Gauge/Gravity Duality and the Black Hole Interior

Joseph Polchinski
Kavli Institute for Theoretical Physics
University of California at Santa Barbara

KIAS-YITP joint workshop
String Theory, Black Holes and Holography
Kyoto, 7/1/13

Note title change

• Same subject but change of emphasis. Not just

   "Is there a firewall?"

but also

   "How complete is our current understanding of quantum gravity via AdS/CFT?"

• Expanded version of Strings 2013 talk

Gauge/gravity duality (BFSS, AdS/CFT, etc.) gives a nonperturbative construction of quantum gravity in spacetimes with special boundary conditions. But how complete is it?

We have complete dictionary for boundary observables (Gubser, Klebanov, Polyakov; Witten). The boundary observer can carry out many kinds of experiment. E.g. they can create a black hole and watch it evaporate, and this must be unitary because the gauge theory is.

However, the dictionary becomes less sharp as we go into the bulk.   Perhaps this the best we can do in quantum gravity, where it is difficult to define precise observables?

The firewall provides a sharp test of the completeness of the theory.

AMPS(S) propose that the black hole interior is in a highly excited state. Can we test this using gauge/gravity duality?

Outline

1. A new version of the AMPS(S) argument, for unentangled AdS black holes, and its relation to the old one.

2. Limits on gauge/gravity duality?

3. EPR = ER for general entanglement?

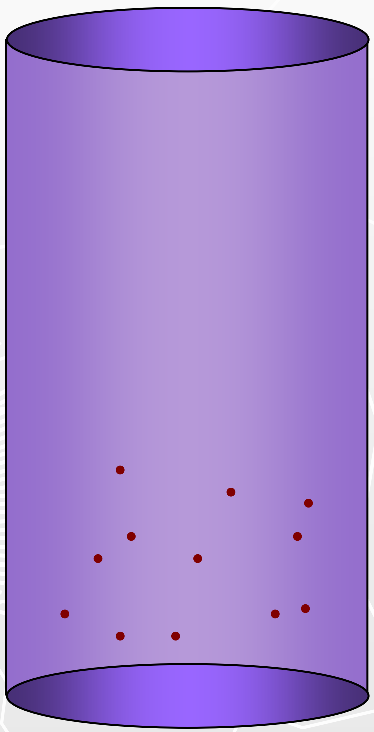
4. If there is a firewall, does the Hawking calculation give the right flux, and why?

1207.3123, 1304.6483 & in preparation.

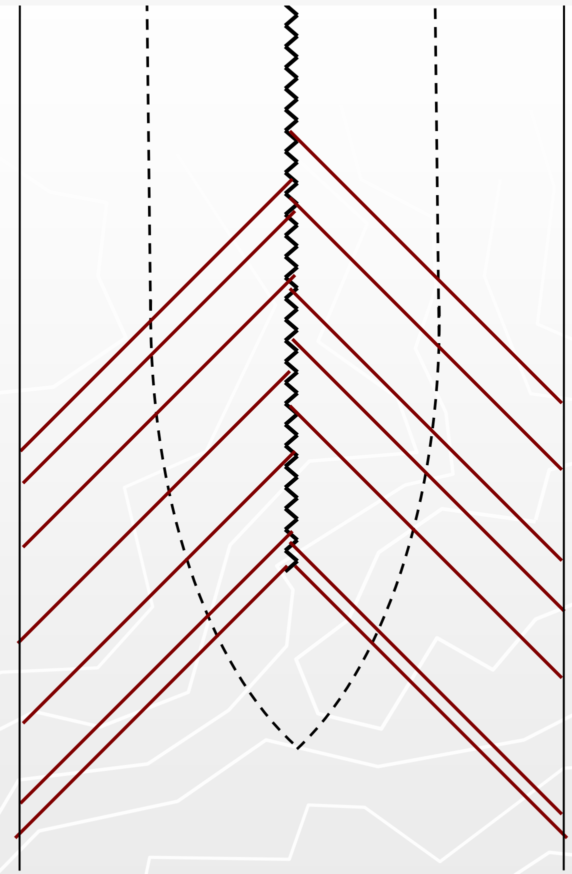
I. An argument for firewalls in typical black holes

Consider $\mathcal{H}_{CFT}$, all CFT states that can be created by products of local operators:

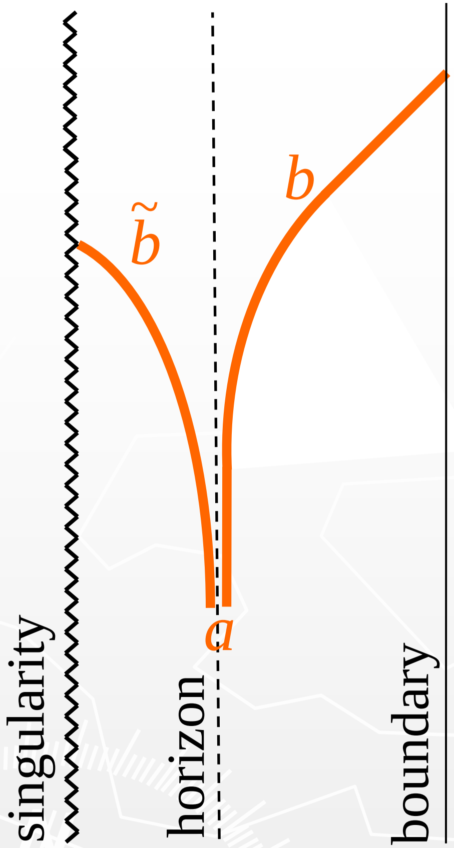
CFT

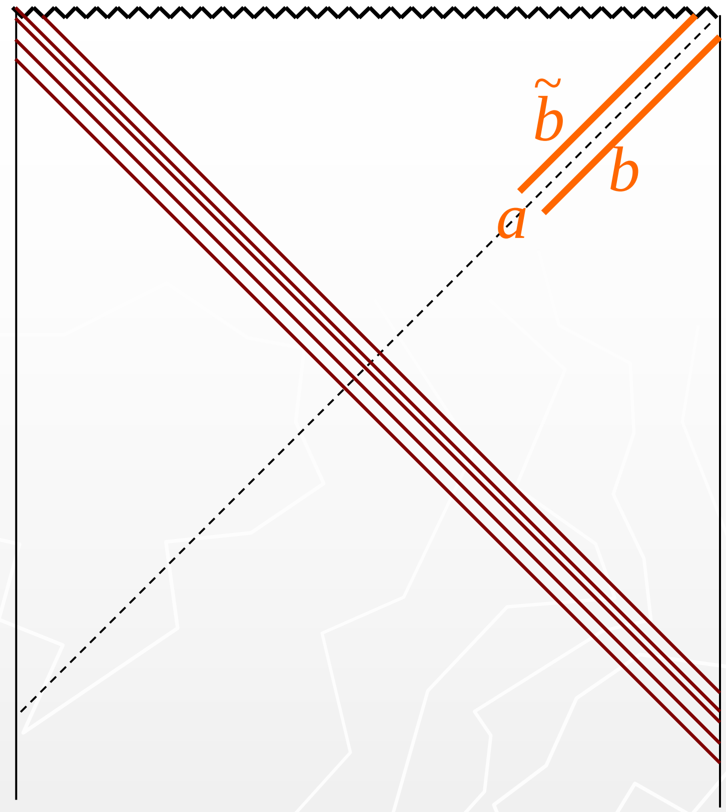


bulk



At high enough energy, these are large (stable) black holes in the bulk. Do they have firewalls?



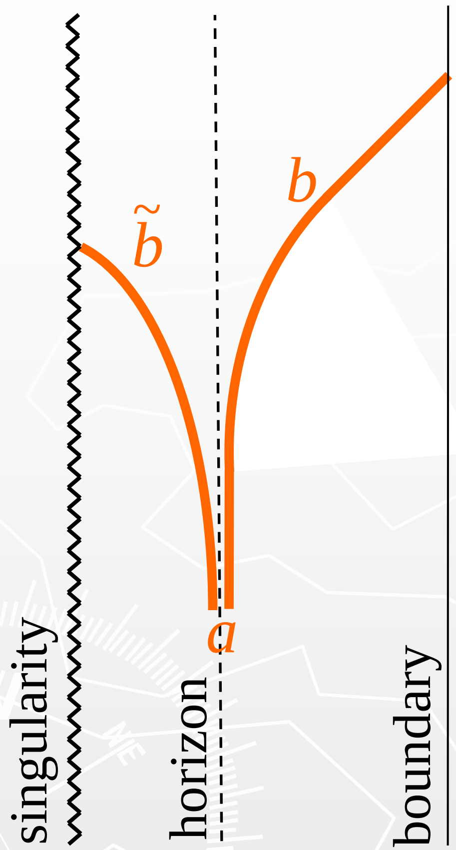


b: outgoing Hawking mode*, centered on CFT frequency ω

$\tilde{b}$: interior partner mode (~ = inside)

a: smooth mode across horizon, used by infalling observer

$$b = Ba + Ca^\dagger \qquad a = B'\, b + C'\, b^\dagger + B''\, \tilde{b} + C''\, b^\dagger$$



