An ensemble-based deterministic and unitary theory of quantum measurements: a case study from Jackiw-Teitelboim gravity as an ensemble theory

Minseong Kim

(Dated: November 5, 2025)

Abstract

A unitary and deterministic theory of quantum measurements is put forward, inspired by the JT/RMT (Jackiw-Teitelboim / random matrix theory) duality, based on the distinction between an ensemble theory and an actual theory. An ensemble theory seeks to probe average behaviors of randomly given actual theories (or Hamiltonians). In ordinary contexts, an ensemble state approximates an actual state sufficiently, but this approximate equality breaks down in some contexts such as black hole evaporation. With the chaotic system assumption and the final mixed ensemble state of some system being maximally mixed while an actual state is guaranteed to be a pure state, a derivation similar to the one in the Zurek's envariance article recovers the Born rule, providing a unitary and deterministic account of quantum measurements. Some possible misconceptions and controversies surrounding JT/RMT and replica wormholes, in particular (non-)unitarity, are analyzed and clarified.

Keywords: JT/RMT, quantum measurement theory, replica wormhole, envariance, quantum gravity, black hole, JT/SYK

CONTENTS

I. INTRODUCTION

Suppose that the Hamiltonian of the universe is given probabilistically, and an ensemble theory, which we assume to be a unitary quantum theory as well, provides the evolution of an ensemble state. An ensemble reflects the average picture of the results of each theory's evolution. In normal circumstances, it is safe to assume that an ensemble state is approximately the same as an actual state. With an initial state equivalent in both an ensemble theory and an actual theory, can we reproduce the Born rule with some setup? In this paper, the question is answered yes with the derivation similar to the Zurek's envariance paper [1]. Therefore, a deterministic and unitary theory of quantum measurements is provided, with some caveats discussed in the final conclusion.

An implicit assumption behind the conclusion is that an ensemble theory is valid because there are degrees of freedom missing in an effective theory in use. One example of this would be the JT/RMT (Jackiw-Teitelboim gravity/random matrix theory) duality [2], which is heavily exploited in this paper. In case of JT/RMT, ensemble states themselves are largely considered unimportant and focus is typically on general properties of an ensemble - for example, what would be the (average) entanglement entropy of actual states in an ensemble? Within this emphasis, JT with modifications to be described in this paper (especially in Section III, represented by replica wormholes) is a different but equivalent expression of random matrix theory and vice versa.

However, that JT is an ensemble theory in JT/RMT leads to complications in interpretations. First, in both JT and RMT with the right density function, we have the same Schwarzian action. Therefore, unmodified JT should be dual to RMT as well in some form. Furthermore, as we will see in Section III, unitarity of 'modifications' to JT is implausible, and this has already been pointed out in the early days of the duality [3] - JT with modifications cannot be a unitary quantum theory. And there is nothing wrong with unmodified JT in JT/RMT if we ask the right question - ask 'what are the properties of ensemble states' then there is always the same one-boundary partition function $Z(\beta)$ such that we have the same unmodified theory, whether we work in JT or RMT with the right density function. If we ask different questions such as 'what are the average properties of individual states in an ensemble' then we need additional contributions to conventional JT calculations but they are not genuine non-perturbative contributions of JT. The reason why these new contributions arise can be explained with a simple statistics (or probability theory) insight in Section III A.

Having established (or rectified) the conceptual issues regarding (the differences between) an ensemble theory (or state) and an actual theory (or state), the lessons of replica wormholes in black hole evaporation can be summarized as follows. Assume that a black hole is a chaotic system. An initial state is approximately the same whether in an ensemble theory or in an actual theory. As evaporation continues, this approximate equality breaks down, and the final ensemble state of the black hole is maximally mixed. However, the final actual state of the black hole is a pure state, though there are different possible pure states consistent with the ensemble state. This arises because effective bulk field theory is missing some degrees of freedom in quantum gravity and therefore probabilistic treatments are necessary.

With an arbitrary initial state and a final maximally mixed (ensemble) state along with unitarity of an ensemble theory and an actual theory (and the chaotic system assumption), the Born rule can be recovered with the derivation similar to [1]. Despite being inspired from black hole physics, an ensemble theory in the quantum measurement theory of this paper does not have to be JT, nor does a system has to be a black hole, as long as the required conditions are satisfied.

Section II discusses the setup of the quantum measurement theory of this paper, similar to that of [3], one of the papers where a replica wormhole was introduced. However, as aforementioned, this JT setup is strictly not necessary and a variety of setups may be feasible.

Section III discusses the interpretational issues for replica wormholes and JT/RMT (in particular unitarity) and why replica wormholes are yet to be universally accepted as being a part of the resolution toward the black hole information problem. Nevertheless, once interpreted within the ensemble interpretation, replica wormholes themselves are not problematic at all. As long as the ensemble interpretation is correctly understood, Section III is all about black hole physics contexts and may be skipped. (Interested readers may find Section III C 1 to be a simple proof of non-unitarity of replica wormhole corrections though again with caveats, and Section III D to be about a potential non-firewall alternative to large bulk EFT corrections that cannot be avoided in a JT gravity approach to black hole evaporation.)

Section IV does the actual exposition and the derivation of the quantum measurement theory of this paper, with the calculations being basically the adapted form of [1].

Note that the claim is not that this measurement process is the actual measurement process, and more works are necessary to completely resolve the issue, especially given the caveats discussed in the final conclusion. The paper only takes a small step in the direction of progress.

II. SETUP

The setup of [3] is almost exactly followed. Explanations regarding replica wormholes are given in Section III. In the end, replica wormholes only serve to set up the case where an ensemble state is given by a maximally mixed state (resembling the Hawking picture of black hole evaporation) while an individual microstate (an actual and realized state) is a pure state, and details regarding replica wormholes are not terribly important.

