

Failure of the split property in gravity and the information paradox

Suvrat Raju

International Centre for Theoretical Sciences, Tata Institute of Fundamental Research, Shivakote, Bengaluru 560089, India.

E-mail: suvrat@icts.res.in

ABSTRACT: In an ordinary quantum field theory, the “split property” implies that the state of the system can be specified independently on a bounded subregion of a Cauchy slice and its complement. This property does not hold for theories of gravity, where observables near the boundary of the Cauchy slice uniquely fix the state on the entire slice. The original formulation of the information paradox explicitly assumed the split property and we follow this assumption to isolate the precise error in Hawking’s argument. A similar assumption also underpins the monogamy paradox of Mathur and AMPS. Finally the same assumption is used to support the common idea that the entanglement entropy of the region outside a black hole should follow a Page curve. It is for this reason that computations of the Page curve have been performed only in nonstandard theories of gravity, which include a nongravitational bath and massive gravitons. The fine-grained entropy at \mathcal{I}^+ does not obey a Page curve for an evaporating black hole in standard theories of gravity but we discuss possibilities for coarse graining that might lead to a Page curve in such cases.

Contents

1	The split property and its failure in gravity	3
2	The error in Hawking’s argument	7
3	The monogamy paradox	10
4	The Page curve	12
5	Conclusion	20

It is sometimes possible to make progress simply by discarding an incorrect assumption. In this essay, we will discuss the incorrect assumption that the state of a system, in a standard theory of quantum gravity, can be specified independently on a bounded region and on its complement.

It is not hard to see why this assumption, although incorrect, is appealing. First, such a property does hold for nongravitational quantum field theories, and can be formalized in terms of the “split property” [1–3] that we review below. The split property relies on the fact that, in nongravitational theories, the algebra of observables on a Cauchy slice can be separated into a subalgebra associated with a bounded region and a commuting subalgebra associated with its complement. Second, in classical gravity it is possible to find solutions that coincide outside a bounded region but differ inside [4]. This has led to a common belief that, even in quantum gravity, observables outside a bounded region do not give us much information about the state inside the region except for the energy and a small number of conserved charges.

However, it has recently been shown [5–7], in several examples, that the split property fails in theories of quantum gravity. Instead such theories obey the principle of “holography of information”, which states that all the information available on a Cauchy slice is also available near its boundary. In particular, this means that once all observables outside a bounded region have been specified there is *no freedom* to specify the state of the system inside the region. So the manner in which quantum gravity localizes information is the opposite of what the split property would suggest.

In this essay, we review how the split property has been assumed repeatedly in discussions of black-hole information — sometimes tacitly and sometimes explicitly — and this has often led to paradoxes. This assumption underpins Hawking’s original

argument for information loss [8, 9]. It was also a key ingredient in the monogamy paradox formulated by Mathur [10] and then elaborated by AMPS [11] that led to discussions of “firewalls” behind the horizon. More interestingly, this assumption also underlies the idea that the entropy of Hawking radiation should obey a “Page curve” [12, 13].

It is because this assumption is false in standard theories of gravity that computations of the Page curve [14–19] have been carried out only in nonstandard theories of gravity that involve a nongravitational bath, contain a massive graviton and do not obey the Gauss law. But these models of gravity do not provide an accurate understanding of the evaporation of black holes in standard gravity. We will point out how simply discarding the assumption of the split property leads to a more robust resolution of Hawking’s paradox.

This essay presents a brief summary of this perspective, with minimal technical details. We provide specific references to the literature in the text and refer the reader to the review [20] for a more detailed discussion and additional references. In section 1, we review the split property and its failure due to gravitational effects. In section 2, we isolate the precise point at which the split property was assumed in Hawking’s original argument for information loss and, in section 3 we describe the significance of this assumption in some later refinements of the paradox. We then explain, in section 4, how the same assumption led to the expectation that the entropy of the black hole exterior should follow the Page curve and review the significance of the nongravitational bath and massive gravitons in recent computations of the Page curve. In section 4, we also discuss some possibilities for coarse-grained sets of observables that might lead to a Page curve in standard theories of gravity.

However, one of the messages of this essay is that the failure of the split property in theories of gravity is not a bug that needs to be avoided. Rather, this unusual localization of quantum information in quantum gravity is an important physical lesson that emerges from a study of black-hole information. By suitable manipulations, it is possible to construct systems that lead to a Page curve and make a theory of gravity resemble nongravitational theories. But this exercise does not appear to be directly relevant for understanding the evaporation of black holes in a standard theory of gravity. In standard theories of gravity, black hole spacetimes store information in a manner that is more interesting than the Page curve would suggest.

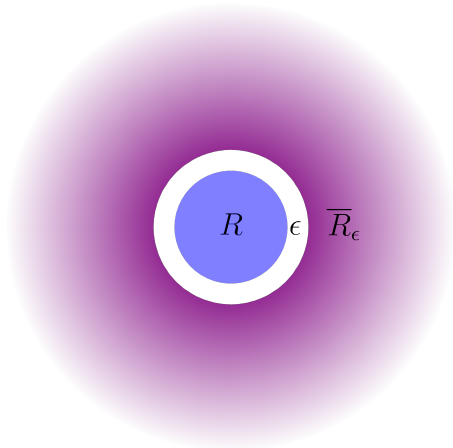


Figure 1: *The configuration of interest. We study a region R (blue), surrounded by a collar region ϵ (white) and the complement of $R \cup \epsilon$, which is denoted by \overline{R}_ϵ (purple). The region, \overline{R}_ϵ extends to infinity. In a nongravitational theory, the state on \overline{R}_ϵ and R can be specified entirely independently but in a theory of gravity the state on \overline{R}_ϵ completely fixes the state on R .*

1 The split property and its failure in gravity

1.1 Nongravitational quantum field theory

We start by briefly reviewing the split property in nongravitational quantum field theories [1–3]. (See also the review [21].) Consider a Cauchy slice in a noncompact spacetime. Let R be a bounded region that is surrounded by a small “collar region”, ϵ . The complement of $R \cup \epsilon$ is denoted by \overline{R}_ϵ as shown in Figure 1. In a nongravitational theory, there is no obstacle to studying local operators, $\phi(x)$ that probe the physics in the region R and are labelled by a point $x \in R$. As usual, it is possible to consider arbitrary polynomials of such operators to generate the algebra of the region R that we denote by \mathcal{A} .

$$\mathcal{A} = \text{span of } \{\phi(x_1), \phi(x_1)\phi(x_2), \dots, \phi(x_1)\phi(x_2) \dots \phi(x_n), \dots\}, \quad x_i \in R \quad (1.1)$$

Similarly, one may generate the algebra of observables in the region \overline{R}_ϵ that we denote by $\overline{\mathcal{A}}$. Then, in a nongravitational quantum field theory, microcausality implies that elements of these two algebras must commute pairwise

$$[A, \overline{A}] = 0, \quad \forall A \in \mathcal{A}, \overline{A} \in \overline{\mathcal{A}}. \quad (1.2)$$

In the context above, the split property states that if ρ_1 and ρ_2 are any two mixed states in the theory, which we represent as density matrices, then it is possible to find a “split state”, ρ , that appears to be like ρ_1 in R and like ρ_2 in \bar{R}_ϵ .

$$\text{Tr}(\rho A \bar{A}) = \text{Tr}(\rho_1 A) \text{Tr}(\rho_2 \bar{A}), \quad \forall A \in \mathcal{A}, \bar{A} \in \bar{\mathcal{A}}. \quad (1.3)$$

The split property tells us that, in a nongravitational theory, even if we are given the value of all observables in \bar{R}_ϵ we are completely ignorant about the state in R .

