Evaporating firewalls

Mark Van Raamsdonk

Department of Physics and Astronomy, University of British Columbia,
6224 Agricultural Road, Vancouver, B.C., V6T 1W9, Canada

E-mail: mav@physics.ubc.ca

Abstract: In this note, we begin by presenting an argument suggesting that large AdS black holes dual to typical high-energy pure states of a single holographic CFT must have some structure at the horizon, i.e. a fuzzball/firewall, unless the procedure to probe physics behind the horizon is state-dependent. By weakly coupling the CFT to an auxiliary system, such a black hole can be made to evaporate. In a case where the auxiliary system is a second identical CFT, it is possible (for specific initial states) that the system evolves to precisely the thermofield double state as the original black hole evaporates. In this case, the dual geometry should include the "late-time" part of the eternal AdS black hole spacetime which includes smooth spacetime behind the horizon of the original black hole. Thus, if a firewall is present initially, it evaporates. This provides a specific realization of the recent ideas of Maldacena and Susskind that the existence of smooth spacetime behind the horizon of an evaporating black hole can be enabled by maximal entanglement with a Hawking radiation system (in our case the second CFT) rather than prevented by it. For initial states which are not finely-tuned to produce the thermofield double state, the question of whether a late-time infalling observer experiences a firewall translates to a question about the gravity dual of a typical high-energy state of a two-CFT system.

Contents

1 Introduction

There has been much recent debate over the fate of an infalling observer at the horizon of a black hole. Classically, the existence of solutions of Einstein's equations in which smooth spacetime continues uninterrupted past the horizon suggests that an infalling observer will notice nothing special at the horizon. Semi-classically, the absence of a large stress-energy tensor at the horizon requires a particular entanglement structure between quantum field theory modes on the two sides of the horizon. Almheiri, Marolf, Polchinski, and Sully (AMPS) have argued that such entanglement is not possible in an old black hole that has become maximally entangled with its Hawking radiation, assuming unitarity of black hole evaporation and validity of effective field theory outside the horizon [1]. Thus, they argue that an infalling observer will experience a firewall (if the assumptions hold).[1]

Firewalls were also proposed in a different context in [43], following [44–53]. There, in the nonperturbative setting of AdS/CFT, it was argued that the existence of any spacetime at all behind a general horizon (including Rindler horizons in pure AdS) requires entanglement between the degrees of freedom associated with the region outside the horizon and some other independent degrees of freedom (see figure 1). Based on these observations, it was proposed in the context of hyperbolic AdS black holes that black hole microstates (typical pure states of a CFT on H^d) are dual to black holes with no spacetime past the horizon, i.e. with a lightlike singularity or firewall at the would-be horizon. As we review in section 2, very similar arguments suggest that large AdS black hole microstates described by typical high-energy states of a single CFT have firewalls (or fuzzballs). An alternative argument for this has been given recently in [37]; our argument seems independent of this

---

[1]See also the closely related earlier work [2–5]. Since [1], there have been many arguments for and against the firewall proposal, including [6–42].
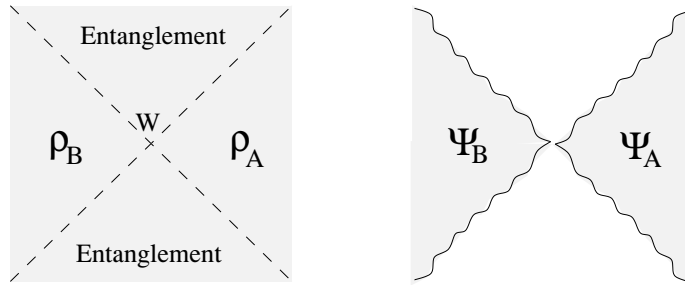
Figure 1. Proposed relation between spacetime structure and entanglement structure [43]. Density matrices for complementary sets of fundamental degrees of freedom A and B determine physics in the left and right wedges formed by light sheets from an extremal surface W. Physics in the upper and lower regions is encoded in the entanglement between A and B. If this entanglement is removed by placing A and B in typical pure states in the ensembles ρ_A and ρ_B, the right and left wedges are almost unchanged while the upper and lower regions are excised. Thus, the left and right wedges are connected by entanglement, while disentangling leaves disconnected wedges ending in a firewall.

and in particular does not make reference to perturbative quantum field theory modes in the bulk.[2]

Since these large AdS microstate black holes are not maximally entangled with their Hawking radiation, they are young black holes in the sense of AMPS. Thus, their firewalls are not forced upon us by the original AMPS argument. To study "old" black holes in the context of AdS/CFT, and thus to evaluate the AMPS argument in a non-perturbative setting, we need to make our large AdS black holes evaporate. To achieve this, we consider in section 3 a thought experiment in which we weakly couple the CFT to an auxiliary "radiation" system.[3] Specifically, we imagine that the CFT (originally in a high-energy pure state) is coupled weakly to another identical CFT, for example by placing a wire between the two spheres on which the CFTs live.

Through this weak interaction, the two-CFT system evolves from a disentangled product state to a state where the two CFTs are nearly maximally entangled (given the constraints of energy conservation). In the dual gravity picture, the interaction allows radiation from the original black hole to leak out at the AdS boundary and enter a second asymptotically AdS spacetime where it collects and eventually forms a second black hole. For a specific choice of initial state, it is possible that the entangled state that the CFTs evolve to is precisely the thermofield double state. In this case, we argue that the gravity dual to the two-CFT system includes a "late-time" region of the eternal AdS black hole spacetime. Thus, an observer falling into the original black hole will find smooth spacetime behind the horizon.

