Quantum Information Scrambling in Black Holes with Symmetry

Navya Gupta

An essay submitted for partial fulfillment of Perimeter Scholars International

June, 2020

Contents

Quantum Information Scrambling in Black Holes with Symmetry

Navya Gupta

Supervisor: Beni Yoshida

Hayden and Preskill studied information processing in a black hole assumed to have unitary and chaotic dynamics [1]. They demonstrated that quantum information deposited into a black hole past its Page point is rapidly released by Hawking radiation. In this essay, we review the use of the decoupling inequality in quantum information for demonstrating information recoverability in the Hayden-Preskill setup. Then we extend this setup to incorporate symmetry preserving dynamics. We show that if the black hole's initial state belongs to a symmetric subspace, information is still rapidly released by the Hawking radiation. Finally, we comment on extending the notion of decoupling to von Neumann algebras.

Statement of original research

This essay contains original research in section 3.2.

1 Introduction

Understanding the information content of the Hawking radiation from an evaporating black hole is of crucial importance. It is believed that black holes are the most powerful information scramblers in the Universe [2]. Thus, one could naively expect that the radiation coming out from the black hole contains little or no information. While this may be true for early radiation, Page has shown that the late radiation from a black hole rapidly releases information [3]. Hayden and Preskill [1] refine this analysis to show that a particular quantum of information thrown into a black hole past its Page time is almost instantly revealed by radiation. It seems that far from hiding information, the scrambling nature of black hole dynamics crucially contributes to the information content of the Hawking radiation. At the heart of this surprising result, lie deep connections between the notions of scrambling, chaos, typicality, and quantum error correction. In this essay, we shall explore the relation between typicality and quantum error correction.

In section 1, we set the stage for the Hayden-Preskill thought experiment by providing a brief overview of the black hole information loss problem and the black hole complementarity paradigm. In section 1.3, we set up important definitions and notation. After introducing the Hayden-Preskill setup in section 2.1, we present the decoupling inequality and discuss information recoverability in section 2.2. In section 3, we motivate symmetric black hole dynamics, derive the corresponding decoupling bound and demonstrate information recoverability for a few examples. In section 4, we sketch some initial thoughts about extending these notions to von Neumann algebras. We conclude with some discussion about future research avenues in section 5.

1.1 A Brief History of Black Hole Information

Black holes [4] are a consequence of Einstein's general relativity. They arise when sufficiently dense concentrations of matter and energy bend spacetime so severely that even light is unable to escape its gravitational field. The event horizon acts as a "one-way causality barrier", which hides a true singularity behind itself. While objects or information can flow past the event horizon towards this singularity, nothing can ever come out of a classical black hole.

Curiously enough, the connections between black holes and information go deeper than these considerations about signalling. From classical general relativity alone, it is possible to derive the four laws of black hole mechanics, which are analogous to the laws of thermodynamics [5]. By the 1970s, it was known that black holes tend to increase horizon surface area when undergoing various transformations (such as the Penrose process in Kerr black holes) [6, 7]. Comparing this monotonicity of horizon area to the monotonicity of thermodynamic entropy, Bekenstein proposed that the entropy of a black hole is proportional to its horizon area [8]. According to the no-hair conjecture [4], all equilibrium black hole solutions to the Einstein and Maxwell field equations are fully characterized by just three classical observables - mass M, electric charge Q, and angular momentum $\vec{L}$. However, a black hole with a given value of M, Q and $\vec{L}$ could have formed in various ways - from the collapse of a normal star or the collapse of a neutron star or perhaps the collapse of a geon. In some sense, while M, Q and $\vec{L}$ characterize the observable macrostates of a black hole, the various histories of collapse characterize different possible (but inaccessible) internal configurations or microstates. Bekenstein proposed that the black hole entropy is a measure of this inaccessible information. When an object (which carries entropy) falls into a black hole, it adds to this inaccessible information and contributes to the growth of horizon area and black hole entropy.

Thus far, the black hole was treated as a classical object. In 1975, Hawking [9] considered the behaviour of vacuum fields on the curved black hole spacetime and predicted that black holes radiate particles like black bodies at some temperature which is determined by the values of M, Q and $\vec{L}$. This calculation enabled Hawking to determine the proportionality constant between black hole entropy and horizon area, thus demonstrating that the relation between entropy and area is not just an analogy but an identity. Since then, this notion of black hole entropy has come to be known as the Bekenstein-Hawking entropy. As a black hole evaporates by giving off Hawking radiation, it loses mass and consequently its horizon area and entropy. Yet the generalized second law introduced by Bekenstein [8] holds as the loss in horizon area is compensated by the entropy of the emergent radiation.

We have noted that the black hole microstates constitute inaccessible information. Furthermore, the outgoing Hawking radiation is thermal and essentially featureless (except that its temperature depends upon the macroscopic parameters of the black hole M,Q and $\vec{L}$). If a black hole were to evaporate completely by giving off Hawking radiation, all information characterizing its matter content would get lost (save for its imprint on the temperature of the Hawking radiation). This loss of information (and hence time reversibility) is at odds with the unitarity of quantum mechanics. Hence, it is claimed that this characterization of black holes is incompatible with quantum mechanics.

The community has expressed a wide spectrum of reactions to this claim. While some see this crisis as pushing us towards a critical re-examination of unitarity in quantum mechanics (which already contains irrevererersible processes such as measurements), others hold that the preservation of information is inviolable and quantum measurements are indeed reversible (as in the Many-Worlds paradigm). To support the latter belief, one must demonstrate that the black hole, in fact, does not destroy information. Of course, this would require one to go beyond Hawking's semiclassical calculations, and to assume the breakdown of these semiclassical calculations (and the necessity to apply a full-fledged theory of quantum gravity) at some point in the evaporation process. Evidence from string theory suggests that the apparently inaccessible information which is encoded into the black hole's internal degrees of freedom ultimately gets transferred to the outgoing radiation [10, 11]. This viewpoint is known as black hole complementarity [12]. Thus, unlike Hawking's prediction of uninformative thermal radiation, black hole complementarity posits that radiation bears information.

1.2 Black Hole Complementarity

According to the black hole complementarity hypothesis [12], a stationary observer far away from the black hole horizon (say Bob) and another observer falling into the black hole (say Alice) have two different (complementary) perspectives as Alice approaches the horizon. Due to the infinite time dilation at the horizon, Bob never sees Alice cross the horizon. Instead, from Bob's vantage point, the information characterizing Alice spreads out uniformly across the stretched horizon, which is a membrane hovering about one Planck length outside the event horizon. As static observers at the stretched horizon have a proper acceleration of order 1 in Planck units, they are surrounded by a bath of Unruh radiation carrying temperature of order 1. The radially outward directed quanta from this bath constitute Hawking radiation and reach infinity with a red-shifted temperature of $1/r_s$, where r_s is the Schwarzschild radius of the black hole. To the observer Bob, the stretched horizon appears as a physical membrane with temperature, entropy, viscosity, and electrical conductivity. Alice's information, which is encoded in the black hole's internal state, eventually leaks out in the form of Hawking radiation.

In contrast, Alice does not encounter anything special at the stretched horizon. She simply falls through and hits the singularity. Thus, information is both reflected from the horizon and passes through the horizon. Apparently, the black hole violates the no-cloning theorem [13]. This problem is dealt with by saying that no physical observer can detect the same information both inside and outside the horizon.

Several natural questions arise. How does Alice's information spread across the stretched horizon? How does radiation encode this information? How long does the information take to come out of the black hole? How do we ensure that no observer can detect cloning of information? We shall evoke the principles of quantum information to address these questions. But before we motivate the focus of this essay - the Hayden-Preskill thought experiment [1] - we introduce some useful notation.

1.3 Notation and Definitions

This section introduces the diagrammatic notation for quantum circuits that is used in this essay. Similar notation finds extensive use in quantum information literature. As shown in fig. 1, circuits run from bottom to top. A system A described by the Hilbert space $\mathcal{H}_A$ is represented by a wire of a particular color. A linear transformation $M : \mathcal{H}_A \rightarrow \mathcal{H}_A$ is represented by a box as shown in fig. 1 (a). If $\{|i\rangle\}$ forms an orthonormal basis for $\mathcal{H}_A$, then M has a decomposition $M_{ij} |i\rangle \langle j|$. Note that the summation convention has been employed to avoid cumbersome notation. As shown in figs. 1 (c) and 1 (d), traces are represented by joining the input and output legs of the tensor. The most general quantum operation is represented by completely positive trace preserving (CPTP) maps $\mathscr{T} : \mathcal{L}(\mathcal{H}_A) \rightarrow \mathcal{L}(\mathcal{H}_B)$. The terms quantum operation, quantum channel, and CPTP map are used interchangeably. The same wire notation is used to represent quantum channels as shown in fig. 1 (e). The sense in which the wire notation is used can be identified by the fact that quantum channels are represented in calligraphic letters.

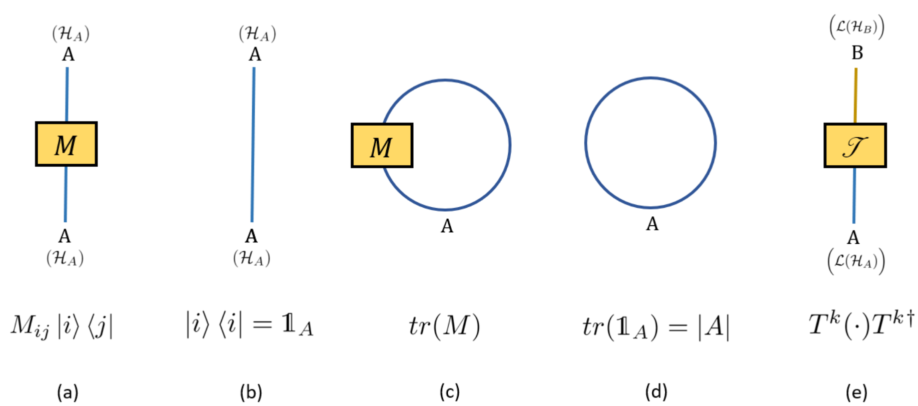

Figure 1: Diagrammatic representation of operators, traces, and channels.

As shown in fig. 2 (a), if the input wire of $M_{ij} |i\rangle \langle j|$ in fig. 1 is bent upwards, one obtains $M_{ij}\left|i\right\rangle\left|j\right\rangle$, the state representation of the operator M over the doubled Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_{A'}$. If the identity wire is bent similarly, one obtains a representation of an EPR state $\left|\phi\right\rangle_{AA'}$ over the systems AA', once the normalization factor has been included as shown in 2 (b). In general, in this notation, states are entities which only have output wires and no input wires. The density matrix $\phi_{AA'}$ corresponding to this EPR state is indicated in fig. 2 (c).

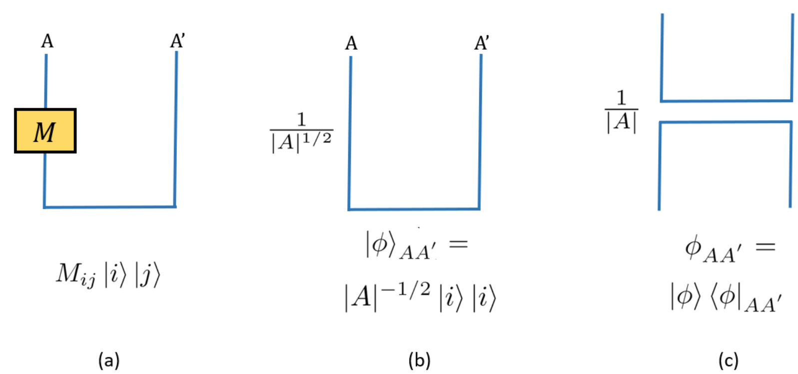

Figure 2: State-operator mapping and EPR pairs

The operator-state map introduced in 2 (a) can be extended to all CPTP maps $\mathscr{T}$. This is formally known as the Choi-Jamiolkowski isomorphism:

Definition 1 (Choi-Jamiolkowski Isomorphism [14, 15]). *The Choi-Jamiolkowski map J takes quantum operations $\mathscr{T} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ to operators $J(\mathscr{T}_{A\to B}) \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B)$. It is defined as:*

$$J(\mathscr{T}_{A\to B}) = (\mathbb{1}_A \otimes \mathscr{T}_{A'\to B'})\phi_{AA'} = \tau_{AB'} \tag{1}$$

where $\phi_{AA'}$ is the EPR state on the doubled Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_{A'}$ introduced above. This map is an isomorphism.

In fig. 3 (a) and (b), we see that both operators and channels can be composed by vertical juxtaposition. Horizontal juxtaposition represents the tensor product between states and operators. Fig. 3 (c) provides notation for Hilbert spaces which do not factorize. Here Q is an operator from the joint Hilbert space over systems AB to that of systems CD. Note that $\mathcal{H}_{AB} \cong \mathcal{H}_{CD}$ and AB, CD represent two different partitions of the same Hilbert space. We will see this in more detail in the computations that follow.

