Complementarity vs Firewalls: Are there surprising quantum effects near black holes?

Donald Marolf, UCSB

10/16/2012

With A. Almheiri, J. Polchinski, and J. Sully

See arxiv: **1207.3132**

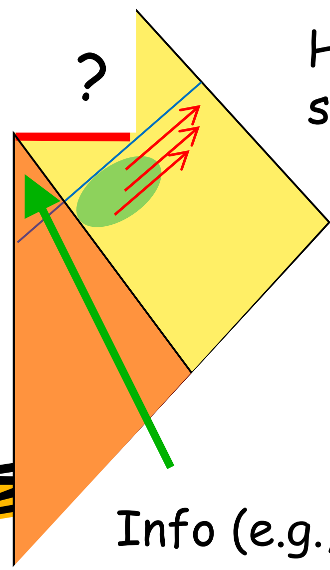
Overview

Hawking (1976): Semiclassical physics (aka QFT in curved space, aka low energy effective field theory) predicts

a) The evaporation of black holes

b) That the resulting radiation is very nearly thermal and contains very little information about the initial black hole.

Robust against small corrections!



Hawking radiation forms outside the black hole, spacelike separated from the infalling info.

Copying the infalling information to the Hawking radiation would violate the no Quantum Xerox thm (which follows from linearity).

Info (e.g., set of spins) travels deep inside the black hole.

A puzzle:

This suggests an effective loss of unitarity, at least on the timescale set by evaporation.

Historically Debated Issues:

i) Banks, Peskin, and Susskind: In QM $U(t) = e^{-iHt}$. Can one violate effective unitarity without violating (effective) energy conservation?

ii) Giddings (early 90's): There are many ways to make a large black hole. If each leads (via Hawking decay) to a distinct internal state of a smaller black hole, then each black hole must have infinitely many states. Shouldn't they be infinitely pair produced?

Maldacena's AdS/CFT (1997+):

Evaporation is unitary on evaporative timescales!

Does this imply large quantum gravity effects near black hole horizons?

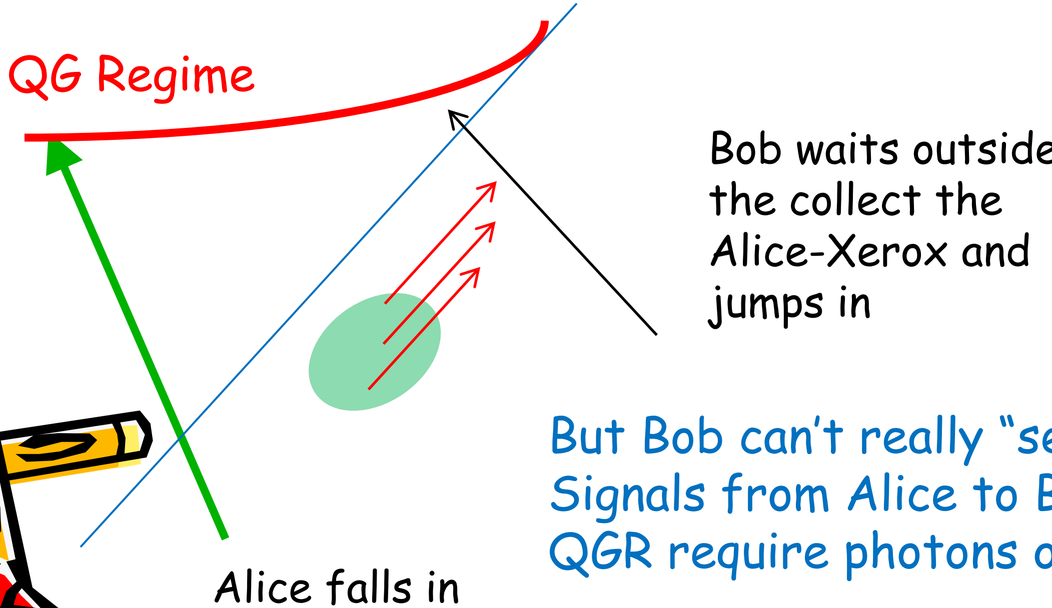
Note: Naively, QG effects are of order $L_{planck}/R_s \sim 10^{-38}$ for solar mass black holes.

