip}}\sim\exp\left(N\log 2-\frac{JNT}{q2^q}\right) \tag{4.159}$$

For large values of rT, we get a different solution $\sigma=\mp1$ or

$$G_{\mathsf{LR}}=\pm\frac{i}{2},\quad \Sigma_{\mathsf{LR}}=\mp\frac{ri}{2^{q-2}} \tag{4.160}$$

For the usual spectral form factor, these solutions give us an action that is independent of T (for large T). However, $r<1$ these solutions predict an exponential decay:

$$\langle Z_1(T)Z_2(-T)\rangle_{\text{ramp}}\sim\exp\left[\frac{-NT(1-r)}{q2^q}\right] \tag{4.161}$$

So we get an exponentially decaying term. Note that this formula is valid when $rT\gg1$.

4.9.2 Journal disk

Now consider the problem we are actually interested in, the Renyi entropy of the journal. Let us start by finding the disk solution. Then we just need to consider $\langle t_rU_J(T)U_{J'}(T)\rangle$. The difference between this setup and (4.155) is the just the choice of boundary conditions. Instead of the second

line in (4.155) computing a trace, we should write

$$2^{N/2}\langle 0|\exp\left(\int H dt\right)|0\rangle = 2^{N/2}\left(e^{\pm\frac{i}{4}T\Sigma}\right)^N \tag{4.162}$$

Note the normalization $2^{N/2} = t_r 1$. We are left with an action

$$\exp\left\{N\left[\frac{\pm iT\Sigma}{4} - \frac{JT}{q2^q} + i^q r\frac{JT}{q}G^q - \frac{T}{2}\Sigma G\right]\right\} \tag{4.163}$$

Extremizing the action gives

$$G = \pm i/2, \quad \Sigma = i^q r J G^{q-1} = \pm rJ \tag{4.164}$$

The on-shell action is

$$Z \sim \exp\left[\frac{-JT(1-r)}{q2^q}\right] \tag{4.165}$$

We see that this disk solution will actually conflict with unitarity at late times T. Note that when $r = 0$, the top and bottom disk are independent. This is equivalent to the usual Brownian disk with $2T$. Indeed, from SSS we see that that the disk has an action

$$Z_{\text{disk}} \sim 2^{N/2}\exp\left\{-\frac{JTN}{2q2^q}\right\} \tag{4.166}$$

To be a bit more explicit, consider a two-point function

$$\begin{aligned} G(T',T) &= t_r\left(U(T')\psi_j U(T)\psi_j\right) = \langle 1|\, U(T')^L\psi_j^L U(T)^L\psi_j^L\,|1\rangle \\ &= -i\,\langle 1|\, U(T')^L\psi_j^L\psi_j^R U(T)^L\,|1\rangle \\ &= -iG_{LR}. \end{aligned} \tag{4.167}$$

The first line is true for any maximally entangled state $|1\rangle$. In the second line, we specialize to the choice $J_-|1\rangle = \left(\psi^L + i\psi^R\right)|1\rangle = 0$.

4.9.3 Wormhole

At long last, we find the wormhole. Actually for $r = 0$, we have already found the wormhole in the form of the Brownian "ramp" solution. More generally, we need to introduce 4 sides L_1, L_2, R_1, R_2

such that L_1 and L_2 are perfectly correlated but $L_1 R_1$ are correlated by an amount r.

We will have a quadratic fermion action

$$\int \mathcal{D}\psi_a^{(L)} \mathcal{D}\psi_a^{(R)} \exp\left\{ -\frac{1}{2} \int_0^T dt \left[\psi_a^{(A)} \partial_t \psi_a^{(A)} - \psi_a^{(L)}(t) \psi_a^{(R)}(t) \Sigma(t) \right] \right\} \tag{4.168}$$

It is possible to diagonalize the anti symmetric Σ matrix, so we could try to find constant solutions. To reduce the number of variables, we could try an ansatz

$$\begin{aligned} \Sigma_{L1R2} &= \Sigma_{R1L2}, \\ \Sigma_{L1L2} &= -\Sigma_{R1R2}, \\ \Sigma_{L1R1} &= \Sigma_{L2R2} \end{aligned} \tag{4.169}$$

This is motivated by the UV property $\psi_L |1\rangle = -i\psi_R |1\rangle, \psi_R |1\rangle = i\psi_L |1\rangle$.