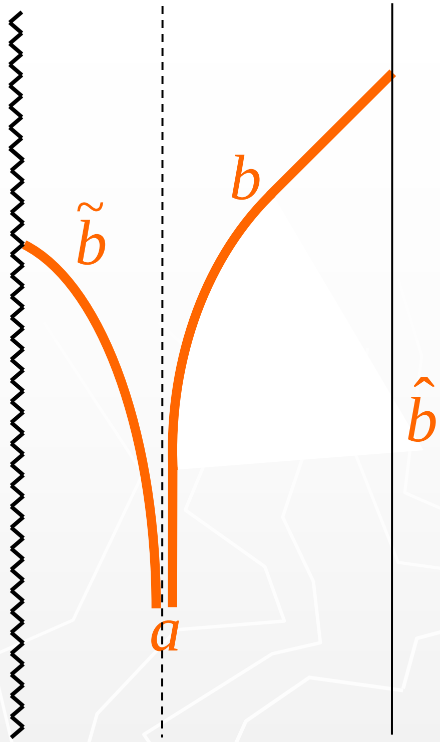
Short derivation of Hawking radiation, in effective field theory:

- *a* has high frequency (>>*T*) as seen by infalling observer, so it must be in its ground state by the adiabatic principle

- $b = Ba + Ca^\dagger$

- *b* is not in its ground state: in the *a*-vacuum *b* is in a thermal state

Bulk dictionary:



b has an image $\hat{b}$ in the CFT, obtained at N^0 by relating mode expansion for bulk field to CFT operator via usual dictionary
(Banks, Douglas, Horowitz, Martinec 1998; Balasubramanian, Kraus, Lawrence, Trivedi 1998; Bena 1999).

Expandable in $1/N$ (Kabat, Lifschytz, Lowe 2011):

$$\hat{b} = \text{single-trace} + \text{double-trace} + \ldots ,$$

where the operators on the RHS are integrated against 'smearing functions' on the boundary

Consider a basis for $\mathsf{H}_{\mathrm{CFT}}$ in which $\hat{N}_b = \hat{b}^\dagger \hat{b}$ is diagonal. Since $\hat{N}_b$ is thermal in the *a* (infalling) vacuum, these basis states are far from the *a*-vacuum

$$\langle \psi | \hat{N}_a | \psi \rangle \geq O(1)$$

in any $\hat{N}_b$ eigenstate. Taking the average over $\mathsf{H}_{\mathrm{CFT}}$,

$$\mathrm{Tr}(\hat{N}_a) / \mathrm{Tr}(1) \geq O(1).$$

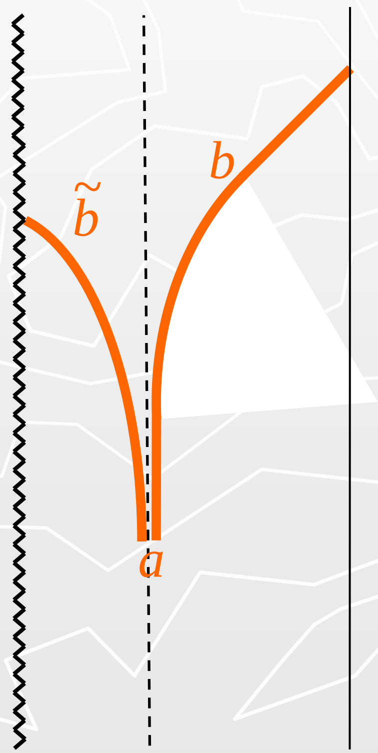
Thus, each infalling mode is excited with probability $O(1)$: typical black hole states have firewalls.

Compare original AMPS argument, for an evaporating black hole after the Page time (~ half-lifetime):

- b is entangled with the early Hawking radiation (purity/non-loss of information)

- Therefore b cannot be in a pure state with the internal mode $\tilde{b}$, so a is excited = firewall. (Hawking, in reverse!)

For evaporating black hole, the firewall takes time to form (at least the scrambling time $R \ln R/l_P$, at most the Page time R^3/l_P^2): this is the time it takes to become 'typical.'



Original AMPS argument vs. new argument

Common assumptions:

- QM for asymptotic observer
- Effective field theory governs evolution of *b* outside the horizon

Difference:

- In the new argument, black holes are not entangled with any external system.