This property just formalizes the notion that, in a lattice regularization, the Hilbert space of a nongravitational theory factorizes into a part associated with R and a part associated with its complement. If the Hilbert space factorizes, we do not even need to introduce a collar region. Say the full Hilbert space can be written as a tensor product of a factor associated with R and another factor associated with its complement \bar{R} : $\mathcal{H} = \mathcal{H}_R \otimes \mathcal{H}_{\bar{R}}$. A split state can then be generated simply by taking

$$\rho = \text{Tr}_{\mathcal{H}_{\bar{R}}}(\rho_1) \otimes \text{Tr}_{\mathcal{H}_R}(\rho_2); \quad \text{split state in a factorized Hilbert space.} \quad (1.4)$$

In the continuum theory, if one does not include the collar region then a product state of the form above would have infinite energy. The collar region ensures that it is possible to preserve the very-short distance correlations across the boundary of R and \bar{R}_ϵ , allow a smooth “transition” from the state ρ_1 in R to the state ρ_2 in \bar{R}_ϵ and also ensure no correlations between R and \bar{R}_ϵ .

If the reader finds the phrase “split property” unduly formal, the reader is free to think in terms of the more familiar but slightly imprecise idea that the Hilbert space “factorizes” into a part associated with a region and another part associated with its complement. Indeed, we will use this formulation below to make contact with the existing literature.

The “split property”, as written in the form (1.3), also holds in nongravitational gauge theories. The existence of conserved charges does not modify (1.3); measuring the charge in \bar{R}_ϵ does not tell us about the charge contained in R since an arbitrary amount of compensating charge can be stuffed into the collar region.

1.2 Quantum gravity

1.2.1 Holography of information

The split property fails in quantum gravity and is replaced by the principle of holography of information that we now review. We do not provide proofs of the results quoted below but include references to the relevant literature.

In a theory of gravity, it is simplest to consider the algebra of asymptotic observables. This is because the gauge redundancy of the theory only comprises diffeomorphisms that die off asymptotically. So asymptotic observables are gauge invariant.

First consider four-dimensional asymptotically flat space. Let us denote the algebra of operators at \mathcal{I}^+ by $\mathcal{A}(\mathcal{I}^+)$. (Similar statements hold for the algebra of operators on \mathcal{I}^- and to avoid repetition, we only discuss \mathcal{I}^+ .) The algebra $\mathcal{A}(\mathcal{I}^+)$ is generated by the fields that describe the asymptotic fluctuations of the metric and of other dynamical fields. The fluctuations of the metric can themselves be divided into the Bondi news tensor $N_{AB}(u, \Omega)$ (A, B run along the directions on \mathcal{I}^+) and the Bondi mass aspect $m(u, \Omega)$ both of which are labeled by a retarded time, u , and a point on the celestial sphere Ω .¹ The algebra $\mathcal{A}(\mathcal{I}^+)$ is then constructed by taking all possible polynomials in these operators precisely as in equation (1.1).

Now, consider the subalgebra formed by operators near the past boundary of \mathcal{I}^+ , which we denote by $\mathcal{A}_{\text{bdry}}$. (See Figure 2.) This is generated just as above except that we restrict to operators whose retarded time is in the range $(-\infty, -\frac{1}{\epsilon})$ for some choice of ϵ . Naively, one might have thought that $\mathcal{A}_{\text{bdry}}$ is a small subalgebra of $\mathcal{A}(\mathcal{I}^+)$. But it can be shown any state of massless particles is *uniquely* specified by observables from this algebra.

$$\text{Tr}(\rho_1 A) = \text{Tr}(\rho_2 A) \quad \forall A \in \mathcal{A}_{\text{bdry}} \Leftrightarrow \rho_1 = \rho_2, \quad (1.5)$$

where ρ_1 and ρ_2 are any two mixed states in the massless Hilbert space. The surprising aspect of the statement above is that equality of the expectation values is *not* imposed for all elements of $\mathcal{A}(\mathcal{I}^+)$ but only for the elements of $\mathcal{A}_{\text{bdry}}$. This result is proved as “Result 1” in [5] and we note that $\mathcal{A}_{\text{bdry}}$ here is denoted by $\mathcal{A}_{-\infty, \epsilon}$ there.

We now turn to global AdS_{d+1} whose conformal boundary is of the form $S^{d-1} \times R$ where the time coordinate runs from $(-\infty, \infty)$. Let us denote the algebra of all observables localized in boundary time $t \in [0, \epsilon]$ by $\mathcal{A}_{\text{bdry}}$. Then a statement of precisely the form above can be established. If ρ_1 and ρ_2 are any two states in the theory then imposing the equality of the expectation value of all elements in $\mathcal{A}_{\text{bdry}}$ imposes the equality of the states just as in (1.5). Such a statement would not be true in a nongravitational theory, where the algebra of boundary observables would be much larger than the algebra $\mathcal{A}_{\text{bdry}}$.

The result above is proved as “Result 5” in [5]. Note that this statement is *not* proved using AdS/CFT. It can be independently derived by a consideration of the operator algebra. In perturbation theory, this statement can also be derived directly through an analysis of the constraints imposed on states by the Wheeler-DeWitt equation [7]. Within perturbation theory, it is also possible to construct a simple “physical

¹In [22], it was argued that the mass aspect is not a well defined operator at finite retarded time. For what follows, we only need the mass aspect near $u \rightarrow -\infty$ which can be defined as described in Appendix B of [5].

protocol” to extract information about the bulk from a small time band on the boundary [6].

Let us now consider the physical interpretation of (1.5). In asymptotically flat space, \mathcal{I}^+ can be thought of as a limit of Cauchy slices on which one can specify data for massless excitations. **Therefore, in asymptotically flat space, (1.5) tells us that the information available on this entire Cauchy slice is available near its boundary, which is precisely the principle of holography of information. Equation (1.5) can be interpreted similarly in AdS. The algebra of operators at the conformal boundary in the time band $[0, \epsilon]$ is similar to the algebra of operators at the time $t = \frac{\epsilon}{2}$ but in the radial range $r \in [\cot \frac{\epsilon}{2}, \infty)$. Therefore, in AdS, (1.5) again tells us that information available anywhere on a Cauchy slice is available near its boundary.**

Proofs of the principle of holography of information have only been given in these cases, although it seems very plausible that this should be a general property of gravitational theories. For related discussion, we refer the reader to [23–25]. However, it is not currently known how this principle applies when spatial slices are compact.

1.2.2 Failure of the split property

We now return to the implication for the split property. We emphasize that our discussion here is limited to bounded regions that are surrounded by their complement as shown in Figure 1. The discussion below does *not* apply when the region and its complement both extend to asymptotic infinity.

In quantum gravity, it is impossible to define local gauge-invariant observables and so the first subtlety in considering the split property is to define what one means by the algebra, \mathcal{A} , associated with a region. One way to proceed is as follows. Even in quantum gravity, it is possible to find *approximately localized* observables that probe the physics in the region R . A common way to obtain such observables is simply to fix gauge. Another, more elaborate procedure, is to “dress” such observables by means of a gravitational Wilson line [26]. However, as the latter picture makes clear, such operators are ultimately nonlocal. Nevertheless, starting with a set of approximately local operator in the region R , it is possible to define an algebra by taking all polynomials of these operators as shown in (1.1). One may similarly associate an algebra, $\overline{\mathcal{A}}$ with the region \overline{R}_ϵ . We can then ask whether these algebras obey some version of the split property.

To answer this question, we note that whatever procedure is adopted for associating an algebra with a region it is clear that, in *all cases*, the algebra $\overline{\mathcal{A}}$ must contain the algebra of all asymptotic operators that we denoted by $\mathcal{A}_{\text{bdry}}$ above. This simple observation immediately implies that the split property must fail in gravity.

This can be easily shown by contradiction. Assume that (1.3) holds in gravity. Then taking $A = 1$ and $\bar{A} \in \mathcal{A}_{\text{bdry}}$ we find that

$$\text{Tr}(\rho\bar{A}) = \text{Tr}(\rho_2\bar{A}), \quad \forall \bar{A} \in \mathcal{A}_{\text{bdry}}. \quad (1.6)$$

But then applying (1.5) we see that we must have $\rho = \rho_2$. This means that for some nontrivial A and for a general ρ_1 it is impossible to satisfy (1.3). Therefore, as claimed, it is impossible to specify the state of the theory inside and outside a bounded region independently in a theory of gravity.

