Not the Measurement Problem's Problem: Black Hole Information Loss with Schrödinger's Cat

Saakshi Dulani, Johns Hopkins University

January 9, 2025

Abstract

Recently, several philosophers and physicists have increasingly noticed the hegemony of unitarity in the black hole information loss discourse and are challenging its legitimacy in the face of the measurement problem. They proclaim that embracing non-unitarity solves two paradoxes for the price of one. Though I share their distaste over the philosophical bias, I disagree with their strategy of still privileging certain interpretations of quantum theory. I argue that information-restoring solutions can be interpretation-neutral because the manifestation of non-unitarity in Hawking's original derivation is unrelated to what's found in collapse theories or generalized stochastic approaches, thereby decoupling the two puzzles.

Contents

1 Introduction

Hawking (1975,1976) shocked the physics world by declaring that the quantum mechanical treatment of black holes is notoriously un-quantum. Conventional quantum systems evolve unitarily, i.e., deterministically from pure-to-pure states. In contrast, black hole spacetimes with quantum matter fields evolve non-unitarily, i.e., indeterministically from pure-to-mixed states. The difference between pure and mixed states is that the former possess maximal information about superpositions, interference effects, nonlocal entanglement correlations, and probabilities, whereas the latter possess incomplete information about these attributes.

Around that time, Wald (1975) refined and extended Hawking's semi-classical calculation. He also corroborated Hawking's visual heuristic on why black holes change the game. Quantum matter fields deplete a black hole's gravitational energy through spontaneous particle creation, causing it to dissipate entirely. In such a fashion, solely the radiation earning Hawking's namesake lingers on. This evolution is non-unitary because the black hole interior is eventually traced out, permanently.[1] Pre-evaporation states are pure since both the interior and exterior regions are included in the description, encoding complete information about the entanglement structure across the event horizon. Post-evaporation states are mixed, however, since the black hole has disappeared, encoding incomplete information about the exterior region as a formerly entangled subsystem. To top it off, the event horizon prevents information about the

[1] For temporal claims, I'm referencing Hawking's 1975 foliation of an evaporation Penrose diagram and bracketing Wald's intuitions about retaining the black hole interior on disconnected Cauchy surfaces (see Maudlin 2017 and Manchak and Weatherall 2018).

black hole interior from escaping to the exterior spacetime. This is why the phenomenon came to be known as black hole information loss and deemed paradoxical by those who swear by standard quantum theory.

For all that Hawking and future generations have been critiqued on misappropriating information talk (see Belot et al. 1999; Maudlin 2017), the connection between non-unitarity and information loss resonates on a visceral level (see Dulani 2024 on recommendations to productively re-purpose that information talk). From the 1990s onwards, the reaction has overwhelmingly been to protect unitarity, and by extension, strict information conservation through whichever tricks of the trade imaginable (for literature reviews, see Page 2005; Mathur 2009; Unruh and Wald 2017; Polchinski 2017; Raju 2022). It's foreseeable then that the hegemony of unitarity would sow the seeds of a counterculture casting doubts on the discourse's legitimacy.

This skepticism is especially nontrivial when unitarity may not survive even in non-relativistic quantum theory under the shadow of the measurement problem. Okon and Sudarsky (2015) express optimism that progress can be made on both fronts by revitalizing dynamical collapse approaches:

> [T]he analysis of the black hole evaporation process changes dramatically once one accepts, at the fundamental quantum level, a departure from unitarity. If unitarity is universally broken, and information is generically lost, then the fact that black holes lose information stops being that surprising and problematic. (p. 466).

Beyond the measurement problem, Oppenheim (2023) surmises that unitarity is no match for the herculean task of reconciling quantum mechanics with general relativity.

Generalized stochastic theories hold promise for this project, and a unified theory is indispensable to advancing our knowledge of black hole evaporation.

