Title: The Black Hole Information Paradox in the Age of Holographic Entanglement Entropy

Speakers: Netta Engelhardt

Series: Colloquium

Date: March 16, 2022 - 2:00 PM

URL: https://pirsa.org/22030037

Abstract: The black hole information paradox -- whether information escapes an evaporating black hole or not -- remains one of the greatest unsolved mysteries of theoretical physics. The apparent conflict between validity of semiclassical gravity at low energies and unitarity of quantum mechanics has long been expected to find its resolution in the deep quantum gravity regime. Recent developments in the holographic dictionary and in particular its application to entanglement, however, have shown that a semiclassical analysis of gravitational physics has a hallmark feature of unitary evolution. I will describe this recent progress and discuss some potential new avenues for working towards a resolution of the information paradox.

The Black Hole Information Paradox in the Age of Holographic Entanglement Entropy

Netta Engelhardt

MIT

Perimeter Institute Colloquium

The Black Hole Information Paradox

1. What is it?

2. Why study it?

3. How to make progress on it?

The Information Paradox: the what

- “semiclassical gravity”: weakly coupled limit of quantum gravity
- Should be valid near event horizons of (large) black holes.
- The information paradox: semiclassical gravity prediction contradicts QM. Fixing it requires large modifications to semiclassical gravity at *at the event horizon of a black hole*.

Image credit: Event Horizon Telescope

Information Paradox: the what

Semiclassical gravity predicts that black holes evaporate

... and along the way violate unitary evolution of closed quantum systems.

The Information Paradox: the why

- Information loss $\Rightarrow$ loss of determinism
- Black holes exist thanks, LIGO, EHT!, so nature has a definitive answer. And furthermore, without an answer to this question, our description of physical phenomena in the universe is incomplete.

Table of Contents

The Information Paradox: Then and Now

Then:

- Quantum gravity is *necessary* to see information conservation.
- Basic idea: semiclassical gravity is inconsistent with unitary evaporation.

Now:

- A purely semiclassical analysis consistent with unitary evaporation. Almheiri, Engelhardt, Marolf, Maxfield; Penington

BH Entropy: Wheeler's Gedankenexperiment

"No Hair Theorem": black holes in classical gravity have no microstates.

"Teacup Experiment"

No microstates ⇒ no entropy. Throw a hot teacup into the black hole – and decrease the entropy of the universe.

Second Law of Thermodynamics

If your theory is found to be against the second law of thermodynamics I can give you no hope; there is nothing for it but to collapse in deepest humiliation.

– Sir Arthur Eddington

Black Hole Entropy

Bekenstein: Black holes are thermal: they have temperature and entropy due to quantum gravity microstates:

$$S_{\mathrm{BH}} = \frac{\mathrm{Area}[\text{ event horizon }]}{4G\hbar}.$$

For a black hole close to equilibrium, we expect a temperature $T \propto M^{-1}$.

This is due to quantum corrections: black holes in classical gravity are perfect absorbers.

Evaporating Black Holes

Temperature ⇒ radiation. If

$$T_{BH} > T_{surroundings}$$

radiation ⇒ mass decreases

$$T \propto M^{-1}$$

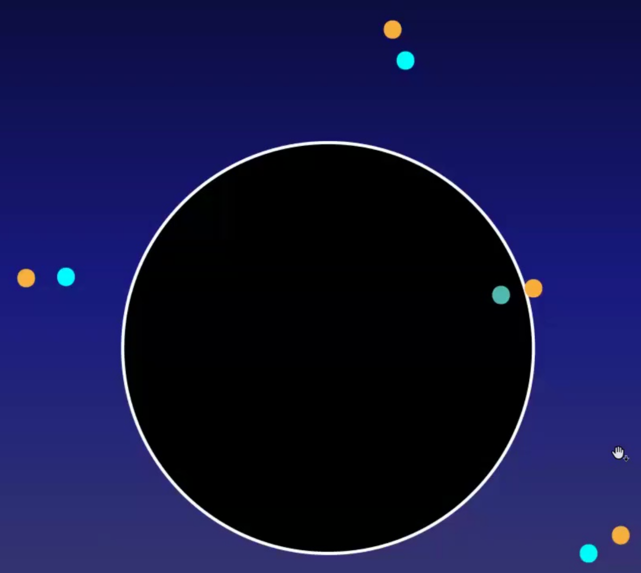
BH evaporates until it "explodes".

