Complementarity plus backreaction is enough

Hui, L.; Yang, I.S.

Link to publication

Complementarity plus backreaction is enough

Lam Hui[*]

ISCAP and Physics Department Columbia University, New York, New York 10027, USA

I-Sheng Yang[†]

IOP and GRAPPA, Universiteit van Amsterdam, Science Park 904, 1090 GL Amsterdam, Netherlands

We investigate a recent development of the black hole information problem, in which a practical paradox has been formulated to show that complementarity is insufficient. A crucial ingredient in this practical paradox is to distill information from the early Hawking radiation within the past light cone of the black hole. By causality this action can backreact on the black hole. Taking this backreaction into account, the paradox could be resolved without invoking any new physics beyond complementarity. This resolution requires a certain constraint on the S matrix to be satisfied. Further insights into the S matrix could potentially be obtained by effective-field-theory computations of the backreaction on the nice slice.

I. INTRODUCTION

The black hole information paradox [1] has always been an inspiring topic. Recent arguments made by Almheiri, Marolf, Polchinski, and Sully (AMPS) [2] (see also [3]) led to a new surge of discussions.[1] They argued that complementarity [5–10], a conjecture previously accepted by most, was not enough to resolve the information paradox. In this paper, we wish to explore the possibility that complementarity is actually sufficient to resolve the problem, once backreaction (from the manipulation of Hawking quanta necessary in practical versions of the paradox) is taken into account.

Let us first review the information problem. Throughout this paper, we adopt the notation where $M_{\rm Planck} = 1$, so M stands for both the mass and also the horizon size of the black hole (or more precisely, half the Schwarzschild radius). Let B be a particular near-horizon mode after Page time [11], A be its interior partner, and R be the early Hawking radiation. Knowing the initial state, the unitary evaporation allows us to distill a minimal subsystem R_B from R such that it is maximally entangled with B. On the other hand, the equivalence principle demands that A and B are maximally entangled as the in-falling vacuum state. These two facts put together violate the monogamy of entanglement. A possible solution is $A = R_B$, or colloquially "in = out," which means that the interior is identified with the early Hawking radiation.

This follows the spirit of complementarity in that the interior and exterior cannot be treated as two independent sets of degrees of freedom. Note that it should only work if A and R_B can never be brought together and compared even

in principle, for example if Bob distills R_B from R at distance $\sim M^3$ away as shown in Fig. 1(left). Here, in = out can be thought of as a consolation to those who insist on thinking globally. As a practical matter, since "in" and "out" can never be brought together and compared, by definition it does not lead to any paradox.

However the key argument in [2] is to point out that there is a practical paradox as shown in Fig. 1(right). After leaving the black hole, Hawking radiation is basically free streaming. It should not make any difference if someone intercepts them earlier. Alice intercepts the radiation earlier, and closer to the black hole. She thus has enough time to distill R_B from R, and then carry it into the black hole to compare with A. $A = R_B$ now becomes a blatant quantum cloning. From our point of view, this practical paradox is the strongest version of the AMPS argument. One single observer, Alice, conducts two experiments and sees conflicting results. Any proposed resolution must directly confront this practical paradox.[2]

First let us revisit Fig. 1. Comparing Bob and Alice, Bob performs distillation further away and later, while Alice does the same closer in and earlier. The fact that the distillation is done at different times does not greatly affect the state of R received by them—R earlier and later are simply related by the unitary transformation associated with free streaming. There is, however, one crucial difference in the distillation processes performed by Bob and Alice, due to their space-time locations in relation to the

---

[*]lhui@astro.columbia.edu
[†]isheng.yang@gmail.com
[1]See [4] for an extensive list of references.

---

[2]Note that for a practical paradox, it is essential to distill R_B, or at least to have a system smaller than R that entangles with R_B (such as performing a classical measurement). That is because Alice cannot bring all $\sim M^2$ qubits in R back into the black hole without dramatically altering the geometry. She must process the information in R one way or another to reduce the amount of information she needs to carry.
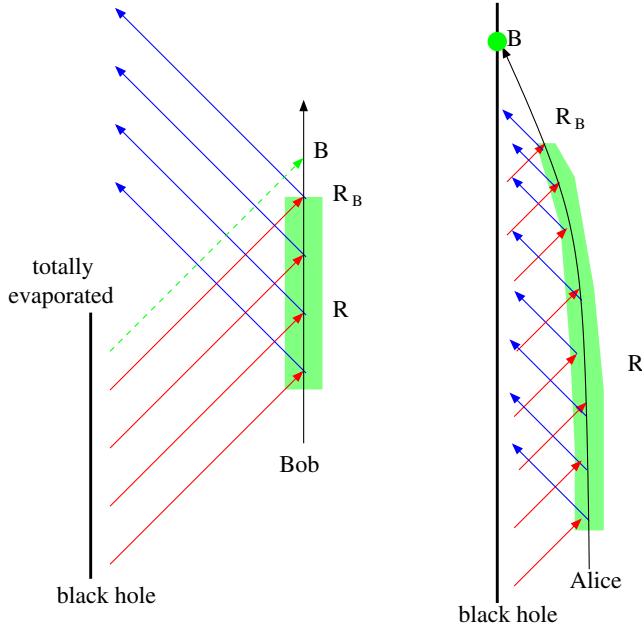
FIG. 1 (color online). Black holes are drawn as thick black lines, with their near-horizon and interior details suppressed (which includes the interior mode A). The left figure shows Bob, who stays very far away. His distiller (operating in the green box) is spacelike separated from the near-horizon origin of the late quantum B, which then propagates to Bob along the green dotted line. Backreaction from his distiller is irrelevant but he also cannot jump into the black hole. The right figure shows Alice, who stays near the black hole since she wants to jump in later. The backreaction from her distiller will affect the late quantum B (the green dot) she encounters later.

black hole. By causality, Bob's action cannot affect the black hole in any way, and there is no paradox for him. On the other hand, Alice can jump into the black hole to witness a potential paradox, but *the same fact ensures that her actions can causally affect the black hole.* If complementarity can survive without additional new physics, the key must be in how the distillation process done by Alice backreacts on the black hole.

