Holographic models of black hole evaporation from the ground up

Galit Anikeeva
Department of Physics, Stanford University, Stanford, CA 94305
(Dated: June 19, 2020)

Submitted as coursework for PH470, Stanford University, Spring 2020

One of the main paradoxes while developing a theory of quantum gravity is how to resolve the tension between the information loss that we expect in black holes with the understanding that no information is lost in a quantum theory. This is the setup where information goes into the black hole at finite time, the black hole evaporates at a large time, and we check if we can recover the information after the black hole disappears. The calculations done by Hawking[1] showed that this is not possible. The holographic principle, and AdS/CFT duality in particular, has emerged recently as the most promising theory of quantum gravity. It allows us to map a quantum gravitational picture into a pure quantum field theory without gravity and vice versa. This makes preparation and analysis simpler, as we have more options to work with. We present a review of Akers et al[2] that provides two holographic models of black hole evaporation. While superficially similar, one of these models predicts information loss, while the other does not. They attempt to clarify why recent models of black holes seem to conserve information, in apparent contradiction to Hawking's famous calculations. In particular, they keep the relation between the bulk geometry and the boundary explicit.

I. BACKGROUND

A. AdS/CFT

AdS/CFT duality[3], and more generally the gauge/gravity duality, is the idea that for every theory of gravity, there is a quantum field theory in the limit, and most concepts on one have a mapping in the other one. In it's most concrete iteration, it shows that we can obtain a conformal field theory (CFT) dual to the Anti-de Sitter (AdS) space. In particular, we have a duality between operators in the CFT and AdS. One way to construct this operator correspondance is by taking the large radius limit of AdS.

Thus we have quantum gravity in asymptotically AdS space in the center (bulk) and a quantum field theory without gravity in the boundary. We will be working with AdS_3/CFT_2, visualized in 1, which means that our theory will have two space dimensions plus one time dimension in the bulk, and one space dimension plus one time dimension in the boundary. A nice characteristic of this model is that gravitational waves do not exist, simplifying a lot of the needed results.

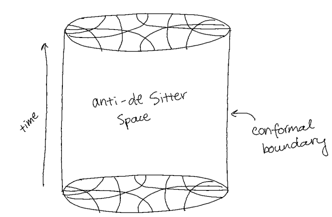

FIG. 1: Visualization of AdS_3/CFT_2 space

B. Black holes in AdS/CFT

Anti de-Sitter space is the maximally symmetric space with constant negative curvature. This curvature induces a scale, called $L = L_{AdS}$. In AdS space there are two kinds of stationary (Schwarschild) black holes. We have small black holes, that is black holes with radius smaller than the AdS scale, which are thermodynamically unstable. They evaporate but do so in an extreme way[4]. This is because, by being smaller than the AdS scale, they are essentially in flat space - which results in its disintegration, since flat space is not able to support black holes. To set up an equivalent scenario as the one proposed by Hawking[1] we need slow dynamics. Additionally, this instability means they have not been studied as much.

We also have large black holes, that is black holes with radius bigger than the AdS scale, which are stable. Since they are bigger than the AdS scale, the curvature starts affecting them. In essence, the asymptotic AdS space acts as a box, and it make the black hole be will be in equilibrium with a radiation bath. Thus the large black holes will reabsorb radiation faster than they evaporate, and reach a thermal equilibrium. They have a positive specific heat and they do not disappear at late times. This stability makes their connection to CFT simpler.

We want to be working with these latter black holes. To induce evaporation, we will couple the CFT with an external system, so that radiation does not bounce back. More precisely, we will be connecting a main CFT additional CFTs to create stable wormholes between them, shifting the thermal equilibrium. We will be manually driving evaporation forward through each coupling step.

An eternal black hole in AdS[5] at temperature T has as a dual in the CFT the thermal density state

$$\rho_{TDS} = \sum_i p_i \left|\Psi_i\right\rangle \left\langle\Psi_i\right| \propto e^{-\beta E_i} \left|\Psi_i\right\rangle \left\langle\Psi_i\right|$$

Here $\beta^{-1} = k_B T$ and k_B is the Boltzmann constant, and $|\Psi_i\rangle$ is the eigenstate of the CFT at energy E_i. Recall that β is inversely proportional to the radius of the black hole.

