The Black Hole Information Paradox: a Resolution on the Horizon?

Netta Engelhardt

MIT

Colloquium, Extreme Universe Project

The Black Hole Information Paradox

The Black Hole Information Paradox

1. What is it?

The Black Hole Information Paradox

1. What is it?

2. Why study it?

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

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: 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 the event horizon of a black hole*.

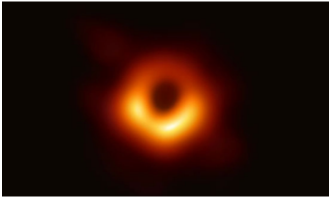


Image credit: Event Horizon Telescope

Information Paradox: the what

Semiclassical gravity predicts that black holes evaporate

Information Paradox: the what

Semiclassical gravity predicts that black holes evaporate

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

Information Paradox: the what

Semiclassical gravity predicts that black holes evaporate

... and along the way violate unitary evolution of closed quantum systems. In direct violation of unitarity of quantum mechanics!

The Information Paradox: the why

- Violations of unitarity of closed quantum systems (information loss) $\Rightarrow$ loss of determinism

The Information Paradox: the why

- Violations of unitarity of closed quantum systems (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.

The Information Paradox: the why

- Violations of unitarity of closed quantum systems (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.
- Ultimately, the information paradox has a lot to teach us about quantum gravity!

The Information Paradox: Then and Now

Then:

- Quantum gravity is *necessary* to see information conservation.

The Information Paradox: Then and Now

Then:

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

The Information Paradox: Then and Now

Then:

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

In 2019:

- A purely semiclassical analysis consistent with unitary evaporation.

  Almheiri, Engelhardt, Marolf, Maxfield; Penington

The Information Paradox: Then and Now

Then:

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

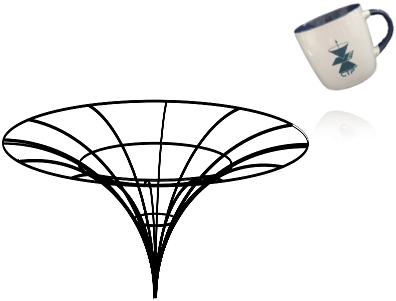
In 2019:

- A purely semiclassical analysis consistent with unitary evaporation. Almheiri, Engelhardt, Marolf, Maxfield; Penington I think it's safe to say that many (most?) of us thought this was impossible.

Overview

BH Entropy: Wheeler's Gedankenexperiment

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

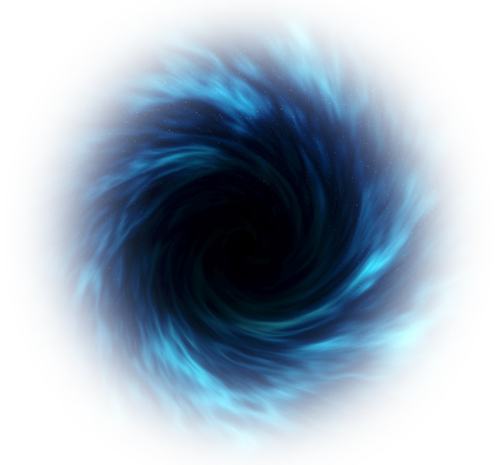


"Teacup Experiment"

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:

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

Black Hole Entropy

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

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

Tjey also have a temperature $T \propto M^{-1}$.

Black Hole Entropy

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

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

Tjey also have a temperature $T \propto M^{-1}$.

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

Evaporating Black Holes

Temperature $\Rightarrow$ radiation. If

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

radiation $\Rightarrow$ mass decreases

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

BH evaporates until it "explodes".

Evaporating Black Holes

Temperature $\Rightarrow$ radiation. If

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

radiation $\Rightarrow$ mass decreases

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

BH evaporates until it "explodes".

Some numbers:

- Sagittarius A*, $T \sim 10^{-14} K$

Evaporating Black Holes

Temperature $\Rightarrow$ radiation. If

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

radiation $\Rightarrow$ mass decreases

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

BH evaporates until it “explodes”.

