Smooth causal patches for AdS black holes

Suvrat Raju[*]

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

We review the paradox of low energy excitations of a black hole in anti–de Sitter space (AdS). An appropriately chosen unitary operator in the boundary theory can create a locally strong excitation near the black hole horizon, whose global energy is small as a result of the gravitational redshift. The paradox is that this seems to violate a general rule of statistical mechanics, which states that an operator with energy parametrically smaller than kT cannot create a significant excitation in a thermal system. When we carefully examine the position dependence of the boundary unitary operator that produces the excitation and the bulk observable necessary to detect the anomalously large effect, we find that they do not both fit in a single causal patch. This follows from a remarkable property of position-space AdS correlators that we establish explicitly and resolves the paradox in a generic state of the system, since no combination of observers can both create the excitation and observe its effect. As a special case of our analysis, we show how this resolves the "Born rule" paradox of Marolf and Polchinski [J. High Energy Phys. 01 (**2016**) 008] and we verify our solution using an independent calculation. We then consider boundary states that are finely tuned to display a spontaneous excitation outside the causal patch of the infalling observer, and we propose a version of causal patch complementarity in AdS/CFT that resolves the paradox for such states as well.

I. INTRODUCTION AND SUMMARY OF RESULTS

Does a typical pure state with high enough energy in a holographic conformal field theory (CFT) correspond to a classical black hole geometry with a smooth interior? This question can be made precise as follows. Consider a CFT with a holographic dual [1], and in such a CFT pick a typical state $|\Psi\rangle$ from the set of states with energy much larger than the central charge of the CFT. A typical state is a linear combination of energy eigenstates with coefficients that have a random magnitude and phase. In such a state, can we map the local field degrees of freedom at a bulk point x_i to some CFT operator $\phi(x_i)$? Moreover, if we compute multilocal correlators in the state $|\Psi\rangle$,

$$\langle\Psi|\phi(x_1)\ldots\phi(x_n)|\Psi\rangle, \qquad (1.1)$$

do these correlators match those that one would obtain by quantizing a bulk field about the classical black hole background?

If the points x_i are all outside the black hole horizon, then this question was answered constructively in [2,3], by finding an explicit set of operators in the CFT, $\phi(x_i)$, whose correlators in a typical state match bulk correlators about a black hole background. In [4,5], we also proposed a representation of the black hole interior for typical states in the CFT. These constructions, taken together, provide strong evidence that the answer to the question above is affirmative—typical high energy states of the CFT do correspond to a smooth black hole solution.

However, a novel feature of the construction of [4,5] was that operators inside the black hole were *state dependent*.

This feature was necessary because the authors of [6,7] pointed out that when some of the points x_i were behind the horizon, it was impossible to find state-independent operators, $\phi(x_i)$, whose correlators matched with the correlators of a bulk black hole geometry with a smooth interior. The authors of [6,7] suggested that typical pure states in the CFT have an exterior geometry that looks exactly like a classical black hole, but have no interior; state dependence provides an alternative that preserves a smooth interior.

A. The paradox of low energy excitations

However, state dependence does not automatically resolve all paradoxes associated with the black hole, and in [8], Marolf and Polchinski, building on previous work by Harlow [9], presented one such paradox. They constructed a unitary operator, U_{MP}, in the boundary theory that rotates the phases of the modes outside the black hole at a very small cost in energy. But since this phase rotation breaks the entanglement between the interior and the exterior, one might naively believe that when some points x_i are behind the horizon, and others are outside, the correlator (1.1) is modified sharply.

This is unusual because in a typical state of a general statistical system, the change in the value of an observable, A_α, under the action of a unitary excitation U, is bounded by

$$\delta\langle A_\alpha\rangle \equiv |\langle\Psi|U^\dagger A_\alpha U|\Psi\rangle - \langle\Psi|A_\alpha|\Psi\rangle| \leq 2\sqrt{\beta\delta E}\sigma_\alpha, \quad (1.2)$$

where δE is the change in the expectation value of the Hamiltonian induced by the unitary, and the inequality holds in the limit where $\beta\delta E \ll 1$. Here, σ_α is a measure of the

[*]suvrat@icts.res.in

fluctuations of the operator A_α. We will review and prove this inequality in Sec. II.

If we take A_α to be the correlator in (1.1) with some points behind the horizon and the unitary to be U_{MP} then it appears that this inequality is violated. This is a sign of what is a more general paradox in AdS/CFT. Consider the bulk dual to the state $U|\Psi\rangle$, where U is some unitary operator, and use the bulk equations of motion to analyze the geometry in the past and the future. Then, as a result of the gravitational redshift, it may happen that even though U only injects a small amount of global energy into the CFT state, the geometry obtained by the procedure above has a locally large excitation near the horizon. This is similar to the setup of [10] although one difference is that we are interested in excitations that inject an energy $\beta\delta E \ll 1$, whereas [10] considered excitations where $\beta\delta E \sim 1$. Then by considering observables along the worldline of an infalling observer it might appear that we could violate (1.2).

This paradox necessarily involves an infalling observer. Intuitively, this is because an observer, who stays outside the horizon and tries to detect the strong excitation, has to go close to the horizon and then return and so cannot differentiate the excitation from the Unruh effect that he experiences [11]. More formally, the bulk-boundary map involves only state-independent operators outside the horizon so all correlators evaluated purely outside the horizon, including the boundary correlators considered in [10], automatically obey the constraints of (1.2).

B. Plan of the resolution

In this paper, we address the paradox for the infalling observer as follows. First, we consider bulk observables that are confined to a single causal patch. We did not impose this restriction in our previous work [4,5,12,13] but it is a natural restriction because while a theory might predict some values for observables that do not fit in any causal patch, these observables cannot actually be measured by any combination of observers.

In *addition*, we divide the allowed excitations into two classes that we deal with separately.

(1) First in Sec. IV, we focus on the class of excitations that can be produced by deforming the CFT Hamiltonian at time t through

$$H \to H + J(t)A_\gamma^J(t), \qquad (1.3)$$

where $A_\gamma^J(t)$ is some simple operator on the boundary that is localized at the same time t.[1] This class

of excitations is more restrictive than the excitations previously considered in [4,5,12,13]. In those papers, while we insisted that the bulk observer could only perform experiments involving simple operators, we did *not* insist that the deformation of the Hamiltonian be local. For example, we did not previously exclude excitations of the form $H \to H + \delta(t) \int dx A_\gamma^J(x) \times \cos(\omega x)$. Such an excitation would involve acting with the *Fourier mode* of A_γ^J. However, physically, an observer at time t does not have access to boundary operators at a later (or earlier) time. In fact, it is not difficult to show that if an observer could act with excitations from a later time, she could engineer violations of the second law of thermodynamics. So we exclude excitations that are nonlocal in time in Sec. IV. With this restriction, we will show that *bulk correlators in a causal patch obey* (1.2) *for arbitrary excitations produced by sources of the form* (1.3).

(2) In Sec. VI, we allow more general excitations where the Hamiltonian at time t is deformed by operators that may be localized in the future of t. This returns us to the class of excitations considered in [4,5,12,13]. This analysis is closely related to the situation where the excitation is produced *spontaneously* in the theory, and the observer simply happens to find herself in a state that is about to undergo an excitation after some time. However, as explained above, this larger set of excitations and states is of limited physical interest: black holes with spontaneous macroscopic excitations occupy an *exponentially small* volume in the Hilbert space, which contains states that violate the second law of thermodynamics. Nevertheless, in Sec. VI, we consider this possibility and advance a tentative proposal of "causal patch complementarity" to ensure that (1.2) is preserved even in this situation.

C. Local sources and unitaries on the boundary of a causal patch

As explained above, the central thrust of the resolution advanced in this paper is that when the excitation is produced by a boundary source dual to local CFT operators, and when the observable in question is confined to a single causal patch, we do not observe any violation of (1.2). However, as we now describe, it is more convenient to frame this result in terms of unitary excitations on the boundary of the causal patch that contains the observable.

The correct Heisenberg operators at time t, which we should use upon adding a term $J(t)A_\gamma^J(t)$ to the Hamiltonian, are given by

$$\phi^J(t, r_*, \Omega) = \bar{\mathcal{T}}\{e^{i\int_\vartheta^t J(x)A_\gamma(x)dx}\}\phi(t, r_*, \Omega)$$
$$\times \mathcal{T}\{e^{-i\int_\vartheta^t J(x)A_\gamma(x)dx}\}, \qquad (1.4)$$

---

[1]Note that we only turn on sources that are dual to boundary operators. This does not involve any loss of generality since all gauge invariant operators in the bulk theory that can be added to the Hamiltonian are contained in the boundary. Indeed, as we discuss below, a bulk source would violate local conservation of energy. Energy can only be thrown in from the boundary and then, in the bulk, it can only be redistributed subject to the Wheeler-de Witt constraints.

where we have separately displayed the time, radial and spherical coordinates in the local field, and $\mathcal{T}$ indicates time ordering, whereas $\bar{\mathcal{T}}$ indicates anti-time ordering. We have assumed that the source has no support before a past cutoff ϑ. This formula follows from Schwinger's in-in formalism [14].

The operators (1.4) are somewhat inconvenient to work with directly. So, rather than working with the deformed Hamiltonian and the operators (1.4), we instead consider correlators of operators evolved using the *undeformed Hamiltonian* but sandwiched in a state excited by a unitary of the form

$$U = \exp\left[-i \int_{\vartheta}^{t_C} A_\alpha(t)dt\right]. \qquad (1.5)$$

This means that we consider correlators

$$\langle\Psi|U^\dagger\phi(x_1)...\phi(x_n)U|\Psi\rangle,$$

where we restrict the Kruskal "U_K" coordinate, corresponding to the points x_i, to $U_K < e^{-\frac{2\pi t_C}{\beta}}$. (We use conventions where $U_K < 0$ outside the horizon, and $U_K > 0$ inside the horizon.) An infalling observer who jumps in at the time t_C cannot reach any value of U_K larger than this limit, even if he travels at the speed of light.

Even though the problem above looks like it is different from the situation where we have turned on a source in the boundary, these two situations are closely related when the bulk observables fit in a single causal patch. We explain the logic briefly here and direct the reader to Sec. III C 1 for further details.

It is a simple consequence of bulk causality that when we turn on a source dual to some local operator on the boundary, only the part of the source that is in the causal past of a bulk point can affect the local field operator there. In particular, if we consider a correlator with insertions in a single causal patch, and the largest U_K value of the insertions is $e^{-\frac{2\pi t_C}{\beta}}$, then we can just ignore the part of the source that acts after time t_C.

This is why the class of observables obtained by considering correlators in a causal patch in the presence of an *arbitrary* simple local source—even one that continues for a long time in the future—is contained in the class of observables obtained by considering correlators sandwiched between a unitary of the form (1.5) and its adjoint that live on the boundary of the causal patch. The precise map between observables in the two possible descriptions is given in Sec. III C 1.

In the discussion below, we will often refer to the case where the "unitary and the observable both fit in the same causal patch." The reader should remember this contains the case where we turn on arbitrary local sources and examine a bulk observable that fits in some single causal patch.

D. Some technical details of the resolution

The central technical result that we prove in this paper is that given a unitary and a product of field operators, all of which fit in a single causal patch, we have

$$|\langle\Psi|U^\dagger\phi(x_1)...\phi(x_n)U|\Psi\rangle - \langle\Psi|\phi(x_1)...\phi(x_n)|\Psi\rangle|$$
$$\leq 2\sqrt{\beta\delta E}\sigma \qquad (1.6)$$

where δE is the energy injected by the unitary U and σ is precisely the measure of fluctuations in this correlator that appears in (1.2). This is a very nontrivial and surprising property of AdS correlation functions, and Sec. IV is devoted to establishing this result. We reiterate that by the argument above, (1.6) implies that correlators of the operators (1.4) restricted to a single causal patch also obey the constraints of (1.2).

The conclusion, therefore, is as follows: no combination of observers with finite powers can start with an equilibrium state, *create* an excitation U, and detect a correlator that violates (1.2). Since typical states of the form $|\Psi\rangle$ take up all but an exponentially small volume of the Hilbert space, and it is very reasonable to restrict to experiments where we turn on a local source and observe its effect, this already resolves the paradox in almost all cases of interest.

In the original formulation of the paradox in [8], this resolution is not evident because both the excitation and the observable to be measured are framed in frequency space. The fact that the excitation and the observable do not fit in a causal patch becomes apparent only when we examine these operators in position space.

E. A specific example

In Sec. V, we explicitly examine the specific unitary U_{MP} and a set of near-horizon correlators through an independent calculation that does not rely on the results of Sec. IV. We consider a position space version of U_{MP},

$$U_{\mathrm{MP}} = \exp\left[i \int_{\vartheta}^{t_C} dt_1 dt_2 \int d\Omega_1 d\Omega_2 \mathcal{O}(t_1, \Omega_1)\mathcal{O}(t_2, \Omega_2)\right.$$
$$\left. \times G(t_1, \Omega_1)G^*(t_2, \Omega_2)\right],$$

where $G(t, \Omega) \propto e^{i\omega_0 t}Y_\ell(\Omega)$ and $\mathcal{O}(t, \Omega)$ is a single trace primary operator on the boundary. Nevertheless, we show that if we restrict correlators in the state $U_{\mathrm{MP}}|\Psi\rangle$ to the causal patch that contains U_{MP}, which, in particular, implies that we consider fields localized at points so that $U_K < e^{-\frac{2\pi t_C}{\beta}}$ then such correlators do not differ significantly from correlators in the state $|\Psi\rangle$.

F. Some speculations on causal patch complementarity

In Sec. VI, we turn to a general class of excitations where the state is excited by an operator from the future before the bulk observer jumps in. Alternately, we can think of the situation where an observer jumps into a CFT state that is about to undergo a spontaneous excitation, but hits the singularity before the light ray from the excitation on the boundary can reach him.

Then it seems that, in principle, a superobserver with the ability to prepare arbitrary states in the CFT could compare the experience of the infaller in such states to the experience of the infaller in states where the excitation is absent altogether and observe a violation of (1.2).

However, the result of Sec. IV lead us to an interesting idea: it is possible to modify the standard map between bulk and boundary operators in such a way as to preserve (1.2) even in such experiments. This is the hypothesis of "causal patch complementarity."

Before we explain the mathematical details, a rough analogy may help to explain this idea. In quantum field theory, a correlation function that is classically well-defined requires the additional specification of a cutoff. The idea of causal patch complementarity is that bulk correlators in spontaneously excited states require additional data—about the perspective of the causal patch from which they are viewed—before they are fixed completely.

The precise proposal is as follows. For a large class of unitaries U that we describe more precisely in Sec. VI C, we suggest that the field operators, $\phi_\mathcal{C}$, that are appropriate for an infalling observer in such a state and in a particular causal patch satisfy

$$\langle\Psi|U^\dagger\phi_\mathcal{C}(x_1)...\phi_\mathcal{C}(x_n)U|\Psi\rangle = \langle\Psi|U^{\mathcal{C}\dagger}\phi(x_1)...\phi(x_n)U^\mathcal{C}|\Psi\rangle,$$
$$(1.7)$$

where on the right-hand side we have the usual field operators constructed using the standard bulk-boundary transfer function. The unitary $U^\mathcal{C}$ is an "attenuated" version of the unitary U. It has the property that it is localized on the intersection of the causal patch with the boundary (which we term $\mathcal{B}_\mathcal{C}$) and for any other simple operator $A_\alpha^\mathcal{C}$ that lives on $\mathcal{B}_\mathcal{C}$, we have $\langle\Psi|U^{\mathcal{C}\dagger}A_\alpha^\mathcal{C}U^\mathcal{C}|\Psi\rangle = \langle\Psi|U^\dagger A_\alpha^\mathcal{C}U|\Psi\rangle$; so the unitaries U and $U^\mathcal{C}$ are indistinguishable using operators only on $\mathcal{B}_\mathcal{C}$. We show that the description using the fields $\phi_\mathcal{C}$ is consistent within the causal patch $\mathcal{C}$ and also that it obeys (1.2). But it is not the same as the description that would be used by another observer who stays on the boundary long enough to see the excitation U. We suggest that these two descriptions are complementary. Our exploration of this proposal is not yet complete but we describe it in this paper, since it seems to point in some very interesting directions.

G. Plan of the Paper

This paper is structured as follows. In Sec. II, we derive the inequality (1.2) that shows that in a general statistical system, it is not possible to definitively excite the system with energy less than kT. Section II follows the ideas of [8] but is self-contained and may be considered a derivation of a simple result in statistical mechanics. In Sec. III, we describe our setup in holography and define our notation and various concepts that we use later. In Sec. IV, we prove that causal-patch correlators in anti–de Sitter space (AdS)

obey the constraints of statistical mechanics (1.2). Section IV is technical, but contains the central results of this paper. In Sec. V, we turn to a specific example to elucidate the calculations of Sec. IV. We examine the "Born rule" paradox of [8] and simply do a direct calculation, without using the results of Sec. IV, to show that this paradox disappears if we restrict the excitation and correlators to a single causal patch. Finally, in Sec. VI, we consider correlators in a causal patch but in a state that is excited with a unitary outside that patch; we explore the use of field operators that obey (1.7) to describe physics in a particular causal patch in this state, and we show that this ensures that these correlators are also stable, in that they are close to correlators in a state where the excitation is absent.

We emphasize, once again, that our calculations in Secs. IV and V are logically independent of the discussion in Sec. VI. The proof that no violation of (1.2) can be seen if we perform experiments where we turn on a local source in an equilibrium state and then perform bulk observations in a causal patch, which is the setting for any reasonable experiment, is an independent result that does not rely on our hypothesis of causal patch complementarity.

The recent literature on the reconstruction of the bulk from the boundary, and the associated paradoxes, starting with the work of Mathur [15] and AMPS [6], is rather extensive [16]. We reviewed some of these developments in detail in [12]. The application of causal patches to the information paradox has also been explored in [17,18], and causal patches are also relevant in the cosmological context [19].

A summary of the results of Sec. IV, with minimal technical details, is presented in [20], which the reader may find helpful.

II. LOW ENERGY EXCITATIONS IN A THERMAL SYSTEM

In this section, we will derive a bound on the effect of a low energy excitation in a thermal system. Our discussion is not specific to AdS/CFT, and this can be taken to be a simple and general result in statistical mechanics. This derivation is motivated by the analysis of [8].

We consider a statistical system with a large number of degrees of freedom. In such a system, we consider the set of energy eigenstates around some energy, E in the energy range $E \pm \Delta$, with $\Delta \ll E$, and we denote the Hilbert space spanned by these states by $\mathcal{H}_E$. In a generic statistical system, we expect that the dimension of this space is given by $\dim(\mathcal{H}_E) = e^S$ with $S \propto E$. We now consider a state $|\Psi\rangle$ that we pick using the Haar measure on this Hilbert space. Then we remind the reader that by using the eigenstate thermalization hypothesis [21], for any coarse-grained observable A_α, we have

$$\langle\Psi|A_\alpha|\Psi\rangle = \frac{1}{Z(\beta)}\text{Tr}(e^{-\beta H}A_\alpha) + \text{O}\left(\frac{1}{\sqrt{S}}\right), \quad (2.1)$$

where $\beta = \frac{\partial S}{\partial E}$ is the inverse of the effective temperature of the state and $Z(\beta)$ is the partition function at that temperature. So, for any coarse grained probe, the pure state $|\Psi\rangle$ is effectively thermal.

We now consider a low energy excitation of such a system. By this, we mean a unitary operator U that has the property that for a generic state $|\Psi\rangle$, we have

$$\langle\Psi|U^\dagger HU|\Psi\rangle - \langle\Psi|H|\Psi\rangle = \delta E. \qquad (2.2)$$

Note that a unitary cannot, on average, decrease the energy of a typical state and so—leaving aside very special unitaries that commute with the Hamiltonian—we have $\delta E > 0$. We are interested in the situation where $\delta E \ll \beta$, and also $\delta E \ll \Delta$; so $\frac{\delta E}{E} = O(\frac{1}{S})$.

We can also write (2.2) in the form

$$U\colon \mathcal{H}_E \to \mathcal{H}_{E+\delta E},$$

which just shows that the unitary maps us from the microcanonical ensemble at energy E into an ensemble at energy $E + \delta E$. The map U is an injective map. On the other hand, it is not surjective because the larger ensemble has dimension

$$\dim(\mathcal{H}_{E+\delta E}) = e^{S+\frac{\partial S}{\partial E}\delta E} = e^{S+\beta\delta E}.$$

The image of $\mathcal{H}_E$ under U forms a subspace of $\mathcal{H}_{E+\delta E}$, and we denote the projector onto this subspace by P_{U_E}. Now, let $|\Psi'\rangle$ be a typical state picked using the Haar measure on $\mathcal{H}_{E+\delta E}$. We note that

$$\langle\Psi'|P_{U_E}|\Psi'\rangle = 1 - \beta\delta E + O((\beta\delta E)^2),$$
$$\langle\Psi'|(1-P_{U_E})|\Psi'\rangle = \beta\delta E + O((\beta\delta E)^2).$$

For use below, we define

$$|\Psi_E\rangle = \langle\Psi'|P_{U_E}|\Psi'\rangle^{\frac{-1}{2}}P_{U_E}|\Psi'\rangle,$$
$$|\Psi_o\rangle = \langle\Psi'|1-P_{U_E}|\Psi'\rangle^{\frac{-1}{2}}(1-P_{U_E})|\Psi'\rangle,$$

so that we can decompose

$$|\Psi'\rangle = \left(1-\frac{\beta\delta E}{2}\right)|\Psi_E\rangle + (\beta\delta E)^{\frac{1}{2}}|\Psi_o\rangle + O((\beta\delta E)^{\frac{3}{2}}). \quad (2.3)$$

We also note that statistically, $|\Psi'\rangle$ is very similar to a typical state picked from $\mathcal{H}_E$. This is because the difference in the temperature of the two ensembles is given by

$$\delta\beta = \frac{\partial\beta}{\partial E}\delta E = -\beta^2\frac{\delta E}{C_V} = 0 + O\left(\frac{1}{S}\right), \qquad (2.4)$$

where C_V is the specific heat at constant volume, which is also macroscopic and scales like S. Then, using (2.1) and (2.4), we see that

$$\langle\Psi'|A_\alpha|\Psi'\rangle = \langle\Psi|A_\alpha|\Psi\rangle + O\left(\frac{1}{\sqrt{S}}\right).$$

This means that the expectation value of a coarse grained observable in a typical state from the larger ensemble is the same as the expectation value in the smaller ensemble.

