Entanglement Entropy of a Radiative Black Hole

Steven Jackson
Advisor: Herman Verlinde
Princeton University

March 13, 2013

Abstract

Motivated by obtaining quantitative answers to subtle information theoretic questions about black holes, we consider a quantum mechanical model for the evolution of thermodynamic black holes and their radiation, akin to a radiative atom. Using a few simple calculational rules we calculate the Renyi entropies in this model and use these to calculate the entanglement entropy in two simple limits with broad applicability. In doing so, we calculate a nearly complete Page curve for our model, providing insight to some of these information theoretic questions.

1 Introduction

The discovery of Hawking radiation [Hawk] led to an important question about what happens to information in black holes. Is it lost forever inside the black hole as it evaporates, violating unitarity of quantum mechanics, or is the information somehow encoded in the Hawking radiation? The AdS/CFT correspondence suggests that the internal degrees of freedom of the black hole store the information, which is then transferred to the radiation over time.

This then raises the questions of precisely how the black hole entropy evolves over time and when and how it subsequently emits information. Moreover, the firewall paradox of Ahlmieri et al. [AMPS] is information theoretic, so sharpening our understanding of black hole thermodynamics might also help resolve such paradoxes. Over the intervening years, there have been several notable attempts at answering these information theoretic questions about black holes.

First, we have the argument made by Page in [Page]. Rather than assuming anything about the black hole, his approach was to consider a large system containing a smaller subsystem, and to see how the average entanglement entropy of the smaller subsystem depends on its size. He found that for small subsystems that the entanglement entropy is very nearly maximal on average. Moreover, if both subsystems are sufficiently large, the entanglement entropy of the smaller subsystem is very nearly maximal, even when the size of each is approximately equal.

To relate this to a black hole, we initially start with a black hole in a pure state with no radiation, so the overall system is in a pure state. This means that the entanglement entropy of the black hole will equal that of the radiation at all times. As the black hole evolves it will become entangled with the radiation, and so as it ages and emits radiation its entanglement entropy also increases. Since the total dimension of the Hilbert space is enormous, it turns out that the entanglement entropy is always within half a bit of the maximum value.

A simplified, but exactly solvable, model of a black hole is the bit model. We start with a black hole containing n_0 qubits, where $2^{n_0} = e^{4\pi M^2}$ is the number of states at that mass. Of these qubits, n_1 will be entangled with other qubits in the black hole, while n_2 will be entangled with qubits outside the black hole, and thus the entanglement entropy will be proportional to n_2. For the dynamics, when $t \to t+1$ we choose an interior qubit randomly for the black hole to emit. With probability n_1/n this will move a type 1 qubit outside, decreasing n_1 by two while increasing n_2 by one. Likewise, with probability n_2/n a type 2 qubit is selected, and n_2 decreases by one. This results in a pair of coupled differential equations

$$\frac{dn_1}{dt} = -2\frac{n_1}{n}; \qquad \frac{dn_2}{dt} = \frac{n_1 - n_2}{n} \tag{1}$$

with initial conditions $n_1 = n_0$ and $n_2 = 0$. Solving these equations, we find that

$$n_1(t) = \frac{(n_0 - t)^2}{n_0}; \qquad n_2(t) = \frac{t(n_0 - t)}{n_0}. \tag{2}$$

Thus, the page curve for this model will be a downwards parabola. This is compared with the Page curve for average subsystems in figure 1.

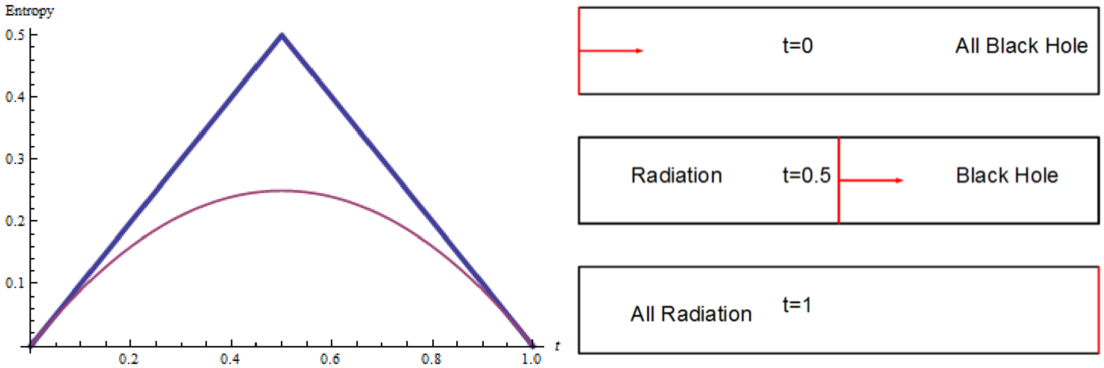

Figure 1: **a.** The Page curve for a generic subspace (blue) and the bit model of the black hole (purple). Time is scaled so that the slope of the generic subspace model is ± 1 and the evaporation time for the black hole is $t = 1$. Entropy is scaled so that the initial Bekenstein-Hawking bound is $S = 1$. **b.** The red bar proceeds to the right at a speed of 1, representing the decay of the black hole into radiation. For $t < .5$ the black hole is the dominant system, while for $t > .5$ the radiation dominates.

Beyond these models for the evolution of the entropy of a black hole, Hayden and Preskill have also considered how well black holes hide information in [HP]. They found that old black holes emit the information almost as quickly as possible. Moreover, a black hole can hide information as long as it is young, but once it becomes old it emits information in much the same way.

In order to answer questions like these, we will formulate a model for the internal dynamics of a black hole that respects the rules of black hole thermodynamics. Moreover, time evolution will be manifestly unitary and the time evolution matrices will be ergodic to allow for fast scrambling. In the most simple sense, one can thing of our model as treating the black hole like a highly excited atom that spontaneously emits atoms as time evolves.

We start by developing this model and applying it to the simplest cases of extremely young and extremely old black holes. We then develop our formalism further to deal with the more complicated calculations necessary to compute the entropy in more general cases. Finally, we consider how to apply this formalism to the Hayden Preskill experiment.