The resolution we wish to explore can thus be described as "complementarity + backreaction": (i) complementarity, in the sense of in = out, addresses situations in which the distillation process is spacelike separated from the near-horizon origin of quantum B; (ii) backreaction on the black hole addresses situations where the near-horizon origin of quantum B is within the forward light cone of the distillation process. Our main goal in this paper is to demonstrate why such a backreaction is plausible, and how it could resolve the paradox.

The idea of backreaction is in a sense a natural one, but there has not been much discussion in the literature on how this could address the AMPS paradox. There are probably several reasons for this.

First, a perhaps common response to the idea of back-reaction is to say, yes, in principle this could happen, but

presumably a clever experimentalist can make the back-reaction sufficiently weak to be negligible. We will argue that the fairly nontrivial distillation process, whereby Alice obtains the complement of B from R, necessarily gives rise to a certain level of backreaction. Tuning the backreaction to be acceptably small could sacrifice one's ability to distill R_B from R.

Second, one might think that since backreaction involves sending signals from outside into the black hole, it seems to go in the wrong direction, i.e. opposite direction from the late outgoing Hawking quantum B which we wish to affect. This is not an issue, as long as one keeps in mind that the backreaction could affect the state of the black hole and therefore its emission as well.

Third, one might think some kind of shield can be set up to prevent backreaction signals from reaching the black hole. From Fig. 1(right), it is apparent that setting up a shield to block off a significant fraction of the backreaction signals would also prevent a fraction of the early Hawking radiation from reaching the distillation apparatus.

The nature of the distillation process of course depends on how the information is encoded in the Hawking radiation, which is determined by both the S matrix and the initial state. For example if only the very first early quantum is entangled with the late quantum B, Alice can simply ignore all other early quanta. Such a trivial distillation needs not induce a meaningful backreaction, but there is no reason to expect such a trivial entanglement structure would emerge out of the S matrix.[3] In other words, our resolution does impose certain constraints on the black hole S matrix. Spelling them out in more detail might lead to a better understanding of the S matrix.

Let us briefly comment on the relation between our resolution to some existing proposals. The ER = EPR proposal of [16] can be viewed as a physical way to enforce in = out by using the wormhole geometry. This is compatible with part of our viewpoint: in = out, i.e. complementarity, when there is no backreaction. On the other hand, when backreaction does occur, our view is that it can occur by propagating signals through the normal part of the space-time, as opposed to through the Einstein-Rosen bridge. The practical paradox may also be resolved by a limitation on the computation time of the distillation process [17]. Here we consider the possibility that this limitation can be somehow circumvented, in which case an alternative resolution of the paradox is required. Let us close this introduction by pointing out that our resolution is

---

[3]States with such a trivial entanglement structure might form a complete basis in the Hilbert space of the Hawking radiation [12–14], but there is no reason to expect the property of basis states to be carried over to a general superposition. The fact that the Hilbert space for Hawking radiation is larger than the Hilbert space of the black hole [15] suggests one should expect a general superposition rather than a very special trivial state for the Hawking radiation.

consistent with the possibility that Alice, with the powerful knowledge of the S matrix and initial state, might be able manipulate the black hole to make some energetic late quantum, i.e. what is commonly referred to as the firewall. However a firewall has no reason to spontaneously develop during an unaltered evaporation process (one that suffers no backreaction), so there is no violation of the equivalence principle.

A. Outline

In Sec. II we construct a model for the distillation process and the backreaction. In Sec. III we show how complementarity + backreaction should be enough to avoid the practical paradox. In addition, the firewall paradox has been formulated in various ways, and some of them do not involve an explicit distillation or measurement process. In Sec. III E we explicitly show that the resolution in one such formulation directly follows from our resolution to the practical paradox. This is not surprising since they are ultimately the same paradox. In Sec. IV we summarize our result and point out the possibility to decode the black hole S matrix by performing computations on the nice slice.

II. BACKREACTION

A. The distillation process

In order to address backreaction, we need a more concrete description of how to extract information from the early Hawking radiation. Our argument focuses on a unitary distillation process, during which R_B is separated from the rest of R. However it should be obvious that it applies to classical measurements, too.[4] We will first argue for the following universal requirement for any distiller.

A distiller that can extract R_B from R must carry some current that interacts with the Hawking radiation. Such a current is designed to match the expected pattern that encodes R_B within R, which is determined by the black hole initial state, the S matrix, and the specific late quantum B.

In order to see this, we model the distillation process as the following unitary evolution:

$$e^{-i \int H dt d^3 x}|R\rangle|\text{distiller}\rangle|0\rangle = |R', \text{distiller}'\rangle|R_B\rangle, \quad (2.1)$$

where a time ordering is implied in the evolution operator. Initially, Alice has a memory stick of one empty qubit, where empty just means it has one unit of fixed and irrelevant information content. She plugs it into the distiller and together form the "distiller system," which interacts with the incoming Hawking radiation R. In the end this memory stick should be loaded with R_B such that Alice can unplug it and bring it into the black hole. After this process

---

[4]By a classical measurement, we mean a nonunitary projection.