As long as a similar ensemble theory setup exists, the unitary quantum measurement model of this paper can work without exclusive reliance on Jackiw-Teitelbom (JT) gravity, which is only assumed for the sake of brevity. Quantum gravity plays a minimal role, serving only as an example in which effective theories do not reflect all degrees of freedom.

The universe consists of two systems (or three systems) : R (exterior radiations or the system to be eventually measured) and B (black hole and the end-of-the-world brane together). It is assumed that the brane configuration k, which is represented by the state $|\psi_k\rangle$ of B, and the state $|k\rangle$ of R are entangled - therefore $|k\rangle$ can be used as a substitute for identifying end-of-the-world (EOW) brane configuration k.

In [3], the initial brane configuration is not directly mentioned, because it is unnecessary to do so. In this paper, we will identify the initial brane configuration with the following state of the universe:

$$|\Psi_i\rangle = \sum_{k=0}^{N-1} a_k |\psi_k\rangle |k\rangle \tag{1}$$

where N roughly refers to the number of possible brane configurations. Note that this initial state will eventually be modified to the state given in Equation (20) of Section IV. This choice is made for explanatory reasons, but it will be straightforward once Section IV is reached.

Now define the Euclidean action of B as:

$$I = I_{JT} + \mu \int_{brane} ds$$
$$I_{JT} = -\frac{S_0}{4\pi}\left[\int_{\mathcal{M}} \sqrt{g}R + 2\int_{\partial\mathcal{M}} \sqrt{h}K\right] - \frac{1}{2}\left[\int_{\mathcal{M}} \sqrt{g}\Phi(R+2) + 2\int_{\partial\mathcal{M}} \sqrt{h}\Phi K\right] \tag{2}$$

where $\int ds$ refers to the integral over the worldline of the brane, and the action is the same as in [3], noting that the last term has $\sqrt{h}\phi K$ instead of $\sqrt{h}\phi(K-1)$. The boundary conditions are given as follows - at the standard asymptotic boundary:

$$ds^2|_{\partial M} = \frac{1}{\epsilon^2}d\tau^2,\ \phi = \frac{1}{\epsilon}, \quad \epsilon \to 0 \tag{3}$$

where τ is the imaginary time coordinate. At the EOW brane, the boundary condition goes as:

$$\partial_n \phi = \mu,\ K = 0 \tag{4}$$

where n in this context (only) refers to the normal to the EOW brane, with $\mu \geq 0$. This completes the setup.

The following result of [3] (which is universally accepted) is central to the analysis of this paper. Regardless of $|\Psi_i\rangle$, if not for replica wormholes (and islands), the entanglement entropy of B (or equivalently R) will reach $\log N$, the maximally mixed state for R. This is eventually identified as the entanglement entropy of R from the point of an *ensemble* state, where we lack knowledge of an actual theory governing microstates. This distinction between an ensemble state and an actual realized state is reviewed in Section III in the sense of the JT/RMT duality [2].

By contrast, as stated in [3], the actual entanglement entropy of R eventually becomes zero with replica wormhole corrections. This implies that the actual final outcome of R is a pure state.

Unitarity preserves information (and the ensemble theory is assumed to be a unitary quantum theory as is for JT gravity), so a_k in Equation (1) has to appear somewhere, while also being consistent with the ensemble state being maximally mixed. This is analyzed in Section IV, which exploits the fact that this is just the envariance setup presented in [1] where the Born rule is derived. This results in a unitary and deterministic model of quantum measurements when viewed from the actual theory point of view.

III. WHY ARE REPLICA WORMHOLES CONTROVERSIAL IN BLACK HOLE PHYSICS?

This section is not essential to the quantum measurement model of this paper, as long as we recognize the following points: both the uncorrected bulk gravity theory without replica wormholes and the corrected theory with replica wormholes are mutually consistent, with each addressing different questions - evolution of an ensemble state of mutually consistent microstates under ignorance of fine-grained quantum gravity details versus average (ensemble) properties of individual microstates, such as (late-time) correlation functions and (late-time) entanglement entropy.

In less gravity-oriented terms, there are differences between an ensemble theory of random Hamiltonians and an actual theory represented by each random Hamiltonian. These theories probe different questions. Corrections to the ensemble theory arise when we are addressing the questions that are supposed to be computed within actual theories (and then averages are taken) but computational tractability leads us to compute in the ensemble theory.

Nevertheless, this section is written to clear up some controversies and confusions regarding replica wormholes in the simplest way possible, since the actual issues are largely conceptual and semantics, not mathematical. Therefore, it is not a substitute for a technical review or overview, which would require a separate full-length paper.

Playing devil's advocates, replica wormholes would largely be approached from the critics point of view, though with sufficient references to the actual arguments of the proponents. The end result is quite subtle to interpret, and one can only try their best to be free of further controversies. The emphasis is placed on the notion that replica wormholes themselves are unproblematic, supported by calculations, and it is only their interpretations and their place on realistic black holes that are often controversial.

The arguments proceed from Section III A, which provides all the intuitions we need - that the necessity of replica wormholes can be understood with basic statistics. Then Section III B suggests that replica wormholes are not genuine non-perturbative contributions (or consequences) of JT but rather are corrections to JT to probe fine-grained details missing in JT, with uncorrected JT being a theory of ensemble states. Then using the 'chaotic system' assumption, Section III B 2 suggests that it is infeasible for an ensemble state to approximate a microstate at late times and therefore ensemble state-wise replica wormhole corrections are implausible (in other words, replica wormholes are not genuine implications of a unitary ensemble theory, and they only arise because we are probing average properties of actual states instead of ensemble state properties).

