Replica wormholes and the entropy of Hawking radiation

Ahmed Almheiri,[a] **Thomas Hartman,**[b] **Juan Maldacena,**[a] **Edgar Shaghoulian**[b]
and Amirhossein Tajdini[b]

[a]*Institute for Advanced Study,*
Princeton, New Jersey, U.S.A.

[b]*Department of Physics, Cornell University,*
Ithaca, New York, U.S.A.

E-mail: almheiri@ias.edu, hartman@cornell.edu, malda@ias.edu,
eshaghoulian@cornell.edu, at734@cornell.edu

ABSTRACT: The information paradox can be realized in anti-de Sitter spacetime joined to a Minkowski region. In this setting, we show that the large discrepancy between the von Neumann entropy as calculated by Hawking and the requirements of unitarity is fixed by including new saddles in the gravitational path integral. These saddles arise in the replica method as complexified wormholes connecting different copies of the black hole. As the replica number $n \to 1$, the presence of these wormholes leads to the island rule for the computation of the fine-grained gravitational entropy. We discuss these replica wormholes explicitly in two-dimensional Jackiw-Teitelboim gravity coupled to matter.

KEYWORDS: 2D Gravity, Black Holes, Models of Quantum Gravity

ARXIV EPRINT: 1911.12333

Contents

1 Introduction

Hawking famously noted that the process of black hole formation and evaporation seems to create entropy [1]. We can form a black hole from a pure state. The formation of the black hole horizon leaves an inaccessible region behind, and the entanglement of quantum fields across the horizon is responsible for the thermal nature of the Hawking radiation as well as its growing entropy.
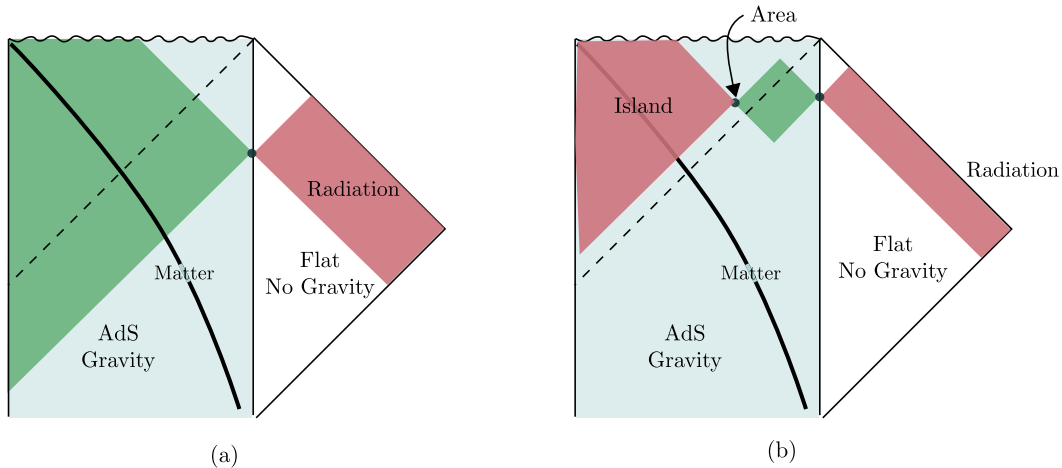
Figure 1. We display an evaporating black hole. The vertical line separates a region on the left where gravity is dynamical from a region on the right where we can approximate it as not being dynamical. The black hole is evaporating into this second region. In red we see the regions associated to the computation of the entropy of radiation and in green the regions computing the entropy of the black hole. (a) Early times. (b) Late times, where we have an island.

A useful diagnostic for information loss is the fine-grained (von Neumann) entropy of the Hawking radiation, $S_R = -\text{Tr}\,\rho_R \log \rho_R$, where ρ_R is the density matrix of the radiation. This entropy initially increases, because the Hawking radiation is entangled with its partners in the black hole interior. But if the evaporation is unitary, then it must eventually fall back to zero following the Page curve [2, 3]. On the other hand, Hawking's calculation predicts an entropy that rises monotonically as the black hole evaporates.

Hawking's computation of the entropy seems straightforward. It can be done far from the black hole where the effects of quantum gravity are small, so it is unclear what could have gone wrong. An answer to this puzzle was recently proposed [4–6] (see also [7–19]). The proposal is that Hawking used the wrong formula for computing the entropy. As the theory is coupled to gravity, we should use the proper gravitational formula for entropy: the gravitational fine-grained entropy formula studied by Ryu and Takayanagi [20] and extended in [21–23], also allowing for spatially disconnected regions, called "islands," see figure 1. Even though the radiation lives in a region where the gravitational effects are small, the fact that we are describing a state in a theory of gravity implies that we should use the gravitational formula for the entropy, including the island rule.

In this paper we consider a version of the information paradox formulated recently in [4, 5] (see also [24]) where a black hole in anti-de Sitter spacetime radiates into an attached Minkowski region. We show that the first principles computation of the fine-grained entropy using the gravitational path integral description receives large corrections from non-perturbative effects. The effects come from new saddles in the gravitational path integral — replica wormholes — that dominate over the standard Euclidean black hole saddle, and lead to a fine-grained entropy consistent with unitarity.

We will discuss the saddles explicitly only in some simple examples related to the in-

formation paradox for eternal black holes in two-dimensional Jackiw-Teitelboim (JT) gravity [25–27], reviewed below, but we can nonetheless compute the effect on the fine-grained entropy more generally. The same answer for the entropy was obtained holographically in [6, 15, 16]. Our goal is to provide a direct, bulk derivation without using holography.

To summarize our approach briefly, we will revisit the calculation of the von Neumann entropy of radiation outside a black hole in AdS glued to flat space, using the replica method. We introduce n copies of the original black hole, analytically continue to non-integer n, and compute the von Neumann entropy as $S_R = -\partial_n \text{Tr}\,(\rho_R)^n|_{n=1}$. Since the theory is coupled to gravity, we must do the gravitational path integral to calculate $\text{Tr}\,(\rho_R)^n$. Under our assumptions about the matter content, this path integral is dominated by a saddlepoint. There is one obvious saddle, in which the geometry is n copies of the original black hole; this saddle leads to the standard Hawking result for the von Neumann entropy, i.e., the entropy of quantum fields in a fixed curved spacetime, see figure 6(a).

There is, however, another class of saddles in which the different replicas are connected by a new geometry. These are the replica wormholes, see figure 6(b), 7. In the examples we consider, whenever the Hawking-like calculation leads to an entropy in tension with unitarity, the replica wormholes start to dominate the gravitational path integral, and resolve the tension.

Our use of the replica trick in a theory coupled to gravity closely parallels the derivation of the Ryu-Takayanagi formula and its generalizations [22, 28–30].

In the rest of the introduction we summarize the main idea in more detail.

Similar ideas are explored independently in a paper by Penington, Shenker, Stanford, and Yang [31].

1.1 The island rule for computing gravitational von Neumann entropies

We begin by reviewing the recent progress on the information paradox in AdS/CFT [4, 5].

The classic information paradox is difficult to study in AdS/CFT, because large black holes do not evaporate. Radiation bounces off the AdS boundary and falls back into the black hole. For this reason, until recently, most discussions of the information paradox in AdS/CFT have focused on exponentially small effects, such as the late-time behavior of boundary correlation functions [32–35].

In contrast, the discrepancy in the Page curve is a large, $O(1/G_N)$, effect. This classic version of the information paradox can be embedded into AdS/CFT by coupling AdS to an auxiliary system that absorbs the radiation, allowing the black hole to evaporate [4, 5] (see also [7, 36, 37]). This is illustrated in figure 1 in the case where the auxiliary system is half of Minkowski space, glued to the boundary of AdS. There is no gravity in the Minkowski region, where effectively $G_N \to 0$, but radiation into matter fields is allowed to pass through the interface.

In this setup, the Page curve of the black hole was calculated in [4, 5]. It is important to note that this calculation gives the Page curve of the black hole, not the radiation, which is where the paradox lies; we return to this momentarily. The entropy of the black hole is given by the generalized entropy of the quantum extremal surface (QES) [23], which is a quantum-corrected Ryu-Takayanagi (or Hubeny-Rangamani-Takayanagi) surface [20, 21].

According to the QES proposal, the von Neumann entropy of the black hole is

$$S_B = \text{ext}_Q \left[\frac{\text{Area}(Q)}{4G_N} + S_{\text{matter}}(B) \right] \tag{1.1}$$

where Q is the quantum extremal surface, and B is the region between Q and the AdS boundary. S_{matter} denotes the von Neumann entropy of the quantum field theory (including perturbative gravitons) calculated in the fixed background geometry. The extremization is over the choice of surface Q. If there is more than one extremum, then Q is the surface with minimal entropy. For dilaton gravity in AdS$_2$, Q is a point, and its 'area' means the value of the dilaton.

The black hole Page curve is the function $S_B(t)$, where t is the time on the AdS boundary where B is anchored. It depends on time because the radiation can cross into the auxiliary system. It behaves as expected: it grows at early times, then eventually falls back to zero [4, 5]. A crucial element of this analysis is that at late times, the dominant quantum extremal surface sits near the black hole horizon, as in figure 1.

This does not resolve the Hawking paradox, which involves the radiation entropy $S_{\text{matter}}(R)$, where R is a region outside the black hole containing the radiation that has come out. Clearly the problem is that neither R nor B includes the region I behind the horizon, called the island, see figure 1. The state of the quantum fields on $R \cup B$ is apparently not pure, and, apparently $S_R \neq S_B$. Only if we assume unitarity, or related holographic input such as entanglement wedge reconstruction [4], can we claim that the QES computes the entropy of the radiation. It does, however, tell us what to aim for in a unitary theory.

With this motivation, in [6], the evaporating black hole in Jackiw-Teitelboim (JT) gravity in AdS$_2$ was embedded into a holographic theory in one higher dimension. The AdS$_2$ black hole lives on a brane at the boundary of AdS$_3$, similar to a Randall-Sundrum model [38, 39], with JT gravity on the brane (see also [10] for an analogous construction on an AdS$_4$ boundary of AdS$_5$). In this setup, [6] derived the QES prescription *for the radiation* using AdS$_3$ holography. It was found that the von Neumann entropy of the radiation in region R, computed holographically in AdS$_3$, agrees with the black hole entropy in (1.1). This led to the conjecture that in a system coupled to gravity, the ordinary calculation of von Neumann entropy should be supplemented by the contribution from "islands" according to the following rule:

$$S(\rho_R) = \text{ext}_Q \left[\frac{\text{Area}(Q)}{4G_N} + S(\tilde{\rho}_{I \cup R}) \right] , \tag{1.2}$$

up to subleading corrections. Here ρ_R is the density matrix of the region R in the full theory coupled to quantum gravity, and $\tilde{\rho}_{I \cup R}$ is the density matrix of the state prepared via the semi-classical path integral on the Euclidean black hole saddle. This is equal to (1.1), since the quantum fields are pure on the full Cauchy slice $I \cup B \cup R$. Thus the tension with unitarity is resolved within three-dimensional holography.

In this paper we explain how the surprising island rule (1.2) follows from the standard rules for computing gravitational fine-grained entropy, without appealing to higher dimensional holography.
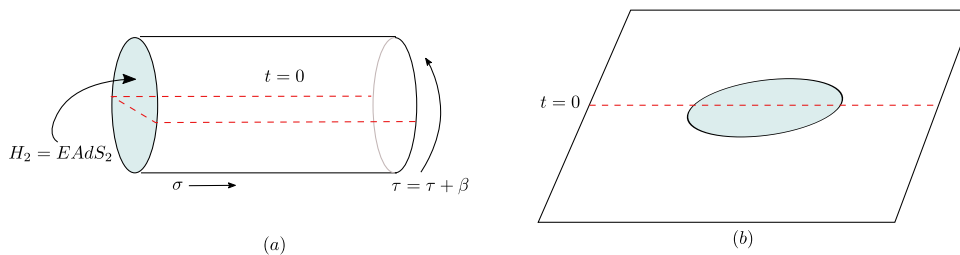
Figure 2. We prepare the combined thermofield double state of the black hole and radiation using a Euclidean path integral. These are two pictures for the combined geometry. In (b) we have represented the outside cylinder as the outside of the disk. By cutting along the red dotted line, we get our desired thermofield double initial state that we can then use for subsequent Lorentzian evolution (forwards or backwards in time) to get the diagram in figure 3.

1.2 Two dimensional eternal black holes and the information paradox

We consider an AdS$_2$ JT gravity theory coupled to a 2d CFT. This CFT also lives in non-gravitational Minkowski regions, and has transparent boundary conditions at the AdS boundary. The dilaton goes to infinity at the AdS$_2$ boundary so it is consistent to freeze gravity on the outside [5, 37]. We will assume that the matter CFT has a large central charge $c \gg 1$, but we will not assume that it is holographic, as all our calculations are done directly in the 2d theory. For example it could be c free bosons. Taking the central charge large is to suppress the quantum fluctuations of the (boundary) graviton relative to the matter sector.

This simple model of an AdS_2 black hole glued to flat space can be directly applied to certain four dimensional black holes. For example, for the near extremal magnetically charged black holes discussed in [40], at low temperatures we can approximate the dynamics as an AdS_2 region joined to a flat space region, and the light fields come from effectively two dimensional fields moving in the radial and time direction that connect the two regions.

We will consider a simple initial state which is the thermofield double state for the black hole plus radiation. This state is prepared by a simple Euclidean path integral, see figure 2. The resulting Lorentzian geometry is shown in figure 3.

Despite its simplicity, this setup exhibits Hawking's information paradox, and the corresponding puzzle with the Page curve [2, 3]. To reach a paradox, we collect Hawking radiation in region R in figure 3. As a function of time, R moves upward on both sides of the Penrose diagram, so this is not a symmetry. Indeed, the von Neumann entropy of the radiation as calculated by Hawking, $S_{\text{matter}}(R(t))$, grows linearly with time, see figure 5. The origin of this growth is the following. At $t = 0$ the radiation modes on the left are entangled with modes on the right. However, as time progresses some of these modes fall into the black holes, others are replaced by black hole modes, see figure 4.

If this growth were to continue forever, it would become larger than the Bekenstein-Hawking entropies of the two black holes, and this is a contradiction. See a related discussion of the critically illuminated black hole in flat spacetime in [41].

In a unitary theory, $S_R(t)$ should saturate at around the twice the Bekenstein-Hawking entropy of each black hole, see figure 5. This was confirmed using the island rule in [14].
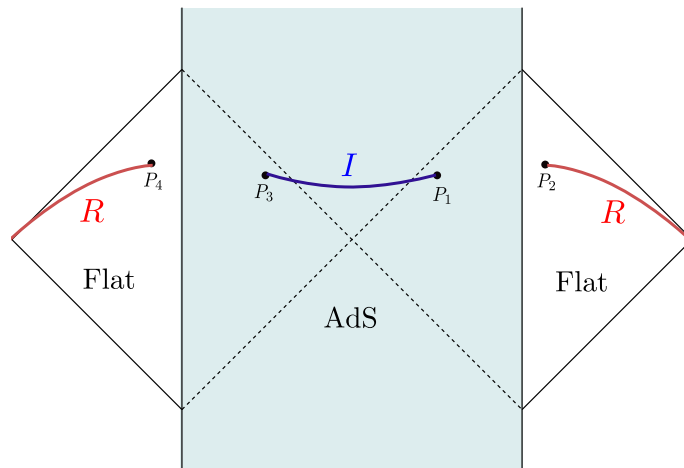
Figure 3. Eternal black hole in AdS_2, glued to Minkowski space on both sides. Hawking radiation is collected in region R, which has two disjoint components. Region I is the island. The shaded region is coupled to JT gravity.
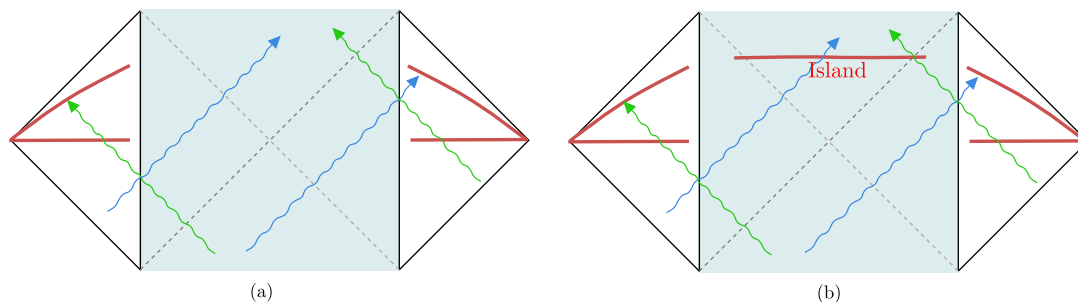

Figure 4. (a) Growing entropy for the radiation for an eternal black hole plus radiation in the thermofield double state. We draw two instants in time. The particles with the same color are entangled. They do not contribute to the entanglement of the radiation region (indicated in red) at $t = 0$ but they do contribute at a later value of t. (b) When the island is included the entanglement ceases to grow, because now both entangled modes mentioned above are included in $I \cup R$.