Some numbers:

- Sagittarius A*, $T \sim 10^{-14} K$
- Solar mass black hole, $T \sim 10^{-8} K$

Evaporating Black Holes

Temperature $\Rightarrow$ radiation. If

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

radiation $\Rightarrow$ 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, $T \sim T_{CMB}$

Evaporating Black Holes

Temperature $\Rightarrow$ radiation. If

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

radiation $\Rightarrow$ mass decreases

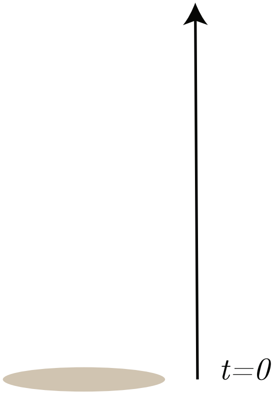
$$T \propto M^{-1}$$

BH evaporates until it "explodes".

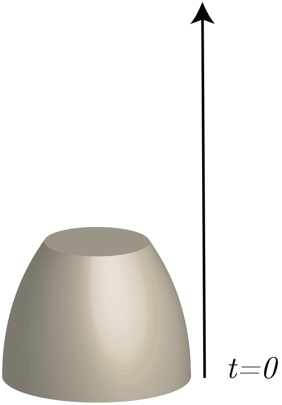
Some numbers:

- 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.

