Title: Observers in quantum mechanics and quantum gravity (Vision Talk)

Speakers: Daniel Lord Harlow

Collection/Series: QIQG 2025

Subject: Quantum Gravity, Quantum Information

Date: June 27, 2025 - 10:00 AM

URL: https://pirsa.org/25060023

Quantum mechanics and observers for gravity in a closed universe

Daniel Harlow

MIT

June 27, 2025

Introduction

In this talk I will grapple with a rather puzzling issue that has come up in the literature on black holes and holography.

- We now have a fairly robust framework for thinking about the black hole information problem and the emergence of the black hole interior.
- The effective field theory of the interior is "non-isometricaly encoded" into the microstate degrees of freedom, leading to a Page curve that is consistent with unitarity via the quantum extremal surface formula.
- It is natural to see what we can learn about quantum cosmology by applying this idea to the universe as a whole.

Unfortunately doing so leads us to a rather shocking conclusion:

$$\dim \mathcal{H}_{closed} = 1$$

Quantum gravity in a closed universe has zero degrees of freedom!

How can such a seemingly absurd claim be true?

- Does this mean that we can experimentally rule out the possibility that we live in a finite universe?
- If not, how can we fit the richness of human experience into a one-dimensional Hilbert space?
- What is wrong with effective field theory in a closed universe, which certainly has more than zero degrees of freedom?
- Is there a mistake in the argument, and if so does it ruin the recent progress on black hole physics?

So far most people in the field seem to be hoping that the problem goes away, and for a while this included me, but one can only be an ostrich for so long.

My current feeling is that this may be an opportunity: to fix such a severe problem we need to do something crazy, and if that crazy thing is correct we will learn something deep about the laws of physics.

- Our main proposal will be that by explicitly including the observer as part of the universe, and forcing them to be classical by erasing their quantum coherence with the rest of the universe, we can recover the predictions of effective field theory up to corrections which are of order $e^{-S_{Ob}}$.
- I emphasize that for reasonable-size observers these corrections are large compared to the expected limits on the validity of effective field theory in cosmology (which are of order $e^{-\ell_{cos}^2/G}$), and in fact these observer-dependent corrections can (and must) persist in the limit $G \to 0$.
- One way to think about this is that we are introducing corrections to quantum mechanics which vanish in the limit of an infinitely big observer.

To spread the blame, I'll mention this is based on Harlow/Usatyuk/Zhao 2025.

Here is the plan for the rest of the talk:

- Arguments for a one-dimensional Hilbert space
- Review on observers and decoherence
- Including the observer as a resolution of $\dim \mathcal{H}_{closed} = 1$.

Arguments for $\dim \mathcal{H}_{closed} = 1$

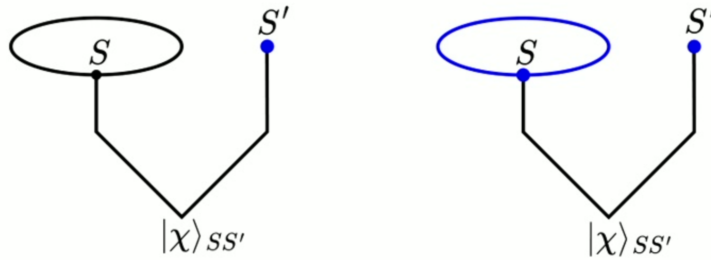
Before discussing the observer, I think it is worth spending some time reviewing some of the arguments that $\dim \mathcal{H}_{closed} = 1$. After all extraordinary claims require extraordinary evidence.

- Perhaps the most insightful argument, but also the least precise, is that in holography the fundamental degrees of freedom should live at the asymptotic spatial boundary. In a closed universe there is no spatial boundary, so there should be no fundamental degrees of freedom.
- Another argument worth mentioning is based on the idea that in quantum gravity there should be no "-1-form global symmetries" McNamara/Vafa 2021.

The most precise arguments however are based on the QES formula, the gravitational path integral, and the AdS/CFT correspondence.