At large times T, we expect that only the smallest eigenvalue will matter.

$$-S = N\frac{T}{2}\sqrt{s_{1L2R}^2 + s_{1L2L}^2 - s_{1L1R}^2} - \frac{2JT}{q2^q} + i^q \frac{2JT}{q} \left(r G_{1L2R}^q + r G_{1L1R}^q + G_{1L2L}^q \right) - T\Sigma_A G_A \tag{4.170}$$

We find two solutions of this action $G_{1L2L} = \pm 1/2$ which leads to an action that is independent of time. For this solution $|G_{1R2R}| = 1/2$ but all other correlators are zero. So it looks the wormhole is "made" of disks, just like the large q SYK wormhole.

Bibliography

[1] H. W. Lin, J. Maldacena, and Y. Zhao, "Symmetries Near the Horizon," *JHEP* **08** (2019) 049, arXiv:1904.12820 [hep-th].

[2] Y. Chen and H. W. Lin, "Signatures of global symmetry violation in relative entropies and replica wormholes," *JHEP* **03** (2021) 040, arXiv:2011.06005 [hep-th].

[3] A. Almheiri and H. W. Lin, "The Entanglement Wedge of Unknown Couplings," arXiv:2111.06298 [hep-th].

[4] A. Kitaev and S. J. Suh, "Statistical mechanics of a two-dimensional black hole," arXiv:1808.07032 [hep-th].

[5] T. G. Mertens, G. J. Turiaci, and H. L. Verlinde, "Solving the Schwarzian via the Conformal Bootstrap," *JHEP* **08** (2017) 136, arXiv:1705.08408 [hep-th].

[6] Z. Yang, "The Quantum Gravity Dynamics of Near Extremal Black Holes," arXiv:1809.08647 [hep-th].

[7] P. Saad, S. H. Shenker, and D. Stanford, "A semiclassical ramp in SYK and in gravity," arXiv:1806.06840 [hep-th].

[8] P. Saad, S. H. Shenker, and D. Stanford, "JT gravity as a matrix integral," arXiv:1903.11115 [hep-th].

[9] P. Saad, "Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity," arXiv:1910.10311 [hep-th].

[10] P. Kraus, H. Ooguri, and S. Shenker, "Inside the horizon with AdS / CFT," *Phys. Rev. D* **67** (2003) 124022, arXiv:hep-th/0212277.

[11] L. Fidkowski, V. Hubeny, M. Kleban, and S. Shenker, "The Black hole singularity in AdS / CFT," *JHEP* **02** (2004) 014, arXiv:hep-th/0306170.

[12] M. Grinberg and J. Maldacena, "Proper time to the black hole singularity from thermal one-point functions," *JHEP* **03** (2021) 131, arXiv:2011.01004 [hep-th].

[13] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, "Replica wormholes and the black hole interior," arXiv:1911.11977 [hep-th].

[14] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, "Replica Wormholes and the Entropy of Hawking Radiation," arXiv:1911.12333 [hep-th].

[15] A. Almheiri and J. Polchinski, "Models of AdS_2 backreaction and holography," *JHEP* **11** (2015) 014, arXiv:1402.6334 [hep-th].

[16] K. Jensen, "Chaos and hydrodynamics near AdS_2," arXiv:1605.06098 [hep-th].

[17] J. Maldacena, D. Stanford, and Z. Yang, "Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space," *PTEP* **2016** no. 12, (2016) 12C104, arXiv:1606.01857 [hep-th].

[18] J. Engelsy, T. G. Mertens, and H. Verlinde, "An investigation of AdS_2 backreaction and holography," *JHEP* **07** (2016) 139, arXiv:1606.03438 [hep-th].