Susskind, Thorlacius, & Uglam (1993):
No, due Black Hole Complementarity

The idea: Perhaps the d.o.f.'s inside the BH are somehow identified with those outside, avoiding a violation of the no Quantum Xerox theorem.

I.e., there is some duality. This duality is commonly assumed to be AdS/CFT-like and, in fact embodied by AdS/CFT itself.

The principle of Black Hole Complementarity (BHC): If no observer can detect the supposed Xeroxing, it need not be a real effect (c.f., the Heisenberg microscope).



But Bob can't really "see" Alice inside the BH. Signals from Alice to Bob before reaching the QGR require photons of energy >> M_{BH}!

Postulates of BHC

Postulate 1: Distant observers describe BH formation and evaporation as a standard quantum process. Infalling matter is related to outgoing Hawking radiation via a unitary S-matrix.

Postulate 2: Outside the stretched horizon of a massive black hole (i.e., more than L_{planck} from the horizon), physics can be described to good approximation by a [standard] set of semi-classical field equations.

Postulate 3: To a distant observer, a black hole appears to be a quantum system with discrete energy levels. The dimension of the subspace of states describing a black hole of mass M is the exponential of the Bekenstein entropy S(M).

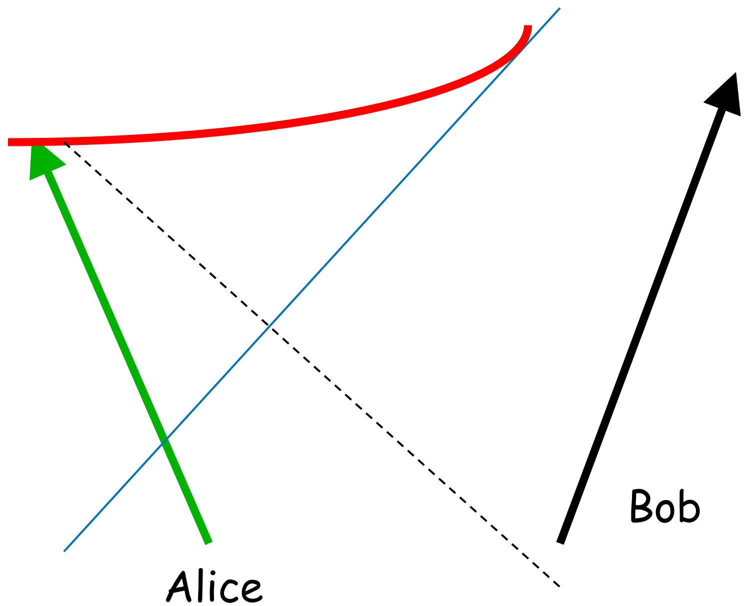
Postulate 4: Freely falling observers experiences nothing out of the ordinary when crossing the horizon. Their experiments are well-described by familiar effective field theory and their probability to encounter a quantum with energy $E \gg 1/R_s$ is very small.

Note: Believed by many who would not call it "complementarity."

But these postulates are mutually inconsistent!

Comparison Principle

Detailed interpretations of the above postulates (and the willingness to add amendments) differ from reader to reader, but…



In their common causal past, the Alice-and the Bob-Theories can disagree only by predictions that are parametrically hard (in ℓ_{Planck}) for Alice to test at times parametrically long (in Planck units) before Alice leaves this region.