> The fact that coupling classical gravity to quantum theory necessarily requires stochasticity is particularly compelling in light of the black-hole information problem. . . The class of theories presented here may allow us to resolve the black-hole information paradox in favor of information loss. (pp. 3-4)

My goal in the ensuing analysis is to investigate potential cross-fertilization between the black hole information loss paradox and the measurement problem. I argue that contrary to some high hopes, black hole information loss is not the measurement problem's problem. Schrödinger's cat may fall down an evaporating black hole, but no proposition about it being dead, alive, neither, or both, pertains to the evolution from pure black hole states to mixed radiation states.

I first present the "unsophisticated" black hole information loss paradox in section 2. I then anchor it in a "two-for-one deal" illustrating how non-unitarity in the black hole context provocatively overlaps with non-unitarity as a response to the measurement problem. However, in section 3, I expose the fallacy of elevating non-unitarity to the status of the holy grail. My assessment of a toy model combining black hole evaporation with dynamical collapse uncovers two sources of non-unitarity reflecting unprecedented kinematics versus modified dynamics. This distinction closes the door to potential cross-fertilization. I conclude in section 4 that a sophisticated resolution to information loss stays neutral about interpretations of quantum mechanics and, therefore, unbiased towards either determinism or indeterminism.

2 Unsophisticated Information Loss: Non-Unitarity

Proponents of non-unitary interpretations of quantum mechanics argue that their preferred solution to the measurement problem concurrently resolves the black hole information loss paradox (henceforth, BHILP). Their claim is predicated on the historically mainstream narrative of BHILP, commonly understood to simply be decrying global non-unitarity. However, as I will demonstrate, this oversimplified perspective obscures critical nuances. In order to navigate the debate and rectify misleading arguments, I've formulated what I contend to be the "unsophisticated" BHILP as the horns of a dilemma: (1) Quantum systems evolve unitarily; (2) Evaporating black holes evolve non-unitarily.

Premise (1) is definitional: A 'quantum' system just is one whose state evolves according to a unitary, fundamental law, like Schrödinger's equation in non-relativistic contexts or the Klein-Gordon equation for relativistic free fields. Despite controversy over the metaphysics of quantum states and their unitary evolution, I'm adopting a realist stance: Quantum states represent a system's present ontology as well as probabilities over future experimental outcomes. Furthermore, conditions for (1), such as the absence of external interactions (like measurement), are satisfied if the universe is treated as a closed system.

Conversely, (2) shows that black hole evaporation challenges the ostensibly lawful behavior of quantum systems. Despite the involvement of quantum matter fields in a semi-classical regime, a universe containing evaporating black holes defies unitary laws due to interference from global features of the spacetime. And this dilemma is difficult to dissolve because it stems from a self-undermining loop: Applying quantum theory to a black hole spacetime spells its own demise.

A noteworthy implication of this dilemma is the fate of information. I've articulated a third, corollary premise that's operative in BHILP: (3) If information is conserved from the initial to final state (and vice versa), then the evolution is unitary. Let's dissect this statement. States encode information by listing the values of degrees of freedom, as part of the kinematics of a theory. The dynamics, in contrast, link successive states into admissible histories of a system's evolution. Whether information is conserved rests on the laws. Susskind and Lindesay (2004) emphasize predictability and retrodictability. In the scattering matrix formalism, every unitary matrix is invertible, allowing the initial state to be reverse-generated from the final state. Maudlin (2017) also highlights the general import of determinism: information is conserved when "the value of the state at any time implies the value of the state at any other time" (p. 3).

So far, we've confronted two nuanced facets of information conservation under unitary dynamics: individual state specification and perfect retrievability between states. Together, these conditions entail a precise one-to-one mapping, where the input uniquely produces the output and vice versa. Maximal information is conserved when the inputs are individual pure states, although some relevant information is still conserved when the inputs are mixed states describing ensembles of pure states.