One important consequence of this result is that there is no simple definition of the entanglement entropy of the complement of a bounded region in gravity. A natural question is whether we can somehow coarse grain the set of observables so as to recover an approximate split property and thereby define an entanglement entropy. We return to this question below although, in general, the answer is state dependent.

But, in many cases, the result (1.5) is important even at leading order in perturbation theory. One striking example is as follows. Consider the vacuum of global AdS. Then, as reviewed in [6], even if we keep track of the metric-degrees of freedom to leading nontrivial order in perturbation theory in $\frac{l_{\text{pl}}}{l_{\text{AdS}}}$ (Planck scale divided by the AdS scale), the entanglement entropy of the complement of a bounded region is trivial. This is because the measurement of the energy from near the boundary is enough to uniquely identify the global vacuum that is the only state with vanishing energy. The entanglement entropy of bounded regions and their complements is often studied in the vacuum of nongravitational field theories but such a quantity appears subtle to define even perturbatively in gravity in global AdS!

In other states, it is harder to extract information about a bounded region from its complement. For instance, it is nonperturbatively difficult to determine a black hole microstate by observations outside the black hole. This allows for a definition of a coarse-grained entropy for the exterior, even in the presence of gravity. Nevertheless, as we describe in section 4, the failure of the split property becomes relevant again if one attempts to define a fine-grained entropy that obeys a Page curve.

2 The error in Hawking's argument

It is commonly stated that to resolve Hawking's original paradox, it is necessary to find the precise error in Hawking's original argument [9]. We now explain how the failure of the split property in gravity allows us to isolate this error.

In [8] and [9], Hawking presented both a calculation and an argument. The calculation was as follows. Let ρ be the final state of a black hole formed from collapse

and let $a_\omega^\dagger, a_\omega$ be creation and annihilation operators that describe outgoing modes of frequency ω for a dynamical field. Then [8] showed that

$$\mathrm{Tr}(\rho a_\omega a_\omega^\dagger) = \frac{1}{1 - e^{-\beta\omega}}, \quad (2.1)$$

where β is the inverse temperature of the black hole. It was also argued that the occupancy of higher powers of $a_\omega a_\omega^\dagger$ would be thermal. (Here we have slightly smeared the operators to avoid writing a delta function in frequency as explained near equation 2.21 of [20].)

By itself, the result (2.1) does *not* amount to a computation of the entanglement entropy of Hawking radiation or an argument that the final state, ρ , must be thermal. This is because low-point correlators are insufficient to fix the entanglement entropy. For instance, if we were to modify the result above to

$$\mathrm{Tr}(\rho a_\omega a_\omega^\dagger) = \frac{1}{1 - e^{-\beta\omega}} + \mathcal{O}\left(e^{-\frac{S_{\mathrm{BH}}}{2}}\right), \quad (2.2)$$

with corrections that are exponentially suppressed in the black hole entropy, S_{BH} , this would be perfectly consistent with ρ being a pure density matrix [27, 28]. Therefore the result (2.1) is not precise enough, by itself, to conclude that ρ must be mixed.

But Hawking also provided a general argument to suggest that the purity of the initial state could not be preserved by small corrections of the form (2.2). This argument was enunciated through, what was termed, the “principle of ignorance” in [9]. In [9], Hawking argued that it was necessary to introduce a “hidden surface” about a black hole background. Hawking then stated that the principle of ignorance would imply that (see p. 2463)

“all field configurations on these hidden surfaces are equally probably provided they are compatible with the conservation of mass, angular momentum, etc. which can be measured by surface integrals at a distance from the hole.”

The argument of [9] proceeded by explicitly introducing a factorization of the final Hilbert space. If H_1 is used to denote the Hilbert space describing data on the initial surface then it was argued that transition amplitudes would require the specification of a state on the “hidden surface” Hilbert space H_2 and a specification of the state on the final surface outside the black hole H_3 . (See unnumbered equation on p. 2463 of [9]) The crux of the argument in [9] was that since data on the hidden surface would be lost to the outside observer, the outside observer would perceive a mixed state.

Figure 2 restates this argument in the context of a nice slice through an evaporating black hole spacetime. In a nongravitational quantum field theory, it is possible to

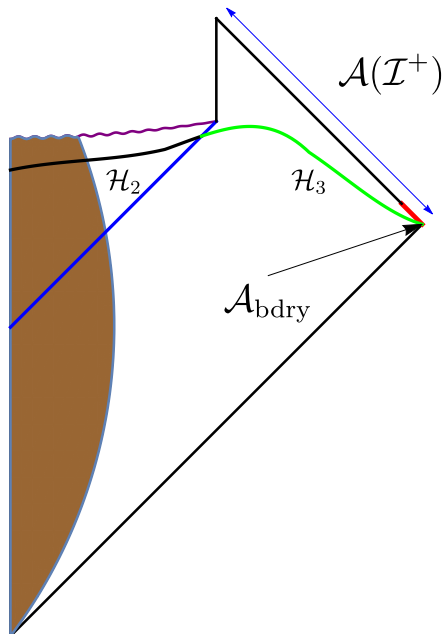


Figure 2: *The extended Penrose diagram of an evaporating black hole. A common incorrect assumption is that the Hilbert space factorizes into a part H_2 associated with the interior and a part H_3 associated with the exterior, up to constraints imposed by mass and global charges. But the Hilbert space does not factorize on a nice slice or even on \mathcal{I}^+ . Information available in the algebra of all operators on \mathcal{I}^+ (marked in blue) is available in the algebra $\mathcal{A}_{\text{bdry}}$ of operators from the red region near \mathcal{I}_-^+ .*

specify data independently on the part of the nice slice behind the horizon and the part outside the horizon. Subject to the ultraviolet subtleties described in section 1.1, the Hilbert space factorizes into a part H_2 associated with the interior and a part H_3 associated with the exterior. In classical gravity, it is possible to specify different configurations in the black hole exterior subject to only the constraints of some conserved charges. So if the split property were to hold in gravity, an observer at \mathcal{I}^+ would lose information about the state of the interior and would have to trace over all possibilities for H_2 leading to a final mixed state at \mathcal{I}^+ .

So we see that the argument of [9] relies explicitly on the split property whose failure in theories of gravity was reviewed above. As explained in section 1, if one examines the full wavefunctional of the theory then it is *not* possible to specify data outside and inside the black hole separately. This quantum mechanical effect cannot

be neglected if we ask fine grained questions about whether or not the final state is pure.

To summarize, there are two specific issues with Hawking’s argument for information loss.

1. The computation of mode-occupancy performed by Hawking is not precise enough to conclude that the final state is mixed.
2. In order to extend the validity of this calculation, Hawking invoked a “principle of ignorance” that explicitly assumed a factorization of the Hilbert space. The principle of holography of information implies that this assumption is incorrect in gravity.

3 The monogamy paradox

The original formulation of the information paradox was refined by Mathur in [10] and Mathur’s argument was later elaborated by AMPS [11]. The idea of [10,11] was to set up a potential contradiction between effective field theory, unitarity and the monogamy of entanglement that would be robust against small corrections of the form (2.2). A detailed review is provided in [20] and we provide a brief summary here to elucidate how the split property was assumed in these paradoxes.

Consider a nice slice in an evaporating black hole spacetime, as shown in Figure 2 and divide it into three parts: A, B, C where A is just outside the horizon, B is just inside the horizon and C is a region near infinity corresponding to an asymptotic observer. We will repeat the original argument of [10], which assumed that, even in gravity, it should be possible to define an entanglement entropy corresponding to these regions and we denote the entropy of the region R by S_R^{ent} below.

Effective field theory tells us that pairs of Hawking quanta in regions A and B must be entangled with each other, which led [10] to the conclusion that

$$S_{AB}^{\text{ent}} < S_B^{\text{ent}}, \quad \text{assuming a smooth horizon,} \quad (3.1)$$

i.e. the region B by itself is more “mixed” than A and B taken together.

