A Co-Boundary Proposal[1]

David Wakeham

The *AdS/CFT correspondence* [1] is the statement that some quantum field theories secretly encode theories of quantum gravity. In this thesis, we will use the correspondence to peer inside black holes and see how they evaporate.

1 Background

This introductory section recalls some basic facts about AdS/CFT, black holes, and entanglement. AdS/CFT tells us that d-dimensional conformal field theories (CFT_d) are dual to quantum gravity in spacetimes which are asymptotically saddle-like, i.e. which asymptotically approach anti-de Sitter space in one dimension higher (AdS_{d+1}), as in Fig. 1 (left). This is a duality of Hilbert spaces: states (including mixed states) in the CFT correspond to states in the Hilbert space of quantum gravity on AdS. In principle, AdS/CFT furnishes a non-perturbative *definition* of quantum gravity, for arbitrary states and CFTs. We focus, however, on *large-N CFTs*,[2] dual to classical gravity, and states whose bulk duals are described by perturbative quantum field theory on a curved background. We have to learn to walk before we can run.

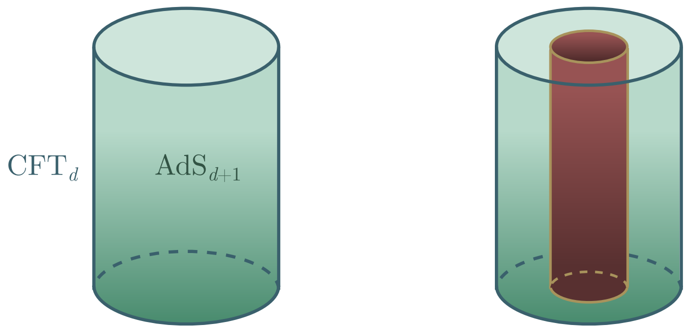

Figure 1. *Left.* The AdS/CFT correspondence on a cylinder. The exterior is the CFT, the interior is AdS. *Right.* An AdS-Schwarzschild black hole, dual to a thermal state.

Walk or run, we will eventually collide with black holes. Bardeen, Carter and Hawking [2] discovered that black holes obey laws analogous to those of classical thermodynamics. In AdS/CFT, this is simply because black holes are dual to thermal states of the CFT (Fig. 1, right). More precisely, consider the canonical ensemble at inverse temperature $\beta = 1/k_B T$, defined by the density matrix $\rho_\beta := Z[\beta]^{-1} e^{-\beta \hat{H}}$. Above the Hawking-Page

[1] *Explanatory note.* Here, the "co-boundaries" refer to AdS/BCFT, to be explained below, and the pun is on the "no-boundary proposal".

[2] Morally, N is the number of local degrees of freedom in the CFT. Immorally, it is some parameter labelling a family of theories such that correlators factorize to leading order in N.

transition temperature [3], the dual geometry is the exterior of an AdS-Schwarzschild black hole at temperature β [4].

Although the exterior is fixed and classical, the interior remains mysterious. This is easily explained. The thermal state is just the canonical probability distribution $p_E \propto e^{-\beta E}$ over all eigenstates $|E\rangle$, peaking around some saddle-point energy $E(\beta)$. There will be many eigenstates around this energy, with dramatically different interior geometries, or no classically describable interior at all. It may seem odd that these non-classical states around $E(\beta)$ do not spoil the exterior geometry. This is related to the famous *Eigenstate Thermalization Hypothesis (ETH)* [5, 6], which conjectures that in chaotic systems, high energy states around $E(\beta)$ give the same answers to simple questions, with differences suppressed by system size. Evidently, simple questions live in the exterior.

Since the mixedness of the canonical ensemble seems to obscure the interior geometry, purifying ρ_β will perhaps reveal it. The simplest method of purification is the *thermofield double (TFD)*, introduced by Israel [7]. Let CFT_1 denote our original CFT, and CFT_2 a second copy of the same system. The TFD is the thermally entangled pure state

$$|\mathrm{TFD}(\beta)\rangle := \frac{1}{Z[\beta]^{-1/2}} \sum_E e^{-\beta H/2} |E_1\rangle |E_2\rangle \ , \tag{1.1}$$

where $|E_i\rangle$ denotes energy eigenstates of system CFT_i. If we partially trace out CFT_2, we recover the thermal state, $\rho_\beta = \mathrm{Tr}_2 |\mathrm{TFD}(\beta)\rangle\langle \mathrm{TFD}(\beta)|$, which describes the AdS-Schwarzschild exterior. Of course, the same statement is true if we trace out CFT_1, so the bulk should include *two* AdS-Schwarzschild exteriors, asymptotic to the two CFTs. The simplest geometry with two black hole exteriors is a wormhole, i.e. the *maximally extended AdS-Schwarzschild solution.* Maldacena gave a formal argument for this identification in [8] (Fig. 2, left). In some mysterious fashion, entanglement seems to build the spacetime of the interior [9].

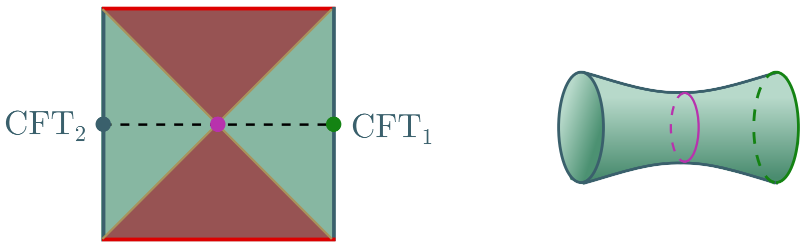

Figure 2. *Left.* The wormhole dual to the TFD state, with sphere $\mathbb{S}^{d-1}$ suppressed. *Right.* The embedding geometry of the dotted timeslice, with the sphere schematically shown as a circle.