[19] S. Sachdev and J.-w. Ye, "Gapless spin fluid ground state in a random, quantum Heisenberg magnet," *Phys. Rev. Lett.* **70** (1993) 3339, arXiv:cond-mat/9212030 [cond-mat].

[20] A. Kitaev, "A simple model of quantum holography." http://online.kitp.ucsb.edu/online/entangled15/kitaev/,http://online.kitp.ucsb.edu/online/entangled15/kitaev2/. Talks at KITP, April 7, 2015 and May 27, 2015.

[21] A. Kitaev and S. J. Suh, "The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual," *JHEP* **05** (2018) 183, arXiv:1711.08467 [hep-th].

[22] J. Maldacena and D. Stanford, "Comments on the Sachdev-Ye-Kitaev model," arXiv:1604.07818 [hep-th].

[23] D. Bagrets, A. Altland, and A. Kamenev, "SachdevYeKitaev model as Liouville quantum mechanics," *Nucl. Phys.* **B911** (2016) 191–205, arXiv:1607.00694 [cond-mat.str-el].

[24] D. Bagrets, A. Altland, and A. Kamenev, "Power-law out of time order correlation functions in the SYK model," *Nucl. Phys.* **B921** (2017) 727–752, arXiv:1702.08902 [cond-mat.str-el].

[25] D. Stanford and E. Witten, "Fermionic Localization of the Schwarzian Theory," *JHEP* **10** (2017) 008, arXiv:1703.04612 [hep-th].

[26] P. Gao, D. L. Jafferis, and A. Wall, "Traversable Wormholes via a Double Trace Deformation," *JHEP* **12** (2017) 151, arXiv:1608.05687 [hep-th].

[27] J. Maldacena and X.-L. Qi, "Eternal traversable wormhole," arXiv:1804.00491 [hep-th].

[28] L. Susskind, "Why do Things Fall?," arXiv:1802.01198 [hep-th].

[29] A. R. Brown, H. Gharibyan, A. Streicher, L. Susskind, L. Thorlacius, and Y. Zhao, "Falling Toward Charged Black Holes," *Phys. Rev.* **D98** no. 12, (2018) 126016, arXiv:1804.04156 [hep-th].

[30] X.-L. Qi and A. Streicher, "Quantum Epidemiology: Operator Growth, Thermal Effects, and SYK," arXiv:1810.11958 [hep-th].

[31] D. A. Roberts, D. Stanford, and A. Streicher, "Operator growth in the SYK model," *JHEP* **06** (2018) 122, arXiv:1802.02633 [hep-th].

[32] R. Jackiw, "Lower Dimensional Gravity," *Nucl. Phys.* **B252** (1985) 343–356.

[33] C. Teitelboim, "Gravitation and Hamiltonian Structure in Two Space-Time Dimensions," *Phys. Lett.* **B126** (1983) 41–45.

[34] M. Henneaux, "Quantum Gravity in Two Dimensions: Exact Solution of the Jackiw Model," *Phys. Rev. Lett.* **54** (1985) 959–962.

[35] D. Louis-Martinez, J. Gegenberg, and G. Kunstatter, "Exact Dirac quantization of all 2-D dilaton gravity theories," *Phys. Lett.* **B321** (1994) 193–198, arXiv:gr-qc/9309018 [gr-qc].

[36] Y. Sekino and L. Susskind, "Fast Scramblers," *JHEP* **10** (2008) 065, arXiv:0808.2096 [hep-th].

[37] K. V. Kuchar, "Geometrodynamics of Schwarzschild black holes," *Phys. Rev.* **D50** (1994) 3961–3981, arXiv:gr-qc/9403003 [gr-qc].

[38] D. Harlow and D. Jafferis, "The Factorization Problem in Jackiw-Teitelboim Gravity," arXiv:1804.01081 [hep-th].

[39] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," *JHEP* **03** (2014) 067, arXiv:1306.0622 [hep-th].