On the other hand, we can also calculate this expectation value using the decomposition (2.3), which leads to

$$\langle\Psi'|A_\alpha|\Psi'\rangle - \langle\Psi_E|A_\alpha|\Psi_E\rangle$$
$$= \sqrt{\beta\delta E}(\langle\Psi_E|A_\alpha|\Psi_o\rangle + \langle\Psi_o|A_\alpha|\Psi_E\rangle) + O(\beta\delta E).$$

Since $|\Psi_E\rangle$ is a state obtained by acting with the unitary on a typical state of $\mathcal{H}_E$, the absolute value of the left-hand side can also be replaced by the change in the expectation value of A_α under the action of the unitary:

$$\delta\langle A_\alpha\rangle \equiv |\langle\Psi|U^\dagger A_\alpha U|\Psi\rangle - \langle\Psi|A_\alpha|\Psi\rangle|.$$

We can also bound the magnitude of the right-hand side by noticing that

$$|\langle\Psi_E|A_\alpha|\Psi_o\rangle|^2 = \langle\Psi_E|A_\alpha|\Psi_o\rangle\langle\Psi_o|A_\alpha^\dagger|\Psi_E\rangle$$
$$= \langle\Psi_E|A_\alpha(|\Psi_o\rangle\langle\Psi_o| + |\Psi_E\rangle\langle\Psi_E|)A_\alpha^\dagger|\Psi_E\rangle$$
$$- |\langle\Psi_E|A_\alpha|\Psi_E\rangle|^2$$
$$\leq \langle\Psi_E|A_\alpha A_\alpha^\dagger|\Psi_E\rangle - |\langle\Psi_E|A_\alpha|\Psi_E\rangle|^2.$$

In the intermediate step, we used the fact that $|\Psi_E\rangle\langle\Psi_E| + |\Psi_o\rangle\langle\Psi_o|$ is a projector when the two vectors are orthogonal and that for any projector, P,

$$\langle\Psi_E|A_\alpha P A_\alpha^\dagger|\Psi_E\rangle = \langle\Psi_E|A_\alpha PP A_\alpha^\dagger|\Psi_E\rangle$$
$$= |PA_\alpha^\dagger|\Psi_E\rangle|^2 \leq |A_\alpha^\dagger|\Psi_E\rangle|^2.$$

We can, of course, derive a similar equality showing that

$$|\langle\Psi_E|A_\alpha|\Psi_o\rangle|^2 \leq \langle\Psi_o|A_\alpha^\dagger A_\alpha|\Psi_o\rangle - |\langle\Psi_o|A_\alpha|\Psi_o\rangle|^2.$$

We expect correlators in $|\Psi_o\rangle$ to be thermal and the analysis above tells us that, to leading order in $\beta\delta E$, correlators in $|\Psi_E\rangle$ are also thermal.

Putting these results together, we find that

$$\delta\langle A_\alpha\rangle \leq 2\sqrt{\beta\delta E}\sigma_\alpha, \qquad (2.5)$$

for $\beta\delta E \ll 1$, where the "deviation" of A_α is defined through

$$\sigma_\alpha^2 \equiv \frac{1}{Z(\beta)} \min[\text{Tr}(e^{-\beta H} A_\alpha^\dagger A_\alpha), \text{Tr}(e^{-\beta H} A_\alpha A_\alpha^\dagger)]$$
$$- \left| \frac{1}{Z(\beta)} \text{Tr}(e^{-\beta H} A_\alpha) \right|^2. \qquad (2.6)$$

Therefore, we see that if $\beta\delta E \ll 1$, the change in the expectation value of A_α is very small. In other words, in terms of the expectation value of coarse-grained operators, the state $U|\Psi\rangle$ is very similar to the state $|\Psi\rangle$ for a typical state $|\Psi\rangle \in \mathcal{H}_E$. This result can be stated in the form of a slogan: it is impossible to definitively excite a thermal system with energy less than kT.

We emphasize that the result above applies to a typical state $|\Psi\rangle$. Given a specific state $|\Psi\rangle$, we can always tailor a unitary operator, U_ψ to excite that particular state so that $U_\psi|\Psi\rangle$ has very different properties from the original state. But the same unitary U_ψ will not work in other states that also belong to the space $\mathcal{H}_E$.

The result that we have proved here is entirely general and applies to any statistical system. We will see in the next few sections how at first sight, it appears that black holes could violate this bound [8] but that, in fact, these putative violations are unobservable. In a gravitational context, we need to be careful about the position dependence of the unitary excitation U and the observable A_α that could lead to a violation of (2.5). When we take this into account, we find that these two never fit into a single causal patch, and so no observer or collection of observers can detect this potential violation.

III. HOLOGRAPHIC SETUP AND CAUSAL PATCHES

We are interested in examining (2.5) in the context of AdS black holes. In this section, we describe our setup. As far as possible, we have maintained the conventions used in our previous papers [4,5,12]. The notion of causal patches, however, has not appeared previously in our work, and so we describe them in our context in some detail in Sec. III B. The proof that AdS correlators, after we account for the constraints of causality, obey (2.5) is provided in the next section.

A. Review of conventions

We will consider a CFT with a holographic dual, and in this dual geometry, we are interested in a large AdS Schwarzschild black hole. The metric is given by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2 d\Omega^2, \quad (3.1)$$

with

$$f(r) = 1 - \frac{c_d M}{r^{d-2}} + r^2,$$

with $c_d = 8(d-1)^{-1} \pi^{\frac{2-d}{2}} \Gamma(d/2)$, and where M is the mass of the black hole. It is convenient to go to tortoise coordinates

$$dr_* = \frac{dr}{f(r)}. \qquad (3.2)$$

We set $r_* \to 0$ at the boundary, and this leads to $r_* \to -\infty$ at the horizon. Note that near the boundary, we have $r_* = \frac{1}{r}$. Near the horizon, we have $f(r) \to \frac{4\pi}{\beta}(r - r_0)$, and therefore in this region, we have $r_* \sim \frac{\beta}{4\pi} \ln(r - r_0)$, where β is the inverse temperature of the black hole.

To cross the horizon, we move to Kruskal coordinates defined by

$$U_K = -e^{\frac{2\pi}{\beta}(r_*-t)}; \quad V_K = e^{\frac{2\pi}{\beta}(r_*+t)}, \quad \text{outside the horizon.}$$
$$(3.3)$$

Once we are behind the horizon, we can introduce a second Schwarzschild patch using the definitions

$$U_K = e^{\frac{2\pi}{\beta}(r_*-t)}; \quad V_K = e^{\frac{2\pi}{\beta}(r_*+t)}, \quad \text{inside the black hole.}$$

In this paper, we are interested in single-sided black holes, and in observers who fall into the black hole given by the metric (3.1) starting around time $t = 0$ on the boundary. Therefore, the region of the Penrose diagram deep in the interior of the black hole, corresponding to $U_K \gg 1$ is not of interest to us, since it is not accessible to any such observer. We are also only concerned with the region where $V_K > 0$.

The black hole metric (3.1) is dual to a state $|\Psi\rangle$ in the boundary CFT, with $\langle\Psi|H|\Psi\rangle \propto N^2$, where N^2 is the central charge of the CFT and H is the CFT Hamiltonian. This is a pure state in the CFT, which is not entangled with any other system. The expectation is that a typical state picked with the Haar measure on the microcanonical ensemble centered around an energy much larger than the central charge will correspond to the AdS-Schwarzschild geometry above.

B. Causal patches

Causal patches will play an important role in this paper. We define them as follows. Consider a point, $\mathcal{P}_\mathcal{C}$, on the singularity. The causal patch is then defined as the set of all bulk points that are in the causal past of this point. To picturise this set, consider past directed null rays emanating from the singularity at all possible angles. This light cone intersects the boundary of AdS on a sphere, at a time that we denote by $t_\mathcal{C}$. We now extend this region into the past. In the past, the black hole may have formed through collapse or through another process, and we are not interested in the details of the formation. So we cut off the causal patch at a value of $V_K = \exp[\frac{2\pi\vartheta}{\beta}]$; this intersects the boundary at a time ϑ in the past. The intersection of the causal patch with the boundary is denoted by $\mathcal{B}_\mathcal{C}$. This extends for a time

$|\mathcal{B}_\mathcal{C}| = t_\mathcal{C} - \vartheta$ along the boundary. It will be convenient for the reader to think of ϑ as a large negative number, and we discuss its magnitude in the next section. We denote the entire causal patch by $\mathcal{C}$. In an AdS Schwarzschild black hole, we have a one-parameter family of patches that can all be labeled by $t_\mathcal{C}$.

The physical motivation for considering causal patches is as follows. Consider the worldlines of an army of observers who all jump in from the boundary and end their lives on the singularity at $\mathcal{P}_\mathcal{C}$. Then the causal patch, as we have defined it, is the set of all spacetime events that could influence these observers before their wordlines terminate on the singularity.

Strictly speaking, as explained in [17], in $d > 2$, an observer who enters the horizon at some angular point Ω can receive signals from all points on the sphere only when he is exactly at the singularity. However, by considering the set of all wordlines that end at a point on the singularity rather than the causal past of a single observer, we obtain a spherically symmetric patch, which simplifies the algebra in the calculations that follow. In other contexts, it may be important to consider the causal patch corresponding to a single observer rather than an army of observers, but, in our context, this distinction does not seem to be important.

Figure 1 illustrates our perspective on the causal patch. The figure on the left shows a causal patch in a black hole formed from a Datt-Oppenheimer-Snyder type collapse [22] in AdS. Note that in the left figure, light rays are not 45 deg lines. So the inner boundary of the patch inside the horizon is a null geodesic but not a straight line. We are interested in the geometry of the causal patches at late time, without being concerned about the details of the collapsing shell. So, it is more convenient to think of an eternal geometry. Such a geometry can be obtained even in the single sided case by considering an orbifold of the eternal black hole geometry. This is shown on the right, where we also show the early time cutoff ϑ.

For future use, it is also useful to define two additional geometrical regions. We denote the causal wedge of $\mathcal{B}_\mathcal{C}$—all points in the bulk with the property that $\mathcal{B}_\mathcal{C}$ contains some part of their causal past and also some part of their causal future—by $\wedge_\mathcal{C}$. The intersection of the causal patch with the interior of the black hole is denoted by $\vee_\mathcal{C}$. These regions are shown in Fig. 2.

We pause to emphasize two points. First, it is possible to consider the causal patch corresponding to the observer who stays outside the black hole at all times. This causal patch, which we can call $\mathcal{C}_\infty$, encompasses the entire region outside the black hole. Second, note that these distinct causal patches only appear in a black hole background. In empty AdS, for example, there is a single causal patch that encompasses all observers.

As we explained above, a causal patch is specified by a point on the singularity. Note that the singularity is located at $r = 0$, and using (3.2), it is located at a value of r_{*s}, given by

$$r_{*s} = \lim_{\delta \to 0^+} \left[\int_{r_0 - \delta}^0 \frac{dr}{f(r)} - \int_{r_0 + \delta}^\infty \frac{dr}{f(r)} \right].$$

As explained in [23], for $d > 2$, we have $r_{*s} < 0$. For $d = 2$, we have $r_{*s} = 0$.

In Kruskal coordinates, the singularity is located at $U_K V_K = e^{\frac{4\pi}{\beta} r_{*s}}$. A point on the singularity can be labeled by a value of time $t_{\sin}$ on the singularity. Corresponding to this value, we have a causal patch, whose boundaries are
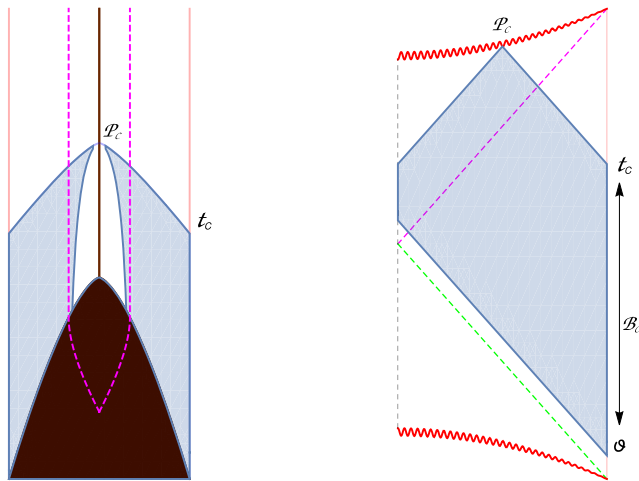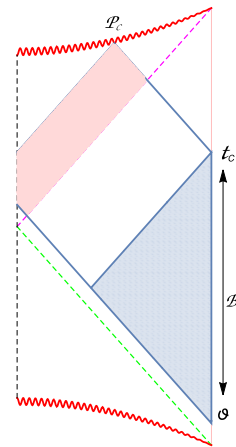

FIG. 1. Two views of a causal patch. On the left, we show a causal patch in a Datt-Oppenheimer-Snyder collapse in AdS. The collapsing star is shaded in brown. On the right, we show a causal patch in an eternal single-sided geometry, and we also show the early time cutoff at ϑ. In both figures, the patch is demarcated in blue gray, and the future horizon is shown by a dashed magenta line.

FIG. 2. The causal wedge, $\wedge_\mathcal{C}$ is shaded in blue outside the horizon; the intersection of the causal patch with the interior, denoted by $\vee_\mathcal{C}$ is shaded pink inside the horizon.

$U_K = e^{\frac{2\pi}{\beta}(r_{*s} - t_{\sin})}$ and $V_K = e^{\frac{2\pi}{\beta}(r_{*s} + t_{\sin})}$. This causal patch intersects the boundary at the time $t_{\mathcal{C}} = r_{*s} + t_{\sin}$. Since $r_{*s} \leq 0$ as we pointed out above, we see that $t_{\sin} \geq t_{\mathcal{C}}$.

C. CFT definitions and conventions

A central role in our analysis is played by generalized free fields, which are light operators in the theory whose correlators factorize. We denote such a field by $\mathcal{O}(t, \Omega)$. For example, in the $N = 4$ SYM theory, both $\mathrm{Tr}(F^2)$ and the stress tensor $T_{\mu\nu}$ are generalized free fields, although in this paper, to lighten the notation, we will drop the tensor indices on all such operators. Just as in [24], we consider angular momentum modes in such fields

$$\mathcal{O}_\ell(t) = \int d\Omega \mathcal{O}(t, \Omega) Y_\ell^*(\Omega), \qquad (3.4)$$

where the integral is over the boundary S^{d-1} and $Y_\ell(\Omega)$ are the spherical harmonics. We then consider the polynomial algebra in these modes

$$\mathcal{A} = \mathrm{span\ of} \{\mathcal{O}_{\ell_1}(t_1), \mathcal{O}_{\ell_2}(t_2)\mathcal{O}_{\ell_3}(t_3), \dots,$$
$$\mathcal{O}_{\ell_4}(t_4)\mathcal{O}_{\ell_5}(t_5)\cdots\mathcal{O}_{\ell_{\mathcal{D}_m}}(t_{\mathcal{D}_m})\}. \qquad (3.5)$$

Here, we have introduced a cutoff $\mathcal{D}_m$ in the highest angular momentum mode that is allowed to appear, with $\mathcal{D}_m \ll N$. As we mentioned above, it is also important to place a long time cutoff on the boundary. This can be done by restricting the times t_i that appear above to lie in $t_i \in [-|\vartheta|, |\vartheta|]$. Here, ϑ can be taken to be a power of N but not exponentially large $|\vartheta| \ll e^{N^2}$. As discussed in [12,13], unless we place such a cutoff on the length of the longest possible time interval in the CFT, there is no meaningful way to speak of a dual geometry.

The reader will notice that our notation differs from the notation of [4,5] in that we have defined the algebra using operators that are local in time, rather than using operators of definite frequency. At an abstract level, these two definitions are equivalent. But our notation reflects one of the messages of this paper: a position space analysis often provides insights that may be missed by a purely frequency-space analysis.

An element of the algebra is denoted by $A_\alpha \in \mathcal{A}$. At times, we would like to consider an element that is localized at a given point in time, and we denote this by $A_\alpha(t)$. Such an operator is a polynomial in local operators, all of which are at the time t. The modes defined in (3.4) are examples of such operators.

We will also consider nonequilibrium states that are related to the simple equilibrium state $|\Psi\rangle$. These excited states are of the form

$$|\Psi^{\mathrm{ne}}\rangle = U|\Psi\rangle,$$

where U is a unitary formed by exponentiating a Hermitian element of the algebra localized around some time T. Sometimes, rather than thinking of the nonequilibrium state as a state in the original theory, it is convenient to think of it as being produced by actively deforming the Hamiltonian on the boundary. This is an elementary point, but since it may be the source of some confusion, we now explain it in some detail.

1. Source deformed correlators as correlators in autonomously excited states

To actively produce a nonequilibrium state, we proceed as follows. We allow the CFT to thermalize so that it goes into an equilibrium state $|\Psi\rangle$. Then we deform the CFT Hamiltonian by turning on a source

$$H^J(t) = H + J(t)A_\alpha^J(t), \qquad (3.6)$$

where $A_\alpha(t)$ is a some element of the algebra made out of local operators at time t. H is the original CFT Hamiltonian. A_α^J are the new Heisenberg operators, and they are related to the original Heisenberg operators through

$$A_\gamma^J(t) = \bar{\mathcal{T}}\{e^{i\int_\vartheta^t J(x)A_\alpha(x)dx}\}A_\gamma(t)\mathcal{T}\{e^{-i\int_\vartheta^t J(x)A_\alpha(x)dx}\}, \quad (3.7)$$

where $\mathcal{T}$ ($\bar{\mathcal{T}}$) is the (anti)time-ordering symbol, and the superscript distinguishes these operators from the original operators. If we consider the expectation value of a single source-deformed Heisenberg observable localized at a time T, this is given just by sandwiching the expression above in a state

$$\langle\Psi|A_\gamma^J(T)|\Psi\rangle = \langle\Psi|\bar{\mathcal{T}}\{e^{i\int_\vartheta^T J(x)A_\alpha(x)dx}\}A_\gamma(T)$$
$$\times \mathcal{T}\{e^{-i\int_\vartheta^T J(x)A_\alpha(x)dx}\}|\Psi\rangle.$$

Therefore, modifying the Hamiltonian through (3.6) functions *as if* we were evaluating correlators in the non-equilibrium state $|\Psi^{\mathrm{ne}}\rangle = \mathcal{T}\{e^{-i\int_\vartheta^T J(x)A_\alpha(x)dx}\}|\psi\rangle$ after the time T. Before that time, the state is just the ordinary equilibrium state.

On the other hand, it is also possible to think of $|\Psi^{\mathrm{ne}}\rangle$ as an autonomous state i.e., a state in the theory with the undeformed Hamiltonian that simply happens to have a spontaneous excitation around the time T. Note that even in this picture, at times $t \ll T$, the state looks close to equilibrium. For example, if we think of the boundary as a fluid, then this is a state where the fluid is initially in equilibrium, but then the microscopic elements of the fluid conspire to throw up an excitation around T that again decays away. An observer on the boundary in such a state would observe a violation of the second law of

thermodynamics, but such states are part of the Hilbert space of the CFT.[2]

This equivalence between the two descriptions generalizes to the case where the observable in question is a product of localized operators in the bulk. As a special case of (3.7), the bulk fields are deformed to

$$\phi^J(t, r_*, \Omega) = \bar{\mathcal{T}}\{e^{i\int_\vartheta^t J(x)A_\alpha(x)dx}\}\phi(t, r_*, \Omega)$$
$$\times \mathcal{T}\{e^{-i\int_\vartheta^t J(x)A_\alpha(x)dx}\}.$$

But now, note that in the time-ordered unitary that appears on the right, the later part of the integral that runs from $[t + r_*, t]$ involves operators that commute with the bulk field. Since $A_\alpha(t)$ is localized on the boundary, by bulk locality, we have $[\phi(t, r_*, \Omega), A_\alpha(x)] = 0$ if $t + r_* < x < t - r_*$. We need only the first part of this inequality since in the expression above, $x \leq t$. The adjoint of the part of the time-ordered unitary that runs from $[t + r_*, t]$ appears on the extreme right of the anti-time-ordered unitary. Therefore,

$$\phi^J(t, r_*, \Omega) = W^\dagger(t + r_*)\phi(t, r_*, \Omega)W(t + r_*),$$

where

$$W(t + r_*) \equiv \mathcal{T}\{e^{-i\int_\vartheta^{t+r_*} J(x)A_\alpha(x)dx}\}.$$

But this means that a correlator of the deformed field can be written as

$$\langle\Psi|\phi^J(t_1, r_{*1}, \Omega_1)...\phi^J(t_n, r_{*n}, \Omega_n)|\Psi\rangle$$
$$= \langle\Psi|W(t_1 + r_{*1})^\dagger\phi(t_1, r_{*1}, \Omega_1)W(t_1 + r_{*1})...$$
$$\times W(t_n + r_{*n})^\dagger\phi(t_n, r_{*n}, \Omega_n)W(t_n + r_{*n})|\Psi\rangle.$$

Furthermore, we can massage this correlator so that it looks precisely of the form (2.5) as follows. Define

$$U = \exp\left[-i\int_\vartheta^{t_C} A_\alpha(t)dt\right],$$

as in (1.5). Then we find that

$$\langle\Psi|\phi^J(t_1, r_{*1}, \Omega_1)...\phi^J(t_n, r_{*n}, \Omega_n)|\Psi\rangle = \langle\Psi|U^\dagger A_\gamma U|\Psi\rangle$$

---

where

$$A_\gamma = \exp\left[-i\int_\vartheta^{t_C} A_\alpha(t)dt\right]W(t_1 + r_{*1})^\dagger\phi(t_1, r_{*1}, \Omega_1)$$
$$W(t_1 + r_{*1})...W(t_n + r_{*n})^\dagger\phi(t_n, r_{*n}, \Omega_n)$$
$$\times W(t_n + r_{*n})\exp\left[i\int_\vartheta^{t_C} A_\alpha(t)dt\right]. \quad (3.8)$$

Now, we note that $(t, r_*, \Omega) \in \mathcal{C} \Rightarrow (t + r_*, 0, \Omega') \in \mathcal{C}, \forall\Omega'$. Therefore, the operator A_γ is entirely localized in the causal patch, and U is localized on the boundary of the patch.

The conclusion is that *a correlator with source-deformed fields can be reduced to a correlator of ordinary fields with insertions only in $\mathcal{C}$ and an excitation on the boundary of $\mathcal{C}$.*

In operational terms, this yields the following protocol to check (2.5) in the situation where we create an excitation *actively* by deforming the Hamiltonian through a local source. The infalling observer can measure $\langle\Psi|\phi^J(t_1, r_{*1}, \Omega_1)...\phi^J(t_n, r_{*n}, \Omega_n)|\Psi\rangle$ and compare this excited expectation value with $\langle\Psi|A_\gamma|\Psi\rangle$ in the unexcited state, where A_γ is specified in (3.8).

On physical grounds, it is more natural to think of $|\Psi^{\text{ne}}\rangle$ as being produced by turning on a local source on the boundary. But, from a technical perspective, it is more convenient to think of the state as an autonomous state since then we do not have to worry about the time ordering in (3.7).

In Sec. IV we will be able to prove that AdS correlators in a causal patch are stable, in that they obey the constraints (2.5) of statistical mechanics, in all possible autonomous states of the form $|\Psi^{\text{ne}}\rangle$ where the unitary and the observables fit in the same causal patch. As we have explained above, this also implies that these correlators are stable if an observer or an army of observers on the boundary turn on a source and subsequently jump into the bulk and cross the horizon. We reiterate that this applies to arbitrary sources dual to local boundary operators—even those that continue to act beyond the causal patch.

No bulk sources While we have considered the effect of sources dual to local operators on the boundary above, we emphasize that we do not consider situations where sources are turned on in the bulk. A bulk source would violate local energy conservation. So, in a theory of quantum gravity, it does not make sense to modify the bulk Hamiltonian and only boundary sources are allowed.

This is a subtle point. If $\phi(f)$ is a smeared Hermitian bulk operator then correlators where local insertions are sandwiched between bulk unitaries, $e^{i\phi(f)}$ and $e^{-i\phi(f)}$, can be expanded in a series of Wightman functions. But the bulk observer can only measure time ordered correlators, and not arbitrary Wightman functions. So the allowed bulk observables are

$$\langle \Psi^{\mathrm{ne}}|\mathcal{T}[\phi(t_1, r_{*1}, \Omega_1)...\phi(t_n, r_{*n}, \Omega_n)]|\Psi^{\mathrm{ne}}\rangle$$
$$= \langle \Psi|U^{\dagger}\mathcal{T}[\phi(t_1, r_{*1}, \Omega_1)...\phi(t_n, r_{*n}, \Omega_n)]U|\Psi\rangle.$$

Notice that the unitaries on the left and the right are outside the time ordering, since they may be produced by spontaneous excitations or boundary sources.

