Title
Quantum information bound on the energy

Journal
Physical Review D, 100(12)

ISSN
2470-0010

Authors
Bousso, Raphael
Shahbazi-Moghaddam, Arvin
Tomašević, Marija

Publication Date
2019-12-15

DOI
10.1103/physrevd.100.126010

Peer reviewed

Quantum information bound on the energy

Raphael Bousso,[1,2] Arvin Shahbazi-Moghaddam,[1,2] and Marija Tomašević[3]
[1]*Center for Theoretical Physics and Department of Physics, University of California, Berkeley, California 94720, USA*
[2]*Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Departament de Física Quàntica i Astrofísica and Institut de Ciències del Cosmos (ICCUB), Universitat de Barcelona, Martí i Franquès 1, E-08028 Barcelona, Spain*

(Received 13 September 2019; published 10 December 2019)

According to the classical Penrose inequality, the mass at spatial infinity is bounded from below by a function of the area of certain trapped surfaces. We exhibit quantum field theory states that violate this relation at the semiclassical level. We formulate a quantum Penrose inequality, by replacing the area with the generalized entropy of the light sheet of an appropriate quantum trapped surface. We perform a number of nontrivial tests of our proposal, and we consider and rule out alternative formulations. We also discuss the relation to weak cosmic censorship.

DOI: 10.1103/PhysRevD.100.126010

I. INTRODUCTION

Semiclassical general relativity allows for quantum matter while keeping the gravitational field classical, by coupling the metric to the expectation value of the stress tensor:

$$G_{ab} = 8\pi G\langle T_{ab}\rangle. \tag{1.1}$$

Since $\langle T_{ab}\rangle$ receives quantum contributions proportional to $\hbar$, this approximation can be organized as a perturbative expansion in $G\hbar$ and solved iteratively. This approach has proven to be quite useful, leading to the discovery of black hole thermodynamics and the associated information paradox.

Numerous theorems in general relativity rely on the null energy condition (NEC), which states that

$$T_{ab}k^a k^b \geq 0 \tag{1.2}$$

at every point in the spacetime, where k^a is any null vector. The NEC underlies the area theorems for event horizons [1] and for future holographic screens [2,3], the focusing theorem [4], and Penrose's singularity theorem [5]. In other theorems, the stress tensor is assumed to obey even stronger conditions, which are nevertheless satisfied by known classical matter and radiation.

However, in relativistic quantum field theories (QFTs) such as the Standard Model, there are states in which $\langle T_{ab}\rangle$ violates the NEC in some regions of spacetime. Hence, none of the classical theorems mentioned above apply at the semiclassical level. The evaporation of a black hole, for example, is accompanied by violations of all of the above theorems. This is possible because the NEC is violated in the vicinity of the horizon.

Remarkably, there is considerable evidence that all of the above theorems admit a conjectural semiclassical extension. The key step to obtaining a viable proposal is to replace the area of surfaces with their generalized entropy. Thus the area theorem becomes the generalized second law (GSL) for event horizons [6–8] and for Q screens [9]. The focusing theorem becomes the quantum focusing conjecture (QFC) [10]; and Penrose's singularity theorem becomes Wall's quantum singularity theorem [11].

Though these are conjectural statements about the semiclassical limit of quantum gravity, they can have interesting nongravitational limits. Some of these limit statements were already known, but others came as completely new and nontrivial results in QFT. The main example is the quantum null energy condition [10], which has since been rigorously proven within QFT, using a variety of methods [12–14]. Thus, the study of semiclassical gravity has had considerable impact in a seemingly unrelated arena.

The present work is inspired by these developments. We will study an important conjecture in classical general relativity, the Penrose inequality [15]. This is a relation between the area of certain marginally trapped surfaces μ in the spacetime and the total mass defined at spatial infinity [16]:

$$m \geq \sqrt{\frac{A[\mu]}{16\pi G^2}}. \tag{1.3}$$

The conjecture can be thought of as a generalization of the positive mass theorem [17]. For either statement, it is clearly essential that matter with negative energy be excluded. This can be implemented by assuming the dominant energy condition (DEC) that for any timelike future-directed vector t^a, $-T^a{}_b t^b$ is timelike and future directed.

The Penrose inequality has not been proven and thus is not a theorem. But no counterexample to the conjecture is known. We will review the classical Penrose inequality in Sec. II, where we provide both the reasoning motivating it and a more careful formulation.

Since quantum matter can violate the NEC, it can also violate the DEC, threatening the validity of the Penrose inequality. It is not immediately obvious that Eq. (1.3) fails, since the stress tensor in QFT cannot be dialed arbitrarily.

In fact, we find in Sec. III that Eq. (1.3) continues to be satisfied in a simple example of black hole formation and evaporation. However, we then provide an explicit counterexample to the classical Penrose inequality, by exploiting the thermal nature of the vacuum state near the horizon. When the thermal state is depleted, the vicinity of the horizon can contribute significant negative energy. This cancels an order one fraction of the black hole's mass, leading to a substantial violation of Eq. (1.3).

We are thus motivated, in Sec. IV, to propose a quantum-corrected version of the Penrose inequality. We introduce the relevant concepts of generalized entropy, quantum expansion, and quantum (marginally) trapped surfaces. We draw some lessons from the failure of the classical Penrose inequality in the semiclassical setting, and we formulate a quantum Penrose inequality (QPI).

In Sec. V, we provide evidence for our proposal. We consider several interesting examples that could challenge the QPI, and we show that our proposal survives these tests. In Sec. VI, we discuss a number of alternative formulations of the QPI. We show why they are either excluded or not ideal. In Sec. VII, we discuss the formulation of a QPI in asymptotically anti–de Sitter spacetimes. This helps us identify subtleties that also affect the original QPI.

In Sec. VIII, we discuss the classical and the nongravitational limits of the QPI.

Penrose's motivation in proposing Eq. (1.3) was as a test of the weak cosmic censorship conjecture (CCC). In Sec. IX, we review this connection and the status of the CCC. We speculate that the quantum Penrose inequality could inform the formulation of a "quantum" CCC that accommodates the known, physically sensible violations of the classical CCC.

In Appendix A, we compute the expansion of outgoing null rays and the positions of classical and quantum marginally trapped surfaces for an evaporating Schwarzschild black hole. In Appendix B, we present a perturbative construction of Q screens [9], which plays a role in our discussion of the quantum Penrose inequality in anti–de Sitter (AdS) spacetimes.

A brief summary of the main results of our investigation has appeared elsewhere [18].

II. CLASSICAL PENROSE INEQUALITY

In this section we describe the (classical) Penrose inequality; see Ref. [19] for a broader review.

A. Formulation

We formulate the classical Penrose inequality as follows:

Let m be the total mass of an asymptotically flat spacetime. Let μ be a trapped surface that has minimal area among all surfaces that enclose it, on some Cauchy surface that contains μ. Then

$$m \geq \sqrt{\frac{A[\mu]}{16\pi G^2}}. \tag{2.1}$$

Next, we provide detailed definitions and explanations of the terms appearing in this formulation.

Let (M, g_{ab}) be a connected Lorentzian spacetime with metric. Let μ be a codimension $1+1$ compact spacelike submanifold (a "surface").[1] Let $\theta_\pm$ be the expansion of the future-directed light rays emanating orthogonally from μ to either side. If $\theta_+ \leq 0$ and $\theta_- \leq 0$, then μ is called *trapped*. If $\theta_+ = 0$ and $\theta_- \leq 0$, then μ is *marginally trapped*.

Now let (M, g_{ab}) be in addition asymptotically flat. Note that we do not require μ to be connected; for example, in a spacetime where multiple black holes are forming, μ could be the union of connected marginally trapped surfaces inside some or all of them.

Suppose that the surface μ has an "outer wedge" O_W that contains a single asymptotic region. By this we mean that μ forms the only boundary of any Cauchy surface of a globally hyperbolic region of space O_W that (in the "unphysical spacetime" or Penrose diagram) contains a single copy of spatial infinity, i_0. This will be the case for trapped surfaces in a spacetime with a single asymptotic region. In the case of "two-sided" black hole solutions, it will hold if μ is homologous[2] to a horizon (with either choice of side), but not if μ is contractible. We will be interested in bounding the Arnowitt-Deser-Misner (ADM) mass at spatial infinity [16] from below.

Finally, we assume that there exists a Cauchy surface Σ of O_W on which μ is the minimal area surface homologous

[1] In the remainder of this paper we will specialize to $(3+1)$-dimensional spacetime, so that μ will be a two-dimensional surface. Generalization to higher dimensions is trivial.

[2] Two cycles (closed submanifolds that are not boundaries of any other submanifolds) are said to be homologous, or equivalently, belong to the same homology class, if they can be continuously deformed into each other.

to large spheres near i_0 (or in the AdS case, homologous to the boundary sphere) [20]. The purpose of this set of assumptions will become clear as we turn to presenting a heuristic argument that the Penrose inequality should hold for μ.

B. Heuristic argument

The Penrose inequality was originally intended as a test of cosmic censorship, which guarantees that an asymptotically flat spacetime with regular initial conditions will be strongly asymptotically predictable [4]. If this latter property holds, then a compelling argument can be given that the Penrose inequality must hold; thus, any regular initial dataset that violates the Penrose inequality would likely exclude cosmic censorship. We now present the argument.

Roughly speaking, strong asymptotic predictability establishes the existence of $\tilde{V}$, a globally hyperbolic open subset of M that contains any black hole horizons and their exterior, $\tilde{V} \supset J^-(\tilde{\mathcal{I}}^+)$. (See Ref. [4] for more details.) The *black hole region* is $B \equiv M - J^-(\mathcal{I}^+)$. The black hole *event horizon* is its boundary $\dot{B}$.

Suppose that

$$R_{ab}k^a k^b \geq 0, \tag{2.2}$$

as would be the case if Einstein's equations hold with matter satisfying the null energy condition. Then any trapped or marginally trapped surface μ must lie in the black hole region:

$$\mu \subset B. \tag{2.3}$$

For a proof, see Propositions 12.2.2 in Ref. [4]. The key technical assumption is that M be strongly asymptotically predictable.[3]

Let $H = \dot{B} \cup \Sigma$ be the slice of the black hole event horizon (possibly with multiple disconnected components), on the Cauchy surface Σ of O_W. Since μ has a minimal area on Σ, it follows that the horizon must be at least as large[4]:

$$A[H] \geq A[\mu]. \tag{2.4}$$

The null curvature condition, Eq. (2.2), and strong asymptotic predictability imply that the area of the event horizon cannot decrease with time [1]. Let $H' = \dot{B} \cup \Sigma'$, where Σ' is a Cauchy surface to the future of Σ. Then

$$A[H'] \geq A[H]. \tag{2.5}$$

Physically, it is reasonable to assume that regular initial data will eventually settle down to a Kerr black hole. (In four dimensions, this follows from the assumption of late-time stationarity, by the Israel-Hawking-Carter theorems [21].) Letting H' be a slice of the horizon at this late time, the formula for the area of a Kerr black hole implies that

$$16\pi G^2 m_{\text{Kerr}}^2 \geq A[H']. \tag{2.6}$$

The spacetime will not be exactly Kerr, however. One expects that massive fields will have fallen into the black hole, but there may be massless fields that propagate to future null infinity. Because this radiation becomes dilute and well separated from the black hole, gravitational binding energy will be negligible. Hence the ADM mass, m, will be given by the sum

$$m = m_{\text{Kerr}} + m_{\text{rad}} \geq m_{\text{Kerr}}. \tag{2.7}$$

Combining the previous four inequalities, we obtain the Penrose conjecture, Eq. (2.1).

We would like to add a second, somewhat independent heuristic argument for Eq. (2.1). A *future holographic screen* is a hypersurface foliated by marginally trapped surfaces called *leaves* [2,22,23]. Assuming the null energy condition, the area of the leaves increases monotonically along this foliation [2,3]. In the spherically symmetric case, the screen eventually asymptotes to the event horizon (from the interior), so its final area will be equal to the late time event horizon area. Thus the screen area theorem implies the Penrose inequality in this case. More generally, given a marginally trapped surface μ, a future holographic screen can be constructed at least in a neighborhood. The Penrose inequality would follow from the stronger assumption that there exists a future holographic screen that interpolates from μ to the late-time event horizon, as in the spherical case.

III. VIOLATION BY QUANTUM EFFECTS

In this section, we will show that there is a need for a quantum generalization of the classical Penrose inequality (CPI). We will construct an explicit counterexample that is based on a Boulware-like state outside a Schwarzschild black hole. It violates the CPI by a substantial, classical amount.

This will be a counterexample to the CPI in the same sense as black hole evaporation is a counterexample to Hawking's area theorem: we identify a physically allowed state in which a key assumption of the classical statement, the null energy condition, does not hold, and we verify that the conclusion fails as well.

[3]The same property, $\nu \subset B$, follows from Proposition 12.2.3 in Ref. [4] for another class of surfaces called *outer trapped*. These would form an alternate starting point from which the classical and quantum Penrose conjectures could be developed along the same lines as we do here for trapped surfaces.

[4]Instead of assuming that μ has minimal area on *some* Cauchy slice of O_W, an alternative way of handling this issue is to replace $A[\mu]$ with the minimal area of all surfaces enclosing μ on a *given* initial Cauchy slice [19]. Verifying this assumption does not require knowledge of more than the initial slice.