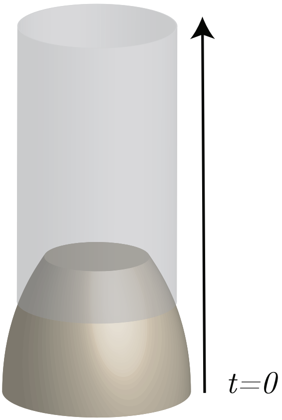
Collapsing Black Holes



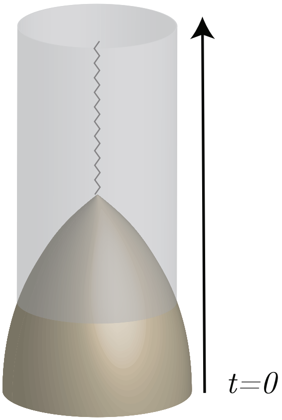
Collapsing Black Holes



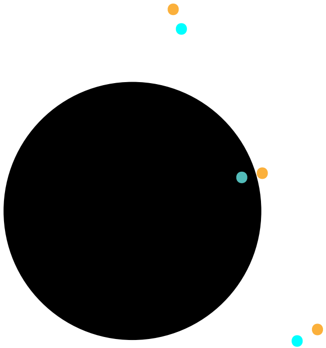
Collapsing Black Holes



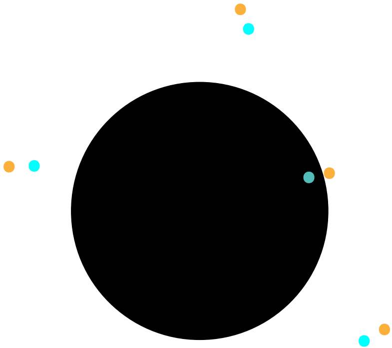
Collapsing Black Holes



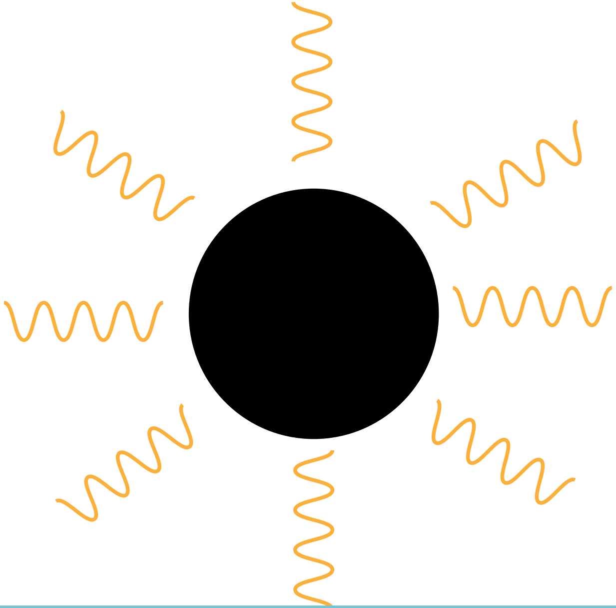
Evaporating Black Holes



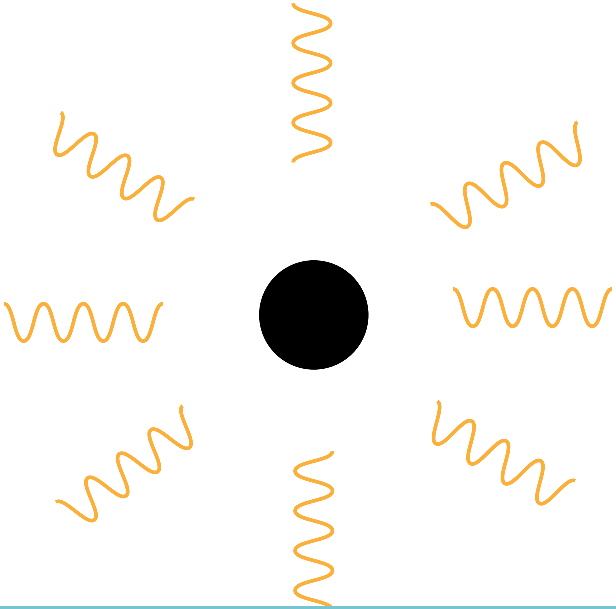
Evaporating Black Holes



Evaporating Black Holes

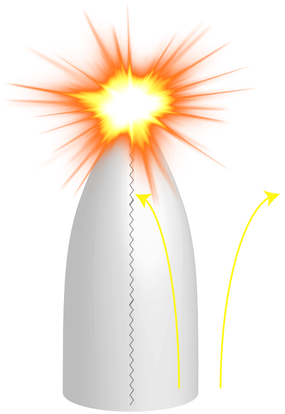


Evaporating Black Holes



Evaporating Black Holes

Evaporating Black Holes

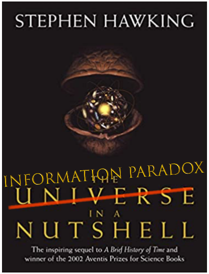


Semiclassical gravity ⇒ information loss during evaporation.

Hawking Radiation In a Nutshell

1. Some pure state collapses into a black hole.
   Before the black hole forms:

$$S_{\mathrm{vN}}[\rho_{\text{outside BH}}] = S_{\mathrm{vN}}[\rho_{\text{universe}}] = 0$$



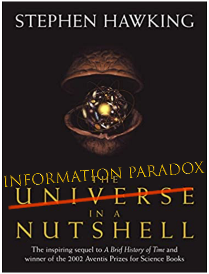
Hawking Radiation In a Nutshell

1. Some pure state collapses into a black hole.
   Before the black hole forms:

$$S_{\mathrm{vN}}[\rho_{\text{outside BH}}] = S_{\mathrm{vN}}[\rho_{\text{universe}}] = 0$$

2. At intermediate times:

$$S_{\mathrm{vN}}[\rho_{\text{outside BH}}] = S_{\mathrm{vN}}[\rho_{\text{rad}}] > 0$$



Hawking Radiation In a Nutshell

1. Some pure state collapses into a black hole. Before the black hole forms:

$$S_{\text{vN}}[\rho_{\text{outside BH}}] = S_{\text{vN}}[\rho_{\text{universe}}] = 0$$

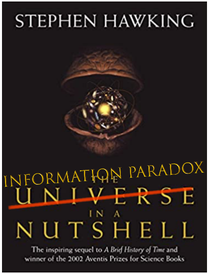
2. At intermediate times:

$$S_{\text{vN}}[\rho_{\text{outside BH}}] = S_{\text{vN}}[\rho_{\text{rad}}] > 0$$

3. After the BH evaporates, we must have

$$S_{\text{vN}}[\rho_{\text{rad}}] = S_{\text{vN}}[\rho_{\text{universe}}] = 0$$

if time evolution is unitary.