We will not need this time ordering in most of this paper since the precise results that we describe below—such as the proof of (1.6)—hold for all Wightman functions, and not just for time-ordered bulk correlators. Nevertheless, the physical restriction to time-ordered bulk observables is important to keep in mind to avoid the appearance of *additional paradoxes* corresponding to bulk sources and bulk unitaries.

D. Transfer function

An important role in the program of reconstructing the bulk from the boundary is played by transfer functions that map bulk operators to boundary operators. This is reviewed in detail in [3]. Here, we review the essential formulas and direct the reader to [3] and references there for further details.

Outside the horizon (corresponding to $U < 0, e^{\frac{2\pi\vartheta}{\beta}} < V < e^{\frac{2\pi}{\beta}t_C}$), the bulk field can be mapped to boundary operators via the mapping

$$\phi(t, r_*, \Omega) = \sum_{\ell,\omega} D^0(\omega, \ell)\mathcal{O}_{\omega,\ell}\zeta_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega) + \mathrm{H.c.},$$

$$(3.9)$$

whereas for the part of the causal patch inside the horizon (corresponding to $0 < U < e^{\frac{2\pi}{\beta}(2r_{*s}-t_C)}, e^{\frac{2\pi}{\beta}} < V < e^{\frac{2\pi}{\beta}t_C}$, the mapping between bulk and boundary operators is given by

$$\phi(t, r_*, \Omega) = \sum_{\ell,\omega} Y_\ell(\Omega)D^0(\omega, \ell)\widetilde{\zeta}^-_{\omega,\ell}(r_*)$$
$$\times (\mathcal{O}_{\omega,\ell}e^{2i\delta_{\omega,\ell}}e^{-i\omega t} + \widetilde{\mathcal{O}}_{\omega,\ell}e^{i\omega t}) + \mathrm{H.c.} \quad (3.10)$$

We have introduced several pieces of notation above, and we explain them now. The radial functions $\zeta_{\omega,\ell}(r_*)$ and $\widetilde{\zeta}^-_{\omega,\ell}(r_*)$ ensure that the full mode satisfies the wave equation both inside and outside the horizon. These functions are chosen to satisfy specific boundary conditions. So, if the bulk field has mass m, then

$$(\Box - m^2)\zeta_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega) = 0,$$
$$\zeta_{\omega,\ell}(r_*) \underset{r_* \to -\infty}{\longrightarrow} e^{i\omega r_*} + e^{2i\delta_{\omega,\ell}}e^{-i\omega r_*},$$

whereas the mode behind the horizon satisfies

$$\widetilde{\zeta}^-_{\omega,\ell}(r_*) \underset{r_* \to -\infty}{\longrightarrow} e^{-i\omega r_*},$$

and this time the limit $r_* \to -\infty$ is taken from behind the horizon.

The phase $\delta_{\omega,\ell}$ and the function $D^0(\omega, \ell)$ are fixed by ensuring that ζ is normalizable at the boundary and the normalizable part has coefficient,

$$\zeta_{\omega,\ell}(r_*) \underset{r_* \to 0}{\longrightarrow} \frac{1}{D^0(\omega, \ell)}(r_*)^\Delta,$$

where Δ is the dimension of the operator $\mathcal{O}$. These radial equations are discussed in some detail in Sec. IV.

Finally, $\mathcal{O}_{\omega,\ell}$ are modes of the boundary operators, and $\widetilde{\mathcal{O}}_{\omega,\ell}$ are the "mirror" operators, which we will review below. As we pointed out in [5], the modes of the boundary operators cannot be defined as infinite time Fourier transforms; instead, it is important to define these modes as Fourier series coefficients corresponding to a compact region on the boundary.

$$\mathcal{O}_{\omega,\ell} = \int d\Omega \int_{-|\vartheta|}^{|\vartheta|} dt\mathcal{O}(t, \Omega)Y_\ell^*(\Omega)e^{i\omega t}, \quad (3.11)$$

where $|\vartheta|$ is the long time cutoff introduced above. Corresponding to this, we see that ω must be quantized in units of $\frac{\pi n}{|\vartheta|}$. We wish to avoid having this cutoff appear in all our formulas. To do this, we will use the following notational tricks. First, it is understood that whenever we write a sum over ω, we really mean

$$\sum_{\omega} f(\omega) \equiv \frac{1}{2|\vartheta|}\sum_{n>0} f(n),$$

where $\omega = \frac{\pi n}{|\vartheta|}$ as above and additionally, the sum over frequencies runs only over positive frequencies. Second, when we write a delta function, we mean

$$\delta_{\omega,\omega'} \equiv 2|\vartheta|\delta_{nn'}.$$

The reader will note that we have already tacitly used this notation in the formulas above.

We now turn to the definition of the operators $\widetilde{\mathcal{O}}$. These operators were described in [4,5,12], and we refer the reader to those papers for details. Here, we simply recall that these operators are defined through a set of linear equations. Given an equilibrium state, we define the mirror operators to satisfy

$$\widetilde{\mathcal{O}}_{\omega,\ell}A_\alpha|\Psi\rangle = e^{-\frac{\beta\omega}{2}}A_\alpha\mathcal{O}^{\dagger}_{\omega,\ell}|\Psi\rangle. \quad (3.12)$$

This completely specifies the action of the operators on the subspace $\mathcal{H}_\Psi = \mathcal{A}|\Psi\rangle$; this is the subspace formed by low energy excitations of the equilibrium state and is the only part of the Hilbert space that is relevant for effective field theory experiments done about the state $|\Psi\rangle$.

IV. PROOF OF THE STABILITY OF AdS CORRELATORS IN A CAUSAL PATCH

In this section, we will now prove that AdS correlators obey the inequality (2.5) whenever the excitation U and the correlation function to be measured fit in a single causal patch. This automatically implies that in the physically important setting where we start with an equilibrium state, and turn on a local source, we do not observe any violation of (2.5). This is because as we see from (1.4) an operator at any point in the bulk cannot be affected by any part of the source that is not in its causal past; therefore it cannot be affected by any part of the source that is outside its causal patch.[3]

Since our proof is somewhat technical, we summarize the main ideas before starting the proof. It is useful to understand why the inequality is in danger of being violated in the first place. Bulk operators defined through (3.9) are simply rewritings of ordinary operators on the boundary. So they must automatically obey the inequality (2.5) which applies to any ordinary operator. On the other hand, the bulk operators defined in (3.10) depend on the operators $\tilde{\mathcal{O}}_{\omega,\ell}$ that are defined through the linear equations (3.12). We remind the reader that these operators are *state dependent*. This means that it is impossible to find a globally defined operator $\tilde{\mathcal{O}}_{\omega,\ell}$ that obeys the Eqs. (3.12) about all equilibrium states $|\Psi\rangle$. The danger is that this state dependence may cause these operators to behave very differently in an excited state $U|\Psi\rangle$ than in a typical equilibrium state $|\Psi\rangle$, even if U is a low energy excitation.

But the modes $\mathcal{O}_{\omega,\ell}$ are an auxiliary device for constructing the local field operators $\phi(t, r_*, \Omega)$. What we will show in this section is that when these operators are localized in the same causal patch as the excitation U, they obey the inequality (2.5) because the anomalous transformations of the individual modes cancel among each other.

In more detail, we proceed as follows. Even though the AdS potential causes backscattering, it is possible to define a set of right-moving and left-moving modes in AdS. The right movers are the modes that behave like $\phi_r \sim e^{i\omega(r_* - t)}$ near the horizon, and the left movers are those that behave like $\phi_l \sim e^{-i\omega(r_* + t)} e^{2i\delta_{\omega,\ell}}$ near the horizon. As we cross the horizon, the left movers just continue smoothly from the region outside the horizon to the region inside the horizon. But the right movers behind the horizon, $\tilde{\phi}_r$, now contain the mirror modes rather than the ordinary modes.

The operators $\tilde{\phi}_r$ are defined so that in an equilibrium state, they can be replaced, up to a conjugation by $e^{\frac{\beta H}{2}}$, with an ordinary operator $\hat{\phi}$. The calculation that we perform is to
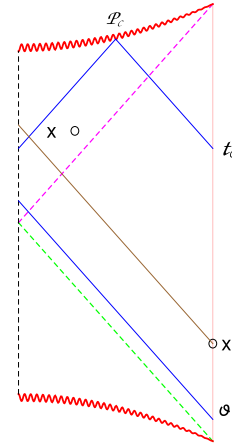

FIG. 3. A boundary excitation at x' commutes with the right moving hatted field at x even though x is in the causal future of x'. The boundaries of the causal patch are in blue. The light cone from x' is marked off in brown.

show that the commutator $[\hat{\phi}(x), \phi(x')]$ vanishes when x is placed in $\vee_{\mathcal{C}}$—which is the part of the causal patch inside the horizon—and x' is placed near the boundary on $\mathcal{B}_{\mathcal{C}}$. (See Fig. 3).

Now, consider two states—an equilibrium state $|\Psi\rangle$ and a near-equilibrium state $|\Psi^{\text{ne}}\rangle = U^{\mathcal{C}}|\Psi\rangle$, where $U^{\mathcal{C}}$ is localized on $\mathcal{B}_{\mathcal{C}}$. As a consequence of this vanishing commutator, a correlator involving $\tilde{\phi}_r(x)$ can be replaced—up to the conjugation by $e^{\frac{\beta H}{2}}$ that leads to an error of only $O(\beta\delta E)$—by a correlator of ordinary operators involving just $\hat{\phi}(x)$ *both* in $|\Psi\rangle$ and $|\Psi^{\text{ne}}\rangle$. If the commutator had not vanished, this would not have been possible since $\tilde{\phi}_r(x)$ commutes with $U^{\mathcal{C}}$ and so it can only be substituted by another operator that also commutes with $U^{\mathcal{C}}$. Now, since this ordinary operator obeys the inequality (2.5), we see that the correlator involving $\tilde{\phi}_r(x)$ also obeys (2.5) as long as x is within the causal patch.

To compute the commutator, first, we analyze the analytic properties of the normalizable mode outside the horizon. This is most conveniently done by building the mode starting at the boundary and then solving inwards in Sec. IV A. Next, we define and analyze the analytic properties of the right-moving component of this mode; this is most conveniently done by specifying the mode at the horizon and solving outwards towards the boundary in Sec. IV B. Finally, we consider the analytic properties of components of the field behind the horizon in Sec. IV C, which are the same as those of the right-moving components outside the horizon. In Sec. IV D, we determine the commutator $[\hat{\phi}(x), \phi(x')]$. A knowledge of the analytic properties of the mode functions is enough to show the vanishing of the commutator in the region of interest. In the last section, we put together all this information to show that AdS correlators are stable under low energy excitations. This means that they obey the inequality

(2.5) and a low energy excitation cannot change the correlator significantly.

A. Analysis from the boundary side

To analyze the AdS wave functions, we will follow the techniques of [25]. In the context of AdS/CFT, these wave functions have been discussed in the context of Schwarzschild-AdS quasinormal modes [26,27]. We review and rephrase this discussion in the context that we need.

The Klein-Gordon wave equation is given by

$$\frac{1}{\sqrt{-g}}\partial_\mu g^{\mu\nu}\sqrt{-g}\partial_\nu\phi - m^2\phi = 0,$$

where the AdS_{d+1} black hole metric is given by (3.1). As above it is convenient to go to the tortoise coordinates defined in (3.2). We remind the reader that $r_* \to 0$ at the boundary and $r_* \to -\infty$ at the horizon.

In these coordinates, we have $\sqrt{-g} = f(r)r^{d-1}$. We also have $g^{r_*r_*} = -g^{tt} = \frac{1}{f(r)}$. Then the wave equation becomes

$$\frac{1}{f(r)r^{d-1}}\partial_{r_*}r^{d-1}\partial_{r_*}\phi - \frac{1}{f(r)}\partial_t^2\phi + \frac{1}{r^2}\Box_\Omega\phi - m^2\phi = 0,$$

$$(4.1)$$

where r is now understood to be a function of r_* and $\Box_\Omega$ is the Laplacian on the sphere.

We can separate variables and write an ansatz for the solution of (4.1) as

$$\phi_{\omega,\ell}(t, r_*, \Omega) = r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega).$$

Note that

$$\partial_{r_*}r^{d-1}\partial_{r_*}r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*) = r^{\frac{d-1}{2}}\partial_{r_*}^2\chi_{\omega,\ell}(r_*) + \frac{1}{4}[(d-3)f(r)$$
$$+ 2rf'(r)](1-d)f(r)r^{\frac{d-5}{2}}\chi_{\omega,\ell}(r_*),$$

and the differential equation for $\chi_{\omega,\ell}$ now becomes

$$\frac{\partial^2\chi_{\omega,\ell}(r_*)}{\partial r_*^2} = V(r_*)\chi_{\omega,\ell}(r_*),$$

where

$$V(r_*) = -\left[\omega^2 + \frac{(d-3)(1-d)}{4}\frac{f(r)^2}{r^2} - m^2f(r)\right.$$
$$\left. - l(l+d-2)\frac{f(r)}{r^2} + \frac{1-d}{2}\frac{f'(r)f(r)}{r}\right].$$

Near the boundary, it is also convenient to write this equation as

$$\frac{\partial^2\chi_{\omega,\ell}(r_*)}{\partial r_*^2} + \gamma^2\chi_{\omega,\ell} - \frac{(m^2 + \frac{d^2-1}{4})}{r_*^2}\chi_{\omega,\ell} = V_{\rm bd}(r_*), \quad (4.2)$$

where

$$\gamma^2 \equiv \omega^2 - l(l+d-2)) - \frac{m^2 + (d-2)(d-1)}{3},$$

and

$$V_{\rm bd}(r_*) = l(l+d-2)\left(\frac{f(r)}{r^2} - 1\right) + \frac{(d-3)(d-1)}{4}$$
$$\times \left(\frac{f^2(r)}{r^2} - \frac{1}{r_*^2} - \frac{4}{3}\right) + m^2\left(f(r) - \frac{1}{r_*^2} - \frac{1}{3}\right)$$
$$+ \frac{d-1}{2}\left(\frac{f'(r)f(r)}{r} - \frac{2}{r_*^2} - \frac{2}{3}\right).$$

We see that $V_{\rm bd}(r_*)$ is now finite at all values of $r_* \in (-\infty, 0]$, which is the range that we are interested in, and that near the boundary, $V_{\rm bd}(r_*) = O(r_*^2)$.

We can solve this equation using Green's functions both near the boundary and near the horizon. The starting point for this is to write a "free" equation near the boundary (neglecting $V_{\rm bd}$),

$$\frac{\partial^2\chi_{\omega,\ell}^0(r_*)}{\partial r_*^2} + \gamma^2\chi_{\omega,\ell}^0(r_*) - \frac{(m^2 + \frac{d^2-1}{4})}{r_*^2}\chi_{\omega,\ell}^0(r_*) = 0. \quad (4.3)$$

Note that at the boundary, since the potential $V_{\rm bd} \to 0$, we have $\chi_{\omega,\ell}^0(r_*) \to_{r_* \to 0} \chi_{\omega,\ell}(r_*)$. Therefore, imposing normalizable boundary conditions at the boundary, the only allowed solution is

$$\chi_{\omega,\ell}^0(r_*) = \frac{1}{\gamma^\nu}\sqrt{r_*}J_\nu(\gamma r_*),$$

with $\nu = \sqrt{m^2 + \frac{d^2}{4}}$. We have normalized the solution so that near the boundary, where $r_* \to \frac{1}{r} \to 0$, the behavior of the original field is

$$\phi_{\omega,\ell}(t, r_*, \Omega) = e^{-i\omega t}Y_\ell(\Omega)r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*) \to_{r_* \to 0} e^{-i\omega t}Y_\ell(\Omega)r_*^{\frac{d}{2}+\nu}.$$

We are ultimately interested in the analytic properties of the mode as we vary ω. From the series expansion of the Bessel function given above, we see that

$$\frac{1}{\gamma^\nu}J_\nu(\gamma r_*) = \sum_{s=0}^\infty \frac{(-1)^s}{\Gamma(s+1)\Gamma(\nu+s+1)}\gamma^{2s}\left(\frac{r_*}{2}\right)^{\nu+2s}.$$

In particular, we see that in the expression for $\chi_{\omega,\ell}^0(r_*)$, γ appears only raised to positive even integer powers. Therefore we see that $\chi_{\omega,\ell}^{(0)}(r_*)$ has no poles or branch cuts in ω at any finite r_* at any finite value of ω.

We now show that the full solution with the potential $V_{\rm bd}$ included also has the same analytic structure in ω. To construct the solution away from the boundary, we need the Green's function for the Eq. (4.2). This can be constructed as follows. Consider the second solution to the free equation (4.3). This can be written as $\gamma^\nu \sqrt{r_*} Y_\nu(\gamma r_*)$. We can then write the Green's function as

$$G(r'_*, r_*) = \frac{\pi}{2}\sqrt{r_* r'_*}(J_\nu(\gamma r_*)Y_\nu(\gamma r'_*)$$
$$- Y_\nu(\gamma r_*)J_\nu(\gamma r'_*))\theta(r'_* - r_*). \quad (4.4)$$

To examine the analytic properties of $G(r'_*, r_*)$ in ω, we note that

$$Y_\nu(x) = \frac{\cos(\nu\pi)J_\nu(x) - J_{-\nu}(x)}{\sin\nu\pi}.$$

This form is valid for all noninteger ν, and for integer ν, it should be understood appropriately as a limit taken from noninteger ν. We see then that a term proportional to $J_\nu(\gamma r_*)J_\nu(\gamma r'_*)$ cancels between the two terms of (4.4), leading to a series expansion

$$G(r'_*, r_*) = \frac{\pi\sqrt{r_* r'_*}}{2\sin(\nu\pi)}\sum_{s,t=0}^\infty \frac{(-1)^{s+t}\gamma^{2(s+t)}(\frac{r'_*}{2})^{2s+\nu}(\frac{r_*}{2})^{2t-\nu}}{\Gamma(s+1)\Gamma(\nu+s+1)\Gamma(t+1)\Gamma(t-\nu+1)} - (r'_* \leftrightarrow r_*), \quad (4.5)$$

for $r'_* > r_*$. The leading nonanalytic term γ^ν cancels in the product of $J_\nu(\gamma r_*)$ and $J_{-\nu}(\gamma r'_*)$ and thereafter, only positive integer powers of γ^2 occur. Therefore, we conclude that (4.4) is also an analytic function of ω with no poles or branch cuts at finite ω.

Using this Green's function and the solution $\chi^0_{\omega,\ell}$, we can proceed to construct the full solution. We write

$$\chi^{(n)}_{\omega,\ell}(r_*) = \int_0^{r_*}\chi^{(n-1)}_{\omega,\ell}(r'_*)G(r'_*, r_*)V_{\rm bd}(r'_*)dr'_*, \quad (4.6)$$

where the range of integration can be restricted to r_* because the Green's function vanishes beyond that. We then write the full solution as

$$\chi_{\omega,\ell}(r_*) = \sum_{n=0}^\infty \chi^{(n)}_{\omega,\ell}(r_*). \quad (4.7)$$

Now, we note that in (4.6), the integrand is finite at all values of r'_*, and each $\chi^{(n)}_{\omega,\ell}(r_*)$ dies off at least as fast as $(r_*)^{\nu+\frac{1}{2}+4n}$ near the boundary. We can prove these facts together using induction. These properties are clearly true for the iteration from $\chi^{(0)}_{\omega,\ell}$ to $\chi^{(1)}_{\omega,\ell}(r_*)$. Now assume that we have performed the iteration till $\chi^{(n-1)}_{\omega,\ell}(r_*)$. Then the Green's function has a singularity which, at worst, goes like $(r'_*)^{\frac{1}{2}-\nu}$ near the boundary. On the other hand, by assumption, $\chi^{(n-1)}_{\omega,\ell}$ dies off like $(r'_*)^{\nu+\frac{1}{2}+4(n-1)}$ near the boundary and, in addition, we can check that $V_{\rm bd}(r'_*) \to_{r'_* \to 0}(r'_*)^2$. So the integrand scales like $(r'_*)^{4n-1}$ as we approach the boundary. We see that after we do the integral up to r_* and note that the Green's function has a leading factor of $(r_*)^{\nu+\frac{1}{2}}$ in (4.5), we find that $\chi^{(n)}(r_*)$ dies off like $r_*^{\nu+\frac{1}{2}+4n}$, which is what we need.

Second, since we are integrating a finite integrand over a finite region, we see that the iteration (4.6) produces no poles at finite ω and finite r_*. Therefore, the full solution (4.7) has no poles or branch cuts at finite ω and finite r_*.

1. Transfer function using boundary-side modes

We now describe how these modes can be combined with the modes of boundary operators to write down a bulk field. Since the function $r^{\frac{1-d}{2}}\chi_{\omega,\ell}$ is normalized to die off like $r^{-\Delta}$ near the boundary, we can immediately write an expression for the bulk field as

$$\phi(t, r_*, \Omega) = \sum_{\ell,\omega}\mathcal{O}_{\omega,\ell}r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega) + {\rm H.c.},$$

where the boundary modes $\mathcal{O}_{\omega,\ell}$ were defined in (3.11).

However, we would like to write the bulk field in terms of ordinary creation and annihilation operators that have simple commutators. To do this, we consider the expansion of the radial mode near the horizon

$$r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*) \underset{r_* \to -\infty}{\longrightarrow} D^0(\omega, \ell)(e^{i\omega r_*} + e^{2i\delta_{\omega,\ell}}e^{-i\omega r_*}), \quad (4.8)$$

where the functions D^0 and $\delta_{\omega,\ell}$ are defined by the relation above. We can now alternately define

$$\zeta_{\omega,\ell}(r_*) = \frac{1}{D^0(\omega, \ell)}r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*),$$

where $\zeta_{\omega,\ell}(r_*)$ is a sum of plane waves near the horizon. Using these wave functions, the bulk field can also be written as

$$\phi(t, r_*, \Omega) = \sum_{\ell,\omega}\frac{a_{\omega,\ell}}{\sqrt{\omega}}\zeta_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega) + {\rm H.c.}, \quad (4.9)$$

where

$$a_{\omega,\ell} = D^0(\omega,\ell)\sqrt{\omega}\mathcal{O}_{\omega,\ell}. \qquad (4.10)$$

With a little algebra, it is not difficult to see that (with an appropriate ω-independent and ℓ-independent normalization for the spherical harmonics) ϕ obeys the canonical commutation relations if the operators $a_{\omega,\ell}$ satisfy

$$[a_{\omega,\ell}, a^{\dagger}_{\omega',\ell'}] = \delta_{\omega,\omega'}\delta_{\ell,\ell'}.$$

Note that this implies that[4]

Now we turn to the analyticity properties of $\zeta_{\omega,\ell}$. At finite ω, since $\chi_{\omega,\ell}$ has no singularities, the singularities of $\zeta(\omega,\ell)$ can only come from the zeroes of $D^0(\omega,\ell)$ since. But these zeroes have a nice physical interpretation. They correspond to the *quasinormal* modes of the black hole. This is because $D^0(\omega,\ell)$ is defined as the coefficient of the outgoing wave, $e^{i\omega r_*}$, in (4.8). So, when $D^0(\omega,\ell)$ vanishes, the mode becomes purely ingoing at horizon.

