Breakdown of locality in quantum gravity

Jefferson, R.A.

Breakdown of Locality in Quantum Gravity

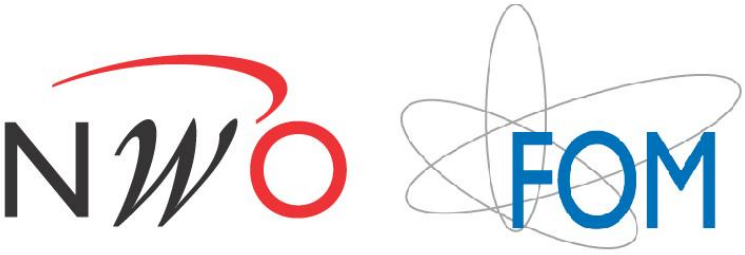
R. A. Jefferson

Breakdown of Locality in Quantum Gravity

R. A. Jefferson

Breakdown of Locality in Quantum Gravity

This work has been conducted as part of the Δ-ITP consortium through the Institute for Theoretical Physics (ITFA) and the center for GRavitation and AstroParticle Physics Amsterdam (GRAPPA) of the University of Amsterdam (UvA), and supported in part by the Foundation for Fundamental Research on Matter (FOM). Both are programs under the Netherlands Organization for Scientific Research (NWO) funded by the Dutch Ministry of Education, Culture, and Science (OCW).



Breakdown of Locality in Quantum Gravity

Academisch Proefschrift

ter verkrijging van de graad van doctor

aan de Universiteit van Amsterdam

op gezag van de Rector Magnificus

prof. dr. ir. K. I. J. Maex

ten overstaan van een door het college voor promoties

ingestelde commissie,

in het openbaar te verdedigen in de Agnietenkapel

op donderdag 7 september 2017, te 12:00 uur

door

Robert Alan Jefferson

geboren te Edmonton, Canada

Promotiecommissie

Promotor

prof. dr. E. P. Verlinde	Universiteit van Amsterdam

Co-Promotor

dr. B. Freivogel	Universitiet van Amsterdam

Overige leden

prof. dr. D. Baumann	Universitiet van Amsterdam
dr. A. Castro	Universitiet van Amsterdam
prof. dr. J. de Boer	Universitiet van Amsterdam
prof. dr. D. Kabat	City University of New York
prof. dr. K. Schalm	Universiteit Leiden

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Publications

This thesis is based on the following publications:

[1] B. Freivogel, R. A. Jefferson, L. Kabir, and I-S. Yang
"Geometry of the Infalling Causal Patch"
Phys. Rev. **D91**, 4 (2015), arXiv:1406.6043 [hep-th].

[2] B. Freivogel, R. A. Jefferson, L. Kabir, B. Mosk, and I-S. Yang
"Casting Shadows on Holographic Reconstruction"
Phys. Rev. **D91**, 8 (2015), arXiv:1203.1036 [hep-th].

[3] B. Freivogel, R. A. Jefferson, and L. Kabir
"Precursors, Gauge Invariance, and Quantum Error Correction in AdS/CFT"
JHEP **04**, 119 (2016), arXiv:1602.04811 [hep-th].

[4] "A. Belin, B. Freivogel, R. A. Jefferson, and L. Kabir
"Sub-AdS Scale Locality in AdS_3/CFT_2"
arXiv:1611.08601. [hep-th].

Preface

This thesis is an exposition of some of the work I conducted as part of my PhD. In broad terms, the subject of interest is quantum gravity—specifically, as alluded in the somewhat provocative title, the indications we've uncovered that at least some degree of nonlocality will be fundamental to the eventual theory thereof. Precisely what this means is context-dependent, and we shall be deliberately vague in defining it as a reflection of the incomplete and multifaceted picture that has emerged in the past few years. In general however, the basic idea is that local physics is encoded in some equivalent, nonlocal description.

The objective of this thesis is to investigate various aspects of this proposal as manifested in two interrelated areas, namely *Black Hole Thermodynamics* and the *Holographic Principle*. We shall begin in chapter 1 with an introduction to the former, starting with the famous discovery of Bekenstein and Hawking that black holes are not quite black after all, but radiate with an approximately thermal spectrum. This leads to a conflict between general relativity and quantum field theory, since an evaporating black hole appears to imply a loss of information, which the latter forbids. In its modern form, this contradictory state of affairs is referred to as the *Firewall Paradox*.

There have been many proposed resolutions, with various consequences for local effective field theory, and chapter 3 discusses my own contribution to the subject. Specifically, a careful analysis of the region which is causally accessible to an observer who falls into the black hole in an attempt to uncover the paradox reveals that she will almost never succeed in doing so. This leaves open the possibility that an earlier resolution by the name of *Black Hole Complementarity* – with its own curious implications – may yet survive the recent assault.

However, perhaps the biggest insight from the physics of black holes in this

context is the discovery of the *Holographic Principle*, particularly its precise realization in the *AdS/CFT Correspondence*. This is a remarkable equivalence between a gravitational theory in $(d+1)$-dimensional spacetime and a theory without gravity on the d-dimensional boundary thereof. We shall give a criminally brief introduction to this subject in chapter 2, first introducing the geometry of anti-de Sitter space (the "bulk"), then the most relevant features of conformal field theory (the "boundary"), and finally the salient aspects of the correspondence itself—in particular the main entries in the so-called "dictionary" that relates quantities on either side of the duality.

In the final three chapters, we shall put this framework to use. We begin in chapter 4 with a consideration of the problem of *bulk reconstruction*: the task of reconstructing local operators in the bulk from CFT data. In particular, we investigate a failure of presently understood bulk probes in non-vacuum geometries. This is due to the existence a region called the *holographic shadow* that appears beyond the reach of known elements of the aforementioned dictionary. We submit that the information in these regions is encoded nonlocally, and making this intuition precise is the subject of ongoing work.

We discuss one approach in this direction in chapter 5, wherein we investigate the degree to which information about a given region of the bulk can be localized on the boundary. Our analysis centers on objects called *precursors*, which are intrinsically nonlocal boundary operators with seemingly prescient knowledge of phenomena in the bulk. Specifically, we shall show that both gauge freedom and the entanglement structure of the field theory can be used to localize precursors to within a particular subregion of the boundary. In the course of doing so, we shall introduce the idea of *quantum error correction*, the use of which alludes to the growing ties between high energy physics and quantum information theory in the nascent endeavor to understand emergent spacetime in the holographic context.

In chapter 6, we turn to the question of locality on sub-AdS scales. This is a particularly subtle issue, since it is not clear whether a sensible semiclassical description extends below this scale. Indeed, this is itself an extension of the question of which CFTs have sensible bulk duals. Certain criteria on the CFT have been proposed in order to satisfy this requirement, and we shall consider this question with the help of an explicit toy model for $\mathrm{AdS}_3/\mathrm{CFT}_2$.

We close with a summary (samenvatting), which discusses this research within the broader context of emergent spacetime. This is an exciting idea that draws on, and has implications for, many different areas of physics, and understanding the extent to which quantum gravity is nonlocal will be a vital component in making this precise. We hope to contribute further to this endeavor in the future.

Notation and conventions

Throughout this thesis, we shall work in civilized units, $\hbar = c = G_N = k_B = 1$, except where explicitly including these constants facilitates the point at hand.

The Lorentzian signature is always taken to be $(-, +, \ldots, +)$, in which the space-time dimension is denoted $D = d + 1$, while d is reserved for the purely spatial component.

Contents

1 Black holes and revelations

In this chapter, we give a brief introduction to the quantum physics of black holes. This is a fascinating subject spanning over four decades of inquiry, and remains an active area of research today. Our exposition will be broadly historical, and focused on those aspects deemed most relevant for the subsequent chapters—namely, the black hole information paradox, and the germ of holography discovered herein.

1.1 Black hole thermodynamics

In 1973, Jacob Bekenstein observed [5] that black holes must be endowed with an entropy in order to preserve the second law of thermodynamics; otherwise, one could decrease the entropy of the universe by simply throwing subsystems with high entropy (e.g., a hot cup of coffee, this thesis) into a black hole. At face value, this is an intuitive proposal: since the information about the degrees of freedom that comprise the hypothetical subsystem would then be hidden behind the event horizon, it makes sense to count them among the microstates of the black hole.

The unintuitive twist (the first of many!) comes from the realization that this naïve bookkeeping is not at all how black holes operate. The entropy of familiar systems scales with the volume thereof, $S \sim V$, which is consistent with simply counting the obvious (particulate) degrees of freedom in the examples above. Black hole entropy, in stark contrast, scales with the *area* of the event horizon, $S \sim A$. Bekenstein's original motivation for this proposal hinged largely on Hawking's 1971 result [6] that the surface area of a black hole cannot decrease in any classical process (the so-called "area theorem"). This lead Bekenstein to propose an analogy between black holes and statistical thermodynamics, which has since been enshrined in the laws of *black hole thermodynamics* for stationary black holes:

Zeroth Law: The surface gravity, κ, on the horizon is constant. This implies that surface gravity is analogous to temperature.

First Law: For a stationary Kerr-Newman black hole, the change in energy under small perturbations is given by

$$\mathrm{d}E = \frac{\kappa}{8\pi}\,\mathrm{d}A + \Omega\,\mathrm{d}J + \Phi\,\mathrm{d}Q\ . \tag{1.1}$$

This is the statement of energy conservation, where the right-hand side is equal to $T\,\mathrm{d}S$.

Second Law: Assuming the weak energy condition holds, the horizon area is non-decreasing,

$$\frac{\mathrm{d}A}{\mathrm{d}t} \geq 0\ . \tag{1.2}$$

This is the aforementioned area theorem, and corresponds (under the instigating observation of Bekenstein above) to the statement that the entropy never decreases.

Third Law: It is not possible to form a black hole with vanishing surface gravity,

$$\kappa > 0\ . \tag{1.3}$$

The third law of ordinary thermodynamics is essentially the statement that a system at absolute zero must be in the state with minimum possible energy. In the usual example of a perfect crystal, this is assumed to be comprised of a single eigenstate, hence the entropy vanishes. The corresponding example here is an extremal black hole, which has $\kappa = 0$.

A more detailed exposition of these laws can be found in, e.g., [7]. However, despite the apparent necessity of ascribing to black holes an entropy proportional to A, thus far black hole thermodynamics is little more than an analogy: classically, black holes do not radiate (hence the name), and therefore have zero temperature and consequently zero thermodynamic entropy. Indeed, Bekenstein's original proposal [5] explicitly views the entropy in an information-theoretic – as opposed to thermodynamic – sense, i.e., as the Shannon entropy measuring the inaccessibility of the internal microstates of a system. General relativity ensures that these degrees of freedom are forever isolated from the external universe, hence an external observer can never extract information, and thus the entropy of the black hole must be non-decreasing. It is worth emphasizing however that, at least at the classical level, this entropy is properly regarded as referring to the equivalence class of black holes with the same mass, charge, and angular momentum, rather than to the temperature of any single black hole.

The situation changed the following year, when Hawking showed [8] that, quantum mechanically, black holes *do* radiate, with temperature

$$T = \frac{\kappa}{2\pi} = \frac{1}{8\pi M} \ , \tag{1.4}$$

and entropy

$$S = \frac{A}{4\ell_{\rm P}^2} \ , \tag{1.5}$$

where we have explicitly included the Planck length, $\ell_{\rm P} = \sqrt{\hbar G/c^3}$, in the latter formula lest the reader be disturbed by the mismatch in dimensions between S and A. The existence of Hawking radiation implies that black holes can evaporate, and thus their surface area A can in fact decrease.[1] This requires a modification of the second law, to the effect that the total entropy of the black hole (still identified with its horizon area) *plus* the entropy of the Hawking radiation is non-decreasing. This is referred to as the *generalized second law*.

With Hawking's discovery that black holes are not completely black after all, black hole thermodynamics went from epistemic to ontic in one fell swoop. The precise nature of the Hawking radiation itself, however, remains muddled to this day.

The vast interpretational quagmire surrounding Hawking radiation is due in no small part to the fact that there are a multitude of seemingly distinct derivations thereof. Hawking's original 1975 calculation considers a black hole that forms from collapse. The mode expansion of a scalar field at past and future null infinity differ, on account of the difference in vacuum state—namely, the Minkowski and Schwarzschild vacua, respectively. One can express the latter in terms of the former by means of a Bogoliubov transformation, which results in a thermal expectation value for the outgoing modes.[2]

However, the collapsing geometry is in fact entirely incidental to the radiation, as Hawking himself observed [8]. Rather, it is the presence of the event horizon that is the key feature [12, 13]. Indeed, it is straightforward to show that an accelerating observer in Minkowski space observes a thermal spectrum associated with the Rindler horizon, which well-approximates the near-horizon region of a large Schwarzschild black hole [11].

[1] In other words, the aforementioned area law was a purely classical statement. Quantum mechanical effects render the weak energy condition – a key assumption – invalid [8].

[2] More technically, the initial Minkowski vacuum $|0_M\rangle$ corresponds to the Kruskal or Hartle-Hawking vacuum $|0_K\rangle$, while the final Schwarzschild vacuum $|0_S\rangle$ is analogous to Rindler space $|0_R\rangle$. While the Kruskal modes are defined on the entire manifold, a Rindler observer, who has access to only the exterior spacetime, will perceive a thermal vacuum corresponding to tracing out the degrees of freedom behind the horizon. This is the mechanism that underlies the *Unruh effect* [9, 10]. For more details on this and other aspects of black hole evaporation, both physical and mathematical, see [11].

The centrality of horizons in this context is elegantly demonstrated by the 1977 paper by Gibbons and Hawking [14], in which they compute the entropy of a black hole from what is essentially a purely geometrical[3] argument. The basic idea is to compute the path integral for the black hole by Wick rotating to Euclidean signature, in which the geometry pinches off smoothly at the horizon. This corresponds to the fixed point of the $U(1)$ symmetry, which we obtain by periodically identifying Euclidean time to avoid a conical deficit. The contribution from the fixed point dominates the path integral Z; and since Z is also the partition function, a simple thermodynamic argument allows one to derive an expression for the entropy in terms of the leading saddle-point, which yields precisely the above, well-known result (1.5).

1.2 The information paradox

The fact that black holes radiate has shattering implications, which Hawking was swift to point out in his subsequent work [15]. Suppose that we form a black hole by collapsing some matter distribution in an initially pure state. After the black hole has completely evaporated, we are left with radiation in a thermal state, which is by definition mixed. But the transformation from a pure state to a mixed state violates unitarity, a fundamental principle of quantum mechanics necessary to ensure conservation of probabilities. In other words, non-unitary evolution would imply that information is lost in the process, which quantum mechanics forbids. Thus it appears that the very quantum mechanical laws which give rise to Hawking radiation are violated as a result! This is the substance of the *black hole information paradox*. As we shall see, it provides perhaps the first hints that our conception of locality may require modification.

It is illuminating to contrast this situation with the apparently pure-to-thermal evolution of normal matter upon incineration, say a burning lump of coal [16]. Supposing this to be in an initially pure state, the final state again involves a thermal bath of radiation, with the apparent loss of information that implies. But we do not concern ourselves with unitarity-violating bbq's. The reason is that subtle correlations between early and late radiation conspire to preserve the purity of the total system. It is only in coarse-graining (or tracing out whatever fraction of coal remains at a given stage) that we perceive a thermal state. It may be impossible to actually recover this information in practice, but in principle, the laws of quantum mechanics survive intact—that is, a sufficiently powerful computer could do it.

The essential difference between the coal and the black hole is that the former

[3] In contrast to the usual jargon, here I mean "geometry" as distinct from "classical gravity", since the presence of $\hbar$ in the path integral technically places us beyond the domain of the latter.

has no horizon. Early "Hawking" modes are entangled with modes inside the coal, which can – via their interactions with other interior modes – imprint this information on the late radiation. In contrast, the presence of a horizon imposes a very specific entanglement structure on the modes that prevents those behind the horizon from transmitting the information in any obvious manner. This follows from the fact that the Minkowski vacuum is in some sense an infinitely entangled state: the correlation function between local field excitations at spacelike-separated points A and B will diverge as $A \to B$. We can make this more precise by considering the Rindler decomposition of the vacuum,

$$|0\rangle = \frac{1}{\sqrt{Z}} \sum_i e^{-\pi\omega_i} |i\rangle_L |i'\rangle_R \ , \tag{1.6}$$

where Z is the Euclidean path integral with no insertions, and the relation between the basis vectors for the left (L) and right (R) wedges is $|i\rangle_L = \Theta^\dagger |i'\rangle_R$, where Θ is the CPT operator.[4] Now consider decomposing a free scalar field into modes of definite boost energy ω $(-\omega)$ in the right (left) Rindler wedge. Then the vacuum state can be equivalently written as a product state over all modes [17]:

$$|0\rangle = \bigotimes_{\omega,k} \sqrt{1 - e^{-2\pi\omega}} \sum_n e^{-\pi\omega n} |n\rangle_{L\omega(-k)} |n\rangle_{R\omega k} \ . \tag{1.7}$$

This *pairwise entanglement* between modes across the horizon is ultimately what prevents the modes from sharing their entanglement as in the lump of coal.

On this point, an important clarification bears mentioning: while the pairwise entangled modes are a characteristic feature of horizons, the popular conception of Hawking radiation as pairwise entangled *particles* is misleadingly invalid, "a cartoon Hawking invented to explain his paper to children" [17]. The wavelength of the modes is of order M^{-1}, the size of the black hole, and thus the particle interpretation breaks down long before one reaches the horizon [11]. It is therefore meaningless to speak of the radiation as being localized in this manner.[5] The related question of where, precisely, the Hawking radiation originates has not been settled, though the evidence suggests that the adjective "precisely" may lose out to nonlocality as well [18].

Despite these difficulties, there have been several attempts to reconcile the apparent information loss by black holes by appealing to subtle correlations in the Hawking radiation. And indeed, in this regard it is worth emphasizing an oft-

[4] This is an antiunitary operator that exists in all QFTs, whose action on a scalar field Φ is $\Theta^\dagger \Phi(t, x, \mathbf{y})\Theta = \Phi^\dagger(-t, -x, \mathbf{y})$. See [17] for a nice exposition of how this enters the picture.

[5] Just to be clear, this of course does not imply that an infalling observer won't see particles as usual in her own reference frame, as per the equivalence principle. It is merely the blueshifting of Hawking modes back from infinity that is ill-defined; the associated divergence is simply the statement that, from the perspective of an external observer, time appears to stop at the horizon.

misstated point, namely that the radiation is *not* exactly thermal in the technical sense of the term—meaning, possessing a Planckian spectrum. Lower bounds on deviations from thermality can be derived from greybody factors, as well as from adiabatic and phase space constraints [19]. The appearance of (exact) thermality in certain calculations of the radiation spectrum (e.g., Hawking's original work [8]) stems from the fact that the Hartle-Hawking state presupposes that the black hole is in thermal equilibrium with the radiation, in which case one inevitably recovers a perfect black body spectrum. That said, the spectrum is thermal to a very good approximation [19], so we shall follow the conventional abuse of terminology and continue to use "thermal" in the colloquial sense, i.e., in reference to a highly mixed state with an approximately, rather than exactly, Planckian spectrum.

Even allowing for small deviations from exact thermality, it has been argued that subtle correlations in the Hawking radiation are insufficient to restore unitarity, and that these would instead have to constitute an $O(1)$ correction, which would destroy the very semiclassical physics they were intended to save [20, 21]. But the possibility of encoding information in such a manner has not been ruled out. In fact, arguments from holography – more specifically the AdS/CFT correspondence – indicate that unitarity is indeed preserved, and consequently the belief that the information is somehow encoded in the Hawking radiation is currently the most popular position [17].

An alternative proposal is that the evaporation process halts with a Planck-scale *remnant*, which contains all the information necessary to purify the radiation. However, whether remnants actually possess such an information capacity has been called into question [22]. Furthermore, even if the issue of unitarity could be resolved (or rather, sidestepped) in this manner, it would require an object on the order of 10^{-35} m to contain an (in principle) infinite number of internal states [23, 24]! This hardly seems a reasonable resolution, and remnants are generally disfavored for these and other reasons [25, 26]. That said, it is worth commenting that once the black hole approaches the Planck scale, semi-classical gravity breaks down, and a full theory of quantum gravity is needed to specify what happens in the final moments of a black hole's life.

A somewhat more fanciful possibility is to suppose that the black hole gives rise to another universe, such that unitarity is preserved in the total system (that is, the resulting multiverse). However, information would still be lost from the perspective of outside observers [27].[6] Additionally, there is ongoing debate as to whether evolution to a mixed state (or in this case, to a state defined on a non-Cauchy surface) violates conservation of energy [30, 29, 31].[7] In any case, this

[6]Certain models suggest that when making measurements on an ensemble, the loss of information to the baby universes is not observable; however, this does not appear to resolve the paradox when restricted (as we are) to a single parent universe [28, 29].

[7]Note that in this context we are considering the evolution of the entire system, as opposed

possibility would seem *qua definitione* beyond observable verification. And as we shall see below, holography provides stronger arguments against black holes acting as "information sinks" (in the manner of [31]), and thus we leave this option aside as well.

1.3 Black hole complementarity

All three of the proposed solutions (or rather, classes of solutions) above suffer drawbacks that, as of yet, have prevented a satisfactory resolution from emerging. However, in the early 90's, Susskind, Thorlacius, and Uglum [33] (see also [34, 35]) argued that that there is in fact no contradiction due to what they termed *black hole complementarity* (BHC). Building on earlier ideas by 't Hooft [36, 37], they proposed what is essentially a radical deviation from locality, whereby the same information is observed in different locations by complementary observers. The adjective here denotes the key restriction that these observers are unable to communicate; both measurements are then equally valid, since the contradiction between them could only be observed by transmitting and comparing.[8]

The postulates of BHC, as introduced in the original paper [33], are as follows:

1. **Unitarity**: Black hole formation and evaporation is described by a unitary S-matrix within the context of standard quantum field theory.

2. **EFT**: Physics outside the horizon is described to a good approximation by effective field theory.

3. **Thermodynamics**: To an external observer, the black hole appears to be a quantum system with discrete energy levels, and the dimension of the subspace is given by e^S.

4. **Equivalence principle**: A freely falling observer experiences "no drama", i.e., no substantial deviation from the predictions of general relativity, when crossing the horizon of a large black hole.[9]

Postulates 1 and 3 follow from the usual demands of quantum mechanics and black hole thermodynamics, respectively, as described above. Postulates 2 and 4

to subsystems, from pure to mixed. The latter is a benign and fundamental feature of quantum mechanics known as *decoherence* [32].

[8]This is not to say that the information is in two places simultaneously, since that would violate the no-cloning principle [38, 39]. Rather, "complementarity" refers to the fundamental feature of quantum mechanics whereby non-commuting observables cannot be simultaneously measured (the most famous example of which are the canonical position and momentum operators).

[9]This was not listed among the original three "postulates", but was explicitly introduced as an "assumption". It is unclear what the authors were attempting to accomplish by denying that these are synonyms.

essentially follow from the fact that the horizon of a large black hole is a region of low curvature, and (insofar as event horizons are global constructs) its presence is not revealed by any local invariant.[10] Indeed, the Earth could be falling through the event horizon of a sufficiently large black hole at this very moment; according to the equivalence principle, we'd be unable to tell. In other words, while new physics, specifically a theory of quantum gravity, is obviously needed for the Planck-scale region near the singularity, one fully expects that semi-classical physics remains valid on (large) horizon scales.

The upshot of BHC is that an observer who remains outside the black hole perceives a hot membrane at the horizon which radiates information, while an infalling observer encounters nothing out of the ordinary as she falls through. The former sees unitary evolution but cannot verify the apparent loss of the equivalence principle, while the situation for the latter is precisely reversed.[11]

It is instructive to ask what prevents the external observer from jumping into the black hole at some later time in order to compare her observations with those of the earlier infaller. If possible, this would violate the no-cloning principle and thereby render BHC invalid. However, as argued in [44], and subsequently refined in [45], the external observer must wait until after the Page time before she can collect any information. If she then attempts to receive an illegal quantum copy from the earlier infaller by subsequently diving into the hole, the message must be sent with more energy than the entire black hole itself contains—otherwise, she'll hit the singularity first. Thus it appears that a careful balance of factors conspires to keep the two frames of reference complementary in the above sense.

BHC is not as far-fetched as it initially sounds. Indeed, the idea that one should only endow observable quantities with ontic status is not only central to relativity, but a core tenet of science in general. Nonetheless, BHC does entail a significant departure from standard quantum mechanics with regards to the interpretation of the Hilbert space on a Cauchy slice that crosses into the interior of the black hole in such a way as to intersect both "copies" of the information. In particular, the question is whether a global Hilbert space can be meaningfully said to exist on these "nice slices."

If one posits a global Hilbert space, it must be the case that spacelike operators – specifically those in the interior and exterior – no longer commute. Otherwise, an observer whose causal past includes both regions would be able to measure

[10] However, there are large *nonlocal* invariants, in particular a large relative boost. In standard quantum field theory, only large local invariants can lead to a breakdown [13] (see also [40]). But highly boosted strings behave differently than point particles, and some recent work has investigated string scattering near the horizon as a means of probing the possible breakdown of locality in effective field theory [41, 42, 43].

[11] As the authors of [33] emphasized, BHC does not contradict the relativistic law that physics is the same in all reference frames, but merely asserts that the description of events in frames "separated by a large boost parameter" may differ.

them simultaneously. In this case, one preserves the usual formulation of quantum mechanics, except that *locality is broken* in such a manner as to make the same piece of information appear differently to different observers—specifically, observers who are complementary in the above sense. This is sometimes referred to as the *weak interpretation* of BHC, in contrast to the alternative below. As we shall see, this interpretation is morally in line with AdS/CFT, which also presumes quantum mechanics (i.e., the existence of a single, global Hilbert space) but is fundamentally nonlocal or *holographic* in nature [13].

Alternatively, one can deny the existence of such a global Hilbert space. In this so-called *strong interpretation* of BHC, the interior and exterior observers have their own separate Hilbert spaces, with some suitable matching conditions on the boundary (namely, the horizon). This preserves locality in the sense that spacelike observables commute as expected within each Hilbert space, but it is unclear whether it is possible to formulate a consistent set of matching conditions.[12] Additionally, as noted in [13], this interpretation still constitutes a "weakening" of local quantum field theory, since it makes the Hilbert space structure subordinate to the causal structure.[13]

1.4 Firewalls: the paradox reloaded

Until recently, BHC was generally the *de facto* (albeit perhaps not entirely satisfactory) solution to the information paradox. In 2012 however, Almheiri, Marolf, Polchinski, and Sully (AMPS) argued that the postulates of BHC are in fact mutually inconsistent [47] (see also [21, 48, 49, 50] for earlier work). This rekindled the information paradox with a vengeance, and the modern, as yet unresolved version is known as the *firewall paradox*.

The AMPS argument can be crudely summarized as follows (see figure 3.1): smoothness of the horizon – i.e., the equivalence principle – requires that a given Hawking mode H and its interior partner P be maximally entangled, as discussed above (more generally, the exterior mode is purified by its interior partner [51]), while purity of the final radiation – i.e., unitarity – requires that H be maximally entangled with the earlier radiation R. But this violates the monogamy of quantum entanglement, and thus it appears that at least one of the assumptions must be modified. AMPS chose the equivalence principle as the least egregious sacrifice. This would imply that an infalling observer indeed encounters the hot membrane perceived by her external collaborator—and is completely incinerated; hence the

[12] For example, insofar as horizons are global properties of the spacetime, the matching conditions would need to be defined nonlocally in time.

[13] This is the inverse of the standard formulation of QFT, wherein locality or "microcausality" is seen to emerge from quantum mechanics in conjunction with special relativity and the clustering property (i.e., factorization of the S-matrix) [46].

name "firewall".

The key difference between the AMPS argument and that from BHC is that the former uses the entangled Hawking modes to phrase the paradox, while the latter relied on recovering the information emitted after thermalization of some previously ingoing bit—which, as mentioned above, does not leave sufficient time for an external observer to detect a contradiction. In the case of Hawking modes however, the external observer can make a measurement of a single mode after the Page time, which must be entangled with the early radiation if the final state is to be pure. She can then immediately jump in and capture the entanglement between the mode and its interior partner, thus violating the monogamy of entanglement.

The aftermath of AMPS was considerable. As of this writing, less than 5 years after their paper's appearance, it has received nearly 700 citations. Accordingly, we shall not even attempt a full synopsis, and instead refer the reader to the following modern reviews and references therein: [17, 13]. However, we will note one aspect of the controversy with relevance to the main theme of this work: a crucial assumption of the AMPS scenario is that all the necessary ingredients can be made to fit within the causal past of a single observer. In chapter 3, we shall see that this is assumption is actually dubious at best. Thus it may be that BHC survives after all.

1.5 The holographic principle

Regardless of AMPS' alleged deposition of BHC, the latter paved the way for a new paradigm, which many believe resolves (or, more correctly, proves the existence of a resolution of) the firewall/information paradox: *the holographic principle*.

Suppose we take a spherical region of space with entropy S and surface area A, and proceed to add information to the region in the form of matter, energy, whatever. Eventually, we will have added so much mass that the region collapses to a black hole. And from black hole thermodynamics, we know that the entropy is proportional to the surface area, $S \sim A$. The key observation is that this represents the maximum density of information, the so-called *holographic bound*. Any attempt to store more information than allowed by this bound would violate the generalized second law, since this would require the entropy to decrease upon collapse.

As mentioned above, the entropy of a black hole cannot be thought of as arising from whatever matter distribution collapsed to form it; general relativity demands that the matter continue to collapse to a singularity, and hence its surface area shrinks to zero. Thus the holographic bound is consistent with our earlier statements that black hole entropy is intimately linked with the presence of the horizon. There are, of course, a number of technical subtleties (for example, the

naïve spatial bound cannot be applied to a homogeneous, infinite universe), for which we defer the interested reader to the literature: [27, 52, 53, 54, 55, 56].

As first emphasized by 't Hooft [57], the holographic bound implies that the dimensionality of the Hilbert space in a given region of space is an exponential function of the surface area, not the volume. This is extremely counter-intuitive. Consider a volume of gas: the degrees of freedom fill the region, and we'd expect the entropy to scale with the volume accordingly. But the lesson from black hole thermodynamics is that in quantum gravity, this is not so: the Hilbert space of any finite region will be exponentially smaller than we'd otherwise expect. Following 't Hooft's original suggestion [57], this idea has since been codified in the holographic principle [58, 54, 59]. And it provides perhaps the most dramatic evidence that *quantum gravity must be fundamentally nonlocal.*

Clearly however, our daily experience is adequately described by local physics, which must therefore emerge from the full theory of quantum gravity in the appropriate limit, much as Newton's laws emerge from Einstein's relativity in the limit $c \to \infty$. Indeed, a swiftly growing area of research is devoted to understanding how *spacetime emerges* in a holographic context [60]. We will have more to say about this later.

To fully explain the impact of holography on the firewall paradox in particular would require us to first introduce the machinery of AdS/CFT, which is the subject of the next chapter. Suffice to say that the holographic theory is unitary, and cannot lose information.[14] Thus it appears that effective field theory must be modified in such a way as to allow the information to leak out in the Hawking radiation. Indeed, as mentioned above, most researchers in the field have come to favour this option for precisely this reason, though there has yet to appear a satisfactory account for how this transpires.

[14] That is, an evaporating black hole in AdS is dual to a unitary process in the CFT.

2 Introduction to AdS/CFT

In the previous chapter, we reviewed how ideas from black hole thermodynamics led to *the holographic principle*. In its most general form, this is the statement that a gravitational theory in $(d+1)$-dimensional spacetime is dual to a non-gravitational theory on the d-dimensional boundary thereof. This notion has since found its most precise formulations in the context of string theory, culminating in Maldacena's celebrated AdS/CFT correspondence [61].

As of this writing, Maldacena's paper [61] has received over 15,000 citations. Thus we will not even attempt a representative survey, nor will we follow the initial lines of development, instead referring the reader to the many excellent resources on the subject, e.g., [62, 63, 64, 65, 66, 17]. Rather, we will simply present the basics of AdS/CFT in its modern form, with a strong focus on those aspects most relevant for the bulk of this work.

2.1 Anti-de Sitter spacetime

Anti-de Sitter (AdS) spacetime is the maximally symmetric[1] vacuum[2] solution to Einstein's equations with negative cosmological constant. Its Euclidean analogue is hyperbolic space, which may be visualized as an isometric embedding of a pseudosphere in one-higher dimensional flat space. Similarly, we may define AdS_{d+1} by an embedding of the Lorentzian analogue of the sphere in $(d+2)$-dimensional Minkowski space $\mathbb{R}^{d,2}$, whose metric is

$$\mathrm{d}s^2 = -\bar{\eta}_{MN}\,\mathrm{d}X^M\,\mathrm{d}X^N \ , \quad \bar{\eta}_{MN} = \mathrm{diag}\,(-,+,\ldots,+,-) \ . \tag{2.1}$$

[1] A spacetime is maximally symmetric if it admits the maximum number of linearly independent Killing vectors, which for an n-dimensional manifold is $n(n+1)/2$ [62].

[2] $T_{\mu\nu} = 0$.

AdS_{d+1} is then defined as the hypersurface

$$\bar{\eta}_{MN} X^M X^N = -\left(X^0\right)^2 + \sum_{i=1}^{d} \left(X^i\right)^2 - \left(X^{d+1}\right)^2 = -\ell_{\mathrm{AdS}}^2 \ , \tag{2.2}$$

where ℓ_{AdS} is the radius of curvature (a.k.a. AdS radius or AdS length-scale), which is related to the cosmological constant via

$$\Lambda = -\frac{d(d+1)}{2\ell_{\mathrm{AdS}}^2} \ , \tag{2.3}$$

i.e., Λ is identified as the scalar curvature $R = R^{\mu}{}_{\mu}$, where $R_{\mu\nu}$ is the Ricci curvature tensor. The latter is, of course, defined by the contraction of the first and third indices of the Riemann curvature tensor, which in the present case may be written

$$R_{\mu\nu\rho\sigma} = -\frac{1}{\ell_{\mathrm{AdS}}^2} \left(g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho}\right) \ . \tag{2.4}$$

Note that the hypersurface (2.2) is invariant under $O(d,2)$ transformations on $\mathbb{R}^{d,2}$. The definition of AdS via this embedding thus has the nice property that it makes the symmetry group manifest. In fact, though it is less obvious from this presentation, the isometry group of AdS_{d+1} is $SO(d,2)$ [62].

There are several common choices of coordinate patches used in discussing AdS. *Global coordinates* (ρ, τ, Ω_i) are defined via

$$\begin{aligned} X^0 &= \ell_{\mathrm{AdS}} \cosh\rho \cos\tau \ , \\ X^{d+1} &= \ell_{\mathrm{AdS}} \cosh\rho \sin\tau \ , \\ X^i &= \ell_{\mathrm{AdS}} \Omega_i \sinh\rho \ , \quad i \in \{1, \ldots, d\} \ , \end{aligned} \tag{2.5}$$

where Ω_i with $\sum_i \Omega_i^2 = 1$ parameterize the unit sphere S^{d-1}. These are so-named because, with $\rho \geq 0$, they cover the entire spacetime exactly once. In these coordinates, the metric (2.1) becomes

$$\mathrm{d}s^2 = \ell_{\mathrm{AdS}}^2 \left(-\cosh^2\rho \, \mathrm{d}\tau^2 + \mathrm{d}\rho^2 + \sinh^2\rho \, \mathrm{d}\Omega_{d-1}^2\right) \ . \tag{2.6}$$

The metric now has a manifest timelike Killing vector ∂_τ, and consequently τ is a global time coordinate on the manifold. However, the form of (2.5) indicates that τ is periodic with period 2π, and thus the metric contains closed timelike curves.[3] To avoid the concomitant pathologies (see, e.g., [67]), we "unwrap" the timelike circle by taking $\tau \in \mathbb{R}$ instead of identifying endpoints. The resulting space is

[3] This is not a failure inherent to our choice of coordinates, but is rather inherited from the extra time dimension in (2.1). It is a general fact that any spacetime with more than one temporal direction will contain closed timelike curves.

properly referred to as the *universal cover* of AdS, but for the sake of conciseness we shall follow convention and take the unqualified "AdS" to refer to the universal covering space of the embedding henceforth.

By making a further change of variables $\tan\theta = \sinh\rho$, with $\theta \in [0, \pi/2]$, and rescaling to remove the conformal factor $\ell_{\text{AdS}}^2/\cos^2\theta$, we obtain the Penrose compactification

$$\mathrm{d}s^2 = -\,\mathrm{d}\tau^2 + \,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\Omega_{d-2}^2 \ . \tag{2.7}$$

This is the *Einstein static universe*, with topology $\mathbb{R} \times S^{d-1}$. However, since $0 \leq \theta \leq \pi/2$, AdS covers only half the spacetime; we shall return to this point momentarily. The conformal boundary at $\theta = \pi/2$ will play a central role in the AdS/CFT correspondence, since this is where the CFT is said to live.

Another useful choice of coordinates is the *Poincaré patch*, defined via

$$\begin{aligned} X^0 &= \frac{\ell_{\text{AdS}}^2}{2r}\left(1 + \frac{r^2}{\ell_{\text{AdS}}^4}\left(\mathbf{x}^2 - t^2 + \ell_{\text{AdS}}^2\right)\right) \ , \\ X^i &= \frac{rX^i}{\ell_{\text{AdS}}} \ , \quad i \in \{1, \ldots, d\} \ , \\ X^d &= \frac{\ell_{\text{AdS}}^2}{2r}\left(1 + \frac{r^2}{\ell_{\text{AdS}}^4}\left(\mathbf{x}^2 - t^2 - \ell_{\text{AdS}}^2\right)\right) \ , \\ X^{d+1} &= \frac{rt}{\ell_{\text{AdS}}} \ , \end{aligned} \tag{2.8}$$

with $t \in \mathbb{R}$, $r \in \mathbb{R}_+$, and $\mathbf{x} = \left(x^1, \ldots, x^d\right)$. In contrast to the global coordinates above, the restriction to $r > 0$ implies that the Poincaré patch covers only half the spacetime, but it has the advantage of being locally equivalent to flat space – parameterized by $t, \mathbf{x}$ – plus an extra warped direction r. The latter is apparent by writing the metric (2.1) in the coordinates (2.8),

$$\mathrm{d}s^2 = \frac{\ell_{\text{AdS}}^2}{r^2}\,\mathrm{d}r^2 + \frac{r^2}{\ell_{\text{AdS}}^2}\left(-\,\mathrm{d}t^2 + \,\mathrm{d}\mathbf{x}^2\right) = \frac{\ell_{\text{AdS}}^2}{r^2}\,\mathrm{d}r^2 + \frac{r^2}{\ell_{\text{AdS}}^2}\eta_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu \ , \tag{2.9}$$

where $\eta_{\mu\nu}$ is the standard Minkowski metric, with $x^0 = t$. In these coordinates, the conformal boundary is at $r \to \infty$, while the opposite limit $r \to 0$ is characterized by a degenerate Killing horizon.[4] This is also referred to as the Poincaré horizon, and is merely a coordinate – as opposed to curvature – singularity: the other side,

[4] A Killing horizon is a null hypersurface defined by $k_\mu k^\mu = 0$ for some Killing vector k_μ. We may define the *surface gravity* κ of a static Killing horizon as the acceleration (measured at infinity) necessary to remain there. Note however that this is only *analogous* to the Newtonian concept of surface gravity as acceleration; the latter is formally infinite on the event horizon of a black hole, and thus the definition via Killing horizons is necessary in relativity. Killing horizons are also significant in that the conserved quantity associated to k_μ is undefined thereupon. "Degenerate" refers to the case where $\kappa = 0$.

$r < 0$, is covered by a second Poincaré patch; the two patches collectively cover the whole spacetime. See figure 2.1.

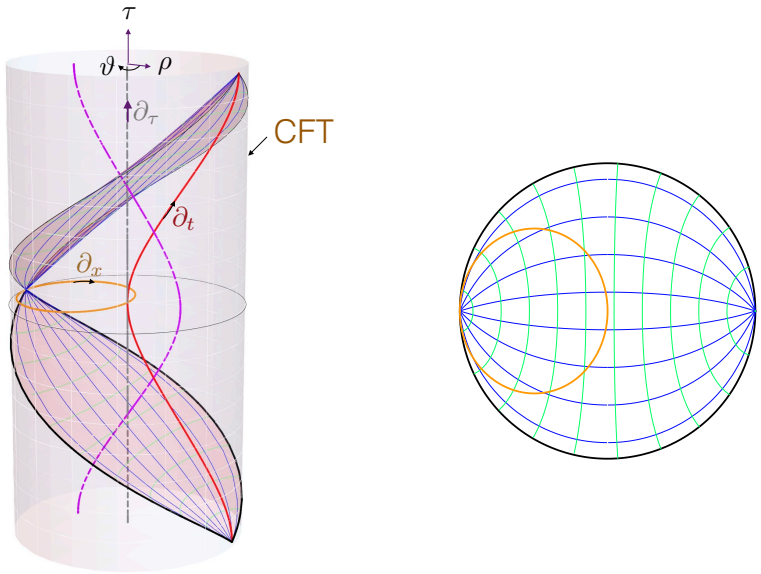


Figure 2.1: AdS_5 in global coordinates. In the left image, the Poincaré patch is the region between the two light sheets that comprise the Poincaré horizon. The orange (∂_x) and red (∂_t) paths illustrate the orbit of Killing fields, while the confining nature of AdS is demonstrated by the purple, oscillating timelike curve. The right image is a constant timeslice, with spacelike geodesics in green and the projection of null geodesics (e.g., those that form the Poincaré horizon) in blue; see also figure 2.2 below. Image source: [64].

The geometry of AdS exhibits two curious properties [68]. First, the topology $\mathbb{R} \times S^{d-1}$ implies that AdS is characterized by a timelike infinity (in contrast to the null and spacelike infinities in Minkowski space). As a consequence, there are no Cauchy surfaces in this spacetime: while any spacelike slice does cover the whole space, there nevertheless exist null geodesics from timelike infinity which do not intersect any point thereupon. Thus, while one can make the Cauchy problem well-posed within the half of the Einstein static universe covered by the coordinates (2.7) by specifying boundary conditions at $\theta = \pi/2$, one cannot predict beyond this region, essentially because new information can "sneak in" from timelike infinity.

A second key feature is that, while null geodesics reach the conformal boundary in finite time, timelike geodesics never do. Instead, timelike geodesics emitted from some point p will be reflected by the curved geometry at some finite distance back to an image point p'. They will then diverge outwards again, and continue executing this simple harmonic motion for all time; see figure 2.1.

2.2 Conformal field theory

In standard quantum field theory, one takes Poincaré transformations as the fundamental symmetry group, augmented with possible internal symmetries of the fields. In Euclidean signature, this can be extended to the *conformal group*, which consists of all angle-preserving transformations. Translating this statement into Minkowski spacetime, the conformal group is the most general set of transformations that locally preserves the causal structure (i.e., spacelike, timelike, and lightlike separated points remain so under conformal transformations). We may thus define conformal transformations as those which leave the metric invariant up to some spacetime-dependent factor,

$$g_{\mu\nu}(x) \to \Omega^2(x) g_{\mu\nu}(x) \ . \tag{2.10}$$

One can show [69] that this corresponds to the follow set of finite transformations $x^\mu \to x'^\mu$, with the associated generators:

$$\text{translation :} \qquad x'^\mu = x^\mu + a^\mu \qquad P_\mu = -i\partial_\mu \tag{2.11}$$

$$\text{dilation :} \qquad x'^\mu = \lambda x^\mu \qquad D = -x^\mu \partial_\mu \tag{2.12}$$

$$\text{rotation :} \qquad x'^\mu = M^\mu{}_\nu x^\nu \qquad L_{\mu\nu} = i\left(x_\mu \partial_\nu - x_\nu \partial_\mu\right) \tag{2.13}$$

$$\text{SCT :} \qquad x'^\mu = \frac{x^\mu - b^\mu \mathbf{x}^2}{1 - 2\mathbf{b}\cdot\mathbf{x} + b^2\mathbf{x}^2} \qquad K_\mu = -i\left(2x_\mu x^\nu \partial_\nu - \mathbf{x}^2 \partial_\mu\right) \tag{2.14}$$

The last of these, *special conformal transformations*, induces the scale factor

$$\Omega(x) = \left(1 - 2\mathbf{b}\cdot\mathbf{x} + b^2\mathbf{x}^2\right)^2 \ , \tag{2.15}$$

and, while somewhat more obscure than the other three, can be equivalently thought of as an inversion $x^\mu \to x^\mu/\mathbf{x}^2$, followed by a translation, and then another inversion:

$$\frac{x'^\mu}{\mathbf{x}'^2} = \frac{x^\mu}{\mathbf{x}^2} - b^\mu \ . \tag{2.16}$$

The set of commutation relations satisfied by the generators P_μ, D, $L_{\mu\nu}$, and K_μ defines the *conformal algebra*. We shall not digress upon it here, except to note that we may define a new set of generators [69]

$$\begin{aligned} J_{\mu\nu} &= L_{\mu\nu} \ , & J_{-1\mu} &= \frac{1}{2}\left(P_\mu - K_\mu\right) \ , \\ J_{-10} &= D \ , & J_{0\mu} &= \frac{1}{2}\left(P_\mu + K_\mu\right) \ , \end{aligned} \tag{2.17}$$

where $J_{ab} = -J_{ba}$ with $a, b \in \{-1, 0, 1, \ldots, d\}$, which satisfy the commutation relations of $SO(d, 2)$,

$$[J_{ab}, J_{bc}] = i\left(\eta_{ad}J_{bc} + \eta_{bc}J_{ad} - \eta_{ac}J_{bd} - \eta_{bd}J_{ac}\right) \ , \tag{2.18}$$

where $\eta_{ab} = \text{diag}\,(-1, -1, 1, \ldots, 1)$. Thus we see that the conformal group is isomorphic to $SO(d, 2)$. As mentioned in the previous subsection, this is precisely the isometry group of AdS_{d+1}, and the agreement between these two symmetry groups is a foundational property of the AdS/CFT correspondence discussed in the next subsection.

While the symmetry group above is in some sense the defining property of a CFT, our interest lies mainly in the definition of states and (local) operators in the quantum theory. *A priori*, these are rather different entities: as explained in [70], states are delocalized over an entire spatial slice, while local operators are defined at a single spacetime point. However, one of the remarkable features of CFTs is the existence of an isomorphism between them, known as the *state-operator correspondence*. As this isomorphism is fundamental to the definition of bulk states from CFT operators, we shall provide a brief overview of the most salient ingredients. Our exposition will closely follow the excellent pedagogical reference [70], to which the interested reader is referred for more details.

One tool from quantum field theory with special importance for CFT is the *operator product expansion (OPE)*, which describes the behaviour of local operators as their spacetime points approach one another—a maneuver that, in QFT, is notoriously fraught with divergences. Denoting local CFT operators $\mathcal{O}_i$, the OPE is defined as

$$\mathcal{O}_i(z, \bar{z})\mathcal{O}_j(w, \bar{w}) = \sum_k C_{ij}^k\left(z - w, \bar{z} - \bar{w}\right)\mathcal{O}_k(w, \bar{w}) \ , \tag{2.19}$$

where C_{ij}^k are a set of functions which (by virtue of translation invariance) depend only on the separation $|z - w|$. It is important to note that (2.19) only holds as an operator equation, i.e., as an insertion within time-ordered correlation functions; but it is common to neglect writing $\langle \ldots \rangle$, and we shall follow this convention.[5]

The OPE is clearly singular as $z \to w$. And in fact, it is precisely this singular behaviour in which we're generally interested. The reason for this stems from the *Ward identities*. These are beautifully derived in Tong's lectures [70], and we will not elaborate upon them here. Suffice to say that these are analogous to Noether's theorem in quantum field theory, and allow one to obtain conservation equations for operator insertions that coincide with the conserved current J associated to some symmetry transformation $\mathcal{O} \to \mathcal{O} + \epsilon\delta\mathcal{O}$. The relevant feature here is that

[5]The radius of convergence is equal to the distance to the nearest other insertion; in other words, the other operators in $\langle \ldots \rangle$ are arbitrary so long as they're displaced beyond $|z - w|$.

the change in such an operator $\mathcal{O}$ is then given by the residue in the OPE between J and $\mathcal{O}$, which one can re-express in terms of the stress tensor as

$$\delta\mathcal{O} = -\mathrm{Res}\left[\epsilon(z)T(z)\mathcal{O}\right] \ , \tag{2.20}$$

where $\delta z = \epsilon(z)$, and similarly for $\delta\bar{z} = \bar{\epsilon}(\bar{z})$. The upshot is that the OPE between an operator and the stress tensor tells us how that operator transforms under the conformal symmetry group.

Accordingly, the OPE of an operator $\mathcal{O}$ with the stress-energy tensor T is of central importance in the study of CFTs. In particular, we define a *primary operator* as one whose OPE with T truncates at order $(z-w)^{-2}$:

$$T(z)\mathcal{O}(w,\bar{w}) \sim h\frac{\mathcal{O}(w,\bar{w})}{(z-w)^2} + \frac{\partial\mathcal{O}(w,\bar{w})}{z-w} \ , \tag{2.21}$$

where "$\sim$" denotes equivalence up to non-singular terms, and h will be defined momentarily. A similar expression holds for $\bar{T}$.

Primary operators have especially simple transformation properties, which one can straightforwardly derive from (2.21) [70]: under a finite conformal transformation $z \to z'$, $\bar{z} \to \bar{z}'$, a primary operator transforms as

$$\mathcal{O}(z,\bar{z}) \to \bar{\mathcal{O}}(z',\bar{z}') = \left(\frac{\partial z'}{\partial z}\right)^{-h}\left(\frac{\partial \bar{z}'}{\partial \bar{z}}\right)^{-\bar{h}}\mathcal{O}(z,\bar{z}) \ . \tag{2.22}$$

where $(h,\bar{h})$ are the *weights* of the operator. These encode information about how the operator transforms under rotations and scalings. In particular, they allow us to define the *spin*, $s = h - \bar{h}$, and *scaling dimension*, $\Delta = h + \bar{h}$; the latter will surface explicitly in our holographic toy models in later chapters. More generally, as we shall see below, the spectrum of weights of primary operators is equivalent to the spectrum of particle masses, and thus contains important information about the CFT.

There is one more specific OPE worth mentioning before we move on: that of the stress tensor with itself,

$$T(z)T(w) \sim \frac{c/2}{(z-w)^4} + \frac{2T(w)}{(z-w)^2} + \frac{\partial T(w)}{z-w} \ , \tag{2.23}$$

where c is the *central charge*, one of the most important numbers characterizing the CFT. It has several simultaneous roles: it is the Casimir energy of the system, appears in Cardy's formula $S(E) \sim \sqrt{cE}$ for the entropy of high-energy states [71, 72],[6] and plays a key role in the *c-theorem* that enables us to understand

[6]The relation between the spectrum and the density of states will be explored in detail in chapter 6.

CFTs as fixed points of the renormalization group flow [73]. It also appears in the algebra of the symmetry generators, to be introduced below.

The relation between states and operators mentioned above relies on the fact that there exists a conformal mapping between the cylinder and the complex plane. Consider the former, parametrized by $w = \sigma + i\tau$ with $\sigma \in [0, 2\pi)$. Then the conformal transformation $z = e^{-iw}$ maps constant timeslices on the cylinder to circles of constant radius in the complex plane. Note that the infinite past $\tau \to -\infty$ is mapped to the origin at $z = 0$; time-evolution on the cylinder (generated by the Hamiltonian $H = \partial_\tau$) then corresponds to radial evolution (generated by the dilatation operator $D = z\partial + \bar{z}\bar{\partial}$) on the plane.

This mapping allows us to define an important class of operators, the *Virasoro generators*. These are obtained by Fourier expanding the stress tensor on the cylinder,

$$T(w) = -\sum_{m=-\infty}^{\infty} L_m e^{imw} + \frac{c}{24} \ , \tag{2.24}$$

which then maps to a Laurent expansion on the plane; one then inverts the expansion by a suitable contour integral to obtain

$$L_n = \frac{1}{2\pi i} \oint \mathrm{d}z z^{n+1} T(z) \ , \tag{2.25}$$

and similarly for $\bar{L}_n$. One can think of these as the conserved (under radial evolution) charges associated to conformal transformations $\delta z = z^{n+1}$ (sim. $\delta\bar{z} = \bar{z}^{n+1}$). Upon quantizing the theory, these conserved charges become generators for the conformal transformations, and are then known as Virasoro generators. For example, L_{-1} generates translations, while L_0 generates scaling and rotations; the latter implies that on the plane, the evolution operator can be represented as $D = L_0 + \bar{L}_0$.

The set of commutation relations satisfied by these conserved charges is the well-known *Virasoro algebra*,

$$[L_m, L_n] = (m-n)L_{m+n} + \frac{c}{12} m\left(m^2 - 1\right)\delta_{m,-n} \ . \tag{2.26}$$

The algebra gives us a great deal of information about the states of the CFT. Consider an eigenstate $|\psi\rangle$ of L_0, $\bar{L}_0$, with

$$L_0 |\psi\rangle = h |\psi\rangle \ , \quad \bar{L}_0 |\psi\rangle = \bar{h} |\psi\rangle \ . \tag{2.27}$$

Since the evolution operator D maps to the Hamiltonian on the cylinder, this

corresponds to an energy eigenstate with

$$\frac{E}{2\pi} = h + \bar{h} - \frac{c+\bar{c}}{24} \ , \tag{2.28}$$

and hence the eigenvalues h, $\bar{h}$ correspond to the energy of the state. This further implies that L_{-n}, L_n with $n > 0$ act as raising and lowering operators, respectively, analagous to the ladder operators in quantum mechanics; e.g.,

$$L_0 L_n \left|\psi\right\rangle = (L_n L_0 - n L_n) \left|\psi\right\rangle = (h-n) \left|\psi\right\rangle \ , \tag{2.29}$$

where we have used (2.26). And just as in quantum mechanics, there must exist a minimum energy state so that the spectrum is bounded from below; such states are annihilated by all L_n with $n > 0$,

$$L_n \left|\psi\right\rangle = \bar{L}_n \left|\psi\right\rangle = 0 \ , \quad \forall n > 0 \ , \tag{2.30}$$

and are called *primary states* (sometimes referred to as "highest weight states" in representation theory). We can then obtain representations of the Virasoro algebra by acting on primary states with L_{-n}. The states that comprise the resulting infinite tower are referred to as *descendants*. For example, beginning with the primary state ψ, we have

$$\begin{gathered} \left|\psi\right\rangle \\ L_{-1} \left|\psi\right\rangle \\ L_{-1}^2 \left|\psi\right\rangle \ , \ L_{-2} \left|\psi\right\rangle \\ L_{-1}^3 \left|\psi\right\rangle \ , \ L_{-1} L_{-2} \left|\psi\right\rangle \ , \ L_{-3} \left|\psi\right\rangle \end{gathered}$$

and so on. The fact that we can build such an irreducible representation of the Virasoro algebra from primary operators in this manner implies that knowing the spectrum of primary operators is tantamount to knowing the (energy) spectrum of the entire CFT.[7]

We are now in a position to elucidate the state-operator correspondence mentioned above—as well as justify the re-use of notation $(h, \bar{h})$ for both the weights of primary operators and the energy of primary states. The correspondence relies crucially on the aforementioned fact that we can conformally map the cylinder to the complex plane. Consider the evolution from an initial state Ψ_i on the cylinder

[7] We are ignoring the subtlety of null states, which are linear combinations of descendants with vanishing norm. See for example [74] for an exposition of their importance in string theory, as well as a great deal more information about the Virasoro algebra in that context.

to some final state Ψ_f,

$$\Psi_f\left[\phi_f(\sigma),\tau_f\right]=\int_{\phi(\tau_i)=\phi_i}^{\phi(\tau_f)=\phi_f}\mathcal{D}\phi e^{-S[\phi]}\psi_i\left[\phi_i(\sigma),\tau_i\right]\ . \tag{2.31}$$

When mapping to the complex plane, this becomes an integral over the annulus defined by these two states at circles of constant radius, r_i and $r_f > r_i$. In particular, changing the initial state is equivalent to changing the boundary condition at $|z| = r_i$. Specifying the initial state in the infinite past then corresponds to a local operator insertion at $z = 0$. Then the path integral over the entire disk $|z| \leq r_f$ defines the corresponding state in the field theory,

$$\Psi\left[\phi_f,r\right]=\int\mathcal{D}\phi e^{-S[\phi]}\mathcal{O}(z=0)\ , \tag{2.32}$$

where the integral is over all field configurations in the disk, with upper boundary condition $\phi(r) = \phi_f$.