Hawking Radiation In a Nutshell

1. Some pure state collapses into a black hole.
   Before the black hole forms:

$$S_{\mathrm{vN}}[\rho_{\text{outside BH}}] = S_{\mathrm{vN}}[\rho_{\text{universe}}] = 0$$

2. At intermediate times:

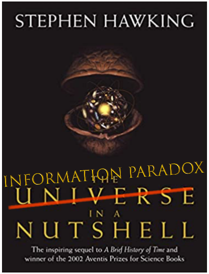
$$S_{\mathrm{vN}}[\rho_{\text{outside BH}}] = S_{\mathrm{vN}}[\rho_{\text{rad}}] > 0$$

3. After the BH evaporates, we must have

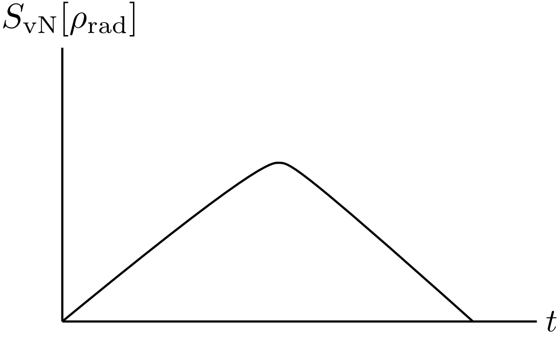
$$S_{\mathrm{vN}}[\rho_{\text{rad}}] = S_{\mathrm{vN}}[\rho_{\text{universe}}] = 0$$

   if time evolution is unitary.

4. So we expect a non-monotonic curve for $S_{\mathrm{vN}}[\rho_{\text{rad}}]$.



The Page Curve

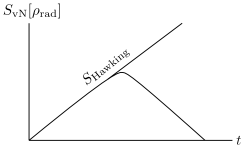


The Paradox

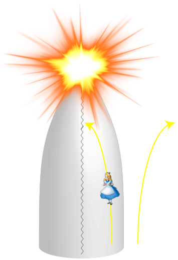
- To compute $S_{\mathrm{vN}}[\rho_{\mathrm{rad}}]$, we need a description of the physics at the black hole event horizon.

- *Semiclassical gravity*: the limit where large quantum gravity effects are suppressed; should be valid at the horizon.

- Semiclassical gravity $\overset{\text{Hawking}}{\Rightarrow}$ $-\mathrm{tr}\rho_{\mathrm{rad}} \ln \rho_{\mathrm{rad}}$ does *not* follow the Page curve:



Caricature of Information Loss



Quick argument for info loss:

$$|\psi\rangle = |+\rangle|-\rangle + |-\rangle|+\rangle$$

After evaporation, Alice disappears from the universe – even though the universe is a closed system.

State after evaporation is just $\rho_{\text{Bob}} = \frac{1}{2}\mathbb{I}$.

Information Loss in Hawking Radiation

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

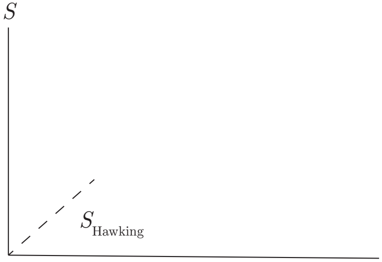
$$S_{\text{initial}} = 0.$$

After the black hole has finished evaporating,

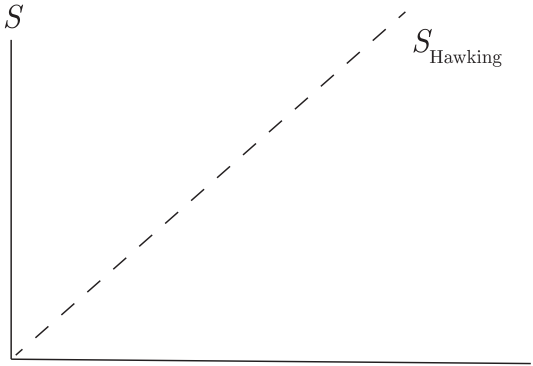
$$S_{\text{final}} = S_{\text{Bob}} = \log 2 \neq 0$$