According to the Bekenstein-Hawking law [10, 11], the entropy of the black hole should be identified with its horizon area $\mathcal{A}_{\mathrm{hor}}$ according to $S = \mathcal{A}_{\mathrm{hor}}/4G$, where G is Newton's constant. But in the wormhole geometry, the horizon is simply the area of the "throat" joining the two exterior regions, i.e. the *minimal surface* between them (Fig. 2, right). Microscopically, the thermal entropy of ρ_β is equal to the *entanglement entropy* between the two CFTs, $S[\mathrm{CFT}_2] := -\mathrm{Tr}[\rho_\beta \log \rho_\beta]$. The entanglement entropy is therefore computed by the area of a bulk minimal surface. This observation is dramatically generalized by the

Ryu-Takayanagi (RT) formula [12, 13], depicted in Fig. 3 (left). This proposes that, for any subregion A of a CFT in state $|\Psi\rangle$,[3] the entanglement entropy of the reduced density $\rho_A := \mathrm{Tr}_{\bar{A}}|\Psi\rangle\langle\Psi|$ is computed by the bulk minimal surface $\mathcal{X}_A$ anchored at A, with

$$S[A] = \frac{\mathcal{A}[\mathcal{X}_A]}{4G} + O(G^0) \ . \tag{1.2}$$

Bekenstein-Hawking is the special case where the "subregion" is an entire CFT!

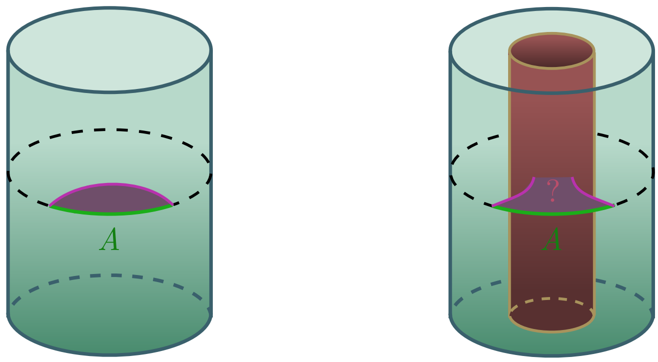

Figure 3. *Left.* The RT formula for a green boundary region A. The minimal surface $\mathcal{X}$ is fuschia, and the entanglement wedge Ξ_A mauve. *Right.* An RT surface probing behind the horizon.

The TFD is an entangled state of two CFTs with a known classical interior. Perhaps, via the RT formula (1.2), we can use entanglement to look inside a *single-sided black hole*, dual to a pure state on a single CFT (Fig. 3, right). The idea would be to identify a high-energy state $|E\rangle$ with some interior geometry, then find a boundary region A whose minimal surface $\mathcal{X}_A$ ventures inside the horizon. If we could find such a state and minimal surface, we would indeed be granted access to behind-the-horizon physics, since AdS/CFT lore [15, 16] tells us that the density matrix ρ_A encodes physics in the *entanglement wedge* Ξ_A, the region[4] between A and $\mathcal{X}_A$. Sadly, these efforts are doomed: the horizon is an impassable barrier to extremal surfaces anchored on the boundary [17]. Another strategy is needed.

2 Boundary state black holes

Waste not, want not. The TFD (1.1) has a perfectly good interior geometry, and we can use it to make a one-sided geometry fit for purpose. The first step is to transform (1.1) into a state of CFT_1, our candidate single-sided black hole. This is most easily achieved by projecting CFT_2 onto a specific state $|B_2\rangle$:

$$\langle B_2|\mathrm{TFD}(\beta)\rangle = \langle B_2|e^{-\beta H/2}|\mathrm{EPR}\rangle =: |B_1(\beta)\rangle \ , \tag{2.1}$$

[3] Of course, the state should be semiclassical so that an area can be computed. The original RT formula applies only to static geometries. Hubeny, Rangamani and Takanagi generalized (1.2) to time-dependent geometries, in which $\mathcal{X}_A$ is a minimal *extremal* surface instead of a minimal surface simpliciter [14].

[4] Technically, its bulk domain of dependence.

where $|\mathrm{EPR}\rangle := \sum_E |E_1\rangle|E_2\rangle$ is the (non-normalizable) analogue of the EPR state. More importantly, it is Choi–Jamiołkowski dual to the identity map [18].[5] Thus, correlators can equivalently be computed in the state $e^{-\beta H/2}|B_2\rangle$:

$$\langle B_1(\beta)|\mathcal{O}_1 \cdots \mathcal{O}_n|B_1(\beta)\rangle = \langle B_2|e^{-\beta H/2}\mathcal{O}_1 \cdots \mathcal{O}_n e^{-\beta H/2}|B_2\rangle \ , \tag{2.2}$$

as demonstrated graphically in Fig. 4. This is precisely the *imaginary time formalism* used in quantum quenches [19]. When the $|B_2\rangle$ is a conformally invariant[6] *boundary state* [21], then (2.2) thermalizes at inverse temperature 2β, in the sense that the correlators give canonically thermal answers [22].

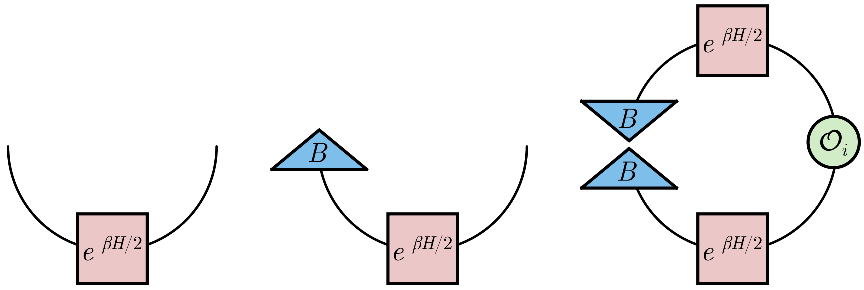

Figure 4. *Left.* A cartoon of the thermofield double state. *Middle.* Projecting the TFD onto the boundary state $|B(\beta)\rangle$. *Right.* Computing correlators in the boundary state.