1.3 Replica wormholes to the rescue

To reproduce the unitary answer directly from a gravity calculation, we will use the replica method to compute the von Neumann entropy of region R. The saddles relevant to the unitary Page curve will ultimately be complex solutions of the gravitational equations. The idea is to do Euclidean computations and then analytically continue to Lorentzian signature.

Consider $n = 2$ replicas. The replica partition function $\mathrm{Tr}\,(\rho_R)^2$ is computed by a Euclidean path integral on two copies of the Euclidean system, with the matter sector sewed together along the cuts on region R. Since we are doing a gravitational path integral, we do not specify the geometry in the gravity region; we only fix the boundary conditions at the edge. Gravity then fills in the geometry dynamically, see figure 6.

We consider two different saddles with the correct boundary conditions. The first is the Hawking saddle, see figure 6(a). The corresponding von Neumann entropy is the usual
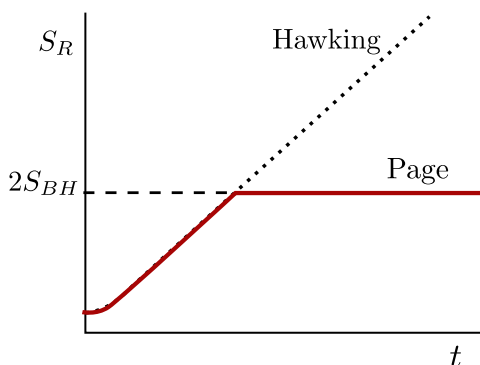
Figure 5. Page curve for the entropy of the radiation, for the model in figure 3. The dotted line is the growing result given by the Hawking computation, and the entropy calculated from the other saddle is dashed. The minimum of the two is the Page curve for this model.
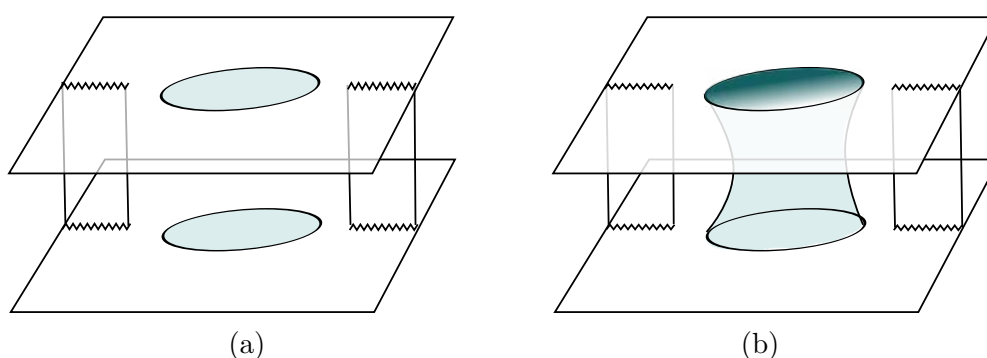

Figure 6. Two different saddlepoint contributions to the two-replica path integral in the presence of gravity in the shaded region. On the left the replicas are sewn together along the branch points, outside of the shaded region, as we would do in an ordinary quantum field theory calculation. These will give the standard QFT answer, as computed by Hawking, which can lead to a paradox. On the right we have a saddle where gravity dynamically glues together the shaded regions. This is the replica wormhole. In the examples considered in this paper, this saddle dominates in the relevant kinematics, leading to a Page curve consistent with unitarity.

answer, $S_{\text{matter}}(R(t))$, which grows linearly forever. The second is the replica wormhole, which, as we will show, reproduces the entropy of the island rule, see figure 6(b). A replica wormhole with higher n is illustrated in figure 7.

Replica wormholes have higher topology, so they are suppressed by factors of e^{-S_0} where S_0 is the genus-counting parameter of JT gravity. At late times, the contribution of the Hawking saddle is heavily suppressed by the kinematics, and this is what makes it possible for the replica wormhole to take over despite the topological suppression. Indeed, the nth wormhole, see figure 7, gives a partition function $Z_n \propto e^{S_0(2-n)}$ which leads to a $2S_0$ contribution to the entropy.

The wormhole topology has a saddle point at finite n. (We will not show this in general, but confirm it explicitly in certain limits; see below for details.) The equations that control this saddle point can be analytically continued to non-integer n, and used to define the replica limit $n \to 1$. To analyze this limit it is most convenient to assume replica
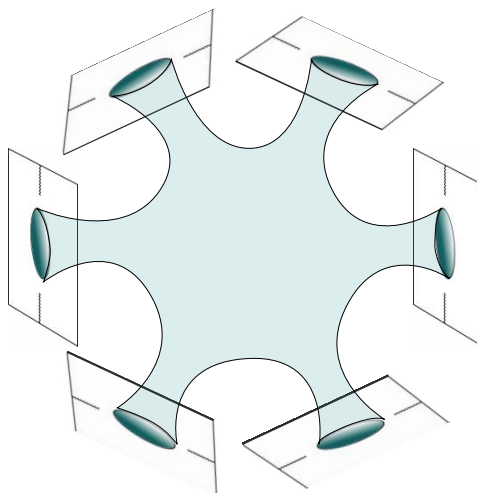
Figure 7. Topology of a replica wormhole with $n = 6$. The sheets are also glued together cyclically along the cuts in the matter region.

symmetry and go to a quotient space which has a simpler topology but contains conical singularities and insertions of twist operators for the matter fields, see figure 9. In the limit $n \to 1$ both of these effects become very small and represent a small perturbation for the geometry, but they give a contribution to the entropy of precisely the same form as the gravitational generalized entropy for regions in the $n = 1$ solution. The boundaries of the regions are specified by the locations of the twist operators. The replica wormholes give rise to the island contributions to the entropy.

The physical picture that descends from accounting for these higher topology saddles in the entropy calculation is as follows. In the initial stages of the black hole evaporation, the quantum state of the Hawking radiation is accurately described by quantum field theory on a fixed background as originally studied by Hawking. This is accurate up to the Page time, defined to be the time when the semi-classical von Neumann entropy of the Hawking radiation becomes equal to the the coarse-grained entropy of the black hole. At later times, a non-perturbative effect in the gravitational path integral results in an $O(1)$ deviation of the evolution of the entropy of the Hawking radiation form the semi-classical result. This is due to an exchange of dominance between the trivial topology saddle and the wormhole saddle in the Renyi entropy calculation. This new saddle suggests that we should think of the inside of the black hole as a subsystem of the outgoing Hawking radiation. Namely, in the $n \to 1$ limit of the the replica trick, most of the black hole interior is included, together with the radiation, in the computation of the entropy. This has the effect that entanglement across the event horizon of the Hawking pairs no longer contributes to the von Neumann entropy of the outgoing part, while at the same time maintaining the necessary entanglement to ensure semi-classical physics at the horizon.

This paper is organized as follows.

In section 2.1 we review and slightly clarify the gravitational derivation of the quantum extremal surface presciption from the replica trick in a general theory [22, 28–30]. The slight

improvement is that we show that the off shell action near $n \sim 1$ becomes the generalized entropy, so that the extremality condition follows directly from the extremization of the action. In section 2.2 we discuss some general aspects of replica manifolds for the case of JT gravity plus a CFT.

In section 3 we discuss the computation of the entropy for an interval that contains the degrees of freedom living at the AdS boundary. In this case the quantum extremal surface is slightly outside the horizon. We set up the discussion of the Renyi entropy computations for this case. We reduce the problem to an integro-differential equation for a single function $\theta(\tau)$ that relates the physical time τ to the AdS time θ. We solve this equation for $n \to 1$ recovering the quantum extremal surface result. We also solve the problem for relatively high temperatures but for any n.

In section 4 we discuss the special case of the zero temperature limit, and we comment on some features of the island in that case.

In section 5 we discuss aspects of the two intervals case, which is the one most relevant for the information problem for the eternal black hole.

In section 6 we make the connection to entanglement wedge reconstruction of the black hole interior.

We end in section 7 with conclusions and discussion.

2 The replica trick for the von Neumann entropy

The replica trick for computing the von Neumann entropy is based on the observation that the computation of $Tr[\rho^n]$ can be viewed as an observable in n copies of the original system [42]. In particular, for a quantum field theory the von Neumann entropy of some region can be computed by considering n copies of the original theory and choosing boundary conditions that connect the various copies inside the interval in a cyclic way, see e.g. [43] for a review. This can be viewed as the insertion of a "twist operator" in the quantum field theory containing n copies of the original system. This unnormalized correlator of twist operators can also be viewed as the partition function of the theory on a topologically non-trivial manifold, $Z_n = Z[\widetilde{\mathcal{M}}_n] = \langle \mathcal{T}_1 \cdots \mathcal{T}_k \rangle$. Then the entropy can be computed by analytically continuing in n and setting

$$S = -\partial_n \left(\frac{\log Z_n}{n} \right) \bigg|_{n=1} \qquad (2.1)$$

We will now review the argument for how this is computed in theories of gravity. Then we will consider the specific case of the JT gravity theory.

2.1 The replicated action for $n \sim 1$ becomes the generalized entropy

In this section we review the ideas in [22, 28–30] for proving the holographic formula for the fine-grained entropy, or von Neumann entropy. We clarify why we get the generalized entropy when we evaluate the off shell gravity action near the $n = 1$ solution.

The replica trick involves a manifold $\widetilde{\mathcal{M}}_n$ which computes the nth Renyi entropy. The geometry of this manifold is completely fixed in the non-gravitational region, where

we define the regions whose entropies we are computing.[1] In the gravitational region we can consider any manifold, with any topology, which obeys the appropriate boundary conditions. The full action for the system is a sum of the gravitational action and the partition function for the quantum fields on the geometry $\widetilde{\mathcal{M}}_n$,

$$\frac{\log Z_n}{n} = -\frac{1}{n}I_{\mathrm{grav}}[\widetilde{\mathcal{M}}_n] + \frac{1}{n}\log Z_{\mathrm{mat}}[\widetilde{\mathcal{M}}_n]. \tag{2.2}$$

This is an effective action for the geometry and we will look for a classical solution of this combined action. In other words, the integral over geometries is evaluated as a saddle point. So the metric is classical, but the equations contain the quantum expectation value of the matter stress tensor on that geometry. Under the assumption of replica symmetry, we can instead consider another manifold $\mathcal{M}_n = \widetilde{\mathcal{M}}_n/Z_n$. This manifold can be viewed as one where n identical copies of the field theory are living. We have twist operators $\mathcal{T}_n$ at the endpoints of the intervals in the non-gravitational region. In the gravitational region we also have twist operators $\mathcal{T}_n$ at the fixed points of the Z_n action, where the manifold $\mathcal{M}_n$ has conical singularities with opening angle $2\pi/n$. Of course, at these points the covering manifold $\widetilde{\mathcal{M}}_n$ is smooth. It is convenient to translate the problem in (2.2) to a problem involving the manifold $\mathcal{M}_n$. We have n copies of the matter theory propagating on this manifold. In the gravitational region we can enforce the proper conical singularities in $\mathcal{M}_n$ by adding codimension-two "cosmic branes" of tension

$$4G_N T_n = 1 - \frac{1}{n}. \tag{2.3}$$

At these cosmic branes we also insert twist operators $\mathcal{T}_n$ for the n copies of the matter theory. In two dimensions these "cosmic branes" are simply points, while in four dimensions they are "cosmic strings." The positions of these cosmic branes are fixed by solving the Einstein equations. We then replace the gravitational part of the action in (2.2) by

$$\frac{1}{n}I_{\mathrm{grav}}[\widetilde{\mathcal{M}}_n] = I_{\mathrm{grav}}[\mathcal{M}_n] + T_n \int_{\Sigma_{d-2}} \sqrt{g}. \tag{2.4}$$

As opposed to [28], here we add the action of these cosmic branes explicity and we also integrate the Einstein term through the singularity, which includes a δ function for the curvature. These two extra terms cancel out so that we get the same final answer as in [28] where no contribution from the singularity was included. We will see that the present prescription is more convenient.[2]

In the part of the manifold where the metric is dynamical the position of these cosmic branes is fixed by the Einstein equations. Also, the reparametrization symmetry implies we cannot fix these points from the outside.

---

[1]If we only had the AdS theory, without an outside region, then the non-gravitational part should be viewed just as the boundary of AdS.

[2]In theories with higher derivatives we would need to add extra terms in the action of the cosmic brane so that they just produce a conical singularity. These presumably lead to an off shell action of the form considered in [44] but we did not check this.

When $n = 1$ we have the manifold $\mathcal{M}_1 = \widetilde{\mathcal{M}}_1$, which is the original solution to the problem. It is a solution of the action I_1^{tot}. In order to find the manifold $\mathcal{M}_n$ for $n \sim 1$ we need to add the cosmic branes. Then the action is

$$\left(\frac{I^{\text{tot}}}{n}\right)_{n \to 1} = I_1 + \delta\left(\frac{I}{n}\right) \tag{2.5}$$

where δI contains extra terms that arise from two effects, both of which are of order $n-1$. The first comes from the tension of the cosmic brane (the second term in (2.4). The second comes from the insertion of the twist fields at the position of this cosmic brane. To evaluate the action perturbatively, we start from the solution $\mathcal{M}_1$, we add the cosmic brane and twist fields, and we also consider a small deformation of the geometry away from $\mathcal{M}_1$, where all these effects are of order $n-1$. Because the $\mathcal{M}_1$ geometry is a solution of the original action I_1 in (2.5), any small deformation of the geometry drops out of the action. For the extra term $\delta(I/n)$ in (2.5), we can consider the cosmic brane action and twist fields as living on the old geometry $\mathcal{M}_1$ since these extra terms are already of order $n-1$.

Then we conclude that the δI term is simply proportional to the generalized entropy

$$\delta\left(\frac{\log Z}{n}\right) = -\delta\left(\frac{I}{n}\right) = (1-n)S_{\text{gen}}(w_i) = (1-n)\left[\frac{\text{Area}}{4G_N} + S_{\text{matter}}\right], \qquad n \sim 1 \tag{2.6}$$

where we emphasized that it depends on the positions of the cosmic branes. We should emphasize that (2.6) is the full off-shell action that we need to extremize to find the classical solution of I_n for $n \sim 1$. In this way, we obtain the quantum extremal surface prescription of [23], and also [20, 21]. Moreover, if we think of the cosmic strings as dynamical objects, then we can pair create them so as to form islands. This pair creation is possible in the gravity region where the tension is finite. In the region without gravity their tension is effectively infinite.