Quantum extremal surface formula

As a first argument, let's consider entangling some degree of freedom S in a closed universe with some external reference system S' in a state $|\chi\rangle_{SS'}$ in the effective description:



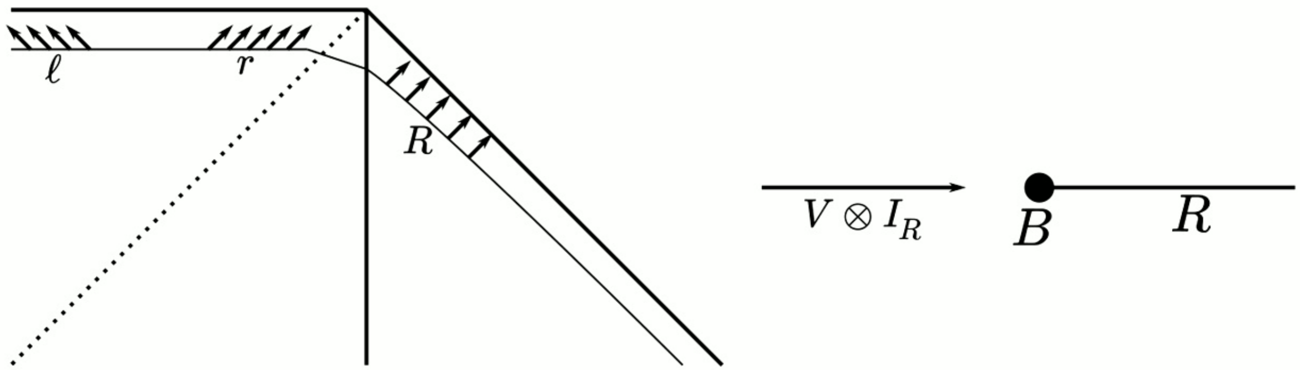
$$S_{gen} = S(\chi_{S'}) \qquad\qquad S_{gen} = 0$$

To compute the entropy of S' in the fundamental description, Engelhardt and Wall tell us to compute the generalized entropy of all possible *entanglement wedges* of S', which in this case are S' itself or S' together with the closed universe, and then choose the one which minimizes S_{gen}. In this case $S_{gen} = 0$ is an option, so it always wins and thus $S_{S'} = 0$. You can never be entangled with a closed universe! Almheiri/Mahajan/Maldacena/Zhao 2019.

Evaporating black hole

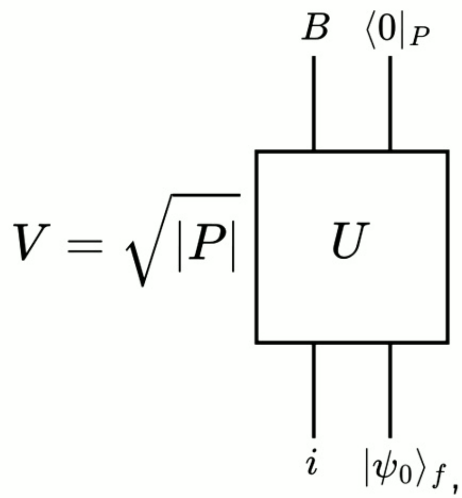
In any quantum system with emergent behavior, we can formulate this emergence mathematically in terms of a linear *encoding map*

$$V : \mathcal{H}_{eff} \to \mathcal{H}_{fund}.$$



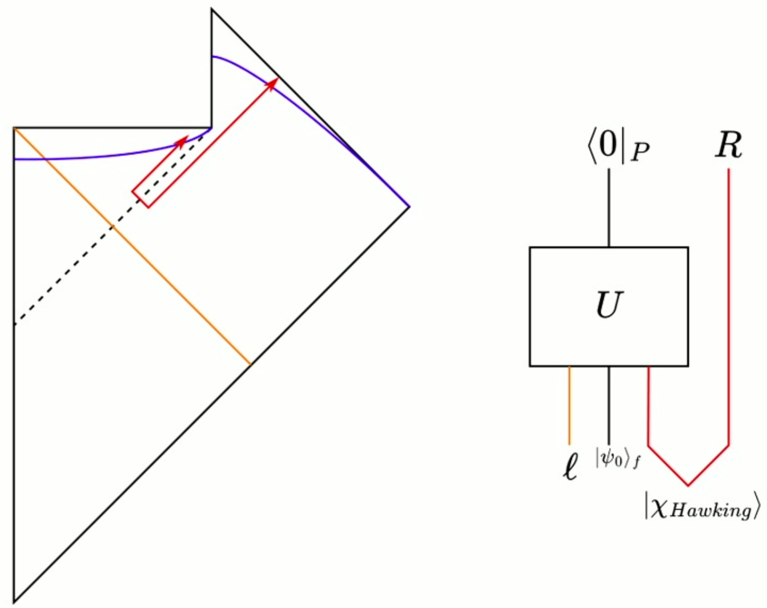
For an evaporating black hole we have an encoding map $V : \mathcal{H}_\ell \otimes \mathcal{H}_r \to \mathcal{H}_B$, where ℓ and r are interior modes and B is the microstate degrees of freedom. We also assume a trivial encoding map for the Hawking radiation R.