Throughout the main body of the paper, equations marked with overlines are to be understood to be statistical relations rather than identities.

2 Hilbert Spaces and Time Evolution

A black hole of mass M was shown by Bekenstein and Hawking to have an entropy of $A/4$. For a Schwarzschild black hole, the radius is $2M$, and consequently the entropy is $S = 4\pi M^2$. The total number of microstates will then be the exponential of this, which we will call N_M, or just N in cases where we are considering a single mass only. The Hilbert space for a black hole of mass M will then be an N dimensional space which we call $\mathcal{H}_M$.

As we will also want to allow for radiational decay of the black hole mass, we want a single Hilbert space for the black hole at any future time. Therefore, the Hilbert space for the black hole will be a direct sum of $\mathcal{H}_M$ for all masses less than our starting mass, with some coarse-graining implied:

$$\mathcal{H}_{\mathrm{BH}} = \bigoplus_{E \leq M} \mathcal{H}_{M-E}. \tag{3}$$

Since the black hole will be emitting radiation over time, we also need to specify a Hilbert space for the radiation. This will just be a standard Hilbert space for radiation, with occupation numbers for each wavelength.

In the simplest case, we begin our evolution in a known state of the black hole at mass M with no radiation, which we denote as $|i\rangle|0\rangle$, where i is the label of one of the states with mass M. After time t this evolves to a new state

$$U(t)|i\rangle|0\rangle = \sum_{n,j} C^i_{jn}|j\rangle|n\rangle. \tag{4}$$

This new state will necessarily be of lower mass, having emitted an energy E_n as photons in the radiation state.

Little is known about the particulars of internal black hole dynamics, but we will make a fast scrambling assumption for the dynamics. That is, we will take a random matrix theory approach and simply assume that the matrices C_n with entries C^i_{jn} are highly mixing ergodic matrices. That is, our dynamics allow the black hole state to become maximally entangled very quickly. Finally it is worth noting that C_n maps down from $\mathcal{H}_M$ to $\mathcal{H}_{M-E_n}$, its Hermitian conjugate maps up from $\mathcal{H}_{M-E_n}$ to $\mathcal{H}_M$.

In addition to the ergodicity assumption, the C^i_{jn} coefficients must satisfy certain relations consistent with black hole thermodynamics. First, the unitarity of time evolution requires that

$$\sum_{j,n}(C^i_{jn})^* C^k_{jn} = \delta^{ik}. \tag{5}$$

Moreover, the density matrix of the radiation should be thermal at the Hawking temperature, leading to the statistical relation

$$\sum_{j}\overline{(C^i_{jm})^* C^i_{jn}} = w_n \delta_{mn}, \tag{6}$$

where w_n is the weighted Boltzmann factor for the radiation state n at energy E_n

$$w_n = \frac{e^{-\beta E_n}}{Z}; \qquad\qquad Z = \sum_n e^{-\beta E_n}, \tag{7}$$

and $\beta = 8\pi M$ is the inverse Hawking temperature. Before moving on, we should note that this holds for any i in the black hole, and that if the states i and k are orthogonal then the result will be too.

Our final computational rules, which will be introduced more thoroughly later. The first is the generalization of (6) mentioned above:

$$\sum_{j}\overline{(C^i_{jn})^* C^k_{jm}} = w_n \delta_{nm}\delta^{ik}. \tag{8}$$

The second is

$$\sum_{i}\overline{C^i_{jn}(C^i_{km})^*} = \frac{\delta_{jk}\delta_{nm}}{Z}. \tag{9}$$

To check that these are sensible, we should consider in what cases we have constructive interference, which should contribute the main terms for each of these, and in which other cases we'll have destructive interference, in which case square-root cancellation should lead to subleading terms.

For the first, we obviously have constructive interference whenever all of the coefficients coincide, which is when $i = k$ and $n = m$. When these conditions are met, there are $Ne^{-\beta E_n}$ values of j that sum constructively. From the unitarity condition, when we sum over all n we simply obtain 1, so a typical term is of order $1/NZ$. We then end up with a total term of $e^{-\beta E_n}/Z = w_n$ multiplying the delta functions.

For the second, we again have constructive interference whenever $j = k$ and $n = m$, and typical terms are again of size $1/NZ$. We are summing over all i, so there are N such terms hence the overall factor multiplying the delta functions is $1/Z$.

Throughout this paper, we will always be operating over small time steps t so that the energy emitted during that time step is sufficiently small compared to the mass, or $E_n \ll M$.

Now that all of our computational rules are laid out, we can consider how the information in the system evolves over time. From the density matrix we can calculate the *von Neumann entropy* of the state ρ:

$$S = -\mathrm{tr}\rho \log \rho. \tag{10}$$

We say a state ρ is *maximally entangled* if it is proportional to the identity matrix. That is, ρ is maximally entangled if $\rho = \frac{1}{N}\mathbb{1}_N$, implying that the entropy is $S(\rho) = \log N$, which is the maximum value possible for a state of dimension N.

If we divide our system into two subsystems A and B and let ρ_{AB} denote the density matrix for the entire system, we can also calculate the entropy of the subsystems. We let ρ_A be the density matrix for A when ignoring correlations with B

$$\rho_A = \mathrm{tr}_B \rho_{AB}, \tag{11}$$

where tr_B denotes a trace over the degrees of freedom of subsystem B. The *entanglement entropy of* A, S_A is the von Neumann entropy of ρ_A:

$$S_A = -\mathrm{tr}\rho_A \log \rho_A. \tag{12}$$

If ρ_{AB} is a pure state, then the entanglement entropy of the two systems will always be equal. Additionally, entanglement entropies satisfy what is called *monogamy of entanglement*, which says that if A is maximally entangled with B then A can share no other entanglement with some other region C. Of course, entanglement between A and B can be transferred to entanglement between A and C over time.

Returning to black hole dynamics, we can consider the more general case of a black hole where the initial state is a uniform superposition over some code subspace of dimension $N_C \leq N$, with corresponding entropy $\log N_C$. In this case, the initial density matrix is given by

$$\frac{1}{N_C}\sum_{\bar{i}} |\bar{i}\rangle\langle\bar{i}| \otimes |0\rangle\langle 0|, \tag{13}$$

where the $\bar{i}$ are the states in our code space.