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

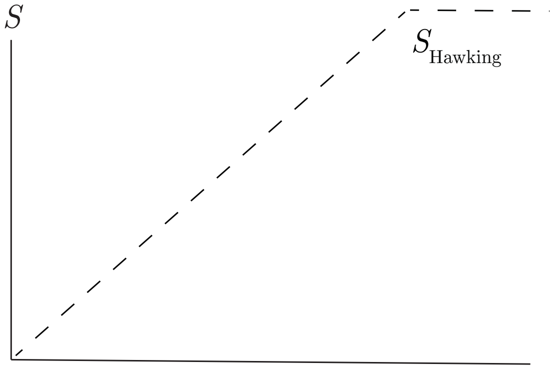
Hawking's Calculation



HAWKING'S CALCULATION



Hawking's Calculation



Overview

Progress: Computing the Entropy of Hawking Radiation

- It is hard to see where Hawking's analysis goes wrong.

Progress: Computing the Entropy of Hawking Radiation

- It is hard to see where Hawking's analysis goes wrong.
- Are there large corrections from quantum gravity in a regime where there shouldn't be?

Progress: Computing the Entropy of Hawking Radiation

- It is hard to see where Hawking's analysis goes wrong.
- Are there large corrections from quantum gravity in a regime where there shouldn't be?
- 3 years ago: a semiclassical calculation of the Page curve, but using a different – holographic – formula for S_{vN}.

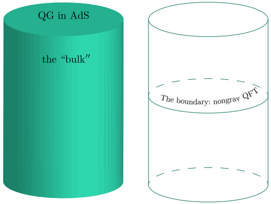
Progress: Computing the Entropy of Hawking Radiation

- It is hard to see where Hawking's analysis goes wrong.
- Are there large corrections from quantum gravity in a regime where there shouldn't be?
- 3 years ago: a semiclassical calculation of the Page curve, but using a different – holographic – formula for S_{vN}.
- This formula is a bit of a black box: can be interpreted as indirect input from quantum gravity

Holography

Holography

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



Colloquially, we call the gravitational theory the bulk, and the non-gravitational theory the boundary.

Holography: a Black Hole is just a Quantum System

This means that a black hole in AdS is just an ordinary, nongravitational quantum system.

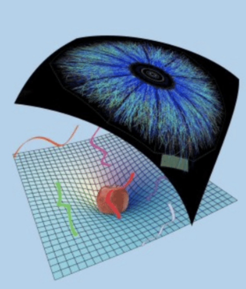


image credit: ESI Programme on AdS Holography and the Quark-Gluon Plasma

... nongravitational (closed) quantum systems evolve unitarily.

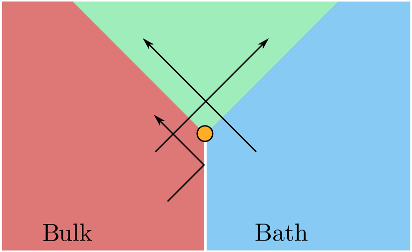
Computing the Page Curve

We know that $S_{\mathrm{vN}}[\rho_{\mathrm{rad}}]$ computed in the nongravitational "boundary" theory follows the Page curve.

AdS/CFT is a dictionary: it relates quantities in the boundary theory to the bulk theory.

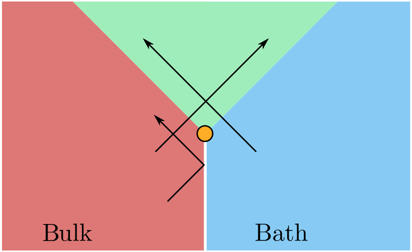
What quantity in the bulk theory translates to $S_{\mathrm{vN}}[\rho_{\mathrm{rad}}]$ as computed by the boundary theory?

Technical Aside: Auxiliary Bath



Technically, AdS black holes don't evaporate due to reflecting boundary conditions at infinity (the radiation just bounces back into the black hole). To force them to evaporate (and make them look more like flat space), we introduce a nongravitating cold bath at infinity.