A. An intuition-building example: an ensemble of pure states

Consider an ensemble state where we have all pure states with an equal probability:

$$\rho_{A,ens,ex1} = \frac{1}{n} \sum_i |\psi_{A,act,i}\rangle\langle\psi_{A,act,i}| \tag{5}$$

where $|\psi_{A,act,i}\rangle$ is a pure state that is an actual microstate of some subregion A. $\rho_{A,ens,ex1}$ has every microstate of subregion A with equal probability of $1/n$ and is a mixed state, but this does not mean that the actual microstate is a mixed state - instead, an actual and realized microstate is a pure state. Therefore, the (ensemble) average of the entanglement entropy of A is zero.

To compute property quantities like this, it is clear that if we are to compute inside the ensemble bulk theory, the conventional path integral method would give us a wrong answer, since the resulting answer would correspond to the property of $\rho_{A,ens,ex1}$, not an (ensemble) average of some property of $|\psi_{A,act,i}\rangle$. Replica wormholes provide corrections such that desired quantities can properly be computed.

B. Common confusions that muddle the interpretational analysis

There are confusions that often muddle the interpretational question for replica wormholes, especially when existing derivations and calculations are referenced in support of replica wormholes. It is true that for each RMT (random matrix theory) calculation (with the appropriate density function), there is a corresponding JT theory calculation, and some of these RMT calculations lead to path integrals in JT theory that involve geometries previously unaccounted for. Therefore, as far as JT is considered as a dual holographic theory of (random ensemble) RMT, the new contributions are 'genuine' non-perturbative additions *to* (as opposed to *of*) JT.

However, this does not directly imply that the uncorrected JT theory *requires* these 'non-perturbative' contributions outside of holography. JT works perfectly as a standalone theory of gravity outside RMT, and when discussed as a non-ensemble bulk theory, these contributions are not needed. The next sections confirm this point. Furthermore, since RMT is set to match the same Schwarzian action as an effective action with JT, the uncorrected JT theory also has some dual to RMT [2].

1. *Penington et al. (2022) (though it should really be 2019)*

In [3], Section 2.3 'Factorization and averaging' confirms that replica wormholes are to be understood as consequences of ensemble averaging and provides the following analysis:

$$\langle \psi_i | \psi_j \rangle = \delta_{ij} + e^{-S_0/2} R_{ij} \ \ , \langle \langle \psi_i | \psi_j \rangle \rangle = \delta_{ij} \tag{6}$$

where the outer $\langle \cdot \rangle$ refers to expectation over theories for $\langle \langle \psi_i | \psi_j \rangle \rangle$ and R_{ij} is a random variable with zero mean. This can indeed be cast as an example of eigenstate thermalization hypothesis (ETH), if $|\psi_i\rangle = |E_n\rangle$, $|E_n\rangle$ being energy eigenstate, as discussed in [4] and [5]. This is why it is possible to have the following:

$$\langle |\langle \psi_i | \psi_j \rangle|^2 \rangle \neq \delta_{ij} \tag{7}$$

where the outer $\langle \cdot \rangle$ again refers to expectation over theories. This confirms the ensemble interpretation of replica wormholes. Appendix D of [3] is essentially the holographic setup that is shared with this paper. (Note that Section 7 of [3] for non-averaged (that is, non-ensemble) theories is left mostly as largely unrealized future works.)

2. *Saad-Shenker-Stanford (2019)*

This exactly follows Section 2 of [2]. First, consider partition function $\mathcal{Z}$ of the random matrix theory:

$$\mathcal{Z} = \int dH \, e^{-L \mathrm{Tr}[V(H)]} \tag{8}$$

with H being a $L \times L$ Hermitian matrix and $V(H)$ being some potential function over H. This $\mathcal{Z}$ is not to be confused with $\langle Z(\beta) \rangle$ which is actually used for duality with bulk partition function $Z_{JT}(\beta)$:

$$Z_{JT}(\beta) = \langle Z(\beta) \rangle = \frac{1}{\mathcal{Z}} \int dH \, Z(\beta) e^{-L \mathrm{Tr}[V(H)]} \tag{9}$$

where $Z(\beta) = \text{Tr } e^{-\beta H}$ is used as an observable in this ensemble context. The problem arises when we ask for two-boundary partition function $Z_{JT}(\beta_1, \beta_2)$ (or in general, multi-boundary partition function). When used in a non-ensemble context, the right equation is:

$$Z_{JT,non-ens}(\beta_1, \beta_2) = \langle Z(\beta_1)\rangle\langle Z(\beta_2)\rangle \tag{10}$$

However, in an ensemble context, the 'right' (though with '') equation is:

$$Z_{JT,ens}(\beta_1, \beta_2) = \int dH\, Z(\beta_1)Z(\beta_2)e^{-L\text{Tr}[V(H)]} \neq \langle Z(\beta_1)\rangle\langle Z(\beta_2)\rangle \tag{11}$$

and the dual path integral for $Z_{JT,ens}$ is done over connected geometries as well, including replica wormholes, expressed as the following bulk path integral for the two-boundary partition function integrated over topologies and boundaries constrained by β_1 and β_2 and Euclidean JT action given by $I_{JT}[g, \phi]$:

$$Z_{JT,ens}(\beta_1, \beta_2) = \int_{disk} \mathcal{D}g\, \mathcal{D}\phi\ e^{-I_{JT}[g,\phi]} + \int_{cylinder} \mathcal{D}g\, \mathcal{D}\phi\ e^{-I_{JT}[g,\phi]} + .. \tag{12}$$

For the disk topology, two boundaries are disconnected (and constitutes the conventional partition function term), while cylinder topologies connect two boundaries ('connected geometries') and thus the term 'replica' wormhole.

a. A major issue or controversy: does $Z_{JT}(\beta)$ require non-perturbative 'connected geometries' corrections? An issue, or a controversy, is whether we should take $Z_{JT,full}(\beta) = \lim_{n\to 1} Z_{JT,ens}(\beta_1, \beta_2, ..\beta_n)$ where $Z_{full}(\beta)$ is understood as the correct non-perturbative result for $Z_{JT}(\beta)$. This is addressed in Section III C, but a brief argument is provided against such an interpretation.