Outline

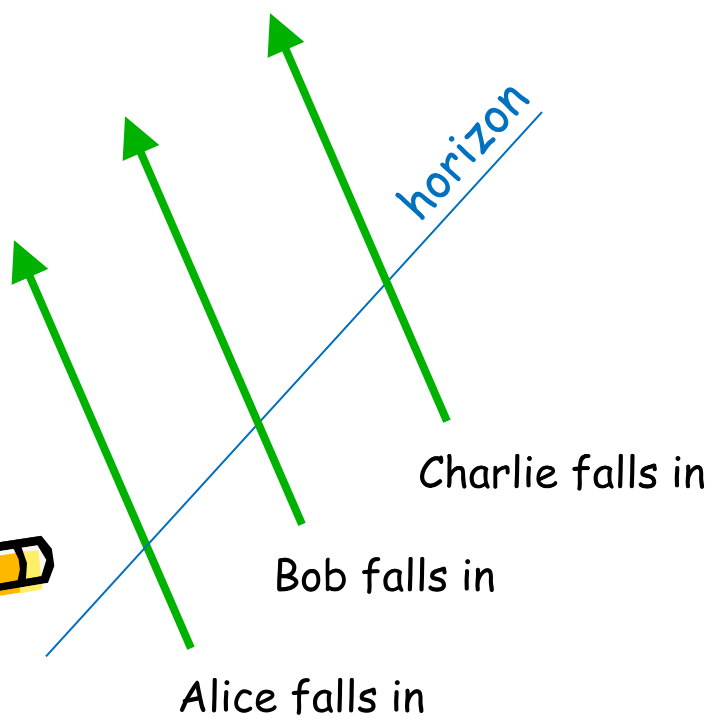
I. Introduction (Done!)
II. Review of the Hawking Effect
III. Our "paradox"
IV. Mining Black Holes
V. Firewalls are natural
VI. Timescales
VII. Common Questions
VIII. Summary

II. Review of the Hawking Effect

1a) The horizon is generated by unstable null geodesics, finely balanced between falling into the BH and escaping to infinity.

1b) Nearby geodesics (and wavepackets) diverge exponentially toward the future and converge exponentially toward the past.

1c) Alice sees any mode as exponentially blue-shifted relative to Charlie.

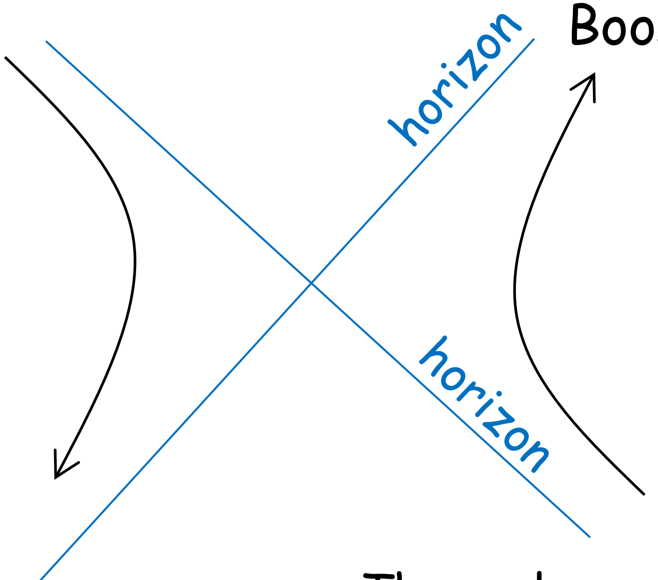


2) Suppose that Charlie studies an outward-moving mode X with
$L_{planck} \ll \lambda \ll R_s$
long after the BH forms.

Unless the BH formation involved particles with exponentially super-Planckian energies, Charlie should describe X as being in its vacuum state.

II. Review of the Hawking Effect

3) Since the BH spacetime is flat on the scale λ, this reduces the problem to the Unruh effect.



Boost Symmetry ξ

Let $|E\rangle_R$, $|E\rangle_L$ be Rinder states in the right and left wedge with energy E wrt ξ.

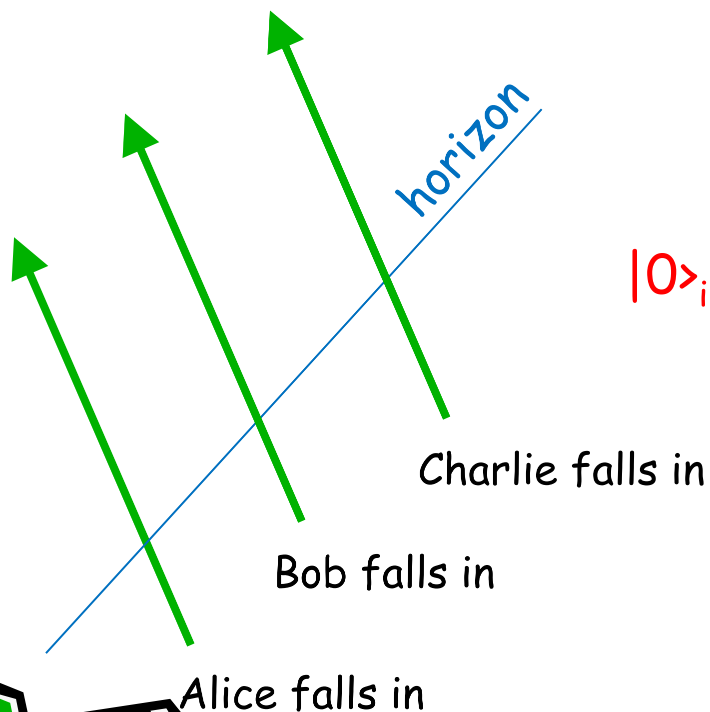
Note: ξ is the only time-translation-like symmetry in the actual BH problem.