At least for the specific initial states that lead to the thermofield double state, our conclusions are that the black hole starts with a firewall but ends with a smooth horizon.

---

[2]A related discussion appears in [38, 39] who argue that the gravity dual of the thermal state of a single CFT has a firewall, though we would say instead that different purifications of this thermal state correspond to different physics past the horizon which may or may not include a firewall.

[3]A similar construction was considered in [37].

Loosely, we could say that the black hole mellows with age and the firewall evaporates. Since the late black hole is maximally entangled with the radiation system, it is an old black hole in the sense of AMPS, and the AMPS arguments should apply. The fact that we find a smooth horizon suggests that the conclusions of AMPS can be avoided, i.e. that one of the assumptions (explicit or hidden) must be incorrect.

The thought experiment here is a specific realization of the recent ideas of Maldacena and Susskind [54]. They conjectured that the Hawking radiation can act like the second CFT in the thermofield double state; that is, the maximal entanglement between an old black hole and its Hawking radiation effectively creates a behind-the-horizon region of the spacetime. In our setup, we have simply made this more concrete by arranging that the Hawking radiation degrees of freedom *are* a second copy of the CFT and that the entangled state *is* the thermofield double state where we already know the dual spacetime. At least in this special case, we see that the spacetime behind the horizon is not forbidden by the entanglement between the black hole and its Hawking radiation; rather, it is a manifestation of this entanglement.

Ending up with the thermofield double state requires a very finely tuned initial state. This is sufficient in order to demonstrate that it is *possible* to end up with a smooth horizon in an old black hole, but leaves open the question of whether the firewalls proposed in section 2 generically disappear through the process of black hole evaporation. We have no definitive answer for this question, but we offer a few comments in section 4.

We conclude in section 5 with some further discussion, including comments on various related works.

2   Fuzzballs/firewalls for black hole microstates in AdS

In this section, we would like to ask what happens to an observer who falls into a large AdS black hole microstate described by a pure state $|\psi\rangle$ in a single CFT on S^d, using the ideas in [43].[4] We assume that this is a typical state in the thermal ensemble at some temperature.

Before reaching the horizon, the observer experiences the same Schwarzschild-AdS geometry, regardless of which black hole microstate we are in. From the CFT point of view, we can explain this by saying that the operators which tell us about the geometry outside the horizon are not sensitive to the fine details of the microstate; they may be computed using the reduced density matrix for a small fraction of the degrees of freedom. Such observables have almost exactly the same value for any typical state in the thermal ensemble (see, for example, [55]).

We would now like to ask what happens when the observer reaches the horizon. The AdS/Schwarzschild geometry has a standard extension to a smooth spacetime extending behind the horizon to a spacelike singularity. A simple possibility is that for any typical microstate, the observer experiences this geometry to a good approximation, at least for

---

[4]This section is based on talks given at the 2012 Amsterdam String Workshop and at Stanford University. I am grateful to the audience members for helpful comments that have refined the argument.

some substantially larger than Planck-scale distance beyond the horizon. We will now argue against this possibility.

Suppose that this hypothesis is correct. In this case, there should be some observable in the CFT that can be used to tell us about the local physics just behind the horizon.[5] According to the hypothesis, this observable will have approximately the same value for any typical state $|\psi\rangle$. In this case, the observable will also have the same value in the thermal state, described by a density matrix

$$\rho = Z^{-1} \sum_i e^{-\beta E_i} |E_i\rangle\langle E_i|$$

since the states $|\psi\rangle$ we are considering are typical states in this ensemble. So we could say that the thermal state ρ corresponds to an AdS-Schwarzchild black hole spacetime including at least some of the region behind the horizon in the extended Schwarzchild geometry. Defining M_ρ as the geometry associated with the density matrix ρ, we can say that M_ρ contains the geometry in common between the duals $M_{|\psi_i\rangle}$ of all the typical states $|\psi_i\rangle$ in the ensemble ρ.

The density matrix ρ can also arise as a reduced density matrix starting from various pure states $|\Psi\rangle$ in theories with degrees of freedom beyond those of the single CFT. The state $|\Psi\rangle$ is a "purification" of the density matrix ρ. In some cases, the state $|\Psi\rangle$ will have a dual gravity interpretation corresponding to a classical spacetime M. Since knowledge of ρ alone (a subset of the information in Ψ) allows us to deduce the existence of a classical geometry M_ρ, we expect that M_ρ should be present as a subset of the geometry M dual to any purification Ψ. Thus, M_ρ cannot be larger than the region of spacetime common to all geometries $M_{|\Psi\rangle}$ dual to the various purifications $|\Psi\rangle$. To summarize, for typical states $|\psi_i\rangle$ in the ensemble ρ and any set of states $|\Psi_i\rangle$ that purify ρ, we should have

$$\cap M_{|\psi_i\rangle} \subset M_\rho \subset \cap M_{|\Psi_i\rangle}.$$

By considering a particular class of purifications, we will now argue that M_ρ cannot be larger than the region outside the horizon of the AdS-Schwarzschild black hole. In the system we consider, the additional degrees of freedom are the degrees of freedom of a second identical CFT on S^d. In this system, one possible purification of ρ is the thermofield double state

$$|\Psi_{\mathrm{TD}}\rangle = Z^{-1/2} \sum_i e^{-\beta E_i/2} |E_i\rangle \otimes |E_i\rangle.$$