The non-perturbative interpretation seems natural when JT is considered from the random matrix theory point of view. However, holographic duality has the other direction, and the random matrix theory can be considered as another way of expressing JT theory. If there is no ensemble in the bulk, then either non-perturbatively or perturbatively there is no reason to integrate over connected geometries even for $Z_{JT}(\beta_1, \beta_2)$.

Alternatively, as discussed in Section III A of this paper, Z_{JT}, or simply JT theory, is about an ensemble state when JT is placed in an ensemble context. Therefore any observable calculation within JT is all about properties of an ensemble state. However, the following integral (Equation (11))

$$\int dH\, Z(\beta_1)Z(\beta_2)e^{-L\text{Tr}[V(H)]} \neq \langle Z(\beta_1)\rangle\langle Z(\beta_2)\rangle$$

has ensemble 'observable' $Z(\beta_1)Z(\beta_2)$ that refer to an individual theory within an ensemble. Therefore, it is trivial to see why $Z_{JT}(\beta_1, \beta_2)$ does not equate to the above integral.

The reason why Equation (11) is more appropriate in an ensemble context is that we rarely want properties of an ensemble state - instead, we want to probe properties of an individual theory within an ensemble. For example, in black hole evaporation, we do not want the coarse-grained entanglement entropy (or 'ensemble state entropy') of a black hole - we want fine-grained entropy instead so that unitarity of black hole evaporation can be checked.

b. Chaotic system assumption In [4] and [5] but not restricted to them, ETH was used to provide evidence for replica wormholes. But in the context of this paper, it is not ETH that is important - the assumption that a black hole is a chaotic system is more crucial. If a black hole is sensitive to initial conditions, then we cannot expect an ensemble state to approximate an individual (theory's) microstate - in other words, we cannot expect the final ensemble state of a black hole to be a pure state.

C. Non-unitarity of replica wormholes

1. *Simple proof*

It is simple to prove non-unitarity of replica wormholes. Suppose that at $t = k\Delta t$, the universe is at state $|\psi_{ord,k}\rangle$. Then given $U = e^{-iH\Delta t}$ where H is universe Hamiltonian, the dynamics are given as:

$$|\psi'_{k+1}\rangle = e^{-iH\Delta t}|\psi_{ord,k}\rangle = a_{0,k+1}|\psi_{ord,k+1}\rangle + a_{1,k+1}|\psi_{worm,k+1}\rangle \tag{13}$$

where $|a_{0,k}| \approx 1 - e^{-S} \approx 1$ and $|a_{1,k}| \approx e^{-S/2} \neq 0$ but $|a_{1,k}| \approx 0$. $a_{1,k}|worm\rangle$ refers to corrections due to replica wormholes and other so-called 'non-perturbative' effects. Therefore, if starting from $|\psi_0\rangle = |\psi_{ord,0}\rangle$ at $t = 0$, then the final state at $t = K\Delta t$ is:

$$|\psi_K\rangle = \left(\prod_{k=1}^{K} a_{0,k}\right)|\psi_{ord,K}\rangle + corrections \tag{14}$$

Then unless $K\Delta t$ is exponential time relative to S, $\left|\prod_{k=1}^{K} a_{0,k}\right| \approx 1$. This implies that as long as unitarity holds, corrections $|\psi_{worm,k+1}\rangle$ cannot have much effect unless a black hole is not just large but exponentially large. Note that the result holds regardless of perturbative or non-perturbative analysis.

This is essentially the small corrections theorem argument of Samir D. Mathur [6] (though with a much simpler argument due to a simpler context), but this is intentionally not stated at the beginning. The intention here is to explore replica wormholes and not whether the semiclassical smooth horizon picture of a black hole is not feasible. Therefore, the standard arguments against the small corrections theorem are not really applicable here.

If we intend to use replica wormholes as part of bulk EFT, then non-unitarity is not really a serious issue as long as it is properly recognized. 'E' in EFT refers to effective, so this is not a fundamental issue. The simple proof also is also consistent with an individual theory in an ensemble being a unitary theory - it only renders the corrected theory of 'bulk ensemble EFT with replica wormholes' non-unitary.

2. Impossible alternative: replica wormholes not small corrections

It may be argued that for some ordinary states (that we would have under JT without replica wormholes) $|\psi_{ord,k}\rangle$, $|a_{0,k}|$ in Equation (13) is not approximately 1, and therefore replica wormholes can remain unitary. However, this implies significant corrections to the bulk EFT not just as arising from build-up of small corrections but as outright corrections. Furthermore, this view is unsupported by actual calculations.

To see what this means, consider a Schwarzschild black hole of Schwarzschild radius $r_{s,2}$. One point about the Page curve is that whether this black hole is in early or late-time evaporation (that evolves from the black hole of radius $r_{s,1} \gg r_{s,2}$) matters quantum mechanically, and accumulated corrections will be very small for early-time evaporation while large for late-time evaporation.

However, when $|a_{0,k}|$ is not approximately 1, large corrections hit regardless of an evaporation process (whether early or late-time). We only need to look at thermal two-point function results for thermofield double states in [4] to realize that this cannot be the case.

$$
\begin{gathered}
G_{2,\beta}(t) \sim e^{S_0} \quad t \sim 1 \\
G_{2,\beta}(t) \sim O(1) \quad t \sim \text{late times}
\end{gathered}
\tag{15}
$$

That is, evaporation duration matters (with decay of two-point function stopping at late times suggesting corrections), and naive large corrections by replica wormhole contributions for some class of bulk states are not what actual derivations suggest.