It seems that $|B(\beta)\rangle$ describes a black hole at temperature 2β, with boundary $|B\rangle$ at imaginary time $\tau = -\beta/2$.[7] This by itself does not guarantee a classical interior geometry, but the natural appearance of boundary states is fortuitous. A CFT with boundary and conformally invariant boundary conditions is called a *BCFT*. The *AdS/BCFT dictionary* [24, 25] proposes that BCFTs are dual to AdS space cut off by a codimension-1 surface called a *brane*. The position of the brane can be determined from boundary data by judiciously modifying the RT formula (1.2). For instance, in a $1+1$-dimensional BCFT on $\{(x,t) \in \mathbb{R}^2 : x \geq 0\}$, the ground state entanglement of the interval $A := [0, L]$ is given by

$$S_A = \frac{c}{6}\log\left(\frac{2L}{\epsilon}\right) + g_B \ , \tag{2.3}$$

where c is the central charge, ϵ is a UV regulator, and $g_B := \log\langle B|0\rangle$ an L-independent constant called the *boundary entropy* [26].

To reproduce this universal result from the RT formula, we must allow the minimal geodesic computing S_A to end on the brane itself. From symmetry, one then learns that the brane is a surface of constant tension (i.e. extrinsic curvature) determined by g_B. In higher dimensions, the story is similar, with universal results for the entanglement entropy of a boundary-centred half-sphere reproduced by minimal surfaces ending on a constant tension brane [27]. The appropriate generalization of (1.2) simply permits $\mathcal{X}_A$ to end on the brane $\mathcal{B}$, an *inclusion condition*[8] we can succinctly write as $\partial\mathcal{X}_A \backslash \mathcal{B} = \partial A$.

[5] Less mysteriously, we identify $\sum_E |E\rangle|E\rangle \to \sum_E |E\rangle\langle E| = I$, using a CPT map on the second factor.

[6] More carefully, it obeys the *kinematic* constraint of preserving an $\mathrm{SO}(d,1)$ subgroup of the full conformal symmetry, and a *dynamic* channel duality constraint called the "boundary bootstrap" [20].

[7] In case of lingering doubt, there is an independent argument from gravitational physics in [23].

[8] So called because the brane is included in the bulk geometry for the purposes of the RT formula [28]. In

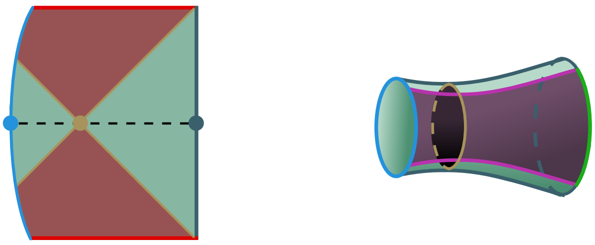

Figure 5. *Left.* Projecting the TFD onto a boundary state of CFT_2. *Right.* The embedding geometry at fixed time is a wormhole cut off by a brane. The RT surface for a large subregion on the right (green) passes through the horizon and ends on the brane.

In [23], we combine these ingredients to construct pure state black holes $|B(\beta)\rangle$ with geometric interiors, and find boundary regions A which encode behind-the-horizon physics (Fig. 5). By the AdS/BCFT dictionary, $|B(\beta)\rangle$ is the wormhole geometry of (1.1), terminating on a constant tension brane. If boundary entropy dictates that the extrinsic curvature is negative, it sits behind the black hole horizon. Finally, one can choose a sufficiently large boundary region A that the RT surface ends on the brane.[9]

3 Entanglement entropy in 2d BCFTs

Using the gravitational ansatz of AdS/BCFT and the "inclusive" version of the RT formula, we seem to be able to peer inside black holes. But how do we know when there is a classical bulk geometry? And even if there is a geometry, is entanglement really captured by our inclusive RT formula? These are questions about the microscopics of quantum gravity. In [29], we answer both in the tractable but nontrivial setting of a 2d BCFT. The basic observation is that the entanglement entropy can be obtained as a limit:

$$S_A = \lim_{n\to 1} S_A^{(n)} := \lim_{n\to 1} \frac{1}{n-1}\mathrm{Tr}[\rho_A^n]\,. \tag{3.1}$$

The $S_A^{(n)}$ are *n-Rényi entropies*, and the trace $\mathrm{Tr}[\rho_A^n] \sim Z^{(n)}$ is the partition function of the replica geometry $\mathcal{R}_n$, consisting of n copies of the system cyclically identified along the cuts A. The answer is analytically continued to arbitrary n, and the limit $n \to 1$ taken, to obtain the entanglement entropy.

Evaluating these replica partition functions explicitly is impossible in higher dimensions. But in a 2d CFT, the replica geometry can be "mocked up" by local operators called *twists* Φ_n (Fig. 6, left), implementing the cyclic boundary conditions [19, 30]. The Rényi entropy is then just a correlator of these local operators. Hartman [31] performs the replica calculation in a 2d CFT without boundary, explicitly matching the predictions of the RT formula given the assumption of large central charge and a sparse spectrum of primary operators. [29] repeats these steps for a 2d BCFT, with appropriate modifications, and microscopically verifies the gravitational results.

---

fact, in the corresponding supergravity solutions, we expect the brane to arise from a smoothly degenerating internal dimension, so that it genuinely is part of the geometry.