[40] A. Kitaev, "Hidden correlations in the hawking radiation and thermal noise," 2014. Talk given at the Fundamental Physics Prize Symposium, Nov. 10, 2014.

[41] A. Kitaev. Talk given at the Fundamental Physics Prize Symposium, Nov. 10, 2014.

[42] J. Maldacena, D. Stanford, and Z. Yang, "Diving into traversable wormholes," *Fortsch. Phys.* **65** no. 5, (2017) 1700034, arXiv:1704.05333 [hep-th].

[43] A. Blommaert, T. G. Mertens, and H. Verschelde, "Clocks and Rods in Jackiw-Teitelboim Quantum Gravity," arXiv:1902.11194 [hep-th].

[44] S. R. Coleman and F. De Luccia, "Gravitational Effects on and of Vacuum Decay," *Phys. Rev.* **D21** (1980) 3305.

[45] M. Dafermos and J. Luk, "The interior of dynamical vacuum black holes I: The C^0-stability of the Kerr Cauchy horizon," arXiv:1710.01722 [gr-qc].

[46] J. Maldacena, "Vacuum decay into Anti de Sitter space," arXiv:1012.0274 [hep-th].

[47] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, "Local bulk operators in AdS/CFT: A Boundary view of horizons and locality," *Phys. Rev.* **D73** (2006) 086003, arXiv:hep-th/0506118 [hep-th].

[48] I. Heemskerk, D. Marolf, J. Polchinski, and J. Sully, "Bulk and Transhorizon Measurements in AdS/CFT," *JHEP* **10** (2012) 165, arXiv:1201.3664 [hep-th].

[49] D. Kabat, G. Lifschytz, and D. A. Lowe, "Constructing local bulk observables in interacting AdS/CFT," *Phys. Rev.* **D83** (2011) 106009, arXiv:1102.2910 [hep-th].

[50] D. L. Jafferis, "Bulk reconstruction and the Hartle-Hawking wavefunction," arXiv:1703.01519 [hep-th].

[51] I. Heemskerk, "Construction of Bulk Fields with Gauge Redundancy," *JHEP* **09** (2012) 106, arXiv:1201.3666 [hep-th].

[52] D. Kabat and G. Lifschytz, "Decoding the hologram: Scalar fields interacting with gravity," *Phys. Rev.* **D89** no. 6, (2014) 066010, arXiv:1311.3020 [hep-th].

[53] A. Streicher. Private Communication.

[54] D. J. Gross and V. Rosenhaus, "The Bulk Dual of SYK: Cubic Couplings," *JHEP* **05** (2017) 092, arXiv:1702.08016 [hep-th].

[55] L. Susskind, "Complexity and Newton's Laws," *Front. in Phys.* **8** (2020) 262, arXiv:1904.12819 [hep-th].

[56] D. Stanford and L. Susskind, "Complexity and Shock Wave Geometries," *Phys. Rev.* **D90** no. 12, (2014) 126007, arXiv:1406.2678 [hep-th].

[57] H. Casini, M. Huerta, and R. C. Myers, "Towards a derivation of holographic entanglement entropy," *JHEP* **05** (2011) 036, arXiv:1102.0440 [hep-th].

[58] W. Donnelly, "Entanglement entropy and nonabelian gauge symmetry," *Class. Quant. Grav.* **31** no. 21, (2014) 214003, arXiv:1406.7304 [hep-th].

[59] L. Susskind, "Computational Complexity and Black Hole Horizons," *Fortsch. Phys.* **64** (2016) 24–43, arXiv:1402.5674 [hep-th].

[60] L. Susskind, "Addendum to Computational Complexity and Black Hole Horizons," arXiv:1403.5695 [hep-th].

[61] A. R. Brown, H. Gharibyan, H. W. Lin, L. Susskind, L. Thorlacius, and Y. Zhao, "Complexity of Jackiw-Teitelboim gravity," *Phys. Rev.* **D99** no. 4, (2019) 046016, arXiv:1810.08741 [hep-th].