$$|0\rangle_{Mink} = \int dE\, e^{-E/T_H}\, |E\rangle_R\, |E\rangle_L$$

Thermal correlations between the left and right wedges, and a thermal spectrum of particles with respect to ξ.

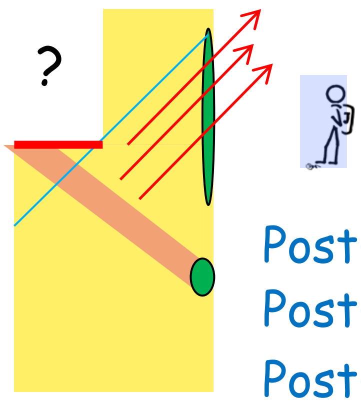
Outward-moving modes propagate to infinity, far from the BH, where spacetime is flat and E becomes the usual energy. So they represent a thermal flux of energy to infinity at $T_H = hc/4\pi R_s k_b$.

Key point for us: If infalling observers see vacuum, the state outside the horizon (right Rindler wedge) must be purified through its entanglement with states inside the horizon (left Rindler wedge).



$$|0\rangle_{infalling} = \int dE\, e^{-E/T_H}\, |E\rangle_R\, |E\rangle_L$$

III. Our "Paradox"



I will give the cleanest (AdS/CFT) argument: No constraints of space, time or gravity outside AdS.

BHC

Postulate 1: Unitarity for Distant observers.

Postulate 2: QFT valid outside the "stretched horizon."

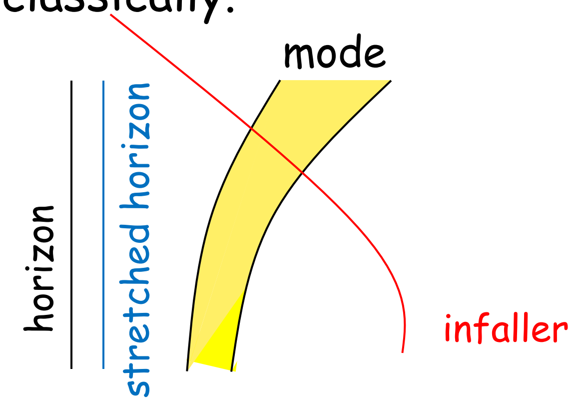
Postulate 3: To a distant observer, # of states = exp[S_{BH}].

Postulate 4: "Lack of drama" for freely falling observers at horizon.

We'll use wavepackets of compact support, already well separated from the stretched horizon when the infaller decies to jump in. This allows the infaller to sample the entire mode and still have time to send signals to the distant observer. [Comparison principle]

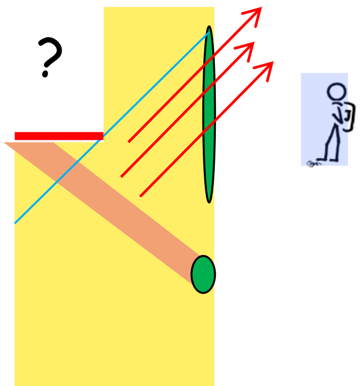
We assume that the stretched horizon acts causally, and so knows nothing about the observer's decision. Thus by #2 any back-reaction effects caused by the observer can be computed semi-classically.

Eddington-Finklestein Diagram:



III. Our "Paradox"

(Abstract version, builds on Mathur et al.)