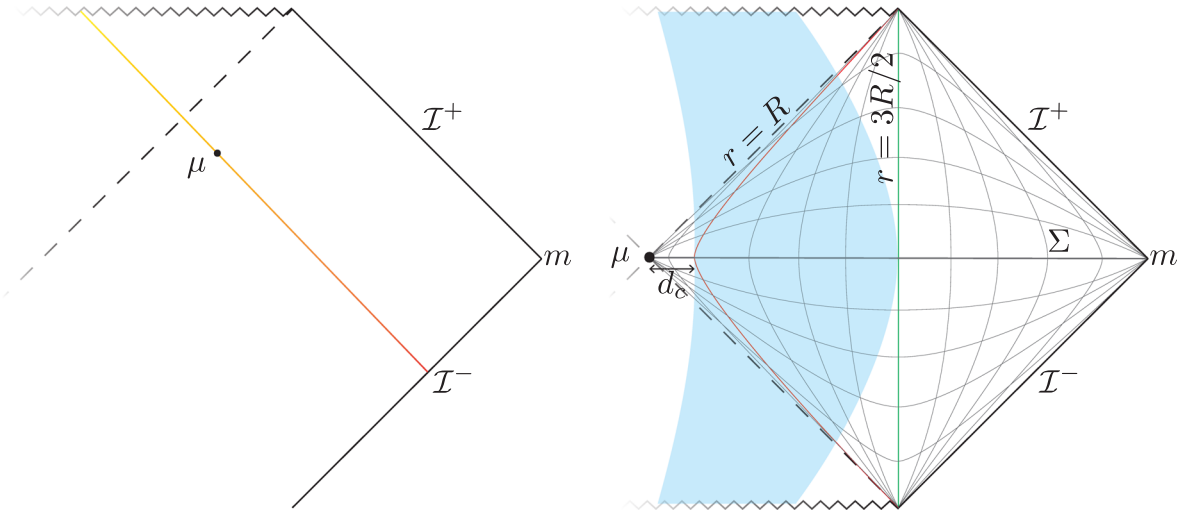

FIG. 1. Left: A null shell collapsing in asymptotically flat spacetime. The classically marginally trapped surface μ is slightly outside of the event horizon due to the evaporation. It is not clear that this example violates the CPI. Right: Initial data that violate the classical Penrose inequality. Here μ is the bifurcation surface of the Schwarzschild (Kruskal) solution. Inside a proper distance d_c, the state is the Hartle-Hawking vacuum. Outside of d_c, it becomes the Boulware vacuum, which has negative energy in the near-horizon zone (blue strip). This lowers the mass at infinity by an $O(1)$ fraction compared to a classical black hole.

However, before we turn to our counterexample, it is worth noting that no obvious violation of the CPI arises in the "normal" formation and evaporation of a black hole in the Unruh state. This is interesting, because in this setting the null energy condition is already violated, and other theorems such as the area theorem or the focusing theorem do fail. In order to have full control and exclude transient effects, let us consider the collapse of a null shell of mass m; see Fig. 1. Then by causality, there are no corrections to the classical solution on the shell and to its past, where the spacetime is a portion of Minkowski space. In particular, the marginally trapped surface on the shell will have the same area as in the classical case, and the CPI will be saturated. (The fact that the event horizon is inside of this surface is irrelevant.) At later times, we expect the apparent horizon area to decrease. Since the mass at infinity does not change during evaporation, the CPI will remain satisfied.

We do not claim that the CPI will hold for all black holes formed from collapse; and even in the above example, its validity may rely on idealizations, such as treating the collapsing null shell as infinitely thin and stable. But we would like to exhibit a situation where the CPI is definitely violated; in order to do this, we will consider a somewhat more artificial (but certainly valid) quantum state.

To demonstrate a violation of the classical PI by quantum effects, we now consider a Boulware-like state [24] of a massless scalar field, on one side of a maximally extended Schwarzschild black hole, at the time-symmetric slice; see Fig. 1. The Boulware vacuum is analogous to the Rindler vacuum. It corresponds to the vanishing occupation number of the modes with support strictly outside the event horizon. This will contribute some negative energy outside of the black hole, in the near-horizon region $R < r < 3R/2$. Far from the black hole, the stress tensor vanishes in the Boulware vacuum.

Note that the classical Penrose inequality, applied to the bifurcation surface, is classically saturated. (That is, it is saturated if the stress tensor vanishes everywhere outside the black hole.) Thus, any net negative energy in the exterior will lead to a violation of Eq. (2.1).

The local stress tensor diverges in the Boulware vacuum as the horizon is approached [24,25]. We regulate this divergence by building wave packets with support strictly outside of a sphere H_c at proper distance $d_c > 0$ from the horizon (in this case, from the bifurcation surface). For full control of the semiclassical expansion, we choose

$$l_P \ll d_c \ll R. \tag{3.1}$$

Roughly speaking, this yields a Hartle-Hawking-like state (vanishing stress tensor) inside of H_c and a Boulware-like state outside of H_c.

Integration of the QFT stress tensor computed in Ref. [25], outside the regulator sphere H_c, yields a QFT contribution to the energy at infinity of order $-(l_P/d_c)^2 M$, where $M = R/2G$ is the mass of the black hole [18]. Here we will go further; instead of naively gluing to QFT states across a surface (which does not generally yield an allowed QFT state), we consider junction effects at H_c. Positivity of the energy for infalling observers requires some positive energy near H_c, which we wish to estimate and show to be negligible.

For this purpose it will be useful to analyze the problem mode by mode. This will allow us to distinguish between two cutoffs that we can freely choose: the angular momentum of the included QFT modes, and d_c. Establishing a small

hierarchy between these cutoffs will give us a control parameter $1/n_{\rm node} \ll 1$, by which the positive energy at H_c is suppressed at infinity, relative to the negative contribution.

We will focus on the most relevant modes in the near-horizon zone, which have an occupation number of order one in the thermal ensemble corresponding to the Hartle-Hawking state. These modes have the property that any wave packet constructed from them has characteristic wavelength comparable to its distance from the horizon. Moreover, increasing the occupation number of the mode by 1 increases the energy at infinity by $\hbar/R$.

This set of modes includes s waves as well as modes with nonzero angular momentum. Here we will use $\ell = 0, 1, \ldots$, for the angular momentum quantum number. The number of modes in the thermal atmosphere can be estimated from the number of nodes in a strictly outgoing Rindler mode in an interval beginning at proper distance d_c from the horizon and ending at a distance R (for the spherical modes, which we approximate as propagating freely) or $R/(\ell+1)$ (for the modes with angular momentum, which we approximate as being reflected by an angular momentum barrier). See Fig. 2. Hence there are

$$n_\ell = (2\ell+1)\log\left(\frac{R/(\ell+1)}{d_c}\right) \tag{3.2}$$

linearly independent modes with angular momentum ℓ.

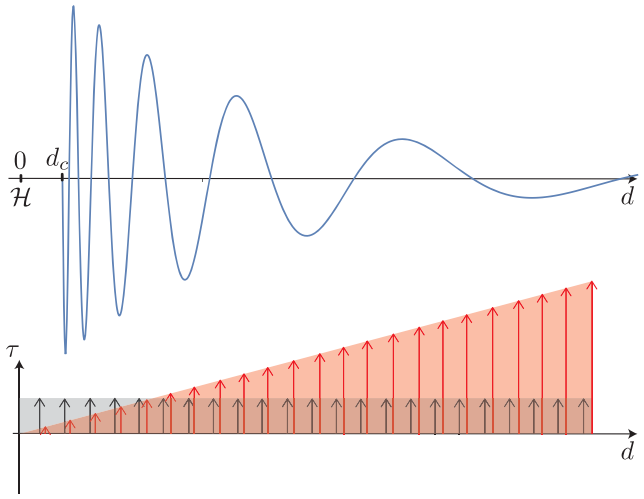

FIG. 2. A typical wave packet mode in the thermal atmosphere of the black hole, regulated to have support outside a sphere a proper distance d_c outside of the horizon. The classical Penrose inequality is violated in a Boulware-like state in which such modes have a zero occupation number and negative energy. In a local inertial frame (black Killing vector field, ∂_τ, where τ is proper time), a large fraction of their energy is concentrated near the cutoff d_c. The total energy must appear positive in this frame; this can be satisfied by adding a comparable amount of positive energy inside of d_c. To an asymptotic observer (red Killing vector field, ∂_t), the negative energy is spread evenly over the mode, due to the greater redshift near the horizon. Thus the positive energy beyond the cutoff has a negligible effect on the ADM mass.

In the Hartle-Hawking state, these modes are all thermally excited with $O(1)$ occupation numbers; this corresponds to vanishing stress tensor near the horizon. In the Boulware-like state, the modes are unoccupied. This corresponds to a negative stress tensor; it contributes an energy at infinity of order $-\hbar/R$, per mode. We choose a cutoff $\ell_{\rm max}$ on the angular momentum such that the angular momentum barrier is somewhat outside the short distance cutoff d_c:

$$\log\log\left(\frac{R/(\ell_{\rm max}+1)}{d_c}\right) \sim O(1), \tag{3.3}$$

where the second log enforces a small hierarchy whose purpose will become clear below. From the previous two equations, the total number of unoccupied modes is

$$n_{\rm total} \equiv \sum_{\ell=0}^{\ell_{\rm max}} n_\ell \sim \frac{R^2}{d_c^2}. \tag{3.4}$$

Thus the total energy at infinity of the quantum field will be

$$E_{\rm neg} \sim -\frac{\hbar}{R} n_{\rm total} \sim -\alpha M, \tag{3.5}$$

where

$$\alpha = \frac{l_P^2}{d_c^2}. \tag{3.6}$$

The presence of a substantial amount of negative energy outside the black hole may seem suspect. However, we note that our construction cannot achieve vanishing or negative total ADM mass. Since the black hole contributes M, the total mass is $(1-\alpha)M$. Making this negative would require taking $d_c \lesssim l_P$, in conflict with Eq. (3.1), and so would take us outside of the semiclassical expansion. Moreover, our result is consistent with positive total matter energy in an appropriate neighborhood of the horizon. This is important since the spacetime can be treated as approximately flat on a distance scale $d_c \ll d_{\rm flat} \ll R$.

To see this, we note that the wave packets we study have approximately constant Killing energy per cycle, where a cycle denotes the portion of a wave packet between two nodes. See Fig. 2. The local proper wavelength of a given mode grows as the distance from the horizon, but this is precisely canceled by the decreasing redshift. Thus from the viewpoint of infinity, each cycle of each mode contributes an ADM energy (per occupation number) of $\hbar/(Rn_{\rm node})$, where

$$n_{\rm node}(\ell) \sim \log\left(\frac{R/(\ell+1)}{d_c}\right) \tag{3.7}$$

is the number of nodes or cycles in the wave packet.

In a local inertial frame, on the other hand, there is no redshift effect. Yet, the proper wavelength grows exponentially away from the horizon, roughly doubling with every cycle. Thus an $O(1)$ fraction of the local energy of a mode is contained in the first phase cycle. In the Boulware-like state, this is the negative energy that must be canceled. To have positive energy in the local frame, it suffices to have compensating positive energy just for this first cycle. The positive energy can be localized, for example, just below d_c.

This positive energy will partially cancel the negative ADM energy of the quantum state, Eq. (3.5). But because all cycles of the wave packet contribute equally to the Killing energy, the correction is parametrically small, of order $|E_{\rm neg}|/n_{\rm node} \ll |E_{\rm neg}|$. In practice, the $n_{\rm node}$ of order a few suffices, so we will not update Eq. (3.6). The purpose of the second log in Eq. (3.3) was to chose the angular momentum cutoff $\ell_{\rm max}$ so as to achieve $n_{\rm node} \sim$ a few, for all modes involved in the construction.

Finally, we note that the location and area of the marginally trapped surface do not receive large enough corrections to rescue the classical Penrose inequality. The bifurcation surface remains marginally trapped when we pass from the classical treatment to the Hartle-Hawking state, since the stress tensor vanishes there. Our construction keeps the Hartle-Hawking state near the bifurcation surface, up to corrections that can be suppressed arbitrarily by dialing $n_{\rm node} \gg 1$.

To summarize, one can reduce the mass at infinity from M (in the Unruh state) to $(1-\alpha)M$ in the Boulware-like state. Since we require that $l_P \ll d_c$ for control, this correction is parametrically small, $\alpha \ll 1$. But since the Penrose inequality is saturated classically for a Schwarzschild black hole, our example violates it.

Moreover, the violation is substantial in the sense that it is not $O(\hbar)$ but $O(1)$. The contribution from each mode is $O(\hbar)$; but the number of available modes in the thermal atmosphere, at fixed control parameter l_P/d_c, is $n_{\rm total} \sim O(\hbar^{-1})$. Thus, the negative energy of the quantum fields can cancel off an $O(1)$ fraction of the black hole's classical mass.

IV. QUANTUM PENROSE INEQUALITY

In this section, we will formulate the QPI. In Sec. IV A, we review various concepts necessary for the quantum generalization of classical statements involving area and null expansion. In Sec. IV B, we draw some conclusions from the failure of the classical Penrose inequality. In Sec. IV C, we formulate our proposal for the QPI.

A. Generalized entropy and quantum expansion

We begin by introducing the notion of generalized entropy and its main properties. We then use the generalized entropy to define certain quantum generalizations of various geometric quantities, necessary for formulating the quantum Penrose inequality; see [10] for more details.

The *generalized entropy* $S_{\rm gen}$ was first introduced by Bekenstein [6,7] as the total entropy of a system consisting of a black hole and its exterior on a given time slice. The definition can be extended to apply not only to the horizon of a black hole but also to any Cauchy-splitting surface σ,

$$S_{\rm gen} \equiv \frac{A[\sigma]}{4G\hbar} + S_{\rm out} + \cdots, \tag{4.1}$$

where $A[\sigma]$ is the area of σ and

$$S_{\rm out} = -{\rm Tr}\rho_{\rm out} \log \rho_{\rm out} \tag{4.2}$$

is the von Neumann entropy of the state of the quantum fields, restricted to one side of σ,

$$\rho_{\rm out} = {\rm Tr}_{\overline{\rm out}}\rho. \tag{4.3}$$

Here, the state ρ is the global quantum state, and the trace is over the complement region, which we define as $\overline{\rm out}$.

The von Neumann entropy $S_{\rm out}$ quantifies the amount of entanglement in the vacuum across σ and, as such, has divergences coming from short-distance entanglement. The leading divergence is given by A/ϵ^2, where ϵ is a short-distance cutoff. However, we can think of the geometric term in Eq. (4.1) as a counterterm. The dots indicate the presence of subleading divergences in $S_{\rm out}$ that come with their own geometric counterterms. It is expected that the divergences coming from the renormalization of G and from short-distance entanglement will cancel out [10], so as to keep $S_{\rm gen}$ a finite and well-defined quantity.