[9]For some regime of time and tension.

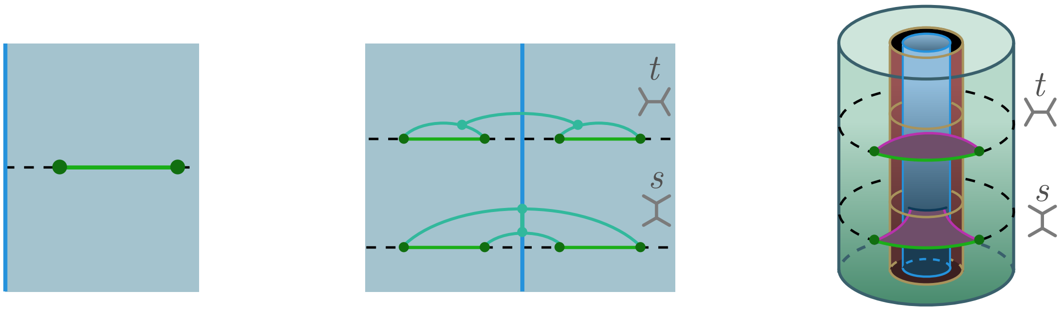

Figure 6. *Left.* The entanglement entropy of an interval is computed by a correlator of twists, dark green. *Middle.* Doubling the twists in an auxiliary CFT. There are two OPE Feynman diagrams. *Right.* The corresponding RT surfaces in the bulk. The brane is a surface behind the horizon.

In more detail, we first note that the correlator (2.2) can be evaluated in imaginary time on a strip of height β and infinite width.[10] This is the image of the half-plane under a conformal transformation, with a boundary at $x = 0$ mapping to the edges of the strip; by symmetry, we can work in this simpler geometry. Consider a collection of k intervals $A = \bigcup_i [x_{2i}, x_{2i+1}]$ on the half-plane. The n-Rényi entropy can be computed using a *doubling trick* (Fig. 6, middle), based on the observation that the representation theory of the BCFT is equivalent to doubling the insertions and placing them in a regular CFT. In this auxiliary CFT, we define the "doubled" interval

$$-A \cup A = \bigcup_i [x_{2i}, x_{2i+1}] \cup [-x_{2i+1}, -x_{2i}].$$

The usual CFT replica calculation gives

$$e^{(1-n)S_A^{(n)}} = \left\langle \prod_i \Phi_n(x_{2i})\bar{\Phi}_n(x_{2i+1})\Phi_n(-x_{2i+1})\bar{\Phi}_n(-x_{2i}) \right\rangle, \tag{3.2}$$

with the boundary entropy encoded into the normalization of the twists [33].

For large central charge, $c \to \infty$, and a gapped spectrum of bulk and boundary operators, this correlator is dominated by the exchange of the identity operator, which we can view as a virtual particle running in a cubic graph which joins insertions, i.e. a Feynman diagram for the operator product expansion (Fig. 6, middle). There are different diagrams, but from a large-c saddle-point expansion, one graph will dominate. In the bulk (Fig. 6, right), this corresponds to the minimal length geodesic pairing of endpoints with each other or the brane, which is precisely our modified RT formula in the limit $n \to 1$. This is a strong consistency check in the following sense. We start by identifying the bulk dual which correctly reproduces the *universal* result for the entanglement entropy of a half-interval via the RT formula. The RT formula then makes some non-universal predictions about entanglement entropy. We find that, for some set of spectral conditions, we can reproduce these non-universal predictions. Our procedure thus determines the quantum gravity microscopics self-consistently encoded by our choice of bulk geometry.

[10]Technically, the CFT is on a spatial circle, but we decompactify it and invoke large-N volume independence [32]. In Fig. 6 (right), we have compactified again.

4 Information radiation

This construction can be brought to bear on a related puzzle of black hole physics. To set the scene, in 1975, Hawking [34] discovered that black holes are not only formally thermal, but emit a blackbody spectrum of *Hawking radiation* from the near-horizon region. Over time, they can dissipate their energy into the environment and disappear. The INFORMATION PARADOX is that once the black hole has gone, only thermal radiation remains; apparently, it takes its secrets to the grave. This destruction of secrets is irreversible, and hence violates *unitarity*, one the basic principles of quantum mechanics.

In more detail, Hawking showed that the total entropy of black hole and radiation increases, in accord with the generalized second law proposed by Bekenstein [35]:

$$\frac{dS_{\text{gen}}}{dt} \geq 0, \quad S_{\text{gen}} = \frac{\mathcal{A}_{\text{hor}}}{4G} + S_{\text{ext}}, \tag{4.1}$$

where $\mathcal{A}_{\text{hor}}$ is the horizon area and S_{ext} is the entropy of matter in the black hole exterior. This is called the *generalized* or *coarse-grained entropy*. The *fine-grained entropy* is the microscopic entanglement entropy of black hole system, $S[\text{BH}] = -\text{Tr}[\rho_{\text{BH}} \log \rho_{\text{BH}}]$. If evaporation is a unitary, highly chaotic reorganization of these degrees of freedom into the radiation system, then $S[\text{BH}]$ should go up as it becomes entangled with the radiation, then down again as the system shrinks, describing the *Page curve* [36]. The Information Paradox can be restated as a homework problem: find a way to compute fine-grained entropy which gives the Page curve and not the monotonically increasing coarse-grained answer. For bonus marks, justify it and explain why Hawking was wrong.

