UC Berkeley
UC Berkeley Previously Published Works

Title
Complementarity for a dynamical black hole

Permalink
https://escholarship.org/uc/item/2jf433hs

Journal
Physical Review D, 110(8)

ISSN
2470-0010

Authors
Concepcion, Benjamin
Nomura, Yasunori
Ritchie, Kyle
et al.

Publication Date
2024-10-15

DOI
10.1103/physrevd.110.086002

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

arXiv:2405.15849v2 [hep-th] 24 Sep 2024

Complementarity for a Dynamical Black Hole

Benjamin Concepcion,[a,b] **Yasunori Nomura,**[a,b,c,d] **Kyle Ritchie,**[a,b] **and Samuel Weiss**[a,b]

[a] *Berkeley Center for Theoretical Physics, Department of Physics, University of California, Berkeley, CA 94720, USA*

[b] *Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA*

[c] *RIKEN Interdisciplinary Theoretical and Mathematical Sciences Program (iTHEMS), Wako, Saitama 351-0198, Japan*

[d] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan*

Abstract

Black hole complementarity posits that the interior of a black hole is not independent from its Hawking radiation. This leads to an apparent violation of causality: the interior can be acausally affected by operators acting solely on the radiation. We argue that this perspective is misleading and that the black hole interior must be viewed as existing in the causal past of the Hawking radiation, despite the fact that they are spacelike separated in the semiclassical description. Consequently, no operation on the Hawking radiation—no matter how complex—can affect the experience of an infalling observer. The black hole interior and the radiation only appear spacelike separated in the semiclassical description because an infalling observer's ability to access complex information is limited; the chaotic dynamics on the horizon, as viewed from the exterior, then converts any effect caused by such an observer to information in the Hawking radiation which cannot be accessed at the semiclassical level. We arrive at the picture described above by considering a unitary exterior description in which the flow of information is strictly causal, which we extend to apply throughout the entire history of black hole evolution, including its formation. This description uses the stretched event horizon as an inner edge of spacetime, on which the information inside is holographically encoded. We argue that the global spacetime picture arises from coarse-graining over black hole microstates, and discuss its relationship with the exterior description.

Contents

1 Introduction **2**

2 Exterior Description **3**

3 Emergence of Global Spacetime **5**

3.1 Effective theory of the interior 6

3.2 Global spacetime from a patchwork 9

4 A Puzzle in the Global Spacetime Picture and Its Resolution **12**

4.1 A problem with the microscopic description of global spacetime 12

4.2 A solution: the interior as an effective description of dissipation on the horizon 14

4.3 Converting relations from timelike to spacelike using complexity 18

5 Causal Structure of Black Hole Formation and Evaporation **20**

5.1 Exterior description with trivial causal structure 21

5.2 Significance of the event horizon 26

5.3 Topology change and quantum global hyperbolicity 27

6 Global Spacetime as a Coarse-Grained Description **28**

6.1 The Page curve from replica wormholes and entanglement islands 29

6.2 Baby universe $\boldsymbol{\alpha}$-states as fine-grained information in Hawking radiation 29

7 Conclusion **31**

1 Introduction

Studying the quantum mechanics of a black hole has revealed surprising facts about the nature of spacetime [1–3]. Of particular interest is that the black hole interior and Hawking radiation are not independent quantum degrees of freedom. This idea, called complementarity [4], has been tested in various simplified setups [5–9], using methods [10–13] developed by analyzing the holographic nature of quantum gravity [14, 15], in particular the AdS/CFT correspondence [16].

This lack of independence allows us to construct operators representing the interior which act solely on the Hawking radiation, as is indeed implied by entanglement wedge reconstruction [17–24] with entanglement islands [5–7]. Since the radiation and the interior are spacelike separated, an exterior observer's ability to affect interior degrees of freedom leads to an apparent violation of causality. This issue was analyzed by Kim, Tang, and Preskill [25], who studied a partially evaporating black hole and proposed that the spacelike nature of the interior and radiation in the semiclassical description is protected because the operators representing the interior effectively commute with any operators that do not have computational complexity exponential in the entropy of the black hole. They, however, did not address how to interpret operators acting on the radiation after the black hole is fully evaporated, which would still be able to affect the interior acausally. Bousso and Penington studied a process involving such operators and concluded that it leads to a breakdown of the semiclassical spacetime [26].

In this paper, we study how complementarity is manifested without the issue of causality described above. We claim that to fully understand this issue, we need to properly analyze how a dynamical black hole behaves under time evolution. For this purpose, we adopt an exterior description [27–32] in which an effective theory describing a portion of the black hole interior is constructed from the horizon (and Hawking radiation) degrees of freedom *at each exterior time*. This portion of the interior is obtained by evolving the modes on the corresponding exterior time slice by the *infalling* Hamiltonian, not by exterior time evolution, and therefore must be viewed as associated with the exterior time at which the effective theory is erected. A sketch of this construction is given in Fig. 3. In this description, the black hole interior is a transient phenomenon obtained by reshuffling the relevant degrees of freedom while an infalling object is being scrambled into the black hole.

Using this construction, we show the following:

- The black hole interior must be viewed as existing in the causal past of the corresponding Hawking radiation despite the fact that they are spacelike separated in the semiclassical picture. In the fundamental description, the relationship between operators corresponding to the interior and their reconstruction in the Hawking radiation is timelike.

- The relationship between these operators becomes effectively spacelike in the semiclassical description because of limitations on an infalling observers' ability to access complex information. With this limitation, the chaotic dynamics on the horizon, as viewed from the exterior, convert any effect caused by an interior observer to information in the Hawking radiation which cannot be accessed at the semiclassical level.

- An operation on the Hawking radiation cannot affect the experience of an observer in the interior no matter how complex it is, providing a solution to the causality issue different from that in [25]. Executing such an operation can only hope to "rewrite the record" to make it appear that the past was different. We show that the same applies also to black hole mining [33, 34], so that the semiclassical spacetime is robust under information extraction.

Note that this provides a simple solution to the problem of possible information cloning in the interior and the radiation. The no-cloning theorem [35] is not violated simply because an infalling object and its information in Hawking radiation are the same timelike-related degrees of freedom.

We will formulate a unitary exterior description in which the flow of information is strictly causal throughout the black hole's formation and evaporation. This is not trivial, especially during the black hole formation era. We will describe the challenge and propose a solution which applies bulk holography to the stretched *event* horizon. In this description, the stretched event horizon provides the inner "edge" of the spacetime, on which the information about the interior is holographically encoded.

The picture presented here implies that a description based on the global spacetime is necessarily coarse-grained. We will discuss the relation between the exterior description and calculations employing the global spacetime description, such as replica wormhole calculation [8, 9], entanglement islands [5–7], and α-states arising from baby universes [36–38] created by the evaporation of a black hole [39, 40].

The organization of this paper is as follows. After laying out the exterior description in Section 2, we discuss in Section 3 how the interior of the black hole and the global spacetime of general relativity emerge from the exterior description. In Section 4, we argue that the global spacetime description cannot be used to define unitary quantum mechanics consistent with interior reconstruction, and show how the framework in Section 3 addresses this issue. In particular, we see that the black hole interior must be viewed as existing in the past of the Hawking radiation containing its information, despite the fact that they are spacelike separated in the semiclassical description. We discuss how the timelike relation can be regarded as spacelike at the semiclassical level because of the limitation of an infaller's ability and the chaotic dynamics of the horizon degrees of freedom.

In Section 5, we extend the framework of Section 3 to the entire history of black hole evolution, including its formation era. This gives a description in which the flow of information is causally trivial throughout the history. Our construction employs bulk holography associated with the stretched event horizon. In Section 6, we discuss the relation of our picture with other works that use the global spacetime description. Section 7 gives conclusions.

Throughout the paper, we adopt natural units $\hbar = c = k = 1$.

2 Exterior Description

Black holes form and then evaporate, and a distant observer expects the evolution to be unitary in agreement with quantum mechanics [4, 41–43]. Unitarity is further supported by the AdS/CFT corre-

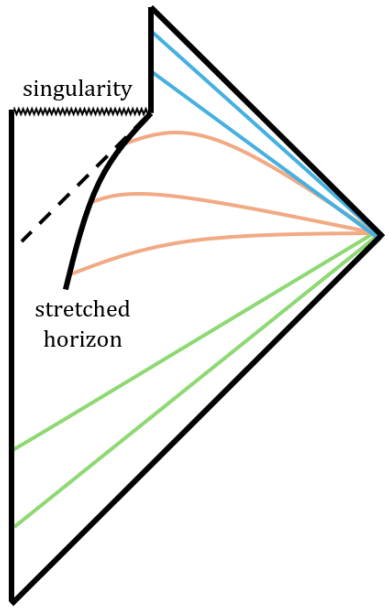

Figure 1: In the exterior description, equal-time slices foliate the spacetime region outside the "inner boundary" of space, which for a (quasi-)static black hole corresponds to the stretched horizon.

spondence [16]. To preserve unitarity, we must follow *all* the relevant degrees of freedom, including the "intrinsically quantum gravitational" (or "Planckian") degrees of freedom associated with the stretched horizon [4]. For a static black hole, this horizon is located at a proper distance of order the cutoff/string length l_s away from the mathematical horizon. The spacetime region relevant for the description is the exterior of the stretched horizon.

A simple way to implement the exterior description is to adopt the canonical formalism of quantum mechanics. The relevant time slices are Schwarzschild-like time slices, i.e. slices that foliate the spacetime region outside the horizon; see Fig. 1. At times when the black hole exists, these time slices cover the region outside the black hole. In the context of holography, these slices can be obtained, for example, by continuously pulling the boundary into the bulk using the method in [44, 45].

A feature of the exterior description is that if the state before the formation of the black hole is pure, then the states with the black hole present and those after it evaporates are also pure. The relevant degrees of freedom in each period are

- **Before the black hole formation:** low energy degrees of freedom whose dynamics can be described by a semiclassical theory. In particular, they include matter that collapses gravitationally to form the black hole.
- **While the black hole is present:** the degrees of freedom outside the horizon. At each time, we can divide these degrees of freedom into three classes (see [32] for more detail):
  - *Horizon modes:* modes on the stretched horizon, which may be viewed as "located" in the

region $r \leq r_{\rm s}$, where $r_{\rm s}$ represents the location of the stretched horizon in the Schwarzschild radial coordinate r (although the classical spacetime picture is not valid in this region).[1]

  - *Zone modes:* modes localized in the region $r_{\rm s} < r \leq r_{\rm z}$, where $r_{\rm z}$ is the location of the outer edge of the near black-hole region called the zone (i.e. the region inside the barrier of the effective gravitational potential generated by the black hole).
  - *Far modes:* modes located in the asymptotic region, $r > r_{\rm z}$.

  This decomposition can be performed at the gauge-fixed level with positive semi-definite inner products between the states, up to the intrinsic ambiguity of separation between horizon and zone modes (we assume that the usual issue of factorizing modes located in different regions in quantum field theory is appropriately taken care of). The dynamics of the horizon modes cannot be described by a low energy semiclassical theory, while those of the zone and far modes can.

- **After the black hole evaporation:** low energy degrees of freedom whose dynamics can be described by the semiclassical theory.[2] In particular, they include Hawking radiation emitted from the black hole.

As has long been argued [4, 41–43], the exterior description should be unitary, since the stretched horizon behaves as a regular material surface as far as the flow of quantum information is concerned. On the other hand, general relativity implies the existence of an interior. While there are proposals otherwise, such as fuzzballs [46–48] and firewalls [49–51], we take the view that nothing special happens at the horizon for an infalling observer. The question then is how we can understand the emergence of a smooth interior for the infaller in the exterior description.

3 Emergence of Global Spacetime

In [27–32], it was shown that an effective theory of a smooth interior can be constructed from the exterior degrees of freedom. In particular, the smooth interior emerges as a secondary concept associated with the modes in the zone describing an object falling into the black hole. Coherent excitations of the rest of the zone and horizon modes (as well as far modes for an old black hole) play the role of the modes in the second exterior of a two-sided geometry, which is emergent in the effective theory. As emphasized in [27–32], it is important to apply this construction to a state *at a fixed time* t in the exterior description. This provides *a portion of* the black hole interior, which can be used to describe the experience of the infalling object at the semiclassical level. To cover the entire black hole interior, we need multiple effective theories erected at different times.

[1] Without an explicit ultraviolet theory, it is scheme dependent if we view these degrees of freedom as additional degrees of freedom beyond the low energy quantum fields or their cutoff-scale excitations near the stretched horizon. In any event, the microscopic dynamics of these degrees of freedom cannot be described by low energy semiclassical theory.

[2] This does not mean that the state of these degrees of freedom is fully specified by operators within the low energy semiclassical theory; see Section 4.

In Section 3.1, we review this construction. We will focus on the main idea, referring the readers to [32] for more details. In Section 3.2, we discuss properties of the construction relevant for our later discussions, especially focusing on how the picture of global spacetime may emerge.

3.1 Effective theory of the interior

For concreteness, we focus on a collapse-formed black hole with negligible spin and charge in 4-dimensional asymptotically flat spacetime. We assume that the black hole has a mass M at some time t, where t is the time in the exterior description.[3] The horizon, zone, and far modes are defined based on their locations *at this time*, according to the criterion given in Section 2. A *vacuum* microstate of the semiclassical black hole at time t, i.e. a state in which excitations described by the semiclassical theory are absent, can then be written as follows.