3. *Digression: Petz map and crossed product construction*

From Section III C 1, it can be noted that if Petz reconstruction map [7] and crossed product construction [8] are understood as corrections to the bulk theory, then these corrections have to be either non-unitary or large in a small time interval. (For the Petz map understanding of replica wormholes, see [3], in particular Section 3.)

The aforementioned conclusion can be avoided if we understand these approaches as not corrections but rather as theory transformations - in other words, they have different Hilbert spaces such that comparisons in Section III C 1 are invalid. For the purpose of precluding the non-perturbative interpretation of replica wormholes, this is all we need - these relatively recent proposals cannot be used to support the non-perturbative interpretation. For the general question of whether these new ideas can resolve the black hole information problem, an additional conceptual question has to be asked, which really has nothing to do with mathematics and pure logical and conceptual questions.

It is true that these methods mathematically demonstrate that either the Page curve [9, 10] or Bekenstein-Hawking entropy [8] can be derived with appropriate theory transformations, which can rigorously be stated within the Tomita-Takesaki theory. But these methods rely on some form of effective observer dependence (think of 'modular flow' as an example) for theory transformations. For example, an exterior observer (relative to a black hole) can see a black hole as purifying, while other observers may see the same black hole as remaining in heavily mixed states.

These methods can definitely be understood information-wise, and as long as all information can be recovered, an argument can go that physical entanglement can be ignored. That is, even if two systems are physically entangled, as long as an observer manages to deduce the pure state of the two systems from information in one system, then entanglement is treated irrelevant. Whether this is a valid path still would require further analysis - in the context of this paper, this is irrelevant.

4. *Firewalls*

A known consequence of replica wormholes is that *firewalls* can arise at late times, so there are significant dramas around the black hole horizon [11–14]. We do not need to define what firewalls exactly refer to, and vaguely define them as resulting in semiclassical gravity and EFT being no longer feasible.

Despite the seemingly alarming result, there is actually no surprise. It is only a significant issue if one does not adhere to the ensemble interpretation of replica wormholes and continues to argue that they are genuine non-perturbative corrections to JT.

Within the ensemble interpretation, it must be recognized that an individual theory should have its own holographic dual theory. Therefore, firewalls in JT EFT (with replica wormholes) only signal that the spacetime described by JT is very different from the actual spacetime. There is no need to sacrifice smooth black hole horizons and no drama in this interpretation.

To potential adversaries (of replica wormholes), this may be thought to hint that replica wormholes alone cannot yet be proposed as resolving the black hole information problem, since we are yet to understand the consistency and stability of the actual bulk theory, instead of some ensemble theory.

5. *Small note/digression: real-time (Lorentzian) replicas*

While replica wormholes are often cast in Euclidean time, they are not solely Euclidean effects. Indeed, bulk path integrals with real-time (Lorentzian) replicas have been performed (which give us essentially the same results) - see [15] for more information.

D. Alternatives to classical dramas and firewalls

A typical concern for large corrections to the bulk theory (this includes the actual bulk theory observationally deviating from the ensemble theory at late times) is that we do not expect drama classically. If there are large EFT corrections, then the classical drama is likely to be observable, which we are yet to witness. In this subsection, a toy model is set up to show how we may not witness classical drama, even when there are quantum dramas and surprises. This does not require the ensemble interpretation and works generally. Readers are advised to skip to Section IV if they are uninterested in purely black hole contexts.

1. *Intuition (1): Classical geometry superposition intuition*

Approximate a quantum black hole as a superposition of classical geometries. If EFT is valid for each classical geometry, then even if the superposition requires significant corrections to EFT, EFT is technically not modified and therefore there is actually no drama around the event horizon. For clarifications, see below.

In quantum computing terms, this is possible when we have the same evolution operator (or quantum operation) for every classical state (or 'computational basis') except for one particular state, which makes the quantum operation conditional. In black holes, there is a natural candidate for this 'particular state', which is the case of an empty black hole.

While each superposition state may accurately be approximated by some background geometry [16], dynamics of the superposition state may deviate from what EFT under this 'superposition state' background geometry predicts. And it is demonstration of this point that this subsection deals with.

2. *Intuition (2): final state anchor intuition for zero final entanglement entropy*

The idea involves using a final black hole state as an anchor, and has some aspects of the Horowitz-Maldacena final state proposal [17], though they are not equivalent. Or one may say that the idea in this sub-subsection is the Page curve derivation made dynamic [1]. Which interpretation one takes does not matter in this paper, since this paper does not intend to propose some model of a black hole and only uses the idea for suggesting that an alternative to a complete EFT drama (including firewalls) is possible.

If a black hole has a single final pure state that is the effective ground state of a black hole, then if every black hole trajectory heads toward complete evaporation, then it eventually reaches this final pure state. Therefore, the entanglement entropy of a black hole must fall to zero.

[1] Thomas Faulkner's comment (personal conversation)

The question is whether this scenario is realistic. This can only partially be resolved in this paper.

3. Toy model (which is not a quantum measurement model)

A qubit model is developed, and this can lead to the following criticisms, especially given that some heavy abstraction is made - the model here does not properly address locality that can only be addressed by the full field theory. While this is technically true, there is not anything this model that would require non-local interactions. The purpose is not to develop a full account of black hole evaporation anyway, so this would not be an issue.

This model is a slight modification of the Jaynes-Cummings model [18], but the presentation in this paper will be self-sufficient for an analysis and understanding.

a. Setup Each kth qubit ($1 \leq k$, $k \in \mathbb{N}$) in R initially in $|0\rangle$ interacts sequentially with the black hole B (the brane does not exist here) for duration of Δt_k ("interaction time") from $t = \sum_{j=1}^{k-1} \Delta t_j$ to $t = \sum_{j=1}^{k} \Delta t_j$. Then the qubit flies away, never interacting with the black hole again. The initial state of the black hole is assumed to be $|N\rangle$ at $t = 0$, with N treated as if representing the number of particles in the black hole. There are M radiation qubits, which are distinguished from interior 'particles'. N and M are re-defined only for this section and will return to the original definition.