R loses the information in R_B and becomes R', which generically will be entangled with the distiller state that is also altered.[5]

Equation (2.1) is in the Schrödinger picture where the states evolve with time according to the full Hamiltonian density H. However only the coupling term H_{int} can be responsible for the transferring of information regarding R_B[6]:

$$H_{\text{int}} = A_\mu J^\mu + \sum (\text{radiation})(\text{current}). \quad (2.2)$$

The "radiation" operator acts on the state of Hawking radiation R while the "current" operator acts on the state of the distiller system. Therefore H_{int} entangles them and transfers information. For simplicity let us focus on the standard interaction term in electromagnetism ($A_\mu J^\mu$), while in general we expect radiation of all types (e.g. gravitons, neutrinos, scalars) each of which is coupled to its corresponding current. Our argument works in the same way, regardless of the spin of the radiation particle.

Note that given the same initial pure state of the black hole, if we are interested in a different late quantum B', we need to distill a different $R_{B'}$ correspondingly. Since the incoming radiation R is still the same state, of course we need a different (initial) distiller state such that J_μ acts on it differently. Similarly, the requisite distiller state also depends on the initial state of the black hole. If we have a different initial black hole pure state with the same macroscopic parameters, R will be in a different state which encodes the information of R_B differently. Aiming for the same B according to a roughly identical evaporation process and classical geometry (for example the 24601st quantum), we will also need a different distiller state. In other words, the required initial distiller state depends on a number of things: it depends on the classical label B and on the state of the expected Hawking quanta, which in turn depends on the initial state of the black hole and the S matrix.

We find it convenient to express the dependence of both the distiller state and the state of the incoming Hawking quanta on these various features of the problem (the initial state of the black hole, S matrix, etc.) using a somewhat unusual form of the Heisenberg picture. In the usual Heisenberg picture, time evolution is transferred to the operators; states do not evolve. Here, we wish to go one step further: we encode features of the initial states in the

---

[5]For classical (projection) measurements, we can use exactly the same system but instead of the full quantum information of R_B, Alice is only allowed to carry away some classical information related to it.

[6]We focus on the physical process represented by this Hamiltonian density, and assume that the distillation in Eq. (2.1) is in principle possible within a reasonable amount of time. As argued in [17], the paradox can be resolved if the required time is too long, but we are looking for another resolution independent from that possibility.

operators as well. It works as follows. Let $t = 0$ be the time where the Schrödinger picture and the (usual) Heisenberg picture agrees. In other words, we say

$$|\Psi_1(t=0)\rangle^S = |\Psi_1\rangle^{H'}; \qquad |\Psi_2(t=0)\rangle^S = |\Psi_2\rangle^{H'}, \quad (2.3)$$

where the superscripts S and H' denote the Schrodinger picture and the (usual) Heisenberg picture, respectively. Here $|\Psi_1\rangle$ and $|\Psi_2\rangle$ denote two different initial states. Let us further define a *modified* Heisenberg picture, denoted by superscript H, as follows:

$$|\Psi_1(t=0)\rangle^S = |\Psi_1\rangle^{H'} = U_1|\Psi_0\rangle^H;$$
$$|\Psi_2(t=0)\rangle^S = |\Psi_2\rangle^{H'} = U_2|\Psi_0\rangle^H, \quad (2.4)$$

where $|\Psi_0\rangle^H$ is some common "ancestor" state (whose particular choice is not important) which is related to $|\Psi_1\rangle$ and $|\Psi_2\rangle$ by the unitary transformations U_1 and U_2, respectively. In our adaptation of the Heisenberg picture, the state is thus always $|\psi_0\rangle^H$, independent of time and independent of initial conditions. All the interesting information about the dynamics and initial conditions are encoded in the operators:

$$(A_\mu)_1^H = U_1^\dagger U^\dagger(t,0)(A_\mu)^S U(t,0)U_1,$$
$$(A_\mu)_2^H = U_2^\dagger U^\dagger(t,0)(A_\mu)^S U(t,0)U_2, \quad (2.5)$$

$$(J_\mu)_1^H = U_1^\dagger U^\dagger(t,0)(J_\mu)^S U(t,0)U_1,$$
$$(J_\mu)_2^H = U_2^\dagger U^\dagger(t,0)(J_\mu)^S U(t,0)U_2. \quad (2.6)$$

Henceforth, we will drop the subscript 1 or 2 which only serves to remind us the operator in question, A_μ^H or J_μ^H, cares about the initial state. We will even drop the superscript H—hereafter when we refer to operators, we mean Heisenberg operators defined in this way. Therefore, we write

$$H_{\text{int}} = A_\mu(\text{black hole})J^\mu(\text{distiller}), \quad (2.7)$$

$$A_\mu(\text{black hole}) = \{S\text{matrix, initial state}\}, \quad (2.8)$$

$$J_\mu(\text{distiller}) = \{A_\mu(\text{black hole}, r_{\text{distiller}}), S\text{matrix}, B\}. \quad (2.9)$$

The content of the Hawking radiation $A_\mu(\text{black hole})$ depends on the black hole S matrix and initial state. The current $J_\mu(\text{distiller})$ required to distill R_B depends on B and also on the S matrix which determines how R_B is encoded in R. There is a trivial dependence on $r_{\text{distiller}}$ (location of the distiller) which determines at what time the radiation arrives at the distiller.

The advantage of our generalized Heisenberg picture is that it allows us to use a language that is almost classical. The operator $A_\mu(\text{black hole})$ can be thought of as the (state of) radiation from the black hole. The operator $J_\mu(\text{distiller})$ can be thought of as the (state of) current of the distiller. Occasionally, we would freely switch picture in our descriptions. For instance, it should be obvious that "changing the current $J_\mu(\text{distiller})$" is in the generalized Heisenberg picture and "changing the state $|\text{distiller}\rangle$" is in the Schrödinger picture—they could even be the same change, expressed in different ways.