We first select a subset of the zone modes that will be relevant for describing a falling object which is assumed to lie in the zone at time t. These are called *hard modes*, which we collectively label as α. The semiclassical vacuum microstate, together with Hawking radiation emitted earlier, can then be written as[4]

$$|\Psi(M)\rangle = \sum_n \sum_{i_n=1}^{e^{S_{\rm bh}(M-E_n)}} \sum_a c_{ni_na} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \tag{3.1}$$

where $|\{n_\alpha\}\rangle$ represent states of the hard modes, specified by the occupation number n_α for each mode α, while $|\psi_{i_n}^{(n)}\rangle$ represent orthonormal states of the horizon and zone modes that are not included in the hard modes, which are called *soft modes* (any basis in the i_n space can be taken for each n). Finally, $|\phi_a\rangle$ are the states of the far modes, which generally include Hawking radiation emitted earlier. Note that since the total energy of the black hole system, consisting of the hard and soft modes, is M, the number of basis states $|\psi_{i_n}^{(n)}\rangle$ appearing with $|\{n_\alpha\}\rangle$ in (3.1) is given by $S_{\rm bh}(M - E_n)$. Here, E_n is the energy carried by $|\{n_\alpha\}\rangle$, $S_{\rm bh}(E) = 4\pi E^2 l_{\rm P}^2$ is the Bekenstein-Hawking entropy density at energy E, and $l_{\rm P}$ is the Planck length, which we take to be not too different from $l_{\rm s}$, assuming that the number of low energy species is not enormous.

Because of fast-scrambling [52, 53] and chaotic [54] dynamics at the horizon, (the relevant portion of) the coefficients c_{ni_na} in (3.1) take random values across all low energy species.[5] (The specific values of c_{ni_na} correspond to a specific black hole microstate.) This implies that the state of the system can

[3] More precisely, we focus on a branch of the state at t in which the black hole has mass M and is located at a fixed position, up to irreducible quantum uncertainties of order the Hawking temperature $T_{\rm H}$ and the proper length $l_{\rm s}$, respectively [27]. Our construction of the effective theory of the interior applies to each such branch separately.

[4] Here we ignore the effect of backreaction of Hawking emission on the occupation numbers of hard modes, which has some implication for the region that can be directly reconstructed from the state $|\Psi(M)\rangle$ without resorting to exterior time evolution; for details, see [32].

[5] More precisely, for each a, the coefficients c_{ni_na} are expected to behave as independent and identically distributed complex Gaussian random variables with mean 0.

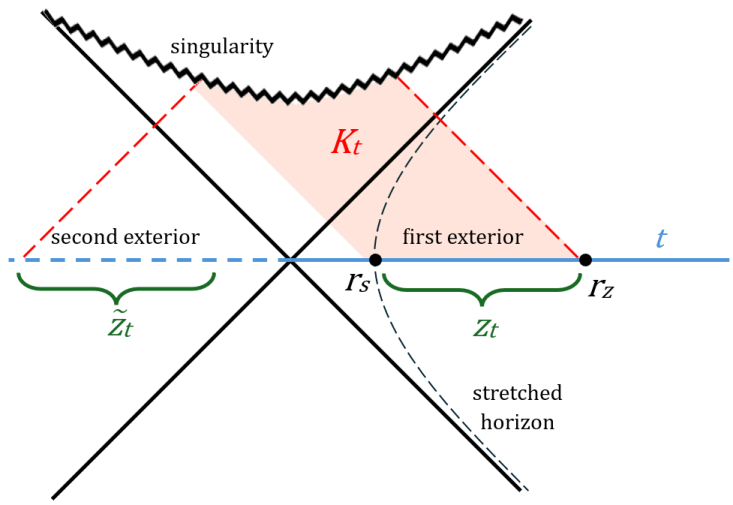

Figure 2: The effective theory of the interior erected at time t gives an emergent two-sided black hole geometry. The degrees of freedom associated with the zone $\tilde{Z}_t$ of the second exterior arise from soft and far mode states coupled to the states in the zone Z_t of the first exterior in the vacuum microstate. The spacetime region described by this effective theory is the region indicated as K_t.

be written in the thermofield double form

$$|\Psi(M)\rangle = \frac{1}{\sqrt{\sum_m e^{-\frac{E_m}{T_{\rm H}}}}} \sum_n e^{-\frac{E_n}{2T_{\rm H}}} |\{n_\alpha\}\rangle \| \{n_\alpha\} \rangle\!\rangle, \tag{3.2}$$

up to exponential accuracy, where $T_{\rm H} = 1/(\partial S_{\rm bh}(E)/\partial E|_{E=M}) = 1/8\pi M l_{\rm P}^2$ is the Hawking temperature. The state with the double-ket symbol is defined as the state of soft and far modes which couples to the state $|\{n_\alpha\}\rangle$ in (3.1)

$$\| \{n_\alpha\} \rangle\!\rangle = \sqrt{\sum_m e^{-\frac{E_m}{T_{\rm H}}}}\, e^{\frac{E_n}{2T_{\rm H}}} \sum_{i_n=1}^{e^{S_{\rm bh}(M-E_n)}} \sum_a c_{n i_n a} |\psi_{i_n}^{(n)}\rangle |\phi_a\rangle, \tag{3.3}$$

where the overall coefficient is determined by the normalization condition, $\langle\!\langle \{n_\alpha\} \| \{n_\alpha\} \rangle\!\rangle = 1$. We see that the double-ket states play the role of the states in the second exterior of the effective two-sided black hole, as depicted in Fig. 2. This allows us to define, along the lines of [55–59], annihilation and creation operators

$$b_\gamma = \sum_n \sqrt{n_\gamma}\, |\{n_\alpha - \delta_{\alpha\gamma}\}\rangle\langle\{n_\alpha\}|, \qquad b_\gamma^\dagger = \sum_n \sqrt{n_\gamma + 1}\, |\{n_\alpha + \delta_{\alpha\gamma}\}\rangle\langle\{n_\alpha\}|, \tag{3.4}$$

$$\tilde{b}_\gamma = \sum_n \sqrt{n_\gamma}\, \| \{n_\alpha - \delta_{\alpha\gamma}\} \rangle\!\rangle \langle\!\langle \{n_\alpha\} \|, \qquad \tilde{b}_\gamma^\dagger = \sum_n \sqrt{n_\gamma + 1}\, \| \{n_\alpha + \delta_{\alpha\gamma}\} \rangle\!\rangle \langle\!\langle \{n_\alpha\} \|, \tag{3.5}$$

from which we can construct operators describing the interior of the black hole, using only the degrees of freedom on the exterior region, i.e. the hard, soft, and far modes defined on the time slice t.[6]

[6] While the operators in (3.5) depend sensitively on the microstate of the black hole, i.e. the precise values of $c_{n i_n a}$ in (3.1), we can promote them to apply to a wide range of microstates to the extent that it is sufficient to ensure the validity of the semiclassical description of the interior [31, 60, 61].

The operators describing semiclassical excitations in the interior region, obtained from Eqs. (3.4) and (3.5), act on all of the hard, soft, and far modes. However, if the black hole is young, i.e. if the combined system of the hard and soft modes is not maximally entangled with far modes at time t, then the interior operators can be constructed without involving far modes [29, 31]. This can be done by projecting the operators acting on all modes onto the Hilbert space of hard and soft modes using the Petz map. Schematically, when the entanglement between the black hole and far modes is negligible, the vacuum microstate can be written as

$$|\Psi(M)\rangle \approx \sum_n \sum_{i_n=1}^{e^{S_{\rm bh}(M-E_n)}} c_{ni_n} |\{n_\alpha\}\rangle |\psi_{i_n}^{(n)}\rangle. \tag{3.6}$$

The double-ket states can then be defined by

$$\|\{n_\alpha\}\rangle\rangle \approx \sqrt{\sum_m e^{-\frac{E_m}{T_{\rm H}}}} e^{\frac{E_n}{2T_{\rm H}}} \sum_{i_n=1}^{e^{S_{\rm bh}(M-E_n)}} c_{ni_n} |\psi_{i_n}^{(n)}\rangle, \tag{3.7}$$

instead of (3.3), giving interior operators acting only on hard and soft modes. For a young black hole, these operators reproduce the same algebra as that obtained using the "full" double-ket states in (3.3), up to exponentially suppressed corrections.

We stress that in the case of either a young or an old black hole, interior operators constructed as above *must* involve soft modes, i.e. modes near the black hole region. This implies, in particular, that the construction does not work when a black hole does not exist at time t. This is different from entanglement wedge reconstruction [17–24], which allows for reconstructing interior operators only using far modes (early radiation) [5–7]. The difference comes from the fact that the construction described here does not use exterior time evolution with respect to t in converting the fundamental modes to infalling modes, while entanglement wedge reconstruction (implicitly) uses it by assuming the knowledge of microscopic time evolution [30, 32].

For an evaporating black hole, the spacetime region K_t that can be obtained by the construction above is a portion of the interior which corresponds to the intersection of the domain of dependence of the union of Z_t and $\tilde{Z}_t$ with the causal future of Z_t [32]:

$$K_t = D(Z_t \cup \tilde{Z}_t) \cap J^+(Z_t), \tag{3.8}$$

where Z_t represents the zone at time t, i.e. the partial Cauchy surface $r_{\rm s} < r \leq r_{\rm z}$ at t, and $\tilde{Z}_t$ is its second exterior mirror in the effective two-sided geometry arising in the effective theory; see Fig. 2. Any operator of the semiclassical theory contained in this region can be constructed using the time evolution operator $\tilde{U}(\tau) = e^{-i\tilde{H}\tau}$ associated with an infalling frame. Here,

$$\tilde{H} = \sum_\xi \Omega_\xi a_\xi^\dagger a_\xi + \tilde{H}_{\rm int}\big(\{a_\xi\}, \{a_\xi^\dagger\}\big) \tag{3.9}$$

is the generator of the infalling time evolution, and τ is the infalling time; a_ξ and $a_\xi^\dagger$ are the annihilation and creation operators for infalling mode ξ, obtained from b_γ, $b_\gamma^\dagger$, $\tilde{b}_\gamma$, and $\tilde{b}_\gamma^\dagger$ by Bogoliubov transformations, and Ω_ξ is the frequency of the mode ξ. Note that this time evolution is *different* from the original

time evolution, which acts on all the microscopic degrees of freedom. In fact, the effective theory of the interior obtained in this way is intrinsically semiclassical, possessing intrinsic ambiguities of order $e^{-S_{\rm tot}/2}$, where $S_{\rm tot}$ is the coarse-grained entropy of the combined black hole and radiation system [31].

The picture presented here implies that a generic state in the Hilbert space of microstates for a collapse-formed, single-sided black hole does not have a firewall. In [51], it was argued that since on a black hole background "arbitrarily many" configurations of matter can be excited without costing energy as viewed from a distance (by putting them near the horizon), a generic state of the black hole Hilbert space has a firewall. These configurations, however, correspond precisely to black hole microstates; specific values of c_{ni_n} in (3.6), or c_{ni_na} in (3.1), correspond simply to a specific *vacuum* microstate. For generic values of these coefficients, the present construction works, giving an effective description in which a falling object goes through the horizon smoothly. In particular, the fate of a semiclassical object in the zone, represented by a set of $b_\gamma^\dagger$ operators *acted on* the vacuum microstate, can be calculated by evolving the system by $\tilde{U}(\tau)$ using the in-in formalism, reproducing the result consistent with general relativity. It is only for an exponentially atypical state (e.g. a state in which one of the coefficients is exponentially larger than the others) that this construction does not work, leading to a firewall at the horizon.[7]

3.2 Global spacetime from a patchwork

From the exterior description, the global spacetime picture is obtained by erecting effective theories of the interior described above at different times t and "patching" them [27], as sketched in Fig. 3. We first highlight two important aspects in this procedure:

(i) There are many different ways to represent the same interior operator in fundamental (hard, soft, and far) degrees of freedom. To see this, consider the state at time $t' > t + O(1/T_{\rm H})$ of the system that was described by (3.1) at time t. This state can be written as

$$|\Psi(M')\rangle \approx \sum_{n'} \sum_{i'_{n'}=1}^{e^{S_{\rm bh}(M'-E_{n'})}} \sum_{a'} c_{n'i'_{n'}a'} |\{n'_{\alpha'}\}\rangle |\psi^{(n')}_{i'_{n'}}\rangle |\phi_{a'}\rangle, \tag{3.10}$$

where $M' < M - O(T_{\rm H})$ is the black hole mass at t', and $|\{n'_{\alpha'}\}\rangle$, $|\psi^{(n')}_{i'_{n'}}\rangle$, and $|\phi_{a'}\rangle$ are the states of the hard, soft, and far modes *defined at that time*. If $t' - t$ is smaller than the signal propagating time between the edge of the zone and the stretched horizon, then the interior portions of K_t and $K_{t'}$ constructed from the two states in Eqs. (3.1) and (3.10) overlap. This allows us to represent the same interior operator in $K_t \cap K_{t'}$ differently—as O and O'—using the microscopic degrees of freedom appearing in two vacuum microstates Eqs. (3.1) and (3.10).

[7] This can be phrased such that a generic single-sided state is mapped onto a special (thermofield double) two-sided state in the effective description. This occurs because of the state-dependent [31, 55–61] definition of infalling mode operators. For refutations against other arguments [49, 50, 62, 63] for firewalls, see [27–32] and references therein.