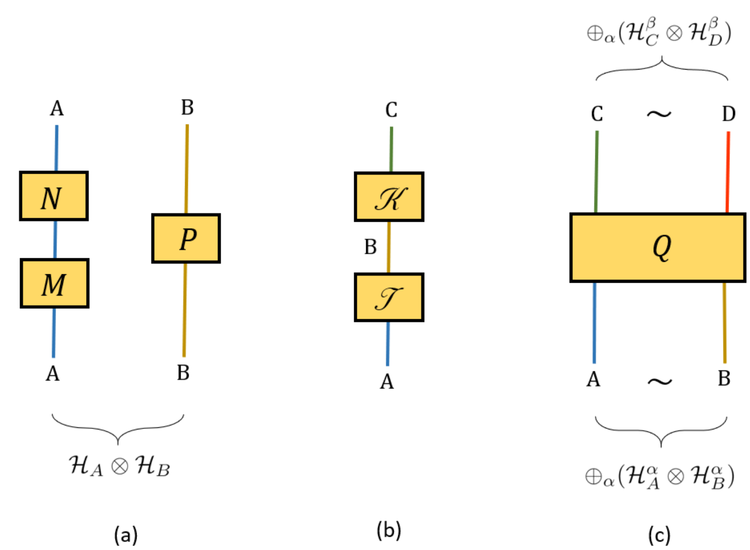

Figure 3: Diagrammatic representation of composition, tensor factors and direct sum decomposition

For a state ρ_{AB} on a Hilbert space $\mathcal{H}_{AB}$, which may or may not factorize into Hilbert spaces over systems A and B, ρ_A and ρ_B denote the reduced states on A and B respectively. In case $\mathcal{H}_{AB} \cong \mathcal{H}_A \otimes \mathcal{H}_B$, the reduced states are obtained by taking partial traces. The process of taking reduced states for the case when $\mathcal{H}_{AB}$ does not tensor factorize will be introduced later. Finally, the space of all density matrices (i.e, normalized positive-definite operators) over some Hilbert space $\mathcal{H}$ will be represented by $\mathcal{D}(\mathcal{H})$.

1.4 Modelling a Black Hole

Now we can begin addressing the questions raised in section 1.2. Recall that we are interested in understanding black hole dynamics in the complementarity paradigm. For the rest of the essay, we focus on one of the two complementary viewpoints - namely, the viewpoint of Bob, the observer located far outside the black hole horizon. Bob observes that the internal state of the black hole is encoded on the stretched membrane. In the spirit of Don Page [3, 16], we assume that the combined state of the black hole B and radiation R_B is described by a pure quantum state. Thus, immediately after gravitational collapse, the black hole starts out in a pure state B_i. Radiation is modelled by the black hole B_i discarding a random subsystem R_B. Equivalent to discarding a random subsystem is the application of a random unitary U followed by the discarding of a fixed subsystem R_B of B as shown in fig. 4. By a random unitary, we mean a unitary U sampled from the uniform Haar distribution over the group of unitaries acting on the Hilbert space of the black hole $\mathcal{H}_{B_i}$.

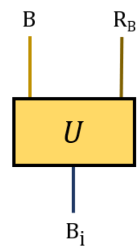

Figure 4: Page's model for black hole evaporation.

.

If $\psi_{BR_B}(U)$ describes the joint state of the black hole and radiation, we can define the following entropies:

1. $s_{BH} = \log |B|$, the Bekenstein-Hawking (thermodynamic) entropy of the black hole B.

2. $s_{R_B} = \log |R_B|$, the thermodynamic entropy of the radiation R_B.

3. $S_B = -\operatorname{tr} \psi_B \log \psi_B$, the von Neumann entropy of the black hole B.

4. $S_{R_B} = -\operatorname{tr} \psi_{R_B} \log \psi_{R_B}$, the von Neumann entropy of the radiation R_B.

Note that $S_B = S_{R_B}$ is also the entanglement entropy for B and R_B. Page defines the radiation's information content I by the deviation of the entanglement entropy from its maximum [3]. The maximum of the entanglement entropy is the thermodynamic entropy.

$$I = \log |R_B| - S_{R_B} \tag{2}$$

Here comes Page's crucial insight [16]: When U is Haar random, then the smaller of the two subsystems B and R_B is very nearly maximally entangled with a subsystem of the larger subsystem. We will come back to this result after we have discussed the decoupling inequality. For now, we use this insight to see that (fig. 5) while the information content of the radiation is zero in the beginning, information emerges rapidly after what is known as the Page time. The black hole reaches Page time when it has given up nearly half of its Bekenstein-Hawking entropy i.e, nearly half of its matter content.

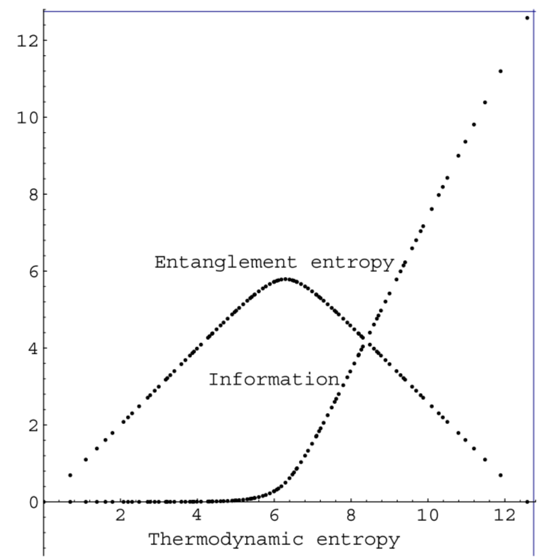

Figure 5: The page curve. This figure is reproduced from Page's 1993 paper. [3]

We can sharpen the question about information recovery by focusing on a special subsystem (say the one corresponding to Alice) in the initial black hole B_i, and asking how much radiation must Bob collect to gain information about Alice. Clearly, Bob must collect radiation till the Page time at the very least. After the Page time, if Bob collects just a little more radiation than the size of Alice's system, he would be able to reconstruct Alice's state with high fidelity. This is the content of the Hayden-Preskill thought experiment which we shall examine in the next section.

Note that all the arguments made above referred to a rather simplistic circuit model of a black hole. How do these assumptions fare vis-à-vis known/conjectured results about black holes? With reference to the picture presented in section 1.2, the random unitary U represents the black hole's dynamics which thermalizes Alice's information across the stretched horizon. It has been shown that black holes are the fastest information scramblers in nature [2, 17, 18]. In other words, information which is initially localized on the stretched horizon quickly gets thoroughly mixed with the entire system. This scrambling is a result of strongly coupled chaotic dynamics. As it turns out, random unitaries are often used to approximate such chaotic dynamics [19]. By encoding quantum information in a highly non-local manner, the black hole protects it from damage. In fact, it acts as an optimal encoder for quantum error correction, which enables information recovery [1]. These deep connections between randomness, chaos, and information scrambling lie at the heart of information processing and recoverability in the complementarity model [19]. We shall defer the discussion about timescales and no-cloning to a later section (2.3).

2 Hayden-Preskill Thought Experiment

In this chapter, we study the Hayden-Preskill (HP) thought experiment [1]. We introduce the setup in 2.1, study the role of decoupling in the HP setting in 2.2 and revisit complementarity in 2.3.

2.1 The Setup

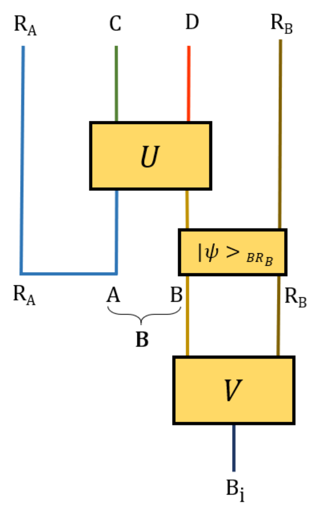

Figure 6: The HP setup with the initial radiation phase included.

We consider a black hole B which has already emitted some Hawking radiation R_B. The joint state of the black hole and radiation is represented by $\psi_{BR_B}(= |\psi\rangle\langle\psi|_{BR_B})$ as discussed in section 1.4. At this stage of radiation, Alice falls into the black hole with her quantum memory A. The black hole's system is now jointly described by the systems A, B and shall be denoted by $\mathbf{B}$. Let the initial state on A, B be represented by $\rho_{AB} = \rho_{\mathbf{B}}$. Now, the black hole's dynamics, represented by the Haar random unitary U, mixes A with B. Subsequently the black hole $\mathbf{B}$ undergoes another phase of Hawking radiation. Let D represent this new Hawking radiation and C represent the final black hole. The final state is represented by $\rho_{\mathbf{B}}(U) = \rho_{CD}$. Suppose that the observer Bob, who is sitting far away from the horizon, has collected both the old radiation R_B and the new radiation D. To what extent can he reconstruct Alice's state?

To address this question, it is convenient to introduce a reference state R_A which is maximally entangled with Alice's state A, as shown in fig. 6. Thus, the initial state over $R_A ABR_B$ is given by $\rho_{R_A ABR_B} = |\phi\rangle_{R_A A} \otimes |\psi\rangle_{BR_B}$, where $\mathrm{tr}_{R_A R_B}\, \rho_{R_A ABR_B} = \rho_{AB}$. Similarly, the final state would be represented by $\rho_{R_A CDR_B}$. We now claim that information about Alice's system A is contained in the purification of the reference system R_A. This statement is trivial for the intial state because the purification of R_A is A. However, after the unitary evolution, the state over R_A develops correlations with the states over C, D and R_B. Therefore, information about A is contained in some suitable subsystem of CDR_B which purifies R_A. For Bob to be able to recover this information from the new and old radiation DR_B alone, no part of R_A's purification should be supported on $\mathcal{H}_C$. This happens when the final state on R_A and C i.e, $\rho_{R_A C}(U)$ separates or decouples i.e, $\rho_{R_A C} \cong \rho_{R_A} \otimes \rho_C$. Using standard calculations in quantum Shannon theory [20], this is also a sufficient condition for information recovery. Furthermore, even if decoupling is achieved only up to a certain tolerance ϵ, Bob can reconstruct Alice's state with a minimal fidelity of $1 - \epsilon$, i.e,

$$\|\rho_{R_A C}(U) - \rho_{R_A} \otimes \rho_C\|_1 \leq \epsilon \implies F(\rho_A, \hat{\rho}_A) \geq 1 - \epsilon \tag{3}$$

Here, the one-norm for some operator M is defined by $\|M\|_1 = \mathrm{tr}\,\sqrt{M^\dagger M}$. The fidelity of two density matrices ρ and σ is defined by $F(\rho, \sigma) = \left[\mathrm{tr}\,\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right]^2$ [14]. The state $\hat{\rho}_A$ is constructed by Bob after applying suitable operations to the systems D, R_B. Thus, ϵ quantifies the degree of recoverability given a choice of initial state and dynamics U. We are now going to use a familiar trick in information theory i.e, compute the average of ϵ over the dynamics U. A low value of this average implies that information can be recovered to a reasonable degree for a typical unitary sampled from the Haar ensemble. The decoupling inequality is the tool that we use for this purpose.

2.2 The Decoupling Inequality

The decoupling inequality characterizes the conditions under which the quantum correlations between two systems get destroyed. Say, we have a bipartite quantum system (such as $\rho_{\mathbf{B}R_A}$ from the previous section). We now apply some quantum operation on one system ($\mathbf{B}$) mapping it to the final system C, while leaving the other system (R_A) untouched. This results in the loss of correlations between the two systems (C and R_A). This is a result of the data processing inequality in quantum information [14, 20].

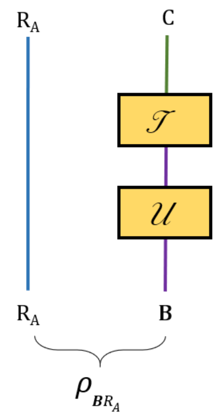

Figure 7: Setup for the decoupling inequality. $\mathscr{U}(\cdot) : \mathcal{L}(\mathcal{H}_{\mathbf{B}}) \to \mathcal{L}(\mathcal{H}_{\mathbf{B}})$, given by $\rho_{\mathbf{B}} \mapsto U\rho_{\mathbf{B}}U^{\dagger}$, is followed by a completely positive trace preserving map $\mathscr{T}$: $\mathcal{L}(\mathcal{H}_{\mathbf{B}}) \to \mathcal{L}(\mathcal{H}_C)$, which is the partial trace over the system D in the HP setting. This is the first instance in this essay where the wire notation is used to represent maps between linear operators (i.e, superoperators) instead of maps between states. Also, note that the system R_B does not enter these considerations.