Technical Aside: Auxiliary Bath



Technically, AdS black holes don't evaporate due to reflecting boundary conditions at infinity (the radiation just bounces back into the black hole). To force them to evaporate (and make them look more like flat space), we introduce a nongravitating cold bath at infinity. In the dual CFT can be viewed as coupling in an auxiliary CFT (the "bath") at lower temperature. The joint CFT+bath system evolves unitarily.

Holographic von Neumann Entropy

$S_{\rm vN}$ in AdS/CFT Ryu-Takayanagi, Hubeny-Rangamani-Takayanagi, Faulkner-Lewkowycz-Maldacena, Engelhardt-Wall

The von Neumann entropy of the radiation is given by:

$$S_{\rm vN}[\rho_{\rm rad}] = \frac{\text{Area}[\chi]}{4} + S_{\rm vN}[\rho_{\text{Out}[\chi]}]$$

where χ is a "quantum extremal surface": if you slightly perturb χ, the sum $\frac{\text{Area}[\chi]}{4} + S_{\rm vN}[\rho_{\text{Out}[\chi]}]$ won't change to leading order in the perturbation.

Holographic von Neumann Entropy

S_{vN} in AdS/CFT Ryu-Takayanagi, Hubeny-Rangamani-Takayanagi, Faulkner-Lewkowycz-Maldacena, Engelhardt-Wall

The von Neumann entropy of the radiation is given by:

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

where χ is a "quantum extremal surface": if you slightly perturb χ, the sum $\frac{\mathrm{Area}[\chi]}{4} + S_{\mathrm{vN}}[\rho_{\mathrm{Out}[\chi]}]$ won't change to leading order in the perturbation.

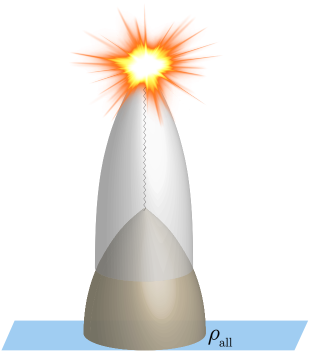
We get $\rho_{\mathrm{out}[\chi]}$ from a bipartition of the system into interior of χ and exterior.

Formula justified from euclidean gravity in evaporating black hole setup by Almheiri, Hartman, Maldacena, Shaghoulian, Tajdini; Penington, Shenker, Stanford, Yang

Quantum Extremal Surfaces

Initially, the QES is the empty set.

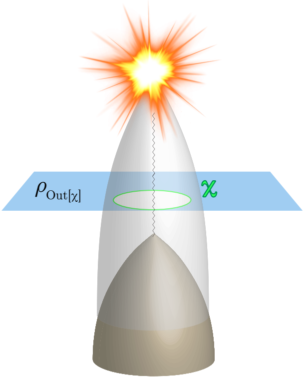
$$S_{vN}[\rho_{bdy}] = \frac{\text{Area}[\varnothing]}{4G\hbar} + S_{\text{vN}}[\rho_{\text{all}}] = S_{\text{vN}}[\rho_{\text{all}}]$$

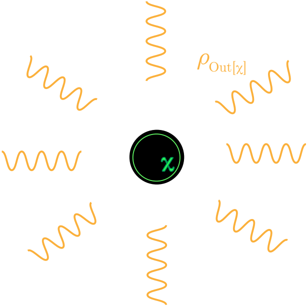


With time $S_{\text{vN}}[\rho_{\text{all}}]$ grows.

QESs in Evaporating BHs Almheiri, Engelhardt, Marolf, Maxfield; Penington

About halfway through evaporation, a new quantum extremal surface appears:

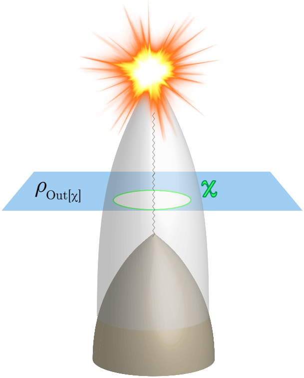


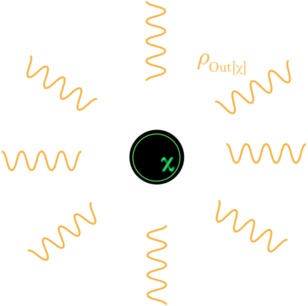


As we evolve forward in time, $S_{\mathrm{gen}}[\chi]$ decreases.