One can interpret $S_{\rm gen}$ in two distinct ways. Following the original motivation, one can view the area term as a (large) "correction" to the entropy of quantum fields. Alternatively, we can define a quantum-corrected area of the surface σ,

$$A_Q[\sigma] \equiv A[\sigma] + 4G\hbar S_{\rm out} + \cdots, \tag{4.4}$$

in a semiclassical expansion in $G\hbar$. Hence, one can use the notion of generalized entropy to incorporate quantum effects into certain geometrical objects that derive from the area of surfaces.

One example is the notion of *quantum expansion*. Recall, the classical expansion of a surface σ at a point $y \in \sigma$ is defined as the trace of the null extrinsic curvature at y. Equivalently, one can define the classical expansion as a functional derivative,

$$\theta[\sigma; y] = \frac{1}{\sqrt{h(y)}} \frac{\delta A[V]}{\delta V(y)}, \tag{4.5}$$

where h represents the area element of the metric restricted to σ, inserted to ensure that the functional derivative is taken per unit proper area, not coordinate area. The function $V(y)$ is used to specify the affine location of σ and nearby surfaces along a congruence of null geodesics

orthogonal to σ. The above definition of the classical expansion is needlessly complicated, in that it invokes the entire surface σ, even though θ depends only on its local extrinsic curvature at y. However, this definition naturally generalizes to the quantum expansion, Θ, which does depend on all of σ:

$$\Theta[\sigma; y] \equiv \frac{4G\hbar}{\sqrt{h(y)}} \frac{\delta S_{\rm gen}[V]}{\delta V(y)}. \tag{4.6}$$

As in the classical case, we can use the notion of expansion to define certain types of surfaces (see Sec. II A). Let $\Theta_\pm$ be the quantum expansion of the future-directed light rays orthogonal to a surface μ_Q. (As before, we take the $+$ label to refer to the direction of spatial infinity.) If $\Theta_+ \leq 0$ ($\Theta_+ = 0$) and $\Theta_- \leq 0$, then we call μ_Q a *quantum (marginally) trapped surface*.

Quantum trapped surfaces, in the semiclassical setting, have some of the properties obeyed by trapped surfaces in the classical setting. For example, trapped surfaces cannot lie outside the black hole, assuming weak cosmic censorship and the null energy condition. When the NEC is violated, they can; however, quantum trapped surfaces must still lie inside or on the horizon [11] (still assuming weak cosmic censorship). This will prove to be important for our formulation of the quantum Penrose inequality.

A *quantum future holographic screen*, or Q screen, is a hypersurface foliated by quantum marginally trapped surfaces. Assuming the quantum focusing conjecture [10], Q screens obey a generalized second law [9].

B. Lessons from the counterexample

The failure of the classical PI in the presence of quantum matter (Sec. III) illustrates the need for a quantum Penrose inequality. It also motivates some of the choices we will make below.

Let us distinguish two different timescales: the time for the negative energy of the Boulware-like state to enter the black hole, and the evaporation time. The former is of order the scrambling time $\Delta t_s \sim R \log(R/l_P)$. The latter is much greater, of order $R^3/G\hbar$.

On the shorter timescale, the process results in an outcome very similar to that invoked in motivating the classical Penrose inequality: a Kerr black hole with area $A_{\rm late}$ and no further evolution. That is, we neglect evaporation since it occurs on a much greater timescale, and by construction, no matter that will ever enter the black hole. Thus, the mass should obey $16\pi G^2 m^2 \geq A_{\rm late}$.

The key difference to the classical case is that the "late" area need not be greater than the area of trapped surfaces at earlier times; indeed, our counterexample shows that it will not be. However, we know that the GSL takes the place of the area theorem in this setting. Thus, we expect that the generalized entropy of earlier quantum trapped surfaces should be less than $A_{\rm late}/4G\hbar$. And so, the generalized entropy of quantum trapped surfaces should replace the area of trapped surfaces when we replace the classical by a quantum Penrose inequality.

This argument is based on the GSL for the event horizon, and so involves an intermediate step where one argues that the generalized entropy of quantum marginally trapped surfaces inside the black hole will not be greater than that of the event horizon. To avoid this step, we can generalize the second heuristic argument for the classical Penrose inequality, which was based on the area theorem for future holographic screens. Q screens obey a GSL that interpolates directly between different marginally quantum trapped surfaces. If a suitable Q screen connects μ_Q to the late-time event horizon, this establishes a quantum Penrose inequality. Of course, this is far from a trivial assumption; our goal here was only to gain some intuition.

In the above heuristic arguments, it was important that the late-time generalized entropy should be given just by $A_{\rm late}$, i.e., that no entropy remains outside of the black hole. However, this will not be the case in general examples. This will motivate our choice, below, that the generalized entropy entering the quantum Penrose inequality should be evaluated on slices that remain inside the black hole. We will discuss this important issue further in Sec. VI A.

C. Formulation

We will now obtain a quantum Penrose inequality from the classical PI, in three steps. First, we replace the area with generalized entropy in Eq. (2.1):

$$A \to 4G\hbar S_{\rm gen} \equiv A + 4G\hbar S_{\rm out}. \tag{4.7}$$

Thus we propose an inequality of the form

$$m \geq \sqrt{\frac{\hbar S_{\rm gen}}{4\pi G}}. \tag{4.8}$$

Second, we must specify the surfaces to which the inequality can be applied. In the classical case, a surface μ has to be trapped for the Penrose inequality to apply, corresponding to criteria satisfied by the classical expansion. For the QPI, it is natural to apply the same criteria to the quantum expansion,

$$\theta \to \Theta. \tag{4.9}$$

Thus in Eq. (4.8), $S_{\rm gen}$ is the generalized entropy of any surface μ_Q that is quantum trapped. We expect that the most interesting bounds will obtain when μ_Q is quantum marginally trapped, and we will only consider this case in all examples below.

Next, we must specify on which achronal hypersurface the generalized entropy appearing in Eq. (4.8) should be computed. As we will explain in Sec. VI A, this *cannot* be

chosen to be a Cauchy surface of the outer wedge. Instead, we will propose that this hypersurface should be entirely contained in the "black hole region" $B \equiv M - J^-(\mathcal{I}^+)$, i.e., inside or on the horizon.

More precisely, we require that $S_{\rm gen}$ should be evaluated on the "future portion" of the boundary of the outer wedge,

$$L(\mu_Q) \equiv \dot{O}_W(\mu_Q) - I^-(O_W(\mu_Q)). \tag{4.10}$$

See Fig. 3. L is generated by the congruence of future-directed outgoing null geodesics orthogonal to μ_Q [4,26]. Their initial quantum expansion is $\Theta_+ = 0$ by construction, so assuming the QFC [10], $\Theta_+ \leq 0$ everywhere on L. Hence L will be a (quantum) light sheet of μ_Q. Assuming an appropriate version of weak cosmic censorship, L will terminate on the singularity inside the black hole. (Strictly, in order to remain in the semiclassical regime, one should terminate L slightly earlier, resulting in a second area term that can be made small by approaching the singularity.)

Note that the surface μ_Q must be quantum trapped with respect to L; it need not be quantum trapped with respect to any other hypersurface, such as a Cauchy surface of $O_W(\mu_Q)$. To find a suitable μ_Q, consider a null hypersurface N inside the black hole, for example, the boundary of the future of an event q inside the black hole; see Fig. 3. Typically the area of N will increase near q and later decrease toward the singularity. Hence the area will have a maximum on some cut of N, and the generalized entropy of cuts of N (computed with respect to the future of the cuts on N) will have a maximum on some nearby cut. This cut will be a suitable quantum marginally trapped surface μ_Q, and later cuts will also be quantum trapped.

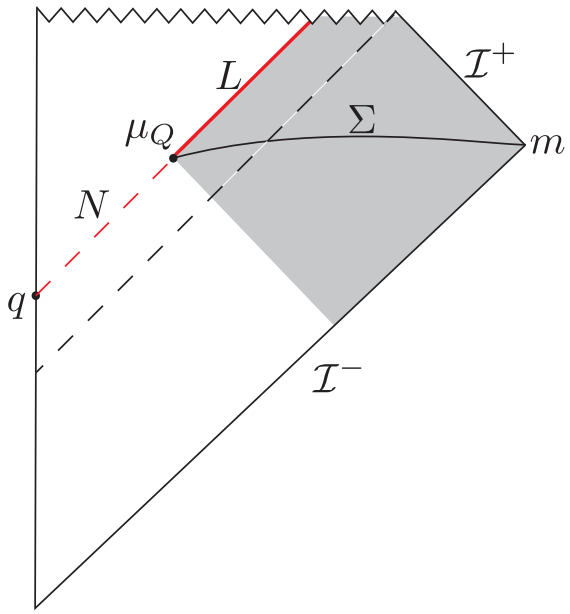

FIG. 3. The quantum Penrose inequality bounds the mass at infinity in terms of the generalized entropy of a quantum marginally trapped surface μ_Q. The generalized entropy must be evaluated on the light sheet L (red line), *not* on a Cauchy surface Σ of the outer wedge $O_W[\mu_Q]$ (shaded region).

Finally, we must impose a requirement analogous to the minimum area condition imposed on μ in the classical case. This condition demanded that there exist a Cauchy surface of O_W on which no surface enclosing μ_Q has area less than μ_Q. Here, we will instead consider the generalized entropy of any surface ν enclosing μ_Q, computed on the partial Cauchy slice $\Sigma(\nu)$ that ends at spatial infinity (similar to slice Σ for μ_Q in Fig. 3). For the QPI to apply to a quantum trapped surface μ_Q, we demand that there exist a Cauchy surface of $O_W[\mu_Q]$ on which no enclosing surface ν satisfies $S_{\rm gen}[\Sigma(\nu)] < S_{\rm gen}[\Sigma]$.

To summarize, we propose that the mass at spatial infinity of an asymptotically flat spacetime satisfies the quantum Penrose inequality

$$m \geq \sqrt{\frac{\hbar S_{\rm gen}[L(\mu_Q)]}{4\pi G}}, \tag{4.11}$$

where $S_{\rm gen}$ is computed on the future-outgoing light sheet of μ_Q and μ_Q is any quantum trapped surface homologous to spatial infinity that has minimal generalized entropy on some Cauchy surface of its outer wedge, in the sense described above.

We close by discussing a subtlety that introduces a small uncertainty in the formulation of the QPI. In Eq. (4.11), we used the classical functional relation between the area and mass of Schwarzschild black holes; we merely replaced the area with the generalized entropy. In fact, there will be a field-content-dependent quantum correction to the functional relation itself. However, this correction is small compared to the difference between our QPI and the classical Penrose inequality.

This is easier to discuss in asymptotically AdS space, where the Schwarzschild black hole can be in thermal equilibrium. Therefore, we will revisit the issue in more detail in Sec. VII. In general, the black hole exterior will have nonzero energy density in equilibrium. This is a kind of Casimir energy associated with the potential well provided by the near horizon zone. It contributes to the total mass at infinity; but since it stays outside the black hole, it will not contribute to $S_{\rm gen}[L(\mu_Q)]$.

By dimensional analysis, one expects each field theory degree of freedom to contribute an amount of order $\hbar/R$ to this Casimir energy. In Eq. (4.11), this is equivalent to changing the area or generalized entropy by $O(c)$, where c is the number of matter quantum fields. For large black holes in AdS, it is possible to determine this correction and include it in the QPI (see Sec. VII). In general, however, we are presently unable to determine it.

Since $S_{\rm gen}$ is $O(\hbar^{-1})$ and c is $O(1)$, the undetermined Casimir term in Eq. (4.11) is subleading. But naively, it is

comparable to the refinement we introduced in passing from the classical Penrose inequality to the QPI. However, the Casimir correction cannot be enhanced by factors proportional to $\hbar^{-1}$. Thus it is much smaller than the violations of the classical Penrose inequality that were exhibited in Sec. III. Because of the $\hbar^{-1}$ enhancement, Eq. (2.1) can be violated by a *classical* amount through quantum effects. Correspondingly, a successful QPI cannot be a small modification of the classical Penrose inequality. Indeed, it is not: as we shall demonstrate in the next section, the counterexample to Eq. (2.1) is evaded by Eq. (4.11). In this and many other interesting examples, the Casimir correction is small compared to the difference between Eq. (2.1) and Eq. (4.11).

V. EVIDENCE FOR THE QUANTUM PENROSE INEQUALITY

We will now analyze the validity of our proposal in a number of examples. In the process, we will gain some intuition about the key quantity that appears in it: $S_{\rm gen}[L]$, the generalized entropy of the future-outgoing light sheet L of a quantum marginally trapped surface μ_Q.

A. Black hole in the Unruh state

As a first example, consider a black hole formed from collapse of a null shell; see Fig. 4. This is the example we analyzed in the context of the classical Penrose inequality, at the beginning of Sec. III. We showed there that the CPI is saturated, since the area of the classically marginally trapped surface μ immediately after the collapse satisfies

$$16\pi G^2 m^2 = A[\mu]. \tag{5.1}$$

Here we are interested in a quantum marginally trapped surface with the largest generalized entropy, for which the QPI provides the greatest lower bound on the mass. The area of (quantum) trapped surfaces decreases along with the event horizon, and the contribution from the entropy term is approximately time independent. Hence we will again choose the earliest possible surface μ_Q, right after the collapse.