Note that there are two notions of entropy here. The first one is the entropy in the CFT picture, von Neumann entropy, which we can calculate from the above state. Recall that for a general state ρ this is given by[6]

$$S_{CFT} = -\operatorname{Tr}(\rho \ln \rho)$$

The second one is the entropy in the geometric picture, given by the Ryu-Takayanagi formula[7]. We can think of this as a variant to the Bekenstein-Hawking entropy. For a region A, this is given by

$$S_{AdS} = \frac{\text{Area}(\gamma)}{4G_N}$$

where γ is the Hubeny, Rangamani, and Takayanagi (HRT) surface, which is the extremal surface of minimal area and G_N is the gravitational constant. For our purposes, this is the area of the smallest cut possible cut that separates the region A from the rest. In our case, we will be separating the main black hole from the rest.

Finally, we have that black holes in AdS_3 satisfy the energy-entropy relation $S = 2\pi\sqrt{\frac{E}{3}}$. In terms of the area, this means that $A = 8G_N\pi\sqrt{\frac{E}{3}}$. Note that "area" for AdS_3 refers to the length of the boundary, since we are dealing with 2 space dimensions[8].

C. Simple Wormholes and multi-boundary wormholes

Wormholes can also be viewed throught the lens of AdS/CFT. In particular, Schwarzschild wormhole in the AdS/CFT view are understood to be duals of the thermofield double state. That is, given two identical CFTs, we can construct the entangled state[9]

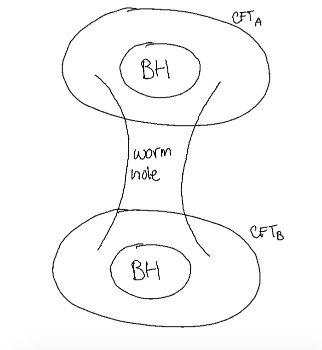

FIG. 2: Example of wormhole

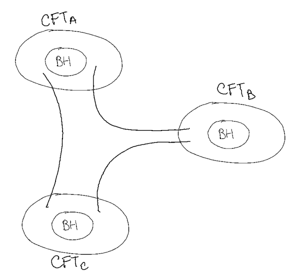

FIG. 3: Example of multi-boundary wormhole

$$|TFD\rangle = \sum_i e^{-\beta E_i} \left|\Psi_i\right\rangle_0 \left|\Psi_i\right\rangle_1$$

where the reduced density matrix on each side is given by the thermal density state. This backs up the interpretation of the state as a wormhole. This interpretation is the core of the conjecture ER=EPR[10]. We have that $\beta^{-1} = \frac{r}{2\pi L_{AdS}}$, where r is the horizon radius[11]. Here we are looking at a wormhole with the same radius on each side. See figure 2.

One of the tools we use to construct one of the two models are multi-boundary wormholes[11], as in 3. The extra mouths of the wormhole, from the point of view of these models, will be created from black hole evaporation. In the CFT picture, this is a multi-CFT generalization of the thermofield double. Note that the above describes a wormhole where each mouth of the wormhole is of the same radius. The corresponding expression for more general geometries is more complicated. Invoking the duality will give us great leverage here.

Going forward, we will work with AdS_3/CFT_2, since it will simplify the analysis. We want to understand how to prepare the different wormholes that we need, as well as the relation between the parameters in the state and the parameters of the black hole.

We will first focus on the case of the simple $|TFD\rangle$ state. To prepare this state, we can take a cylinder with a CFT in each side, and compute an euclidean path integral. The β parameter will be given by the conformal modulus of the cylinder. In general, we can consider a general Riemann surface connecting several CFTs, and then take an euclidean path integral along the boundaries to create a multi-boundary black hole, as proposed in[11].

We note that for our purposes, we only care that the state can be prepared. Whether we do this with an euclidean path integral or with another method will not affect our results.

II. MODELS

We present two alternative models for the information paradox in black holes. One of our models will preserve information and the other will lose it. Recall that the relevant metric to measure is the entropy of the state outside the black hole. In both cases this state will start pure. In the information conservation case this will be approximately zero at late times, while in the information loss case this will be high at late times.

We avoid direct entropy calculation, and instead strive to get a result purely from the geometry, more precisely from the comparison of few HRT alternatives. For this, we work directly within AdS_3/CFT_2. As mentioned before, the lack of gravitational waves simplifies the results. In particular, quantum gravitational corrections go away. Our plan is to describe the relevant state in, describe the evolution of the state, and the mapping through time in the gravitational picture.

It bears emphasizing that we manipulate the CFT, and we analyze the AdS. Both states start at the same state, and both black holes end up being evaporated. However, the path they take to get there, and in particular the geometry of the entanglement, determines if information is lost or not.