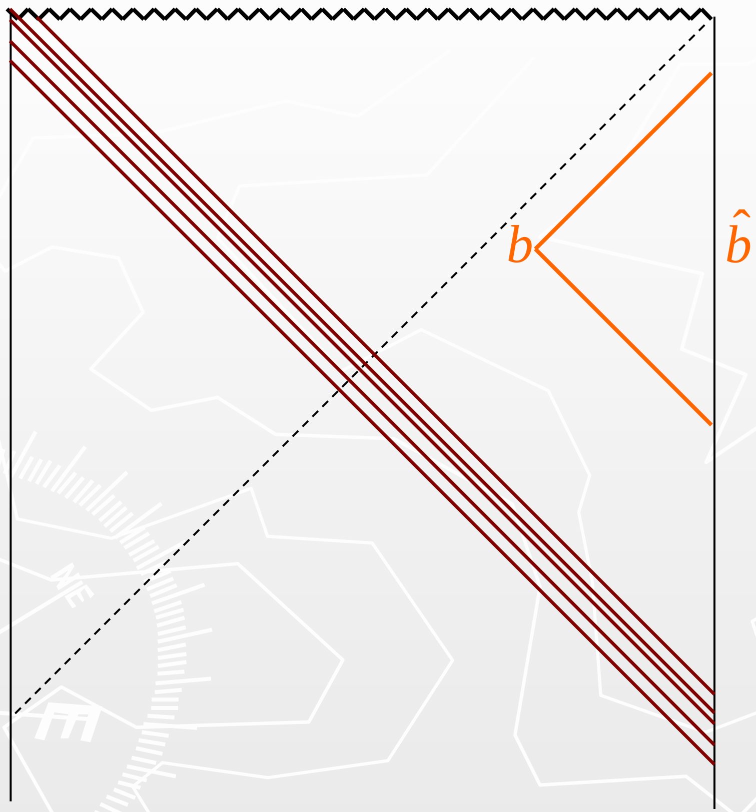
II. Limits on seeing the interior

If gauge/gravity duality were as complete as we might hope, we could test this reasoning by identifying the CFT operator $\hat{T}_{\mu\nu}(x)$ dual to the matter energy-momentum tensor $T_{\mu\nu}(x)$ at some point in the black hole interior, and calculating its expectation in these CFT states.

Obvious problem: what is the dictionary?

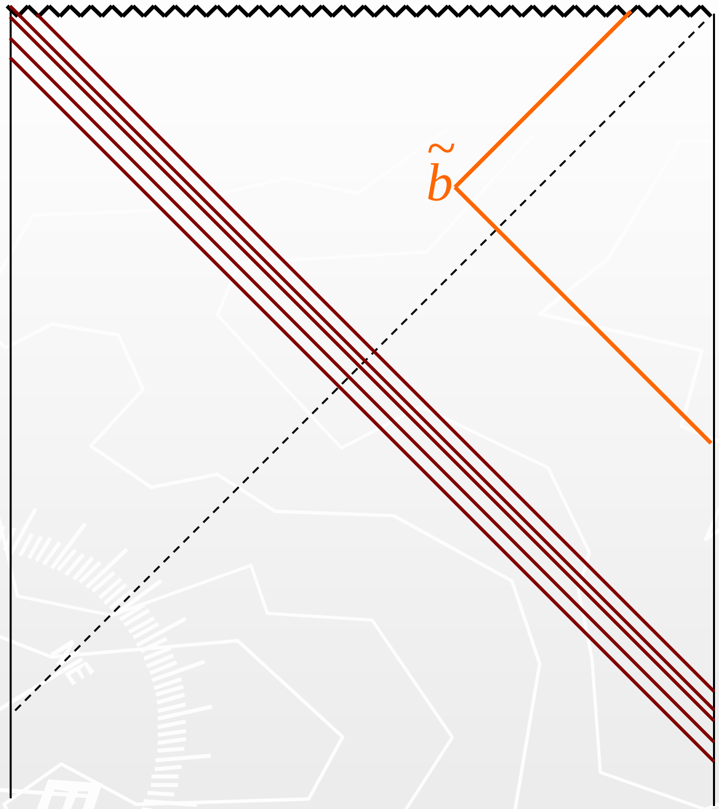
Less obvious problem: there may be no such dictionary!

The dictionary $b \to \hat{b}$ is essentially obtained by integrating in a spacelike direction to the boundary:



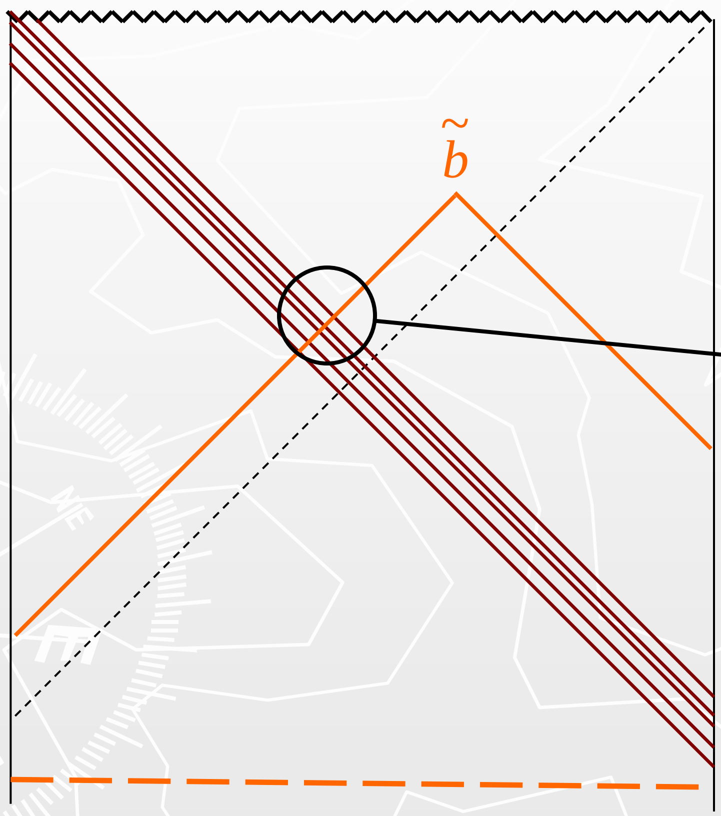
This is overdetermined, but OK because boundary data is constrained by AdS/CFT.

For points behind the horizon this runs into the singularity.



Alternately, integrate back in time to before the black hole formed, then outward to the boundary (Freivogel, Susskind 2004; Heemskerk, Marolf, Polchinski, Sully 2012).

Alternately, integrate back in time to before the black hole formed, then outward to the boundary (Freivogel, Susskind 2004; Heemskerk, Marolf, Polchinski, Sully 2012).



Problem: exponential blue shift leads to trans-Planckian collision, presumable singularity, for $\tilde{b}$ after the scrambling time $R \ln R$.

A more basic problem: there can be *no* operator with the properties

$$[\hat{\tilde{b}}^\dagger, \hat{\tilde{b}}] = -1, \quad [H, \hat{\tilde{b}}^\dagger] = -\omega\, \hat{\tilde{b}}^\dagger$$

(tilde = behind horizon, hat = CFT image).
That is, this *lowers* the energy, where the original Hawking mode is narrowly centered at frequency ω.
Proof: consider all states $|i\rangle$ such that

$$M < E < M+\delta\,.$$

Then for $\hat{\tilde{b}}^\dagger|i\rangle$,

$$M-\omega < E < M-\omega+\delta\,.$$

The number of such states is smaller by a factor $e^{-\beta\omega} = e^{-O(1)}$.
So $\hat{\tilde{b}}^\dagger$ has a kernel. But it cannot, it is a raising operator.

Four possible interpretations of this result (I don’t know which is right, maybe it is ‘none of the above’):

1. No interior

2. Excited interior

3. Strong complementarity

4. The quantum donkey

Possible interpretation 1:

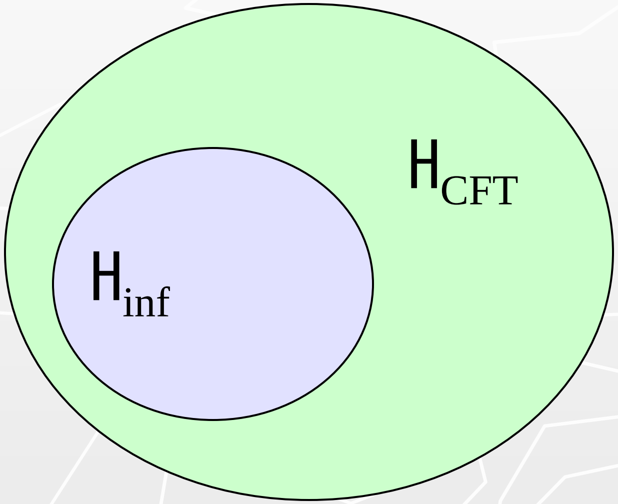
There is no $\hat{\tilde{b}}$ because there is no interior.

Possible interpretation 2:

The properties $[\hat{\tilde{b}}^\dagger, \hat{\tilde{b}}] = -1$, $[H, \hat{\tilde{b}}^\dagger] = -\omega \hat{\tilde{b}}^\dagger$ might have large corrections for highly excited states, evading the argument. OK, but it implies that almost all states are highly excited.