The free Hamiltonian H_{rad} of each qubit is given as, with qubit states satisfying $\langle 0|1\rangle = 0$:

$$H_{rad} = \omega_a \sigma_+ \sigma_- = \omega_a |1\rangle\langle 1| \tag{16}$$

This is simply about only the excited state of the qubit having non-zero energy, with $\sigma_+ = |1\rangle\langle 0|$ and $\sigma_- = |0\rangle\langle 1|$.

The free Hamiltonian H_{BH} of the black hole is given as:

$$H_{BH} = \omega_a a^\dagger a \tag{17}$$

with commutation relations $[a, a^\dagger] = 1$ and $[a, a] = [a^\dagger, a^\dagger] = 0$. This gives us $|n\rangle$ with particle number operator $\hat{n} = a^\dagger a$ such that $\hat{n}|n\rangle = n|n\rangle$ for $n \in \mathbb{N}$. The initial state then satisfies $\hat{n}|N\rangle = N|N\rangle$.

Interaction Hamiltonian H_I that couples a *single* qubit in R and the black hole B (when under interaction) goes as:

$$H_I = g(\sigma_+ a + \sigma_- a^\dagger) \tag{18}$$

Since we have same ω_a in the free Hamiltonians, this interaction is all about trading non-interaction energy: if the black hole loses energy ω_a, then the qubit obtains ω_a and vice versa. The total number of 'particles' and non-interaction energy are conserved in the model. The value of ω_a is irrelevant for the remainder of the discussions.

b. Simplification The toy model can further be approximated and simplified as follows - for the purpose of this paper, the simplified approximation would suffice.

$$\begin{gathered} U|0,n\rangle = \frac{1}{\sqrt{2}}\left(|0,n\rangle + |1,n-1\rangle\right) \quad (0<n) \\ U|0,0\rangle = |0,0\rangle \end{gathered} \tag{19}$$

The first $|0\rangle$ refers to a radiation qubit which after U is applied never interacts with the black hole interior $|n\rangle$ again. The argument is that U, which we may take as a representative (or an 'analogy') for EFT, applies for every classical interior configuration $|n\rangle$ $n > 0$, and therefore technically there is zero correction to EFT. As such, the interior entropy follows monotonic entanglement entropy increases. What stops this monotonic increase is $n = 0$, where no further entanglement can be generated.

Start from the initial interior state $|N\rangle$ with all qubits in $|0\rangle$. Eventually with probability of 1, every interior outcome collapses to $|n = 0\rangle$, which automatically purifies the black hole interior.

Therefore, it is possible to have no classical EFT drama while there are quantum EFT dramas.

IV. THE ENVARIANCE CALCULATION WITHIN JT/RMT BLACK HOLES

A. Terminology

The terms described in the below table are thoroughly explained throughout the section and the readers may skip to the derivation.

- R: refers to the system being measured at final time $t = t_f$. Analogous to black hole radiations (but black hole contexts are not necessary).

- B: rest of the universe excluding R. Analogous to a black hole (but black hole contexts are not necessary).

- $R \cup B$: universe.
- $|\Psi_i\rangle$: Initial ensemble and actual state of the universe, and $|\Psi_i\rangle = |\Psi^i_{ens}\rangle = |\Psi^i_{actual}\rangle$. (Equation (20))
- $|\psi_{\phi_j}\rangle$: an initial individual outcome of B entangled with coarse-grained state $|\phi_j\rangle$ of R. (Equation (20))
- $|\phi_j\rangle$: an initial individual outcome of R. (Equation (20)) Deriving final measurement probability as corresponding to $|a_j|^2$ associated with $|\phi_j\rangle$ is the goal of this section.
- $|k\rangle$ ($k \in \mathbb{N}$): an actual final microstate of R at final time $t = t_f$. (Equation (23))
- $|\psi_k\rangle$: an actual final outcome of B, entangled with final outcome $|k\rangle$ of R. (Equation (23))
- $|k_j\rangle$: For each initial (coarse-grained) outcome $|\phi_j\rangle$ of R, an associated final outcome would be $|k_j\rangle$ for R, with the number of possible $|k_j\rangle$ for $|\phi_j\rangle$ in final ensemble state $|\Psi^f_{ens}\rangle$ (or final coarse-grained outcome $|\Phi^f_j\rangle$ of the universe) being $\#(\phi_j)$. (Equation (25) and for $\#(\phi_j)$, see Equation (24)) This is just $|k\rangle$ with $k = k_j$.
- $|\psi_{k_j}\rangle$: In final ensemble state $|\Psi^f_{ens}\rangle$ (or final coarse-grained outcome $|\Phi^f_j\rangle$ of the universe), $|\psi_{k_j}\rangle$ is an outcome of B entangled with $|k_j\rangle$ of R. (Equation (25).)
- $|\Phi^f_j\rangle$: The final 'coarse-grained' outcome (and state) of the universe $R+B$ if the initial state of R is $|\phi_j\rangle$ in an ensemble theory. (Equation (25))
- $|\Psi^f_{ens}\rangle$: The final ensemble state of the universe. In the ensemble theory, initial universe state $|\Psi_i\rangle$ evolves to final universe state $|\Psi^f_{ens}\rangle$. (Equation (23) and as an equality, see also Equation (26) that is written in terms of $|\Phi^f_j\rangle$)
- $\#(\phi_j)$: The number of final outcomes of R consistent with initial outcome $|\phi_j\rangle$ of R.
- U_{ens}: The evolution propagator of an ensemble theory from initial time to final time. Roughly $e^{-iH\Delta t}$ where H is the universe Hamiltonian.
- N: The number of fine-grained microstates in R.
- M: The number of initial coarse-grained states in R. (Equation (20))

- a_j: The initial amplitude of each coarse-grained outcome of R. It is final measurement probability $|a_j|^2$ that we seek to recover. (Equation (20))
- b_k: $|b_k|^2 = 1/N$, with b_k being the final ensemble state amplitude for each fine-grained outcome $|k\rangle$ of R in Equation (23).
- b_{k_j}: b_k with $k = k_j$. (Equation (25))

B. Derivation

While JT gravity is used as a reference theory, as long as a final ensemble state of some measured system R is a maximally mixed state with an actual microstate being a pure state, with an initial microstate and ensemble state being approximately the same state, the calculations in this section are consistent with any theory of a chaotic system.