The argument in [10] then proceeded by noting that if the entanglement entropy of the radiation follows a Page curve then for an old black hole the entropy of C must start falling with time. But since the radiation in C at a given time is the union of the radiation in A and in C at an earlier time we must have

$$S_{AC}^{\text{ent}} < S_C^{\text{ent}}, \quad \text{assuming a Page curve for } S_C^{\text{ent}} \text{ and a old black hole.} \quad (3.2)$$

These two claims are in contradiction with the strong subadditivity of entropy that reads

$$S_{AB}^{\text{ent}} + S_{AC}^{\text{ent}} \geq S_B^{\text{ent}} + S_C^{\text{ent}}. \quad (3.3)$$

In [10], this was used to argue that (3.1) must break down and the interior of the black hole must be replaced by a “fuzzball” and in [11] the same argument was used to suggest the presence of “firewall” at the horizon.

But, a simple resolution to this paradox is obtained through the following observation. The strong subadditivity inequality (3.3) holds *only* if the operator algebras corresponding to subsystems A, B, C commute. Therefore, we see that a key assumption — unstated but tacit — that goes into the monogamy paradox of [10] and [11] is that the algebra of operators can be separated into commuting factors associated with bounded geometric regions. As explained in section 1 this assumption is false in theories of gravity. This failure cannot be ignored if we wish to make reference to fine-grained quantities like the entanglement entropy. So the purported refinement of Hawking’s original paradox is based on the same erroneous assumption as the original paradox.

In the context of the discussion of the firewall, the issue of factorization was discussed in [29, 30] and it was then observed in [28] that the paradox could be resolved by recognizing the failure of the independence of operator-algebras in gravity. Similar ideas were explored in [31].

Although, as presented above, the original paradox of [10] and [11] was phrased using the entanglement entropy, it is possible to rephrase the paradox in terms of Bell correlators. These correlators avoid the ambiguity in the definition of the entanglement entropy and can be defined in a theory of gravity. This makes the paradox sharper, although the resolution remains precisely the same as the one outlined above: the operator-algebra of the theory does not contain commuting factors associated with A, B, C . This reformulation and resolution can be seen particularly cleanly in a toy version of the paradox both in AdS [32] and in flat space [33].

3.1 Other paradoxes

The original firewall paradox should be distinguished from subsequent paradoxes outlined in [34, 35]. The question addressed in the paradoxes of [34, 35] is whether “typical states in AdS/CFT have a smooth interior.” These paradoxes are reviewed and sharpened in [36]. They cannot be resolved through the simple observation that the Hilbert space does not factorize in quantum gravity. These paradoxes, and similar paradoxes that arise for the eternal black hole [37] can be addressed through a state-dependent reconstruction of the black hole interior. The issue of the consistency of state dependence

is further discussed in [38, 39]. However, it is not clear if these issues are directly relevant for evaporating black holes since evaporating black holes are never typical states in the Hilbert space as discussed in [20].

4 The Page curve

We now turn to the common idea that the entanglement entropy of the region outside the black hole should follow a Page curve. We explain why this idea also involves the same assumption about the split property.

This assumption is stated explicitly in the original argument provided by Page and Lubkin [12, 40]. The argument of [12] explicitly considered a factorized Hilbert space

$$H = H_A \otimes H_B; \quad \dim(H_A) = n; \quad \dim(H_B) = m. \quad (4.1)$$

It was then argued that for $|\Psi\rangle$ a pure state in the large Hilbert space, the entropy of the reduced density matrix of a subsystem

$$\rho_A = \text{Tr}_B |\Psi\rangle\langle\Psi| \quad (4.2)$$

would obey

$$- \langle \text{Tr}_A(\rho_A \log \rho_A) \rangle_{\text{Haar}} = \min(\log(n), \log(m)), \quad (4.3)$$

up to corrections suppressed by $\min(n, m)/\max(n, m)$ and where the expectation value is taken with respect to the Haar measure on pure states of the larger system. Since almost all states are Haar-typical this implies that for a “generic pure state” the entropy of a subsystem is given by the formula above.

The argument applies perfectly well to nongravitational systems and has been useful in that context [41]. However, this argument is also often applied to an evaporating black hole. In [13], where this result was first applied to black holes (see paragraph above Eqn. 1 in the published version), Page assumed that

“the black hole subsystem has dimension n . . . the radiation subsystem has dimension m . . . [and] these subsystems form a total system in a pure state in a Hilbert space of dimension mn .”

If one could indeed treat the black hole and radiation as parts of a larger factorized Hilbert space, the assumptions of [12, 40] would be met, and the entanglement entropy of the radiation would follow a Page curve.

However, this is yet another restatement of the flawed assumption that we have discussed above. As explained in section 1, if one thinks of the black hole as occupying

a bounded region on a Cauchy slice, and the radiation as occupying the complementary region then the state of the black hole is completely fixed once we specify all observables in the radiation region. Therefore, at least if we consider the fine-grained entropy, there is no reason to expect that the entropy of the radiation region will follow a Page curve.

It is not even true that the Page curve appears at \mathcal{I}^+ . Let $\mathcal{A}(-\infty, u)$ denote the algebra of observables at \mathcal{I}^+ of the segment that extends from its past boundary to some cut at retarded time u . Then, as is standard [42], one may define the entanglement entropy $S^{\text{ent}}(u)$ with respect to this algebra. It can then be shown (see Result 3 in [5]) that

$$\frac{\partial S^{\text{ent}}(u)}{\partial u} = 0, \quad (4.4)$$

which is the statement that the entanglement entropy is constant along \mathcal{I}^+ and does not obey a Page curve.

4.1 The difference between ordinary statistical systems and black holes

At this point the reader may wonder why the effect above can be neglected for ordinary statistical systems: although we live in a world where gravity is presumably quantized, we still expect that the entanglement entropy of ordinary systems will obey a Page curve.

This can be understood as follows. The difference between two typical microstates of an ordinary statistical system with thermodynamic entropy S is suppressed by $O\left(e^{-\frac{S}{2}}\right)$ [43]. This follows simply from the fact that the number of relevant microstates is given by e^S . Therefore to meaningfully discuss the entanglement entropy of a subsystem, we need to keep track of its observables to at least this accuracy. On the other hand, gravitational effects are suppressed by a factor of GE^{d-2} in d dimensions, where E is the typical energy scale associated with excitations in the system. For ordinary systems, we can consider a limit where $GE^{d-2} \rightarrow 0$ while S remains finite.

However, such a limit does not exist for black holes. For a black hole, if we take the energy of a typical excitation to be $E = T$, where T is the Hawking temperature then, for a Schwarzschild black hole in d dimensions we find that

$$GE^{d-2} = \frac{\alpha}{S_{\text{BH}}}, \quad \alpha = \frac{\pi^{\frac{3-d}{2}} \left(\frac{d-3}{4}\right)^{d-2}}{2\Gamma\left(\frac{d-1}{2}\right)} \quad (4.5)$$

where S_{BH} is the entropy of the black hole. So there is no simple limit, where it is possible to keep the entropy of the black hole finite while neglecting the gravitational effects described in section 1.

Even in a theory of gravity, there is no obstruction to thinking of the black hole radiation as a subsystem, in a coarse-grained sense. We can simply focus on observables

that are insensitive to effects that are suppressed by powers of S_{BH} . For instance, the classical expectation value of the metric and other geometric quantities are all of this form. However, within this approximation, it is not possible to speak meaningfully of a fine-grained entropy for the radiation. It is possible that there is some intermediate level of “coarse graining” where a Page curve appears for black holes, and we discuss this possibility in section 4.4.

4.2 Computations of the Page curve

We now turn to the computations of the Page curve in the recent literature [14–19]. We first discuss the nature of “information transfer” in these computations. We then describe why these computations involve massive gravity, and their lessons are inapplicable to standard theories of gravity.

4.2.1 Nature of information transfer

In popular descriptions [44], it is sometimes stated that the recent computations of the Page curve indicate how information “emerges” from a black hole. However, this is misleading.

Computations of the Page curve have been performed by coupling a holographic gravitational system in AdS to a nongravitational bath. Therefore the Page curve that is computed only measures how information is transferred from one section of this nongravitational system to another. These computations are perfectly consistent with the results explained above, which state that information is always available near the boundary of the gravitational spacetime.