For this state, the dual geometry M is the maximally extended Schwarzschild-AdS black hole [45]. However, there are many other possible purifications of ρ in the two-CFT system. The most general purification may be written as $(U \otimes \mathbb{1})|\Psi_{\mathrm{TD}}\rangle$, where U is a unitary

---

[5]Crucially, we assume here that the same operator can be used for each CFT state, i.e. that the operator is not state-dependent. Thus, the argument in this section can be taken as a "proof" that the absence of firewalls for large AdS black holes dual to pure states of a CFT implies that the spacetime behind the horizon must be reconstructed using state-dependent operators. See section 5 for more discussion of this point.
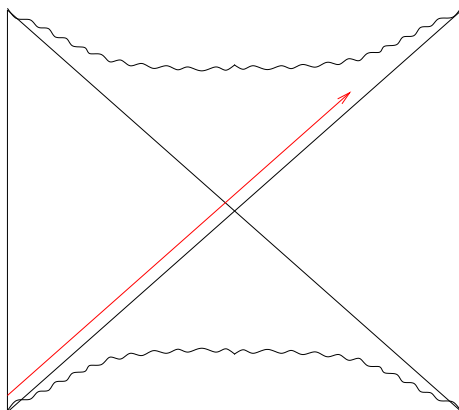
Figure 2. Gravity dual of state $(U \otimes \mathbb{1})|\Psi_{\text{TD}}\rangle$ for unitary U close to the identity. Perturbations originating in the second asymptotic region propagate past the horizon. For general perturbations, only the exterior of the black hole in the first asymptotic region (on the right) is unaffected.

operator that acts only on the auxiliary CFT (corresponding to the left factor in the tensor product).

It is easy to check that the reduced density matrix for the original CFT is ρ for all these states. On the other hand, observables involving fields in the second CFT will generally differ depending on the operator U. From the perspective of the dual spacetime, we can think of acting with the operator U as causing perturbations to the second asymptotic region.[6] These will propagate past the horizon and in general affect the full spacetime everywhere except the first asymptotic region outside the horizon, as shown in figure 2. For a generic state $U|\Psi\rangle$ we expect that the spacetime past the horizon of the first asymptotic region will differ in an arbitrarily large way from the spacetime corresponding to $|\Psi_{\text{TD}}\rangle$. Thus, the region in common to the spacetimes M corresponding to the various purifications of ρ in our two-CFT system is strictly the exterior of the AdS Schwarzschild black hole.

In summary, the hypothesis that gravity dual geometries for typical microstates in the thermal ensemble have a common smooth region behind the horizon suggests that the thermal state itself can be associated with a geometry M_ρ including such a region. But M_ρ should be a subset of any spacetime M dual to a purification $|\Psi\rangle$. From the two-CFT example, it is clear that the only common region in such geometries M is the exterior of the black hole. Thus, there can be no common region behind the horizon of geometries corresponding to typical microstates.

Our conclusion essentially leads to the fuzzball picture of black hole microstates [2] i.e. that black hole microstates are geometries for which the region behind the black hole horizon is replaced by some structure that is different for different microstates.[7] For some

---

[6]For example, U might be a local or nearly local CFT operator whose effect is to perturbs the spacetime near some boundary point.

[7]A further qualitative argument for our conclusion is the following. From the CFT perspective, the observer falling into the horizon corresponds to some initially localized information being added to the CFT state (placing an observer near the boundary) and getting scrambled. It seems natural to identify the time when the observer hits the horizon with the scrambling time. At this time, the information cannot be

special microstates, this structure may correspond to a smooth, weakly curved geometry, but in general, we expect that the structure is Planck-scale. The simplest interpretation is that classical spacetime simply ends at the would-be horizon, as we argued for the hyperbolic black hole microstates in [43].[8] In this case, an observer falling into the black hole will effectively encounter a lightlike singularity at the horizon; the fuzzball acts as a firewall.[9] Our conclusion is the same as in the recent work [37], though the arguments here appear to be rather different.

It is important to note that the black holes we consider are not old black holes maximally entangled with their Hawking radiation. Thus, the firewalls are apparently not the "old age" firewalls of AMPS, but rather the "purity" firewalls of [43]. To make contact with AMPS, we consider in the next section a thought experiment where we force one of these pure black holes to evaporate. Surprisingly, we find that this evaporation process can in some cases eliminate the firewall and give rise to smooth spacetime behind the horizon.

3 Evaporating large AdS black holes

Large black holes in Anti-de-Sitter space typically do not evaporate. They are in equilibrium with their Hawking radiation, which reaches the boundary of AdS and returns to the black hole in finite time. However, by adjusting the boundary conditions, we can allow some of the radiation to leak out to an auxiliary system (as discussed recently in [37]). In this case, the black hole can evaporate. After some time, the black hole subsystem (which we will take to be a holographic CFT) becomes maximally entangled with the auxiliary system (which we will take to be a second holographic CFT) that contains the radiation. We can then ask what happens to an observer who falls into the black hole.

Our specific setup is the following. We start with a single CFT on S^d in a typical high-energy (pure) state $|\psi_0\rangle$ with $\langle H \rangle = E$. We consider a second identical CFT on S^d that is initially in its vacuum state. We can imagine that these CFTs are living on spheres in a lab, as in the recent discussion of [54]. We now connect the two spheres by a "wire" so that some UV degrees of freedom in each CFT are weakly coupled (figure 3).[10]

---

recovered by looking at any small subset of degrees of freedom in the system. To recover the information that there was an observer, we need to use observables that are sensitive to most of the accessible degrees of freedom of the system at once. These are exactly the kind of observables we need to distinguish the individual microstates. Thus, any observables that tell us about the existence or experience of the observer after they have reached the horizon can be expected to give different results for different microstates.