QESs in Evaporating BHs Almheiri, Engelhardt, Marolf, Maxfield; Penington

About halfway through evaporation, a new quantum extremal surface appears:





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

This alternative computation of S_{vN} in gravity *does* yield a unitary Page curve.

Takeaways

- We've computed the unitary Page curve using semiclassical gravity!
- What does this tell us?
- For example, how do we see that information has escapes, and that the black hole interior is now *encoded* in the radiation?

The Entanglement Wedge

To answer this, we'll take a detour by asking another question: what is the significance of the region $\mathrm{Out}[\chi]$?

Reconstruction in AdS/CFT?

AdS/CFT is a quantum error correcting code see work beginning with Almheiri, Dong, Harlow. How much of the bulk can we reconstruct from the state ρ_{bdy} of the nongravitational dual?

The Entanglement Wedge

To answer this, we'll take a detour by asking another question: what is the significance of the region Out$[\chi]$?

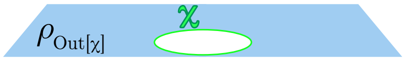
Reconstruction in AdS/CFT?

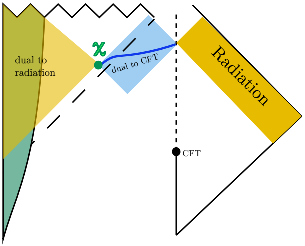
AdS/CFT is a quantum error correcting code see work beginning with Almheiri, Dong, Harlow. How much of the bulk can we reconstruct from the state ρ_{bdy} of the nongravitational dual?

Everything in Out$[\chi]$.*

The Entanglement Wedge of the Radiation

After the Page time, this is most of the black hole interior: information that fell into the black hole is now available in the radiation.





How do we use this to find a resolution of the black hole information paradox?

Overview

To recap:

Problem

The semiclassical approximation is:

1. Wrong: it is inconsistent with unitarity, i.e. we get the "Hawking" entropy curve from it.
2. But not *that* wrong: we can still use the effective description to compute, e.g. the unitary Page curve Page; Penington; Almheiri, NE, Marolf, Maxfield, without resorting to nonperturbatively suppressed effects – but by using the QES formula NE, Wall rather than the standard $-\mathrm{tr}\rho \ln \rho$.

So, just how seriously are we supposed to take semiclassical gravity?

Wrong and Right at the Same Time

- The information paradox: an apparent breakdown of the effective field theory where naively semiclassical gravity should be valid.
- Possible conclusion: don't take EFT seriously near the event horizon...
- The QES formula: take EFT seriously, but not *too* seriously.
- Route towards resolution: understanding in what ways EFT is right and in what ways it's wrong in a black hole setting.

Accessibility of Information

- If you throw a diary into the black hole, some time later its information should be accessible in the radiation.

Accessibility of Information

- If you throw a diary into the black hole, some time later its information should be accessible in the radiation.

- **Hayden-Preskill:** Assume BH dynamics drawn from unitary 2-design. Take an old black hole B and its radiation E. Throw $|i\rangle$ into it. The info is recoverable from the radiation in a time logarithmic in S_{BH}:
$$V_{EL}\,|\Psi\rangle_{B'EL} = |i\rangle_1 \otimes |\phi\rangle_{B',2}$$
where E is the early radiation, L is the late radiation (after $|i\rangle$ is thrown in) and B' is the black hole after L has radiated (and 1 and 2 are factors of $\mathcal{H}_{LE}$). But:

Accessibility of Information

- If you throw a diary into the black hole, some time later its information should be accessible in the radiation.