This section mostly follows the strategy of [1]. The initial state of the universe is now modified from Section II such that instead of R tracking the fine-grained configuration of the EOW brane, R tracks the configuration of the brane in a coarse-grained way with $M \ll N$ (and N is assumed to be sufficiently large so that we can effectively treat it as infinity):

$$|\Psi_i\rangle = \sum_{j=0}^{M-1} a_j |\psi_{\phi_j}\rangle |\phi_j\rangle \tag{20}$$

where $|\phi_j\rangle$ is some coarse-grained state of R with $\langle\phi_j|\phi_i\rangle = \delta_{ij}$. Note that this initial state abstracts away the entanglement process between the brane and R, arising from the following pure state of R:

$$\sum_{j=0}^{M-1} a_j |\phi_j\rangle \tag{21}$$

The universe consists of R and B. The final ensemble state of R goes as:

$$\rho^f_{R,ens} = \frac{1}{N} \sum_{k=0}^{N-1} |k\rangle\langle k| \tag{22}$$

with $\langle k|j\rangle = \delta_{jk}$, and the final ensemble state of the universe is therefore:

$$|\Psi^f_{ens}\rangle = \sum_{k=0}^{N-1} b_k |\psi_k\rangle |k\rangle \tag{23}$$

where $|b_k|^2 = \frac{1}{N}$.

Both an ensemble theory and an individual theory in an ensemble are unitary. Furthermore, an actual final microstate is known to be a pure state, and B (analogous to a black hole) is a chaotic system, which implies that different final microstates will be very close to being mutually orthogonal. Therefore, a_j information in Equation (20) has to be captured by $|k\rangle$ of Equation (23).

This is only possible when multiple $|k\rangle$, each with probability of $1/N$, correspond to the same initial coarse-grained state $|\phi_j\rangle$ of probability $|a_j|^2$. The number of final microstates ($|k\rangle$) that correspond to initial coarse-grained state $|\phi_j\rangle$ is therefore given as:

$$\#(\phi_j) \approx |a_j|^2 N \tag{24}$$

Since $M \ll N$ and N is sufficiently large as to be approximated as infinity, the above equation is effectively an equality.

To be mathematically precise, the above is all about defining a final coarse-grained state $|\Phi^f_{j,ens}\rangle$:

$$\begin{aligned} |\Phi^f_{j,ens}\rangle &= U_{ens}|\psi_{\phi_j}\rangle|\phi_j\rangle, \quad (U_{ens} \approx e^{-iH_{ens}t}) \\ |\Phi^f_{j,ens}\rangle &= \sqrt{\frac{N}{\#(\phi_j)}} \sum_{k_j} b_{k_j}|\psi_{k_j}\rangle|k_j\rangle \end{aligned} \tag{25}$$

where each $|k_j\rangle$ roughly corresponds to initial coarse-grained state $|\phi_j\rangle$ and there are $\#(\phi_j)$ such k_j's. Therefore,

$$|\Psi^f_{ens}\rangle = \sum_{j=0}^{M-1} |a_j||\Phi^f_j\rangle \tag{26}$$

We could therefore say that initial outcome $|\psi_{\phi_j}\rangle|\phi_j\rangle$ evolved to $|\Phi^f_j\rangle$ in an ensemble theory, maintaining unitarity.

In an actual (non-ensemble) state point of view, initial state $|\Psi_i\rangle$ evolves to one of fine-grained microstate $|\psi_k\rangle|k\rangle$ (in Equation (23)), which implies that the state of R is pure state $|k\rangle$. If there is only access to coarse-grained observables, then all we can verify is each $|\Phi^f_j\rangle$, for which $\langle\psi_k, k|\Phi^f_j\rangle \neq 0$.

Therefore, final outcome probability of $|a_j|^2$ is recovered for each initial outcome $|\phi_j\rangle$ of R.

V. CONCLUSION

This paper exploited an ensemble theory that can also be re-interpreted as a non-ensemble unitary quantum theory to develop a unitary and deterministic account of quantum measurements. In this understanding of quantum measurements, their probabilistic illusions arise because of our ignorance of which theory is actual within an ensemble theory. In non-measurement and/or non-black hole contexts, an ensemble theory sufficiently approximates an actual theory such that they are not distinguishable.

An ensemble theory does not have to be Jackiw-Teitelboim (JT) gravity, as long as conditions that mimic black hole evaporation are satisfied. From some initial quantum state of R to be measured, individual outcomes of R get entangled with the rest (B) of the universe. Then the universe evolves in the ensemble theory such that the ensemble state of R becomes maximally mixed. Assuming B to be a chaotic system, in an individual and actual theory, only one of the outcomes in the ensemble state is realized, which gives us a pure actual state for R. As long as a maximally mixed ensemble state is assumed to imply equal probability, we obtain the Born rule for probability of an individual outcome, with the derivation similar to [1].