[8]We emphasize that it is still possible for certain special pure states to have gravity duals which contain some smooth spacetime behind a horizon. For example, we expect that state corresponding to the collapse of a null shell to form a black hole will temporarily include some smooth locally AdS spacetime inside the horizon that forms due to the collapsing shell.

[9]It is not necessary that the observer burns up in a some classical way. Indeed, if it were the case that the dual of a typical microstate includes a bulk energy density that diverges in a particular way, this diverging behavior would appear also in the thermal state and thus also for the eternal black hole. We thank Don Marolf for emphasizing this point.

[10]Mathematically, we add an interaction term to the Hamiltonian. This could be of the form $H_{\text{int}} = \int_1 \int_2 \mathcal{O}_1(x_1)\mathcal{O}_2(x_2)$, where $\mathcal{O}_1$ and $\mathcal{O}_2$ have support in a small region of their respective spheres. We could also use many wires distributed around the sphere so that the interaction is approximately spherically symmetric.
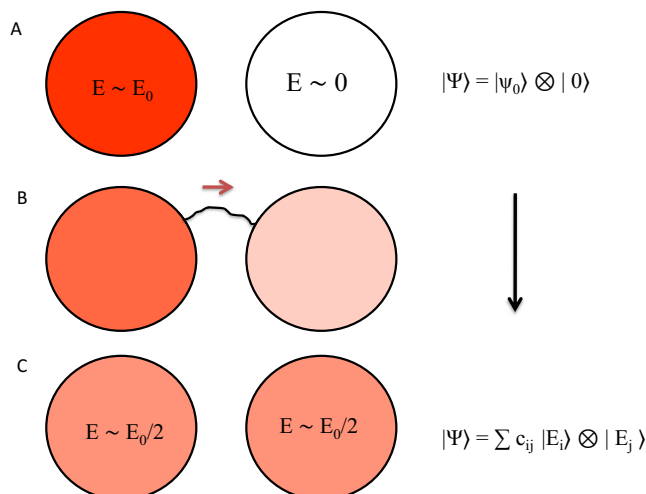
Figure 3. CFT picture: when CFTs are weakly coupled, the product state evolves to a highly entangled state.

The first CFT in the state $|\psi_0\rangle$ corresponds to a large AdS black hole microstate in a spacetime that is asymptotically global AdS. This black hole is in equilibrium with its Hawking radiation. When we attach the wire, a localized region of the AdS boundary experiences a change in boundary conditions, such that some of the radiation no longer bounces back, but is lost to an external system. Initially, there is no radiation coming in from this external system, so the black hole begins to evaporate slowly. As time passes, radiation begins to enter from the external system, and eventually the rate of outgoing radiation matches the rate of incoming radiation and the black hole (now smaller) stops evaporating.

The second CFT, initially in its vacuum state, corresponds to an empty global AdS spacetime. However, when we attach the wire, a localized region of the spatial boundary of this AdS space experiences a change in boundary conditions, and radiation begins to enter the spacetime from this region of the boundary. Initially, this radiation gives rise to a gas of gravitons in AdS space, but eventually this collapses to form a black hole that grows in size. At some point the flux of energy out through the boundary matches the flux of energy in from the boundary, so the black hole reaches an equilibrium size.

In summary, the bulk picture is that after a long time, we have two black holes in different asymptotically global AdS spacetimes. Though the two CFTs start in a disentangled state $|\psi_0\rangle \otimes |0\rangle$, the second law of thermodynamics ensures that they will evolve to a state of the form

$$|\Psi\rangle = \sum c_{ij}|E_i\rangle_1 \otimes |E_j\rangle_2 \qquad (3.1)$$

where the entanglement entropy is close to the maximum possible value given the total energy of the system [56, 57]. While the individual states $|E_i\rangle_1 \otimes |E_j\rangle_2$ in the superposition (3.1) correspond to disconnected spacetimes, a quantum superposition of the form (3.1) with near maximal entanglement can (at least in special cases) correspond to a spacetime in which the two asymptotic regions are connected by a smooth spacetime behind the

horizons of their respective black holes [45, 48, 54]. In this case, we might expect that an observer falling into the original black hole at these late times will find some smooth spacetime behind the horizon.

For a general highly entangled state of the form (3.1), we do not know what the dual spacetime is, so it is difficult to evaluate this suggestion. However, at least for certain initial states $|\psi_0\rangle$ and certain choices for the weak interaction Hamiltonian, we can ensure that the state of the two-CFT system after some time is precisely the thermofield double state[11]

$$|\Psi_{\text{TFD}}\rangle = Z^{-1/2} \sum e^{-\beta E_i/2} |E_i\rangle_1 \otimes |E_i\rangle_2 \,. \tag{3.2}$$

In this case, the dual geometry for late times matches with the eternal black hole in AdS, and we can be certain that an observer falling into the original black hole will not see a firewall. A more detailed argument for this is presented in the following subsection.

Achieving exactly the thermofield double state clearly requires severe fine-tuning of the initial state, but a finely-tuned example suffices to show that an observer falling into a black hole that has become maximally entangled with its Hawking radiation does *not necessarily* encounter a firewall. Thus, the conclusions of AMPS can apparently be avoided. Apparently in this case the process of a black hole becoming maximally entangled with its Hawking radiation removes a firewall rather than creating it. As we mentioned in the introduction, our argument here is closely related to the recent paper by Maldacena and Susskind [54]; the construction here provides a concrete example where entanglement between a black hole and its Hawking radiation can indeed lead to smooth spacetime behind the horizon.