For one, perfect retrievability extends to symmetric many-to-many maps, where an ensemble of inputs uniquely produces an ensemble of outputs of equal size. This relationship between preserving ensemble size and conserving information, implied by unitarity, parallels the classical Liouville's theorem. Additionally, probability distributions in mixed states are conducive to uncertainty measures. If possibilities are weighted uniformly, larger ensembles increase uncertainty about the system's actual state, whereas smaller ensembles reduce it. This inverse correlation is recognized as "negentropy" (Brillouin, 2013). But even before, as a precursory step, identifying the full set of appropriate degrees of freedom is imperative to state construction. For instance, quantum degrees of freedom correspond to the coordinates of a complex Hilbert space. The total number of coordinates determines the dimensionality of that space's orthonormal basis – its maximum "information storage capacity" (Susskind and Lindesay, 2004).

Armed with this analysis, it should be apparent that shifting to non-unitary evolution prompts information loss. But which facets suffer? Indeterministic dynamics, expressed mathematically as non-invertible, asymmetric maps, sacrifice perfect retrievability and allow ensembles of states to change over time. Consequently, information about individual state specification is always lost, although information about ensemble size and uncertainty usually fluctuates. Degrees of freedom, conversely, are expected to be conserved, since it's typical for state values to evolve, not state variables. As it were, this assumption is at risk of being subverted, for reasons explored in section 3.

Suffice it to say from (3) that non-unitary black hole evaporation violates information conservation in some shape or form. But on its own, this conclusion is not paradoxical. Remember, the contradiction lies between (1), quantum unitarity, and (2), semi-classical non-unitarity. Therefore, dissolving the dilemma warrants grabbing one of the horns. A decision in favor of either horn belies one's antecedent beliefs about how information should behave in physics as well as one's metaphysical preferences about laws of nature.

Okon and Sudarsky (2014) observe how most proposals stress the implausibility of the second horn due to what Crowther (2018) identifies as shared guiding principles in the development and provisional acceptance of a final theory of quantum gravity.

Though unitarity in and of itself is revered as a guiding principle by a sector of the theoretical physics community, other guiding principles also converge on the recovery of unitarity as a serendipitous side effect.

The idea is that the semi-classical framework breaks down when it ventures into high-energy, short-distance Planckian territory. One candidate is the extremely warped interior region surrounding the central singularity. Another candidate is the near-horizon exterior region piling on infinitely blue-shifted Hawking modes. Hawking's calculation also falters as the evaporating black hole becomes inveritably small and radiates at infernally hot temperatures. These blow ups are anticipated to be smoothed over by a more fundamental theory, potentially recovering unitarity. New opportunities emerge if Planck-scale corrections get amplified at lower energies, disrupting the original derivation while the evaporating black hole is still macroscopic. Such modifications underscore the impending role of quantum gravity, placing string theory and loop quantum gravity at the forefront of proposals.

Opting to take the road less traveled, Okon and Sudarsky (2017) contemplate suggestive similarities between black hole evaporation and dynamical collapse. They consider non-unitary variants of quantum theory and re-evaluate the plausibility of the first horn of BHILP. Indeed, they realize that dismissing unitarity simultaneously rejects a premise in the measurement problem, inspiring them to embrace information loss.

The measurement problem (henceforth, MP) has long been formulated in philosophy of physics as the horns of a trilemma (adapted from Maudlin 1995): (1) Quantum systems evolve unitarily; (2) Quantum states are descriptively complete; (3) Measurement outcomes are determinate. Take Schrödinger's cat. It's frowned upon for our beloved feline to be trapped inside a closed box, hovering permanently in some mystical state superposing death and life. The remedy then is to open the box. Discovering at this moment that Schrödinger's cat is indeed dead or alive entails the falsity of at least one premise. Perhaps we obtained a determinate outcome upon opening the box (denial of (1)). Alternatively, our beloved feline could've been definitively dead or alive even when the box was closed (denial of (2)). Finally, we might've entered a superposition along with it upon opening the box, though presumably without noticing (denial of (3)).