Note that the vanishing of $D^0(\omega,\ell)$ does not mean that the entire mode in (4.8) vanishes. The phase factor can be written as $e^{2i\delta_{\omega,\ell}} = \frac{D^0(-\omega,\ell)}{D^0(\omega,\ell)}$, and so it has poles corresponding to the complex zeroes of $D^0(\omega,\ell)$.

The quasinormal modes have the property that they all have $\text{Im}(\omega) < 0$. (See [26] for a proof.) Therefore, we see that $\zeta_{\omega,\ell}$ has poles when ω is in the lower-half plane, and it is regular in the upper-half plane.

One subtle point has to do with the behavior of $D^0(\omega,\ell)$ when ω approaches infinity through the upper-half plane. In this limit, where $|\omega| \gg 1$, while $\text{Im}(\omega) > 0$, the potential V_{bd} becomes irrelevant, and we can read off $D^0(\omega,\ell)$ from $\chi^0_{\omega,\ell}(r_*)$, and we see that it approaches a constant and does not go to 0.

B. Analysis near the horizon

We now consider the bulk modes from a second perspective starting with the horizon and solving outwards. This allows us to define what we mean by the "right moving" modes and analyze their analytic properties.

---

[4]This relates a property of the bulk wave function specified in (4.8) to the commutator of boundary operators in a thermal state; this is a prediction of AdS/CFT and can be verified in several cases. (See, for example, Appendix A of [3].)

$$\langle\Psi|[\mathcal{O}_{\omega,\ell}, \mathcal{O}^{\dagger}_{\omega',\ell'}]|\Psi\rangle = \frac{1}{Z(\beta)}\text{Tr}(e^{-\beta H}[\mathcal{O}_{\omega,\ell}, \mathcal{O}^{\dagger}_{\omega',\ell'}])$$

$$= \frac{1}{\omega|D^0(\omega,\ell)|^2}\delta_{\omega,\omega'}\delta_{\ell,\ell'}. \qquad (4.11)$$

Near the horizon, it is convenient to write

$$V_{\text{hor}}(r_*) = -\left[\frac{(d-3)(1-d)}{4}\frac{f(r)^2}{r^2} - m^2 f(r) - l(l+d-2)\frac{f(r)}{r^2} + \frac{1-d}{2}\frac{f'(r)f(r)}{r}\right],$$

and write the differential equation for the radial mode as

$$\frac{\partial \zeta^{\pm}_{\omega,\ell}(r_*)}{\partial r_*^2} + \omega^2\zeta^{\pm}_{\omega,\ell}(r_*) = V_{\text{hor}}(r_*)\zeta^{\pm}_{\omega,\ell}(r_*), \qquad (4.12)$$

where the superscript $\pm$ distinguishes the two independent solutions of this equation.

We now construct $\zeta^+_{\omega,\ell}(r_*)$, which corresponds to the right-moving mode. The construction of the left-movers is similar. In the absence of the potential, the Eq. (4.12) is solved by

$$\zeta^{+0}_{\omega,\ell}(r_*) = e^{i\omega r_*},$$

and $\zeta^+_{\omega,\ell}(r_*)$ is the completion of this solution.

The Green's function for (4.12) can be written as

$$G_{\text{hor}}(r'_*, r_*) = \frac{1}{\omega}\sin(\omega(r_* - r'_*))\theta(r_* - r'_*).$$

We then again write an iterative solution for this wave function as

$$\zeta^{+,(n)}_{\omega,\ell}(r_*) = \int_{-\infty}^{r_*} \zeta^{+,(n-1)}_{\omega,\ell}(r'_*)V_{\text{hor}}(r'_*)G_{\text{hor}}(r'_*, r_*)dr'_*,$$

with the full solution being given by

$$\zeta^+_{\omega,\ell}(r_*) = \sum_{n=0}^{\infty}\zeta^{+,(n)}_{\omega,\ell}(r_*).$$

Now, we note that the potential can be expanded as

$$V_{\text{hor}}(r_*) = \sum_{m>0}V_m e^{\frac{4\pi m r_*}{\beta}}. \qquad (4.13)$$

To see this, it is most convenient to switch to the Kruskal coordinates. The potential only depends on the product $U_K V_K$ and not on the ratio V_K/U_K. Near the horizon, we can explicitly derive an expansion of the form (4.13). We now expect this horizon to converge till the nearest pole of V_{hor}. The closest poles of V_{hor} occur at the singularity where $|U_K V_K| = e^{\frac{4\pi r_{*s}}{\beta}}$, and we expect this series expansion to be valid till this point.

In the region where (4.13) is valid, we can easily determine the analytic form of $\zeta^{+,(n-1)}_{\omega,\ell}(r_*)$. For example, $\zeta^{+,(1)}_{\omega,\ell}(r_*)$ is simply given by

$$\zeta_{\omega,\ell}^{+,(1)}(r_*) = \frac{1}{2i\omega}\int_{-\infty}^{r_*}dr'_* e^{i\omega r'_*}(e^{i\omega(r_*-r'_*)} - e^{i\omega(r'_*-r_*)})V_m e^{\frac{4\pi m r'_*}{\beta}}$$

$$= \frac{e^{i\omega r_*}}{2i\omega}\sum_{m>0}V_m e^{\frac{4\pi m r_*}{\beta}}\left(\frac{\beta}{4\pi m} - \frac{1}{2i\omega + \frac{4\pi m}{\beta}}\right)$$

$$= \sum_m \frac{\beta V_m e^{(i\omega + \frac{4\pi m}{\beta})r_*}}{4\pi m(2i\omega + \frac{4\pi m}{\beta})}.$$

Thus, we only get poles at $\omega = \frac{2\pi i m}{\beta}$ for m a positive integer. We can now easily prove this by induction for all m. Let us say that we have proved that $\zeta_{\omega,\ell}^{+,(n)}(r_*) = \sum D_m^{+,(n)}(\omega,\ell)e^{(i\omega + \frac{4\pi m}{\beta})r_*}$, where $D_m^{+,(n)}(\omega,\ell)$ has poles only for $\omega = \frac{2\pi i m}{\beta}$. Then we see that

$$\zeta_{\omega,\ell}^{+,(n+1)}(r_*) = \sum_{m,m'}D_m^{+,(n)}(\omega,\ell)\frac{\beta V_{m'} e^{(i\omega + \frac{4\pi m}{\beta} + \frac{4\pi m'}{\beta})r_*}}{4\pi(m+m')(2i\omega + \frac{4\pi(m+m')}{\beta})},$$

and therefore, it also has poles only for $\omega = \frac{2\pi i q}{\beta}$, where q is a positive integer.

The conclusion is that the full solution $\zeta_{\omega,\ell}^+(r_*)$ can be written as

$$\zeta_{\omega,\ell}^+(r_*) = e^{i\omega r_*}D_{\omega,\ell}^+(r_*),$$

where the function $D_{\omega,\ell}^+(r_*)$ has poles only when $\frac{-i\omega\beta}{2\pi}$ is a positive integer.

We can obtain another solution $\zeta_{\omega,\ell}^-(r_*)$ by repeating the procedure above, through

$$\zeta_{\omega,\ell}^{-,0}(r_*) = e^{-i\omega r_*};$$

$$\zeta_{\omega,\ell}^{-,(n)}(r_*) = \int_{-\infty}^{r_*}\zeta_{\omega,\ell}^{-,(n-1)}(r'_*)V_{\text{hor}}(r'_*)G_{\text{hor}}(r'_*,r_*)dr'_*;$$

$$\zeta_{\omega,\ell}^-(r_*) = \sum_{n=0}^{\infty}\zeta_{\omega,\ell}^{-,(n)}(r_*).$$

This is simply related to the solution above through $\zeta_{\omega,\ell}^-(r_*) = (\zeta_{\omega,\ell}^+(r_*))^*$, and the mode $\zeta_{\omega,\ell}^-$ only has poles when $\frac{i\omega\beta}{2\pi}$ is a positive integer.

Transfer function using horizon modes We can now write a transfer function using the modes $\zeta_{\omega,\ell}^\pm(r_*)$ above. They are related quite simply to the mode $\zeta_{\omega,\ell}(r_*)$ that we introduced above through[5]

$$\zeta_{\omega,\ell}(r_*) = \zeta_{\omega,\ell}^+(r_*) + e^{2i\delta_{\omega,\ell}}\zeta_{\omega,\ell}^-(r_*).$$

---

[5]The careful reader may note that, in terms of the factors of $e^{i\delta_{\omega,\ell}}$, we are using a different convention from [4,5,13]. This is because we want to normalize the coefficient of the "right-moving" wave function to 1. As a result, our treatment of the left and right movers is not symmetric.

The field outside the horizon can therefore be written as

$$\phi(t,r_*,\Omega) = \phi_l(r_*,t,\Omega) + \phi_r(r_*,t,\Omega),$$

where

$$\phi_r(t,r_*,\Omega) = \sum_{\ell,\omega}\frac{a_{\omega,\ell}}{\sqrt{\omega}}\zeta_{\omega,\ell}^+(r_*)e^{-i\omega t}Y_\ell(\Omega) + \text{H.c.},$$

$$\phi_l(t,r_*,\Omega) = \sum_{\ell,\omega}\frac{a_{\omega,\ell}}{\sqrt{\omega}}e^{2i\delta_{\omega,\ell}}\zeta_{\omega,\ell}^-(r_*)e^{-i\omega t}Y_\ell(\Omega) + \text{H.c.}$$

$$(4.14)$$

We would like to make two important comments. First, while the factor of $e^{2i\delta_{\omega,\ell}}$ can be fixed by comparing the solutions above to $\zeta_{\omega,\ell}(r_*)$, it can also be fixed, simply by requiring the normalizability condition

$$\lim_{r_*\to 0}[\phi_l(t,r_*,\Omega) + \phi_r(t,r_*,\Omega)] = 0.$$

Second, note that the functions $\zeta_{\omega,\ell}^+(r_*)$ and $\zeta_{\omega,\ell}^-(r_*)$ individually have different analytic properties from $\zeta_{\omega,\ell}(r_*)$. While the analytic properties of $\zeta_{\omega,\ell}^+(r_*)$ will also important below, the reader should keep in mind that the analytic properties of the full wave function outside are controlled by those of $\zeta_{\omega,\ell}(r_*)$ and not by the components $\zeta_{\omega,\ell}^\pm(r_*)$.

C. Crossing the horizon

Finally, we turn to the construction of the bulk field behind the horizon.

As we reviewed in detail in [5,12], behind the horizon, the field again consists of left and right movers. By continuity, the left movers from outside the horizon cross over smoothly to the black hole interior. More precisely, we can write the field as

$$\phi(t,r_*,\Omega) = \widetilde{\phi}_l(r_*,t,\Omega) + \widetilde{\phi}_r(r_*,t,\Omega), \quad \text{behind the horizon,}$$

where, the left and right moving elements are

$$\widetilde{\phi}_l(t,r_*,\Omega) = \sum_{\omega,\ell}\frac{a_{\omega,\ell}}{\sqrt{\omega}}e^{2i\delta_{\omega,\ell}}e^{-i\omega t}Y_\ell(\Omega)\widetilde{\zeta}_{\omega,\ell}^-(r_*) + \text{H.c.},$$

$$\widetilde{\phi}_r(t,r_*,\Omega) = \sum_{\omega,\ell}\frac{\widetilde{a}_{\omega,\ell}}{\sqrt{\omega}}e^{i\omega t}Y_\ell^*(\Omega)\widetilde{\zeta}_{\omega,\ell}^-(r_*) + \text{H.c.}$$

Note that the operator modes $a_{\omega,\ell}$ that appear in $\widetilde{\phi}_l(t,r_*,\Omega)$ are the same as those that appear outside the horizon. This is because the left-moving part of the field is continuous across the horizon. Given the operators defined in (3.12), these operators are defined through

$$\widetilde{a}_{\omega,\ell} = (D^0(\omega,\ell))^*\sqrt{\omega}\widetilde{\mathcal{O}}_{\omega,\ell}.$$

The $\widetilde{\mathcal{O}}_{\omega,\ell}$ operators are defined and discussed in detail in [4,5,12,13].

The functions $\widetilde{\zeta}^- = (\widetilde{\zeta}^+)^*$ that appear in the field above have the near-horizon expansion from behind the horizon

$$\widetilde{\zeta}^-_{\omega,\ell}(r_*) = (\widetilde{\zeta}^+_{\omega,\ell})^*(r_*) \underset{r_* \to -\infty}{\to} e^{-i\omega r_*}.$$

We can analyze these modes using the same techniques as the previous section and the same Green's functions. This is because the "free wave equation" both in front and behind the horizon is the same. So, the starting point of the iteration is just given by the functions $e^{\pm i\omega r_*}$ and the Green's function that is used to construct the full solution is the same both in front and behind the horizon.

The potential can again be expanded in a series of the form (4.13). As we mentioned above, this is really an expansion in the variable $(U_K V_K)$; the value of $U_K V_K$ at a point (t, r_*) behind the horizon is simply the negative of its value at a point with the same Schwarzschild coordinates outside the horizon. Since V_{hor} is analytic across $U_K V_K = 0$, we can simply analytically continue the series to find that

$$V_{\text{hor}}(r_*) = \sum_{m>0} (-1)^m V_m e^{\frac{4\pi m r_*}{\beta}}, \qquad \text{behind the horizon,}$$

(4.15)

where the coefficients V_m are the same as those that appear in (4.15) but we have an additional factor of $(-1)^m$.

From this point, the analysis of the analytic properties of $\widetilde{\zeta}^-$ and $\widetilde{\zeta}^+$ is exactly the same as in the previous subsection. The additional factors of $(-1)^m$ above make no difference to the analytic properties. So we find that

$$\widetilde{\zeta}^\pm_{\omega,\ell}(r_*) = e^{\pm i\omega r_*}\widetilde{D}^\pm_{\omega,\ell}(r_*),$$

where the function $\widetilde{D}^-_{\omega,\ell}(r_*)$ has poles only when $\frac{i\omega\beta}{2\pi}$ is a positive integer and the function $\widetilde{D}^+_{\omega,\ell}(r_*)$ only has poles when $\frac{-i\omega\beta}{2\pi}$ is a positive integer.

The key property that we will need here is the following. Consider a correlator involving an insertion of a right moving part of the field behind the horizon. Then using the definition of $\widetilde{a}_{\omega,\ell}$ and the results above, we find that

$$\langle\Psi|\widetilde{\phi}_r(t, r_*, \Omega)A_\alpha|\Psi\rangle = \langle\Psi|A_\alpha\widetilde{\phi}_r(t, r_*, \Omega)|\Psi\rangle$$
$$= \langle\Psi|A_\alpha e^{-\frac{\beta H}{2}}\widehat{\phi}(t, r_*, \Omega)e^{\frac{\beta H}{2}}|\Psi\rangle,$$

(4.16)

where

$$\widehat{\phi}(t, r_*, \Omega) = \sum_{\omega,\ell} \frac{a_{\omega,\ell}}{\sqrt{\omega}}\widetilde{\zeta}^+_{\omega,\ell}(r_*)e^{-i\omega t}Y_\ell(\Omega) + \text{H.c.}$$

Note that $\widehat{\phi}$ is made up entirely of ordinary operators. To derive this we simply used the definition of the mirror operators in (3.12) to first commute the mirror field through the ordinary field and then substitute the modes with a

conjugated version of the ordinary modes. Note that whenever we use the field $\widehat{\phi}(t, r_*, \Omega)$ it is understood that the coordinates are localized in $\vee_{\mathcal{C}}$.

D. Commutators

We are now in a position to prove an important intermediate technical result. If $U^{\mathcal{C}}$ is a unitary made out of simple operators localized in $\mathcal{B}_{\mathcal{C}}$ and $(t_1, r_{*1}, \Omega_1) \in \vee_{\mathcal{C}}$ then inside any correlator

$$U^{\mathcal{C}\dagger}\widehat{\phi}(t_1, r_{*1}, \Omega_1)U^{\mathcal{C}} = \widehat{\phi}(t_1, r_{*1}, \Omega_1).$$

This can be proved by showing that for any point $(t_2, \Omega_2) \in \mathcal{B}_{\mathcal{C}}$, we have

$$[\widehat{\phi}(t_1, r_{*1}, \Omega_1), \mathcal{O}(t_2, \Omega_2)] = 0. \qquad (4.17)$$

To see (4.17), we use the field expansions (4.14) and (4.9) and the mode commutators to find that the commutator of two field operators can be written as

$$[\widehat{\phi}(t_1, r_{*1}, \Omega_1), O(t_2, \Omega_2)]$$
$$= \lim_{r_{*2}\to 0}(r_{*2})^{-\Delta}[\widehat{\phi}(t_1, r_{*1}, \Omega_1), \phi(t_2, r_{*2}, \Omega_2)]$$
$$= \lim_{r_{*2}\to 0}(r_{*2})^{-\Delta}\sum_{\ell,\omega} e^{-i\omega(t_1-t_2)}Y_\ell(\Omega_1)$$
$$\times Y_\ell^*(\Omega_2)\omega^{-1}\widetilde{\zeta}^+_{\omega,\ell}(r_{*1})\zeta_{\omega,\ell}(r_{*2}) - \text{H.c.}$$

As a consequence of the definitions of $\zeta_{\omega,\ell}$ and $\widetilde{\zeta}^+_{\omega,\ell}$, we can substitute

$$\zeta_{-\omega,\ell}(r_*) \equiv (\zeta_{\omega,\ell}(r_*))^*, \quad \widetilde{\zeta}^+_{-\omega,\ell}(r_*) \equiv (\widetilde{\zeta}^+_{\omega,\ell}(r_*))^*. \quad (4.18)$$

Second, recall that the spherical harmonic can be conjugated by reversing some of the angular momentum quantum numbers, which we denote by $Y_{-\ell}(\Omega) = Y_\ell(\Omega)^*$. We also note that

$$\zeta_{\omega,-\ell}(r_*) = \zeta_{\omega,\ell}(r_*), \qquad \widetilde{\zeta}^+_{\omega,-\ell}(r_*) = \widetilde{\zeta}^+_{\omega,\ell}(r_*). \quad (4.19)$$

Using (4.18) and (4.19) and by converting the sum over ω to an integral, we obtain an integral that runs over $\omega \in (-\infty, \infty)$,

$$[\widehat{\phi}(t_1, r_{*1}, \Omega_1), O(t_2, \Omega_2)]$$
$$= \lim_{r_{*2}\to 0} r_{*2}^{-\Delta}\sum_\ell Y_\ell(\Omega_1)Y_\ell^*(\Omega_2)$$
$$\times \int \frac{d\omega}{\omega}\widetilde{\zeta}^+_{\omega,\ell}(r_{*1})\zeta_{-\omega,\ell}(r_{*2})e^{-i\omega(t_1-t_2)}$$
$$= \sum_\ell Y_\ell(\Omega_1)Y_\ell^*(\Omega_2)\int \frac{d\omega}{\omega}e^{-i\omega(t_1-t_2-r_{*1})}\frac{\widetilde{D}^+_{\omega,\ell}(r_*)}{D^0(-\omega,\ell)}.$$

The integral around $\omega = 0$ should be understood in the sense of a principal value; alternately, the reader may simply wish to differentiate with respect to t_1 or t_2 to make the integrand regular there. In either case, we can neglect the potential pole there. Now we recall that $\widetilde{D}^{+}_{\omega,\ell}(r_*)$ has singularities only when ω is in the upper-half plane. $[D^0(\omega, \ell)]^{-1}$ has poles at the quasinormal mode frequencies when ω is in the lower-half plane, and therefore $[D^0(-\omega, \ell)]^{-1}$, which appears above has poles only in the upper-half plane.

Moreover, with the points localized so that $(t_1, r_{*1}, \Omega_1) \in \vee_{\mathcal{C}}$ and $(t_2, \Omega_2) \in \mathcal{B}_{\mathcal{C}}$, we also have

$$t_1 - t_2 - r_{*1} \geq 0.$$

So the integrand is regular in the lower half ω plane, and we can complete the contour through that region to prove (4.17).

E. Stability of correlators

We are now in a position to analyze and prove the stability of position space AdS correlators under perturbations in the causal patch. Consider a correlator involving the mirror operators in a state that has been excited by a unitary $U^{\mathcal{C}}$ localized in the causal patch and another insertion of some local operators that we denote by A_α below. By the definition of the mirror operators, we have

$$\langle \Psi | U^{\mathcal{C}^\dagger} \widetilde{\phi}_r(t, r_*, \Omega) A_\alpha U^{\mathcal{C}} | \Psi \rangle$$
$$= \langle \Psi | U^{\mathcal{C}^\dagger} A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) U^{\mathcal{C}} | \Psi \rangle$$
$$= \langle \Psi | U^{\mathcal{C}^\dagger} A_\alpha U^{\mathcal{C}} e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} | \Psi \rangle. \quad (4.20)$$

To analyze this correlator, we first note that

$$e^{-\frac{\beta H}{2}} U e^{\frac{\beta H}{2}} | \Psi \rangle = U | \Psi \rangle + \mathrm{O}(\beta \delta E), \quad (4.21)$$

where δE is the change in the expectation value of the Hamiltonian under the action of U that was defined in (2.2). As the notation indicates, this holds whether or not U is localized on $\mathcal{B}_{\mathcal{C}}$. Although this is an intuitive relation, we prove this relation in Appendix B.

Since the relation above is an equality on states, using (2.1), we see that for any coarse grained operator A_γ

$$\mathrm{Tr}(e^{-\beta H} A_\gamma U) = Z(\beta) \langle \Psi | A_\gamma U | \Psi \rangle$$
$$= Z(\beta) \langle \Psi | A_\gamma e^{-\frac{\beta H}{2}} U e^{\frac{\beta H}{2}} | \Psi \rangle + \mathrm{O}(\beta \delta E)$$
$$= \mathrm{Tr}(e^{-\frac{\beta H}{2}} A_\gamma e^{-\frac{\beta H}{2}} U) + \mathrm{O}(\beta \delta E),$$

where we have used the cyclicity of the trace in the last step.