In section 4, we return to the question of whether firewalls exist in old black holes for generic initial states.

3.1 Gravity duals for CFTs with changing Hamiltonians

In the previous section, we have considered a state of a two CFT system in which one CFT is originally in a typical high-energy pure state while the other CFT is in the vacuum state. At some common field theory time, the two CFTs are coupled and the state evolves to the thermofield double state. At this point, the two CFTs are decoupled again, and the entangled state evolves as usual. What can we say about the gravity dual of this system with a changing Hamiltonian?

Generally speaking, suppose that in the Schrodinger picture of a holographic quantum theory with Hamiltonian H we have a state $|\Psi(t)\rangle$ whose gravity dual is known to be some spacetime M. Now suppose we consider a Hamiltonian H' that differs from H after time T and a state $|\Psi'(t)\rangle$ that differs from $|\Psi(t)\rangle$ only after time T. In the gravity dual picture, changing the Hamiltonian at time T corresponds to changing the boundary conditions some boundary time. The effects of this change propagate into the bulk from

---

[11]In fact, for any generic weak interaction between the two CFTs, the reverse time evolution of the thermofield double state will eventually (on recurrence timescales) lead to a state where almost all the energy is in the first CFT, and there is almost no entanglement between the two CFTs. Taking this to be the initial black hole state in our setup will ensure that we end up with exactly the thermofield double state at some time.
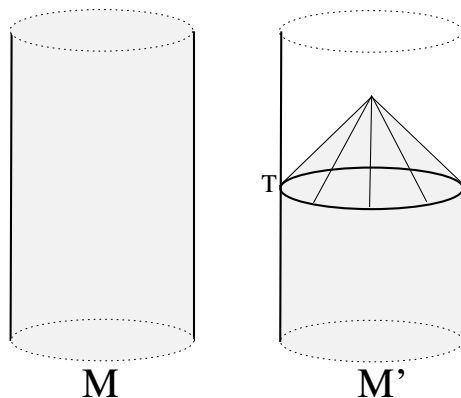
Figure 4. Gravity dual M' of a state $|\Psi'(t)\rangle$ that is the vacuum state before time T but changes after time T due to a perturbation of the Hamiltonian. The spacetime M' matches pure global AdS in the past of the future light sheet from the boundary surface at time T (shaded).

the boundary surface at time T. Thus, the spacetime M' dual to $|\Psi'(t)\rangle$ should match the spacetime M in the past of the future light sheet of the boundary surface at time T. This is illustrated in figure 4 for the example where M is pure AdS.

Similarly, if we have a state $|\Psi(t)\rangle$ dual to M and a state $|\Psi'(t)\rangle$ that differs from $|\Psi(t)\rangle$ for all times after time T', the gravity dual M' to $|\Psi'\rangle$ should match with M at all points in the future of the past light sheet of the boundary surface at time T'.

Applying these requirements to the state we are interested in, we conclude that the dual spacetime to the two CFT system that is weakly coupled between times T and T' should include early-time portions of pure AdS and pure-state black hole geometries, and late-time portions of the eternal AdS black hole geometry, as shown in figure 5. In particular, while the whole spacetime does not match the eternal black hole spacetime, the smooth spacetime region behind the horizon at late times will be present, as we require for the argument in the previous section. A schematic picture of a possible spacetime with these features is shown in figure 6, though the details of this picture should not be taken too seriously.

4   Generic case

Ending up with the thermofield double state requires a finely tuned initial state. This is sufficient in order to demonstrate that it is *possible* to have a smooth horizon in an old black hole, but leaves open the question of whether a late-time observer in our setup will generically experience a firewall.

In general, the state of the system after a long time is a highly entangled state

$$|\Psi\rangle = \sum c_{ij}|E_i\rangle_1 \otimes |E_j\rangle_2. \tag{4.1}$$

Assuming that the energy of the initial black hole is in $[E, E + dE]$ and the interaction is very weak, we can think of this final state as a typical state in the microcanonical ensemble for the two-CFT systems with total energy $E_{\text{TOT}} \in [E, E + dE]$. To answer the question
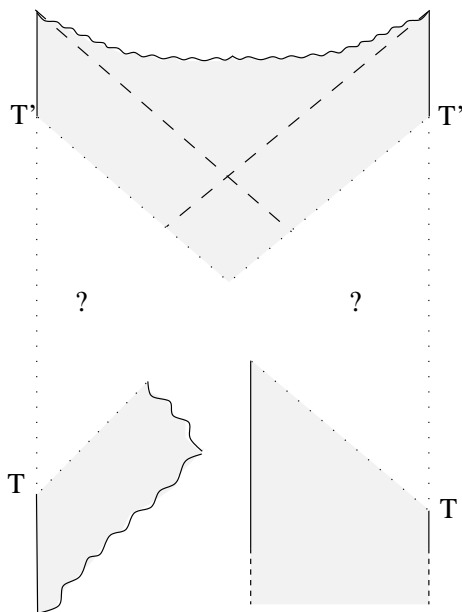
Figure 5. Gravity dual of a state of a two-CFT system where the two CFTs are coupled between time T and T'. Before time T and after time T', the CFT states are identical to states with known dual geometries.
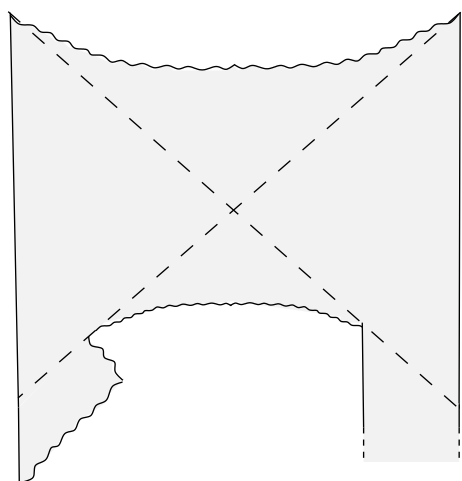