I concede that this formulation of MP neither excludes measurement-induced collapse nor cleanly categorizes its alternatives.[2] Nonetheless, I'm maintaining Okon and Sudarsky's preferred formulation to engage with their argument. Before proceeding, it's worth considering whether the resources of hidden variables or Everettian approaches could explain away information loss. The answer is no, not unilaterally, as recovering unitarity in these interpretations relies primarily on novel, Planck-scale physics.[3] Yet interestingly enough, dynamical collapse and stochastic models are touted as exceptions. Okon and Sudarsky (2017) and Oppenheim (2023) are optimistic that non-unitarity enables the semi-classical framework to remain self-consistent in the limit of a more

[2]For instance, I'd classify generalized stochastic theories, akin to Oppenheim's proposal, as both non-unitary and hidden variable approaches based on the realist commitments outlined earlier. Yet some proponents argue that unitarity is retained instrumentally, as a predictive tool for aggregate statistics (see Barandes 2023).

[3]Let me address the suggestion of Bao et al. (2018) that Everettianism might successfully defuse the firewall paradox. This is not truly a counterexample for BHILP, as the firewall paradox arises only after recovering global unitarity, particularly by purifying Hawking radiation.

fundamental theory.

To that end, they promote a buy-one-get-one-free incentive. Okon and Sudarsky (2017) notice that both BHILP and MP hinge on unitary evolution, so they endorse relinquishing it to swiftly bypass both paradoxes. They profess that (1) in each is thereby the weakest link. It's fruitful to frame their reasoning as demystifying a particular instance by appealing to a universal phenomenon. Below, I distill a "two-for-one deal" in the spirit of their argument (also alluded to by Oppenheim 2023), with a more faithful reconstruction following in section 4. In essence, embracing non-unitary evolution for all quantum systems offers an elegant explanation of black hole evaporation, with information loss as a byproduct.

The Two-For-One Deal:

1. Quantum systems evolve non-unitarily.

2. Black holes are quantum systems.

3. Therefore, black holes evolve non-unitarily.

As cogent as their reasoning sounds, something has led us astray. Rejecting the same premise in two arguments doesn't make the revised premises mutually reinforcing. Moreover, we only procure two solutions for the price of one if 'non-unitarity' refers to a single phenomenon, which it doesn't. As I'll explain shortly, non-unitarity isn't confined to dynamical modifications of unitary laws; it can also reflect kinematic instabilities. Simply put, black hole evaporation isn't non-unitary in the same sense as collapse or stochastic models. Therefore, any version of a two-for-one deal is a specious argument. In the next section, I will evaluate a proposal that combines black hole evaporation with dynamical collapse to further expose the fallacy of non-unitarity as the holy grail.

3 The Fallacy of Non-Unitarity as the Holy Grail

Let's begin by examining the non-unitary factors behind dynamical collapse and black hole evaporation separately. According to the projection postulate behind dynamical collapse, a superposition state abruptly pivots to a determinate outcome, with relative frequencies over repeated instances obeying the Born rule. A single superposition state can collapse onto any one of its probable outcomes, manifesting a one-to-many map that fails to predict an exact future state. Conversely, multiple superposition states with overlapping outcomes but different probability assignments can collapse onto the same outcome, manifesting a many-to-one map that fails to retrodict an exact past state.[4] That said, the target system always actualizes a pure state, but indeterminism introduces and compounds uncertainty, so mixed states function as a convenient bookkeeping device for tracking ensembles of possible pure states.

Let's move on to black hole evaporation. The universe transitions from an initial pure state, describing the entangled interior and exterior regions, to a final mixed state, describing just the exterior region. Any perturbation to the initial state that leaves unchanged the black hole's mass, angular momentum, and charge repeatedly churns out the same final state of Hawking radiation. This is because Hawking radiation is thermal. Its energy spectrum depends only on temperature, which is a function of those macroscopic parameters. Crucially, more detailed information about the matter forming the black hole can't be retrieved from Hawking radiation. The past is thus washed away.

To accommodate this global time-irreversibility, Hawking (1976) invents the

[4]To be careful, dynamical collapse models can't project onto position eigenstates without violating energy conservation due to noncommutativity with momenum eigenstates.