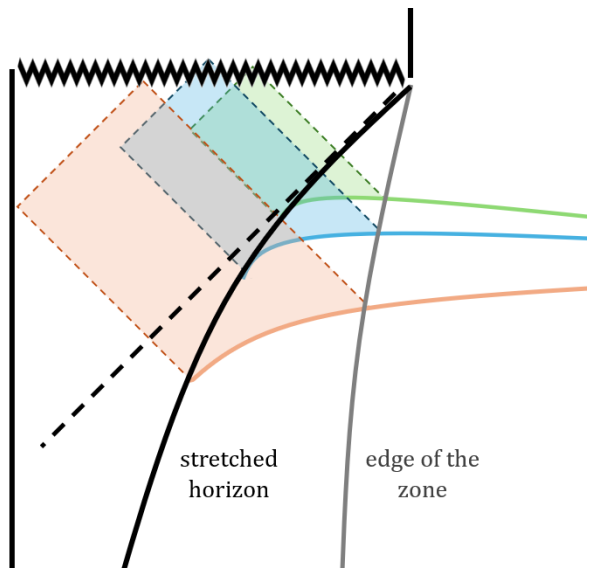

Figure 3: The effective theory of the interior erected at an exterior time t describes the region of the black hole interior corresponding to the interior of the effective two-sided black hole composed of the zone at t and its second exterior mirror. To obtain the whole black hole interior, we need to patch the pictures obtained by effective theories erected at different exterior times.

For $M - M' > O(T_{\rm H})$, the hard mode states appearing in Eqs. (3.1) and (3.10) are orthogonal (possibly up to exponentially small overlaps), since the black holes at times t and t' are macroscopically distinguishable; the same is true for the soft and far mode states. It follows that O and O' are not related simply by the original exterior time evolution $U(t', t)$:

$$O' \neq U(t', t)\, O\, U^\dagger(t', t). \tag{3.11}$$

This can be seen, for example, from the fact that some of the hard modes at time t evolve into far modes at t', and that for a young black hole, interior operators can be constructed without involving far mode degrees of freedom. The argument here demonstrates that a version of quantum error correction discussed in [64] also applies here in constructing interior operators from the bulk degrees of freedom, i.e. the hard, soft, and far modes.

(ii) Different operators in the interior region, in particular operators that are independent at the semiclassical level, involve the same microscopic degrees of freedom. This is, in a sense, the converse of the phenomenon described in (i). This violation of independence comes from the fact that soft modes at some time t are not necessarily independent from hard modes at some other time t', as they can be related by time evolution $U(t, t')$. However, since the number of independent soft mode and radiation states is exponentially larger than that of hard mode states, the degree of violation arising in this way is suppressed exponentially by the factor $e^{-S_{\rm tot}/2}$.

These two features reinforce the conclusion that the picture of the black hole interior is intrinsically semiclassical: it is defined only up to the precision of order $e^{-S_{\rm tot}/2}$.

We now discuss the third aspect, which is particularly relevant for our discussion:

(iii) If the state at time t is given by (3.1), then the operator $\Phi_\Gamma(x)$ that acts *on this state* and represents a local Heisenberg-picture operator at $x \in K_t$ in (3.8) is given by

$$\Phi_\Gamma(x) = \sum_{s,\Omega,\mathbf{L}} \left(f_s(\Omega,\mathbf{L})\, \varphi_{\Omega,\mathbf{L}}(x)\, a_\xi + g_s(\Omega,\mathbf{L})\, \varphi^*_{\Omega,\mathbf{L}}(x)\, a^\dagger_{\xi^c} \right). \tag{3.12}$$

Here, we have decomposed index ξ of the infalling mode operators a_ξ and $a^\dagger_\xi$ into Γ, s, Ω, and $\mathbf{L} = \{\ell, m\}$ which represent species, spin, frequency, and orbital angular momentum quantum numbers, respectively, with ξ^c representing the CPT conjugate of ξ; $f_s(\Omega,\mathbf{L})$ and $g_s(\Omega,\mathbf{L})$ are the standard factors providing Lorentz representation of the field (Dirac spinors, polarization vectors, etc), and $\varphi_{\Omega,\mathbf{L}}(x)$ are the mode functions. Note that since a_ξ and $a^\dagger_\xi$ are given by Bogoliubov transformations of b_γ, $b^\dagger_\gamma$, $\tilde{b}_\gamma$, and $\tilde{b}^\dagger_\gamma$ in Eqs. (3.4) and (3.5), these operators act on the hard, soft, and far mode states defined at t.

As depicted in Fig. 3, in order to cover the entire black hole interior I, we need to erect the effective theory of the interior in at least every time interval of order $1/2\pi T_{\rm H}$. In other words, in order to obtain the complete set of interior operators $\{\Phi_\Gamma(x)\,|\,x \in I\}$, we need to use a_ξ and $a^\dagger_\xi$ defined at different times t_x in constructing operators at different locations x, since not all $x \in I$ belong to the same K_t for some t. While operators obtained in this way act generally on the hard, soft, and far mode states at different times t_x, we can time evolve them to obtain the set of operators that all act on the radiation state at time $t_{\rm r}$ after the black hole is fully evaporated, i.e. $t_{\rm r} > t_{\rm evap}$:

$$\mathcal{I}_{t_{\rm r}} = \left\{ \mathcal{O}_{\Gamma,x}(t_{\rm r}) \equiv U(t_{\rm r}, t_x)\Phi_\Gamma(x)U^\dagger(t_{\rm r}, t_x) \,\middle|\, x \in I \right\}. \tag{3.13}$$

As discussed in [30, 32], this is how entanglement wedge reconstruction [17–24] represents all the interior operators in the final state Hawking radiation. We stress that from this perspective, entanglement wedge reconstruction of interior operators from Hawking radiation involves evolution *backward in time*, despite the fact that the interior region and the region in which the radiation exists are spacelike separated in the global spacetime picture.

The operators $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ acting on final-state Hawking radiation represent local operators in the interior of the black hole, a region *spacelike separated* from the region where the radiation resides. This seems to imply that an operation on the Hawking radiation corresponding to the application of $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ can affect the experience of an observer falling into the black hole, potentially violating the causality of general relativity.[8] This issue was studied in [25] (see also [65, 66]) for radiation before complete evaporation (the case in footnote 8), wherein it was shown that the operation of acting $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ is so

[8] This argument can be extended to radiation emitted from an old black hole, i.e. to the case $t_{\rm Page} < t_{\rm r} < t_{\rm evap}$, where $t_{\rm Page}$ is the Page time [43]. In this case, the amount of information in the relevant interior region, $(\bigcup_{t<t_{\rm r}-t_{\rm scr}} K_t) \cap I$, accessible in the radiation is limited and depends on the location in this region, where $t_{\rm scr}$ is the scrambling time [52, 53]. See Fig. 13 of [32] and discussion associated with it for more details.

highly complex that locality in the semiclassical description (i.e. the spacelike separated nature between the interior and radiation regions) is preserved for all practical purposes.

However, as we will see in the next section, this leaves a major puzzle associated with causality in the global spacetime picture. We will see that the construction given in [27–32] and described here addresses this puzzle. In particular, we will see that in our picture, unlike what was envisioned in earlier works [25,26,67], one *cannot* affect the experience of an observer who has fallen into the black hole at an earlier time by performing an (even complex) operation on radiation emitted from the black hole. This is because, at the microscopic level, the region inside the black hole should be regarded as being *timelike separated* from the region in which the radiation resides; more specifically, the former is in the past of the latter. The fact that operators like $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ exist simply reflects the standard statement that there are operators at a time t corresponding to changing the *record* of what happened at times earlier than t. Together with the fact that the $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ effectively commute with a class of simple operators acting on radiation—the feature arising from the involvement of complex horizon dynamics—this allows us to treat the past event represented by $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ as occurring in a spacelike separated region in semiclassical spacetime.

4 A Puzzle in the Global Spacetime Picture and Its Resolution

In this section, we discuss a problem with global spacetime when it is used to formulate standard quantum mechanics at the fundamental level. We will then show that by using the description of the black hole interior presented in the previous section, this problem is avoided. In this description, the Hawking radiation containing information about the interior always lies in the causal future of the corresponding interior region. It is only in the semiclassical description that the interior appears causally disconnected from the radiation region.

We will discuss in detail how this apparent change of the description arises from the complex dynamics at the horizon, as viewed from the exterior. By the semiclassical description, we mean the description including only finite products of local operators in the limit that the Newton constant is sent to zero, $G_{\rm N} \to 0$. For operators acting on the final-state Hawking radiation, this corresponds to operators that act on a number of degrees of freedom which does not scale with the entropy of the radiation.[9] We will show that these operators all effectively commute with operators describing the interior, so that they can be viewed as spacelike separated at the semiclassical level.

4.1 A problem with the microscopic description of global spacetime

We argue that the following three postulates cannot be satisfied at the same time:

(i) Diffeomorphism invariance persists throughout inextendible (global) spacetime, i.e. the region in

[9] This implies that the semiclassical theory does not describe *everything* about the low energy degrees of freedom.

which the classical description of spacetime is available in general relativity.

(ii) The canonical formalism of quantum mechanics can be built using Cauchy surfaces in global spacetime, at least when they satisfy certain "niceness" properties [68–70].

(iii) An observer in the exterior region can reconstruct local operators in the interior using only Hawking radiation degrees of freedom.

We will show this by demonstrating that the above three postulates, taken together, lead to a violation of the no-cloning theorem of quantum mechanics.

To see this, consider a setup in which a detector staying outside the horizon throughout the history of the black hole evolution extracts information about the interior at some time $t_{\rm r}$ after the Page time $t_{\rm Page}$.[10] For simplicity, we focus on the case in which $t_{\rm r}$ is after the full evaporation of the black hole, $t_{\rm r} > t_{\rm evap}$, though the extension to the case with $t_{\rm Page} < t_{\rm r} < t_{\rm evap}$ is straightforward (see footnote 8).

While the operation performed by the detector to achieve the information extraction needs to be highly nonlocal, acting on most of the radiation degrees of freedom, we can fine-tune the initial configuration of the detector so that it reads off information corresponding to a semiclassical excitation in the interior.[11] As a concrete example, we can set up the initial configuration of the detector such that it interacts with the degree of freedom in the Hawking radiation which is represented by a product of operators $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ for some spacetime point $x \in I$. Note that this requires an exponential fine-tuning of the configuration of the detector, which must be determined by the microscopic theory using the knowledge of the initial state of collapsing matter forming the black hole. This is difficult in practice but possible in principle, because we can have as much time as we need to prepare the detector before actually forming the black hole.

With this setup, we find it problematic to consider the global spacetime as a platform for a regular quantum mechanical system. In particular, we cannot employ the canonical formalism of quantum mechanics using a series of Cauchy surfaces, such as nice slices [68–70], in which the information extracted by the detector is removed from the future of x in the interior and appears in the future of r in the exterior region. This can be seen by considering a foliation of spacetime accommodating a time slice that goes through both the past of x and the future of the region in which the detector extracts the information; see Fig. 4. On this slice, the information extracted by the detector is cloned, since the destruction of the same information in the interior occurs at x and hence has not yet happened on this time slice. This contradicts the no-cloning theorem, hence violating the superposition principle of

[10] Strictly speaking, $t_{\rm r}$ is not exactly a single time but has a "width," since it will take some time for the detector to extract the information from Hawking radiation through interactions with it; but we will loosely refer to the period of time in which the detector extracts the information as $t_{\rm r}$.

[11] The corresponding operator acting on the radiation comprises a product of $S_{\rm rad}$ local operators, where $S_{\rm rad}$ is the coarse-grained entropy of the radiation, so that the required fine-tuning is of order $e^{-O(S_{\rm rad})}$. We do not consider such an operator to be a part of the semiclassical theory. Thus, the problem discussed here is not that of the semiclassical description but of the fundamental description applied to the global spacetime.

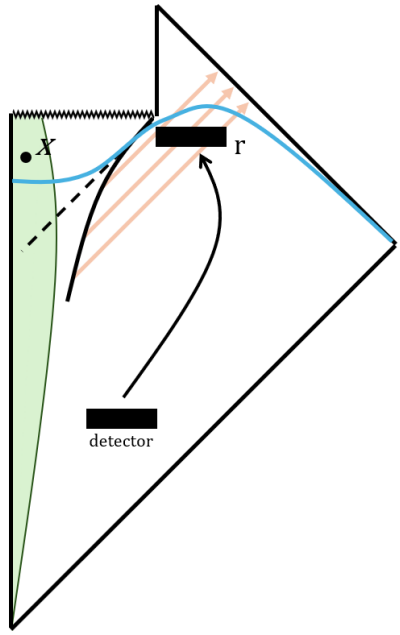

Figure 4: A detector can be set up such that it extracts from Hawking radiation the information that is located at a spacetime point x in the interior of the black hole. Interpreting this as being caused by nonlocal interactions between x and the spacetime region r in which the radiation resides, we encounter the problem of the existence of a time slice (blue line) that goes through both the future of r and the past of x. On such a slice, the information in question is cloned, leading to a contradiction with the no-cloning theorem of quantum mechanics.

quantum mechanics. We stress that this problem occurs even when we only use slices that are nice in the semiclassical description.

One might think that this problem is avoided by claiming that the microscopic description of the global spacetime forces us to use a special foliation in which the time slice going through t_{r} *must* also go through x. We find it difficult, however, to make sense of such a claim given that the arbitrariness of choosing time foliation, i.e. diffeomorphism invariance, plays an important role in ensuring the consistency of theories with dynamical gravity, even at the perturbative level.[12]

4.2 A solution: the interior as an effective description of dissipation on the horizon

The above argument implies that, if we are to keep (iii), we must relinquish (i) or (ii). The understanding of the black hole interior described in Section 3 addresses the puzzle by abandoning postulate (ii). In this framework, the interior spacetime—in fact only a portion of it—emerges as an effective description associated with a given exterior time. The effective time evolution foliating the interior region is different from the more fundamental, exterior time evolution, which foliates only the spacetime region outside

[12] This is a difficulty which we encounter in addition to the fact that the time t_{r} has a width, so that the relevant time slice would generally be ill-defined.