In the HP setting, this quantum operation consists of the unitary transformation U (the black hole thermalization) followed by a partial trace over system D (Hawking radiation). Composed together, these map a state on $\mathbf{B}$ (the black hole together with Alice's system) to a state on C (the remaining black hole) as shown in fig. 7. The system R_A, on the other hand, is left untouched. The decoupling inequality quantifies the extent to which the final state on R_AC decouples. Unitary transformations preserve quantum correlations. The partial trace operation is responsible for decoupling in this case. In fact, we can investigate decoupling under the application of any arbitrary CPTP map $\mathscr{T} : \mathcal{L}(\mathcal{H}_{\mathbf{B}}) \to \mathcal{L}(\mathcal{H}_C)$. The statement of the non-smoothed [1] decoupling inequality (Theorem 3.3 in [15]) is:

[1]For more details about smoothing, please refer to [15]. Some comments about the smooth decoupling inequality are also mentioned in section 5

Theorem 2. *Let $\rho_{\mathbf{B}R_A} \in \mathcal{D}(\mathcal{H}_{\mathbf{B}R_A})$, $U \in \mathcal{L}(\mathcal{H}_{\mathbf{B}})$, and let $\mathscr{T}_{\mathbf{B}\to C}$ be a completely positive map with the Choi-Jamiolkowski representation $\tau_{\mathbf{B}'C} = J(\mathscr{T})$. Then,*

$$\mathbf{E}_U\Big(\|\mathscr{T}_{\mathbf{B}\to C}(U\rho_{\mathbf{B}R_A}U^\dagger) - \tau_C \otimes \rho_{R_A}\|_1\Big) \leq 2^{-\frac{1}{2}H_2(\mathbf{B}|R_A)_\rho} 2^{-\frac{1}{2}H_2(\mathbf{B}|C)_\tau} \tag{4}$$

where $\mathbf{E}_U(\cdot)$ is the average over the Haar measure over the unitary group on $\mathcal{H}_{\mathbf{B}}$, and $H_2(A|B)_\rho$ is the conditional collision entropy defined by:

$$H_2(A|B)_\rho = \sup_{\sigma_B \in \mathcal{D}(\mathcal{H}_B)} -\log \operatorname{tr}\left[\Big((\mathbb{1}_A \otimes \sigma_B^{-1/4})\rho_{AB}(\mathbb{1}_A \otimes \sigma_B^{-1/4})\Big)^2\right] \tag{5}$$

The interested reader is referred to [15] for a proof of this inequality. For now, we just mention that the invariance properties of the Haar measure together with the Schur-Weyl duality are key ingredients in this proof. Let us digest the theorem's statement by breaking it down:

1. The left hand side of the inequality is the average trace distance between the state $\mathscr{T} \circ \mathscr{U}(\rho_{\mathbf{B}R_A})$ obtained as a result of dynamics specified by U and the fully decoupled state $\tau_C \otimes \rho_{R_A}$. While $\rho_{R_A} = \operatorname{tr}_{\mathbf{B}} \rho_{\mathbf{B}R_A}$, τ_C is completely determined by the channel $\mathscr{T}_{\mathbf{B}\mapsto C}$.

2. The right hand side (call it Δ) is the average value of the tolerance ϵ that we mentioned in the previous section. As a measure of correlations between R_A and C, Δ breaks into two factors:

   (a) $\Delta_\rho = 2^{-\frac{1}{2}H_2(\mathbf{B}|R_A)_\rho}$, which is determined by the initial state $\rho_{\mathbf{B}R_A}$ alone and quantifies the the correlations between R_A and $\mathbf{B}$ that we begin with.

   (b) $\Delta_\tau = 2^{-\frac{1}{2}H_2(\mathbf{B}|C)_\tau}$, which is determined by the channel $\mathscr{T}_{\mathbf{B}\mapsto C}$ alone, and quantifies the extent to which the channel preserves these correlations.

3. The conditional collision entropy of A given B i.e, $H_2(A|B)_\rho \geq -\log\big(|A|\big)$. The lower bound is achieved by maximally entangled states. Thus, the smaller the value of this entropy, the larger the quantum correlations between A and B. The negativity of quantum conditional entropy is a signature of quantum correlations. [14].

Now we are ready to compute the right hand side of the decoupling inequality for the HP setup described in section 2.1. We will denote the decoupling bounds for the HP setting by $\Delta^{HP} = \Delta_\rho^{HP}\Delta_\tau^{HP}$.

Computing the channel term Δ_τ^{HP}

Recall that in the HP setting, the quantum channel $\mathscr{T}^{HP} : \mathcal{L}(\mathcal{H}_{\mathbf{B}}) \to \mathcal{L}(\mathcal{H}_C)$ in fig. 7 is given by $\rho_{\mathbf{B}} \mapsto \mathrm{tr}_D\, \rho_{\mathbf{B}}$ in the HP setting. The CJ (Choi-Jamiolkowski) representation for this channel is given by (1):

$$J(\mathscr{T}^{HP}) = \tau_{\mathbf{B}'C} = (\mathbb{1}_{\mathbf{B}'} \otimes \mathscr{T}^{HP}_{\mathbf{B}\to C})(\phi_{\mathbf{B}'\mathbf{B}}) = (\mathbb{1}_{C'D'} \otimes \mathrm{tr}_C)(\phi_{C'D'CD}) \tag{6}$$

where $\phi_{\mathbf{B}'\mathbf{B}}$ is the EPR state over $\mathcal{H}_{\mathbf{B}'} \otimes \mathcal{H}_{\mathbf{B}}$. A graphical representation of this state is given in fig. 8.

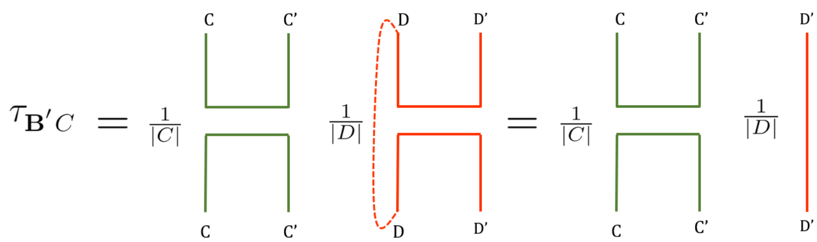

Figure 8: Choi-Jamiolkowski representation for the partial trace channel

The reduced state on C, which features on the left hand side of the decoupling inequality is given by:

$$\tau_C = \mathrm{tr}_{\mathbf{B}'}\, \tau_{\mathbf{B}'C} = \frac{\mathbb{1}_C}{|C|} = \rho_C^{\max} \tag{7}$$

where $\rho_C^{\max}$ refers to the maximally mixed state on $\mathcal{H}_C$. We compute Δ_τ^{HP} as follows:

$$\begin{aligned}\Delta_\tau^{HP} = 2^{-H_2(\mathbf{B}|C)_\tau} &= \inf_{\sigma_C\in\mathcal{D}(\mathcal{H}_C)} \mathrm{tr}\left[\left((\mathbb{1}_{C'D'}\otimes\sigma_C^{-1/4})\tau_{\mathbf{B}'C}(\mathbb{1}_{C'D'}\otimes\sigma_C^{-1/4})\right)^2\right]\\ &= \inf_{\sigma_C\in\mathcal{D}(\mathcal{H}_C)} \frac{1}{|D|}\frac{1}{|C|^2}\left[\mathrm{tr}\left(\sigma_C^{-1/2}\right)\right] = \frac{|C|}{|D|}\end{aligned} \tag{8}$$

The infimum is achieved for the maximally mixed state $\sigma_C = \frac{\mathbb{1}_C}{|C|}$.

Computing the initial state term Δ_ρ^{HP}

Recall from fig. 6 that the intial state in the HP thought experiment is given by $\rho_{R_AABR_B} = \phi_{R_AA} \otimes |\psi\rangle_{BR_B}$. As promised, we now reproduce Page's result about black hole evaporation (presented in section 1.4) by using the decoupling inequality. If we say that $\psi_{BR_B}(V) = |\psi\rangle\langle\psi|_{BR_B}$ and $\psi_B^{\max}$, $\psi_{R_B}^{\max}$ are the maximally mixed states over B,R_B respectively, then:

$$\begin{aligned} \mathbf{E}_V\Big(\|\operatorname{tr}_{R_B}\psi_{BR_B}(V) - \psi_B^{\max}\|_1\Big) &\leq \frac{|B|}{|R_B|} \quad \text{(Tracing out } R_B\text{)} \\ \mathbf{E}_V\Big(\|\operatorname{tr}_{B}\psi_{BR_B}(V) - \psi_{R_B}^{\max}\|_1\Big) &\leq \frac{|R_B|}{|B|} \quad \text{(Tracing out } B\text{)} \end{aligned} \tag{9}$$

Notice that in the above applications of the decoupling inequality, there is no reference system. But we will still refer to the bound on the right as a "decoupling bound". We see that when $|R_B| < |B|$, ψ_{R_B} is nearly maximally mixed and when $|B| < |R_B|$, ψ_B is nearly maximally mixed. This is equivalent to Page's result. With this, we compute the initial state term Δ_ρ^{HP} as follows:

$$\begin{aligned} \Delta_\rho^{HP} &= 2^{-H_2(\mathbf{B}|R_A)_\rho} = 2^{-H_2(AB|R_A)_\rho} \\ &= \inf_{\sigma_{R_A}\in\mathcal{D}(\mathcal{H}_{R_A})} \operatorname{tr}\left[\Big((\mathbb{1}_{AB}\otimes\sigma_{R_A}^{-1/4})\rho_{ABR_A}(\mathbb{1}_{AB}\otimes\sigma_{R_A}^{-1/4})\Big)^2\right] \\ &= \inf_{\sigma_{R_A}\in\mathcal{D}(\mathcal{H}_{R_A})} \operatorname{tr}\left[\Big((\mathbb{1}_{AB}\otimes\sigma_{R_A}^{-1/4})(\phi_{AR_A}\otimes\psi_B)(\mathbb{1}_{AB}\otimes\sigma_{R_A}^{-1/4})\Big)^2\right] \\ &= \inf_{\sigma_{R_A}\in\mathcal{D}(\mathcal{H}_{R_A})} \operatorname{tr}\left[\Big((\sigma_{R_A}^{-1/4}\phi_{AR_A}\sigma_{R_A}^{-1/4})\otimes\psi_B\Big)^2\right] \\ &= \inf_{\sigma_{R_A}\in\mathcal{D}(\mathcal{H}_{R_A})} \operatorname{tr}\left[(\sigma_{R_A}^{-1/4}\phi_{AR_A}\sigma_{R_A}^{-1/4})^2\right]\operatorname{Tr}\big(\psi_B^2\big) \\ &= |A|\operatorname{tr}\big(\psi_B^2\big) = \begin{cases} \frac{|A|}{|R_B|}, & \text{if } |R_B| < |B| \text{ (Before Page time)} \\ \frac{|A|}{|B|}, & \text{if } |R_B| \geq |B| \text{(After Page time)} \end{cases} \end{aligned} \tag{10}$$

Decoupling bound for Hayden-Preskill Thought Experiment

Now we are ready to put together the expressions obtained in 8 and 10 to obtain the decoupling bound Δ^{HP}.

$$\mathbf{E}_U\Big(\| \operatorname{tr}_D U\rho_{\mathbf{B}R_A}U^\dagger - \rho_C^{\max}\otimes\rho_{R_A}\|_1\Big) \leq \begin{cases} \sqrt{\frac{|C|}{|D|}\frac{|A|}{|R_B|}} = \frac{|A|}{|D|}\sqrt{\frac{|B|}{|R_B|}}\,, & |R_B| < |B| \\ \sqrt{\frac{|C|}{|D|}\frac{|A|}{|B|}} = \frac{|A|}{|D|}\,, & |R_B| \geq |B| \end{cases} \tag{11}$$

If we consider that each system R in this problem is constituted by n_R qubits, then $\Delta^{HP} = 2^{n_A - n_D}$ after the Page time (i.e, when $|R_B| \geq |B|$). This bound acquires a very small value when n_D exceeds n_A by just a few more $O(1)$ qubits. Thus, an old enough black hole spits out Alice's information rather quickly. On the other hand, for a black hole before the Page time, the bound is enhanced by a factor of $\sqrt{\frac{|B|}{|R_B|}}$. This number is large when $|R_B| < |B|$, and Bob must wait for the black hole to at least reach its Page time before any decoupling (and information leakage) happens [2].

To summarize, Page showed that nearly all information about a black hole's internal state is carried by the radiation that it emits after half its lifetime. Hayden and Preskill refined this result by showing that information about an n_A-qubit quantum system thrown into a black hole past its Page time is revealed by just a few more than n_A qubits of radiation. Thus, they demonstrated the role of the black hole's age in hiding information. Whereas new black holes conceal information for at least the Page time, old black holes "mirror" the information almost instantly.