In principle, AdS/CFT does our homework for us, since black holes are dual to thermal states in a closed, manifestly unitary system. There are two problems with this purported resolution. First, AdS black holes do not evaporate, but are in equilibrium with the Hawking radiation bouncing back in from the boundary; second, even if the black hole could be made to evaporate, the RT formula will not give a unitary answer![11]

In [38, 39], both problems are elegantly resolved. To make black holes evaporate, AdS/CFT is dunked in a flat-space bath, liberating the quanta from the prison of reflecting boundary conditions. The fine-grained entropy is calculated using an older proposal of Engelhardt and Wall [40]. This hacks generalized entropy (4.1) into an entanglement functional the same way Ryu and Takayanagi hacked the Bekenstein-Hawking term. The *quantum extremal surface* is the minimal extremum of

$$S_{\text{gen}}[A] = \frac{\mathcal{A}[\mathcal{X}_A]}{4G} + S_{\text{bulk}}[\Xi_A], \tag{4.2}$$

where $S_{\text{bulk}}[\Xi_A]$ is the entanglement entropy of bulk fields in the entanglement wedge Ξ_A. The quantum extremal surfaces for boundary regions in this shrinking black hole reproduce the Page curve and other expected features of unitary evaporation. Thus, information is saved due to bulk discounts in entanglement entropy.

[11] A more sophisticated guess is the FLM formula [37], consisting of minimal surface plus bulk entanglement in the entanglement wedge. For an evaporating black hole, this just gives the generalized entropy up to small corrections. Note that FLM is different from (4.2) since it extremizes the first term only.

This does not entirely resolve the paradox. The question remains: how does the information get out when it is stuck behind a horizon? In toy models, the escape is beautifully geometrized by an emergent higher dimension.[12] In the AMMZ model [42], a 2d black hole is dual to a 1d system, which is itself the boundary of a 2d holographic BCFT. The BCFT thus provides a flat-space bath and 3d bulk spacetime, as in Fig. 7. Quantum extremal surfaces are replaced by ordinary RT surfaces in the 3d bulk which can end on the lower-dimensional black hole or a "Cardy brane", introduced by hand to model the evaporation. When the minimal surface for a half-infinite interval (Fig. 7, middle) hits the interior, it forms an "island", and provides a channel for the information to escape.

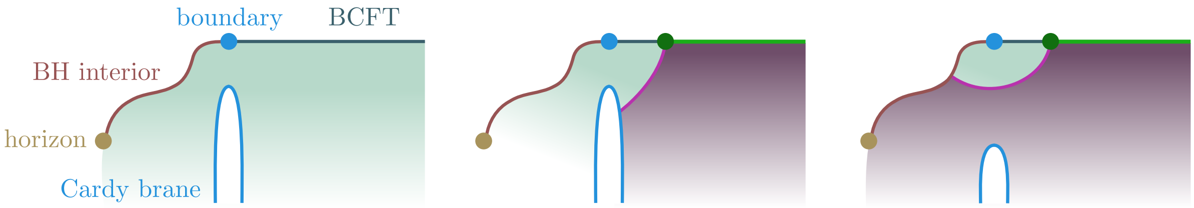

Figure 7. *Left.* Time slice of AMMZ [42]. *Middle.* At early times, the Cardy brane obstructs the minimal surface for a half-infinite region of the BCFT, so it does not include any of the black hole interior. *Right.* Like a lowering sluice, the Cardy brane opens up a channel to the interior.

In [43] (in progress when AMMZ was published), we repurposed the boundary state black hole to give a simpler setup with explicit microscopic control. The basic idea is to map the half-plane BCFT to the plane with a disk removed. Analytically continuing[13] to real time gives a boundary theory of two half-lines, accelerating away from each other on Rindler trajectories, and joined in the 3d bulk by a brane (Fig. 8, left). The brane itself has causal horizons and can therefore be regarded as a black hole.

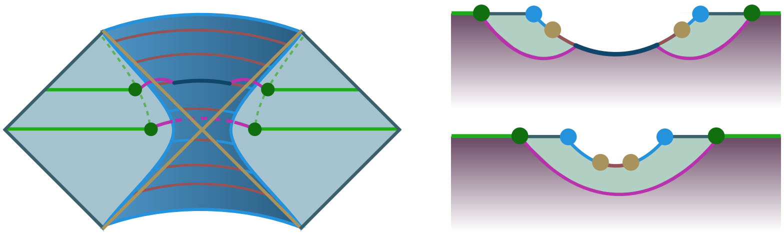

Figure 8. *Left.* The bulk spacetime of [43], analytically continued from the plane minus a disk. *Right.* The inevitable formation of an island, with the same coloring scheme as Fig. 7.

Instead of computing minimal surfaces for a single half-line, we have two *symmetric* half-lines, $A = (-\infty, -x] \cup [x, \infty]$, and can compute their entanglement entropy as a function of time. Although these minimal surfaces initially skirt around the interior (Fig. 8, bottom right), they inevitably transition to form an island (Fig. 8, top right). Once again, the information sails out through an emergent dimension.

[12]This serves as a concrete realization of Susskind and Maldacena's ER=EPR conjecture [41].

[13]This BCFT has Cartesian coordinates $z = x + i\tau$, with disk at $|z| = R$. Analytic continuation means performing the inverse Wick rotation $t = -i\tau$ from the $\tau = 0$ surface, both in the boundary and the bulk.

5 Wormholes, averages and eigenstates

Our toy models teach us that islands restore unitarity. But they don't tell why Hawking was wrong and we are right. We might expect this to depend on the ultraviolet details of quantum gravity, and the original setting of AdS/CFT, where we have a stringy UV completion by fiat, seems to reinforce this. But the island construction can be generalized to black holes in *flat space* [44]. For better or worse, we need not invoke the arcane secrets of quantum gravity to understand how a black hole evaporates.