Some numbers:

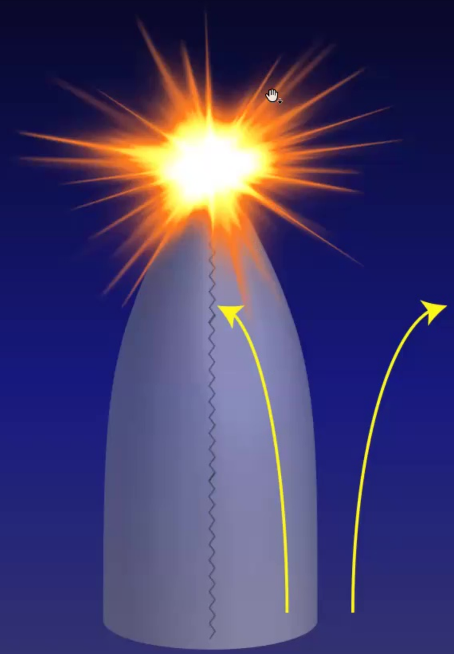
- Sagittarius A*, $T \sim 10^{-14} K$
- Solar mass black hole, $T \sim 10^{-8} K$
- Lunar mass black hole, $T \sim T_{CMB}$
- Coronavirus-sized black hole, $T \sim$ room temperature.

Collapsing Black Holes

Evaporating Black Holes



Evaporating Black Holes



The paradox: semiclassical gravity $\Rightarrow$ information loss during evaporation.

What Quantity can Diagnose (Non)Unitarity?

Define the von Neumann entropy

$$S_{vN}[\rho] = -\mathrm{tr}\,(\rho \log \rho)$$

Properties of S_{vN}:

- It is invariant under unitary evolution.
- It vanishes for a pure state (and only for a pure state): $S_{vN}[|\psi\rangle\langle\psi|] = 0$
- It is bounded from above by the thermal entropy (also called the coarse-grained entropy)

Information Loss

Rough intuition: the information in ρ can be quantified in terms of

$$S_{\text{thermal}} - S_{\text{vN}}.$$

The thermal state has essentially **no information content**. If a system evolves to larger S_{vN}, **information is lost**.

This happens all the time in open systems!

But the entire universe is not an open system.

A Toy Model

Alice falling down the ~~rabbit~~ black hole

$$|\psi\rangle = \frac{1}{\sqrt{2}}\left(|+\rangle|+\rangle + |-\rangle|-\rangle\right)$$

Bob by himself is in a density matrix:

$$\rho_{\mathrm{Bob}} = \frac{1}{2}\left(|+\rangle\langle+| + |-\rangle\langle-|\right) = \frac{1}{2}\mathbb{1}$$

obtained by tracing out Alice.

Information Loss in Hawking Radiation

In toy model of the EPR pair, the initial state of the pair is pure:

$$S_{\text{initial}} = 0.$$

After the black hole has finished evaporating,

$$S_{\text{final}} = S_{\text{Bob}} = \log 2 \neq 0$$

The entropy has increased: the system was not evolving unitarily.

Information Loss in Hawking Radiation

In toy model of the EPR pair, the initial state of the pair is pure:

$$S_{\mathrm{initial}} = 0.$$

After the black hole has finished evaporating,

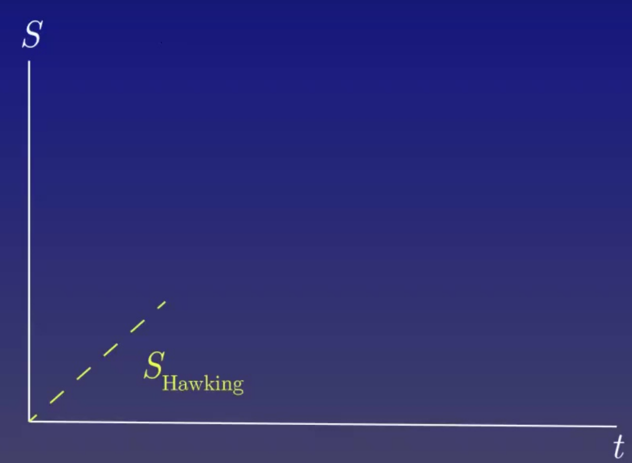
$$S_{\mathrm{final}} = S_{\mathrm{Bob}} = \log 2 \neq 0$$

The entropy has increased: the system was not evolving unitarily.