We assume we have $n+1$ CFTs, with large enough n, labeled from 0 to n. We have a black hole of radius $r >> L_{AdS}$ in CFT_0 and the vacuum state in the rest. Our operations won't change the energy of the collection, so we will have a constant energy through the process. This means that the square of the length of the HRT surfaces will be constant, by the energy-entropy relation.

A. Information conservation

In the information conservation, at each time step the black hole evaporation process produces an additional boundary in a wormhole. The possible HRT surfaces are "entrance", where the wormhole has the opening towards the main black hole, and the alternative is the union of the additional boundaries. Through time, the minimum between these two first increases, then decreases and reaches 0 again. We interpret this as information conservation at late times. We can see a visualization of this in 4.

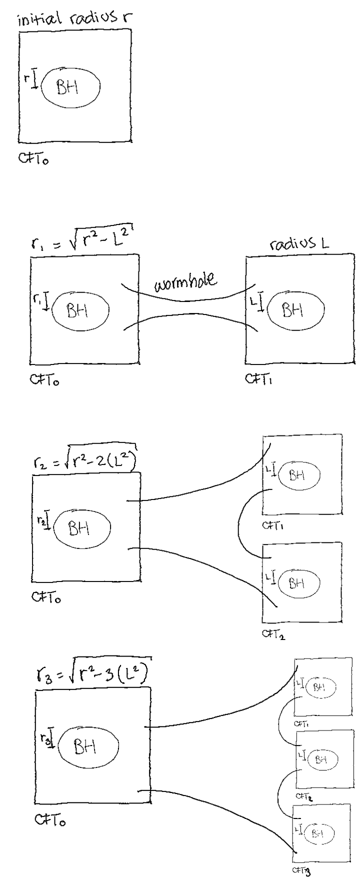

FIG. 4: Evolution of the information conservation model. Note that there is just a single multi-boundary wormhole.

In higher detail, consider n CFTs, one of them dual to a black hole of radius $r >> L_{AdS}$ and the rest in vacuum state. We label the CFT with the original black hole as CFT_0.

By coupling CFT_0 with CFT_1, we can create a wormhole connecting the two spaces. We choose the coupling so that the radius of the wormhole in CFT_1 is just above L_{AdS}. We now want to compute the entropy of the initial black hole. Suppose the new radius of the black hole in CFT_0 is r_1. Then we have that $r_1^2 + L_{AdS}^2 = r^2$. We have that $r_1 = \sqrt{r^2 - L_{AdS}^2}$. Since the entropy will be the area of the extremal surface separating CFT_0 from the rest, it will be the minimum between r_1 and L_{AdS}. Since we assume $r >> L_{AdS}$, we have $r_1 > L_{AdS}$, so we have that the entropy a this point will be $L_{AdS} \cdot \frac{2\pi}{4G_N}$.

We then couple this wormhole with CFT_2, creating a multi-boundary wormhole between CFT_0,CFT_1, and CFT_2. We do the coupling so that CFT_2 has a radius of length L_{AdS}. Suppose now that the black hole in CFT_0 has radius r_2. Then we have $r_2^2 + 2L_{AdS}^2 = r^2$. Thus $r_2 = \sqrt{r^2 - 2L_{AdS}^2}$. There are two options for separating CFT_0 from CFT_1 and CFT_2. Either the boundaries around the black holes in CFT_1 and CFT_2 or the bound-

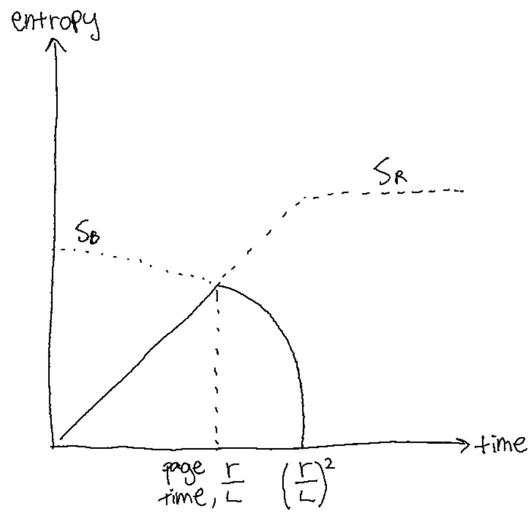

FIG. 5: Page curve of the information conservation model. The dotted line is the alternative page curve, and the diverging point.

ary of the black hole in CFT_2. The minimum of this is $2\pi(2L_{AdS})$. Thus the entropy will be $(2L_{AdS}) \cdot \frac{2\pi}{4G_N}$.