Given the current $J_\mu(\text{distiller})$, the discussion of back-reaction is straightforward. A current that can respond to incoming radiation is also a source itself. As we have focused on the EM part of the Hawking radiation, we can describe that (in Lorenz gauge) simply by

$$\partial_\mu\partial^\mu A_\nu(\text{distiller}) = J_\nu(\text{distiller}). \quad (2.10)$$

Similar equations hold for other types of radiation, such as a massless scalar.[7] The backreaction we are interested in refers to how the black hole is affected by this $A_\mu(\text{distiller})$, or in general the field sourced by the distiller. We shall analyze it from two different perspectives. In Sec. II B we follow the membrane paradigm [18] and treat the black hole as some object of size $\sim M$. In Sec. II C we analyze the Hawking process on the nice slice [19]. Both pictures show that the backreaction can modify the late Hawking quantum B.

B. The membrane paradigm

From the outside point of view, the black hole can be thought of as a membrane that sources Hawking radiation. Its effective current $J_\nu(\text{black hole})$ can be worked out from

$$\partial_\mu\partial^\mu A_\nu(\text{black hole}) = J_\nu(\text{black hole}). \quad (2.11)$$

This is a rather indirect way to describe the evaporation process. The analysis in the next section II C will give a more concrete picture. Here the goal is to put the distiller and the black hole on the same footing, and make plausible the notion that a nontrivial transformation of the distiller (i.e. the distillation process) entails a nontrivial transformation of the black hole (i.e. backreaction).

For the same reason that the distiller is influenced by the interaction between its own current and the radiation from the black hole (previously H_{int}),

$$H_{\text{distiller}} = A_\mu(\text{black hole})J^\mu(\text{distiller}), \quad (2.12)$$

the black hole is also influenced by the radiation from the distiller,

---

[7]This type of equation assumes the particle corresponding to the radiation, absent interaction with a current, is free. This does not hold, for instance, for gluons. The story for such particles would be more involved, but should be similar in spirit to the one we are telling.

$$H_{\text{black hole}} = A_\mu(\text{distiller})J^\mu(\text{black hole}). \qquad (2.13)$$

This is the natural backreaction of any distiller Alice wishes to employ.

Our notation requires some explanation. From the quantum field theory point of view, there is only one current operator J_μ and one photon operator A_μ. We can split J_μ into two halves, one nonzero only in the space-time region containing the distiller, i.e. $J_\mu(\text{distiller})$, and the other nonzero only in the region containing the black hole, i.e. $J_\mu(\text{black hole})$. The full A_μ sourced by the full J_μ can thus also be split into two halves: $A_\mu(\text{distiller})$ sourced by $J_\mu(\text{distiller})$ and $A_\mu(\text{black holes})$ sourced by $J_\mu(\text{black hole})$. Out of the full product $A_\mu J^\mu$, Eqs. (2.12) and (2.13) are the cross terms which describe the influence of one system on the other. The diagonal terms describe how each system evolves on its own, which is not what we are interested in.

It should already be obvious that Eqs. (2.12) and (2.13) are closely related. We can see that relation more clearly by integrating out A_μ, utilizing the Green's function $\mathcal{G}^{\mu\nu}$:

$$\int dt dx^3 H_{\text{distiller}} = \int dt dx^3 \int dt' dx'^3$$
$$\mathcal{G}^{\mu\nu}(\text{from black hole to distiller})$$
$$J_\mu(\text{distiller})J_\nu(\text{black hole})', \qquad (2.14)$$

$$\int dt dx^3 H_{\text{black hole}} = \int dt dx^3 \int dt' dx'^3$$
$$\mathcal{G}^{\mu\nu}(\text{from distiller to black hole})$$
$$J_\nu(\text{distiller})'J_\mu(\text{black hole}). \qquad (2.15)$$

In the first equation the (x, t) integral goes over the distiller while the (x', t') integral goes over the black hole. In the second equation it is the other way round. Presenting them side by side makes it clear that just like the black hole influences the distiller (first equation), the distiller can influence the black hole (second equation). The two expressions are not identical, however, due to the fact that the retarded Green's function is not symmetric, because of causality.

The entire early Hawking radiation lasts a long time $\sim M^3$, which is usually much larger than the distance between the distiller and the black hole if Alice wishes to jump in later. Therefore the difference caused by the Green's function is small, and Eqs. (2.14) and (2.15) are expected to give fairly similar results. We put forward the following conjecture.

Given a typical black hole initial state and a typical late quantum B, the effective current $J_\mu(\text{black hole})$ that sources the Hawking radiation and the required current $J_\mu(\text{distiller})$ to distill R_B—both depending on the S matrix—take forms which make Eqs. (2.14) and (2.15) comparable.

The main substance of this conjecture is that nearly *all* the early radiation has to be processed by the distiller

(in order to obtain R_B), so the black hole is backreacted for roughly the same duration as the distiller functions. This places some nontrivial constraint on the black hole S matrix. This conjecture implies that backreaction cannot be tuned arbitrarily small. Recall that the state of Alice's memory stick has to be changed from the empty initial state $|0\rangle$ to the intended information content R_B. This change requires a transfer of information effected by $H_{\text{distiller}}$, and a nontrivial change requires

$$\left\langle \int dt dx^3 H_{\text{distiller}} \right\rangle \gtrsim 1. \qquad (2.16)$$

The conjecture then implies

$$\left\langle \int dt dx^3 H_{\text{black hole}} \right\rangle \gtrsim 1, \qquad (2.17)$$

meaning that informationwise, the change of the black hole state is similarly bounded from below.

It should be stressed that we do not have a definitive proof that backreaction is non-negligible. Indeed, we can be accused of assuming the answer we want by positing the conjecture. Nonetheless, the form of the black-hole-distiller interaction expressed in Eqs. (2.14) and (2.15) is rather suggestive. It suggests that the backreaction on the black hole cannot be avoided however clever the design of the distiller. The source of backreaction is the distiller current given by Eq. (2.9). It is dictated by our desire to distill R_B, given the S matrix and the initial state. Trying to reduce its backreaction also diminishes the ability to distill R_B, and thus there is a lower bound.