The evolution for each state is still the same as above, so this initial density matrix will evolve to

$$\begin{aligned}\rho_{AB} &= U(t)\frac{1}{N_C}\sum_{\bar{i}} |\bar{i}\rangle\langle\bar{i}| \otimes |0\rangle\langle 0| U(t)^\dagger \\ &= \frac{1}{N_C}\sum_{\bar{i},j,k,n,m} C^{\bar{i}}_{jn}|j\rangle|n\rangle\langle k|\langle m|(C^{\bar{i}}_{km})^* \\ &= \frac{1}{N_C}\sum_{\bar{i},n,m} C_n|\bar{i}\rangle\langle\bar{i}|C^\dagger_m, \end{aligned}\tag{14}$$

where the latter is compact notation so we can more easily write the density matrix. Tracing over B to obtain ρ_A then sets $m = n$ so we obtain

$$\rho_A = \sum_{\bar{i},n} \frac{C_n|\bar{i}\rangle\langle\bar{i}|C^\dagger_n}{N_C}. \tag{15}$$

2.1 Mutual Information for Black Holes Young and Old

For our purposes, the entanglement entropy is most useful for calculating the *mutual information* between regions A and B, which is proportional to the number of bits of information that can be learned about one region through measurements the other. It is defined by

$$I(A;B) = S_A + S_B - S_{AB}, \tag{16}$$

and is necessarily non-negative.

As a very simple example that will become important later in the paper, an EPR pair is a pure state, so $S_{s_1} = S_{s_2} = \log 2$, where s_1 and s_2 are the two spins in the EPR pair. The mutual information is therefore $2\log 2$, reflecting the fact that one bit of information can be learned about one of the qubits by measuring the other.

For a young black hole for which $N_C = 1$, we start with a pure state so we have that $S_{AB} = 0$, so $S_A = S_B$ and $I(A;B) = 2S_B$. Calculating S_B is simplest since the radiation density matrix is thermal, meaning we know its eigenvalues w_n. Then

$$S_B = S_A = -\sum_n w_n \log w_n = \log Z + \beta \bar{E}, \tag{17}$$

where $\bar{E}$ is the average energy of the radiation state. Thus, the mutual information between the black hole and the radiation is given by $I(A;B)_{\text{young}} = 2(\log Z + \beta\bar{E})$. That is, the entropy of the black hole initially increases proportional to the emitted energy while it is young, showing that the Page curve of a black hole initally increases linearly in time.

For an old black hole, our starting state will be an even superposition over all states of that mass (i.e. $N_C = N$), so we immediately have that $S_{AB} = \log N$. The calculation for S_B goes through as for a young black hole since the radiation density matrix is always thermal, hence $S_B = \log Z + \beta\bar{E}$.

The calculation for S_A is slightly more complicated. Since we have assumed that our evolution operators are ergodic, the black hole should remain in a maximally mixed state, but its mass must decrease in accordance with the photons that it emits. If the mass decreases from M to $M - E_n$, the dimension of the black hole Hilbert space will decrease from N to $Ne^{-\beta E_n}$. We will then be averaging over

$$\sum_n N e^{-\beta E_n} = NZ \tag{18}$$

states. The density matrix for the system isn't exact [2V], but suffices for the black hole calculation and is given by

$$\overline{\rho_{AB}} = \frac{1}{NZ} \sum_n \mathbb{1}_{M-E_n} \delta_{nm}. \tag{19}$$

Since $n \neq m$ can still have $E_n = E_m$, when we trace over B we get a factor of the density of states at energy E, $d_B(E)$, so $\rho_A = \sum_E \mathbb{1}_{M-E} d_B(E)/(NZ)$. Moreover, we must have that $\sum_E e^{-\beta E} d_B(E) = Z$. Then

$$\begin{aligned} S_A &= \frac{d_B(E)}{NZ} \sum_E N e^{-\beta E_n} \log \frac{NZ}{d_B(E)} \\ &= \log N - \beta \bar{E}, \end{aligned} \tag{20}$$

where we have used $e^{-\beta\bar{E}}\overline{d_B(E)} = Z$. Therefore, the mutual information between the old black hole and the radiation is $I(A;B)_{\text{old}} = \log Z$. This shows that the entropy of an old black hole decreases in a manner proportional to the energy emitted, so the Page curve decreases as old black holes that have already attained the Bekenstein-Hawking bound decay through Hawking radiation.

2.2 Subdividing by Energy

While using the radiation states n was fine for most of the calculation above, it was inadequate for calculating S_A for the old black hole. We now further develop this notation to handle the more complicated calculations with intermediate N_C needed to do more general calculations of the Page curve and perform the Hayden-Preskill thought experiment.

For the radiation density matrix, our new notation is (summing only over states n and m satisfying $E_n = E_m = E$)

$$\rho_B = \sum_{E,n,m} \rho^B_{nm}(E)|n\rangle\langle m|, \tag{21}$$

where we have defined

$$\overline{\rho^B_{nm}(E)} = \frac{1}{N_C}\sum_{\bar{i},j} \overline{C^{\bar{i}}_{jn}(C^{\bar{i}}_{jm})^*} = \frac{d_A(M-E)}{d_A(M)Z}\delta_{nm} = w_n\delta_{nm}, \tag{22}$$

where $d_A(M-E)$ is the density of states of the black hole at mass $M-E$, or $Ne^{-\beta E}$. The trace of this density matrix must be one, which simply says that the partition function is now $Z = \sum_E e^{-\beta E} d_B(E)$. Alternatively, we can write this as $\sum_E w_E d_B(E) = 1$, defining the new quantity $w_E = e^{-\beta E}/Z$ as the weighted Boltzmann factor for a single state of energy E.