[62] L. Susskind, "Dear Qubitzers, GR=QM," arXiv:1708.03040 [hep-th].

[63] D. Harlow, "Wormholes, Emergent Gauge Fields, and the Weak Gravity Conjecture," *JHEP* **01** (2016) 122, arXiv:1510.07911 [hep-th].

[64] S. B. Giddings and A. Strominger, "Loss of Incoherence and Determination of Coupling Constants in Quantum Gravity," *Nucl. Phys. B* **307** (1988) 854–866.

[65] R. Kallosh, A. D. Linde, D. A. Linde, and L. Susskind, "Gravity and global symmetries," *Phys. Rev. D* **52** (1995) 912–935, arXiv:hep-th/9502069.

[66] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, "The String landscape, black holes and gravity as the weakest force," *JHEP* **06** (2007) 060, arXiv:hep-th/0601001 [hep-th].

[67] T. Banks and N. Seiberg, "Symmetries and strings in field theory and gravity," *Physical Review D* **83** no. 8, (2011) 084019.

[68] D. Harlow and H. Ooguri, "Symmetries in quantum field theory and quantum gravity," arXiv:1810.05338 [hep-th].

[69] D. Harlow and H. Ooguri, "Constraints on Symmetries from Holography," *Phys. Rev. Lett.* **122** no. 19, (2019) 191601, arXiv:1810.05337 [hep-th].

[70] S. W. Hawking, "Particle Creation by Black Holes," *Commun. Math. Phys.* **43** (1975) 199–220. [,167(1975)].

[71] L. Susskind, "Trouble for remnants," arXiv:hep-th/9501106.

[72] D. N. Page, "Information in black hole radiation," *Phys. Rev. Lett.* **71** (1993) 3743–3746, arXiv:hep-th/9306083 [hep-th].

[73] G. Penington, "Entanglement Wedge Reconstruction and the Information Paradox," *JHEP* **09** (2020) 002, arXiv:1905.08255 [hep-th].

[74] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, "The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole," *JHEP* **12** (2019) 063, arXiv:1905.08762 [hep-th].

[75] S. Ryu and T. Takayanagi, "Holographic derivation of entanglement entropy from AdS/CFT," *Phys. Rev. Lett.* **96** (2006) 181602, arXiv:hep-th/0603001 [hep-th].

[76] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A Covariant holographic entanglement entropy proposal," *JHEP* **07** (2007) 062, arXiv:0705.0016 [hep-th].

[77] T. Faulkner, A. Lewkowycz, and J. Maldacena, "Quantum corrections to holographic entanglement entropy," *JHEP* **11** (2013) 074, arXiv:1307.2892 [hep-th].

[78] N. Engelhardt and A. C. Wall, "Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime," *JHEP* **01** (2015) 073, arXiv:1408.3203 [hep-th].

[79] A. Almheiri, R. Mahajan, J. Maldacena, and Y. Zhao, "The Page curve of Hawking radiation from semiclassical geometry," *JHEP* **03** (2020) 149, arXiv:1908.10996 [hep-th].

[80] D. L. Jafferis, A. Lewkowycz, J. Maldacena, and S. J. Suh, "Relative entropy equals bulk relative entropy," *JHEP* **06** (2016) 004, arXiv:1512.06431 [hep-th].

[81] A. Lewkowycz and J. Maldacena, "Generalized gravitational entropy," *JHEP* **08** (2013) 090, arXiv:1304.4926 [hep-th].

[82] L. Abbott and M. B. Wise, "Wormholes and Global Symmetries," *Nucl. Phys. B* **325** (1989) 687–704.

[83] P. Hayden and J. Preskill, "Black holes as mirrors: Quantum information in random subsystems," *JHEP* **09** (2007) 120, arXiv:0708.4025 [hep-th].

[84] D. Harlow and E. Shaghoulian, "Global symmetry, Euclidean gravity, and the black hole information problem," arXiv:2010.10539 [hep-th].