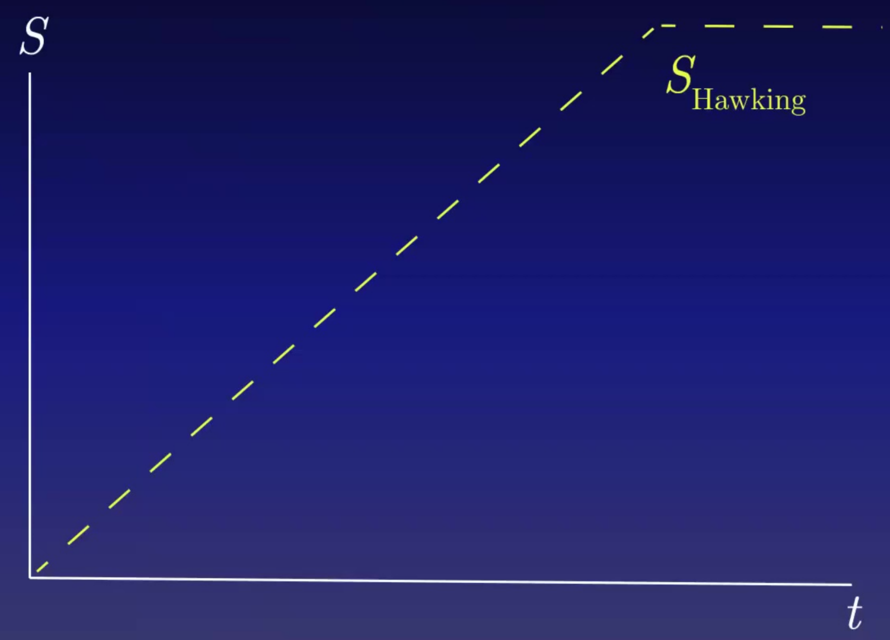
Hawking's Calculation

More accurately:

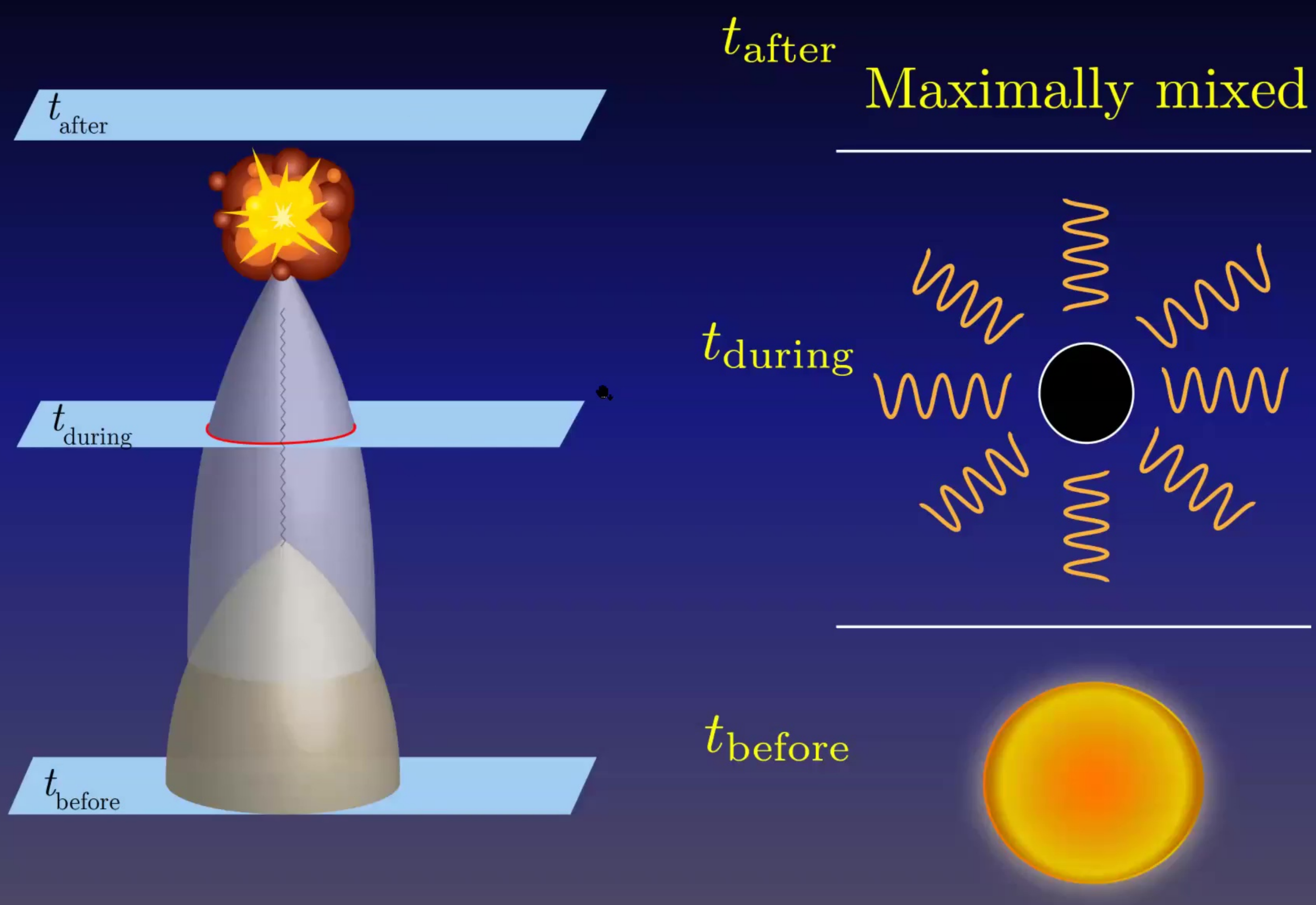
1. Universe is in some pure quantum state $|\psi\rangle$.
2. Form a black hole in this universe.
3. Compute $S_{\mathrm{vN}}[\mathrm{BH\ exterior}]$