Turning to the correlator in (4.20), and dropping terms of $\mathrm{O}(\beta \delta E)$, we see that

$$Z(\beta) \langle \Psi | U^{\mathcal{C}} A_\alpha U^{\mathcal{C}^\dagger} e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} | \Psi \rangle$$
$$= \mathrm{Tr}(e^{-\frac{\beta H}{2}} U^{\mathcal{C}} A_\alpha U^{\mathcal{C}^\dagger} e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega))$$
$$= \mathrm{Tr}(e^{-\beta H} \widehat{\phi}(t, r_*, \Omega) e^{-\frac{\beta H}{2}} U^{\mathcal{C}} A_\alpha U^{\mathcal{C}^\dagger} e^{\frac{\beta H}{2}})$$
$$= \mathrm{Tr}(e^{-\beta H} \widehat{\phi}(t, r_*, \Omega) e^{-\frac{\beta H}{2}} U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} U^{\mathcal{C}^\dagger} e^{\beta H})$$
$$= \mathrm{Tr}(e^{-\frac{\beta H}{2}} U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} U^{\mathcal{C}^\dagger} \widehat{\phi}(t, r_*, \Omega))$$
$$= \mathrm{Tr}(e^{-\beta H} U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) U^{\mathcal{C}^\dagger} e^{\frac{\beta H}{2}})$$
$$= \mathrm{Tr}(e^{-\beta H} U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} U^{\mathcal{C}^\dagger}).$$

In this sequence of manipulations, we have repeatedly used (4.21), the cyclicity of the trace and (4.17). The end result means that

$$\langle \Psi | U^{\mathcal{C}} A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) U^{\mathcal{C}^\dagger} | \Psi \rangle$$
$$= \langle \Psi | U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} U^{\mathcal{C}^\dagger} | \Psi \rangle + \mathrm{O}(\beta \delta E).$$

In particular, the *change* in the correlator under the excitation $U^{\mathcal{C}^\dagger}$ is given by

$$\langle \Psi | U^{\mathcal{C}} A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) U^{\mathcal{C}^\dagger} | \Psi \rangle - \langle \Psi | A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) | \Psi \rangle$$
$$= \langle \Psi | U^{\mathcal{C}} A_\alpha e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} U^{\mathcal{C}^\dagger} | \Psi \rangle$$
$$- \langle \Psi | A_\alpha e^{-\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{\frac{\beta H}{2}} | \Psi \rangle + \mathrm{O}(\beta \delta E). \quad (4.22)$$

It is possible to derive a similar relation by moving the mirror operators to the left and replacing them with a hatted operator

$$\langle \Psi | U^{\mathcal{C}} A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) U^{\mathcal{C}^\dagger} | \Psi \rangle - \langle \Psi | A_\alpha \widetilde{\phi}_r(t, r_*, \Omega) | \Psi \rangle$$
$$= \langle \Psi | U^{\mathcal{C}} e^{\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{-\frac{\beta H}{2}} A_\alpha U^{\mathcal{C}^\dagger} | \Psi \rangle$$
$$- \langle \Psi | e^{\frac{\beta H}{2}} \widehat{\phi}(t, r_*, \Omega) e^{-\frac{\beta H}{2}} A_\alpha | \Psi \rangle + \mathrm{O}(\beta \delta E). \quad (4.23)$$

But the right-hand sides of (4.22) and (4.23) represent the change in ordinary (i.e., state independent) operators. These obviously obey the constraints of (2.5). We now show using these constraints that the original correlator involving the mirrors also obeys these constraints. The last factor we need to check is the "deviation" that appears in (2.5).

To obtain a finite deviation, we need to smear the position space operator. We define

$$\widehat{\phi}(g) = \int \widehat{\phi}(t, r_*, \Omega) g(t, r_*, \Omega) dt \, dr_* \, d\Omega;$$
$$\widehat{\phi}^\dagger(g) = \int \widehat{\phi}(t, r_*, \Omega) g^*(t, r_*, \Omega) dt \, dr_* \, d\Omega,$$

where g is some function.

If we replace the operator on the right-hand side of (4.22) with its smeared version, then its deviation is not larger than $\widehat{\sigma}_1$ where

$$(\widehat{\sigma}_1)^2 = \frac{1}{Z(\beta)} \text{Tr}(\widehat{\phi}^{\dagger}(g) e^{-\frac{\beta H}{2}} A_{\alpha}^{\dagger} A_{\alpha} e^{\frac{-\beta H}{2}} \widehat{\phi}(g))$$
$$- \left| \frac{1}{Z(\beta)} \text{Tr}(e^{-\frac{\beta H}{2}} A_{\alpha} e^{-\frac{\beta H}{2}} \widehat{\phi}(g)) \right|^2.$$

The deviation could be smaller than this due to the other possible ordering in (2.6).

Similarly, by replacing the operator on the right-hand side of (4.23) with its smeared version, we see that its deviation is not larger than $\widehat{\sigma}_2$, where

$$(\widehat{\sigma}_2)^2 = \frac{1}{Z(\beta)} \text{Tr}(\widehat{\phi}(g) e^{-\frac{\beta H}{2}} A_{\alpha} A_{\alpha}^{\dagger} e^{-\frac{\beta H}{2}} \widehat{\phi}^{\dagger}(g))$$
$$- \left| \frac{1}{Z(\beta)} \text{Tr}(e^{-\frac{\beta H}{2}} A_{\alpha} e^{-\frac{\beta H}{2}} \widehat{\phi}(g)) \right|^2.$$

The minimum of $\widehat{\sigma}_1$ and $\widehat{\sigma}_2$ matches precisely with the deviation of the operator $A_{\alpha} \widetilde{\phi}_r(g)$. That deviation is given by $\min(\widetilde{\sigma}_1, \widetilde{\sigma}_2)$, where

$$(\widetilde{\sigma}_1)^2 = \langle \Psi | \widetilde{\phi}_r^{\dagger}(g) A_{\alpha}^{\dagger} A_{\alpha} \widetilde{\phi}_r(g) | \Psi \rangle - |\langle \Psi | A_{\alpha} \widetilde{\phi}_r(g) | \Psi \rangle|^2$$
$$= (\widehat{\sigma}_1)^2,$$
$$(\widetilde{\sigma}_2)^2 = \langle \Psi | A_{\alpha} \widetilde{\phi}_r(g) \widetilde{\phi}_r^{\dagger}(g) A_{\alpha}^{\dagger} | \Psi \rangle - |\langle \Psi | A_{\alpha} \widetilde{\phi}_r(g) | \Psi \rangle|^2$$
$$= (\widehat{\sigma}_2)^2.$$

The relations above follow immediately by applying (4.16).

Now applying (2.5) to the right-hand sides of (4.23) and (4.22), we see that as $\beta \delta E \to 0$, we have

$$\langle \Psi | U^{\mathcal{C}} A_{\alpha} \widetilde{\phi}_r(t, r_*, \Omega) U^{\mathcal{C}\dagger} | \Psi \rangle - \langle \Psi | A_{\alpha} \widetilde{\phi}_r(t, r_*, \Omega) | \Psi \rangle$$
$$\leq 2 \min(\widetilde{\sigma}_1, \widetilde{\sigma}_2) \sqrt{\beta \delta E},$$

which is exactly what we need.

To summarize, the conclusion is as follows. If we consider correlators and boundary excitations within a single causal patch, then the correlators involving the tildes can be converted to correlators of ordinary operators both in equilibrium and in nonequilibrium states up to errors that are negligible when the nonequilibrium state has only slightly higher energy than the equilibrium state. Since the correlators of ordinary operators clearly obey the ordinary constraints of statistical mechanics, so do the correlators of the tildes.

This brings us to the end of our somewhat intricate proof of the stability of AdS correlators. in the causal patch when the excitations are also localized on the boundary of the patch. We have therefore proved that correlators about an AdS black hole obey (2.5) and the low-energy paradoxes that one might naively expect go away when we carefully analyze the constraints from causality.

We remind the reader that as an immediate corollary, if we start with an equilibrium state and then turn on arbitrary sources, bulk AdS correlators are stable and (2.5) holds.

V. THE BORN RULE PARADOX AND ITS RESOLUTION

The calculations of Sec. IV are rather involved. So, as an example, we now consider a concrete situation where we can examine a specific low energy excitation and calculate the response of the black hole geometry. This example demonstrates the key aspects of the paradox and our resolution. We will make some simplifications to facilitate the analysis and make it more accessible than the analysis of the previous section; the trade-off is that our results will be less precise.

We will consider the specific paradox described in [8] by Marolf and Polchinski (MP), who argued as follows. We see from the relation (3.12) that

$$\langle \Psi | \widetilde{a}_{\omega',\ell'} a_{\omega,\ell} | \Psi \rangle = \frac{e^{-\frac{\beta\omega}{2}}}{1 - e^{-\beta\omega}} \delta_{\omega\omega'} \delta_{\ell\ell'}.$$

This value of the correlator is a precondition for a smooth horizon. MP now considered the unitary operator $U_{\text{MP}} = e^{i\theta N}$, where $N \approx a_{\omega,\ell} a_{\omega,\ell}^{\dagger}$ is proportional to the number operator for a given mode outside the horizon. Then, $[N, H] \approx 0$, but yet U_{MP} introduces a phase $e^{i\theta}$ in the correlator above because it rotates $a_{\omega,\ell}$ but not $\widetilde{a}_{\omega',\ell'}$. Thus, we seem to have a contradiction with (2.5).

MP termed this paradox a violation of "Born rule" because (2.5) holds only for linear operators. So, MP suggested that to ensure that generic states have smooth horizons, the operators (3.12) would have to be significantly nonlinear, which would create difficulties for the quantum mechanical interpretation of the interior. The other alternative, that MP have advocated in previous papers, was that generic black hole states in AdS/CFT do not have an interior at all, or have a firewall, so that the operators (3.12) are not relevant for generic states.

The results of Sec. IV imply that if we only consider observables that respect the constraints of bulk causality, the violations of the "Born rule" should vanish. To see this directly, we translate the paradox above to position space by considering a unitary operator

$$U_{\text{MP}} = e^{i\theta N};$$
$$N = \int_{\mathcal{B}_{\mathcal{C}}} \mathcal{O}(t_1, \Omega_1) \mathcal{O}(t_2, \Omega_2) G(t_1, \Omega_1)$$
$$\times G^*(t_2, \Omega_2) dt_1 dt_2 d\Omega_1 d\Omega_2,$$

where the integral runs only over $\mathcal{B}_{\mathcal{C}}$ because the excitation is supported only on the boundary of the causal patch. We demand that the function G be sharply peaked in frequency

space; for example, $G(t,\Omega) = |\mathcal{B}_\mathcal{C}|^{\frac{-1}{2}} e^{i\omega_0 t} Y_\ell(\Omega)$, where the normalization ensures that with $\theta = O(1)$, the unitary increases the energy only by a small amount. The specific form of G will not be important below.

In this section, we will check that if we consider near-horizon correlators of bulk fields, $\phi(t, r_*, \Omega)$, in the causal patch, which in particular implies that $t - r_* > t_C$, then such correlators are *not* affected strongly by U_{MP} and change only by order $\sqrt{\beta\delta E}$ as predicted by (2.5). This result, of course, already follows from the general analysis of Sec. IV but here we show it through a simple and direct calculation.

As above, this has two related but distinct implications. The main result it that starting with an equilibrium state, it is impossible for any combination of observers to act with U_{MP} and also observe the strong bulk excitation that would naively be predicted by examining the frequency-space effect of this operation. Second, if we consider a state where the excitation U_{MP} appears spontaneously, and we constrain the infalling observer's worldline with the requirement that at least one other observer should be able to detect the excitation fully on the boundary, and report it to the infalling observer, we find that the infaller finds no violations of (2.5).

Before proceeding, we wish to point out another significant difficulty with the argument of [8]. The excitation U_{MP} cannot be produced in an active sense by simply modifying the Hamiltonian on the boundary as shown in (3.6). The reason for this is that N is a double-trace excitation on the boundary that is *nonlocal* in time. So, "acting" with such an operator would require the observer at one point of time to have the ability to act with operators localized at a different point of time. We might imagine that such an excitation could be produced by coupling the CFT to an external system. But then such a coupling would generate entanglement with that system, and not leave the CFT in a pure state.

Nevertheless, we proceed with U_{MP} in the form above, since this is the unitary operator considered in [8]. It does not affect our resolution, the important aspect in the analysis above is causality—namely that an observer, or army of observers who interact with the boundary till a time t_C can produce boundary excitations supported on the past of t_C but not those supported on the future of t_C.

We define

$$F(\omega, \ell) = \int_{\mathcal{B}_\mathcal{C}} G(t, \Omega) e^{-i\omega t} Y_\ell(\Omega) dt d\Omega;$$

$$F(-\omega, \ell) = \int_{\mathcal{B}_\mathcal{C}} G(t, \Omega) e^{i\omega t} Y_\ell(\Omega) dt d\Omega.$$

Note that we have reversed the sign conventions in the Fourier transform compared to (3.11) to avoid having to carry some minus signs below. We consider the case where $G(t, \Omega)$ is largely made up of positive frequencies and has only a negligible negative frequency component. This means that we approximate

$$G(t, \Omega) \approx \sum_{\ell,\omega} F(\omega, \ell) e^{i\omega t} Y_\ell^*(\Omega),$$

where the sum runs only over positive frequencies as per our summation conventions. Conversely, $G(t,\Omega)^*$ effectively comprises only negative frequencies. Note that since $G(t,\Omega)$ vanishes outside $\mathcal{B}_\mathcal{C}$, then we cannot make $F(-\omega, \ell)$ exactly zero since the Fourier transform of a function with compact support always has both positive and negative frequencies. But even in that case, we can make $\frac{|F(-\omega,\ell)|}{|F(\omega,\ell)|} \sim |\mathcal{B}_\mathcal{C}|^{-1}$ for the typical frequencies on which F has support. This is small enough for our purposes. The second approximation that we will need below is that $F(\omega, \ell)$ is sharply peaked about some particular frequency ω_0 and dies off rapidly as we move away from ω_0. This ensures that N behaves like a number operator for ω_0. We will denote the spread in $F(\omega, \ell)$ by $\delta\omega$.

Within this approximation, we can write

$$N = \sum_{\ell_i, \omega_i} \mathcal{O}_{\omega_1, \ell_1} \mathcal{O}^\dagger_{\omega_2, \ell_2} F(\omega_1, \ell_1) F^*(\omega_2, \ell_2),$$

where we remind the reader that following our summation conventions $\omega_1 > 0$, $\omega_2 > 0$ in the sum above.

First, let us work out the change in the expectation value of the Hamiltonian. To first order, we are interested in

$$[N, H] = \sum_{\ell_i, \omega_i} F(\omega_1, \ell_1) F^*(\omega_2, \ell_2)([\mathcal{O}_{\omega_1, \ell_1}, H] \mathcal{O}^\dagger_{\omega_2, \ell_2}$$
$$+ \mathcal{O}_{\omega_1, \ell_1}[\mathcal{O}^\dagger_{\omega_2, \ell_2}, H])$$
$$= \sum_{\ell_i, \omega_i} F(\omega_1, \ell_1) F^*(\omega_2, \ell_2)(\omega_1 - \omega_2) \mathcal{O}_{\omega_1, \ell_1} \mathcal{O}^\dagger_{\omega_2, \ell_2}.$$

In a typical state, by time-translational invariance,

$$\langle\Psi|\mathcal{O}_{\omega_1, \ell_1} \mathcal{O}^\dagger_{\omega_2, \ell_2}|\Psi\rangle = G_{\omega_1, \ell_1} \delta_{\omega_1, \omega_2} \delta_{\ell_1 \ell_2},$$

where $G_{\omega,\ell}$—the two point function of boundary modes—is defined by the equation above. The expectation value of the commutator above vanishes because when $\omega_1 = \omega_2$, the summand vanishes.

$$\langle\Psi|[N, H]|\Psi\rangle = 0.$$

To work out the second order term, we need to compute

$$[N, [N, H]] = \sum_{\omega_i, \ell_i} F(\omega_1, \ell_1) F(\omega_2, \ell_2)^* (\omega_3 - \omega_4) F(\omega_3, \ell_3)$$
$$\times F^*(\omega_4, \ell_4)[\mathcal{O}_{\omega_1, \ell_1} \mathcal{O}^\dagger_{\omega_2, \ell_2}, \mathcal{O}_{\omega_3, \ell_3} \mathcal{O}^\dagger_{\omega_4, \ell_4}].$$

The commutator evaluates to

$$[\mathcal{O}_{\omega_1,\ell_1}\mathcal{O}^\dagger_{\omega_2,\ell_2},\mathcal{O}_{\omega_3,\ell_3}\mathcal{O}^\dagger_{\omega_4,\ell_4}]$$

$$= \mathcal{O}_{\omega_1,\ell_1}[\mathcal{O}^\dagger_{\omega_2,\ell_2},\mathcal{O}_{\omega_3,\ell_3}]\mathcal{O}^\dagger_{\omega_4,\ell_4}$$

$$+ [\mathcal{O}_{\omega_1,\ell_1},\mathcal{O}_{\omega_3,\ell_3}]\mathcal{O}^\dagger_{\omega_2,\ell_2}\mathcal{O}^\dagger_{\omega_4,\ell_4}$$

$$+ \mathcal{O}_{\omega_3,\ell_3}[\mathcal{O}_{\omega_1,\ell_1},\mathcal{O}^\dagger_{\omega_4,\ell_4}]\mathcal{O}^\dagger_{\omega_2,\ell_2}$$

$$+ \mathcal{O}_{\omega_3,\ell_3}\mathcal{O}_{\omega_1,\ell_1}[\mathcal{O}^\dagger_{\omega_2,\ell_2},\mathcal{O}^\dagger_{\omega_4,\ell_4}]$$

$$= C_{\omega_1,\ell_1}\mathcal{O}_{\omega_3,\ell_3}\mathcal{O}^\dagger_{\omega_2,\ell_2}\delta_{\omega_1,\omega_4}\delta_{\ell_1,\ell_4}$$

$$- C_{\omega_2,\ell_2}\mathcal{O}_{\omega_1,\ell_1}\mathcal{O}^\dagger_{\omega_4,\ell_4}\delta_{\omega_2,\omega_3}\delta_{\ell_2,\ell_3},$$

where we have used

$$[\mathcal{O}_{\omega_1,\ell_1},\mathcal{O}^\dagger_{\omega_2,\ell_2}] \equiv C_{\omega_1,\ell_1}\delta_{\omega_1,\omega_2}\delta_{\ell_1,\ell_2},$$

and all other commutators vanish. The form of $C_{\omega,\ell}$ was given in (4.11), but the precise expression will not be important here. We now find that

$$\langle\Psi|[N,[N,H]]|\Psi\rangle = \sum_{\omega_i,\ell_i}|F(\omega_1,\ell_1)|^2|F(\omega_3,\ell_3)|^2(\omega_3-\omega_1)$$

$$\times [C_{\omega_1,\ell_1}G_{\omega_3,\ell_3} - C_{\omega_3,\ell_3}G_{\omega_1,\ell_1}].$$

The Kubo-Martin-Schwinger (KMS) relation relates the expectation value of the commutator and the two point function through

$$\mathrm{Tr}(e^{-\beta H}\mathcal{O}_{\omega_1,\ell_1}\mathcal{O}^\dagger_{\omega_2,\ell_2}) = \mathrm{Tr}(\mathcal{O}^\dagger_{\omega_2,\ell_2}e^{-\beta H}\mathcal{O}_{\omega_1,\ell_1})$$

$$= e^{\beta\omega_2}\mathrm{Tr}(e^{-\beta H}\mathcal{O}^\dagger_{\omega_2,\ell_2}\mathcal{O}_{\omega_1,\ell_1})$$

$$= e^{\beta\omega_2}\mathrm{Tr}(e^{-\beta H}\mathcal{O}_{\omega_1,\ell_1}\mathcal{O}^\dagger_{\omega_2,\ell_2})$$

$$- Z(\beta)e^{\beta\omega_2}C_{\omega_1,\ell_1}\delta_{\omega_1,\omega_2}.$$

Therefore,

$$G_{\omega,\ell} = \frac{1}{1-e^{-\beta\omega}}C_{\omega,\ell}.$$

This allows us to evaluate

$$\langle\Psi|U^\dagger_{\mathrm{MP}}HU_{\mathrm{MP}}|\Psi\rangle = \langle\Psi|H|\Psi\rangle - i\theta\langle\Psi|[N,H]|\Psi\rangle$$

$$- \frac{\theta^2}{2}\langle\Psi|[N,[N,H]]|\Psi\rangle + \mathrm{O}(\theta^3),$$

and using the relations above, we see that

$$\delta E = \langle\Psi|U^\dagger_{\mathrm{MP}}HU_{\mathrm{MP}}|\Psi\rangle - \langle\Psi|H|\Psi\rangle$$

$$= \sum_{\omega_1,\omega_3}C_{\omega_1,\ell_1}C_{\omega_3,\ell_3}|F(\omega_1,\ell_1)|^2|F(\omega_3,\ell_3)|^2(\omega_3-\omega_1)$$

$$\times \left(\frac{1}{1-e^{-\beta\omega_1}} - \frac{1}{1-e^{-\beta\omega_3}}\right),$$

neglecting terms of $\mathrm{O}(\theta^3)$ and higher.

We see that this change in energy is rather small. In particular, when $F(\omega)$ is a sharply peaked function, with a spread in frequencies given by $(\delta\omega)$, we have $\delta E \propto \beta(\delta\omega)^2$.

As the observable in (2.5), we take the two point function of the field. In the large-N limit, this is the key observable, since all local correlators factorize into products of two point functions. As we explained in the Introduction, and in the previous section, the only two-point function that could be potentially problematic has one point behind the horizon and another point in front of the horizon. This requires entanglement between the degrees of freedom outside and inside, which the unitary operator above is intended to break.

We will work in the near-horizon region in this section, since that is where the effect of the unitary is the strongest. Of course, we emphasize that the proof of stability will work even without these approximations as we showed in Sec. IV, and here we take this physically relevant limit only to simplify the algebra and permit explicit computations.

In the near horizon limit it is convenient to move to Kruskal coordinates (3.3). Then we see from the previous section that in the near-horizon limit the mode functions of Sec. IV simplify greatly and become proportional to $\zeta^+(r_*)e^{-i\omega t} \to |U_K|^{\frac{i\beta\omega}{2\pi}}$; $\zeta^-(r_*)e^{-i\omega t} \to V_K^{\frac{-i\beta\omega}{2\pi}}$. The physics in the near-horizon limit becomes effectively two dimensional and so we consider derivatives of the scalar field to get smooth correlators. The product of two field operators near the horizon but on opposite sides, so that their Kruskal U_K coordinates satisfy $U_1 < 0$ and $U_2 > 0$, can be written as

$$\partial_{U_1}\phi(U_1,V_1,\Omega_1)\partial_{U_2}\phi(U_2,V_2,\Omega_2)$$

$$= \frac{\beta^2}{4\pi^2 U_1 U_2}\sum_{\omega_i,\ell_i}(\omega_1\omega_2)^{\frac{1}{2}}\mathcal{I}_{\omega_1,\omega_2,\ell_1,\ell_2},$$

with

$$\mathcal{I}_{\omega_1,\omega_2,\ell_1,\ell_2} \equiv a_{\omega_1,\ell_1}\widetilde{a}_{\omega_2,\ell_2}Y_{\ell_1}(\Omega_1)Y^*_{\ell_2}(\Omega_2)(-U_1)^{\frac{i\beta\omega_1}{2\pi}}(U_2)^{\frac{-i\beta\omega_2}{2\pi}}$$

$$+ a^\dagger_{\omega_1,\ell_1}\widetilde{a}^\dagger_{\omega_2,\ell_2}(-U_1)^{\frac{-i\beta\omega_1}{2\pi}}(U_2)^{\frac{i\beta\omega_2}{2\pi}}Y^*_{\ell_1}(\Omega_1)Y_{\ell_2}(\Omega_2)$$

$$+ \cdots.$$

Here the $\cdots$ indicate terms with products like $a_{\omega,\ell}\widetilde{a}^\dagger_{\omega',\ell'}$ that have no expectation value and will not pick up one under the action of U_{MP}.