It is worth emphasizing that the backreaction is a backreaction on the state of the black hole. In other words, the radiation sourced by the distiller is not by itself interesting—indeed, that radiation enters, as opposed to emerges out of, the black hole horizon, and so appears to go in the wrong direction compared to the direction of the late outgoing Hawking quantum B—it is interesting only because the black hole is radiating to begin with; i.e. the backreaction is effected by the *product* of the distiller and black hole currents. The physical effect of the backreaction is an alteration of the black hole state, or can be thought of as a modulation of the evaporating process.[8] *If Alice performs distillation for a long duration* $\sim M^3$, *the black hole's emission is also modulated by a comparable duration, so the state of a particular late quantum is*

---

[8]This is the general interaction between two sources emitting the same type of radiation [20]. The leading order effect is not that of one sending something to another, but a mutual modulation: one source appears to be emitting more or less under the influence of the other. This is equivalent to an interference effect between their emissions. In certain cases the original emission greatly amplifies the influence of the incoming radiation and serves as a good detector [21].

accordingly modified. The same physical interpretation will emerge from our analysis in the next section.

C. The nice slice

There is another way to understand how A_μ(distiller) affects the evaporation process. Instead of treating the black hole as a source with an effective current, let us follow another description of the Hawking radiation. We construct the "nice slice" which are spacelike surfaces going through the horizon [19]. Curvature remains small everywhere, while the near-horizon region has an expanding geometry. We can use quantum field theory to describe the comoving modes whose wavelengths are being stretched by the expansion. When the physical wavelength of a comoving mode is short, it should be in the vacuum state of a locally flat region. Without outside influence, it stays in that comoving vacuum state while being stretched. As the wavelength becomes longer and the modes migrate into the asymptotic region of the nice slice, this comoving vacuum disagrees with the asymptotic vacuum. This pair production in the expanding background is the source of Hawking radiation in the outgoing modes.

When Alice operates her distiller, the above stretching process proceeds with the near-horizon region permeated by A_μ(distiller). As depicted in Fig. 2, this modifies the dynamics. The comoving modes are affected by A_μ(distiller) during the stretching process and do not have to stay in the comoving vacuum.

Note that on the nice slice, the objection of a "wrong direction" (see discussion in Secs. I and II B) does not hold. The proper time progresses very slowly in the interior portion of the nice slices. So A_μ(distiller) actually cannot cross the expanding near-horizon region and indeed lingers around to modify the later Hawking radiation. This leads to a spectrum different from the original Hawking radiation. *The current J_μ(distiller) to distill R_B leads to an external influence A_μ(distiller) that constantly modulates the evaporation process by acting on the comoving modes near the horizon in the nice slice.* The process of how A_μ(distiller) effectively alters the state of the late radiation should be treatable with standard field theory techniques. The problem is fairly analogous to computing inflationary perturbations under the influence of disturbances that cause excitations above the Bunch-Davies state.

III. COMPLEMENTARITY PREVAILS

In this section we go over a few different scenarios and demonstrate how the effect of backreaction can help resolve the practical paradox. They suggest that complementarity is self-consistent, if backreaction is correctly taken into account.

A. Bob's picture

First let us consider Bob, who has no intention, nor possibility, of jumping into the black hole. He can check
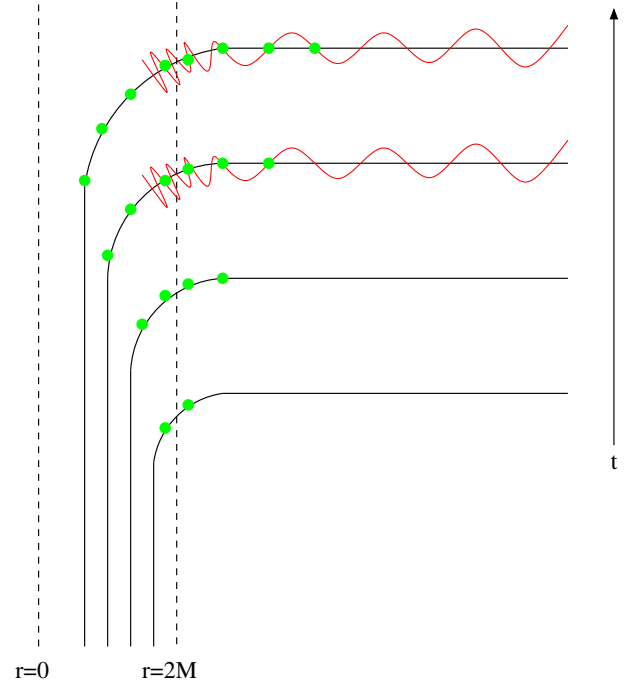

FIG. 2 (color online). The green dots represent pair-produced particles in the near-horizon region which are being pulled apart by the expanding geometry. The horizontal and vertical directions represent schematically the Schwarzschild r and t. The first two (lower) slices show the pair production from vacuum, which leads to the Hawking radiation from an unaltered evaporation process. The second two (higher) slices are the same process under the effect of backreaction from a distiller, which fills the near-horizon region with A_μ(distiller) drawn as the red waves. They are blueshifted and piled up near the horizon. The pair production is no longer from vacuum, so the evaporation process is altered.

whether the evaporation process is unitary while staying very far away, $r_{\text{Bob}} \gtrsim M^3$. He will operate his distiller with the current given by Eq. (2.9). The black hole will not be there when the backreaction from Bob arrives, so he has no reason to worry about anything we said in the previous section. When the radiation and backreactions cross path in empty regions away from both the black hole and the distiller, there is basically no interaction. After distilling R_B from R, Bob will receive the late quantum B and confirms a unitary evaporation process.