This can be understood more clearly by considering the two-step process shown in Figure 3. In the first step (Figure 3a) we prepare a black hole in a theory of a gravity in an asymptotically AdS spacetime. According to the discussion of section 1, all information about the state is always present near the boundary of the spacetime. In the second step (Figure 3b), we couple a nongravitational bath to the conformal boundary of AdS. Information then flows into the nongravitational bath.

Since the bath is nongravitational, it does obey the “split property”. Therefore, if one draws an imaginary interface in this bath, the information transfer across this interface is controlled by a Page curve as one would expect on general grounds in any nongravitational system. This is the Page curve that has been computed.

Moreover, the computation of this Page curve is also unnecessary to establish the unitarity of the original process of black-hole formation. If the black hole is formed from a pure state, unitarity follows from the fact that the entropy of the boundary remains zero after the black hole forms. The Page curve is a check that the theory remains unitary after it is coupled to a bath. But this latter question is a question

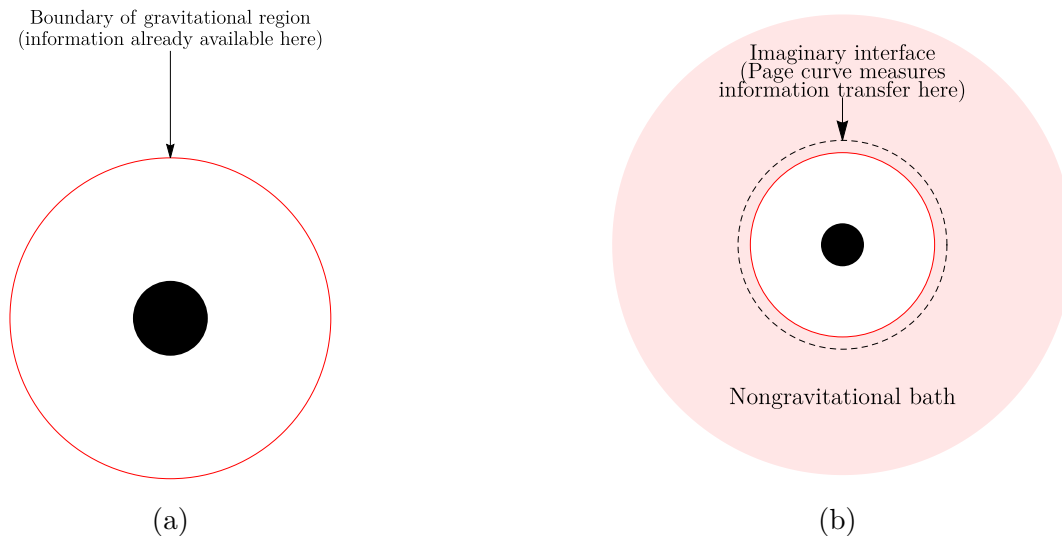


Figure 3: A two step process clarifying the nature of information transfer in computations of the Page curve. First, we prepare a black hole in AdS (left subfigure). According to the principle of holography of information, information about the microstate resides near the boundary of the spacetime. Second, we couple the system to a nongravitational bath (pink shaded region in the right subfigure). Information flows across a nongravitational interface (dashed line) and this information transfer is described by the Page curve.

about the nongravitational bath and its coupling to the original system. This question is interesting in its own right but it is not a question that is important for the original gravitational theory.

Several computations of the Page curve use “doubly holographic” settings where the system of the black hole and bath can be embedded in a higher-dimensional AdS. This is achieved by inserting a Karch-Randall [45, 46] brane in AdS, and the entropy of a part of the bath can then be studied by the standard Ryu-Takayanagi prescription [47–49]. (See [17, 50, 51] for an elaboration of this perspective.) The authors of [52] generalized this setup by considering an AdS space with two Karch-Randall branes inserted in a black-string geometry. This serves as a model of a system where the bath itself is gravitating.

It was then found in [52] that there was no nontrivial Ryu-Takayanagi surface (excluding the black-string horizon itself) that extended from one brane to the other with symmetric Neumann boundary conditions imposed on both branes. Therefore, the Page curve disappears when the bath itself is gravitating. In [52], it was shown that this conclusion is true even if gravity on the “bath brane” is much weaker than

gravity on the “system brane”. This lends evidence to the idea that systems with “weak gravity” should not be conflated with systems with “no gravity.”

4.3 Massive gravity and islands.

As explained above, extant Page curve computations address a nongravitational problem. What is interesting, nevertheless, is that this nongravitational problem can be studied using gravitational techniques, including an elegant “island rule” [19], that give answers consistent with the Page curve.

While the emergence of the Page curve in this nongravitational context is natural, a closer examination of the gravitational computation reveals that the “island rule” appears to be in tension with the principle of holography of information. The island rule suggests that after the bath has been coupled to the holographic system, the state of an “island” inside the gravitational system, is captured by part of the bath and the state of its complement in the gravitational region is captured by the complementary part of the bath. Since the nongravitational bath obeys a split property, this implies that it is also possible to set up a split state in the gravitational bulk. This seems to be in contradiction with the result of section 1.

The authors of [53] turned this into a sharp puzzle for theories of gravity in higher than $1 + 1$ dimensions. This puzzle does not require the full power of the principle of holography of information. Rather, it can be shown, using only perturbation theory, that islands are inconsistent in a theory of gravity that obeys the Gauss law.

The puzzle established in [53] exploits the observation that correlators of simple operators in a typical black hole microstate, $|\Psi_{\text{typ}}\rangle$ are approximately time independent. Now consider a black hole that contains an island, and consider creating an excitation in the island using an operator of the form $U = e^{i\lambda\phi(P)}$, where $\phi(P)$ is a weakly coupled propagating field smeared near a point P in the island. A simple perturbative computation reveals that

$$\frac{\partial}{\partial\lambda}\langle\Psi_{\text{typ}}|U^\dagger H \frac{\partial\phi(P')}{\partial t} U|\Psi_{\text{typ}}\rangle \neq 0, \quad (4.6)$$

where H is the integral of the metric near the AdS boundary that measures the total energy and $\phi(P')$ is an insertion of the same scalar field smeared near a point P' at the asymptotic boundary and t is the boundary time. The right hand side can be computed explicitly in perturbation theory. (See section 3 in [53].) But this contradicts the idea that simple unitary operators in an entanglement wedge should leave correlators outside the wedge invariant.

Note that in equation (4.6) one is *not* attempting to identify the original black hole microstate, which would not be possible within perturbation theory. Rather, equation

(4.6) exploits the fact that an excitation on top of a typical microstate leads to an atypical state. It is the difference between this atypical state and the original typical state that can be detected in perturbation theory by simple correlators outside the horizon [54].

The physical observation underlying the puzzle is simply that islands are unique among entanglement wedges because islands are surrounded by their complements. This brings islands within the purview of section 1; in contrast, standard entanglement wedges in AdS/CFT always contain a piece of the asymptotic region and so the discussion of section 1 does not apply to them. In standard entanglement wedges, it makes sense to consider a little Hilbert space (as introduced in [55] or, equivalently, a “code subspace” [56]) and one can study excitations within the entanglement wedge that leave all correlators in the complement of the wedge invariant. This is because the excitation can be “dressed” to infinity through the asymptotic part of the wedge. Since islands do not contain such a region, in a standard theory of gravity, there are no excitations that leave the complement of the island invariant. So, in standard gravity, islands cannot constitute entanglement wedges, at least in the usual sense of the term. The computation (4.6) shows this explicitly.

This does not mean that computations based on the island rule are wrong. Rather the resolution to the puzzle is that (in higher than 1 + 1 dimensions) the coupling the gravitational system to a bath causes the bulk graviton to pick up a mass. Therefore almost all computations based on islands have been carried out in massive gravity [57]. Massive gravity is quite different from standard gravity; it does not have a Gauss law, and so the principle of holography of information does not apply and islands are consistent in such theories. Before we explain how this resolves the puzzle we first explain why the nongravitational bath induces a mass for the graviton.