2.2 The two dimensional JT gravity theory plus a CFT

In this section we specify in more detail the theory under consideration. We have the Jackiw-Teitelboim gravity theory describing a nearly AdS_2 spacetime coupled to a matter theory that is a CFT. In addition, we have the same CFT living in an exterior flat and rigid geometry with no gravity. Since the interior and the exterior involve the same CFT we can impose transparent boundary conditions at the boundary, see figure 8. In other words, we have the action

$$\log Z^{\text{tot}} = \frac{S_0}{4\pi}\left[\int_{\Sigma_2} R + \int_{\partial\Sigma_2} 2K\right] + \int_{\Sigma_2}\frac{\phi}{4\pi}(R+2) + \frac{\phi_b}{4\pi}\int_{\partial\Sigma_2} 2K + \log Z_{CFT}[g] \tag{2.7}$$

where the CFT action is defined over a geometry which is rigid in the exterior region and is dynamical in the interior region. We are setting $4G_N = 1$ so that the area terms in the entropies will be just given by the value of ϕ, $\frac{\text{Area}}{4G_N} = S_0 + \phi$.

In this theory, we want to consider the replica manifolds described above, see figure 7. Because we consider replica symmetric solutions, it is convenient to quotient by Z_n and discuss a single manifold with n copies of the matter theory on it. In other words, we go
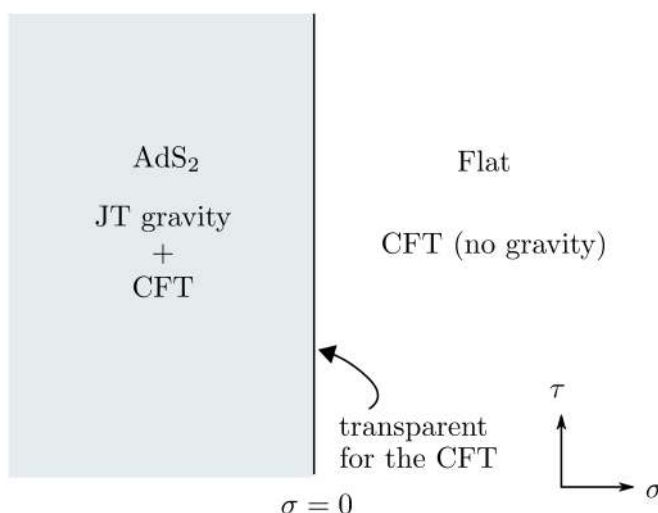
Figure 8. We consider nearly-AdS_2 gravity with a matter CFT. The same CFT lives in an exterior flat space with no gravity. We have transparent boundary conditions for the CFT.

from the action (2.7) on $\widetilde{\mathcal{M}}_n$ to a problem on $\mathcal{M}_n = \widetilde{\mathcal{M}}_n/Z_n$. We find that this simplifies a bit the description of the manifold, see figure 9. Namely, the manifold $\mathcal{M}_n$ can be viewed as a disk with conical singularities and with twist operators for the matter theory inserted at these singularities. These are the cosmic branes discussed in section 2.1. The final gravitational action is as in (2.7) but with an additional factor of n and extra terms that produce the conical singularities

$$-\frac{1}{n}I_{\text{grav}} = \frac{S_0}{4\pi}\left[\int_{\Sigma_2} R + \int_{\partial\Sigma_2} 2K\right] + \int_{\Sigma_2} \frac{\phi}{4\pi}(R+2)$$
$$+ \frac{\phi_b}{4\pi}\int_{\partial\Sigma_2} 2K - \left(1-\frac{1}{n}\right)\sum_i [S_0 + \phi(w_i)] \tag{2.8}$$

where w_i are the positions of the conical singularities, or cosmic branes (which are just instantons or -1 branes). We can consider (2.8) as a new gravity theory and add n copies of the CFT. In addition, we put twist fields at the positions w_i of the cosmic branes. It might look like we are breaking reparametrization invariance when when add these terms. Reparametrization symmetry is restored because w_i are dynamical variables which can be anywhere on the manifold and will be fixed by the equations of motion.

We treat the CFT as a quantum theory and evaluate its partition function. Then we solve the classical equations for the metric and dilaton inserting the quantum expectation value of the stress tensor. This approximation is particularly appropriate when the central charge is large $c \gg 1$. So we imagine that we are in that regime for the simple euclidean solutions we discuss here. The approximation can also be justified in other regimes where the entanglement entropy of matter is large for kinematical reasons. However, this description is *not* correct when we need to include the quantum aspects of gravity. That computation should be done in the original manifold and the fact that the fluctuations can break the replica symmetry is important.
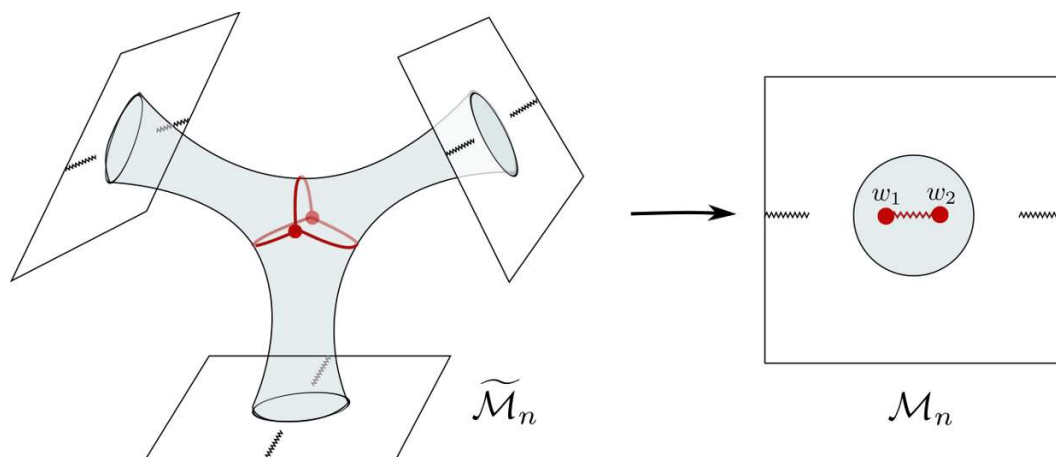
Figure 9. Here we display the replica manifold, $\widetilde{\mathcal{M}}_3$, and also the manifold $\mathcal{M}_3 = \widetilde{\mathcal{M}}_3/Z_3$ which has the topology of the disk with conical singularities at two points w_1 and w_2 which corresponds to the fixed points of the Z_3 action on $\widetilde{\mathcal{M}}_3$. We parametrize this disk in terms of the holomorphic coordinate w. The exterior regions of $\widetilde{\mathcal{M}}_n$ are also glued together cyclically along the cuts.

We can define an interior complex coordinate w where the metric for the manifold $\mathcal{M}_n$ in the gravitational region is

$$ds^2 = e^{2\rho}dwd\bar{w}, \qquad \text{with} \ \ |w| \le 1. \tag{2.9}$$

The boundary of AdS_2 is at $|w| = 1$, or $w = e^{i\theta}$. (2.9) is a constant curvature metric on the disk $|w| \le 1$ with conical singularities at certain values w_i with opening angle $2\pi/n$. This type of metric is enforced by the dilaton equation of motion in (2.8)

$$-4\partial_w\partial_{\bar{w}}\rho + e^{2\rho} = 2\pi\left(1 - \frac{1}{n}\right)\sum_i \delta^2(w - w_i) \tag{2.10}$$

On this space we have n copies of the CFT and we have twist fields inserted at the conical singularities. Notice that once we impose this equation, the contributions in (2.8) from the delta functions in the curvature cancel against the explicit cosmic brane action terms, as we anticipated in section 2.1.

This metric should be joined to the flat space outside. We consider a finite temperature configuration where $\tau \sim \tau + 2\pi$. For general temperatures, all we need to do is to rescale $\phi_r \to 2\pi\phi_r/\beta$. In other words, the only dimensionful scale is ϕ_r, so the only dependence on the temperature for dimensionless quantities is through ϕ_r/β. We define the coordinate $v = e^y$. So the physical half cylinder $\sigma \ge 0$ corresponds to $|v| \ge 1$. At the boundary we have that $w = e^{i\theta(\tau)}$, $v = e^{i\tau}$. Unfortunately, we cannot extend this to a holomorphic map in the interior of the disk. However, we can find another coordinate z such that there are holomorphic maps from $|w| \le 1$ and $|v| \ge 1$ to the coordinate z, see figure 10.

In other words, it is possible to find two functions G and F such that

$$z = G(w), \qquad \text{for} \ \ |w| \le 1 \tag{2.11}$$

$$z = F(v), \qquad \text{for} \ \ |v| \ge 1 \tag{2.12}$$

$$G(e^{i\theta(\tau)}) = F(e^{i\tau}), \qquad \text{for} \ \ |w| = |v| = 1. \tag{2.13}$$
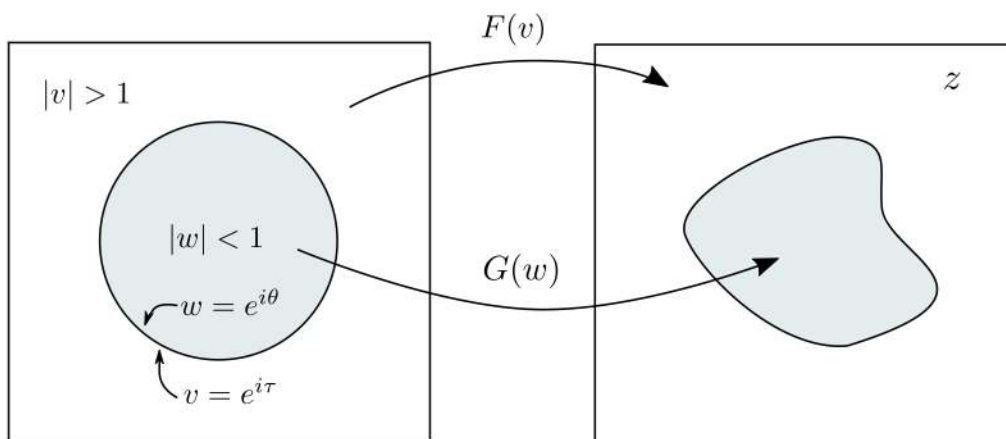
Figure 10. The conformal welding problem. We are given two disks, one parametrized by $|w| \leq 1$ and another by $|v| \geq 1$ with their boundaries glued in terms of a given function $\theta(\tau)$ where $w = e^{i\theta}$ and $v = e^{i\tau}$. Then we need to find holomorphic maps of each disk to a region of the complex z plane so that they are compatible at the boundary. The functions F and G are only required to be holomorphic inside their respective disks.

The functions F and G are holomorphic in their respective domains (they do not have to be holomorphic at the boundary). The problem of finding F and G given $\theta(\tau)$ is called the "conformal welding problem," see [45] for a nice discussion.[3] F and G end up depending non-locally on $\theta(\tau)$ and they map the inside and outside disks to the inside and outside of some irregular region in the complex plane, see figure 10. In our problem, $\theta(\tau)$ arises as the reparametrization mode, or "boundary graviton" of the nearly-AdS_2 gravity theory [37, 46, 47].

When $n = 1$, we have a trivial stress tensor in the z plane. We then insert the twist operators in the outside region, and also in the inside region. We are free to insert as many conical singularities and twist fields in the inside as we want. This amounts to considering various numbers of islands in the gravity region. We will only discuss cases with one or two inside insertions in the subsequent sections. This gives us a non-trivial stress tensor $T_{zz}(z)$ and $T_{\bar{z}\bar{z}}(\bar{z})$. We can then compute the physical stress tensor that will appear in the equation of motion using the conformal anomaly,

$$T_{yy} = \left(\frac{dF(e^{iy})}{dy}\right)^2 T_{zz} - \frac{c}{24\pi}\{F(e^{iy}), y\} \tag{2.14}$$

and a similar expression for $T_{\bar{y}\bar{y}}$. The expression for the physical stress tensor in the w plane involves the function G and also a conformal anomaly contribution from ρ in the metric (2.9).

Let us now turn to the problem of writing the equations of motion for the boundary reparametrization mode. Naively we are tempted to write the action just as $\{e^{i\theta}, \tau\}$. This

---

[3]We thank L. Iliesiu and Z. Yang for discussions on this problem, and A. Lupsasca for pointing out the connection to [45].

would be correct if there were no conical singularities in the interior. However, the presence of those conical singularities implies that the metric (2.9) has small deviations compared to the metric of a standard hyperbolic disk

$$ds^2 = e^{2\rho}dwd\bar{w}\,, \qquad e^{2\rho} = \frac{4}{(1-|w|^2)^2}e^{2\delta\rho} \tag{2.15}$$

where $\delta\rho$ goes as

$$\delta\rho \sim -\frac{(1-|w|)^2}{3}U(\theta)\,, \qquad \text{as} \quad |w| \to 1\,. \tag{2.16}$$

The function U depends on the positions of the conical singularities and therefore also on the moduli of the Riemann surface. This then implies that the Schwarzian term, and the full equation of motion can now be written as

$$\frac{\phi_r}{2\pi}\frac{d}{d\tau}\left[\{e^{i\theta},\tau\} + U(\theta)\theta'^2\right] = i(T_{yy} - T_{\bar{y}\bar{y}}) = T_{\tau\sigma}\,. \tag{2.17}$$

The term in brackets is proportional to the energy. This equation relates the change in energy to the energy flux from the flat space region. Here the flux of energy on the right hand side is that of one copy, or the flux of the n copies divided by n. The action can be derived from the extrinsic curvature term in the same way that was discussed in [37, 46, 47], see appendix A, where we also discuss the explicit derivation of the equation of motion (2.17).

There are also equations that result from varying the moduli of the Riemann surface, or the positions of the conical singularities. They have the form

$$-\left(1 - \frac{1}{n}\right)\partial_w\phi(w_i) + \partial_{w_i}\left(\frac{\log Z_n^{\text{mat}}}{n}\right) = 0\,, \tag{2.18}$$

where we used that the w_i dependence of the gravitational part of the action comes only from the last term in (2.8).

In the $n \to 1$ limit we can replace the $n = 1$ value for the dilaton in (2.18). Similarly the value of $\log Z_n^{\text{mat}}/n$ near $n = 1$ involves the matter entropy. Therefore (2.18) reduces to the condition on the extremization of the generalized entropy, as we discussed in general above.

For general n, we need to compute the dilaton by solving its equations of motion in order to write (2.18). This can be done using the expression for the stress tensor in the interior of the disk. We have not attempted to simplify it further. However, we should note that for the particular case of one interval, discussed in section 3, there is only one point and there are no moduli for the Riemann surface. Therefore this equation is redundant and in fact, it is contained in (2.17) as will be discussed in section 3.

Next we apply this general discussion to the calculation of the entropy of various subregions of the flat space CFT. The goal is to understand how configurations of the gravity region contribute to the entropy of those CFT regions.