This provides the relation between primary states and operators. Let us take the operator $\mathcal{O}$ in the above path integral to be primary, and denote the corresponding state (the l.h.s.) by $|\mathcal{O}\rangle$. Acting with (2.25), we have[8]

$$L_n\left|\mathcal{O}\right\rangle=\oint\frac{\mathrm{d}z}{2\pi i}z^{n+1}T(z)\mathcal{O}(z=0)\sim\oint\frac{\mathrm{d}z}{2\pi i}z^{n+1}\left(\frac{h\mathcal{O}}{z^2}+\frac{\partial\mathcal{O}}{z}\right) \tag{2.33}$$

where we have used (2.21). From this expression, we deduce the following: first, note that if $n > 0$, the r.h.s. vanishes, and thus we recover the condition for $|\mathcal{O}\rangle$ to be a primary state, (2.30),

$$L_n\left|\mathcal{O}\right\rangle=0\ ,\quad\forall n>0\ . \tag{2.34}$$

If $n = 0$, the $1/z$ singularity drops out, and we retain only the leading-order term, hence

$$L_0\left|\mathcal{O}\right\rangle=h\left|\mathcal{O}\right\rangle\ , \tag{2.35}$$

and thus the weights indeed correspond to the (energy and angular momentum) spectrum of states in the CFT. We shall consider the spectrum of a particular CFT in detail in chapter 6. Finally, note that for $n = -1$,

$$L_{-1}\left|\mathcal{O}\right\rangle=\left|\partial\mathcal{O}\right\rangle\ , \tag{2.36}$$

which confirms our earlier claim that L_{-1} generates translations.

[8] Recall that the OPE is only valid within correlation functions, and thus the path integral $\int\mathcal{D}\phi e^{-S[\phi]}$ is implicit in this expression.

2.3 Elements of the correspondance

Having assembled the two basic ingredients, we may now state the correspondence.

AdS/CFT was originally discovered by studying the low-energy limit of a system of D-branes in string theory [61]. Specifically, one considers a stack of D3-branes in type IIB superstring theory in ten dimensions, which has a description in terms of both closed and open strings. In the low-energy limit ($E \ll l_s^{-1}$), the open string description reduces to $\mathcal{N} = 4$ super Yang-Mills (SYM), while the closed strings backreact in such a manner as to produce $\mathrm{AdS}_5 \times S^5$. For more details on how these dual pictures emerge, the interested reader is referred to section 5.2 of [62]. The upshot is that the duality of AdS/CFT can be seen to arise as a consequence of the duality between closed and open strings [75].[9]

The field theory is conformal, and thus SYM is referred to as the CFT or "boundary" side of the correspondence. It is dual to string theory living on AdS, referred to as the "bulk". The latter terminology arises from comparing the symmetries of the two systems. As we saw in the previous subsection, the isometry group of AdS_5 is $SO(4,2)$, which is precisely the conformal group.[10] Furthermore, we emphasized that the conformal boundary of AdS is timelike with topology $\mathbb{R} \times S^{d-1}$ (with $\mathbb{R}$ temporal), which is conformally equivalent to Euclidean $\mathbb{R}^d$. Since the state-operator correspondence allows us to map operators on the latter to states on the former, we speak of the CFT as living on the boundary S^{d-1}, while the dual string theory lives in the interior of the cylinder—i.e., the bulk. As emphasized in [17] however, one should take care not to interpret this terminology too literally, namely as suggesting a single Hilbert space encompassing both bulk and boundary, with edge modes on the latter. Rather, we have separate Hilbert spaces for each theory, which are dual in the true sense of the word [76]. This is the meaning behind the statement below that the two theories are "dynamically equivalent": both the bulk and the boundary contain a complete description of the same physics.

It is in this sense that AdS/CFT provides a concrete realization of the holographic principle [77], in that the complete physical description of a $(d+1)$-dimensional spacetime is mapped to an equivalent description on the d-dimensional boundary. What is even more remarkable is that the duality relates a (bulk) theory with gravity to a (boundary) theory without it. Thus AdS/CFT also provides a precise realization of the emergent spacetime paradigm; understanding precisely how gravity emerges along with the extra spatial (radial) dimension in the bulk is an active area of current research.[11]

[9] That said, it is possible to argue for the correspondence independently of string theory [64].

[10] We are suppressing the $SO(6)$ symmetry of the S^5, since this just corresponds to a global symmetry in the field theory [64].

[11] Implicit in the claim that the bulk spacetime emerges from the (gravitationless) boundary,

The above sketch corresponds to the most well-studied form of the correspondence, which states that $\mathcal{N} = 4$ SYM with gauge group $SU(N)$ is dynamically equivalent to type IIB superstring theory on $AdS_5 \times S^5$ [62]. The pertinent field theory parameters are the Yang-Mills coupling constant $g_{\rm YM}$ and the degree of the gauge group N, while those on the string theory side are the string length $l_s = \sqrt{\alpha'}$, coupling $g_{\rm s}$, and radius of curvature $\ell_{\rm AdS}$. They are related as follows:

$$g^2_{\rm YM} = 2\pi g_{\rm s} \ , \quad 2g^2_{\rm YM} N = \ell^4_{\rm AdS}/\alpha'^2 \ . \tag{2.37}$$

As alluded above, the large N limit features in the correspondence in an essential way. Considering the relationship between parameters (2.37), taking the low-energy ($l_{\rm s} \to 0$) limit is equivalent to[12]

$$\lambda \equiv g^2_{\rm YM} N \to \infty \ , \tag{2.38}$$

where λ is called the *'t Hooft coupling*, and (2.38) is referred to as the *'t Hooft limit*. The name hails back to early work by 't Hooft, who showed that a gauge theory is equivalent to string theory in the limit $N \to \infty$ with λ held fixed [78]. In particular, the perturbative expansion of a non-Abelian gauge theory in $1/N$ corresponds to a loop expansion in string theory with $g_{\rm s} \sim 1/N$. See, e.g., [62] or [63] for a pedagogical exposition of this relationship.

However, the gauge theory is valid at any coupling, and thus a natural question is to what extent the CFT gives rise to a gravity theory with a sensible semiclassical description. Turning this question around, the requirement of semiclassical gravity in the bulk imposes certain constraints on the CFT, essentially such that the perturbative expansion in $1/N$ is well-defined. We shall return to this issue in detail in chapter 6.

Before proceeding with the details of the correspondence, two other key features bear mentioning. First, one of the most useful aspects of AdS/CFT is that it is a *strong/weak* duality: the 't Hooft limit $\lambda \to \infty$ corresponds to a strongly coupled field theory on the boundary, while in the bulk it implies that the string scale $l_{\rm s}$ is much smaller than the AdS curvature scale $\ell_{\rm AdS}$. Thus strongly coupled SYM is dual to supergravity on weakly curved spacetime. In fact, this is sometimes referred to as the *weak form* of the AdS/CFT correspondence; the *strong form* asserts that

rather than the reverse, is the consideration of the CFT as epistemologically prior. That is, we have no alternative non-perturbative description of string theory in the bulk, and thus the correspondence can be viewed as providing a definition (via the boundary CFT) of quantum gravity in AdS spacetime, at least in the semi-classical limit [17]. This is also the reason that Maldacena's "conjecture" is hard to prove.

[12] On the supergravity side, the dimensionless parameter is $\ell_{\rm AdS}/l_{\rm s}$. The limit thus corresponds to making the AdS length scale sufficiently larger than the string scale such that one is in the weak curvature regime, where the low-energy approximation is valid [62].

the duality holds true at any value of N and λ.[13] Nevertheless, the strong/weak duality continues to hold by virtue of the relationship (2.37): a general feature of gauge/gravity duality is that a strongly (weakly) coupled theory on the boundary is mapped to a weakly (strongly) coupled theory in the bulk. The reason this is particularly useful is that we understand string theory quite well in the weakly coupled regime, $g_s \ll 1$, and thus we can use the bulk side to understand strongly coupled field theories (which are generally hard). Conversely, a weakly coupled CFT allows us to gain insight into string theory in the non-perturbative regime, where no alternative tractable description is available.

A second fundamental feature, first emphasized in [77], is that AdS/CFT serves as a *UV/IR* duality. Specifically, the authors of [77] showed that a UV cutoff in the boundary corresponds to an IR cutoff in the bulk. One can think of the radial direction as parameterizing this relationship, insofar as the radial coordinate r in (2.9) is associated with the energy scale of the gauge theory; the boundary limit $r \to \infty$ corresponds to UV physics in the CFT. It is relatively easy to visualize this relationship in terms of geodesics. Consider two spacelike separated points on the boundary. Because AdS is negatively curved, the (properly regularized) geodesic that connects them through the bulk is actually shorter than the one on the boundary. The larger the separation on the boundary (that is, the further into the IR), the deeper this geodesic will penetrate into the bulk (see figure 2.2). This is a fact which will enter heavily in chapter 4. For now, suffice it to say that that UV/IR duality is also referred to as a *scale/radius* duality for this reason.

The fact that the radial direction in the bulk is associated with the energy scale in the dual field theory is a fundamental feature that lies at the heart of the holographic renormalization group [79, 80, 81, 82], and indeed there have been efforts to understand AdS/CFT as a holographic RG flow [83, 84]. We will not have much to say about this, except to note that the emergence of the radial direction in this context [85, 86] is similar in spirit to the notion of emergent spacetime in the more general sense [60, 87].

Throughout the above, we have been alluding to a map between these dual theories. This detailed mapping is the purview of the so-called AdS/CFT dictionary, which provides a precise relationship between entries on either side. The dictionary is far from complete, and indeed the search for missing elements serves as a motivating force for much of this work. Accordingly, in the remainder of this section we shall review the main entries in the dictionary, in order to lay the groundwork for what follows.

We have already alluded to the matching of symmetry groups. In particular, the symmetry generators of the $SO(d, 2)$ in the CFT correspond to the symmetry

[13] To avoid any possible confusion: use of "weak" and "strong" in this latter sense simply means that the strong form implies the weak form; the adjectives here have nothing to do with their previous reference to the strengths of various couplings.

generators of AdS. The Hamiltonians of the two sides also match, and therefore any bulk quantity that depends on the space of states (e.g., the thermal partition function) is given (that is, computable) in the CFT by the same. However, as one might have predicted by the theme of this thesis, defining local bulk operators is far more subtle. Indeed, insofar as AdS/CFT is a prescription for quantum gravity, strictly speaking such operators do not exist.[14] However, one can make precise the notion of the boundary limit of local bulk fields:[15]

$$\lim_{z\to 0} z^{-\Delta}\phi\left(t,z,\Omega\right) = \mathcal{O}\left(t,\Omega\right) \ , \tag{2.39}$$

where z is the radial coordinate (with the boundary at $z=0$), ϕ is a bulk scalar field with scaling dimension Δ, and $\mathcal{O}$ is a CFT scalar primary. The scaling dimension is given in terms of the mass of the bulk scalar m by $\Delta\left(\Delta-d\right)=m^2$, and hence there are two possible solutions [90],

$$\Delta_{\pm} = \frac{d}{2} \pm \sqrt{\frac{d^2}{4}+m^2} \ . \tag{2.40}$$

In fact, it has been shown [91] that for $m^2 > -d^2/4+1$, only Δ_+ is admissible, while for masses in the range

$$-\frac{d^2}{4} < m^2 < -\frac{d^2}{4}+1 \ , \tag{2.41}$$

both $\Delta_{\pm}$ lead to consistent solutions. Given (2.39), the choice of scaling dimension determines the fall-off of the wave function as $z\to 0$; thus in this context, a consistent solution is one that gives rise to normalizable bulk modes [92]. This will be a subtle yet important point in the toy model put forth in chapter 5.

The lower limit in (2.41) is the so-called *Breitenlohner-Freedman (BF) bound*. Note that unlike in flat space, AdS permits the existence of states with negative mass squared without the usual tachyonic pathologies, provided they aren't below the BF bound. Very crudely, one can think of the negatively curved spacetime of AdS (there's a factor of $\ell_{\text{AdS}}^2=1$ in the denominator in (2.41)) as "compensating" for the backreaction, so that small excitations do not induce the usual vacuum instability.

Eqn. (2.39) is known as the *extrapolate dictionary* [77, 93]. Properly speaking,

[14] This is merely the statement that there are no local diffeomorphism-invariant operators in quantum gravity, since the backreaction of the operator on the spacetime is not gauge-invariant. In other words, one cannot satisfy the Gauss constraint since any local excitation carries non-zero energy. See [88, 89] and references therein for some interesting work in this vein.

[15] As pointed out in [17], this is analagous to how one defines states in Minkowski space in terms of free fields at infinity.

the equivalence is stated between correlation functions in the bulk and boundary:

$$\lim_{z\to 0} z^{-n\Delta} \langle \phi(\mathbf{x}_1, z) \dots \phi(\mathbf{x}_n, z) \rangle = \langle \mathcal{O}(\mathbf{x}_1) \dots \mathcal{O}(\mathbf{x}_n) \rangle \ . \tag{2.42}$$

where $\mathbf{x}_i = (t_i, \Omega_i)$. There is also an alternative mapping, the *differentiate dictionary* [94, 95], which is stated as an equivalence of partition functions:

$$Z_{\text{bulk}} = Z_{\text{CFT}} \ , \tag{2.43}$$

where Z_{bulk} refers to the on-shell bulk action at large N. These two prescriptions were shown to be equivalent in [96]; in particular, one can compute the same correlation functions (2.42) by the usual variation of the path integral, where the sources in Z_{CFT} correspond to the boundary insertions of bulk fields. In this thesis, we will generally have in mind the extrapolate form of dictionary, but the differentiate version is very useful for other purposes.[16]

We note in passing that, while we stated the extrapolate dictionary for scalar fields, it holds more generally. For example, the energy-momentum tensor $T_{\mu\nu}$ in the CFT is dual to the metric tensor in AdS, and the conserved currents arising (via Noether's theorem) from any global symmetries of the field theory are dual to gauge fields in the bulk [17].

While the extrapolate dictionary provides a prescription for bulk fields in the boundary limit, points deeper in the bulk are necessarily described nonlocally. This essentially follows from the freedom in reconstructing a given bulk operator from multiple boundary regions [101, 3], a fact which will feature centrally in chapter 5. And while multiple prescriptions have been proposed for reconstructing general bulk operators, the most popular and relevant is the eponymous *HKLL prescription* of Hamilton, Kabat, Lifschytz, and Lowe [102, 103, 104], and further developed by Kabat and Lifschytz [105, 106, 107, 108] (see also [109] for earlier work in this vein). The basic idea is that, given a local operator $\mathcal{O}$ dual to the boundary limit of a bulk field as in (2.39), a local bulk field at some finite distance from the boundary is dual to a nonlocal operator defined by integrating over all spacelike-separated operators $\mathcal{O}$:

$$\phi(t, \mathbf{x}) = \int \mathrm{d}\mathbf{x}' K(\mathbf{x}'|z, x)\, \mathcal{O}(\mathbf{x}') \tag{2.44}$$

where here we absorb both the angular and temporal dependence into $\mathbf{x}$, and the prime denotes coordinates on the boundary. The integration kernel K is called the *smearing function*, the explicit form of which requires solving the bulk equations of motion for the field in question, $\left(\Box - m^2\right)\phi = 0$. A pedagogical explanation

[16] For example, identifying the bulk and boundary partition functions features crucially in the derivation of the Ryu-Takayanagi prescription [97] by Lewkowycz and Maldacena [98, 99, 100].

of this procedure can be found in [110], which also mentions various subtleties therein that we will encounter in chapter 5.

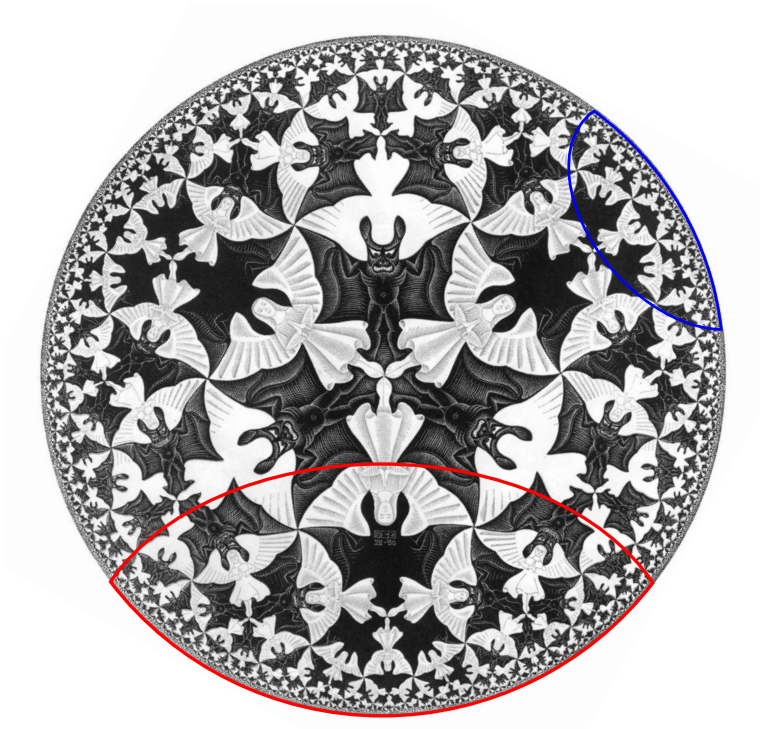


Figure 2.2: A constant timeslice of global AdS_{d+1}, with $d-2$ dimensions suppressed. The geometry of the hyperbolic disc is nicely illustrated by Escher's *Circle Limit IV*: the devils are all the same size. We have sketched two boundary regions and their corresponding Ryu-Takayanagi surfaces. Note that the larger boundary region corresponds to a deeper bulk probe. This is a reflection of the UV/IR duality mentioned above, and will feature prominently in the main text, especially in chapter 4.

The dictionary also contains intrinsically nonlocal, geometric entries. These play an especially important role in this work. The most well-known is the *Ryu-Takayanagi* proposal [97, 111], subsequently generalized to the time-dependent case in [112]. This states that the entanglement entropy associated to some subregion A of the boundary,

$$S_A = -\mathrm{tr}\rho_A \ln \rho_A \ , \tag{2.45}$$

is given by the area $A(\Sigma)$ of the codimension 2 minimal surface Σ in the bulk,

with $\partial\Sigma = \partial A$:[17]

$$S_A = \frac{A(\Sigma)}{4} \ . \tag{2.46}$$

See figure 2.2. The similarity to the black hole entropy formula (1.5) is both deliberate and striking. Indeed, multiple lines of evidence point towards entanglement as being a fundamental ingredient for bulk reconstruction, and we shall have more to say about this later.

A superficially similar prescription is called *causal holographic information* [113, 114], in which the area of an extremal surface in the bulk, in Planck units, is identified as quantifying the information content of the corresponding boundary region A. The key difference between these surfaces and those employed by Ryu-Takayanagi is that the former are defined by the bulk domain of dependence given by the causal wedge of A; see figure 2.3. This implies that the causal surfaces are delimited by null rays, while the Ryu-Takayanagi surfaces are defined as minimal spacelike geodesics. This difference will feature crucially in chapter 4.

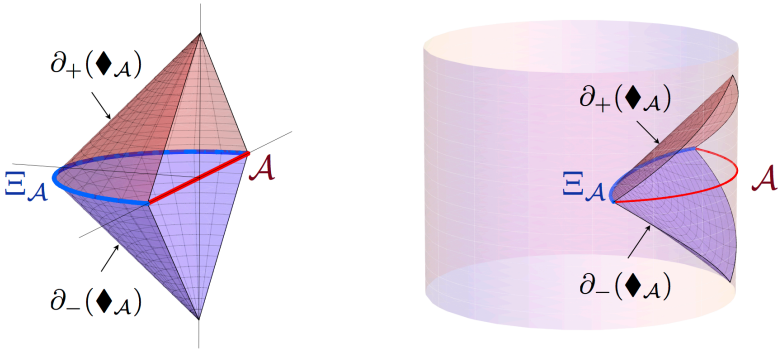


Figure 2.3: Sketch of the causal wedge associated with causal holographic information in planar (left) and global (right) AdS_3. $\blacklozenge_{\mathcal{A}}$ denotes the bulk causal wedge defined by the boundary domain of dependence of the region $\mathcal{A}$, while $\Xi_{\mathcal{A}}$ is the causal surface determined by the intersection of $\blacklozenge_{\mathcal{A}}$ with the latter's timeslice. Image source: [115].

A third and final probe of relevance to us is given by *Wilson loops* [116, 117, 118],

$$\mathcal{W}_{\mathcal{C}} = \frac{1}{N}\mathrm{tr}\mathcal{P}\exp\left(i\oint_{\mathcal{C}} A\right) \tag{2.47}$$

where $\mathcal{C}$ is a closed loop on the boundary, and the trace is over the fundamental representation of the gauge group. Physically, $\mathcal{W}_{\mathcal{C}}$ can be thought of as the non-Abelian phase factor associated with parallel transporting an electrically charged

[17] Technically, Σ is the minimum among all such extremal surfaces which are homologous to A.

particle around the closed path $\mathcal{C}$. Since the trace is invariant under cyclic permutations, $\mathcal{W}_\mathcal{C}$ is manifestly gauge-invariant. And since any physical quantity must be gauge-invariant, the formulation of gauge theories in terms of Wilson loop variables provides a natural description of the dual string theory, albeit at the cost of locality.

The bulk dual of a Wilson loop is given by the action of the string worldsheet,

$$S = \frac{1}{2\pi\alpha'} \int \mathrm{d}\tau\, \mathrm{d}\sigma \sqrt{\det G_{MN} \partial_\alpha X^M \partial_\beta X^N} \,, \tag{2.48}$$

where, for example, G_{MN} is the metric of Euclidean $AdS_5 \times S^5$. The worldsheet extends into the bulk as shown in figure 2.4. Wilson loops thus provide a third nonlocal probe of physics in the bulk, and we shall consider it alongside extremal and causal surfaces in chapter 4.

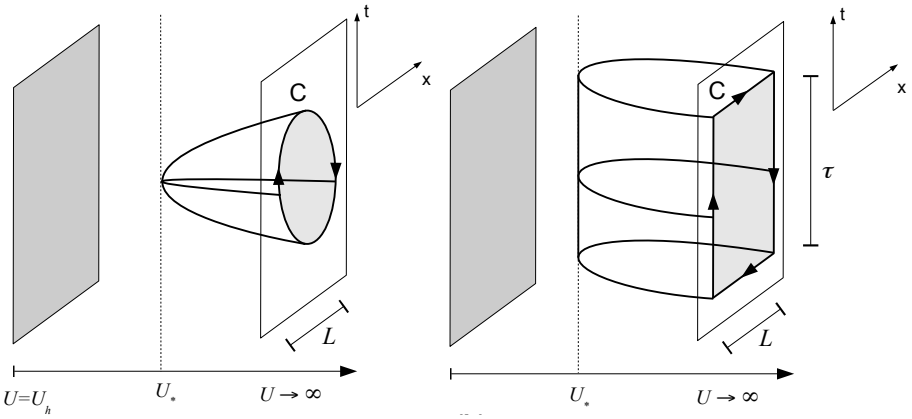


Figure 2.4: Sketch of a Wilson loop in the Poincaré patch for an arbitrary path (left) and in the rectangular limit (right), where $U = r$ (cf. eqn. (2.9)). The path C encloses a region on the boundary at $U \to \infty$, while the plane at U_h represents the Poincaré horizon. Image source: [119].

Ideally, the complete dictionary will allow us to compute any bulk quantity of interest in the CFT. However, an arbitrary CFT is not guaranteed to have a well-defined semi-classical dual: as we mentioned above, it is only in the large N limit that we recover gravity in the bulk. Thus to properly flesh-out the rather bare-bones dictionary above, we must also consider the conditions under which the CFT has such a limit. In particular, this requires imposing the following additional set of constraints [17]:

- There is a finite set of single-trace primary operators $\mathcal{O}_i$ with spin ≤ 2, with equality only for the stress tensor (with $\Delta = d$).

- Given any set of single-trace operators $\{\mathcal{O}_{i_1}, \ldots, \mathcal{O}_{in}\}$, there exists a multi-trace operator $\mathcal{O}_{i_1} \ldots \mathcal{O}_{i_n}$ with $\Delta = \Delta_{i_1} + \ldots + \Delta_{i_n} + O(1/N)$.
- If the 2-point functions of multi-trace operators are normalized to $O(1)$, then all higher correlation functions are suppressed by powers of $1/N$, unless their components can be matched in pairs.[18] In the latter case, the correlation function is the sum over all possible contractions. This important property is called *large N factorization*, and is reminiscent of, though technically distinct from, Wick's theorem for scalar fields [120, 121].
- All $O(1)$ operators are either single-trace primaries, multi-trace primaries, or their descendants.

The importance of these criteria can be summarized as follows [17]: the fact that higher-point correlation functions are suppressed by $1/N$ ensures, via the state-operator correspondence discussed above, that the low-energy spectrum of the CFT indeed corresponds to semiclassical (weakly-coupled) EFT in the bulk. We will investigate these criteria in detail in chapter 6.

[18] So, to take the example in [17], $\langle \mathcal{O}_i(x)\mathcal{O}_j(y)\mathcal{O}_i\mathcal{O}_k(z)\rangle$ is $O(1)$ if $k = j$, and $O(1/N)$ otherwise.

3

Constraining firewalls with causality

This chapter is based on [1].

In this chapter, we shall analyze the causal structure of black hole spacetimes in order to determine whether all the necessary ingredients for the AMPS firewall paradox fit within a single observer's causal patch. We particularly focus on the question of whether the interior partner modes of the outgoing Hawking quanta can, in principle, be measured by an infalling observer. Since the relevant modes are spread over the entire sphere, we answer a simple geometrical question: can any observer see an entire sphere behind the horizon? As we shall see, this will lead to questions concerning the localization of information that will resurface again in subsequent chapters.

We find that for all static black holes in 3+1 and higher dimensions, with any value of the cosmological constant, no single observer can see both the early Hawking radiation and the interior modes with low angular momentum. We present a detailed description of the causal patch geometry of the Schwarzschild black hole in 3+1 dimensions, where an infalling observer comes closest to being able to measure the relevant modes.

3.1 Introduction

Recently, Almheiri, Marolf, Polchinski and Sully (AMPS) [47] identified a remarkable conflict between fundamental physical principles. Consider an "old" black hole—one that has already emitted more than half of the Hawking quanta—and focus on the emission of the next Hawking photon H. The equivalence principle requires that the region near the horizon should look locally like the Minkowski vacuum, requiring that H be strongly entangled with its "partner mode" P be-

hind the horizon (see fig. 3.1). However, unitarity requires that H be strongly entangled with the radiation R that has already been emitted. The monogamy of entanglement prohibits H from being maximally entangled with two distinct systems, and locality dictates that P and H are independent.

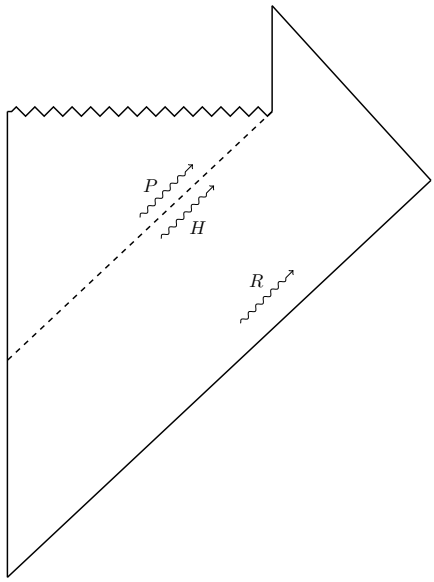


Figure 3.1: Penrose diagram depicting the near-horizon Hawking mode H, its behind-the-horizon partner P, and the early radiation R.

The essential conflict stated here was already present in Hawking's original work [15], and was phrased clearly in terms of entanglement by [21, 122]. However, before the work of AMPS, the information paradox could be addressed with black hole complementarity (BHC) [33]. In a nutshell, the postulates of BHC simply state that no observer ever witnesses a violation of any physical law, since causality restricts access to all the necessary information. Observers who remain outside the black hole have access to H and R and can thus confirm the unitarity of black hole evaporation, while an infalling observer has access to H and P and can verify the equivalence principle.

One key innovation of AMPS was to consider the causal patch of an observer who falls into an old black hole. Such an observer would seem to have access to all three ingredients necessary for the paradox. If that is the case, then black hole complementarity is no longer sufficient to resolve the information paradox. However, a closer inspection of the geometrical limitations of a causal patch may reveal deeper issues in need of investigation, such as those proposed in [123, 124].

In this chapter, we analyze another geometric question: can any single observer

see the entire sphere behind the horizon? This is a very relevant question because the simplest and most robust version of the paradox requires that the Hawking quantum H actually escapes from the black hole. Due to the angular momentum barrier, Hawking radiation occurs almost exclusively in modes with low angular momentum ℓ. Furthermore, one encounters subtleties when trying to address the issue of high-ℓ modes, which we comment on below. Hence we focus our attention on the s-wave firewall, as this version of the AMPS paradox is both the most fundamental conceptually as well as least ambiguous mathematically. In this model, H should be spread over the entire sphere near the horizon, so its entangled partner mode P is also spread over the entire sphere. Therefore, an observer who cannot see the entire behind-the-horizon sphere will have difficulties recognizing the entanglement between these two modes. Thus, in the context of complementarity, classical considerations are necessary to determine whether any observer can identify the relevant quantum state.

We analyze this question for static black holes in all dimensions, in spacetimes with positive, negative, and zero cosmological constant. For static black holes in asymptotically Minkowski spacetime, an infalling observer cannot receive signals from the entire sphere behind the horizon before hitting the singularity. The most interesting case is 3+1 dimensions, where an observer can see nearly the entire sphere, but with an important caveat: there is a trade-off between the radial and angular extent of the causal patch, as we describe. In higher dimensions, less than half the sphere fits within one causal patch.

Adding a negative cosmological constant decreases the region that is causally accessible; for large black holes in asymptotically anti-de Sitter spacetime, in 3+1 and higher dimensions, an infalling observer can only see a small fraction of the horizon sphere, with physical size of order the AdS radius. This result is potentially important for the AdS version of the firewall paradox [125, 126], which some consider to be the most robust against the concerns of computation time [127] and backreaction [128]. Since an infalling observer can only see sub-AdS scales near the horizon, the subtle issue of reconstructing these modes from CFT data can play an important role in the firewall paradox [129, 110, 130].

Adding a positive cosmological constant increases the angular size of the causal patch. However, we show that once the cosmological constant is large enough to allow an observer to collect information from the entire sphere, the information contained in the Hawking radiation cannot fit within the cosmological horizon. In other words, as the cosmological constant is increased, an infalling observer begins to be able to measure P but loses the ability to measure R.

These geometrical results motivate a possible resolution of the firewall paradox: even for an old black hole, some degrees of freedom that are smeared over the entire sphere in the near-horizon zone are entangled with the early radiation, while localized modes in the near-horizon zone are entangled with their partners behind

the horizon. This would then avoid an observable conflict between the equivalence principle and unitarity.

There are reasons to think that the AMPS paradox can be reformulated to only refer to modes within a single causal patch. However, existing arguments tend to assume that the geometry allows the measurement of any desired mode. As we endeavor to show, this is often not the case. Causal patch considerations must be taken into account in order to formulate the paradox as cleanly as possible. Our results thus serve as motivation for such a reformulation.

The organization of this chapter is as follows: in Sec. 3.2 we discuss results for various static black holes, except for Schwarzschild black holes in $3+1$ dimensions, which we treat separately in Sec. 3.3. The reason for this separation is that with the exception of the latter, it is clear that the geometry of the causal patch alone offers an escape from the firewall paradox. In the case of $(3+1)$-dimensional Schwarzschild black holes however, a more detailed analysis is required which occupies the bulk of this work. Then, in Sec. 3.4, we discuss the consequences for entropy and information in the context of the casual patch considerations in the $(3+1)$-dimensional Schwarzschild background.

Cases and issues not addressed here:

BTZ black holes [black holes in (2+1)-dimensional AdS spacetime], named for Bañados, Teitelboim, and Zanelli, are an exception: in this case an infalling observer can collect information from the entire sphere behind the horizon. The physics of black holes in 2+1 dimensions is rather different than in higher dimensions – for example, there are no black holes in asymptotically flat space in 2+1 dimensions. We leave them aside for the purpose of this analysis, but it may be interesting to further consider this case.

We do not treat rotating black holes in this chapter. In this case, there is no spherical symmetry, so it is less obvious which sphere must be contained within the causal patch in order to formulate the paradox. Additionally, due to the presence of a nearly null inner horizon, light rays may be able to travel farther before hitting the singularity. We leave this analysis for future work.

An additional issue concerns black hole mining. AMPS argued that the high-ℓ modes must also be entangled with the early radiation. Their arguments involved "mining" black holes: inserting a device such as a string that collects radiation from deep in the zone and transports it to the exterior. Brown [131] derived a number of interesting constraints on black hole mining, including the constraint that the mining equipment must be smaller than the local thermal wavelength of the Hawking radiation. Furthermore, in order to successfully extract energy and information from the black hole, the mining device must be nearly static. But clearly the presence of such a device can disrupt the entanglement between the

relevant Hawking mode that is mined and its partner behind the horizon. The process of deploying this mining device may also disrupt the entanglement between the late quanta and the early radiation. We regard mining as an interesting direction for future work. Here, we restrict our analysis to unmined black holes, where the outgoing radiation is almost exclusively in the modes with low angular momentum on the sphere.

3.2 Static black holes in higher dimensions

In this section, we consider arbitrary dimensional static black holes in spacetimes with positive, negative, and zero cosmological constant. We postpone a detailed discussion of the critical $(3+1)$-dimensional static black hole to the next section, as the geometry of the causal patch and its implications for the firewall discussion are more subtle in this case.

3.2.1 Black holes in asymptotically Minkowski spacetime

To address the question of how much of the sphere an infalling observer can see, we need to calculate the maximum angle a light ray can travel between the horizon and the singularity. For static black holes in $D > 3$ spacetime dimensions, the metric is

$$\mathrm{d}s^2 = -f(r)\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f(r)} + r^2\,\mathrm{d}\Omega_{D-2}^2 \tag{3.1}$$

where

$$f(r) = \left[1 - \left(\frac{r_-}{r}\right)^{D-3}\right]\left[1 - \left(\frac{r_+}{r}\right)^{D-3}\right] \tag{3.2}$$

with

$$r_\pm = \frac{1}{2}\left(r_s \pm \sqrt{r_s^2 - 4r_Q^2}\right) \tag{3.3}$$

Here r_+ and r_- are the radii of the outer and inner horizons, respectively; the parameter r_Q is determined by the charge of the black hole, and is given by $r_Q^2 = Q^2 G/\left(4\pi\epsilon_0 c^4\right)$. For uncharged black holes, $r_Q = 0$ and the above reduces to the Schwarzschild solution ($r_- \to 0$, $r_+ \to r_s$) with Schwarzschild radius r_s. For the Reissner-Nordstrom solution ($Q^2 > 0$), the inner horizon is believed to be unstable to perturbations, so the natural question is how far light rays can travel between the outer horizon and inner horizon in the angular direction.

Inside the outer horizon, the r and t coordinates switch roles, such that r is temporal and t is spatial. Hence to move the maximum distance along the sphere,

the ray should not move in the t direction. Therefore the null ray that travels the maximal angle satisfies

$$r^2\,\mathrm{d}\theta^2 = -\frac{\mathrm{d}r^2}{f(r)} \tag{3.4}$$

and the angle is given by

$$\Delta\theta = \int_{r_-}^{r_+} \frac{\mathrm{d}r}{r\sqrt{-f(r)}} = \frac{\pi}{D-3} \tag{3.5}$$

Thus for higher-dimensional black holes, it is impossible for a single observer to see the entire horizon, and therefore such an observer will have difficulty identifying the quantum state necessary to formulate the paradox in the global framework.[1] For the limiting case $D = 4$, there is at most just enough time for the information to be collected at a point, but no time for it to be processed. If the same property holds for all black holes, it suggests a principle: a freely falling observer cannot access the entire horizon sphere, and therefore cannot measure modes of definite angular momentum.

3.2.2 Black holes in de Sitter

One can ask about the effect of a nonzero cosmological constant on the above calculation. In this section we show that introducing a positive cosmological constant increases $\Delta\theta$, allowing the infalling observer to fit the entire infalling sphere inside her causal diamond. However, at the same time the cosmological horizon moves closer to the black hole. We find that by the time the cosmological constant is large enough to allow the infalling observer to see the entire sphere, the cosmological horizon is too small to allow for the early radiation to be collected.

3.2.2.1 3+1 dimensions

Introducing a positive cosmological constant will change the metric so that now

$$f(r) = 1 - \frac{M}{r} - \frac{r^2}{b^2} \tag{3.6}$$

where M is the black hole mass and $b^2 \equiv 3/\Lambda$. We want to know how this affects the angle computed above – will putting black holes in de Sitter space allow the infalling observer to see the entire horizon sphere?

Using again (3.5) for the angle, we get

$$\Delta\theta = \int_0^{r_H} \frac{\mathrm{d}r}{r\sqrt{-1+\frac{M}{r}+\frac{r^2}{b^2}}} = b\int_0^{r_H} \frac{\mathrm{d}r}{\sqrt{r}\sqrt{(r-r_1)(r-r_2)(r-r_3)}} \tag{3.7}$$

[1] Although existing versions of the paradox rely on a global picture, it may be possible to formulate a local version of the paradox, which might allow one to evade such concerns.

where the r_i are the three roots of the equation $f(r) = 0$. If we assume that $M < M_c \equiv \frac{2b}{3\sqrt{3}}$, these three roots are the black hole horizon r_H, the cosmological horizon r_c, and a third negative root $r_3 = -r_H - r_c$. Defining a dimensionless variable $u = r/r_H$ and rearranging gives

$$\Delta\theta = b \int_0^1 \frac{\mathrm{d}u}{\sqrt{u-u^2}\sqrt{r_c(r_c+r_H) - r_H^2(u+u^2)}} \tag{3.8}$$

Note that in the limit that the dS radius is much bigger than the black hole, $r_c \approx b$ and the second factor approaches 1, giving the flat space result.

We would like to approximate the formula for $r_H \ll r_c$. First we use that the product of the three roots is $\prod_i r_i = -Mb^2$, so

$$r_c(r_c + r_H) = \frac{Mb^2}{r_H} \tag{3.9}$$

so that

$$\Delta\theta = \int_0^1 \frac{\mathrm{d}u}{\sqrt{u-u^2}\sqrt{\frac{M}{r_H} - \frac{r_H^2}{b^2}(u+u^2)}} \tag{3.10}$$

Now, perturbatively solving (3.6) for r_H and taking the limit where $r_H \approx M$ yields

$$\frac{M}{r_H} = 1 - \frac{M^2}{b^2} + \dots \tag{3.11}$$

so that finally the integral of interest is

$$\Delta\theta \approx \int_0^1 \frac{\mathrm{d}u}{\sqrt{u-u^2}\sqrt{1 - \frac{r_H^2}{b^2}(1+u+u^2)}} \approx \pi + \frac{15\pi}{16}\frac{r_H^2}{b^2} \tag{3.12}$$

A nice way to summarize this result is to write it in terms of the entropy of the two horizons:

$$\Delta\theta = \pi + \frac{15\pi}{16}\frac{S_{BH}}{S_{dS}} \tag{3.13}$$

This shows that in principle an observer inside has access to the entire horizon sphere in some location. Now suppose that we want to collect the information at least a Planck distance from the singularity – then instead of integrating all the way to $r = 0$ we should integrate to the location $r = r_p$ where

$$\int_0^{r_P} \frac{\mathrm{d}r}{\sqrt{-f(r)}} \approx \int_0^{r_P} \mathrm{d}r\sqrt{r/M} = \frac{2}{3}\frac{r_P^{3/2}}{\sqrt{M}} \equiv l_P \tag{3.14}$$

so that $r_P = \left(\frac{3}{2}\right)^{2/3} l_P^{2/3} M^{1/3}$, giving a lower cutoff on the u integral of $u_P =$

$r_P/r_H \approx \left(\frac{3}{2}\right)^{2/3} \left(\frac{l_P}{M}\right)^{2/3}$, where we used that $r_H \approx M$. This corrects the angle by about

$$\int_0^{u_P} \frac{\mathrm{d}u}{\sqrt{u}} = 2\sqrt{u_P} \approx 12^{1/3}\pi^{1/6} S_{BH}^{-1/6} \tag{3.15}$$

So overall, the angular distance that light can travel behind the horizon of a Schwarzschild black hole in de Sitter space before reaching regions of Planckian curvature is

$$\Delta\theta = \pi + \frac{15\pi}{16}\frac{S_{BH}}{S_{dS}} - 12^{1/3}\pi^{1/6} S_{BH}^{-1/6} \tag{3.16}$$

and, at this level of analysis, we can see the entire horizon as long as

$$S_{dS} < S_{BH}^{7/6} \tag{3.17}$$

where we have neglected order 1 factors. However, the amount of information that can be stored inside the horizon in any ordinary system is [56]

$$S_R < S_{dS}^{3/4} \tag{3.18}$$

Since we need to be able to collect a number of bits comparable to the black hole entropy, $S_R \sim S_{BH}$. Therefore, the combined constraints on the size of the cosmological horizon give

$$S_{BH}^{4/3} < S_{dS} < S_{BH}^{7/6} \ . \tag{3.19}$$

But since S_{dS} is larger than 1, $S_{BH}^{4/3} > S_{BH}^{7/6}$, so the combined inequality cannot be satisfied.

Therefore, whenever the cosmological constant is large enough to allow the infalling observer to see the partner modes behind the horizon, the AMPS paradox cannot be constructed for another reason: the Hawking radiation will not fit inside the cosmological horizon.

3.2.2.2 Higher dimensions

For dS black holes in arbitrary dimensions, (3.18) becomes $S_R < S_{dS}^{(D-1)/D}$. This means that for large black holes whose radiation can be collected within the causal patch, the cosmological horizon b is much larger than the black hole horizon r_H. In this limit, the higher-order corrections to the flat space result $\Delta\theta = \frac{\pi}{D-3}$ are small, so they do not change the conclusion that the observer is missing an order 1 fraction of the sphere. Therefore, as long as S_R fits inside the cosmological horizon, the infalling observer cannot see the entire horizon sphere.

3.2.3 Black holes in anti-de Sitter

For AdS-Schwarzschild black holes, the result is very interesting. In this case we will work in general D-dimensional spacetime, where $D \geq 4$. The metric function for an AdS black hole is given by

$$f(r) = 1 + \frac{r^2}{b^2} - \frac{R_S^{D-3}}{r^{D-3}} \tag{3.20}$$

where for AdS we have $b^2 \equiv -3/\Lambda > 0$. The relevant integral is

$$\Delta\theta = \int_0^{r_H} \frac{\mathrm{d}r}{r\sqrt{-f(r)}} \,. \tag{3.21}$$

For a large black hole with horizon radius much larger than the AdS radius, it is important to ask how large the part of the horizon is that fits inside one causal patch: is it many AdS radii, or not? Taking the large black hole limit, we get

$$\Delta\theta \approx \int_0^{r_H} \frac{\mathrm{d}r}{r\sqrt{\frac{R_S^{D-3}}{r^{D-3}} - \frac{r^2}{b^2}}} \tag{3.22}$$

$$= \int_0^{r_H} \frac{\mathrm{d}r}{R_S^{\frac{D-3}{2}} r^{\frac{5-D}{2}} \sqrt{1 - \frac{r^{D-1}}{R_S^{D-3} b^2}}} \tag{3.23}$$

In the $b^2 \ll r_H^2$ limit we can use that $r_H^{D-1} \approx R_S^{D-3} b^2$ and change variables to get the dependence on parameters outside the integral, giving

$$\Delta\theta = \frac{b}{r_H} \int_0^1 \mathrm{d}u \frac{u^{\frac{D-5}{2}}}{\sqrt{1-u^{D-1}}} \sim \frac{b}{r_H} \tag{3.24}$$

where the integral can be evaluated exactly to give an $O(1)$ number for $D = 4$ which is monotonically decreasing with increasing D. This shows that for a big black hole in AdS, only a small fraction of the horizon fits inside the causal patch of an infalling observer. The corresponding physical length along the horizon that fits in one causal patch is

$$\Delta x \sim b. \tag{3.25}$$

We can conclude that an observer falling into a large AdS-Schwarzschild black hole in a D-dimensional spacetime has access to only a small part of the horizon, with physical size of order one AdS radius.

This fact may have important consequences for the AdS/CFT version of the firewall argument [125]. Existing techniques for mapping bulk to boundary encounter interesting complications when applied to fields localized to less than one

AdS radius in the near-horizon region [132, 129]. It is very intriguing that the arguments for a firewall in AdS black holes must focus on phenomena within a single AdS radius. It is precisely in this regime that the AdS/CFT duality is not well understood, and there are obstacles to reconstructing the bulk physics from the CFT.

3.3 Black holes in 3+1 dimensions

As indicated by (3.5), for (3+1)-dimensional black holes in asymptotically Minkowski spacetime, an infalling observer can see the entire sphere just as she hits the singularity. This case calls for a more detailed analysis of the causal patch.

A full analysis requires the inclusion of both interior and exterior s-wave partners, and thus we must identify a spacelike slice that crosses the horizon of the black hole. We want to know about the physics of observers who fall in to the black hole from infinity. The Gullstrand-Painlevé (GP, a.k.a. "rain-frame") coordinates are ideally suited for such purposes: the GP time variable T is the proper time along the worldline of observers falling into the black hole, starting from rest at infinity. The slices of constant T are thus orthogonal to such observers, and have the additional appeal of being spatially flat. Therefore, analyzing the entanglement in this frame is directly relevant to the question of whether an infalling observer detects any violation of the equivalence principle, as the geometrical properties of the GP coordinates precisely reflect the causal evolution along an infalling trajectory.

The GP coordinates are defined as follows [133]: Beginning with the Schwarzschild metric, define a new coordinate

$$T = t + r_s \left(2\sqrt{\frac{r}{r_s}} + \ln \left| \frac{\sqrt{\frac{r}{r_s}} - 1}{\sqrt{\frac{r}{r_s}} + 1} \right| \right) \tag{3.26}$$

called the Gullstrand-Painlevé time, with which the metric may be rewritten

$$\mathrm{d}s^2 = -f\,\mathrm{d}T^2 + 2\sqrt{\frac{r_s}{r}}\,\mathrm{d}T\mathrm{d}r + \mathrm{d}r^2 + r^2\,\mathrm{d}\Omega^2 \tag{3.27}$$

which has the appeal of being regular at $r = r_s$. See fig. 3.2 for a depiction of the constant T slices.

We want to determine the causal structure, so we need the equation for null

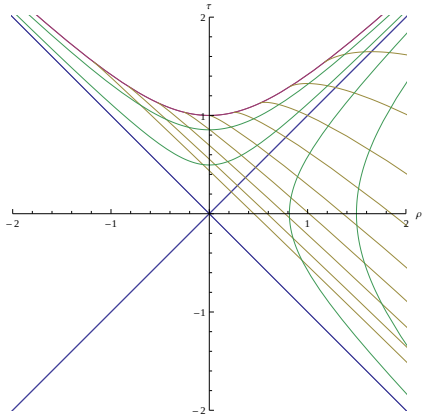


Figure 3.2: Schwarzschild black hole in Gullstrand-Painlevé coordinates, with singularity at $r = 0$ (top-most curved line, red), showing constant r slices (curved lines, green), and constant T slices (slanted lines, yellow). The vertical and horizontal axes are Kruskal-Szekeres time and radius, respectively, while the Schwarzschild radius has been set to $r = 1$.

geodesics in these coordinates. The conserved quantities for the GP metric are

$$E = f\dot{T} - \sqrt{\frac{r_s}{r}}\dot{r} \tag{3.28}$$

$$\ell = r^2\dot{\theta} \tag{3.29}$$

where the dot denotes differentiation with respect to some affine parameter. By using the second of these to replace $\dot{\theta}$ in the null geodesic equation $\mathrm{d}s^2 = 0$, and using the resulting expression for $\dot{T}$ in (3.28), one obtains a third conservation expression:

$$E^2 = \dot{r}^2 + \frac{f}{r^2}\ell^2 \tag{3.30}$$

which we may use to eliminate the affine parameter and obtain an expression for the angular distance traversed by an arbitrary null geodesic:

$$\frac{\dot{\theta}}{\dot{r}} = \frac{\mathrm{d}\theta}{\mathrm{d}r} \implies \Delta\theta = \int_0^{r'} \frac{\pm \mathrm{d}r}{\sqrt{\epsilon^2 r^4 + r^2 f}} \tag{3.31}$$

where $\epsilon \equiv E/\ell$, and the $\pm$ sign selects the polar direction in which the null ray travels. Note the fundamental difference between this expression and (3.5): our null rays are no longer constrained to move along constant Schwarzschild t-slices in the black hole interior.

Similarly, we obtain an expression for the Gullstrand-Painlevé time difference

corresponding to (3.31):

$$\frac{\dot{T}}{\dot{r}} = \frac{\mathrm{d}T}{\mathrm{d}r} \implies \Delta T = \int_0^{r'} \frac{1}{f}\left(\sqrt{\frac{r_s}{r}} \pm \frac{\epsilon r}{\sqrt{\epsilon^2 r^2 - f}}\right)\mathrm{d}r \tag{3.32}$$

Henceforth we will absorb the $\pm$ sign in our expressions for ΔT into ϵ by allowing the latter to take negative values.

Now we would like to determine which part of the constant time surface fits within a single causal patch. We fix a single observer, who determines the causal patch, just above the singularity at Schwarzschild time $t = 0$, at the north pole of the sphere, $\theta = 0$. This observer will collect measurements transmitted to her from an infalling distributed measuring device – say, a ring of probes spread around the horizon. At some specified GP time T, the probes will perform a measurement of the interior s-wave and transmit this information to the observer to be collected for analysis. The intersection of the observer's past light cone with this T-slice determines the causal patch under consideration (see fig. 3.3).

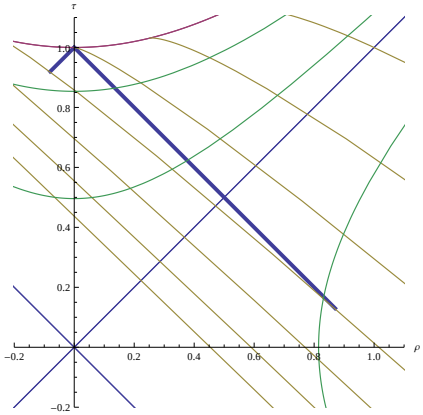
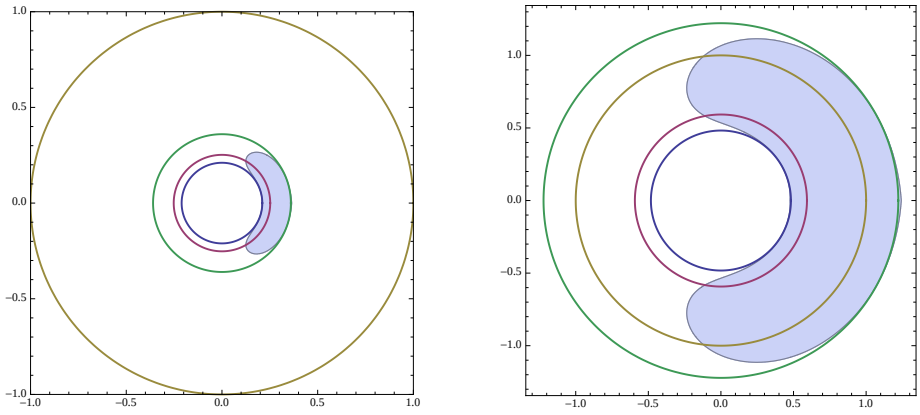


Figure 3.3: Past light cone (bold blue) of an observer hovering just above the singularity at $(t, r) \approx (0, 0)$. The interior and exterior radial null rays (left and right cone sides, respectively) intersect the T-slice at $r_{\epsilon\to\infty}$, $r_{\epsilon\to-\infty}$, respectively. The geometry of the patch is determined by evaluating $\Delta\theta$ along the T-slice for the null rays between these two radial extremes.

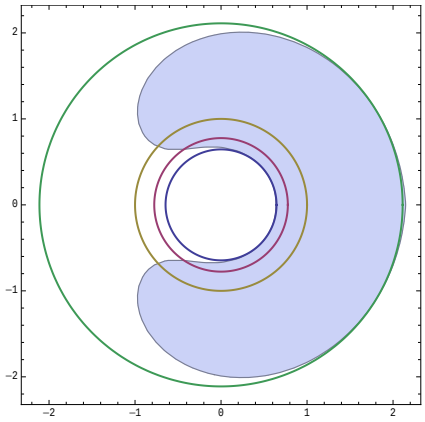
The Schwarzschild time of the observer ($t = 0$) intersects this T-slice at $r = r_0$. We wish to know the geometry of this causal patch as a function of the choice of T (equivalent to considering observers who fall in at different Schwarzschild times), which requires numerically evaluating (3.31) along the T-slice.