2.3 Revisiting Complementarity

In section 1.2, we had mentioned that black holes are great scramblers or delocalizers of quantum information. A Haar random dynamics reasonably approximates such a mixing dynamics. Another important consideration is the timescale for thermalization and information retention. So far, we have only stated the information retention time in terms of the number of qubits that must be released before information is revealed. We must map this to physical time. Further, for the Hayden-Preskill setup to make sense in the first place, the thermalization timescale must be shorter than the time interval between the radiation of two successive quanta. Following the discussions

[2] In saying this, we assume that decoupling does not happen when the decoupling bound is too big. In other words, we assume that a small value of the decoupling bound is not just a sufficient condition for decoupling but also a necessary one. This is expected to be true in a certain sense which is discussed in section 5

in [1], we make a few brief comments about these issues.

The unitary dynamics corresponds to the uniform spreading of Alice's information on the stretched horizon. Classically, if an electric charge droplet is introduced on the stretched horizon, it spreads uniformly and the charge density decays exponentially like e^{-t/r_s}, where t is the Schwarzschild time and r_s is the Schwarzschild radius. This means that the droplet uniformly envelops the horizon in a time of order $r_s \log r_s$. If we model information spreading like the spreading of the charge droplet, we see that it does thermalize fairly rapidly.

The time interval between the emission of successive radiation quanta is of the order of r_s in Schwarzschild time. Thus, information about n_A qubits thrown into a black hole past its Page time is revealed on a timescale of $O(n_A r_s)$. Since information is released fairly rapidly by old black holes, it is important to check that the information retention time is still long enough for the complementarity hypothesis to hold i.e, Bob cannot verify the cloning of quantum information by first reconstructing Alice's state outside the horizon and then comparing it with Alice's state inside the horizon [3]. Computations in [1] show that this is just barely true.

3 The Thought Experiment with Symmetries

In this section, we consider the Hayden-Preskill scenario for black holes with symmetry. By symmetry, we mean that the dynamics of the black hole conserves the value of some observable like angular momentum, charge, or energy. Specifically, we will consider a $U(1)$ symmetry in the following section. This $U(1)$ symmetry could represent charge, spin, or it could even approximate energy subspaces as in [21]. Due to the symmetry constraint, we expect the dynamics to be less mixing. Given that information recovery relies upon the thorough mixing of degrees of freedom, it is important to check the extent to which it can be recovered in the symmetric case. We shall see that the introduction of symmetries results in Hilbert spaces that do not factorize. Decoupling on non-factorizable Hilbert spaces is an interesting conceptual question. Symmetries naturally give rise to algebraic descriptions of quantum systems. A description of decoupling in an algebraic language would be interesting in

[3]This statement assumes that Bob's reconstruction protocol is fast or even instantaneous. But in reality, obtaining the purification of the reference system from its highly diffused form in the radiation can have immense computational complexity. And this complexity of decoding may be enough to justify that the complementarity hypothesis cannot be violated.

light of making contact with quantum field theory. Some very preliminary ideas about expressing decoupling in the language of Type I von Neumann algebras are presented in section 4.

There are several motivations for studying symmetric black holes. They provide an interesting case study for symmetries in quantum gravity [22, 23]. In [21], the imposition of a $U(1)$ symmetry allowed the resolution of an apparent tension between Hayden-Preskill recoverability and the Almheiri-Marolf-Polchinski-Sully (AMPS) thought experiment [24]. Charged black holes feature in the formulation of the weak gravity conjecture [25]. It is also of interest to study symmetric black hole models to compare them with charged SYK models [23].

3.1 The Setup

Recall that after Alice falls into the black hole, the black hole is represented by the Hilbert space $\mathcal{H}_{\mathbf{B}} \cong \mathcal{H}_{AB} \cong \mathcal{H}_{CD}$. In other words, Alice's system A and the initial black hole B on the one hand, and the remaining black hole C and the new radiation D on the other, represent two different ways of partitioning the Hilbert space of the total black hole system $\mathbf{B}$. In the cases considered so far, both these partitions involved tensor factors i.e, $\mathcal{H}_{AB} \cong \mathcal{H}_A \otimes \mathcal{H}_B$ and $\mathcal{H}_{CD} \cong \mathcal{H}_C \otimes \mathcal{H}_D$. The chaotic black hole dynamics was represented by a Haar random unitary $U \in \mathcal{L}(\mathcal{H}_{\mathbf{B}})$. As we shall see, the introduction of a symmetry requires us to modify both these assumptions. For simplicity, we illustrate this fact using a concrete example with qubits and a $U(1)$ symmetric dynamics.

Each system R will be described by n_R qubits. The corresponding Hilbert space will be denoted by $\mathcal{H}_R$. Let us say that the Hilbert space of a qubit is spanned by the vectors $|0\rangle$ and $|1\rangle$. We define the charge operator on the qubit Hilbert space as:

$$\hat{Q} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{12}$$

The joint Hilbert space of n_R qubits is spanned by vectors of the form $|a_1, a_2, ..., a_{n_R}\rangle$, where $a_i \in \{0, 1\}$. The charge operator on this n_R-qubit Hilbert space is defined by the sum of the charge operators on individual qubit spaces: $\hat{Q}_R = \sum_{i=1}^{n_R} \hat{Q}_i$ (relevant tensor factorizations by identity are implicit). Thus, the charge on the state $|a_1, a_2, ..., a_{n_R}\rangle$ is given by the number of 1's in the string $(a_1, a_2, ..., a_{n_R})$. We define $\mathcal{H}_R^q \subseteq \mathcal{H}_R$ to be the subspace spanned by those n_R qubit states which carry a total charge of q. The dimension of $\mathcal{H}_R^q$ will be represented by $|R^q|$. With this notation, the total black hole Hilbert space can be expressed as:

$$\mathcal{H}_{\mathbf{B}} \cong \oplus_{Q=0}^{n_{\mathbf{B}}} \mathcal{H}_{\mathbf{B}}^Q \cong \oplus_{Q=0}^{n_{\mathbf{B}}} \mathcal{H}_{AB}^Q \cong \oplus_{Q=0}^{n_{\mathbf{B}}} \mathcal{H}_{CD}^Q \tag{13}$$

Now, let us suppose that the black hole dynamics U conserves charge i.e, $[U, \hat{Q}_{\mathbf{B}}] = 0$. This means that U has a block-diagonal decomposition of the form $U = \oplus_{Q=0}^{n_{\mathbf{B}}} U_Q$, where each $U_Q \in \mathcal{L}(\mathcal{H}_{\mathbf{B}}^Q)$ is Haar random.

Assuming that $n_B \geq n_A$ and $n_C \geq n_D$, the Hilbert spaces $\mathcal{H}_{AB}^Q$ and $\mathcal{H}_{CD}^Q$ further decompose as:

$$\begin{aligned} \mathcal{H}_{AB}^Q &\cong \oplus_{q=0}^{n_A} \mathcal{H}_A^q \otimes \mathcal{H}_B^{Q-q} \cong \oplus_{q=0}^{n_A} \mathcal{H}_{AB}^{q,Q-q} \\ \mathcal{H}_{CD}^Q &\cong \oplus_{q=0}^{n_D} \mathcal{H}_C^{Q-q} \otimes \mathcal{H}_D^q \cong \oplus_{q=0}^{n_A} \mathcal{H}_{CD}^{Q-q,q} \end{aligned} \tag{14}$$

The symmetries of the initial state, if any, must be specified independently of the symmetry of the dynamics. The symmetries of the final state, on the other hand, would depend both upon the symmetries of the initial state and the symmetry of the dynamics. Depending upon these specifications, several cases arise, and they correspond to restricting our attention to different subspaces $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$ of $\mathcal{H}_{AB}$ and $\mathcal{H}_{CD}$ respectively:

1. Both the initial black hole B and Alice's system A have fixed charges q_A and q_B respectively, with $q_A + q_B = Q$. Then, $\mathcal{H}_{in} \cong \mathcal{H}_A^{q_A} \otimes \mathcal{H}_B^{q_B}$ and $\mathcal{H}_{out} \cong \mathcal{H}_{CD}^Q \cong \oplus_{q=0}^{n_D} \mathcal{H}_C^{Q-q} \otimes \mathcal{H}_D^q$.

2. The initial black hole B has a fixed charge q_B while Alice's system can have any charge in $\{0, 1, ..., n_A\}$. Then, $\mathcal{H}_{in} \cong (\oplus_{q_A=0}^{n_A} \mathcal{H}_A^{q_A}) \otimes \mathcal{H}_B^{q_B}$ and $\mathcal{H}_{out} \cong \oplus_{q_A=0}^{n_A} \mathcal{H}_{CD}^{q_A+q_B} \cong \oplus_{q_A=0}^{n_A} \oplus_{q=0}^{n_D} \mathcal{H}_C^{q_A+q_B-q} \otimes \mathcal{H}_D^q$.

3. The total charge of A and B is equal to Q. Then, $\mathcal{H}_{in} \cong \mathcal{H}_{AB}^Q \cong \oplus_{q=0}^{n_A} \mathcal{H}_A^q \otimes \mathcal{H}_B^{Q-q}$ and $\mathcal{H}_{out} \cong \mathcal{H}_{CD}^Q \cong \oplus_{q=0}^{n_A} \mathcal{H}_C^q \otimes \mathcal{H}_D^{Q-q}$.

4. There are no restrictions. Then, $\mathcal{H}_{in} \cong \mathcal{H}_{AB}$ and $\mathcal{H}_{out} \cong \mathcal{H}_{CD}$.

In the next section, we shall compute the specific form of Dupuis et al.'s [15] non-smooth decoupling inequality (4) for case 1. All other cases require generalizations of the inequality. In cases 2, 3 and 4, we would need the notion of partial decoupling on non-factorizable Hilbert spaces. For studying cases 2 and 4, one would also need to compute an average over unitaries with a special block diagonal decomposition. The non-randomized partial decoupling inequality developed in [26] encompasses all these cases.

Cases 1 and 2 were considered in [21], where it was concluded that $n_D >> n_A$ is a sufficient condition for recovery in case 1, while in case 2, there are certain states for which recovery is not possible. While [21] illustrated these results with some direct computations of Haar averages, we wish to reproduce a subset of those results i.e, case 1 by applying the decoupling inequality.

Note: Now that we are restricting attention to the subspace $\mathcal{H}_{\mathbf{B}}^{Q}$ alone, we shall often drop the Q superscript/subscript and restore it whenever it is convenient. We shall set $Q-q=\tilde{q}$. Thus $\mathcal{H}_{out} \cong \oplus_{q=0}^{n_A} \mathcal{H}_C^{\tilde{q}} \otimes \mathcal{H}_D^{q} \cong \oplus_{q=0}^{n_A} \mathcal{H}_{CD}^{\tilde{q},q}$.

3.2 Modified Decoupling Bound

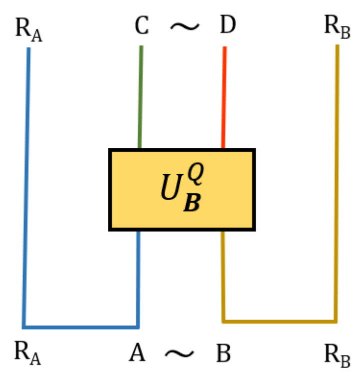

Figure 9: Hayden-Preskill setup for $U(1)$ symmetric dynamics with fixed charge Q.

Fig. 9 presents the setup for the HP experiment with symmetric dynamics. The input and output Hilbert spaces are given by $\mathcal{H}_{in} \cong \mathcal{H}_{AB}^{Q} \cong \mathcal{H}_{out} \cong \mathcal{H}_{CD}^{Q}$. Although we identify $\mathcal{H}_{in}$ to be $\mathcal{H}_{AB}^{Q}$ to match the dimensions of $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$, we shall eventually restrict ourselves to initial states on $\mathcal{H}_A^{qA} \otimes \mathcal{H}_B^{qB} \subseteq \mathcal{H}_{in}$, as in case 1. The initial state on AR_A is the EPR state defined on $\mathcal{H}_A^{q_A} \otimes \mathcal{H}_{R_A}$ (where $|R_A| = |A^{q_A}|$) i.e, $\phi_{A^{q^A} R^A}$. The black hole is considered to be past the Page point and it is assumed that it has a maximally mixed state on the space $\mathcal{H}_B^{q^B}$. More will be said about the initial state of the black hole later in this section. We will refer to the decoupling bound for the symmetric case as Δ^{sym}. As before, we will split it into $\Delta^{sym} = \Delta_\rho^{sym} \Delta_\tau^{sym}$, where Δ_ρ^{sym} and Δ_τ^{sym} are the terms corresponding to the initial state and channel respectively.