Form the BH from a pure state.

$S_{HR}(t)$:= Entropy of Hawking radiation *outside* AdS at time t.

By #2: $S_{HR}(t) \Longrightarrow 0$ as $v \Longrightarrow \infty$.

So S_{HR} is non-increasing for at least one t_0 .

Let A = HR before t_0, B = HR near t_0, and C be Hawking partner modes behind horizon (as described by infaller for which $\ell_{Planck} \ll \lambda \ll Rs$).

Note that $S_A \geq S_{AB}$. (Both observers).

Subadditivity* (infaller): $S_{ABC} \geq S_A - S_{BC}$

Strong subadditivity*: $S_{AB} + S_{BC} \geq S_B + S_{ABC}$

$\cancel{S_A} + S_{BC} \geq S_B + \cancel{S_A} - S_{BC}$

But S_{BC}/S_B is small by #4!

BHC

1: Unitarity for Distant observers.

2: QFT valid outside the "stretched horizon."

3: To a distant observer, # of states = $\exp[S_{BH}]$.

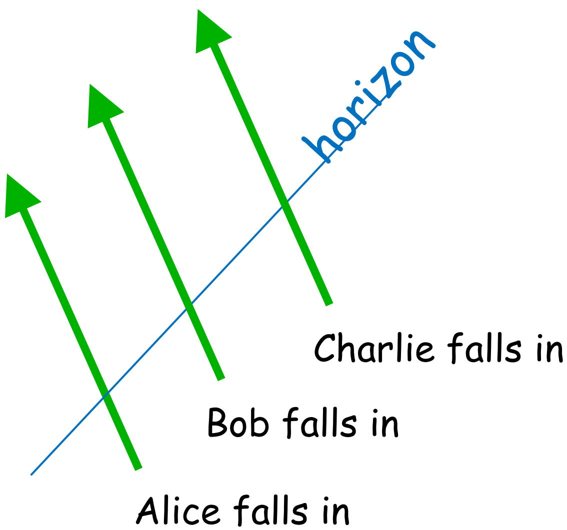
4: "Lack of drama" for freely falling observers at horizon.

* Since A,B,C are distinct.

A firewall?

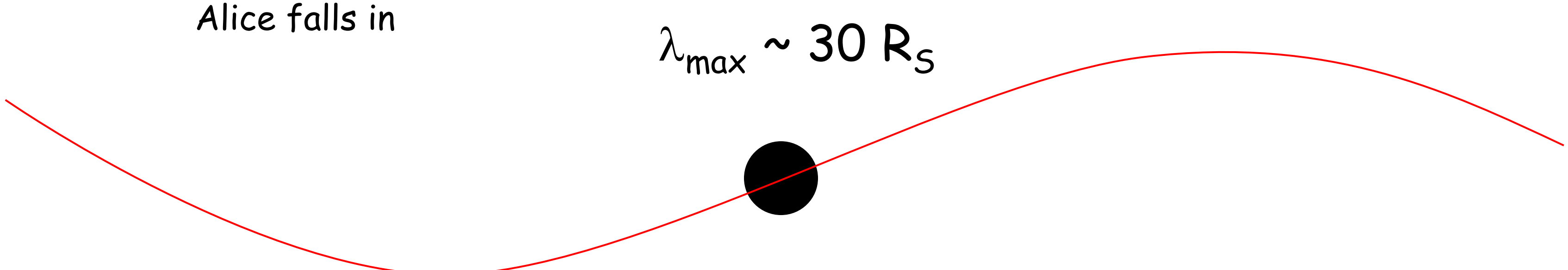
Presenting the argument as above suggests that the infaller encounters extra particles *at** the horizon. Firewall.

(Aside: might be avoided through other failures of #2, #4.)



Due to the usual blueshift, some of these will be Planck scale in (say) Alice's frame.

But so far we discuss only modes that would propagate to infinity.