Hawking's Calculation



Computing $S_{\mathrm{vN}}[\mathrm{exterior}(t)]$



Post-Evaporation

Semiclassical Gravity: Post-evaporation state is mixed

$$S_{\mathrm{vN}}[\mathrm{exterior}] > 0$$

Standard Quantum: Evolution of the *entire* universe – a closed system – is unitary:

$$|\psi(t)\rangle = e^{-iHt}|\psi\rangle$$

So:

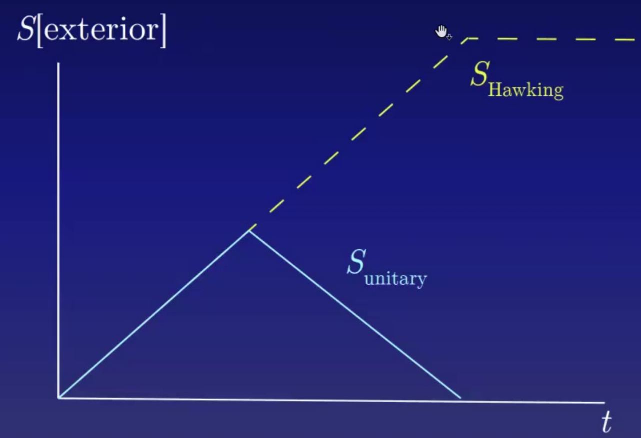
$$S_{\mathrm{vN}}[\mathrm{entire\ universe}] = S_{\mathrm{vN}}[|\psi(t)\rangle] = S_{\mathrm{vN}}[|\psi\rangle] = 0$$

Conflict: Post-evaporation exterior = entire universe.

$S_{\mathrm{vN}}[\mathrm{universe}]$ should vanish.

The Page Curve

What we should get vs what we actually get:



Which is the correct curve for QG? How do we compute it?

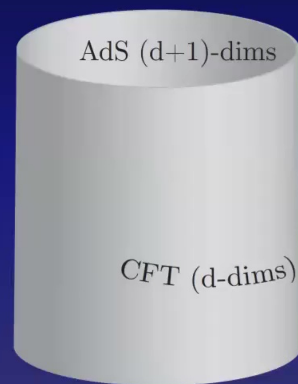
Preview: the New Developments

- Hard to see where Hawking's analysis goes wrong.
- Need input from QG to figure this out?
- New developments: a completely semiclassical calculation of the Page curve – with indirect input from QG.
- Specifically, from holography.

Holography

AdS/CFT aka Holography

Quantum gravity "in a box" (with Anti-de Sitter boundary conditions) is dual to a lower-dimensional nongravitational QFT.



Colloquially, we call the AdS quantum gravity the bulk, and the QFT the boundary.

Holography: a Black Hole is just another Quantum System

$\Rightarrow$ a black hole in AdS is just an ordinary, nongravitational quantum system.

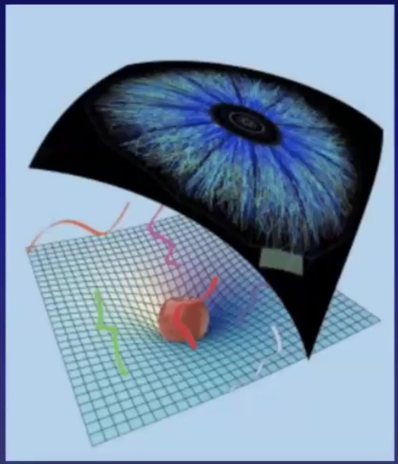


image credit: ESI Programme on AdS Holography and the Quark-Gluon Plasma

Unitary BH Evaporation

So

Black hole $=$ ordinary nongrav closed quantum system

And nongrav isolated quantum systems evolve unitarily.