Computing the channel term Δ_τ^{sym}

Since $\mathcal{H}_{CD}^{Q}$ does not simply tensor factor into Hilbert spaces representing systems C and D, we are going to introduce a generalized partial trace operation denoted by $\hat{\mathrm{Tr}}_D$. Given a state $\sigma_{CD} = \oplus_q \sigma_{CD}^q$, it is defined by $\hat{\mathrm{Tr}}_D \sigma_{CD} = \oplus_q \mathrm{tr}_{D^q} \sigma_{CD}^q$, where tr_{D_q} is the usual partial trace. In general, the state $\rho_{CD}(U_{\mathbf{B}}^Q)$ need not have a block diagonal decomposition. Hence we shall define:

$$\tilde{\rho}_{CD}(U_{\mathbf{B}}^Q) = \sum_{q=0}^{n_D} \pi_{CD}^q \, \rho_{CD}(U_{\mathbf{B}}^Q) \, \pi_{CD}^q = \oplus_{q=0}^{n_D} \Big(\rho_{CD}(U_{\mathbf{B}}^Q)\Big)^q \tag{15}$$

where $\pi_{CD}^q \in \mathcal{L}(\mathcal{H}_{\mathbf{B}}^Q)$ is the projector onto the Hilbert space $\mathcal{H}_{CD}^{\tilde{q},q}$ and $\Big(\rho_{CD}(U_{\mathbf{B}}^Q)\Big)^q$ represents the (unnormalised) state on the q^{th} block. We define the state of the remaining black hole C as:

$$\rho_C = \hat{\mathrm{Tr}}_D \tilde{\rho}_{CD}(U_{\mathbf{B}}^Q) = \sum_{q=0}^{n_D} \hat{\mathrm{Tr}}_D\Big(\pi_{CD}^q \, \rho_{CD}(U_{\mathbf{B}}^Q) \, \pi_{CD}^q\Big) = \oplus_{q=0}^{n_D} \mathrm{tr}_{D^q} \Big(\rho_{CD}(U_{\mathbf{B}}^Q)\Big)^q \tag{16}$$

Finally, we define the CPTP map $\mathscr{T}^s : \mathcal{L}\Big(\mathcal{H}_{\mathbf{B}}^Q\Big) \to \mathcal{L}\Big(\oplus_{q=0}^{n_D} \mathcal{H}_C^{\tilde{q}}\Big)$ as:

$$\mathscr{T}^s(\rho) = \sum_{q=0}^{n_D} \hat{\mathrm{Tr}}_D\Big(\pi_{CD}^q \, \rho \, \pi_{CD}^q\Big) \text{ for } \rho \in \mathcal{L}(\mathcal{H}_{\mathbf{B}}^Q) \tag{17}$$

The 's' superscript is used to distinguish this from the simple partial trace channel $\mathscr{T}^{HP}$ considered before. We now compute the CJ representation for the channel $\mathscr{T}^s$:

$$
\begin{aligned}
J\big(\mathscr{T}^s\big) &= \tau^s_{\mathbf{B}'C} = \tau^s_{(C'D')C} = (\mathbb{1}_{\mathbf{B}'} \otimes \mathscr{T}^s)(\phi_{\mathbf{B}'\mathbf{B}}) \\
&= \sum_{q=0}^{n_D} \hat{\mathrm{Tr}}_D\Big(\pi^q_{CD}\ \phi_{\mathbf{B}'\mathbf{B}}\ \pi^q_{CD}\Big) \\
&= \sum_{q=0}^{n_D} \hat{\mathrm{Tr}}_D\Big(\pi^q_{CD}\ \phi_{(C'D')(CD)}\ \pi^q_{CD}\Big) \\
&\overset{a}{=} \sum_{q=0}^{n_D} \frac{|\mathcal{H}^{q,\tilde{q}}_{CD}|}{|\mathcal{H}^Q_{\mathbf{B}}|} \hat{\mathrm{Tr}}_D\Big(\phi_{(C'^{\tilde{q}}D'^q)(C^{\tilde{q}}D^q)}\ \ \oplus_{m\neq n} 0_{(C'^{\tilde{m}}D'^m)(C^{\tilde{n}}D^n)}\Big) \\
&= \oplus_{q=0}^{n_D} \frac{|\mathcal{H}^{q,\tilde{q}}_{CD}|}{|\mathcal{H}^Q_{\mathbf{B}}|}\Big(\ \mathrm{tr}_{D^q}\, \phi_{(C'^{\tilde{q}}D'^q)(C^{\tilde{q}}D^q)}\Big)\ \ \oplus_{m\neq n} 0_{(C'^{\tilde{m}}D'^m)(C^{\tilde{n}})} \\
&\overset{b}{=} \oplus_{q=0}^{n_D}\ p_q\ \tau^q_{(C'D')(C)}\ \ \oplus_{m\neq n}\ \ 0_{(C'^{\tilde{m}}D'^m)(C^{\tilde{n}})}
\end{aligned}
\tag{18}
$$

Equality (a) is obtained by the fact that conjugation by the projector π^q_{CD} sifts out only those terms in the EPR density matrix $\phi_{\mathbf{B}'\mathbf{B}} = \phi_{(C'D')(CD)}$ that belong to the subspace $\mathcal{L}(\mathcal{H}^{q,\tilde{q}}_{C'D'} \otimes \mathcal{H}^{q,\tilde{q}}_{CD})$, giving the EPR state over $\mathcal{H}^{q,\tilde{q}}_{C'D'} \otimes \mathcal{H}^{q,\tilde{q}}_{CD}$ denoted by $\phi_{(C'^{\tilde{q}}D'^q)(C^{\tilde{q}}D^q)}$ together with the normalization factor $\frac{|\mathcal{H}^{q,\tilde{q}}_{CD}|}{|\mathcal{H}^Q_{\mathbf{B}}|}$. The appended 0 matrices represent the subspaces which have been annihilated by the projector. For equality (c), the normalization factor is redefined as p_q while the normalized state $\mathrm{tr}_{D^q}\, \phi_{(C'^{\tilde{q}}D'^q)(C^{\tilde{q}}D^q)}$ is defined as $\tau^q_{(C'D')(C)}$. Note that $\tau^q_{(C'D')(C)}$ is precisely the Choi-Jamiolkowski representation of the partial trace channel that was obtained in section 2.2.

Recall that in the left hand side of the decoupling inequality, we must also specify the τ^s_C, the reduced state on C in the decoupled state. In this case we have:

$$
\tau^s_{C^Q} = \mathrm{tr}_{\mathbf{B}'}\, \tau^s_{\mathbf{B}'^Q C^Q} = \oplus_{q=0}^{n_D}\ p^q\ \tau^q_C = \oplus_{q=0}^{n_D}\ p^q\ \rho^{q,\max}_C
\tag{19}
$$

$\rho^{q,\max}_C$ is the maximally mixed state on $\mathcal{H}^{\tilde{q}}_C$.

Blockwise Conditional Collision Entropy

Recall the definition of the conditional collision entropy (eq. 5) which is reproduced below:

$$\begin{aligned} H_2(C'D'|C)_{\tau^s} &= \sup_{\sigma_C \in \mathcal{D}(\mathcal{H}_C)} -\log \operatorname{Tr}\left[\left((\mathbb{1}_{C'D'} \otimes \sigma_C^{-1/4})\tau^s_{(C'D')C}(\mathbb{1}_{C'D'} \otimes \sigma_C^{-1/4})\right)^2\right] \\ &= \sup_{\sigma_C \in \mathcal{D}(\mathcal{H}_C)} H_2(C'D'|C)_{\tau^s|\sigma} \end{aligned} \tag{20}$$

where the definition of $H_2(C'D'|C)_{\tau^s|\sigma}$ is obvious from the above expressions. The state τ^s, computed in eq. 18, has a very special block-diagonal form. This motivates us to obtain an upper bound to the Δ_τ factor in the decoupling inequality as follows. In the symmetric case, Δ_τ is given by:

$$\Delta_{\tau^s} = 2^{-H_2(\mathbf{B}|C)_{\tau^s}} \;=\; 2^{-\sup_{\sigma_C \in \mathcal{D}(\mathcal{H}_C)} \; H_2(C'D'|C)_{\tau^s|\sigma}} \tag{21}$$

An upper bound for this expression can be obtained by taking the supremum only over a subset of $\mathcal{D}(\mathcal{H}_C)$. We consider the subset to be the collection of those σ_C which have a block-diagonal decomposition of the form $\sigma_C = \oplus_{q=0}^{n_D} r^q \, \sigma_C^q$. We will call this subset $\mathcal{D}'(\mathcal{H}_C)$ Here, σ_C^q are normalized states over $\mathcal{H}_C^{\tilde{q}}$ and $\{r_q\}$ are probability weights across the q blocks. Therefore we define:

$$H_2^B(C'D'|C)_{\tau^s} = \sup_{\sigma_C \in D'(\mathcal{H}_C)} H_2(C'D'|C)_{\tau^s|\sigma} \tag{22}$$

We will refer to this entity as the "blockwise" conditional collision entropy because it involves taking the infimum over block diagonal states. H_2^B is a lower bound for H_2, and hence can be used for defining an upper bound for Δ_{τ^s}. We will denote this upper bound by Δ_τ^{sym}.

$$\Delta_\tau^{sym} = 2^{-H_2^B(\mathbf{B}|C)_{\tau^s}} \geq \Delta_{\tau^s} \tag{23}$$

We now proceed to compute $H_2(C'D'|C)_{\tau^s|\sigma}$ for $\sigma_C \in D'(\mathcal{H}_C)$:

$$\begin{aligned} 2^{-H_2(C'D'|C)_{\tau^s|\sigma}} &= \operatorname{tr}\left[\left((\mathbb{1}_{C'D'} \otimes \sigma_C^{-1/4})\tau^s_{(C'D')C}(\mathbb{1}_{C'D'} \otimes \sigma_C^{-1/4})\right)^2\right] \\ &= \operatorname{tr}\left[\;\oplus_{q=0}^{n_D}\;\left((\mathbb{1}^q_{C'D'} \otimes (r^q\sigma_C^q)^{-1/4})p^q\tau^q_{(C'D')C}(\mathbb{1}^q_{C'D'} \otimes (r^q\sigma_C^q)^{-1/4})\right)^2\right] \\ &= \sum_{q=0}^{n_D} \frac{p_q^2}{r_q}\;\operatorname{tr}\left[\left(\mathbb{1}^q_{C'D'} \otimes (\sigma_C^q)^{-1/4})\tau^q_{(C'D')C}(\mathbb{1}^q_{C'D'} \otimes (\sigma_C^q)^{-1/4})\right)^2\right] \end{aligned} \tag{24}$$

Now we take the infimum:

$$
\begin{aligned}
2^{-H_2^B(C'D'|C)_{\tau^s}} &= \inf_{\sigma_C \in \mathcal{D}'(\mathcal{H}_C)} 2^{-H_2(C'D'|C)_{\tau^s|\sigma}} = \inf_{\{r_q\}} \inf_{\{\sigma_C^q\}} 2^{-H_2(C'D'|C)_{\tau^s|\sigma}} \\
&= \inf_{\{r_q\}} \inf_{\{\sigma_C^q\}} \sum_{q=0}^{n_D} \frac{p_q^2}{r_q} \operatorname{tr}\left[\left((\mathbb{1}_{C'D'}^q \otimes (\sigma_C^q)^{-1/4})\tau_{(C'D')C}^q(\mathbb{1}_{C'D'}^q \otimes (\sigma_C^q)^{-1/4})\right)^2\right] \\
&\overset{a}{=} \inf_{\{r_q\}} \sum_{q=0}^{n_D} \frac{p_q^2}{r_q} \inf_{\sigma_C^q} \operatorname{tr}\left[\left((\mathbb{1}_{C'D'}^q \otimes (\sigma_C^q)^{-1/4})\tau_{(C'D')C}^q(\mathbb{1}_{C'D'}^q \otimes (\sigma_C^q)^{-1/4})\right)^2\right] \\
&\overset{b}{=} \inf_{\{r_q\}} \sum_{q=0}^{n_D} \frac{p_q^2}{r_q} \frac{|C^{\tilde{q}}|}{|D^q|} \\
&= \left(\sum_{q=0}^{n_D} p_q \sqrt{\frac{|C^{\tilde{q}}|}{|D^q|}}\right)^2 = \left(\Delta_\tau^{sym}\right)^2
\end{aligned}
\tag{25}
$$

Equality (a) follows from the fact that σ_C^q are independent of each other for different choices of q. Equality (b) follows because the minimization problem over σ_C^q is precisely the same as the one considered in eq. 8. Note that when systems C and D are tensor factors there is only one term in the sum, and we retrieve the result for the non-symmetric case (eq. 8).