[85] P.-S. Hsin, L. V. Iliesiu, and Z. Yang, "A violation of global symmetries from replica wormholes and the fate of black hole remnants," arXiv:2011.09444 [hep-th].

[86] B. Czech, J. L. Karczmarek, F. Nogueira, and M. Van Raamsdonk, "The Gravity Dual of a Density Matrix," *Class. Quant. Grav.* **29** (2012) 155009, arXiv:1204.1330 [hep-th].

[87] A. C. Wall, "Maximin Surfaces, and the Strong Subadditivity of the Covariant Holographic Entanglement Entropy," *Class. Quant. Grav.* **31** no. 22, (2014) 225007, arXiv:1211.3494 [hep-th].

[88] M. Headrick, V. E. Hubeny, A. Lawrence, and M. Rangamani, "Causality & holographic entanglement entropy," *JHEP* **12** (2014) 162, arXiv:1408.6300 [hep-th].

[89] A. Almheiri, R. Mahajan, and J. Maldacena, "Islands outside the horizon," arXiv:1910.11077 [hep-th].

[90] R. Bousso and M. Tomašević, "Unitarity From a Smooth Horizon?," *Phys. Rev. D* **102** no. 10, (2020) 106019, arXiv:1911.06305 [hep-th].

[91] A. Almheiri, R. Mahajan, and J. E. Santos, "Entanglement islands in higher dimensions," *SciPost Phys.* **9** no. 1, (2020) 001, arXiv:1911.09666 [hep-th].

[92] C. Akers, N. Engelhardt, and D. Harlow, "Simple holographic models of black hole evaporation," arXiv:1910.00972 [hep-th].

[93] M. Rozali, J. Sully, M. Van Raamsdonk, C. Waddell, and D. Wakeham, "Information radiation in BCFT models of black holes," arXiv:1910.12836 [hep-th].

[94] F. F. Gautason, L. Schneiderbauer, W. Sybesma, and L. Thorlacius, "Page Curve for an Evaporating Black Hole," *JHEP* **05** (2020) 091, arXiv:2004.00598 [hep-th].

[95] T. Anegawa and N. Iizuka, "Notes on islands in asymptotically flat 2d dilaton black holes," *JHEP* **07** (2020) 036, arXiv:2004.01601 [hep-th].

[96] T. Hartman, E. Shaghoulian, and A. Strominger, "Islands in Asymptotically Flat 2D Gravity," *JHEP* **07** (2020) 022, arXiv:2004.13857 [hep-th].

[97] V. Balasubramanian, A. Kar, O. Parrikar, G. Sárosi, and T. Ugajin, "Geometric secret sharing in a model of Hawking radiation," arXiv:2003.05448 [hep-th].

[98] K. Hashimoto, N. Iizuka, and Y. Matsuo, "Islands in Schwarzschild black holes," *JHEP* **06** (2020) 085, arXiv:2004.05863 [hep-th].

[99] H. Geng and A. Karch, "Massive islands," *JHEP* **09** (2020) 121, arXiv:2006.02438 [hep-th].

[100] H. Z. Chen, R. C. Myers, D. Neuenfeld, I. A. Reyes, and J. Sandor, "Quantum Extremal Islands Made Easy, Part I: Entanglement on the Brane," *JHEP* **10** (2020) 166, arXiv:2006.04851 [hep-th].

[101] Y. Chen, V. Gorbenko, and J. Maldacena, "Bra-ket wormholes in gravitationally prepared states," arXiv:2007.16091 [hep-th].

[102] T. Hartman, Y. Jiang, and E. Shaghoulian, "Islands in cosmology," arXiv:2008.01022 [hep-th].

[103] V. Balasubramanian, A. Kar, and T. Ugajin, "Islands in de Sitter space," arXiv:2008.05275 [hep-th].

[104] W. Sybesma, "Pure de Sitter space and the island moving back in time," arXiv:2008.07994 [hep-th].

[105] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, "The entropy of Hawking radiation," *Rev. Mod. Phys.* **93** no. 3, (2021) 035002, arXiv:2006.06872 [hep-th].