A concrete model of a non-isometric encoding map can be contructed using a random unitary matrix U:



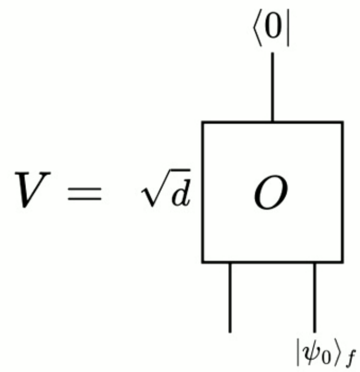
where we combined ℓ and r into i (for "interior" or "island"). This simple model can already reproduce the famous "Page curve" calculations of Penington, Almheiri/Engelhardt/Marolf/Maxfield 2019, and with some work one can show that generically the non-isometric nature of the code cannot be detected by any observer who can't perform operations whose complexity is exponential in the black hole entropy Akers/Engelhardt/Harlow/Penington/Vardhan 2022.

It is instructive to consider in particular the limit of a completely evaporated black hole:



In this limit the encoding map is a rank-one projection, and feeding in a pure state of the infalling degrees of freedom ℓ results in a pure state of the hawking radiation R (reducing to Horowitz/Maldacena 2003).
If we look at the interior alone, which is a closed universe, we have $\dim \mathcal{H}_{closed} = 1$!

Leaving behind the black hole and considering now a smooth closed universe (without a Planckian region at the evaporation point), we can model the encoding map like this:



We take O to be orthogonal to reflect the fact that in a closed universe $\mathcal{CPT}$ symmetry should be gauged, leading to a real Hilbert space Harlow/Numasawa 2023, see also Witten 2025.

This kind of encoding map clearly won't do a good job of preserving the quantum mechanics of the effective description, and we can quantify this by looking at what it does to the inner product on average in the Haar measure on $O(d)$:

$$\int dO \langle\phi|V^\dagger V|\psi\rangle = \langle\phi|\psi\rangle$$

$$\int dO |\langle\phi|V^\dagger V|\psi\rangle - \langle\phi|\psi\rangle|^2 = 1 + |\langle\phi|\psi\rangle|^2 + \mathcal{O}(1/d).$$

On average the inner product is preserved, but with large fluctuations in any particular choice of O. What can we do about this?
What I will try to argue in the remainder of the talk is that by careful treatment of the observer who is doing the measurements, we can suppress these fluctuations by $e^{-S_{Ob}}$.

Contact with the path integral

Before discussing the observer, I will briefly mention how we can reach similar conclusions using the gravitational path integral in a simple model of 2D gravity:

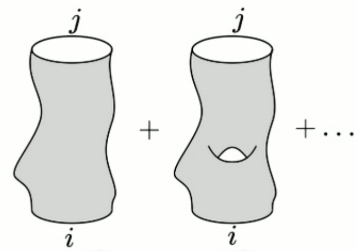
$$I_E[M] = -\frac{S_0}{4\pi}\int_M R = -S_0\chi[M],$$

with wordline matter carrying a species index i. Marolf/Maxfield 2019, Usatyuk/Wang/Zhao 2024

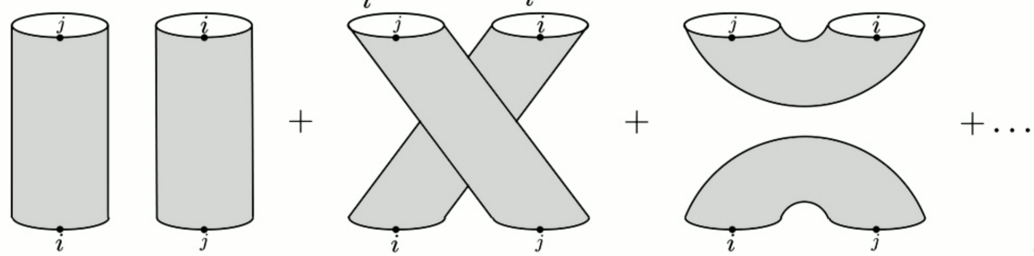
The "dynamics" for the matter index are simply that it must agree at the start and end of a worldline.