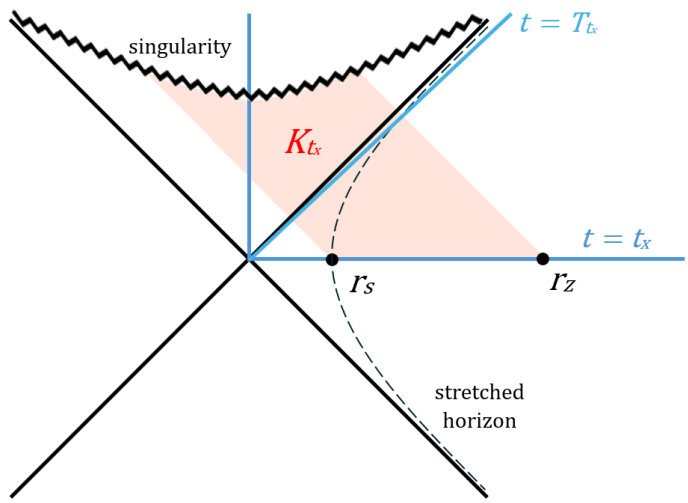

Figure 5: The physics occurring in the region described by the effective theory erected at time t_x cannot be obtained by any effective theory erected after $T_{t_x} = t_x + 2r_+ \ln(r_+/l_{\rm P})$, implying that it must be viewed as having occurred before T_{t_x} in the exterior description.

the horizon. Consequently, any information in the effective interior is contained in modes *outside the horizon*, and hence exists in the causal past of the radiation. By viewing the interior this way, we need not demand that the interior is foliable by global time slices, and hence do not take postulate (ii). The property in postulate (i) emerges only effectively after patching the pictures obtained by effective theories erected at different times. This results in postulate (iii) not implying that an operation on Hawking radiation can affect the physics in the interior, avoiding the problem described in Section 4.1.

Recall that in the framework of Section 3, physics occurring at the interior point x is described by an effective theory erected at a time t_x ($< t_{\rm evap}$) satisfying

$$x \in K_{t_x}, \tag{4.1}$$

where K_t is given by (3.8).[13] As can be seen in Fig. 5, the spacetime region K_{t_x} described by the effective theory contains (a portion of) the zone, $r_{\rm s} < r < r_{\rm z}$, only up to the time

$$T_{t_x} = t_x + 2r_+ \ln \frac{r_+}{l_{\rm P}}, \tag{4.2}$$

where r_+ is the horizon radius of the black hole *at time* t_x, and we have ignored the term of order r_+ which is not enhanced by the logarithm $\ln(r_+/l_{\rm P})$. (We will also do the same below.) This implies that in the exterior description, the physics occurring in the interior portion of K_{t_x} should be viewed

[13] It is important that K_{t_x} only contains the future of the region *outside* the stretched horizon; it does not contain the region which is in the future of $r < r_{\rm s}$ at t_x but not in the future of $r > r_{\rm s}$ at t_x, as illustrated in Fig. 5. This ultraviolet cutoff is a manifestation of the fact that the effective theory of the interior is intrinsically semiclassical, without which all the radial null geodesics entering the stretched horizon before t_x would pile up along the horizon of the second exterior in the effective theory causing large backreaction on the spacetime [27].

as occurring in the time frame

$$t_x < t < T_{t_x}, \tag{4.3}$$

since the effective theory simply reorganizes correlations between the degrees of freedom associated with the time slices in (4.3) such that they can be interpreted as arising from the infalling time evolution generated by $\tilde{H}$ in (3.9), *without* involving a further exterior time evolution. The earliest effective theory to which x can be associated is the one erected at time $t_{\min} = \min\{t_x : x \in K_{t_x}\}$. The latest effective theory which x can be associated with is similarly the one erected at $t_{\max} = \max\{t_x : x \in K_{t_x}\}$. Consequently, x must be associated with exterior modes between $t_{\min}$ and $T_{t_{\max}}$. In other words, the physics occurring in the interior spacetime point x must be viewed as occurring in the time frame

$$t_{\min} < t < T_{t_{\max}} \tag{4.4}$$

in the exterior description.

Importantly, for any point x in the interior of the black hole

$$T_{t_{\max}} < t_{\text{evap}}. \tag{4.5}$$

This implies that in the exterior description, the physics which an interior observer experiences at the spacetime point x must have *already happened* before the black hole evaporates. Thus, whatever the detector does to the final state Hawking radiation, it *cannot affect* their experience at x nor events occurring in its future. Exerting operators $\mathcal{O}_{\Gamma,x}(t_{\text{r}})$ on the radiation at t_{r} can *reconstruct* the event at x, or even *rewrite the record* of what happened at x and in its future; however, it cannot change the actual experience which the interior observer has had inside the black hole.

In this respect, the situation is not so different from the standard situation in which a complicated operation can be applied to the system "after the fact" so that it can change the record of what has happened before the operation.[14] In our setup, if we reconstruct from Hawking radiation what happened in the interior of the black hole before applying the complex operation, i.e. at $t < t_{\text{r}}$, then we would conclude that nothing extraordinary happened at x in the interior. On the other hand, if we do such a reconstruction after the detector exerts the complex operation, i.e. at $t > t_{\text{r}}$, then we would conclude that an event occurred at x "out of the blue" which did affect the future of x in the interior. (The past of x is unaffected because the boundary condition for the interior is determined by the exterior state at t_x, fixing the state in the past light cone of x.) This "nonlocal influence" by the operation, however, is an illusion—what a physical observer *actually experiences* in the interior is not affected by the operation of the detector.

In the exterior description, the event x must thus be viewed as having occurred *in the past of* r, despite the fact that they are spacelike separated in the semiclassical description. This picture becomes

[14] Throughout the paper, we assume that things happen when they actually occur through time evolution. In other words, we consider that a measurement occurs when the relevant information corresponding to the measurement outcome (which is not full quantum information) is amplified into various subsystems [71–73]. In particular, we do not take an extreme (and what we think an exotic) standpoint [74] that a measurement occurs only when the relevant information reaches to a boundary or horizon region.

evident because we have employed the construction in [27–32], which makes manifest the dynamical nature of interior reconstruction, i.e., how the exterior time evolution is involved in the reconstruction. In fact, this feature is not visible in a model without appropriately taking into account time evolution, for example in the case of a tensor network corresponding to a single global time slice.

An alternative way of viewing all of this is that to correctly reconstruct the interior, we need to include the agent, e.g. a detector, which has exerted operators $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ on the radiation. We should then evolve the *combined* system of the agent and radiation backward in time and use the construction of [27–32] at time t_x. By doing this, rather than simply adopting the entanglement wedge reconstruction protocol to the radiation after $\mathcal{O}_{\Gamma,x}(t_{\rm r})$'s are operated, we will indeed reconstruct the interior in which nothing extraordinary happens at x. The inclusion of the agent affects the interior significantly, despite the fact that it comprises a much smaller number of degrees of freedom than those of the radiation. This is because the standard process of black hole formation and evaporation is one in which the coarse-grained entropy increases [75, 76], and hence the reverse time evolution employed during interior reconstruction is a highly fine-tuned, fragile process.

The analysis here can be extended to the case in which we act operators $\mathcal{O}_{\Gamma,x}(t_{\rm r})$ before the black hole is fully evaporated, which is possible if $t_{\rm r} > t_{\rm Page}$. To see this, consider an infalling observer who reaches the stretched horizon at some exterior time $t_{\rm f}$. While the experience of the observer shortly after crossing the horizon is described by the effective theory of the interior erected at $t_{\rm f}$, this does not necessarily mean that all the experiences of this observer are described in this effective theory. For example, the observer may accelerate toward the horizon right after they cross it, eventually leaving $K_{t_{\rm f}}$; see Fig. 6. Even in this case, however, the experience of the infallen observer should be viewed as fully completed by

$$t_{\rm comp} = t_{\rm f} + t_{\rm scr} \tag{4.6}$$

in the exterior description, where

$$t_{\rm scr} = 4r_+ \ln \frac{r_+}{l_{\rm P}} \tag{4.7}$$

is the scrambling time. One way to see this is to notice that the late trajectory of the observer is fully contained in the effective theory erected at $t'_{\rm f} = t_{\rm f} + 2r_+ \ln(r_+/l_{\rm P})$, even in the worst case that the observer starts moving outward at the speed of light one cutoff time after crossing the mathematical horizon. This is indicated by the arrow in Fig. 6. The argument that led to (4.3) then says that any experience of the observer in the interior must have occurred by

$$T_{t'_{\rm f}} = t_{\rm f} + t_{\rm scr}. \tag{4.8}$$

On the other hand, the general argument in [52] (or entanglement wedge reconstruction [5]) says that the information about an object that fell into the black hole at $t_{\rm f}$ emerges in Hawking radiation only after $t_{\rm f} + t_{\rm scr}$. Thus, again any operation on Hawking radiation—even if it is performed before $t_{\rm evap}$, and no matter how complex it is—cannot affect the experience of an observer in the interior. This conclusion

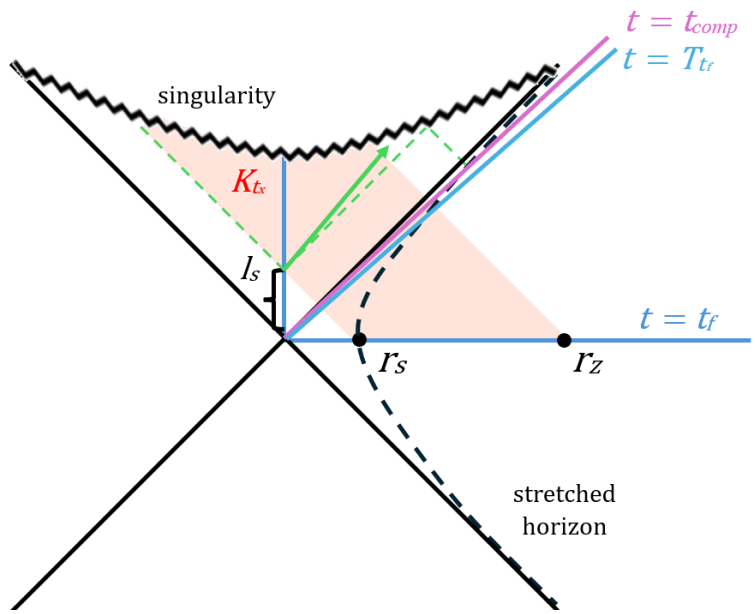

Figure 6: The physics associated with an object that fell into the horizon at time $t_{\rm f}$ cannot be described by any effective theory erected after $t_{\rm comp} = t_{\rm f} + t_{\rm scr}$. This implies that it must be viewed as having occurred before $t_{\rm comp}$ in the exterior description.

holds even if we mine [33, 34] radiation from the Hawking cloud in the zone around the black hole.[15]

Finally, our analysis yields the following interpretation [27] of the black hole's singularity. From an exterior perspective, an object does not immediately thermalize when it encounters the stretched horizon. Instead, it manifests as a distinct excitation on the horizon for some duration. Over the scrambling time, $t_{\rm scr}$, this excitation will gradually dissipate into the horizon. During this period, the dissipation process has an effective description as an object traversing the near-empty interior of the black hole. Here, coherent "quasi-particle" horizon modes—and beyond the Page time, early radiation modes as well—denoted by $\tilde{b}_\gamma$, serve as modes in the effective second exterior of the black hole. The singularity of the semiclassical picture signifies the limit of the applicability of this effective description. After the object's excitation is fully integrated into the horizon, the effective theory describing the dissipation process is no longer available.

4.3 Converting relations from timelike to spacelike using complexity

It is highly nontrivial to regard two sets of degrees of freedom that are not independent with each other as spacelike separated at the semiclassical level. If interior reconstruction indeed works, Hawking radiation for an old black hole is not independent of the interior degrees of freedom. To regard these as spacelike separated, while preserving causality at the semiclassical level, an operator acting on one set

[15] To reconstruct the interior of the black hole at an earlier time t from radiation at $t_{\rm r}$ ($> t$, $< t_{\rm evap}$), we need to include all semiclassically distinguishable branches (with different black hole masses and positions) that have resulted from the backreaction [77, 78] of Hawking emission occurring after t.

must commute with that acting on the other, possibly up to corrections suppressed exponentially in a macroscopic entropy. Here we argue that our framework indeed implies that a process involving only low energy degrees of freedom cannot affect the semiclassical physics in a region that can be viewed as spacelike separated.