Computing the initial state term Δ_ρ^{sym}

Recall that in section 2.2, we had justified the Page curve by invoking the decoupling inequality. Now, we will present similar arguments for a charged black hole which is evaporating under the conservative dynamics described previously. Fig. 10 shows the circuit diagram for the evaporation of a charged black hole. The initial state of the black hole B_i is assumed to be pure and to carry a charge Q_0. Since the black hole's dynamics U is assumed to be charge conserving, the total charge on the radiation system R_B and remaining black hole B is given by Q_0. Therefore, as discussed before, the Hilbert space $\mathcal{H}_{BR_B}^{Q_0}$ has a direct sum decomposition, with each term in the direct sum corresponding to different ways of distributing the charge Q_0 over the systems B and R_B. Let the final state on BR_B be given by ψ_{BR_B}.

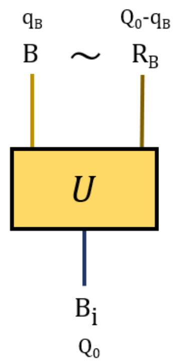

Figure 10: Page's black hole evaporation model for a charged black hole with charge-conserving dynamics.

Using the results of the previous section, the decoupling inequality after the black hole B_i has given up atleast half its degrees of freedom (i.e, when $|B| \leq |R_B|$) is given by:

$$\mathbf{E}_U\left(\|\hat{\mathrm{Tr}}_{R_B}\psi_{BR_B}(U) - \oplus_{q_B=0}^{n_B}\, w_{q_B}\psi_B^{q_B,\max}\|_1\right) \leq \sum_{q_B=0}^{n_B}\, w_{q_B}\sqrt{\frac{|B^{q_B}|}{|R_B^{Q_0-q_B}|}} = \Delta_{evap} \tag{26}$$

where $w_{q_B} = \frac{|B^{q_B}||R_B^{Q_0-q_B}|}{|B_i^{Q_0}|}$ are the probability weights across the different blocks in the decomposition of the Hilbert space for BR_B. A plot of these probability weights as a function of q_B for various values of the charge Q_0 is given in fig. 11 (a). The value of Δ_{evap} as a function of n_{R_B} for an initial black hole with 300 qubits and with various values of charge Q_0 is plotted in fig. 11 (b). The corresponding decoupling bound for the evaporation of the chargeless black hole discussed earlier is also shown in the same figure. We see that the decoupling bound crosses the value of 1 when half of the qubits in the initial black hole B_i have been given up in both the chargeless case and the case where the black hole has a charge Q_0 of 150. For the other Q_0 values of 125 and 170, this crossing happens when a few more than 150 qubits have been given up. Fig. 11 (c) shows the number of radiation qubits n_{R_B} that must be given up by the black hole B_i before the corresponding Δ_{evap} value crosses 1 as a function of Q_0. We note that this value of n_{R_B} is symmetric about $Q_0 = 150$, which corresponds to 50 % charge density. The approach to the sectorwise maximally mixed state is fastest for $Q_0 = 150$, and slows down the farther we get away from this value in either direction.

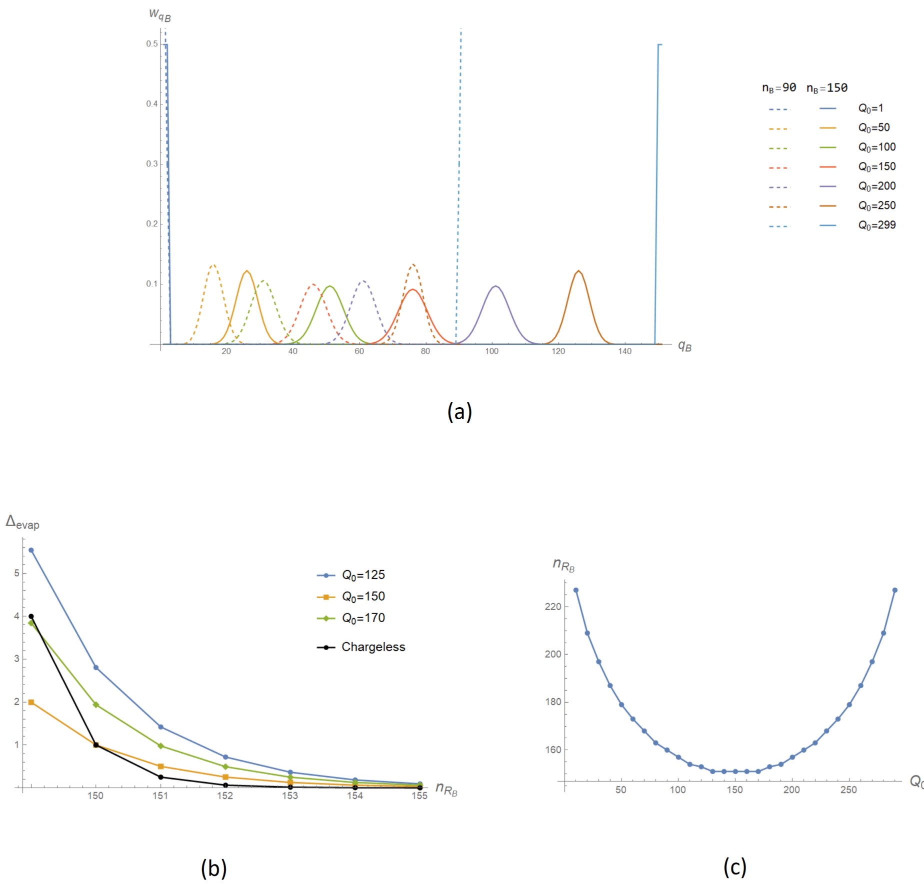

Figure 11: Evaporation of a charged black hole with initial size n_{B_i}=300 qubits and for various values of initial charge Q_0. (a) probability weights $\{w_{q_B}\}$ for $n_B = 90$ (dashed) and $n_B = 150$. (b) Decoupling bound Δ_{evap} as a function of n_{R_B}, the number of radiated qubits. (c) Value of n_{R_B} at which Δ_{evap} crosses 1 as a function of the initial charge Q_0.

Thus, we see that the black hole tends towards a sectorwise maximally mixed state as it gives off more and more radiation. The probability distribution across q_B sectors (fig. 11 (a)) is such that the average charge density on both B and R_B is the same as the charge density on B_i. The average charge on B is given by:

$$q_B^{av} = \sum_{q_B=0}^{n_B} w_{q_B} q_B \tag{27}$$

For the ease of computation, we approximate the sectorwise maximally mixed state on B by a maximally mixed state on the averaged charged q_B^{av} sector. A similar approximation was used in [27]. Thus, we can follow exactly the same calculations as in section 2.2 to compute:

$$\Delta_\rho^{sym} = 2^{-H_2(AB|R_A)_\rho} = \frac{|A^{q_A}|}{|B^{q_B^{av}}|} \tag{28}$$

Henceforth, we drop the "av" superscript from q_B^{av}. The sharper the peak representing the distribution $\{w_{q_B}\}$, the better justified this approximation. As we can see in fig. 11 (a), the peaks get sharper as the average charge density gets farther away from the 50 % charge density. But at such charge densities, the "decoupling" also gets worse as shown in fig. 11 (c). So, there is a trade-off here as we require both a small decoupling bound and a sharp probability distribution for justifying our choice of the initial black hole state. Fig. 11 (a) also shows that the peaks get sharper as n_B falls. As n_B falls, the decoupling bound also falls. Thus, both the decoupling bound and the averaged charge approximation get better with the black hole's age. This discussion spells out the regimes in which the results presented in the following section are applicable.

The final decoupling inequality

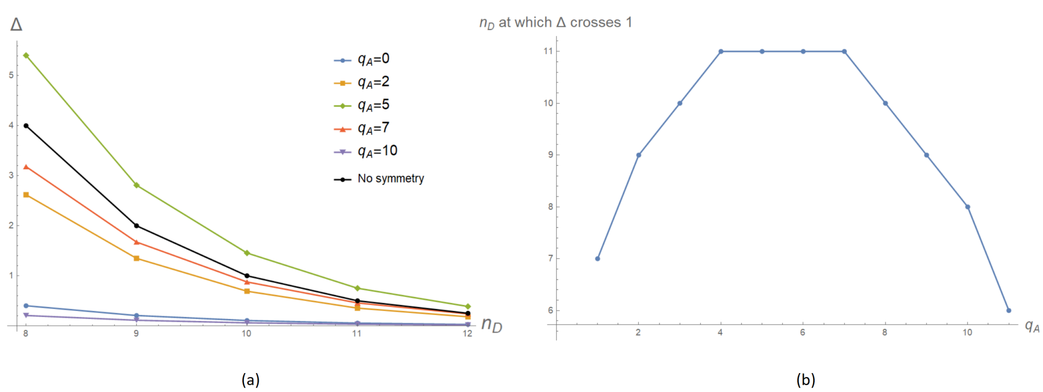

Figure 12: Decoupling when Alice's system with $n_A = 10$ quubits is dropped into a black hole with $n_B = 90$ qubits and charge $q_B = 50$. (a) Decoupling bound as a function of the number of new radiation qubits n_D for various values of q_A. (b) Value of n_D at which the decoupling bound crosses 1 as a function of q_A for the same setup.

Putting everything together, the final decoupling inequality is given by:

$$\mathbf{E}_U\left(\|\mathscr{T}^s_{\mathbf{B}^Q\to C^Q}(U^Q_{\mathbf{B}}\rho_{\mathbf{B}R_A}U^{Q\dagger}_{\mathbf{B}})-\tau^s_{C^Q}\otimes\rho_{R_A}\|_1\right)^2 \leq \frac{|A^{q_A}|}{|B^{q_B}|}\left(\sum_{q=0}^{n_D} p_q\sqrt{\frac{|C^{\tilde{q}}|}{|D^q|}}\right)^2 \quad (29)$$

where $p_q = \frac{|\mathcal{H}^{q,\tilde{q}}_{CD}|}{|\mathcal{H}^Q_{\mathbf{B}}|}$ and $q_A + q_B = Q$.

In fig. 12 (a), we plot the decoupling bound as a function of n_D for various values of q_A, with $n_B = 90$, $q_B = 50$ and $n_A = 10$. We also plot the decoupling bound for the non-symmetric case for the same value of n_A. We note that the bound falls rapidly in all these cases. Fig. 12 (b) plots the value of n_D at which the bound crosses 1 as a function of q_A. We see that the "slowest" decoupling happens when A has a 50 % charge density, and the decoupling gets "faster" as we move away from this charge density in either direction.

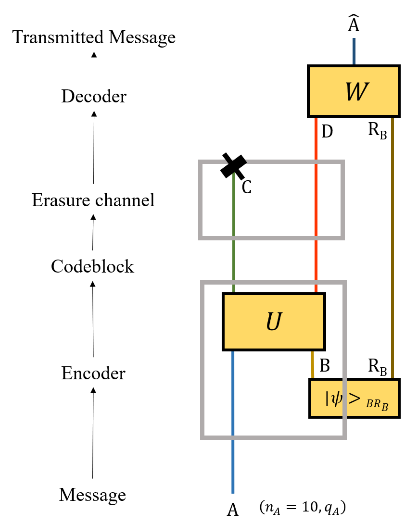

Figure 13: The HP setup as a standard quantum communication protocol.

This behaviour can be understood from the perspective of the quantum channel coding theorem [14]. Fig. 13 depicts the HP setup as a communication protocol which involves the transmission of an encoded message across an entanglement assisted quantum erasure channel.

The rate of transmission R across the channel is defined as the number of message qubits per encoded qubit i.e, $R = \frac{\log|A^{q_A}|}{\log|\mathbf{B}^Q|}$. The quantum capacity of the entanglement assisted erasure channel is given by $1 - p_{erase}$, where p_{erase} is the probability of erasure given by:

$$p_{erase} = \sum_{q=0}^{n_D} p_q \frac{|C^{Q-q}|}{|D^q|} \tag{30}$$

Fig. 14 plots the rate as a function of q_A as well as the values of the channel capacity corresponding to different numbers of radiated qubits n_D. We note that the n_D value corresponding to the decoupling bound crossing the value of 1 as shown in fig. 12 (b) corresponds to precisely that value of n_D at which the channel capacity first exceeds the required information transmission rate. This is a result of the fact that the even though the erasure channel has been modified slightly to preserve charge, the encodings are still random and hence optimal.