In this model we can compute the coarse-grained inner product and inner-product fluctuation using the path integral:

$$\overline{\langle j|i\rangle} =$$



$$\overline{|\langle j|i\rangle|^2} =$$



with results

$$\overline{\langle j|i\rangle} = \delta_{ij} + \mathcal{O}(e^{-2S_0})$$
$$\overline{|\langle j|i\rangle - \overline{\langle j|i\rangle}|^2} = 1 + \delta_{ij} + \mathcal{O}(e^{-2S_0}).$$

These precisely match our results from the code model with a random O, note in particular the $\mathcal{O}(1)$ deviation from preserving the inner product.

Including the observer

Let's now talk about observers. I will assume three things about them:

(1) An observer has a Hilbert space $\mathcal{H}_{Ob}$ with $\dim \mathcal{H}_{Ob} = e^{S_{Ob}}$, and they can only do experiments up to errors which are exponentially small in S_{Ob}.

(2) An observer is effectively classical, in the sense that there is a "pointer basis" of states $|a\rangle_{Ob}$ which are stable under interaction with their environment E:

$$\mathrm{Tr}_E \left(U_{int}^{\dagger} \left(|a\rangle\langle b|_{Ob} \otimes \frac{I_E}{|E|} \right) U_{int} \right) \approx \delta_{ab} |a\rangle\langle a|.$$

(3) The observer is entangled with their environment E in a state $|\omega\rangle_{Ob,E}$, with an entanglement entropy of order S_{Ob}.

See Zurek 2003 for more discussion of these assumptions.

This interaction with the environment is an example of what is called a *quantum-to-classical channel*, which in general acts in some basis as

$$\mathcal{C}(|a\rangle\langle b|) = \delta_{ab}|a\rangle\langle a|.$$

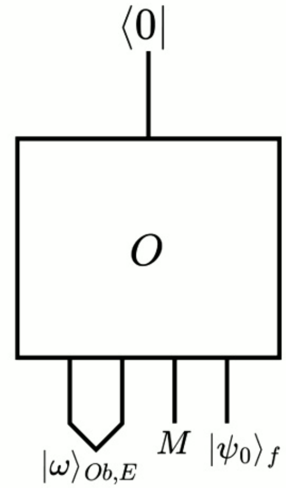
Essentially what the environment does is erase the off-diagonal elements of the density matrix in the pointer basis, which ensures that we can interpret the quantum evolution as inducing an evolution of classical probability distributions.

$$\sum_a C_a|a\rangle_{Ob} \to \sum_a C_a|a\rangle_{Ob'}|a\rangle_{Ob} \to \sum_a |C_a|^2|a\rangle\langle a|_{Ob}$$

A nice general way to implement this channel is by cloning in the pointer basis and then tracing out the clone.

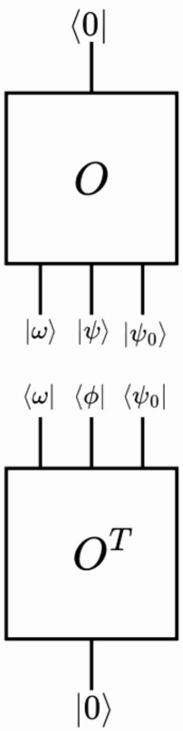
Let's now include the observer explicitly in our closed universe code:

$$V|\omega\rangle_{Ob,E} = \sqrt{d}$$



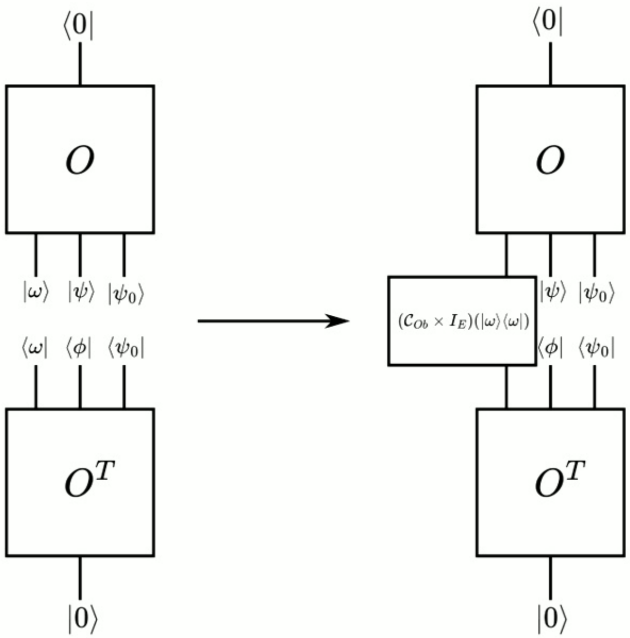
Here we have split the EFT degrees of freedom into three parts: the observer Ob, their environment E, and some "matter" degrees of freedom M that they wish to do experiments on. We have restricted the inputs to ensure that Ob and E are entangled in some fixed state $|\omega\rangle_{Ob,E}$.

Let's now compute the inner product between two states of this type:



This inner product on M is still rank one, so we still have a problem.

We claim that the problem with this inner product is that it still treats the observer as a fully quantum system, the same as the rest of the matter.



Our proposal is that in this inner product we need to act on *Ob* with a quantum-to-classical channel, in order to remove the coherence of their entanglement with their environment. In equations

$$\mathrm{Tr}\left(V|\omega\psi\rangle\langle\omega\psi|V^\dagger\right) \to \mathrm{Tr}\left(V\mathcal{C}_{Ob}(|\omega\psi\rangle\langle\omega\phi|)V^\dagger\right).$$

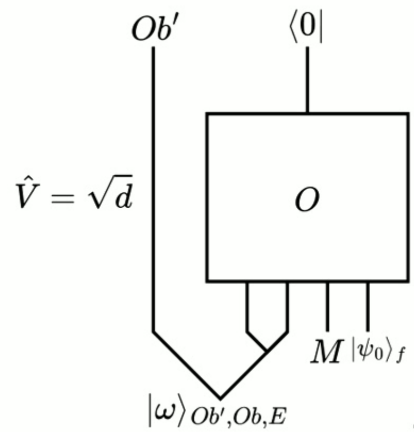
Some comments about this rule:

- This breaks the factorization of the inner product! So we now have a chance at a nontrivial inner product.
- Our explicit criterion for the successful emergence of the bulk spacetime is

$$\mathrm{Tr}\left(V\mathcal{C}_{Ob}(|\omega\psi\rangle\langle\omega\phi|)V^{\dagger}\right) \approx \langle\phi|\psi\rangle,$$

  with the approximation holding up to errors of order $e^{-S_{Ob}}$.
- You can think of the action by $\mathcal{C}_{Ob}$ as a mathematical realization of the "Heisenberg cut" between quantum and classical from the Copenhagen interpretation of quantum mechanics. Normally we say this decoherence is induced automatically by some external environment; the proposal here is that when you are talking about the whole universe the cut must be imposed by the laws of physics themselves.
- This inner product depends on which observer you use, so the laws of physics depend on the observer - a form of complemtarity (see Elliot's talk for more).

We can make the effect of this channel look more familiar by using the cloning construction to define a new encoding map:



.

In this presentation we define $\hat{V} : \mathcal{H}_M \to \mathcal{H}_{Ob'}$ by cloning the observer in their pointer basis and then acting with $I_{Ob'} \otimes V$, and our modified inner product is given by

$$\langle \phi | \hat{V}^\dagger \hat{V} | \psi \rangle.$$

By forcing the observer to be classical, we have "created" a fundamental Hilbert space Ob' for the system, and at least some quantum information from M can now be transmitted to Ob'.