The quantum marginally trapped surface μ_Q must lie inside the event horizon [11], whereas μ lies outside. Therefore

$$A[\mu_Q] < A[\mu]. \tag{5.2}$$

We now turn to estimating $S_{\rm gen}[L]$. Strictly, $S_{\rm gen}[L]$ should be computed from the quantum state on a global Cauchy surface Σ that contains L. One would first compute the (divergent) field theory entropy $S[L]$ by tracing over the complement of L on Σ. One would then add the gravitational counterterms whose leading contribution is $A[\mu_Q]$. Locally, in a vacuum state, one expects $S_{\rm gen} \approx A[\mu_Q]/4G\hbar$,

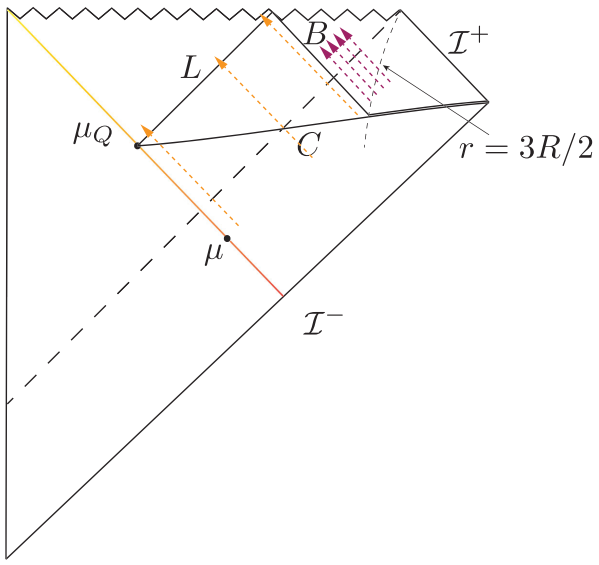

FIG. 4. Black hole formed from the collapse of a null shell (orange line). The classically marginally trapped surface μ lies a Planckian distance outside of the event horizon. The quantum marginally trapped surface μ_Q lies a Planckian distance inside the horizon. The light sheet $L(\mu_Q)$ captures $\sim\log(R/l_P)$ infalling Hawking modes (orange dashed lines); in the Unruh states these modes are unoccupied and so contribute negative entropy on L, compared to the Hartle-Hawking state. L ends at the singularity and does not encounter any later infalling modes (purple dashed lines). The entropy on L can also be computed using the mutual information, $S_L = S_C - S_B + I(L:B)$.

where G is the "infrared" value of Newton's constant that would be observed at large distances.

However, the state on L is not a standard vacuum state. L nearly coincides with the black hole horizon for a time $t \ll \Delta t_s$, where Δt_s is the scrambling time. The vacuum state on the horizon is the Hartle-Hawking state, which contains ingoing radiation. The ingoing radiation on L is entangled with modes on the other side of L. This contribution must be canceled by the counterterm so as to obtain $S_{\rm gen} \approx A[\mu_Q]/4G\hbar$ in the Hartle-Hawking state.

The actual state we consider here is the Unruh state, which does not have this ingoing radiation. As a result, the light sheet will contain less entropy than in the vacuum state. Thus

$$S_{\rm gen}[L] < \frac{A[\mu_Q]}{4G\hbar}. \tag{5.3}$$

Combined with Eqs. (5.1) and (5.2) this establishes that the QPI is satisfied (and not saturated) in this example.

We would like to go further and estimate the "gap" by which the QPI fails to be saturated in this example,

$$\Delta \equiv \frac{4\pi G}{\hbar} m^2 - S_{\rm gen}[L]. \tag{5.4}$$

We will be interested only in the order of magnitude of this gap and so will make a number of approximations. We refer to Sec. III for notation and conventions.

First, we will assume that the higher angular momentum modes, $\ell > 0$, in the near-horizon zone completely reflect off of the angular momentum barrier and so will behave as if they were in the Hartle-Hawking state. In this approximation, the Unruh state differs only through the spherical ($\ell = 0$) modes, which we treat as having no angular momentum barrier at all. We also assume that the ingoing and outgoing s waves do not interact.

A Planck sized, radially outgoing wave packet starting a Planck distance from the horizon will be redshifted in such a way that its proper distance from the horizon remains comparable to its proper wavelength, while it propagates in the near horizon zone, $r \lesssim 3R/2$. Thus, the number of independent ingoing s-wave modes captured by L is of order $\log(R/l_P)$, as shown in Fig. 4. In other words, L "sees" what enters the black hole in the first scrambling time after infalling geodesics that would have crossed μ_Q (see also Appendix A 3).

Every such mode would contribute $O(1)$ entropy in the Hartle-Hawking state but is pure in the Unruh state (since it is in the ground state). The missing entropy, and the gap to saturating the QPI, is thus

$$\Delta \sim \log \frac{R}{l_P}. \tag{5.5}$$

The entropy on null surfaces can have surprising and counterintuitive properties [27]. As a check on the above arguments, we now verify this result by evaluating $S_{\text{gen}}[L]$ using an alternative method, in which von Neumann entropies are evaluated only on spacelike hypersurfaces.[5]

The mutual information of any two systems is defined in terms of the von Neumann entropies of the individual and joint systems as follows:

$$I(L:B) \equiv S_L + S_B - S_{LB}. \tag{5.6}$$

Here we consider the light sheet L and the partial Cauchy surface B shown in Fig. 4. We take B to be null until it meets the end of the near horizon zone, $r = 3R/2$, and to coincide approximately with a constant t hypersurface outside of this radius. To stay in the semiclassical regime, one can terminate L slightly before the singularity. We can choose this terminal surface to have area cl_P^2, where $1 \ll c \ll \log(R/l_P)$. The second inequality ensures that its contribution will be subleading to our result.

Note that the joint system LB is equivalent by unitary evolution to the purely spacelike Cauchy surface C. We can thus evaluate the von Neumann entropy on L as

$$S_L = S_C - S_B + I(L:B). \tag{5.7}$$

Moreover, L and C have the same boundary, μ_Q, whereas B has a boundary of negligible area. It follows that

$$S_{\text{gen}}[L] = S_{\text{gen}}[C] - S_B + I(L:B). \tag{5.8}$$

We chose μ_Q to be just after black hole formation, so there will be no outgoing Hawking radiation present on C. In the Unruh state, the ingoing spherical modes in the near-horizon zone are unoccupied, which reduces the entropy by $\log(R/l_P)$ compared to the Hartle-Hawking value. Hence

$$S_{\text{gen}}[C] - \frac{A[\mu_Q]}{4G\hbar} \sim \log \frac{R}{l_P}. \tag{5.9}$$

In our approximation, B captures the same outgoing modes as C, but none of the ingoing modes that cross L, so $S_B = 0$. There are no data on L that are entangled with data on B, so $I(L:B) = 0$. Hence Eq. (5.7) implies $S_{\text{gen}}[L] = S_{\text{gen}}[C]$ in our example. Since $16\pi G^2 m^2 = A[\mu] = A[\mu_Q] + O(l_P^2)$, we recover Eq. (5.5).

Note that the Planck length enters Eq. (5.5) through the position of the quantum marginally trapped surface μ_Q, which is a proper distance of order l_P inside of the event horizon (or of μ). It would appear, therefore, that Δ could be minimized if one could arrange for μ_Q to lie a distance comparable to R inside the horizon. However, this requires a large perturbation of the black hole, to which the current analysis does not apply. We will revisit this question in Sec. V C.

B. Near saturation of the QPI

In the previous subsection, we found that in a newly formed Schwarzschild black hole with no exterior matter, the QPI will be satisfied but not quite saturated, with a gap of $\Delta \sim \log(R/l_P)$. The gap is only logarithmic, but it still becomes arbitrarily large for large black holes. Here we show that the logarithmic gap can be eliminated. Thus, the QPI can be saturated up to a fixed gap of order a Planck area, which we do not have full control over.

The simplest way to accomplish this is to time reverse the state of the semiclassical fields on the partial Cauchy surface C shown in Fig. 4. In our approximation, this will not affect the $\ell > 0$ modes, but it will put the spherical waves in a time-reversed Unruh state. That is, the outgoing modes will be unoccupied and the ingoing modes will be occupied, reversing the situation considered in the previous subsection. Crucially, this modification will not change the mass m at infinity, so we still have

$$16\pi G^2 m^2 = A[\mu] = A[\mu_Q] + O(l_P^2). \tag{5.10}$$

Because of the restriction to semiclassical modes, there is a cutoff near μ_Q at least of order l_P. Thus, while the

[5]We thank Aron Wall for suggesting this approach.

initial conditions we now impose are somewhat unnatural, they will persist only for one scrambling time $\Delta t_s \sim R\log(R/l_P)$. After this time, the black hole will begin to evaporate. In particular, unlike the full Boulware state, there is no singularity at the horizon. Note also that this state differs from the one we considered in Sec. III in that the $\ell > 0$ modes are not in the Boulware vacuum.

The light sheet L is sensitive only to the ingoing part of the radiation, so its generalized entropy will be the same as it would be in the Hartle-Hawking state:

$$S_{\rm gen}[L] = \frac{A[\mu_Q]}{4G\hbar}. \tag{5.11}$$

Thus we find that the QPI is nearly saturated:

$$\Delta \equiv \frac{4\pi G}{\hbar} m^2 - S_{\rm gen}[L] \sim O(1). \tag{5.12}$$

C. Perturbative regime: QPI from the GSL

Next, we will consider the more general case where matter enters into the black hole after its formation. We consider the same formation process as above. We will again focus on μ_Q right after formation so as to obtain the tightest bound. But now we will allow for a nontrivial quantum state outside of the black hole. This could be an ordinary matter system carrying some thermodynamic entropy. It could also be a quantum state with negative energy, such as the Boulware-like state that we considered in Sec. III as a counterexample to the CPI.

The future-outgoing light sheet L of μ_Q will only receive matter that falls into the black hole within the first scrambling time after μ_Q; see Fig. 4. To be precise, consider a family of radially infalling geodesics that are initially at rest at some large radius $r \gg R$. The geodesics are all at the same angle but shifted in time. It is easy to check that the geodesic that passes through μ_Q and the last geodesic that reaches L are separated at large radius by a time of order $\Delta t_s \sim R\log(R/l_P)$. Any matter that falls in later will hit the singularity before reaching Σ. This statement does not depend on the initial radius, and it also holds for ingoing null geodesics; see Appendix A 3.

In the following subsection, we will consider the effects of matter that falls in after the first scrambling time and so does not reach L. However, now we will focus on matter that can be registered on L. By the above argument, we can take this matter to reside within the near-horizon zone, $R < r < 3R/2$, on the partial Cauchy surface C. Let H be the portion of the event horizon to the future of C, and let $S_{\rm gen}[H]$ be its generalized entropy.

We begin by making a simplifying assumption that will be relaxed below, that all of the matter that falls across the horizon will also cross L (as opposed to passing through the portion of B inside the black hole). The quantum marginally trapped surface μ_Q and the boundary of H have approximately the same area, so there is a simple relationship between the entropy on H and L,

$$S_{\rm gen}[L] = S_{\rm gen}[H] - \Delta S[H_{\rm late}] + \mathcal{O}(1), \tag{5.13}$$

where $H_{\rm late}$ is the portion of the horizon above a sufficiently late Cauchy slice, when the black hole has relaxed to equilibrium, but early enough that negligible Hawking radiation has been produced.

We have assumed a state in which there is negligible mutual information between L and $H_{\rm late}$. For example, if the black hole simply evaporates with no further matter falling in, $\Delta S[H_{\rm late}]$ is the (negative) renormalized entropy that exists on the horizon in the Unruh state (due to the missing infalling modes when compared to the Hartle-Hawking state).

From

$$S_{\rm gen}[H_{\rm late}] - \Delta S[H_{\rm late}] = \frac{A_{\rm late}}{4G\hbar} \tag{5.14}$$

and Eq. (5.13), the QPI follows:

$$\begin{aligned} S_{\rm gen}[L] &= S_{\rm gen}[H] - \Delta S[H_{\rm late}] \\ &\le S_{\rm gen}[H_{\rm late}] - \Delta S[H_{\rm late}] = \frac{A_{\rm late}}{4G\hbar} \le \frac{4\pi G}{\hbar} m^2. \end{aligned} \tag{5.15}$$

The first inequality in this sequence is the GSL for event horizons. Note that we have ignored the $\mathcal{O}(1)$ additive uncertainty in Eq. (5.13) in light of the discussion at the end of Sec. IV.

This argument establishes the QPI for a large class of examples, including the Boulware-like state that served as a counterexample to the classical Penrose inequality in Sec. III. In this case, $A_{\rm late}$ (which sets the mass) will be significantly smaller than the area of the trapped surface μ. Here we use the quantum trapped surface μ_Q, but its area is almost the same as that of μ. What saves the QPI is the contribution of the entropy on L, which is negative in this example. Specifically, the GSL guarantees that the lower bound, $S_{\rm gen}[L]$, is smaller than the area of μ_Q by a sufficient amount for the QPI to hold.

In the case where positive entropy registers on H and L, our QPI is stronger than the classical Penrose inequality. The light sheet "knows" that more matter will enter the black hole after μ_Q, and the GSL "knows" that this will result in an area increase. Effectively, this larger area becomes the lower bound on the mass.

D. Failed counterexample: Negative energy that misses the light sheet

In the previous subsection, we considered the case where all matter outside the quantum trapped surface μ_Q crosses its light sheet L. Here we generalize to discuss matter for

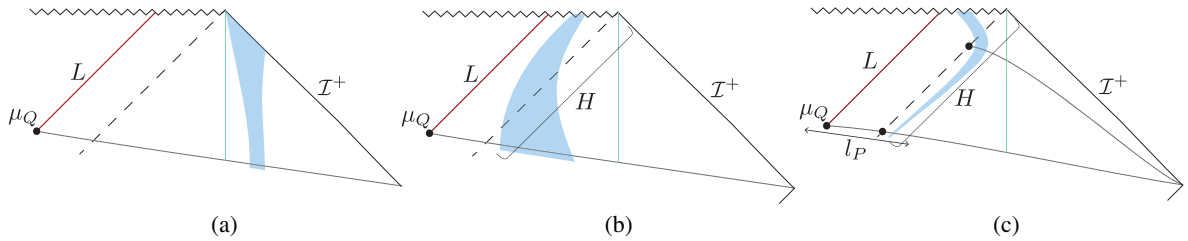

FIG. 5. The QPI is threatened by any negative energy (blue world volume) that fails to register on the light sheet L. We analyze three possibilities but find that none of them leads to a violation of the QPI. (a) Negative energy outside of the near horizon zone (vertical green line). (b) Negative energy that enters the black hole soon after μ_Q but evades L by accelerating outward. (c) Negative energy that remains near the black hole for more than a scrambling time.

which this does not happen. In this case, we cannot use the GSL for the event horizon to constrain the relation between $S_{\text{gen}}[L]$ and the mass at infinity. However, we will give some plausibility arguments for the validity of the QPI.