The Idea

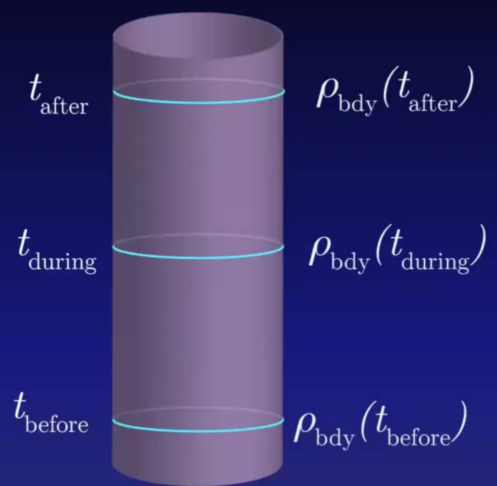
But this isn't satisfactory! How does quantum gravity preserve unitarity??

Know: the usual formula in the bulk:

$$S_{\mathrm{vN}}[\rho_{\mathrm{bulk}}] = -\mathrm{tr}\,\rho_{\mathrm{bulk}} \ln \rho_{\mathrm{bulk}}$$

gives "Hawking curve": information loss (here ρ_{bulk} is the density matrix describing the quantum fields in the bulk portion outside of the black hole).

Meanwhile, on the back of the soupcan...



Also know: in the CFT via the usual formula:

$$S_{\text{vN}}[\rho_{\text{bdy}}] = -\text{tr}\rho_{\text{bdy}} \ln \rho_{\text{bdy}}$$

gives a unitary result: information is conserved.

The Idea

So: there's a nongrav computation of a unitary entropy and a grav computation of a non-unitary entropy.

So what if instead of using

$$-\mathrm{tr}\rho_{\mathrm{bulk}} \ln \rho_{\mathrm{bulk}}$$

We use the **holographic prescription** for computing

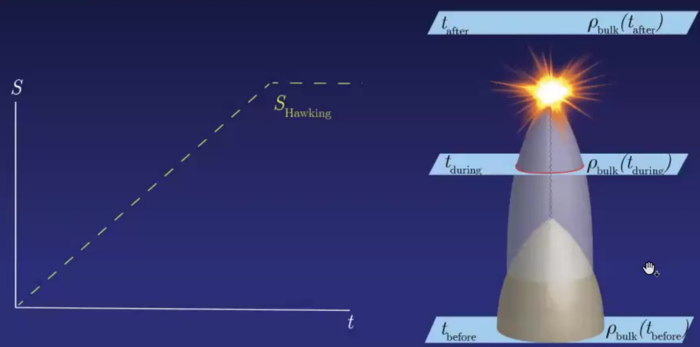
$$-\mathrm{tr}\rho_{\mathrm{bdy}} \ln \rho_{\mathrm{bdy}}$$

This could give a **gravitational** calculation of a **unitary** entropy!

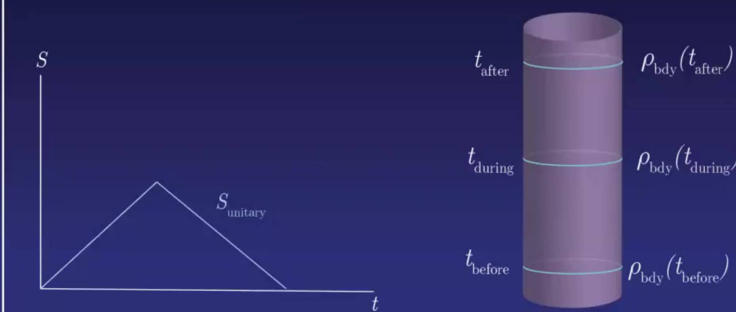
Remarkably, it does!

The Different Angles

Option 1: Geometric, Non-Unitary



Option 2: Non-geometric, Unitary



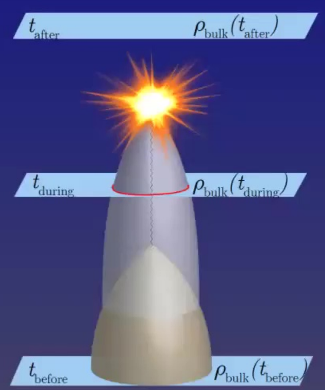
OR...

Option 3: use holographic dictionary! Do the calculation in option 2 in the geometric language of the bulk.