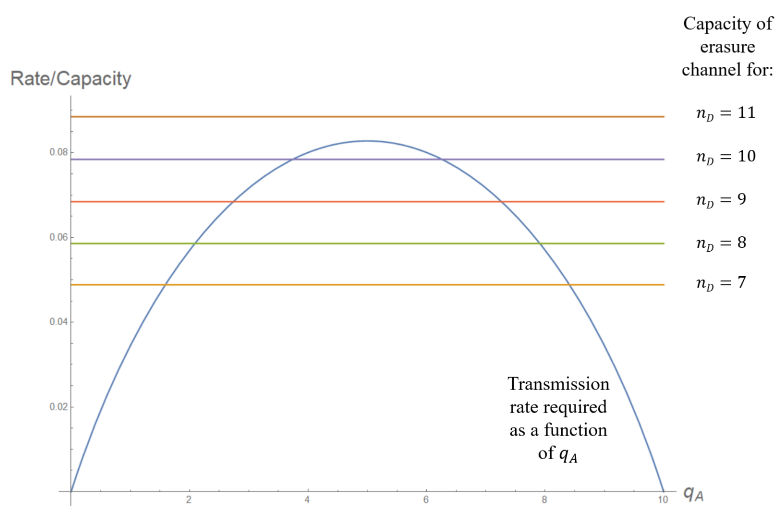

Figure 14: Channel capacity and rate. Blue curve shows the rate of message transmission as a function of q_A. Horizontal lines show the quantum capacity of the entanglement-assisted erasure channel for various values of n_D. Message transmission with low error is possible at capacities which are greater than or equal to the rate.

4 Towards an Algebraic Description

One important feature of symmetric dynamics is the non-factorizability of Hilbert spaces. Usually, systems are identified by some tensor factor of the

Hilbert space. In order to describe systems for non-factorizable Hilbert spaces, one needs to use the language of operator algebras. This section presents some preliminary ideas about expressing the decoupling inequality for type I von Neumann algebras on finite dimensional Hilbert spaces [4] [28]. Only those aspects of von Neumann algebras are introduced which are directly relevant to this report, and no rigorous definitions or proofs are provided. This section is meant as a motivation for future calculations. The presentation is inspired by Appendix A in [29].

A von Neumann algebra $\mathscr{A}$ on a Hilbert space $\mathcal{H}$ is a subset $\mathscr{A} \subseteq \mathcal{L}(\mathcal{H})$ which is:

1. Unital i.e, it contains the identity operator.
2. A vector space i.e, closed under addition and scalar multiplication.
3. Closed under multiplication and Hermitian conjugation.

The classification theorem for von Neumann algebras states that given any von Neumann algebra $\mathscr{A}$ on $\mathcal{H}$, there exists a decomposition of $\mathcal{H} \cong \oplus_\alpha \left(\mathcal{H}_1^\alpha \otimes \mathcal{H}_2^\alpha\right)$ such that:

$$\mathscr{A} = \{\oplus_\alpha \left(O_1^\alpha \otimes \mathbb{1}_2^\alpha\right) \mid O_1^\alpha \in \mathcal{L}(\mathcal{H}_1^\alpha)\ \forall\ \alpha\} \tag{31}$$

From this explicit representation of a von Neumann algebra, we clearly see that it satisfies all the properties listed in the definition above. The commutant of $\mathscr{A}$ is defined to be the algebra which contains all elements of $\mathcal{L}(\mathcal{H})$ which commute with all elements of $\mathscr{A}$. It is denote by $\mathscr{A}'$:

$$\mathscr{A}' = \{\oplus_\alpha \left(\mathbb{1}_1^\alpha \otimes O_2'^\alpha\right) \mid O_2'^\alpha \in \mathcal{L}(\mathcal{H}_2^\alpha)\ \forall\ \alpha\} \tag{32}$$

Finally, the centre $Z_{\mathscr{A}}$ corresponding to the algebra $\mathscr{A}$ is defined as the algebra $Z_{\mathscr{A}} = \mathscr{A} \cap \mathscr{A}'$:

$$Z_{\mathscr{A}} = \{\oplus_\alpha\ \lambda^\alpha(\mathbb{1}_1^\alpha \otimes \mathbb{1}_2^\alpha) \mid \lambda^\alpha \in \mathbb{C}\ \forall\ \alpha\} \tag{33}$$

Conversely, if we are given a particular Hilbert space decomposition such as $\mathcal{H}_{out} \cong \mathcal{H}_{CD} \cong \oplus_{q=0}^{n_D} \mathcal{H}_C^{\tilde{q}} \otimes \mathcal{H}_D^q$, we can deduce the corresponding algebras. For instance, the algebra representing the remaining black hole system C is given by:

[4] Henceforth, we drop the qualifiers type I and finite-dimensional Hilbert spaces because these always apply in this section. We will also often drop the qualifier von Neumann and simply say "algebra".

$$\mathscr{A}_C = \left\{ \oplus_{q=0}^{n_A} \ (O_C^{\tilde{q}} \otimes \mathbb{1}_D^q) \mid O_C^{\tilde{q}} \in \mathcal{L}(\mathcal{H}_C^{\tilde{q}}) \ \forall \ q \in \{0,1,..,n_D\} \right\} \tag{34}$$

This means that the set $\mathscr{A}_C$ contains the relevant observables for describing the system C. Similarly, the new radiation system D is described by the algebra $\mathscr{A}_D = \mathscr{A}_C'$. We clearly see that the centre of this algebra i.e, $Z_{\mathscr{A}_C} = \mathscr{A}_C \cap \mathscr{A}_D$ is non-trivial. In fact, the non-triviality of this centre is directly responsible for the non-factorizability of $\mathcal{H}_{CD}$. Likewise, using the decomposition $\mathcal{H}_{in} \cong \mathcal{H}_{AB} \cong \ \oplus_{q=0}^{n_A} \mathcal{H}_A^q \otimes \mathcal{H}_B^{\tilde{q}}$, we can deduce the algebras $\mathscr{A}_B$ and $\mathscr{A}_A$ representing the initial black hole and Alice's system respectively. We also define the algebras $\mathscr{A}_{AB} =< \mathscr{A}_A \cup \mathscr{A}_B >$ and $\mathscr{A}_{CD} =< \mathscr{A}_C \cup \mathscr{A}_D >$. Here $< ... >$ refers to the completion of the set under addition, multiplication, and Hermitian conjugation. More explicitly, we have:

$$\mathscr{A}_{CD} = \left\{ \oplus_{q=0}^{n_A} \ O_{CD}^{\tilde{q};q} \mid O_{CD}^{\tilde{q};q} \in \mathcal{L}(\mathcal{H}_C^{\tilde{q}} \otimes \mathcal{H}_D^q) \ \forall \ q \in \{0,1,..,n_D\} \right\} \tag{35}$$

$\mathscr{A}_{BC}$ is defined similarly. Note that even though $\mathcal{H}_{in} \cong \mathcal{H}_{out}$, $\mathscr{A}_{AB}$ and $\mathscr{A}_{CD}$ are two different algebras with distinct block diagonal decompositions.

We now introduce the notion of a state on an algebra. Say, we are given an arbitrary state $\rho \in \mathcal{L}(\mathcal{H})$. In general ρ contains more information than is needed for computing the expectation values for elements of the algebra $\mathscr{A}$ on $\mathcal{H}$. Therefore, we can define a unique state $\rho_{\mathscr{A}} \in \mathscr{A}$ such that $\mathrm{tr}(\rho A) = \mathrm{tr}(\rho_{\mathscr{A}} A)$ for all $A \in \mathscr{A}$ as follows:

$$\rho_{\mathscr{A}} = \int_{U \in \mathscr{A}'} dU \ U \rho U^\dagger = \oplus_\alpha \left(p_\alpha \rho_1^\alpha \otimes \frac{\mathbb{1}_2^\alpha}{|\mathcal{H}_2^\alpha|} \right) \tag{36}$$

Here the integral is carried out over the unitary elements of $\mathscr{A}'$ with the Haar measure. Here ρ_1^α is defined by considering the ρ^α block of ρ $\big[$which is an element of $\mathcal{L}(\mathcal{H}_1^\alpha \otimes \mathcal{H}_2^\alpha)$ $\big]$, tracing out the part on $\mathcal{H}_2^\alpha$, and dividing by the normalisation factor p_α. That is,

$$\begin{aligned} \rho_1^\alpha &= \frac{1}{\mathrm{tr}(\rho^\alpha)} \, \mathrm{tr}_{\mathcal{H}_2^\alpha} \, \rho^\alpha \\ p_\alpha &= \mathrm{tr}(\rho^\alpha) \end{aligned} \tag{37}$$

Though it is closely related to the partial trace, the operation of taking a state on an algebra is quite different. In particular, note that $\rho_{\mathscr{A}}$ is defined over the full Hilbert space $\mathcal{H}$ and not just the Hilbert space $\mathcal{H}_1 (= \oplus_\alpha \mathcal{H}_1^\alpha)$. Note that $\rho_{\mathscr{A}}$ contains a "classical part" represented by the probability distribution $\{p_\alpha\}$ and a "quantum part" represented by the normalized states $\left\{\rho_1^\alpha \otimes \frac{\mathbb{1}_2^\alpha}{|\mathcal{H}_2^\alpha|}\right\}$.

For the case in which input and output Hilbert spaces do not factorize, the analog of tracing out or discarding the system D is taking the state of the system on the algebra describing system C i.e, $\mathscr{A}_C$. For more clarity, we perform this operation in two steps as described below.

We denote the state on $\mathbf{B}$ after the application of the unitary $U_{\mathbf{B}}$ by $\rho_{AB}(U)$. At this stage, let us also remind ourselves that our states and dynamics are restricted to the charge Q sector and we have suppressed the Q index implicit in $U_{\mathbf{B}}^Q$. In the first step, we compute the state $\rho_{AB}(U)$ on the algebra $\mathscr{A}_{CD}$:

$$\Big(\rho_{AB}(U_{\mathbf{B}})\Big)^{\mathscr{A}_{CD}} = \int_{U \in \mathscr{A}'_{CD}} dU\; U \rho_{AB}(U_{\mathbf{B}}) U^\dagger = \sum_{q=0}^{n_D} \Big(\pi_{CD}^q\; \rho_{AB}(U_{\mathbf{B}})\; \pi_{CD}^q\Big) \tag{38}$$

Here $\pi_{CD}^q \in \mathcal{L}(\mathcal{H}_{CD})$ is the projector onto the Hilbert space $\mathcal{H}_{CD}^{\tilde{q},q}$. Note that the state $\Big(\rho_{AB}(U_{\mathbf{B}})\Big)^{\mathscr{A}_{CD}}$ has the natural block-diagonal decomposition of $\mathcal{H}_{CD}$. In the next step, we can readily compute the state over $\mathscr{A}_C$:

$$\Big(\rho_{AB}(U_{\mathbf{B}})\Big)^{\mathscr{A}_C} = \int_{U \in \mathscr{A}'_C} dU\; U \Big(\rho_{AB}(U_{\mathbf{B}})\Big)^{\mathscr{A}_{CD}} U^\dagger = \oplus_{q=0}^{n_D} \left(p_q \rho_C^q \otimes \frac{\mathbb{1}_D^q}{|D^q|}\right) \tag{39}$$

where ρ_C^q are normalized states over $\mathcal{H}_C^{\tilde{q}}$ and $p_q = \mathrm{tr}\Big(\pi_{CD}^q\; \rho_{AB}(U_{\mathbf{B}})\; \pi_{CD}^q\Big)$. Compare this to the definition of ρ_C introduced in the previous section:

$$\rho_C = \hat{\mathrm{Tr}}_D \tilde{\rho}_{CD}(U_{\mathbf{B}}^Q) = \sum_{q=0}^{n_D} \hat{\mathrm{Tr}}_D \Big(\pi_{CD}^q\; \rho_{CD}(U_{\mathbf{B}}^Q)\; \pi_{CD}^q\Big) = \oplus_{q=0}^{n_D} p_q \rho_C^q \tag{40}$$

Thus, ρ_C goes a step further from defining the state on the algebra $\mathscr{A}_C$. It also involves taking the partial trace (using the operator $\hat{\mathrm{Tr}}_D$ introduced before). To describe things in an algebraic language, we would have to develop a notion for the decoupling of states on different algebras.