Instead, we must probe the secrets of the Euclidean path integral. To compute entropy in a canonical state $S = -k_B \partial_T (T \log Z[\beta])$, we first compute the partition function $Z[\beta]$. In gravity, the partition function is a sum over geometries subject to periodic boundary conditions in imaginary time. The most elegant version of Hawking's argument is due to Gibbons and Hawking [45], who showed that in the semiclassical limit, the partition function is dominated by a saddle-point "disk" geometry with action I_{disk}, and an exactly thermal spectrum for the matter fields:

$$Z_{\text{grav}}[\beta] = \int_\beta \mathcal{D}g\, e^{-S[g]} \to e^{-I_{\text{disk}}} Z_{\text{matter}}[\beta]. \tag{5.1}$$

In this canonically thermal sea, there are no islands in sight. The resulting entropy is the coarse-grained entropy $S_{\text{gen}} = -k_B \partial_T (T \log Z_{\text{grav}}[\beta])$, which always increases in accord with the generalized second law.

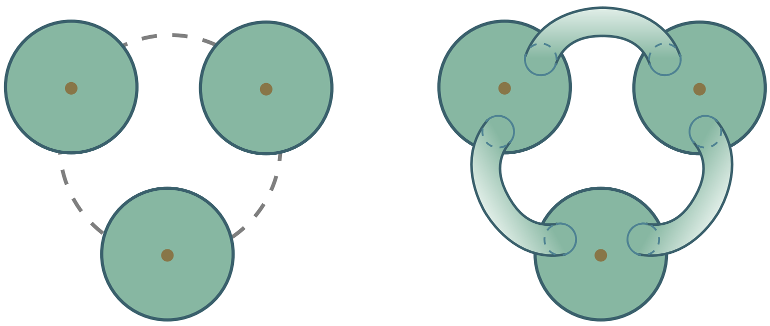

Figure 9. *Left.* The disconnected saddle for $n = 3$, reproducing Hawking's result. *Right.* A connected Euclidean wormhole. If these wormholes survive as $n \to 1$, they can produce islands.

But we have seen a different prescription for fine-grained entropy. The replica trick (3.1) instructs us to take the $n \to 1$ limit of the partition function $Z^{(n)}_{\text{grav}}[\beta]$ on n replicas. If the geometries are disconnected, the partition function factorizes as $Z^{(n)}_{\text{grav}}[\beta] = Z^n_{\text{grav}}[\beta]$, and we recover Hawking's result as $n \to 1$. The landmark papers [46, 47] point out that we should not only sum over geometries on each copy, but also the *topologies* connecting them. We provide an example for $n = 3$ in Fig. 9. Islands form when these connected contributions survive and dominate Hawking's saddle in the $n \to 1$ limit.

The connected topologies are called *Euclidean* or *replica wormholes*, and they pose a curious factorization puzzle. In AdS/CFT, the gravity partition function can be dually calculated via a generating functional on the CFT side [48, 49]. For n copies, we simply have

n identical and disconnected CFTs, and the generating functional trivially factorizes. How is our putative sum over topologies consistent with this? One possibility is that gravity secretly computes an *ensemble average*. To illustrate, consider states $|i\rangle, |j\rangle$ which are orthogonal in the effective gravitational theory, $\langle i|j\rangle \approx \delta_{ij}$, but also have small corrections R_{ij} the gravity path integral cannot see:

$$\langle i|j\rangle = \delta_{ij} + e^{-S/2} R_{ij}. \tag{5.2}$$

If these states are considered part of a statistical ensemble, then R_{ij} is a random matrix which vanishes on average, $\overline{R_{ij}} = 0$, but could have non-zero variance $\overline{R_{ij}R_{ji}} = \sigma^2$. Then averaging over states not only agrees with the path integral in a single copy, $\overline{\langle i|j\rangle} = \delta_{ij}$, but gives off-diagonal contributions for two copies, e.g.

$$\overline{|\langle i|j\rangle|^2} = \overline{(\delta_{ij} + R_{ij})(\delta_{ji} + R_{ji})} = \delta_{ij} + e^{-S}\sigma^2. \tag{5.3}$$

The first term is Hawking's saddle, and $e^{-S}\sigma^2$ a replica wormhole [46].

Unfortunately, this ensemble averaged theory is not unitary, the very property we sought to preserve. In [50], we point out a simple mechanism for getting ensemble-averaged answers without sacrificing unitarity: work with Haar-typical states[14] in a microcanonical energy band, similar in spirit to the ETH. These are pure states of a unitary theory, but they *self-average* in the appropriate way to give connected contributions as in (5.3). Our diagrammatic techniques apply to *any* suitably chaotic theory, so replica wormholes play a role not only in evaporating a black hole, but boiling a kettle or lasing a heavy nucleus.

6 Future directions

We finish by sketching some ongoing projects and directions for future research. We also note that the papers [51–53] are unlikely to appear in the thesis.

6.1 Looking for a bulk brane

Proving the RT formula for a 2d BCFT [29] shows that the bulk tells a self-consistent story about entanglement. A natural question is whether the story remains self-consistent when we add matter, i.e. perturbative scalar fields. Utilizing CFT bootstrap techniques *à la* [20, 54, 55], it appears that holographic BCFTs are much more finely tuned than their unboundaried counterparts, requiring independent constraints for each matter field to ensure a sensible bulk spacetime. This is likely to be completed by early 2021.

6.2 Quantum tasks and islands

The emergent spacetime of AdS/CFT provides an apparently new set of resources for performing quantum tasks. Duality means these are simply the old resources in disguise,

[14] Let $\mathcal{H}_E$ be the Hilbert space of the microcanonical band. If $\mathcal{G} = \mathrm{U}(\mathcal{H}_E)$ is the group of unitary operators on the band, the uniform or group-invariant probability measure $P_\mathcal{G}$ is called the *Haar measure*. Haar-typical states $U|\psi_0\rangle$ are generated by applying a unitary U, chosen according to $P_\mathcal{G}$, to a reference state $|\psi_0\rangle \in \mathcal{H}_E$.

motivating a task-based interpretation of holography [56]. In this vein, the *connected wedge theorem* [57] uses the focusing theorem from classical gravity to prove that a "bulk shortcut" for a position-based version of the BB84 protocol [58] can only be performed when the entanglement wedge for the relevant boundary regions is connected.