Let us examine the change in the value of this correlator under the rotation above. In principle, this correlator is in danger of changing significantly because the ordinary mode gets rotated by U_{MP}, whereas the mirror mode does not. We need to verify that this effect vanishes within the causal patch. Notice that

$$[N,a_{\omega_1,\ell_1}\widetilde{a}_{\omega_2,\ell_2}] = \sum_{\omega_3}F^*(\omega_1,\ell_1)\frac{-1}{D^0(\omega_1,\ell_1)^*\sqrt{\omega_1}}$$

$$\times F(\omega_3,\ell_3)\mathcal{O}_{\omega_3,\ell_3}\widetilde{a}_{\omega_2,\ell_2},$$

$$[N,a^\dagger_{\omega_1,\ell_1}\widetilde{a}^\dagger_{\omega_2,\ell_2}] = \sum_{\omega_4}F(\omega_1,\ell_1)\frac{1}{D^0(\omega_1,\ell_1)\sqrt{\omega_1}}$$

$$\times F(\omega_4,\ell_4)^*\mathcal{O}^\dagger_{\omega_4,\ell_4}\widetilde{a}^\dagger_{\omega_2,\ell_2}. \qquad (5.1)$$

Here, we have used the commutators that follow from (4.10)

$$[\mathcal{O}_{\omega,\ell}, a^\dagger_{\omega',\ell'}] = \frac{1}{D^0(\omega,\ell)\sqrt{\omega}}\delta_{\omega\omega'}\delta_{\ell\ell'}.$$

Moreover, we see that

$$\langle\Psi|\mathcal{O}_{\omega_3,\ell_3}\widetilde{a}_{\omega_2,\ell_2}|\Psi\rangle = \frac{1}{D^0(\omega_2,\ell_2)\sqrt{\omega_2}}\frac{e^{-\frac{\beta\omega_2}{2}}}{1-e^{-\beta\omega_2}}\delta_{\omega_2,\omega_3}\delta_{\ell_2,\ell_3},$$

$$\langle\Psi|\mathcal{O}^\dagger_{\omega_4,\ell_4}\widetilde{a}^\dagger_{\omega_2,\ell_2}|\Psi\rangle = \frac{1}{D^0(\omega_2,\ell_2)^*\sqrt{\omega_2}}\frac{e^{-\frac{\beta\omega_2}{2}}}{1-e^{-\beta\omega_2}}\delta_{\omega_2,\omega_4}\delta_{\ell_2,\ell_4}.$$

Therefore we see that, to first order in θ, the change in the correlator under the action of U_{MP} is given by

$$\langle\Psi|U_{\text{MP}}\partial_{U_1}\phi(U_1,V_1,\Omega_1)\partial_{U_2}\phi(U_2,V_2,\Omega_2)U^\dagger_{\text{MP}}|\Psi\rangle$$
$$- \langle\Psi|\partial_{U_1}\phi(U_1,V_1,\Omega_1)\partial_{U_2}\phi(U_2,V_2,\Omega_2)|\Psi\rangle$$
$$= \frac{i\beta^2\theta}{4\pi^2 U_1 U_2}\sum_{\omega_1,\ell_1}\frac{Y^*_{\ell_1}(\Omega_1)}{D^0(\omega_1,\ell_1)}F(\omega_1,\ell_1)(-U_1)^{\frac{-i\beta\omega_1}{2\pi}}$$
$$\times \sum_{\omega_2,\ell_2}\frac{Y_{\ell_2}(\Omega_2)}{D^0(\omega_2,\ell_2)^*}F^*(\omega_2,\ell_2)(U_2)^{\frac{i\beta\omega_2}{2\pi}}\frac{e^{-\frac{\beta\omega_2}{2}}}{1-e^{-\beta\omega_2}} - \text{H.c.}$$

The relative minus sign between the term and its Hermitian conjugate comes from the relative minus sign in the first and second line of (5.1).

Now, we recall from the previous section that $D^0(\omega,\ell)$ has zeroes only for $\text{Im}(\omega) < 0$ corresponding to the quasinormal frequencies of the black hole and is also bounded at large positive $\text{Im}(\omega)$. As a result, we can write

$$\frac{1}{D^0(\omega,\ell)} = \int_{x>0} d(x,\ell)e^{ix\omega}dx,$$

for some $d(x,\ell)$. Using this, the first sum, over ω_1, above is easy to deal with.

$$\sum_{\omega_1}\frac{1}{D^0(\omega_1,\ell_1)}F(\omega_1,\ell_1)(-U_1)^{\frac{-i\beta\omega_1}{2\pi}}$$
$$= \int_{x>0}d(x,\ell_1)\sum_{\omega_1}F(\omega_1,\ell_1)(-U_1)^{-\frac{i\beta\omega_1}{2\pi}}e^{ix\omega_1}$$
$$= \int_{x>0}dxd\Omega d(x,\ell_1)G\left(x-\frac{\beta\ln(-U_1)}{2\pi},\Omega\right)Y_{\ell_1}(\Omega).$$

We see that this term vanishes unless $x - \frac{\beta\ln(-U_1)}{2\pi} < t_\mathcal{C}$. Since $x > 0$, this requires $-U_1 > e^{\frac{-2\pi t_\mathcal{C}}{\beta}}$; therefore, point 1 must be in $\wedge_\mathcal{C}$—the causal wedge of $\mathcal{B}_\mathcal{C}$. This result, in fact, follows from a simple extension of the results obtained in the previous section. Since the fields ϕ_r have analytic properties very similar to $\widehat{\phi}$, the effect of a unitary localized

on $\mathcal{B}_\mathcal{C}$ vanishes on right movers that are localized between the causal wedge and the horizon.

The second sum, over ω_2, is more interesting. To understand this sum, recall that we have assumed that $F(\omega)$ is narrowly peaked about some frequency ω_0 with a spread $\delta\omega$. Now in this sum, we see that since we are in the near-horizon region so that $U_2 \to 0$, the function $U_2^{\frac{i\omega_2\beta}{2\pi}}$ is rapidly varying, but on the other hand, we can substitute the slowly varying $e^{-\beta\omega} \to e^{-\beta\omega_0}$ within the sum while making an error of only $\text{O}(\beta\delta\omega)$. So,

$$\sum_{\omega_2}\frac{1}{D^0(\omega_2,\ell_2)^*}F^*(\omega_2,\ell_2)(U_2)^{\frac{i\beta\omega_2}{2\pi}}\frac{e^{-\frac{\beta\omega_2}{2}}}{1-e^{-\beta\omega_2}}$$
$$= \frac{e^{-\beta\omega_0}}{1-e^{-\beta\omega_0}}\int_{x>0}dxd\Omega\left[d(x,\ell_2)G\left(x-\frac{\beta}{2\pi}\ln(U_2),\Omega\right)\right.$$
$$\left.\times Y_{\ell_2}(\Omega)\right]^* + \text{O}(\beta(\delta\omega)).$$

But now we see that since the second point is also in the causal patch, therefore $U_2 < e^{-\frac{2\pi t_\mathcal{C}}{\beta}}$. Since G has no support for times larger than $t_\mathcal{C}$, the change in the correlator vanishes up to $\text{O}(\beta\delta\omega)$.

This leads us to the central result that we wanted to demonstrate: an observer in the causal patch cannot access the deep region where the effect of the unitary operator U_{MP} is strong if the observer also has time to observer the excitation on the boundary. An immediate corollary is that if an observer, or a set of observers, in an equilibrium state, attempt to act with U_{MP} by turning on a source and then jump into the bulk, they do not observe any anomalously large effect.

We would like to emphasize that it is not as if U_{MP} has no effect on the correlator. As we see above, there is an effect at order $\beta\delta\omega = \sqrt{\beta\delta E}$.[6] This is precisely what is expected from (2.5). So, there is no issue of a "frozen vacuum" [28]: a low energy unitary on the boundary does have a nonzero effect on correlators in the causal patch, but this effect is bounded by the square root of the energy that the unitary injects into the system.

VI. CAUSAL PATCH COMPLEMENTARITY

In previous sections, we have proved that if one considers observables and excitations within a single causal patch, the paradoxes associated with low energy states vanish. As we explained, this has two implications. First, it implies that, within effective field theory, starting with a

---

[6]We remind the reader that in the previous section, we proved that for arbitrary unitary excitations, which include U_{MP}, that the coefficient of this $\text{O}(\sqrt{\beta\delta E})$ term is also what is expected for the change in the correlation function of ordinary operators. But we will not repeat this calculation in this simplified setting.

black hole formed from collapse, no observer can excite the black hole with a local source and then observe such a paradox. But, more interestingly, our results also apply to states that are spontaneously excited. In such states, if we consider a setup where a combination of observers attempts to both confirm the presence of the excitation on the boundary and also observe bulk correlators, then the only bulk correlators accessible to such observers precisely obey the limit (2.5) derived from statistical mechanics.

In this section—which is exploratory in nature and somewhat outside the main line of this paper—we again focus on states that display a spontaneous excitation but the important new element is that we consider observers who jump into the black hole and reach the singularity before the excitation appears on the boundary. More precisely, we consider states of the form

$$|\Psi^{\text{ne}}\rangle = U(\tau)|\Psi\rangle, \qquad (6.1)$$

where $|\Psi\rangle$ is a typical equilibrium state and $U(\tau)$ indicates a unitary that is localized around $\tau > t_{\mathcal{C}}$. In the state (6.1), we now want to consider correlation functions of fields localized at some set of points $(t_i, r_{*i}, \Omega_i) \in \mathcal{C}$.

In such states, it is easy to see that the description of physics in terms of the usual fields ϕ is problematic; the proof of the stability of correlators of ϕ that we provided in Sec. IV breaks down since its assumptions are not met in this situation. In fact, the reader can check, by extending the calculation in Sec. V, that if $\tau > t_{\mathcal{C}}$ then the difference,

$$\langle\Psi^{\text{ne}}|\phi(t_1, r_{*1}, \Omega_1)\dots\phi(t_n, r_{*n}, \Omega_n)|\Psi^{\text{ne}}\rangle$$
$$- \langle\Psi|\phi(t_1, r_{*1}, \Omega_1)\dots\phi(t_n, r_{*n}, \Omega_n)|\psi\rangle,$$

may not obey (2.5) with $\delta E = \langle\Psi|U^\dagger(\tau)HU(\tau)|\Psi\rangle - \langle\Psi|H|\Psi\rangle \ll \frac{1}{\beta}$.

Now, if a superobserver has the ability to prepare arbitrary states in the CFT and can compare the experience of the infaller in the state $|\Psi\rangle$ with the experience of the infaller in the state $U(\tau)|\Psi\rangle$, then such a superobserver would detect that black holes violate the bound (2.5). One possibility, of course, is to simply ignore this problem since it requires the superobserver to either have the ability to prepare the exponentially unlikely states $U(\tau)|\Psi\rangle$, or else have the ability to act with $U(\tau)$ before the time τ. Such a superobserver would also observe a violation of the second law of thermodynamics, and so we may neglect the violation of (2.5) as a relatively minor issue.

However, states where the spontaneous excitation is inside the causal patch are also exponentially unlikely, but we were able to find a consistent description of physics in the causal patch in such states. So it is natural to ask whether we can extend this to states where the excitation is outside the causal patch.

A. The idea

The idea is as follows. For states of the form (6.1), we postulate that the correct mapping between the bulk and the boundary requires us to specify not only the state but also the *causal patch* that we wish to examine in the bulk. We make this precise by describing a new set of bulk observables $\phi_{\mathcal{C}}$, whose relation to the boundary depends on the causal patch $\mathcal{C}$. These observables have the virtue that (2.5) is met in states of the form (6.1) as well. On the other hand, since these operators $\phi_{\mathcal{C}}$ depend on the causal patch, the price that we pay is that we lose the global picture of the black hole geometry that could simultaneously describe all causal patches.

We do not prove, in this paper, that the operators $\phi_{\mathcal{C}}$ are the correct operators for the infalling observer to use, and therefore, we do not prove that this obstruction in the global reconstruction of the geometry exists. But, at the end of this section, we discuss briefly why we might expect such an obstruction to appear in spontaneously excited states.

B. The proposal

To make the idea above precise, first, we define the algebra of simple polynomials localized on $\mathcal{B}_{\mathcal{C}}$.

$$\mathcal{A}^{\mathcal{C}} = \text{span of } \{\mathcal{O}_{\ell_1}(t_1), \mathcal{O}_{\ell_2}(t_2)\mathcal{O}_{\ell_3}(t_3), \dots,$$
$$\mathcal{O}_{\ell_4}(t_4)\mathcal{O}_{\ell_5}(t_5)\cdots\mathcal{O}_{\ell_{\mathcal{D}_m}}(t_{\mathcal{D}_m})\}, \qquad t_i \in \mathcal{B}_{\mathcal{C}}.$$

This is the set of simple polynomials of single trace operators that are localized on $\mathcal{B}_{\mathcal{C}}$. The cutoff $\mathcal{D}_m$ is the same as the cutoff that appeared in (3.5). However, the important difference between $\mathcal{A}^{\mathcal{C}}$ and $\mathcal{A}$ is that the times t_i that appear above are restricted to lie in $\mathcal{B}_{\mathcal{C}}$. The reader will note that this is very similar to the small algebra $\mathcal{A}_{\text{small}}(B)$ that was defined in [24].

Next, we note that to leading order in $\frac{1}{N}$ the full algebra $\mathcal{A}$ can be decomposed into two commuting subalgebras,

$$\mathcal{A} = \mathcal{A}^{\mathcal{C}} \otimes \bar{\mathcal{A}}^{\mathcal{C}}, \qquad [\mathcal{A}^{\mathcal{C}}, \bar{\mathcal{A}}^{\mathcal{C}}] = 0.$$

The simplest way to understand the decomposition is through the bulk geometry. Note that the relation (3.9) gives us an explicit and $1-1$ mapping of all operators in $\mathcal{A}$ to bulk operators outside the horizon. Now, using the bulk equations of motion, all bulk operators can be mapped to the Cauchy slice, ABC, outside the horizon in Fig. 4. The operators in the segment AB all commute with operators in $\mathcal{A}^{\mathcal{C}}$ because AB is spacelike to $\mathcal{B}_{\mathcal{C}}$. On the other hand, the algebra of operators on BC is isomorphic to $\mathcal{A}^{\mathcal{C}}$. This gives us a decomposition of the boundary algebra into two commuting subalgebras. As emphasized in [24], this decomposition holds only to leading order in $\frac{1}{N}$. So, our discussion here is only a leading order discussion.

We caution the reader that if we think of elements of $\mathcal{A}$ as smeared versions of local operators, then the decomposition
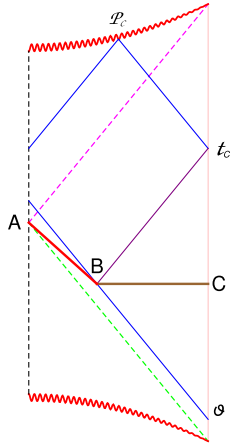
FIG. 4. The slice ABC gives a Cauchy slice for the entire geometry. BC provides complete initial data for the causal wedge $\wedge_C$. In the large-N limit, operators on AB and BC commute.

of an operator into an element of $\mathcal{A}^C$ and $\bar{\mathcal{A}}^C$ does *not* correspond to simply selecting the part of the smearing function with support on $\mathcal{B}_C$. This is because the commutator of single-trace operators is nonzero even for finite time separation on the boundary.

For example, consider a single trace operator $\mathcal{O}(f) = \int O(t, \Omega) f(t, \Omega) dt d\Omega$, where $f(t, \Omega)$ has support both on $\mathcal{B}_C$ and outside $\mathcal{B}_C$. Then this operator can be decomposed into a sum of an element of $\mathcal{A}^C$ and another element of $\bar{\mathcal{A}}^C$ as follows. In Fig. 4, let the coordinates of B be (t_B, r_{*B}). (B represents an entire S^{d-1} sphere.) Consider the time slice t_B that passes through the point B.

Then we can write

$$\mathcal{O}(f) = \int_{r_{*B}}^{0} dr_* \int d\Omega \phi(t_B, r_*, \Omega) g(r_*, \Omega)$$
$$+ \int_{-\infty}^{r_{*B}} dr_* \int d\Omega \phi(t_B, r_*, \Omega) g(r_*, \Omega),$$

which represents the decomposition of this operator as a sum of an operator in $\mathcal{A}^C$ and $\bar{\mathcal{A}}^C$, respectively. Here,

$$g(r_*, \Omega) = -ir^{d-1} \left[\mathcal{O}(f), \frac{\partial \phi}{\partial t} (t_B, r_*, \Omega) \right],$$

which follows just from using the canonical commutation relations for ϕ.

Corresponding to the decomposition of algebras above, we consider a unitary operator that can be decomposed as

$$U(\tau) = U_\tau^C \widehat{V}_\tau^C,$$

where $U_\tau^C \in \mathcal{A}^C$ and $\widehat{V}_\tau^C \in \bar{\mathcal{A}}^C$. We will restrict ourselves to such unitaries in this section. This does not cover all unitaries $U(\tau)$ but it provides a broad class and the proposal below can easily be generalized to other cases. Note that as

we explained above, just because the unitary $U(\tau)$ is largely localized outside $\mathcal{B}_C$ does not mean that $U_\tau^C = 1$.

The unitary U_τ^C is the "part" of the unitary that is visible in the causal patch. This is because

$$\langle \Psi | U(\tau)^\dagger A_\alpha^C U(\tau) | \Psi \rangle = \langle \Psi | U_\tau^{C\dagger} A_\alpha^C U_\tau^C | \Psi \rangle, \quad \forall \, \alpha.$$

So for an observer who conducts experiments on $\mathcal{B}_C$, before jumping into the bulk, it *appears* as if the state has been excited just with U_τ^C.

Now the idea of causal patch complementarity is that in the causal patch C it is consistent to simply neglect $\widehat{V}_\tau^C$. More precisely, we suggest that physics in the state $|\Psi^{ne}\rangle$ in the causal patch C can be described through correlation functions of operators $\phi_C(t, r_*, \Omega)$ that satisfy

$$\langle \Psi^{ne} | \phi_C(t_1, r_{*1}, \Omega_1) \ldots \phi_C(t_n, r_{*n}, \Omega_n) | \Psi^{ne} \rangle$$
$$= \langle \Psi | U_\tau^{C\dagger} \phi(t_1, r_{*1}, \Omega_1) \ldots \phi(t_n, r_{*n}, \Omega_n) U_\tau^C | \Psi \rangle. \quad (6.2)$$

It is possible to solve (6.2) to obtain the action of ϕ_C on states of $\mathcal{H}_\Psi$. Note that the operators ϕ_C are now state dependent even outside the horizon.

If any of the points $(t_i, r_{*i}, \Omega_i) \notin C$, then our proposal has nothing to say about such observables. In such a situation, it may be possible to shift to a different causal patch C' so that all the points fit in that patch and use the field $\phi_{C'}$ to describe these correlators. But, it is important to note that it is impermissible to mix field operators defined with respect to different causal patches. So, the observables

$$\langle \Psi^{ne} | \phi_C(t_1, r_{*1}, \Omega_1) \phi_{C'}(t_2, r_{*2}, \Omega_2) \ldots | \Psi^{ne} \rangle,$$

where C and C' are two different causal patches are meaningless, except in some special cases that we describe below.

Let us mention some properties of this proposal, which also serve as basic consistency checks. First, it is evident that in all equilibrium states, or states with an excitation only on $\mathcal{B}_C$, correlators of ϕ_C coincide with correlators of ϕ. We write this as

$$\phi_C(t, r_*, \Omega) \doteq \phi(t, r_*, \Omega), \text{ in equilibrium states,}$$

where the symbol $\doteq$ implies that correlators of ϕ_C coincide with correlators of ϕ. Second, we note that even in arbitrary nonequilibrium states, the two operators coincide in the causal wedge of $\mathcal{B}_C$.

$$\phi_C(t, r_*, \Omega) \doteq \phi(t, r_*, \Omega), \qquad (t, r_*, \Omega) \in \wedge_C.$$

This is because when the point (t, r_*, Ω) is in $\wedge_C$, the local field $\phi(t, r_*, \Omega) \in \mathcal{A}^C$. So, it automatically commutes with $\widehat{V}_\tau^C$ in (6.2).

Next, we see that as a trivial consequence of this, $\phi_{\mathcal{C}}$ obeys the correct boundary conditions,

$$\lim_{r_* \to \infty} r_*^{-\Delta} \phi_{\mathcal{C}}(t, r_*, \Omega) \doteq O(t, \Omega), \quad \text{for } (t, \Omega) \in \mathcal{B}_{\mathcal{C}}. \quad (6.3)$$

It is also clear from (6.2) that $\phi_{\mathcal{C}}(t, r_*, \Omega)$ also obey the same equations of motion as $\phi(t, r_*, \Omega)$.

$$(\square - m^2)\phi_{\mathcal{C}}(t, r_*, \Omega) \doteq 0. \quad (6.4)$$

Perturbatively, if we add local interactions or sources in the action for ϕ then the equation of motion for $\phi_{\mathcal{C}}$ will also be modified appropriately to match the equation of motion of ϕ.

The other property that is important to note is that the use of the fields $\phi_{\mathcal{C}}$ may preclude a global description of the geometry in states $|\Psi^{ne}\rangle$. If we only want to describe the exterior of the black hole then, as we mentioned in Sec. III, we may use the special causal patch $\mathcal{C}_{\infty}$ that contains only the exterior of the black hole. For this causal patch, the algebra $\mathcal{A}^{\mathcal{C}}$ coincides with $\mathcal{A}$, and the field $\phi_{\mathcal{C}}$ coincides with ϕ.

However no single causal patch accommodates the entire black hole geometry. In an equilibrium state or in an equilibrium state deformed with a source, since $\phi_{\mathcal{C}} \doteq \phi$, it is possible to consider correlation functions across different causal patches. But in a general nonequilibrium state, our proposal does not extend to correlators that do not fit in a single causal patch. In sacrificing a global description, we do not sacrifice any quantity that is observable for a bulk observer since, by construction, correlators that do not fit in any single causal patch cannot be measured by any combination of observers.

C. Stability of causal-patch correlators in spontaneously excited states

An important result is that if we use the operators specified by (6.2), we find that causal patch correlators are stable even in spontaneously excited states of the form (6.1).

We note that the change in the energy induced by the unitary $U(\tau)$ in (6.1) is given by

$$\delta E = \delta E_{\mathcal{C}} + \delta \widehat{E}_{\mathcal{C}}, \quad (6.5)$$

where

$$\delta E = \langle \Psi | U(\tau)^{\dagger} H U(\tau) | \Psi \rangle - \langle \Psi | H | \Psi \rangle,$$

$$\delta E_{\mathcal{C}} = \langle \Psi | U_{\tau}^{\mathcal{C}\dagger} H U_{\tau}^{\mathcal{C}} | \Psi \rangle - \langle \Psi | H | \Psi \rangle,$$

$$\delta \widehat{E}_{\mathcal{C}} = \langle \Psi | \widehat{V}_{\tau}^{\mathcal{C}\dagger} H \widehat{V}_{\tau}^{\mathcal{C}} | \Psi \rangle - \langle \Psi | H | \Psi \rangle.$$

This relation is rather intuitive. The action of unitaries on a typical state always tends to inject a small positive amount of energy. The two unitaries acting together above simply inject energy one after the other. Nevertheless, we check the relation (6.5) explicitly in Appendix B 1.

But we have already proved in Sec. IV that

$$\langle \Psi | U_{\tau}^{\mathcal{C}\dagger} \phi_{\mathcal{C}}(t_1, r_{*1}, \Omega_1) \ldots \phi_{\mathcal{C}}(t_n, r_{*n}, \Omega_n) U_{\tau}^{\mathcal{C}} | \Psi \rangle$$
$$- \langle \Psi | \phi_{\mathcal{C}}(t_1, r_{*1}, \Omega_1) \ldots \phi_{\mathcal{C}}(t_n, r_{*n}, \Omega_n) | \Psi \rangle$$
$$= \langle \Psi | U_{\tau}^{\mathcal{C}\dagger} \phi(t_1, r_{*1}, \Omega_1) \ldots \phi(t_n, r_{*n}, \Omega_n) U_{\tau}^{\mathcal{C}} | \Psi \rangle$$
$$- \langle \Psi | \phi(t_1, r_{*1}, \Omega_1) \ldots \phi(t_n, r_{*n}, \Omega_n) | \Psi \rangle$$
$$\leq 2\sqrt{\beta \delta E_{\mathcal{C}}} \sigma,$$

where σ is the deviation of the product of fields as defined in (2.6). Here we have used the fact that for correlators in an equilibrium state involving insertions only in the causal patch, correlators of $\phi_{\mathcal{C}}$ are the same as correlators of ϕ.