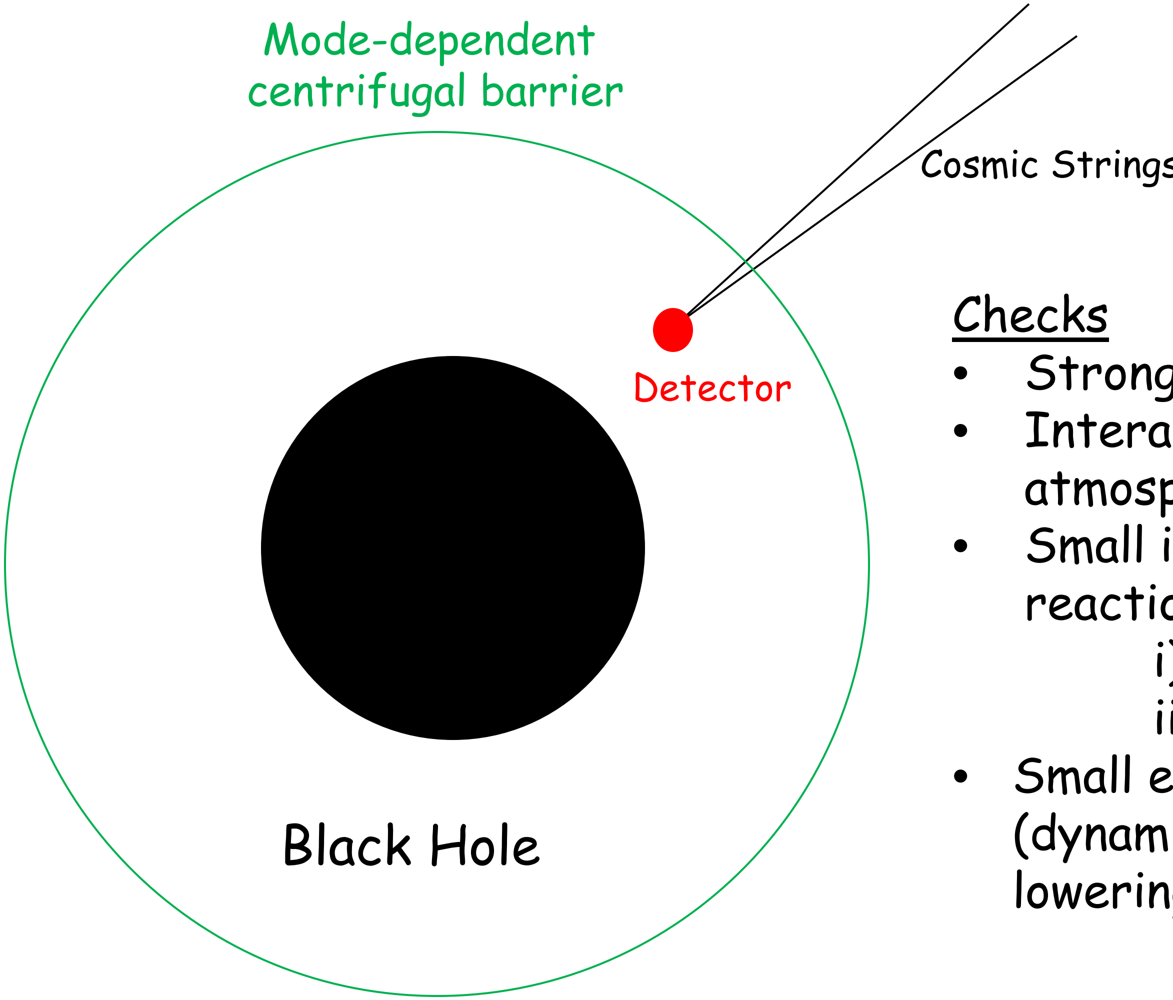


Just s-waves due to grey body factors!

\# of sub-Planckian particles encountered by observer of finite size ~ $\ln(R_S)/R_S^2$

*Could be completely invisible outside.

IV. Black Hole Mining



Checks

- Strong enough string
- Interacts weakly with thermal atmosphere
- Small instantaneous back reaction
  - i) From string.
  - ii) From detector (tides)
- Small entropy production (dynamical tides from lowering/raising)

We find no problem mining typical modes down to ~ ℓ_{planck} from the horizon.

Suggests $T_{firewall} = T_{Planck}$

VI. When does it set in?

Page showed that entropy could be monotonically increasing only up ot the "Page time" ~ R_S^3/ℓ_{Planck}^2 for 3+1 Asymptically Flat case.)