In the previous subsections, we argued that the QPI will hold true if all matter outside of μ_Q passes through L. We can think of the present situation as a complication where we add matter that does not satisfy this property. Since this cannot affect $S[L]$, the only way that the QPI can now be violated is if the matter we added contributes negative mass at infinity. We will now argue that this is impossible in the semiclassical regime.

Matter outside of μ_Q can fail to register on L for any of the following three reasons (see Fig. 5):

(1) The matter never enters the black hole.
(2) The matter enters the black hole during the first scrambling time after C but escapes through the portion of B inside the black hole.
(3) The matter enters the black hole later than a scrambling time after C.

In the first case, the matter can be approximately treated as isolated from the black hole. But the total mass of isolated systems is positive, so distant systems can never cause violations of the QPI. (This does not rule out regions with negative energy, but it implies that sufficient positive energy must be present nearby.)

In the second case, the matter system can be initially near the black hole and so could have regions of negative energy density (as in the example of Sec. III). However, in order to miss L, it would have to accelerate outwards after crossing the horizon. This requires positive energy. We will not attempt to demonstrate here that this always results in a net positive mass contribution; our goal is only to note that the QPI is not obviously violated in this setup. This question merits further study.

In the third case, we again must choose the matter system to be close to the horizon if we wish to give it negative energy. For example, the Boulware-like state of Sec. III would qualify. However, by assumption this state would have to be present more than one scrambling time after C. Moreover, the modes for which it is possible to obtain net negative energy are those that make up the thermal atmosphere of the black hole; these modes evolve exponentially close to the horizon under backward time evolution. Thus the state on C would contain trans-Planckian energy density (similar to a firewall). The initial state would not be a semiclassical state. This argument is robust and rules out an entire class of what naively seemed like promising counterexamples. We view this as nontrivial evidence in favor of our proposal.

VI. ALTERNATIVE PROPOSALS

In this section we consider various alternative conjectures for the QPI. In Sec. VI A we give counterexamples to proposals that might otherwise seem natural. In Sec. VI B we discuss modifications of our proposal that appear viable, and we explain why we are not currently advocating for them.

A. Nonviable alternatives

We will now discuss several alternative conjectures for a QPI that we considered in the process of this work. Our goal is to explain our choice in Sec. IV and to illustrate that the problem is rather constrained. This proves neither that our formulation is unique nor that it is correct. But we will see that it is remarkably difficult to find any alternative statement of the QPI that is not immediately ruled out.

1. Cauchy surfaces that reach spatial infinity

First, we explain why we do not allow $\Sigma[\mu_Q]$ to reach outside the black hole. This prohibition is motivated by the asymptotically flat case, to which we will specialize for now. Let Σ_∞ be a Cauchy surface of $O_W[\mu_Q]$, in violation of our requirements. An example is the black slice in Fig. 6. Let $S_{\text{gen}}[\Sigma_\infty(\mu_Q)]$ be the generalized entropy evaluated on Σ_∞. The alternative QPI thus would take the form

$$m \overset{?}{\geq} \sqrt{\frac{\hbar}{4\pi G} S_{\text{gen}}[\Sigma_\infty(\mu_Q)]}\,. \tag{6.1}$$

But it is easy to find a counterexample to Eq. (6.1): an arbitrary amount of matter entropy can be placed in regions

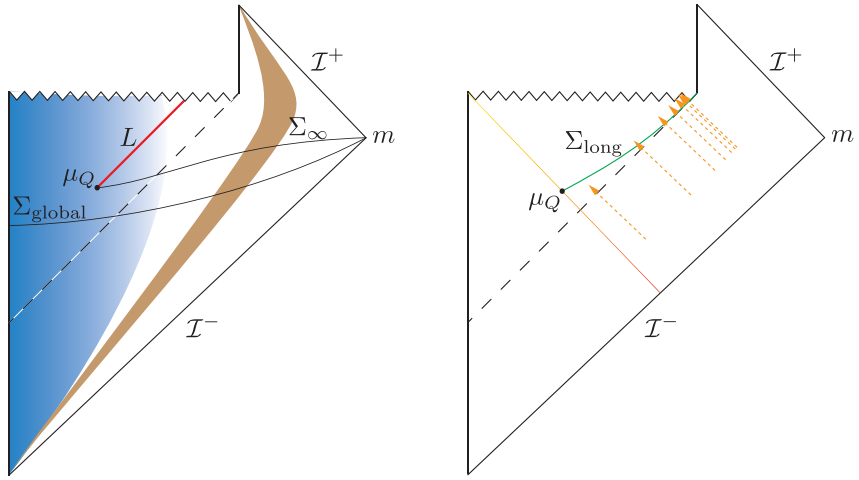

FIG. 6. Left: The generalized entropy on the slice Σ_∞ can be dominated by distant soft particles (brown) and so does not yield a viable lower bound on the mass. The global Cauchy surface Σ_{global} plays a role in an alternative proposal discussed in the main text. Right: The long slice Σ_{long} captures all of the missing infalling Hawking modes.

far from the black hole, at arbitrarily little cost in mass. We now discuss this in detail.

Consider a dilute gas of N photon wave packets, each of characteristic size λ. Each photon occupies a region of volume λ^3, so the photons can be dilute if they occupy a region of volume $N\lambda^3$. We can take each photon to be in a mixed state (say, of polarizations), and in a product state with respect to the rest of the universe. Then the gas contributes of order N to the generalized entropy on Σ.

We take the gas to be very far from the black hole or any other matter, so that gravitational binding energy to other objects is negligible. The gravitational binding energy of the photon cloud itself will be negligible if $NG\hbar/\lambda \ll N^{1/3}\lambda$, so we shall take $\lambda \gg N^{1/3} l_P$, where $l_P \equiv (G\hbar)^{1/2}$ is the Planck length. Then the gas of photons contributes a mass of order $N\hbar/\lambda$ to the ADM mass. This mass contribution can be taken to be arbitrarily small by taking $\lambda \to \infty$ at fixed N without violating any of the previous assumptions.

We are still free to choose N to take any value we like. Thus we have found a family of initial data with bounded m but unbounded $S_{\text{gen}}[\mu_Q] \approx c_1 + c_2 N$, where c_1 and c_2 are independent of N. For large enough N, this leads to a violation of Eq. (6.1).

2. Area of marginally quantum trapped surfaces

A second alternative conjecture would be to use only the area of μ_Q, not its generalized entropy:

$$m \overset{?}{\geq} \sqrt{\frac{A[\mu_Q]}{16\pi G^2}}. \tag{6.2}$$

That is, one would conjecture that Eq. (2.1) holds if A is taken to be the area of a quantum trapped surface. This possibility is attractive because the entropy of distant soft radiation would never contribute to the lower bound in the first place.

However, Eq. (6.2) is ruled out (among other reasons) by the Boulware-like counterexample to the classical Penrose inequality. This is because the area of the bifurcation surface will receive only a correction that can be made parametrically small. This follows from the remarks concerning the classically marginally trapped surface at the end of Sec. III. The same argument implies that the marginally quantum trapped surface area receives only a parametrically small correction, which cannot compete with the large decrease in mass.

3. Subtracting global entropy; interior generalized entropy

Let us revisit the proposal of Sec. VI A and consider the generalized entropy $S_{\text{gen}}[\Sigma_\infty(\mu_Q)]$ of a marginally trapped surface μ_Q, evaluated on a Cauchy surface that reaches outside of the black hole all the way to spatial infinity. This proposal suffered from the problem that distant soft modes can contribute unbounded entropy with bounded energy, so $S_{\text{gen}}[\Sigma_\infty(\mu_Q)]$ is unrelated to any lower bound on the mass.

A natural idea is to subtract the von Neumann entropy on a global Cauchy surface (see Fig. 6):

$$m \overset{?}{\geq} \sqrt{\frac{\hbar(S_{\text{gen}}[\Sigma_\infty(\mu_Q)] - S[\Sigma_{\text{global}}]}{4\pi G}}. \tag{6.3}$$

If the distant soft modes have the same entropy in the global state as in the generalized entropy, then their dangerous contribution will cancel out.

However, this need not be the case. Consider a collapsing star that forms a Schwarzschild black hole of area A. The entropy of the star can be of order $S_{\text{star}} \sim (A/G\hbar)^{3/4}$ or even $S_{\text{star}} \sim A/G\hbar$ [28]. We can chose the global state to

contain only distant soft radiation that purifies the star, so that $S[\Sigma_{\text{global}}] = 0$ and

$$m = \sqrt{\frac{A[\mu_Q]}{16\pi G^2}} + \epsilon, \tag{6.4}$$

where ϵ can be arbitrarily small. But then

$$S_{\text{gen}}[\Sigma_\infty(\mu_Q)] \approx \frac{A[\mu_Q]}{4G\hbar} + S_{\text{star}}, \tag{6.5}$$

so that Eq. (6.3) is violated.

The violation in our example remains bounded, since S_{star} cannot exceed $A[\mu_Q]/4G\hbar$ by the GSL. One might consider absorbing this violation by adding a correction factor of $1/2$ to the right-side of Eq. (6.3). But by considering initial data with a second asymptotic region, one can arrange $S[\Sigma_{\text{global}}] = 0$ with unbounded $S_{\text{gen}}[\Sigma_\infty(\mu_Q)]$ at fixed m, leading to unbounded violations.

A variation of this idea is to use the generalized entropy in the interior (not the exterior) of the surface μ_Q. It is easy to check that it fails for the same reasons.

B. Possible modifications of the QPI

We will now discuss an alternative formulation of the QPI that we cannot currently rule out, and we comment on some of its properties that have led us to reject it as our main proposal.

The basic idea is to consider partial Cauchy surfaces other than L, still bounded by μ_Q and remaining inside the black hole. For example, we could assert that

$$m \geq \sqrt{\frac{\hbar S_{\text{gen}}[\Sigma]}{4\pi G}} \tag{6.6}$$

holds for any achronal hypersurface $\Sigma \subset B \cap O_W[\mu_Q]$ whose only boundary is μ_Q. This class includes the light sheet L, so this conjecture would be strictly stronger than our main proposal. It is clear that the heuristic arguments in support of QPI in Sec. V also apply to this family of slices.

There are some clear downsides to this choice. The region B and therefore this family of slices are defined teleologically. Furthermore, it is not clear to us how one would formulate a minimality requirement in this case, analogous to the requirement that the classically trapped surface minimize the area on some Cauchy surface.

A variation would be to insist on a Cauchy surface that is as "long" as possible, i.e., which does not have any end point on the future singularity. Roughly, this means it ends on the future end points of the horizon generators; see Σ_{long} in Fig. 6. This proposal is weaker than the previous one and is neither stronger nor weaker than our main proposal. We will now argue that for an evaporating black hole this results in a less stringent bound than the one obtained from L.

As discussed in Sec. V, in the Unruh state there is negative entropy falling across the horizon, due to the missing ingoing modes compared to the Hartle-Hawking state. The long slice will capture this negative entropy through the entire process of evaporation. (Here we are assuming that the semiclassical expansion is valid until the black hole area is Planckian in size.) The generalized entropy on this slice is

$$S_{\text{gen}}[\Sigma_{\text{long}}] = \frac{A[\mu_Q]}{4G\hbar} - \gamma\frac{A[\mu_Q]}{4G\hbar}, \tag{6.7}$$

where $\gamma \geq 1$ by the GSL, and the second term arises from the contribution of the missing ingoing modes on Σ.

It is difficult to compute γ exactly. If $\gamma > 1$, then S_{gen} will be negative. This renders (6.2) ill defined. Negative S_{gen} is also conceptually in conflict with the interpretation of S_{gen} as an entropy in the fundamental theory of quantum gravity. This suggests that a careful computation will reveal that $\gamma = 1$, in which case Eq. (6.2) reduces back to the statement of the positivity of the ADM mass. Along with the downsides mentioned earlier, this conundrum shows that such long slices are not ideal for formulating the QPI.

VII. QUANTUM PENROSE INEQUALITY IN ANTI–DE SITTER SPACE

The classical Penrose inequality was motivated by the heuristic argument that a Schwarz-schild black hole with no exterior matter should have the smallest possible mass for a given trapped surface area. In Eq. (2.1) we assumed a vanishing cosmological constant Λ. An analogous argument for asymptotically Anti–de Sitter spacetimes with curvature scale $L = (-\Lambda/3)^{1/2}$ yields the classical inequality

$$m \geq f_{\text{AdS}}(A[\mu]), \tag{7.1}$$

where

$$f_{\text{AdS}}(A) \equiv \left(\frac{A}{16\pi G^2}\right)^{1/2} + \left(\frac{A}{16\pi G^2}\right)^{3/2}\frac{G^2}{L^2} \tag{7.2}$$

and μ is again a trapped surface satisfying an appropriate minimality condition (see Sec. II).