To perform this evaluation requires specification of ϵ. For each point in the causal patch, there intersects in principle an infinite number of possible null rays, parameterized by ϵ, only one of which will have the correct trajectory to be collected by the observer. Furthermore, this value of ϵ is dependent on the upper

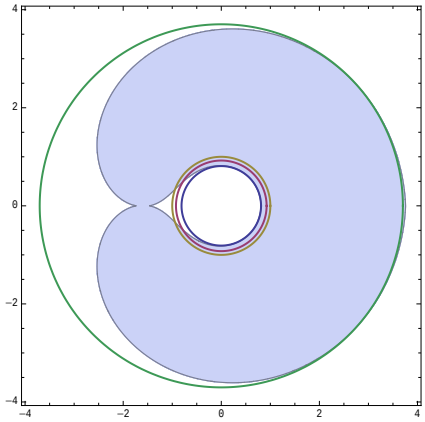


(a) $|\Delta T| = 0.1$

(b) $|\Delta T| = 0.5$



(c) $|\Delta T| = 1$



(d) $|\Delta T| = 2$

Figure 3.4: Causal patch geometry for several values of ΔT. The shaded region depicts the portion of the spacelike T-slice [as $r(\theta)$] visible to the observer. The concentric rings show the horizon $r_s = 1$ (yellow), $r_{\epsilon\to-\infty}$ (green), r_0 (red), and $r_{\epsilon\to\infty}$ (blue). (Note that the axes are rescaled between images). Increasing $|\Delta T|$ corresponds to selecting a T-slice closer to the past horizon in fig. 3.3.

limit of integration, i.e. on the r-position along the T-slice: $\epsilon = 0$ corresponds to $\ell \to \infty$, for which (3.31) reduces to (3.5), while $\epsilon \to \pm\infty$ corresponds to radial null rays with $\ell = 0$, whose intersections with the T-slice give the minimal (at $r = r_{\epsilon\to\infty}$) and maximal ($r = r_{\epsilon\to-\infty}$) radii of the casual patch.

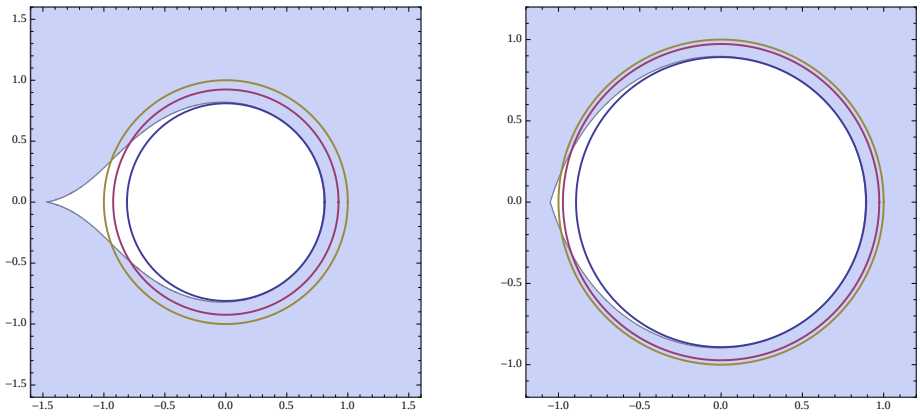
The distance between the observer and our chosen T-slice, denoted T_*, is given by $\Delta T = T_* - T(r = 0, t = 0) = T_*$. Thus we may numerically obtain the values of ϵ for radii along $T = T_*$ by finding the root of $T_* - \Delta T(\epsilon)$, where $\Delta T(\epsilon)$ is given by (3.32), with ϵ as the free parameter. With these values of ϵ in hand, we may

proceed to the numerical evaluation of (3.31). Results are shown in fig. 3.4.

As $|\Delta T|$ is increased, the observer sees less of the interior and more of the exterior of the black hole. This is consistent with an inspection of the geometry in fig. 3.3: as T_* becomes more and more negative, $r_{\epsilon\to\infty}$ approaches the horizon radius, while $r_{\epsilon\to-\infty}$ increases without bound; conversely, as T_* approaches $t = 0$, both $r_{\epsilon\to\infty}$ and $r_{\epsilon\to-\infty}$ shrink, allowing the observer to see more of the black hole interior at the cost of her external view.

In order to try to fit all the ingredients necessary for the firewall paradox inside a single causal patch, we wish to examine a causal patch that contains both an outgoing Hawking quantum and its interior partner mode. Hence for our purposes, the regime of interest is when $|\Delta T|$ becomes large, which allows the observer to maximize both her internal and external angular visibility, and hence affords the best chance of measuring both an outgoing s-wave and its entangled interior partner. However, as pointed out in [123], the wavelength of the interior mode may pose some difficulty to fitting it inside such a patch. In particular, because of the aforementioned trade-off between angular and radial depth visibility, it may not be possible to keep the wavelength of the interior mode above the Planck scale while effecting sufficient angular resolution.

For $|\Delta T|$ sufficiently large to close the exterior visibility region, the exclusion region resembles a raindrop (see fig. 3.5). In the limit of large $|\Delta T|$, $r_{\epsilon\to\infty}$ approaches r_s, and the radial depth available to interior s-wave modes vanishes. Since the energy is $\sim \lambda^{-1}$, this places a lower limit on the energy of the measurable modes, namely $E \gtrsim (r_s - r_{\epsilon\to\infty})^{-1}$.



(a) $|\Delta T| = 2$ (b) $|\Delta T| = 3$

Figure 3.5: Rain in the rain frame: close-up of exclusion regions. The pointed end of the raindrop diminishes, and the droplet approaches a circular region with radius $r \to r_s$, in the limit of large ΔT.

Although an analytical expression for the droplet geometry is not available, it is possible to obtain an approximation in the large $|\Delta T|$ limit, where the droplet begins to look like that in fig. 3.5 for $|\Delta T| = 3$. By approximating (3.31) and (3.32) in the small-ϵ limit, we find

$$\Delta\theta \approx \pi - 2\sqrt{1 - r + \epsilon^2} \tag{3.33}$$

$$\Delta T \approx 2\sqrt{r} + 2\ln\left(\frac{1-\sqrt{r}}{\epsilon + \sqrt{1 - r + \epsilon^2}}\right) + 2\ln\left(\epsilon + \sqrt{1-\epsilon^2}\right) \tag{3.34}$$

The derivation of these expressions is detailed in the Appendix. Note that $\Delta T < 0$ (consistent with an infalling observer, since we integrated outwards from the singularity $r = 0$).

These results can be plotted against the numerical exclusion region (i.e. the droplet) by solving (3.34) for ϵ, and substituting the result into (3.33) to obtain an expression for $\Delta\theta(r)$. We find

$$\Delta\theta \approx \pi - 2\sqrt{\frac{\left(-1 + r + \sqrt{r}\sinh\left(\Delta T/2 - \sqrt{r}\right) + \cosh\left(\Delta T/2 - \sqrt{r}\right)\right)^2}{2 - r - 2\sqrt{r}\sinh\left(\Delta T/2 - \sqrt{r}\right) - 2\cosh\left(\Delta T/2 - \sqrt{r}\right)}} \tag{3.35}$$

Two example cases which serve to demonstrate the validity of this result are shown in fig. 3.6.

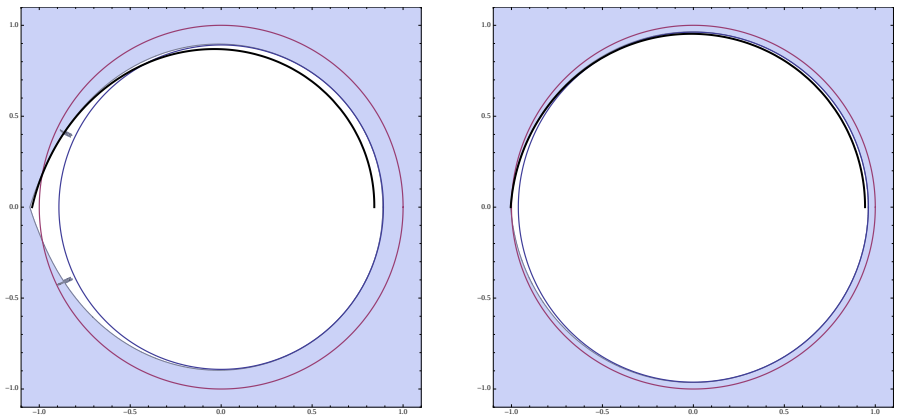


Figure 3.6: $r(\theta)$ (thick black curve), determined by (3.35), plotted against the droplet for $|\Delta T| = 3$ (left) and 5 (right), showing improvement as $|\Delta T|$ is increased. The concentric circles are r_s (outer ring, red) and $r_{\epsilon\to\infty}$ (inner ring, blue). The tick marks in the left image are merely due to a rendering glitch.

We may obtain a more aesthetically pleasing approximation to (3.34) by ex-

panding in the near-horizon region. We find (see appendix)

$$\boxed{\pi - \Delta\theta \approx \sqrt{h} + \frac{1-r}{\sqrt{h}}} \tag{3.36}$$

where

$$\sqrt{h} \equiv 2e^{\Delta T/2-1} \tag{3.37}$$

our choice of the notation "$\sqrt{h}$" will become clear shortly. The accuracy of (3.36) is comparable to (3.35) near the horizon (and hence also on the tip for sufficiently large $|\Delta T|$), but cannot be used along the rest of the droplet body.

At the horizon itself ($r = 1$), the second term in (3.36) vanishes and we obtain an approximation for the angular width of the droplet tip at the Schwarzschild radius as a function of GP time:

$$\pi - \Delta\theta \approx \sqrt{h} = 2e^{\Delta T/2-1} \tag{3.38}$$

Two other droplet parameters are of interest: the height of the tip above the horizon, and the depth of the antipodal point within. The former is defined by $\Delta\theta = \pi$; hence $\epsilon = \sqrt{r-1}$ and (3.34) becomes

$$\Delta T \approx 2\sqrt{r} + 2\log\left(\frac{1-\sqrt{r}}{\sqrt{r-1}}\right) \tag{3.39}$$

where we have discarded the negligible third term. Defining the height of the tip $h \equiv r - 1 > 0$, and expanding around $h = 0$, we find

$$\Delta T \approx 2 + \frac{h}{2} - \ln(4) + \ln(h) \tag{3.40}$$

We may then drop the term linear in h relative to the log, and solve:

$$h \approx 4e^{\Delta T-2} \tag{3.41}$$

cf. (3.37). To obtain a similar expression for the depth of the antipodal point requires a formula valid in the limit $\epsilon \to \infty$. From (3.61) it follows that

$$\lim_{\epsilon\to\infty} \Delta T = 2\sqrt{r} + r + 2\ln\left(1-\sqrt{r}\right) \tag{3.42}$$

Defining the depth $d \equiv 1 - r > 0$ and expanding, we find

$$\Delta T \approx 3 - \frac{3}{2}d - \ln(4) + 2\ln(d) \tag{3.43}$$

As before, we drop the linear d term and solve:

$$d \approx 2e^{(\Delta T-3)/2} = e^{-1/2}\sqrt{h} \tag{3.44}$$

We summarize our results for the droplet parameters in fig. 3.7.

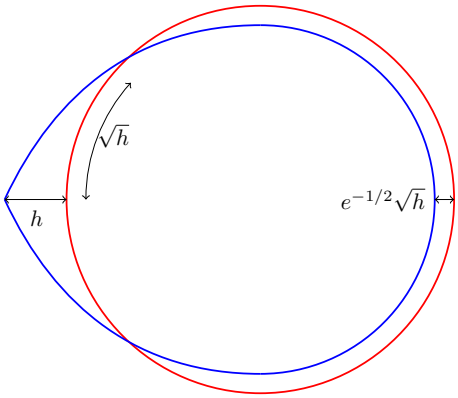


Figure 3.7: Sketch of a heavily distorted droplet (blue) against the horizon $r_s = 1$ (red) with parameters of interest labeled: height $h = 4e^{\Delta T-2}$, width $\sqrt{h}$, and depth $e^{-1/2}\sqrt{h}$. Note that distances are not to scale, although the height is indeed less than the width for $h << 1$ ($|\Delta T|$ large).

3.4 Entropy and information

Having obtained a geometric picture of the infalling observer's causal patch in the case of Schwarzschild black holes in $3+1$ dimensions, we now wish to ask what this implies for the AMPS paradox. We appear to have a trade-off between the energy scale of the measurable modes and the angular resolution; i.e., one has large angular visibility only for interior modes that are highly radially localized near the horizon (see fig. 3.4). What can we then conclude about the entanglement of the partner modes?

For an infalling observer, the entanglement entropy across the horizon may be thought of as being organized into localized Bell pairs, each of which contains a single bit of entanglement entropy [134]. Consider the total number of bits within the droplet $m = \theta^2_{\text{missing}}/\lambda^2$, out of a total $N = 1/\lambda^2$ bits distributed over the entire circle. The wavelength of measurable quanta is limited by the distance between the droplet and the horizon, which for partner modes must be equal inside and

outside the black hole. Hence we have $\lambda \leq h$ with $\Delta\theta_{\text{missing}} \sim \sqrt{h}$, and therefore

$$m = \frac{(\Delta\theta_{\text{missing}})^2}{\lambda}\left(\frac{1}{\lambda}\right) \sim \frac{h}{\lambda}\sqrt{N} \implies m \gtrsim \sqrt{N} \tag{3.45}$$

where $h/\lambda \geq 1$. Thus we find that a single observer is always missing at least about $\sqrt{N}$ out of N bits. Insofar as N is proportional to λ^{-2}, only high-energy modes stand a chance of reducing the missing fraction to the point where collection of sufficient information is possible. Another obvious though important consequence is that, since one cannot speak of trans-Planckian modes in the absence of a full theory of quantum gravity, m will *never* be zero: even the most determined observer is missing at least one bit.

We may also compute the entropy associated with this missing area. Computing the solid angle in the small h approximation, we find

$$A_{\text{missing}} \approx \pi r^2 h \implies \tag{3.46}$$

$$S_{\text{missing}} = \frac{A_{\text{missing}}}{4l_P^2} \approx \frac{\pi r^2}{4l_P^2}h \tag{3.47}$$

where we have taken $k_B = 1$. Via (3.41), this can be written

$$S_{\text{missing}} \approx \frac{\pi r^2}{l_P^2}e^{\Delta T - 2} \tag{3.48}$$

Thus, an observer who wishes to measure a mode with wavelength of order $\lambda \sim h \approx 4e^{\Delta T - 2}$ does so at an entropy cost given by (3.48), which we may think of as the entropy associated with the missing $\sqrt{N}$ bits.

It is interesting to note the consequences for Bousso's double-purity argument[51] in the context of the casual patch considerations above. Essentially, the standard argument is as follows: let X be the interior Hawking quanta, Y the outgoing partner mode, and Z the early Hawking radiation. Then the strong subadditivity of entanglement entropy

$$S_{XYZ} + S_Y \leq S_{XY} + S_{YZ} \tag{3.49}$$

prevents both XY (the infalling vacuum) and YZ (the final out-state) from being pure. That is, $\forall Z : X \cap Z = \varnothing$,

$$S_{XY} = 0 \wedge S_Y > 0 \implies S_{YZ} \neq 0 \tag{3.50}$$

$$S_{YZ} = 0 \implies \nexists X : (S_{XY} = 0 \wedge S_Y > 0) \tag{3.51}$$

Alternatively, as shown in [135], (3.49) is equivalent to

$$S_X + S_Z \leq S_{XY} + S_{YZ} \tag{3.52}$$

from which Bousso's conclusion follows immediately.

However, one can only claim the validity of (3.49) if one has access to the global field theory. In contrast, here one only has access to some subset of the degrees of freedom, and one can obtain the corresponding entropy inequality within a single causal patch as follows. Define $\tilde{X} \subset X$ as the portion that the infalling observer can see, i.e. $X \equiv \tilde{X} \cup D$ where D is the portion obscured by the droplet at the horizon. Similarly for $\tilde{Y}$. We must also consider that only a small portion of the near-horizon radiation Y ($\tilde{Y}$, D) will evolve though the angular momentum barrier to contribute to the late Hawking radiation. Call this subset Y_R ($\tilde{Y}_R$, D_R). With these definitions in hand, strong subadditivity can only be formulated in the infalling patch (for the external observer cannot see any of X) as

$$S_{\tilde{X}\tilde{Y}Z} + S_{\tilde{Y}} \leq S_{\tilde{X}\tilde{Y}} + S_{\tilde{Y}Z} \tag{3.53}$$

and the desired double-purity is really

$$S_{\tilde{X}\tilde{Y}} = 0 \wedge S_{Y_R Z} = 0 \tag{3.54}$$

with $S_{\tilde{Y}} > 0$ and $S_{Y_R} > 0$. In contrast to the original argument above, it is by no means obvious that the both the infalling vacuum $\tilde{X}\tilde{Y}$ and the final out-state $Y_R Z$ cannot be pure. That is, when the limitations of the causal patch geometry are taken into account, it may still be possible for both the infalling and external observers to see a pure state without violating the monogamy of entanglement.

An outstanding question is precisely how much of the horizon area—equivalently, how many bits m—the infalling observer can afford to lose before measurement of the ingoing Hawking mode becomes impossible. Questions of reconstructing information from some subset of bits are considered in quantum information theory in the context of (k, n) threshold schemes [136], in which a quantum "secret" is divided into n shares such that any $k \leq n$ of those shares can be used to reconstruct the original secret, but any $k - 1$ or fewer cannot. The authors of [136] demonstrated that the only general constraint on such threshold schemes is due to monogamy: one must have $n < 2k$ or else the quantum no-cloning theorem is violated.

Consider, as above, an s-wave immediately behind the horizon with an outgoing partner mode directly outside, with the entanglement information distributed in N localized Bell pairs. Further suppose that the information necessary to reconstruct the entangled state is encoded in a (k, n) threshold scheme ($n = N$). The question at hand is then: what is the value of k needed to reconstruct the state?

If reconstruction requires the full N bits ($k = n = N$), then our results imply that doing so is impossible, since one reaches the Planck scale in wavelength before the missing number of bits $m \to 0$. Conversely, if the information can be retrieved from some sufficiently large fraction $\frac{N-m}{N}$, then the infalling observer may still be able to reconstruct high-energy modes. In the absence of a precise statement about how black holes encode their secrets, the general bounds $k \leq n < 2k$ are not sufficiently strict to rule out the possibility that an infalling observer could reconstruct the state despite missing a large number of bits.

However, this still involves a trade-off between the energy scale of the measurable modes and the angular occlusion. It may be that one can only effect sufficient angular resolution for modes whose energy exceeds some critical value, $\lambda_{\text{crit}}^{-1}$, in which case the $\mathcal{O}(1)$ corrections to high-energy modes purported by AMPS—in contradiction to BHC—would only be detectable for very high-energy modes indeed. More work is needed to determine precisely how small the fraction m/N need be.

3.5 Conclusions

We have shown that for static black holes in 3+1 and higher dimensions, there does not exist a causal patch that contains all the ingredients necessary to construct the firewall paradox at the level of s-wave Hawking quanta. A possible exception to this principle arises when considering the Schwarzschild black hole in $3 + 1$ dimensions, and we presented a detailed analysis of the infalling geometry for this case. Our results indicate that the infalling observer is always missing some finite amount of information about the s-wave. Though it remains to show precisely how much angular resolution the observer can afford to lose before reconstruction of the partner mode becomes impossible in principle, our analysis suggests that it is at best difficult in practice.

We focused on the situation for s-waves, as this version of the firewall paradox is the simplest and most robust in our view. Although it would be interesting to consider the consequences for high-ℓ modes, this requires a more thorough understanding of the degree to which the mining apparatus disrupts the entanglement of the quantum state. A more detailed analysis of the localization of partner modes may shed more light on this direction, but we leave this for future work.

We conclude that for static black holes in $3+1$ and higher dimensions, BHC is sufficient to evade at least the simplest version of the firewall paradox. Schwarzschild black holes in $3 + 1$ dimensions nearly allow the paradox to arise within one causal patch, and it is possible that the firewall arguments in that case can be improved, violating complementarity. For rotating black holes and discussions of high-ℓ modes using mining, more work is needed.

3.A Approximations

In this appendix we derive the approximate expressions for $\Delta\theta$ (3.33), ΔT (3.34), and $\Delta\theta\,(r, \Delta T)$ (3.36). We begin with eqns. (3.31) and (3.32) for the Schwarzschild metric:

$$\Delta\theta = \int_0^{r_f} \frac{\pm \mathrm{d}r}{\sqrt{\epsilon^2 r^4 + r(r_s - r)}}$$
$$\Delta T = \int_0^{r_f} \frac{\epsilon r^3 + \sqrt{r_s r}\sqrt{\epsilon^2 r^4 + r(r_s - r)}}{(r_s - r)\sqrt{\epsilon^2 r^4 + r(r_s - r)}}\mathrm{d}r$$

where $\epsilon \in (-\infty, 0]$ for $r \geq r_0$, and $\epsilon \in [0, \infty)$ for $r \leq r_0$, with r_0 denoting the angular limit where $\ell \to \infty \implies \epsilon \to 0$. Note that $r_0 < r_s < r_f$, but both r_0 and r_f approach r_s asymptotically as $|\Delta T|$ increases. Note that in our convention, $\Delta T < 0$.

Beginning with the θ integral: for simplicity of notation, consider only the positive case (the negative is merely a mirror image about the x-axis). Observe that

$$\lim_{\epsilon \to \infty} \Delta\theta = 0 \tag{3.55}$$

and hence a suitable approximation can be obtained by evaluating the integral for small ϵ.[2] Now suppose there exists an r' such that

$$\epsilon^2 r'^4 << r'(r_s - r') \tag{3.56}$$
$$r_s - r' << r_s \tag{3.57}$$

Intuitively, the first of these says that the distance to the horizon dominates over the contribution from ϵ, while the second says that we are still sufficiently close to the horizon that ϵ has not yet become large. (These conditions are easily seen to be consistent with the small ϵ regime, as they can be combined to yield $\epsilon^2 r'^3 << r_s$, which for the near-horizon region reduces to $\epsilon << r_s^{-1}$.)

Thus we can break the integral into two regions:

$$\Delta\theta \approx \int_0^{r'} \frac{\mathrm{d}r}{\sqrt{r(r_s - r)}} + \int_{r'}^{r_f} \frac{\mathrm{d}r}{\sqrt{\epsilon^2 r_s^4 + r_s(r_s - r)}} \tag{3.58}$$

where in the second term we have expanded to first order in $\delta = r_s - r << 1$.

[2] This is to be expected, since $\epsilon \to \infty$ corresponds to the radial limit, in which the angular distance vanishes.

Evaluating (3.58) yields

$$\Delta\theta \approx 2\arcsin\left(\sqrt{\frac{r'}{r_s}}\right) - 2\sqrt{1-\frac{r_f}{r_s}+\epsilon^2 r_s^2} + 2\sqrt{1-\frac{r'}{r_s}+\epsilon^2 r_s^2} \tag{3.59}$$

It now remains to eliminate the r' parameter. In the limit that $r' \to r_s$, $\arcsin\left(\sqrt{\frac{r'}{r_s}}\right) \approx \frac{\pi}{2} - \sqrt{1-\frac{r'}{r_s}}$, and this second term cancels with the last term in (3.59) after dropping the negligible ϵ^2 contribution. Hence, setting r_s and dropping the subscript on r_f,

$$\Delta\theta \approx \pi - 2\sqrt{1-r+\epsilon^2} \tag{3.60}$$

which is (3.33).

Turning now to the T integral, we first observe that

$$\lim_{\epsilon\to\infty} \Delta T = \int \frac{\sqrt{r}}{\sqrt{r}-1}\mathrm{d}r \tag{3.61}$$

and thus one would not expect the same small ϵ approximation to suffice for the entire droplet. However, it so happens that the region of large ϵ is confined relatively close to—that is, has a small angular deviation from—the base of the droplet where $\epsilon \to \infty$, and as we shall see, the small ϵ approximation is perfectly adequate elsewhere.

Performing a similar split as in (3.58) yields

$$\Delta T \approx \int_0^{r'} \frac{\sqrt{r_s r}}{r-r_s}\mathrm{d}r + \int_{r'}^{r_f} \frac{\epsilon r_s^3 + r_s\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}{(r-r_s)\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}\mathrm{d}r \tag{3.62}$$

where the second term has again been expanded to first order in the near-horizon region. Rather than integrate immediately however, we first analytically eliminate the r' parameter by extending the integration regions and subtracting compensating terms:

$$\begin{aligned}\Delta T \approx &\int_0^{r_f} \frac{\sqrt{r_s r}}{r-r_s}\mathrm{d}r + \int_0^{r_f} \frac{\epsilon r_s^3 + r_s\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}{(r-r_s)\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}\mathrm{d}r \\ &- \int_{r'}^{r_f} \frac{\sqrt{r_s r}}{r-r_s}\mathrm{d}r - \int_0^{r'} \frac{\epsilon r_s^3 + r_s\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}{(r-r_s)\sqrt{\epsilon^2 r_s^4 + r_s(r_s-r)}}\mathrm{d}r\end{aligned}$$

Note that the third term is now entirely in the region where $r_s \sim r$, while the

fourth is in the regime where $r_s - r$ dominates over ϵ. Hence:

$$\begin{aligned}\Delta T &\approx \int_0^{r_f} \frac{\sqrt{r_s r}}{r - r_s}\mathrm{d}r + \int_0^{r_f} \frac{\epsilon r_s^3 + r_s\sqrt{\epsilon^2 r_s^4 + r_s(r_s - r)}}{(r - r_s)\sqrt{\epsilon^2 r_s^4 + r_s(r_s - r)}}\mathrm{d}r \\ &\quad - \int_{r'}^{r_f} \frac{r_s}{r - r_s}\mathrm{d}r - \int_0^{r'} \frac{r_s}{r - r_s}\mathrm{d}r \\ &= \int_0^{r_f} \frac{\sqrt{r_s r}}{r - r_s}\mathrm{d}r + \int_0^{r_f} \frac{\epsilon r_s^3 + r_s\sqrt{\epsilon^2 r_s^4 + r_s(r_s - r)}}{(r - r_s)\sqrt{\epsilon^2 r_s^4 + r_s(r_s - r)}}\mathrm{d}r - \int_0^{r_f} \frac{r_s}{r - r_s}\mathrm{d}r\end{aligned}$$

The advantage of this seemingly roundabout exercise is that now each of the above terms can be integrated indefinitely, and the result rearranged prior to plugging in limits in order to obtain a finite, real result. Setting r_s to 1 for simplicity, we find

$$\begin{aligned}\Delta T &\approx 2\sqrt{r} - 2\,\mathrm{arctanh}(\sqrt{r}) - 2\,\mathrm{arctanh}\left(\frac{1}{\epsilon}\sqrt{1 - r + \epsilon^2}\right) + \ln(1 - r) - \ln(r - 1) \\ &= 2\sqrt{r} - \ln\left(\frac{1 + \sqrt{r}}{1 - \sqrt{r}}\right) - \ln\left(\frac{\epsilon + \sqrt{1 - r + \epsilon^2}}{\epsilon - \sqrt{1 - r + \epsilon^2}}\right) + \ln(1 - r) - \ln(r - 1) \\ &= 2\sqrt{r} - \ln\left(\frac{1 + \sqrt{r}}{1 - \sqrt{r}}\right) - 2\ln\left(\epsilon + \sqrt{1 - r + \epsilon^2}\right) \\ &\quad + \ln\left((\epsilon - \sqrt{1 - r + \epsilon^2})(\epsilon + \sqrt{1 - r + \epsilon^2})\right) + \ln(1 - r) - \ln(r - 1) \\ &= 2\sqrt{r} - \ln\left(\frac{1 + \sqrt{r}}{1 - \sqrt{r}}\right) - 2\ln(\epsilon + \sqrt{1 - r + \epsilon^2}) + \ln\left((1 + \sqrt{r})(1 - \sqrt{r})\right) \\ &= 2\sqrt{r} + 2\ln(1 - \sqrt{r}) - 2\ln(\epsilon + \sqrt{1 - r + \epsilon^2}) \\ &= 2\sqrt{r} + 2\ln\left(\frac{1 - \sqrt{r}}{\epsilon + \sqrt{1 - r + \epsilon^2}}\right)\end{aligned}$$

Thus, plugging in limits of integration (again dropping the subscript on r_f),

$$\Delta T = 2\sqrt{r} + 2\ln\left(\frac{1 - \sqrt{r}}{\epsilon + \sqrt{1 - r + \epsilon^2}}\right) + 2\ln\left(\epsilon + \sqrt{1 - \epsilon^2}\right) \tag{3.63}$$

which is (3.34).

Eliminating ϵ in order to combine (3.60) and (3.63) leads to the full expression for the droplet body given in the main text, (3.35). Here we obtain a simpler expression, which is still reasonably accurate away from the droplet base where ϵ becomes large. Defining $x \equiv 1 - r$ in the small-ϵ regime, we have, to first order,

$$\Delta T \approx 2 - x + 2\ln\left(\frac{x}{2(\epsilon+\sqrt{x+\epsilon^2})}\right) + 2\epsilon \implies$$
$$\frac{\Delta T}{2} - 1 \approx \ln\left(\frac{x}{2(\epsilon+\sqrt{x+\epsilon^2})}\right) \tag{3.64}$$

where we have dropped the linear $x << 1$ and $\epsilon << 1$ terms. Then, from (3.33), we have

$$2\sqrt{x+\epsilon^2} \approx \pi - \Delta\theta \equiv \alpha \implies \epsilon \approx \sqrt{\alpha^2/4 - x} \tag{3.65}$$

which we substitute into (3.64):

$$xe^{-(\Delta T/2-1)} \approx \sqrt{\alpha^2 - 4x} + \alpha \implies 2\alpha e^{-(\Delta T/2-1)} \approx xe^{-2(\Delta T/2-1)} + 4 \implies$$
$$\alpha \approx \frac{x}{2}e^{-(\Delta T/2-1)} + 2e^{\Delta T/2-1}$$

Thus we obtain

$$\pi - \Delta\theta \approx \sqrt{h} + \frac{x}{\sqrt{h}} \tag{3.66}$$

where

$$x \equiv 1 - r, \qquad \sqrt{h} \equiv 2e^{\Delta T/2-1} \tag{3.67}$$

which is (3.36). As shown in the main text, the notation "$\sqrt{h}$" was chosen so as to write the various droplet parameters in terms of the height h of the tip above the horizon.

4 Holographic shadows

This chapter is based on [2].

In this chapter, we put aside firewalls in order to investigate the issue of recovering information in a different context, namely bulk reconstruction in AdS/CFT. Specifically, we study several holographic probes that relate information about the bulk spacetime to CFT data. The best-known example is the relation between minimal surfaces in the bulk and entanglement entropy of a subregion in the CFT. Building on earlier work, we identify "shadows" in the bulk: regions that are not illuminated by any of the bulk probes we consider, in the sense that the bulk surfaces do not pass through these regions. We quantify the size of the shadow in the near horizon region of a black hole and in the vicinity of a sufficiently dense star. The existence of shadows motivates further study of the bulk-boundary dictionary in order to identify CFT quantities that encode information about the shadow regions in the bulk. We speculate on the interpretation of our results from a dual field theory perspective.

4.1 Introduction

Despite many remarkable advances in our understanding of the AdS/CFT correspondence, significant obstacles remain in reconstructing local bulk physics from the CFT. These obstacles prevent us from answering elementary questions of enormous importance for our understanding of quantum gravity, such as whether an observer falling into an AdS-Schwarzschild black hole encounters a "firewall" [**?**, 125, 126, 137]. A particularly important and difficult question is how to extract physics on scales short compared to the AdS radius near the black hole horizon.

One powerful tool in reconstructing bulk physics comes from the Ryu-Takayanagi

proposal [97]. It directly links the area of minimal bulk surfaces to the entanglement entropy of spatial regions in the boundary field theory, and thereby provides a quantitative relationship between entanglement in holographic CFTs and spacetime geometry [97, 138, 60, 139, 85, 140, 98].[1] In some cases, it has been shown that the entanglement entropy data alone is sufficient to completely determine the bulk solution [113]. This supports the ambitious claim that the spacetime is emergent and can be reconstructed from the boundary CFT [87, 60, 86, 139, 85, 142].

However, there is an obstacle to performing this type of reconstruction in more general geometries. In general, the bulk contains "shadows", or regions that are skipped over by the minimal surfaces. One reason for the existence of shadows is phase transition behavior: a given boundary region may have multiple bulk surfaces that are all local minima of the area. But the global minimum, with which the CFT quantity is associated, may switch from one branch of local minima to another, and thus the boundary dual skips over some bulk region [143, 144]. In asymptotically global AdS spacetime, it is possible that a region of the bulk is always skipped over no matter which boundary regions we choose. In [145], such regions into which no minimal area surface can probe were dubbed "entanglement shadows." When shadows exist, it is obvious that the boundary data in question does not provide sufficient information to construct a unique bulk geometry.[2]

The most obvious way to overcome this obstacle is to find a better probe, i.e., one that reaches deeper into the bulk and penetrates the shadow. With this situation in mind, we present a generalized framework for determining the "holographic shadows" associated with extremal geometric objects.

Predictably, the interior of a static black hole lies within the entanglement shadow [146, 147], and is likely also part of the holographic shadow for any similar probe. Somewhat more surprising is the fact that, at least in all cases of which we are aware, holographic shadows always extend beyond the horizon. Furthermore, they are determined by the phase transition behaviour mentioned above, and are *not* directly related to the presence of the black hole. Indeed, holographic probes can suffer shadows even in globally regular geometries, and we emphasize this by presenting an explicit example of an entanglement shadow in the case of a star in AdS_{2+1}.

[1] Note that this method of spacetime reconstruction is quite different from, and more ambitious than, programs involving the entire bulk wedge (e.g., the smearing functions of [102]), which presume the existence of a background geometry. This distinction must be kept in mind to avoid seemingly contradictory conclusions (cf. [141]). See section 4.7.1 for further discussion.

[2] Together with other input, such as the full set of bulk equations of motion, such reconstruction might be possible. This is an interesting, though ambitious, future direction, as a full reconstruction must work at the non-perturbative level. At the perturbative level, one can assume that a background bulk geometry exists, and it seems reasonable to assert that the boundary theory knows about some bulk region as long as it is contained within the minimal surface[141]. In this scenario, crossing over a shadow results in an abrupt increase in the amount of information accessible to the boundary region.

	Minimal Area	Wilson Loop	Causal
$d = 2$, $r_H \ll \ell_{\text{AdS}}$	$\sim \ell_{\text{AdS}}$	$\sim \ell_{\text{AdS}}$	$\sim \ell_{\text{AdS}}$
$d = 2$, $r_H \gg \ell_{\text{AdS}}$	$\sim e^{-\#r_H/\ell_{\text{AdS}}}$	$\sim r_H$	$\sim e^{-\#r_H/\ell_{\text{AdS}}}$
$d > 2$, $r_H \ll \ell_{\text{AdS}}$	$\sim r_H$	$\sim (r_H \ell_{\text{AdS}})^{1/2}$	$\sim e^{-\#\ell_{\text{AdS}}/r_H}$
$d > 2$, $r_H \gg \ell_{\text{AdS}}$	$\sim e^{-\#r_H/\ell_{\text{AdS}}}$	$\sim r_H$	$\sim e^{-\#r_H/\ell_{\text{AdS}}}$

Table 4.1: Shadow summary for various probes of AdS-Schwarzschild; d is the spatial dimension. The value listed is the distance from the black hole hole horizon r_H. The # symbol denotes an order one constant, which may depend on the spatial dimension; both this and the overall proportionality are determined explicitly in the main text.

In the case of singular spacetimes, the question "how close to the black hole horizon can we probe?" is both interesting and important. Thus, building on earlier work [148], we conduct a comparison of three distinct holographic probes in AdS-Schwarzschild geometries: minimal area surfaces, static Wilson loops, and causal information surfaces [113, 149, 150]. Our results are summarized in table 4.1. We find that in several cases, the causal information surfaces probe deepest into the bulk. In particular, for small black holes in higher dimensions, causal information surfaces get exponentially close to the horizon, while other probes remain of order one horizon distance away.

Our finding would seem to conflict the common impression that minimal area surfaces reach deeper than causal surfaces. To resolve this apparent conflict, it is important to distinguish local vs. global comparisons. The minimal area surface associated with a fixed boundary region does indeed reach deeper into the bulk than the corresponding causal surface [113]. However, the shadow is defined by the smallest radius accessible by *any* bulk probe, i.e., the maximum depth among all possible boundary regions. In particular, the causal surfaces are not subject to the aforementioned switchover effect, which allows them to gain the advantage over minimal area surfaces despite being locally worse. It is in this second, global sense that we mean a given surface is "better" or reaches deeper, since having a smaller shadow is the more relevant standard for the purpose of holographic reconstruction.

There is an additional, slightly more subtle consideration that may be important for bulk reconstruction. To retrieve complete information about a given bulk region, we might require a probe to not only reach every bulk point, but to do so with every possible orientation. Indeed, this is precisely the requirement of the hole-ographic construction of [142]. Thus, we also identify "partial shadows"—regions of the bulk which are accessible by a given probe, but with only partial coverage of the tangent space. We have only conducted a preliminary investigation of partial shadows, but we regard them as a potentially interesting aspect for future work.

Finally, we should emphasize that in higher than two spatial dimensions, our

results strictly speaking do *not* prove the existence of holographic shadows. We have studied only boundary disks, rather than fully arbitrary boundary regions. Although it is natural to expect that more complicated boundary shapes cannot reduce the shadow size (since these tend to suffer from additional phase transition limitations), we have not succeeded in finding a general proof.

The organization of this chapter is as follows: In sec. 4.2 we present the general framework for using extremal bulk surfaces as probes. We introduce and prove two "coverage theorems" in the interest of formalizing the conditions under which a spacetime exhibits holographic shadows. Then, in sec. 4.3, we use these theorems to demonstrate the existence of entanglement shadows for globally well-defined geometries. In sections 4.4, 4.5, and 4.6, we extend our analysis to AdS-Schwarzschild geometries with three different probes: minimal area surfaces, static Wilson loops, and causal information surfaces. We present a comparison of these probes in the discussion, sec. 4.7, and close with a summary and some comments on future directions. Appendix 4.A contains proofs of some general properties of extremal surfaces. Some additional computational details may be found in the appendices of [2].

4.2 Properties of minimal surfaces

In this section, we present some general properties, terminology, and theorems that will prove useful in the analysis of holographic shadows that follows.

4.2.1 Minimal area surfaces

Let us first review the Ryu-Takayanagi proposal that relates bulk minimal surfaces to entanglement entropy on the boundary CFT [97, 138]. Consider a constant time slice in static, asymptotically AdS_{d+1} spacetime. Let the set of all bulk points be $\mathcal{B}$, and let $\mathcal{A}$ be all points on the asymptotic boundary S_{d-1}. The proposal relates the entanglement entropy for a boundary region $a \subseteq \mathcal{A}$ to the area of a dual bulk surface $b \subset \mathcal{B}$ if (1) b has the smallest area among all surfaces with $\partial b = \partial a$, and (2) b can be continuously deformed to a (more precisely, a must be homologous to b). This proposal has many interesting aspects, but in this chapter we will focus on one property with particular relevance for holographic reconstruction:

The Strong Coverage Property (SCP):
$\forall x \in \mathcal{B},\ \forall v \in T_x\mathcal{B},\ \exists a \subset \mathcal{A}$ whose dual minimal surface b intersects x with tangent vector along v.

Intuitively, this says that the entire bulk and its tangent bundle are "scanned over" by the minimal surfaces b of all possible boundary regions a. This is satisfied

by empty AdS, and also holds up to small perturbations thereof. In $(2+1)$ dimensions, SCP is equivalent to the condition for boundary rigidity [151], which means that knowing the entanglement entropy for every boundary region a uniquely determines the bulk geometry. SCP is also a necessary condition for the "hole-ographic" reconstruction of [142] (see also [145]). However, the requirement that one covers the entire tangent bundle is quite strong, and is not *a priori* obviously necessary for a successful reconstruction scheme. We will therefore also consider a weaker property:

The Weak Coverage Property (WCP):
$\forall x \in \mathcal{B}$, $\exists a \subset \mathcal{A}$ whose dual minimal surface b intersects x.

This simply means that every bulk point is covered by the minimal surface b of some boundary region a, but not necessarily scanning over all orientations in its tangent space. Note that this is not sufficient for boundary rigidity in 2 dimensions, nor for the aforementioned "hole-ographic" reconstruction. Nevertheless, this should be a minimal requirement for any attempt to reconstruct the bulk using this particular geometric dual.

It is worth pointing out that in the case of a disjoint boundary region $a = \bigcup_i a_i$ with dual minimal surface $b = \bigcup_j b_j$, there need not be a direct correspondence between a_i and b_j. This is illustrated in the case of two disconnected boundary subregions in figure 4.1. There are two ways for the two bulk curves to end on the four boundary points that specify ∂a without crossing, so there are (at least) two different local minima of their total area. Since the Ryu-Takayanagi proposal specifies b as possessing the smallest area of all bulk surfaces with $\partial b = \partial a$, the choice of which of these two bulk possibilities to employ is determined by comparing their respective areas.

As illustrated in figure 4.1, as the boundary subregions a_i are continuously increased, the bulk dual surfaces b_j are pushed inwards until, at some critical point, b switches over to the other possible combination of b_j, which are then pushed outwards towards the boundary as the a_i continue to grow. This provides a simple example of a key concept underlying holographic shadows: rather than mirror the continuous deformation of the boundary, the bulk dual surface may undergo a discontinuous switchover in order to ascribe to the global minimum. This is a phase transition from the boundary point of view [143], but here we will focus on the bulk implication. This switchover leaves out the middle region, and thereby limits the region of the bulk that can be probed.

Even without disconnected boundary regions, such switchovers can still occur. It has been examined in detail in the work of Hubeny in the context of AdS black holes [148], and also in geometries with a conical defect [145]. In all of the above examples, one is tempted to ascribe this behaviour to nontrivial topology:

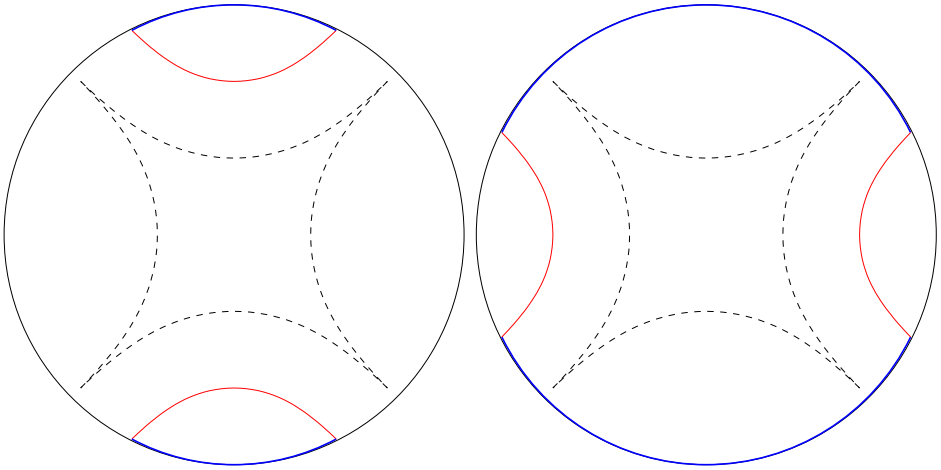


Figure 4.1: The left figure shows a disconnected boundary region $a = \bigsqcup_i a_i$ (blue) and the corresponding disjoint minimal surface $b = \bigsqcup_j b_j$ in the bulk (red). As the boundary region is continuously increased, the bulk surfaces b_j are pushed towards the dashed curve, at which point b discontinuously switches to the new global minimum $b = \bigsqcup_j b'_j$ shown in the right figure. The region inside the dashed curves cannot be probed with this particular choice of bulk dual.

either the boundary region is not simply connected, or the bulk has a horizon or a singularity. But in fact, topology is not the real problem. Given globally well-defined manifolds and simply connected boundary regions, the weak coverage property can still be violated[3]. Thus we will begin by studying the general behavior of SCP/WCP violation in spacetimes without horizons or singularities, and then proceed to analyze singular geometries.

4.2.2 Generalized minimal surfaces

Before proceeding, we shall first introduce a more general formulation of minimal bulk surfaces. In particular, one can formally take the Ryu-Takayanagi proposal as a special case of the following general prescription:

- Let $b \subset \mathcal{B}$ be an $n < d$ dimensional surface in the bulk, and define the geometric quantity

$$L(b) = \int_b \left| \mathrm{d}^n \vec{B} \right| \; F(g_{\mu\nu}) \; . \tag{4.1}$$

[3]Some have tried to establish that a globally regular, WCP-violating geometry is unstable and should collapse into a black hole [152, 153]. However, in this chapter we will show that in $(3+1)$ dimensions, a star of radius 5 times its mass in Planck units – e.g., neutron stars – can already violate WCP. General stability issues are only a serious concern when the radius is near $2M$ [154, 155], which is the Schwarzschild radius. Hence we find no reason to doubt that stable, regular geometries can indeed violate WCP.

Over this surface, we integrate the area element and the function F which only depends on the local geometry. This is then a very intuitive probe of the bulk geometry, as it does not care about the shape of b, but rather only about where b reaches.

- For an n dimensional boundary region a (or its boundary ∂a), one finds an observable Q associated with the minimal value of the above geometric quantity:

$$Q(a) = \text{Min}[L(b)]\Big|_{\partial b = \partial a} . \tag{4.2}$$

When $n = (d-1)$ and $F = 1$, this reduces to the Ryu-Takayanagi proposal with $L =$ area and $Q =$ entanglement entropy. In addition, when $n = 1$ and $F = g_{tt}$, this reduces to the action of certain Wilson loops. According to the form of eqn. (4.1), one should always be interested in a minimum. A maximum is ill-defined as one can always arbitrarily deform the surface along the null directions. In this chapter, we will also limit ourselves to quantities with $F > 0$ and

$$\lim_{b \to a} L(b) = \infty . \tag{4.3}$$

In other words, $L(b)$ is a positive definite quantity which diverges as one deforms b toward the boundary. It is therefore very natural to expect the minimal surface to reach into the bulk. This is related to boundary observables which have UV divergences and need to be regulated.

We can now study the failure of the coverage properties above, and the consequent "holographic shadows," in a more general manner not limited to minimal area surfaces vis-à-vis Ryu-Takayanagi. Other holographic duals can suffer from exactly the same obstacle, namely that the bulk probes fail to cover the entire manifold, thus placing a geometric limit on such reconstruction efforts. Our generalization makes it easier to compare different holographic probes and see which one is better, in the sense of which probe casts the smallest shadow.

4.2.3 Seeking shadows

In this chapter, we will limit ourselves to $O(d)$ symmetric bulk geometries and $O(n)$ symmetric, simply connected boundary regions (disks). In such cases we can specify a bulk point p by its radial distance to the origin, r_*. This point will be the $O(n)$ fixed point of a unique, $O(n)$ symmetric n dimensional surface $b(r_*)$ (modulo the remaining $SO(d-n)$ rotation) such that the first order variation of eqn. (4.1) is zero.[4]

[4]One might intuitively treat r_* as the minimal radius reached by this critical surface, but there is no *a priori* reason for this identification to hold for an arbitrary positive function F in

Proceeding from r_*, we follow the surface $b(r_*)$ to the boundary at $r = \infty$ to find the $(n-1)$-dimensional boundary sphere a on which it ends, $\partial a = \partial b$. We define the interior of a to be the side closer to the initial bulk point p. In other words, one can deform from b to a without going through $r = 0$. Denote the radius of this boundary ball a as $\theta_\infty(r_*)$.[5] We know two special values of this function: $\theta_\infty(\infty) = 0$ and $\theta_\infty(0) = \pi/2$. The first is due to a surface $b(\infty)$ that effectively never leaves the boundary, while the second comes from symmetry: it is basically the surface that cuts the bulk into two halves.

This function is straightforward to compute (at least numerically), and possesses a number of useful properties. First of all, there is a condition which guarantees that a holographic reconstruction scheme will work:

Theorem 1: *The set of all simply-connected, $O(n)$ symmetric boundary regions (balls) satisfies the Strong Coverage Property if $\theta_\infty(r_*) \in (0, \pi/2)$ is monotonic as r_* goes from 0 to ∞.*

Conversely, there is also a condition which guarantees that holographic reconstruction will fail:

Theorem 2: *If $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* > 0$ as $r_* \to 0$, then the weak coverage property fails for the set of all simply-connected, $O(n)$ symmetric boundary regions (balls).*

In this section, we will prove these two theorems using the following lemmas:

Lemma 1: *For a boundary sphere ∂a, the bulk surface b that minimizes L in eqn. (4.1) with $\partial b = \partial a$ must be spherically symmetric.*

Lemma 2: *If the boundary anchors ∂b and $\partial b'$ do not cross each other, but the corresponding bulk surfaces b and b' do, then b and b' cannot both be minimal surfaces.*

Proofs of these Lemmas will be given in appendix 4.A.

Proof of Theorem 1

Monotonicity of the boundary angle implies that every $b(r_*)$ is the unique global minimum for the boundary ball a of radius $\theta_\infty(r_*)$. Lemma 1 then implies that the bulk can be foliated by a family of nonintersecting minimal surfaces anchored

eqn. (4.1). We will be very careful not to assume this identification in the proofs that follow.

[5] There might be cases where some critical surfaces $b(r_*)$ do not reach the boundary, so θ_∞ is not well-defined. This is exactly what happens when there is a horizon, but such cases may be more general.

on the corresponding family of concentric boundary spheres, as illustrated in fig. 4.2. Note that this is sufficient to satisfy WCP; for the strong coverage property, we need also demonstrate coverage of the bulk tangent bundle.

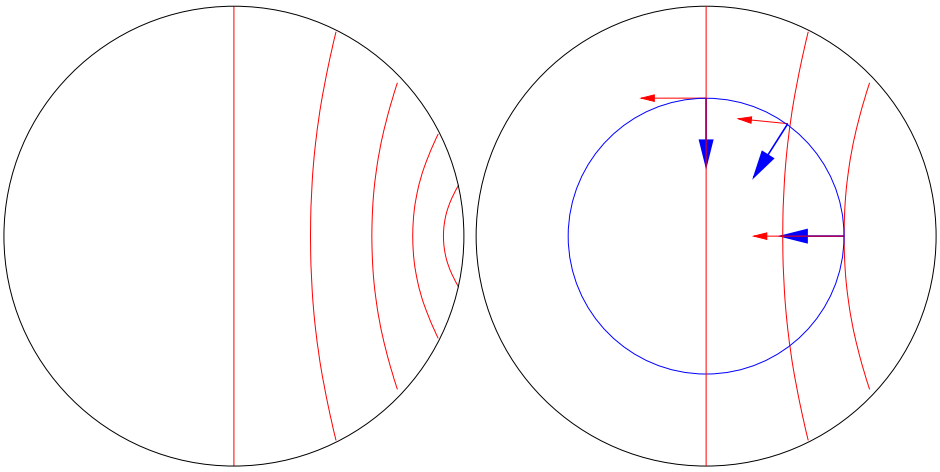


Figure 4.2: The left figure shows a continuous foliation of minimal n-dimensional surfaces (red) on an $(n+1)$-dimensional equatorial slice of the bulk. The right figure shows how the angle between an n-sphere (blue circle) in the bulk and the foliation surfaces changes continuously from 0 to $\pi/2$. Note that although the rightmost red surface is tangent to the blue circle at precisely r_* in this plot, the proof does not rely on this.

Consider a sphere with finite radius R in the bulk. As shown in fig. 4.2, it intersects $b(0)$ at an angle of $\pi/2$ between their normal vectors. As r_* increases, $b(r_*)$ will eventually stop intersecting this sphere. If we follow the intersection point during this process, the angle between the two normal vectors must continuously drop to 0. Thus $b(r_*)$ can cover the full tangent space of a point at radius R. Since R is arbitrary, we have covered the full tangent bundle. QED

Note that the inverse of Theorem 1 is not generally true. That is, a non-monotonic $\theta_\infty(r_*)$ does not guarantee the violation of SCP.[6] But this is not so concerning. We have stipulated SCP as a *sufficient* condition for a successful holographic reconstruction scheme; violating SCP does not necessarily imply that all schemes will fail. Thus, the more physically meaningful "inverse" statement is rather our Theorem 2, about the violation of WCP. Insofar as WCP is a *necessary* condition, this indeed rules out holographic reconstruction (using the set of all boundary disks). Also note that Theorem 2 provides a sufficient condition to violate WCP. While WCP might be violated by other conditions, the condition Theorem 2 provides seems to be the most natural.

[6]The inverse of Theorem 1 can be proven if we use the additional assumption that r_* is the minimal radius reached by the surface $b(r_*)$, which happens to be true in many examples.

Proof of Theorem 2

If $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* > 0$ when $r_* \to 0$, then since $\theta_\infty(0) = \pi/2$ we can find some $r' > 0$ such that $\theta_\infty(r_*) \geq \pi/2$ for all $0 \leq r_* \leq r'$. According to Lemma 2, none of the critical surfaces $b(r_*)$ in this range can be the global minimum of the corresponding boundary sphere ∂b, because they always intersect their own mirror image.

If for all minimal surfaces $b(r_*)$, r_* is the minimal radius reached, then no minimal surfaces can probe the region $r < r'$. On the other hand, if a point $p \in b(r_*)$ with radius $r_p < r_*$ is allowed, one still cannot allow $r_p \to 0$. As shown in fig. 4.3, such a surface can be pinched-off to one with smaller L, which contradicts the assumption that the original surface is a global minimum. Thus in this case there must be a lower bound r'' with $0 < r'' < r'$ beyond which these minimal surfaces cannot probe. QED

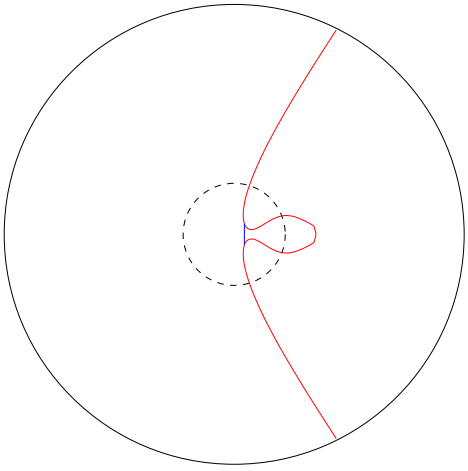


Figure 4.3: A minimal surface (red) with its symmetric point sitting at a finite radius r_* cannot have other points approach arbitrarily close to $r = 0$. Otherwise, a pinched-off version (blue) will have even smaller area.

In this chapter, we will explore the simplest examples where $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* > 0$ for $r_* < r'$ and $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* < 0$ for $r_* > r'$. Additionally, in all the examples we study, we find that r_* is the minimal radius reached by the surface $b(r_*)$. Hence, in the rest of this chapter we will adhere to the notation that r_* refers to the minimal radius reached for a fixed boundary region, while $r_{\min}$ denotes the minimum r_* among all possible boundary regions, i.e., the global minimum. Thus, $r_{\min}$ is also the size of the holographic shadow.