B. A careless Alice

Alice would like to replicate Bob's process at a closer distance $r_{\text{Alice}} \ll M^3$, such that she can carry the result R_B into the black hole. She might believe that a simple modification of Bob's distillation is enough. Before the early quanta R reach Bob, she can intercept them with her own distiller with the current also given by Eq. (2.9). The difference in the variable r takes care of the unitary transformation between the state of R at her position and Bob's, so she expect to get the same result R_B.

Following this simple logic, Alice's plan is doomed to fail. Since $r_{\text{Alice}} \ll M^3$, the backreaction from her distiller starts to affect the black hole soon after distillation begins, and continues through almost the entire process. Since the state of the black hole is altered, it will not emit the same Hawking radiation. The difference might be small, but the distillation is a very delicate process where $J_\mu(\text{distiller})$ is designed to match a particular content of $A_\mu(\text{black hole})$. Using the same current on a different radiation content, both B and R_B will be modified to something totally unexpected, and her check of unitarity will simply fail.

C. A careful Alice

Alice can try to be more careful. Now that she knows about the backreaction, she realizes that distilling R_B given by an unaltered evaporation process is a fool's errand at her location. Her action changes the black hole and she needs to take that into account. In principle, she can do a much more involved calculation to keep track of how both the black hole and her memory stick are simultaneously affected. She can then construct a device that manipulates both systems for a long time $\sim M^3$. The information in a late quantum B will not be the same as the naturally evaporating black hole, and it can be entangled with her memory stick.

Assuming that Alice performs the calculation correctly, she can indeed confirm the desired entanglement. However, this entanglement has little to do with the S matrix of the unaltered evaporation of this black hole. Alice is simply checking the result of her manipulation. It may be surprising, but not against any physical law that the outcome is dynamically enforced. Since she deliberately manipulated the state of B, there is no reason that it is still purified by A.[9] So when she jumps in, she will see those high-energy quanta she created. This comes from her impressive ability to guide the black hole through a very special evaporation process during a long time $\sim M^3$, and does not violate the equivalence principle or any other low energy physical laws.

D. Combining Alice and Bob

Now let us have both Alice and Bob. Alice still follows her fixed trajectory to cross the horizon at some time, say t_{cross}. However she will not attempt to do anything to the Hawking radiation. If she did, the black hole would be altered by backreaction, and neither she nor Bob would have access to an unaltered evaporation process. Alice will just fall through the horizon and check the equivalence principle, namely the smoothness of horizon as the entanglement between the interior mode A and the exterior mode B.

Bob is in charge of checking the unitary evaporation. He knows about Alice's schedule, and agrees to check the state of a near-horizon mode B at t_{cross}. He operates his detector to distill R_B, and then later observes B directly as a quantum in the late Hawking radiation. Now let us examine if the experiences of Alice and Bob can contradict each other.

Let Bob start at $r_{\text{Bob}}(t = t_{\text{cross}}/2) = t_{\text{cross}}/2$, such that the backreaction cannot reach the horizon before Alice falls through. This way Bob's distiller cannot causally change Alice's experience. If Bob just stays there, he can eventually see B as a late quantum and confirm that an unaltered evaporation is unitary. Although he is slightly closer to the black hole than the ideal distance described in Sec. III A, most of the backreaction signals still cross path with the late quantum mostly in empty space. We should allow Bob to use the trivial J_μ in Eq. (2.9) and treat B as unaltered. So, staying at this large distance, Bob should confirm the entanglement between B and R_B given by the unaltered black hole S matrix.[10] Now Bob's observation is in conflict with the normal horizon experienced by Alice. Fortunately they cannot compare their results (if Bob stays put), so by the assumption of complementarity this is not a paradox.

1. Fast-approaching Bob

In order for Alice and Bob to compare their results, Bob cannot stay this far all the time. He has to approach the black hole very fast and cross the horizon at some $t < t_{\text{cross}} + \Delta t$. Note that the t defined here are the Schwarzschild time when the observers cross some fixed distance outside but near the horizon, for example $r = 2M + 1$ cm. Depending on Δt, there is a chance that he can still communicate with Alice. This situation is drawn in Fig. 3.

The first constraint on such a scenario is that the Hawking radiation should not be too blueshifted for Bob:

$$v = \frac{M^3}{M^3 + \Delta t} \approx 1 - \frac{\Delta t}{M^3}, \qquad (3.1)$$

$$\gamma \sim \sqrt{\frac{M^3}{\Delta t}}, \qquad (3.2)$$

$$E_{R,\text{Bob}} \sim M^{-1}\gamma = \sqrt{\frac{M}{\Delta t}}. \qquad (3.3)$$

Bob would like the Hawking quanta to remain sub-Planckian in his frame, and therefore Δt needs to be much larger than M. This, on the other hand, makes Alice's job quite difficult. First of all, there is a geometric constraint that if Bob jumps in Δt later than Alice, then Alice has to send a message within

---

[9]In principle the state of A is also changed by what Alice did, but the exact change does not matter in our argument. Alice enforced an observable entanglement on B and R_B, which must exclude another observable entanglement.