Following our QPI proposal for asymptotically flat space, it would appear natural to propose the following QPI in asymptotically AdS spacetimes:

$$m \overset{?}{\geq} \left(\frac{\hbar S_{\text{gen}}}{4\pi G}\right)^{1/2} + \left(\frac{\hbar S_{\text{gen}}}{4\pi G}\right)^{3/2}\frac{G^2}{L^2} \tag{7.3}$$

in asymptotically AdS spacetimes with curvature scale L. Here S_{gen} is defined with respect to slices defined in Sec. IV C; see Fig. 7. However, because of $\mathcal{O}(1)$ subtleties discussed at the end of Sec. IV C, it is not clear that

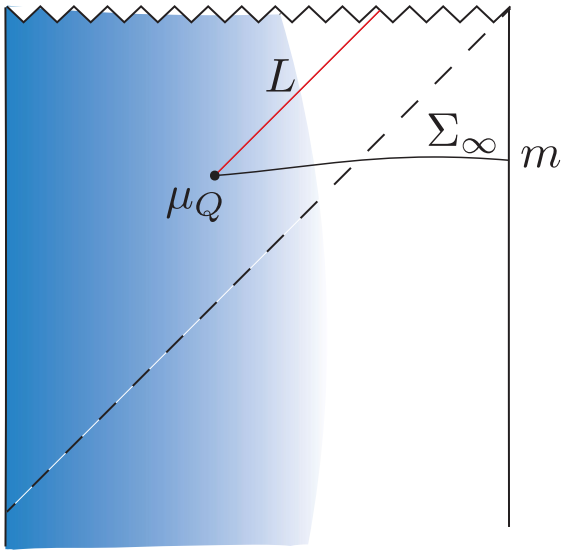

FIG. 7. Different choices of slices anchored to the surface μ_Q on which one could compute $S_{\rm gen}$. The red light sheet L is defined analogously to the asymptotically flat case. Since distant soft modes do not exist for large black holes in AdS, one could also consider computing $S_{\rm gen}$ on the black slice Σ_∞ that ends on the asymptotic boundary.

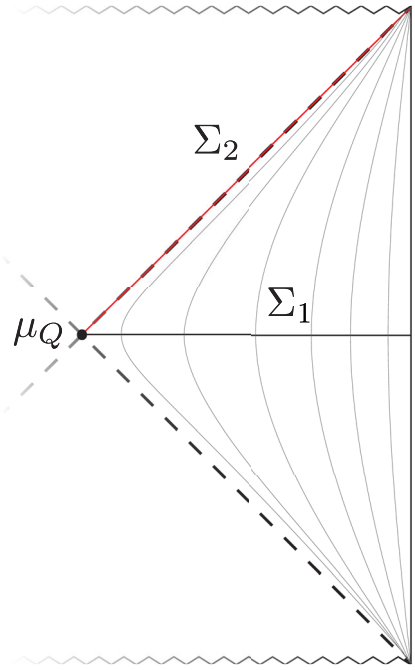

FIG. 8. The Hartle-Hawking state is essential for our definition of $f^q_{\rm AdS}$ via $m = f^q_{\rm AdS}(S_{\rm gen})$. Here m is the ADM mass including the quantum corrections associated with the radiation mass. $m_{\rm rad}$ is computed on the black slice Σ_1 with respect to the timelike Killing field t^μ whose orbits are shown in the figure. $S_{\rm gen}$ is computed on the red null slice Σ_2 on the horizon that ends on the bifurcation surface μ_Q.

Eq. (7.3) will hold exactly in the AdS Hartle-Hawking state (referred to as σ henceforth). The issue is the radiation mass outside of the black hole which could be negative, lowering the left-hand side of Eq. (7.3) to violation. As we will discuss here, in asymptotically AdS spacetimes one could fix this $\mathcal{O}(1)$ issue. Note that the quantum-corrected ADM mass in this state is

$$m = \left(\frac{A}{16\pi G}\right)^{1/2} + \left(\frac{A}{16\pi G}\right)^{3/2}\frac{G^2}{L^2} + m_{\rm rad}, \tag{7.4}$$

with

$$m_{\rm rad} = \int_{\Sigma_1} d\Sigma^\nu t^\mu \langle T_{\mu\nu}\rangle_\sigma, \tag{7.5}$$

where $\langle T_{\mu\nu}\rangle_\sigma$ is the renormalized stress tensor in σ, Σ_1 is a Cauchy slice stretching from the bifurcation surface to the boundary of AdS, and t^μ is the Killing field in Schwarzschild-AdS that is timelike at infinity. Also, note that the area term in Eq. (7.4) is not the quantum-corrected area. Furthermore, based on formulation in Sec. IV C, $S_{\rm gen}$ in the σ is computed on the part of the horizon in the future of the bifurcation surface μ_Q; see Fig. 8. The quantum stress tensor $\langle T_{\mu\nu}\rangle$ has been computed in the Hartle-Hawking state in $2+1$ dimensions with different choices of boundary conditions [29]. One finds that $m_{\rm rad}$ depends on the field content and the boundary conditions; moreover, $m_{\rm rad}$ does not have a definite sign [29]. Explicit calculations in $2+1$ dimensions show that $m_{\rm rad}$ can be negative. We do not expect that the entropy of the matter on Σ_2 and the quantum corrections to the area term would compensate for this negative value of $m_{\rm rad}$ so as to uphold Eq. (7.3). Therefore, we expect that Eq. (7.3) can be violated in the Hartle-Hawking state. Furthermore, the nonuniversality of $m_{\rm rad}$ seems to suggest that the correct formulation of QPI for large AdS black holes must depend on various factors that $m_{\rm rad}$ depends on (e.g., the field content and the boundary conditions).

Here we propose a way to introduce this dependence into a quantum Penrose inequality for asymptotically anti–de Sitter spacetimes. Let $f^q_{\rm AdS}$ be a function such that in the Hartle-Hawking state,

$$m = f^q_{\rm AdS}(S_{\rm gen}[\Sigma_2]), \tag{7.6}$$

where m is the quantum-corrected ADM mass and $S_{\rm gen}$ is associated with the future portion Σ_2 of the horizon; see Fig. 8. Now, we propose

$$m \geq f^q_{\rm AdS}(S_{\rm gen}[L]), \tag{7.7}$$

for any marginally trapped surface μ_Q in an asymptotically AdS spacetime with a large AdS black hole. A heuristic argument for Eq. (7.7) is as follows: First, the above

inequality will follow from the classical Penrose inequality unless we are in a state perturbatively close to Kerr-AdS. In that limit, it can be shown (see Appendix B) that given any quantum marginally trapped surface, there exists a Q screen that approaches the horizon of the Kerr-AdS at late times and has the quantum marginally trapped surface as a leaf. As discussed in Sec. IV, Q screens are known to satisfy a generalized second law [9]. The QPI would then follow from

$$\begin{aligned} S_{\text{gen}}|_{\text{early}} &\leq S_{\text{gen}}|_{\text{Kerr-AdS}} \\ &\Rightarrow f^q_{\text{AdS}}(S_{\text{gen}}|_{\text{early}}) \leq f^q_{\text{AdS}}(S_{\text{gen}}|_{\text{Kerr-AdS}}) \\ &\leq f^q_{\text{AdS}}(S_{\text{gen}}|_\sigma) = m, \end{aligned} \tag{7.8}$$

where the first inequality on the second line follows from the generalized second law of Q screens and the second inequality follows from assuming f^q_{AdS} is a monotonic function.

In general we could have states where the AdS black hole is not large enough to reach stable thermal equilibrium with the asymptotic boundary of the spacetime, so a few words about the case of these small AdS black holes are in order. For such black holes, we cannot define the function f^q_{AdS} as above. Our proposal would then follow more closely our proposal for asymptotically flat spacetimes, where we formulate our conjecture using the function f appearing in the classical Penrose inequality for AdS,

$$m \geq f_{\text{AdS}}(4G\hbar S_{\text{gen}}[L]), \tag{7.9}$$

where f_{AdS} is defined in Eq. (7.2). The phase transition for (in)stability of AdS black holes happen around ADM mass L/G, so our proposal changes for the mass above and below the phase transition point. The exact value of mass associated with a phase transition depends on the choice of boundary conditions and the field content.

An important difference between QPI for large AdS black holes and flat space black holes is the absence of the challenge associated with soft modes. As discussed in Sec. VI, in asymptotically flat space, one can add entropy far away from the black hole at negligible cost to the ADM mass. This prevents any formulation of QPI where the generalized entropy is computed on partial Cauchy slices approaching spatial infinity in asymptotically flat spacetimes.

However, in asymptotically AdS spacetimes and in the presence of a large black hole, excitations require considerable energy to remain outside of the black hole, so the arguments of Sec. VI do not go through and matter entropy outside of the black hole has an energy cost. Therefore, in the presence of large AdS black holes the slice on which S_{gen} is evaluated could end on the asymptotic boundary of AdS. This possibility was discussed in the context of AdS/CFT in [20]. To define the function f^q_{AdS} in this version of QPI, we need to consider the Hartle-Hawking state and the generalized entropy on the spatial slice Σ_1 of Fig. 8,

$$m = f^q_{\text{AdS}}(S_{\text{gen}}[\Sigma_1]). \tag{7.10}$$

The quantum extremal surface prescription [30] equates $S_{\text{gen}}[\Sigma_1]$ with the von Neumann entropy of the dual CFT in the thermofield double state. Therefore, this definition of the function f^q_{AdS} has a very natural interpretation from the CFT perspective

$$\langle \text{H} \rangle_{\text{TFD}} = f^q_{\text{AdS}}(S_{\text{CFT}}[\text{TFD}]), \tag{7.11}$$

where $\langle H \rangle_{\text{TFD}}$ is the expectation value of the CFT Hamiltonian in the thermofield double state.

VIII. CLASSICAL AND NONGRAVITATIONAL LIMITS

In this section we discuss two interesting limits of the QPI: the classical limit, $\hbar \to 0$; and the nongravitational limit, $G \to 0$.

In the $\hbar \to 0$ of QPI, we recover the classical Penrose inequality. This is easy to see. The amount of matter entropy on L is $\mathcal{O}((G\hbar)^0)$, and therefore

$$\lim_{\hbar \to 0} 4G\hbar S_{\text{gen}}[L] = A[\mu_Q]. \tag{8.1}$$

Furthermore, the surface μ_Q is perturbatively close to a (classically) marginally trapped surface such that their area difference is due to quantum corrections and therefore of order $G\hbar$ and can be neglected. Last, any $\hbar$ corrections to the function f can trivially be ignored in the $\hbar \to 0$ limit. We therefore have the desired implication:

$$f^q(4G\hbar S_{\text{gen}}[\mu_Q]) \leq m \overset{\hbar \to 0}{\Rightarrow} f^c(A[\mu]) \leq m. \tag{8.2}$$

We turn to the $G \to 0$ limit of the QPI. This is of interest because some semiclassical conjectures yield nontrivial and novel implications about QFT in this limit. For example, the quantum null energy condition (QNEC) was first discovered by taking the $G \to 0$ limit of the QFC in a particular setting [10]. In order to sidestep the small "Casimir uncertainty" discussed in Sec. IV C, we will consider the QPI in AdS. We further restrict to two complementary scenarios.

First, consider a perturbation to a Hartle-Hawking state such that in a finite amount of time the state settles back down to a Hartle-Hawking state (with a different temperature). In this case, Eq. (7.8) shows that the QPI is equivalent to the GSL. The nongravitational limit of the GSL is the monotonicity of relative entropy. This is a nontrivial but well-known statement in quantum information theory, which applies in particular in QFT.

The second scenario is when the perturbation does not relax to equilibrium. This means that the excitation that takes the state away from the Hartle-Hawking state remains outside of the black hole. Therefore such excitations do not change the generalized entropy on L or the geometry of the event horizon. Let δm be the change in the ADM mass caused by this perturbation. Since the QPI is saturated in the Hartle-Hawking state, it reduces to

$$\delta m \geq 0 \tag{8.3}$$

in the $G \to 0$ limit. Here $\delta m = \int_{\Sigma_1} d\Sigma^\mu \xi^\nu T_{\mu\nu}$ (see Fig. 8), and t^μ is the timelike Killing vector field outside the black hole. This makes physical sense: if the field excitations are isolated from the black hole, they need to satisfy their own positive energy condition.

IX. COSMIC CENSORSHIP CONJECTURE

In this section, we consider the current status of the cosmic censorship conjecture (CCC) and its relation to the Penrose inequality. We argue that there is a need for a quantum generalization of the CCC, and we suggest that the proposed quantum Penrose inequality may inform the formulation of a quantum CCC.

The formation of singularities in gravitational collapse is guaranteed by classical [5] and quantum [11] singularity theorems. However, it is not clear that the formation of a singularity implies the formation of a black hole.

The weak CCC asserts that singularities (regions of arbitrarily high curvature) will not be visible to a distant observer.[6] A precise statement of the conjecture can be formulated as follows [4]: Let $(\Sigma,\ h_{\mu\nu},\ K_{\mu\nu})$ be an asymptotically flat initial dataset for Einstein's equation with $(\Sigma,\ h_{\mu\nu})$ a complete Riemannian manifold. Let the matter sources be such that $T_{\mu\nu}$ satisfies the dominant energy condition and the coupled Einstein-matter field equations are of the form $\Box\phi(x) = F(x, \phi, \nabla_\mu\phi)$, where F is a smooth function of its variables. In addition, let the initial data for the matter fields on Σ satisfy appropriate asymptotic falloff conditions at spatial infinity. Then the maximal Cauchy evolution of these initial data is an asymptotically flat, strongly asymptotically predictable spacetime.

The CCC has not been proven. Indeed, there are a number of known "mild" violations that we will discuss shortly. The (classical) Penrose inequality is only a necessary condition for the CCC, as explained in Sec. II. Even this weaker statement has not been proven; but as a quantitative relation between mass and area, it has been extensively explored. The fact that no counterexample has been found can be viewed as indirect evidence that some version of the CCC may indeed hold.

Let us now discuss the mild violations mentioned in the previous paragraph. A black string in $4+1$ dimensions suffers from the Gregory-Laflamme instability [31]. Further evolution causes the string to become arbitrarily thin in some regions [32,33] and so arbitrarily high curvatures become visible to a distant observer.

In $3+1$ dimensions, there exist fine-tuned initial datasets such that the solution exhibits a self-similar behavior near the threshold of formation of a black hole [34–37]. At the threshold, a naked singularity forms. In some examples, the naked singularity propagates out to $\mathcal{I}^+$.