For the black hole density matrix, we find that

$$\rho_A = \sum_{E,j,k} (\rho_A(E))_{jk}|j\rangle\langle k| = \sum_E \rho_A(E). \tag{23}$$

Then, if the black hole state begins as a uniform superposition over a code subspace of dimension N_C, then to obtain the correct ρ_A we find that $\rho_A(E)$ must have the matrix components

$$(\rho_A(E))_{jk} = \frac{1}{N_C}\sum_{\bar{i},n} C^{\bar{i}}_{jn}\left(C^{\bar{i}}_{kn}\right)^*. \tag{24}$$

Therefore, we see that the trace is given by

$$\mathrm{tr}(\rho_A(E)) = \sum_j (\rho_A(E))_{jj} = w_E d_B(E). \tag{25}$$

Returning to our calculation for an old black hole, in this notation we find that statistically

$$\overline{\rho_A(E)} = \frac{1}{N_C} \sum_{\bar{i},n} \overline{C_n|\bar{i}\rangle\langle\bar{i}|C_n^\dagger} = \sum_E \frac{d_B(E)}{NZ} \mathbb{1}_{M-E}, \tag{26}$$

which yields the correct entropy since $Z = \overline{d_B(E)} e^{-\beta\bar{E}}$.

3 Calculating the Entropy of $\rho(E)$

A useful method for calculating the entropy when faced with more complicated density matrices is to use the *Renyi entropy*, defined by

$$S_n = \frac{1}{1-n} \log \mathrm{tr}(\rho^n). \tag{27}$$

Calculating the Renyi entropy for integer n and then analytically continuing to $n = 1$ is called the *replica trick*, and is done by invoking L'hopital's rule

$$\begin{aligned} \lim_{n\to 1} \frac{1}{1-n} \log \mathrm{tr}(\rho^n) &= -\frac{1}{\mathrm{tr}\rho} \frac{\partial}{\partial n} \left[\mathrm{tr}\rho^n\right]_{n=1} \\ &= -\mathrm{tr}(\rho \log \rho) = S(\rho). \end{aligned} \tag{28}$$

This is often done much more easily than a direct calculation of the von Neumann entropy.

Since $\rho = \sum_E \rho(E)$ and the $\rho(E)$ are orthogonal to each other, we see that

$$\rho^2 = \sum_E \rho(E) \sum_{E'} \rho(E') = \sum_E \rho(E)^2, \tag{29}$$

and similarly for general integer n. Therefore, applying the replica trick to ρ in this case yields

$$\begin{aligned} S(\rho) &= -\mathrm{tr} \frac{\partial}{\partial n} \left[\sum_E \rho(E)^n \right]_{n=1} \\ &= -\sum_E \mathrm{tr}\rho(E) \log \rho(E) = \sum_E S(E). \end{aligned} \tag{30}$$

Thus, to calculate the von Neumann entropy of ρ_A, we have to find an analytic expression for $\rho(E)^n$, differentiate this with respect to n, evaluate at $n = 1$, and finally sum over E.

However, calculating powers of $\rho(E)$ requires quite a bit of work. In the following we develop a recursion relation for powers of the normal ordered density matrices, and then use this recursion relation to find the Renyi entropies, before finally calculating the von Neumann entropy.

To begin, we define the normal ordered density matrix. As usual, the normal ordered density matrix is simply the density matrix minus all of its Wick contractions, and normal orderings of higher powers are defined similarly. Since the available contractions simply yield constants times the density matrix, so we know that each normal ordering will yield a polynomial in $\rho(E)$. Restating this in terms of equations, we have

$$: \rho(E)^n := \rho^n - \text{Wick}(\rho^n) = P_n(\rho(E)). \tag{31}$$

We now determine the form of the polynomial P_n using recursion.

For the first power of the density matrix, the only Wick contraction available is the one in (26), so we find that

$$: \rho(E) := \rho(E) - \frac{d_B(E)}{NZ} \mathbb{1}_{M-E} = \rho(E) - a\mathbb{1}_{M-E}. \tag{32}$$

Note that the trace of the normal ordered density matrix is zero, since the trace of $a\mathbb{1}_{M-E}$ is precisely $w_E d_B(E)$.

Another relation that will be useful in the following is the thermal relation (6), which in our new notation we write as

$$\sum_{n,m} \frac{\overline{C_n^\dagger C_m}}{N_C} = \frac{w_E}{N_C} \delta_{nm} \mathbb{1}_M = b\delta_{nm}\mathbb{1}_M. \tag{33}$$

Schematically, we may write these contractions as $CC^\dagger = a\mathbb{1}_{M-E}$ and $C^\dagger C = b\mathbb{1}_M$.

We can then calculate the normal ordering of the higher powers of the density matrix by noting that

$$: \rho(E)^2 :=: \rho(E) :: \rho(E)^1 : -\text{Wick}(: \rho(E) :: \rho(E)^1 :). \tag{34}$$

For the first term, we simply rewrite $: \rho(E) :: \rho(E)^1 := (\rho(E) - a\mathbb{1}_{M-E}) : \rho(E)^1 :$, while for the second we contract across the two (a $C_n^\dagger C_m$ contraction, so we use (33)) and then normal order the result. In the schematic notation introduced above, this is

$$\begin{aligned}\mathrm{Wick}(:\rho(E)::\rho(E)^1:) &=: C(C^\dagger :: C)C^\dagger : \\ &= b : CC^\dagger : \\ &= b(\rho(E) - a\mathbb{1}_{M-E}).\end{aligned} \tag{35}$$

We then find that

$$:\rho(E)^2 := (\rho(E) - a - b) : \rho(E)^1 : -ab : \rho(E)^0 : . \tag{36}$$

This procedure generalizes immediately to normal orderings of higher powers of the density matrix, leading to the following recursion relation among normal ordered powers of the density matrix:

$$:\rho(E)^n := (\rho(E) - a - b) : \rho(E)^{n-1} : -ab : \rho(E)^{n-2} : . \tag{37}$$

With our initial conditions for $:\rho(E)^0:$ and $:\rho(E):$, this leads to a polynomial recursion relation

$$\begin{aligned}&P_n(\rho(E)) = (\rho(E) - a - b)P_{n-1}(\rho(E)) - abP_{n-2}(\rho(E)) \\ &P_0(\rho(E)) = \mathbb{1}_{M-E} \\ &P_1(\rho(E)) = \rho(E) - a\mathbb{1}_{M-E}.\end{aligned} \tag{38}$$