3 Single interval at finite temperature

We begin with the simple case of a single interval that contains one of the AdS$_2$ boundaries, as shown in figure 11(a). This is the interval $B \equiv [0, b]$.
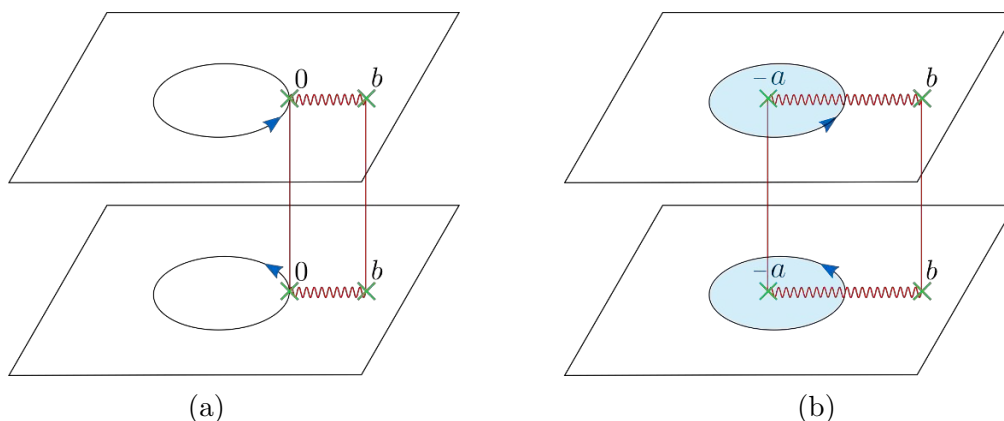
Figure 11. (a) We have a flat space field theory on the exterior of the disk. The disk is hollow in this picture, and will be filled in with gravitational configurations subject to the boundary conditions on the unit circle. This boundary is connected into a single long circle n times longer than the original one. This is indicated by the blue arrow which tells you how to go around the cut. (b) The disk is filled in with a gravitational configuration with the topology of a disk which ends on the elongated unit circle. This configuration can be represented by adding a branch point inside. Note that the local geometry at the branch point "$-a$" is completely smooth.
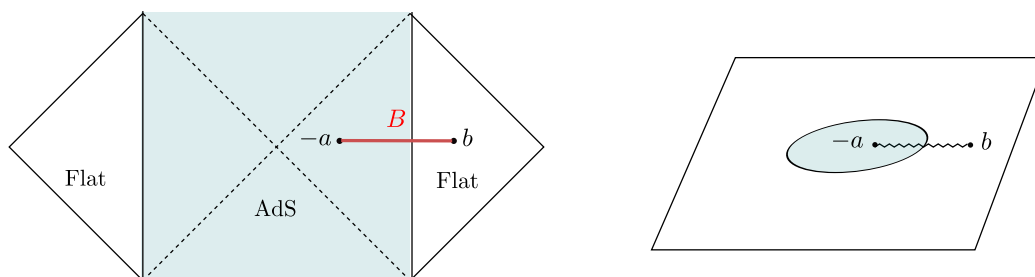

Figure 12. The single interval configuration in Lorentzian signature (left) and in Euclidean signature (right).

To compute the entropy of this region we must consider the Euclidean path integral that evaluates the trace of powers of the density matrix $\mathrm{Tr}[\rho_B^n]$. This is given by the path integral on n copies of the theory identified across the region B, as shown in figure 11. The crucial point is that the presence of the branch point on the unit circle, which is where the asymptotic AdS boundary lives, elongates this circle by a factor of n. The Euclidean gravity configurations we must consider are all smooth manifolds with a single boundary that is identified with this elongated AdS boundary.

The simplest configuration to consider will be that with the topology of a disk. All other higher genus manifolds will be subleading since each extra handle will come with a cost of e^{-S_0}. Filling out the gravity region has the effect of extending the identification across the different sheets into the gravity region, which ends on some point "$-a$" in figure 11. The location of the point "$-a$" will be dynamically determined by the saddle point of the path integral.

We will now construct replica wormholes explicitly for a single interval in the eternal black hole in AdS$_2$. The Lorentzian and Euclidean geometries are shown in figure 12. We will first review of the result of the QES calculation [14], then proceed to derive it from replica wormholes.

3.1 Geometry of the black hole

The metric of eternal black hole, glued to flat space on both sides, is

$$ds_{\text{in}}^2 = \frac{4\pi^2}{\beta^2} \frac{dy d\bar{y}}{\sinh^2 \frac{\pi}{\beta}(y + \bar{y})}, \qquad ds_{\text{out}}^2 = \frac{1}{\epsilon^2} dy d\bar{y}, \tag{3.1}$$

$$y = \sigma + i\tau, \qquad \bar{y} = \sigma - i\tau, \qquad \tau = \tau + \beta . \tag{3.2}$$

The subscript 'in' refers to the gravity zone, and 'out' refers to the matter zone.[4] The interface is along the circle $\sigma = -\epsilon$. Lorenztian time t is $\tau = -it$. The welding maps of figure 10 are trivial and we have

$$z = v = w = e^{2\pi y/\beta}, \qquad y = \frac{\beta}{2\pi} \log w . \tag{3.3}$$

The Euclidean solution is therefore the w-plane with gravity inside the unit disk, $|w| < 1 - \frac{2\pi\epsilon}{\beta}$. The metric is

$$ds_{\text{in}}^2 = \frac{4 dw d\bar{w}}{(1 - |w|^2)^2}, \qquad ds_{\text{out}}^2 = \frac{\beta^2}{4\pi^2 \epsilon^2} \frac{dw d\bar{w}}{|w|^2} . \tag{3.4}$$

The dilaton, which is defined only on the inside region, is rotationally invariant on the w-plane,

$$\phi = \frac{2\pi\phi_r}{\beta} \frac{1 + |w|^2}{1 - |w|^2} = -\frac{2\pi\phi_r}{\beta} \frac{1}{\tanh \frac{2\pi\sigma}{\beta}} . \tag{3.5}$$

with $\phi = \phi_r/\epsilon$ at the boundary. In what follows, we will usually set $\epsilon = 0$, and rescale the exterior coordinate by ϵ so that $ds_{\text{out}}^2 = dy d\bar{y}$.

3.2 Quantum extremal surface

We now review the computation of the entropy of the region $B = [0, b]$ which includes the AdS_2 boundary, see figure 11. In gravity this will involve an interval $[-a, b]$, with $a, b > 0$, see figure 12.

The generalized entropy of the region $[-a, b]$ is

$$S_{\text{gen}} = S_0 + \phi(-a) + S_{\text{CFT}}([-a, b]) . \tag{3.6}$$

---

[4]The Poincare coordinates are $x = \tanh \frac{\pi y}{\beta}$, $ds_{\text{in}}^2 = 4 dx d\bar{x}/(x + \bar{x})^2$. The Schwarzschild coordinates are

$$y = \frac{\beta}{2\pi} \log \frac{r}{\sqrt{r(r + 4\pi/\beta)}} + i\tau, \qquad ds_{\text{in}}^2 = r\left(r + \frac{4\pi}{\beta}\right) d\tau^2 + \frac{dr^2}{r(r + \frac{4\pi}{\beta})} .$$

The entanglement entropy of a CFT on the interval $[w_1, w_2]$ in the metric $ds^2 = \Omega^{-2} dw d\bar{w}$ is

$$S_{\text{CFT}}(w_1, w_2) = \frac{c}{6} \log \left(\frac{|w_1 - w_2|^2}{\epsilon_{1,UV} \epsilon_{2,UV} \Omega(w_1, \bar{w}_1) \Omega(w_2, \bar{w}_2)} \right) . \tag{3.7}$$

Using the map $w = e^{2\pi y/\beta}$ and the conformal factors in (3.4) this becomes

$$S_{\text{CFT}}([-a, b]) = \frac{c}{6} \log \left(\frac{2\beta \sinh^2 \left(\frac{\pi}{\beta}(a + b) \right)}{\epsilon_{a,UV} \epsilon_{b,UV} \pi \sinh \left(\frac{2\pi a}{\beta} \right)} \right) \tag{3.8}$$

Then, using the dilaton in (3.5), (3.6) becomes

$$S_{\text{gen}}([-a, b]) = S_0 + \frac{2\pi \phi_r}{\beta} \frac{1}{\tanh \left(\frac{2\pi a}{\beta} \right)} + \frac{c}{6} \log \left(\frac{2\beta \sinh^2 \left(\frac{\pi}{\beta}(a + b) \right)}{\pi \epsilon \sinh \left(\frac{2\pi a}{\beta} \right)} \right) . \tag{3.9}$$

The UV divergence $\epsilon_{a,UV}$ was absorbed into S_0 and we dropped the outside one at point b. The quantum extremal surface is defined by extremizing S_{gen} over a

$$\partial_a S_{\text{gen}} = 0 \quad \rightarrow \quad \sinh \left(\frac{2\pi a}{\beta} \right) = \frac{12\pi \phi_r}{\beta c} \frac{\sinh \left(\frac{\pi}{\beta}(b + a) \right)}{\sinh \left(\frac{\pi}{\beta}(a - b) \right)} \tag{3.10}$$

This is a cubic equation for $e^{2\pi a/\beta}$. For $b \gtrsim \frac{\beta}{2\pi}$ and $\phi_r/(\beta c) \gtrsim 1$, the solution is

$$a \approx b + \frac{\beta}{2\pi} \log \left(\frac{24\pi \phi_r}{\beta c} \right) , \quad \text{or} \quad e^{-\frac{2\pi a}{\beta}} \approx \frac{\beta c}{24\pi \phi_r} e^{-\frac{2\pi b}{\beta}} \tag{3.11}$$

Since we've restricted to one side of the black hole in this calculation, the configuration is invariant under translations in the Schwarzschild t direction. Therefore the general extremal surface at $t \neq 0$ is related by a time translation; for an interval that starts at t_b and $\sigma_b = b$, the other endpoint is at $t_a = t_b$ and $\sigma_a = -a$, with a as in (3.10).

3.3 Setting up the replica geometries

We will do the replica calculation in Euclidean signature, with a, b real. We set $\beta = 2\pi$, and reintroduce it later by dimensional analysis.

The replica wormhole that we seek is an n-fold cover of the Euclidean black hole, branched at the points a and b, see figure (12). This manifold will have a nontrivial gluing at the unit circle (unlike the black hole itself), so it is more convenient to introduce different coordinates on the inside and outside. We use w, with $|w| < 1$, for the inside and $v = e^y$, with $|v| > 1$ for the outside. The gluing function is $\theta(\tau)$, with $w = e^{i\theta}$, $v = e^{i\tau}$, as in (2.11). We write the branch points as

$$w = A = e^{-a}, \quad v = B = e^b . \tag{3.12}$$

The Schwarzian equation is simplest in a different coordinate,

$$\widetilde{w} = \left(\frac{w - A}{1 - Aw} \right)^{1/n} . \tag{3.13}$$

This coordinate uniformizes n copies of the unit disk, so here we have the standard hyperbolic metric,

$$ds_{\text{in}}^2 = \frac{4|d\tilde{w}|^2}{(1 - |\tilde{w}|^2)^2} \ . \tag{3.14}$$

Defining $\tilde{w} = e^{i\tilde{\theta}}$ at the boundary, the Schwarzian equation is

$$\frac{\phi_r}{2\pi} \partial_\tau \{e^{i\tilde{\theta}}, \tau\} = i(T_{yy}(i\tau) - T_{\bar{y}\bar{y}}(-i\tau)) \ . \tag{3.15}$$

We can now return to the w-disk using the Schwarzian composition identity

$$\{e^{i\tilde{\theta}}, \tau\} = \{e^{i\theta}, \tau\} + \frac{1}{2}\left(1 - \frac{1}{n^2}\right) R(\theta) \,, \tag{3.16}$$

with

$$R(\theta) = -\frac{(1 - A^2)^2 (\partial_\tau \theta)^2}{|1 - Ae^{i\theta}|^4} \ . \tag{3.17}$$

This puts the equation of motion (3.15) into exactly the form of equation (2.17), which we have just derived by a slightly different route. In appendix A we show that they are equivalent.

The stress tensor appearing on the right-hand side of (3.15) is obtained through the conformal welding. That is, we define the z coordinate by the map G on the inside and F on the outside as in (2.11). These maps each have an ambiguity under $\mathrm{SL}(2, \mathbb{C})$ transformations of z, which we may use to map the twist operator at $w = A$ to $z = 0$, and the twist operator at $v = B$ to $z = \infty$. We further discuss the symmetries of the conformal welding problem in appendix B.

The z-coordinate covers the full plane holomorphically. It has twist points at the origin and at infinity, which can be removed by the standard mapping, $\tilde{z} = z^{1/n}$. On the $\tilde{z}$ plane, the stress tensor vanishes, so on the z-plane,

$$T_{zz}(z) = -\frac{c}{24\pi}\{z^{1/n}, z\} = -\frac{c}{48\pi}\left(1 - \frac{1}{n^2}\right)\frac{1}{z^2} \ . \tag{3.18}$$

Finally the stress tensor T_{yy} comes from inverting the conformal welding map to return to the v-plane, and using $v = e^y$:

$$T_{yy}(y) = e^{2y}\left[F'(v)^2 T_{zz} - \frac{c}{24\pi}\{F, v\}\right] - \frac{1}{2} \ . \tag{3.19}$$

Putting it all together, the equation of motion (3.15) is

$$\frac{24\pi\phi_r}{c\beta}\partial_\tau\left[\{e^{i\theta(\tau)}, \tau\} + \frac{1}{2}\left(1 - \frac{1}{n^2}\right)R(\theta(\tau))\right]$$
$$= ie^{2i\tau}\left[-\frac{1}{2}\left(1 - \frac{1}{n^2}\right)\frac{F'(e^{i\tau})^2}{F(e^{i\tau})^2} - \{F, e^{i\tau}\}\right] + \text{c.c.} \tag{3.20}$$

This equation originated on the smooth replica manifold $\widetilde{\mathcal{M}}_n$, but has now been written entirely on the quotient manifold $\mathcal{M}_n = \widetilde{\mathcal{M}}_n/\mathbf{Z}_n$. We have restored the nontrivial temperature dependence.[5] In particular, note that $\theta(\tau + 2\pi) = \theta + 2\pi$. The $\tau \to -\tau$ symmetry of the insertions allows us to choose a function $\theta(\tau) = -\theta(-\tau)$ which will automatically obey $\theta(0) = 0$, $\theta''(0) = 0$. In addition, we should then impose $\theta(\pi) = \pi$ and $\theta''(\pi) = 0$. The problem now is such that n appears as a continuous parameter and there is no difficulty in analytically continuing in n.

This is our final answer for the equation of motion at finite n. It is quite complicated, because the welding map F depends implicitly on the gluing function $\theta(\tau)$. We will solve it in two limits: $\beta \to 0$ at any n, and $n \to 1$ at any β.

3.4 Replica solution as $n \to 1$

We will now show that the equation of motion (3.20) reproduces the equation for the quantum extremal surface.

We start with the solution for $n = 1$. In this case the welding problem is trivial and we can set $w = v$ everywhere. It is convenient to set

$$z = F(v) = \frac{v - A}{B - v} = G(w), \qquad w = v \qquad (3.21)$$

At $n = 1$ any choice of A can do. Different choices of A can be related by an $\mathrm{SL}(2, \mathbb{R})$ transformation that acts on w. It will be convenient for us to choose A so that when we go to $n \sim 1$, it corresponds to the position of the conical singularity.