By modifying the task to incorporate boundary information, we have been able to prove a version of the connected wedge theorem for BCFTs. The time-reversed protocol can be interpreted in terms of island formation, rescuing information stranded in a Rindler black hole on the brane. The paper should appear later this year.

6.3 Pseudorandomness and the ETH

In [50], we took the viewpoint that the ETH is the statement that high energy eigenstates of a chaotic theory are typical. But this typicality only holds with respect to a set of "simple" low-point operators; in the context of black holes, these are the operators which probe the exterior. Thus, high-energy eigenstates are only *pseudorandom*, and it is desirable to give a characterization of this pseudorandomness, both in terms of states and operators. This may connect to the appearance of pseudorandomness in wormhole growth [59] and the complexity of decoding Hawking radiation [60].

A related question is how to build a gravitational effective theory using the typical states of [50]. Bounds on near-orthogonal packings of vectors lead to stronger constraints on the overlap matrix R_{ij} from (5.2), and hence implications for effectively-averaged replica wormholes. In this setting, various ambiguities in the replica prescription (see e.g. [61]), related to subtleties of averaging, might be fruitfully addressed. This project is in its preliminary stages.

6.4 A boundary bounty

A plethora of open questions remain about AdS/BCFT. In [23, 43], we claimed to access behind-the-horizon when the entanglement wedge Ξ_A ends on the brane. But entanglement wedge reconstruction [16] depends on equivalence of bulk and boundary modular flow [62], which in turn depends on the FLM formula [37], none of which have been shown in the context of AdS/BCFT. Modular flow in a 2d BCFT, and the special role of boundary states in algebraic QFT, were discussed in [63], and have yet to be exploited in the holographic context.

A related issue is how the quantum error-correcting properties of AdS/CFT [64, 65] are modified in the context of AdS/BCFT. Progress was initiated in [66], which studied a coarse notion of error correction for boundary state black holes, but a finer-grained approach which deals with spatial subregions is wanting. So there is plenty to do!

References

[1] J. M. Maldacena, *The Large N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113 [hep-th/9711200].

[2] J. M. Bardeen, B. Carter and S. W. Hawking, *The four laws of black hole mechanics*, *Comm. Math. Phys.* **31** (1973) 161.

[3] S. W. Hawking and D. N. Page, *Thermodynamics of Black Holes in anti-De Sitter Space*, *Commun. Math. Phys.* **87** (1983) 577.

[4] E. Witten, *Anti-de Sitter space, thermal phase transition, and confinement in gauge theories*, *Adv. Theor. Math. Phys.* **2** (1998) 505 [hep-th/9803131].

[5] J. M. Deutsch, *Quantum statistical mechanics in a closed system*, *Phys. Rev. A* **43** (1991) 2046.

[6] M. Srednicki, *Chaos and quantum thermalization*, *Phys. Rev. E* **50** (1994) 888.

[7] W. Israel, *Thermo field dynamics of black holes*, *Phys. Lett.* **A57** (1976) 107.

[8] J. M. Maldacena, *Eternal Black Holes in AdS*, *Journal of High Energy Physics* **2003** (2001) 021+ [0106112].

[9] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [1005.3035].

[10] S. W. Hawking, *Gravitational radiation from colliding black holes*, *Phys. Rev. Lett.* **26** (1971) 1344.

[11] J. D. Bekenstein, *Black holes and entropy*, *Phys. Rev. D* **7** (1973) 2333.

[12] S. Ryu and T. Takayanagi, *Aspects of Holographic Entanglement Entropy*, *JHEP* **08** (2006) 045 [hep-th/0605073].

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

[14] V. E. Hubeny, M. Rangamani and T. Takayanagi, *A Covariant holographic entanglement entropy proposal*, *JHEP* **07** (2007) 062 [0705.0016].

[15] B. Czech, J. L. Karczmarek, F. Nogueira and M. Van Raamsdonk, *Rindler Quantum Gravity*, *Class. Quant. Grav.* **29** (2012) 235025 [1206.1323].

[16] 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 [1601.05416].

[17] N. Engelhardt and A. C. Wall, *Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime*, *JHEP* **01** (2015) 073 [1408.3203].

[18] A. Jamiołkowski, *Linear transformations which preserve trace and positive semidefiniteness of operators*, *Reports on Mathematical Physics* **3** (1972) 275 .

[19] P. Calabrese and J. Cardy, *Quantum quenches in 1 + 1 dimensional conformal field theories*, *J. Stat. Mech.* **1606** (2016) 064003 [1603.02889].

[20] P. Liendo, L. Rastelli and B. C. Van Rees, *The bootstrap program for boundary CFT d*, *Journal of High Energy Physics* **2013** (2013) [1210.4258].

[21] J. L. Cardy, *Conformal Invariance and Surface Critical Behavior*, *Nucl. Phys.* **B240** (1984) 514.

[22] J. Cardy, *Quantum Quenches to a Critical Point in One Dimension: some further results*, *J. Stat. Mech.* **1602** (2016) 023103 [1507.07266].

[23] S. Cooper, M. Rozali, B. Swingle, M. Van Raamsdonk, C. Waddell and D. Wakeham, *Black hole microstate cosmology*, *Journal of High Energy Physics* **2019** (2019) .

[24] M. Fujita, T. Takayanagi and E. Tonni, *Aspects of AdS/BCFT*, *JHEP* **11** (2011) 043 [1108.5152].