In [25], Kim, Tang, and Preskill (KTP) analyzed this issue of causality, assuming that the black hole's dynamics can be represented by pseudorandom unitaries [25, 79, 80]. Since it is expected that chaotic dynamics generally produce pseudorandom states sufficiently after the scrambling time, we assume that the same applies to the dynamics of our horizon modes as viewed from the exterior. KTP studied a partially evaporated black hole and concluded that if an operator $\mathcal{O}_{\rm rad}$ acting on the radiation can be modeled by a quantum circuit with size polynomial in the entropy of the black hole, i.e. $\mathcal{C}(\mathcal{O}_{\rm rad}) \sim {\rm poly}(S_{\rm BH})$, then $\mathcal{O}_{\rm rad}$ commutes with operators $\mathcal{O}(x)$ describing the interior up to corrections suppressed exponentially in $S_{\rm BH}$

$$\|[\mathcal{O}_{\rm rad}, \mathcal{O}(x)]\| \sim O(e^{-S_{\rm BH}}). \tag{4.9}$$

Applying the same argument to the final-state Hawking radiation, we could conclude that an operator $\mathcal{O}'_{\rm rad}$ acting only on a portion $S'_{\rm rad}$ ($< S_{\rm rad}$) of the radiation whose complexity goes as

$$\mathcal{C}(\mathcal{O}'_{\rm rad}) \sim {\rm poly}\left(S_{\rm rad} - S'_{\rm rad}\right) \tag{4.10}$$

would effectively commute with operators describing the interior of the black hole: $\|[\mathcal{O}'_{\rm rad}, \mathcal{O}(x)]\| \sim O(e^{-(S_{\rm rad}-S'_{\rm rad})})$.

However, this by itself does not solve the issue, since one would still be able to act on an operator on the entire final-state radiation ($S'_{\rm rad} \approx S_{\rm rad}$) and affect the interior; indeed in that case the operator can be simple.[16] This would lead to the problem discussed in Section 4.1. As we have seen, however, the Hawking radiation containing information about the interior is, in fact, *timelike* related with the interior degrees of freedom. The former is in the future of the latter. Therefore, one cannot change the physics that has already occurred in the interior by acting with an operator on the Hawking radiation, *no matter how complex it is.*

This leaves the question about the effect of an interior operator (past) on Hawking radiation (future). In order to preserve causality, we want the semiclassical physics associated with the radiation not to be affected by an operation in the interior. By semiclassical physics, we mean the collection of processes represented by operators which act on degrees of freedom whose numbers do not scale with $S_{\rm rad}$, which is proportional to the initial entropy of the black hole and hence to $1/G_{\rm N}$. To study this question, we can employ a variant of the KTP analysis; the interior operator in KTP is replaced with a semiclassical operator acting on the radiation, and the combined system of the black hole and radiation in KTP is

[16] Here we have assumed, following [25], that the state of final-state Hawking radiation is efficiently generated. It seems possible to us that the generation of final-state Hawking radiation could be hard, which can happen if, for example, the computational basis of Hawking radiation (the basis determined by spatial locality) and the natural basis of the horizon dynamics are related by a generic unitary. If this is the case, then the operator affecting the interior must be complex, with the complexity $\mathcal{C}$ scaling exponentially in $S_{\rm rad}$.

replaced with the black hole degrees of freedom at the time when the effective theory of the interior is erected.[17] The number of degrees of freedom representing semiclassical physics in the interior is much smaller than that representing the black hole, $S_{\rm BH}$. The KTP argument then implies that an interior operator $\mathcal{O}_{\rm int}$ corresponding to a circuit of size

$$\mathcal{C}(\mathcal{O}_{\rm int}) \sim \text{poly}\,(S_{\rm BH}) \tag{4.11}$$

effectively commutes with semiclassical operators $\tilde{\mathcal{O}}_{\rm rad}$ acting on the radiation, up to corrections suppressed exponentially in $S_{\rm BH}$.

$$\left\|[\mathcal{O}_{\rm int}, \tilde{\mathcal{O}}_{\rm rad}]\right\| \sim O(e^{-S_{\rm BH}}). \tag{4.12}$$

In other words, in order to affect the semiclassical physics of the radiation, an operation performed in the interior must have complexity exponential in $S_{\rm BH}$. This is, however, impossible because any physical "experimenter" in the interior hits the singularity in polynomial time before any such operation can be performed. Note that we have not imposed any constraints on the computational complexity of $\tilde{\mathcal{O}}_{\rm rad}$; we have just imposed the restriction on the number of degrees of freedom on which it acts.

Any physical operation in the black hole interior, therefore, cannot affect semiclassical physics occurring in the radiation at later times, except for corrections suppressed exponentially in the entropy of the black hole. The only operators acting on the radiation which do not commute with interior operators are those acting on a large (macroscopic) number of degrees of freedom scaling as $S_{\rm rad}$. Of course, these noncommutativities are what allow us to decode the information about the interior from the Hawking radiation.

Finally, we note that the analogous statement cannot be made for reverse time evolution. We *can* perform a physical operation on the (now initial-state) radiation which affects semiclassical physics in the interior (of the white hole). In this sense, we may say that the Penrose diagram of Fig. 1 faithfully represents physics for the standard black hole evolution, a process in which the coarse-grained entropy increases, but not for the reverse evolution, a process with decreasing coarse-grained entropy caused by a finely-tuned initial condition.

5 Causal Structure of Black Hole Formation and Evaporation

One virtue of the exterior picture is that the causal flow of information is clear. When an object falls into a black hole, its information stays on the stretched horizon, whose spacetime trajectory is timelike, and later comes back to ambient space in the form of Hawking radiation. One may thus wonder if the causal structure of the entire spacetime can in fact be trivial in the exterior description (or at least in some version of the exterior description). We depict how such a description would look in the right panel of Fig. 7.

[17] The relation between the degrees of freedom representing the interior and Hawking radiation always involves the horizon dynamics, which we assume to be pseudorandom. The analysis of KTP can thus be adopted in the present context.

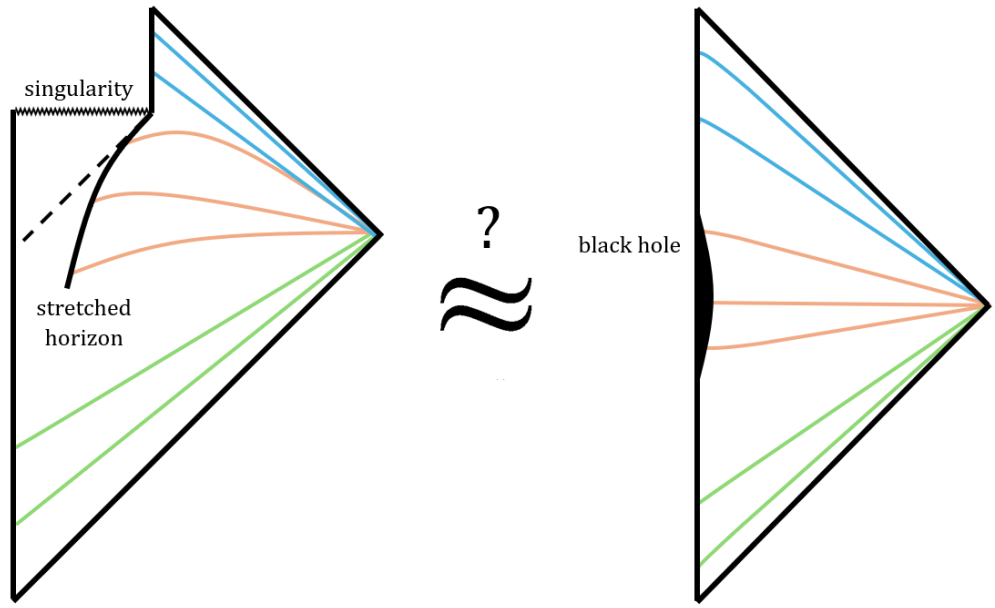

Figure 7: In the exterior description, the causal flow of information is obvious after the black hole is formed. The question is if there exists an analogous picture throughout the history of the black hole evolution, including the times during its formation.

This would indeed be the case if the "stretched horizon," at which the spacetime terminates, follows a timelike trajectory *throughout* the history of the black hole evolution, including during formation. Taking this possibility seriously, however, raises an interesting question. To see this, imagine that a diary has existed at $r = 0$ for an extended period of time in near-empty spacetime. Then, we fire a null shell of particles toward $r = 0$, creating a black hole of mass M; see Fig. 8. Before the null ray, the spacetime is Minkowski. After the null ray, the spacetime is Schwarzschild. Given the unitarity of the S-matrix, the information about the diary must eventually be emitted from the stretched horizon in the form of Hawking radiation. In order to avoid a mysterious spacelike motion of the information from the interior to the exterior, the timelike trajectory of the stretched horizon must extend to the spacetime region before the null ray, as depicted by a curved dashed line in Fig. 8. However, the spacetime in this region is Minkowski, and there is no obvious locally defined surface which can play the role of the stretched horizon.

5.1 Exterior description with trivial causal structure

We argue that the description corresponding to the right panel of Fig. 7, however, is still possible. In particular, we will provide below a surface which plays the role of the stretched horizon before the null shell, which in the case of Fig. 8 is the appropriately stretched horizon of a spherical Rindler observer. Before the black hole's formation, the information about the diary stays in modes near this horizon, and it will eventually come out in a causal way in the form of Hawking radiation. This is possible because

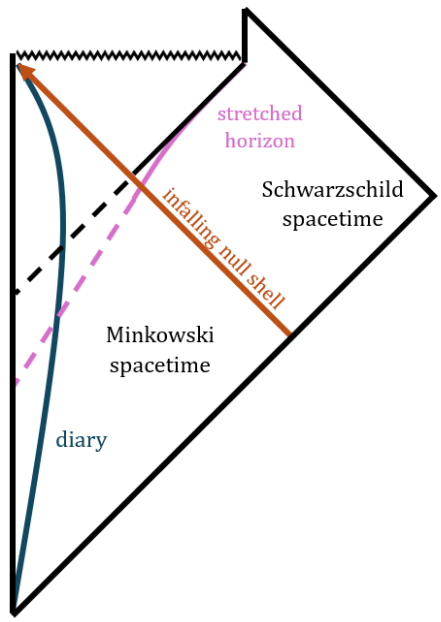

Figure 8: Consider a diary in the region near $r = 0$ far in the past. We then fire a null shell of infalling particles toward $r = 0$, creating a black hole of mass M. The unitarity of the S-matrix implies that the information about the diary must eventually be emitted from the stretched horizon in the form of Hawking radiation. An exterior description with trivial causality, the right panel of Fig. 7, is obtained by extending the stretched horizon toward the past using the event horizon.

in quantum gravity the spatial location of information is not uniquely determined.

In quantum gravity, the physics associated with some spacetime region can be encoded in the degrees of freedom on the boundary surrounding it, if certain conditions are met [14, 15, 81]. In particular, for a closed codimension-2 surface Σ, the physics occurring in the domain of dependence of a partial Cauchy surface bounded by Σ can fully be described using the intrinsically quantum gravitational degrees of freedom on Σ and the degrees of freedom outside of it.

We apply this to each Cauchy surface before the formation of the black hole, with Σ identified as the (stretched) *event horizon*. Specifically, in the setup of Fig. 8, we may adopt the ingoing Vaidya metric

$$ds^2 = -\left(1 - \frac{2M(v)}{r}\right) dv^2 + 2dvdr + r^2 d\Omega^2, \tag{5.1}$$

where

$$M(v) = \begin{cases} 0 & \text{for } v < v_0 \\ M & \text{for } v > v_0 \end{cases}, \tag{5.2}$$

with v_0 the ingoing time at which the null shell is shot. For $v > v_0$, this corresponds to time slices of the exterior description, as can be seen from the fact that the coordinate transformation

$$v = t + r + 2Ml_{\mathrm{P}}^2 \ln \frac{r - 2Ml_{\mathrm{P}}^2}{2Ml_{\mathrm{P}}^2} \tag{5.3}$$

gives the Schwarzschild metric. For $v < v_0$, on the other hand, the metric also covers a region inside the event horizon: by the coordinate transformation

$$v = t + r, \tag{5.4}$$

we obtain the flat space metric

$$ds^2 = -dt^2 + dr^2 + r^2\Omega^2. \tag{5.5}$$

The trajectory of the event horizon in this metric is given by

$$r = t + 4Ml_{\rm P}^2 - v_0 \tag{5.6}$$

for $v_0 - 4Ml_{\rm P}^2 \le t \le v_0 - 2Ml_{\rm P}^2$.

We may convert the metric (5.5) into that corresponding to the "exterior picture," by adopting the spherical Rindler coordinates [82] tailored to the present setup

$$\tau = \operatorname{arctanh}\frac{t - t_*}{r}, \qquad \rho = \sqrt{r^2 - (t - t_*)^2}, \tag{5.7}$$

where

$$t_* = v_0 - 4Ml_{\rm P}^2 \tag{5.8}$$

is the time at which the event horizon forms. This gives the metric

$$ds^2 = -\rho^2 d\tau^2 + d\rho^2 + \rho^2 \cosh^2\tau \, d\Omega^2, \tag{5.9}$$

applicable in the region $v < v_0$, i.e. $\rho < 4Ml_{\rm P}^2 e^{-\tau}$, and $\tau \ge 0$. The Rindler horizon agrees with the event horizon (5.6), which connects smoothly with the black hole horizon at $v = v_0$. While the system is not in the global thermal equilibrium as the metric (5.9) is not static, observers at constant ρ follow trajectories of constant acceleration and see Unruh radiation [83–85]. The stretched horizon can thus be defined to be located where the local (Tolman) Unruh temperature $1/2\pi\rho$ becomes the string scale $1/2\pi l_{\rm s}$ [30, 86]:

$$\rho = \rho_{\rm s} \approx l_{\rm s}. \tag{5.10}$$

This surface connects smoothly with the stretched horizon of the black hole at $v = v_0$. At the time of the formation, the area of the stretched horizon is of the cutoff-scale size, $\mathcal{A}(0) \approx 4\pi l_{\rm s}^2$. The nonstretched and stretched horizons are depicted by dashed lines in Fig. 8.