Possible interpretation 3: ‘Strong Complementarity’
(Banks & Fischler; Bousso; Harlow & Hayden; Page)

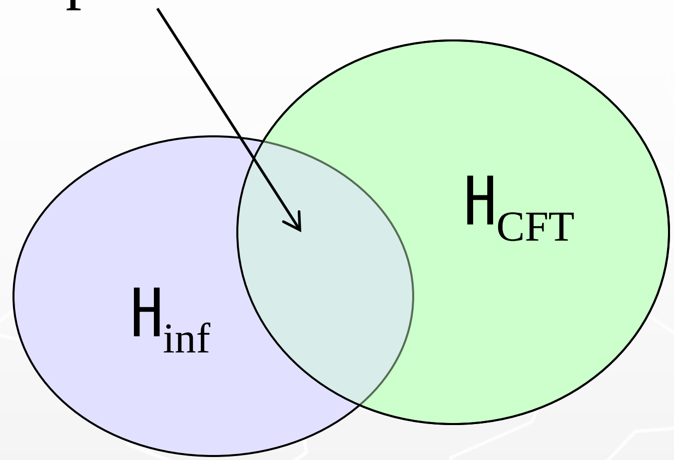
Suppose that the interior exists. Consider the effective field theory of the infalling observer. In the standard picture of Black Hole complementarity, this Hilbert space is contained in that of the CFT:



But this can’t be, because $\mathcal{H}_{\mathrm{inf}}$ contains $\tilde{b}$ and $\mathcal{H}_{\mathrm{CFT}}$ does not!

Rather, it must be that $\mathcal{H}_{\text{CFT}}$ contains only that subspace of $\mathcal{H}_{\text{inf}}$ states that can form in collapse.

tentative picture:



Interior observer has their own Hilbert space.

In retrospect this should have been obvious, because in the standard understanding of black holes only the infalling vacuum state forms, by the adiabatic principle.

With the firewall argument, it is a different subspace that forms.

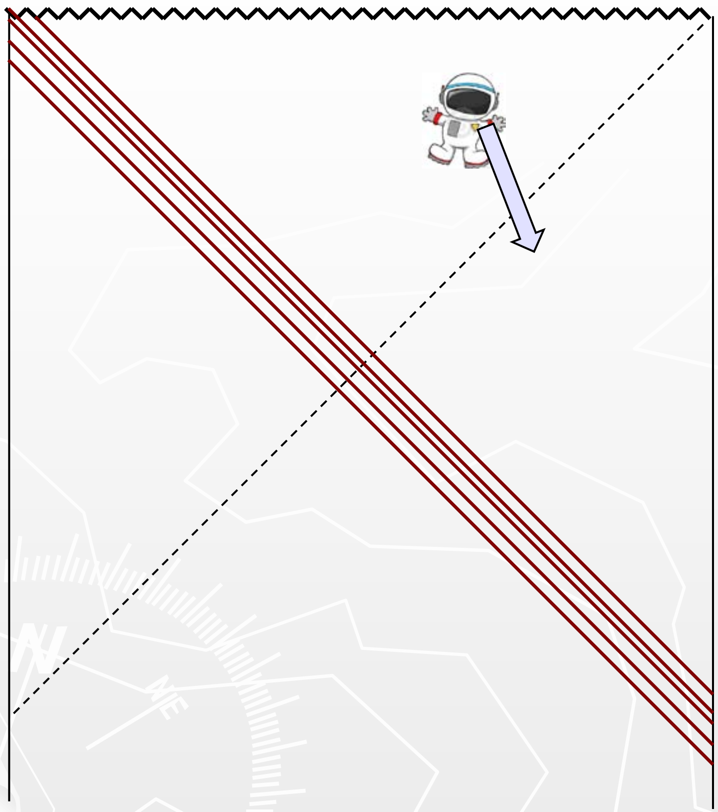
This is subtle. If the CFT Hilbert space contains ‘everything that can happen,’ why do we need anything else?

Because we would like to make measurements like

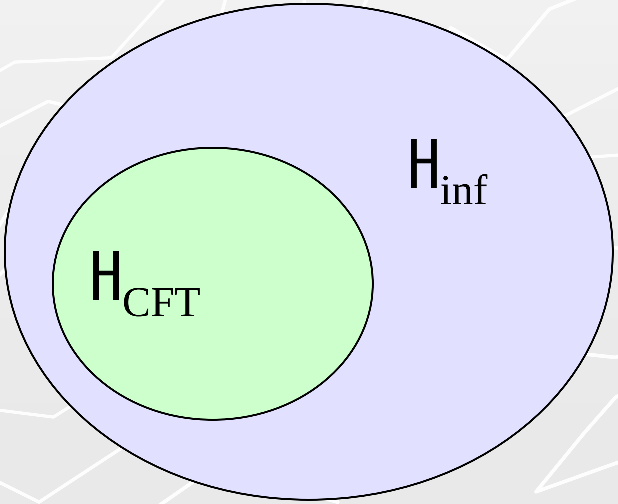
$$\langle \text{state with specific internal excitations} \,|\, \text{state of system} \rangle$$

For example, we would like to ask, what is the probability to find infalling vacuum. The ket $|\,\rangle$ is in $\mathcal{H}_{CFT}$, but the bra $\langle\,|$ is not.