We now go near $n \sim 1$ and expand

$$e^{i\theta} = e^{i\tau} + e^{i\tau} i\delta\theta(\tau), \qquad (3.22)$$

where $\delta\theta$ is of order $n - 1$. We aim to solve (3.20) for $\delta\theta$. The first step is to find the welding map perturbatively in $(n-1)$. In appendix B, we show that

$$e^{2i\tau}\{F, e^{i\tau}\} = -\delta\{e^{i\theta}, \tau\}_- = -(\delta\theta''' + \delta\theta')_- \qquad (3.23)$$

where we used

$$\delta\{e^{i\theta}, \tau\} \equiv \{e^{i\tau + i\delta\theta}, \tau\} - \{e^{i\tau}, \tau\} = \delta\theta''' + \delta\theta' \qquad (3.24)$$

The minus subscript indicates that this is projected onto negative-frequency modes. This can be written neatly using the Hilbert transform, H, which is defined by the action $\mathrm{H} \cdot e^{im\tau} = -\mathrm{sgn}(m)e^{im\tau}$ (and $\mathrm{H} \cdot 1 = 0$). Then

$$e^{2i\tau}\{F, e^{i\tau}\} = -\frac{1}{2}(1 + \mathrm{H})(\delta\theta''' + \delta\theta'). \qquad (3.25)$$

Wherever else F appears in (3.20), it is multiplied by $(n-1)$, so there we can set $F = \frac{v-A}{B-v}$, as in (3.21). Therefore the equation of motion for the perturbation is

$$\partial_\tau(\delta\theta''' + \delta\theta') + \frac{ic}{12\phi_r}\mathrm{H} \cdot (\delta\theta''' + \delta\theta') = (n-1)\left[\frac{c}{12\phi_r}\mathcal{F} - \partial_\tau R(\tau)\right] \qquad (3.26)$$

---

[5]The trivial temperature dependence is restored by $\tau \to \frac{2\pi}{\beta}\tau_{phys}$, with τ_{phys} the physical Euclidean time with period β.

where

$$\mathcal{F} = -i\frac{e^{2i\tau}(A-B)^2}{(e^{i\tau}-A)^2(e^{i\tau}-B)^2} + \text{c.c.} \ . \tag{3.27}$$

Equation (3.26) is nonlocal, due to the Hilbert transform. We can solve it by expanding both sides in a Fourier series. The important observation is that, due to the structure of derivatives in each term of the left hand side of (3.26), the terms with Fourier modes of the form $e^{ik\tau}$ for $k = 0, \pm 1$ are automatically zero in the left hand side. Therefore, in order to solve this equation, we must impose the same condition on the right-hand side. The $k = 1$ mode requires

$$\int_0^{2\pi} d\tau e^{-i\tau}\left(\frac{c}{12\phi_r}\mathcal{F} - \partial_\tau R(\tau)\right) = 0 \ . \tag{3.28}$$

Doing the integrals, this gives the condition

$$\frac{c}{6\phi_r}\frac{\sinh\frac{a-b}{2}}{\sinh\frac{b+a}{2}} = \frac{1}{\sinh a} \ . \tag{3.29}$$

This matches the equation for the quantum extremal surface (3.10) that came from the derivative of the generalized entropy. The term with $k = 0$ is automatically zero in the right hand side, as $\partial_\tau R$ is explicitly a total derivative and $\int_0^{2\pi} d\tau\mathcal{F} = 0$.

Thus we have reproduced the QES directly from the equations of motion. Once the QES condition is imposed, it is straightforward to solve for the rest of the the Fourier modes of $\delta\theta$ to confirm that there is indeed a solution.

The Hilbert transform that appeared in the equations of motion (3.26) has a natural interpretation in Lorentzian signature as the term responsible for dissipation of an evaporating black hole into Hawking radiation. This is elaborated upon in appendix C.

3.5 Entropy

To calculate the entropy, we must evaluate the action to leading order in $n - 1$. By the general arguments of section 2.1, this will reproduce the generalized entropy in the bulk. Here we will check this explicitly.

The gravitational action (2.8) in terms of the Schwarzian is

$$-I_{\text{grav}} = S_0 + \frac{\phi_r}{2\pi}n\int_0^{2\pi} d\tau\left(\{e^{i\theta},\tau\} + \frac{1}{2}\left(1-\frac{1}{n^2}\right)R(\theta)\right) \ . \tag{3.30}$$

The first term is $-S_0$ times the Euler characteristic of the replica wormholes, $\chi = 1$ in this case. After normalizing, the contribution to $-\log\text{Tr}(\rho_R)^n$ for $n \approx 1$ is

$$-I_{\text{grav}}(n) + nI_{\text{grav}}(1) \approx (1-n)S_0 + (n-1)\frac{\phi_r}{2\pi}\int_0^{2\pi} d\tau R(\tau)$$
$$+ (n-1)\frac{\phi_r}{2\pi}\int_0^{2\pi} d\tau\partial_n\{e^{i\theta},\tau\} \ . \tag{3.31}$$

The first two terms give the area term in the generalized entropy. The second term is the dilaton at the branch point,

$$\frac{\phi_r}{2\pi}\int_0^{2\pi} d\tau R(\tau) = -\frac{\phi_r}{\tanh a} \tag{3.32}$$

The leading term in the matter action is the von Neumann entropy of the CFT,[6] plus a contribution from an order $(n-1)$ change in the metric

$$\log Z_n^{\mathrm{mat}} - n \log Z_1^{\mathrm{mat}} = -(n-1)S_{\mathrm{bulk}}([-a,b]) + \delta_g \log Z_M .$$ (3.33)

The matter action is evaluating on the manifold with the dynamical twist point in the gravity region, so the bulk entropy includes the island, I. By the equation of motion at $n=1$, the last term in (3.31) cancels the last term in (3.33), leading to

$$\log \mathrm{Tr}(\rho_B)^n \approx (1-n)S_{\mathrm{gen}}([-a,b]) \quad \rightarrow \quad S([0,b]) = S_{\mathrm{gen}}([-a,b]) ,$$ (3.34)

as predicted by the general arguments reviewed in section 2.1 [30].

3.6 High-temperature limit

For general n is is convenient to write the equation as follows. The problem has an $\mathrm{SL}(2,\mathbb{R})$ gauge symmetry that acts on w and A. We can use it to gauge fix $A = 0$. Then the equation (3.20) becomes

$$\partial_\tau \{e^{i\theta(\tau)/n}, \tau\} = \kappa i e^{2i\tau} \left[-\frac{1}{2}\left(1 - \frac{1}{n^2}\right)\frac{F'(e^{i\tau})^2}{F(e^{i\tau})^2} - \{F, e^{i\tau}\} \right] + \mathrm{c.c.}$$ (3.35)

Where we introduced

$$\kappa \equiv \frac{c\beta}{24\pi\phi_r}$$ (3.36)

This is proportional to the ratio of c and the near extremal entropy of the black hole $S - S_0$. When this parameter is small, the equations simplify. This essentially corresponds to weak gravitational coupling. In this section we will study the equations for small $\kappa \ll 1$.

To leading order, we can ignore the effects of welding and set $F = G$ with

$$F(v) = \frac{v}{B - v} , \qquad G(w) = \frac{w}{B - w}$$ (3.37)

This eliminates all the effects of welding, so the equation of motion is a completely explicit differential equation for $\theta(\tau)$. We expand

$$\theta(\tau) = \tau + \delta\theta(\tau) ,$$ (3.38)

with $\delta\theta$ of order κ. The equation (3.35) is

$$\partial_\tau \left(\delta\theta''' + \frac{1}{n^2}\delta\theta' \right) = \frac{\kappa}{2}\left(1 - \frac{1}{n^2}\right)\mathcal{F}$$ (3.39)

with

$$\mathcal{F} = -i\left(1 - \frac{e^{i\tau}}{B}\right)^{-2} + \mathrm{c.c.} .$$ (3.40)

---

[6]This is derived in the standard way, for example by integrating the CFT Ward identity for $\partial_b \log Z_M$ [48].

We can expand this in a power series. The constant Fourier mode is absent in the right hand side of (3.39). After solving (3.39) in Fourier space we get

$$\delta\theta = -i\frac{\kappa}{2}\left(1 - \frac{1}{n^2}\right)\sum_{m=1}^{\infty}\frac{(m+1)}{m^2(m^2 - \frac{1}{n^2})}\frac{e^{im\tau}}{B^m} + \text{c.c.}. \tag{3.41}$$

This is the solution to this order. Inserting this into the action we can compute the Renyi entropies. We can go to higher orders by solving the conformal welding problem for $\theta = \tau + \delta\theta$, as explained in [45], computing the flux to next order, and solving again the Schwarzian equation to find the next approximation for $\theta(\tau)$. In this way we can systematically go to any order we want.

As a check of (3.41), we can consider the $n \to 1$ limit. In this case all Fourier coefficients of (3.41) go to zero except $m = \pm 1$ so that we get

$$\delta\theta = -i\frac{\kappa}{B}(e^{i\tau} - e^{-i\tau}) \tag{3.42}$$

In order to compare with the results of the quantum extremal surface calculation we should recall that we have gauge fixed A to be zero. Indeed the final solution (3.42) looks like an infinitesimal $\text{SL}(2,\mathbb{R})$ transformation of the $\theta = \tau$ solution. This is precisely what results from the transformation

$$e^{i\theta} \sim e^{i\tau}(1 + i\delta\theta) \sim \frac{e^{i\tau} - A}{1 - Ae^{i\tau}} \sim e^{i\tau}(1 - Ae^{-i\tau} + Ae^{i\tau}), \qquad A \sim \frac{\kappa}{B} \ll 1 \tag{3.43}$$

for small A as in (3.11). This shows that the finite-n solution at high temperatures has the right $n \to 1$ limit.

4 Single interval at zero temperature

There is a very simple version of the information paradox at zero temperature [14]. Consider the region R in figure 13. Ignoring gravity, the von Neumann entropy of the quantum fields on this region is infrared divergent. This is the Hawking-like calculation of the entropy using quantum field theory on a fixed background.

The state of the quantum fields on a full Cauchy slice is pure. However, the AdS_2 region is supposed to be a quantum system with e^{S_0} states. This is a contradiction, because it is impossible for the finite states in the AdS_2 region to purify the IR-divergent entropy of region R. The UV divergence is not relevant to this issue because it is purified by CFT modes very close to the endpoint.

This is resolved by including an island, as in figure 13 [14]. We will describe briefly how this is reproduced from a replica wormhole. This doesn't require any new calculations because we can take the limit $\beta \to \infty$ in the finite temperature result. The pictures, however, are slightly different, because the replica geometries degenerate in this limit and the topology changes.

4.1 Quantum extremal surface

The metric and dilaton for the zero-temperature solution are

$$ds^2_{\text{in}} = \frac{4dy d\bar{y}}{(y + \bar{y})^2}, \qquad \phi = -\frac{2\phi_r}{y + \bar{y}}, \qquad y = \sigma + i\tau \tag{4.1}$$
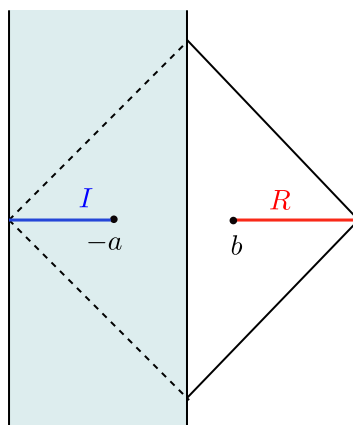
Figure 13. An information puzzle at zero temperature, with AdS_2 on the left and flat space on the right. The naive calculation of matter entropy in region R is infrared-divergent, but this cannot be purified by quantum gravity in AdS_2. This is resolved by including the island, I.

with $\sigma < 0$. As before we glue it to flat space $dy d\bar{y}$ at $\sigma = 0$. The region R and the island I are the intervals

$$I: \quad y \in (-\infty, -a], \qquad R: \quad y \in [b, \infty) \tag{4.2}$$

at $t = 0$. The generalized entropy, including the island, is

$$S_{\text{gen}}(I \cup R) = \frac{\phi_r}{a} + \frac{c}{6} \log \frac{(a+b)^2}{a} \ . \tag{4.3}$$

Setting $\partial_a S_{\text{gen}} = 0$ gives the position of the QES,

$$a = \frac{1}{2}(k + b + \sqrt{b^2 + 6bk + k^2}), \quad k \equiv \frac{6\phi_r}{c} \ . \tag{4.4}$$

4.2 Replica wormholes at zero temperature

The replica partition function $\text{Tr}(\rho_A)^n$ is given by the path integral in figure 14. The boundary condition for the gravity region is n copies of the real line. The Hawking saddle fills in the gravity region with n independent copies of $\mathbb{H}_2$. The replica wormhole, shown in the figure, fills in the gravity region with a single copy of $\mathbb{H}_2$. To see all n sheets of the gravity region, we go to the uniformizing coordinate

$$\widetilde{w} = \left(\frac{a+y}{a-y}\right)^{1/n} \ . \tag{4.5}$$

This maps the full gravity region to a single hyperbolic disk, $|\widetilde{w}| < 1$. This disk is a wormhole connecting n copies of flat space. The nth copy is glued to the segment with arg $\widetilde{w} \in [-\frac{\pi}{n}, \frac{\pi}{n}]$.

The equation of motion, and the answer for the position of the QES, is found by taking $\beta \to \infty$ in the results of section 3. This of course agrees with (4.4). (It is also possible to solve this problem directly at zero temperature, but we found it easier to treat the welding problem at finite temperature where the gluing is compact. In the end, the welding effects drop out in the determination of the position of the QES, as we saw below (3.26).)
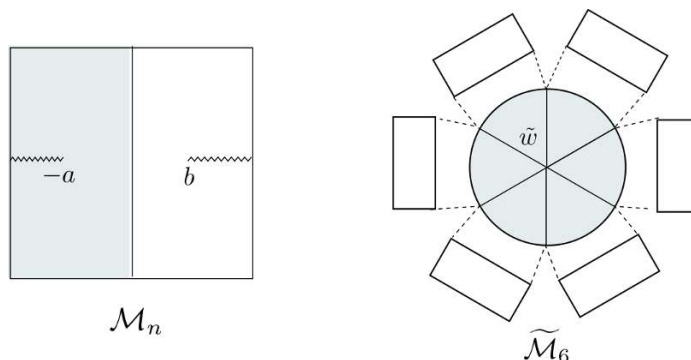
Figure 14. Replica wormhole at zero temperature. On the right, the disk is glued to n copies of the half-plane, as indicated by the dashed lines.

5 Two intervals in the eternal black hole

We now turn to the information paradox in the eternal black hole [14], described in the introduction and pictured in figure 3. In the late-time regime relevant to the information paradox, the generalized entropy, including the island, is simply twice the answer for a single interval. We would like to understand how this is reproduced from wormholes. This is essentially just putting together the general discussion of section 2 with the single-interval results of section 3, so we will be brief. We will only discuss the saddles near $n = 1$; it would be nice to have a more complete understanding of the finite-n wormholes in this setup.

5.1 Review of the QES

We set $\beta = 2\pi$. The points in figure 3 have (σ, t) coordinates

$$P_1 = (-a, t_a), \quad P_2 = (b, t_b), \quad P_3 = (-a, -t_a + i\pi), \quad P_4 = (b, -t_b + i\pi) . \tag{5.1}$$

The radiation region is

$$R = [P_4, \infty_L) \cup [P_2, \infty_R), \tag{5.2}$$

and the island is

$$I = [P_3, P_1] . \tag{5.3}$$

The CFT state is pure on the full Cauchy slice, so

$$S_{\text{CFT}}(I \cup R) = S_{\text{CFT}}([P_4, P_3] \cup [P_1, P_2]) . \tag{5.4}$$

This entropy is non-universal; it depends on the CFT. In the theory of c free Dirac fermions [49], the entanglement entropy of the region

$$[x_1, x_2] \cup [x_3, x_4], \tag{5.5}$$

with metric $ds^2 = \Omega^{-2} dx d\bar{x}$, is

$$S_{\text{fermions}} = \frac{c}{6} \log \left[\frac{|x_{21} x_{32} x_{43} x_{41}|^2}{|x_{31} x_{42}|^2 \Omega_1 \Omega_2 \Omega_3 \Omega_4} \right] . \tag{5.6}$$

where we dropped the UV divergences. With our kinematics and conformal factors, this gives

$$S_{\text{fermions}}(I \cup R) = \frac{c}{3} \log \left[\frac{2 \cosh t_a \cosh t_b \left| \cosh(t_a - t_b) - \cosh(a + b) \right|}{\sinh a \cosh(\frac{a+b-t_a-t_b}{2}) \cosh(\frac{a+b+t_a+t_b}{2})} \right] \tag{5.7}$$

In a general CFT, the two-interval entanglement entropy is a function of the conformal cross-ratios $(z, \bar{z})$ which agrees with (5.7) in the OPE limits $z \to 0$ and $z \to 1$. For concreteness we will do the calculations for the free fermion, but the regime of interest for the information paradox will turn out to be universal.