4.3 Stellar shadows

We begin our analysis by applying the above framework to identify shadows in globally regular geometries—namely, stars in AdS spacetimes. Specifically, we shall present analytical results for stars in AdS_{2+1}, which provides an explicit example of a non-singular bulk geometry that nonetheless exhibits an entanglement shadow.[7] For our purposes, it is not necessary to specify the matter distribution; we assume only constant density.

The case we consider is that of an ideal (constant density) star of radius R embedded in AdS_{2+1}, for which a physically reasonable metric is:

$$ds^2 = g_{tt}(r)\,dt^2 + \frac{dr^2}{f(r)} + r^2\,d\theta^2, \quad f(r) = \begin{cases} r^2 + 1 - GM, & r > R \\ r^2 + 1 - GM\frac{r^2}{R^2}, & r \leq R \end{cases} \tag{4.4}$$

where the AdS radius ℓ_{AdS} is set to 1, and g_{tt} depends on the particular matter distribution. Since the metric admits the Killing vector ∂_t, we can analyze extremal surfaces associated to entanglement entropy on constant-time slices. We thus limit our example to entanglement surfaces, since an analysis of both Wilson loops and causal information surfaces would require explicit knowledge of the g_{tt} component. Though a direct comparison of probes in this geometry would be interesting, the result for entanglement surfaces alone suffices to make our point: holographic shadows are general phenomena not limited to singular or topologically nontrivial geometries.

In what follows, we take $GM > 1$, and try to solve for θ_∞ as a function of r_*[8]. In the exterior region ($r > R$) the spatial part of the metric is identical to that of the BTZ metric (cf. (4.15)) with the identification $r_H^2 \equiv GM - 1$. Thus for $r_* > R$, $\theta_\infty(r_*)$ is identical to the BTZ solution as we will demonstrate later in eqn. (4.24). For $r_* \leq R$, $\theta_\infty(r_*)$ is obtained by smoothly matching the $r < R$ segment and the $r > R$ segment.

The length of a spacelike geodesic may be written:

$$L = \int dr \sqrt{\frac{1}{f(r)} + r^2\theta'^2} \tag{4.5}$$

where the prime denotes differentiation with respect to r. Extremizing via Euler-

[7] These results can be numerically extended to higher dimensions; see [2] for the explicit case of AdS_{3+1}.

[8] The $GM < 1$ case corresponds to the conical defect geometry, for which the analysis proceeds along precisely similar lines.

Lagrange, we have

$$\frac{r^2\theta'}{\sqrt{f(r)^{-1}+r^2\theta'^2}} = \frac{\delta\mathcal{L}}{\delta\theta'} \equiv r_*$$

where the minimum radius r_* for this geodesic is, in our units, equivalent to the associated conserved angular momentum. Solving this expression for θ', we obtain

$$\frac{\mathrm{d}\theta}{\mathrm{d}r} = \frac{r_*}{r\sqrt{f(r)(r^2-r_*^2)}} \,. \tag{4.6}$$

We may then perform an indefinite integral in the exterior ($r > R$), with $f(r) = r^2 + 1 - GM$, to find

$$\theta_E(r) = \frac{1}{2\sqrt{GM-1}} \cosh^{-1}\left(\frac{-2r_*^2(GM-1)+r^2(r_*^2+GM-1)}{r^2(r_*^2-GM+1)}\right) + g(r_*) \tag{4.7}$$

with constant of integration $g(r_*)$, and in the interior ($r < R$), with $f(r) = r^2 + 1 - GMr^2/R^2$, to find

$$\theta_I(r) = \frac{1}{2}\cos^{-1}\left(\frac{2r_*^2+r^2\left(-1+r_*^2\left(1-\frac{GM}{R^2}\right)\right)}{r^2\left(1+r_*^2\left(1-\frac{GM}{R^2}\right)\right)}\right) \tag{4.8}$$

where the subscripts E and I distinguish these functions as valid in the exterior and interior, respectively. For θ_I, the constant of integration has been fixed to 0 by the symmetry assumption that demands that the minimum r_* occurs at $\theta = 0$, i.e. $\theta_I(r_*) = 0$. To fix the constant of integration $g(r_*)$ in θ_E, we demand continuity in both the function and its first derivative at the stellar boundary $r = R$. The latter condition is satisfied automatically by the conserved angular momentum r_*, thus we simply solve $\theta_I(R) = \theta_E(R)$ for $g(r_*)$:

$$\begin{aligned} g(r_*) =& \frac{1}{2}\cos^{-1}\left(\frac{2r_*^2+R^2\left(-1+r_*^2\left(1-\frac{GM}{R^2}\right)\right)}{R^2+R^2r_*^2\left(1-\frac{GM}{R^2}\right)}\right) \\ &- \frac{1}{2\sqrt{GM-1}}\cosh^{-1}\left(\frac{-2r_*^2(GM-1)+R^2\left(r_*^2+GM-1\right)}{R^2\left(r_*^2-GM+1\right)}\right) \end{aligned} \tag{4.9}$$

which we may substitute into (4.7). The function $\theta_\infty(r_*)$ is then obtained by

taking the $r \to \infty$ limit of the result. Dropping the subscript E, we at last obtain

$$\begin{aligned}\theta_\infty(r_*) = &\frac{1}{2\sqrt{GM-1}} \cosh^{-1}\left(\frac{r_*^2+GM-1}{r_*^2-GM+1}\right) \\ &+ \frac{1}{2}\cos^{-1}\left(\frac{2r_*^2+R^2\left(-1+r_*^2\left(1-\frac{GM}{R^2}\right)\right)}{R^2+R^2r_*^2\left(1-\frac{GM}{R^2}\right)}\right) \\ &- \frac{1}{2\sqrt{GM-1}} \cosh^{-1}\left(\frac{-2r_*^2(GM-1)+R^2\left(r_*^2+GM-1\right)}{R^2\left(r_*^2-GM+1\right)}\right)\end{aligned} \tag{4.10}$$

for the minimal geodesics extending from $r_* \leq R$ to the boundary at infinity.

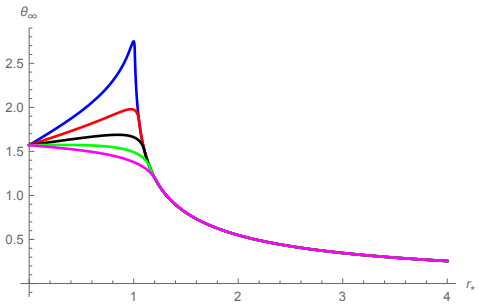
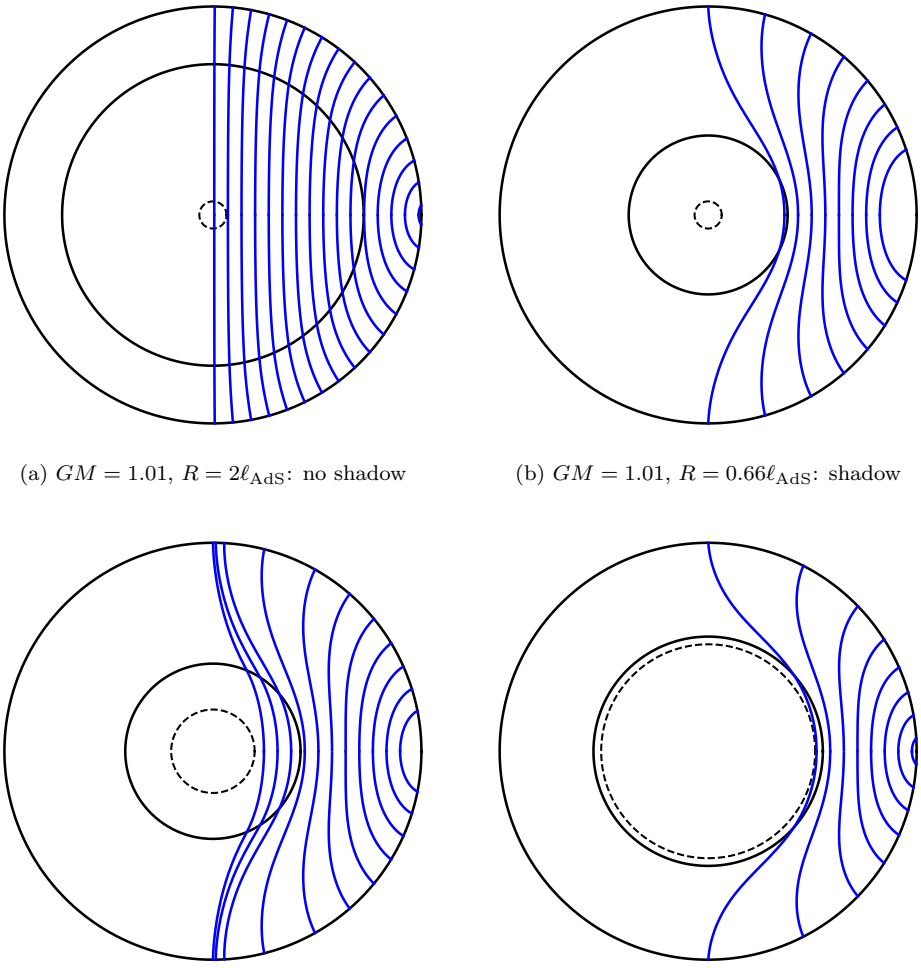


Figure 4.4: $\theta_\infty(r_*)$ for $GM = 2$ and stellar radii $R = 1.01\ell_{\text{AdS}}$ (blue), $1.05\ell_{\text{AdS}}$ (red), $1.1\ell_{\text{AdS}}$ (black), $1.15\ell_{\text{AdS}}$ (green), and $1.2\ell_{\text{AdS}}$ (magenta). The case $R = 1.2\ell_{\text{AdS}}$ is insufficiently dense, and hence exhibits a monotonic function with no shadows. But the other cases, with $R < \sqrt{4/3}\ \ell_{\text{AdS}}$ (cf. (4.11)), have a single maximum at finite radius $r_{\min}$, within which an entanglement shadow exists.

This function is plotted for a range of stellar parameters in fig. 4.4. Clearly, for insufficiently dense stars, $\theta_\infty(r_*)$ is monotonically decreasing, thus SCP is satisfied. However, for sufficiently dense stars, $d\theta_\infty(r_*)/dr_* > 0$ as $r_* \to 0$, thus WCP is violated, implying the existence of a shadow within some $r_{\min}$. Note that in many cases the shadow extends beyond the stellar boundary, $r_{\min} > R$; this is because, for the BTZ geometry in the exterior, we already have $\theta_\infty(\pi/2)$ independent of the stellar mass distribution. In such cases even the assumption of constant density is irrelevant: a shadow will exist as long as enough mass sits within some finite radius R. See fig. 4.5 for plots of the minimal surfaces for a range of stellar densities; the shadow region is easily seen by rotating the surfaces about the center.

We can obtain an expression for the density range that supports shadows from the condition that $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* > 0$ at $r_* = 0$, or from demanding the existence of a real solution to $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* = 0$. Either condition implies:

$$GM - 1 < R^2 < \frac{GM^2}{GM+1}\ . \tag{4.11}$$



(a) $GM = 1.01$, $R = 2\ell_{\text{AdS}}$: no shadow

(b) $GM = 1.01$, $R = 0.66\ell_{\text{AdS}}$: shadow

(c) $GM = 1.1$, $R = 0.75\ell_{\text{AdS}}$: shadow

(d) $GM = 2$, $R = 1.12\ell_{\text{AdS}}$: shadow

Figure 4.5: Plots of extremal surfaces (blue) for stars of varying density. The solid black circle is the stellar radius R; the smaller, dotted black circle is the would-be horizon radius r_H. Note that in the first case, which is outside the range (4.11), there is no restriction against covering the entire bulk.

Note that the l.h.s. is simply r_H^2. Thus the inequality (4.11) effectively imposes a lower limit on the density for which one can satisfy SCP: stars of a given mass whose radius falls below the right-hand side will exhibit shadows.

4.4 Minimal area surfaces in Schwarzschild-AdS geometries

We now turn our attention to singular geometries, in particular AdS with a black hole in the center. Obviously, $\theta_\infty(r_*)$ is undefined if r_* falls within the horizon radius, hence from now on $r_* \geq r_H$ is always implied.

A key point worth emphasizing is that, for AdS black holes, the phase transition ("switchover") behaviour is modified. Previously, the global minimum switched solution branches when

$$A(\theta_\infty) = A(\pi - \theta_\infty) \tag{4.12}$$

where A is the area of the associated surface (or set of surfaces). In other words, one switches from a given bulk surface to the complement when the two have equal area, cf. fig. 4.1. In the case of a black hole however, the complement must include the horizon area [97]. This modifies the above area condition to:

$$A(\theta_\infty) = A(\pi - \theta_\infty) + A_{\rm BH} \tag{4.13}$$

where $A_{\rm BH}$ is the portion that wraps the black hole.

We will present our results in three separate subsections. In sec. 4.4.1 we analytically solve for minimal spacelike geodesics in the BTZ geometry. We then move to higher-dimensional considerations of boundary disks in global AdS in sec. 4.4.2, which we split into large and small black holes to obtain suitable approximations. Although the associated spherically symmetric codimension-1 bulk surfaces are the most natural higher-dimensional generalizations of the lower-dimensional geodesics, we also present a similar analysis of boundary strips in planar/Poincaré-AdS in sec. 4.4.3, as the latter allow for a more straightforward approximation. As we shall see, for large black holes, boundary disks and strips perform almost equally well in the sense that both exhibit exponentially small shadows.[9] For small black holes however, strips suffer from more complicated phase transition behaviour that makes them worse boundary shapes than disks, whose associated shadow is of order r_H.

[9] When referring to AdS-Schwarzschild, we shall speak of the size of the shadow relative to the horizon radius. Thus an exponentially small shadow is one which for which $r_{\rm min} - r_H \sim e^{-\# r_H/\ell_{\rm AdS}}$, with $\#$ some order one constant.

4.4.1 BTZ black holes

The bulk quantity dual to the von Neumann entropy of a boundary subregion $\mathcal{A}$ has been conjectured to be given by [97, 112]:

$$S(\mathcal{A}) = \frac{\text{Area}\left(\mathfrak{E}(\partial\mathcal{A})\right)}{4G} \tag{4.14}$$

where $\mathfrak{E}(\partial\mathcal{A})$ is the extremal bulk surface that ends on $\partial\mathcal{A}$ and has minimal proper area among surfaces continuously deformable to $\mathcal{A}$. When the global state of the boundary is pure, the von Neumann entropy gives a quantitative estimate for the entanglement between the subregion and its complement, called the entanglement entropy. When the global boundary state is mixed, this is no longer necessarily true, although we use the terms von Neumann entropy and entanglement entropy interchangeably in this chapter.

A static BTZ black hole is described by the metric

$$\mathrm{d}s^2 = -(r^2 - r_H^2)\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{r^2 - r_H^2} + r^2\,\mathrm{d}\theta^2. \tag{4.15}$$

To determine the shadow, it is sufficient to consider constant time slices.[10] In $d = 2$ the boundary is a circle, and the subsystem $\mathcal{A}$ an interval on the circle. The bulk extremal surface associated with the entanglement entropy is then simply a geodesic anchored at the two points that comprise $\partial\mathcal{A}$. We consider as a boundary region the interval $(-\theta_\infty, \theta_\infty)$, where the subscript ∞ indicates that the boundary corresponds to $r \to \infty$ in our coordinates (4.15).

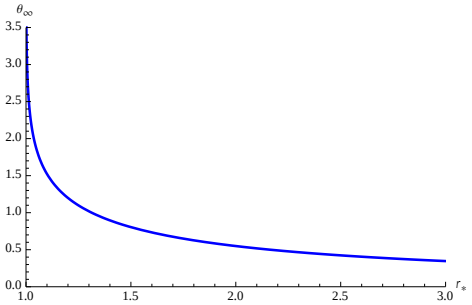


Figure 4.6: $\theta_\infty(r_*)$ for a static BTZ black hole with $r_H = 1$.

The Lagrangian describing such a bulk extremal surface is given by

$$\mathcal{L} = \sqrt{\frac{r'^2}{r^2 - r_H^2} + r^2}\,, \qquad r' \equiv \frac{\mathrm{d}r}{\mathrm{d}\theta}\,. \tag{4.16}$$

[10] One can show that for $\mathrm{d}t \neq 0$ subregions, the entanglement shadow is even larger; see [2].

Since the Lagrangian does not depend on θ, there is a conserved momentum due to translation invariance in θ. Hence:

$$\frac{\delta \mathcal{L}}{\delta r'} r' - \mathcal{L} = \text{constant} \ . \tag{4.17}$$

We may fix the constant by the demanding that the surface reaches its minimal value r_* when $r' = 0$. This leads to the first-order equation of motion

$$\frac{\mathrm{d}r}{\mathrm{d}\theta} = \frac{r}{r_*} \sqrt{r^2 - r_*^2} \sqrt{r^2 - r_H^2} \tag{4.18}$$

which may be integrated to obtain

$$\theta_\infty = \int_{r_*}^{\infty} \mathrm{d}r \frac{\mathrm{d}\theta}{\mathrm{d}r} = \frac{1}{2r_H} \cosh^{-1} \left(\frac{r_*^2 + r_H^2}{r_*^2 - r_H^2} \right) \ . \tag{4.19}$$

This curve is plotted in figure 4.6. Note that it diverges when $r_* \to r_H$, and decreases monotonically with increasing r_*.

We may invert (4.19) to obtain:

$$r_* = \frac{r_H}{\tanh\left(\theta_\infty r_H\right)}. \tag{4.20}$$

which is plotted in figure 4.14. One clearly sees that that there are geodesics that wind around the black hole one or more times as r_* approaches the horizon. But a surface that intersect itself cannot correspond to a local minimum of the area functional (intuitively, the kinks in the intersection can be infinitesimally smoothed out to reduce the area). Thus for the purpose of identifying the appropriate bulk probe, we only care about the range $\theta_\infty \leq \pi$, since a switchover must occur before θ_∞ reaches this value. The alternative global minimum is then a surface with two disconnected components: a geodesic connecting the endpoints at $\pm\theta_\infty$ on the opposite side of the black hole, and a separate part that encircles the horizon; see figure 4.7.

We denote the critical angle at which this switchover happens by θ_{switch}, which is given by (4.13):

$$l(\theta_{\text{switch}}) = l(\pi - \theta_{\text{switch}}) + 2\pi r_H, \tag{4.21}$$

where $l(\theta_\infty)$ is the length of the geodesic connecting the boundary points $\pm\theta_\infty$ and $2\pi r_H$ is the length of the curve that wraps the horizon.

We can compute the length $l(\theta_\infty)$ by integrating the Lagrangian

$$l(\theta_\infty) = 2 \int_{r_*}^{\infty} \sqrt{\frac{1}{r^2 - r_H^2} + r^2 \left(\frac{\mathrm{d}\theta}{\mathrm{d}r} \right)^2} = 2 \int_{r_*}^{\infty} \frac{r \, \mathrm{d}r}{\sqrt{r^2 - r_H^2} \sqrt{r^2 - r_*^2}} \tag{4.22}$$

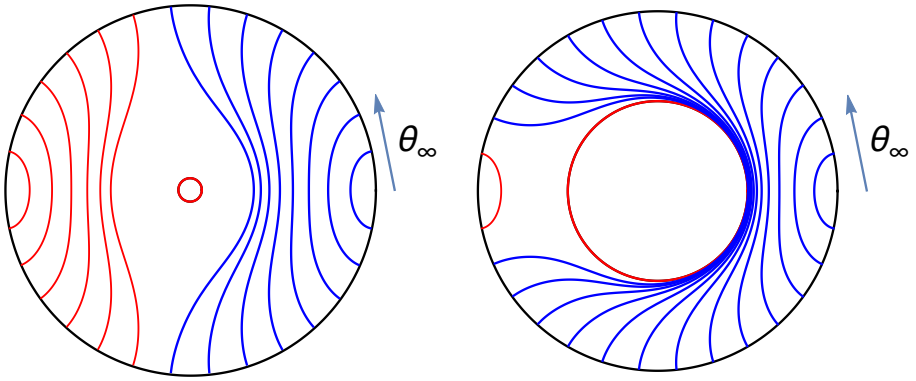


Figure 4.7: Minimal surfaces for boundary intervals of varying size θ_∞, for a black hole of radius (red circle) $r_H = 0.1\ell_{\rm AdS}$ (left) and $r_H = \ell_{\rm AdS}$ (right). The switchover to the disconnected solution (red curves) takes place near $\theta_\infty = \pi/2$ for small black holes (left), and approaches π for large black holes (right).

where we used (4.18), with r_* given by (4.20). The integral is divergent, but the divergent parts on the left- and right-hand side of (4.21) cancel, and the finite parts yield:

$$\theta_{\rm switch} = \frac{\pi}{2} + \frac{1}{2r_H}\ln\left(\cosh(\pi r_H)\right) \ . \tag{4.23}$$

For small black holes ($r_H \ll \ell_{\rm AdS}$) we have that $\theta_{\rm switch} \approx \pi/2$, because the area contribution from the black hole in eqn. (4.21) is close to zero. Conversely, one sees that for large black holes ($r_H \gg \ell_{\rm AdS}$), $\theta_{\rm switch} \approx \pi$. See figure 4.7 for an explicit plot of both cases.

The shadow radius $r_{\rm min}$, within which no extremal surface associated to entanglement entropy can reach, is finally determined by substituting the value of $\theta_{\rm switch}$ into (4.20):

$$r_{\rm min} = \frac{r_H}{\tanh(\pi r_H)} + \frac{r_H e^{-\pi r_H}}{\sinh(\pi r_H)} \ . \tag{4.24}$$

This curve is plotted in figure 4.8. However, since the black hole is always within the shadow region, the shadow may be more conveniently expressed as

$$\Delta r_0 \equiv r_{\rm min} - r_H = \frac{2r_H e^{-\pi r_H}}{\sinh\left(\pi r_H\right)} \tag{4.25}$$

which is plotted in figure 4.9. When referring to the "size" of the shadow, we shall implicitly mean the relative quantity (4.25) unless otherwise noted.

From either eqn. (4.25) or fig. 4.7, one sees that the shadow is exponentially

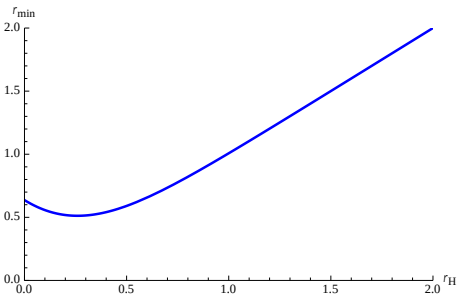


Figure 4.8: Shadow radius $r_{\min}$ as a function of horizon radius r_H for a static BTZ black hole.

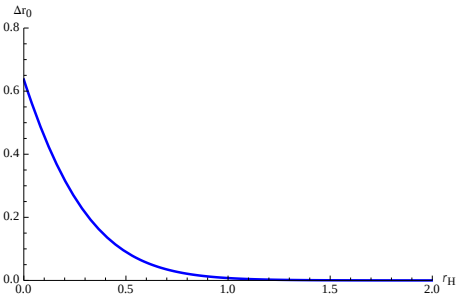


Figure 4.9: Relative shadow size Δr_0 as a function of horizon radius r_H for a static BTZ black hole.

small for large black holes, but remains an order one (AdS radius) distance from the horizon for small black holes. This behavior is easily explained by considering the switchover effect: a large black hole incurs a greater cost from the horizon component in the area condition (4.13), which allows the global minimum to remain on the original (connected) solution branch for larger values of θ_∞.

It may seem strange that that the shadow radius $r_{\min}$ does not go to zero for vanishing horizon radius. This is due to the mass gap in AdS_3: letting $r_H \to 0$ in the BTZ metric (4.15) will not yield the empty AdS_3 metric, but a conical defect. It was previously shown in [145] that the conical defect geometry exhibits entanglement shadows; we shall comment further on this issue in the discussion, sec. 4.7.

4.4.2 Global $SAdS_{d+1}$ with $d \geq 3$

We now wish to ask how this result changes for higher-dimensional black holes. Unlike the BTZ case, in which the boundary interval was completely specified by the angle θ_∞, we will now consider the entanglement entropy of a $(d-1)$-dimensional region in the boundary CFT, which in principle can have an arbitrarily complicated shape (indeed, it need not even be simply connected). This allows for much richer phase transition structure when deforming the region. Hence for simplicity, we will generally assume that the boundary region of interest is $O(d-1)$ symmetric, i.e., we consider minimal surfaces of the form $r(\theta)$.

Note that, among boundary regions of different shapes but equal area, it seems very plausible that these highly symmetric surfaces will maximize the reach into the bulk [146, 156]. However, this does not directly imply that asymmetric regions cannot have minimal surfaces that penetrate the shadows we find herein. This is because, as we have stressed, shadows arise from the switchover behavior, and it is difficult to study such behavior for less symmetric surfaces. Nevertheless,

we believe that even if less symmetric surfaces do probe deeper in some cases, it will not eliminate shadows, and probably will not deviate much from the bounds obtained from these highly symmetric surfaces.

Even when restricting to $O(d-1)$-symmetric surfaces, higher dimensions still allow various interesting new switchover effects. Contrast figures 4.10 and 4.11 below. In figure 4.10, we consider a spherical boundary region, analogous to the BTZ case above. As the radius of this boundary "disk" increases, the global minimum will eventually switch to a disconnected bulk solution consisting of the spherical cap on the far side of the black hole and a part that wraps the horizon. In figure 4.11, we instead consider a band around the boundary sphere. As we increase the width of this "strip", the dual minimal surface will again undergo a switchover, but now from a single connected piece to two hemispherical caps plus the horizon component.

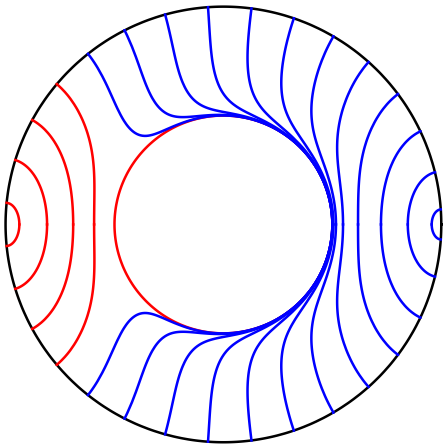


Figure 4.10: Transition between two different boundary disks for a black hole with horizon $r_H = \ell_{\text{AdS}}$ in AdS_5.

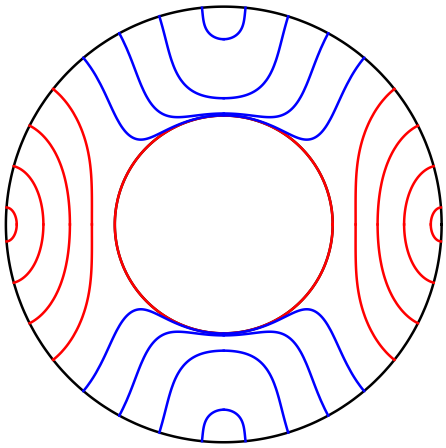


Figure 4.11: Transition between a boundary strip and two disks for a black hole with horizon $r_H = \ell_{\text{AdS}}$ in AdS_5.

In order to study the size of the shadows in these higher dimensional geometries, we will proceed as above, by constructing the function $\theta_\infty(r_*)$ that encodes information about how well the boundary entanglement entropy can reconstruct the bulk. One of the major differences from the AdS_3 case is that in higher dimensions the equations of motion describing the minimal surfaces cannot be solved analytically. We rely instead on numerical methods. Results for a black hole with $r_H = \ell_{\text{AdS}}$ are displayed in figure 4.12. At first sight, it looks qualitatively very similar to the BTZ case, cf. figure 4.6. However, zooming in on the near horizon region, as shown in fig. 4.13, reveals a crucial difference: $\theta_\infty(r_*)$ is *not* actually monotonic. In fact, although not clearly visible in fig. 4.13, it will oscillate an

infinite number times as $r_* \to r_H$ [148]. The difference is due to the fact that in the BTZ geometry the minimal surfaces are geodesics which in principle can self intersect, whereas in higher d the surfaces instead fold into multiple layers around the black hole. See figures 4.14 and 4.15 for an explicit illustration of these two behaviours.

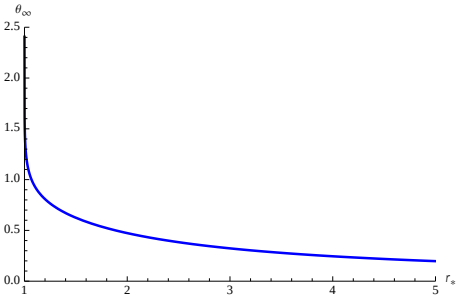


Figure 4.12: $\theta_\infty(r_*)$ for a $SAdS_5$ black hole with $r_H = l_{AdS}$.

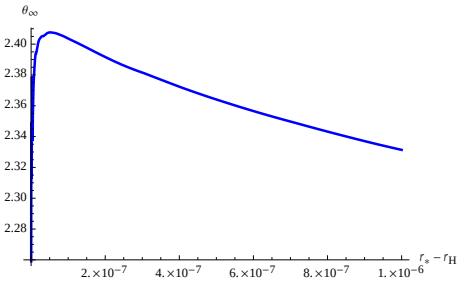


Figure 4.13: Close up of $\theta_\infty(r_*)$ for a $SAdS_5$ black hole with $r_H = l_{AdS}$ near $r \approx r_H$.

To find the shadow, we must study the switchover behavior. Note that while θ_∞ remains finite as $r_* \to r_H$, there are values of r_* for which $\theta_\infty(r_*) \geq \frac{\pi}{2}$, which makes switchovers likely. We know from Lemma 3 in appendix 4.A that values of r_* for which $\mathrm{d}\theta_\infty / \mathrm{d}r_* < 0$ cannot be minimal surfaces. Additionally, the critical surfaces for which θ_∞ undergoes oscillations (e.g. the red curve in figure 4.15) will fold around the black hole and intersect their mirror image. Hence by Lemma 2, they cannot be minimal either. Therefore, we again only need to find the largest value of r_* for which the switchover condition (4.13) is satisfied. This r_* then corresponds to the shadow radius $r_{\min}$ for the symmetric surfaces under consideration.

In the limiting case of a large and small[11] black hole in AdS, we can analytically approximate the size of the shadow Δr_0 in arbitrary dimension as follows. The metric for SAdS_{d+1} is given by

$$\mathrm{d}s^2 = -f(r)\,\mathrm{d}t^2 + \frac{\mathrm{d}r^2}{f(r)} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\Omega_{d-2}^2\right) \tag{4.26}$$

where

$$f(r) = r^2 + 1 - \frac{r_H^{d-2}}{r^{d-2}}\left(r_H^2 + 1\right) \ . \tag{4.27}$$

[11] Although small black black holes have negative heat capacity in $d \geq 3$, they can still describe stable solutions in the microcanonical ensemble for some range of masses [157].

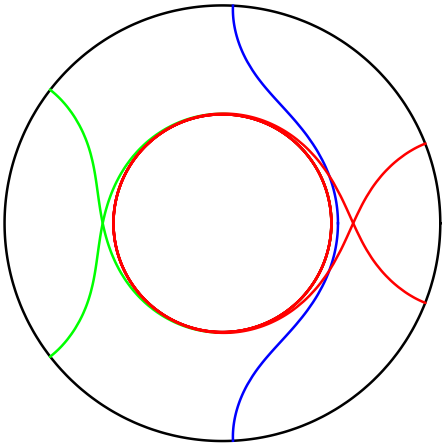


Figure 4.14: Extremal entangling surfaces in BTZ with horizon $r_H = \ell_{\text{AdS}}$ and $\Delta r_0 = 10^{-1}$ (blue), 10^{-3} (green), and 10^{-11} (red). The red surface wraps the horizon four times.

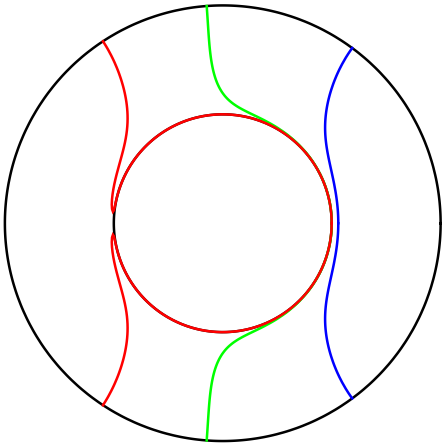


Figure 4.15: Extremal entangling surfaces in AdS_5 with horizon $r_H = \ell_{\text{AdS}}$ and $\Delta r_0 = 10^{-1}$ (blue), 10^{-3} (green), and 10^{-11} (red). Note the folding behaviour in the red surface.

From the Lagrangian describing a $O(d-1)$ minimal surface,

$$\mathcal{L} = (r(\theta)\sin\theta)^{d-2}\sqrt{\frac{r'(\theta)^2}{f(r)} + r(\theta)^2} \tag{4.28}$$

we can write down the Euler-Lagrange equation of motion and expand it:

$$r''(\theta) = (d-1)\left[r_H^2 d + (d-2)\right](r - r_H) - (d-2)\cot(\theta)r' + O(r')^2 + O(r-r_H)^2 \,. \tag{4.29}$$

where as usual the prime denotes differentiation with respect to θ. Assuming we are in a regime where it is permissible to drop the higher order terms (which is near the tip of the surface and close to the horizon), the above may be written

$$r''(\theta) = -(d-2)\cot(\theta)r' + (d-1)r_H\partial_r f(r_H)(r-r_H)\,; \qquad r \approx r_H,\ r' \ll 1. \tag{4.30}$$

This equation can be solved analytically for all d, but in $d = 4$ it takes the particularly simple form

$$r(\theta) = r_H + \frac{\Delta r_0}{\lambda}\frac{\sinh(\theta\lambda)}{\sin\theta}\,, \qquad \lambda \equiv \sqrt{12 r_H^2 + 5}\,. \tag{4.31}$$

The approximation is plotted on top of the exact solution in figure 4.16 for various values of Δr_0.

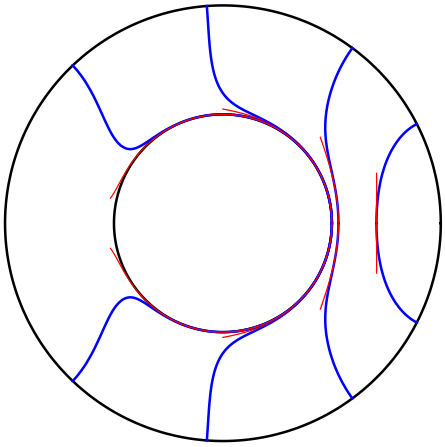


Figure 4.16: The approximation (4.31) (red) rendered atop the exact minimal surfaces (blue) for a black hole with horizon $r_H = \ell_{\text{AdS}}$ in AdS_5. The surfaces are plotted for $\Delta r_0 = 10^{-6}$, 10^{-3}, 10^{-1}, and 1.

4.4.2.1 Large black holes: $r_H \gg \ell_{\text{AdS}}$

For concreteness, we continue our study of shadows for large black holes in AdS_5, but our conclusions will remain valid for large black holes in arbitrary dimension.

In the large black hole limit $r_H \gg 1$, our approximate solution (4.31) reduces to

$$r(\theta) - r_H \approx \frac{\Delta r_0}{2\sqrt{12} r_H} \frac{e^{\sqrt{12} r_H \theta}}{\sin\theta} \ . \tag{4.32}$$

This solution $r(\theta)$ locally solves the minimal area equations in the near horizon geometry. If we pick the parameter Δr_0 to be small, this solution will cover an order one angle θ before the approximation breaks down (see figure 4.16). At this point, the surface is a distance $O(r_H)$ away from the horizon, and one could extend the approximation by matching it to a solution in empty AdS anchored to the boundary. While we don't need to know the exact solution in this regime to estimate the shadow, we can show that the rest of the minimal surface will be quite boring in the sense that it is almost going radially outward towards the boundary. To be more precise, we shall show that the amount of angle $\Delta\theta$ that the minimal surface covers when leaving this near horizon regime will be small in the large black hole limit.

We start with Lagrangian (4.28) and approximate $\sin\theta$ to be constant. We then take $f(r) \approx r^2$ since we are relatively far from the black hole. As the Lagrangian no longer depends on θ, there is a conserved quantity C associated to translations

in the angular direction, hence:

$$\frac{\delta\mathcal{L}}{\delta r'}r' - \mathcal{L} = C \implies r^{2d} = C^2\left(\frac{r'^2}{r^2} + r^2\right) . \tag{4.33}$$

The constant C can subsequently be determined by matching, at $r = 2r_H$, to our near horizon solution. Specifying henceforth to $d = 4$, this yields $C \approx 6r_H^3$. Plugging this into the above, we find

$$\Delta\theta = \int_{2r_H}^{\infty} \frac{\mathrm{d}r}{r^2\sqrt{\frac{r^6}{C^2} - 1}} \approx \frac{0.1}{r_H} \ll 1 \tag{4.34}$$

which confirms that the minimal surfaces are going approximately radially outward outside $r = 2r_H$. Thus we may match our near-horizon solution at a distance r_H from the horizon at some order-one matching angle θ_m to conclude:

$$\Delta r_0 \approx 2\sqrt{12}r_H^2\sin(\theta_m)e^{-\sqrt{12}\theta_m r_H} \propto r_H^2 e^{-\#r_H} \tag{4.35}$$

where # is an $O(1)$ number. Thus we find that the shadow region for minimal surfaces is exponentially small for large black holes. Although this particular result has been obtained for SAdS_5, one can show that it holds in any dimension; see sec. 4.4.3 below.

We must note that in (4.31) we choose as a boundary condition $r_* = r(\theta = 0)$, which corresponds to disk-shaped boundary regions. In contrast, the aforementioned boundary strips would require $r_* = r(\theta = \frac{\pi}{2})$. The analysis for the strip is precisely analogous, and also results in an exponentially small shadow. In section 4.4.3, we explicitly show that the shadow is exponentially small for all d in planar-SAdS_{d+1}, but we first turn to an analysis of small black holes in global SAdS_5.

4.4.2.2 Small black holes: $r_H \ll \ell_{\mathrm{AdS}}$

For small black holes, we rely on a different argument to estimate the size of the shadow. Since the horizon area is small in eqn. 4.13, the switchover angle must be approximately $\pi/2$. Additionally, as explained above, the minimal surface must remain in a single hemisphere, with no folds. The shadow size will therefore be determined by a simple minimal surface at the switchover point. Starting from the boundary at $\theta = \pi/2$, this surface will dive almost radially inward until it is an order $r_H \ll \ell_{\mathrm{AdS}}$ away from the black hole horizon. Here it can be matched to our approximate solution (4.31) in the $r_H \ll \ell_{\mathrm{AdS}}$ limit:

$$r(\theta) = r_H + \frac{\Delta r_0}{\lambda}\frac{\sinh(\theta\lambda)}{\sin\theta} , \qquad \lambda \approx \sqrt{5} . \tag{4.36}$$

If we make Δr_0 too small, the solution will remain in the near horizon regime and the angle traversed will exceed $\pi/2$. Hence, to find the smallest allowed Δr_0, we must pick it in such way that our approximation breaks down and can be matched onto the radially outward piece at almost $\pi/2$. To estimate (and bound from below) this value of Δr_0, we let $r(\theta) - r_H \approx r_H$ and take $\theta = \pi/2$ in our approximation (4.31):

$$\Delta r_0 \approx \frac{r_H}{\sinh\left(\sqrt{5}\frac{\pi}{2}\right)} = \# r_H \tag{4.37}$$

where # is again an $O(1)$ number. We conclude that for a small black hole in AdS_5, the shadow size is $O(r_H)$. A similar analysis confirms that for every $d \geq 3$ the property $\Delta r_0 \propto r_H$ holds, with the coefficient of proportionality decreasing for increasing d. As for the large black hole above, it is important to keep in mind that we presented only disk-shaped boundary regions. It is of course also possible to consider a strip on the boundary, but the small horizon area in this case ensures that the switchover to disconnected surface containing two disks will happen quite soon, which makes strips have even larger shadows.

While these results conclude our analysis of shadows for small black holes in AdS_{d+1}, we would like to end with a parenthetical remark which concerns extending these results to AdS_{d+1} times a compact manifold, as is often the case in concrete realizations of the holographic principle. For example, when considering a small black hole in $AdS_5 \times S^5$ (smeared uniformly over the S^5), one might be inclined to think that the correct minimal surface will be the AdS_5 solution as described above, uniformly wrapping the five-sphere. However, when the size of the AdS black hole is small w.r.t. the compact manifold, one can show that these black holes are Gregory-Laflamme unstable to localizing on the sphere [158, 159]. This means that the black hole will be an effective ten-dimensional one, and to find the associated minimal surfaces one should analyze it in the appropriate $10d$ background—interpolating between a $10d$ Schwarzschild geometry close to the black hole and an $AdS_5 \times S^5$ geometry far away. Although we did not analyze this case in detail, we expect that it will not qualitatively alter the above results.

4.4.3 Planar $SAdS_{d+1}$ with $d \geq 3$

To show that the shadow for a large black hole is exponentially small in any dimension $d \geq 3$, we can perform the analysis in a Poincaré patch of Schwarzschild-AdS_{d+1}, which is an excellent approximation in the large black hole limit. If we furthermore restrict ourselves to boundary strips, the enhanced symmetry of the problem will allow for an analytical treatment which confirms the exponential size of the shadow for arbitrary $d \geq 3$.

To proceed, we make the change of variables $z = r_H/r$ in the metric (4.26) and

consider the $r_H \gg 1$ limit:

$$\mathrm{d}s^2 \approx (1-z^d)\frac{-\,\mathrm{d}t^2}{z^2} + \frac{\mathrm{d}z^2}{z^2(1-z^d)} + \frac{r_H^2 d\Omega_{d-1}^2}{z^2}\ . \tag{4.38}$$

For boundary length scales $\theta_\infty \ll r_H$, we can take the boundary metric as approximately flat, $r_H^2\,\mathrm{d}\Omega_{d-1}^2 \approx \mathrm{d}\mathbf{x}_{d-1}^2$. We consider the strip with width $\theta_\infty = a r_H$ with $a \ll 1$ and assume that the strip is sufficiently wide that the deepest point to which the associated bulk minimal reaches, z_*, penetrates the near-horizon region, i.e., $z_* - 1 \ll 1$.

The action is given by:

$$S = \int \mathrm{d}^{d-2}\mathbf{x} \int \frac{\mathrm{d}x_1}{z^d}\sqrt{1+\left(\frac{\mathrm{d}z}{\mathrm{d}x_1}\right)^2 \frac{1}{1-z^d}}\ , \tag{4.39}$$

where x_1 is the transverse direction. This leads to the equation of motion:

$$\left(\frac{\mathrm{d}z}{\mathrm{d}x_1}\right)^2 = (1-z^d)\left(1-\left(\frac{z_*}{z}\right)^{2(d-1)}\right) \tag{4.40}$$

for which the width of the bulk probe is

$$\frac{\theta_\infty}{2} = \int_0^{\theta_\infty/2} \mathrm{d}x_1 = \int_0^{z_*} \mathrm{d}z \left|\frac{\mathrm{d}x_1}{\mathrm{d}z}\right| = \int_0^{z_*} \mathrm{d}z \left(\frac{z}{z_*}\right)^d \left[\left(1-z^d\right)\left(1-\frac{z^{2(d-1)}}{z_*^{2(d-1)}}\right)\right]^{-1/2}$$

which we may solve approximately by making the change of variables $u \equiv 1 - z/z_*$ and expanding for small u:

$$\begin{aligned}
\theta_\infty &= 2z_* \int_0^1 \mathrm{d}u (1-u)^d \left[\left(1-z_*^d(1-u)^d\right)\left(1-(1-u)^{2(d-1)}\right)\right]^{-1/2} \\
&\approx 2z_* \int_0^1 \mathrm{d}u \left[2u(d-1)\left(1-z_*^d+dz_*^d u\right)\right]^{-1/2} \\
&= \frac{2z_*}{\sqrt{2(d-1)z_*^d}} \cosh^{-1}\left(\frac{2dz_*^d+1-z_*^d}{1-z_*^d}\right)\ .
\end{aligned}$$

For ease of comparison with the higher-d solution in global SAdS (4.35), we make the further approximation $z_* \approx 1$,[12] under which the above expression simplifies

[12] This approximation is valid if θ_∞ is sufficiently large; this can be accomplished without violating $\theta_\infty \ll r_H$ by taking the large black hole limit, $r_H \gg 1$, which is precisely our current regime.

to:

$$z_* \approx 1 - 2d^2 \operatorname{sech}\left(\theta_\infty \sqrt{\frac{d-1}{2}}\right) \implies \Delta r_0 \approx 4 d r_H e^{-a r_H \sqrt{\frac{d-1}{2}}} \ . \tag{4.41}$$

We emphasize that this result is only valid for $\theta_\infty = a r_H \gg 1$ with $a \ll 1$ and $r_H \gg 1$. Although the calculation was done for a boundary strip and not a disk, the result (4.41) supports our claim that the shadow is exponentially small for large black holes in all $d \geq 3$.

4.5 Wilson loops

In this section, we turn to another bulk probe: static worldsheets arising from certain Wilson loops in the boundary CFT. The bulk dual of the expectation value of a Wilson loop $\mathcal{W}(\mathcal{C})$ evaluated in the supergravity limit is proposed to be [116]:

$$\mathcal{W}(\mathcal{C}) \sim e^{-S} \tag{4.42}$$

where S is the proper area of a fundamental string ending on the boundary loop $\mathcal{C}$; see figure 2.4. To simplify our analysis, we will consider rectangular Wilson loops that extend far into the past and future time-directions. Such a Wilson loop with temporal "height" T and spatial width $2\theta_\infty$ can be interpreted as the potential between a quark and an anti-quark [116, 160]. We assume sufficiently large T that the worldsheet may be considered invariant under time translations. The action for such a static worldsheet is given by

$$S = 2T \int_0^{\theta_\infty} \mathrm{d}\theta \sqrt{(\partial_\theta r)^2 + r^2 f(r)} \ . \tag{4.43}$$

Note that in static spacetimes this quantity takes the standard form of eqn. (4.1) with $F \propto \sqrt{-g_{tt}}$, thus we may treat it as a holographic probe similar to minimal area surfaces.

The action (4.43) does not explicitly depend on θ, so there is a conserved quantity that we shall use to write the equation of motion as a first order differential equation. We will find it convenient to distinguish two types of solutions to this equation:

∪-shaped worldsheets are smooth worldsheets anchored on the boundary that do not reach the black hole horizon, instead turning smoothly such that $\partial_\theta r|_{r=r_*} = 0$ at some finite $r_* > r_H$ (see figure 4.17).

⊔-shaped worldsheets consist of two straight segments that extend from the boundary to the black hole, joined discontinuously by a third segment that

partially wraps the horizon (see figure 4.17).

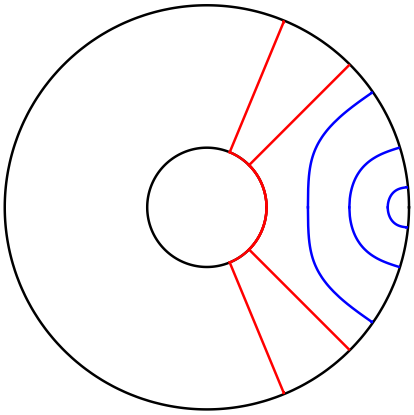


Figure 4.17: Worldsheets corresponding to different boundary angles for a BTZ black hole of radius $r_H = 0.5\ell_{\mathrm{AdS}}$. The ∪-shaped worldsheets are rendered in blue; ⊔-shaped, in red.

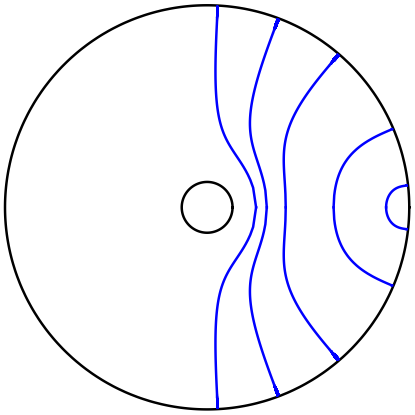


Figure 4.18: Worldsheets corresponding to different boundary angles for a BTZ black hole of radius $r_H = 0.2\ell_{\mathrm{AdS}}$. Small black holes in $d = 2$ are special, because the ∪-shaped worldsheet constitutes the leading saddle point for all values of θ_∞.

For a given boundary angle θ_∞, multiple solutions to the equation of motion may exist. Evaluation of the area functional is therefore necessary to determine which worldsheet constitutes the leading saddle point. Generally, we find that a switchover or phase transition occurs from ∪-shaped to ⊔-worldsheets, as illustrated in fig. 4.17. We discuss this behaviour in more detail below.

We first consider the smooth ∪-shaped solutions to the equation of motion. We can express the conserved charge in terms of the minimal/turning radius r_*. This allows us to find an implicit expression for θ_∞ in terms of r_* by integrating the equation of motion:

$$\theta_\infty(r_*) = \int_{r_*}^{\infty} \mathrm{d}r \frac{1}{r\sqrt{f(r)}} \frac{1}{\sqrt{\frac{r^2 f(r)}{r_*^2 f(r_*)} - 1}}. \tag{4.44}$$

Note that this formula only depends on the number of dimensions via $f(r)$, which is given by eqn. (4.27). $\theta_\infty(r_*)$ is plotted for the BTZ metric (cf. (4.15)) in figure 4.19. The function is characterized by a single maximum, and decreases monotonically for large r_*. Near the horizon however, $\mathrm{d}\theta_\infty/\,\mathrm{d}r_* < 0$, and hence by Lemma 3 (see appendix 4.A) there cannot exist any local minima of the area functional in this range. The ∪-shaped worldsheets thus suffer a shadow that

extends some finite distance from the horizon, but we postpone further discussion of shadows until after considering ⊔-shaped solutions as well.

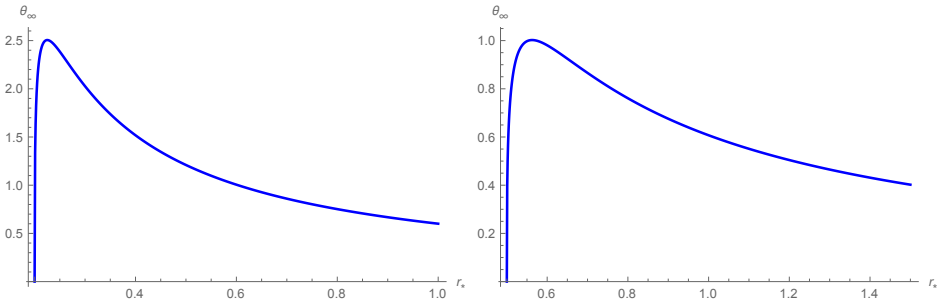


Figure 4.19: $\theta_\infty(r_*)$ for Wilson loops for a black hole of radius $r_H = 0.2\ell_{\rm AdS}$ (left) and $r_h = 0.5\ell_{\rm AdS}$ (right).

As an aside, we note that for $d = 2$, θ_∞ can be much larger than $\pi/2$. The solutions with $\theta_\infty > \pi/2$ correspond to strings that wind one or more times around the black hole; see figure 4.20. However, as mentioned above, strings that cross themselves fail to be minimal, so we can discard these solutions in what follows.

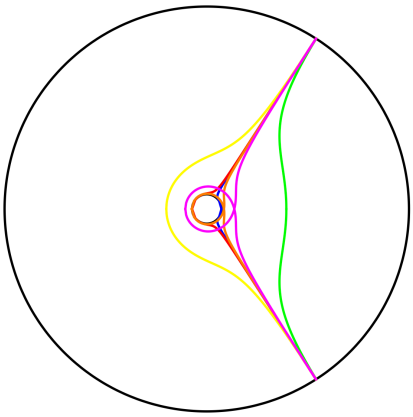


Figure 4.20: Extrema for $\theta_\infty = 1$ for a BTZ black hole with horizon radius $r_H = 0.1\ell_{\rm AdS}$. Only one of these saddle points – that with zero winding number (green) – corresponds to a global minimum of the proper area of the worldsheet.

We now turn to the ⊔-shaped solutions, which consist of two radial segments connecting the boundary and the horizon at $\pm\theta_\infty$, and a segment that wraps the horizon (see figure 4.17). The segment that wraps the horizon does not contribute to the area since the pullback of the metric vanishes. In stark contrast, the radial segments have divergent area; but this is associated with the unrenormalized self-energy of a quark-antiquark pair, and thus the Wilson loops associated to these ⊔-shaped strings do not encode information about the bulk. Nonetheless, because these ⊔-shaped solutions exist for all boundary angles, evaluation of the area functional is necessary to determine when the ∪-shaped solutions constitute the global minimum.

We find that ∪-shaped solutions have minimal area up to some critical angle θ_{switch}, beyond which ⊔-shaped solutions dominate. In general, this switchover will always occur for sufficiently large $\theta_\infty < \frac{\pi}{2}$. The only only exception is a small BTZ black hole, for which the minimal area worldsheets are ∪-shaped for all θ_∞. As shown in [2], for $d \geq 3$ one always has $\theta_{\text{switch}} < \frac{\pi}{2}$.

Denote the smallest radius to which the ∪-shaped worldsheets reach before the switchover by r_{s}. Then the switchover angle θ_{switch} and associated switchover radius r_{s} are determined by the equality of the areas of the ∪-shaped and ⊔-shaped solutions:

$$S_\cup(r_{\text{s}}) = S_\sqcup \ , \qquad \theta_\infty(r_{\text{s}}) \equiv \theta_{\text{switch}} \ . \tag{4.45}$$

The ∪-shaped worldsheet corresponding to the largest possible boundary angle θ_∞ penetrate deepest into the bulk. The switchover angle θ_{switch} is the largest angle for which the ∪-shaped solutions have minimal area, so the shadow radius r_{min} is determined by:

$$r_{\text{min}} = \text{Max}\left[\theta_\infty^{-1}(\pi/2),\ r_{\text{s}}\right] \ . \tag{4.46}$$

We can solve for the value of r_{s} by solving the area condition (4.45):

$$\int_{r_{\text{s}}}^{r_c} \frac{\mathrm{d}r}{\sqrt{1-\frac{r_{\text{s}}^2}{r^2}\frac{f(r_{\text{s}})}{f(r)}}} = \int_{r_H}^{r_c} \mathrm{d}r \implies \int_{r_{\text{s}}}^{\infty} \mathrm{d}r \left(\frac{1}{\sqrt{1-\frac{r_{\text{s}}^2}{r^2}\frac{f(r_{\text{s}})}{f(r)}}} - 1\right) = r_{\text{s}} - r_H \ . \tag{4.47}$$

where r_c is a large radial cutoff, necessitated by the fact that both actions are linearly divergent. The dimensional dependence is encapsulated in $f(r)$. For the BTZ metric, we can solve (4.47) exactly by taking $r_{\text{s}} = \lambda r_H$:

$$\lambda - 1 = \lambda \int_1^\infty \mathrm{d}x \left(\frac{1}{\sqrt{1-\frac{1}{x^2}\frac{\lambda^2-1}{x^2\lambda^2-1}}} - 1\right) \ , \tag{4.48}$$

which evaluates to $\lambda \approx 1.38$. We emphasize again that the BTZ metric is exceptional in the sense that there is no switchover for small black holes $r_H \lesssim 0.26\ell_{\text{AdS}}$.