Figure 6. Schematic of possible dual spacetime to product state of a high-energy state of one CFT with the vacuum of another CFT evolving via a weak coupling to the thermofield double state.

of whether a late-time infalling observer in our setup experiences a firewall, we need to understand the gravity dual description of such a typical state in the two-CFT system. This is beyond the scope of the present work, but we will make a few observations.

We note first that very small perturbations to the thermofield double state can lead to a firewall. Consider a perturbation that adds a small number of low-energy particles falling into the black hole on one side. As emphasized recently in [58], an observer falling into the black hole who encounters these quanta (e.g. observer A in figure 7) will observe
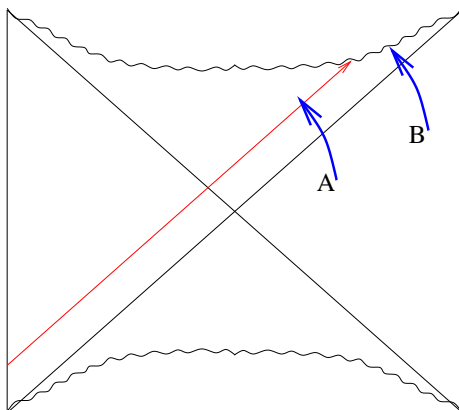
Figure 7. Quanta falling into the black hole on one side will be observed exponentially blue-shifted by observer A falling into the black hole on the other side. Observer B who falls in late enough does not observe these. Back-reaction is not taken into account in the diagram, but this does not affect the conclusion.

them to be exponentially blue-shifted. On the other hand, if the observer falls in late enough (observer B in figure 7), the quanta will not be observed. Thus, for these small perturbations of the thermofield double state, the question of whether an infalling observer will encounter high-energy quanta behind the horizon is related to whether particles entered the horizon of the auxiliary black hole at early enough times.

The existence of not of these destructive infalling particles should be visible directly using observables in the auxiliary CFT (i.e. by knowing the density matrix of the second CFT), since they correspond to physics outside the horizon. This physics is generally believed to be encoded in "simple" observables that do not involve a large fraction of the degrees of freedom of the CFT. Thus, such observables can be calculated using a reduced density matrix for a small fraction of the degrees of freedom of the two-CFT system. But such a density matrix will be almost the same for any typical states of the two-CFT system. Further, it will be almost identical to the reduced density matrix for the case where the two-CFT system is in a thermal state. For this case, the single-CFT density matrix will also be thermal. Thus, we can say that for a typical state of the form (4.1), any simple observable in a single CFT will be almost identical to that observable in a thermal state of the single CFT. We conclude that for the gravity dual of a typical state two-CFT state of the form (4.1), the geometry outside the horizon in each asymptotically AdS region will be almost identical to an unperturbed black hole geometry.

Thus, we seem to have two competing effects: typicality ensures that the physics outside the horizon on each side is very close to that of the unperturbed black hole, but small perturbations on one side appear exponentially amplified to observers falling in to the black hole on the other side (if they fall in early enough). We leave a more detailed investigation of these competing effects to future work.

The preceding comments are only relevant to states that are close to the thermofield double state, so that the geometrical picture in 7 is approximately valid. For more typical states of the form (4.1), there is no reason to believe that something falling behind the

horizon of the second black hole will be able to affect the experience of an observer who falls into the first black hole. Indeed, for a general state, the mutual information between any localized degrees of freedom in one CFT with localized degrees of freedom in the other CFT will be very small.[12] According to the general expectations in [48], this means that the two asymptotic regions are very far apart; alternatively, the lack of correlations between local operators in the two CFTs suggests that there will be no spacelike geodesics connecting the two asymptotic regions. But whether or not observers from the two sides can meet behind the horizon (i.e. the two asymptotic regions are close to one another) is independent from the question of whether each observer will see smooth spacetime behind the horizon.

5 Discussion

If the conclusions of AMPS can be avoided, it must be that either one of the basic postulates is false,[13] or there is a flaw in the argument. In this paper, we do not try to decide which of these is true. We simply present our observations as evidence (building on the recent discussion of Maldacena and Susskind) that "no drama" at the horizon is a logical possibility for old black holes. At the same time, our comments on single-CFT pure state black holes suggest that firewalls (or fuzzballs) do play a role in black hole physics, at least for large black holes in Anti-de-Sitter space. If the pure-state firewalls exist, it means that measuring the full state of the black hole in the auxiliary CFT will project the original black hole into a state with a firewall. Equivalently, measuring all the Hawking radiation should keep the firewall intact.