“superscattering” operator. Unlike unitary scattering operators, superscattering operators can reliably implement pure-to-mixed transitions because of their ability to project across Hilbert subspaces without dimensional restrictions. Pure states correspond to one-dimensional vectors whose total number of coordinates coincides with the dimensionality of the overarching Hilbert space. Mixed states, however, correspond to collections of generic vectors spanning multi-dimensional subspaces embedded in the overarching Hilbert space. The application of superscattering operators in black hole evaporation also intimates an unorthodox, many-to-one mapping. We can replicate the pure-to-mixed transition for all compatible black hole states, sending each one to an identical radiation state. Thus, the initial Hilbert subspace signifies an ensemble of possible pure states, as opposed to the final Hilbert subspace, which reveals a unique and new type of quantum state. The final mixed state is thereby the most complete and objective representation of the post-evaporation universe, having formerly been an entangled region.

Now, how would black hole evaporation unfold if non-unitarity were a universal quantum phenomenon? Penrose (1981) proposes that spacetime curvature induces spontaneous dynamical collapse, in which the probability of collapse per unit time couples with the magnitude of curvature. The stronger the curvature and gravitational field – say, close to the singularity of a black hole – the higher the chance of collapse. Thus, the numerous collapse events during black hole evaporation purportedly explain global non-unitary evolution, given that both processes employ many-to-one maps.

Nevertheless, there’s a telltale dissimilarity between both processes. Unlike black hole evaporation, dynamical collapse admits of one-to-many maps as well. And because projections are apathetic to any temporal orientation, it may seem that the fate of information actually fares worse under dynamical collapse models, considering that individual state specification is rendered futile for prediction and retrodiction alike.[5]

However, Okon and Sudarsky (2014) take a leaf out of Penrose's book and assure us that the fate of information fares significantly better. Their rationale is that the one-to-many maps of dynamical collapse compensate for the many-to-one maps of black hole evaporation. Over extended time periods, approximately symmetric many-to-many maps actually conserve information about ensemble size and uncertainty, albeit effectively. Although they contest unitarity with the purpose of jointly evading MP and BHILP, they attenuate the severity of information loss by offering effective information conservation as a suitable substitute.

Yet, a couple of thorny issues emerge. Dynamical collapses compensate for one-to-many maps with their own many-to-one maps even without black hole evaporation. Additionally, non-unitary maps in this context stochastically project one pure state onto another pure state. A mixed state here designates an ensemble of pure states with built-in uncertainty. But in black hole evaporation, mixed radiation states don't designate ensembles of pure states, since they're nondegenerate. So, the non-unitary maps of black hole evaporation project an ensemble of pure states onto an objectively mixed state. The respective connotations of 'mixed state' thus don't align.

Therefore, Penrose and, to some extent, Okon and Sudarsky appear to be conflating

[5]'Temporal orientation' doesn't imply a metaphysically-robust direction of time, like the Second Law of Thermodynamics. Any relationship between the Second Law and dynamical collapse (a compelling motivation for Penrose 2004 and his Weyl Curvature Hypothesis) concerns coarse-grained (macrolevel) information loss, not the fine-grained (microlevel) information loss being discussed here.

disparate sources of non-unitarity. The species of non-unitarity pertinent to MP directly modifies the equations of motion such that it guarantees determinate outcomes. On the contrary, the species of non-unitarity pertinent to evaporating black holes doesn't touch the equations of motion. It's an implication of black holes as regions of no escape until the very end of their lives.

Black hole information loss is better understood as a kinematic consequence of a topological discontinuity (see Belot et al. 1999). At the outset, elements of the ontology crossing the event horizon are doomed to annihilation by the singular curvature blow up. Besides that, reaching the singularity coincides with complete evaporation. Nothing that went inside the black hole and sealed its fate will ever be resurrected. So, global spacetime structure is – quite literally – throwing a curveball. The disappearance of the black hole interior, including the spacetime region itself, consequently triggers a non-unitary interruption.