[10]Hawking quanta coming out at even later times are modified by the backreaction.
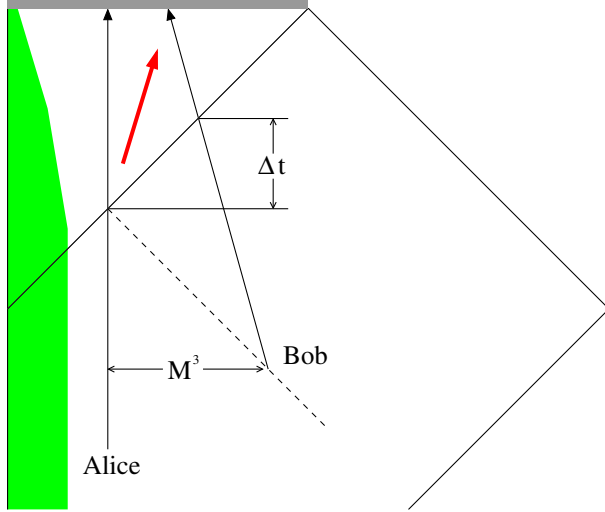
FIG. 3 (color online).   Bob starts to distill R_B from R outside the past light cone of the point which Alice crosses the horizon, which means that he has to be $\sim M^3$ away in the beginning. He approaches the black hole and follows Alice into the horizon Δt later. If Δt is too small, the early Hawking quanta have Planckian wavelength such that Bob cannot read them. If Δt is too big, then the message from Alice (the red, thick arrow) will be Planckian such that Bob cannot read it.

$$\Delta\tau_{\text{Alice}} < Me^{-\Delta t/M}, \qquad (3.4)$$

such that Bob can receive it before crashing into the singularity [22]. Note that $\Delta\tau$ here refers to the proper time for Alice, and the exponential relation between this time and the Schwarzschild time is the key of this argument. She will only have $\Delta\tau$ to explore the black hole interior before sending the information, so the wavelength of the photon she sends that can contain the information about the interior mode A must be bounded by the same quantity:

$$\lambda_{\text{Alice}} < \Delta\tau_{\text{Alice}}. \qquad (3.5)$$

This message will again be somewhat blueshifted to Bob:

$$E_{\text{message from Alice}} = \frac{\gamma}{\lambda_{\text{Alice}}} > \sqrt{\frac{M^3}{\Delta t}}M^{-1}e^{\Delta t/M} = \sqrt{\frac{M}{\Delta t}}e^{\Delta t/M}. \qquad (3.6)$$

We have used the same γ as in Eq. (3.3), and here it is only an approximation. However, the point is that when $\Delta t > M$, the exponential behavior dominates and the exact γ will not compensate its effect. We see that there is no way to make Eqs. (3.3) and (3.6) simultaneously small. Bob either has to deal with Planck scale Hawking radiation, or a Planck scale message from Alice. *If we want to avoid backreaction, there is a fundamental obstruction to observing A and R_B together.*

E. Unitary evolution

The practical paradox (Alice in Fig. 1) is the strongest version of the AMPS argument. Following the footsteps of how we resolved it, the resolution to other versions becomes transparent. For example, one way to phrase the AMPS paradox requires no distillation since it only discusses the validity of a quantum mechanical description. Let $|\phi\rangle$ be the state of an infinite spatial slice before its matter collapses into a black hole. We can follow the quantum mechanical evolution to a later time where A, B and R are all present. In order to maintain a normal horizon between A and B, and a unitary evaporation process relating B and R, A and R must contain the same information. The information paradox can be stated as the violation of the unitarity evolution from $|\phi\rangle$ to $|A\rangle|B\rangle|R\rangle$.

Complementarity claims that this is not a paradox, unless the spatial surface describing this evolution can fit into the causal patch of some observer. Usually we picture the situation that R is very far away and cannot fit into the same causal patch with A so it is not a problem. One might instead try to fit A, B and R into one causal patch as is shown in Fig. 4. However following the discussion of a fast-approaching Bob in Sec. III D 1, we can see that such a fit is problematic. The top of this causal patch can be treated as where a fast-approaching Bob eventually ends up. Thus the early radiation R on this slice is exactly the one Bob reads on his journey, and its wavelength is given by Eq. (3.3). For the same reason, the physical size of A on this slice is related to the message sent by Alice as in Eq. (3.6). So following the math in the previous section, either the interior mode A or the early radiation R has to be a Planckian quantity. Complementarity is intrinsically a
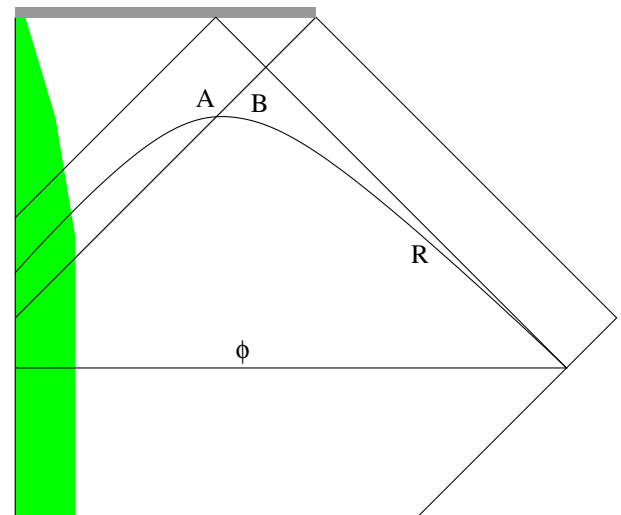

FIG. 4 (color online).   The attempt to fit the interior partner A, the late quanta B, and early Hawking radiation R into one causal patch. Inevitably either A or R has to be a Planckian quantity and stops us from using low energy quantum mechanics to evolve from $|\phi\rangle$ to $|A\rangle|B\rangle|R\rangle$.

statement on low energy physics, so there is a natural and implicit fine print: *Not only do we need A, B and R to fit into a causal patch, all of them should be low energy quantities.* This is violated here, and thus cannot qualify as a paradox for complementarity. A more thorough analysis of this simple geometrical fact can be found in [23].