He also argued that, after this time, the black hole becomes maximally entangled with the emitted Hawking radiation. (And our argument applies immediately whenever this is true.)

If property is intrinsic to the BH, true for maximally mixed ρ_{BH}, and thus for any "typical" state.

Much like phenomenon of "thermalization."

$$t_{\text{Thermalization of a BH}} \sim R_S \ln(R_S/\ell_{planck})$$

Hayden & Preskill, Susskind: BHs may also "scramble" on this timescale;
i.e., timescale to evolve to a generic state
as sampled by "typical" operators.

Suggests that firewall may turn on as early as
~ $R_S \ln (R_S/\ell_{planck})$ after formation of BH.

We view this as an open question; see also Susskind.

VII. Common Questions

Do firewalls violate the correspondence principle as ℓ_{Planck} -> 0?
No. Appear only after time parametrically large in ℓ_{Planck}.

Is the entropy "observable?"
Once the info has left the AdS space, there are no constraints of time, space, or gravity.

Can one hide the effect to make it parametrically hard to observe?
We don't see how. Seems unlikely.

Might the info be imprinted on the Hawking radiation only at some macroscopic distance from the horizon?
But energy can be mined at microscopic distances and can be manipulated on the way out to make an arbitrary state of the mining equipment. We argue that no unitary process can do the job.

Are there non-firewall alternatives?
Sure, but should prevent black hole mining. We *conjecture* that they are equally "deadly."

More Questions

What about black holes with 2 asymptotic regions?
E.g., maximal analytic continuation of AdS/Schwarzschild?
$|\psi\rangle = \Sigma\ |E\rangle|E\rangle$ Not typical state! [DM & A. Wall]

What about Rindler or cosmological horizons?
Again atypical states. (Also Rindler scrambling time is infinite.)

Can Dualities save the horizon?
We don't see how. Just run our argument in whatever effective theory actually describes the observers experiences.

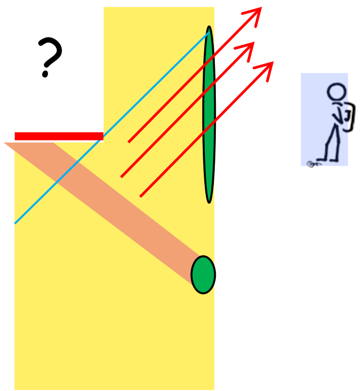
Are Firewalls Fuzzballs?
[Mathur & Turton; Avery, Chowdhuri, & Puhm, etc.]
Maybe, though advertised version appears instead to drop postulate 2; i.e., novel physics can be seen even far from the horizon (at least over time scales $\sim R_s$). Visible to LIGO? Can infallers survive?

Do firewalls avoid xeroxing without Complementarity?
Conjecture: there are no low-curvature Cauchy surfaces that register xeroxing but avoid the firewall.

VIII Summary

- The axioms of BHC are not self-consistent.
- Suggests a firewall at the horizon, but other failures of local QFT may be equally plausible/implausible.
  Are these equally deadly to humans & human-built devices?
- Friendly amendments to BHC do not seem to suffice.
- Open questions, but answers seem to require input from a microscopic theory. [e.g. fuzzball program]
  Toy models may also be useful.

III. "Physical version"
(Builds on Hayden-Preskill)



Form the BH from a pure state.

Full state $|\Psi\rangle$ of all outgoing radiation must then be pure.

Divide into early- and late-time radiation (E & L), with E having $(\frac{1}{2} + \varepsilon)$ of the coarse-grained entropy.

⟹ dim(E) >> dim(L) for a large BH.

Model $|\Psi\rangle$ with a random state in $E \otimes L$ (Page, HP).

Then for any basis $|L_i\rangle$ of L, can show for large BHs that

$$|\Psi\rangle \approx \Sigma \; |E_i\rangle \, |L_i\rangle$$

with $|E_i\rangle$ orthonormal.

⟹ Measurements* of E can project onto any desired pure state of L, as described by both observers!

Removes any entanglement with Hawking partners. (infaller)

* The term "measurement" is just illustrative. State is already decohered.