[106] Y. Chen, "Pulling Out the Island with Modular Flow," *JHEP* **03** (2020) 033, arXiv:1912.02210 [hep-th].

[107] H. Casini and M. Huerta, "Reduced density matrix and internal dynamics for multicomponent regions," *Class. Quant. Grav.* **26** (2009) 185005, arXiv:0903.5284 [hep-th].

[108] X. Dong and H. Wang, "Enhanced corrections near holographic entanglement transitions: a chaotic case study," arXiv:2006.10051 [hep-th].

[109] D. Marolf, S. Wang, and Z. Wang, "Probing phase transitions of holographic entanglement entropy with fixed area states," arXiv:2006.10089 [hep-th].

[110] C. Akers and G. Penington, "Leading order corrections to the quantum extremal surface prescription," arXiv:2008.03319 [hep-th].

[111] N. Lashkari, "Relative Entropies in Conformal Field Theory," *Phys. Rev. Lett.* **113** (2014) 051602, arXiv:1404.3216 [hep-th].

[112] D. Stanford, "More quantum noise from wormholes," arXiv:2008.08570 [hep-th].

[113] J. M. Maldacena and L. Maoz, "Wormholes in AdS," *JHEP* **02** (2004) 053, arXiv:hep-th/0401024.

[114] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, "Black Holes and Random Matrices," *JHEP* **05** (2017) 118, arXiv:1611.04650 [hep-th]. [Erratum: JHEP 09, 002 (2018)].

[115] J. Pollack, M. Rozali, J. Sully, and D. Wakeham, "Eigenstate Thermalization and Disorder Averaging in Gravity," *Phys. Rev. Lett.* **125** no. 2, (2020) 021601, arXiv:2002.02971 [hep-th].

[116] D. Marolf and H. Maxfield, "Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information," *JHEP* **08** (2020) 044, arXiv:2002.08950 [hep-th].

[117] S. B. Giddings and G. J. Turiaci, "Wormhole calculus, replicas, and entropies," *JHEP* **09** (2020) 194, arXiv:2004.02900 [hep-th].

[118] J. McNamara and C. Vafa, "Baby Universes, Holography, and the Swampland," arXiv:2004.06738 [hep-th].

[119] R. Bousso and E. Wildenhain, "Gravity/ensemble duality," *Phys. Rev. D* **102** no. 6, (2020) 066005, arXiv:2006.16289 [hep-th].

[120] S. R. Coleman, "Black Holes as Red Herrings: Topological Fluctuations and the Loss of Quantum Coherence," *Nucl. Phys. B* **307** (1988) 867–882.

[121] S. B. Giddings and A. Strominger, "Baby Universes, Third Quantization and the Cosmological Constant," *Nucl. Phys. B* **321** (1989) 481–508.

[122] J. Polchinski, "Monopoles, duality, and string theory," *Int. J. Mod. Phys. A* **19S1** (2004) 145–156, arXiv:hep-th/0304042.

[123] R. E. Arias, H. Casini, M. Huerta, and D. Pontello, "Entropy and modular Hamiltonian for a free chiral scalar in two intervals," *Phys. Rev.* **D98** no. 12, (2018) 125008, arXiv:1809.00026 [hep-th].

[124] J. Polchinski, "Chaos in the black hole S-matrix," arXiv:1505.08108 [hep-th].

[125] P. Saad, S. H. Shenker, D. Stanford, and S. Yao, "Wormholes without averaging," arXiv:2103.16754 [hep-th].

[126] X.-L. Qi, Z. Shangnan, and Z. Yang, "Holevo Information and Ensemble Theory of Gravity," arXiv:2111.05355 [hep-th].

[127] R. Renner and J. Wang, "The black hole information puzzle and the quantum de Finetti theorem," arXiv:2110.14653 [hep-th].