The easiest way to understand the graviton mass [58] is to again consider the sequence shown in Figure 3, where a black hole is first prepared in a holographic system and later coupled to a bath. In Figure 3a, before the system is coupled to a bath, the holographic stress tensor on the boundary of AdS is conserved:

$$\partial_\mu T^{\mu\nu} = 0, \quad (\text{in Figure 3a before coupling to bath}). \quad (4.7)$$

where μ, ν run along the boundary coordinates. After the coupling is turned on, we now have

$$\partial_\mu T^{\mu\nu} \neq 0. \quad (\text{in Figure 3b after coupling to bath}). \quad (4.8)$$

This must be true since, by construction, the coupling allows energy to leak into the bath.

In conformal representation theory [59–62], the conservation of the stress tensor is simply the statement that the stress tensor lives in a “short representation” of the

conformal algebra. This allows its conformal dimension to be exactly d where d is the boundary dimension. The non-conservation of the stress tensor implies that, in the theory with the bath, the stress tensor is in a “long representation” but this means that its conformal dimension must exceed d , and so it must pick up an anomalous dimension. By the standard rules of AdS/CFT [63, 64] this means that the bulk graviton must have a mass. This can be verified through a direct bulk computation [65, 66].

The constraints of massive gravity are different from those of standard massless gravity. Most importantly, the Hamiltonian of the theory of massive gravity is *not* a boundary term. Therefore, the principle of holography of information does not apply and even the computation leading up to (4.6) is modified.

The simplest setting to see this difference is to compare linearized standard gravity with linearized massive gravity. Writing the metric on a spatial slice as $\delta_{ij} + h_{ij}$, we find that the metric fluctuation is constrained in terms of the local energy density, ρ as

$$\frac{1}{16\pi G} (\partial_j \partial_i h_{ij} - \partial_j \partial_j h_{ii}) = \rho, \quad (4.9)$$

in standard gravity. Upon integration this leads to the familiar ADM expression [67] for the total energy as a surface integral.

$$H = \frac{1}{16\pi G} \int_{S_\infty} n_j (\partial_i h_{ij} - \partial_j h_{ii}), \quad (4.10)$$

where the integral is performed over the sphere at asymptotically large r with normal n_j . Of course, (4.10) is the *definition* of the energy even when one goes beyond the linearized approximation to the full theory of general relativity (see section 20.4 in [68]) and also in the quantum theory [69]. On the other hand, in linearized massive gravity we find that the analogous constraint is [70]

$$\frac{1}{16\pi G} (\partial_j \partial_i h_{ij} - \partial_j \partial_j h_{ii} + m^2 h_{ii}) = \rho. \quad (4.11)$$

Since the left hand side is not a total divergence, the integral of the energy density is not captured by an integral at infinity even in the linearized theory.

In [53], it is shown that, at least, in some well-understood examples of massive theories of gravity, such as those that appear in the Karch-Randall model, there does not appear to be any any immediate obstacle to preparing split states.²

To summarize: existing derivations of the Page curve pertain to information transfer across a nongravitational interface and do not indicate the transfer of information

²Here, we should caution the reader that the question of how information is localized in theories of massive gravity has not been studied in detail.

within the gravitational region. Moreover, the models that have been used to compute the Page curve involve rather nonstandard theories of gravity. These models differ qualitatively from standard gravity, and it does not appear that the island proposal or other results from such models can be generalized to black holes in standard gravity.

4.4 Page curves in standard gravity

In the section above, we criticized the expectation that the radiation from a black hole in standard gravity should follow a Page curve. However, since this question has historically been of interest, one may ask whether it is possible to find an appropriate quantity, in a standard theory of gravity, that does follow a Page curve. Here, we outline some possibilities in this direction.

Page curves in flat space. First, consider four-dimensional flat space. Then the algebra of operators at \mathcal{I}^+ that we denoted in section 1 as $\mathcal{A}(\mathcal{I}^+)$ can be separated into an algebra of “dynamical operators” (generated by the so-called “Bondi news” operators and tails of other massless degrees of freedom at \mathcal{I}^+) and a set of “constrained operators”, which include the ADM Hamiltonian and other soft charges. These constrained modes are what cause the split property on \mathcal{I}^+ to fail. On the other hand, since the operator algebra of dynamical modes at \mathcal{I}^+ is otherwise free, arbitrary products of dynamical operators do not generate constrained operators. The algebra of dynamical operators can therefore be separated into an algebra corresponding to a segment of \mathcal{I}^+ and a commuting algebra corresponding to its complement.

This is still not sufficient to ensure that the entropy with respect to the dynamical algebra follows a Page curve; it is necessary to further assume that, in a generic state, the entanglement between the hard modes and the different soft vacua [71, 72] can be neglected. The validity of this assumption has not been investigated in detail and the proposal of [73] was that this entanglement is significant. But if this assumption holds then it is expected that the entropy of a state with respect to the coarse-grained algebra above will follow a Page curve.

Although the coarse-graining above is mathematically consistent, it is somewhat artificial from a physical point of view. The observer must be sensitive to very small fluctuations of the metric, including those with a magnitude $O(e^{-S})$ but is not allowed to include the mass of the black hole and its correlators with other degrees of freedom in the algebra of observables.

There are other possibilities for obtaining a Page curve at \mathcal{I}^+ all of which boil down to the coarse-graining above. These include considering the algebra of operators in the retarded-time interval (u, ∞) — which extends from a cut at finite retarded time till the future boundary of future null infinity — or equivalently considering the algebra of

operators in the infinitesimal retarded-time interval $(u, u + \epsilon)$. We refer the reader to section 5.4.5 of [20] for additional discussion.

Page curves for small AdS black holes. It is also interesting to consider small black holes in AdS [74]. If the radius of the black hole is much smaller than the AdS scale, it will evaporate entirely with no need to couple the system to an external bath. Of course, in this case we know from AdS/CFT — and entirely consistent with the discussion in section 1 — that information about the black-hole microstate is present at the boundary of AdS at all times. Therefore if the black hole is formed from the collapse of a pure state, the fine-grained entropy of the boundary remains zero at all times.

On the other hand, if one considers a small set of coarse-grained observables — such as generalized free-fields and products of a small number of such operators — then one expects to see an entropy that simply rises according to Hawking’s original prediction.

In order to see a Page curve, one needs to find an intermediate level of accuracy — an algebra that keeps enough information that it can detect that the state is pure after the black hole has evaporated but not so much information that it knows that the state is pure at all times.

Although, currently, there is no known subalgebra of the full boundary algebra that satisfies these properties, it appears plausible that such an “intermediate-level” algebra can be found. However, it is unclear whether such an algebra will be natural from a physical viewpoint or of intrinsic interest separate from reproducing the Page curve.

5 Conclusion

In this essay, we have pointed out how a familiar physical assumption is, in fact, false in quantum gravity. In nongravitational theories, the algebra of observables can be separated into a subalgebra associated with a bounded region and another commuting subalgebra associated with its complement. This property fails in a theory of quantum gravity.

For the purposes of defining a fine-grained entanglement entropy, this issue can be neglected for those systems where it is possible to take a limit in which the degeneracy of microstates remains finite while gravitational effects become arbitrarily weak. However, the limit above does not exist for black holes, and so this gravitational effect is important if we study the fine-grained entanglement entropy of the radiation. In a

fine-grained sense, this entropy is trivial and the region outside a black hole always has information about the interior.

This physical effect was not accounted for in Hawking’s original formulation of the information paradox, which assumed that the Hilbert space in gravity should factorize the way it does in nongravitational theories. Our current understanding allows us to pinpoint the precise error in Hawking’s argument. Ironically, the same assumption of factorization also underlies the expectation that the entropy of the radiation should obey a Page curve.

The point of this essay is not that it is impossible to find a Page curve for an evaporating black hole. Indeed, if we are willing to suitably change the question that is being addressed, it is possible to find an appropriate question to which the answer is the Page curve. In several extant computations in the literature, the Page curve has been derived by considering a nongravitational bath and then computing the entanglement between two regions of this bath. The gravitational description of this system involves a theory of massive gravity that appears to localize information like ordinary quantum field theories. Even in standard theories of gravity, it appears plausible that a prescription for coarse-graining the entropy or somehow neglecting gravitational effects will cause a Page curve to emerge although such a computation has not been carried out.