[25] T. Takayanagi, *Holographic Dual of BCFT*, *Phys. Rev. Lett.* **107** (2011) 101602 [1105.5165].

[26] I. Affleck and A. W. W. Ludwig, *Universal noninteger 'ground state degeneracy' in critical quantum systems*, *Phys. Rev. Lett.* **67** (1991) 161.

[27] K. Jensen and A. O'Bannon, *Holography, Entanglement Entropy, and Conformal Field Theories with Boundaries or Defects*, *Phys. Rev.* **D88** (2013) 106006 [1309.4523].

[28] M. Headrick, *General properties of holographic entanglement entropy*, *JHEP* **03** (2014) 085 [1312.6717].

[29] J. Sully, M. V. Raamsdonk and D. Wakeham, *BCFT entanglement entropy at large central charge and the black hole interior*, 2004.13088.

[30] P. Calabrese and J. Cardy, *Entanglement entropy and conformal field theory*, *J. Phys.* **A42** (2009) 504005 [0905.4013].

[31] T. Hartman, *Entanglement Entropy at Large Central Charge*, 1303.6955.

[32] E. Shaghoulian, *Emergent gravity from Eguchi-Kawai reduction*, *JHEP* **03** (2017) 011 [1611.04189].

[33] J. Cardy and E. Tonni, *Entanglement hamiltonians in two-dimensional conformal field theory*, *J. Stat. Mech.* **1612** (2016) 123103 [1608.01283].

[34] S. W. Hawking, *Particle creation by black holes*, *Comm. Math. Phys.* **43** (1975) 199.

[35] J. D. Bekenstein, *Generalized second law of thermodynamics in black hole physics*, *Phys. Rev. D* **9** (1974) 3292.

[36] D. N. Page, *Average entropy of a subsystem*, *Physical Review Letters* **71** (1993) 1291–1294.

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

[38] G. Penington, *Entanglement Wedge Reconstruction and the Information Paradox*, 1905.08255.

[39] 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 [1905.08762].

[40] N. Engelhardt and A. C. Wall, *Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime*, *Journal of High Energy Physics* **2015** (2015) .

[41] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortsch. Phys.* **61** (2013) 781 [1306.0533].

[42] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, *The Page curve of Hawking radiation from semiclassical geometry*, 1908.10996.

[43] M. Rozali, J. Sully, M. Van Raamsdonk, C. Waddell and D. Wakeham, *Information radiation in BCFT models of black holes*, *JHEP* **05** (2020) 004 [1910.12836].

[44] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *The entropy of Hawking radiation*, 2006.06872.

[45] G. W. Gibbons and S. W. Hawking, *Cosmological event horizons, thermodynamics, and particle creation*, *Phys. Rev. D* **15** (1977) 2738.

[46] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, *Replica wormholes and the black hole interior*, 1911.11977.

[47] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, *Replica Wormholes and the Entropy of Hawking Radiation*, 1911.12333.

[48] S. S. Gubser, I. R. Klebanov and A. M. Polyakov, *Gauge theory correlators from noncritical string theory*, *Phys. Lett.* **B428** (1998) 105 [hep-th/9802109].

[49] E. Witten, *Anti-de Sitter space and holography*, *Adv. Theor. Math. Phys.* **2** (1998) 253 [hep-th/9802150].

[50] J. Pollack, M. Rozali, J. Sully and D. Wakeham, *Eigenstate thermalization and disorder averaging in gravity*, *Physical Review Letters* **125** (2020) .

[51] S. Cooper, D. Neuenfeld, M. Rozali and D. Wakeham, *Brane dynamics from the first law of entanglement*, *Journal of High Energy Physics* **2020** (2020) .

[52] N. Bao, J. Pollack, D. Wakeham and E. Wildenhain, *Microstate distinguishability, quantum complexity, and the Eigenstate Thermalization Hypothesis*, 2009.00632.

[53] D. Wakeham, *Why is a soap bubble like a railway?*, 2008.09611.

[54] I. Heemskerk, J. Penedones, J. Polchinski and J. Sully, *Holography from conformal field theory*, *Journal of High Energy Physics* **2009** (2009) 079–079.

[55] J. Maldacena, D. Simmons-Duffin and A. Zhiboedov, *Looking for a bulk point*, *JHEP* **01** (2017) 013 [1509.03612].

[56] A. May, *Quantum tasks in holography*, *Journal of High Energy Physics* **2019** (2019) .

[57] A. May, G. Penington and J. Sorce, *Holographic scattering requires a connected entanglement wedge*, *Journal of High Energy Physics* **2020** (2020) .

[58] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, *Theoretical Computer Science* **560** (2014) 7 .

[59] A. Bouland, B. Fefferman and U. Vazirani, *Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality*, 1910.14646.

[60] I. Kim, E. Tang and J. Preskill, *The ghost in the radiation: Robust encodings of the black hole interior*, *JHEP* **06** (2020) 031 [2003.05451].

[61] N. Engelhardt, S. Fischetti and A. Maloney, *Free energy from replica wormholes*, 2007.07444.

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

[63] K. C. Hannabuss and M. Semplice, *Boundary conformal fields and tomita–takesaki theory*, *Journal of Mathematical Physics* **44** (2003) 5517.

[64] A. Almheiri, X. Dong and D. Harlow, *Bulk Locality and Quantum Error Correction in AdS/CFT*, *JHEP* **04** (2015) 163 [1411.7041].

[65] D. Harlow, *The Ryu–Takayanagi Formula from Quantum Error Correction*, *Commun. Math. Phys.* **354** (2017) 865 [1607.03901].

[66] A. Almheiri, *Holographic quantum error correction and the projected black hole interior*, 1810.02055.