In the above two examples, the initial data satisfy the dominant energy condition, as required by the CCC. The black string is not asymptotically flat, but one expects that it can be truncated at a sufficiently great length so that local evolution far from the ends still leads to a naked singularity.

Let us now add a third example, which is physically relevant but does not obey the dominant energy condition: a black hole that evaporates completely. In this case, treating the spacetime as a classical manifold, a naked singularity is inevitable [38,39].

Only the last example explicitly involves quantum effects. But it points to a resolution of all three violations: clearly, it makes no sense to treat the spacetime as a classical manifold near the end point of evaporation (i.e., arbitrarily close to the naked singularity). When the curvature formally exceeds the Planck curvature, the semiclassical expansion breaks down, and a classical geometric description of the spacetime need not exist.

But this observation also applies to the other known examples of CCC violation. One would expect a black string to pinch off before it becomes thinner than a Planck length. Similarly, one would expect quantum effects to smooth out the fine-tuned initial data, or at least the singularities they lead to.

Naively, all three examples violate the spirit of the CCC: starting from a highly classical regime, evolution produces an outcome in which quantum gravity is required to maintain predictability. But in an important sense, the violation is "small" in each case. The energies involved are likely no greater than the Planck mass, and we can "guess" a plausible future evolution without having a full quantum gravity theory. For example, the Planck-sized black hole will probably decay into a few more particles, and the black string will simply pinch off.

It would be of interest to formulate a quantum version of the CCC that accounts for these physically reasonable phenomena, i.e., one that is not formally violated by them.[7] We expect the quantum Penrose inequality to play a role analogous to the classical one: as a necessary condition for the quantum CCC, and thus as a useful test. Perhaps more

[6] We will consider only the weak CCC here. The strong form of the CCC states, roughly, that no observer can see a singularity. In all cases, one assumes regular initial data.

[7] A specific proposal will be studied in forthcoming work.

importantly, the quantum Penrose inequality may be of some use in identifying the correct formulation of a quantum CCC in the first place.

ACKNOWLEDGMENTS

We thank Roberto Emparan, Netta Engelhardt, Gary Horowitz, Donald Marolf, Robert Myers, and Aron Wall for discussions. This work was supported in part by the Berkeley Center for Theoretical Physics; by the Department of Energy, Office of Science, Office of High Energy Physics under QuantISED Award No. DE-SC0019380 and Contract No. DE-AC02-05CH11231; and by the National Science Foundation under Grant No. PHY1820912. M. T. acknowledges financial support coming from the innovation program under ERC Advanced Grant No. GravBHs-692951.

APPENDIX A: (QUANTUM) TRAPPED SURFACES IN THE SCHWARZSCHILD GEOMETRY

1. Classical solution and semiclassical corrections

The Schwarzschild metric is

$$ds^2 = -\left(1-\frac{R}{r}\right)dt^2 + \frac{dr^2}{1-R/r} + r^2 d\Omega^2, \quad \text{(A1)}$$

where $R = 2GM$ is the Schwarzschild radius. In ingoing Eddington-Finkelstein coordinates,

$$ds^2 = -\left(1-\frac{R}{r}\right)dv^2 + 2dvdr + r^2 d\Omega^2, \quad \text{(A2)}$$

where

$$v = t + r_*, \qquad r_* = r + R\log\left|\frac{r}{R}-1\right|, \qquad \frac{dr}{dr_*} = 1-\frac{R}{r}. \quad \text{(A3)}$$

Ingoing radial null congruences are at constant v, so $dv = 0$. Outgoing null congruences satisfy $dv = 2dr_*$, so

$$v = 2r_* + \text{const}. \quad \text{(A4)}$$

We are interested in their expansion,

$$\theta = \frac{dA/d\lambda}{A} \quad \text{(A5)}$$

in terms of a convenient affine parameter, λ.

To find λ, first note that r is an affine parameter. This follows because $A = 4\pi r^2$, so

$$\theta = \frac{2}{r}\frac{dr}{d\lambda}; \quad \text{(A6)}$$

and Raychaudhuri's equation in the vacuum, for spherical symmetry, reduces to

$$\frac{d\theta}{d\lambda} + \frac{1}{2}\theta^2 = 0. \quad \text{(A7)}$$

This implies that $dr/d\lambda$ must be constant for any affine λ. We can take that constant to be 1 if we like, and choose another constant of integration so that $r = \lambda$.

However, this choice is not convenient for outgoing light rays, because we are interested in radial null congruences near and on the event horizon,

$$|r - R| \ll R. \quad \text{(A8)}$$

Intuitively, the radius r does not change much for these congruences, so small changes in r correspond to large motions along the congruence. On the horizon, r is degenerate, and inside the black hole, r runs toward the past.

To remedy this, let us consider the coordinate distance $c = r - R$ from the horizon. We will work in the near-horizon limit of Eq. (A8), i.e., to first order in $c/R \ll 1$. For example, $r_* = R + R\log(|c|/R)$ in this approximation; and by Eq. (A4), an outgoing congruence satisfies $v = 2R\log(|c|/R) + \text{const}$. Inverting this, we find

$$c = c_0 e^{v/2R}, \quad \text{(A9)}$$

where c_0 is the coordinate distance from the horizon at $v = 0$. This is the quantity that vanishes on the horizon and goes negative inside, so we can define a nondegenerate, always future-directed parameter by choosing $\lambda = c/c_0$. This is affine since $\lambda = (r-R)/c_0$ and r is affine.

To summarize, we choose the affine parameter

$$\lambda = e^{v/2R} \quad \text{(A10)}$$

on outgoing null geodesics near the horizon. By Eq. (A6), the expansion of any such congruence is given by

$$\theta = \frac{2c_0}{R}, \quad \text{(A11)}$$

where we again used $r - R \ll R$. All surfaces on the event horizon have $c_0 = 0$ and hence $\theta = 0$; they are marginally outer trapped. It is easy to check that these are the only such surfaces.

Any null vector tangent to the outgoing congruences must be proportional to $\partial_t + \partial_{r_*}$. Let k^a be the particular null vector associated with the affine parameter λ. From Eq. (A10) we have

$$k = \frac{d}{d\lambda} = \frac{2R}{\lambda}\frac{d}{dv}\bigg|_{\text{cong}} = \frac{R}{\lambda}(\partial_t + \partial_{r_*}). \quad \text{(A12)}$$

For the second equality, we used that on the outgoing congruence $t = (v + \text{const})/2$, $r_* = (v - \text{const})/2$.

For all ingoing spherical congruences in the region covered by the ingoing Eddington-Finkelstein coordinates, $-r$ is a future-directed nondegenerate affine parameter. Thus Eq. (A6) implies that their expansion, θ_l, is everywhere negative. This establishes that every spherical cut of the event horizon is marginally trapped, i.e., satisfies $\theta = 0$ and $\theta_l \leq 0$.

To treat quantum matter as a small perturbation, we expand the Einstein equation, $G_{ab} = 8\pi G\langle T_{ab}\rangle$, in powers of $G\hbar$, to first order. (We drop the expectation value symbol below.) In this approximation, we can compute matter effects on the expansion of congruences by integrating the Raychaudhuri equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \varsigma^2 - 8\pi G T_{kk}. \tag{A13}$$

Here $T_{kk} = T_{ab}k^a k^b$, and $k^a = (\frac{d}{d\lambda})^a$ is the affine tangent vector to the null congruence. The shear term vanishes for the spherical congruences we consider. In general, the θ^2 term will be $O((G\hbar)^0)$ and thus dominant.

However, here we will be interested in surfaces where classical and quantum effects compete. Such surfaces must have $\theta \sim O(G\hbar)$ classically. By Eq. (A11) they are found in a neighborhood $|c| \leq O(G\hbar)$ of the event horizon. Hence $\theta^2 \sim O((G\hbar)^2)$ will be negligible in the region of interest, and Eq. (A13) reduces to

$$\theta(\lambda) - \theta(\lambda_0) = -8\pi G \int_{\lambda_0}^{\lambda} T_{kk}. \tag{A14}$$

2. Classically trapped surfaces during evaporation

We will now compute the effect of the quantum stress tensor for the Unruh state [25] on the position of (marginally) trapped surfaces in the Schwarzschild geometry.

The renormalized stress tensor in the Unruh vacuum takes the form

$$\langle U|T^b_a|U\rangle_{\text{ren}} \xrightarrow{r\to 2M} \frac{L}{4\pi R^2}\begin{pmatrix} f^{-1} & -1 \\ f^{-2} & -f^{-1}\end{pmatrix}, \tag{A15}$$

where $f = (1 - R/r)$, $R = 2M$, a and b range over t and r, and

$$L \sim \frac{\hbar}{R^2} \tag{A16}$$

is the luminosity of the black hole. Lowering indices we find

$$\langle U|T_{ab}|U\rangle_{\text{ren}} \xrightarrow{r\to 2M} \frac{L}{4\pi R^2}\begin{pmatrix} -1 & -f^{-1} \\ -f^{-1} & -f^{-2}\end{pmatrix}. \tag{A17}$$

Using

$$\partial_{r*} = \frac{dr}{dr*}\partial_r = \left(1 - \frac{R}{r}\right)\partial_r, \tag{A18}$$

we can express the null vector k in (t, r) coordinates,

$$k = \frac{R}{\lambda}\left(\partial_t + \left(1 - \frac{R}{r}\right)\partial_r\right) = k^t\partial_t + k^r\partial_r, \tag{A19}$$

and we obtain

$$\begin{aligned}\langle T_{\mu\nu}k^\mu k^\nu\rangle &= \langle T_{tt}k^t k^t\rangle + \langle T_{rr}k^r k^r\rangle + 2\langle T_{tr}k^t k^r\rangle \\ &= -\frac{L}{\pi\lambda^2} = -\frac{\hbar}{\pi R^2\lambda^2}.\end{aligned} \tag{A20}$$

Next we compute the change in the expansion induced by the above quantum stress tensor. We consider a black hole at the onset of evaporation, for which there is no Hawking radiation outside the near horizon zone yet. Thus we expect the geometry to revert to the classical vacuum Schwarzschild solution far from the black hole. And so, to find the corrected expansion, we integrate backwards from $\lambda = \infty$ to find the shift:

$$\begin{aligned}\delta\theta \equiv \theta(\lambda) - \theta(\infty) &= -8\pi G\int_{\infty}^{\lambda}\langle T_{\mu\nu}k^\mu k^\nu\rangle d\lambda' \\ &= 8\pi G\int_{\lambda_0}^{\lambda}\frac{\hbar}{\pi R^2\lambda'^2}d\lambda' = -\frac{8G\hbar}{R^2\lambda}.\end{aligned} \tag{A21}$$

To find the (classically) marginally trapped surfaces in the Unruh state, we solve

$$\theta^{(0)} + \delta\theta = 0, \tag{A22}$$

where $\theta^{(0)}$ is the uncorrected classical expansion given in Eq. (A11). Using $c = c_0\lambda$, we find that the classical marginally trapped surfaces are located at

$$c_{\text{MTS}} \sim \frac{G\hbar}{R} \tag{A23}$$

in the quantum-corrected geometry. Very near the horizon, we can treat the radial coordinate to be essentially R to zeroth order.

An alternative useful notion of distance is the proper radial distance from the horizon, ℓ, which satisfies

$$d\ell = \frac{dr}{\sqrt{1 - \frac{R}{r}}} \simeq \sqrt{R}\frac{dr}{\sqrt{r - R}} \to \ell \simeq 2\sqrt{R(r - R)} \sim (Rc)^{1/2}. \tag{A24}$$

Since $G\hbar = l_p^2$, we see that the trapped surfaces are about a Planck length outside the horizon:

$$\ell_{\rm MTS} \sim \mathcal{O}(l_p). \tag{A25}$$

Thus, the area of the classical marginally trapped surface is increased by the quantum correction, by

$$\Delta A_{\rm MTS} \sim G\hbar = l_P^2. \tag{A26}$$

3. Quantum trapped surfaces during evaporation

We still consider the quantum-corrected geometry in the Unruh state, so the classical expansion is given by

$$\theta = \theta^{(0)} + \delta\theta \sim \frac{c_0}{R} - \frac{G\hbar}{R^2\lambda}. \tag{A27}$$

The generalized entropy is

$$S_{\rm gen} = \frac{A}{4G\hbar} + S, \tag{A28}$$

where $S = -{\rm Tr}\rho\log\rho$ and ρ is the quantum state in the region exterior to the Cauchy-splitting sphere. The quantum expansion Θ is ($4G\hbar$ times) the rate of change of the generalized entropy, per unit area, under shape deformations. In the spherically symmetric case,

$$\Theta = \theta + \frac{4G\hbar}{A}\frac{dS}{d\lambda}. \tag{A29}$$

Quantum marginally trapped surfaces are characterized by $\Theta = 0$.

The GSL states that any outgoing radial congruence on or outside the event horizon must satisfy $\Theta \geq 0$, so the quantum marginally trapped surfaces must lie inside the horizon [11]. By Eq. (A27), $\theta < 0$ on and inside the horizon. We see from Eq. (A29) that the GSL requires

$$\frac{4G\hbar}{A}\frac{dS}{d\lambda} = -\alpha\theta|_{\mathcal{H}}, \tag{A30}$$

where $\mathcal{H}$ refers to the horizon. We take $\alpha - 1 \sim O(1)$, in line with Page's explicit calculation for an evaporating black hole in the Unruh state [40].

Combining these results and neglecting factors of order unity where appropriate, we find

$$\Theta = \theta - \alpha\theta|_{\mathcal{H}} = \frac{c}{R\lambda} - \frac{G\hbar}{R^2\lambda} + \alpha\frac{G\hbar}{R^2\lambda}. \tag{A31}$$

Setting $\Theta = 0$ yields

$$\frac{c}{R\lambda} = -(\alpha - 1)\frac{G\hbar}{R^2\lambda} \to c \sim -\frac{G\hbar}{R}. \tag{A32}$$

Using the proper area, we find

$$\Delta A_{\rm QMTS} \sim -l_P^2. \tag{A33}$$

Thus, the quantum marginally trapped surfaces are a proper distance of order the Planck length inside of the horizon.