In the argument of section 2, it was assumed that if typical AdS black hole microstates have a common region of smooth spacetime behind the horizon, there should be some CFT operator that probes the physics in this region. Another possibility is that there is some common smooth spacetime behind the horizon of typical black hole microstate, but there is no single operator in the CFT that would tell us about it. In other words, while we can use a standard set of operators to probe the physics outside the horizon, the way to extract the physics behind the horizon is state-dependent. To avoid the firewall conclusion it must be that the state-dependent procedure applied to the individual microstates must give a different answer when applied to the mixed state describing the thermal ensemble of these microstates, or not be applicable at all. Otherwise, we would again have the conclusion that the thermal state implies some smooth spacetime behind the horizon, which we have argued cannot be correct. To summarize, the arguments of section 2 imply that either I) large AdS black holes dual to pure CFT states have firewalls OR II) reconstructing the physics behind the horizon of such a black hole requires a state-dependent construction.

---

[12]Similar comments were made in [54] and in the talk by Joe Polchinski at Strings 2013.

[13]Presumably, the postulate in question is the validity of effective field theory for describing local physics outside the horizon. By assuming the validity of AdS/CFT (and further assuming that a pure state of some holographic degrees of freedom contains complete information about the corresponding dual spacetime physics) we are assuming the unitarity postulate.

Acknowledgments

References

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

[2] S.D. Mathur, *The information paradox: a pedagogical introduction*, *Class. Quant. Grav.* **26** (2009) 224001 [arXiv:0909.1038] [INSPIRE].

[3] S.L. Braunstein, S. Pirandola and K. Życzkowski, *Better late than never: information retrieval from black holes*, *Phys. Rev. Lett.* **110** (2013) 101301 [arXiv:0907.1190] [INSPIRE].

[4] S.B. Giddings, *Models for unitary black hole disintegration*, *Phys. Rev.* **D 85** (2012) 044038 [arXiv:1108.2015] [INSPIRE].

[5] S.B. Giddings, *Black holes, quantum information and unitary evolution*, *Phys. Rev.* **D 85** (2012) 124063 [arXiv:1201.1037] [INSPIRE].

[6] R. Bousso, *Complementarity is not enough*, *Phys. Rev.* **D 87** (2013) 124023 [arXiv:1207.5192] [INSPIRE].

[7] Y. Nomura, J. Varela and S.J. Weinberg, *Complementarity endures: no firewall for an infalling observer*, *JHEP* **03** (2013) 059 [arXiv:1207.6626] [INSPIRE].

[8] S.D. Mathur and D. Turton, *Comments on black holes I: the possibility of complementarity*, *JHEP* **01** (2014) 034 [arXiv:1208.2005] [INSPIRE].

[9] B.D. Chowdhury and A. Puhm, *Is Alice burning or fuzzing?*, *Phys. Rev.* **D 88** (2013) 063509 [arXiv:1208.2026] [INSPIRE].

[10] L. Susskind, *Singularities, firewalls and complementarity*, arXiv:1208.3445 [INSPIRE].

[11] T. Banks and W. Fischler, *Holographic space-time does not predict firewalls*, arXiv:1208.4757 [INSPIRE].

[12] A. Ori, *Firewall or smooth horizon?*, arXiv:1208.6480 [INSPIRE].

[13] I. Bena, A. Puhm and B. Vercnocke, *Non-extremal black hole microstates: fuzzballs of fire or fuzzballs of fuzz?*, *JHEP* **12** (2012) 014 [arXiv:1208.3468] [INSPIRE].

[14] A. Giveon and N. Itzhaki, *String theory versus black hole complementarity*, *JHEP* **12** (2012) 094 [arXiv:1208.3930] [INSPIRE].

[15] R. Brustein, *Origin of the blackhole information paradox*, *Fortschr. Phys.* **62** (2014) 255 [arXiv:1209.2686] [INSPIRE].

[16] L. Susskind, *The transfer of entanglement: the case for firewalls*, arXiv:1210.2098 [INSPIRE].

[17] S. Hossenfelder, *Comment on the black hole firewall*, arXiv:1210.5317 [INSPIRE].

[18] S.G. Avery, B.D. Chowdhury and A. Puhm, *Unitarity and fuzzball complementarity: "Alice fuzzes but may not even know it!"*, *JHEP* **09** (2013) 012 [arXiv:1210.6996] [INSPIRE].

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

[20] D.-i. Hwang, B.-H. Lee and D.-h. Yeom, *Is the firewall consistent? Gedanken experiments on black hole complementarity and firewall proposal*, *JCAP* **01** (2013) 005 [arXiv:1210.6733] [INSPIRE].

[21] S.K. Rama, *Remarks on black hole evolution a la firewalls and fuzzballs*, arXiv:1211.5645 [INSPIRE].

[22] Y. Nomura and J. Varela, *A note on (no) firewalls: the entropy argument*, *JHEP* **07** (2013) 124 [arXiv:1211.7033] [INSPIRE].

[23] S.B. Giddings, *Nonviolent nonlocality*, *Phys. Rev.* **D 88** (2013) 064023 [arXiv:1211.7070] [INSPIRE].

[24] K. Larjo, D.A. Lowe and L. Thorlacius, *Black holes without firewalls*, *Phys. Rev.* **D 87** (2013) 104018 [arXiv:1211.4620] [INSPIRE].

[25] M. Saravani, N. Afshordi and R.B. Mann, *Empty black holes, firewalls and the origin of Bekenstein-Hawking entropy*, arXiv:1212.4176 [INSPIRE].

[26] T. Jacobson, *Boundary unitarity and the black hole information paradox*, *Int. J. Mod. Phys.* **D 22** (2013) 1342002 [arXiv:1212.6944] [INSPIRE].