Results

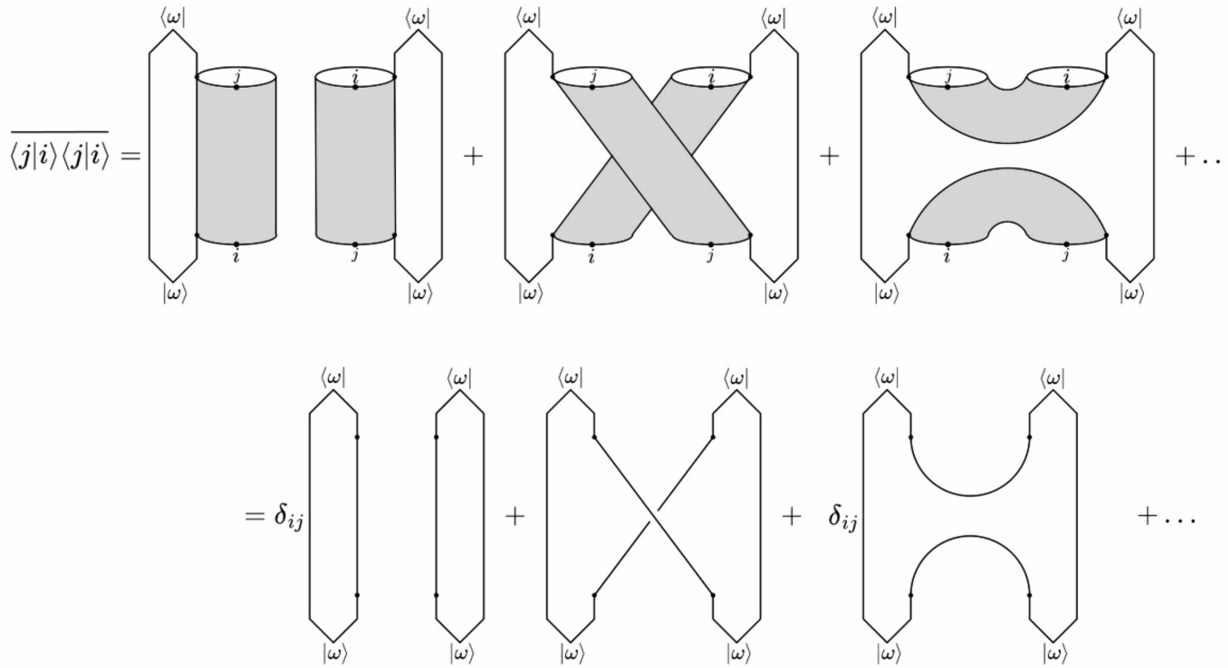
Let's now see how this rule changes our previous calculations:

$$\int dO \langle \phi | \hat{V}^\dagger \hat{V} | \psi \rangle = \langle \phi | \psi \rangle$$

$$\int dO \langle | \langle \phi | \hat{V}^\dagger \hat{V} | \psi \rangle - \langle \phi | \psi \rangle |^2 = e^{-S_2(\omega_{Ob})} \left(1 + | \langle \phi | \psi \rangle |^2 \right) + \mathcal{O}(1/d).$$

The large fluctuations in the inner product are now suppressed by $e^{-S_2(\omega_{Ob})}$, which by assumption is of order $e^{-S_{Ob}}$!

We can also apply this rule to the gravitational path integral in our topological model:



This gives

$$\overline{|\langle j|i\rangle - \overline{\langle j|i\rangle}|^2} = e^{-S_2(\omega_{Ob})}\left(1+\delta_{ij}\right) + \mathcal{O}(e^{-2S_0}),$$

just as we found in the code!

Discussion

There is a lot more I could say here, but I'll close with some general comments:

- The problematic terms in the inner product came from "cross connection" geometries in the gravitational path integral, and these were suppressed by the observer. If we treat the observer as an idealized worldline and take $S_{Ob} \to \infty$ to make them eternal then we get a simplified rule where we just ignore observer cross connections, see Abdalla/Antonini/Iliesiu/Levine, Akers/Bueller/DeWolfe/Higginbotham/Reinking/Rodriguez 2025.
- You might ask why we consider states other than $O^T|0\rangle$ as input, since that is the "unique" state of the theory. The idea here is we need to condition on the part of the wave function consistent with what we know about the universe before doing quantum mechanics.
- This rule is a modification of QM: in addition to knowing the state, to do physics we need to know V and pick an observer. I don't find this too bad, as anyways QM needs to be modified for finite observers.

Thanks!