This caustic disruption ends up eliminating degrees of freedom. The initial Hilbert space factorizes into entangled subspaces, the black hole interior and exterior: $\mathcal{H}_i = H_{int} \otimes H_{ext}$. After evaporation, however, the interior subspace vanishes, reducing the final Hilbert space through a partial trace: $\mathcal{H}_f = H_{ext}$. Thus, $\mathcal{H}_f$ is smaller than $\mathcal{H}_i$, in terms of its span. Its dimensionality is lower if $\mathcal{H}_i$ is finite. Or, if $\mathcal{H}_i$ is infinite and $\mathcal{H}_f$ remains so, it still is but a subspace of $\mathcal{H}_i$.[6] Unsurprisingly then, we forego maps between pre- and post-evaporation states that are even remotely symmetric when a subset of degrees of freedom is removed from the total quantum description.

This scenario – of time-irreversible kinematic instabilities – directly undermines

[6] See Ruetsche 2011 and Susskind and Lindesay 2004 for a review of (in)finite-dimensional Fock space and Planck-scale cutoffs.

effective information conservation. For counterbalancing to occur between many-to-one and one-to-many maps, the overarching Hilbert space must be held fixed. Recall from section 2 that stabilization is imperative to construct comparable states and meaningfully track ensemble size. Ensemble size here is captured by the dimensionality of the Hilbert *subspace* associated with possible pure states. This subspace shrinks when projections temporarily screen off degrees of freedom, and it grows when they become accessible later. Because dynamical collapse doesn't distort the overarching Hilbert space, the full set of degrees of freedom is always fair game. In contrast, black hole evaporation permanently distorts the overarching Hilbert space. A partial trace eliminates degrees of freedom, rendering them inaccessible. Therefore, these two types of maps – screening versus eliminating – are mathematically incommensurate. Any idea of net compensation between them is untenable.

It's intriguing to consider whether a weaker constraint could suffice for effective information conservation. Instead of preserving the original degrees of freedom, perhaps we could just fix the dimensionality of the overarching Hilbert space. As mentioned earlier, $\mathcal{H}$ might remain infinite even as degrees of freedom are eliminated. Alternatively, if $\mathcal{H}$ is finite, its dimensionality might be continually refreshed through processes that create and destroy degrees of freedom at an equal rate. Note, however, that dynamical collapse plays no role in either explanation. It neither contributes to the total dimensionality nor unilaterally accounts for global non-unitarity. Furthermore, kinematic instabilities would still lead to the irrevocable loss of information about which degrees of freedom participate in state construction throughout the system's history.

In light of dynamical collapse being irrelevant, one can now see that the two-for-one-deal is incapable of resolving BHILP. The crux of this paradox is that information-bearing degrees of freedom get obliterated. Suggesting otherwise invites the absurd claims that black hole evaporation is non-unitary in virtue of Schrödinger's cat entering a definitively dead or alive state, and moreover, that this collapse event generates compensatory information for the vanishing interior. The reality is that once our beloved feline crosses the event horizon, no amount of collapse events can redress the inevitable and intractable information loss. It's thus prudent to ask: Is this situation physically reasonable, and if not, how should we alleviate the pathology?

By identifying the core question, we grasp how Penrose's and related proposals are question-begging. For dynamical collapse or generalized stochasticity to exhaustively account for global non-unitarity is to presuppose that degrees of freedom are conserved from $\mathcal{H}_i$ to $\mathcal{H}_f$. Without further justification, the two-for-one deal appears to implicitly conjure up some sort of remnant strategy, such that $\mathcal{H}_f = H_{int} \otimes H_{ext}$.[7] The persistence of the black hole interior inherently modifies the semi-classical framework at the Planck scale, through singularity resolution or other novel physics prior to evaporation (see e.g., Hossenfelder and Smolin 2010). Ultimately, black hole information loss is more accurately construed as a kinematic problem demanding a kinematic solution.