We claim that when the theory is quantized canonically, there is a description which keeps only the portion of the spacetime outside the stretched horizon, $\rho > \rho_{\rm s}$, and in which the stretched horizon carries the coarse-grained entropy (the number of effective degrees of freedom)

$$S_{\rm h}(\tau) = \frac{\mathcal{A}(\tau)}{4l_{\rm P}^2}, \tag{5.11}$$

where

$$\mathcal{A}(\tau) = 4\pi\rho_{\rm s}^2 \cosh^2\tau \tag{5.12}$$

is the area of the stretched horizon at time τ (≥ 0).[18] When the stretched horizon intersects the null shell at

$$\tau = \ln \frac{v_0 - t_*}{\rho_s} = \ln \frac{4Ml_P^2}{\rho_s}, \tag{5.13}$$

$S_h(\tau)$ and $\mathcal{A}(\tau)$ agree with the entropy and the area of the stretched horizon of the black hole of mass M, respectively. We postulate that, as suggested by the holographic principle, the degrees of freedom associated with the stretched horizon, represented by (5.11), encode semiclassical physics in the domain of dependence of a partial Cauchy surface bounded by the stretched horizon at time τ

$$D(\Xi), \qquad \Xi : r < \rho_s \cosh\tau,\ t = t_* + \rho_s \sinh\tau. \tag{5.14}$$

Thus, when the diary hits this past-extended stretched horizon, its information is stored in these degrees of freedom. This information will then be transferred to the black hole horizon degrees of freedom when the stretched horizon is hit by the incoming null shell.

To develop intuition about our picture, we can employ a tensor network model [87–89]. In a tensor network, a semiclassical excitation, in this case the diary, can be regarded as a perturbation on the network. We can insert this perturbation on a specific location in the network, but the same information can then be "pushed" through the network and be viewed as existing in external legs of the network, which typically are identified as the boundary degrees of freedom in holography. The location of this information, however, need not be either the original location of the perturbation or the external legs of the network. In particular, we can cut the network into two at a closed surface Σ which encloses the original location of the perturbation; see Fig. 9. The information about the perturbation can then be viewed as existing on legs that go through Σ. In fact, if all the information about the inner portion of the network can be represented in this way on the cut legs of the outer portion, then we need not even keep the inner portion to describe the system.

Our description conforms with the recent proposal of Bousso and Penington [26, 67] which extends the concept of entanglement wedge reconstruction to bulk subregions. According to their criterion, the max-entanglement wedge $e_{\max}$ [26] of the partial Cauchy surface outside the extended stretched horizon at any time is the entire spacetime. This is consistent with our claim, given that entanglement wedge reconstruction assumes the full knowledge about the (exterior) time evolution, so that it can reconstruct the region $D(\Xi)$ (5.14) for all t.

Our description may also be obtained by "pulling in" the boundary of the (putative) holographic theory [44, 45, 90]. Imagine that there is a "screen" in the asymptotic region $r \to \infty$ on which the holographic theory lives. (In asymptotically AdS spacetime, this can be the conformal boundary.) Each time slice of the screen is called a leaf. We can then renormalize the leaf at each time, which moves the leaf inward in the bulk. As discussed in [45], when doing this at the quantum level, we need to keep the information outside the renormalized leaf $\sigma(\lambda)$, where λ is the flow parameter. Specifically, the flow is given by the equation

$$\frac{dx^\mu}{d\lambda} = \frac{1}{2}(\Theta_k l^\mu + \Theta_l k^\mu), \tag{5.15}$$

[18] The calculation of [82] gives the von Neumann entropy, which is zero in the present setup.

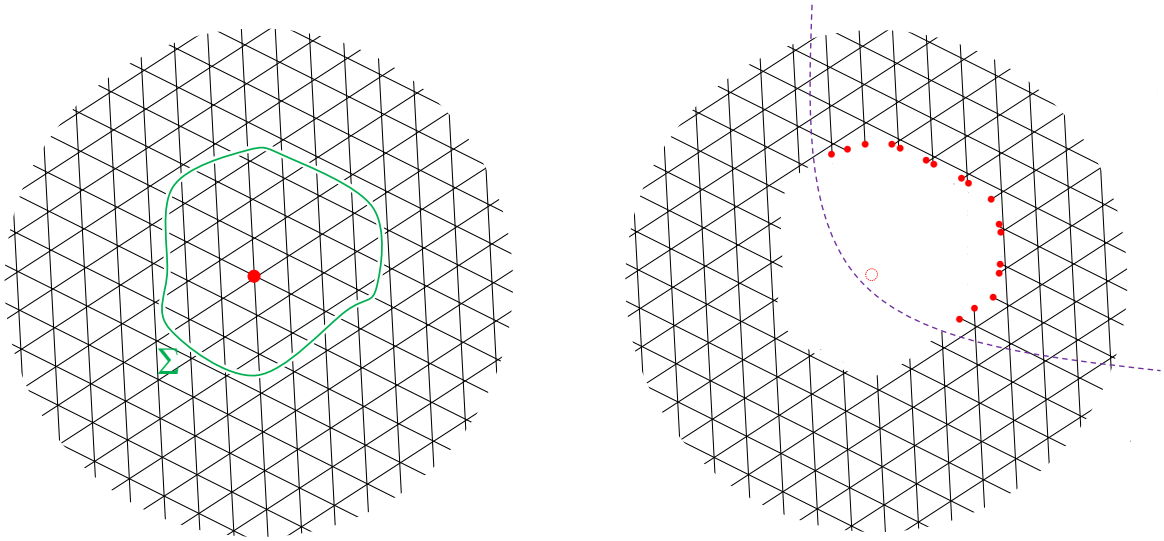

Figure 9: An excitation in a semiclassical theory can be viewed as a perturbation on the network of entanglement representing a spatial slice (left). The network can be cut at a closed surface Σ surrounding the perturbation and be replaced with an equivalent network in which the interior of Σ is removed (right). An operator corresponding to the original excitation can be represented by a nonlocal operator acting on the internal legs of the new network, obtained by pushing the original operator through the network.

where x^μ are the embedding coordinates of the renormalized leaf σ, and $\{k^\mu, l^\mu\}$ are the future-directed null vectors orthogonal to σ. Here, $\Theta_{k,l}$ are the modified versions [45] of the quantum expansions [91] in the corresponding directions, whose computation requires knowledge about the state outside σ. In the Schwarzschild region the flow terminates near the horizon, but in the Minkowski region it can go all the way toward $r = 0$. Our description can, thus, be viewed as the holographic theory in which each leaf is renormalized down to the extended stretched horizon $r = r_{\rm s}(t)$ in the Minkowski region. The region $r < r_{\rm s}(t)$ is then obviously encoded in the renormalized theory defined in the region $r \geq r_{\rm s}(t)$. For a sketch of the trajectory of the renormalized leaves, see Fig. 10.

We expect that the exterior description presented here can be generalized to the case of arbitrary (not necessarily spherically symmetric) collapsing matter. The basic idea is to stretch the event horizon in such a way that the system of the stretched and nonstretched horizons appears locally Rindler. One possible procedure to extend the stretch horizon to earlier times is as follows: consider a congruence of past-directed timelike curves $x^\mu_p(\tau)$, each of which starts from a point p on the stretched horizon at a late time t_0 in the direction orthogonal to the equal-time hypersurface $t = t_0$. Here, τ is the proper time along each curve, with $\tau = 0$ on the stretched horizon. We require that each curve maintains proper acceleration a^μ of

$$\sqrt{a^\mu a_\mu} \approx \frac{1}{l_{\rm s}} \tag{5.16}$$

in the direction orthogonal to the hypersurface given by the spacetime trajectory of the congruence. An equal-time hypersurface in the exterior spacetime region can be chosen in such a way that it agrees

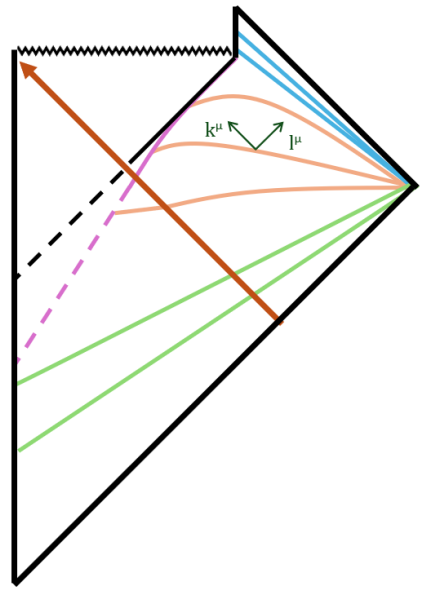

Figure 10: An equal-time hypersurface of the holographic boundary, called a leaf, can be pulled inward by a renormalization flow. The trajectory of each leaf, depicted in the figure, then provides an equal-time hypersurface in the bulk called a holographic slice in [44, 45]. We can terminate the flow when the renormalized leaf reaches the stretched event horizon, giving the manifestly causal exterior description.

with a constant τ surface on the congruence. If caustics develop in the congruence, we terminate the corresponding curves there. We expect that a procedure like this works, at least, for simple spacetimes, but we leave a detailed study of the stretching of the event horizon for future work.

5.2 Significance of the event horizon

The discerning reader may note that the argument presented in Section 5.1 relies primarily on the holographic principle, and may be applied to any timelike surface outside the event horizon, not just the stretched event horizon. For example, one might consider in the setup of Fig. 8 an extension of the stretched horizon defined by some time-dependent radius $r_{\rm t}(t)$, whose spacetime trajectory is timelike and which agrees with the black hole horizon at the infalling null shell, such as

$$\frac{dr_{\rm t}(t)}{dt} > 0, \qquad r_{\rm t}(t_{\rm null}) = r_{\rm s}. \tag{5.17}$$

Here, $t_{\rm null}$ is the time at which the trajectory $r_{\rm t}(t)$ of the surface intersects the null shell, and $r_{\rm s}$ is the radius of the stretched horizon of the black hole of mass M. We call this trajectory the "tube." One can imagine that the tube is capped off in the past at some time t_0: $r_{\rm t}(t_0) = 0$.

Such a description may indeed exist, but the dynamics on the tube would then be very different from the dynamics of the horizon. To see this, consider a trajectory of a particle which enters and then leaves the tube in the global spacetime picture. In the exterior picture, this should be described as a process in which a particle hits the surface of the tube at some time and then some time later the

same particle is spit out from a different point on the surface. This implies that, unlike the case of the stretched horizon, the dynamics of the surface are not scrambling and chaotic. We do, however, expect that these properties are associated with the (stretched) null surface, since a particle inside the surface becomes causally disconnected from the exterior.

What is the significance, then, of the event horizon compared with other null surfaces? For these other surfaces, such as the horizons of the standard and spherical Rindler spaces, the area of the horizon is either infinite or increases without bound. We thus expect that information absorbed into the corresponding stretched horizon will not be sent back to the ambient space, since the horizon degrees of freedom will not be fully scrambled. On the other hand, for an event horizon, its stretched surface becomes the stretched horizon of the black hole, whose area will eventually shrink to zero due to Hawking radiation. Any information contained within the stretched event horizon, therefore, will be transferred to ambient space at late times, either in the form of Hawking radiation or through mining.

The expectation given above is consistent with the fact that neither standard nor spherical Rindler space contains a singularity. In these spacetimes, an "exterior" observer cannot retrieve information about an object which crossed the horizon at an earlier time. Thus, there is no problem of cloning even if the observer jumps into the horizon and catches up with the original object, which is generally possible if there is no singularity. On the other hand, for an event horizon, an exterior observer at late times *can* retrieve information behind the horizon. Therefore, an observer must not be able to jump into the horizon and obtain a second copy of the information they have retrieved. This requires that every causal curve in the interior region ends at a singularity, as is indeed the case for a Schwarzschild black hole (and presumably for other unstable black holes as well, after appropriate quantum effects are taken into account).

5.3 Topology change and quantum global hyperbolicity

In the exterior description given here, the flow of information is causally trivial. This description, however, does not include the region inside the black hole. This requires that the spatial topology changes as time progresses. For example, if the black hole is formed in $(d+1)$-dimensional asymptotically flat spacetime, the spatial topologies in the periods before the formation of the stretched event horizon, during the presence of the stretched (event and black hole) horizon, and after the black hole evaporation are $\mathbb{R}^d$, $\mathbb{R}^d \setminus B^d$, and $\mathbb{R}^d$, respectively. These spacetimes are not globally hyperbolic classically. This is due to the singularity at the endpoint of black hole evaporation, and the fact that equal-time hypersurfaces do not enter into the interior of the stretched horizon.

The unitarity of time evolution in the exterior description, however, implies that the spacetime is "globally hyperbolic" at the fully quantum level, i.e. if we include intrinsically quantum gravitational degrees of freedom on the horizon and use the microscopic Hamiltonian for time evolution. In this sense, equal-time hypersurfaces in the exterior description play the role of Cauchy surfaces in quantum gravity, and may be called quantum Cauchy surfaces.

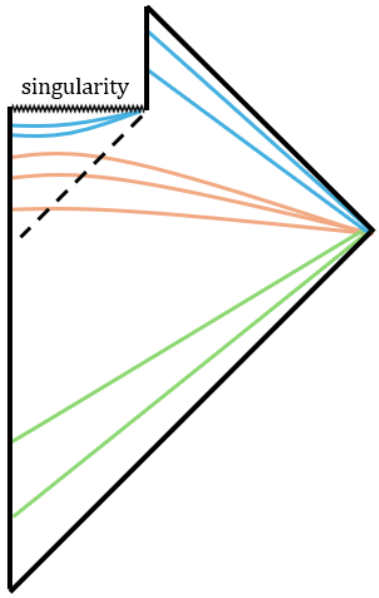

Figure 11: The global spacetime description can be obtained by evolving the initial equal-time hypersurface in such a way that at each time, the equal-time hypersurface is a Cauchy slice. This description, however, is consistent only if we disallow a process involving an exponentially fine-tuned detector configuration accessing information about the interior; see Section 4.