The Idea

Compute

using the gravitational language of



Holographic Prescription for Computing $S_{\mathrm{vN}}[\rho_{\mathrm{bdy}}]$

$$S_{\mathrm{vN}}[\rho_{\mathrm{bdy}}] = \frac{\mathrm{Area}[\chi]}{4G\hbar} + S_{\mathrm{vN}}[\rho_{\mathrm{out}[\chi]}] \equiv S_{\mathrm{gen}}[\chi]$$

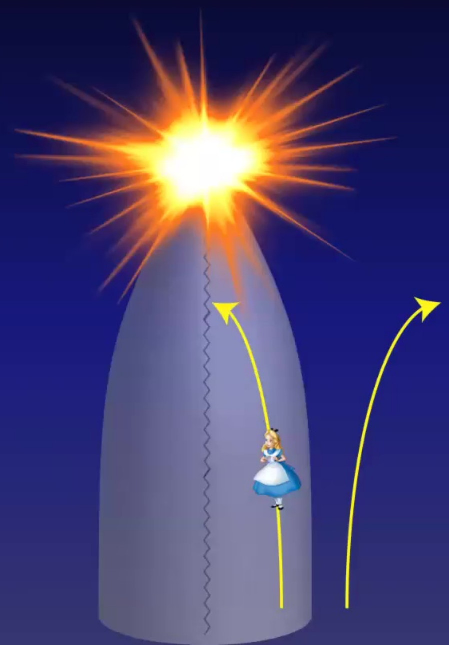
where χ is a distinguished type of surface we called "quantum extremal".

Engelhardt, Wall; without quantum corrections: Ryu-Takayanagi, Hubeny-Rangamani-Takayanagi, and Faulkner-Lewkowycz-Maldacena

If there are multiple quantum extremal surfaces, use the minimal one.

Evaporating Black Holes

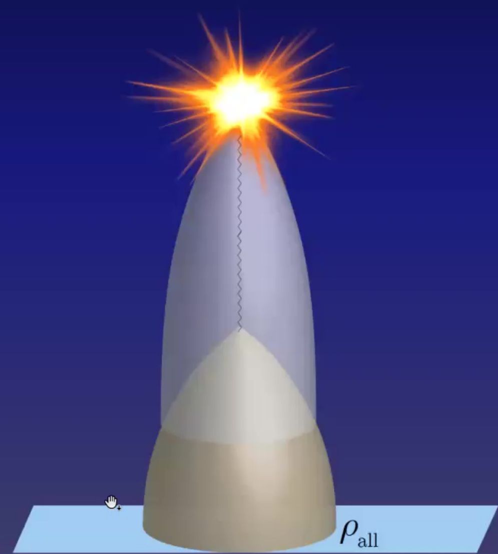
The next few slides are based on Almheiri, Engelhardt, Marolf, Maxfield; Penington.



Quantum Extremal Surfaces

Initially, the QES is the empty set.

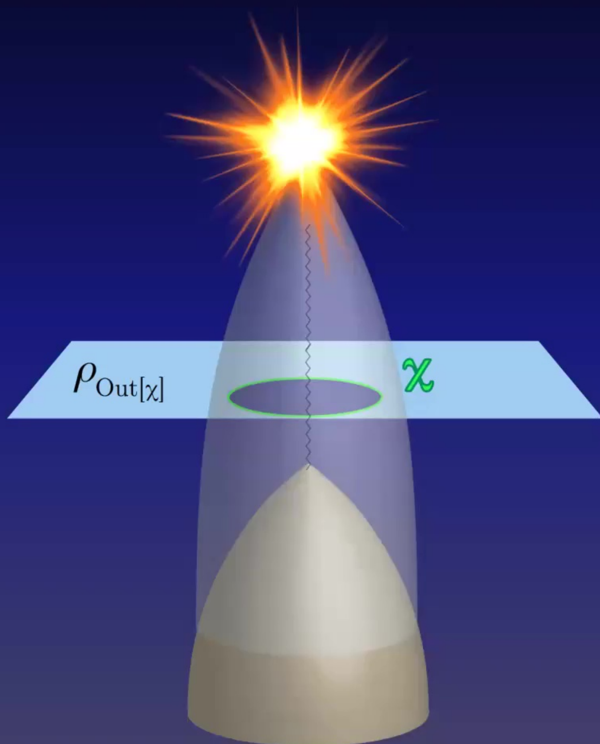
$$S_{vN}[\rho_{bdy}] = \frac{\text{Area}[\varnothing]}{4G\hbar} + S_{\text{vN}}[\rho_{\text{all}}] = S_{\text{vN}}[\rho_{\text{all}}]$$