Moreover, we can turn this recursion relation for the polynomials into a recursion relation for its coefficients. Writing each polynomial in terms of its coefficients c_{nm},

$$P_n(\rho(E)) = \sum_{m\leq n} c_{nm}\rho(E)^m, \tag{39}$$

applying (38) at the coefficient level, we find that

$$\begin{aligned}&c_{nm} = c_{n-1,m-1} - (a+b)c_{n-1,m} - abc_{n-2,m} \\ &c_{0,0} = 1, \qquad c_{1,0} = -a, \qquad c_{1,1} = 1.\end{aligned} \tag{40}$$

For the special cases where $m = 0$, we of course use $c_{n-1,-1} = 0$. Similarly, we always have use that $c_{n-1,n} = 0$.

With these polynomials in hand, we now want to work our way to the Renyi entropies. To do this, we simply take the trace of the polynomail.

$$\mathrm{tr}(P_n(\rho(E))) = \sum_{m\leq n} c_{nm}\mathrm{tr}(\rho(E)^m). \tag{41}$$

Moreover, because $P_n(\rho(E))$ is the polynomial resulting from a normal ordered density matrix for which we subtracted off a statistical average from the density matrix, we have that its trace is zero. From the coefficient recursion relation it is easy to see that $c_{nn} = 1$ for all n, which we can use this to find $\mathrm{tr}(\rho(E)^n)$ in terms of the trace of lower powers:

$$\mathrm{tr}(\rho(E)^n) = -\sum_{m<n} c_{nm}\mathrm{tr}(\rho(E)^m). \tag{42}$$

Since we know the trace of the identity and $\rho(E)$, by using our recursion relation for the coefficients (40), we can use the trace recursion relation (42) to bootstrap our way to traces of higher powers.

We have not yet solved the above recursion relations in general (though we do have some numerical work which we mention below), we can solve them for the special cases where either $a \ll b$ or $b \ll a$. While we can't find the exact form of the interpolating function, its form is constrained by these limits.

If $b \ll a$, we can simply ignore it entirely. Our recursion relation reduces to a single term $P_{n+1}(\rho(E)) = (\rho(E) - a)P_n(\rho(E))$. Together with the initial conditions, this just says that

$$P_n(\rho(E)) = (\rho - a)^n. \tag{43}$$

Since the trace of this quantity is zero, this then implies that

$$\mathrm{tr}(\rho(E)^n) = -\sum_{m<n} (-a)^m\mathrm{tr}(\rho^m). \tag{44}$$

One can show that because of the initial condition $\mathrm{tr}(\rho(E)) = a\mathrm{tr}(\mathbb{1}_{M-E})$ this is simply given by

$$\mathrm{tr}(\rho(E)^n) = a^{n-1}\left(a\mathrm{tr}(\mathbb{1}_{M-E})\right). \tag{45}$$

Subsequently, the replica trick yields the partial von Neumann entropy

$$S(\rho(E)) = -\log(a)\left(a\mathrm{tr}(\mathbb{1}_{M-E})\right) = w_E d_B(E)\log\left(\frac{NZ}{d_B(E)}\right), \tag{46}$$

which is the same as we found for an old black hole in (20).

If $a \ll b$, we can't simply ignore it because our initial condition is still $P_1(\rho(E)) = \rho - a\mathbb{1}_{M-E}$, which results in $\mathrm{tr}(\rho(E)) = a\mathrm{tr}(\mathbb{1}_{M-E})$. However, we can keep the main recursion relation and then drop subleading terms. If we look at P_2 and take the trace, then we find that

$$\mathrm{tr}((\rho - b - a)(\rho - a) - ab\mathbb{1}) = \mathrm{tr}(\rho^2 - (b+a)\rho + (a^2 + ab)\mathbb{1} - ab\mathbb{1}) = 0, \tag{47}$$

which implies that

$$\begin{aligned} \mathrm{tr}(\rho^2) &= (b+a)\mathrm{tr}(\rho) + a^2\mathrm{tr}(\mathbb{1}) \\ &\approx b\mathrm{tr}(\rho) = b(a\mathrm{tr}(\mathbb{1})). \end{aligned} \tag{48}$$

From here, it is easy to see that the leading term will always pick up a single factor of b, so more generally we find that

$$\mathrm{tr}(\rho(E)^n) = b^{n-1}(a\mathrm{tr}(\mathbb{1})). \tag{49}$$

This result could also be obtained by noting that because the recursion relation is symmetric in a and b these polynomials must be symmetric in a and b as well, implying this result. Applying the replica trick then yields a partial entropy of

$$S(\rho(E)) = -w_E d_B(E) \log b. \tag{50}$$

Since $b = w_E/N_C = e^{-\beta E}/N_C Z$, we see that summing over all E gives us the same result as for the young black hole (17) in the case $N_C = 1$ after summing over all energies.

While these two limits perform useful checks for our work thus far, they also allow us to guess at form of the trace of $\rho(E)^n$ for intermediate values of a and b. Both of the expressions always had $\mathrm{tr}(a\mathbb{1}) = w_E d_B(E)$ multiplied by a polynomial. Basically, we know that we must have a function that is asymptotic to a^{n-1} or b^{n-1} in the limit $b \to 0$ or $a \to 0$, and interpolates smoothly between these limits. Thus, we expect that