6 Global Spacetime as a Coarse-Grained Description

We have argued that the global spacetime is only emergent at the level of semiclassical gravity and cannot be used to formulate unitary quantum mechanics at the fundamental level.[19] On the other hand, the concept of global spacetime itself is well defined in the sense that the physics occurring there is robust under physical processes, including nonperturbative information extraction processes performed on Hawking radiation. In this section, we discuss how the description based on the global description can emerge as a coarse-grained description, and explain why global methods nevertheless yield important results, including hints of unitarity.

The global spacetime description can be obtained at the semiclassical level by evolving the initial equal-time hypersurface in such a way that at each time, the equal-time hypersurface is a Cauchy slice; see Fig. 11.[20] The famous calculation of Hawking [92] suggests that this corresponds to one which is coarse-grained over black hole microstates. This results in the final state of Hawking radiation being thermal. In other words, the global spacetime description regards states after the black hole formation

[19] In the context of AdS/CFT, this implies that any putative nonperturbative bulk theory should not be formulated on a global slice (or should not include a smooth interior), unless an intricate nonlocal mechanism reducing the redundancies on an equal-time hypersurface is involved.

[20] Strictly speaking, equal-time hypersurfaces after black hole evaporation are not Cauchy surfaces at the classical level because of the singularity at the endpoint of the evaporation; see the discussion in Section 5.3.

as the (maximal) mixture over black hole microstates [30]. We will show that this interpretation is consistent with results obtained using replica wormholes [8, 9], entanglement islands [5–7], and baby universes [39,40], which see some traits of unitarity despite the fact that they employ the global spacetime description.

6.1 The Page curve from replica wormholes and entanglement islands

While the global spacetime description represents an ensemble average over microstates, it can still allow us to calculate the ensemble averages of many different quantities by changing the boundary conditions.[21] In particular, by adopting the replica boundary conditions [8,9], gravitational path integrals in the global spacetime can calculate the ensemble averages of the traces of powers of the density matrix of Hawking radiation $\overline{\mathrm{Tr}\,\rho_R^n}$ for different n. From this, the ensemble average of the von Neumann entropy of the radiation can be obtained as

$$\overline{S_R} = -\lim_{n\to 1}\frac{\partial}{\partial n}\overline{\mathrm{Tr}\,\rho_R^n} \sim \min\{S_{\mathrm{rad}}, S_{\mathrm{bh}}\}, \tag{6.1}$$

where S_{rad} and S_{bh} are the coarse-grained entropies of the radiation and the black hole. Since this quantity is the ensemble average of microscopic von Neumann entropies, which obey the Page curve [43] for all members of the ensemble, it also obeys the Page curve [30, 95–101]. This is how the replica wormhole calculation in [8,9] yields the Page curve, despite the fact that it employs the global spacetime description, which is not unitary.[22]

The replica prescription for the gravitational path integral is virtually equivalent to the entanglement island prescription [5–7] for calculating entropies. Thus, the success of the calculation in [5–7], which adopts the global spacetime description, can also be interpreted in the same way. Note that because of the ensemble averaging, Hawking radiation in a single copy of spacetime does not contain information about the microstate by itself. This is why the semiclassical entropy must be used for Hawking radiation when applying the island formula (in a single copy of the global spacetime).

6.2 Baby universe α-states as fine-grained information in Hawking radiation

The transition from connected to disconnected Cauchy surfaces depicted in Fig. 11 can be viewed as the creation of a baby universe [36–38, 40], which will be separate from the ambient space at late times. This picture was studied by Marolf and Maxfield in [39]. Assuming that the singularity at the end of the black hole evaporation is mild, we may envision the process as in Fig. 12, in which a closed universe corresponding to the black hole interior is separated from the ambient space describing the exterior of the black hole.

[21] In lower dimensional gravity, the relevant ensemble can consist of different microscopic boundary theories [93, 94].

[22] To recover the full unitary content of the theory, one would calculate the ensemble averages of all of the moments, $\overline{\mathrm{Tr}\,\rho_R^n}$ for $1 \le n \le e^{S_{\mathrm{rad}}}$. Such a calculation, however, would not be able to be performed within the framework of semiclassical gravity.

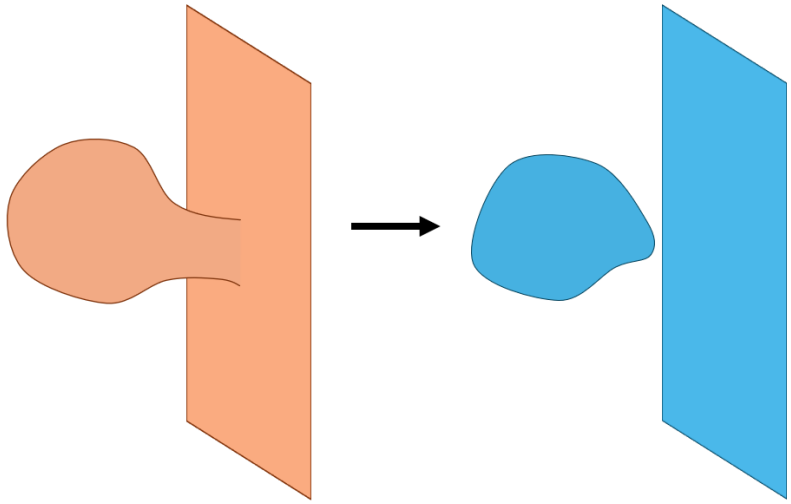

Figure 12: Assuming that the singularity at the endpoint of black hole evaporation is sufficiently weak, the evaporation of a black hole can be viewed as the creation of a baby universe in (but only in) the effective global spacetime description.

Marolf's and Maxfield's interpretation is that after the black hole evaporates, its microscopic information is encoded in the α-states of the baby universes arising in the system of replicated black holes. For a single black hole, different α sectors comprise superselection sectors, and hence its information cannot be accessed in a single experiment involving a single black hole. Our analysis implies that this is an artifact of the global spacetime description, in which the Hawking radiation involves the ensemble average over the microstates. In the full exterior description, the microscopic information *can* be accessed even in a single experiment by an appropriately fine-tuned detector, as depicted in Fig. 4. In other words, α sectors in the global spacetime description are not superselection sectors in the exterior description, which is unitary and hence, unlike [39], gives a pure final state.

This interpretation is consistent with the conjecture [102, 103] that in a complete theory of quantum gravity, the Hilbert space of baby universes is trivial

$$\dim \mathcal{H}_{\mathrm{BU}} = 1. \tag{6.2}$$

A nontrivial Hilbert space for baby universe states emerges only in an effective description in which certain microscopic information in the fundamental description is coarse-grained.

We expect that this conclusion applies more generally. For example, we may imagine that states representing a semiclassical closed universe can be embedded, possibly like bound states [104, 105], in a theory accommodating an asymptotic region (such as null infinities of flat space and the conformal boundary of AdS space). Different α sectors associated with the semiclassical closed universe would then correspond to different microstates in the putative boundary theory in which this universe is embedded.

7 Conclusion

In this paper we have described a consistent view of black hole complementarity. In this picture, the fundamental description of the spacetime ends at the stretched horizon. A semiclassical picture of spacetime including the interior can be obtained only effectively, by the mechanism described in Section 3. As discussed there, the interior emerges as a way of reorganizing the degrees of freedom participating in the transient phenomenon of an infalling object dissipating into the stretched horizon. We argued using this framework that no operation performed on Hawking radiation, however complex, can affect the experience of an infalling observer. Rather, reconstructions of interior operators from Hawking radiation should be interpreted as acting on a historical record of the effective interior, which has causally propagated from the vicinity of the stretched horizon (and early Hawking radiation). While in the effective global picture the interior operators appear spacelike separated from the exterior operators used for reconstruction, in the fundamental picture they are timelike related. We explained this phenomenon as a consequence of the chaotic dynamics of the horizon modes.

It is natural to ask whether this construction can be extended to describe a black hole formed from collapse. To do so, it would be necessary to extend the stretched horizon before the formation of the black hole. In Section 5, we discussed a scheme for doing this, in which the extended stretched horizon is defined as a fixed radius surface in spherical Rindler coordinates, where the Unruh temperature reaches the string scale. We speculate, by analogy to the black hole case, that degrees of freedom in the interior might be constructed using the Unruh modes near this extended stretched horizon. We leave exploration of this and extension to spacetimes without spherical symmetry for future work.

Acknowledgments

We thank Masahiro Hotta and Tomonori Ugajin for useful discussions. This work was supported in part by the Department of Energy, Office of Science, Office of High Energy Physics under QuantISED award DE-SC0019380 and contract DE-AC02-05CH11231 and in part by MEXT KAKENHI grant number JP20H05850, JP20H05860.

References

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

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

[3] S. W. Hawking, "Particle creation by black holes," Commun. Math. Phys. **43**, 199 (1975) [Erratum: Commun. Math. Phys. **46**, 206 (1976)].

[4] L. Susskind, L. Thorlacius and J. Uglum, "The stretched horizon and black hole complementarity," Phys. Rev. D **48**, 3743 (1993) [arXiv:hep-th/9306069].

[5] G. Penington, "Entanglement wedge reconstruction and the information paradox," JHEP **09**, 002 (2020) [arXiv:1905.08255 [hep-th]].

[6] 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**, 063 (2019) [arXiv:1905.08762 [hep-th]].

[7] A. Almheiri, R. Mahajan, J. Maldacena and Y. Zhao, "The Page curve of Hawking radiation from semiclassical geometry," JHEP **03**, 149 (2020) [arXiv:1908.10996 [hep-th]].

[8] G. Penington, S. H. Shenker, D. Stanford and Z. Yang, "Replica wormholes and the black hole interior," JHEP **03**, 205 (2022) [arXiv:1911.11977 [hep-th]].

[9] A. Almheiri, T. Hartman, J. Maldacena, E. Shaghoulian and A. Tajdini, "Replica wormholes and the entropy of Hawking radiation," JHEP **05**, 013 (2020) [arXiv:1911.12333 [hep-th]].

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

[11] V. E. Hubeny, M. Rangamani and T. Takayanagi, "A covariant holographic entanglement entropy proposal," JHEP **07**, 062 (2007) [arXiv:0705.0016 [hep-th]].

[12] T. Faulkner, A. Lewkowycz and J. Maldacena, "Quantum corrections to holographic entanglement entropy," JHEP **11**, 074 (2013) [arXiv:1307.2892 [hep-th]].

[13] N. Engelhardt and A. C. Wall, "Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime," JHEP **01**, 073 (2015) [arXiv:1408.3203 [hep-th]].

[14] G. 't Hooft, "Dimensional reduction in quantum gravity," in *Salamfestschrift*, edited by A. Ali, J. Ellis, and S. Randjbar-Daemi (World Scientific, Singapore, 1994), p. 284 [arXiv:gr-qc/9310026].

[15] L. Susskind, "The world as a hologram," J. Math. Phys. **36**, 6377 (1995) [arXiv:hep-th/9409089].

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

[17] B. Czech, J. L. Karczmarek, F. Nogueira and M. Van Raamsdonk, "The gravity dual of a density matrix," Class. Quant. Grav. **29**, 155009 (2012) [arXiv:1204.1330 [hep-th]].

[18] A. C. Wall, "Maximin surfaces, and the strong subadditivity of the covariant holographic entanglement entropy," Class. Quant. Grav. **31**, 225007 (2014) [arXiv:1211.3494 [hep-th]].

[19] M. Headrick, V. E. Hubeny, A. Lawrence and M. Rangamani, "Causality & holographic entanglement entropy," JHEP **12**, 162 (2014) [arXiv:1408.6300 [hep-th]].

[20] D. L. Jafferis, A. Lewkowycz, J. Maldacena and S. J. Suh, "Relative entropy equals bulk relative entropy," JHEP **06**, 004 (2016) [arXiv:1512.06431 [hep-th]].

[21] X. Dong, D. Harlow and A. C. Wall, "Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality," Phys. Rev. Lett. **117**, 021601 (2016) [arXiv:1601.05416 [hep-th]].

[22] T. Faulkner and A. Lewkowycz, "Bulk locality from modular flow," JHEP **07**, 151 (2017) [arXiv:1704.05464 [hep-th]].

[23] J. Cotler, P. Hayden, G. Penington, G. Salton, B. Swingle and M. Walter, "Entanglement wedge reconstruction via universal recovery channels," Phys. Rev. X **9**, 031011 (2019) [arXiv:1704.05839 [hep-th]].

[24] C. Akers and G. Penington, "Leading order corrections to the quantum extremal surface prescription," JHEP **04**, 062 (2021) [arXiv:2008.03319 [hep-th]].

[25] I. Kim, E. Tang and J. Preskill, "The ghost in the radiation: robust encodings of the black hole interior," JHEP **06**, 031 (2020) [arXiv:2003.05451 [hep-th]].

[26] R. Bousso and G. Penington, "Holograms in our world," Phys. Rev. D **108**, 046007 (2023) [arXiv:2302.07892 [hep-th]].

[27] Y. Nomura, "Reanalyzing an evaporating black hole," Phys. Rev. D **99**, 086004 (2019) [arXiv:1810.09453 [hep-th]].

[28] Y. Nomura, "Spacetime and universal soft modes — black holes and beyond," Phys. Rev. D **101**, 066024 (2020) [arXiv:1908.05728 [hep-th]].