Does the *b* argument from part 1 still apply?



The infalling observer can see the matter that formed the black hole. So the Hilbert space H_{inf} that describes all possible observations must contain H_{CFT}. So it seems we actually have:



The *b* argument then pushes forward to H_{inf}.

Possible interpretation 4: Nonlinear state dependence

(Papadodimas+Raju 1211.6767, Verlinde[2] 1211.6913)

In a typical CFT state, the distribution of $\hat{N}_b$ is thermal. If we *assume* that some particular such state $|\psi\rangle$ is infalling vacuum, for which $N_b = N_{\tilde{b}}$, then by expanding in $\hat{N}_b$ eigenstates we identify $\hat{N}_{\tilde{b}}$ eigenstates:

$$|\psi\rangle = \sum_n |n\rangle_b |\psi_n\rangle_{\mathrm{BH}}$$

We can thus construct the internal CFT.

(Can't cover whole of $\mathsf{H}_{\mathrm{CFT}}$, by previous argument).

To avoid firewall: just assume that whatever state the black hole is in is $|\psi\rangle$?!

This makes quantum mechanics nonlinear --- observables depend on the choice of this base state:

Normal QM operator: $\psi_i \to \mathcal{O} \to \psi_f$

$$\begin{tikzcd} & \psi_{\text{base}} \arrow[d] & \\ \text{Here:} \quad \psi_i \arrow[r] & \mathcal{O} \arrow[r] & \psi_f \end{tikzcd}$$

Not the same as normal background dependence, like

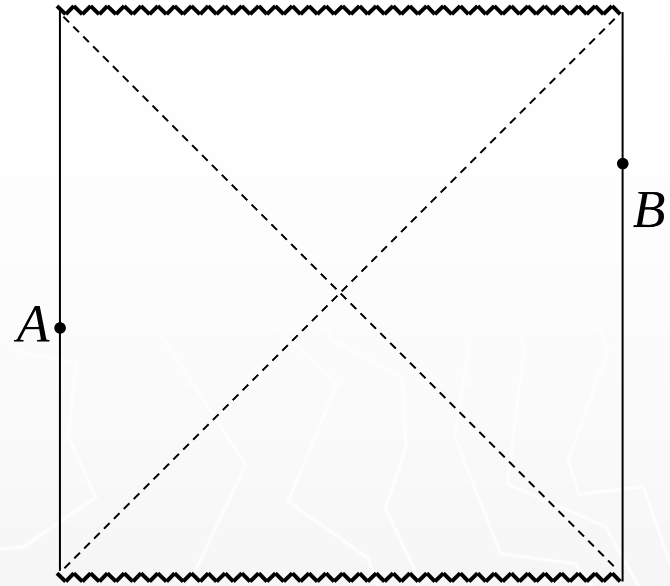
$$b = \int d^dx\, K_1(x)\mathcal{O}(x) + \int d^dx\, d^dy\, K_2(x,y)\mathcal{O}(x)\mathcal{O}(y) + \dots$$

When the single trace operator $\mathcal{O}(x)$ gets a vev, this adjusts, but it is still a linear operator.

H. Verlinde: choice of base state is “pinning the tail on the quantum donkey.”

“God not only plays dice with the world, She also plays pin the tail on the quantum donkey.”

III. A comment on EPR = ER
(Maldacena and Susskind)



Maldacena 2001: two-sided AdS geometry calculates two-CFT correlators in thermofield state