[128] X. Dong, D. Harlow, and A. C. Wall, "Reconstruction of Bulk Operators within the Entanglement Wedge in Gauge-Gravity Duality," *Phys. Rev. Lett.* **117** no. 2, (2016) 021601, arXiv:1601.05416 [hep-th].

[129] X.-L. Qi, "Entanglement island, miracle operators and the firewall," arXiv:2105.06579 [hep-th].

[130] J. L. Cardy, "Boundary conformal field theory," arXiv:hep-th/0411189.

[131] H. T. Lam, T. G. Mertens, G. J. Turiaci, and H. Verlinde, "Shockwave S-matrix from Schwarzian Quantum Mechanics," *JHEP* **11** (2018) 182, arXiv:1804.09834 [hep-th].

[132] D. Stanford, "More quantum noise from wormholes," arXiv:2008.08570 [hep-th].

[133] A. Streicher, "SYK Correlators for All Energies," *JHEP* **02** (2020) 048, arXiv:1911.10171 [hep-th].

[134] A. Eberlein, V. Kasper, S. Sachdev, and J. Steinberg, "Quantum quench of the Sachdev-Ye-Kitaev Model," *Phys. Rev. B* **96** no. 20, (2017) 205123, arXiv:1706.07803 [cond-mat.str-el].

[135] A. Goel, H. T. Lam, G. J. Turiaci, and H. Verlinde, "Expanding the Black Hole Interior: Partially Entangled Thermal States in SYK," *JHEP* **02** (2019) 156, arXiv:1807.03916 [hep-th].

[136] G. Penington, S. H. Shenker, D. Stanford, and Z. Yang, "Replica wormholes and the black hole interior," arXiv:1911.11977 [hep-th].

[137] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black Holes: Complementarity or Firewalls?," *JHEP* **02** (2013) 062, arXiv:1207.3123 [hep-th].

[138] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, "An Apologia for Firewalls," *JHEP* **09** (2013) 018, arXiv:1304.6483 [hep-th].

[139] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, "Replica Wormholes and the Entropy of Hawking Radiation," *JHEP* **05** (2020) 013, arXiv:1911.12333 [hep-th].

[140] J. M. Maldacena, "Eternal black holes in anti-de Sitter," *JHEP* **04** (2003) 021, arXiv:hep-th/0106112 [hep-th].

[141] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," arXiv:1503.01409 [hep-th].

[142] R. Penrose, "Gravitational collapse: The role of general relativity," *Nuovo Cimento Rivista Serie* **1** (1969) 252.

[143] D. Bak, M. Gutperle, and S. Hirano, "A Dilatonic deformation of AdS(5) and its field theory dual," *JHEP* **05** (2003) 072, arXiv:hep-th/0304129.

[144] 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** (2005) 066003, arXiv:hep-th/0407073.

[145] A. Clark and A. Karch, "Super Janus," *JHEP* **10** (2005) 094, arXiv:hep-th/0506265.

[146] D. Gaiotto and E. Witten, "Janus Configurations, Chern-Simons Couplings, And The theta-Angle in N=4 Super Yang-Mills Theory," *JHEP* **06** (2010) 097, arXiv:0804.2907 [hep-th].

[147] D. Bak, M. Gutperle, and S. Hirano, "Three dimensional Janus and time-dependent black holes," *JHEP* **02** (2007) 068, arXiv:hep-th/0701108.

[148] C. Bachas and V. Papadopoulos, "Phases of Holographic Interfaces," *JHEP* **04** (2021) 262, arXiv:2101.12529 [hep-th].

[149] P. Simidzija and M. Van Raamsdonk, "Holo-ween," *JHEP* **12** (2020) 028, arXiv:2006.13943 [hep-th].

[150] A. May, P. Simidzija, and M. Van Raamsdonk, "Negative energy enhancement in layered holographic conformal field theories," arXiv:2103.14046 [hep-th].

[151] M. Oshikawa and I. Affleck, "Boundary conformal field theory approach to the critical two-dimensional Ising model with a defect line," *Nucl. Phys. B* **495** (1997) 533–582, arXiv:cond-mat/9612187.