[29] Y. Nomura, "The interior of a unitarily evaporating black hole," Phys. Rev. D **102**, 026001 (2020) [arXiv:1911.13120 [hep-th]].

[30] K. Langhoff and Y. Nomura, "Ensemble from coarse graining: reconstructing the interior of an evaporating black hole," Phys. Rev. D **102**, 086021 (2020) [arXiv:2008.04202 [hep-th]].

[31] Y. Nomura, "Black hole interior in unitary gauge construction," Phys. Rev. D **103**, 066011 (2021) [arXiv:2010.15827 [hep-th]].

[32] C. Murdia, Y. Nomura and K. Ritchie, "Black hole and de Sitter microstructures from a semiclassical perspective," Phys. Rev. D **107**, 026016 (2023) [arXiv:2207.01625 [hep-th]].

[33] W. G. Unruh and R. M. Wald, "Acceleration radiation and generalized second law of thermodynamics," Phys. Rev. D **25**, 942 (1982).

[34] A. R. Brown, "Tensile strength and the mining of black holes," Phys. Rev. Lett. **111**, 211301 (2013) [arXiv:1207.3342 [gr-qc]].

[35] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," Nature **299**, 802 (1982).

[36] S. Coleman, "Black holes as red herrings: topological fluctuations and the loss of quantum coherence," Nucl. Phys. B **307**, 867 (1988).

[37] S. B. Giddings and A. Strominger, "Loss of incoherence and determination of coupling constants in quantum gravity," Nucl. Phys. B **307**, 854 (1988).

[38] S. B. Giddings and A. Strominger, "Baby universes, third quantization and the cosmological constant," Nucl. Phys. B **321**, 481 (1989).

[39] D. Marolf and H. Maxfield, "Observations of Hawking radiation: the Page curve and baby universes," JHEP **04**, 272 (2021) [arXiv:2010.06602 [hep-th]].

[40] J. Polchinski and A. Strominger, "A possible resolution of the black hole information puzzle," Phys. Rev. D **50**, 7403 (1994) [arXiv:hep-th/9407008].

[41] G. 't Hooft, "On the quantum structure of a black hole," Nucl. Phys. B **256**, 727 (1985).

[42] G. 't Hooft, "The black hole interpretation of string theory," Nucl. Phys. B **335**, 138 (1990).

[43] D. N. Page, "Information in black hole radiation," Phys. Rev. Lett. **71**, 3743 (1993) [arXiv:hep-th/9306083].

[44] Y. Nomura, P. Rath and N. Salzetta, "Pulling the boundary into the bulk," Phys. Rev. D **98**, 026010 (2018) [arXiv:1805.00523 [hep-th]].

[45] C. Murdia, Y. Nomura and P. Rath, "Coarse-graining holographic states: a semiclassical flow in general spacetimes," Phys. Rev. D **102**, 086001 (2020) [arXiv:2008.01755 [hep-th]].

[46] O. Lunin and S. D. Mathur, "AdS/CFT duality and the black hole information paradox," Nucl. Phys. B **623**, 342 (2002) [arXiv:hep-th/0109154].

[47] O. Lunin, J. M. Maldacena and L. Maoz, "Gravity solutions for the D1-D5 system with angular momentum," arXiv:hep-th/0212210.

[48] S. D. Mathur, A. Saxena and Y. K. Srivastava, "Constructing 'hair' for the three charge hole," Nucl. Phys. B **680**, 415 (2004) [arXiv:hep-th/0311092].

[49] A. Almheiri, D. Marolf, J. Polchinski and J. Sully, "Black holes: complementarity or firewalls?," JHEP **02**, 062 (2013) [arXiv:1207.3123 [hep-th]].

[50] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford and J. Sully, "An apologia for firewalls," JHEP **09**, 018 (2013) [arXiv:1304.6483 [hep-th]].

[51] D. Marolf and J. Polchinski, "Gauge/gravity duality and the black hole interior," Phys. Rev. Lett. **111**, 171301 (2013) [arXiv:1307.4706 [hep-th]].

[52] P. Hayden and J. Preskill, "Black holes as mirrors: quantum information in random subsystems," JHEP **09**, 120 (2007) [arXiv:0708.4025 [hep-th]].

[53] Y. Sekino and L. Susskind, "Fast scramblers," JHEP **10**, 065 (2008) [arXiv:0808.2096 [hep-th]].

[54] J. Maldacena, S. H. Shenker and D. Stanford, "A bound on chaos," JHEP **08**, 106 (2016) [arXiv:1503.01409 [hep-th]].

[55] K. Papadodimas and S. Raju, "An infalling observer in AdS/CFT," JHEP **10**, 212 (2013) [arXiv:1211.6767 [hep-th]].

[56] E. Verlinde and H. Verlinde, "Black hole entanglement and quantum error correction," JHEP **10**, 107 (2013) [arXiv:1211.6913 [hep-th]].

[57] Y. Nomura and J. Varela, "A note on (no) firewalls: the entropy argument," JHEP **07**, 124 (2013) [arXiv:1211.7033 [hep-th]].

[58] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," Fortsch. Phys. **61**, 781 (2013) [arXiv:1306.0533 [hep-th]].

[59] K. Papadodimas and S. Raju, "State-dependent bulk-boundary maps and black hole complementarity," Phys. Rev. D **89**, 086010 (2014) [arXiv:1310.6335 [hep-th]].

[60] K. Papadodimas and S. Raju, "Remarks on the necessity and implications of state-dependence in the black hole interior," Phys. Rev. D **93**, 084049 (2016) [arXiv:1503.08825 [hep-th]].

[61] P. Hayden and G. Penington, "Learning the alpha-bits of black holes," JHEP **12**, 007 (2019) [arXiv:1807.06041 [hep-th]].

[62] S. D. Mathur, "The information paradox: a pedagogical introduction," Class. Quant. Grav. **26**, 224001 (2009) [arXiv:0909.1038 [hep-th]].

[63] R. Bousso, "Violations of the equivalence principle by a nonlocally reconstructed vacuum at the black hole horizon," Phys. Rev. Lett. **112**, 041102 (2014) [arXiv:1308.3697 [hep-th]].

[64] A. Almheiri, X. Dong and D. Harlow, "Bulk locality and quantum error correction in AdS/CFT," JHEP **04**, 163 (2015) [arXiv:1411.7041 [hep-th]].

[65] D. Harlow and P. Hayden, "Quantum computation vs. firewalls," JHEP **06**, 085 (2013) [arXiv:1301.4504 [hep-th]].

[66] A. R. Brown, H. Gharibyan, G. Penington and L. Susskind, "The python's lunch: geometric obstructions to decoding Hawking radiation," JHEP **08**, 121 (2020) [arXiv:1912.00228 [hep-th]].

[67] R. Bousso and G. Penington, "Entanglement wedges for gravitating regions," Phys. Rev. D **107**, 086002 (2023) [arXiv:2208.04993 [hep-th]].

[68] D. A. Lowe, J. Polchinski, L. Susskind, L. Thorlacius and J. Uglum, "Black hole complementarity versus locality," Phys. Rev. D **52**, 6997 (1995) [arXiv:hep-th/9506138].

[69] S. B. Giddings, "Black hole information, unitarity, and nonlocality," Phys. Rev. D **74**, 106005 (2006) [arXiv:hep-th/0605196].

[70] R. Bousso and A. Shahbazi-Moghaddam, "Quantum singularities," Phys. Rev. D **107**, 066002 (2023) [arXiv:2206.07001 [hep-th]].

[71] Y. Nomura, "Quantum mechanics, spacetime locality, and gravity," Found. Phys. **43**, 978 (2013) [arXiv:1110.4630 [hep-th]].

[72] H. Ollivier, D. Poulin and W. H. Zurek, "Objective properties from subjective quantum states: environment as a witness," Phys. Rev. Lett. **93**, 220401 (2004) [arXiv:quant-ph/0307229].

[73] R. Blume-Kohout and W. H. Zurek, "Quantum Darwinism: entanglement, branches, and the emergent classicality of redundantly stored quantum information," Phys. Rev. A **73**, 062310 (2006) [arXiv:quant-ph/0505031].

[74] R. Bousso and L. Susskind, "The multiverse interpretation of quantum mechanics," Phys. Rev. D **85**, 045007 (2012) [arXiv:1105.3796 [hep-th]].

[75] W. H. Zurek, "Entropy evaporated by a black hole," Phys. Rev. Lett. **49**, 1683 (1982).

[76] D. N. Page, "Comment on 'Entropy evaporated by a black hole'," Phys. Rev. Lett. **50**, 1013 (1983).

[77] D. N. Page, "Is black hole evaporation predictable?," Phys. Rev. Lett. **44**, 301 (1980).

[78] Y. Nomura, J. Varela and S. J. Weinberg, "Black holes, information, and Hilbert space for quantum gravity," Phys. Rev. D **87**, 084050 (2013) [arXiv:1210.6348 [hep-th]].

[79] A. Bouland, B. Fefferman and U. Vazirani, "Computational pseudorandomness, the wormhole growth paradox, and constraints on the AdS/CFT duality," arXiv:1910.14646 [quant-ph].

[80] N. Engelhardt, Å. Folkestad, A. Levine, E. Verheijden and L. Yang, "Cryptographic censorship," arXiv:2402.03425 [hep-th].

[81] R. Bousso, "The holographic principle," Rev. Mod. Phys. **74**, 825 (2002) [arXiv:hep-th/0203101].

[82] V. Balasubramanian, B. Czech, B. D. Chowdhury and J. de Boer, "The entropy of a hole in spacetime," JHEP **10**, 220 (2013) [arXiv:1305.0856 [hep-th]].

[83] W. G. Unruh, "Notes on black hole evaporation," Phys. Rev. D **14**, 870 (1976).

[84] S. A. Fulling, "Nonuniqueness of canonical field quantization in Riemannian space-time," Phys. Rev. D **7**, 2850 (1973).

[85] P. C. W. Davies, "Scalar particle production in Schwarzschild and Rindler metrics," J. Phys. A **8**, 609 (1975).

[86] M. Parikh and A. Svesko, "Einstein's equations from the stretched future light cone," Phys. Rev. D **98**, 026018 (2018) [arXiv:1712.08475 [hep-th]].

[87] B. Swingle, "Entanglement renormalization and holography," Phys. Rev. D **86**, 065007 (2012) [arXiv:0905.1317 [cond-mat.str-el]].

[88] F. Pastawski, B. Yoshida, D. Harlow and J. Preskill, "Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence," JHEP **06**, 149 (2015) [arXiv:1503.06237 [hep-th]].

[89] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter and Z. Yang, "Holographic duality from random tensor networks," JHEP **11**, 009 (2016) [arXiv:1601.01694 [hep-th]].

[90] M. Miyaji and T. Takayanagi, "Surface/state correspondence as a generalized holography," PTEP **2015**, 073B03 (2015) [arXiv:1503.03542 [hep-th]].

[91] R. Bousso, Z. Fisher, S. Leichenauer and A. C. Wall, "Quantum focusing conjecture," Phys. Rev. D **93**, 064044 (2016) [arXiv:1506.02669 [hep-th]].

[92] S. W. Hawking, "Breakdown of predictability in gravitational collapse," Phys. Rev. D **14**, 2460 (1976).

[93] P. Saad, S. H. Shenker and D. Stanford, "JT gravity as a matrix integral," arXiv:1903.11115 [hep-th].

[94] D. Stanford and E. Witten, "JT gravity and the ensembles of random matrix theory," Adv. Theor. Math. Phys. **24**, 1475 (2020) [arXiv:1907.03363 [hep-th]].

[95] R. Bousso and E. Wildenhain, "Gravity/ensemble duality," Phys. Rev. D **102**, 066005 (2020) [arXiv:2006.16289 [hep-th]].

[96] N. Engelhardt, S. Fischetti and A. Maloney, "Free energy from replica wormholes," Phys. Rev. D **103**, 046021 (2021) [arXiv:2007.07444 [hep-th]].

[97] R. Renner and J. Wang, "The black hole information puzzle and the quantum de Finetti theorem," arXiv:2110.14653 [hep-th].

[98] X.-L. Qi, Z. Shangnan and Z. Yang, "Holevo information and ensemble theory of gravity," JHEP **02**, 056 (2022) [arXiv:2111.05355 [hep-th]].

[99] A. Almheiri and H. W. Lin, "The entanglement wedge of unknown couplings," JHEP **08**, 062 (2022) [arXiv:2111.06298 [hep-th]].

[100] A. Blommaert, L. V. Iliesiu and J. Kruthoff, "Gravity factorized," JHEP **09**, 080 (2022) [arXiv:2111.07863 [hep-th]].

[101] J. Chandra and T. Hartman, "Coarse graining pure states in AdS/CFT," JHEP **10**, 030 (2023) [arXiv:2206.03414 [hep-th]].

[102] J. McNamara and C. Vafa, "Baby universes, holography, and the swampland," arXiv:2004.06738 [hep-th].

[103] D. Marolf and H. Maxfield, "Transcending the ensemble: baby universes, spacetime wormholes, and the order and disorder of black hole information," JHEP **08**, 044 (2020) [arXiv:2002.08950 [hep-th]].

[104] Y. Nomura, "The static quantum multiverse," Phys. Rev. D **86**, 083505 (2012) [arXiv:1205.5550 [hep-th]].

[105] J. Maltz and L. Susskind, "de Sitter space as a resonance," Phys. Rev. Lett. **118**, 101602 (2017) [arXiv:1611.00360 [hep-th]].