The generalized entropy, including the island, is

$$S_{\text{gen}}(I \cup R) = 2S_0 + \frac{2\phi_r}{\tanh a} + S_{\text{fermions}}(I \cup R), \tag{5.8}$$

Without an island, the entropy is the CFT entropy on the complement of R, the interval $[P_4, P_2]$, which is

$$S_{\text{gen}}^{\text{no island}} = S_{\text{fermions}}(R) = \frac{c}{3} \log \left(2 \cosh t_b \right) \tag{5.9}$$

At $t = 0$,

$$S_{\text{gen}}^{\text{island}} = 2S_0 + \frac{2\phi_r}{\tanh a} + \frac{c}{3} \log \left(\frac{4 \tanh^2 \frac{a+b}{2}}{\sinh a} \right) . \tag{5.10}$$

The extremality condition $\partial_a S_{\text{gen}}^{\text{island}} = 0$ at $t_a = t_b = 0$ gives

$$\frac{6\phi_r}{c} \sinh(a + b) = 2 \sinh^2 a - \sinh a \cosh a \sinh(a + b) . \tag{5.11}$$

Whether this has a real-valued solution depends on the parameters b and ϕ_r/c. For example, if $b = 0$, then it has a real solution minimizing $S_{\text{gen}}^{\text{island}}$ when ϕ_r/c is small, but not otherwise.

At late times, the extremality condition $\partial_a S_{\text{gen}}^{\text{island}} = 0$ always has a real solution. The true entropy, according to the QES prescription, is

$$S(R) = \min \left\{ S_{\text{gen}}^{\text{no island}}, S_{\text{gen}}^{\text{island}} \right\} . \tag{5.12}$$

The island always exists and dominates the entropy at late times, because the non-island entropy grows linearly with t, see figure 5. This solution is in the OPE limit where we can approximate the entanglement entropy by twice the single-interval answer,

$$S_{\text{matter}}(I \cup R) \approx 2S_{\text{matter}}([P_1, P_2]) = \frac{c}{3} \log \left(\frac{2 |\cosh(a + b) - \cosh(t_a - t_b)|}{\sinh a} \right) . \tag{5.13}$$

and the QES condition sets $t_a = t_b$.

5.2 Replica wormholes

We would like to discuss some aspects of the wormhole solutions that lead to the island prescription.

For general n these are wormholes which have the topology shown in figures 6(b), 7. Already from these figures we can derive the S_0-dependent contribution (2.7) since it involves only the topology of the manifold. The replica wormhole that involves nontrivial connections, see figure (7), has the topology of a sphere with n holes. This gives a contribution going like $Z_n \propto e^{S_0(2-n)}$ and a contribution of $2S_0 = (1 - n\partial_n) \log Z_n|_{n=1}$ for the von Neumann entropy. This is good, since the island contribution indeed had such a term (5.10).

It is useful to assume replica symmetry and view the Riemann surface as arising from a single disk with n copies of the matter theory and with pairs of twist operators that connect all these n copies in a cyclic fashion, see figure 9. In order to find the full answer, we need to solve the equations (2.17) (2.18). The important point is that, at this stage, we have that n appears purely as a parameter and we can analytically continue the equations in n. We have not managed to solve the equations for finite n. But let us discuss some properties we expect. In the limit of large $c\beta/\phi_r$, it is likely that solutions exist in Euclidean signature.[7] We can put points P_2 and P_4 at $v = \pm Be^{\pm i\varphi}$. Once this solution is found, we can analytically continue $\varphi \to -it$ to generate the Lorentzian solution. That Lorentzian solution at late times t is expected to exist even for low values of $c\beta/\phi_r$. In principle, it should be possible, and probably easier, to analyze directly the late-times Lorentzian equation. In fact, we expect that there should be a way to relate the single interval solution to the two interval solution in this regime. The intuitive reason is that at late times the distance between the two horizons is increasing and so the distance between the two cosmic branes is increasing. We have an external source cosmic brane outside the gravitational region, at the tip of region R. The cosmic brane has some tension, as well as a twist operator on it. For the Hawking saddle, the one without the replica wormholes, the twist operators, and the topological line operators[8] that connect them, generate a contribution that grows linearly in time, due to the behavior of Renyi entropies for the matter quantum field theory, as well as the fact that the wormhole length grows with time. At late times the topological line operator can break by pair producing cosmic branes, with their twist operators. The cost of creating a pair of cosmic branes is finite in the gravitational region, because the dilaton is finite. This cost would be infinite in the non-gravitational region. But once the external cosmic brane is screened by the cosmic brane that appeared in the gravity region we expect to have two approximately independent single interval problems. The reason is that the distance between the left and right sides is growing with time. This is somewhat analogous to two point charges that generate a two dimensional electric field. As one separates the charges it might be convenient to create a pair of charges that screens the electric field. For this it is important that the charges one creates have finite mass.

In the $n \to 1$ limit we can analyze the solution and we get the generalized entropy. This is not too surprising since the arguments in [30] say that this should always work. Here the non-trivial input is the ansatz for the configuration of intervals which follows from the structure of the Riemann surfaces. As discussed in section 2.1, the effective action reduces

---

[7]For low values of $c\beta/\phi_r$ we have already seen, in (5.11), that near $n \sim 1$ the solutions can be complex.

[8]These topological line operators exchange the n copies in a cyclic way. They are represented by red lines in figure 9(b).
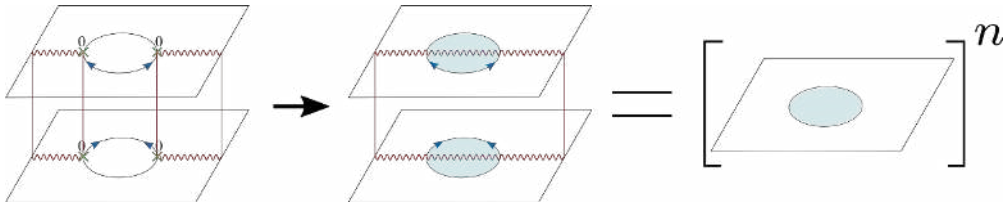
Figure 15. The computation of the entropy of the entire flat space regions including the boundary points. The dominant gravitational saddle connects consecutive sheets. This factorizes into n separate sheets and produces a vanishing entropy consistent with the purity of the flat space region union the endpoints. The blue arrows indicate how the unit circle is identified across the cut.

to the action of certain cosmic branes which are manifestly very light in the $n \to 1$ limit. So in this case, the argument of the previous paragraph can be explicitly checked and one indeed obtains that we get the sum over the two single interval problems [14].

5.3 Purity of the total state

One can take the perspective that our model is defined via a quantum theory living on the flat space region including its boundary endpoints. The global pure state we consider should be a pure state of this region, and a natural question is whether this is captured in the gravity description. Replica wormholes do indeed capture this feature.

The computation of the entropy of this region is given by evaluating the path integral on the manifold shown in figure 15. The branch cuts split the entire flat space region including its boundaries, identifying one half of one sheet with the other half of the next sheet. The most obvious gravitational saddle is the one that connects these consecutive sheets and thereby naturally extending the branch cut through the entire gravity region. A simple rearranging of these sheets shows that this contribution to the Renyi entropy factorizes. This disconnected saddle satisfies $Z_n = Z_1^n$, and evaluating the on shell action on this configuration will give vanishing entropy since

$$\text{Tr}\,\rho^n = \frac{Z_n}{Z_1^n} = 1\,. \tag{5.14}$$

This saddle clearly dominates over all other configurations.

Since the different sheets are not coupled at all in the flat space region, it's plausible that this disconnected saddle is the only saddle that exists. Other off-shell contributions can indeed exist, but we speculate they should give a vanishing contribution in a model with a definite Hamiltonian with no averaging.

6 Comments on reconstructing the interior

The island contribution to the entropy of a flat space region R indicates there is a dictionary between the island I and R in the sense of entanglement wedge reconstruction in AdS/CFT. We could discuss this in general but for concreteness consider the two interval case discussed in the previous section. Let's take the state at late times such that the entropy of R has plateaued and its entropy receives a contribution from the island as shown in figure 16.

The first step to establishing a dictionary is to define a subspace of states which have the same "entanglement wedge" or island. This defines what we will call the code subspace $\mathcal{H}_{\text{code}}$, which we imagine can be prepared via the Euclidean path integral with possible operator insertions. By having the same island we mean that the leading saddle points in the Renyi computations are only modified perturbatively. This naturally puts restrictions on the size of the allowed code subspace for which the statements of this section hold, see for example [50, 51].

We assume that the full Hilbert space of our model is that of the two flat space regions including their boundary, which we write as $\mathcal{H}_{\text{Left}} \otimes \mathcal{H}_{\text{Right}}$. The region R that we are considering is a tensor factor of this Hilbert space, where we can write

$$\mathcal{H}_{\text{Left}} \otimes \mathcal{H}_{\text{Right}} = \mathcal{H}_R \otimes \mathcal{H}_{\bar{R}} \tag{6.1}$$

where $\bar{R}$ is the complement of the region R in the flat space region including the boundary points.

The code subspace $\mathcal{H}_{\text{code}}$ is a subspace of $\mathcal{H}_{\text{Left}} \otimes \mathcal{H}_{\text{Right}}$. However, the code subspace also has a simpler description in terms of the combined description of gravity plus the flat space region as that of effective field theory on a Cauchy slice of the full spacetime. This is the description where the state is prepared using the semi-classical saddle via the Euclidean black hole solution. The code subspace should be thought of as isomorphic to this. Therefore, the code subspace admits the decomposition[9]

$$\mathcal{H}_{\text{code}} \cong \mathcal{H}_R \otimes \mathcal{H}_D \otimes \mathcal{H}_I \tag{6.2}$$

where the region D is the complement of $R \cup I$ on the Cauchy slice. The decomposition is shown in figure 16. It is within this effective description that for any state in the code subspace $|i\rangle \in \mathcal{H}_{\text{code}}$, we have

$$S(\rho_R^i) = S(\tilde{\rho}_{RI}^i) + \frac{\text{Area}[\partial I]}{4G_N} \tag{6.3}$$

where ρ_R^i is what you get by tracing out $\bar{R}$ in the full quantum description $\mathcal{H}_{\text{Left}} \otimes \mathcal{H}_{\text{Right}}$, and $\tilde{\rho}_{RI}^i$ is the density matrix obtained by tracing out the complement of RI in the semi-classical description consisting of quantum fields on a classical geometry.

The validity of the island formula (for a fixed island) within the code subspace implies the equivalence of the relative entropy in the exact state and the semi-classical state:

$$S_{\text{Rel}}(\rho_R|\sigma_R) = S_{\text{Rel}}(\tilde{\rho}_{RI}|\tilde{\sigma}_{RI}) \tag{6.4}$$

A similar observation in the context of AdS/CFT [52] was key in proving entanglement wedge reconstruction [53] using the quantum error correction interpretation of the duality [54]. The same line of argument can be applied here to establish the dictionary. In particular, one can show that for any operator $\mathcal{O}_I$ (and its Hermitian conjugate) acting

---

[9]This should be understood as approximate up to usual issues of the non-factorizability of continuum QFT.
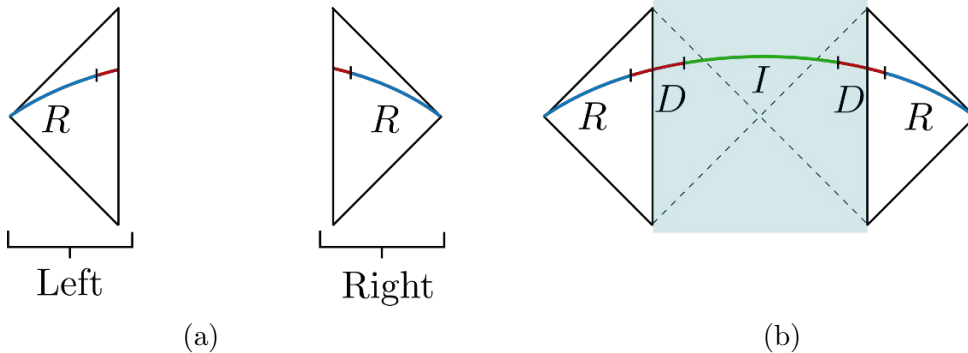
Figure 16. (a) The full Hilbert space is the product Hilbert space of the entire left and right flat space regions including the boundary points. The region R we are interested in is a union of two subregions in the two flat space regions. (b) The effective state used in the island prescription is the semi-classical state defined on the Cauchy slice of the full system. R is the same region in the flat space region whose exact entropy we are computing, I is the island, and D is the complement of the two.

within the $\mathcal{H}_{\text{code}}$ and supported on the island one can find an operator supported on R such that:

$$\mathcal{O}_I|i\rangle = \mathcal{O}_R|i\rangle \tag{6.5}$$

$$\mathcal{O}_I^\dagger|i\rangle = \mathcal{O}_R^\dagger|i\rangle \tag{6.6}$$

The operator $\mathcal{O}_R$ is given by a complicated operator on R involving the matrix elements of $\mathcal{O}_I$ within the code subspace.

In summary, we are using the fine grained entropy formula to understand how the interior is encoded in the full Hilbert space. The relative entropy equality (6.4) tells us that distinguishable states in the interior (the island) are also distinguishable in the radiation, within the full exact quantum description.

7 Discussion

In this paper, we have exhibited non-perturbative effects that dramatically reduce the late time von Neumann entropy of quantum fields outside a black hole.

The computation of the Renyi entropies corresponds to the expectation value of a swap or cyclic permutation operator in n copies of the theory. Systems with very high entropy have very small, exponentially small, expectation values for this observable. This means that non-perturbative effects can compete with the naive answers. In particular, the Hawking-like computation of the Renyi entropies of radiation corresponds to a computation on the leading gravitational background. A growing entropy corresponds to an exponentially decreasing expectation value for the cyclic permutation operator. It decreases exponentially as time progresses. For this reason, we need to pay attention to other geometries, with other topologies. These other topologies give exponentially small effects, but they do not continue decreasing with time for long times. Said in this way,

the effects are vaguely similar to the ones discussed for corrections of other exponentially small effects [32–35]. Though the Renyi entropies are small, the von Neumann entropy is large and the new series of saddles gives rise to a constant von Neumann entropy at late times. More precisely, we can think of the computation of the Renyi entropies in the two interval case as an insertion of a pair of external cosmic branes in the non-gravitational region. As time progresses these are separated further and further through the wormhole. Eventually the dominant contribution is one where a pair of cosmic branes is created in the gravitational region that "screen" the external ones, giving an entropy which is the same as that of two copies of the single interval entropy.

These other topologies are present as subleading saddles also at short times (perhaps as complex saddles) where we can analyze them using Euclidean methods and then analytically continue. We have only done this analytic continuation for the von Neumann entropies, not the Renyi entropies. It would be interesting to do it more explicitly for the Renyi entropies.

There have been discussions on whether small corrections to the density matrix, of order $e^{-S_{BH}}$, could or could not restore unitarity. These results suggest that they interfere constructively to give rise to the right expression for the entropy.