However, the reason for studying the black-hole information paradox is that it teaches us about new physical effects in gravity. As such, one of the lessons that the paradox teaches us is that gravity localizes information unusually. This is a striking effect that persists in the low-energy theory.

A computation of the Page curve is not necessary to resolve the information paradox and, historically, this idea was based on a flawed physical analysis. Although it is possible to modify the system of an evaporating black hole in a standard theory of gravity so as to force a Page curve upon it, this tends to obscure the interesting physics that we learn from the paradox.

Acknowledgments

I am grateful to Subhro Bhattacharjee, Tuneer Chakraborty, Joydeep Chakravarty, Chandramouli Chowdhury, Hao Geng, Victor Godet, Andreas Karch, Alok Laddha, Carlos Perez-Pardavila, Kyriakos Papadodimas, Olga Papadoulaki, Priyadarshi Paul, Siddharth Prabhu, Lisa Randall, Marcos Riojas, Sanjit Shashi and Pushkal Shrivastava for helpful discussions. This work was partially supported by a Swarnajayanti fellowship, DST/SJF/PSA-02/2016-17, of the Department of Science and Technology.

Research at ICTS-TIFR is supported by the government of India through the Department of Atomic Energy grant RTI4001.

References

- [1] H. Roos, “Independence of local algebras in quantum field theory,” *Communications in Mathematical Physics* **16** no. 3, (1970) 238–246.
- [2] D. Buchholz, “Product states for local algebras,” *Communications in Mathematical Physics* **36** no. 4, (1974) 287–304.
- [3] R. Haag, *Local quantum physics: Fields, particles, algebras*. Springer, 1992.
- [4] J. Corvino and R. M. Schoen, “On the asymptotics for the vacuum Einstein constraint equations,” *J. Diff. Geom.* **73** no. 2, (2006) 185–217, [arXiv:gr-qc/0301071](#).
- [5] A. Laddha, S. G. Prabhu, S. Raju, and P. Shrivastava, “The Holographic Nature of Null Infinity,” *SciPost Phys.* **10** (2021) 041, [arXiv:2002.02448](#) [[hep-th](#)].
- [6] C. Chowdhury, O. Papadoulaki, and S. Raju, “A physical protocol for observers near the boundary to obtain bulk information in quantum gravity,” *SciPost Phys.* **10** no. 5, (2021) 106, [arXiv:2008.01740](#) [[hep-th](#)].
- [7] C. Chowdhury, V. Godet, O. Papadoulaki, and S. Raju, “Holography from the Wheeler-DeWitt equation,” [arXiv:2107.14802](#) [[hep-th](#)].
- [8] S. Hawking, “Particle Creation by Black Holes,” *Commun. Math. Phys.* **43** (1975) 199–220. [Erratum: *Commun.Math.Phys.* 46, 206 (1976)].
- [9] S. Hawking, “Breakdown of Predictability in Gravitational Collapse,” *Phys. Rev. D* **14** (1976) 2460–2473.
- [10] S. D. Mathur, “The Information paradox: A Pedagogical introduction,” *Class. Quant. Grav.* **26** (2009) 224001, [arXiv:0909.1038](#) [[hep-th](#)].
- [11] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, “Black Holes: Complementarity or Firewalls?,” *JHEP* **02** (2013) 062, [arXiv:1207.3123](#) [[hep-th](#)].
- [12] D. N. Page, “Average entropy of a subsystem,” *Phys. Rev. Lett.* **71** (1993) 1291–1294, [arXiv:gr-qc/9305007](#).
- [13] D. N. Page, “Information in black hole radiation,” *Phys. Rev. Lett.* **71** (1993) 3743–3746, [arXiv:hep-th/9306083](#).

- [14] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, “The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole,” *JHEP* **12** (2019) 063, [arXiv:1905.08762 \[hep-th\]](#).
- [15] G. Penington, “Entanglement Wedge Reconstruction and the Information Paradox,” *JHEP* **09** (2020) 002, [arXiv:1905.08255 \[hep-th\]](#).
- [16] A. Almheiri, R. Mahajan, J. Maldacena, and Y. Zhao, “The Page curve of Hawking radiation from semiclassical geometry,” *JHEP* **03** (2020) 149, [arXiv:1908.10996 \[hep-th\]](#).
- [17] A. Almheiri, R. Mahajan, and J. E. Santos, “Entanglement islands in higher dimensions,” *SciPost Phys.* **9** no. 1, (2020) 001, [arXiv:1911.09666 \[hep-th\]](#).
- [18] A. Almheiri, R. Mahajan, and J. Maldacena, “Islands outside the horizon,” [arXiv:1910.11077 \[hep-th\]](#).
- [19] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian, and A. Tajdini, “The entropy of Hawking radiation,” *Rev. Mod. Phys.* **93** no. 3, (2021) 035002, [arXiv:2006.06872 \[hep-th\]](#).
- [20] S. Raju, “Lessons from the Information Paradox,” [arXiv:2012.05770 \[hep-th\]](#).
- [21] C. J. Fewster, “The split property for quantum field theories in flat and curved spacetimes,” [arXiv:1601.06936 \[math-ph\]](#).
- [22] R. Bousso, V. Chandrasekaran, I. F. Halpern, and A. Wall, “Asymptotic Charges Cannot Be Measured in Finite Time,” *Phys. Rev. D* **97** no. 4, (2018) 046014, [arXiv:1709.08632 \[hep-th\]](#).
- [23] D. Marolf, “Asymptotic flatness, little string theory, and holography,” *JHEP* **03** (2007) 122, [arXiv:hep-th/0612012](#).
- [24] D. Marolf, “Unitarity and Holography in Gravitational Physics,” *Phys. Rev. D* **79** (2009) 044010, [arXiv:0808.2842 \[gr-qc\]](#).
- [25] T. Jacobson and P. Nguyen, “Diffeomorphism invariance and the black hole information paradox,” *Phys. Rev. D* **100** no. 4, (2019) 046002, [arXiv:1904.04434 \[gr-qc\]](#).
- [26] W. Donnelly and S. B. Giddings, “Diffeomorphism-invariant observables and their nonlocal algebra,” *Phys. Rev. D* **93** no. 2, (2016) 024030, [arXiv:1507.07921 \[hep-th\]](#). [Erratum: *Phys. Rev. D* **94**, no. 2, 029903 (2016)].

- [27] S. Lloyd and J. Preskill, “Unitarity of black hole evaporation in final-state projection models,” *JHEP* **08** (2014) 126, [arXiv:1308.4209 \[hep-th\]](#).
- [28] K. Papadodimas and S. Raju, “An Infalling Observer in AdS/CFT,” *JHEP* **10** (2013) 212, [arXiv:1211.6767 \[hep-th\]](#).
- [29] R. Bousso, “Complementarity Is Not Enough,” *Phys. Rev. D* **87** no. 12, (2013) 124023, [arXiv:1207.5192 \[hep-th\]](#).
- [30] L. Susskind, “The Transfer of Entanglement: The Case for Firewalls,” [arXiv:1210.2098 \[hep-th\]](#).
- [31] T. Jacobson, “Boundary unitarity and the black hole information paradox,” *Int. J. Mod. Phys. D* **22** (2013) 1342002, [arXiv:1212.6944 \[hep-th\]](#).
- [32] S. Raju, “A Toy Model of the Information Paradox in Empty Space,” *SciPost Phys.* **6** no. 6, (2019) 073, [arXiv:1809.10154 \[hep-th\]](#).
- [33] T. Chakraborty, J. Chakravarty, and P. Paul, “Monogamy paradox: A toy model in flat space,” [arXiv:2107.06919 \[hep-th\]](#).
- [34] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, “An Apologia for Firewalls,” *JHEP* **09** (2013) 018, [arXiv:1304.6483 \[hep-th\]](#).
- [35] D. Marolf and J. Polchinski, “Gauge/Gravity Duality and the Black Hole Interior,” *Phys. Rev. Lett.* **111** (2013) 171301, [arXiv:1307.4706 \[hep-th\]](#).
- [36] K. Papadodimas and S. Raju, “Remarks on the necessity and implications of state-dependence in the black hole interior,” *Phys. Rev. D* **93** no. 8, (2016) 084049, [arXiv:1503.08825 \[hep-th\]](#).
- [37] K. Papadodimas and S. Raju, “Local Operators in the Eternal Black Hole,” *Phys. Rev. Lett.* **115** no. 21, (2015) 211601, [arXiv:1502.06692 \[hep-th\]](#).
- [38] D. Marolf and J. Polchinski, “Violations of the Born rule in cool state-dependent horizons,” *JHEP* **01** (2016) 008, [arXiv:1506.01337 \[hep-th\]](#).
- [39] S. Raju, “Smooth Causal Patches for AdS Black Holes,” *Phys. Rev. D* **95** no. 12, (2017) 126002, [arXiv:1604.03095 \[hep-th\]](#).
- [40] E. Lubkin, “Entropy of an n-system from its correlation with a k-reservoir,” *Journal of Mathematical Physics* **19** no. 5, (1978) 1028–1031.
- [41] C. Nadal, S. N. Majumdar, and M. Vergassola, “Phase transitions in the distribution of bipartite entanglement of a random pure state,” *Physical review letters* **104** no. 11, (2010) 110501.