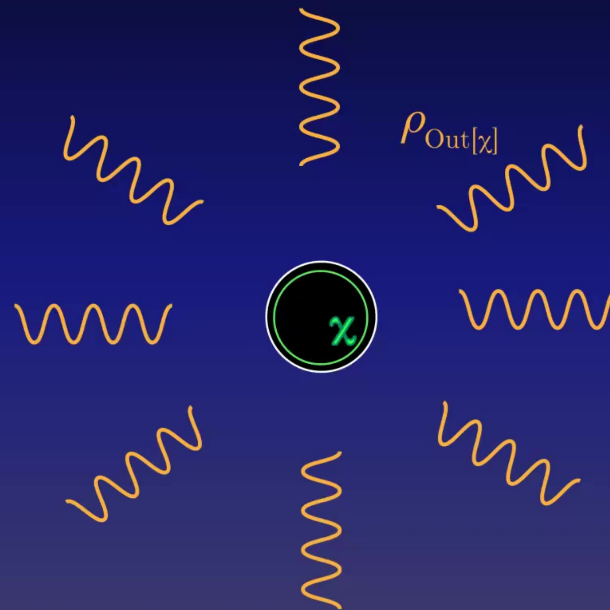


With time $S_{\text{vN}}[\rho_{\text{all}}]$ grows.

Quantum Extremal Surfaces

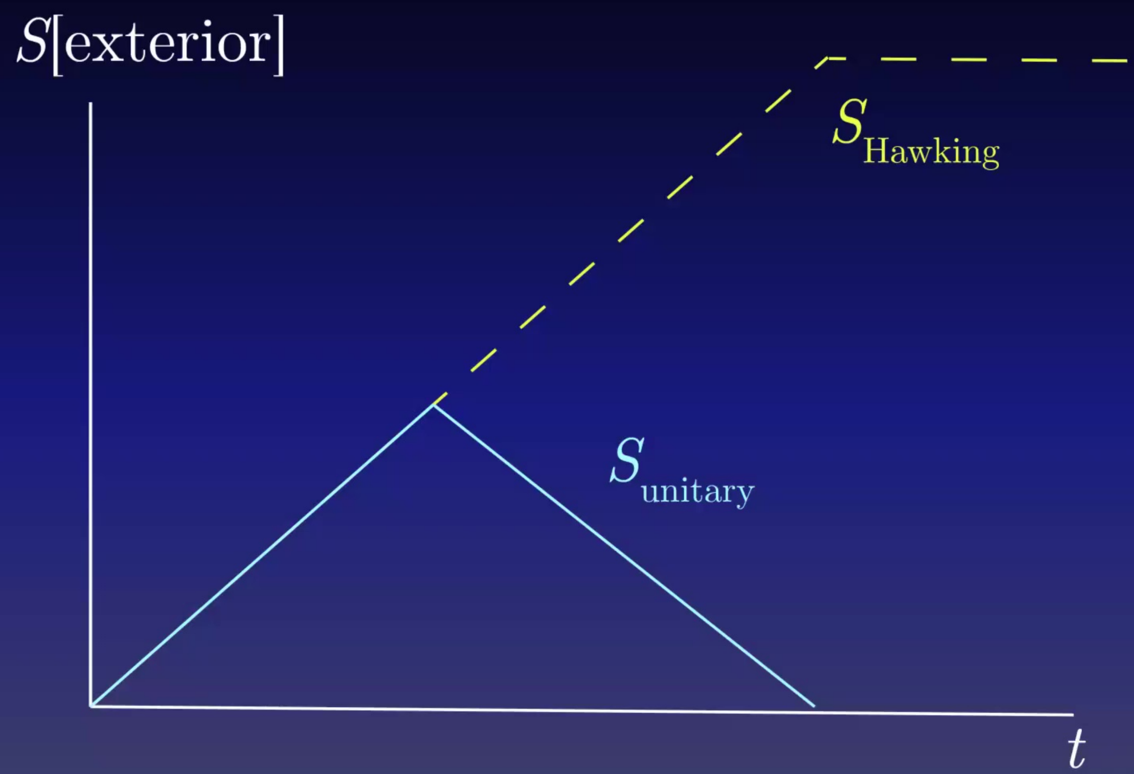
About halfway through evaporation, a new quantum extremal surface:





As we evolve forward in time, $S_{\mathrm{gen}}[\chi]$ decreases.

The Page Curve



The Page curve for unitary evolution is obtained using only semiclassical physics!

Taking a Step Back

What happened here?

$$S_{vN}[\rho_{\mathrm{bdy}}] = \frac{\mathrm{Area}[\chi]}{4G\hbar} + S_{vN}[\rho_{\mathrm{Out}[\chi]}]$$

The contribution from χ – the quantum extremal surface – saves the day.