An important assumption is that each time a black hole-like universe is there, an actual theory differs and is given probabilistically according to an ensemble theory. This can be justified on the ground that there are degrees of freedom that we are yet to understand in quantum gravity and effective Hamiltonians of the universe differ from one instance to another. (Alternatively, we can look toward examples like JT/SYK holographic duality [19–22] to support the viability of randomness.)

It is unclear whether the quantum measurement theory of this paper would be anyhow realistic, nor was it the intention of this paper. Rather, the aim was only to provide a unitary and deterministic theory of quantum measurements that is agnostic to quantum foundation questions.

Some questions remain. It may be argued that the very possibility of a deterministic quantum measurement theory is totally implausible (for which the author has to respectfully disagree) such that any such demonstration only means either that 1) the ideas like replica wormholes are somehow flawed or 2) a true non-perturbative and non-ensemble interpretation of replica wormholes is necessary despite issues seen in Section III. There are reasons to probe both possibilities further, and the author of this paper does not yet intend to put an end to these questions in this paper, though it must be recognized that the natural and logical interpretation of replica wormholes so far is an ensemble interpretation, and non-perturbative unitarity of replica wormholes is very questionable, as explored in Section III. It would be best to interpret the quantum measurement theory of this paper as a consequence of or hinted from the ensemble interpretation of black hole evaporation (or replica wormholes).

ACKNOWLEDGMENTS

This paper partially originates from a personal talk with Thomas Faulkner, who provided invaluable suggestions, especially with regard to (and in defense of) unitarity of replica wormholes. I tried my best to replicate what proponents and supporters of replica wormholes have argued and would have argued.

DATA AVAILABILITY AND DECLARATION OF INTERESTS

The author(s) have no funding source to declare. Furthermore, there is no conflict of interests.

---

[1] Wojciech Hubert Zurek, "Environment-Assisted Invariance, Entanglement, and Probabilities in Quantum Physics," Phys. Rev. Lett. **90**, 120404 (2003).

[2] Phil Saad, Stephen H. Shenker, and Douglas Stanford, "JT gravity as a matrix integral," arXiv e-prints (2019), arXiv:1903.11115 [hep-th].

[3] Geoff Penington, Stephen H. Shenker, Douglas Stanford, and Zhenbin Yang, "Replica wormholes and the black hole interior," Journal of High Energy Physics **2022**, 205 (2022).

[4] Phil Saad, "Late Time Correlation Functions, Baby Universes, and ETH in JT Gravity," arXiv e-prints (2019), 10.48550/arXiv.1910.10311, arXiv:1910.10311 [hep-th].

[5] Alexandre Belin and Jan de Boer, "Random statistics of OPE coefficients and Euclidean wormholes," Classical and Quantum Gravity **38** (2021), 10.1088/1361-6382/ac1082.

[6] Samir D. Mathur, "The information paradox: a pedagogical introduction," Classical and Quantum Gravity **26**, 224001 (2009).

[7] Jordan Cotler, Patrick Hayden, Geoffrey Penington, Grant Salton, Brian Swingle, and Michael Walter, "Entanglement wedge reconstruction via universal recovery channels," Physical Review X **9** (2019), 10.1103/physrevx.9.031011.

[8] Edward Witten, "Gravity and the crossed product," Journal of High Energy Physics **2022** (2022), 10.1007/jhep10(2022)008.

[9] Eyoab Bahiru and Niloofar Vardian, "Explicit reconstruction of the entanglement wedge via the Petz map," Journal of High Energy Physics **2023** (2023), 10.1007/jhep07(2023)025.

[10] Niloofar Vardian, "Black hole interior Petz map reconstruction and Papadodimas-Raju proposal," Journal of High Energy Physics **2023** (2023), 10.1007/jhep10(2023)024.

[11] Douglas Stanford and Zhenbin Yang, "Firewalls from wormholes," arXiv e-prints (2022), arXiv:2208.01625 [hep-th].

[12] Andreas Blommaert, Chang-Han Chen, and Yasunori Nomura, "Firewalls at exponentially late times," Journal of High Energy Physics **2024** (2024), 10.1007/JHEP10(2024)131.

[13] Amir A. Khodahami and Azizollah Azizi, "Replica wormhole and amps firewall," Physics Letters B **859** (2024), 10.1016/j.physletb.2024.139130.

[14] Hamed Zolfi, "Firewalls from wormholes in higher genus," Journal of High Energy Physics **2024** (2024), 10.1007/JHEP05(2024)039.

[15] Sean Colin-Ellerin, Xi Dong, Donald Marolf, Mukund Rangamani, and Zhencheng Wang, "Real-time gravitational replicas: formalism and a variational principle," Journal of High Energy Physics **2021** (2021), 10.1007/JHEP05(2021)117.

[16] Suvrat Raju, "Lessons from the information paradox," Physics Reports **943**, 1–80 (2022).

[17] Gary T. Horowitz and Juan Maldacena, "The black hole final state," Journal of High Energy Physics **2004** (2004), 10.1088/1126-6708/2004/02/008.

[18] E.T. Jaynes and F.W. Cummings, "Comparison of quantum and semiclassical radiation theories with application to the beam maser," Proceedings of the IEEE **51**, 89–109 (1963).

[19] Gustavo J. Turiaci and Herman Verlinde, "On CFT and quantum chaos," Journal of High Energy Physics **2016** (2016), 10.1007/jhep12(2016)110.

[20] Kristan Jensen, "Chaos in ads$_2$ holography," Phys. Rev. Lett. **117** (2016), 10.1103/PhysRevLett.117.111601.

[21] Juan Maldacena, Douglas Stanford, and Zhenbin Yang, "Conformal symmetry and its breaking in two-dimensional nearly anti-de Sitter space," Progress of Theoretical and Experimental Physics **2016** (2016), 10.1093/ptep/ptw124.

[22] Gabor Sarosi, "AdS_2 holography and the SYK model," Proceedings of Science **Modave2017**, 001 (2018).