$$\mathrm{tr}(\rho(E)^n) = f_n(a,b) w_E d_B(E). \tag{51}$$

Based upon numerical work, we find that

$$f_n(a,b) = \sum_{m=0}^{n-1} N(n, m+1) a^{n-1-m} b^m, \tag{52}$$

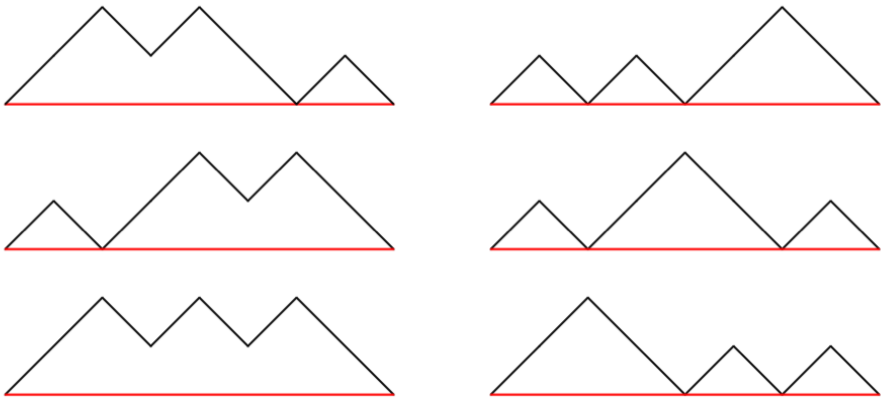

Figure 2: The $N(4,3) = 6$ Dyck paths of 8 steps with 3 peaks that are constrained to always remain above the base.

where $N(n,m)$ is a Narayana number defined by

$$N(n,m) = \frac{1}{n}\binom{n}{m}\binom{n}{m-1}. \tag{53}$$

With this, the Renyi entropies of the density matrix ρ_A are

$$\begin{aligned} S_n(\rho_A) &= \frac{1}{1-n}\log\left[\sum_E \operatorname{tr}(\rho(E)^n)\right] \\ &= \frac{1}{1-n}\log\left[\sum_E \sum_{m=0}^{n-1} N(n,m+1)a(E)^{n-1-m}b(E)^m\right], \end{aligned} \tag{54}$$

where we have made the dependence of a and b on the energy explicit.

These numbers arise combinatorially from Dyck paths, paths which are bounded above the starting point. There are $N(n,m)$ such paths consisting of $2n$ steps and m peaks. An example of the $N(4,3) = 6$ Dyck paths is shown in 2.

The generating function for the Narayana numbers is

$$\sum_{k,j=0}^{\infty} N(k,j)x^k t^j = \frac{1-x-tx-\sqrt{(1-x-tx)^2-4x^2t}}{2x}. \tag{55}$$

4 Summary of Results and Outlook

4.1 Summary

We model black hole dynamics using fast scrambling matrices C_n that manifestly respect unitarity and black hole thermodynamics. Using these matrices we then calculated how the black hole entropy evolves for the simple cases of young and old black holes, which are consistent with the Page curve for an average subsystem and serve as useful checks for our later calculations.

To handle more complicated cases, we switched to considering partial density matrices at a particular energy. We then had $\rho = \sum_E \rho(E)$, and subsequently that

$$S(\rho) = -\sum_E \mathrm{tr}(\rho(E) \log \rho(E)). \tag{56}$$

We then attempted to calculate $\mathrm{tr}(\rho(E) \log \rho(E))$ using the replica trick.

To start, we considered the normal ordered density matrices obtained by subtracting all Wick contractions. This yielded a recursion relation which we could use to calculate these as a polynomial in $\rho(E)$. Since the normal ordered density matrices are traceless, this lets us calculate $\mathrm{tr}(\rho(E)^n)$ in terms of lower powers of the trace. Using our numerical calculations, we found that in terms of

$$a = \frac{d_B(E)}{NZ}; \qquad b = \frac{w_E}{N_C}, \tag{57}$$

the trace of $\rho(E)^n$ is given by

$$\mathrm{tr}(\rho(E)^n) = w_E d_B(E) \left(\sum_{m=0}^{n-1} N(n, m+1) a^{n-1-m} b^m \right), \tag{58}$$

where $N(n, m)$ is a Narayana number.

Using this relation, we can see that in the regions where $a \gg b$ or $b \gg a$, the Page curve for our model follows the average subsystem model quite closely. The exception is in the small turnover region, where a and b are similar in magnitude.

4.2 Outlook

Now that we have an expression for $\mathrm{tr}(\rho(E)^n)$ in (52), we will try to obtain an expression for the Renyi entropy that we can take the derivative of and analytically continue to $n = 1$. This will then allow us from obtaining an expression for the von Neumann entropy in the transitional region where a and b are of comparable size. Once we have the von Neumann entropy, we will be able to calculate the Page curve for this fast scrambling model in its most interesting region, where it transitions from (when properly scaled for comparison with the generic Page curve and the bit model) increasing linearly to decreasing linearly with the energy emitted.

We can also make an estimate of the size of the transitional region when a and b are of similiar size. If we consider where $a = b$, we find that this implies that

$$\frac{e^{-\beta E}}{d_B(E)} = \frac{N_C}{N}. \tag{59}$$

The density of states $d_B(E)$ can be obtained by counting the number of ways to add photon energies to E, and subsequently $d_B(E) \approx 2^{\beta E}$. Since everything is exponential and $\beta = 8\pi M$ is a big number, a change in E of $1/M$ is enough to take us from a regime where b is dominant to a region where a is dominant. Thus, we can conclude that for our model the Page curve is quite close to the average subsystem model pursued by Page itself.

For the Hayden-Preskill thought experiment, we currently have done enough work to make some qualitative arguments, but our future work will allow us to get quantitative data. Based on our previous work with a flawed version of this model, we know that when we compare the entanglement entropies, we are essentially just comparing two different points on the Page curve. If we throw in a single qubit, then the dimension of the initial black hole Hilbert space N doubles and for one of the entanglement entropies N_C also doubles, but for the other it does not. The information then comes out in the region where the two slopes differ, in which case the two cases are straddling the transitional region. Thus, information comes out almost immediately for old black holes, whereas for younger black holes we must wait until the black hole first attains the Bekenstein-Hawking bound. When we have calculated the details of the transitional region, we will be able to calculate exactly how the information comes out, and whether our model is in accord with the findings of [HP].

There are also a number of checks that we could run on our results. An example of an information theoretic bound is the Lindblad-Uhlmann monotinicity, which for the Hayden-Preskill thought experiment says that once we have thrown our information into the black hole the information that the black hole contains can never increase via time evolution. Another check that could be performed through thought experiment might involve the strong subadditivity of the von Neumann entropy.