[27] L. Susskind, *Black hole complementarity and the Harlow-Hayden conjecture*, arXiv:1301.4505 [INSPIRE].

[28] D. Harlow and P. Hayden, *Quantum computation vs. firewalls*, *JHEP* **06** (2013) 085 [arXiv:1301.4504] [INSPIRE].

[29] W. Kim, B.-H. Lee and D.-H. Yeom, *Black hole complementarity and firewall in two dimensions*, *JHEP* **05** (2013) 060 [arXiv:1301.5138] [INSPIRE].

[30] S.D.H. Hsu, *Macroscopic superpositions and black hole unitarity*, arXiv:1302.0451 [INSPIRE].

[31] S.B. Giddings, *Nonviolent information transfer from black holes: a field theory parametrization*, *Phys. Rev.* **D 88** (2013) 024018 [arXiv:1302.2613] [INSPIRE].

[32] H. Kawai, Y. Matsuo and Y. Yokokura, *A self-consistent model of the black hole evaporation*, *Int. J. Mod. Phys.* **A 28** (2013) 1350050 [arXiv:1302.4733] [INSPIRE].

[33] B.-H. Lee and D.-h. Yeom, *Status report: black hole complementarity controversy*, *Nucl. Phys. Proc. Suppl.* **246-247** (2014) 178 [arXiv:1302.6006] [INSPIRE].

[34] T. Banks and W. Fischler, *No firewalls in holographic space-time or matrix theory*, arXiv:1305.3923 [INSPIRE].

[35] D.A. Lowe and L. Thorlacius, *Pure states and black hole complementarity*, *Phys. Rev.* **D 88** (2013) 044012 [arXiv:1305.7459] [INSPIRE].

[36] D.N. Page, *Excluding black hole firewalls with extreme cosmic censorship*, *JCAP* **06** (2014) 051 [arXiv:1306.0562] [INSPIRE].

[37] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford and J. Sully, *An apologia for firewalls*, *JHEP* **09** (2013) 018 [arXiv:1304.6483] [INSPIRE].

[38] S.G. Avery and B.D. Chowdhury, *Firewalls in AdS/CFT*, arXiv:1302.5428 [INSPIRE].

[39] B.D. Chowdhury, *Black holes versus firewalls and thermo-field dynamics*, *Int. J. Mod. Phys.* **D 22** (2013) 1342011 [arXiv:1305.6343] [INSPIRE].

[40] K. Papadodimas and S. Raju, *An infalling observer in AdS/CFT*, *JHEP* **10** (2013) 212 [arXiv:1211.6767] [INSPIRE].

[41] E. Verlinde and H. Verlinde, *Passing through the firewall*, arXiv:1306.0515 [INSPIRE].

[42] E. Verlinde and H. Verlinde, *Black hole information as topological qubits*, arXiv:1306.0516 [INSPIRE].

[43] B. Czech, J.L. Karczmarek, F. Nogueira and M. Van Raamsdonk, *Rindler quantum gravity*, *Class. Quant. Grav.* **29** (2012) 235025 [arXiv:1206.1323] [INSPIRE].

[44] W. Israel, *Thermo field dynamics of black holes*, *Phys. Lett.* **A 57** (1976) 107 [INSPIRE].

[45] J.M. Maldacena, *Eternal black holes in anti-de Sitter*, *JHEP* **04** (2003) 021 [hep-th/0106112] [INSPIRE].

[46] B. Freivogel et al., *Inflation in AdS/CFT*, *JHEP* **03** (2006) 007 [hep-th/0510046] [INSPIRE].

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

[48] M. Van Raamsdonk, *Comments on quantum gravity and entanglement*, arXiv:0907.2939 [INSPIRE].

[49] M. Van Raamsdonk, *Building up spacetime with quantum entanglement*, *Gen. Rel. Grav.* **42** (2010) 2323 [arXiv:1005.3035] [INSPIRE].

[50] S.D. Mathur, *The information paradox and the infall problem*, *Class. Quant. Grav.* **28** (2011) 125010 [arXiv:1012.2101] [INSPIRE].

[51] S.D. Mathur and C.J. Plumberg, *Correlations in Hawking radiation and the infall problem*, *JHEP* **09** (2011) 093 [arXiv:1101.4899] [INSPIRE].

[52] H. Casini, M. Huerta and R.C. Myers, *Towards a derivation of holographic entanglement entropy*, *JHEP* **05** (2011) 036 [arXiv:1102.0440] [INSPIRE].

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

[54] J. Maldacena and L. Susskind, *Cool horizons for entangled black holes*, *Fortschr. Phys.* **61** (2013) 781 [arXiv:1306.0533] [INSPIRE].

[55] V. Balasubramanian et al., *Typicality versus thermality: an analytic distinction*, *Gen. Rel. Grav.* **40** (2008) 1863 [hep-th/0701122] [INSPIRE].

[56] D.N. Page, *Average entropy of a subsystem*, *Phys. Rev. Lett.* **71** (1993) 1291 [gr-qc/9305007] [INSPIRE].

[57] S. Popescu, A.J. Short and A. Winter, *Entanglement and the foundations of statistical mechanics*, *Nat. Phys.* **2** (2006) 754.

[58] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [arXiv:1306.0622] [INSPIRE].

[59] D. Marolf and A.C. Wall, *Eternal black holes and superselection in AdS/CFT*, *Class. Quant. Grav.* **30** (2013) 025001 [arXiv:1210.3590] [INSPIRE].