In general, at the k-th step, we will have a multi-boundary wormhole between CFT_0, CFT_1, ..., and CFT_k. We will have that CFT_i will be have black hole of radius L_{AdS} for $i > 0$, and CFT_0 will have a black hole of radius $r_k = \sqrt{r^2 - kL_{AdS}^2}$. The entropy will be $\min(r_k, kL_{AdS}) \cdot \frac{2\pi}{4G_N}$.

For early times, roughly $k < \frac{r}{L_{AdS}}$, we will have that the minimum between these two lengths will be kL_{AdS}. Thus the entropy will be $kL_{AdS} \cdot \frac{2\pi}{4G_N}$. Note how this is increasing in k.

For late times, roughly $k > \frac{r}{L_{AdS}}$, the latter term becomes dominant, so the entropy becomes $r_k \cdot \frac{2\pi}{4G_N} = \sqrt{r^2 - kL_{AdS}^2} \cdot \frac{2\pi}{4G_N}$. This is decreasing in k. So the entropy eventually starts decreasing.

We can continue until $r_k \approx L_{AdS}$, which happens at $k \approx \frac{r^2}{L_{AdS}^2}$. At the end, we end up with a multi-partite wormhole between $k_f \approx \frac{r^2}{L_{AdS}^2}$ CFTs, each containing a minimally viable large black hole. The entropy at this point is just over $L_{AdS} \cdot \frac{2\pi}{4G_N}$, which is essentially zero for our purposes. So CFT_0 ends up (essentially) pure at the end.

Thus evaporating the model in this way, we recover (essentially) all the information. In summary, we have the following page curve[12] in figure 5.

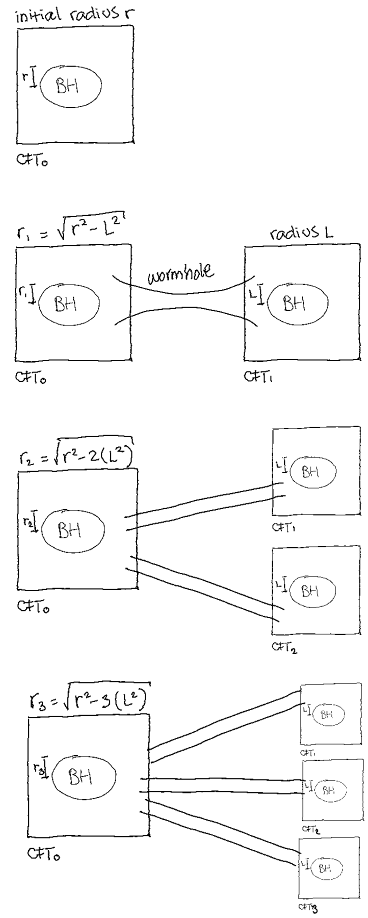

FIG. 6: Evolution of the information loss model. Note that the wormholes are disjoint.

B. Information loss

In the information loss model, at each time step the black hole evaporation produces an additional copy of the thermofield double state. The relevant HRT is the union of the cross-section of each thermofield double. The total area is constant through time. We interpret this as the traditional information loss computation. We can see a visualization of this process in 6.

In higher detail, consider n CFTs, one of them dual to a black hole of radius $r >> L_{AdS}$ and the rest in vacuum state. We label the CFT with the original black hole as CFT_0. At this point, before any operation, this CFT is in a pure state, so it has an entropy of zero.

By coupling CFT_0 with CFT_1, we can create a wormhole connecting the two spaces. We again choose the coupling so that the radius of the wormhole in CFT_1 is just above L_{AdS}. At this point, the situation is identical to the information conservation case, so we get an entropy of $L_{AdS} \cdot \frac{2\pi}{4G_N}$.