This is evidence that including nonperturbative gravitational effects can indeed lead to results compatible with unitarity. However, we emphasize that this is not a full microscopic resolution of the information paradox. We have not given a gravitational description for the S-matrix describing how infalling matter escapes into the radiation. In this sense, these results are on a footing similar to the Bekenstein-Hawking calculation of the entropy, which uses a Euclidean path integral to compute the right answer but does not give an explicit Hilbert space picture for what it is counting. In contrast, the Strominger-Vafa computation of the entropy [55] gives us an explicit Hilbert space, but not a detailed description of the microstates in the gravity variables. Something similar can be said of the CFT description in AdS/CFT. Hopefully these results will be useful for providing a more explicit map.

It is amusing to note that wormholes were initially thought to destroy information [56–58]. But more recently the work of [34, 35], as well as the present discussion, and [31], suggests that the opposite is true. Wormholes are important for producing results that are compatible with unitarity. For earlier work in this direction see also [59–61].

We assumed that $c \gg 1$ as a blanket justification for analyzing the equations classically. However, even for small $c \sim 1$, the basic picture for the Page curve can be justified. The basic point is simple. First consider the single interval computation. In that case for $c \sim 1$ we see that the correction to the black hole solution is very small, for all the Renyi entropies. In other words, we find that A is small, and we can probably not distinguish such a small value of A from zero but that does not matter, the geometries and the entropies are basically those of a black hole. Now when we go to two intervals, and we consider the late time situation, then all that really matters is that we can do an OPE-like expansion of the twist operator insertions. The important observation is that the twist operator insertions in the interior of the black hole are very far from each other. This is the fact that the wormhole is getting longer [62, 63]. Then the solution becomes similar to two non-interacting copies of the single interval solution. The fact that c is small only implies

that we will have to wait longer for the island solution to dominate. We just have to wait a time of order the entropy, $t \propto \beta(S - S_0)/c$ for it to dominate.

In [34], it was argued that pure JT gravity should be interpreted in terms of an average over Hamiltonians. In addition, higher genus corrections were precisely matched. This has raised the question of whether the corrections we are discussing in this paper crucially involve an average over Hamiltonians, or whether they would also apply to a system which has a definite Hamiltonian. Though JT gravity plus a CFT probably does not define a complete quantum gravity theory, it seems likely that well defined theories could be approximated by JT gravity plus a CFT. For example, we could imagine an AdS/CFT example that involves an extremal black hole such that it also has a CFT on its geometry. All we need is this low energy description, the theory might have lots of other massive fields which will not drastically participate in the discussion. They might lead to additional saddles, but it seems that they will not correct the saddles we have been discussing. And we have the seen that the saddles we discussed already give an answer consistent with unitarity, at least for the entropy. In contrast with [34], we are not doing the full path integral, we are simply using a saddle point approximation, so the JT gravity plus CFT only needs to be valid around these saddles.

As we mentioned in the introduction, the setup in this paper can be viewed as an approximation to some magnetically charged near extremal four dimensional black holes [40]. But one could analyze more general asymptotically flat black holes and wonder how to define either exactly or approximately the various entropies involved. In particular, to have a sharp definition of the entropy of radiation it seems important to go to null infinity.

Another interesting question is whether we can give a Lorentzian interpretation to the modification of the density matrix implied by the existence of replica wormholes.

It has been pointed out that a black hole as seen from outside looks like a system obeying the laws of hydrodynamics. For this reason, it is sometimes thought that gravity is just an approximation that intrinsically loses information. Here we see that if we include the black hole interior, and we do a more complete gravity computation, we can get results compatible with unitarity. The fact that gravity is more than dissipative hydrodynamics is already contained in the Ryu-Takayanagi formula for the fine grained entropy, which shows that the geometry of the interior can discriminate between pure and mixed states for a black hole.

Acknowledgments

A Derivation of the gravitational action

In this appendix we derive the action that leads to the equation of motion (2.17).

We start with the expansion of the metric near the boundary (2.15) (2.16)

$$ds^2 = \frac{4dwd\bar{w}}{(1-|w|^2)^2} \left(1 - \frac{2}{3}(1-|w|)^2 U(\theta) + \cdots \right). \tag{A.1}$$

We now write in terms of the variables $w = e^{-\gamma}e^{i\theta}$ and expand it in powers of γ as

$$ds^2 = \frac{d\theta^2}{\gamma^2} + \frac{d\gamma^2}{\gamma^2} - \frac{2}{3}d\theta^2 U(\theta). \tag{A.2}$$

We now equate this to $ds = \frac{d\tau}{\epsilon}$, we set $\theta = \theta(\tau)$ and solve for γ in a power series

$$\gamma = \epsilon\theta' \left[1 + \epsilon^2 \left(\frac{1}{2}\frac{\theta''^2}{\theta'^2} - \frac{1}{3}U(\theta)\theta'^2 \right) + \cdots \right]. \tag{A.3}$$

We can now compute the tangent vector to the curve t^μ and the normal vector n^μ and compute the extrinsic curvature from

$$K = t^\mu t^\nu \nabla_\mu n_\nu = 1 + \epsilon^2 \left[\{\theta, \tau\} + \left(\frac{1}{2} + U(\theta) \right) \theta'^2 \right]. \tag{A.4}$$

Up to the purely topological term, the gravitational action (2.8) reduces to the extrinsic curvature term

$$- I_{\text{grav}} = \frac{1}{4\pi}\frac{\phi_r}{\epsilon} \int \frac{d\tau}{\epsilon} 2K = \frac{2\phi_r}{4\pi\epsilon^2} \int d\tau + \frac{\phi_r}{2\pi} \int d\tau \left[\{\theta(\tau), \tau\} + \left(\frac{1}{2} + U(\theta) \right) \theta'^2 \right] + o(\epsilon). \tag{A.5}$$

The first term is a purely local divergence that can be viewed as the correction to the vacuum energy. We should also remark that we can always choose a coordinate x where the metric locally looks like the standard Poincare coordinates. In those coordinates the action is simply $\{x, \tau\}$. However, we will have a nontrivial identification for x as we move from $\tau \to \tau + 2\pi$. Here we simplified the boundary condition, it is just $\theta = \theta + 2\pi$, but we complicated a bit the action. Notice that we can think of $U(\theta)$ as a stress tensor, the change of coordinates is basically the same that we use to transform this stress tensor to zero. In other words, $x(\theta)$ is a function which obeys $\{x, \theta\} = \frac{1}{2} + U(\theta)$.

The conserved energy of the system is given by

$$E = \frac{\phi_r}{2\pi} \left[\{\theta(\tau), \tau\} + \left(\frac{1}{2} + U(\theta) \right) \theta'^2 \right]. \tag{A.6}$$

We now compute $U(\theta)$ for the case when we put a conical defect at point A in the w plane. We have the metric (3.14) and the change of coordinates (3.13) which imply that

$$ds^2 = \left|\frac{d\tilde{w}}{dw}\right|^2 \frac{4|dw|^2}{(1-|\tilde{w}|^2)^2}$$

$$= \frac{4|dw|^2}{(1-|w|^2)^2}\left[1 - \frac{2}{3}(1-|w|)^2 U(\theta) + \cdots\right], \quad \text{as } |w| \to 1 \qquad (A.7)$$

with

$$U(\theta) = -\frac{1}{2}\left(1 - \frac{1}{n^2}\right)\frac{(1-A^2)^2}{(e^{i\theta} - A)^2(e^{-i\theta} - A)^2}, \qquad (A.8)$$

which leads to the same action as (3.17)

We now would like to derive the equations of motion for this action. In particular, we would like to see that as $\theta \to \theta + \delta\theta$ we get the right equations of motion. The change in gravitational action is simple, we just have

$$-\delta I_{\text{grav}} = -\frac{\phi_r}{2\pi}\int d\tau \frac{\left[\{\theta(\tau), \tau\} + (\frac{1}{2} + U(\theta))\theta'^2\right]'}{\theta'}\delta\theta. \qquad (A.9)$$

Now, let us do the variation of the CFT part. Imagine that we choose locally complex coordinates so that

$$\log w = s + i\theta \qquad (A.10)$$

We also have the outside coordinates $y = \sigma + i\tau$ and we can locally think of the relation between the two in terms of $\log w = i\theta(-iy)$. Now imagine that we do a small change $\theta(\tau) \to \theta + \delta\theta$ with $\delta\theta$ with compact support. This would change the relation between the two sides. However, let us imagine we instead keep the relation fixed, set by $\theta(\tau)$ and we redefine the outside coordinate by an infinitesimal reparametrization, $\tilde{y} = y + \zeta^y$ in such a way that the relation between the new variables is the same as the old one

$$\log w = i\theta(-i\tilde{y}) = i\theta(-iy) + i\delta\theta(-iy) = i\theta(-iy) + \theta'(-iy)\zeta^y \quad \to \quad \zeta^y = i\frac{\delta\theta}{\theta'} \quad (A.11)$$

and we have the complex conjugate expression for $\zeta^{\bar{y}}$. We can then extend this reparametrization in a non-holomorphic way in the region outside, defining

$$\tilde{\zeta}^y = i\frac{\delta\theta(-iy)}{\theta'(-iy)}h(\sigma), \qquad \tilde{\zeta}^{\bar{y}} = -i\frac{\delta\theta(i\bar{y})}{\theta'(i\bar{y})}h(\sigma) \qquad (A.12)$$

where $h(\sigma)$ is one for $\sigma = 0$ and quickly goes to zero at σ increases. An example is $h(\sigma) = \theta(\sigma_0 - \sigma)$ for a small σ_0. This change of coordinates is equivalent to a change in metric

$$ds^2 = dyd\bar{y} = d\tilde{y}d\bar{\tilde{y}} - 2\partial_\alpha\zeta^\beta d\tilde{y}^\alpha d\tilde{y}^\beta, \qquad \delta g_{\alpha\beta} = -2\partial_{(\alpha}\zeta_{\beta)} \qquad (A.13)$$

This differs from the original metric by some terms that are localized near the point where we are doing the variation. The relation between $\log w$ and the $\tilde{y}$ variable was the same as it was before we did the variation, due to our choice of $\tilde{y}$ variable in (A.11). Furthermore, far from the region where we are doing the variation, both variables coincide. Thus, the

only thing we are doing is locally changing the metric of the outside region. Using the definition of the stress tensor, $T_{\alpha\beta} = -\frac{2}{\sqrt{g}}\frac{\delta}{\delta g^{\alpha\beta}}\log Z$, we get

$$
\begin{aligned}
\delta \log \hat{Z}_M &= -\frac{1}{2}\int d\varphi d\sigma (T_{yy}\delta g^{yy} + T_{\bar{y}\bar{y}}\delta g^{\bar{y}\bar{y}}) \\
&= -2\int d\varphi d\sigma (T_{yy}\partial_{\bar{y}}\zeta^y + T_{\bar{y}\bar{y}}\partial_y\zeta^{\bar{y}}),
\end{aligned}
\tag{A.14}
$$

where we used that the background metric is flat and that the trace of the stress tensor is zero. We now use evaluate the derivatives

$$
\partial_{\bar{y}}\zeta^y = \frac{i}{2}\frac{\delta\theta(-iy)}{\theta'(-iy)}h'(\sigma), \qquad \partial_y\zeta^{\bar{y}} = -\frac{i}{2}\frac{\delta\theta(i\bar{y})}{\theta'(i\bar{y})}h'(\sigma), \qquad h' = -\delta(\sigma - \sigma^0). \tag{A.15}
$$

Here we used that the arguments of $\delta\theta$ and θ' are holomorphic or antiholomorphic, so the derivative receives only a contribution from h, which is just a delta function. Inserting this into (A.14), integrating over σ, and taking $\sigma^0 \to 0$, we get

$$
\delta \log \hat{Z}_M = i\int d\tau (T_{yy} - T_{\bar{y}\bar{y}})\frac{\delta\theta}{\theta'}. \tag{A.16}
$$

Using (A.9) we get the appropriate equation (2.17) after cancelling the $1/\theta'$ factor from both sides.

B Linearized solution to the welding problem

Let us start with a discussion of the symmetries of the welding problem (2.11). First we can imagine doing $SL(2,\mathbb{C})$ transformations of the z plane. These move around the point at infinity, and we would need to allow a pole in the functions F or G. If we fix that $F(\infty) = \infty$, then we can then impose that the functions are holomorphic everywhere, with no poles, and this group is reduced to just translations, scalings and rotations of the plane z. None of these transformations change the data for the welding problem which is $\theta(\tau)$. In addition, we have two $SL(2,\mathbb{R})$ transformations, one acting on w and one acting on v, both preserving the circles $|w| = 1$ and $|v| = 1$. These change the data of the welding problem by an $SL(2,\mathbb{R})$ transformation of $e^{i\theta}$ or $e^{i\tau}$ respectively. They map a solution of a welding problem with $\theta(\tau)$ to a solution of a different welding problem given by the transformed function. In our combined gravity plus CFT problem, we are integrating over $\theta(\tau)$, so we can look for symmetries that change $\theta(\tau)$. It turns out that the $SL(2,\mathbb{R})_v$ that acts on the v plane is *not* a symmetry. It changes the Schwarzian action, for example. On the other hand, the $SL(2,\mathbb{R})_w$ is actually a gauge symmetry, when we also act with the $SL(2,\mathbb{R})$ transformation on the possible locations, w_i, of the conical singularities.

Consider a plane with coordinate w inside the unit disk, and v outside, as in figure 10. The plane is glued along the unit circle with a gluing function $\theta(\tau)$, where $w = e^{i\theta}$ and $v = e^{i\tau}$. The solution to the welding problem is a pair of functions

$$
z = G(w) \qquad \text{(inside)} \tag{B.1}
$$

$$
z = F(v) \qquad \text{(outside)} \tag{B.2}
$$

where G is holomorphic inside the disk, and F is holomorphic outside the disk. In this appendix we will solve for F, G perturbatively, assuming the gluing is close to the identity, $\theta(\tau) = \tau + \delta\theta(\tau)$. Here we are considering $\delta\theta(\tau)$ to be a fixed input to the problem of finding F and G.

Expand in Fourier modes,

$$\theta(\tau) = \tau + \sum_{m=-\infty}^{\infty} c_m e^{im\tau}, \quad G(w) = w + \sum_{\ell=0}^{\infty} g_\ell w^\ell, \quad F(v) = v + \sum_{\ell=-\infty}^{2} f_\ell v^\ell. \quad \text{(B.3)}$$

Here c_m, d_1^ℓ, and d_2^ℓ are considered small. There is an SL(2) ambiguity in the zeroth order solution, which we have gauge-fixed to set these maps to the identity. (Note that this is different from the choice in the main text around eqn (3.37).) The matching condition on the unit circle is

$$G(e^{i\theta(\tau)}) = F(e^{i\theta}) . \quad \text{(B.4)}$$

At the linearized level, this sets

$$f_{\ell+1} = ic_\ell \qquad\qquad (\ell \leq -2) \qquad\qquad \text{(B.5)}$$
$$g_{\ell+1} = -ic_\ell \qquad\qquad (\ell \geq 2) \qquad\qquad \text{(B.6)}$$

and

$$ic_{-1} = f_2 - g_2, \quad ic_0 = f_1 - g_1, \quad ic_1 = f_2 - g_2 . \quad \text{(B.7)}$$

There is an ambiguity by a small $\text{SL}(2,\mathbb{C})$ action on the z plane. We can fix it by setting $G(0) = 0$, $F(v) = v+$constant, as $v \to \infty$. This amounts to three complex conditions that set

$$g_0 = f_2 = f_1 = 0 \quad \text{(B.8)}$$

This now implies that we get a unique solution for the remaining coefficients in terms of the c_m

$$f_l = ic_{\ell-1}, \quad \text{for} \ \ \ell \leq 0 ; \qquad\qquad g_\ell = -ic_{\ell-1}, \quad \text{for} \ \ \ell > 0 . \quad \text{(B.9)}$$

From here we can calculate

$$v^2\{F, v\} = \sum_{\ell=-\infty}^{-2} \ell(\ell^2 - 1)ic_\ell v^\ell . \quad \text{(B.10)}$$

Comparing to $\{w, \tau\} = \{e^{i\theta}, \tau\}$ gives the relation used in the main text,

$$e^{2i\tau}\{F, v\} = -\delta\{w, \tau\}_- = -(\delta\theta''' + \delta\theta')_- . \quad \text{(B.11)}$$

C The equation of motion in Lorentzian signature

The Hilbert transform appearing in the equation of motion (3.26) has a nice interpretation in Lorentzian signature. It is responsible for the dissipation of energy into the thermal

bath outside. This makes contact with the Schwarzian equation for black hole evaporation studied in [5, 37].