- **Hayden-Preskill:** Assume BH dynamics drawn from unitary 2-design. Take an old black hole B and its radiation E. Throw $|i\rangle$ into it. The info is recoverable from the radiation in a time logarithmic in S_{BH}:
$$V_{EL}\,|\Psi\rangle_{B'EL} = |i\rangle_1 \otimes |\phi\rangle_{B',2}$$
where E is the early radiation, L is the late radiation (after $|i\rangle$ is thrown in) and B' is the black hole after L has radiated (and 1 and 2 are factors of $\mathcal{H}_{LE}$). But:

- **Harlow-Hayden**: decoding the radiation is *exponentially complex* in the entropy of the black hole. **Aaronson:** decoding the radiation is polynomially hard $\Rightarrow$ there are no injective one-way functions.

A Proposal

Akers, NE, Harlow, Penington, Vardhan; see also earlier work starting with Harlow-Hayden

If you don't have access to exponentially complex information, you'd conclude the information isn't there!

Complexity as a Cutoff

Semiclassical gravity effective field theory accurately describes all the low-curvature physics that can be measured by any observer whose computational power is bounded to be subexponential in S_{BH}.

A Proposal

Akers, NE, Harlow, Penington, Vardhan; see also earlier work starting with Harlow-Hayden

If you don't have access to exponentially complex information, you'd conclude the information isn't there!

Complexity as a Cutoff

Semiclassical gravity effective field theory accurately describes all the low-curvature physics that can be measured by any observer whose computational power is bounded to be subexponential in S_{BH}. There **are** violations of EFT for exponential observables.

A Proposal

Akers, NE, Harlow, Penington, Vardhan; see also earlier work starting with Harlow-Hayden

If you don't have access to exponentially complex information, you'd conclude the information isn't there!

Complexity as a Cutoff

Semiclassical gravity effective field theory accurately describes all the low-curvature physics that can be measured by any observer whose computational power is bounded to be subexponential in S_{BH}. There **are** violations of EFT for exponential observables. These violations are critical for resolving the black hole information problem, but *they cannot be detected by any subexponential observer.*

Does this work?

We want to show that it is possible to resolve the tension between the three fundamental aspects of the black hole information paradox:

1. A unitary S-matrix;
2. A finite black hole entropy with a state-counting interpretation;
3. A black hole interior which is described to a good approximation by gravitational effective field theory.

By modifying the last bulletpoint so it becomes...

Does this work?

We want to show that it is possible to resolve the tension between the three fundamental aspects of the black hole information paradox:

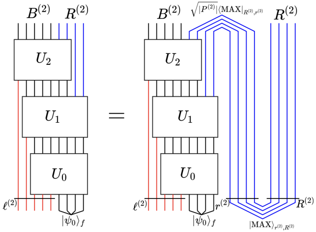
1. A unitary S-matrix;
2. A finite black hole entropy with a state-counting interpretation;
3. A black hole interior which is described to a good approximation by gravitational effective field theory **for subexponentially complex measurements.**

Tension Can be Resolved in a Tensor Network Model

Model the map between semiclassical gravity and QG as a non-isometric encoding from one Hilbert space to the next:

$$V : \mathcal{H}_{EFT} \to \mathcal{H}_{QG}$$

Can accommodate all of the requisite elements: there *are* significant violations of EFT at the event horizon of a black hole, but this is completely undetectable to a semiclassical observer.



In this model, the entropy of the radiation is computed by QESs.

The Non-Isometric Map

The map V annihilates a large number of “null states” that have no image in $\mathcal{H}_{QG}$, resolving the tension between the spuriously large Hilbert space of EFT and the much smaller $\mathcal{H}_{QG}$.

Fortunately – and surprisingly – this can be done without being detected by subexponential observers (due to measure concentration arguments).

Of course, even though the non-isometric nature of the map is hidden from an observer in the EFT, the null states are absolutely critical for this model to provide a toy model for a resolution of the black hole information paradox.

Overview

Summary

- The black hole information problem is a longstanding paradox in modern theoretical physics
- In a little over three years, we've come to understand how to use semiclassical gravity to make progress on it and propose a path towards a resolution
- We've found that QESs are an essential aspect of understanding how and when semiclassical gravity is valid.
- ... and we now have a proposal for resolving the paradox.

Thank you!