References

[AMPS] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully. *Black Holes: Complementarity or Firewalls?*. arXiv:1207.3123

[HP] P. Hayden and J. Preskill. *Black holes as mirrors: Quantum information in random subsystems.* JHEP 0709 (2007) 120, 0708.4025

[Hawk] S. Hawking. *Particle Creation by Black Holes.* Commun.Math.Phys. 43 (1975) 199220

[Page] D. Page. *Average Entropy of a Subsystem*. gr-qc/9305007v2

[STU] L. Susskind, L. Thorlacius, and J. Uglum. *The Stretched horizon and black hole complementarity*. Phys.Rev. D48 (1993) 37433761, hep-th/9306069

[2V] E. Verlinde and H. Verlinde. *Black Hole Entanglement and Quantum Error Correction*

1 Page Curve from Random Matrices

It is instructive to consider the calculation of the von Neumann entropy of the CFT and the radiation from the perspective of the random matrix description of the respective density matrices in terms of the Kraus operators

$$\boldsymbol{C} = \sum_{a,n} |n\rangle \, C_{an} \, \langle a| \qquad ; \qquad \boldsymbol{C}^\dagger \, = \sum_{a,n} |a\rangle \, C^*_{na} \, \langle n| \tag{1.1}$$

This calculation is a slight generalization of the original analysis by Page.[1]

After the evaporation process, the density matrices of the CFT and radiation can then be written in short-hand notation as follows

$$\rho_{\textsf{CFT}} \, = \, \boldsymbol{C}^\dagger \boldsymbol{C} \qquad ; \qquad \rho_{\textsf{Rad}} \, = \, \boldsymbol{C}\boldsymbol{C}^\dagger. \tag{1.2}$$

Here $\boldsymbol{C}$ is viewed as the mapping from the CFT Hilbert space to the radiation Hilbert space with random matrix elements taken from a gaussian ensemble with uniform variance.

The calculation of the Renyi entropies proceed via straightforward application of Wick's theorem. There are two elementary types of Wick contractions defined by taking the ensemble average of both density matrices. Assuming that energy is the only macroscopic conserved quantity, the respective ensemble averages take the general form

$$\overline{\rho_{\textsf{CFT}}} \, = \, \overbracket{\boldsymbol{C}^\dagger \boldsymbol{C}} \, = \, \sum_E q_{{}_E} \, \mathbb{1}_{\textsf{CFT}}(M - E) \tag{1.3}$$

$$\overline{\rho_{\textsf{Rad}}} \, = \, \overbracket{\boldsymbol{C}\boldsymbol{C}}^\dagger \, = \, \sum_E p_{{}_E} \, \mathbb{1}_{\textsf{Rad}}(E) \tag{1.4}$$

Here M is the energy of the initial state and E the total energy in the radiation. Here $\mathbb{1}_{\textsf{CFT}}(M - E)$ and $\mathbb{1}_{\textsf{Rad}}(E)$ are the projections on the CFT Hilbert subspace $\mathcal{H}_{\textsf{CFT}}(M - E)$ with energy $M - E$ and on the radiation Hilbert space $\mathcal{H}_{\textsf{Rad}}(E)$ with energy E, respectively. It is useful to introduce the ratio

$$z_{{}_E} \, = \, \frac{p_{{}_E}}{q_{{}_E}} \, = \, \frac{\mathrm{tr}(\mathbb{1}_{\textsf{CFT}}(M - E))}{\mathrm{tr}(\mathbb{1}_{\textsf{Rad}}(E))} \, = \, \frac{\dim \mathcal{H}_{\textsf{CFT}}(M - E)}{\dim \mathcal{H}_{\textsf{Rad}}(E)} \tag{1.5}$$

where we used that $\mathrm{tr}(\rho_{\textsf{CFT}}) = \mathrm{tr}(\rho_{\textsf{Rad}}) = 1$. Before the Page transition, this ratio $z_{{}_E}$ is very large, whereas after the Page time, it is very small. The situation is symmetric between the CFT and radiation, and corresponding the pre- and post-Page time discussions are also identical. So for definiteness, let us concentrate on the regime after the Page time, so that $z_{{}_E} < 1$. We will work to all order in $z_{{}_E}$, so our treatment will extend into the regime very close

---

[1] This note overlaps with Steven Jackson's prethesis, adding some graphical notation and details.

to the Page transition. For notational convenience, we will assume that the energy resolution is fine-grained enough so that the CFT Hilbert subspace at given E is one-dimensional. The dimension of the radiation Hilbert space at fixed E is defined in a statistical sense, as the ratio between the level density of the radiation and the CFT.

As simple warm up examples, let us compute the second and third Renyi entropy. Trivial book keeping shows that the ensemble averages of the second and third power of the density matrix are given by the sum of two and five Wick contractions, respectively

$$\operatorname{tr}\Big(\rho^2_{\text{CFT}}\Big) = \operatorname{tr}\Big(\boldsymbol{C}^\dagger\boldsymbol{C}\boldsymbol{C}^\dagger\boldsymbol{C}\Big) = \overline{\boldsymbol{C}^\dagger\boldsymbol{C}}\,\overline{\boldsymbol{C}^\dagger\boldsymbol{C}} + \overline{\boldsymbol{C}^\dagger\overline{\boldsymbol{C}\boldsymbol{C}^\dagger}\boldsymbol{C}} \tag{1.6}$$

$$\operatorname{tr}(\rho^3_{\text{CFT}}) = \overline{\boldsymbol{C}^\dagger\boldsymbol{C}}\,\overline{\boldsymbol{C}^\dagger\boldsymbol{C}}\,\overline{\boldsymbol{C}^\dagger\boldsymbol{C}} + \overline{\boldsymbol{C}^\dagger\boldsymbol{C}}\,\overline{\boldsymbol{C}^\dagger\overline{\boldsymbol{C}\boldsymbol{C}^\dagger}\boldsymbol{C}} + \overline{\boldsymbol{C}^\dagger\overline{\boldsymbol{C}\boldsymbol{C}^\dagger}\,\overline{\boldsymbol{C}\boldsymbol{C}^\dagger}\boldsymbol{C}} + \overline{\boldsymbol{C}^\dagger\overline{\boldsymbol{C}\boldsymbol{C}^\dagger}\boldsymbol{C}}\,\overline{\boldsymbol{C}^\dagger\boldsymbol{C}} + \overline{\boldsymbol{C}^\dagger\overline{\boldsymbol{C}\overline{\boldsymbol{C}^\dagger\boldsymbol{C}}\boldsymbol{C}^\dagger}\boldsymbol{C}}$$

Here we used the large N condition that Wick contractions only take place between operators with a common summation index. The five Wick contractions for the computation of the third Renyi entropy are illustrated in the figure below.