The distinction comes in step 2 and forward. We now couple the black hole in CFT_0 with CFT_2, creating a

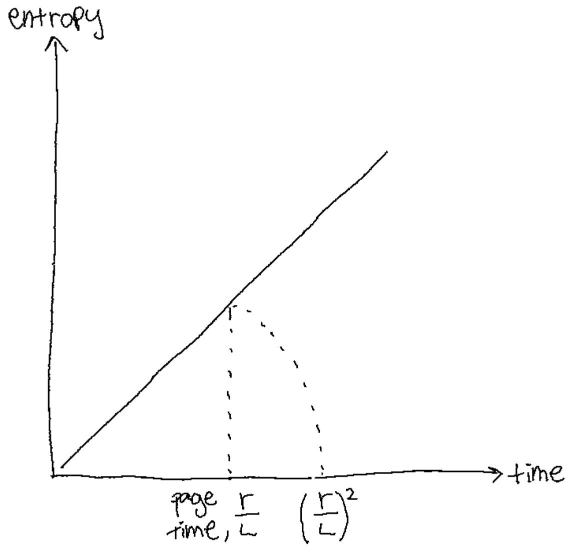

FIG. 7: Page curve of the information loss model. The dotted line is the alternative page curve, and the diverging point.

wormhole between CFT_0 and CFT_2, in addition, and more importantly disjoint, to the existing wormhole between CFT_0 and CFT_2. Since they are disjoint, the minimal surface separating CFT_0 from the rest is just the union of the surfaces separating CFT_0 from CFT_1 and CFT_0 from CFT_2. Individually each area is $2\pi L_{AdS}$. Together the areas will be $4\pi L_{AdS}$. So we get an entropy of $2L_{AdS} \cdot \frac{2\pi}{4G_N}$

In general, at the k-th step, we will have a k disjoint wormholes between CFT_0 and CFT_i, for $i = 1, \ldots, k$. The minimal surface separating CFT_0 from the rest is the addition of the the minimal surfaces separating CFT_0 from CFT_i. Individually each will be $2\pi L_{AdS}$, so in total we will have $2\pi(kL_{AdS})$. Thus the entropy will be $kL_{AdS} \cdot \frac{2\pi}{4G_N}$.

Note that this answer only depends linearly on k. The behavior with respect to k be the same regardless if we are at early or at a late time. As before, we can continue until $k \approx \frac{r^2}{L_{AdS}^2}$. This means that the entropy at the end will be approximately $\frac{r^2}{L_{AdS}} \cdot \frac{2\pi}{4G_N}$.

Thus we do not recover the entropy, and CFT_0 ends up significantly mixed at the end. In summary, we have the following page curve in figure 7.

III. DISCUSSION

Despite the technical complexity necessary to consider this model, the ideas are in essence very simple. Whether information is lost or not depends on the geometry of entanglement. Is radiation going out in a way that is independent each time-step, or is it entangled throughout?

[1] S. W. Hawking, Communications in Mathematical Physics **43**, 199 (1975).
[2] C. Akers, N. Engelhardt, and D. Harlow, *Simple holographic models of black hole evaporation* (2019), 1910.00972.
[3] J. Maldacena, International Journal of Theoretical Physics **38**, 1113–1133 (1999), ISSN 0020-7748, URL http://dx.doi.org/10.1023/A:1026654312961.
[4] N. Jokela, A. Pönni, and A. Vuorinen, Physical Review D **93** (2016), ISSN 2470-0029, URL http://dx.doi.org/10.1103/PhysRevD.93.086004.
[5] J. Maldacena, Journal of High Energy Physics **2003**, 021–021 (2003), ISSN 1029-8479, URL http://dx.doi.org/10.1088/1126-6708/2003/04/021.
[6] I. Bengtsson and K. Zyczkowski, *Geometry of Quantum States: An Introduction to Quantum Entanglement* (Cambridge University Press, 2006).
[7] S. Ryu and T. Takayanagi, Physical Review Letters **96** (2006), ISSN 1079-7114, URL http://dx.doi.org/10.1103/PhysRevLett.96.181602.
[8] N. Tetradis, Journal of High Energy Physics **2012** (2012), ISSN 1029-8479, URL http://dx.doi.org/10.1007/JHEP02(2012)054.
[9] W. Cottrell, B. Freivogel, D. M. Hofman, and S. F. Lokhande, Journal of High Energy Physics **2019** (2019), ISSN 1029-8479, URL http://dx.doi.org/10.1007/JHEP02(2019)058.
[10] J. Maldacena and L. Susskind, Fortschritte der Physik **61**, 781–811 (2013), ISSN 0015-8208, URL http://dx.doi.org/10.1002/prop.201300020.
[11] V. Balasubramanian, P. Hayden, A. Maloney, D. Marolf, and S. F. Ross, Classical and Quantum Gravity **31**, 185015 (2014), ISSN 1361-6382, URL http://dx.doi.org/10.1088/0264-9381/31/18/185015.
[12] D. N. Page, Physical Review Letters **71**, 3743–3746 (1993), ISSN 0031-9007, URL http://dx.doi.org/10.1103/PhysRevLett.71.3743.