[7] I'm basing this inference on Oppenheim's 2023 proposal and personal conversations with Sudarsky. But an alternative to remnants – holography – refutes the tensor product in $\mathcal{H}_i$ to avoid over-counting degrees of freedom when the black hole is present (see e.g., Susskind and Lindesay 2004; Raju 2022).

4 Sophisticated Information Conservation: A Kinematic Solution

At this juncture, I've elucidated how the proposed interface between MP and BHILP is flawed. Before illuminating paths forward, I wish to preempt resistance by Okon and Sudarsky (2015). Departing from Penrose's program, they aspire to invert the explanatory direction of the two-for-one deal by attributing non-unitarity across quantum systems to black hole evaporation. In this view, Schrödinger's cat collapses onto a definitively dead or alive state in virtue of microscopic, virtual black holes,[8] whose cumulative evaporation catalyzes ubiquitous, mundane collapse events. Their reimagined two-for-one deal proceeds as follows.

The New Two-For-One Deal:

1. Black holes evolve non-unitarily.
2. The evolution of quantum systems is mediated by virtual black holes.
3. Therefore, quantum systems evolve non-unitarily.

However, my prior objection still stands. It's opaque how the dynamical source of non-unitarity could arise from the kinematic source, especially when post-collapse states are pure while post-evaporation states are mixed. More importantly, Okon and Sudarsky would have to renounce Penrose's notion of effective information conservation. No quantum system could sustain a fixed Hilbert space throughout its evolution, since degrees of freedom would be eliminated even during routine particle collisions.

[8]Virtual processes are weighted in the path integral method/sum over histories approach to calculate scattering amplitudes.

Given these conflicts, one might insist on forbidding pathological black hole evolution without abandoning dynamical collapse or generalized stochasticity as solutions to MP. Thus, decoupling the two puzzles is optimal, which also facilitates dropping the prefix 'un' in the unsophisticated BHILP. In order to make progress, we'd benefit from deconstructing unitarity as a multi-pronged constraint so as to clarify nuances in violations. In this case study, I've already distinguished between two types of violations: those that vary ensemble size (driven by dynamics) and those that vary degrees of freedom (driven by kinematics). In future work, I aim to refine this distinction and integrate entanglement by deploying the information-theoretic machinery of entropy (see also, Dulani 2024). Furthermore, this recourse broadens the operative notion of information conservation. Despite my criticisms of the two-for-one deal, I support the insight of Penrose, Okon, and Sudarsky that approximately preserving ensemble size, as seen in dynamical collapse, should still count as information conservation.

The beauty of revealing distinct species of non-unitarity is that we learn why it was naive to articulate BHILP as a vague plea for unitarity in the first place. This confusion led to its misguided conflation with MP. A sophisticated treatment retains interpretation-neutrality by staying agnostic about the equations of motion, and by extension, metaphysical preferences regarding deterministic versus indeterministic laws of nature. So in the end, black hole information loss with Schrödinger's cat is not, and never has been, the measurement problem's problem.

5 Conclusion

My primary motivation for disambiguating the black hole information loss paradox from the measurement problem has been to loosen the grip of unitarity on the discourse. It's better for everyone to frame BHILP as logically independent to promote plurality and inclusion, especially since it should handle proposals notwithstanding their interpretation of quantum mechanics. I've arrived at this conclusion by discerning between kinematic and dynamical sources of non-unitarity and relegating them to their respective puzzle. I contend that pursuit-worthy proposals at this stage should prioritize alleviating the kinematic tensions of black hole evaporation by conserving degrees of freedom. This strategy accommodates both deterministic and indeterministic dynamics, as well as strict and effective information conservation. Consequently, any pursuit-worthy proposal can be adapted to reflect progress on MP without presupposing the ideal interpretation of quantum mechanics. Mitigating the risk of premature adjudication has been, after all, the call to action by Okon and Sudarsky (2017).