The only place where quantum gravity appeared is in the **interpretation** of $S_{\mathrm{gen}}[\chi]$ as the entropy of the radiation.

Why does this work?

Why does quantum gravity repackage entropy in this way?

What microscopic Planckian physics is responsible for the success of the quantum extremal surface prescription?

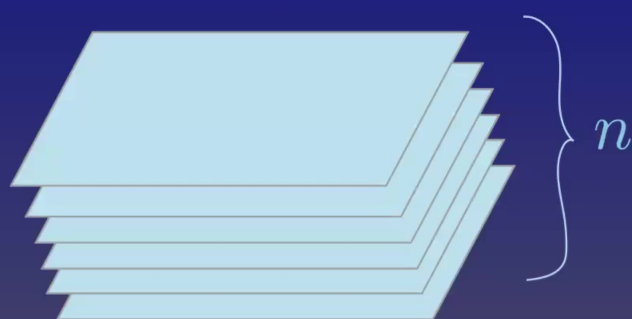
Some Progress on this Front

How do we generally compute entropies?

The Replica Trick

$$S_{\mathrm{vN}}[\rho] = -\mathrm{tr}\rho \ln \rho = -\lim_{n\to 1}\left(\frac{\partial}{\partial n}\log \mathrm{Tr}[\rho^n]\right)$$

where ρ^n is the state ρ on n independent copies ("replicas") of this nongravitational theory.



Some Progress on this Front

The Replica Trick in Holography:

- In AdS/CFT, we know how to relate $\mathrm{Tr}(\rho^n)$ to the "gravitational path integral": Lewkowycz, Maldacena

$$\int Dg e^{S}$$

- Doing this, Penington et al; Almheiri et al justify the new QES that leads to unitarity. But at cost of these n replicas having nontrivial correlations.

Some Progress on this Front

The Replica Trick in Holography:

- In AdS/CFT, we know how to relate $\mathrm{Tr}(\rho^n)$ to the "gravitational path integral": Lewkowycz, Maldacena

$$\int Dg e^{S}$$

- Doing this, Penington et al; Almheiri et al justify the new QES that leads to unitarity. But at cost of these n replicas having nontrivial correlations.

- But these are n-independent copies of a single quantum theory! How can they be correlated?

Comparing Semiclassical Calculations

- We have two different calculations that use the tools of semiclassical gravity:
  1. Hawking's calculation: inconsistent with a pure out state.
  2. QES calculation: consistent with a pure out state.
- How do we bridge the gap between the two?
- **One perspective**: Hawking's calculation corresponds to using the empty set: the wrong QES.
- **Another perspective**: Even if the radiation purifies itself, it takes an exponentially (in G_N^{-1})powerful quantum computer to detect that: if you threw a diary into the black hole, and the information in the diary eventually gets radiated out, you would need an extremely powerful quantum computer to read the diary contents from the radiation Harlow, Hayden.

Comparing Semiclassical Calculations

- We have two different calculations that use the tools of semiclassical gravity:

  1. Hawking's calculation: inconsistent with a pure out state.
  2. QES calculation: consistent with a pure out state.

- How do we bridge the gap between the two?

- **One perspective**: Hawking's calculation corresponds to using the empty set: the wrong QES.

- **Another perspective**: Even if the radiation purifies itself, it takes an exponentially (in G_N^{-1})powerful quantum computer to detect that: if you threw a diary into the black hole, and the information in the diary eventually gets radiated out, you would need an extremely powerful quantum computer to read the diary contents from the radiation Harlow, Hayden.

A Hypothesis

If you were an observer computationally restricted to sub-exponentially complicated tasks, you would never know that the radiation purifies itself!

A Hypothesis

Hawking's calculation coarse-grains over exponential complexity.

How does this mesh with comparing the QES calculation to Hawking's calculation?

Complexity and QESs

The holographic avatar of exponential complexity is nonminimal QESs Brown et al. Using the "wrong" QES is equivalent to coarse-graining over exponential complexity. NE, Penington, Shahbazi-Moghaddam

Current State of the Information Paradox

Spoiler: **we haven't solved it yet.**

- What entropy is the quantum extremal surface/gravitational path integral calculating? How do we compute it directly from the formula $-\mathrm{tr}\rho\ln\rho$?
- What does an observer outside the black hole actually measure?
- Is the path integral doing an averaging of some sort?
- What are the internal dynamics of an evaporating black hole?
- If the bridge between the unitary calculation and Hawking's calculation is computational complexity, how do we understand it in the language of Hawking's original computation?

Thank you!