We will now show that the "duration" of the light sheet L of a quantum marginally trapped surface μ_Q is of order of scrambling time

$$\Delta t_s \sim R\log\frac{R}{l_P}. \tag{A34}$$

This assumes that μ_Q is about one Planck length inside of the event horizon, as would be the case for an isolated, slowly evaporating black hole. Of course, the points on L are null or spacelike separated. What we mean by the "duration" of L is the amount of time, as measured at large radius r, for which it will be the case that matter falling in radially from this radius will cross L (see Fig. 9).

We will approximate the infalling matter as ingoing radial null geodesics; the result would be the same for timelike geodesics starting at rest at a large radius. Let the earliest geodesic crossing L be at $v = v_1$ in the Eddington-Finkelstein coordinates defined in Appendix A 1. It will meet L at μ_Q, whose radius satisfies $R - r_{\mu_Q} \sim l_P^2/R$. The last geodesic that meets L will do so where L hits the singularity, at $r = 0$. The light sheet L is characterized by $u = {\rm const}$, where u is the ingoing Eddington-Finkelstein coordinate, $u \equiv t - r_*$. Here r_* is the tortoise coordinate defined in Eq. (A3). Since r_* depends only on r, we have

$$\Delta t = t_2 - t_1 = r_*(r_{\mu_Q}) - r_*(0) = r_{\mu_Q} + R\log\frac{R}{l_P^2/R} \sim \Delta t_s. \tag{A35}$$

A similar analysis demonstrates that the scrambling time is how long it takes a geodesic to propagate from about a

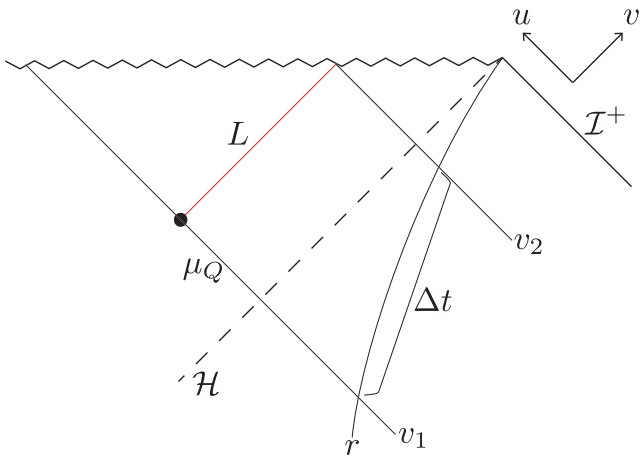

FIG. 9. The future outgoing light sheet of μ_Q (top red line) is crossed by two ingoing radial null geodesics at v_1 (at μ_Q) and v_2 (at the singularity). Their Schwarzschild time difference at fixed r is the scrambling time, Δt_s.

Planck distance outside the horizon to the edge of the near-horizon zone, at $r = 3R/2$.

APPENDIX B: PERTURBATIVE CONSTRUCTION OF Q SCREENS

Let μ_Q be a quantum marginally trapped surface near a perturbed Killing horizon that approaches the Hartle-Hawking state in the future. Then there exists a Q screen that approaches the Killing horizon in the future and contains μ_Q as a leaf.

This fact is useful in sketching a heuristic argument for our conjectured QPI in asymptotically AdS spacetime, following Eq. (7.8). We will now demonstrate this claim by explicit construction.

Consider an event horizon $\mathcal{H}$ which is a perturbation to a Killing horizon caused by matter excitations $T_{\mu\nu} \sim \mathcal{O}(\hbar)$ such that in the far future $\mathcal{H}$ settles down to a Killing horizon in the Hartle-Hawking state. Furthermore, assume that there exists a quantum marginally trapped surface near $\mathcal{H}$. It is known [11] that quantum marginally trapped surfaces are behind event horizons, so μ_Q will be a small distance in the inward direction l from $\mathcal{H}$. Given any codimension two surface in this spacetime, k and l, respectively, represent the outward and inward null vectors perpendicular to the surface. Let y parametrize the transverse position of the surface; see Fig. 10.

For the construction of the Q screen, we start by emanating a past outwards-directed null plane from μ_Q and mark its intersection with the horizon as μ_H. Now, we can pick a foliation of the horizon that starts from μ_H and continues toward the future of $\mathcal{H}$ such that it eventually approaches the preferred foliation of the Killing horizon. Mark the leaves of this foliation by λ such that $\lambda = 0$ is μ_H and λ grows along the future leaves. We construct the Q screen by shooting null future-directed inward null planes from the leaves μ_H and on that null plane look for a quantum marginally trapped surface.

Suppose that a given leaf of our foliation of $\mathcal{H}$ (marked by λ) has a quantum expansion $\Theta_k(\lambda; y)$ at a given transverse position y. By the generalized second law, $\Theta_k \geq 0$. Then, perturbatively we can find the location of a quantum marginally trapped surface as

$$\Theta_k(\lambda; y) + \delta U(\lambda; y)(\partial_l \Theta_k(\lambda; y)) = 0, \tag{B1}$$

where δU is the amount of affine parameter in the l direction we need to venture to find a quantum marginally trapped surface and $\Theta_k = \mathcal{O}(G\hbar)$.

We need to solve for a function $\delta U(y)$ and show that it approaches zero as we go toward higher values of λ. From the definition of quantum expansion it follows that

$$\partial_l \Theta_k = \partial_l \theta_k + 4G\hbar \partial_l \partial_k S_{\text{out}}. \tag{B2}$$

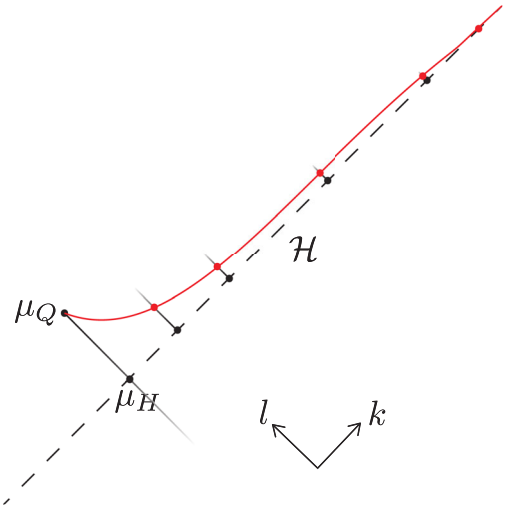

FIG. 10. A quantum marginally trapped surface μ_Q in the vicinity of a perturbed Killing horizon $\mathcal{H}$. We construct a Q screen containing μ_Q that asymptotes to the Killing horizon at late times. We first fire a null plane toward $\mathcal{H}$ that intersects it on μ_H. We then foliate $\mathcal{H}$ starting from μ_H. At every leaf of this foliation, we fire null planes inwards and to the future. On each null plane, we find a quantum marginally trapped surface at an affine distance δU from $\mathcal{H}$. The Q screen is the union of these quantum marginally trapped surfaces.

The cross-focusing equation is

$$\partial_l \theta_k = -\frac{1}{2}\mathcal{R} - \theta_l \theta_k + \nabla \cdot \chi + \chi^2 + 8\pi G T_{kl}, \tag{B3}$$

where $\mathcal{R}$ is the intrinsic Ricci scalar of the leaf and χ is its twist [41]. From Eq. (B1), we see that in order to solve for δU to first nontrivial order in $G\hbar$, we only need the leading order expression for $\partial_l \Theta_k$. The leading order term is

$$\partial_l \Theta_k = -\frac{1}{2}\mathcal{R}^{(0)} + \mathcal{O}(G\hbar), \tag{B4}$$

where $\mathcal{R}^{(0)}$ is the (y-independent) intrinsic Ricci scalar of the leaf on the unperturbed Killing horizon. For a 2-sphere $\mathcal{R}^{(0)} = 2$. Combining the above equations with (B1), we can solve for δU to the first nontrivial order in $\mathcal{O}(G\hbar)$:

$$\delta U(y; \lambda) = \Theta_k(\lambda; y). \tag{B5}$$

Since by assumption $\mathcal{H}$ approaches a Killing horizon in the Hartle-Hawking state in the future, we have

$$\lim_{\lambda \to \infty} \Theta_k(\lambda; y) = 0 \Rightarrow \lim_{\lambda \to \infty} \delta U(\lambda; y) = 0, \tag{B6}$$

where the implication follows from Eq. (B5). This means that the leaves of the Q screen start at μ_Q and approach the late times of the event horizon, which is what we set out to show.

[1] S. W. Hawking, Gravitational Radiation from Colliding Black Holes, Phys. Rev. Lett. **26**, 1344 (1971).

[2] R. Bousso and N. Engelhardt, New Area Law in General Relativity, Phys. Rev. Lett. **115**, 081301 (2015).

[3] R. Bousso and N. Engelhardt, Proof of a new area law in general relativity, Phys. Rev. D **92**, 044031 (2015).

[4] R. M. Wald, *General Relativity* (University of Chicago Press, Chicago, 1984).

[5] R. Penrose, Gravitational Collapse and Space-Time Singularities, Phys. Rev. Lett. **14**, 57 (1965).

[6] J. D. Bekenstein, Black holes and the second law, Nuovo Cimento Lett. **4**, 737 (1972).

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

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

[9] R. Bousso and N. Engelhardt, Generalized second law for cosmology, Phys. Rev. D **93**, 024025 (2016).

[10] R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall, Quantum focusing conjecture, Phys. Rev. D **93**, 064044 (2016).

[11] A. C. Wall, The generalized second law implies a quantum singularity theorem, Classical Quantum Gravity **30**, 165003 (2013).

[12] R. Bousso, Z. Fisher, J. Koeller, S. Leichenauer, and A. C. Wall, Proof of the quantum null energy condition, Phys. Rev. D **93**, 024017 (2016).

[13] J. Koeller and S. Leichenauer, Holographic proof of the quantum null energy condition, Phys. Rev. D **94**, 024026 (2016).

[14] S. Balakrishnan, T. Faulkner, Z. U. Khandker, and H. Wang, A general proof of the quantum null energy condition, J. High Energy Phys. 09 (2019) 020.

[15] R. Penrose, Naked singularities, Ann. N.Y. Acad. Sci. **224**, 125 (1973).

[16] R. Arnowitt, S. Deser, and C. W. Misner, The dynamics of general relativity, in *Gravitation: An Introduction to Current Research*, edited by L. Witten (Wiley, New York, 1962), pp. 227–265.

[17] R. Schoen and S. T. Yau, Proof of the positive mass theorem. II, Commun. Math. Phys. **79**, 231 (1981).

[18] R. Bousso, A. Shahbazi-Moghaddam, and M. Tomasevic, Quantum Penrose Inequality, companion Letter, Phys. Rev. Lett. **123**, 241301 (2019).

[19] M. Mars, Present status of the Penrose inequality, Classical Quantum Gravity **26**, 193001 (2009).

[20] N. Engelhardt and G. T. Horowitz, A holographic argument for the penrose inequality in AdS, Phys. Rev. D **99**, 126009 (2019).

[21] S. W. Hawking and G. F. R. Ellis, *The Large Scale Stucture of Space-Time* (Cambridge University Press, Cambridge, England, 1973).

[22] R. Bousso, A covariant entropy conjecture, J. High Energy Phys. 07 (1999) 004.

[23] R. Bousso, Holography in general space-times, J. High Energy Phys. 06 (1999) 028.

[24] D. G. Boulware, Quantum field theory in Schwarzschild and Rindler spaces, Phys. Rev. D **11**, 1404 (1975).

[25] P. Candelas, Vacuum polarization in Schwarzschild spacetime, Phys. Rev. D **21**, 2185 (1980).

[26] C. Akers, R. Bousso, I. F. Halpern, and G. N. Remmen, Boundary of the future of a surface, Phys. Rev. D **97**, 024018 (2018).

[27] R. Bousso, H. Casini, Z. Fisher, and J. Maldacena, Entropy on a null surface for interacting quantum field theories and the Bousso bound, Phys. Rev. D **91**, 084030 (2015).

[28] R. Bousso, B. Freivogel, and S. Leichenauer, Saturating the holographic entropy bound, Phys. Rev. D **82**, 084024 (2010).

[29] A. R. Steif, The Quantum stress tensor in the three-dimensional black hole, Phys. Rev. D **49**, R585 (1994).

[30] N. Engelhardt and A. C. Wall, Quantum extremal surfaces: Holographic entanglement entropy beyond the classical regime, J. High Energy Phys. 01 (2015) 073.

[31] R. Gregory and R. Laflamme, Black Strings and p-Branes Are Unstable, Phys. Rev. Lett. **70**, 2837 (1993).

[32] R. Gregory, The gregory-laflamme instability, arXiv:1107.5821.

[33] L. Lehner and F. Pretorius, Final state of gregory-laflamme instability, arXiv:1106.5184.

[34] R. M. Wald, Gravitational collapse and cosmic censorship, arXiv:gr-qc/9710068.

[35] C. Gundlach and J. M. Martín-García, Critical phenomena in gravitational collapse, Living Rev. Relativity **10**, 5 (2007).

[36] M. W. Choptuik, Universality and Scaling in Gravitational Collapse of a Massless Scalar Field, Phys. Rev. Lett. **70**, 9 (1993).

[37] D. Christodoulou, Violation of cosmic censorship in the gravitational collapse of a dust cloud, Commun. Math. Phys. **93**, 171 (1984).

[38] H. Kodama, Inevitability of a naked singularity associated with the black hole evaporation, Prog. Theor. Phys. **62**, 1434 (1979).

[39] R. M. Wald and S. Christensen, *Black Holes, Singularities and Predictability* (Adam Hilger Limited, Bristol, UK, 1984).

[40] D. N. Page, Particle emission rates from a black hole. II. Massless particles from a rotating hole, Phys. Rev. D **14**, 3260 (1976).

[41] R. Bousso and M. Moosa, Dynamics and observer-dependence of holographic screens, Phys. Rev. D **95**, 046005 (2017).