$$\langle\psi|A_L(-t)B_R(t')|\psi\rangle$$

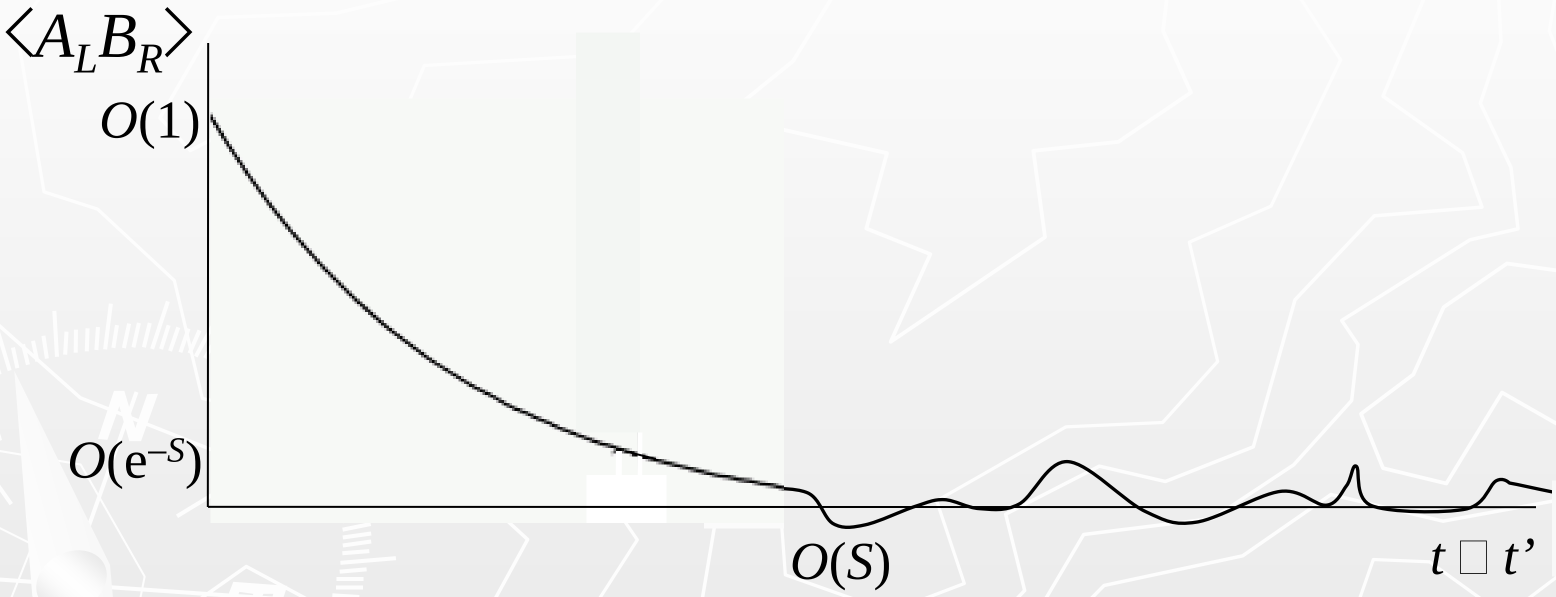
$$= \sum_{\alpha,\beta,\gamma,\delta} e^{-itE_{\delta\gamma}-it'E_{\alpha\beta}}\psi^*_{\delta\beta}\psi_{\gamma\alpha}A_{\gamma\delta}B_{\beta\alpha}$$

$$\psi_{\alpha\gamma} = Z^{-1/2}\delta_{\alpha\gamma}e^{-\beta E_\alpha/2}$$ (Energy eigenbasis)

Does the generic highly entangled state, e.g. one produced thermally, have a geometric interpretation? (cf. Shenker and Stanford, van Raamsdonk)

What do we mean by a geometric interpretation?

For the thermofield state, the time dependence of the correlator is of the form



Exponential falloff given by AdS calculation ≡ geometric
Long-term $O(e^{-S})$ dominated by random phases ≡ non-geometric

$$\langle A_L B_R \rangle = \sum_{\alpha,\beta,\gamma,\delta} e^{-itE_{\delta\gamma}-it'E_{\alpha\beta}} \psi^*_{\delta\beta}\psi_{\gamma\alpha} A_{\gamma\delta} B_{\beta\alpha}$$

Form of matrix element in chaotic systems:

$$A_{\alpha\beta} = \mathsf{A}(E)\delta_{\alpha\beta} + e^{-S(E)/2} f(E) R^A{}_{\alpha\beta}$$

Eigenstate Thermalization Hypothesis (Deutsch, Srednicki).
A, S, f are smooth functions, $R_{\alpha\beta}$ is random and $O(1)$, with

$$R^A{}_{\alpha\beta} R^B{}_{\gamma\delta} = \text{smooth} \times \delta_{\alpha\delta}\delta_{\beta\gamma} + \text{random}$$

With this, the opposite-side correlator is exponentially small and dominated by random phases at *all* times, so apparently no geometric interpretation, (except for states that are diagonal in energy).

IV. If there is a firewall, why should the Hawking calculation give the right flux?

• The Hawking flux is determined by the density matrix for *b*

• This is the same in every microstate, up to exponentially small corrections, as it is in the thermofield state (by the ETH). (cf. Mathur)

• The thermofield state is described by EFT across the horizon, so the Hawking calculation holds in every microstate.

• Unlike the usual derivation of the flux, this does not imply the same fine-grained result and does not assume infalling vacuum for the general microstate.

Conclusion

The fact that we depend so much on logical arguments rather than first-principles derivations strongly suggests that our current theory of quantum gravity is incomplete, even in AdS.

We need a better nonperturbative construction of the bulk theory.