In this appendix we set $n = 1$, but allow for CFT operators inserted in the non-gravitational region. The perturbative Schwarzian equation in Euclidean signature is

$$\partial_\tau S + i\kappa \mathrm{H} \cdot S = i\kappa \mathcal{F} \tag{C.1}$$

where $S = \delta\{e^{i\theta}, \tau\}$ and

$$\mathcal{F} = T_{yy}(i\tau) - T_{yy}(-i\tau) \ . \tag{C.2}$$

We separate this into positive and negative frequencies on the Euclidean τ-circle,

$$\partial_\tau S_+ - i\kappa S_+ = i\kappa \mathcal{F}_+ \tag{C.3}$$

$$\partial_\tau S_- + i\kappa S_- = i\kappa \mathcal{F}_- \ . \tag{C.4}$$

Here the '+' terms include only the non-negative powers of e^y, and the '−' terms have the negative powers. Now continuing to Lorentzian signature with $\tau = it$, this becomes

$$\partial_t S_\pm \pm \kappa S_\pm = -\kappa \mathcal{F}_\pm \tag{C.5}$$

This is the Lorentzian equation of motion. As an example, consider a state with two scalar operators $\mathcal{O}(y_1)\mathcal{O}(y_2)$ inserted at

$$y_1 = L + i\delta, \qquad y_2 = \bar{y}_1 = L - i\delta \,, \tag{C.6}$$

with $0 < \delta \ll L$. This creates a shockwave that falls into the AdS region at time $t \approx L$. The state is time-symmetric, so there is also a shockwave exiting the AdS region at $t \approx -L$. The stress tensor is

$$T_{yy}(y) = -\frac{h_O}{2\pi} \frac{v^2(v_1 - v_2)^2}{(v - v_1)^2(v - v_2)^2} \,, \tag{C.7}$$

with $v = e^y$. The projections onto positive and negative Euclidean frequencies are

$$\mathcal{F}_+ = -\frac{h_O}{2\pi} \frac{v^2(v_1 - v_2)^2}{(v - v_1)^2(v - v_2)^2} \,, \qquad \mathcal{F}_- = \frac{h_O}{2\pi} \frac{v^2(v_1 - v_2)^2}{(1 - v_1 v)^2(1 - v_1 v)^2} \ . \tag{C.8}$$

In Lorentzian signature this becomes

$$\mathcal{F}_+ = \frac{h_O \sin^2 \delta}{2\pi(\cos \delta - \cosh(L + t))^2} \tag{C.9}$$

$$\mathcal{F}_- = -\frac{h_O \sin^2 \delta}{2\pi(\cos \delta - \cosh(L - t))^2} \tag{C.10}$$

As $\delta \to 0$, these vanishes away from the singularities, leading to

$$\partial_t S_+ + \kappa S_+ = -\kappa E_O \delta(t + L) \tag{C.11}$$

$$\partial_t S_- - \kappa S_- = \kappa E_O \delta(t - L) \,, \tag{C.12}$$

where $E_O = h_O/\delta$. The delta functions are the shockwaves exiting and entering the AdS region. The signs here, and in particular the extra minus sign from the Hilbert transform,

ensure that there is a sensible solution for the Schwarzian, which is time-symmetric and goes to zero as $t \to \pm\infty$. The solution is

$$S_+ = \Theta(-t - L)\kappa E_0 e^{\kappa(t+L)}, \qquad S_- = \Theta(t - L)\kappa E_0 e^{\kappa(L-t)}. \qquad \text{(C.13)}$$

For $t > 0$, this is essentially the same solution as the evaporating black hole in [5], which had a shockwave produced by a joining quench rather than an operator insertion.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](#)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

[1] S.W. Hawking, *Breakdown of predictability in gravitational collapse*, *Phys. Rev.* **D 14** (1976) 2460 [INSPIRE].

[2] D.N. Page, *Information in black hole radiation*, *Phys. Rev. Lett.* **71** (1993) 3743 [hep-th/9306083] [INSPIRE].

[3] D.N. Page, *Time dependence of Hawking radiation entropy*, *JCAP* **09** (2013) 028 [arXiv:1301.4995] [INSPIRE].

[4] G. Penington, *Entanglement wedge reconstruction and the information paradox*, arXiv:1905.08255 [INSPIRE].

[5] A. Almheiri, N. Engelhardt, D. Marolf and H. Maxfield, *The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole*, *JHEP* **12** (2019) 063 [arXiv:1905.08762] [INSPIRE].

[6] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, *JHEP* **03** (2020) 149 [arXiv:1908.10996] [INSPIRE].

[7] T.G. Mertens, *Towards black hole evaporation in Jackiw-Teitelboim gravity*, *JHEP* **07** (2019) 097 [arXiv:1903.10485] [INSPIRE].

[8] C. Akers, S. Leichenauer and A. Levine, *Large breakdowns of entanglement wedge reconstruction*, *Phys. Rev.* **D 100** (2019) 126006 [arXiv:1908.03975] [INSPIRE].

[9] U. Moitra, S.K. Sake, S.P. Trivedi and V. Vishal, *Jackiw-Teitelboim model coupled to conformal matter in the semi-classical limit*, *JHEP* **04** (2020) 199 [arXiv:1908.08523] [INSPIRE].

[10] A. Almheiri, R. Mahajan and J.E. Santos, *Entanglement islands in higher dimensions*, arXiv:1911.09666 [INSPIRE].

[11] Z. Fu and D. Marolf, *Bag-of-gold spacetimes, Euclidean wormholes and inflation from domain walls in AdS/CFT*, *JHEP* **11** (2019) 040 [arXiv:1909.02505] [INSPIRE].

[12] P. Zhang, *Evaporation dynamics of the Sachdev-Ye-Kitaev model*, *Phys. Rev.* **B 100** (2019) 245104 [arXiv:1909.10637] [INSPIRE].

[13] C. Akers, N. Engelhardt and D. Harlow, *Simple holographic models of black hole evaporation*, arXiv:1910.00972 [INSPIRE].

[14] A. Almheiri, R. Mahajan and J. Maldacena, *Islands outside the horizon*, arXiv:1910.11077 [INSPIRE].

[15] M. Rozali et al., *Information radiation in BCFT models of black holes*, `arXiv:1910.12836` [ɪɴSPIRE].

[16] H.Z. Chen et al., *Information flow in black hole evaporation*, *JHEP* **03** (2020) 152 [`arXiv:1911.03402`] [ɪɴSPIRE].

[17] R. Bousso and M. Tomašević, *Unitarity from a smooth horizon?*, `arXiv:1911.06305` [ɪɴSPIRE].

[18] D.L. Jafferis and D.K. Kolchmeyer, *Entanglement entropy in Jackiw-Teitelboim gravity*, `arXiv:1911.10663` [ɪɴSPIRE].

[19] A. Blommaert, T.G. Mertens and H. Verschelde, *Eigenbranes in Jackiw-Teitelboim gravity*, `arXiv:1911.11603` [ɪɴSPIRE].

[20] S. Ryu and T. Takayanagi, *Holographic derivation of entanglement entropy from AdS/CFT*, *Phys. Rev. Lett.* **96** (2006) 181602 [`hep-th/0603001`] [ɪɴSPIRE].

[21] V.E. Hubeny, M. Rangamani and T. Takayanagi, *A covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [`arXiv:0705.0016`] [ɪɴSPIRE].

[22] T. Faulkner, A. Lewkowycz and J. Maldacena, *Quantum corrections to holographic entanglement entropy*, *JHEP* **11** (2013) 074 [`arXiv:1307.2892`] [ɪɴSPIRE].

[23] N. Engelhardt and A.C. Wall, *Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime*, *JHEP* **01** (2015) 073 [`arXiv:1408.3203`] [ɪɴSPIRE].

[24] S.D. Mathur, *What is the dual of two entangled CFTs?*, `arXiv:1402.6378` [ɪɴSPIRE].

[25] R. Jackiw, *Lower dimensional gravity*, *Nucl. Phys.* **B 252** (1985) 343 [ɪɴSPIRE].

[26] C. Teitelboim, *Gravitation and Hamiltonian structure in two space-time dimensions*, *Phys. Lett.* **126B** (1983) 41 [ɪɴSPIRE].

[27] A. Almheiri and J. Polchinski, *Models of AdS$_2$ backreaction and holography*, *JHEP* **11** (2015) 014 [`arXiv:1402.6334`] [ɪɴSPIRE].

[28] A. Lewkowycz and J. Maldacena, *Generalized gravitational entropy*, *JHEP* **08** (2013) 090 [`arXiv:1304.4926`] [ɪɴSPIRE].

[29] X. Dong, A. Lewkowycz and M. Rangamani, *Deriving covariant holographic entanglement*, *JHEP* **11** (2016) 028 [`arXiv:1607.07506`] [ɪɴSPIRE].

[30] X. Dong and A. Lewkowycz, *Entropy, extremality, Euclidean variations and the equations of motion*, *JHEP* **01** (2018) 081 [`arXiv:1705.08453`] [ɪɴSPIRE].

[31] G. Penington, S.H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, `arXiv:1911.11977` [ɪɴSPIRE].

[32] J.M. Maldacena, *Eternal black holes in Anti-de Sitter*, *JHEP* **04** (2003) 021 [`hep-th/0106112`] [ɪɴSPIRE].

[33] P. Saad, S.H. Shenker and D. Stanford, *A semiclassical ramp in SYK and in gravity*, `arXiv:1806.06840` [ɪɴSPIRE].

[34] P. Saad, S.H. Shenker and D. Stanford, *JT gravity as a matrix integral*, `arXiv:1903.11115` [ɪɴSPIRE].

[35] P. Saad, *Late time correlation functions, baby universes and ETH in JT gravity*, `arXiv:1910.10311` [ɪɴSPIRE].

[36] J.V. Rocha, *Evaporation of large black holes in AdS: coupling to the evaporon*, *JHEP* **08** (2008) 075 [arXiv:0804.0055] [INSPIRE].

[37] J. Engelsöy, T.G. Mertens and H. Verlinde, *An investigation of AdS_2 backreaction and holography*, *JHEP* **07** (2016) 139 [arXiv:1606.03438] [INSPIRE].

[38] L. Randall and R. Sundrum, *An alternative to compactification*, *Phys. Rev. Lett.* **83** (1999) 4690 [hep-th/9906064] [INSPIRE].

[39] A. Karch and L. Randall, *Locally localized gravity*, *JHEP* **05** (2001) 008 [hep-th/0011156] [INSPIRE].

[40] J. Maldacena, A. Milekhin and F. Popov, *Traversable wormholes in four dimensions*, arXiv:1807.04726 [INSPIRE].

[41] T.M. Fiola, J. Preskill, A. Strominger and S.P. Trivedi, *Black hole thermodynamics and information loss in two-dimensions*, *Phys. Rev.* **D 50** (1994) 3987 [hep-th/9403137] [INSPIRE].

[42] C.G. Callan Jr. and F. Wilczek, *On geometric entropy*, *Phys. Lett.* **B 333** (1994) 55 [hep-th/9401072] [INSPIRE].

[43] H. Casini and M. Huerta, *Entanglement entropy in free quantum field theory*, *J. Phys.* **A 42** (2009) 504007 [arXiv:0905.2562] [INSPIRE].

[44] X. Dong, *Holographic entanglement entropy for general higher derivative gravity*, *JHEP* **01** (2014) 044 [arXiv:1310.5713] [INSPIRE].

[45] E. Sharon and D. Mumford, *2D-shape analysis using conformal mapping*, *Int. J. Computer Vision* **70** (2006) 55.

[46] K. Jensen, *Chaos in AdS_2 holography*, *Phys. Rev. Lett.* **117** (2016) 111601 [arXiv:1605.06098] [INSPIRE].

[47] J. Maldacena, D. Stanford and Z. Yang, *Conformal symmetry and its breaking in two dimensional nearly Anti-de-Sitter space*, *PTEP* **2016** (2016) 12C104 [arXiv:1606.01857] [INSPIRE].

[48] P. Calabrese and J.L. Cardy, *Entanglement entropy and quantum field theory*, *J. Stat. Mech.* **0406** (2004) P06002 [hep-th/0405152] [INSPIRE].

[49] H. Casini, C.D. Fosco and M. Huerta, *Entanglement and alpha entropies for a massive Dirac field in two dimensions*, *J. Stat. Mech.* **0507** (2005) P07007 [cond-mat/0505563] [INSPIRE].

[50] P. Hayden and G. Penington, *Learning the alpha-bits of black holes*, *JHEP* **12** (2019) 007 [arXiv:1807.06041] [INSPIRE].

[51] P. Hayden and G. Penington, *Approximate quantum error correction revisited: introducing the alpha-bit*, arXiv:1706.09434.

[52] D.L. Jafferis, A. Lewkowycz, J. Maldacena and S.J. Suh, *Relative entropy equals bulk relative entropy*, *JHEP* **06** (2016) 004 [arXiv:1512.06431] [INSPIRE].

[53] X. Dong, D. Harlow and A.C. Wall, *Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality*, *Phys. Rev. Lett.* **117** (2016) 021601 [arXiv:1601.05416] [INSPIRE].

[54] A. Almheiri, X. Dong and D. Harlow, *Bulk locality and quantum error correction in AdS/CFT*, *JHEP* **04** (2015) 163 [arXiv:1411.7041] [INSPIRE].

[55] A. Strominger and C. Vafa, *Microscopic origin of the Bekenstein-Hawking entropy*, *Phys. Lett.* **B 379** (1996) 99 [hep-th/9601029] [ɪɴSPIRE].

[56] S.W. Hawking, *Quantum coherence down the wormhole*, *Phys. Lett.* **B 195** (1987) 337 [ɪɴSPIRE].

[57] G.V. Lavrelashvili, V.A. Rubakov and P.G. Tinyakov, *Disruption of quantum coherence upon a change in spatial topology in quantum gravity*, *JETP Lett.* **46** (1987) 167 [ɪɴSPIRE].

[58] S.B. Giddings and A. Strominger, *Axion induced topology change in quantum gravity and string theory*, *Nucl. Phys.* **B 306** (1988) 890 [ɪɴSPIRE].

[59] S.R. Coleman, *Black holes as red herrings: topological fluctuations and the loss of quantum coherence*, *Nucl. Phys.* **B 307** (1988) 867 [ɪɴSPIRE].

[60] S.B. Giddings and A. Strominger, *Loss of incoherence and determination of coupling constants in quantum gravity*, *Nucl. Phys.* **B 307** (1988) 854 [ɪɴSPIRE].

[61] J. Polchinski and A. Strominger, *A possible resolution of the black hole information puzzle*, *Phys. Rev.* **D 50** (1994) 7403 [hep-th/9407008] [ɪɴSPIRE].

[62] T. Hartman and J. Maldacena, *Time evolution of entanglement entropy from black hole interiors*, *JHEP* **05** (2013) 014 [arXiv:1303.1080] [ɪɴSPIRE].

[63] L. Susskind, *Entanglement is not enough*, *Fortsch. Phys.* **64** (2016) 49 [arXiv:1411.0690] [ɪɴSPIRE].