In this case the ⊔-shaped worldsheets never constitute the leading saddle point of the area functional, even for $\theta_\infty > \pi/2$, and we find numerically that $r_{\min} \sim \ell_{\text{AdS}}$.

For $d > 2$ we cannot exactly solve (4.47) for the switchover radius, but we can obtain an approximation for large and small black holes. The former is especially well-motivated, since for large black holes there is a natural interpretation of the switchover as a "confinement-deconfinement" phase transition [161, 160, 162]. However, the analysis of these two cases is somewhat tedious; the interested reader is referred to [2], especially appendix C thereof. Here, we shall simply state the results, which are summarized in 4.2.

We note that locally, i.e., for a given $\theta_\infty < \theta_{\text{switch}}$, Wilson loops probe more deeply into the bulk than the corresponding minimal surface due to the extra factor of $\sqrt{-g_{tt}}$ in the action (4.43). But since the shadow radius $r_{\min}$ is the infimum of the collection of $r_*(\theta_\infty)$ from ∪-shaped worldsheets, we have to take into account the switchover effect in order to make the more appropriate global comparison.

	$d = 2$	$d = 3$	$d > 3$
$r_H \ll \ell_{\text{AdS}}$	O(1)	$\sim \sqrt{2r_H/\pi}$	$\sim \sqrt{4r_H/\pi}$
$r_H \gg \ell_{\text{AdS}}$	$\sim \lambda_2 r_H$, $\lambda_2 \approx 1.38$	$\sim \lambda_3 r_H$, $\lambda_3 \approx 1.46$	$\sim \lambda_d r_H$, $\lambda_d \gtrsim 1.52$

Table 4.2: Leading-order approximation of the shadow size $r_{\min} - r_H$ for Wilson loops. The proportionality constants are determined numerically via eqn. (4.47). .

4.6 Causal information surfaces

The third and final bulk probe we shall consider is the causal information surface [113], whose associated boundary quantity is dubbed "causal holographic information". This differs from the previous two probes in two ways. Firstly, its boundary CFT interpretation is unclear, although suggestions have been made in [150, 149]. Secondly, it does not take the general form we described in sec. 4.2.2 as a minimal geometric object. Nevertheless, it is still natural to define $\theta_\infty(r_*)$ for this probe. Thus we can study this probe alongside those above, and later make a comparison of their respective shadows.

The formal definition of the causal information surface is as follows: given a boundary region a, we first find its boundary causal diamond $\Diamond_a$, defined as the union of the boundary future and past domains of dependence of a:

$$\Diamond_a = D^+(a) \cup D^-(a) \ . \tag{4.49}$$

The causal information surface[13] $\Xi_{\mathcal{A}}$ is then defined as the intersection of the

[13]The geometry of causal information surfaces has been discussed in detail in [114]. In particular, note that for small AdS-Schwarzschild black holes and sufficiently large $\theta_\infty > \pi/2$, the

boundaries of the bulk future and past domains of influence $J^{\pm}(\Diamond_a)$[113]:

$$\Xi_a = \partial J^+(\Diamond_a) \cap \partial J^-(\Diamond_a) \ . \tag{4.50}$$

See figure 2.3.

In static, spherically symmetric spacetimes, we can understand this by reversing the construction. Start from a point in the bulk at radial coordinate r_*, and construct the two radially outgoing light rays to the future and past. These will end on two boundary points, $p_a^{\pm}$. The past boundary lightcone from p_a^+ and the future boundary lightcone from p_a^- enclose a causal diamond. The waist of diamond is exactly a boundary ball of radius θ_∞ that sits on the same timeslice as the initial bulk point. In other words,

$$\theta_\infty(r_*) = \int_{r_*}^{\infty} \mathrm{d}r \left|\frac{\mathrm{d}t}{\mathrm{d}r}\right| = \int_{r_*}^{\infty} \mathrm{d}r \sqrt{-\frac{g_{rr}}{g_{tt}}} = \int_{r_*}^{\infty} \frac{\mathrm{d}r}{f(r)} \ . \tag{4.51}$$

However, this is only true when $\theta_\infty < \pi$. When $\theta_\infty \geq \pi$, the ball covers the entire asymptotic boundary, and its domain of dependence is the entire spacetime. Therefore, there is an effective phase transition at $\theta_\infty = \pi$, and the shadow radius is given by

$$r_{\min} = \theta_\infty^{-1}(\pi) \ , \tag{4.52}$$

if this inverse exists. Otherwise there is no shadow.

In spacetimes with a horizon at r_H, $f(r) \to 0$ linearly as $r \to r_H$, thus $\theta_\infty \to \infty$, and such spacetimes will always exhibit shadows. For example, for the BTZ geometry with $f(r) = r^2 - r_H^2$, we have from (4.51),

$$\pi = \int_{r_{\min}}^{\infty} \frac{\mathrm{d}r}{r^2 - r_H^2} = \frac{1}{r_H} \operatorname{arccoth}\left(\frac{r_{\min}}{r_H}\right) \implies r_{\min} = \frac{r_H}{\tanh(r_H \pi)} \ . \tag{4.53}$$

Note that this is precisely the first term of (4.24)! In light of the earlier work by Hubeny [113], this similarity is not surprising. In the BTZ background, the causal information surface $\Xi_{\mathcal{A}}$ coincides with the extremal surface for a given boundary subregion. The only difference between their respective shadows is that the minimal area surfaces encounter a phase transition at some $\theta_\infty < \pi$ determined by the area matching condition (4.13). In particular, the phase transition for minimal area surfaces with a small black hole occurs when $\theta_\infty \sim \pi/2$, which makes a significant difference from the causal information surfaces. For large black holes, the minimal surface transition occurs at $\theta_\infty \lesssim \pi$, so these two probes agree with each other in this limit.

The situation is more complicated in higher dimensions [114]. For $d \geq 3$

surface consists of two parts, only one of which is connected to the boundary while the other encloses the black hole.

the integral in eqn. (4.51) is slightly more involved, but since we are primarily interested in knowing how close the surface gets to the black hole, a near-horizon approximation will suffice. Thus we assume $r_* - r_H \ll 1$ and expand the integrand in terms of $(r - r_H)$. For large black holes ($r_H \gg 1$), the near horizon contribution dominates θ_∞, so the phase transition happens when

$$\pi \approx \int_{r_{\min}}^{r_{\min}+a} \frac{\mathrm{d}r}{f'(r_H)(r - r_H)} = \frac{1}{f'(r_H)} \ln\left(\frac{r_{\min} - r_H + a}{r_{\min} - r_H}\right) \ . \tag{4.54}$$

where $a \lesssim r_H$ is some constant, and $f(r)$ is given by (4.27). Solving for $r_{\min}$, we find

$$r_{\min} \approx r_H + a e^{-d\pi r_H} \ . \tag{4.55}$$

Thus for large black holes, the causal information surfaces probe exponentially close to the horizon.

For small black holes ($r_H \ll 1$), the left-most side of (4.54) is instead $\pi/2$. This is because far from the horizon, the empty AdS region already contributes almost $\pi/2$ to the integral in (4.51). The solution is then

$$r_{\min} \approx r_H + a e^{-\frac{\pi(d-2)}{2r_H}} \ . \tag{4.56}$$

Thus causal surfaces also probe exponentially close to small black holes, which is dramatically better than minimal area surfaces in this limit (cf. (4.37)).

4.7 Discussion

4.7.1 Comparison of probes in AdS-Schwarzschild

In this section, we summarize our results by comparing the three probes – minimal area surfaces, Wilson loops, and causal information surfaces – for static black holes in asymptotically AdS space.

For $d = 2$, the calculation was sufficiently simple that we were able to obtain exact analytical results in all three cases; see figure 4.21 (left panel). As noted earlier, the shadow persist even when $r_H = 0$ due to the mass gap in AdS_{2+1}. The horizon radius is related to the ADM mass by $r_H^2 = GM - 1$, so a vanishing horizon does not recover empty AdS. In the right panel of figure 4.21, we extend the parameter range below the mass gap to include the conical defect. Then as $GM \to 0$, all shadows indeed disappear.

We can see clearly that causal information surfaces almost always leave the smallest shadow. This conclusion appears to hold in higher dimensions as well, as

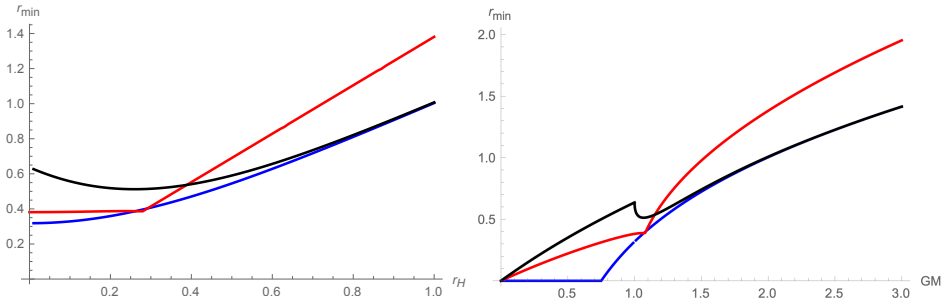


Figure 4.21: Shadow radius $r_{\min}$ as a function of the black hole radius r_H (left) and mass GM (right) for the different bulk probes: entanglement entropy (black), Wilson loops (red), and causal information (blue). The kink in the Wilson loops curves are due to the transition from ∪-shaped to ⊔-shaped worldsheets. The kink in the minimal area surface curve in the right panel is exactly at the horizon $r_H = 0$, at which point the phase transition angle becomes fixed at $\pi/2$.

indicated by our numerical results and approximations for both small and large black hole (see [2] for details). More quantitatively, both causal information and minimal area surfaces can probe exponentially close to the horizon of a large black hole, but the former can also probe exponentially close to a small black hole in $d \geq 3$. This fact, and more generally the relative shadow size between probes, can be understood be recalling their respective phase transition behaviours:

Minimal area surfaces encounter a phase transition for small black holes when $\theta_\infty \sim \pi/2$, so in this case are significantly worse than causal information surfaces. For large black holes, their phase transition angle approaches π, so they become comparable to causal information surfaces.

Static Wilson loops encounter a phase transition at exactly $\pi/2$ for small black holes in $d = 2$, and are thus comparable to minimal surfaces in this case. For large black holes or in higher dimensions, they encounter a deconfining phase transition when $\theta_\infty < \pi/2$, and thus suffer a larger shadow.

Causal information surfaces only encounter a phase transition when $\theta_\infty = \pi$. This enables them to probe most deeply into the bulk.

It is perhaps worth remarking on the comparison between casual information and extremal surfaces in relation to the earlier work [141]. There it was shown that if the bulk metric obeys the null energy condition, then the extremal surface anchored on a given boundary region b will lie outside (that is, deeper in the bulk than) the corresponding causal surface. In particular, this implies that the entanglement wedge covers more of the bulk than the casual wedge for the same boundary subregion. At first glance, this suggests that the entanglement wedge offers a stronger, or more complete reconstruction scheme that seems at odds with

our conclusion above. However, a key point of our analysis is that we are only interested in the surface of this bulk region, in the framework of the generalized minimal surfaces discussed in section 4.2, not with the entire bulk wedge. Although suggestions have been made for how one might reconstruct the spacetime within the entanglement wedge (see for example [101]), such reconstruction schemes are rather different from the geometric surface prescriptions considered here, and we leave them for another study.

It is interesting to note that for a point at radius $r_{\min}$, it may be that a given probe can only reach it with a specific orientation, implying a restriction on the accessibility of the bulk tangent space. Empty AdS satisfies the Strong Coverage Property, i.e., the entire tangent space of any point is covered, and indeed this property is necessary for certain reconstruction schemes [142, 145]. It is thus interesting to ask how much of the tangent space one loses due to the presence of a black hole. The reader is referred to [2] for a brief discussion of this issue.

4.7.2 Perspectives

A holographic duality such as AdS/CFT is an intriguing notion. In principle, every property of the bulk spacetime can be reconstructed from the combination of all boundary data. In practice, one seeks simple properties of the bulk that can be associated with particular observables in some subset of the boundary. The generalized geometric probe we defined in sec. 4.2.2 provides a continuous, infinite family of such associations between bulk codimension-1 surfaces and boundary regions. Two examples among them – the area of minimal surfaces and the action of Wilson loops – are known to have specific boundary observables.

In empty AdS space, these geometric probes faithfully scan through the entire bulk with full coverage of the tangent space at every point. We encapsulated this complete coverage in the Strong Coverage Property, which is a requisite for some specific reconstruction programs, such as recovering Einstein's equations or constructions relying on arbitrary shapes [60, 142, 163, 140]. However, when coverage of the bulk is incomplete – either through failure to cover the entire bulk or some portion of the tangent space – such reconstruction proposals fail.

Black holes are known to create unreachable regions, which we generically referred to as holographic shadows. In particular, these shadows are not limited to the black hole interior, but extend well beyond the horizon. Therefore, even if one replaces the black hole by a sufficiently dense (e.g., neutron) star, such shadows will persist. Proposals to reconstruct the bulk using smearing functions [102] in Lorentzian AdS/CFT encounter similar obstructions in the presence of trapped null geodesics [110]. In general, it appears that sufficient deviations from pure AdS will pose difficulties for straightforward attempts to completely cover the bulk, even for topologically trivial spacetimes.

It is very interesting to contemplate the implications of these holographic shadows in the context of AdS/CFT. Consider a minimal surface and a bulk field operator $\phi(x)$ inside the region demarcated by the surface, that is, between the surface and the boundary. It is widely believed that this bulk operator $\phi(x)$ can be described in terms of a CFT operator $O(x)$ which has support only in the boundary region defined by the endpoints of this minimal surface. However, if the spacetime exhibits shadows, then the CFT dual of any bulk operators located within the shadow region is less clear.

One can interpret this scenario in various ways. One possibility is that the CFT degrees of freedom that correspond to bulk operators within the shadow region are completely spread out over the boundary sphere. The shadow for a particular geometry would then imply a characteristic nonlocality in the boundary field theory below some IR cutoff. An alternative is that these degrees of freedom are encoded in a quantum secret sharing scheme [136, 101], an interpretation that follows from the switchover effect.[14] To see this, let us assume for concreteness that the shadow is caused by the presence of a black hole. The disconnected component that wraps the black hole in principle contains the entire bulk geometry down to the horizon, and one could hope that the CFT must therefore capture all the bulk physics between this surface and the boundary (notably including the shadow). In this picture, the boundary abruptly gains access to all bulk degrees of freedom in the shadow region (the "secret") after the phase transition, but contains no information before the switchover. We shall discuss quantum secret sharing in the context of bulk reconstruction further in the next chapter. Finally, one could conclude that the dual CFT simply does not capture everything that happens in the bulk. This would be the most radical point of view, and also the most unsatisfactory, since it would seem to imply that holographic reconstruction techniques, at least as presently understood, will always be incomplete.

To our knowledge, the only extant proposal that may have no shadows involves the bulk "entwinement" surfaces defined in [145]. However, these are dramatically different from the above geometric probes. The boundary data required to reconstruct entwinement surfaces is highly nonlocal, and cannot be associated with a particular subregion of the boundary. Aside from special cases in which the spacetime happens to be an integer quotient of pure AdS, the precise definition of this boundary data is hard to visualize. In light of our results, it seems appropriate to ask whether such explicitly nonlocal observables are necessarily required for holographic reconstruction, or whether there exists some simple geometric probe within our generalized framework that nonetheless leaves no shadow.

The bulk surfaces within this general class are naturally associated with boundary subregions, and hence to observables that are guaranteed to satisfy strong

[14] We thank Aron Wall for stimulating discussions on this issue.

sub-additivity. If there are indeed some probes that cast no shadows in the bulk, then we will have a transparent picture of emergent spacetime in this context. If on the other hand, one can prove that shadows are truly general features of such probes, then we have motivation to conclude that nonlocality will be an intrinsic feature of any successful holographic reconstruction scheme. Indeed, this will be an underlying theme in our study of precursors in the next chapter.

4.A Proofs

In this appendix, we present proofs of the two lemmas used in support of our coverage theorems. Note that Lemma 1 is not limited to globally regular geometries, while the form of Lemma 2 in the main text is. However, we will prove a more general version of Lemma 2 that is applicable to geometries with horizons and/or singularities. We also introduce and prove a third lemma, from which the coverage properties are independent, but which finds utility in the main text.

Lemma 1:

For a boundary sphere ∂a, the bulk surface b that minimizes L in eqn. (4.1) *with $\partial b = \partial a$ must be spherically symmetric.*

Proof:

If the minimal surface b is not spherically symmetric, one can rotate it to get a degenerate minimum b' of the same boundary region, with $\partial b = \partial b' = \partial a$. As shown in the left panel of fig. 4.22, b and b' must intersect, but it follows from the uniqueness theorem that their normal vectors cannot agree at the intersection. Thus they must intersect with a "kink". We assume for simplicity that this kink separates the surfaces into two regions each, but the generalization to multiple intersections is straightforward. Let b be separated into regions 1 and 2, and b' into 3 and 4 as depicted in fig. 4.22. By symmetry, regions 1 and 3 contribute the same amount to the geometric quantity L in eqn. (4.1), which we denote L_{13}. Similarly, we denote the contribution from regions 2 and 4 by L_{24}.

If $L_{24} > L_{13}$, then we could construct a new surface from regions 1 and 3 with the same boundary, thereby contradicting the assumption that both b and b' are minima. Similarly for $L_{13} > L_{24}$. If instead $L_{13} = L_{24}$, then both of the newly constructed surfaces have the same L as b and b'. But these new surfaces will not be smooth due to the kink at the intersection, so neither can be a local minimum of L. This again contradicts our assumption. QED

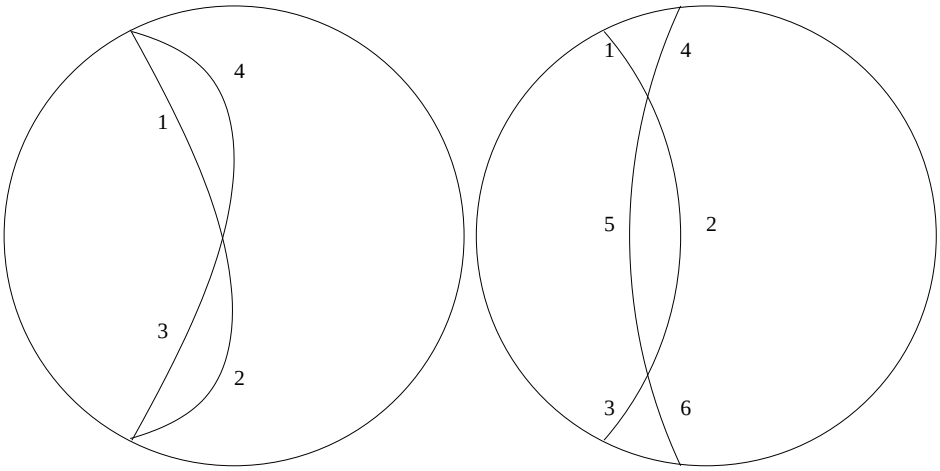


Figure 4.22: The left panel shows two non-spherically symmetric bulk surfaces, $b = (1 + 2)$ and $b' = (3 + 4)$, ending on the same spherical boundary, $\partial b = \partial b' = \partial a$. The right panel shows two intersecting bulk surfaces, $b = (1 + 2 + 3)$ and $b' = (4 + 5 + 6)$, whose corresponding boundary anchors do not intersect.

Lemma 2:

If the boundary anchors ∂b and $\partial b'$ do not cross each other, but the corresponding bulk surfaces b and b' do, and at least one connected region between b and b' does not contain a geometric obstruction, then b and b' cannot both be minimal surfaces.

Proof:

For this proof, we define a geometric obstruction as any object, defined purely by the metric, through which a bulk surface cannot be deformed without leaving a disconnected piece that wraps the obstruction; this wrapping piece should furthermore have a nonzero contribution to L in (4.1). (In other words, they are essentially generalizations of the black hole horizon in the case of minimal area surfaces.)

Refer to right panel of figure 4.22. Let $b = (1 + 2 + 3)$, $b' = (4 + 5 + 6)$, and assume there is no geometric obstruction within the volume enclosed between 2 and 5. We denote the contribution of region 5 as L_5, and the contribution of region 2 as L_2. If $L_2 > L_5$, then surface $(1 + 2 + 3)$ fails to be the minimum since surface $(1 + 5 + 3)$ has even smaller L. Similarly for $L_5 > L_2$. If $L_2 = L_5$, the uniqueness theorem again guarantees that the surface $(1 + 5 + 3)$ is not smooth, and thus we still arrive at a contradiction. Hence both b and b' cannot be global minima. QED

Lemma 3:

If $\mathrm{d}\theta_\infty / \mathrm{d}r_* > 0$, then the surface $b(r_*)$ cannot be a local minimum.

Proof:

By continuity, if $b(r_*)$ is a local minimum, there must be an infinitesimal δr such that $b(r_* + \delta r)$ is also a local minimum. Since $\mathrm{d}\theta_\infty / \mathrm{d}r_* > 0$, the corresponding boundary regions $a(r_* + \delta r)$ and $a(r_*)$ intersect exactly as in the right panel of fig. 4.22. Applying Lemma 2 to these two surfaces then implies that they cannot both be local minima. QED

5

Localizing precursors

This chapter is based on [3].

The issues of locality in the context of bulk reconstructed alluded to in the previous chapter are perhaps no better illustrated than by *precursors*. These embody the puzzling fact that a single bulk operator can be mapped to multiple different boundary operators, and we shall investigate this in the present chapter. By improving upon a recent model of Mintun, Polchinski, and Rosenhaus, we demonstrate explicitly how this ambiguity arises in a simple model of the field theory. In particular, we show how gauge invariance in the boundary theory manifests as a freedom in the smearing function used in the bulk-boundary mapping, and explicitly show how this freedom can be used to localize the precursor in different spatial regions. We also show how the ambiguity can be understood in terms of quantum error correction, by appealing to the entanglement present in the CFT. The concordance of these two approaches suggests that gauge invariance and entanglement in the boundary field theory are intimately connected to the reconstruction of local operators in the dual spacetime.

5.1 Introduction

In AdS/CFT, much interest has focused on the emergence of the bulk spacetime from boundary CFT data, but a complete understanding of bulk locality remains elusive. The boundary dual of a local bulk field Φ located a finite distance z away from the boundary has a remarkably simple formula in terms of an integral of the corresponding local CFT operator $\mathcal{O}$ over space and time:

$$\Phi(t,x,z) = \int \mathrm{d}x'dt' K\left(t,x,z|x',t'\right)\mathcal{O}\left(x',t'\right) + O(1/N) \tag{5.1}$$

where the kernel K is called the *smearing function*. In the cases where K exists and can be computed, its support on the boundary is a measure for what subregion of the boundary stores the information of a given bulk point. The above construction is often referred to as the HKLL construction, following the extensive work by the eponymous authors [103, 102, 104, 105].

A perplexing feature of this procedure is that there is a freedom in choosing the smearing function K, allowing for a family of different CFT operators corresponding to a given bulk operator. These different CFT operators, when evolved back to one time[1], can even have support in different spatial regions of the CFT (see figure 5.1). We refer to these CFT operators as "precursors", because in general they contain information about bulk events before signals from these events have had time to reach the boundary [164, 165, 166].

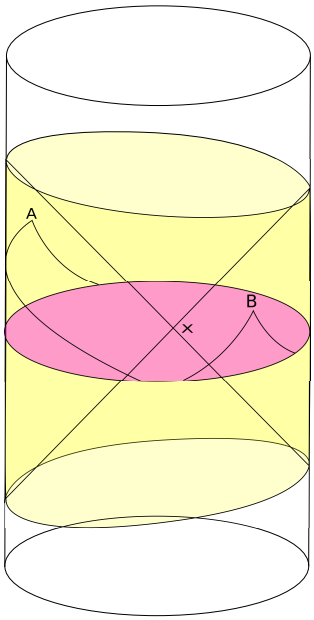


Figure 5.1: Global AdS_3, showing the lightcone for a bulk point x, which defines a spacelike separated region on the boundary (shaded, yellow online). The corresponding nonlocal boundary operator is defined à la (5.1) as an integral over this region. The local CFT operators can be time-evolved to a single Cauchy slice (shaded, pink online). This is illustrated schematically for points A and B, where we've indicated the null lines on the boundary. In our model, the boundary operators factorize along the lightcone directions, and are trivially evolved to bilocals at the $t = 0$ Cauchy slice.

[1] As mentioned in chapter 2, there are strictly speaking no Cauchy surfaces in AdS. However, the Cauchy problem is still well-posed within the domain of dependence set by the boundary of the surface, and the "Cauchy slices" discussed in this chapter are to be understood in this limited (slightly abusive) sense.

Almheiri, Dong, and Harlow [101] pointed out that these different CFT operators cannot be really equal as operators unless the field theory violates the time slice axiom, which is believed to be a fundamental property of physically relevant quantum field theories [167]. These authors proposed that the different CFT operators are only equivalent when acting on a certain subclass of states (the "code subspace"), casting the bulk reconstruction problem in the language of quantum error correction (QEC) and quantum secret sharing [136], in which increasing radial depth into the bulk is interpreted as improved resilience of the boundary theory against local quantum erasures. This idea has been beautifully implemented in several tensor network models [168, 169, 170].

Subsequently, Mintun, Polchinski, and Rosenhaus ("MPR") [171] argued that the structure of QEC emerges naturally when one considers the gauge invariance of the boundary field theory. MPR reconcile the representation of a local bulk operator by a number of different CFT operators by pointing out that an operator can be modified by a "pure-gauge" contribution that changes its support on the boundary without changing its action on physical states. This suggests that the emergence of local operators in the dual spacetime may be deeply connected with gauge symmetries in the CFT.[2]

In this work, we clarify the relationship between quantum error correction, gauge freedom, and the localization of precursors, in the context of an explicit bulk reconstruction scheme in AdS_3/CFT_2. We will first point out a shortcoming of the MPR model: the particular boundary conditions specified by MPR lead to a theory with no bulk dynamics[3]. This difficulty is easily fixed by choosing different boundary conditions, and we revise their model of the CFT accordingly in section 5.2. We show that with these revisions, the MPR model works as advertised, and provides a nice, tractable model for understanding the CFT encoding of bulk information, including such issues as the role of quantum error correction.

We show in section 5.3 that the known ambiguity in the choice of smearing function arises from the gauge freedom in the $N \to \infty$ limit, and explicitly show how to use this freedom to localize the precursor in different spatial regions. We begin with a standard representation of a local bulk operator spread over the entire CFT, as illustrated in figure 5.1, and show that the gauge freedom allows us to localize the precursor within a single boundary Rindler wedge. This result agrees with the claims of MPR, but now in a model with genuine bulk dynamics. We find that this localization procedure works when the bulk field is located inside the corresponding entanglement wedge[4], consistent with general expectations for bulk reconstruction. This result is independent of the weakly coupled CFT model,

[2] Similar ideas have also been emphasized in [172] and related works.

[3] We thank Ian Morrison for discussions on this point.

[4] For most of this work, the entanglement and causal wedges agree, and we use the terms interchangeably. We will address the crucial difference between them in section 5.5.

and relies only on the freedom in the choice of smearing function.

In section 5.4, we instead take a quantum error correction approach to localizing the precursor in a boundary region A: we use the entanglement of the ground state to map operators acting on the complement $\bar{A}$ into operators acting on A. We point out how this procedure can fail, and show that it is successful when the above condition is satisfied: the bulk point must lie in the entanglement wedge corresponding to the boundary region under consideration.

Finally, we conclude in section 5.5 with some discussion on the relationship between quantum error correction and gauge freedom in light of our results, and speculate on how our model may be generalized to disconnected boundary regions, where the distinction between the causal and entanglement wedges is significant.

5.2 Improved toy model of the bulk-boundary correspondence

In this section, we describe the MPR model [171] along with our improvements. In the former, the CFT consists of free massless scalars ϕ^i in two dimensions, where i is a global $O(N)$ index. The global $O(N)$ symmetry is a simple model for the gauge invariance of the full theory, so "gauge-invariant" operators are defined to be operators that are invariant under global $O(N)$ transformations. MPR consider a massless bulk field Φ in AdS_3, and from the two possible consistent quantization schemes [90],

$$\Delta_\pm = \frac{d}{2} \pm \sqrt{\frac{d}{2} + m^2} \tag{5.2}$$

they choose boundary conditions such that the bulk field is dual to a $\Delta_- = 0$ operator, which they take to be $\phi_i \phi^i$.

The choice $\Delta = 0$ is unfortunate for a number of related reasons. From the CFT point of view, $\Delta = 0$ saturates the unitarity bound. In any dimension, an operator $\mathcal{O}$ saturating the unitarity bound must obey the *boundary* equation of motion $\Box \mathcal{O} = 0$, meaning that it acts like a free field on the boundary[5]. From the bulk point of view, when we impose the boundary condition $\Phi \propto z^{\Delta_-}$ as $z \to 0$, with $\Delta_- = 0$, there are no solutions to the bulk equation of motion except for the special modes satisfying the boundary wave equation. Furthermore, $\phi_i \phi^i$ isn't really a local operator in the CFT, since it's correlation functions are logarithmic. Therefore, this field does not have true bulk dynamics and is not a good setting to discuss bulk reconstruction; see [173] for a more detailed discussion of the $\Delta = 0$ limit.

This problem is easily fixed: we simply choose the other boundary condition

[5] We thank Ian Morrison for pointing this out.

$\Phi \to z^2$ ($\Delta_+ = 2$), and take the boundary operator dual to the bulk field to be

$$\mathcal{O} = \partial_\mu \phi^i \partial^\mu \phi_i \ , \tag{5.3}$$

where the ϕ^i are free massless scalar fields as in the MPR model. Strictly speaking, this is also a poor model for perturbative bulk physics, since the CFT is weakly coupled. However, at the level of two-point functions it suffices to capture the salient features. This improved model is almost identical to [166], which in turn was closely related to [164].

In the following we suppress the $O(N)$ index i and use lightcone coordinates $x_\pm = t \pm x$ in the boundary, so we can write simply

$$\mathcal{O} = \partial_+ \phi \partial_- \phi \ . \tag{5.4}$$

We expand the CFT field ϕ in terms of creation and annihilation operators as

$$\phi(x_+, x_-) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\nu_+}{\nu_+} \alpha_{\nu_+} e^{-i\nu_+ x_+} + \int_{-\infty}^{\infty} \frac{\mathrm{d}\nu_-}{\nu_-} \tilde{\alpha}_{\nu_-} e^{-i\nu_- x_-} \ . \tag{5.5}$$

where α and $\tilde{\alpha}$ correspond to the right and left movers, respectively. This then yields a simple formula for the "primary" operator $\mathcal{O}$,

$$\mathcal{O}(x_+, x_-) = -\int_{-\infty}^{\infty} \mathrm{d}\nu_+ \, \mathrm{d}\nu_- e^{-i(\nu_+ x_+ + \nu_- x_-)} \alpha_{\nu_+} \tilde{\alpha}_{\nu_-} \ . \tag{5.6}$$

In the large N limit, MPR pointed out that the global $O(N)$ gauge invariance includes the freedom to add to any operator a linear combination of operators of the form $\alpha_{\nu_+} \tilde{\alpha}_{\nu_-}$ as long as $\nu_+ \nu_- < 0$.[6]

5.3 Localizing the precursor via gauge freedom

The freedom identified by MPR at first appears distinct from the freedom in the choice of smearing function, but we will show that they are in fact identical. We will then show explicitly how this freedom can be used to localize the precursor within a given boundary region, in an effort to make more precise the role that gauge invariance plays in the localization and non-uniqueness of boundary data.

The precursor for a local bulk field Φ is defined with support on the entire boundary by eqn. (5.1), in the $N \to \infty$ limit,

$$\Phi(t, x, z) = \int \mathrm{d}x' \, \mathrm{d}t' K\left(t, x, z | x', t'\right) \mathcal{O}\left(x', t'\right) \ . \tag{5.7}$$

[6]Note that this construction implicitly assumes a restriction to the low-energy subspace of states. We shall return to this point in sec. 5.5.

The smearing function K in Poincaré-AdS$_3$ for a field with conformal dimension $\Delta = 2$ is given by [103]

$$K(t, x, z|t', x') = \log\left(\frac{|z^2 + (x - x')^2 - (t - t')^2|}{2z}\right) \equiv K. \tag{5.8}$$

The ambiguity in the smearing function consists of the freedom to add a function δK which in Fourier space satisfies $\nu_+\nu_- < 0$. This can be understood from the fact that satisfying the bulk wave equation, $\Box\Phi = 0$, in global AdS implies that modes with frequency ω and boundary momentum κ satisfying $\omega^2 < \kappa^2$ (or in lightcone coordinates, $\nu_+\nu_- < 0$) are disallowed. Hence the dual operator has no support on the space of these modes, $\int \mathrm{d}^2x\ \mathcal{O}\ \delta K = 0$.

Focusing on a particular Fourier mode, the change in the smearing function is

$$\delta K(x_+, x_-) = e^{i(\nu_+ x_+ + \nu_- x_-)}\ . \tag{5.9}$$

The corresponding change in the precursor is therefore

$$\delta\Phi = \int \mathrm{d}x_+\, \mathrm{d}x_- e^{i(\nu_+ x_+ + \nu_- x_-)}\mathcal{O}\ . \tag{5.10}$$

Plugging in the expansion for the field in terms of creation and annihilation operators (5.6) then gives

$$\delta\Phi = -\int \mathrm{d}x_+\, \mathrm{d}x_- e^{i(\nu_+ x_+ + \nu_- x_-)} \int \mathrm{d}\nu_+{}'\, \mathrm{d}\nu_-{}' e^{-i(\nu_+{}' x_+ + \nu_-{}' x_-)}\alpha_{\nu_+{}'}\tilde{\alpha}_{\nu_-{}'}\ . \tag{5.11}$$

The spatial integrals can be performed, yielding

$$\delta\Phi = -\alpha_{\nu_+}\tilde{\alpha}_{\nu_-}\ . \tag{5.12}$$

This demonstrates that the freedom identified in MPR corresponds precisely to the freedom in the choice of smearing function. In this sense, we will refer to the function δK satisfying $\nu_+\nu_- < 0$ as "pure gauge" henceforth.

We are now prepared to investigate the idea of MPR in the context of an explicit HKLL construction [103], by demonstrating that the gauge freedom can be used to localize the precursor to within a single boundary Rindler wedge. In Poincaré lightcone coordinates, the metric for Rindler-AdS$_3$ is

$$\mathrm{d}s^2 = \frac{-\,\mathrm{d}x_+\, \mathrm{d}x_- + \mathrm{d}z^2}{z^2}\ , \tag{5.13}$$

which naturally leads to a bulk Rindler horizon at $x_+ = x_- = 0$. This horizon defines the bulk Rindler or causal wedge, and our aim is to localize the precursor

for a given field in this wedge within the corresponding boundary region. This requires finding the most general pure-gauge function δK that we can add such that the new smearing function, $\hat{K} = K + \delta K$, only has support within that region.

To proceed, we need to know how the pure-gauge mode functions (that is, the Poincaré modes with $\nu_+\nu_- < 0$) look in the various Rindler wedges. This can be done by studying the analyticity of the mode functions in the complex plane.

It is convenient to work in terms of the Rindler modes[7]

$$x_+^{i\omega_+} x_-^{i\omega_-} \ . \tag{5.14}$$

where $x_\pm$ are the Poincaré lightcone coordinates, as above. The Rindler plane is sketched in fig. 5.2. The above Rindler modes (5.14) are then defined as-is in the northern quadrant, where $x_+ > 0$, $x_- > 0$. We would then like to know what this looks like in the remaining three quadrants. However, getting there requires navigating the branch cuts at $x_+ = 0$ and/or $x_- = 0$.

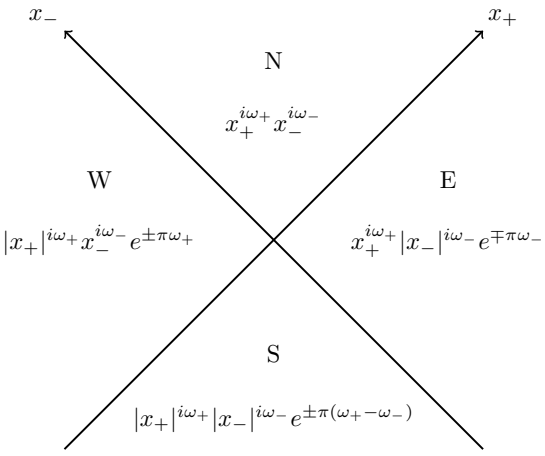


Figure 5.2: Rindler plane in lightcone coordinates, indicating the phase changes in the mode functions (5.14) when crossing the branch cuts at $x_\pm = 0$. The sign choice is arbitrary, but must be consistent across all four quadrants in order to obtain a pure-gauge Poincaré mode. We refer to these quadrants throughout as the northern (N), southern (S), eastern (E), and western (W) wedges, labelled in the obvious manner.

Consider moving into the western wedge. We have a choice of contour upon crossing the branch cut at $x_+ = 0$. Suppose we take the function to be analytic in the lower-half complex x_+ plane. Then the transformation from the northern

[7]In fact, working in terms of the Rindler modes is more than a convenience, it is a necessity, because the resulting smearing function can only be written in Fourier space; it cannot be transformed to position space.

wedge ($x_+ > 0$) across $x_+ = 0$ into the western wedge ($x_+ < 0$), is $x_+ \to |x_+|e^{-i\pi}$, where the minus sign in the exponential corresponds to our choice of contour. The Rindler mode changes as

$$x_+^{i\omega_+} x_-^{i\omega_-} \to |x_+|^{i\omega_+} x_-^{i\omega_-} e^{\pi\omega_+} \ . \tag{5.15}$$

Since we chose x_+ to be analytic in the lower half-plane, our mode is a superposition of positive frequency Poincaré modes $\nu_+ > 0$.[8] Had we made the opposite choice for the analyticity of the function, we would take the opposite sign of ν_+. Hence the general transformation across the $x_+ = 0$ branch cut into the western wedge is

$$x_+^{i\omega_+} x_-^{i\omega_-} \to |x_+|^{i\omega_+} x_-^{i\omega_-} e^{\pm\pi\omega_+} \qquad (\mathrm{N} \to \mathrm{W}) \tag{5.16}$$

where the upper sign is for $\nu_+ > 0$, lower for $\nu_+ < 0$. From this relation one immediately writes down the transformation from the northern quadrant across $x_- = 0$ into the east ($x_- < 0$):

$$x_+^{i\omega_+} x_-^{i\omega_-} \to x_+^{i\omega_+} |x_-|^{i\omega_-} e^{\mp\pi\omega_-} \qquad (\mathrm{N} \to \mathrm{E}) \tag{5.17}$$

where the upper sign is for $\nu_- < 0$, lower for $\nu_- > 0$. Similarly, the transformation of the precursor into the southern quadrant, with two branch crossings, is

$$x_+^{i\omega_+} x_-^{i\omega_-} \to |x_+|^{i\omega_+} |x_-|^{i\omega_-} e^{\pm\pi\omega_+ \mp\pi\omega_-} \qquad (\mathrm{N} \to \mathrm{S}) \ . \tag{5.18}$$

The crucial fact is that the above, with a consistent sign choice (upper or lower), corresponds to a pure-gauge function in Poincaré, since we have $\nu_+\nu_- < 0$ by construction. The nice feature of this method is that we're guaranteed this without having to explicitly work with Poincaré modes, where the meaning of $\nu_+\nu_- < 0$ in the various quadrants is not readily visualized.

From this analyticity analysis, we can immediately write down the general form of the pure-gauge function δK:

$$\delta K = \int \mathrm{d}\omega_+ \, \mathrm{d}\omega_- \left(c_{\omega_+\omega_-} f^{\text{upper}}_{\omega_+\omega_-} + d_{\omega_+\omega_-} f^{\text{lower}}_{\omega_+\omega_-} \right) \tag{5.19}$$

with, from (5.16), (5.17), and (5.18),

$$\begin{aligned} f_{\omega_+\omega_-} &= x_+^{i\omega_+} x_-^{i\omega_-} \Theta(x_+)\Theta(x_-) + |x_+|^{i\omega_+} x_-^{i\omega_-} e^{\pm\pi\omega_+} \Theta(-x_+)\Theta(x_-) \\ &\quad + x_+^{i\omega_+} |x_-|^{i\omega_-} e^{\mp\pi\omega_-} \Theta(x_+)\Theta(-x_-) \\ &\quad + |x_+|^{i\omega_+} |x_-|^{i\omega_-} e^{\pm\pi\omega_+ \mp\pi\omega_-} \Theta(-x_+)\Theta(-x_-) \ . \end{aligned} \tag{5.20}$$

[8] Any function $f(x)$ built out of positive frequency Fourier modes (that is $e^{-i\nu x}$ with $\nu > 0$) must be analytic in the lower half of the complex x-plane, and vice versa.

The labels "upper" and "lower" on the functions f in (5.19) indicate choosing the upper or lower signs in the exponentials in (5.20), and the coefficients c and d are undetermined functions of the momenta.

We may extract from this general expression the pure-gauge function in momentum space, $\delta\tilde{K}$, in each of the four quadrants:

$$\begin{aligned}
\delta\tilde{K}_N &= c + d \\
\delta\tilde{K}_W &= e^{\pi\omega_+}c + e^{-\pi\omega_+}d \\
\delta\tilde{K}_S &= e^{\pi(\omega_+ - \omega_-)}c + e^{-\pi(\omega_+ - \omega_-)}d \\
\delta\tilde{K}_E &= e^{-\pi\omega_-}c + e^{\pi\omega_-}d
\end{aligned} \tag{5.21}$$

where we have suppressed the $\omega_\pm$ subscripts on c and d to minimize clutter.

In order to localize support for the precursor within a single Rindler wedge, we must choose the coefficients c and d such that $\hat{\tilde{K}} = \tilde{K} + \delta\tilde{K}$ is zero in the other three regions. Let us attempt to localize the precursor in the east, so that only $\hat{\tilde{K}}_E \neq 0$. Then the coefficients must be chosen such that

$$\delta\tilde{K}_N = -\tilde{K}_N \ , \quad \delta\tilde{K}_W = -\tilde{K}_W \ , \quad \text{and} \quad \delta\tilde{K}_S = -\tilde{K}_S \tag{5.22}$$

where $\tilde{K}_X$ with $X \in \{E, N, W, S\}$ is the Fourier transform of the smearing function (5.8) in the specified wedge,

$$\tilde{K}_X\left(\omega_+, \omega_-\right) \equiv \iint_X \mathrm{d}x_+ \, \mathrm{d}x_- K(0, a, z|x_+, x_-)|x_+|^{-i\omega_+ - 1}|x_-|^{-i\omega_- - 1} \ . \tag{5.23}$$

We have chosen the bulk field to be located at time $t = 0$, radial coordinate z, and a distance a into the eastern wedge of the bulk.

At a glance, the system (5.22) appears overdetermined, as we have three equations and only two unknowns, c and d. However, we shall find that the system does indeed have a consistent solution, provided that the bulk point lies within the bulk extension (the causal or entanglement wedge) of the boundary Rindler wedge in which we attempt to localize the smearing function, in this case the east. We shall return to this requirement below.

In the course of solving this system, we rely on the following relations between the Fourier transforms of the smearing function, which we prove in appendix 5.A:

$$\begin{aligned}
\tilde{K}_N &= \cosh(\pi\omega_+)\tilde{K}_W \\
\tilde{K}_S &= \cosh(\pi\omega_-)\tilde{K}_W \\
\tilde{K}_E &= \cosh\left(\pi(\omega_+ - \omega_-)\right)\tilde{K}_W
\end{aligned} \tag{5.24}$$

Note that the singularities in the smearing function (5.8), which occur when the

argument of the logarithm is zero, do not extend into the western quadrant. This is a consequence of the fact that we chose the bulk point to be in the eastern Rindler wedge. The benefit of these relations is that they allow us to rewrite everything in terms of the Fourier transform $\tilde{K}_W$, which is well-defined.

With the relations (5.24) in hand, one can show that the system (5.22) is solved by

$$c = -\frac{1}{2}e^{-\pi\omega_+}\tilde{K}_W \qquad d = -\frac{1}{2}e^{\pi\omega_+}\tilde{K}_W \ . \tag{5.25}$$

and therefore that the only non-zero portion of the momentum space smearing function, $\hat{\tilde{K}}_E$, is

$$\begin{aligned}\hat{\tilde{K}}_E \equiv \tilde{K}_E + \delta\tilde{K}_E &= \cosh\left(\pi(\omega_+ - \omega_-)\right)\tilde{K}_W + \left(e^{-\pi\omega_-}c + e^{\pi\omega_-}d\right) \\ &= -2\sinh(\pi\omega_+)\sinh(\pi\omega_-)\tilde{K}_W \ .\end{aligned} \tag{5.26}$$

It then remains to obtain an explicit expression for $\tilde{K}_W$, which we can do by computing the Fourier transform of (5.8) in the western Rindler wedge. The integration is performed in appendix 5.B. Substituting the result into (5.26), we have

$$\hat{\tilde{K}}_E = -2\pi^2\left(\frac{z}{a}\right)^2 a^{-i(\omega_+ + \omega_-)}\,{}_2F_1\left(1 + i\omega_+, 1 + i\omega_-, 2, \frac{-z^2}{a^2}\right) \ , \tag{5.27}$$

which is consistent with results found in the literature [102]. We therefore find that the smeared bulk operator (5.7) at $t = 0$, $x = a > 0$, and radial distance z, with support localized entirely within the eastern Rindler wedge, is given by

$$\begin{aligned}\Phi(0, a, z) = &-2\pi^2\left(\frac{z}{a}\right)^2 \int \mathrm{d}\omega_+\,\mathrm{d}\omega_- a^{-i(\omega_+ + \omega_-)} \\ &\times {}_2F_1\left(1 + i\omega_+, 1 + i\omega_-, 2, \frac{-z^2}{a^2}\right)\tilde{\mathcal{O}}^E_{\omega_+,\omega_-} \ ,\end{aligned} \tag{5.28}$$

where $\tilde{\mathcal{O}}^E_{\omega_+,\omega_-}$ is the momentum-space boundary operator, with support in the eastern wedge. We will write this explicitly in Rindler modes (cf. the Poincaré expression (5.6)) in the next section, but forgo unnecessary details here.

The action of the precursor (5.28) is UV-sensitive, and only well-defined when acting on an appropriate subclass of states. As we show explicitly in appendix 5.C, its vacuum two-point function reproduces the correct bulk correlator in the near-horizon limit.

As mentioned previously, a condition on the success of our procedure is that the bulk point be located in the entanglement wedge of the boundary region in which we attempt to localize the precursor. A natural question to ask is whether

the gauge freedom in the smearing function can still be used to reconstruct a bulk point located outside the entanglement wedge. As we placed our bulk point in the eastern wedge, this would amount to trying to set the smearing function to zero in the eastern quadrant instead of the western quadrant as we did above. The set of conditions on the pure-gauge function $\delta\tilde{K}$ is then

$$\delta\tilde{K}_N = -\tilde{K}_N \ , \quad \delta\tilde{K}_E = -\tilde{K}_E \ , \quad \text{and} \quad \delta\tilde{K}_S = -\tilde{K}_S \ . \tag{5.29}$$

One can simply check that the overdetermined system of equations (5.21), (5.24) and (5.29) is now inconsistent: there no longer exists a solution for c and d. Hence we conclude that our model is consistent with the current understanding of bulk reconstruction, namely that it succeeds when the bulk point is inside – and fails when the point is outside – the causal/entanglement wedge. We shall comment more on this in section 5.5, and elaborate on the distinction between the two types of bulk wedges, but first we turn to an alternative approach of localizing the bulk field, appealing instead to the entanglement structure in the CFT.

5.4 Localizing the precursor via entanglement mapping

In this section we will present an alternative method for localizing the precursor. As before, our starting point is the smeared operator in Poincaré coordinates (5.7), which has non-zero support on the entire boundary and can be time-evolved to bilocals at $t = 0$. Instead of using the gauge freedom to manipulate the support of the smearing function K however, we will now use entanglement in the field theory to map all bilocal operators into the eastern Rindler wedge. We will explicitly show that this gives the same result as that obtained in the previous section, thereby establishing that the freedom in the smearing function from gauge invariance can be equivalently understood from an entanglement perspective.

In eqn. (5.5), we expanded the CFT field ϕ in terms of Poincaré modes. We may equivalently write the mode expansion in terms of Rindler creation ($\omega < 0$) and annihilation ($\omega > 0$) operators $\beta_{\omega_\pm}$ with left- and right-moving[9] Rindler momenta $\omega_\pm$. These satisfy

$$[\beta_{\omega_\pm}, \beta_{\omega'_\pm}] = \omega_\pm \delta(\omega_\pm + \omega'_\pm) \qquad \text{and} \qquad \beta^\dagger_{\omega_\pm} = \beta_{-\omega_\pm} \ . \tag{5.30}$$

In lightcone coordinates $x_\pm \equiv t \pm x$, the Rindler expansion of the field in the

[9]In this section, to avoid clutter, we denote right movers by β_{ω_+} and left movers by β_{ω_-}, with no tilde on the left movers. Left and right movers commute.

eastern and western wedges (cf. fig. 5.2) is, respectively,

$$\phi^E(t,x) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega_+}{\omega_+}\, \beta^E_{\omega_+} |x_+|^{-i\omega_+} + \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega_-}{\omega_-}\, \beta^E_{\omega_-} |x_-|^{i\omega_-} \tag{5.31}$$

$$\phi^W(t,x) = \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega_+}{\omega_+}\, \beta^W_{\omega_+} |x_+|^{i\omega_+} + \int_{-\infty}^{+\infty} \frac{\mathrm{d}\omega_-}{\omega_-}\, \beta^W_{\omega_-} |x_-|^{-i\omega_-} \ , \tag{5.32}$$

where the Rindler mode functions are chosen such that they are positive frequency with respect to Rindler time, which we take to run upwards in both the eastern and western wedge. The lightcone derivatives are

$$\partial_+\phi^E = -i \int_{-\infty}^{+\infty} \mathrm{d}\omega_+ \beta^E_{\omega_+} |x_+|^{-i\omega_+ - 1} \qquad \partial_-\phi^E = i \int_{-\infty}^{+\infty} \mathrm{d}\omega_- \beta^E_{\omega_-} |x_-|^{i\omega_- - 1} \tag{5.33}$$

$$\partial_+\phi^W = i \int_{-\infty}^{+\infty} \mathrm{d}\omega_+ \beta^W_{\omega_+} |x_+|^{i\omega_+ - 1} \qquad \partial_-\phi^W = -i \int_{-\infty}^{+\infty} \mathrm{d}\omega_- \beta^W_{\omega_-} |x_-|^{-i\omega_- - 1} \tag{5.34}$$

and are manifestly purely left/right-moving. As a consequence, their time evolution becomes trivial:

$$\partial_+\phi(t,x) = \partial_+\phi(0, x+t) \qquad \partial_-\phi(t,x) = \partial_-\phi(0, x-t) \ . \tag{5.35}$$

This was to be expected, since ϕ satisfies the 1+1-dimensional wave equation $\Box\phi = 0$. This factorization along the null directions allows us to write the precursor, for a bulk operator shifted a distance a into the east, as a bilocal at $t = 0$:

$$\Phi(t = 0, x = a > 0, z) = \int \mathrm{d}x_+\, \mathrm{d}x_- K(0, a, z | x_+, x_-) \partial_+\phi(0, x_+) \partial_-\phi(0, -x_-) \ , \tag{5.36}$$

where the smearing function (5.8), in lightcone coordinates, is

$$K(0, a, z | x_+, x_-) = \log\left(\frac{|z^2 - (x_+ - a)(x_- + a)|}{2z} \right) \equiv K \ . \tag{5.37}$$

Using (5.33) and (5.34), we can explicitly decompose the integral (5.36) over all

four wedges:

$$\begin{aligned}\Phi(0,a,z) = &-\iint_N \mathrm{d}x_+\,\mathrm{d}x_- \iint \mathrm{d}\omega_+\,\mathrm{d}\omega_- K|x_+|^{-i\omega_+-1}|x_-|^{-i\omega_--1}\beta^E_{\omega_+}\beta^W_{\omega_-}\\ &-\iint_S \mathrm{d}x_+\,\mathrm{d}x_- \iint \mathrm{d}\omega_+\,\mathrm{d}\omega_- K|x_+|^{i\omega_+-1}|x_-|^{i\omega_--1}\beta^E_{\omega_-}\beta^W_{\omega_+}\\ &+\iint_E \mathrm{d}x_+\,\mathrm{d}x_- \iint \mathrm{d}\omega_+\,\mathrm{d}\omega_- K|x_+|^{-i\omega_+-1}|x_-|^{i\omega_--1}\beta^E_{\omega_-}\beta^E_{\omega_+}\\ &+\iint_W \mathrm{d}x_+\,\mathrm{d}x_- \iint \mathrm{d}\omega_+\,\mathrm{d}\omega_- K|x_+|^{i\omega_+-1}|x_-|^{-i\omega_--1}\beta^W_{\omega_-}\beta^W_{\omega_+}\ .\end{aligned} \tag{5.38}$$

We may write this more succinctly in terms of the Fourier transform (5.23), paying careful attention to the signs of the momenta:

$$\begin{aligned}\Phi(0,a,z) = \iint \mathrm{d}\omega_+\,\mathrm{d}\omega_- \Big(&-\tilde{K}_N(\omega_+,\omega_-)\,\beta^E_{\omega_+}\beta^W_{\omega_-} - \tilde{K}_S(-\omega_+,-\omega_-)\,\beta^W_{\omega_+}\beta^E_{\omega_-}\\ &+\tilde{K}_E(\omega_+,-\omega_-)\,\beta^E_{\omega_+}\beta^E_{\omega_-} + \tilde{K}_W(-\omega_+,\omega_-)\,\beta^W_{\omega_+}\beta^W_{\omega_-}\Big)\ .\end{aligned} \tag{5.39}$$

5.4.1 Mapping the precursor into the eastern Rindler wedge

From the expression (5.39), one sees that upon time-evolving the boundary operator $\mathcal{O} = \partial_+\phi\partial_-\phi$ to the $t = 0$ Cauchy slice, one or both parts of the resulting bilocal may have support in the western wedge (indicated by β^W). We now demonstrate that the entanglement present in the Minkowski vacuum can be used to map these parts into the east. The set-up is illustrated schematically in fig. 5.3.