Acknowledgments

I'm very grateful to numerous mentors and colleagues who have helped me sharpen and polish the arguments of this paper, leading up to and during the 2024 Biennial PSA Conference. First, I'd like to thank Daniel Sudarsky for conversations over the years that unbeknownst to him, inspired this paper. I'd also like to acknowledge Jacob Barandes, Sam Bysh, Sean Carroll, Nick Huggett, Lucy James, Emilia Margoni, Elizabeth Miller, Dominic Ryder, and David Wallace for their invaluable feedback. Finally, this paper significantly built on my dissertation research, which was graciously supported and funded by the U.S. Fulbright Program, Swiss Excellence Scholarship, Cogito Foundation, and John Templeton Foundation (as part of the Beyond Spacetime Project). The views expressed in this paper do not belong to any of the aforementioned funding bodies.

References

Bao, N., S. M. Carroll, A. Chatwin-Davies, J. Pollack, and G. N. Remmen (2018, June). Branches of the black hole wave function need not contain firewalls. *Physical Review D 97*(12).

Barandes, J. A. (2023). The stochastic-quantum theorem.

Belot, G., J. Earman, and L. Ruetsche (1999). The Hawking information loss paradox: The anatomy of controversy. *British Journal for the Philosophy of Science 50*(2), 189–229.

Brillouin, L. (2013). *Science and Information Theory* (2 ed.). Dover Publications, Inc.

Crowther, K. (2018). Defining a crisis: the roles of principles in the search for a theory of quantum gravity. *Synthese*.

Dulani, S. (2024). *Black Hole Paradoxes: A Unified Framework for Information Loss.* Ph. D. thesis, University of Geneva.

Hawking, S. W. (1975). Particle creation in black holes. *Communications in Mathematical Physics 43*(6), 199–220.

Hawking, S. W. (1976, November). Breakdown of predictability in gravitational collapse. *Phys. Rev. D 14*, 2460–2473.

Hossenfelder, S. and L. Smolin (2010, March). Conservative solutions to the black hole information problem. *Physical Review D 81*(6).

Manchak, J. and J. O. Weatherall (2018). (Information) paradox regained? A brief comment on Maudlin on black hole information loss. *Foundations of Physics 48*(6), 611–627.

Mathur, S. D. (2009). The information paradox: a pedagogical introduction. *Classical and Quantum Gravity 26*(22), 224001.

Maudlin, T. (1995). Three measurement problems. *Topoi 14*(1), 7–15.

Maudlin, T. (2017). (Information) paradox lost.

Okon, E. and D. Sudarsky (2014). Benefits of objective collapse models for cosmology and quantum gravity. *Foundations of Physics 44*(2), 114–143.

Okon, E. and D. Sudarsky (2015). The black hole information paradox and the collapse of the wave function. *Foundations of Physics 45*(4), 461–470.

Okon, E. and D. Sudarsky (2017, January). Black holes, information loss and the measurement problem. *Foundations of Physics 47*(1), 120–131.

Oppenheim, J. (2023, December). A postquantum theory of classical gravity? *Physics Review X 13*, 041040.

Page, D. N. (2005, September). Hawking radiation and black hole thermodynamics. *New Journal of Physics 7*, 203–203.

Penrose, R. (1981). Time asymmetry ad quantum gravity. In C. J. Isham, R. Penrose, and D. W. Sciamma (Eds.), *Quantum Gravity II*, pp. 244–272.

Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe*. Random House.

Polchinski, J. (2017). The black hole information problem. In *New Frontiers in Fields and Strings*, Chapter 6, pp. 353–397.

Raju, S. (2022). Lessons from the information paradox. *Physics Reports 943*, 1–80. Lessons from the information paradox.

Ruetsche, L. (2011). *Interpreting Quantum Theories: The Art of the Possible*. Oxford University Press.

Susskind, L. and J. Lindesay (2004). *An Introduction to Black Holes, Information and the String Theory Revolution: The Holographic Universe*. World Scientific Publishing Co Pte Ltd.

Unruh, W. G. and R. M. Wald (2017). Information loss. *Reports on Progress in Physics 80*(9), 092002.

Wald, R. M. (1975). On particle creation by black holes. *Communications in mathematical physics 45*(1), 9–34.