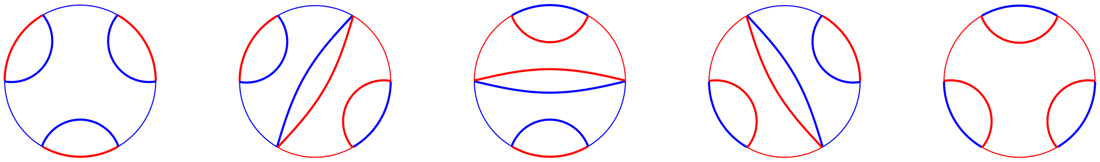

Figure 1: The five Wick contractions of $\operatorname{tr}(\rho^3_{\text{CFT}})$. The blue lines represent the radiation states, the red lines the CFT states and the black lines are the initial matter states.

Plugging in the result (1.3) and (1.4) for the elementary Wick contractions, we find that the expression decomposes into a sum over energies

$$\operatorname{tr}\Big(\rho^2_{\text{CFT}}\Big) = \sum_E \Big(q^2_E + q_E p_E\Big), \qquad \operatorname{tr}\Big(\rho^3_{\text{CFT}}\Big) = \sum_E \Big(q^3_E + 3q^2_E p_E + q_E p^2_E\Big)\,. \tag{1.7}$$

This simple structure follows from the fact that both elementary Wick contractions produce density matrices that are diagonal in the energy basis. This structure persists for all higher Renyi entropies.

The higher Renyi entropies can be computed in an identical manner. The combinatorics is straighforward. For the n-th Renyi entropy, the sum over Wick contractions splits into a

$$\operatorname{tr}\Big(\rho^n_{\text{CFT}}\Big) = \sum_E \sum_{m=0}^{n-1} N(n,m)\, q_E^{n-m} p_E^m \tag{1.8}$$

where $N(n,m)$ counts the number of terms with $n-m$ Wick contractions of the type (1.3) and m Wick contractions of the type (1.4). It turns out that this is a well known combinatorical

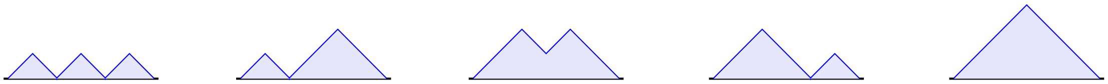

Figure 2: The five Wick contractions of $\mathrm{tr}(\rho^3_{\mathsf{CFT}})$ represented as five Dyck paths.

problem. We can graphically represent the five Wick contractions for $\mathrm{tr}(\rho^3_{\mathsf{CFT}})$ in terms of so-called Dyck paths as shown in the figure below A Dyck path can take steps up or down, but is required to stay above the starting point. A step up is the start and a step down is the end of a Wick contraction. The number of possible Dyck paths of $2n$ steps with m peaks are given by the Narayana numbers $N(n, m)$

$$N(n,m) = \frac{1}{n}\binom{n}{m-1}\binom{n}{m}. \tag{1.9}$$

As the next step, we write $\mathrm{tr}\Big(\rho^n_{\mathsf{CFT}}\Big)$ in terms of the Narayana polynomial

$$\mathrm{tr}\Big(\rho^n_{\mathsf{CFT}}\Big) = \sum_E q^n_{{}_E}\, \mathcal{N}_n(z_{{}_E}) \qquad\qquad \mathcal{N}_n(z) = \sum_{m=1}^{n} N(n,m) z^{m-1} \tag{1.10}$$

Using the fact that the Narayana polynomials can be written as an integral transform of the Legendre polynomials as $\mathcal{N}_n(z) = (z-1)^n \int_0^1 dx\, P_n(\frac{2zx}{z-1} - 1)$, and using the familiar the integral representation of Legendre functions, this yields the following expression for the nth Renyi entropy

$$S_n(\rho_{\mathsf{CFT}}) \;=\; \frac{1}{1-n}\log\left[\sum_E \frac{1}{2\pi}\int_{-\pi}^{\pi}\!d\theta \int_0^1\!dx\, \Big(q_{{}_E} f(x,\theta,z_{{}_E})\Big)^n\right] \tag{1.11}$$

with

$$f(x,\theta,z) \;=\; 1-z+2zx+\sqrt{4zx(1-z+zx)}\,\cos\theta \tag{1.12}$$

Finally, we take the $n \to 1$ limit, where we invoke L'Hopital's rule, we obtain a general expression for the von Neumann entropy,

$$S_{\mathrm{vN}}(\rho_{\mathsf{CFT}}) \;=\; -\sum_E \int_0^1 d\lambda_{{}_E} D(\lambda_{{}_E})\lambda_{{}_E}\log\lambda_{{}_E} \tag{1.13}$$

Here $D(\lambda_{{}_E})$ is the entanglement spectrum of the density matrix, given by

$$D(\lambda_{{}_E}) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\!d\theta\int_0^1\!dx\,\delta\Big(\lambda_{{}_E} - q_{{}_E} f(x,\theta,z_{{}_E})\Big) \tag{1.14}$$

$$= \frac{1}{2\pi\lambda_{{}_E}}\sqrt{4q_{{}_E}p_{{}_E} - (q_{{}_E}+p_{{}_E}-q_{{}_E}p_{{}_E}\lambda_{{}_E})^2} \tag{1.15}$$

plus a delta function contribution at $\lambda_{{}_E} = 0$. This is Page's result.