IV. DISCUSSION

We argue that the practical paradox formulated by AMPS [2] can be resolved without new physics. A quick summary of this practical paradox is for Alice to

(1) distill R_B from the early Hawking radiation R,
(2) bring R_B to check its entanglement with the corresponding late near-horizon mode B,
(3) cross the horizon to check the entanglement between B and its interior partner A.

We show that the "distillation" process, if carried out in the causal past of the point Alice jumps in, modifies the black hole through backreaction. We make a plausible argument that a successful (nontrivial) distillation implies a minimum level of backreaction.

In light of this, there are two possible outcomes for Alice and neither is paradoxical. If Alice is not careful enough and believes that she can distill R_B from R just like Bob (who is $\sim M^3$ away) does, her check of unitary evaporation will fail. That is because the backreaction makes the black hole and her distiller interact. Without taking this into account, the subtle entanglement she tries to verify would not exist. She can perform the distillation also from $\sim M^3$ away to prevent backreaction (i.e. the Bob situation). In that case she will confirm the unitary evaporation but will be unable to jump in and observe A, thus no practical paradox.

If Alice is very careful, she can stay close and perform something that is much more involved than a simple distilling. She can design a device to coherently affect the black hole together with a qubit she holds. This ensures that the particular late outgoing quantum does end up being maximally entangled with her qubit. However doing so is exactly using her powerful knowledge to disrupt the entanglement between A and B. While crossing the horizon, Alice will see some high energy quanta due to the mismatch in $A - B$, but it is not a spontaneous violation of the equivalence principle. Alice has been manipulating the black hole for a long time and eventually has to taste her own medicine. In short, Alice herself creates the firewall.

This backreaction is causal, and we present a model of how it works using the standard theory of radiation-source response. Furthermore, we argue that, viewed in the nice slices, breaking the cross-horizon entanglement is the natural interpretation of the physical effect of this

backreaction. Our resolution of the AMPS paradox requires no new physics and is perhaps the most conservative one proposed so far. The starting point of our argument is a very general statement: Eq. (2.1) must effectively describe any practical method to distill R_B from R. For a more detailed implementation, we picked the least intrusive method, that the distiller simply waits for the Hawking radiation to come. Since there is already backreaction in such a minimal setup, it seems unlikely that more intrusive/elaborate methods, such as mining [24,25], can evade backreaction.

In fact we would like to make a stronger statement by employing the following general definition of the distillation process. Whenever and wherever Hawking radiation stops free streaming in vacuum and interacts with something else, those space-time regions are considered part of the "distiller."[11] Equation (2.1) continues to describe such a distillation process, and our backreaction argument should apply.

Although our resolution is conservative, it has interesting implications. We provide two qualitative descriptions of how the backreaction modulates the evaporation process in Sec. II, according to the membrane paradigm and the expanding geometry on the nice slices. The latter picture is particularly interesting, since in principle the calculation is similar to the one carried out in the context of inflation, in the presence of external disturbances that cause nontrivial excitations. It is possible that we can employ those techniques and try to learn something about the black hole S matrix.

ACKNOWLEDGMENTS

---

[11]For example one can try to arrange mirrors that reflect radiations to shield the black hole from the backreaction. These mirrors will inevitably affect the normal outgoing radiation as well—mirrors that perfectly transmit in one direction and reflect in the other do not exist. We can just treat them effectively as part of the distiller. There is a proposed setup that includes an auxiliary system [4] which has no clear interpretation purely in the bulk, so we do not consider it as constituting a practical paradox.

[1] S. W. Hawking, Phys. Rev. D **14**, 2460 (1976).
[2] A. Almheiri, D. Marolf, J. Polchinski, and J. Sully, J. High Energy Phys. 02 (2013) 062.
[3] S. L. Braunstein, S. Pirandola, and K. Życzkowski, Phys. Rev. Lett. **110**, 101301 (2013).
[4] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford, and J. Sully, J. High Energy Phys. 09 (2013) 018.
[5] G. 't Hooft, Nucl. Phys. **B 335**, 138 (1990).
[6] L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **48**, 3743 (1993).
[7] L. Susskind, Phys. Rev. Lett. **71**, 2367 (1993).
[8] E. Verlinde, arXiv:hep-th/9503120.
[9] D. A. Lowe, J. Polchinski, L. Susskind, L. Thorlacius, and J. Uglum, Phys. Rev. D **52**, 6997 (1995).
[10] R. Bousso, Phys. Rev. D **79**, 123524 (2009).
[11] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[12] S. G. Avery, B. D. Chowdhury, and A. Puhm, J. High Energy Phys. 09 (2013) 012.
[13] B. D. Chowdhury, J. High Energy Phys. 10 (2013) 034.
[14] R. Bousso, Phys. Rev. D **88**, 084035 (2013).

[15] D. N. Page, New J. Phys. **7**, 203 (2005).
[16] J. Maldacena and L. Susskind, Fortschr. Phys. **61**, 781 (2013).
[17] D. Harlow and P. Hayden, J. High Energy Phys. 06 (2013) 085.
[18] R. H. Price and K. S. Thorne, *The membrane paradigm* (Yale University Press, 1986).
[19] S. D. Mathur, Classical Quantum Gravity **26**, 224001 (2009).
[20] L. Hui and I. S. Yang (private communication).
[21] L. Hui, S. T. McWilliams, and I.-S. Yang, Phys. Rev. D **87**, 084009 (2013).
[22] Y. Sekino and L. Susskind, J. High Energy Phys. 10 (2008) 065.
[23] I. Ilgin and I.-S. Yang, Phys. Rev. D **89**, 044007 (2014).
[24] W. Unruh and R. Wald, Gen. Relativ. Gravit. **15**, 195 (1983).
[25] A. R. Brown, Phys. Rev. Lett. **111**, 211301 (2013).