The key observation is that acting on the Minkowski vacuum with a Rindler operator, we have

$$\beta^W_{\omega_\pm}|0\rangle = e^{-\pi\omega_\pm}\beta^E_{-\omega_\pm}|0\rangle \tag{5.40}$$

which one can see by writing $|0\rangle \propto \bigotimes_\omega \sum_n e^{-\pi\omega n}|n\rangle_W \otimes |n\rangle_E$. We shall use this fact to write (5.39) entirely in terms of operators in the eastern wedge, β^E. For this mapping between western and eastern operators to succeed, we require only that both their left- and right-action on the vacuum state agree,

$$\Phi\,|0\rangle = \mathcal{O}_E\,|0\rangle\ , \quad \text{and} \quad \langle 0|\,\Phi = \langle 0|\,\mathcal{O}_E\ , \tag{5.41}$$

which is enough to ensure that 2-pt correlators are preserved. Our strategy is to satisfy the left equation by construction, and then check whether the right equation is also satisfied.

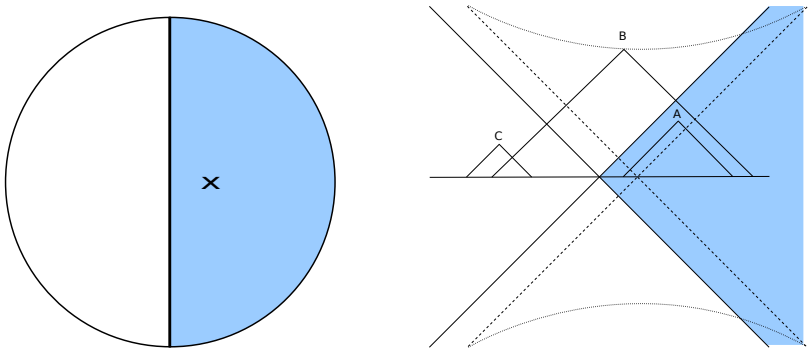


Figure 5.3: Left: $t = 0$ Cauchy slice, with a bulk point x displaced slightly into the eastern Rindler wedge (shaded). Right: time-evolution of local boundary operators to bilocals at $t = 0$. The dashed axes show the lightcone of the bulk point; the future and past singularities in the smearing function are indicated by the dotted lines. Point A falls entirely within the eastern wedge, while one leg of B, and both legs of C, must be mapped into the east using using the entanglement of the Rindler vacuum. Note that with the bulk point as shown, at most one singular leg must be mapped, but this potential divergence is exactly cancelled by a decaying exponential arising from (5.40), so the resulting expression remains well-defined.

Performing this mapping allows us to write (5.39) as

$$\begin{aligned}
&\Phi(0,a,z)|0\rangle \\
&= \iint \mathrm{d}\omega_+ \,\mathrm{d}\omega_- \Big[-\tilde{K}_N\left(\omega_+,\omega_-\right)\beta^E_{\omega_+}\beta^E_{-\omega_-}e^{-\pi\omega_-} - \tilde{K}_S\left(-\omega_+,-\omega_-\right)\beta^E_{-\omega_+}\beta^E_{\omega_-}e^{-\pi\omega_+} \\
&\qquad\qquad + \tilde{K}_E\left(\omega_+,-\omega_-\right)\beta^E_{\omega_+}\beta^E_{\omega_-} + \tilde{K}_W\left(-\omega_+,\omega_-\right)\beta^E_{-\omega_+}\beta^E_{-\omega_-}e^{-\pi(\omega_++\omega_-)}\Big]|0\rangle \\
&= \iint \mathrm{d}\omega_+ \,\mathrm{d}\omega_- \Big[-\cosh(\pi\omega_+)e^{-\pi\omega_-} - \cosh(\pi\omega_-)e^{\pi\omega_+} \\
&\qquad\qquad + \cosh(\pi(\omega_+-\omega_-)) + e^{\pi(\omega_+-\omega_-)}\Big]\tilde{K}_W\left(\omega_+,\omega_-\right)\beta^E_{\omega_+}\beta^E_{-\omega_-}|0\rangle \\
&= -2\iint \mathrm{d}\omega_+ \,\mathrm{d}\omega_- \sinh(\pi\omega_+)\sinh(\pi\omega_-)\tilde{K}_W\left(\omega_+,\omega_-\right)\beta^E_{\omega_+}\beta^E_{-\omega_-}|0\rangle \ ,
\end{aligned} \tag{5.42}$$

where have used the relations (5.24). Substituting in the explicit form of $\tilde{K}_W$, (5.66), we find

$$\begin{aligned}
\Phi|0\rangle = &-2\pi^2\left(\frac{z}{a}\right)^2\int \mathrm{d}\omega_+\,\mathrm{d}\omega_- a^{-i(\omega_++\omega_-)} \\
&\times {}_2F_1\left(1+i\omega_+,1+i\omega_-,2,-\frac{z^2}{a^2}\right)\beta^E_{\omega_+}\beta^E_{-\omega_-}|0\rangle \ ,
\end{aligned} \tag{5.43}$$

which is precisely (5.28), with $\tilde{\mathcal{O}}^E_{\omega_+,\omega_-} = \beta^E_{\omega_+}\beta^E_{-\omega_-}$. One can check that this operator Φ satisfies the condition (5.41). This demonstrates that the entanglement

structure of Minkowski space can be used to localize the precursor entirely within a single Rindler wedge, thus providing an alternative realization of the approach based on gauge freedom discussed above.

5.4.2 Mapping the precursor into the "wrong" Rindler wedge

To further explore this link between precursors and entanglement, let us now ask what happens if we instead attempt to map the bilocal operator into the western wedge. Since our bulk point is located in the east, we would naïvely expect this to fail, as this would correspond to reconstructing the bulk operator located outside the causal/entanglement wedge (cf. the end of section 5.3). Hence we refer to this as mapping the precursor into the *wrong* wedge.

The set-up is illustrated in fig. 5.4. Following the same procedure as in the previous subsection, one obtains

$$\begin{aligned}
&\Phi(0,a,z)|0\rangle \\
&\quad = -2\int \mathrm{d}\omega_+\,\mathrm{d}\omega_- e^{\pi(\omega_- - \omega_+)}\sinh(\pi\omega_+)\sinh(\pi\omega_-)\tilde{K}_W\left(\omega_+,\omega_-\right)\beta^W_{-\omega_+}\beta^W_{\omega_-}|0\rangle \\
&\quad \propto \int \mathrm{d}\omega_+\,\mathrm{d}\omega_- a^{-i(\omega_- + \omega_+)}e^{\pi(\omega_- - \omega_+)}{}_2F_1\left(1+i\omega_+, 1+i\omega_-, 2, -\frac{z^2}{x^2}\right)\beta^W_{-\omega_+}\beta^W_{\omega_-}|0\rangle\ .
\end{aligned} \tag{5.44}$$

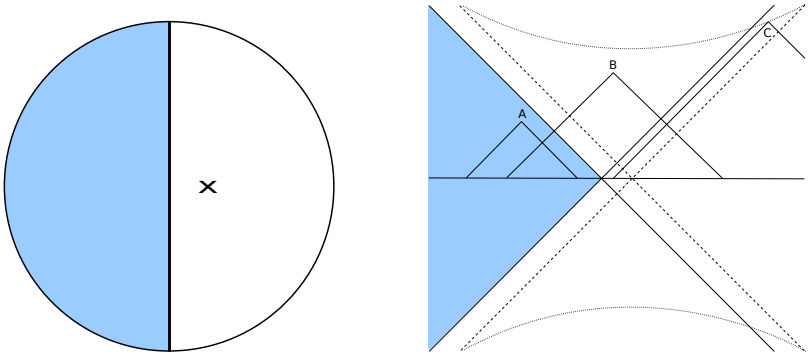


Figure 5.4: Left: $t = 0$ Cauchy slice, with a bulk point x displaced slightly into east as before, but reconstruction attempted in the western (wrong) wedge Rindler wedge (shaded). Right: time-evolution of local boundary operators to bilocals at $t = 0$. Note that while A and B can be mapped without difficulty, as discussed in the previous section, there are now points like C with two divergent legs, both of which must be mapped into the western wedge. This is one more exponential in momentum than we are capable of taming, and thus localization of the associated bulk point fails.

But upon conjugating (5.44), and taking $\omega_\pm \to -\omega_\pm$ under the integral, we

find

$$\Phi^\dagger |0\rangle \propto \int \mathrm{d}\omega_+ \, \mathrm{d}\omega_- a^{-i(\omega_- + \omega_+)} e^{-\pi(\omega_- - \omega_+)} \times {}_2F_1(1 + i\omega_+, 1 + i\omega_-, 2, -\frac{z^2}{x^2}) \beta^W_{-\omega_+} \beta^W_{\omega_-} |0\rangle \ , \tag{5.45}$$

in clear violation of the condition (5.41). Thus our entanglement mapping condition fails when the bulk operator lies in the complement of the selected boundary region.

The importance of this condition was recently emphasized in [174], who phrased it as the requirement of hermiticity. In particular, they proved that in order to satisfy (5.41), the field Φ must lie within the bulk entanglement wedge of the boundary region that contains the operator $\mathcal{O}$. Our model may therefore be taken as an explicit demonstration of this principle. Specifically, if one attempts to localize the boundary representation of a bulk operator in the complement, the resulting operator will be non-hermitian. In order to construct a well-defined precursor, the localization must be attempted within the entanglement wedge that includes the bulk field in question.

One can see that the wrong-wedge operator (5.44) is manifestly ill-behaved when acting on the Minkowski vacuum: in the limit $\omega_+ \gg 1$ and $\omega_- \ll -1$, we have two Rindler creation operators acting on $|0\rangle$, with a coefficient which grows exponentially. This means we create a state which is highly UV-sensitive (note the singular legs that must be mapped in fig. 5.4). Indeed, one can show that the two-point function $\langle 0|\Phi\Phi^\dagger|0\rangle$ diverges using the wrong-wedge operator Φ. The fact that UV-divergences occur in the same circumstance as when hermiticity is lost is suggestive, but we have not found a clear conceptual link between the two.

5.5 Discussion

The boundary duals of operators deep in the bulk have highly nonlocal representations in the CFT, known as "precursors". Following the HKLL construction, these can be localized to the boundary region of an AdS-Rindler wedge that contains the bulk field [102]. This immediately raises the question of redundant boundary duals: as illustrated in [101], a bulk field that falls within multiple boundary wedges must have multiple, different boundary representations. We are therefore left with the problem of how inequivalent precursors can all give rise to the same bulk operator.

In this chapter, we presented an improved version of the model in [171], wherein it was argued that the non-uniqueness of precursors is a simple consequence of boundary gauge invariance. We have provided an explicit demonstration of this proposal, using the gauge freedom in the smearing function to localize precursors

within a single Rindler wedge. This supports the claim [171] that gauge invariance may be deeply connected to the emergence of the dual spacetime. In section 5.3, this was accomplished without any mention of boundary entanglement. Rather, it relied only on the freedom to add pure-gauge modes to the precursor.

In contrast, entanglement is essential for a quantum error correction scheme to succeed [101]. Indeed, it has been postulated that the entanglement between boundary regions plays a crucial role in the emergence of the bulk spacetime, and there are reasons to believe that the entanglement – as opposed to the causal – wedge is the more natural bulk dual for holographic reconstruction [175, 98, 174, 176, 141, 177, 178, 179]. In the interest of further exploring the link between entanglement and localization, we showed explicitly in section 5.4 that the entanglement between boundary Rindler wedges can likewise be used to localize information to within a single region, in agreement with the approach from gauge freedom above.

In considering the concordance of these seemingly disparate approaches, it is worth emphasizing that in writing the pure-gauge operators in a particular form (in terms of creation and annihilation operators, cf. (5.6)), MPR [171] make use of the large N approximation, which includes an assumption about the class of states. So although the freedom to add these operators was demonstrated by MPR to be linked to gauge invariance, their explicit form, and thus the resulting localized precursor, does rely on an assumption that we act within the low-energy subspace of the theory.[10] We hope that we have shed some light on the physics, if not the linguistics, by explicitly calculating the resulting operators.

Our model may also be useful for diagnosing proposals for the description of operators behind the black hole horizon, such as [181], since the bulk spacetime we considered does have a Rindler horizon. In addition, it may clarify subtleties in the CFT operators dual to bulk fields outside the black hole horizon, which have the same properties as our Rindler precursors.

It is interesting to ask whether our model continues to agree with expectations about the full AdS/CFT correspondence when we consider more complicated boundary regions, such as disconnected intervals. The analogous set-up for a disconnected boundary region is shown in figure 5.5. The shaded region is the entanglement wedge for the given, disconnected boundary region. When this region becomes sufficiently large, the bulk minimal surface transitions to the new global minimum, whereupon the entanglement wedge suddenly includes the bulk point [148, 2]. The question we wish to ask is whether our model generalizes to agree with the corresponding reconstruction prescription. Specifically, can the precursor corresponding to a bulk point within the shaded bulk region be localized within the (disconnected) boundary of this region?

[10] In fact, this issue arises already at the level of the smearing function. As discussed in [110, 129, 180, 130], there are subtleties in attempting to construct an HKLL-type precursor in non-trivial geometries, e.g., in the presence of horizons.

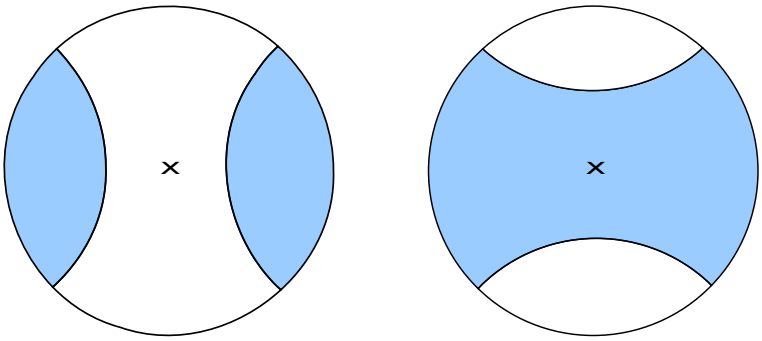


Figure 5.5: Entanglement wedge for a disconnected region, shaded. If the region is sufficiently small (left), the bulk point labelled x will not be included, and hence the shaded boundary region contains no information about it. However, as the boundary region is increased, the bulk geodesics that define the entanglement wedge eventually transition to a new global minimum (right), whereupon the shaded boundary region abruptly gains information about the given bulk point. Intuitively, one needs "enough" of the boundary to reconstruct the bulk.

Generalizing our above results to multiple, disconnected boundary regions requires either an explicit formula for the pure-gauge smearing function δK, or a general prescription for when a particular bilocal can be mapped into a given wedge. We do not present a general solution here, but instead comment on what one might expect given the above results, in the interest of comparing them with reconstruction proposals involving the entanglement wedge [174] and quantum error correction [101].

Figure 5.6 demonstrates the potential problem. Naïvely generalizing our results above for the case of a single boundary region, we suspect that bilocals with both points outside the boundary region, which are in addition integrated against non-smooth functions, cannot be mapped to healthy operators within the given CFT region. The smearing function K is smooth except at the boundary points that are lightlike connected to the bulk point, indicated by the dotted hyperbolas in the figure. Our suggested criterion, then, is that the precursor cannot be localized within the boundary wedge if some bilocals that are evolved back from the lightcone singularity have both points outside our region of the CFT.

Referring to the figure, one can see that even a bulk field within the entanglement wedge (the right image in 5.5) leads to such divergent bilocals that we cannot map into the correct boundary region. These are indicated by points A and B in fig. 5.6. Therefore, if our guess is correct for when the precursor can be localized, our simple model fails to reproduce the expected result, namely that bulk operators in the entanglement wedge can be mapped to precursors in the corresponding boundary region.

This should perhaps not be too surprising, since expectations about the entanglement wedge are based on the Ryu-Takayanagi formula for the entanglement

entropy. However, it is known that a simple free field model on the boundary will not reproduce the correct RT formula for the entanglement entropy of multiple intervals after a quench [182]. So it may simply be that our weakly-coupled model does not preserve the requisite entanglement between subregions upon evolving to bilocals along a single Cauchy slice.

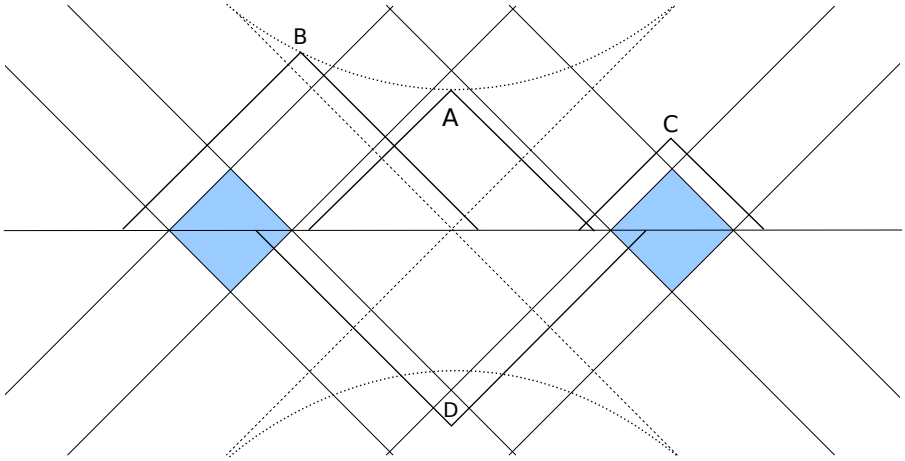


Figure 5.6: Disconnected region shown in fig. 5.5, with the bulk point in the center. Points A and B each have two singular legs, and cannot be mapped into the correct (shaded) boundary region with the naïve extension of our model. Which of C or D requires mapping is highly model-specific. In our model, with information localized along the edge of the light cone, only C requires mapping. If one instead devises a model in which information is smeared out along spacelike slices of the light cone, then most of C would instead fall into the correct wedge, but D would require a (presumably more complicated) mapping procedure.

One is therefore led to ask whether our model can be improved to capture the entire bulk entanglement wedge. Consider point C in fig. 5.6. In our model, this is time-evolved to bilocals lying entirely outside the entanglement wedge. However, one could imagine a different model in which the information about the local operator at C becomes smeared out along the intersection of the backwards lightcone of C and the $t = 0$ Cauchy slice, such that this point is still captured – that is, requires no potentially problematic mapping – provided some minimum amount of information falls within the wedge, perhaps evoking some quantum secret sharing scheme [136].

Unfortunately, such a modification creates other problems. Consider instead point D, which is time-evolved to bilocals lying entirely within the disconnected wedge. If we instead adopt this modification to our model, then most of the information about D would fall outside this region. Thus, in terms of mapping difficulty, we've only succeeded in trading C for D, and the underlying problem remains.

Performing the localization via a pure-gauge smearing function δK likewise appears impossible, although it would be interesting to try to extend our techniques to this case. Recall from section 5.3 that the ability to fix $\hat{K} = K + \delta K = 0$ relied on nontrivial relationships between the smearing function in different wedges in order to reduce the number of conditions (one per wedge) from three down to two, the number of undetermined coefficients. While we have not proven it, a quick glance at the many wedges of fig. 5.6 – which contains four Rindler-type axes – suggests that fixing $\hat{K}$ to zero in all but the two shaded regions would require a miraculous conspiracy of conditions.

Thus, while our model appears to generalize naturally to disconnected *causal* wedges, there is no obvious generalization that would correctly reproduce the *entanglement* wedge prescription. However, one fully expects that in the latter case a localized operator satisfying the condition (5.41) exists. Understanding precisely how the entanglement structure, or the gauge freedom, conspires to produce localized precursors for more general boundary regions would be illuminating.

5.A Relating Fourier transforms of the smearing function

In this appendix, we will prove the relations (5.24):

$$\begin{aligned}
\tilde{K}_N(\omega_+, \omega_-) &= \cosh(\pi\omega_+)\tilde{K}_W(\omega_+, \omega_-) \\
\tilde{K}_S(\omega_+, \omega_-) &= \cosh(\pi\omega_-)\tilde{K}_W(\omega_+, \omega_-) \\
\tilde{K}_E(\omega_+, \omega_-) &= \cosh\left(\pi(\omega_+ - \omega_-)\right)\tilde{K}_W(\omega_+, \omega_-)
\end{aligned} \tag{5.46}$$

where the Fourier transform of the smearing function, $\tilde{K}_W$, is given by (5.23), with K written in lightcone coordinates as in (5.37):

$$\begin{aligned}
&\tilde{K}_W(\omega_+, \omega_-) \\
&\quad = \int_{-\infty}^{0} \mathrm{d}x_+ \int_{0}^{\infty} \mathrm{d}x_- \log\left(\frac{|z^2 - (x_+ - a)(x_- + a)|}{2z}\right) |x_+|^{-i\omega_+ - 1} |x_-|^{-i\omega_- - 1} \\
&\quad = \int_{0}^{\infty} \mathrm{d}x_+ \int_{0}^{\infty} \mathrm{d}x_- \log\left(\frac{|z^2 + (x_+ + a)(x_- + a)|}{2z}\right) |x_+|^{-i\omega_+ - 1} |x_-|^{-i\omega_- - 1} \\
&\quad = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(\frac{|z^2 + (e^u + a)(e^v + a)|}{2z}\right) e^{-i\omega_+ u} e^{-i\omega_- v} \,,
\end{aligned} \tag{5.47}$$

where in the last step we've made the change of variables $x_+ = e^u$, $x_- = e^v$. Note that the logarithm does not become singular in the western quadrant, as we shifted the bulk point into the east (cf. fig. 5.3). For convenience, we may rescale the zero

mode to remove the constant factor in the denominator of the argument of the logarithm. Hence, suppressing the ω_+ and ω_- subscripts, the explicit expression for δK in each of the four wedges may be written

$$\tilde{K}_W = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(|z^2 + (e^u + a)(e^v + a)|\right) e^{-i\omega_+ u} e^{-i\omega_- v} \tag{5.48}$$

$$\tilde{K}_N = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(|z^2 - (e^u - a)(e^v + a)|\right) e^{-i\omega_+ u} e^{-i\omega_- v} \tag{5.49}$$

$$\tilde{K}_E = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(|z^2 + (e^u - a)(e^v - a)|\right) e^{-i\omega_+ u} e^{-i\omega_- v} \tag{5.50}$$

$$\tilde{K}_S = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(|z^2 - (e^u + a)(e^v - a)|\right) e^{-i\omega_+ u} e^{-i\omega_- v}. \tag{5.51}$$

Let us begin by relating $\tilde{K}_W$ and $\tilde{K}_N$. Define the function $f(u)$ for $u \in \mathbb{C}$ as[11]

$$f(u) \equiv \int_{-\infty}^{\infty} \mathrm{d}v \log\left(z^2 + (e^u + a)(e^v + a)\right) e^{-i\omega_+ u} e^{-i\omega_- v} \ . \tag{5.52}$$

Note that integrating f over the real u-axis gives $\tilde{K}_W$ (since $a > 0$), while integrating $f(u \pm i\pi)$ is of the same basic form as $\tilde{K}_N$,

$$f(u \pm i\pi) = e^{\pm \pi \omega_+} \int_{-\infty}^{\infty} \mathrm{d}v \log\left(z^2 - (e^u - a)(e^v + a)\right) e^{-i\omega_+ u} e^{-i\omega_- v} \ , \tag{5.53}$$

up to a factor of $e^{\pm\pi\omega_+}$, and ambiguities due to the singularities in the logarithm. In particular, the argument of the log is negative when $z^2 < (e^u - a)\,(e^v + a)$, so there is a branch point at

$$u^* = \log\left(\frac{z^2}{e^v + a} + a\right) \tag{5.54}$$

and branch cuts running horizontally at $u \pm i\pi$ for $u > u^*$. Now, imagine a rectangular contour in the complex u-plane running from $-\infty$ to ∞ along the real axis, and then back the other way along $u \pm i\pi$ (that is, just inside the complex region prescribed by the branch cuts). Since we enclose no poles, the total contour

[11] For simplicity we included the v-integral in the definition of f. For the reader worried about its convergence, the following contour argument can still be made, relating the v-integrands, by defining $f(u) \equiv e^{-i\omega_- v} \log\left(z^2 + (e^u + a)(e^v + a)\right) e^{-i\omega_+ u}$ for fixed v.

integral vanishes, and we may write

$$0 = \int_{-\infty}^{\infty} \mathrm{d}u \, (f(u) - f(u \pm i\pi)) + \int_{\infty}^{\infty \pm i\pi} \mathrm{d}u f(u) + \int_{-\infty \pm i\pi}^{-\infty} \mathrm{d}u f(u) \tag{5.55}$$

where the last two terms are the vertical side contributions for the function evaluated at $u \to \pm\infty$ from the real axis to $\pm i\pi$. One takes the upper signs in (5.55) for the contour in the upper half-plane, which runs counter-clockwise, and the lower signs for the clockwise contour in the lower half. We then observe that the side contributions can be made to vanish by suitably deforming the contour off the real axis. Hence, dropping these terms and writing the above expression in terms of $\tilde{K}$, we have

$$0 = \tilde{K}_W - e^{\pm\pi\omega_+} \tilde{K}_N - e^{\pm\pi\omega_+} \int_{u^*}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \, (\pm i\pi) \, e^{-i\omega_+ u} e^{-i\omega_- v} \tag{5.56}$$

where we've taken the principle value of the complex logarithm in (5.53), $\mathrm{Log}(x) = \log|x| \pm i\pi$, where the upper/lower sign corresponds to approaching the negative real axis from the upper/lower half-plane, respectively (i.e., our choice of contour). By a linear combination of the two equations in (5.56), the third term on the r.h.s. cancels, and one obtains

$$\tilde{K}_N = \cosh(\pi\omega_+) \tilde{K}_W \tag{5.57}$$

which is the desired result. Similarly, one can show

$$\tilde{K}_S = \cosh(\pi\omega_-) \tilde{K}_W \ . \tag{5.58}$$

The derivation of the third relation, between $\tilde{K}_W$ and $\tilde{K}_E$, follows a similar contour argument, but requires a slight change of coordinates. In particular, we first write (5.50) as

$$\tilde{K}_E = \int_{-\infty}^{\infty} \mathrm{d}t \int_{-\infty}^{\infty} \mathrm{d}x \log\left(|z^2 + (e^{t+x} - a)(e^{t-x} - a)|\right) e^{-i\omega t} e^{-ikx} \tag{5.59}$$

where we defined $\omega \equiv \omega_+ + \omega_-$ and $k \equiv \omega_- - \omega_+$, and similarly for $\tilde{K}_W$. We then define a function $g(x)$ for $x \in \mathbb{C}$,

$$g(x) \equiv \int_{-\infty}^{\infty} \mathrm{d}t \log\left(z^2 + (e^{t+x} - a)(e^{t-x} - a)\right) e^{-i\omega t} e^{-ikx} \ , \tag{5.60}$$

which will be related to (5.59) upon integrating along the x-axis, and observe that

the integral of

$$g(x \pm i\pi) \equiv \int_{-\infty}^{\infty} \mathrm{d}t \log\left(z^2 + (e^{t+x} + a)(e^{t-x} + a)\right) e^{-i\omega t} e^{-ikx} e^{\pm \pi k} \tag{5.61}$$

yields $\tilde{K}_W$.

We can now apply essentially the same argument as before. The argument of the logarithm in (5.59) is negative when $z^2 + a^2 + e^{2t} < 2ae^t \cosh x$, implying branch points at

$$x^* = \pm \cosh^{-1}\left(\frac{z^2 + a^2 + e^{2t}}{2ae^t}\right) . \tag{5.62}$$

We choose the branch cuts running out horizontally to infinity. The integration contours are then restricted to the rectangular region between the x-axis and $x \pm i\pi$, given an expression analogous to (5.55). Analytically continuing the logarithm to complex values as above, and dropping the side contributions, we have

$$0 = \tilde{K}_E - e^{\pm \pi k} \tilde{K}_W + \left(\int_{-\infty}^{-x^*} \mathrm{d}x + \int_{x^*}^{\infty} \mathrm{d}x\right) \int_{-\infty}^{\infty} \mathrm{d}t (\pm i\pi) e^{-i\omega t} e^{-ikx} . \tag{5.63}$$

Taking a linear combination of these two equations, we obtain

$$\tilde{K}_E = \cosh(\pi(\omega_- - \omega_+)) \tilde{K}_W , \tag{5.64}$$

as desired.

5.B Evaluating the smearing function

In this appendix we evaluate the Fourier integral of the smearing function in the western Rindler wedge, $\tilde{K}_W$ (5.48),

$$\tilde{K}_W = \int_{-\infty}^{\infty} \mathrm{d}u \int_{-\infty}^{\infty} \mathrm{d}v \log\left(z^2 + (e^u + a)(e^v + a)\right) e^{-i\omega_+ u} e^{-i\omega_- v} , \tag{5.65}$$

where the argument of the log is always positive by virtue of our having shifted the bulk point into the east, as described in the main text. Integrating by parts

twice, this becomes

$$\begin{aligned}\tilde{K}_W = &-\frac{1}{\omega_+\omega_-}e^{-i\omega_+ u}e^{-i\omega_- v}\ln\left(z^2+(e^u+a)(e^v+a)\right)\Big|_{u,v=-\infty}^{\infty}\\ &+\frac{1}{\omega_+\omega_-}\int_{-\infty}^{\infty}\mathrm{d}u e^{-i\omega_+ u}e^{-i\omega_- v}\frac{e^u(e^v+a)}{z^2+(e^u+a)(e^v+a)}\Big|_{v=-\infty}^{\infty}\\ &+\frac{1}{\omega_+\omega_-}\int_{-\infty}^{\infty}\mathrm{d}v e^{-i\omega_+ u}e^{-i\omega_- v}\frac{(e^u+a)e^v}{z^2+(e^u+a)(e^v+a)}\Big|_{u=-\infty}^{\infty}\\ &-\frac{z^2}{\omega_+\omega_-}\int_{-\infty}^{\infty}\mathrm{d}u\,\mathrm{d}v e^{-i\omega_+ u}e^{-i\omega_- v}\frac{e^{u+v}}{\left(z^2+(e^u+a)(e^v+a)\right)^2}\ .\end{aligned}$$

The first three (boundary) terms can be made to vanish by a suitable contour deformation. The remaining double integral (the fourth term) can be evaluated to yield

$$\tilde{K}_W = -\pi^2\left(\frac{z}{a}\right)^2 a^{-i(\omega_+ +\omega_-)}\operatorname{csch}(\pi\omega_+)\operatorname{csch}(\pi\omega_-){}_2F_1\left(1+i\omega_+,1+i\omega_-,2,\frac{-z^2}{a^2}\right)\ . \tag{5.66}$$

5.C Computing the two-point function

As an extra check of our formalism, we include an explicit calculation of the two-point function, and show that it reduces to the correct AdS_{2+1} correlator in the near-horizon limit. This will serve as a diagnostic of whether our expression for the bulk field constructed from boundary data entirely in the eastern wedge, (5.43)

$$\begin{aligned}\Phi(0,a,z) = &-2\pi^2\left(\frac{z}{a}\right)^2\int \mathrm{d}\omega_+\,\mathrm{d}\omega_- a^{-i(\omega_+ +\omega_-)}\\ &\times {}_2F_1\left(1+i\omega_+,1+i\omega_-,2,\frac{-z^2}{a^2}\right)\beta^E_{\omega_+}\beta^E_{-\omega_-}\ ,\end{aligned}$$

is well-defined. Here $\beta^E_{\omega_\pm}$ are the Rindler creation ($\omega<0$) and annihilation ($\omega>0$) operators in the eastern wedge, as defined in the main text. Since we work entirely in the eastern wedge in what follows, we shall henceforth suppress the superscript E to minimize clutter.

Inside the two-point function, we will have left/right moving Rindler operators acting on the Minkowski vacuum. As the left- and right-movers commute, the

four-β correlator is

$$\langle 0|\beta_{\omega_+}\beta_{-\omega_-}\beta_{\omega'_+}\beta_{-\omega'_-}|0\rangle = \delta(\omega_+ + \omega'_+)\delta(\omega_- + \omega'_-)\left(\frac{\omega_+}{1-e^{-2\pi\omega_+}}\right)\left(\frac{\omega_-}{e^{2\pi\omega_-}-1}\right) . \tag{5.67}$$

The bulk two-point function we seek to examine is therefore written explicitly as

$$\begin{aligned}
&\langle \Phi(a_1,z_1)\,\Phi(a_2,z_2)\rangle \\
&\qquad = 4\pi^4\left(\frac{z_1 z_2}{a_1 a_2}\right)^2 \int_{-\infty}^{\infty} \mathrm{d}\omega_+\,\mathrm{d}\omega_-\,\mathrm{d}\omega'_+\,\mathrm{d}\omega'_-\,\delta(\omega_+ + \omega'_+)\delta(\omega_- + \omega'_-) \\
&\qquad \times a_1^{-i(\omega_+ + \omega_-)} a_2^{-i\left(\omega'_+ + \omega'_-\right)} \left(\frac{\omega_+}{1-e^{-2\pi\omega_+}}\right)\left(\frac{-\omega_-}{1-e^{2\pi\omega_-}}\right) \\
&\qquad \times {}_2F_1\left(1+i\omega_+, 1+i\omega_-, 2, \frac{-z_1^2}{a_1^2}\right) {}_2F_1\left(1+i\omega'_+, 1+i\omega'_-, 2, \frac{-z_2^2}{a_2^2}\right) .
\end{aligned}$$

By virtue of the delta functions, the integrals over primed frequencies are trivial:

$$\begin{aligned}
&\langle \Phi(a_1,z_1)\,\Phi(a_2,z_2)\rangle \\
&\qquad = \pi^4\left(\frac{z_1 z_2}{a_1 a_2}\right)^2 \int_{-\infty}^{\infty} \mathrm{d}\omega_+\,\mathrm{d}\omega_-\left(\frac{a_1}{a_2}\right)^{-i(\omega_+ + \omega_-)} \\
&\qquad \times \omega_+\omega_-\left(\coth(\pi\omega_+)+1\right)\left(\coth(\pi\omega_-)-1\right) \\
&\qquad \times {}_2F_1\left(1+i\omega_+, 1+i\omega_-, 2, \frac{-z_1^2}{a_1^2}\right) {}_2F_1\left(1-i\omega_+, 1-i\omega_-, 2, \frac{-z_2^2}{a_2^2}\right) .
\end{aligned} \tag{5.68}$$

Unfortunately, we have not succeeded in evaluating the remaining integrals exactly. However, we can investigate the behaviour in the near-horizon limit, equivalent to taking $z_1/a_1, z_2/a_2 \to \infty$. To avoid subtleties associated with the branch cut at infinity, we first performing a $z \to 1/z$ transform,

$$\begin{aligned}
F(a,b,c;z) = &\frac{\Gamma(c)\Gamma(b-a)}{\Gamma(b)\Gamma(c-a)}(-z)^{-a}F(a,a-c+1,a-b+1,1/z) \\
&+ \frac{\Gamma(c)\Gamma(a-b)}{\Gamma(a)\Gamma(c-b)}(-z)^{-b}F(b,b-c+1,-a+b+1,1/z) ,
\end{aligned} \tag{5.69}$$

which allows us to expand in the limit where the fourth argument of the hypergeometric function vanishes. Applying this to the product of hypergeometric functions

in (5.68), and then expanding around $z/a \to \infty$ yields, to first order,

$$
\begin{aligned}
{}_2F_1\left(1+i\omega_+, 1+i\omega_-, 2, \frac{-z_1^2}{a_1^2}\right) {}_2F_1\left(1-i\omega_+, 1-i\omega_-, 2, \frac{-z_2^2}{a_2^2}\right) &= \left(\frac{a_1 a_2}{z_1 z_2}\right)^2 \\
\times \left[\left(\frac{z_2}{a_2}\right)^{2i\omega_-} \frac{\Gamma\left(i(\omega_- - \omega_+)\right)}{\Gamma(1+i\omega_-)\Gamma(1-i\omega_+)} + \left(\frac{z_2}{a_2}\right)^{2i\omega_+} \frac{\Gamma\left(-i(\omega_- - \omega_+)\right)}{\Gamma(1-i\omega_-)\Gamma(1+i\omega_+)}\right] \\
\times \left[\left(\frac{z_1}{a_1}\right)^{-2i\omega_+} \frac{\Gamma\left(i(\omega_- - \omega_+)\right)}{\Gamma(1+i\omega_-)\Gamma(1-i\omega_+)} + \left(\frac{z_1}{a_1}\right)^{-2i\omega_-} \frac{\Gamma\left(-i(\omega_- - \omega_+)\right)}{\Gamma(1-i\omega_-)\Gamma(1+i\omega_+)}\right] .
\end{aligned}
\tag{5.70}
$$

Without loss of generality, we shall assume $z_2 > z_1$. Substituting this expansion into the two point function yields

$$
\langle 0|\Phi(0, a_1, z_1)\Phi(0, a_2, z_2)|0\rangle = \int \mathrm{d}\omega_+ \, \mathrm{d}\omega_- (U + L) \tag{5.71}
$$

where we've defined

$$
\begin{aligned}
U \equiv & \, \omega_-\omega_+ e^{\pi(\omega_+ - \omega_-)} \mathrm{csch}(\pi\omega_-)\mathrm{csch}(\pi\omega_+) \left(\frac{a_1}{a_2}\right)^{-i(\omega_- + \omega_+)} \\
& \times \left[\frac{\left(\frac{z_1^2}{a_1^2}\right)^{-i\omega_-} \Gamma(i\omega_+ - i\omega_-)}{\Gamma(1-i\omega_-)\Gamma(i\omega_+ + 1)} + \frac{\left(\frac{z_1^2}{a_1^2}\right)^{-i\omega_+} \Gamma(i\omega_- - i\omega_+)}{\Gamma(i\omega_- + 1)\Gamma(1-i\omega_+)}\right] \left[\frac{\left(\frac{z_2^2}{a_2^2}\right)^{i\omega_-} \Gamma(i\omega_- - i\omega_+)}{\Gamma(i\omega_- + 1)\Gamma(1-i\omega_+)}\right]
\end{aligned}
\tag{5.72}
$$

$$
\begin{aligned}
L \equiv & \, \omega_-\omega_+ e^{\pi(\omega_+ - \omega_-)} \mathrm{csch}(\pi\omega_-)\mathrm{csch}(\pi\omega_+) \left(\frac{a_1}{a_2}\right)^{-i(\omega_- + \omega_+)} \\
& \times \left[\frac{\left(\frac{z_1^2}{a_1^2}\right)^{-i\omega_-} \Gamma(i\omega_+ - i\omega_-)}{\Gamma(1-i\omega_-)\Gamma(i\omega_+ + 1)} + \frac{\left(\frac{z_1^2}{a_1^2}\right)^{-i\omega_+} \Gamma(i\omega_- - i\omega_+)}{\Gamma(i\omega_- + 1)\Gamma(1-i\omega_+)}\right] \left[\frac{\left(\frac{z_2^2}{a_2^2}\right)^{i\omega_+} \Gamma(i\omega_+ - i\omega_-)}{\Gamma(1-i\omega_-)\Gamma(i\omega_+ + 1)}\right] .
\end{aligned}
\tag{5.73}
$$

We will first perform the integral over ω_-, by viewing U and L as functions on the complex ω_--plane. One can then show the following:

- U and L have simple poles at $\omega_- = \omega_+ \pm ni$, for $n \in \mathbb{Z}$.
- $|U(i\omega_-)| \to 0$ and $|L(-i\omega_-)| \to 0$ in the limit $\omega_- \gg 1$.
- $U + L$ has no poles on the real ω_- axis.

With these properties in hand, the integral can be performed via the residue theorem, where we close the contour in the upper/lower half-plane for U/L, re-

spectively:

$$\begin{aligned}\langle\Phi\Phi\rangle &= \int \mathrm{d}\omega_+ \int \mathrm{d}\omega_-(U+L) \\ &= \int \mathrm{d}\omega_+ 2\pi i \left[-\mathrm{Res}(L,\omega_+) + \sum_{n=1}^{\infty}\left(\mathrm{Res}(U,\omega_+ + ni) - \mathrm{Res}(L,\omega_+ - ni)\right)\right] \\ &\approx \int \mathrm{d}\omega_+ \frac{-2}{\pi}\left(\frac{z_1}{z_2}\right)^{-2i\omega_+}\left[2\log\left(\frac{a_2}{z_2}\right) + 2\gamma + \psi(i\omega_+) + \psi(-i\omega_+)\right] \ ,\end{aligned} \tag{5.74}$$

where $\psi(z) = \Gamma'(z)/\Gamma(z)$ and $\gamma = -\psi(1)$. In evaluating the residues we used that $z_i/a_i \gg 1$ for $i = 1, 2$.

The integral over ω_+ is evaluated in a similar fashion. Viewing the integrand as a function in the complex ω_+-plane, one can see that it is well behaved on the real axis, and goes to zero at $+i\infty$. Closing the integration contour in the upper half-plane, the residue theorem yields

$$\begin{aligned}\langle\Phi\Phi\rangle &\approx \int \mathrm{d}\omega_+ \frac{-2}{\pi}\left(\frac{z_1}{z_2}\right)^{-2i\omega_+}\left(2\log\left(\frac{a_2}{z_2}\right) + 2\gamma + \psi(i\omega_+) + \psi(-i\omega_+)\right) \\ &= 2\pi i \sum_{n=1}^{\infty}\left(\frac{-2iz_1^{2n}}{\pi z_2^{2n}}\right) \propto \frac{z_1^2}{z_2^2 - z_1^2} \ ,\end{aligned} \tag{5.75}$$

which one can recognise as the correct two-point function for a massless scalar in AdS_{2+1}, i.e.,

$$\langle\Phi(0,a_1,z_1)\Phi(0,a_2,z_2)\rangle = \frac{1}{e^S\sinh(S)} \propto \frac{z_1^2}{z_2^2 - z_1^2} \tag{5.76}$$

where the geodesic distance S in the near-horizon limit is given by $S = \log(z_2/z_1)$.

6

Locality on sub-AdS scales

This chapter is based on [4].

Throughout the preceding chapters, we have emphasized nonlocality without much qualification of the term. However, as alluded in the introduction, locality on sub-AdS scales is particularly non-trivial. In this chapter, we investigate sub-AdS scale locality in a weakly coupled toy model of the AdS_3/CFT_2 correspondence. We find that this simple model has the correct density of states at low and high energies to be dual to Einstein gravity coupled to matter in AdS_3. The bulk correlation functions also have the correct behavior at leading order in the large N expansion, but deviations appear at order $1/N$. We interpret this as evidence for nonlocality of the theory, which is consistent with the presence of an infinite tower of massless higher-spin fields. Finally, we conjecture that any large N CFT_2 that is both modular invariant, and exhibits the correct low-energy density of states, is dual to a gravitational theory with sub-AdS scale locality.

6.1 Introduction

The AdS/CFT correspondence has enabled tremendous progress in our understanding of quantum gravity. However, many important questions remain unanswered. Which CFTs are dual to bulk theories of Einstein gravity, with or without matter fields? What is the simplest CFT that reproduces the basic features of Einstein gravity? How does sub-AdS scale locality emerge in AdS/CFT? The goal of the present chapter is to address these questions in the context of an explicit toy model.

We will focus on AdS_3/CFT_2, where it is simplest to obtain precise answers to these rather grand questions. Indeed, the AdS_3/CFT_2 duality is a particu-

larly constrained example of holography. Einstein gravity is topological in three dimensions, so there are no propagating gravitons. Additionally, two-dimensional CFTs are highly constrained by the presence of the additional Virasoro symmetry. Nevertheless, many important features of quantum gravity, for example aspects of black hole physics, are still captured in three-dimensional gravity. The more constrained 3-dimensional framework thus provides a tractable environment amenable to precise results, while yielding insights that generalize to higher dimensions.

In the strongest interpretation of the AdS/CFT correspondence, every two-dimensional CFT is dual to a theory of quantum gravity in AdS_3. In some sense, the CFT defines the theory of quantum gravity in the bulk. The CFT data, namely the full set of correlation functions, can be interpreted as scattering amplitudes in the dual theory. The central charge is given by the AdS radius in Planck units [183],

$$c = \frac{3\ell_{\text{AdS}}}{2G_N} \ . \tag{6.1}$$

However, a generic CFT will not correspond to a theory of weakly coupled gravity. Rather, there exists a set of conditions the field theory must satisfy in order for it to have a well-behaved geometric dual. Identifying this list of necessary and/or sufficient conditions has been the focus of much recent effort [184, 86, 185, 186, 187, 188, 189, 190, 191]. Here we briefly summarize the important constraints that will be relevant to the present work. We start with the weakest assumption, and incrementally carve out a smaller and smaller subset of the space of all two-dimensional CFTs.

1. **The large N criterion.** First, the relation (6.1) makes it clear that a weakly coupled gravitational theory requires large central charge. The large N limit in the CFT is thus equivalent to the semi-classical limit of the gravitational theory.

2. **The convergence criterion.** To obtain a sensible semi-classical limit, further constraints must be imposed. Chief among them is the requirement that the spectrum of the theory remains well-defined in the large N limit [188, 190, 189]. Specifically, we require that the density of states $\rho(\Delta)$ remains finite in the $N \to \infty$ limit at fixed energy Δ. This criterion can be seen as demanding that perturbation theory remains valid in the bulk, since the latter requires a finite number of bulk fields at every given energy.

   It is important to note that this is only a criterion on the perturbative spectrum of the gravitational theory, and therefore it says nothing about black holes; as $N \to \infty$, the energy Δ of the lightest black hole diverges.

3. **The sparseness criterion.** The phase structure of Einstein gravity in AdS_3 is such that there are two saddle points that dominate the finite temperature

partition function at low and high temperature, respectively: thermal AdS and the BTZ black hole. These saddles exchange dominance in the Hawking-Page phase transition at the self-dual temperature $\beta = 2\pi$. In [186], it was shown that in order for a CFT to reproduce this phase structure in the large N limit, the density of light operators must be bounded by

$$\rho(\Delta) \lesssim \exp(2\pi\Delta) , \qquad \Delta \leq \frac{c}{12} . \tag{6.2}$$

We refer to this as the sparseness criterion. However, this is a rather weak constraint, since it corresponds to a Hagedorn growth typical of string theories in which the string and AdS scales are equal. Thus it allows for theories that are drastically different from Einstein gravity, and in particular theories that are nonlocal on sub-AdS scales. The fact that such string theories can reproduce the phase structure of Einstein gravity is a peculiarity of AdS_3 (see [192] for a discussion of higher dimensions). It is therefore necessary to impose a stronger constraint on the CFT in order to ensure that we recover a bulk dual that is local on sub-AdS scales, which motivates the fourth and final criterion on our list:

4. **The locality criterion.** If the perturbative sector of the bulk theory is to behave as a local quantum field theory in AdS, then the CFT must satisfy the following condition on the density of states:

$$\rho(\Delta) \sim \exp\left(\gamma \Delta^{\frac{D-1}{D}}\right) , \qquad 1 \ll \Delta \ll N , \tag{6.3}$$

where γ is some order-one coefficient, and D is a (positive) integer with a natural interpretation: it is the total number of bulk dimensions whose sizes are comparable to the AdS radius. The free energy resulting from such a density of states will be compatible with bulk thermodynamics of a local quantum field theory in D dimensions, namely $F \propto V_D T^{D+1}$, with a proportionality constant that depends on γ. This criterion is therefore necessary to reproduce the correct bulk thermodynamics at low temperatures.

One may wonder, after carving out this subspace of field theories, whether these four criteria are in fact sufficient to ensure locality on sub-AdS scales. In this chapter, we will show that they are not, by investigating sub-AdS scale locality in a weakly coupled toy model. Despite its simplicity, our model reproduces a surprising number of the desired features of a theory dual to Einstein gravity coupled to matter in AdS_3. This includes the correct density of states at both low and high energies, as well as the correct bulk correlation functions at leading order in the large N expansion. Nonlocal effects are seen to emerge at order $1/N$. This is supported by the presence of an infinite tower of massless higher-spin fields in

the bulk, which renders the theory nonlocal in the sense that the effective bulk Lagrangian contains interactions with an unbounded number of derivatives.

However, a deeper pathology of our toy model is the lack of modular invariance; indeed, any attempt to restore modular invariance would add too many states to the low lying spectrum, violating the sparseness criterion and thereby displacing us beyond the subspace of holographic CFTs we so carefully circumscribed above. For this reason, we are led to the following conjecture:

Sub-AdS Locality Conjecture:

At large N, every CFT_2 that satisfies the locality criterion, and has modular invariance, is dual to a bulk gravitational theory with sub-AdS scale locality.

The evidence for this conjecture is essentially experimental, based largely on orbifold CFTs. The basic reasoning is as follows: starting from a large N theory with a global symmetry and many low lying states, one can try to project out states until the bound (6.3) is satisfied. In order to preserve modular invariance, twisted sectors must be added in proportion to the severity of the projection. In [188, 190], it was shown that for any orbifold by a permutation group $G \subseteq S_N$, the locality criterion cannot be satisfied. This leaves the possibility that a projection by a bigger group such as $O(N)$ could achieve this criterion. However, although this works for the untwisted sector, modular invariance forces the inclusion of so many twisted sectors that the spectrum grows even faster than Hagedorn [193, 194]. None of the extant orbifold constructions seem to work, even for non-discrete groups. Thus we would like to emphasize that the role of modular invariance is to constrain the set of theories one can consider in light of the criteria above. We are not claiming that restoring modular invariance in a theory that originally was not will automatically ensure bulk locality.

Of course, the absence of known counterexamples does not constitute a proof of our conjecture, though it would be interesting to try to construct one. Conversely, the CFT data that could most likely be used to disprove our conjecture are the OPE coefficients. Upon imposing (6.3) and demanding modular invariance, one could try to constrain the OPE coefficients using bootstrap techniques along the lines of [185, 187] (see also [195]). It would also be interesting to understand how our conjecture relates to other criteria, such as the gap in the operator dimensions given in [184]. We leave such attempts for future work, and instead focus here on the properties and consequences of this particular model.

6.1.1 Summary of results

In this chapter, we investigate the aforementioned criteria, and in particular the question of sub-AdS scale locality, by exploring the detailed properties of an explicit toy model for holography. The model, originally introduced in [196] and refined in [3], consists of N massless free bosons restricted to the singlet sector of the global $O(N)$ symmetry. This model can be thought of as the two-dimensional version of the GKPY duality [197, 198]. The theory has a scalar operator $\mathcal{O}$ dual to a massless scalar field in the bulk, defined as

$$\mathcal{O} = \partial\phi^I \bar{\partial}\phi^I. \tag{6.4}$$

In [196, 3], the connection between gauge invariance and quantum error correction [101] was investigated in the context of holographic reconstruction; this was the subject of chapter 5, where the model was used to explicitly show how one can localize bulk operators within a given spatial region. In this chapter, we will investigate more refined properties of the model, including its spectrum and $1/N$ effects in correlation functions. We will see that the spectrum of the theory is given by

$$\rho(\Delta) \sim \begin{cases} \exp\left(\gamma \Delta^{\frac{2}{3}}\right) , & 1 \ll \Delta \lesssim N \\ \exp\left(2\pi\sqrt{\frac{N}{3}\Delta}\right) , & \Delta \gg N \ . \end{cases} \tag{6.5}$$

The high energy spectrum is given by the Cardy formula. This is actually surprising, since the theory is not modular invariant. The projection to $O(N)$ singlets breaks modular invariance, and hence Cardy's formula does not *a priori* apply. However, we will argue – based on an explicit proof for $SO(3)$ – that this projection is only a subleading effect at energies much larger than N. Note that because modular invariance is broken, the regime of validity of the Cardy formula does not extend to $\Delta \sim N$ even though the growth of the low energy spectrum (6.5) satisfies the sparseness criterion. In the intermediate range, the spectrum will interpolate smoothly between the two regimes in (6.5).

The low-energy spectrum is compatible with a local quantum field theory in AdS_3. However, the spectrum contains an infinite tower of higher-spin fields which ultimately cause the breakdown of sub-AdS scale locality. We demonstrate this breakdown from properties of the Lorentzian four point function of the operator $\mathcal{O}$. In particular, there is no divergence at order $\mathcal{O}(1/N)$ when the boundary points form a bulk Landau diagram [199, 200, 201, 184, 202]. Furthermore, the bulk theory is a Vasiliev higher-spin theory [203], and the effective Lagrangian contains interactions with an unbounded number of derivatives. In fact, it turns out that this model is equivalent to a sector of the coset models described in [204, 205],

with a $\mathcal{W}_\infty^{(e)}$ symmetry at $\lambda = 1$.

Our model demonstrates that the locality criterion on the spectrum is actually not a sufficient condition for sub-AdS scale locality. However, the model was constructed by taking a modular invariant theory and projecting out many states. The result is manifestly not modular invariant, and restoring it with the addition of twisted sectors would completely destroy the sparseness of the low lying states. This was shown in a similar context in [193]. Our theory can therefore not satisfy both the locality criterion and modular invariance simultaneously. We believe that these arguments extend beyond our specific toy model, which leads us to the sub-AdS scale locality conjecture above.

The remainder of the chapter is organized as follows: in section 6.2, we discuss properties of the spectrum of our toy model at both low and high energies. In section 6.3, we comment on properties of correlation functions at leading and subleading order in the $1/N$ expansion. Explicit expressions for the first few single-trace primaries are collected in appendix 6.A.

6.2 A toy model for holography

6.2.1 The model

The model we consider, introduced in chapter 5, was defined in [3] as a refinement of an earlier version proposed in [196]. The CFT consists of N free massless scalars in two dimensions. The action is

$$S = \int d^2x \, \partial_\mu \phi^I \partial^\mu \phi^I \ , \tag{6.6}$$

where the scalars ϕ^I transform in the fundamental representation of a global $O(N)$ symmetry. The Hilbert space of such a theory is given by

$$\mathcal{H}_N = \mathcal{H}^{\otimes N} \ , \tag{6.7}$$

where $\mathcal{H}$ is the Hilbert space of a single free boson. We wish to consider the subspace of states that are invariant under the $O(N)$ symmetry, namely the singlet sector. Therefore the relevant Hilbert space is

$$\mathcal{H}_{\text{singlet}} = \mathcal{H}^{\otimes N}/O(N) \ . \tag{6.8}$$

It is important to specify the procedure by which we impose such a constraint. In general field theories, the way to do so with local dynamics is by gauging the symmetry. This will enforce Gauss' Law and project to the singlet sector. However, preserving conformal invariance in the process is more subtle. In three dimensions,

this has been accomplished by weakly gauging the global symmetry and bestowing Chern-Simons dynamics on the gauge field. If the topology is trivial, one obtains the singlet projection without the introduction of additional states. On non-trivial topologies however, the holonomies of the gauge field come into play and appear to give rise to many new degrees of freedom [206].