But since $\delta E_{\mathcal{C}} \leq \delta E$, we see using (6.2) that we automatically have

$$\langle \Psi | U(\tau)^{\dagger} \phi_{\mathcal{C}}(t_1, r_{*1}, \Omega_1) \ldots \phi_{\mathcal{C}}(t_n, r_{*n}, \Omega_n) U(\tau) | \Psi \rangle$$
$$- \langle \Psi | \phi_{\mathcal{C}}(t_1, r_{*1}, \Omega_1) \ldots \phi_{\mathcal{C}}(t_n, r_{*n}, \Omega_n) | \Psi \rangle$$
$$\leq 2\sqrt{\beta \delta E} \sigma. \quad (6.6)$$

This proves that if we use the fields $\phi_{\mathcal{C}}$ then the inequality (2.5) is satisfied even in spontaneously excited states, and correlators in the causal patch are stable under low energy excitations. The result (6.6) is the one of the main reasons to believe that the operators (6.2) are the right operators to use in a causal patch.

An example may help to clarify the mechanism by means of which this stability under low-energy excitations is achieved. In the state $|\Psi^{ne}\rangle$, let us consider three causal patches $\mathcal{C}, \mathcal{C}', \mathcal{C}''$, where $t_{\mathcal{C}} \ll t'_{\mathcal{C}} < \tau < t''_{\mathcal{C}}$.

The description using $\phi_{\mathcal{C}''}$ is appropriate for an observer who jumps in after the time τ.[7] This observer sees the full unitary $U(\tau)$ but he can only reach a limited "depth" (in terms of the maximum U_K coordinate) inside the black hole, and so the correlators he measures all obey (2.5). Such an observer may also naively *infer* that the observer who jumped in at time $t_{\mathcal{C}}$ had a strong interaction with a high-energy shock wave.

But the description using the fields $\phi_{\mathcal{C}''}$ is a poor description for the observer who jumps into the bulk around $t_{\mathcal{C}}$ because $\mathcal{C}''$ contains only a very small part of his worldline behind the horizon. In the description provided by the fields $\phi_{\mathcal{C}}$, the state $|\Psi^{ne}\rangle$ is essentially an equilibrium state, and the observer simply falls through the horizon without encountering any particles.

The description using $\phi_{\mathcal{C}'}$ is appropriate for an observer who jumps into the black hole, a little before the excitation appears on the boundary. This observer sees only part of the full unitary. So, the description using $\phi_{\mathcal{C}'}$ is slightly out of

---

[7]Here we are using the "experience of the observers" as shorthand to describe the results of correlators computed using (6.2) for the three causal patches.

equilibrium, but this observer also does not observe any strong excitation at the horizon.

Thus, we see that all three descriptions obey (2.5) but for different reasons. This is similar to the spirit of the original proposal of black hole complementarity [11,29,30] that suggested that we should use different descriptions for the observer who stays outside the black hole and the infalling observer. Here, we have attempted to sharpen this idea by specifying the precise local correlation functions that are appropriate for a specific causal patch.

D. Overlapping patches

We now turn to another aspect of our proposal: a single spacetime point belongs to multiple causal patches: Fig. 5 shows the overlap of two causal patches $\mathcal{C}$ and $\mathcal{C}'$. Now if a point belongs to both patches, we may ask: "what is the correct field to use at such a point? Is it $\phi_\mathcal{C}$ or $\phi_{\mathcal{C}'}$?"

We remind the reader that this question is not relevant in equilibrium states or even in nonequilibrium states where the point $(t, r_*, \Omega) \in \wedge_\mathcal{C} \cap \wedge_{\mathcal{C}'}$; correlators of the fields $\phi_\mathcal{C}$ and $\phi_{\mathcal{C}'}$ coincide in such cases because both of them coincide with ϕ.

But these correlators do differ in a class of states. For example, if we consider a state that has a spontaneous excitation on the boundary in between $t_\mathcal{C}$ and $t'_\mathcal{C}$, and also consider points that are not in the intersection of the causal wedges $\wedge_\mathcal{C} \cap \wedge_{\mathcal{C}'}$ then the correlators given by the two fields may differ. Note that this is not a contradiction: even though notionally both these operators correspond to the "field", $\phi_\mathcal{C}$ and $\phi_{\mathcal{C}'}$ are different operators and can have different correlators. We say that the operators $\phi_\mathcal{C}$ and $\phi_{\mathcal{C}'}$ are giving complementary, but internally self-consistent descriptions of the same physics. So, as long as we consistently use the description provided by $\mathcal{C}$ or consistently use the one provided by $\mathcal{C}'$ both descriptions are valid about the point P. We emphasize that cross-correlators of the form $\langle \Psi^{\mathrm{ne}} | \phi_\mathcal{C}(t_1, r_{*1}, \Omega_1) \phi_{\mathcal{C}'}(t_2, r_{*2}, \Omega_2) | \Psi^{\mathrm{ne}} \rangle$ do *not* have any physical significance in this situation.
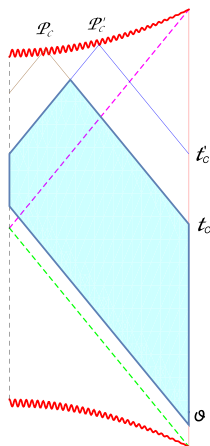

FIG. 5. The overlap of two causal patches.

For example, if we act on the state $|\Psi^{\mathrm{ne}}\rangle$ with an additional unitary so that $U_{\mathrm{obs}}|\Psi^{\mathrm{ne}}\rangle$ can be interpreted as the state plus an observer, then our proposal allows us to reconstruct the bulk geometry including the observer in a causal patch. The description in terms of $\phi_\mathcal{C}$ and the description in terms of $\phi_{\mathcal{C}'}$ give internally consistent but complementary descriptions of the experience of the observer.

What is important for us, is that provided we use each set of correlators in its respective domain of validity then *each* of these descriptions obeys effective field theory because of (6.4), the right relation with boundary fields because of (6.3) and the constraints of statistical mechanics because of (6.6).

The reader may be a little concerned because in empty AdS, or flat-space quantum field theory, we do not usually have multiple possible descriptions of bulk physics. The reason that this ambiguity appears here is as follows. Usually, in quantum field theory, we proceed with the following tacit operational understanding. First, we do a set of experiments to fix the operators that we intend to use. Then in future experiments, we always tacitly refer back to this calibration.

The contrast of the setting that we are considering with empty AdS may help clarify this point. In empty AdS, if a bulk observer observes the boundary for a time band of length larger than π (in units where the light-crossing time of AdS is π) and calibrates his fields to satisfy (6.3) then this completely removes the ambiguity in any of his future measurements. We can make this precise by recalling the transfer function in empty AdS.

$$\phi(t, r, \Omega) = \sum_{n,\ell} \mathcal{O}_{n,\ell} e^{-i(2n+\ell+\Delta)t} Y_\ell(\Omega) \chi^\Omega_{n,\ell}(r) + \mathrm{H.c.},$$

where $\chi^\Omega_{n,\ell}(r)$ is a specific hypergeometric function that the reader can find in Sec. IV of [24]. Now the key point here is that the modes $\mathcal{O}_{n,\ell}$ may be defined entirely in the causal past of the point (t, r, Ω) through

$$\mathcal{O}_{n,\ell} = \int_I dt \int_{S^{d-1}} d^{d-1}\Omega \mathcal{O}(t, \Omega) e^{i(2n+\ell+\Delta)t} Y^*_\ell(\Omega),$$

where I is any interval of length π that can be placed entirely in the past of (t, r, Ω). So, if the bulk observer has ensured that his field variable obeys (6.3) on the interval I then this leaves no ambiguity in his future measurements.

On the other hand, the transfer function (3.9) is not *causal*. Unlike empty AdS, it cannot be written in a form that involves boundary modes only in the past. Therefore, the definition of ϕ in terms of boundary operators is necessarily teleological and not causal in the presence of a black hole. And we see from (6.2) that the fields $\phi_\mathcal{C}$ and ϕ differ only in their response to *future* spontaneous excitations. So, at a point in the bulk from a causal perspective, there is no reason to privilege the description in terms of ϕ over the description in terms of $\phi_\mathcal{C}$ of $\phi_{\mathcal{C}'}$.

Nevertheless, we could still ask: "what will the bulk observer see subjectively?" This is an imprecise question. Often, in situations like this, the question of what the

observer will "see" is answered by simply picking the most convenient description. For example, consider an observer who couples himself to a system that is described by $N = 4$ Yang-Mills theory at large 't Hooft coupling. Then, one description, of course, is that the observer turns into some mixture of strongly coupled glueballs. But usually, we simply declare that the observer will "perceive himself" as falling into AdS_5. Both these descriptions are correct, but the description in terms of gravity is more convenient.

This happens in the case of causal patches as well. For example, consider once again the nonequilibrium state $|\Psi^{ne}\rangle$ with $\tau \gg t_C$. Then the correlators of ϕ_C are equilibrium correlators. On the other hand, if we consider the causal patch C' with $t'_C \sim \tau$ then correlators of the field $\phi_{C'}$ are out of equilibrium since in this patch the effect of the unitary is easier to see. If we consider a freely falling observer jumps into the bulk around t_C and continues onto the singularity, then clearly the description in terms of C is more convenient since it allows us to describe a larger part of the observer's worldline and involves an undisturbed Schwarzschild black hole.

The general question of which description is the "preferred" one in an arbitrary state goes beyond the scope of this paper. However, we mention that there is an interesting sense in which this question may be made well-defined.[8] Given a state $|\Psi^{ne}\rangle$ and a point P can we find the causal patch C_P so that the physics described using the fields ϕ_{C_P} is as "classical" as possible [31] around P? If so, it may be useful to declare that the fields ϕ_{C_P} define the "subjective experience" of the observer and that this description is the "preferred description." However, to explore this question further requires a quantification of the notion of "more classical" and a more precise dynamical description of an observer, and we do not address it in this paper.

VII. CONCLUSION

In this paper, we described and resolved the paradox of low energy excitations about black holes. The paradox arises because a robust property of statistical mechanics appears to be in conflict with properties of the black hole geometry. Statistical mechanics tells us that low energy excitations in thermal systems have small effects on physical observables. This property can be quantified, and we did so in (2.5). A paradox arises in the black hole geometry, because a naive analysis suggests that a locally large distortion of the geometry near the horizon should raise the Arnowitt-Deser-Misner (ADM) energy by only a small amount. Another way to understand this is that if we consider a state that displays an excitation on the boundary at some time, and backcalculate the geometry in the past, then this geometry has a blueshifted excitation near the horizon.

We showed how this paradox was resolved by considerations of causality. To implement this constraint, our analysis in this paper focused on the properties of AdS correlators in position space. These correlators provide insights that are sometimes missed in the usual frequency-mode analysis of these problems. In our case, we found that if we consider a correlator where the excitation and insertions all fit in a single causal patch, then a delicate combination of factors ensures that this correlator obey the standard constraints of statistical mechanics. This is not obvious, *a priori*, since the operators inside the black hole are state dependent. Although we established this by means of a detailed calculation in Sec. IV, it is not difficult to state the final result.

The most common reason that a black hole state may be excited at late times is because we start with an equilibrium state and then turn on a source on the boundary. In such a situation, the past of the state is just an ordinary equilibrium state and so the blueshifted excitation that could have created a problem does not exist.

A far more unlikely but technically more interesting situation arises if we consider a state that displays a spontaneous excitation at some point of time. Now we may consider two observers—one of them jumps into the black hole at an early time, while the other waits near the boundary to confirm that there is indeed a late-time excitation and then jumps into the black hole. The remarkable fact that we pointed out in this paper is as follows. If the first observer follows a trajectory that allows him to meet with the second observer, then he cannot go too "deep" into the black hole. But, on the other hand, the first observer can only detect the blueshifted excitation by freely falling through the horizon for at least some time and probing the geometry in the interior to some depth. It turns out the constraint imposed by the fact that he must meet the second observer precisely prevents him from probing the black hole deep enough to be able to detect the excitation.

These conclusions already resolve the paradox of low energy excitations in the vast majority of states. Starting in an equilibrium state, no combination of observers can excite the state on the boundary and observe a correlator that violates (2.5). But, even in spontaneously excited states, no combination of observers can verify the presence of the excitation and also observe a bulk correlator that shows a violation of (2.5).

Spontaneously excited states are exponentially unlikely. Nevertheless, we may ask what happens in such a state if an observer simply continues to fall freely through the horizon without bothering about meeting the second observer who waits on the boundary to verify the presence of the excitation. In this situation, the unitary operator that produces the excitation does not fit into the causal patch of the first observer. This is the situation that we considered in Sec. VI. Here we proposed that it is consistent to use a description where the first observer only sees the part of the unitary that fits on the boundary of his causal patch and,

therefore, using the results of Sec. IV does not see an anomalously large effect. This description is not consistent with the experience of the second observer. But now since the worldlines of the two observers do not fit into a single patch, the descriptions that we must use for them are complementary and do not have to be reconciled.

In a sense, such a notion of complementarity is inbuilt into AdS/CFT and was used tacitly in [4,5]. This is because, if we wait for an exponentially long time, every pure state will eventually see a spontaneous excitation. If we take the bulk geometry in the presence of this excitation seriously, then the geometry involves an extraordinarily blueshifted shock wave that lurks close to the horizon for a very long time, only to emerge much later, hit the boundary and fall back into the black hole. Evidently, a more sensible perspective is simply to put an early and late time cutoff on the boundary and ignore these excitations that occur beyond these cutoffs. Our proposal here simply makes this more precise and generalizes this notion of complementarity to settings where the spontaneous excitation is just slightly in the future of the causal patch of interest, rather than being exponentially far away.

Although we have now proposed a boundary to bulk map that ensures that a description in terms of fields restricted to a single causal patch is consistent, how do we know that these are the fields that the infalling observer will "observe"? This is also an open question for the construction of [4,5]. The mirror operators provide a consistent description of the interior, but why are they the right operators to use? So, in a strict sense, the results of this paper should be added to the results of [4,5] as an existence proof: it is possible to find a consistent description of the black hole interior that avoids various potential paradoxes and ensures that the horizon is smooth in generic states.

To go further, we may need to construct a dynamical model of the observer and then analyze the full system including the observer. In a rough sense, the situation here is analogous to the situation that arises in quantum field theory in curved spacetime. There it was realized that particle number was an ambiguous concept. However, by constructing a dynamical model of a particle detector, Unruh [32] and DeWitt [33] were able to specify, among the infinity of possible definitions of particle number, which definition would be most convenient for an observer moving along a given trajectory. Here, we are asked to go a step further and specify which "field variable" would be most convenient for an infalling observer to use. This is a very interesting question, and we do not address it in this paper, although we hope to return to it in future work.

There are various technical questions that should be amenable to simple analysis. For example, while we have dealt with pure states here, it would be interesting to understand situations where the CFT is entangled with another system. The only potential complication is the interplay between the superposition of geometries and the

causal structure. However, geometries that have significant overlap differ only by quantum fluctuations, and so have the same causal structure; geometries with different causal structures have an exponentially small overlap and therefore, behave like a classical statistical mixture. So, it should not be difficult to extend these results to entangled states in the spirit of [12].

Next, our analysis is currently only a leading order analysis, and it would be quite interesting to understand how it is modified by $\frac{1}{N}$ corrections. In gravity, these corrections prevent any notion of microcausality, beyond leading order, and so it would be interesting to understand how our analysis here that relies strongly on the light cone structure of the geometry must be generalized. We hope to address this question in future work.

APPENDIX A: MODES IN THE BTZ BLACK HOLE

Here, we explicitly work out the wave functions in the BTZ black hole and verify the claims made in Sec. IV.

The metric of the BTZ black hole is just (3.1), at $d = 2$

$$ds^2 = -(r^2 - r_0^2)dt^2 + (r^2 - r_0^2)^{-1}dr^2 + r^2 d\theta^2.$$

Consider the bulk wave equation $(\Box - m^2)\phi = 0$, and to be consistent with the conventions of Sec. IV, we take the ansatz $\phi_{\omega,\ell} = r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r)e^{-i\omega t}e^{i\ell\theta}$. The equation for χ then becomes

$$\frac{1}{r}\partial_r(r^2 - r_0^2)r\partial_r r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r)$$
$$+ \left[\frac{\omega^2}{r^2 - r_0^2} - \frac{\ell^2}{r^2} - m^2\right]r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r) = 0.$$

It is convenient to work directly with r rather than the tortoise coordinate for the BTZ black hole. Near the boundary, the radial wave equation is solved by a combination of hypergeometric functions

$$r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r) = r^{-\Delta}r_0^{\Delta-\frac{i\omega}{r_0}}\left(1-\frac{r_0^2}{r^2}\right)^{-\frac{i\omega}{2r_0}}$$

$$\times\left[A_1{}_2F_1\left(\frac{\Delta}{2}+i\frac{\ell-\omega}{2r_0},\frac{\Delta}{2}-i\frac{(\ell+\omega)}{2r_0};\Delta;\frac{r_0^2}{r^2}\right)\right.$$

$$+A_2r_0^2r^{2\Delta-2}{}_2F_1\left(1-\frac{\Delta}{2}+i\frac{\ell-\omega}{2r_0},\right.$$

$$\left.\left.1-\frac{\Delta}{2}-i\frac{\ell+\omega}{2r_0};2-\Delta;\frac{r_0^2}{r^2}\right)\right].$$

By imposing normalizability at the boundary, we see that $A_2 = 0$. To ensure that the field scales like $r^{-\Delta}$ near the boundary with no other factors, we set $A_1 = r_0^{\frac{i\omega}{r_0}-\Delta}$.

Next recall the series expansion of the hypergeometric function

$$_2F_1(a,b;c;x) = \sum_{n=0}^{\infty}\frac{(a)_n(b)_n}{(c)_n}\frac{x^n}{n!},$$

where $(a)_n \equiv a(a+1)\ldots(a+n-1)$. From this expansion, we see that, as claimed, the function $\chi_{\omega,\ell}$ has no poles at finite values of ω and for $r_0 < r < \infty$.

By using hypergeometric transformations, we can put the solution in a form appropriate for the near-horizon expansion. We can rewrite

$$r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r) = A_1\Gamma(\Delta)r_0^{\frac{-i\omega}{r_0}}\left(\frac{r}{r_0}\right)^{-\frac{i\ell}{r_0}}\left[\frac{(\frac{r^2}{r_0^2}-1)^{\frac{i\omega}{2r_0}}\Gamma(-\frac{i\omega}{r_0}){}_2F_1(\frac{\Delta}{2}-i\frac{(\ell-\omega)}{2r_0},1-\frac{\Delta}{2}-i\frac{\ell+\omega}{2r_0};\frac{i\omega}{r_0}+1;1-\frac{r^2}{r_0^2})}{\Gamma(\frac{i\ell+r_0\Delta-i\omega}{2r_0})\Gamma(-\frac{i(\ell+ir_0\Delta+\omega)}{2r_0})}\right.$$

$$\left.+\frac{(\frac{r^2}{r_0^2}-1)^{\frac{-i\omega}{2r_0}}\Gamma(\frac{i\omega}{r_0}){}_2F_1(1-\frac{\Delta}{2}-i\frac{\ell+\omega}{2r_0},-\frac{\Delta}{2}-i\frac{(\ell+\omega)}{2r_0};1-\frac{i\omega}{r_0};1-\frac{r^2}{r_0^2})}{\Gamma(-\frac{i(\ell+ir_0\Delta-\omega)}{2r_0})\Gamma(\frac{i(\ell-ir_0\Delta+\omega)}{2r_0})}\right].$$

Recall that in this case, the tortoise coordinate is given simply by

$$r_* = \frac{1}{2r_0}\log\frac{r-r_0}{r+r_0}.$$

Therefore, we see that (4.8) holds, and we can write

$$r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r_*)\underset{r_*\to-\infty}{\longrightarrow}D^0(\omega,\ell)(e^{i\omega r_*}+e^{2i\delta_{\omega,\ell}}e^{-i\omega r_*}),$$

with

$$D^0(\omega,\ell) = \left(\frac{r}{r_0}\right)^{-\frac{i\ell}{r_0}}2^{\frac{i\omega}{r_0}}\frac{\Gamma(\Delta)}{r_0^{\Delta}}\frac{\Gamma(-\frac{i\omega}{r_0})}{\Gamma(\frac{\Delta}{2}-\frac{i\omega}{2r_0}-\frac{i\ell}{2r_0})\Gamma(\frac{\Delta}{2}-\frac{i\omega}{2r_0}+\frac{i\ell}{2r_0})},$$

where we have now substituted the value of A_1 calculated above. The phase factors in (4.8) is just

$$e^{2i\delta_{\omega,\ell}} = \frac{D^0(-\omega)}{D^0(\omega)}.$$

As advertised, the function D^0 has zeroes when the gamma functions in the denominator have poles: $\frac{-i\omega}{2r_0}\pm\frac{i\ell}{2r_0}+\frac{\Delta}{2}\in\mathcal{Z}^-$. Therefore, the zeros are

$$\omega = \pm\ell - i\Delta r_0 - 2inr_0,$$

for n a positive integer. Clearly these occur only for $\text{Im}(\omega) < 0$, and indicate the positions of the quasinormal modes. The function D^0 is regular in the upper half ω plane; using Stirling's approximation, we can check that this is also true as we approach infinity through the upper-half plane. And we also see that for $\omega \in \mathcal{R}$, the phase δ is real.

Finally, defining $\zeta(r) = r^{\frac{1-d}{2}}\chi_{\omega,\ell}(r)(D^0(\omega,\ell))^{-1}$, we can write

$$\zeta(r) = \zeta^+(r) + e^{2i\delta_{\omega,\ell}}\zeta^-(r),$$

where

$$\zeta^+ = \left(\frac{r}{r_0}\right)^{-\frac{i\ell}{r_0}}\left(\frac{r-r_0}{r+r_0}\right)^{\frac{i\omega}{2r_0}}{}_2F_1\left(\frac{\Delta}{2}-i\frac{(\ell-\omega)}{2r_0},\right.$$

$$\left.1-\frac{\Delta}{2}-i\frac{\ell+\omega}{2r_0};\frac{i\omega}{r_0}+1;1-\frac{r^2}{r_0^2}\right)$$

is the function relevant for the "right moving" wave. This is again consistent with our analysis in Sec. IV. Using the expansion of the hypergeometric function, we see that $\zeta^+(r)$ has poles only for $\frac{\omega}{r_0} = \frac{\beta\omega}{2\pi}\in i\mathcal{Z}^+$, which is exactly what we expect. The same function for $r < r_0$ also describes the interior right moving mode; as expected, this has the same analytic properties as the right movers outside the horizon.