- [42] H. Casini, M. Huerta, and J. A. Rosabal, “Remarks on entanglement entropy for gauge fields,” *Phys. Rev. D* **89** no. 8, (2014) 085012, [arXiv:1312.1183 \[hep-th\]](#).
- [43] S. Lloyd, *Black Holes, Demons and the Loss of Coherence: How complex systems get information, and what they do with it*. PhD thesis, Rockefeller University, 1988.
- [44] G. Musser, “The Most Famous Paradox in Physics Nears Its End,” *Quanta Magazine* (2020) . <https://www.quantamagazine.org/the-most-famous-paradox-in-physics-nears-its-end-20201029/>.
- [45] A. Karch and L. Randall, “Open and closed string interpretation of SUSY CFT’s on branes with boundaries,” *JHEP* **06** (2001) 063, [arXiv:hep-th/0105132](#).
- [46] A. Karch and L. Randall, “Locally localized gravity,” *JHEP* **05** (2001) 008, [arXiv:hep-th/0011156](#).
- [47] S. Ryu and T. Takayanagi, “Aspects of Holographic Entanglement Entropy,” *JHEP* **08** (2006) 045, [arXiv:hep-th/0605073](#).
- [48] S. Ryu and T. Takayanagi, “Holographic derivation of entanglement entropy from AdS/CFT,” *Phys. Rev. Lett.* **96** (2006) 181602, [arXiv:hep-th/0603001](#).
- [49] V. E. Hubeny, M. Rangamani, and T. Takayanagi, “A Covariant holographic entanglement entropy proposal,” *JHEP* **07** (2007) 062, [arXiv:0705.0016 \[hep-th\]](#).
- [50] H. Z. Chen, R. C. Myers, D. Neuenfeld, I. A. Reyes, and J. Sandor, “Quantum Extremal Islands Made Easy, Part I: Entanglement on the Brane,” *JHEP* **10** (2020) 166, [arXiv:2006.04851 \[hep-th\]](#).
- [51] C. F. Uhlemann, “Islands and Page curves in 4d from Type IIB,” [arXiv:2105.00008 \[hep-th\]](#).
- [52] H. Geng, A. Karch, C. Perez-Pardavila, S. Raju, L. Randall, M. Riojas, and S. Shashi, “Information Transfer with a Gravitating Bath,” [arXiv:2012.04671 \[hep-th\]](#).
- [53] H. Geng, A. Karch, C. Perez-Pardavila, S. Raju, L. Randall, M. Riojas, and S. Shashi, “Inconsistency of Islands in Theories with Long-Range Gravity,” [arXiv:2107.03390 \[hep-th\]](#).
- [54] K. Papadodimas, “A class of non-equilibrium states and the black hole interior,” [arXiv:1708.06328 \[hep-th\]](#).
- [55] K. Papadodimas and S. Raju, “State-Dependent Bulk-Boundary Maps and Black Hole Complementarity,” *Phys. Rev. D* **89** no. 8, (2014) 086010, [arXiv:1310.6335 \[hep-th\]](#).

- [56] A. Almheiri, X. Dong, and D. Harlow, “Bulk Locality and Quantum Error Correction in AdS/CFT,” *JHEP* **04** (2015) 163, [arXiv:1411.7041 \[hep-th\]](#).
- [57] H. Geng and A. Karch, “Massive islands,” *JHEP* **09** (2020) 121, [arXiv:2006.02438 \[hep-th\]](#).
- [58] O. Aharony, A. B. Clark, and A. Karch, “The CFT/AdS correspondence, massive gravitons and a connectivity index conjecture,” *Phys. Rev. D* **74** (2006) 086006, [arXiv:hep-th/0608089](#).
- [59] G. Mack, “All Unitary Ray Representations of the Conformal Group SU(2,2) with Positive Energy,” *Commun. Math. Phys.* **55** (1977) 1.
- [60] S. Minwalla, “Restrictions imposed by superconformal invariance on quantum field theories,” *Adv. Theor. Math. Phys.* **2** (1998) 781–846, [arXiv:hep-th/9712074](#).
- [61] V. K. Dobrev, “Characters of the positive energy UIRs of $D = 4$ conformal supersymmetry,” *Phys. Part. Nucl.* **38** (2007) 564–609, [arXiv:hep-th/0406154](#).
- [62] V. K. Dobrev, “Positive energy unitary irreducible representations of $D = 6$ conformal supersymmetry,” *J. Phys.* **A35** (2002) 7079–7100, [hep-th/0201076](#).
- [63] E. Witten, “Anti-de Sitter space and holography,” *Adv. Theor. Math. Phys.* **2** (1998) 253–291, [arXiv:hep-th/9802150](#).
- [64] S. Gubser, I. R. Klebanov, and A. M. Polyakov, “Gauge theory correlators from noncritical string theory,” *Phys. Lett. B* **428** (1998) 105–114, [arXiv:hep-th/9802109](#).
- [65] M. Porrati, “Mass and gauge invariance 4. Holography for the Karch-Randall model,” *Phys. Rev. D* **65** (2002) 044015, [arXiv:hep-th/0109017](#).
- [66] M. Porrati, “Higgs phenomenon for the graviton in ADS space,” *Mod. Phys. Lett. A* **18** (2003) 1793–1802, [arXiv:hep-th/0306253](#).
- [67] R. L. Arnowitt, S. Deser, and C. W. Misner, “The Dynamics of general relativity,” *Gen. Rel. Grav.* **40** (2008) 1997–2027, [arXiv:gr-qc/0405109](#).
- [68] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*. Macmillan, 1973.
- [69] T. Regge and C. Teitelboim, “Role of Surface Integrals in the Hamiltonian Formulation of General Relativity,” *Annals Phys.* **88** (1974) 286.
- [70] K. Hinterbichler, “Theoretical Aspects of Massive Gravity,” *Rev. Mod. Phys.* **84** (2012) 671–710, [arXiv:1105.3735 \[hep-th\]](#).

- [71] A. Strominger, “On BMS Invariance of Gravitational Scattering,” *JHEP* **07** (2014) 152, [arXiv:1312.2229 \[hep-th\]](#).
- [72] A. Ashtekar, M. Campiglia, and A. Laddha, “Null infinity, the BMS group and infrared issues,” *Gen. Rel. Grav.* **50** no. 11, (2018) 140–163, [arXiv:1808.07093 \[gr-qc\]](#).
- [73] S. W. Hawking, M. J. Perry, and A. Strominger, “Soft Hair on Black Holes,” *Phys. Rev. Lett.* **116** no. 23, (2016) 231301, [arXiv:1601.00921 \[hep-th\]](#).
- [74] D. A. Lowe and L. Thorlacius, “AdS / CFT and the information paradox,” *Phys. Rev. D* **60** (1999) 104012, [arXiv:hep-th/9903237](#).