In two dimensions, there is a very natural way to enforce a singlet constraint while preserving conformal invariance: orbifolding. The orbifolding procedure (which is usually done for a discrete group) enforces the singlet constraint, but also adds new operators to the theory from the twisted sectors. Indeed, a CFT_2 orbifold should really be thought of as a discrete gauge theory in two dimensions, where the twisted sectors are the degrees of freedom arising from the holonomies of the gauge field. Note that the inclusion of the twisted sector states comes from demanding that the theory is modular invariant on the torus. Projecting to the singlet sector without adding twisted sectors manifestly breaks modular invariance.

Throughout this chapter, we will only consider the untwisted sector, which is tantamount to imposing the singlet constraint by hand. As a consequence, our theory will not be modular invariant. This has some important ramifications, some of which we address when we discuss the high energy spectrum below. That said, we wish to emphasize that the singlet sector nonetheless retains many desirable properties. For example, the sector is closed: only singlet operators appear in the OPE of any two singlet operators. This implies in particular that the four point function of any singlet operators obeys the crossing relations.

6.2.2 Spectrum of primaries

In this section, we describe the spectrum of singlet operators in our CFT. We will be particularly interested in the single-trace Virasoro primaries, since every such operator is dual to a new bulk field, while multi-trace primaries correspond to multi-particle states (single particle states with additional boundary gravitons can also be viewed as multi-particle states in some broader sense).

The spectrum of the theory is characterized by the appearance of one new single-trace Virasoro primary at every even level $h, \bar{h} \geq 4$, in each of the holomorphic and anti-holomorphic sectors. The general expression for these operators may be written [207]

$$W^s(z) = \frac{2^{s-3} s!}{(2s-3)!!} \sum_{l=1}^{s-1} \frac{(-1)^l}{s-1} \binom{s-1}{l} \binom{s-1}{s-l} \partial^l \phi^I \partial^{s-l} \phi^I + \mathcal{O}\left(\frac{1}{N}\right) . \tag{6.9}$$

Note that these operators are not exactly single trace, but their double trace components are suppressed by powers of $1/N$. We give explicit expressions to all orders in $1/N$ for the holomorphic primaries up to level 12 in appendix 6.A,

and find that the multi-trace components are indeed always suppressed by higher powers of N. These fields correspond to higher-spin currents, and have been shown to generate a non-linear $\mathcal{W}_\infty^{(e)}[\lambda = 1]$ algebra [207]. In the mixed sector, the theory contains one single-trace scalar operator,

$$\mathcal{O} = \partial\phi^I \bar{\partial}\phi^I \ , \tag{6.10}$$

with dimension $(h, \bar{h}) = (1, 1)$. This operator is also a $\mathcal{W}_\infty^{(e)}$ primary, and naturally induces an infinite tower of multi-trace $\mathcal{W}_\infty^{(e)}$ operators given schematically by

$$\mathcal{O}^k_{n_i,\bar{n}_i} =: \sum_{n_i,\bar{n}_i} a_{n_1\ldots n_k \bar{n}_1\ldots\bar{n}_k} \partial^{n_1}\bar{\partial}^{\bar{n}_1}\mathcal{O}\partial^{n_2}\bar{\partial}^{\bar{n}_2}\mathcal{O}....\partial^{n_k}\bar{\partial}^{\bar{n}_k}\mathcal{O} : +\mathcal{O}\left(\frac{1}{N}\right) , \tag{6.11}$$

for an appropriate choice of coefficients $a_{n_i,\bar{n}_i}$. A generic choice of these coefficient will not lead to a primary, since the global descendants of the lower dimensional operators must be subtracted out. Along with their global and $\mathcal{W}_\infty^{(e)}$ descendants, the operators (6.11) generate the entire spectrum of the theory in the limit $N \to \infty$. At finite N, there are new primary operators that appear at $\Delta = N$. These will play an important role when we discuss the high energy part of the spectrum.

It is worth mentioning that we do not include zero modes. The standard vertex operators $e^{ik^I\phi^I}$ are not invariant under the $O(N)$ symmetry and are thus projected out. However, this still allows for operators of the form $e^{\lambda\phi^I\phi^I}$. We will not consider such operators, and instead implicitly further project to states that are invariant under $ISO(N)$ symmetries $\phi^I \to R^{IJ}\phi^J + C^I$.

6.2.3 Density of states

Low energies: $1 \ll \Delta \ll N$

We first compute the asymptotic density of perturbative states, i.e., states whose energy is parametrically smaller than N. States whose energy scales with N are typically associated to non-perturbative objects such as a black holes, and will be mentioned below.

We will consider free bosons on the cylinder, where the excitations are given by oscillators a^I_{-j}. The index j denotes the energy of the oscillator, hence a single-oscillator state would have $h = j$. In order to compute the density of perturbative states $\rho(\Delta)$, we consider $n < N$ oscillators a^I, each in the fundamental representation of $O(N)$. The singlet constraint forces us to contract all indices to form an invariant state. If n is even, this can be done in $(n-1)!!$ different ways, while if n is odd, the singlet constraint implies $\rho(\Delta) = 0$. The density of states for an

n-oscillator state can therefore be estimated as

$$\rho_n(\Delta) \sim (n-1)!! \cdot \frac{1}{n!} \int_0^\Delta d\Delta_1 \dots \int_0^\Delta d\Delta_n \; \delta\left(\Delta - \sum_i \Delta_i\right) \tag{6.12}$$

$$= (n-1)!! \frac{\Delta^{n-1}}{n!(n-1)!} = \frac{\Delta^{n-1}}{n!(n-2)!!} \; , \tag{6.13}$$

where the factor of $1/n!$ in (6.12) approximates the number of ways of distributing the energy Δ over n oscillators. The total density of states is then

$$\rho(\Delta) \sim \sum_{n=1}^{\Delta} 2^n \rho_n(\Delta) \; , \quad \rho_n(\Delta) \approx e^{n \log \Delta - \frac{3}{2} n \log n} \quad \text{for} \quad n \gg 1 \; , \tag{6.14}$$

where the factor of 2^n accounts for the inclusion of both left- and right-movers. We may evaluate this sum by performing a saddle-point approximation on n. The dominant saddle is at $n_0 = (2\Delta)^{\frac{2}{3}} e^{-1}$, which yields

$$\rho(\Delta) \sim e^{\gamma \Delta^{\frac{2}{3}}} \; , \quad 1 \ll \Delta \ll N \; . \tag{6.15}$$

Note that, in addition to the saddle point, we made two other approximations in the course of obtaining this result: the factor of 2^n from the choice of a or $\bar{a}$, and the double factorial $(n-1)!!$ from pair contractions. These two factors are only exact when all the oscillators have different momenta, otherwise one should include an appropriate symmetrization factor. Our approximations thus yield an overcounting of the total number of states, but are subdominant in the regime under consideration. This is the reason for the undetermined coefficient γ in (6.15), which cannot be determined from this analysis. Nonetheless, this is sufficient to demonstrate that our theory satisfies the locality criterion.

Asymptotically high energies: $1 \ll N \ll \Delta$

The density of states at asymptotically high energies $\Delta \gg N$ exhibits Cardy growth. We shall only comment briefly on the proof of this statement here; the interested reader is referred to [4] for details. The idea is to show that the density of states in this regime has the same leading asymptotics as the product theory and the correction is only polynomial in the energy. Explicitly, the proof consists of showing that

$$e^{2\pi\sqrt{\frac{N}{3}\Delta}} \sim \rho_{\text{product}}(\Delta) \geq \rho_{\text{singlet}}(\Delta) \geq \frac{\rho_{\text{product}}(\Delta)}{\Delta^p} \, , \qquad \Delta \gg N \; . \tag{6.16}$$

This result may seem surprising, since it implies that the Cardy formula also

holds asymptotically in the singlet theory, even though the theory is not modular invariant. This is a consequence of the nature of the projection, which preserves certain properties of the full theory even though modular invariance is lost. To see this, consider an orbifold by a discrete group G. The singlet sector (equivalently, the untwisted sector) partition function is given by

$$Z_N(q,\bar{q}) = \frac{1}{|G|}\sum_{g\in G} \mathrm{Tr}_{\mathcal{H}^{\otimes N}}\left[\, g\; q^{L_0-c/24}\bar{q}^{\bar{L}_0-\bar{c}/24}\right] . \tag{6.17}$$

The term in this sum where the group element g is the identity will be

$$Z_N(q,\bar{q}) = \frac{1}{|G|}\mathrm{Tr}_{\mathcal{H}^{\otimes N}}\left[\, q^{L_0-c/24}\bar{q}^{\bar{L}_0-\bar{c}/24}\right] = \frac{1}{|G|}Z(q,\bar{q})^N , \tag{6.18}$$

where Z is the partition function of one free boson. For any discrete group, $|G|$ is a finite number, and will constitute only a small correction for sufficiently large temperatures. Performing an inverse Laplace transform to obtain the density of states, one finds that the growth is Cardy up to some subleading correction from $|G|$. This shows that for any discrete orbifold, even the non-modular invariant singlet theory still has a Cardy growth. Unfortunately, such an argument fails for projections by continuous groups. However, the analogue of the correction coming from $|G|$ can still be calculated in our $O(N)$ example. It is no longer constant in the energy, but it is still subleading compared to the Cardy growth.

6.3 Bulk locality

6.3.1 Locality and reconstruction

In this section, we review how bulk locality emerges in the model, and probe the breakdown thereof. We shall work in Lorentzian signature in the CFT. The field theory contains an operator $\mathcal{O} = \partial_+\phi^I\partial_-\phi^I$ with conformal dimension $\Delta = 2$, which is dual to a massless scalar Φ in AdS_3. In [3], this holographic toy model was used to investigate bulk locality and reconstruction of Φ in the large N limit. At leading order in $1/N$, the bulk field is free, and can be reconstructed on the boundary by integrating the CFT operator against a suitable smearing function [102]

$$\Phi(X) = \int \mathrm{d}x\,\mathrm{d}t\; K(X|x,t)\mathcal{O}(x,t) + O(\frac{1}{N}) . \tag{6.19}$$

This prescription correctly reproduces the bulk two-point function from the CFT.

We now demonstrate explicitly how bulk locality emerges at large N in this

model. Expanding the bulk field Φ into mode functions in Poincaré AdS_3, we have

$$\Phi(t,x,z) = \int \mathrm{d}\omega\, \mathrm{d}k\ \left(\alpha_{\omega k} g_{\omega k}(t,z,x) + \text{h.c.}\right). \tag{6.20}$$

A local bulk field should satisfy the equal-time commutation relations ,

$$[\Phi(x,z), \Phi(x',z')] = [\Pi(x,z), \Pi(x',z')] = 0\ , \tag{6.21}$$

$$[\Phi(x,z), \Pi(x',z')] \sim \delta(x-x')\delta(z-z')\ , \tag{6.22}$$

which in turn require

$$[\alpha_{\omega k}, \alpha_{\omega' k'}] = [\alpha^\dagger_{\omega k}, \alpha^\dagger_{\omega' k'}] = 0\ , \tag{6.23}$$

$$[\alpha_{\omega k}, \alpha^\dagger_{\omega' k'}] \sim \delta(\omega-\omega')\delta(k-k')\ . \tag{6.24}$$

Via the extrapolate dictionary, we can relate the bulk creation and annihilation operators above to the those in the CFT by demanding that $\lim_{z\to 0} z^{-\Delta}\Phi(t,x,z) \leftrightarrow \partial_+\phi^I \partial_-\phi^I$. This implies

$$\alpha_{\omega k} \sim \frac{a^I_{\omega+k}\tilde{a}^I_{\omega-k}}{\sqrt{N}} = \frac{a^I_{\omega_+}\tilde{a}^I_{\omega_-}}{\sqrt{N}}\ , \tag{6.25}$$

where the a's are the left- and right-moving Fourier modes of the boundary fields ϕ^I. Equation (6.25) is essentially the statement that a bulk particle corresponds to a pair of left- and right-moving excitations in the CFT. Note that $\omega_\pm < 0$ corresponds to a creation operator, and that $a^\dagger_{\omega_\pm} = a_{-\omega_\pm}$. Translating the bulk commutation relations (6.23) and (6.24) into the CFT using $[a^I_\omega, a^J_{\omega'}] = \omega\delta(\omega+\omega')\delta^{IJ}$ yields

$$\begin{aligned}\frac{1}{N}[a^I_{\omega_+}\tilde{a}^I_{\omega_-}, a^J_{\omega'_+}\tilde{a}^J_{\omega'_-}] &= \omega_+\omega_-\delta(\omega_+ + \omega'_+)\delta(\omega_- + \omega'_-) \\ &\quad + \frac{1}{N}\left(\omega_- a^I_{\omega'_+} a^I_{\omega_+}\delta(\omega_- + \omega'_-) + \omega_+\tilde{a}^I_{\omega'_-}\tilde{a}^I_{\omega_-}\delta(\omega_+ + \omega'_+)\right)\ ,\end{aligned} \tag{6.26}$$

which becomes local when N is large (i.e., when the last two terms can be dropped).

6.3.2 3- and 4-point correlation functions

At next-to-leading order in $1/N$, we expect the bulk dual of our CFT to be nonlocal, despite having the density of states of a local quantum field theory in $2+1$-dimensions. As detailed in section 6.2.2, the bulk contains massless higher-spin fields, which strongly suggests locality violation since the effective Lagrangian will be unbounded in the number of derivatives. To quantify the nonlocality, we

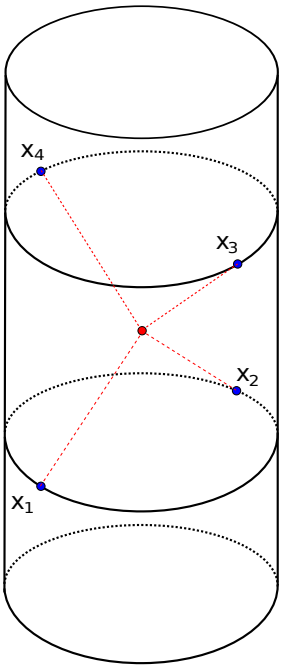


Figure 6.1: Four CFT insertions that are not lightlike separated in the CFT, but whose bulk lightcones intersect at a point.

calculate the 3- and 4-point functions of our primary field $\mathcal{O}$. As explained in [199, 200, 201, 184, 202], the 4-point functions provide a strong test of bulk locality. Any theory with a non-trivial S-matrix *in the flat space limit* must have certain lightcone singularities in the 4-point function. These singularities arise when the bulk interaction point is lightlike connected to all 4 boundary points, none of which are lightlike separated in the boundary theory; see fig. 6.1. Such singularities do not occur in a CFT at finite N, but they can appear in the large N limit.

The 3-point function of the operator $\mathcal{O}$ is zero,

$$\langle \mathcal{O}\mathcal{O}\mathcal{O} \rangle = 0 \ . \tag{6.27}$$

This is easily seen since each $\mathcal{O}$ contains one left-mover and one right-mover, so the 3-point function contains 3 left-movers. Since the boundary theory is free, the vacuum expectation value of an odd number of left-movers is zero.

The 4-point function contains a factorized piece, which dominates at large N, and a subleading connected piece. Defining the operator $\mathcal{O}$ with a normalization

that makes the 2-point function order one in N-scaling,

$$\mathcal{O} = \frac{1}{\sqrt{N}} \partial_+ \phi^I \partial_- \phi^I \ , \tag{6.28}$$

the 4-point function is

$$\begin{aligned} &\langle \mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle = \\ &\quad \frac{1}{N^2}\langle \partial_+\phi^I(x_1)\partial_-\phi^I(x_1)\partial_+\phi^J(x_2)\partial_-\phi^J(x_2)\partial_+\phi^K(x_3)\partial_-\phi^K(x_3)\partial_+\phi^L(x_4)\partial_-\phi^L(x_4)\rangle . \end{aligned} \tag{6.29}$$

We can then use the fact that

$$\langle \partial_+\phi^I(x_1)\partial_+\phi^J(x_2)\rangle = \frac{\delta^{IJ}}{(x_1^+ - x_2^+)^2} \ , \quad \langle \partial_+\phi^I(x_1)\partial_-\phi^J(x_2)\rangle = 0 \ , \tag{6.30}$$

to obtain

$$\begin{aligned} \langle \mathcal{O}(x_1)\mathcal{O}(x_2)\mathcal{O}(x_3)\mathcal{O}(x_4)\rangle = \ &\text{disconnected} \\ + \frac{1}{N}&\frac{1}{(x_1^+ - x_2^+)^2(x_1^- - x_3^-)^2(x_2^- - x_4^-)^2(x_3^+ - x_4^+)^2} \ + \ \text{permutations} \ , \end{aligned} \tag{6.31}$$

where, with our normalization conventions, the disconnected piece is of order N^0.

Examining this expression for the full 4-point function, it is clear that singularities arise only when some pair of points are lightlike separated on the boundary, such that they have the same value of x^+ or x^-. There are no additional singularities, which would appear if the bulk theory were truly described by Einstein gravity coupled to matter. This leaves us with two non-exclusive possibilities: the bulk theory is either nonlocal, or has a trivial S-matrix in the flat-space limit[1]. There is some evidence for the latter on general grounds (see, e.g., [215] and related work), so we cannot conclude directly from the singularity structure that the bulk theory is nonlocal. However, we have found above that the bulk theory contains an infinite tower of massless higher-spin fields, indicating that it is nonlocal in the sense that the Lagrangian contains an arbitrarily large number of derivatives. It would be interesting to better quantify the degree of nonlocality in the bulk (see for example [213]), and to determine whether the commutators can be corrected order-by-order in $1/N$. We leave these questions for future work.

[1] Determining the space of permissible field redefinitions that reveals the S-matrix to be trivial despite the presence of interaction terms is an open area of research. We will not attempt to address the issue here, but refer the reader to the higher-spin literature, e.g., [208, 209, 210, 211, 212, 213, 214].

6.A Holomorphic primaries

Here we give explicit expressions for the holomorphic Virasoro primaries at *finite* N, up to $h = 12$. We will work on the cylinder and discuss primary states. The comparison with the operators on the plane can be performed via the state-operator correspondence; e.g., the spin 4 operator is given in [207]. To see that our states are single-trace in the large N limit, some care is needed in the estimation of the magnitude of a given term. Terms with more oscillators naturally weigh more since they have several sums. Each oscillator carries an effective weight of $N^{1/4}$, which follows from considering any normalized state,

$$\mathcal{N} a_1 .... a_k \left|0\right\rangle \sim N^{-k/4} a_1 .... a_k \left|0\right\rangle \ . \tag{6.32}$$

The states below are given up to an overall normalization.

The computation at each level proceeds as follows. First, one identifies all descendant states obtained by acting with L_{-n} on lower-level primaries, including the vacuum state $|0\rangle$, and writes these in terms of the creation ($n < 0$) and annihilation ($n > 0$) operators a_n. The mode expansion of L_n for arbitrary n may be written

$$L_n = \frac{1}{2} \sum_{m=-\infty}^{\infty} : a^I_{m-n} a^I_m : \ , \tag{6.33}$$

where $: \ldots :$ denotes normal-ordering, and the modes satisfy

$$\left[a^I_m, a^J_n\right] = m \delta_{m,-n} \delta_{IJ} \ , \tag{6.34}$$

where I, J are $O(N)$ indices, for which Einstein's summation convention applies; similarly for the antiholomorphic sector. One can then write out all general linear combinations of modes that generate a state at a given level. For a suitable choice of coefficients, some of these will be precisely the descendant states obtained by acting with the Virasoro operators L_{-n}, while the remainder are identified as new primary states.

For example, at level $h = 2$, the only descendant states are obtained by acting on the vacuum with L_{-2}, $\bar{L}_{-2}$. However, there are 3 unique linear combinations of the relevant modes: a^2_{-1}, $\bar{a}^2_{-1}$, and $a^I_{-1} \bar{a}^I_{-1}$ (where $a^2 \equiv a^I a^I$). Writing out the mode expansion for L_{-2} explicitly, we find

$$L_{-2} \left|0\right\rangle = \frac{1}{2} a^2_{-1} \left|0\right\rangle \ ,$$

and similarly for $\bar{L}_{-2}$. Thus, two of these linear combinations are accounted for by the descendants, and the third, $a^I_{-1} \bar{a}^I_{-1}$, is identified as a new primary state at level 2, which we denote $|h, \bar{h}\rangle = |1, 1\rangle$.

Repeating this analysis at level 3 reveals no new primaries, while at level 4, we have:

	modes	descendants	primaries $\|h,\bar{h}\rangle$
$\Delta = J$	$a_{-3}a_{-1} + \frac{1}{2}a_{-2}^2$	$L_{-4}\|0\rangle$	
	$a_{-3}a_{-1} + \frac{1}{4}a_{-1}^4$	$L_{-2}^2\|0\rangle$	
	$4a_{-3}a_{-1} - 3a_{-2}^2 - \frac{6}{N+2}a_{-1}^4$		$\|4,0\rangle$
	$\tilde{a}_{-3}\tilde{a}_{-1} + \frac{1}{2}\tilde{a}_{-2}^2$	$\bar{L}_{-4}\|0\rangle$	
	$\tilde{a}_{-3}\tilde{a}_{-1} + \frac{1}{4}\tilde{a}_{-1}^4$	$\bar{L}_{-2}^2\|0\rangle$	
	$4\tilde{a}_{-3}\tilde{a}_{-1} - 3\tilde{a}_{-2}^2 - \frac{6}{N+2}\tilde{a}_{-1}^4$		$\|0,4\rangle$
mixed	$2a_{-3}\tilde{a}_{-1}$	$L_{-1}^2\|1,1\rangle$	
	$a_{-3}\tilde{a}_{-1} + \frac{1}{2}a_{-1}^3\tilde{a}_{-1}$	$L_{-2}\|1,1\rangle$	
	$2\tilde{a}_{-3}a_{-1}$	$\bar{L}_{-1}^2\|1,1\rangle$	
	$a_{-1}\tilde{a}_{-3} + \frac{1}{2}a_{-1}\tilde{a}_{-1}^3$	$\bar{L}_{-2}\|1,1\rangle$	
$J = 0$	$a_{-2}\tilde{a}_{-2}$	$L_{-1}\bar{L}_{-1}\|1,1\rangle$	
	$a_{-1}^I a_{-1}^I \tilde{a}_{-1}^J \tilde{a}_{-1}^J$	$4L_{-2}\bar{L}_{-2}\|1,1\rangle$	
	$a_{-1}^I a_{-1}^J \tilde{a}_{-1}^I \tilde{a}_{-1}^J - \frac{1}{N}a_{-1}^I a_{-1}^I \tilde{a}_{-1}^J \tilde{a}_{-1}^J$		$\|2,2\rangle$

and thus we find a single new primary in the holomorphic sector for $h = 4$, which we denote

$$\mathcal{W}_4 = a_{-1}^I a_{-3}^I - \tfrac{3}{4}a_{-2}^I a_{-2}^I - \tfrac{3}{2(N+2)}a_{-1}^I a_{-1}^I a_{-1}^J a_{-1}^J \ .$$

As the reader will no doubt appreciate, this process rapidly becomes intractably tedious as the number of states per level increases. So, we have written a *Mathematica* code to compute the holomorphic primaries as above. We find that the next new primary appears at level $h = 6$:

$$\begin{aligned}\mathcal{W}_6 = {} & a_{-1}^I a_{-5}^I - \tfrac{5}{2}a_{-2}^I a_{-4}^I + \frac{5}{3}a_{-3}^I a_{-3}^I + \tfrac{5(8N+7)}{4(N-1)(N+2)}a_{-1}^I a_{-1}^I a_{-2}^J a_{-2}^J \\ & + \tfrac{5(N-16)}{4(N-1)(N+2)}a_{-1}^I a_{-1}^J a_{-2}^I a_{-2}^J - \tfrac{15}{N+2}a_{-1}^I a_{-1}^J a_{-1}^J a_{-3}^I \\ & + \tfrac{15}{(N+2)(N+4)}a_{-1}^I a_{-1}^I a_{-1}^J a_{-1}^J a_{-1}^K a_{-1}^K \ .\end{aligned}$$

There are 2 new primaries at $h = 8$. An orthogonal basis can be chosen such that one of these becomes single-trace at large N, while the other remains multi-trace. The former may be written:

$$\begin{aligned}\mathcal{W}_8 = &- \tfrac{N+2}{28} a^I_{-1}a^I_{-7} + \tfrac{N+2}{8} a^I_{-2}a^I_{-6} - \tfrac{N+2}{4} a^I_{-3}a^I_{-5} + \tfrac{5(N+2)}{32} a^I_{-4}a^I_{-4} + a^I_{-1}a^J_{-1}a^J_{-1}a^I_{-5} \\ &- \tfrac{45}{32} a^I_{-2}a^I_{-2}a^I_{-2}a^J_{-2} - \tfrac{4N+3}{2(N-1)} a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-4} + \tfrac{5(3N+8)}{12(N-1)} a^I_{-1}a^I_{-1}a^J_{-3}a^J_{-3} \\ &- \tfrac{N-8}{2(N-1)} a^I_{-1}a^J_{-1}a^J_{-2}a^I_{-4} - \tfrac{28-N}{4(N-1)} a^I_{-1}a^J_{-2}a^I_{-2}a^J_{-3} - \tfrac{5(5N+6)}{12(N-1)} a^I_{-1}a^J_{-1}a^I_{-3}a^J_{-3} \\ &+ \tfrac{14N+13}{4(N-1)} a^I_{-1}a^J_{-2}a^J_{-2}a^I_{-3} - \tfrac{5}{4(N-1)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-2} \\ &+ \tfrac{5}{4(N-1)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^I_{-2}a^J_{-2} \ .\end{aligned}$$

Similarly, there are 3 new primaries at $h = 10$, only one of which will be single-trace at large N:

$$\begin{aligned}\mathcal{W}_{10} = &- \tfrac{N^3+5N^2+2N-8}{105(N+104)} a^I_{-1}a^I_{-9} + \tfrac{3(N^3+5N^2+2N-8)}{70(N+104)} a^I_{-2}a^I_{-8} - \tfrac{4(N^3+5N^2+2N-8)}{35(N+104)} a^I_{-3}a^I_{-7} \\ &+ \tfrac{N^3+5N^2+2N-8}{5(N+104)} a^I_{-4}a^I_{-6} - \tfrac{3(N^3+5N^2+2N-8)}{25(N+104)} a^I_{-5}a^I_{-5} + \tfrac{3(N^2+3N-4)}{7(N+104)} a^I_{-1}a^I_{-7}a^J_{-1}a^J_{-1} \\ &+ \tfrac{3(N^2-8N-48)}{7(N+104)} a^I_{-1}a^J_{-2}a^I_{-2}a^J_{-5} + \tfrac{3(6N^2+29N+20)}{7(N+104)} a^I_{-1}a^J_{-2}a^J_{-2}a^I_{-5} \\ &- \tfrac{3N^2+16N+16}{N+104} a^I_{-1}a^J_{-1}a^J_{-3}a^I_{-5} + \tfrac{3(11N^2+76N+128)}{40(N+104)} a^I_{-1}a^J_{-1}a^I_{-4}a^J_{-4} \\ &+ \tfrac{4(48N^2+253N+224)}{35(N+104)} a^I_{-1}a^J_{-2}a^I_{-3}a^J_{-4} + \tfrac{2N^2+13N+20}{N+104} a^I_{-1}a^I_{-1}a^J_{-3}a^J_{-5} \\ &- \tfrac{N^2+26N+88}{5(N+104)} a^I_{-1}a^I_{-2}a^J_{-3}a^J_{-4} - \tfrac{9N^2+82N+184}{7(N+104)} a^I_{-1}a^J_{-2}a^J_{-3}a^I_{-4} \\ &- \tfrac{28(N^2+3N-4)}{9(N+104)} a^I_{-1}a^J_{-3}a^J_{-3}a^I_{-3} - \tfrac{5N^2-12N-128}{14(N+104)} a^I_{-1}a^J_{-1}a^J_{-2}a^I_{-6} \\ &- \tfrac{16N^2+75N+44}{14(N+104)} a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-6} + \tfrac{7N^2-48N-304}{15(N+104)} a^I_{-2}a^J_{-2}a^I_{-3}a^J_{-3} \\ &+ \tfrac{38N^2+183N+124}{15(N+104)} a^I_{-2}a^I_{-2}a^J_{-3}a^J_{-3} - \tfrac{3(16N^2+9N+108)}{40(N+104)} a^I_{-1}a^I_{-1}a^J_{-4}a^J_{-4} \\ &- \tfrac{9(N^2+3N-4)}{2(N+104)} a^I_{-2}a^J_{-2}a^J_{-2}a^I_{-4} - \tfrac{1}{3} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-3}a^K_{-3} \\ &+ a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^I_{-2}a^K_{-3} + \tfrac{3(24N+71)}{4(N+104)} a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-2}a^K_{-2}a^K_{-2} \\ &+ \tfrac{3(N-96)}{4(N+104)} a^I_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^I_{-2}a^J_{-2} - \tfrac{N-96}{N+104} a^I_{-1}a^J_{-1}a^K_{-1}a^J_{-2}a^K_{-2}a^I_{-3} \\ &- \tfrac{50(N+3)}{N+104} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^I_{-3} + \tfrac{101N+4}{3(N+104)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^I_{-3}a^K_{-3} \\ &+ \tfrac{50(N+5)}{(N+6)(N+104)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-2}a^L_{-2} \\ &- \tfrac{200(N-1)}{3(N+6)(N+104)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^I_{-3} \\ &+ \tfrac{40(N-1)}{(N+6)(N+8)(N+104)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^M_{-1}a^M_{-1} \\ &- \tfrac{300}{(N+6)(N+104)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^I_{-2}a^J_{-2} \ .\end{aligned}$$

There are new (multi-trace) primaries at odd levels as well, beginning with one multi-trace primary at $h = 11$:

$$\begin{aligned}
\mathcal{W}_{11} = &- \tfrac{2(N+4)}{5} a^I_{-1}a^J_{-2}a^I_{-3}a^J_{-5} + \tfrac{2(N^2+3N-4)}{3(3N+4)} a^I_{-2}a^J_{-3}a^J_{-3}a^I_{-3} \\
&+ \tfrac{9(N^2+3N-4)}{10(3N+4)} a^I_{-2}a^J_{-2}a^J_{-2}a^I_{-5} + \tfrac{3(N^2+4N)}{2(3N+4)} a^I_{-1}a^J_{-2}a^J_{-4}a^I_{-4} \\
&+ \tfrac{2(N^2+4N)}{3(3N+4)} a^I_{-1}a^J_{-1}a^J_{-3}a^I_{-6} + \tfrac{2(N^2+7N+12)}{3(3N+4)} a^I_{-1}a^J_{-3}a^I_{-3}a^J_{-4} \\
&+ \tfrac{2(N^2+8N+16)}{5(3N+4)} a^I_{-1}a^J_{-2}a^J_{-3}a^I_{-5} - \tfrac{N^2+3N-4}{3N+4} a^I_{-2}a^J_{-2}a^J_{-3}a^I_{-4} \\
&- \tfrac{N^2+3N-4}{2(3N+4)} a^I_{-2}a^I_{-2}a^J_{-3}a^J_{-4} - \tfrac{N^2+5N+4}{2(3N+4)} a^I_{-1}a^J_{-2}a^J_{-2}a^I_{-6} \\
&- \tfrac{3N^2+17N+20}{3(3N+4)} a^I_{-1}a^J_{-3}a^J_{-3}a^I_{-4} - \tfrac{3(N^2+4N)}{5(3N+4)} a^I_{-1}a^I_{-1}a^J_{-4}a^I_{-5} \\
&+ a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^J_{-3}a^I_{-4} + \tfrac{3(3N+7)}{5(3N+4)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-5} \\
&+ \tfrac{2(N+6)}{3N+4} a^I_{-1}a^J_{-1}a^J_{-1}a^I_{-2}a^K_{-3}a^K_{-3} + \tfrac{3(N-1)}{2(3N+4)} a^I_{-1}a^I_{-1}a^J_{-2}a^K_{-2}a^K_{-2}a^J_{-3} \\
&+ \tfrac{6(N-6)}{5(3N+4)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^J_{-2}a^I_{-5} - \tfrac{N-1}{3N+4} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^I_{-6} \\
&- \tfrac{N+6}{3N+4} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-3}a^K_{-4} + \tfrac{N-26}{2(3N+4)} a^I_{-1}a^J_{-1}a^K_{-2}a^I_{-2}a^J_{-2}a^K_{-3} \\
&- \tfrac{3(2N+1)}{2(3N+4)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^I_{-2}a^K_{-4} + \tfrac{11N+14}{2(3N+4)} a^I_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^J_{-2}a^I_{-3} \\
&- \tfrac{8(N+1)}{3(3N+4)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-3}a^I_{-3} - \tfrac{2(5N+8)}{3(3N+4)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-2}a^I_{-3}a^J_{-3} \\
&- \tfrac{3(N-10)}{4(3N+4)} a^I_{-1}a^J_{-1}a^K_{-1}a^J_{-2}a^K_{-2}a^I_{-4} - \tfrac{9(N-1)}{4(3N+4)} a^I_{-1}a^J_{-2}a^J_{-2}a^K_{-2}a^K_{-2}a^I_{-2} \\
&- \tfrac{3(2N+1)}{4(3N+3)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^I_{-4} - \tfrac{2(N+4)}{3(3N+4)} a^I_{-3}a^I_{-6}a^J_{-1}a^J_{-1} \\
&+ \tfrac{3(N+4)}{5(3N+4)} a^I_{-1}a^I_{-1}a^J_{-4}a^J_{-5} + \tfrac{4(N+4)}{5(3N+4)} a^I_{-1}a^I_{-2}a^J_{-3}a^J_{-5} \\
&+ \tfrac{N+4}{3N+4} a^I_{-1}a^J_{-2}a^I_{-2}a^J_{-6} - \tfrac{3(N+4)}{2(3N+4)} a^I_{-1}a^I_{-2}a^J_{-4}a^J_{-4} \\
&+ \tfrac{1}{3N+4} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-2}a^L_{-3} + \tfrac{3}{2(3N+4)} a^I_{-1}a^J_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^I_{-2}a^J_{-2}a^K_{-2} \\
&- \tfrac{1}{3N+4} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^J_{-2}a^I_{-3} - \tfrac{3}{2(3N+4)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-2}a^L_{-2}a^I_{-2} \,.
\end{aligned}$$

Finally, there are 6 new primaries at $h = 12$, only one of which is single-trace at large N:

$$
\begin{aligned}
\mathcal{W}_{12} = {} & \frac{\left(-21980N^3+168221N^2+691811N+534898\right)}{4158(785N-3927)}a^I_{-11}a^I_{-1} + \frac{\left(-21980N^3+168221N^2+691811N+534898\right)}{2968812-593460N}a^I_{-10}a^I_{-2} \\
& + \frac{5\left(-21980N^3+168221N^2+691811N+534898\right)}{1134(785N-3927)}a^I_{-3}a^I_{-9} + \frac{5\left(21980N^3-168221N^2-691811N-534898\right)}{504(785N-3927)}a^I_{-4}a^I_{-8} \\
& + \frac{\left(-21980N^3+168221N^2+691811N+534898\right)}{63(785N-3927)}a^I_{-5}a^I_{-7} + \frac{\left(-21980N^3+168221N^2+691811N+534898\right)}{424116-84780N}a^I_{-6}a^I_{-6} \\
& + \frac{\left(-21980N^2+212181N+267449\right)}{247401-49455N}a^I_{-1}a^I_{-9}a^J_{-1}a^J_{-1} + \frac{5\left(306935N^2-3887313N-8239196\right)}{567(785N-3927)}a^I_{-3}a^I_{-3}a^J_{-3}a^J_{-3} \\
& + \frac{\left(-1305455N^2+15914271N+34603814\right)}{336(785N-3927)}a^I_{-2}a^J_{-2}a^J_{-2}a^I_{-6} + \frac{\left(-26690N^3+324523N^2+667124N-2139592\right)}{21(N-1)(785N-3927)}a^I_{-1}a^J_{-1}a^J_{-4}a^I_{-6} \\
& + \frac{5\left(305365N^3-4092910N^2-8201257N+4471138\right)}{336(N-1)(785N-3927)}a^I_{-1}a^I_{-3}a^J_{-4}a^J_{-4} + \frac{\left(510250N^3-4443925N^2-27596383N+52673488\right)}{1134(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^J_{-3}a^I_{-6} \\
& + \frac{\left(545575N^3-6683525N^2-15872458N+18721430\right)}{315(N-1)(785N-3927)}a^I_{-1}a^J_{-1}a^I_{-5}a^J_{-5} + \frac{\left(769300N^3-7202610N^2-11642125N+21364413\right)}{315(N-1)(785N-3927)}a^I_{-1}a^I_{-1}a^J_{-5}a^J_{-5} \\
& + \frac{5\left(1023640N^3-12610589N^2-38988835N-21781732\right)}{1008(N-1)(785N-3927)}a^I_{-2}a^J_{-2}a^I_{-4}a^J_{-4} + \frac{\left(1923250N^3-23419885N^2+46206677N-3566612\right)}{2646(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^I_{-2}a^J_{-7} \\
& + \frac{\left(9708095N^3-137201834N^2-57820355N+100740374\right)}{10584(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^J_{-2}a^I_{-7} + \frac{\left(-32970N^3+87089N^2+3039615N+8652616\right)}{42(N-1)(785N-3927)}a^I_{-1}a^I_{-2}a^J_{-3}a^J_{-6} \\
& + \frac{\left(-87920N^3+787494N^2+1769837N-1294776\right)}{21(N-1)(785N-3927)}a^I_{-1}a^I_{-1}a^J_{-4}a^J_{-6} + \frac{\left(-182905N^3+1912276N^2+12438979N+18721430\right)}{294(N-1)(785N-3927)}a^I_{-1}a^J_{-1}a^J_{-3}a^I_{-7} \\
& + \frac{\left(-184475N^3+1585790N^2+8866187N+7352023\right)}{189(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^I_{-4}a^J_{-5} + \frac{5\left(-268470N^3+3426506N^2+8284521N+1478428\right)}{189(N-1)(785N-3927)}a^I_{-2}a^J_{-3}a^J_{-3}a^I_{-4} \\
& + \frac{\left(-439600N^3+3957880N^2+8107333N+3476837\right)}{378(N-1)(785N-3927)}a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-8} + \frac{\left(-731620N^3+9702664N^2+20738839N-5747329\right)}{189(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^J_{-4}a^I_{-5} \\
& + \frac{\left(-2990065N^3+38646588N^2+90510931N-40654026\right)}{378(N-1)(785N-3927)}a^I_{-1}a^J_{-3}a^I_{-3}a^J_{-5} + \frac{\left(-15499825N^3+174967830N^2+711969205N+575713110\right)}{4032(N-1)(785N-3927)}a^I_{-2}a^I_{-2}a^J_{-4}a^J_{-4} \\
& - \frac{\left(-384650N^3+3237065N^2+10118963N+3473512\right)}{147(N-1)(785N-3927)}a^I_{-1}a^I_{-1}a^J_{-3}a^J_{-7} - \frac{\left(-1135895N^3+13711834N^2+36977579N+12467210\right)}{168(N-1)(785N-3927)}a^I_{-2}a^I_{-2}a^J_{-3}a^J_{-5} \\
& - \frac{5\left(38465N^3-767742N^2+3932638N+9717624\right)}{189(N-1)(785N-3927)}a^I_{-2}a^J_{-3}a^I_{-3}a^J_{-4} - \frac{\left(-339120N^3+4412329N^2+16084509N+22598996\right)}{189(N-1)(785N-3927)}a^I_{-1}a^J_{-3}a^J_{-3}a^I_{-5}
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{\left(153860N^3-2364467N^2+6615097N+10697960\right)}{378(N-1)(785N-3927)}a^I_{-1}a^J_{-1}a^J_{-2}a^I_{-8}-\frac{\left(-9682975N^3+11284057 0N^2+4247584 51N+14867371 4\right)}{2268(N-1)(785N-3927)}a^I_{-1}a^J_{-2}a^I_{-3}a^J_{-6}\\
&+\frac{1639}{224}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-4}a^I_{-4}+a^I_{-1}a^J_{-1}a^J_{-1}a^I_{-2}a^K_{-3}a^I_{-4}-\frac{339}{28}a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-3}a^K_{-5}+\frac{\left(1551109-103620N\right)}{7(785N-3927)}a^I_{-1}a^J_{-1}a^K_{-2}a^J_{-3}a^I_{-4}\\
&+\frac{\left(1551109-103620N\right)}{109956-21980N}a^I_{-1}a^J_{-1}a^K_{-1}a^I_{-4}a^J_{-4}+\frac{15\left(213222N^2-2080765N-6198962\right)}{448(785N^2-787N-15708)}a^I_{-2}a^I_{-2}a^J_{-2}a^J_{-2}a^K_{-2}a^K_{-2}\\
&+\frac{\left(-894115N^2+6945993N+15518692\right)}{196(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^I_{-1}a^I_{-7}+\frac{\left(785N^3+1944083N^2-31525652N+22197364\right)}{126(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-1}a^J_{-2}a^K_{-2}a^I_{-5}\\
&+\frac{\left(1112345N^3-5658604N^2-3680153N+67293772\right)}{168(N-1)(785N^2-787N-15708)}a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-6}-\frac{\left(39250N^3-507485N^2+1436813N+6414842\right)}{7(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^I_{-2}a^K_{-5}\\
&-\frac{5\left(36895N^4-512348N^3-4696816N^2+50411840N+144883494\right)}{189(N-2)(N-1)(785N^2-787N-15708)}a^I_{-1}a^I_{-1}a^J_{-2}a^K_{-2}a^J_{-3}a^K_{-3}\\
&-\frac{\left(-2281995N^4+19494269N^3+102794444N^2+402192820N+899108812\right)}{378(N-2)(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^I_{-2}a^J_{-2}a^K_{-3}a^K_{-3}\\
&+\frac{\left(-1413785N^3+13582987N^2+69249656N+58866122\right)}{126(N-1)(785N^2-787N-15708)}a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-2}a^J_{-2}a^I_{-5}\\
&+\frac{\left(-6984080N^3+43464553N^2+363334769N+560029358\right)}{336(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-2}a^J_{-2}a^K_{-2}a^K_{-2}a^I_{-3}\\
&+\frac{5\left(5785545N^4-14988647N^3+39274346N^2+310925180N+225350496\right)}{378(N-2)(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-2}a^J_{-2}a^K_{-3}a^I_{-3}\\
&+\frac{\left(25311935N^4-99984139N^3-122403220 4N^2-232303718 0N+100062748 8\right)}{1512(N-2)(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^I_{-3}a^J_{-3}\\
&-\frac{\left(-196250N^3+2232715N^2+2754743N+2592212\right)}{21(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^J_{-2}a^I_{-6}\\
&+\frac{5\left(-16320 15N^4+36063 01N^3+14385853 0N^2+41725775 6N+37460376 8\right)}{1512(N-2)(N-1)(785N^2-787N-15708)}a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-2}a^K_{-3}a^K_{-3}\\
&-\frac{\left(-408985N^3+4521077N^2+19620676N+13184332\right)}{28(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^J_{-2}a^I_{-4}-\frac{5\left(64559N^3+594207N^2-829586N-21979440\right)}{126(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^K_{-1}a^I_{-3}a^J_{-3}a^K_{-3}\\
&-\frac{\left(-1694030N^3+13612621N^2+84831098N+87835811\right)}{126(N-1)(785N^2-787N-15708)}a^I_{-1}a^I_{-1}a^J_{-2}a^J_{-2}a^K_{-2}a^J_{-4}-\frac{25\left(9577N^3-268385N^2+753488N+3935372\right)}{126(N-1)(785N^2-787N-15708)}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-3}a^I_{-3}a^I_{-3}
\end{aligned}
$$

$$
\begin{aligned}
&- \frac{\left(1305455N^3-29882431N^2+89007952N+419491324\right)}{168(N-1)(785N^2-787N-15708)} a^I_{-1}a^J_{-2}a^K_{-2}a^K_{-2}a^I_{-2}a^J_{-3} \\
&- \frac{25\left(70807N^3-187421N^2-9485554N-18454828\right)}{252(N-1)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^K_{-2}a^I_{-2}a^J_{-2}a^K_{-4} \\
&+ \frac{\left(109115N^2+12863447N-8778882\right)}{28(N-1)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^L_{-1}a^L_{-2}a^I_{-2}a^J_{-3} + \frac{\left(423115N^2-1936593N-7873432\right)}{21(N+6)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^I_{-5} \\
&+ \frac{5\left(434781N^2-1266368N-4272058\right)}{56(N+6)(N+8)(N+10)(785N^2-787N-15708)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^M_{-1}a^M_{-1}a^N_{-1}a^N_{-1} \\
&+ \frac{\left(-167445N^2-4229162N+11780027\right)}{28(N-1)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-2}a^L_{-2}a^I_{-2}a^I_{-3} + \frac{\left(3140N^3+16452N^2-612344N-1387347\right)}{9(N-1)(785N-3927)} a^I_{-1}a^I_{-2}a^J_{-4}a^J_{-5} \\
&+ \frac{5\left(3140N^3-38741N^2-7691N+311780\right)}{12(N-1)(785N-3927)} a^I_{-1}a^J_{-3}a^J_{-4}a^I_{-4} + \frac{\left(42390N^3-876973N^2+4572009N+11767756\right)}{42(N-1)(785N-3927)} a^I_{-2}a^J_{-2}a^J_{-3}a^I_{-5} \\
&+ \frac{5\left(289447N^3-901229N^2+5046122N-26584600\right)}{84(N-1)(N+6)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^I_{-3}a^J_{-3} \\
&+ \frac{5\left(707285N^3-3818293N^2-633298888N-206745644\right)}{1008(N-1)(N+1)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^K_{-1}a^L_{-1}a^I_{-2}a^J_{-2}a^K_{-2}a^L_{-2} \\
&+ \frac{5\left(894115N^3-2783791N^2-212988284N+89639284\right)}{252(N-1)(N+6)(785N^2-787N-15708)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-3}a^L_{-3} \\
&+ \frac{5\left(1729552N^3+8569597N^2-126974758N-429697471\right)}{2016(N-1)(N+1)(785N^2-787N-15708)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-2}a^K_{-2}a^L_{-2}a^L_{-2} \\
&+ \frac{\left(-1199480N^3+2023961N^2+27419135N-72544136\right)}{42(N-1)(N+6)(785N^2-787N-15708)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-2}a^L_{-4} \\
&+ \frac{5\left(423115N^3-6377731N^2-571552N+13909588\right)}{84(N-1)(N+6)(N+8)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^M_{-1}a^M_{-1}a^I_{-2}a^J_{-2} \\
&+ \frac{5\left(2220569N^3+10200583N^2-24765002N-17189830\right)}{336(N-1)(N+6)(N+8)(785N^2-787N-15708)} a^I_{-1}a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^M_{-2}a^M_{-2} \\
&- \frac{5\left(434781N^2-1266368N-4272058\right)}{28(N+6)(N+8)(785N^2-787N-15708)} a^I_{-1}a^J_{-1}a^J_{-1}a^K_{-1}a^K_{-1}a^L_{-1}a^L_{-1}a^M_{-1}a^M_{-1}a^I_{-3} \,.
\end{aligned}
$$

Outlook

The investigations in the preceding chapters leave many questions unanswered.

The Firewall remains unextinguished. While restricting our observations to within a single causal patch certainly causes problems for the s-wave formulation of the paradox as discussed in chapter 3, other variants – such as those involving the mining of high-ℓ modes – have not been completely disarmed. Furthermore, our constraints are most lenient for Schwarzschild black holes in 3+1 dimensions—arguably the case of maximal interest! We speculated that an external observer could still perform the necessary measurement using some quantum secret sharing scheme, but to our knowledge there is no compelling evidence either for or against this possibility.

More generally, no convincing model exists for precisely how information escapes in the Hawking radiation, as required for unitarity of the S-matrix. One perspective is that our conception of locality will require modification, such as in the so-called nonviolent nonlocality proposal of [216, 217], but a satisfactory resolution still eludes us. And locality is not the only tenet in need of reassessment. For example, nearly every discussion in this context involves an assumption about the entanglement structure at the horizon, namely, that the Hilbert space factorizes into a tensor product (e.g., Rindler space). But this is not true in gauge theory, let alone gravity. The deepening connections between entanglement and spacetime geometry uncovered in recent years may shed light on this issue, and it is one to which we hope to return.

Lessons from holography suggest that nonlocality will also play a central role in encoding bulk physics in the boundary field theory. Both the Ryu-Takayanagi conjecture and the HKLL prescription, collectively considered the current state-of-the-art in bulk reconstruction, are intrinsically nonlocal. But as demonstrated

in chapter 4, even these objects do not allow one to fully reconstruct the bulk due to the presence of holographic shadows. Some efforts [218] have been made to endow the subleading entwinement surfaces, which do penetrate these shadow regions, with a concrete boundary interpretation, but more work is needed before they can be deemed illuminated.

The incompleteness of the holographic dictionary in this context becomes even more pronounced if we wish to make inquiries of the black hole interior. Of course, insofar as an event horizon is a global property of the spacetime, it is technically possible for extremal surfaces to reach behind the horizon in collapsing geometries; but we regard it as misleading to claim to have probed within the black hole in this sense. Consider instead an example beloved of holographers: the eternal black hole in AdS, which is dual to the thermofield double state (TFD). The CFT must contain a complete description of physics in the bulk, but we have as yet no means of reading within the interior of the black hole.

In fact, the interior continues to evolve for long after the thermalization time. This, in conjunction with the above arguments, lead Susskind to claim that "entanglement is not enough" to completely capture the physics of black holes in the bulk [219]. He proposed holographic complexity as the field-theoretic entity that tracks this continued evolution. On the gravitational side, this is conjectured to correspond to either the volume of the Einstein-Rosen bridge, or the action of the Wheeler-DeWitt patch. However, at present there is no satisfactory definition of complexity in the CFT, and thus it is far too soon to say whether this notion will bear fruit; we are currently investigating whether progress can be made in this direction.

The precursors discussed in chapter 5 present another nonlocal puzzle. A glance at fig. 5.1 begs the question: if a tree falls in the bulk and no local observers are around to hear it, does it make a sound in the CFT? Since the boundary contains a complete description of physics in the bulk, the answer is certainly yes. The real question is how information from such a hypothetical bulk event is encoded in the spacelike separated region of the boundary. We discussed some tantalizing connections between bulk reconstruction and quantum error correction (QEC) in the context of localizing this boundary data in this thesis, but we do not yet have a general prescription for how QEC or related ideas can be used to explicitly recover bulk physics.

The above makes it apparent that there are still many gaps in the dictionary, on both sub-AdS scales and beyond. Elucidating the former served as partial motivation for chapter 6, where we examined the question of which CFTs have well-defined bulk duals. Particularly in the context of bulk reconstruction, AdS/CFT is often taken as providing a definition of quantum gravity in the bulk. Strictly speaking however, the duality is formulated in the limit $N \to \infty$, and thus it is relevant to ask how much of what we've learned survives as we move to finite

N. Understanding the role that locality plays in formulating a general theory of quantum gravity is a subject of ongoing work.

As alluded above, ideas from quantum information theory have played an increasingly large role in efforts to understand evaporating black holes, in bulk reconstruction, and in high-energy theory in general. The *It from Qubit* collaboration, funded by the Simons Foundation, is a testament to the growing connections between these fields. Indeed, recent investigations suggest that entanglement is not only fundamental to the encoding of information in AdS/CFT, but may provide the foundation for spacetime itself [60]. A demonstration of this idea is obtained by starting from the aforementioned TFD, and sending in shockwaves to create a wormhole whose interior is causally disconnected from both boundaries [220]. If the state in the CFT is to remain well-defined, then information about this interior region must be encoded in the entanglement structure of the perturbed TFD. In other words, understanding the wormhole's deep interior in the CFT is tantamount to isolating the degrees of freedom that sew the bulk spacetime together. Thus one expects that their description in the CFT will tell us how this spacetime – i.e., gravity – emerges from elements in the boundary. The underlying idea, that spacetime emerges from quantum entanglement, is a fascinating concept which we hope to investigate further.

Bibliography

[1] B. Freivogel, R. A. Jefferson, L. Kabir, and I.-S. Yang, "Geometry of the Infalling Causal Patch," *Phys. Rev.* **D91** no. 4, (2015) 044036, arXiv:1406.6043 [hep-th].

[2] B. Freivogel, R. A. Jefferson, L. Kabir, B. Mosk, and I.-S. Yang, "Casting Shadows on Holographic Reconstruction," *Phys. Rev.* **D91** no. 8, (2015) 086013, arXiv:1412.5175 [hep-th].

[3] B. Freivogel, R. A. Jefferson, and L. Kabir, "Precursors, Gauge Invariance, and Quantum Error Correction in AdS/CFT," *JHEP* **04** (2016) 119, arXiv:1602.04811 [hep-th].

[4] A. Belin, B. Freivogel, R. A. Jefferson, and L. Kabir, "Sub-AdS Scale Locality in AdS_3/CFT_2," arXiv:1611.08601 [hep-th].

[5] J. D. Bekenstein, "Black holes and entropy," *Phys. Rev. D* **7** (Apr, 1973) 2333–2346. https://link.aps.org/doi/10.1103/PhysRevD.7.2333.

[6] S. W. Hawking, "Gravitational radiation from colliding black holes," *Phys. Rev. Lett.* **26** (May, 1971) 1344–1346. https://link.aps.org/doi/10.1103/PhysRevLett.26.1344.

[7] T. Jacobson, "Introductory Lectures on Black Hole Thermodynamics,". http://www.physics.umd.edu/grt/taj/776b/lectures.pdf.

[8] S. W. Hawking, "Particle creation by black holes," *Comm. Math. Phys.* **43** no. 3, (1975) 199–220. http://projecteuclid.org/euclid.cmp/1103899181.

[9] W. G. Unruh, "Notes on black-hole evaporation," *Phys. Rev. D* **14** (Aug, 1976) 870–892. https://link.aps.org/doi/10.1103/PhysRevD.14.870.

[10] P. C. W. Davies, "Scalar particle production in Schwarzschild and Rindler metrics," *J. Phys.* **A8** (1975) 609–616.

[11] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*. Cambridge Monographs on Mathematical Physics. Cambridge Univ. Press, Cambridge, UK, 1984. http://www.cambridge.org/mw/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-fields-curved-space?format=PB.

[12] T. Jacobson and R. Parentani, "Horizon entropy," *Found. Phys.* **33** (2003) 323–348, arXiv:gr-qc/0302099 [gr-qc].

[13] J. Polchinski, "The Black Hole Information Problem," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 353–397. 2017. arXiv:1609.04036 [hep-th].

[14] G. W. Gibbons and S. W. Hawking, "Action integrals and partition functions in quantum gravity," *Phys. Rev. D* **15** (May, 1977) 2752–2756. https://link.aps.org/doi/10.1103/PhysRevD.15.2752.