4.1 Decoupling of Algebras

Consider a bipartite state $\rho_{MN} \in \mathcal{L}(\mathcal{H}_{MN})$, where $\mathcal{H}_{MN} \cong \mathcal{H}_M \otimes \mathcal{H}_N$. The state ρ_{MN} is decoupled when $\rho_{MN} = \rho_M \otimes \rho_N$, where $\rho_M = \mathrm{tr}_N\, \rho_{MN}$ and $\rho_N = \mathrm{tr}_M\, \rho_{MN}$. An equivalent characterization of decoupling can be given in terms of the von Neumann entropy, which is defined as:

$$S(\rho) = -\,\mathrm{tr}\,\rho \log \rho \tag{41}$$

The state ρ_{MN} is said to decouple when its von Neumann entropy is additive with respect to its susbsystems i.e,

$$S(\rho_{MN}) = S(\rho_M) + S(\rho_N) \tag{42}$$

This characterization of decoupling motivates a definition for the decoupling of von Neumann algebras. The entropy of a state ρ on a von Neumann algebra $\mathscr{A}$ is defined as follows [29]:

$$S(\rho, \mathscr{A}) = -\sum_{\alpha} p_\alpha \log p_\alpha + \sum_{\alpha} p_\alpha S(\rho_1^\alpha) = H(p_\alpha) + \sum_{\alpha} p_\alpha S(\rho_1^\alpha) \tag{43}$$

We recall that p_α and ρ_1^α feature in the definition of the state on an algebra, which is reproduced below:

$$\rho_{\mathscr{A}} = \int_{U \in \mathscr{A}'} dU\; U \rho U^\dagger = \oplus_\alpha \left(p_\alpha \rho_1^\alpha \otimes \frac{\mathbb{1}_2^\alpha}{|\mathcal{H}_2^\alpha|} \right) \tag{44}$$

The first term in the definition 43 is the classical Shannon entropy $H(p_\alpha)$ for the probability distribution across the algebra blocks indexed by α. The second piece is the von Neumann entropy of each block averaged over this classical probability distribution. We now propose the definition for the decoupling of a state on an algebra and its commutant:

Definition 3. *Given an algebra $\mathscr{A}_M$ and its commutant $\mathscr{A}_N$ on a Hilbert space $\mathcal{H}_{MN}$, such that they induce a decomposition $\mathcal{H}_{MN} = \oplus_\alpha\, \mathcal{H}_M^\alpha \otimes \mathcal{H}_N^\alpha$, and a state $\rho \in \mathcal{L}(\mathcal{H}_{MN})$, the state is said to decouple when:*

$$S(\rho, \mathscr{A}_{MN}) = S(\rho, \mathscr{A}_M) + S(\rho, \mathscr{A}_N) - H(p_\alpha) \tag{45}$$

Basically, if the above condition holds, then density matrices decouple on each block in the conventional sense. The extra Shannon entropy piece is a result of non-factorizability. It disappears when $\mathscr{A}_M$ and $\mathscr{A}_N$ define tensor factors. The above definition makes explicit use of the Hilbert space decomposition and we would like to make it more abstract. Once we have a nice formulation of decoupling, we would need to define a notion of distance between states on an algebra and go on to formulate the steps in the proof provided in [15] in an algebraic language. Given the lack of time, we defer these computations to future research.

5 Discussion

In this essay, we reviewed the decoupling approach to information recovery in evaporating black holes. After presenting Hayden and Preskill's original argument, we extended it to black holes with symmetric input states and symmetry preserving dynamics. We approximated the black hole's sectorwise maximally mixed state by a maximally mixed state over the average charge sector. Under these conditions, we showed that information is still recoverable. In fact, this recoverability is an outcome of the quantum channel coding theorem just as in Hayden and Preskill's original setup [1]. We now mention some avenues for further study.

First, while the decay of the decoupling bound to zero is clearly a sufficient condition for decoupling, it is important to ask whether this is also a necessary condition. In other words, we must understand how tight are the bounds given by the non-smooth decoupling theorem [15] that we have employed here. The smooth decoupling inequality introduced by Dupuis et al. in [15] is more appropriate for the one-shot setting that concerns us here. In fact, Dupuis et al. have shown that the condition for decoupling, when described in terms of smooth entropies, is necessary and sufficient provided the ϵ-smooth conditional min- and max-entropy for the channel $\mathscr{T}$ (and $\mathscr{T}^s$) coincide. In the asymptotic setting, where we consider an infinite number of copies of the setup, both the ϵ-smooth conditional min- and max-entropies converge to the conditional entropy $H(A|B)_\rho = H(AB)_\rho - H(B)_\rho$. Thus, the asymptotic decoupling inequality, which is given below, is optimal:

$$\lim_{n\to\infty}\left[\mathbf{E}_U\left(\|\mathscr{T}^{\otimes n}_{\mathbf{B}\to C}(U^{\otimes n}_{\mathbf{B}}\rho^{\otimes n}_{\mathbf{B}R_A}U^{\otimes n\dagger}_{\mathbf{B}}) - \tau^{\otimes n}_C \otimes \rho^{\otimes n}_{R_A}\|_1\right)\right]^{\frac{1}{n}} \leq 2^{-\frac{1}{2}H(\mathbf{B}|R_A)_\rho}2^{-\frac{1}{2}H(\mathbf{B}|C)_\tau} \tag{46}$$

where $\mathbf{E}_U$ represents the n-fold average. Even though we are not concerned with the i.i.d setting in this case, we present a plot of both the asymptotic and non-smooth decoupling bounds in fig. 15. We note that both are in good agreement. One must still perform a careful analysis of the smooth entropies.

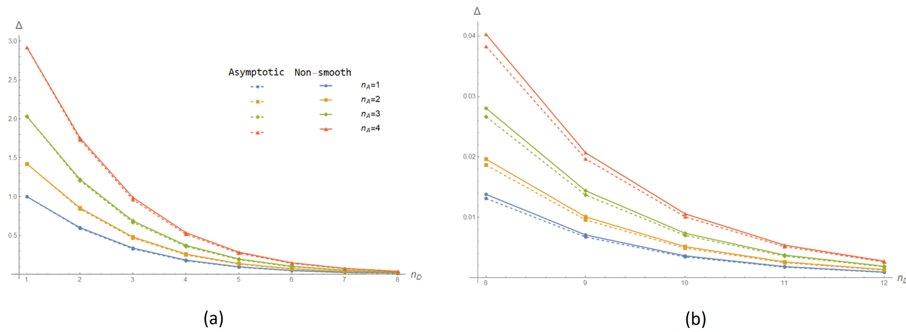

Figure 15: A plot of non-smooth and asymptotic decoupling bounds as a function of n_D for various values of n_A, $n_B = 100 - n_A$ and $q_B = 50$.

Second, we restricted our attention to initial states on a symmetric subspace i.e, a subspace carrying a fixed total charge Q. In fact, we went on further to restrict the individual charges of q_A and q_B. As shown in [21], things are not so straightforward once we allow non-symmetric input states. Interesting physics is revealed as information recovery fails for certain classes of input states. All this motivates the use of the partial decoupling inequality demonstrated in [26] for studying information recoverability in full generality. The generalization is two fold - Haar averaging is replaced by averaging over unitaries with special block decomposition, and complete decoupling over the factorizable Hilbert space $\mathcal{H}_{R_AC}$ is replaced by partial decoupling over a non-factorizable Hilbert space $\mathcal{H}_{R_AC}$.

The criterion for the recoverability of information was that the state on system R_AC decouples i.e, the purification of R_A be contained in the radiation system DR_B. This merely demonstrates the potentiality for recovering information from R_AB (hence the use of the word "recoverability"). It is a whole new ball game to describe actual recovery protocols which would require Bob to not only have access to the old and new radiation, but to also have knowledge of the black hole's dynamics, and the ability to extract the purification of R_A from its highly non-local distribution in the radiation. This raises several interesting questions about the time, physical resources, and computational power required to retrieve information. Explicit recovery protocols have been discussed in [21, 30].

Haar random dynamics is ultimately an approximation for the highly chaotic black hole dynamics. A signature of chaos is the scrambling or rapid delocalization of previously localized quantum information. While the decay of out-of-time-ordered correlators quantifies chaos [18, 31], the negativity of tripartite information quantifies scrambling [32]. The behaviour of these measures has been studied for the Hayden-Preskill setting in [21, 30]. It would be interesting to see how their behaviour changes once we introduce a conservative dynamics.

We sketched some preliminary ideas about extending the notion of decoupling to von Neumann algebras. There is a lot of work which needs to be done towards accomplishing this reformulation, but this looks like a very exciting avenue for future research. Stating decoupling algebraically would allow us to break free from the notion of qubits and factorizable Hilbert spaces, allowing a whole new realm of possibility. It could offer a path for generalizing the decoupling inequality to infinite dimensions and for making contact with quantum field theory.

Ultimately, this research paradigm is inspired by quantum gravity. It is amongst the many conceptual curiosities that arise at the interplay of quantum mechanics and general relativity. Thus, it is important to relate various entities in this toy model to physical entities. [1] presents an interesting discussion on the implications of the Hayden-Preskill thought experiment to the information loss problem. But as they emphasize, there are still plenty of missing pieces in this puzzle.

6 Acknowledgements

I am grateful to Beni Yoshida for his insightful supervision throughout this project. I thank the PSI fellows, course instructors, and my fellows PSIons for nurturing my love for physics. I am thankful to Debbie, Diana, Jamie and all those who have worked tirelessly to make PSI such a memorable experience.

Bibliography

[1] P. Hayden and J. Preskill, "Black holes as mirrors: quantum information in random subsystems," *Journal of high energy physics* **2007** no. 09, (2007) 120.

[2] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," *Journal of High Energy Physics* **2016** no. 8, (2016) 106.

[3] D. N. Page, "Information in black hole radiation," *Physical review letters* **71** no. 23, (1993) 3743.

[4] C. W. Misner, K. S. Thorne, J. A. Wheeler, and W. Gravitation, "Freeman and company," *San Francisco* (1973) 891.

[5] J. M. Bardeen, B. Carter, and S. W. Hawking, "The four laws of black hole mechanics," *Communications in mathematical physics* **31** no. 2, (1973) 161–170.

[6] D. Christodoulou, "Reversible and irreversible transformations in black-hole physics," *Physical Review Letters* **25** no. 22, (1970) 1596.

[7] S. W. Hawking, "Gravitational radiation from colliding black holes," *Physical Review Letters* **26** no. 21, (1971) 1344.

[8] J. D. Bekenstein, "Black holes and entropy," *Physical Review D* **7** no. 8, (1973) 2333.

[9] S. W. Hawking, "Particle creation by black holes," *Communications in Mathematical Physics* **43** no. 3, (1975) 199–220.

[10] A. Strominger and C. Vafa, "Microscopic origin of the bekenstein-hawking entropy," *Physics Letters B* **379** no. 1-4, (1996) 99–104.

[11] J. Maldacena, "The large-n limit of superconformal field theories and supergravity," *International journal of theoretical physics* **38** no. 4, (1999) 1113–1133.

[12] L. Susskind, L. Thorlacius, and J. Uglum, "The stretched horizon and black hole complementarity," *Physical Review D* **48** no. 8, (1993) 3743.

[13] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature* **299** no. 5886, (1982) 802–803.

[14] M. M. Wilde, *Quantum information theory.* Cambridge University Press, 2013.

[15] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, "One-shot decoupling," *Communications in Mathematical Physics* **328** no. 1, (2014) 251–284.

[16] D. N. Page, "Average entropy of a subsystem," *Physical review letters* **71** no. 9, (1993) 1291.

[17] Y. Sekino and L. Susskind, "Fast scramblers," *Journal of High Energy Physics* **2008** no. 10, (2008) 065.

[18] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," *Journal of High Energy Physics* **2014** no. 3, (2014) 67.

[19] D. A. Roberts and B. Yoshida, "Chaos and complexity by design," *Journal of High Energy Physics* **2017** no. 4, (2017) 121.

[20] J. Preskill, "Quantum shannon theory," *arXiv preprint arXiv:1604.07450* (2016) .

[21] B. Yoshida, "Soft mode and interior operator in the Hayden-Preskill thought experiment," *Physical Review D* **100** (2019) .

[22] D. Harlow and H. Ooguri, "Symmetries in quantum field theory and quantum gravity," *arXiv preprint arXiv:1810.05338* (2018) .

[23] J. Liu, "Scrambling and decoding the charged quantum information," *arXiv preprint arXiv:2003.11425* (2020) .

[24] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, "Black holes: complementarity or firewalls?" *Journal of High Energy Physics* **2013** no. 2, (2013) 62.

[25] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, "The string landscape, black holes and gravity as the weakest force," *Journal of High Energy Physics* **2007** no. 06, (2007) 060.

[26] E. Wakakuwa and Y. Nakata, "One-shot randomized and nonrandomized partial decoupling," in *2019 IEEE International Symposium on Information Theory (ISIT)*, pp. 3007–3011, IEEE. 2019.

[27] L. Piroli, C. Sünderhauf, and X.-L. Qi, "A random unitary circuit model for black hole evaporation," *Journal of High Energy Physics: JHEP* no. 4, (2020) .

[28] V. F. Jones, *von Neumann algebras in mathematics and physics.* American Mathematical Society, 1990.

[29] D. Harlow, "The ryu–takayanagi formula from quantum error correction," *Communications in Mathematical Physics* **354** no. 3, (2017) 865–912.

[30] B. Yoshida and A. Kitaev, "Efficient decoding for the hayden-preskill protocol," *arXiv preprint arXiv:1710.03363* (2017) .

[31] D. A. Roberts, D. Stanford, and L. Susskind, "Localized shocks," *Journal of High Energy Physics* **2015** no. 3, (2015) 51.

[32] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels," *Journal of High Energy Physics* **2016** no. 2, (2016) 4.