APPENDIX B: LOW ENERGY UNITARIES

In this appendix, we prove (4.21). We note that with $V = e^{\frac{-\beta H}{2}} U e^{\frac{\beta H}{2}}$,

$$\langle \Psi | V^\dagger V | \Psi \rangle = \langle \Psi | e^{\frac{\beta H}{2}} U^\dagger e^{-\beta H} U e^{\frac{\beta H}{2}} | \Psi \rangle$$

$$= \frac{1}{Z(\beta)} \text{Tr} \left(e^{-\beta H} e^{\frac{\beta H}{2}} U^\dagger e^{-\beta H} U e^{\frac{\beta H}{2}} \right) = 1.$$

So, we will be done if we can show that

$$\langle \Psi | U^\dagger V | \Psi \rangle = 1 - \text{O}(\beta \delta E). \tag{B1}$$

Our strategy is to expand the unitary as

$$U = e^{i\theta \sum_\omega (A_\omega + A_\omega^\dagger)},$$

where $A_\omega \in \mathcal{A}$ are elements of the algebra with definite energy so that $[H, A_\omega] = -\omega A_\omega$ plus possible $\text{O}(\frac{1}{N})$ terms that we neglect. First, let us work out δE under this unitary. We see that

$$U^\dagger H U = H - i\theta \sum_\omega \omega (A_\omega - A_\omega^\dagger)$$

$$- \frac{\theta^2}{2} \sum_{\omega', \omega} \omega [A_{\omega'} + A_{\omega'}^\dagger, A_\omega - A_\omega^\dagger] + \text{O}(\theta^3).$$

Taking the expectation value in a typical state, we see that

$$\delta E = \langle \Psi | U^\dagger H U | \Psi \rangle - \langle \Psi | H | \Psi \rangle = \theta^2 \sum_\omega \omega C_\omega^A,$$

where $\langle \Psi | [A_\omega, A_{\omega'}^\dagger] | \Psi \rangle \equiv C_\omega^A \delta_{\omega\omega'}$. This appears because in an equilibrium state only expectation values of operators with "zero energy" are nonvanishing, and we neglect terms of order θ^3 and higher.

Now we compare this with (B1). We see that

$$V = e^{-\frac{\beta H}{2}} U e^{\frac{\beta H}{2}}$$

$$= 1 + i\theta \sum_\omega (e^{\frac{\beta\omega}{2}} A_\omega + e^{-\frac{\beta\omega}{2}} A_\omega^\dagger)$$

$$- \frac{\theta^2}{2} \sum_{\omega, \omega'} (e^{\frac{\beta\omega}{2}} A_\omega + e^{-\frac{\beta\omega}{2}} A_\omega^\dagger)(e^{\frac{\beta\omega'}{2}} A_{\omega'} + e^{-\frac{\beta\omega'}{2}} A_{\omega'}^\dagger)$$

$$+ \text{O}(\theta^3),$$

while

$$U^\dagger = 1 - i\theta \sum_\omega (A_\omega + A_\omega^\dagger)$$

$$- \frac{\theta^2}{2} \sum_{\omega, \omega'} (A_\omega + A_\omega^\dagger)(A_{\omega'} + A_{\omega'}^\dagger) + \text{O}(\theta^3).$$

Now we see that to first order in θ

$$\langle \Psi | U^\dagger V | \Psi \rangle = 1 - i\theta \sum_\omega \langle \Psi | (e^{-\beta\omega} - 1) A_\omega + (e^{\frac{\beta\omega}{2}} - 1) A_\omega^\dagger | \Psi \rangle,$$

the first order term vanishes because the only operators with a nonzero expectation value are those with $\omega = 0$, but at that point, the coefficients of these operators in the sum vanish. Moving to the next term, which is second order in θ, we see that

$$\langle \Psi | U^\dagger V | \Psi \rangle = 1 + \theta^2 \sum_\omega [G_\omega^A (e^{-\frac{\beta\omega}{2}} - 1) + (G_\omega^A - C_\omega^A)(e^{\frac{\beta\omega}{2}} - 1)],$$

where $\langle \Psi | A_\omega A_{\omega'}^\dagger | \Psi \rangle = \langle \Psi | A_\omega^\dagger A_{\omega'} | \Psi \rangle + C_\omega^A \equiv G_\omega^A \delta_{\omega, \omega'}$. We can use the KMS relations to convert this to a form that is similar to δE. We note that, first

$$\text{Tr}(e^{-\beta H} A_\omega^\dagger A_\omega) = e^{-\beta\omega} \text{Tr}(e^{-\beta H} A_\omega A_\omega^\dagger),$$

by using the cyclicity of the trace and the commutator of the negative frequency operator with the Hamiltonian. Second, from here we see that

$$\text{Tr}(e^{-\beta H} [A_\omega, A_\omega^\dagger]) = (1 - e^{-\beta\omega}) \text{Tr}(e^{-\beta H} A_\omega A_\omega^\dagger),$$

or $C_\omega = (1 - e^{-\beta\omega}) G_\omega$. Therefore,

$$\langle \Psi | U^\dagger V | \Psi \rangle = 1 - \theta^2 \sum_\omega (1 - e^{-\frac{\beta\omega}{2}})^2 G_\omega$$

$$= 1 - \theta^2 \sum_\omega \frac{(1 - e^{-\frac{\beta\omega}{2}})^2}{1 - e^{-\beta\omega}} C_\omega.$$

Noting that $(1 - e^{-\frac{\beta\omega}{2}})^2 < \omega(1 - e^{-\beta\omega})$ for $\omega > 0$, we see that

$$1 - \langle \Psi | U^\dagger V | \Psi \rangle < (\beta \delta E),$$

where we have used the fact that $\langle \Psi | [A_\omega, A_\omega^\dagger] | \Psi \rangle > 0$, and so both terms involve a sum over positive numbers, and the sum involving the change in energy is larger when examined term by term.

This proves the result (4.21). Since the result is intuitive and easy to state, it is possible that we have missed a simpler argument. However, we briefly mention one complication that leads us to reject some other seemingly simple arguments. We see that

$$U^\dagger H U = H + \delta H,$$

where we already know that $\langle \Psi | \delta H | \Psi \rangle = \delta E$, but we do not have much control over the operator (δH) itself from this relation. As a result of this, even though the unitary U raises the energy by δE, we see that $e^{-\frac{\beta H}{2}} U e^{\frac{\beta H}{2}} \neq e^{-\frac{\beta \delta E}{2}} U$. We see

that such a relation cannot hold because unitary $U^\dagger$ *also* raises the energy, and so the conjugate of this relation is manifestly untrue $e^{\frac{\beta H}{2}} U^\dagger e^{-\frac{\beta H}{2}} \neq e^{\frac{\beta \delta E}{2}} U^\dagger$.

This tells us that the fact that δH cannot be substituted by its expectation value is important. It is to get control over this operator that we have to examine the expansion of the unitary in terms of an exponential of components of definite frequency as we did above.

1. Multiple unitaries

Using the results above, it is also easy to prove the result (6.5). We expand both unitaries $U_{t_C}^{\mathcal{C}}$ and $\widehat{V}_{t_C}^{\mathcal{C}}$ as

$$U_\tau^{\mathcal{C}} = e^{i\theta_1 \sum_\omega (A_\omega^{\mathcal{C}} + A_\omega^{\mathcal{C}\dagger})}; \qquad \widehat{V}_\tau^{\mathcal{C}} = e^{i\theta_2 \sum (\bar{A}_\omega^{\mathcal{C}} + \bar{A}_\omega^{\mathcal{C}\dagger})}.$$

We now see that

$$\begin{aligned}
\delta E &= \langle \Psi | \widehat{V}_\tau^{\mathcal{C}\dagger} U_\tau^{\mathcal{C}\dagger} H U^{\mathcal{C}} \widehat{V}^{\mathcal{C}\dagger} | \Psi \rangle - \langle \Psi | H | \Psi \rangle \\
&= \theta_1^2 \sum_{\omega,\omega'} \omega \langle \Psi | \widehat{V}_\tau^{\mathcal{C}\dagger} [A_\omega^{\mathcal{C}}, A_{\omega'}^{\mathcal{C}\dagger}] \widehat{V}_\tau^{\mathcal{C}} | \Psi \rangle \\
&\quad + \theta_2^2 \sum_{\omega,\omega'} \omega \langle \Psi | U_\tau^{\mathcal{C}\dagger} [\bar{A}_\omega^{\mathcal{C}}, \bar{A}_{\omega'}^{\mathcal{C}\dagger}] U_\tau^{\mathcal{C}} | \Psi \rangle + \mathrm{O}(\theta_1^3) + \mathrm{O}(\theta_2^3) \\
&= \delta E_{\mathcal{C}} + \delta \bar{E}_{\mathcal{C}}.
\end{aligned}$$

Here we have just used the fact that $U_\tau^{\mathcal{C}}$ commutes with elements of $\mathcal{A}$ whereas $\widehat{V}^{\mathcal{C}}$ commutes with elements of $\bar{\mathcal{A}}$.

[1] J. M. Maldacena, The large N limit of superconformal field theories and supergravity, Int. J. Theor. Phys. **38**, 1113 (1999); S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, Gauge theory correlators from noncritical string theory, Phys. Lett. B **428**, 105 (1998); E. Witten, Anti-de Sitter space and holography, Adv. Theor. Math. Phys. **2**, 253 (1998).

[2] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, Local bulk operators in AdS/CFT: A boundary view of horizons and locality, Phys. Rev. D **73**, 086003 (2006); Holographic representation of local bulk operators, Phys. Rev. D **74**, 066009 (2006); Local bulk operators in AdS/CFT and the fate of the BTZ singularity, AMS/IP Stud. Adv. Math. **44**, 85 (2008); Local bulk operators in AdS/CFT: A holographic description of the black hole interior, Phys. Rev. D **75**, 106001 (2007).

[3] K. Papadodimas and S. Raju, An infalling observer in AdS/CFT, J. High Energy Phys. 10 (2013) 212,

[4] K. Papadodimas and S. Raju, Black Hole Interior in the Holographic Correspondence and the Information Paradox, Phys. Rev. Lett. **112**, 051301 (2014).

[5] K. Papadodimas and S. Raju, State-dependent bulk-boundary maps and black hole complementarity, Phys. Rev. D **89**, 086010 (2014).

[6] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, Black holes: Complementarity or firewalls?, J. High Energy Phys. 02 (2013) 062.

[7] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, An apologia for firewalls, J. High Energy Phys. 09 (2013) 018; D. Marolf and J. Polchinski, Gauge/Gravity Duality and the Black Hole Interior, Phys. Rev. Lett. **111**, 171301 (2013).

[8] D. Marolf and J. Polchinski, Violations of the Born rule in cool state-dependent horizons, J. High Energy Phys. 01 (2016) 008.

[9] D. Harlow, Aspects of the Papadodimas-Raju proposal for the black hole interior, J. High Energy Phys. 11 (2014) 055.

[10] S. H. Shenker and D. Stanford, Black holes and the butterfly effect, J. High Energy Phys. 03 (2014) 067; Multiple shocks, J. High Energy Phys. 12 (2014) 046.

[11] L. Susskind and L. Thorlacius, Gedanken experiments involving black holes, Phys. Rev. D **49**, 966 (1994).

[12] K. Papadodimas and S. Raju, Comments on the necessity and implications of state-dependence in the black hole interior, Phys. Rev. D **93**, 084049 (2016).

[13] K. Papadodimas and S. Raju, Local Operators in the Eternal Black Hole, Phys. Rev. Lett. **115**, 211601 (2015).

[14] S. Weinberg, Quantum contributions to cosmological correlations, Phys. Rev. D **72**, 043514 (2005).

[15] S. D. Mathur, The information paradox: A pedagogical introduction, Classical Quantum Gravity **26**, 224001 (2009).

[16] R. Bousso, Complementarity is not enough, Phys. Rev. D **87**, 124023 (2013); L. Susskind, The transfer of entanglement: The case for firewalls, arXiv:1210.2098; S. D. Mathur and D. Turton, Comments on black holes I: The possibility of complementarity, J. High Energy Phys. 01 (2014) 034; B. D. Chowdhury and A. Puhm, Is Alice burning or fuzzing?, Phys. Rev. D **88**, 063509 (2013); L. Susskind, Singularities, firewalls, and complementarity, arXiv:1208.3445; I. Bena, A. Puhm, and B. Vercnocke, Non-extremal black hole microstates: Fuzzballs of fire or fuzzballs of fuzz?, J. High Energy Phys. 12 (2012) 014; A. Giveon and N. Itzhaki, String theory versus black hole complementarity, J. High Energy Phys. 12 (2012) 094; T. Banks and W. Fischler, Holographic space-time does not predict firewalls, arXiv:1208.4757v3; A. Ori, Firewall or smooth horizon?, Gen. Relativ. Gravit. **48**, 9 (2016); S. Hossenfelder, Comment on the black hole firewall, arXiv:1210.5317; D.-i. Hwang, B.-H. Lee, and D.-h. Yeom, Is the firewall consistent?: Gedanken experiments on black hole complementarity and firewall proposal, J. Cosmol. Astropart. Phys. 01 (2013) 005; S. G. Avery, B. D. Chowdhury, and A. Puhm, Unitarity and fuzzball

complementarity: 'Alice fuzzes but may not even know it!', J. High Energy Phys. 09 (2013) 012; K. Larjo, D. A. Lowe, and L. Thorlacius, Black holes without firewalls, Phys. Rev. D **87,** 104018 (2013); S. K. Rama, Remarks on black hole evolution a la firewalls and fuzzballs, arXiv:1211.5645v4; D. N. Page, Hyper-entropic gravitational fireballs (grireballs) with firewalls, J. Cosmol. Astropart. Phys. 04 (2013) 037; M. Saravani, N. Afshordi, and R. B. Mann, Empty black holes, firewalls, and the origin of Bekenstein-Hawking entropy, Int. J. Mod. Phys. D **23,** 1443007 (2014); T. Jacobson, Boundary unitarity and the black hole information paradox, Int. J. Mod. Phys. D **22,** 1342002 (2013); L. Susskind, Black hole complementarity and the Harlow-Hayden conjecture, arXiv:1301.4505; W. Kim, B.-H. Lee, and D.-H. Yeom, Black hole complementarity and firewall in two dimensions, J. High Energy Phys. 05 (2013) 060; I. Y. Park, On the pattern of black hole information release, Int. J. Mod. Phys. A **29,** 1450047 (2014); S. D. H. Hsu, Macroscopic superpositions and black hole unitarity, arXiv:1302.0451; S. B. Giddings, Nonviolent information transfer from black holes: A field theory parametrization, Phys. Rev. D **88,** 024018 (2013); B.-H. Lee and D.-h. Yeom, Status report: Black hole complementarity controversy, Nucl. Phys. B, Proc. Suppl. **246–247,** 178 (2014); S. G. Avery and B. D. Chowdhury, Firewalls in AdS/CFT, J. High Energy Phys. 10 (2014) 174; B. Kang, Bulk cluster decomposition in AdS/CFT and a no-go theorem for correlators in microstates of extremal black holes, arXiv:1305.2797; B. D. Chowdhury, Black holes versus firewalls and thermo-field dynamics, Int. J. Mod. Phys. D **22,** 1342011 (2013); D. N. Page, Excluding black hole firewalls with extreme cosmic censorship, J. Cosmol. Astropart. Phys. 06 (2014) 051; M. Axenides, E. G. Floratos, and S. Nicolis, Modular discretization of the AdS_2/CFT_1 holography, J. High Energy Phys. 02 (2014) 109; M. Gary, Still no Rindler firewalls, arXiv:1307.4972; B. D. Chowdhury, Cool horizons lead to information loss, J. High Energy Phys. 10 (2013) 034; A. de la Fuente and R. Sundrum, Holography of the BTZ black hole, inside and out, J. High Energy Phys. 09 (2014) 073; J. L. F. Barbn and E. Rabinovici, Conformal Complementarity Maps, J. High Energy Phys. 12 (2013) 023; S. Lloyd and J. Preskill, Unitarity of black hole evaporation in final-state projection models, J. High Energy Phys. 08 (2014) 126; S. D. H. Hsu, Factorization of unitarity and black hole firewalls, arXiv:1308.5686v2; D. N. Page, Time dependence of Hawking radiation entropy, J. Cosmol. Astropart. Phys. 09 (2013) 028; S. B. Giddings and Y. Shi, Effective field theory models for nonviolent information transfer from black holes, Phys. Rev. D **89,** 124032 (2014); S. D. Mathur, What does strong subadditivity tell us about black holes?, Nucl. Phys. B, Proc. Suppl. **251–252,** 16 (2014); S. D. Mathur and D. Turton, The flaw in the firewall argument, Nucl. Phys. **B884,** 566 (2014); V. Balasubramanian, M. Berkooz, S. F. Ross, and J. Simon, Black holes, entanglement and random matrices, Classical Quantum Gravity **31,** 185009 (2014); B. Freivogel, Energy and information near black hole horizons, J. Cosmol. Astropart. Phys. 07 (2014) 041; R. Akhoury, Unitary S matrices with long-range correlations and the quantum black hole, J. High Energy Phys. 08 (2014) 169; N. Lashkari and J. Simn, From state distinguishability to effective bulk locality,

J. High Energy Phys. 06 (2014) 038; J. L. F. Barbon and E. Rabinovici, Geometry and quantum noise, Fortschr. Phys. **62,** 626 (2014); D. Kabat and G. Lifschytz, Asymmetric interiors for small black holes, J. High Energy Phys. 08 (2016) 097; M. Guica and D. L. Jafferis, On the construction of charged operators inside an eternal black hole, arXiv:1511.05627; F. M. Haehl, R. Loganayagam, and M. Rangamani, The Fluid Manifesto: Emergent symmetries, hydrodynamics, and black holes, J. High Energy Phys. 01 (2016) 184; J. Lin, Ph.D. thesis, Chicago University, 2015 [arXiv:1510.02367]; N. Bao, J. Pollack, and G. N. Remmen, Wormhole and entanglement (non-)detection in the ER = EPR correspondence, J. High Energy Phys. 11 (2015) 126; S. Emelyanov, Holography versus correspondence principle: eternal Schwarzschild-anti-de Sitter geometry, Phys. Rev. D **95,** 064044 (2017); S. R. Roy and D. Sarkar, Hologram of a pure state black hole, Phys. Rev. D **92,** 126003 (2015); D. Engelhardt, B. Freivogel, and N. Iqbal, Electric fields and quantum wormholes, Phys. Rev. D **92,** 064050 (2015); S. Emelyanov, Observing quantum gravity in asymptotically AdS space, Phys. Rev. D **92,** 124062 (2015); E. Verlinde and H. Verlinde, Black hole information as topological qubits, arXiv:1306.0516; Black hole entanglement and quantum error correction, J. High Energy Phys. 10 (2013) 107; Passing through the firewall, arXiv:1306.0515; Behind the horizon in AdS/CFT, arXiv:1311.1137; Y. Nomura, J. Varela, and S. J. Weinberg, Complementarity endures: No firewall for an infalling observer, J. High Energy Phys. 03 (2013) 059; R. Brustein, Origin of the blackhole information paradox, Fortschr. Phys. **62,** 255 (2014); R. Brustein and A. J. M. Medved, Phases of information release during black hole evaporation, J. High Energy Phys. 02 (2014) 116; Semiclassical black holes expose forbidden charges and censor divergent densities, J. High Energy Phys. 09 (2013) 108; Restoring predictability in semiclassical gravitational collapse, J. High Energy Phys. 09 (2013) 015; E. Silverstein, Backdraft: String creation in an old Schwarzschild black hole, arXiv:1402.1486v2; Y. Nomura, J. Varela, and S. J. Weinberg, Black holes, information, and Hilbert space for quantum gravity, Phys. Rev. D **87,** 084050 (2013); Y. Nomura and J. Varela, A note on (no) firewalls: The entropy argument, J. High Energy Phys. 07 (2013) 124; Y. Nomura, J. Varela, and S. J. Weinberg, Black holes or firewalls: A theory of horizons, Phys. Rev. D **88,** 084052 (2013); S. Braunstein, S. Pirandola, and K. Życzkowski, Better Late than Never: Information Retrieval from Black Holes, Phys. Rev. Lett. **110,** 101301 (2013); D. Harlow and P. Hayden, Quantum computation vs. firewalls, J. High Energy Phys. 06 (2013) 085; L. Susskind, New concepts for old black holes, arXiv:1311.3335; A. Sivaramakrishnan, Localized excitations from localized unitary operators, Ann. Phys. (Amsterdam) **381,** 41 (2017).

[17] B. Freivogel, R. A. Jefferson, L. Kabir, and I.-S. Yang, Geometry of the infalling causal patch, Phys. Rev. D **91,** 044036 (2015).

[18] I. Ilgin and I.-S. Yang, Causal patch complementarity: The inside story for old black holes, Phys. Rev. D **89,** 044007 (2014); T. Banks, W. Fischler, S. Kundu, and J. F. Pedraza,

Holographic space-time and black holes: Mirages as alternate reality, arXiv:1401.3341.

[19] R. Bousso, Holographic Probabilities in Eternal Inflation, Phys. Rev. Lett. **97,** 191302 (2006); R. Bousso, B. Freivogel, and I.-S. Yang, Eternal inflation: The inside story, Phys. Rev. D **74,** 103516 (2006).

[20] S. Raju, The sensitivity of black holes to low energy excitations, arXiv:1703.10159.

[21] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43,** 2046 (1991); M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50,** 888 (1994); The approach to thermal equilibrium in quantized chaotic systems, J. Phys. A **32,** 1163 (1999).

[22] B. Datt, On a class of solutions of the gravitation equations of relativity, Gen. Relativ. Gravit. **31,** 1619 (1999); Über eine klasse von lösungen der gravitationsgleichungen der relativität, Z. Phys. **108,** 314 (1938); J. R. Oppenheimer and H. Snyder, On continued gravitational contraction, Phys. Rev. **56,** 455 (1939).

[23] L. Fidkowski, V. Hubeny, M. Kleban, and S. Shenker, The black hole singularity in AdS/CFT, J. High Energy Phys. 02 (2004) 014.

[24] S. Banerjee, J.-W. Bryan, K. Papadodimas, and S. Raju, A toy model of black hole complementarity, J. High Energy Phys. 05 (2016) 004.

[25] B. P. Jensen and P. Candelas, The Schwarzschild radial functions, Phys. Rev. D **33,** 1590 (1986).

[26] G. T. Horowitz and V. E. Hubeny, Quasinormal modes of AdS black holes and the approach to thermal equilibrium, Phys. Rev. D **62,** 024027 (2000).

[27] E. Berti, V. Cardoso, and A. O. Starinets, Quasinormal modes of black holes and black branes, Classical Quantum Gravity **26,** 163001 (2009); K. D. Kokkotas and B. G. Schmidt, Quasinormal modes of stars and black holes, Living Rev. Relativ. **2,** 2 (1999).

[28] R. Bousso, Violations of the Equivalence Principle by a Nonlocally Reconstructed Vacuum at the Black Hole Horizon, Phys. Rev. Lett. **112,** 041102 (2014).

[29] G. 't Hooft, On the quantum structure of a black hole, Nucl. Phys. **B256,** 727 (1985).

[30] L. Susskind, L. Thorlacius, and J. Uglum, The Stretched horizon and black hole complementarity, Phys. Rev. D **48,** 3743 (1993); Y. Kiem, H. L. Verlinde, and E. P. Verlinde, Black hole horizons and complementarity, Phys. Rev. D **52,** 7053 (1995).

[31] E. Joos, H. D. Zeh, C. Kiefer, D. J. Giulini, J. Kupsch, and I.-O. Stamatescu, *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer Science and Business Media, New York, 2013).

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

[33] B. S. DeWitt, in *General Relativity: An Einstein Centenary Survey* (Cambridge University Press, Cambridge, England, 1980), p. 680.