[15] S. W. Hawking, "Breakdown of predictability in gravitational collapse," *Phys. Rev. D* **14** (Nov, 1976) 2460–2473. https://link.aps.org/doi/10.1103/PhysRevD.14.2460.

[16] A. Alonso-Serrano and M. Visser, "On burning a lump of coal," *Phys. Lett.* **B757** (2016) 383–386, arXiv:1511.01162 [gr-qc].

[17] D. Harlow, "Jerusalem Lectures on Black Holes and Quantum Information," *Rev. Mod. Phys.* **88** (2016) 15002, arXiv:1409.1231 [hep-th]. [Rev. Mod. Phys.88,15002(2016)].

[18] S. B. Giddings, "Hawking radiation, the Stefan-Boltzmann law, and unitarization," *Phys. Lett.* **B754** (2016) 39–42, arXiv:1511.08221 [hep-th].

[19] M. Visser, "Thermality of the Hawking flux," *JHEP* **07** (2015) 009, arXiv:1409.7754 [gr-qc].

[20] S. D. Mathur and C. J. Plumberg, "Correlations in Hawking radiation and the infall problem," *JHEP* **09** (2011) 093, arXiv:1101.4899 [hep-th].

[21] S. D. Mathur, "The Information paradox: A Pedagogical introduction," *Class. Quant. Grav.* **26** (2009) 224001, arXiv:0909.1038 [hep-th].

[22] J. D. Bekenstein, "Entropy bounds and black hole remnants," *Phys. Rev. D* **49** (Feb, 1994) 1912–1921. https://link.aps.org/doi/10.1103/PhysRevD.49.1912.

[23] Y. Aharonov, A. Casher, and S. Nussinov, "The unitarity puzzle and planck mass stable particles," 1987.

[24] S. B. Giddings, "Black holes and massive remnants," *Phys. Rev.* **D46** (1992) 1347–1352, arXiv:hep-th/9203059 [hep-th].

[25] J. Preskill, "Do black holes destroy information?," in *International Symposium on Black holes, Membranes, Wormholes and Superstrings Woodlands, Texas, January 16-18, 1992*, pp. 22–39. 1992. arXiv:hep-th/9209058 [hep-th]. http://alice.cern.ch/format/showfull?sysnb=0157265.

[26] L. Susskind, "Trouble for remnants," arXiv:hep-th/9501106 [hep-th].

[27] J. D. Bekenstein, "Black holes and information theory," *Contemp. Phys.* **45** (2003) 31–43, arXiv:quant-ph/0311049 [quant-ph].

[28] S. B. Giddings and A. Strominger, "Loss of Incoherence and Determination of Coupling Constants in Quantum Gravity," *Nucl. Phys.* **B307** (1988) 854–866.

[29] J. Polchinski and A. Strominger, "A possible resolution of the black hole information puzzle," *Phys. Rev.* **D50** (1994) 7403–7409, arXiv:hep-th/9407008 [hep-th].

[30] T. Banks, L. Susskind, and M. E. Peskin, "Difficulties for the Evolution of Pure States Into Mixed States," *Nucl. Phys.* **B244** (1984) 125–134.

[31] W. G. Unruh and R. M. Wald, "Information Loss," arXiv:1703.02140 [hep-th].

[32] J. Preskill, "Lecture notes for physics 229: Quantum information and computation,". http://www.theory.caltech.edu/people/preskill/ph229/.

[33] L. Susskind, L. Thorlacius, and J. Uglum, "The Stretched horizon and black hole complementarity," *Phys. Rev.* **D48** (1993) 3743–3761, arXiv:hep-th/9306069 [hep-th].

[34] D. A. Lowe, J. Polchinski, L. Susskind, L. Thorlacius, and J. Uglum, "Black hole complementarity versus locality," *Phys. Rev.* **D52** (1995) 6997–7010, arXiv:hep-th/9506138 [hep-th].

[35] Y. Kiem, H. L. Verlinde, and E. P. Verlinde, "Black hole horizons and complementarity," *Phys. Rev.* **D52** (1995) 7053–7065, arXiv:hep-th/9502074 [hep-th].

[36] G. 't Hooft, "On the Quantum Structure of a Black Hole," *Nucl. Phys.* **B256** (1985) 727–745.

[37] G. 't Hooft, "The black hole interpretation of string theory," *Nucl. Phys.* **B335** (1990) 138–154.

[38] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature* **299** (1982) 802–803.

[39] D. Dieks, "Communication by EPR devices," *Phys. Lett.* **A92** (1982) 271–272.

[40] S. B. Giddings, "Locality in quantum gravity and string theory," *Phys. Rev.* **D74** (2006) 106006, arXiv:hep-th/0604072 [hep-th].

[41] E. Silverstein, "Backdraft: String Creation in an Old Schwarzschild Black Hole," arXiv:1402.1486 [hep-th].

[42] M. Dodelson and E. Silverstein, "String-theoretic breakdown of effective field theory near black hole horizons," arXiv:1504.05536 [hep-th].

[43] M. Dodelson and E. Silverstein, "Longitudinal nonlocality in the string S-matrix," arXiv:1504.05537 [hep-th].

[44] L. Susskind and L. Thorlacius, "Gedanken experiments involving black holes," *Phys. Rev.* **D49** (1994) 966–974, arXiv:hep-th/9308100 [hep-th].

[45] P. Hayden and J. Preskill, "Black holes as mirrors: Quantum information in random subsystems," *JHEP* **09** (2007) 120, arXiv:0708.4025 [hep-th].

[46] A. Duncan, *The Conceptual Framework of Quantum Field Theory*. Oxford Univ. Press, Oxford, UK, 2012. https://global.oup.com/academic/product/the-conceptual-framework-of-quantum-field-theory-9780199573264?cc=ca&lang=en&.

[47] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black Holes: Complementarity or Firewalls?," *JHEP* **02** (2013) 062, arXiv:1207.3123 [hep-th].

[48] S. B. Giddings, "Models for unitary black hole disintegration," *Phys. Rev.* **D85** (2012) 044038, arXiv:1108.2015 [hep-th].

[49] S. G. Avery, "Qubit Models of Black Hole Evaporation," *JHEP* **01** (2013) 176, arXiv:1109.2911 [hep-th].

[50] S. L. Braunstein, S. Pirandola, and K. Życzkowski, "Better Late than Never: Information Retrieval from Black Holes," *Phys. Rev. Lett.* **110** no. 10, (2013) 101301, arXiv:0907.1190 [quant-ph].

[51] R. Bousso, "Firewalls from double purity," *Phys. Rev.* **D88** no. 8, (2013) 084035, arXiv:1308.2665 [hep-th].

[52] J. D. Bekenstein, "Universal upper bound on the entropy-to-energy ratio for bounded systems," *Phys. Rev. D* **23** (Jan, 1981) 287–298. https://link.aps.org/doi/10.1103/PhysRevD.23.287.

[53] R. Bousso, "A covariant entropy conjecture," *JHEP* **07** (1999) 004, arXiv:hep-th/9905177 [hep-th].

[54] R. Bousso, "The holographic principle," *Rev. Mod. Phys.* **74** (2002) 825–874, arXiv:hep-th/0203101 [hep-th].

[55] A. Strominger and D. M. Thompson, "A quantum Bousso bound," *Phys. Rev.* **D70** (2004) 044007, arXiv:hep-th/0303067 [hep-th].

[56] R. Bousso, B. Freivogel, and S. Leichenauer, "Saturating the holographic entropy bound," *Phys. Rev.* **D82** (2010) 084024, arXiv:1003.3012 [hep-th].

[57] G. 't Hooft, "Dimensional reduction in quantum gravity," in *Salamfest 1993:0284-296*, pp. 0284–296. 1993. arXiv:gr-qc/9310026 [gr-qc].

[58] L. Susskind, "The World as a hologram," *J. Math. Phys.* **36** (1995) 6377–6396, arXiv:hep-th/9409089 [hep-th].

[59] D. Bigatti and L. Susskind, "TASI lectures on the holographic principle," in *Strings, branes and gravity. Proceedings, Theoretical Advanced Study Institute, TASI'99, Boulder, USA, May 31-June 25, 1999*, pp. 883–933. 1999. arXiv:hep-th/0002044 [hep-th]. http://alice.cern.ch/format/showfull?sysnb=2175061.

[60] M. Van Raamsdonk, "Building up spacetime with quantum entanglement," *Gen. Rel. Grav.* **42** (2010) 2323–2329, arXiv:1005.3035 [hep-th]. [Int. J. Mod. Phys.D19,2429(2010)].

[61] J. M. Maldacena, "The Large N limit of superconformal field theories and supergravity," *Int. J. Theor. Phys.* **38** (1999) 1113–1133, arXiv:hep-th/9711200 [hep-th]. [Adv. Theor. Math. Phys.2,231(1998)].

[62] M. Ammon and J. Erdmenger, *Gauge/gravity duality*. Cambridge Univ. Pr., Cambridge, UK, 2015. http://www.cambridge.org/de/academic/subjects/physics/theoretical-physics-and-mathematical-physics/gaugegravity-duality-foundations-and-applications.

[63] O. Aharony, S. S. Gubser, J. M. Maldacena, H. Ooguri, and Y. Oz, "Large N field theories, string theory and gravity," *Phys. Rept.* **323** (2000) 183–386, arXiv:hep-th/9905111 [hep-th].

[64] V. E. Hubeny, "The AdS/CFT Correspondence," *Class. Quant. Grav.* **32** no. 12, (2015) 124010, arXiv:1501.00007 [gr-qc].

[65] J. Polchinski, "Introduction to Gauge/Gravity Duality," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June 1-25, 2010*, pp. 3–46. 2010. arXiv:1010.6134 [hep-th].

[66] E. D'Hoker and D. Z. Freedman, "Supersymmetric gauge theories and the AdS / CFT correspondence," in *Strings, Branes and Extra Dimensions: TASI 2001: Proceedings*, pp. 3–158. 2002. arXiv:hep-th/0201253 [hep-th].

[67] S. W. Hawking, "The Chronology protection conjecture," *Phys. Rev.* **D46** (1992) 603–611.

[68] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2011.

[69] P. Di Francesco, P. Mathieu, and D. Senechal, *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer-Verlag, New York, 1997. http://www-spires.fnal.gov/spires/find/books/www?cl=QC174.52.C66D5::1997.

[70] D. Tong, "String Theory," arXiv:0908.0333 [hep-th].

[71] J. L. Cardy, "Operator Content of Two-Dimensional Conformally Invariant Theories," *Nucl. Phys.* **B270** (1986) 186–204.

[72] E. P. Verlinde, "On the holographic principle in a radiation dominated universe," arXiv:hep-th/0008140 [hep-th].

[73] A. B. Zamolodchikov, "Irreversibility of the Flux of the Renormalization Group in a 2D Field Theory," *JETP Lett.* **43** (1986) 730–732. [Pisma Zh. Eksp. Teor. Fiz.43,565(1986)].

[74] M. B. Green, J. H. Schwarz, and E. Witten, *SUPERSTRING THEORY. VOL. 1: INTRODUCTION*. Cambridge Monographs on Mathematical Physics. 1988. http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/superstring-theory-volume-1.

[75] J. de Boer, "Introduction to the AdS/CFT correspondence," in *Supersymmetry and unification of fundamental interactions. Proceedings, 10th International Conference, SUSY'02, Hamburg, Germany, June 17-23, 2002*, pp. 512–527. 2002. http://www-library.desy.de/preparch/desy/proc/proc02-02/Proceedings/pl.6/deboer_pr.pdf.

[76] J. Polchinski, "Dualities of Fields and Strings," arXiv:1412.5704 [hep-th].

[77] L. Susskind and E. Witten, "The Holographic bound in anti-de Sitter space," arXiv:hep-th/9805114 [hep-th].

[78] G. 't Hooft, "A Planar Diagram Theory for Strong Interactions," *Nucl. Phys.* **B72** (1974) 461.

[79] J. de Boer, "The Holographic renormalization group," *Fortsch. Phys.* **49** (2001) 339–358, arXiv:hep-th/0101026 [hep-th].

[80] J. de Boer, E. P. Verlinde, and H. L. Verlinde, "On the holographic renormalization group," *JHEP* **08** (2000) 003, arXiv:hep-th/9912012 [hep-th].

[81] E. T. Akhmedov, "A Remark on the AdS / CFT correspondence and the renormalization group flow," *Phys. Lett.* **B442** (1998) 152–158, arXiv:hep-th/9806217 [hep-th].

[82] V. Balasubramanian and P. Kraus, "Space-time and the holographic renormalization group," *Phys. Rev. Lett.* **83** (1999) 3605–3608, arXiv:hep-th/9903190 [hep-th].

[83] A. Mukhopadhyay, "Understanding the holographic principle via RG flow," *Int. J. Mod. Phys.* **A31** no. 34, (2016) 1630059, arXiv:1612.00141 [hep-th].

[84] S.-S. Lee, "Quantum Renormalization Group and Holography," *JHEP* **01** (2014) 076, arXiv:1305.3908 [hep-th].

[85] B. Swingle, "Constructing holographic spacetimes using entanglement renormalization," arXiv:1209.3304 [hep-th].

[86] S. El-Showk and K. Papadodimas, "Emergent Spacetime and Holographic CFTs," *JHEP* **10** (2012) 106, arXiv:1101.4163 [hep-th].

[87] N. Seiberg, "Emergent spacetime," in *The Quantum Structure of Space and Time: Proceedings of the 23rd Solvay Conference on Physics. Brussels, Belgium. 1 - 3 December 2005*, pp. 163–178. 2006. arXiv:hep-th/0601234 [hep-th].

[88] W. Donnelly and S. B. Giddings, "Diffeomorphism-invariant observables and their nonlocal algebra," *Phys. Rev.* **D93** no. 2, (2016) 024030, arXiv:1507.07921 [hep-th]. [Erratum: Phys. Rev.D94,no.2,029903(2016)].

[89] S. B. Giddings, D. Marolf, and J. B. Hartle, "Observables in effective gravity," *Phys. Rev.* **D74** (2006) 064018, arXiv:hep-th/0512200 [hep-th].

[90] I. R. Klebanov and E. Witten, "AdS / CFT correspondence and symmetry breaking," *Nucl. Phys.* **B556** (1999) 89–114, arXiv:hep-th/9905104 [hep-th].

[91] P. Breitenlohner and D. Z. Freedman, "Stability in Gauged Extended Supergravity," *Annals Phys.* **144** (1982) 249.

[92] V. Balasubramanian, P. Kraus, and A. E. Lawrence, "Bulk versus boundary dynamics in anti-de Sitter space-time," *Phys. Rev.* **D59** (1999) 046003, arXiv:hep-th/9805171 [hep-th].

[93] T. Banks, M. R. Douglas, G. T. Horowitz, and E. J. Martinec, "AdS dynamics from conformal field theory," arXiv:hep-th/9808016 [hep-th].

[94] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, "Gauge theory correlators from noncritical string theory," *Phys. Lett.* **B428** (1998) 105–114, arXiv:hep-th/9802109 [hep-th].

[95] E. Witten, "Anti-de Sitter space and holography," *Adv. Theor. Math. Phys.* **2** (1998) 253–291, arXiv:hep-th/9802150 [hep-th].

[96] D. Harlow and D. Stanford, "Operator Dictionaries and Wave Functions in AdS/CFT and dS/CFT," arXiv:1104.2621 [hep-th].

[97] 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].

[98] A. Lewkowycz and J. Maldacena, "Generalized gravitational entropy," *JHEP* **08** (2013) 090, arXiv:1304.4926 [hep-th].

[99] X. Dong, "Holographic Entanglement Entropy for General Higher Derivative Gravity," *JHEP* **01** (2014) 044, arXiv:1310.5713 [hep-th].

[100] R. A. Jefferson, "Gravitational entropy: from black holes to holography," in *PhD String Seminars, University of Amsterdam.* Sept. 29th, 2016.

[101] A. Almheiri, X. Dong, and D. Harlow, "Bulk Locality and Quantum Error Correction in AdS/CFT," *JHEP* **04** (2015) 163, arXiv:1411.7041 [hep-th].

[102] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, "Holographic representation of local bulk operators," *Phys. Rev.* **D74** (2006) 066009, arXiv:hep-th/0606141 [hep-th].

[103] 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].

[104] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, "Local bulk operators in AdS/CFT: A Holographic description of the black hole interior," *Phys. Rev.* **D75** (2007) 106001, arXiv:hep-th/0612053 [hep-th]. [Erratum: Phys. Rev.D75,129902(2007)].

[105] 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].

[106] D. Kabat and G. Lifschytz, "Bulk equations of motion from CFT correlators," *JHEP* **09** (2015) 059, arXiv:1505.03755 [hep-th].

[107] D. Kabat and G. Lifschytz, "Locality, bulk equations of motion and the conformal bootstrap," *JHEP* **10** (2016) 091, arXiv:1603.06800 [hep-th].

[108] D. Kabat and G. Lifschytz, "Local bulk physics from intersecting modular Hamiltonians," arXiv:1703.06523 [hep-th].

[109] I. Bena, "On the construction of local fields in the bulk of AdS(5) and other spaces," *Phys. Rev.* **D62** (2000) 066007, arXiv:hep-th/9905186 [hep-th].

[110] S. Leichenauer and V. Rosenhaus, "AdS black holes, the bulk-boundary dictionary, and smearing functions," *Phys. Rev.* **D88** no. 2, (2013) 026003, arXiv:1304.6821 [hep-th].

[111] T. Nishioka, S. Ryu, and T. Takayanagi, "Holographic Entanglement Entropy: An Overview," *J. Phys.* **A42** (2009) 504008, arXiv:0905.0932 [hep-th].

[112] V. E. Hubeny, M. Rangamani, and T. Takayanagi, "A Covariant holographic entanglement entropy proposal," *JHEP* **07** (2007) 062, arXiv:0705.0016 [hep-th].

[113] V. E. Hubeny and M. Rangamani, "Causal Holographic Information," *JHEP* **06** (2012) 114, arXiv:1204.1698 [hep-th].

[114] V. E. Hubeny, M. Rangamani, and E. Tonni, "Global properties of causal wedges in asymptotically AdS spacetimes," *JHEP* **10** (2013) 059, arXiv:1306.4324 [hep-th].

[115] V. E. Hubeny, M. Rangamani, and E. Tonni, "Thermalization of Causal Holographic Information," *JHEP* **05** (2013) 136, arXiv:1302.0853 [hep-th].

[116] J. M. Maldacena, "Wilson loops in large N field theories," *Phys. Rev. Lett.* **80** (1998) 4859–4862, arXiv:hep-th/9803002 [hep-th].

[117] S.-J. Rey and J.-T. Yee, "Macroscopic strings as heavy quarks in large N gauge theory and anti-de Sitter supergravity," *Eur. Phys. J.* **C22** (2001) 379–394, arXiv:hep-th/9803001 [hep-th].

[118] Y. Makeenko, "A Brief Introduction to Wilson Loops and Large N," *Phys. Atom. Nucl.* **73** (2010) 878–894, arXiv:0906.4487 [hep-th].

[119] S. I. Finazzo and J. Noronha, "Estimates for the Thermal Width of Heavy Quarkonia in Strongly Coupled Plasmas from Holography," *JHEP* **11** (2013) 042, arXiv:1306.2613 [hep-ph].

[120] Y. Makeenko, "Large N gauge theories," *NATO Sci. Ser. C* **556** (2000) 285–354, arXiv:hep-th/0001047 [hep-th].

[121] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*. 1995. http://www.slac.stanford.edu/spires/find/books/www?cl=QC174.45%3AP4.

[122] S. L. Braunstein, H.-J. Sommers, and K. Zyczkowski, "Entangled black holes as ciphers of hidden information," arXiv:0907.0739 [quant-ph].

[123] I. Ilgin and I.-S. Yang, "Causal Patch Complementarity: The Inside Story for Old Black Holes," *Phys. Rev.* **D89** no. 4, (2014) 044007, arXiv:1311.1219 [hep-th].

[124] B. Freivogel, "Energy and Information Near Black Hole Horizons," *JCAP* **1407** (2014) 041, arXiv:1401.5340 [hep-th].

[125] D. Marolf and J. Polchinski, "Gauge/Gravity Duality and the Black Hole Interior," *Phys. Rev. Lett.* **111** (2013) 171301, arXiv:1307.4706 [hep-th].

[126] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, "An Apologia for Firewalls," *JHEP* **09** (2013) 018, arXiv:1304.6483 [hep-th].

[127] D. Harlow and P. Hayden, "Quantum Computation vs. Firewalls," *JHEP* **1306** (2013) 085, arXiv:1301.4504 [hep-th].

[128] L. Hui and I.-S. Yang, "Complementarity plus backreaction is enough," *Phys. Rev.* **D89** no. 8, (2014) 084011, arXiv:1308.6268 [hep-th].

[129] R. Bousso, B. Freivogel, S. Leichenauer, V. Rosenhaus, and C. Zukowski, "Null Geodesics, Local CFT Operators and AdS/CFT for Subregions," *Phys. Rev.* **D88** (2013) 064057, arXiv:1209.4641 [hep-th].

[130] S.-J. Rey and V. Rosenhaus, "Scanning Tunneling Macroscopy, Black Holes, and AdS/CFT Bulk Locality," *JHEP* **07** (2014) 050, arXiv:1403.3943 [hep-th].

[131] A. R. Brown, "Tensile Strength and the Mining of Black Holes," *Phys. Rev. Lett.* **111** no. 21, (2013) 211301, arXiv:1207.3342 [gr-qc].

[132] 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].

[133] K. Martel and E. Poisson, "Regular coordinate systems for Schwarzschild and other spherical space-times," *Am. J. Phys.* **69** (2001) 476–480, arXiv:gr-qc/0001069 [gr-qc].

[134] L. Susskind, "Black Hole Complementarity and the Harlow-Hayden Conjecture," arXiv:1301.4505 [hep-th].

[135] H. Araki and E. H. Lieb, "Entropy inequalities," *Comm. Math. Phys.* **18** no. 2, (1970) 160–170. http://projecteuclid.org/euclid.cmp/1103842506.

[136] R. Cleve, D. Gottesman, and H.-K. Lo, "How to share a quantum secret," *Phys. Rev. Lett.* **83** (1999) 648–651, arXiv:quant-ph/9901025 [quant-ph].

[137] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," *Fortsch.Phys.* **61** (2013) 781–811, arXiv:1306.0533 [hep-th].

[138] S. Ryu and T. Takayanagi, "Aspects of Holographic Entanglement Entropy," *JHEP* **08** (2006) 045, arXiv:hep-th/0605073 [hep-th].

[139] E. Bianchi and R. C. Myers, "On the Architecture of Spacetime Geometry," *Class. Quant. Grav.* **31** (2014) 214002, arXiv:1212.5183 [hep-th].

[140] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, "Gravitation from Entanglement in Holographic CFTs," *JHEP* **03** (2014) 051, arXiv:1312.7856 [hep-th].

[141] M. Headrick, V. E. Hubeny, A. Lawrence, and M. Rangamani, "Causality & holographic entanglement entropy," *JHEP* **12** (2014) 162, arXiv:1408.6300 [hep-th].

[142] V. Balasubramanian, B. D. Chowdhury, B. Czech, J. de Boer, and M. P. Heller, "Bulk curves from boundary data in holography," *Phys. Rev.* **D89** no. 8, (2014) 086004, arXiv:1310.4204 [hep-th].

[143] S. A. Hartnoll and R. Mahajan, "Holographic mutual information and distinguishability of Wilson loop and defect operators," *JHEP* **02** (2015) 100, arXiv:1407.8191 [hep-th].

[144] B. Czech, X. Dong, and J. Sully, "Holographic Reconstruction of General Bulk Surfaces," *JHEP* **11** (2014) 015, arXiv:1406.4889 [hep-th].

[145] V. Balasubramanian, B. D. Chowdhury, B. Czech, and J. de Boer, "Entwinement and the emergence of spacetime," *JHEP* **01** (2015) 048, arXiv:1406.5859 [hep-th].

[146] V. E. Hubeny, "Extremal surfaces as bulk probes in AdS/CFT," *JHEP* **07** (2012) 093, arXiv:1203.1044 [hep-th].

[147] N. Engelhardt and A. C. Wall, "Extremal Surface Barriers," *JHEP* **03** (2014) 068, arXiv:1312.3699 [hep-th].

[148] V. E. Hubeny, H. Maxfield, M. Rangamani, and E. Tonni, "Holographic entanglement plateaux," *JHEP* **08** (2013) 092, arXiv:1306.4004 [hep-th].

[149] B. Freivogel and B. Mosk, "Properties of Causal Holographic Information," *JHEP* **09** (2013) 100, arXiv:1304.7229 [hep-th].

[150] W. R. Kelly and A. C. Wall, "Coarse-grained entropy and causal holographic information in AdS/CFT," *JHEP* **03** (2014) 118, arXiv:1309.3610 [hep-th].

[151] L. Pestov and G. Uhlmann, "Two dimensional compact simple Riemannian manifolds are boundary distance rigid," *Ann. Math.* **161** no. 2, (2003) , arXiv:0305280 [math].

[152] F. Nogueira, "Extremal Surfaces in Asymptotically AdS Charged Boson Stars Backgrounds," *Phys. Rev.* **D87** (2013) 106006, arXiv:1301.4316 [hep-th].

[153] S. A. Gentle and M. Rangamani, "Holographic entanglement and causal information in coherent states," *JHEP* **01** (2014) 120, arXiv:1311.0015 [hep-th].

[154] H. Bondi, "Massive Spheres in General Relativity," *Proceedings of the Royal Society of London Series A* **282** (Nov., 1964) 303–317.

[155] M. Kleban, A. Lawrence, M. M. Roberts, and S. Storace, "Metastability and instability in holographic gauge theories," *JHEP* **06** (2014) 152, arXiv:1312.1312 [hep-th].

[156] A. F. Astaneh, G. Gibbons, and S. N. Solodukhin, "What surface maximizes entanglement entropy?," *Phys. Rev.* **D90** no. 8, (2014) 085021, arXiv:1407.4719 [hep-th].

[157] G. T. Horowitz, "Comments on black holes in string theory," *Class. Quant. Grav.* **17** (2000) 1107–1116, arXiv:hep-th/9910082 [hep-th].

[158] R. Gregory, "The Gregory-Laflamme instability," in *Black holes in higher dimensions*, G. T. Horowitz, ed., pp. 29–43. 2011. arXiv:1107.5821 [gr-qc].

[159] T. Hollowood, S. P. Kumar, and A. Naqvi, "Instabilities of the Small Black Hole: A View from N=4 SYM," *JHEP* **01** (2007) 001, arXiv:hep-th/0607111 [hep-th].

[160] A. Brandhuber, N. Itzhaki, J. Sonnenschein, and S. Yankielowicz, "Wilson loops in the large N limit at finite temperature," *Phys. Lett.* **B434** (1998) 36–40, arXiv:hep-th/9803137 [hep-th].

[161] D. S. Ageev and I. Ya. Aref'eva, "Holography and nonlocal operators for the BTZ black hole with nonzero angular momentum," *Theor. Math. Phys.* **180** (2014) 881–893, arXiv:1402.6937 [hep-th]. [Teor. Mat. Fiz.180,no.2,147(2014)].

[162] S.-J. Rey, S. Theisen, and J.-T. Yee, "Wilson-Polyakov loop at finite temperature in large N gauge theory and anti-de Sitter supergravity," *Nucl. Phys.* **B527** (1998) 171–186, arXiv:hep-th/9803135 [hep-th].

[163] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, "Gravitational dynamics from entanglement 'thermodynamics'," *JHEP* **04** (2014) 195, arXiv:1308.3716 [hep-th].

[164] J. Polchinski, L. Susskind, and N. Toumbas, "Negative energy, superluminosity and holography," *Phys. Rev.* **D60** (1999) 084006, arXiv:hep-th/9903228 [hep-th].

[165] S. B. Giddings and M. Lippert, "Precursors, black holes, and a locality bound," *Phys. Rev.* **D65** (2002) 024006, arXiv:hep-th/0103231 [hep-th].

[166] B. Freivogel, S. B. Giddings, and M. Lippert, "Toward a theory of precursors," *Phys. Rev.* **D66** (2002) 106002, arXiv:hep-th/0207083 [hep-th].

[167] B. Chilian and K. Fredenhagen, "The Time Slice Axiom in Perturbative Quantum Field Theory on Globally Hyperbolic Spacetimes," *Communications in Mathematical Physics* **287** no. 2, (2009) 513–522. arXiv:0802.1642.

[168] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence," *JHEP* **06** (2015) 149, arXiv:1503.06237 [hep-th].

[169] Z. Yang, P. Hayden, and X.-L. Qi, "Bidirectional holographic codes and sub-AdS locality," *JHEP* **01** (2016) 175, arXiv:1510.03784 [hep-th].

[170] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, "Holographic duality from random tensor networks," *JHEP* **11** (2016) 009, arXiv:1601.01694 [hep-th].

[171] E. Mintun, J. Polchinski, and V. Rosenhaus, "Bulk-Boundary Duality, Gauge Invariance, and Quantum Error Corrections," *Phys. Rev. Lett.* **115** no. 15, (2015) 151601, arXiv:1501.06577 [hep-th].

[172] F. Ferrari, "Gauge Theories, D-Branes and Holography," *Nucl. Phys.* **B880** (2014) 247–289, arXiv:1310.6788 [hep-th].

[173] T. Andrade and D. Marolf, "AdS/CFT beyond the unitarity bound," *JHEP* **01** (2012) 049, arXiv:1105.6337 [hep-th].

[174] 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].

[175] 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].

[176] T. Faulkner, A. Lewkowycz, and J. Maldacena, "Quantum corrections to holographic entanglement entropy," *JHEP* **1311** (2013) 074, arXiv:1307.2892.

[177] N. Bao and I. H. Kim, "Precursor problem and holographic mutual information," arXiv:1601.07616 [hep-th].

[178] 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].

[179] 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].

[180] I. A. Morrison, "Boundary-to-bulk maps for AdS causal wedges and the Reeh-Schlieder property in holography," *JHEP* **05** (2014) 053, arXiv:1403.3426 [hep-th].

[181] M. Guica and D. L. Jafferis, "On the construction of charged operators inside an eternal black hole," arXiv:1511.05627 [hep-th].

[182] S. Leichenauer and M. Moosa, "Entanglement Tsunami in (1+1)-Dimensions," *Phys. Rev.* **D92** (2015) 126004, arXiv:1505.04225 [hep-th].

[183] J. D. Brown and M. Henneaux, "Central Charges in the Canonical Realization of Asymptotic Symmetries: An Example from Three-Dimensional Gravity," *Commun. Math. Phys.* **104** (1986) 207–226.

[184] I. Heemskerk, J. Penedones, J. Polchinski, and J. Sully, "Holography from Conformal Field Theory," *JHEP* **10** (2009) 079, arXiv:0907.0151 [hep-th].

[185] A. L. Fitzpatrick and J. Kaplan, "AdS Field Theory from Conformal Field Theory," *JHEP* **02** (2013) 054, arXiv:1208.0337 [hep-th].

[186] T. Hartman, C. A. Keller, and B. Stoica, "Universal Spectrum of 2d Conformal Field Theory in the Large c Limit," *JHEP* **09** (2014) 118, arXiv:1405.5137 [hep-th].

[187] A. L. Fitzpatrick, J. Kaplan, and M. T. Walters, "Universality of Long-Distance AdS Physics from the CFT Bootstrap," *JHEP* **08** (2014) 145, arXiv:1403.6829 [hep-th].

[188] A. Belin, C. A. Keller, and A. Maloney, "String Universality for Permutation Orbifolds," *Phys. Rev.* **D91** no. 10, (2015) 106005, arXiv:1412.7159 [hep-th].

[189] F. M. Haehl and M. Rangamani, "Permutation orbifolds and holography," *JHEP* **03** (2015) 163, arXiv:1412.2759 [hep-th].

[190] A. Belin, C. A. Keller, and A. Maloney, "Permutation Orbifolds in the large N Limit," arXiv:1509.01256 [hep-th].

[191] E. Shaghoulian, "Emergent gravity from Eguchi-Kawai reduction," arXiv:1611.04189 [hep-th].

[192] A. Belin, J. de Boer, J. Kruthoff, B. Michel, E. Shaghoulian, and M. Shyani, "Universality of Sparse d>2 Conformal Field Theory at Large N," arXiv:1610.06186 [hep-th].

[193] S. Banerjee, A. Castro, S. Hellerman, E. Hijano, A. Lepage-Jutier, A. Maloney, and S. Shenker, "Smoothed Transitions in Higher Spin AdS Gravity," *Class. Quant. Grav.* **30** (2013) 104001, arXiv:1209.5396 [hep-th].

[194] M. R. Gaberdiel, R. Gopakumar, and M. Rangamani, "The Spectrum of Light States in Large N Minimal Models," *JHEP* **01** (2014) 116, arXiv:1310.1744 [hep-th].

[195] P. Kraus and A. Maloney, "A Cardy Formula for Three-Point Coefficients: How the Black Hole Got its Spots," arXiv:1608.03284 [hep-th].

[196] E. Mintun, J. Polchinski, and V. Rosenhaus, "Bulk-Boundary Duality, Gauge Invariance, and Quantum Error Corrections," *Phys. Rev. Lett.* **115** no. 15, (2015) 151601, arXiv:1501.06577 [hep-th].

[197] I. Klebanov and A. Polyakov, "AdS dual of the critical O(N) vector model," *Phys.Lett.* **B550** (2002) 213–219, arXiv:hep-th/0210114 [hep-th].

[198] S. Giombi and X. Yin, "The Higher Spin/Vector Model Duality," *J. Phys.* **A46** (2013) 214003, arXiv:1208.4036 [hep-th].

[199] S. B. Giddings, "Flat space scattering and bulk locality in the AdS / CFT correspondence," *Phys. Rev.* **D61** (2000) 106008, arXiv:hep-th/9907129 [hep-th].

[200] M. Gary, S. B. Giddings, and J. Penedones, "Local bulk S-matrix elements and CFT singularities," *Phys. Rev.* **D80** (2009) 085005, arXiv:0903.4437 [hep-th].

[201] M. Gary and S. B. Giddings, "The Flat space S-matrix from the AdS/CFT correspondence?," *Phys. Rev.* **D80** (2009) 046008, arXiv:0904.3544 [hep-th].

[202] J. Maldacena, D. Simmons-Duffin, and A. Zhiboedov, "Looking for a bulk point," arXiv:1509.03612 [hep-th].

[203] M. A. Vasiliev, "Higher spin gauge theories in four-dimensions, three-dimensions, and two-dimensions," *Int. J. Mod. Phys.* **D5** (1996) 763–797, arXiv:hep-th/9611024 [hep-th].

[204] M. R. Gaberdiel and C. Vollenweider, "Minimal Model Holography for SO(2N)," *JHEP* **08** (2011) 104, arXiv:1106.2634 [hep-th].

[205] C. Candu, M. R. Gaberdiel, M. Kelm, and C. Vollenweider, "Even spin minimal model holography," *JHEP* **01** (2013) 185, arXiv:1211.3113 [hep-th].

[206] S. Banerjee, S. Hellerman, J. Maltz, and S. H. Shenker, "Light States in Chern-Simons Theory Coupled to Fundamental Matter," *JHEP* **03** (2013) 097, arXiv:1207.4195 [hep-th].

[207] M. R. Gaberdiel, K. Jin, and W. Li, "Perturbations of W(infinity) CFTs," *JHEP* **10** (2013) 162, arXiv:1307.4087 [hep-th].

[208] M. A. Vasiliev, "Star-Product Functions in Higher-Spin Theory and Locality," *JHEP* **06** (2015) 031, arXiv:1502.02271 [hep-th].

[209] E. D. Skvortsov and M. Taronna, "On Locality, Holography and Unfolding," *JHEP* **11** (2015) 044, arXiv:1508.04764 [hep-th].

[210] M. Taronna, "Pseudo-local Theories: A Functional Class Proposal," in *Proceedings, International Workshop on Higher Spin Gauge Theories: Singapore, Singapore, November 4-6, 2015*, pp. 59–84. 2017. arXiv:1602.08566 [hep-th].

[211] M. A. Vasiliev, "Current Interactions, Locality and Holography from the 0-Form Sector of Nonlinear Higher-Spin Equations," arXiv:1605.02662 [hep-th].

[212] M. Taronna, "A Note on Field Redefinitions and Higher-Spin Equations," *EPJ Web Conf.* **125** (2016) 05025, arXiv:1607.04718 [hep-th].

[213] X. Bekaert, J. Erdmenger, D. Ponomarev, and C. Sleight, "Bulk quartic vertices from boundary four-point correlators," in *Proceedings, International Workshop on Higher Spin Gauge Theories: Singapore, Singapore, November 4-6, 2015*, pp. 291–303. 2017. arXiv:1602.08570 [hep-th].

[214] C. Sleight, *Interactions in Higher-Spin Gravity: a Holographic Perspective*. PhD thesis, Munich U., 2016. arXiv:1610.01318 [hep-th].

[215] C. Sleight and M. Taronna, "Higher-Spin Algebras, Holography and Flat Space," arXiv:1609.00991 [hep-th].

[216] S. B. Giddings, "Nonviolent nonlocality," *Phys. Rev.* **D88** (2013) 064023, arXiv:1211.7070 [hep-th].

[217] S. B. Giddings and Y. Shi, "Effective field theory models for nonviolent information transfer from black holes," *Phys. Rev.* **D89** no. 12, (2014) 124032, arXiv:1310.5700 [hep-th].

[218] J. Lin, "A Toy Model of Entwinement," arXiv:1608.02040 [hep-th].

[219] L. Susskind, "Entanglement is not enough," *Fortsch. Phys.* **64** (2016) 49–71, arXiv:1411.0690 [hep-th].

[220] S. H. Shenker and D. Stanford, "Multiple Shocks," *JHEP* **12** (2014) 046, arXiv:1312.3296 [hep-th].

Contribution to publications

The publications on which this thesis is based were undertaken in close collaboration with my coauthors, to whom joint credit for intellectual content is due. Group discussions, both conceptual and technical, were a significant part of our collaboration. That said, in this section I give a brief summary of my main contributions.[2]

[1] B. Freivogel, R. A. Jefferson, L. Kabir, and I-S. Yang
"Geometry of the Infalling Causal Patch"
Phys. Rev. **D91**, 4 (2015), arXiv:1406.6043 [hep-th].

In terms of both intellectual content and writing, I was responsible for sections 3, 4, and the appendix. I also made all plots and figures, and conducted the requisite analysis for each.

[2] B. Freivogel, R. A. Jefferson, L. Kabir, B. Mosk, and I-S. Yang
"Casting Shadows on Holographic Reconstruction"
Phys. Rev. **D91**, 8 (2015), arXiv:1203.1036 [hep-th].

I wrote most of sections 1, 3, and 7, and revised the entire paper during the final editing stages to ensure a cohesive story emerged from the contributions of multiple authors, culminating in my chosen title. I was responsible for the stellar counter-example in section 3.1, was involved in the calculations in sections 4 and 5, and verified the proofs in section 2 and the appendix.

[2] Note that within our field of theoretical physics, authors are listed alphabetically, and hence the order of names does not reflect the extent of an individual's contribution.

[3] B. Freivogel, R. A. Jefferson, and L. Kabir
"Precursors, Gauge Invariance, and Quantum Error Correction in AdS/CFT"
JHEP **04**, 119 (2016), arXiv:1602.04811 [hep-th].

I wrote most of this paper, made all figures, and was responsible for final editing, abstract, title, and submission. In terms of intellectual content, sections 3 and 5 are predominantly my own work, as are much of the calculations in the appendices.

[4] "A. Belin, B. Freivogel, R. A. Jefferson, and L. Kabir
"Sub-AdS Scale Locality in AdS_3/CFT_2"
arXiv:1611.08601. [hep-th].

I participated in group discussions of all calculations, computed the density of states at low energies in section 2, and clarified the exposition of the tower of constraints in the introduction. I wrote the *Mathematica* code that produced the holomorphic primaries up to level 12 in the appendix.

Summary

It has now been 100 years since Einstein's discovery of general relativity; and yet after a century of scrutiny, it continues to mystify us. It resists unification with the other three fundamental forces into a theory of quantum gravity. And when combined with quantum mechanics, as in the case of black holes, general relativity breaks down entirely.

Recently, we've gained a new perspective on this issue. A number of interesting developments in our field suggest that spacetime is emergent. In other words, that gravity is not a fundamental force of nature, but can be explained by (or "emerges" from) some deeper quantum mechanical phenomena. This is the remarkable idea that underlies the Holographic Principle, which asserts that a D-dimensional theory with gravity is equivalent to a (D-1)-dimensional theory without it. Another way to state this is that the information required to completely specify some system scales, not with its volume as one would expect, but with its boundary surface area. This intuition-defying notion has its roots in Hawking's work on Black Hole Thermodynamics, and has since been elaborated and expanded considerably with ideas from String Theory, culminating in Maldacena's celebrated AdS/CFT correspondence.

While holography has withstood intensive scrutiny and gained broad acceptance in the theoretical physics community, it is only in recent years that we've truly begun to explore the strange and remarkable implications for quantum gravity. The idea that space and time themselves are emergent concepts implies a fundamental change in our understanding of the Universe. And as with all great paradigm shifts, this revolution will likely require us to rethink even our most deeply held theories, and to develop mathematical tools capable of casting light where our meagre intuition cannot tread. Indeed, my own research during my PhD

has unearthed a number of questions that cannot be answered with present ideas. And although recent investigations have made great progress with some of the more technical questions in this endeavour, thus far the most profound is also the least concrete, namely: how does spacetime — by which we mean, a gravitational theory obeying Einstein's General Relativity — emerge from fundamental principles of quantum field theory? The overarching ambition of this thesis is to make this question more precise by focusing on the breakdown of locality in quantum gravity.

The concepts of locality and causality are intimately linked, and are so fundamental to our mode of thought that we rarely think to question them—and our intuition has immense difficulty when we do! But as alluded above, physics has provided us with ample evidence that, for all their comforting familiarity, they do not appear as entries in Nature's rule book. There are many interesting examples in this vein, but here I will limit myself to two main avenues of inquiry to which I contributed during my PhD, and which are presented in detail in the present work.

The Black Hole Information Paradox

Also known as the Firewall Paradox in its modern form, this refers to an apparent inconsistency between foundational axioms in quantum mechanics and general relativity. It was recently shown that combining these two theories to understand black holes forces us to give up locality, unitarity, or the equivalence principle. (In brief, the original argument hinges on a property of quantum entanglement known as monogamy, which the adoption of these three axioms can be shown to violate). Yet despite the subsequent flurry of activity — the five-year-old article has been cited nearly 700 times – we have yet to find a satisfying resolution. My own contributions to this endeavour are detailed in chapter 3, which examines precisely what one would be able see if one were to fall into a black hole in pursuit of the answer. One fundamental lesson from my research seems to be that the concept of locality, while seemingly intact at everyday human scales, simply does not survive in the extreme environment of a black hole.

The Holographic Principle

Or, in its most precise form, the AdS/CFT Correspondence, is an isomorphism between a gravitational theory in anti-de Sitter (AdS) spacetime, called the "bulk", and a conformal field theory (CFT) living on the "boundary". One convenient visualization is to imagine a can of soup; the bulk is the soup, while the CFT lives on the sides of the can. Time runs upwards, with the infinite past at the bottom and the infinite future at the top. A particular horizontal slice through the can thus corresponds to a snapshot of space at an instant, with 2 dimensions in the

bulk (the filled-in circle of soup) and only 1 dimension on the boundary (the circle formed by the can). The remarkable aspect of this construction is that it's what physicists call a "duality", which means (to push the analogy a bit further) that the gravitational dynamics of the soup are somehow encoded in the quantum state of the can. The million-dollar question is: how?!

A main theme of my research in holography has been the role that nonlocality must play in the reconstruction of the bulk from the boundary. In particular, one of my important contributions, discussed in chapter 4, was to show that none of the present reconstruction schemes is sufficient! This is due to an effect I called "holographic shadows": regions of the bulk that simply cannot be reached with any boundary data. The essential problem is that current holographic probes do not adequately capture the nonlocal features (e.g., quantum entanglement) necessary for a complete reconstruction of non-trivial spacetimes. In this sense, my shadows are harbingers of the breakdown of locality in AdS/CFT, and no concrete approach yet exists for moving beyond this obstacle. Subsequently, in chapter 5, I approached this issue from a different perspective, with an examination of how gauge invariance vs. quantum error correction conspire to localize information in subregions of the boundary. Then, in chapter 6, I consider a toy model of holography that aims to pinpoint exactly where locality goes awry. The underlying lesson from my research seems to be that further progress in understanding how the bulk spacetime emerges from the boundary field theory will require a better understanding of how AdS/CFT encodes the nonlocal information required.

Both black holes and holography are especially promising areas of study, because they highlight aspects of quantum gravity in unique and illuminating ways, and my research into both areas has revealed nonlocality as a core underlying feature (for example, both demand a better understanding of how information is encoded nonlocally on the horizon/boundary). In this sense, they are the ideal laboratories for investigating the breakdown of locality, and hence in making progress towards understanding emergent spacetime and an eventual theory of quantum gravity. However, it must be emphasized that this issue of (non)locality is quite fundamental to many areas of physics, and indeed there are many other examples that speak of its importance, including gravitational dressing, entanglement in lattice gauge theories, quantum information theory, precursors, and de Sitter (i.e., cosmological) spacetimes. This widespread importance is further reflected in the increasingly interdisciplinary nature of the field, which now draws insight and experts from, e.g., computer science and quantum information theory, lattice gauge theory and condensed matter physics, algebraic quantum field theory, and string theory.

Understanding how our concept of locality may be modified, and more generally how spacetime emerges in a theory of quantum gravity, are among the most exciting and challenging open questions in theoretical physics today. We have

endeavored in this thesis to take a few small steps in this direction. Yet, despite all that black holes and holography have taught us, one cannot help the inspiring impression that we have only begun to learn.

Samenvatting

Het is reeds 100 jaar geleden dat Einstein zijn algemene relativiteitstheorie ontdekte; en zelfs na een eeuw van kritisch onderzoek blijft deze ons verbazen. Het laat namelijk geen unificatie toe met de drie andere fundamentele natuurkuurkrachten tot een theorie van kwantumgravitatie. Bovendien wanneer het gecombineerd wordt met kwantummechanica, zoals nodig is bij zwarte gaten, stort de algemene relativiteitstheorie helemaal in elkaar.

Onlangs hebben we een nieuw perspectief gekregen op dit probleem. Een aantal interessante ontwikkelingen in ons vakgebied suggereert dat ruimtetijd emergent is. Met andere woorden, zwaartekracht is geen fundamentele natuurkracht, maar kan verklaard worden ("emergeert" uit) een dieper onderliggend kwantummechanisch fenomeen. Dit merkwaardige idee ligt aan de basis van het holografisch principe, wat stelt dat een D-dimensionale theorie van de zwaartekracht equivalent is aan een (D-1)-dimensionale theorie zonder zwaartekracht. Een andere manier om dit te zeggen is dat de informatie die nodig is om een systeem volledig te beschrijven niet met het volume schaalt, zoals men zou verwachten, maar met de oppervlakte die het systeem omsluit. Dit tegen-intu"itieve concept komt voort uit Hawking's werk over de thermodynamica van zwarte gaten, en is sindsdien aangevuld met ideeën uit de snaartheorie, en culmineerde uiteindelijk in Maldacena's gevierde AdS/CFT correspondentie.

Terwijl holografie de test van het kritische onderzoek goed heeft doorstaan en algemeen geaccepteerd wordt onder theoretisch natuurkundigen, is het pas sinds enkele jaren dat we de vreemde en merkwaardige implicaties for kwantumgravitatie beginnen te begrijpen. Het idee dat ruimte en tijd emergente concepten zijn impliceert een fundamentele verandering in ons begrip van het universum. En zoals met alle grote paradigma veranderingen, zal deze revolutie hoogst waar-

schijnlijk vragen dat we onze meest dierbare theorieën herinterpreteren, en dat we nieuwe mathematische concepten moeten ontwikkelen om inzicht te krijgen waar onze intuïtie niet meer volstaat. Inderdaad, mijn onderzoek tijdens mijn PhD heeft een aantal vragen proberen op te lossen die niet met de huidige ideeën te beantwoorden vallen. En terwijl huidig onderzoek veel vooruitgang geboekt heeft met sommige meer technische vragen, is de meest diepe vraag ook de minst concrete, namelijk: hoe kan een ruimtetijd – waarmee we bedoelen, een theorie van de zwaartekracht die door Einstein's algemene relativiteitstheorie beschreven wordt – emergent zijn uit de fundamentele principes van een kwantumvelden theorie? Het overkoepelende doel van deze thesis is om deze vraag concreet te maken door te focussen op het afwezig zijn van lokaliteit in kwantumgravitatie.

De concepten lokaliteit en causaliteit zijn intiem verbonden, en zo fundamenteel voor onze manier van denken dat we zelden deze in vraag stellen – onze intuïtie heeft het dan ook bijzonder lastig wanneer we dat doen! Maar zoals reeds vermeld, de natuurkunde geeft ons bewijs dat deze concepten niet in het wetboek van moeder natuur staan. Er zijn verschillende interessante voorbeelden, maar hier zal ik me beperken om twee voorbeelden te geven waartoe ik heb bijgedragen in mijn PhD, en welke in detail beschreven staan in deze thesis.

De Zwarte Gaten Informatie Paradox

Ook wel gekend als de firewall paradox in zijn moderne vorm, welke refereert naar een ogenschijnlijke tegenstelling tussen fundamentele axioma's in de kwantummechanica en de algemene relativiteitstheorie. Er werd recent aangetoond that het combineren van deze twee theorieën om zwarte gaten te begrijpen, leidt tot het opgeven van lokaliteit, unitariteit, of het equivalentie principe. (Kort gezegd, het originele argument gebruikt een eigenschap van de verstrengelingsentropie die gekend staat als monogamie, welke geschonden wordt door de drie axioma's tezamen.) Ondanks de daaropvolgende opstoot van activiteit – het vijf jaar oude artikel is meer dan 700 keer geciteerd – is een bevredigende oplossing nog steeds zoek. Mijn eigen bijdrage in deze zoektocht staat uiteengezet in hoofdstuk 3, waarin onderzocht wordt wat men precies zal waarnemen indien men in een zwart gat zou springen met de onderliggende vraag in het achterhoofd. Een fundamentele les van mijn onderzoek is dat het concept van de lokaliteit, hoewel onbetwistbaar op alledaagse menselijke afstanden, gewoonweg niet meer stand houdt in de omgeving nabij een zwart gat.

Het Holografisch Principe

Oftewel, in zijn meest precieze vorm, de AdS/CFT correspondentie, is een isomorfisme tussen een gravitationele theorie in anti-de Sitter (AdS) ruimte, ook wel de "bulk"genoemd, en een conforme velden theorie (CFT) die op de "rand" leeft. Een

geschikte voorstelling is een blik soep; de bulk is de soep terwijl de CFT leeft op de rand van het blik. Tijd loopt van onder naar boven, met een oneindige verleden op de bodem en een oneindige toekomst op de top. Een zekere horizontale snede door het blik correspondeert dus met een momentopname, met 2 dimensies in de bulk (de ingevulde cirkel soep) en slechts 1 dimensie op de rand (de cirkel gevormd door het blik). Het merkwaardige aspect van deze constructie is wat natuurkundigen een "dualiteit"noemen, wat betekent (om de analogie verder voort te zetten) dat de zwaartekrachtsdynamica van de soep gecodeerd is in de kwantumtoestand van het blik. De hamvraag is nu: hoe?!

Een belangrijk thema in mijn onderzoek van de holografie is de rol die niet-lokaliteit speelt in de reconstructie van de bulk vanuit de rand. In het bijzonder, een van mijn bijdragen zoals bediscussieerd in hoofdstuk 4, bestaat er uit om aan te tonen dat geen van de huidige reconstructietechnieken adequaat is! Dit is omwille van een effect dat ik "holografische schaduwen" gedoopt heb: regio's in de bulk die gewoonweg niet bereikt kunnen worden met data op de rand. Het essentiële probleem bestaat erin dat de huidige holografische sondes onvoldoende de niet-lokale eigenschappen (e.g. verstrengeling) bevatten die nodig zijn voor een volledige reconstructie van een niet-triviale ruimtetijd. In deze zin, zijn mijn schaduwen voorlopers van het afwezig zijn van lokaliteit in AdS/CFT, en er bestaat voorlopig geen concrete aanpak om dit probleem te overwinnen. Vervolgens, in hoofdstuk 5,heb ik dit probleem vanuit een ander perspectief benaderd, door een studie van hoe ijkinvariantie vs. kwantumfouten correctie samenzweren om informatie op de rand te lokaliseren. Daarna, in hoofdstuk 6, beschouw ik een eenvoudig holografisch model dat exact probeert aan te wijzen waar het misloopt met lokaliteit. De onderliggende les van mijn onderzoek suggereert dat verdere vooruitgang in het begrijpen hoe bulk ruimtetijd emergent is uit de velden theorie op de rand, een beter begrip van hoe AdS/CFT niet-lokale informatie codeert nodig heeft.

Zowel zwarte gaten en holografie zijn bijzonder interessante studiedomeinen, omdat ze de aspecten van kwantumgravitatie in de verf zetten in een unieke en verhelderende manier, en mijn onderzoek in beide domeinen heeft aangetoond dat niet-lokaliteit een onderliggende kerneigenschap is (bijvoorbeeld, beide domeinen vragen een beter begrip van hoe informatie niet-lokaal opgeslagen is op de horizon/rand). In deze zin zijn beiden ideale laboratoria om de afwezigheid van niet-lokaliteit te bestuderen,en om vooruitgang te maken richting het beter begrijpen van emergente ruimtetijd, en uiteindelijk een theorie van kwantumzwaartekracht. Hoewel, het moet worden benadrukt dat het probleem van niet-lokaliteit fundamenteel is aan verschillende deelgebieden van de natuurkunde, en er zijn inderdaad talrijke voorbeelden die dit onderstrepen, zoals gravitationele dressing, verstrengeling in rooster ijktheorieën, kwantum informatie theorie, precursoren, en de Sitter (i.e. kosmologische) ruimtes. Het wijdverspreide belang is verder weerspiegeld

in de toenemende mate van interdisciplinariteit van het veld, welke nu inzichten en experten aantrekt uit o.a. computerwetenschappen en kwantum informatie theorie, rooster ijktheorie en gecondenseerde materie natuurkunde, algebraïsche kwantumvelden theorie, en snaartheorie.

Begrijpen hoe ons concept van lokaliteit gewijzigd kan worden, en meer algemeen hoe ruimtetijd emergent is in een theorie van kwantumgravitatie, zijn een van de meest opwindende en uitdagende open vragen in de theoretische natuurkunde vandaag. In deze thesis hebben we een poging ondernomen om enkele kleine stappen in deze richting te zetten. Ondanks alles wat zwarte gaten en holografie ons geleerd hebben, kan men zich niet van de indruk ontdoen dat we maar net begonnen zijn met alles te begrijpen.