A collapse mechanism may not save semiclassical gravity: a black hole thought experiment

Minseong Kim*

(Dated: June 22, 2025)

Abstract

Typical thought experiments against a variety of semiclassical gravity theories rely on measurements or assume specific wavefunction collapse mechanisms to state uncertainty principle violations or superluminal signaling. This makes them susceptible to counterarguments based on other wavefunction collapse and measurement theories. This paper puts forward a qubit model thought experiment that does not rely on measurements - it rather checks the entanglement entropy of a black hole as it completely evaporates away. If gravity is quantized, then correcting for the empty black hole outcome is sufficient to obtain zero entropy at the end of evaporation. If not, then the entropy has to increase monotonically due to the original Hawking paradox plus the small corrections theorem.

Keywords: black hole information problem, semiclassical gravity, quantization of gravity, qubit model

* mkimacad@gmail.com; ORCiD:0000-0003-2115-081X

CONTENTS

I. INTRODUCTION

The question of whether gravity needs to be quantized is a theoretically unsettled question [1–4], although there is a general consensus that gravity does need to be quantized.

The main issue with any no-go thought experiment against semiclassical gravity is that whenever a quantum field theory (QFT) on curved background spacetime is well-defined, it forms a valid unitary theory. Therefore, most no-go thought experiments against semiclassical gravity rely on measurements or some wavefunction collapse mechanism in order to state uncertainty principle violations and superluminal signaling.

These thought experiments are inevitably susceptible to counterarguments based on specific collapse or measurement theories, and disproving these theories may require bringing down entire interpretations of quantum mechanics. This likely involves more than thought experiments against semiclassical gravity.

QFT on curved background spacetime is not without issues. While algebraic quantum field theory (AQFT) is widely used for rigor and the benefits it provides for curved backgrounds, whether it can reproduce all empirical successes of canonical QFT is yet unsettled [5]. Furthermore, while largely compatible, there are different approaches to AQFT [6, 7], along with the extension of path integrals to curved spacetime and gravity. Not all backgrounds admit quantization, which is often limited to globally hyperbolic spacetime [7].

Furthermore, more problematically, there is the backreaction issue. Nevertheless, once backreaction is incorporated properly such that some sort of equilibrium is reached, even Møller-Rosenfeld semiclassical gravity ('conventional semiclassical gravity'), with Einstein field equations coupled to average stress-energy tensor, is consistent by design, if we are careful enough to avoid divergences. Depending on a theory of semiclassical gravity, mean-field backreaction may not even be relevant, if there is any backreaction [8].

In this paper, we offer a thought experiment based on a black hole that avoids invoking measurements. Rather we start with the conventional semiclassical Hawking picture of a black hole that leaves the black hole exterior heavily entangled even after complete evaporation. Then we correct this picture with the empty black hole outcome not producing any lasting long-range entanglement at asymptotic infinity, just as expected for Minkowski spacetime. This correction requires gravity to be quantized, since we separate cases of non-empty black holes and the empty black hole. In such a quantum gravity case, the unitarity of the black hole exterior can be restored. This suggests that black hole consistency requires gravity to be quantized.

The conclusion of course depends on the validity of conventional expectations regarding black holes, which are discussed later in this paper. If one of these assumptions turns out to not be the case, then the thought experiment breaks down.

We first review some important no-go thought experiments against semiclassical gravity and counterarguments against them. As mentioned above, the lesson is that the necessity of quantizing gravity remains unsettled, though the arguments are aesthetically and logically appealing. In contrast to other reviews [1–4], we focus only on the physics behind no-go thought experiments and not other philosophical considerations. Then the qubit model setup that captures important aspects of black hole evaporation is laid out, leading to a no-go thought experiment against semiclassical gravity if conventional expectations regarding black holes are correct.

II. A REVIEW OF NO-GO THOUGHT EXPERIMENTS AGAINST SEMICLASSICAL GRAVITY

A. On the definition of semiclassical gravity

There is no one definition of semiclassical gravity, with only one feature standing as a common ground - that gravity is classical, without quantum correlations. As we see in some theories of semiclassical gravity such as postquantum gravity [9], semiclassical gravity can even be classically probabilistic. For this reason, we distinguish whether a particular no-go thought experiment works generally or only for a particular theory of semiclassical gravity. Most no-go thought experiments apply for deterministic semiclassical gravity with a Copenhagen-style instantaneous collapse mechanism.

Typically, Møller-Rosenfeld semiclassical gravity is assumed whenever semiclassical gravity is mentioned, though we take a caution not to use the word 'semiclassical gravity' this way. This theory states that gravity operates by the following Einstein field equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} \left\langle \hat{T}_{\mu\nu} \right\rangle_{\psi} \tag{1}$$

where $G_{\mu\nu}$ is the Einstein tensor, G the gravitational constant and ψ is some quantum state.

Note also that while quantum discussions of black holes assume that gravity is to be quantized, the Møller-Rosenfeld semiclassical gravity picture is understood to be sufficient for most calculations. This leads to the black hole information problem, and there is no consensus as to where conventional calculations go wrong. Some do argue that quantum black holes cannot be well-approximated by the semiclassical picture and some structures like fuzzballs appear, though they remain a significant minority - see [10] for a defense of the semiclassical approximation in mainstream black hole physics.

B. Feynman-Bohr-Rosenfeld (FBR) argument

1. Bohr-Rosenfeld (BR) argument

The BR argument [11] adapted to gravity is one of the earliest and primitive arguments against semiclassical gravity [12]. It also remains as an influential intuition behind the common consensus on the necessity of quantizing gravity.

Consider some massive source particle (of significant mass M) of initial superposition of positions:

$$|\psi_M\rangle = \frac{1}{\sqrt{2}}(|d_0\rangle + |d_1\rangle) \tag{2}$$

where d_0, d_1 represent the distances from the initial position of the test probe particle, with $d_1 \gg d_0$. The test probe particle is initially in some acceleration eigenstate.

If the gravitational field is determined as the deterministic function of $|\psi_M\rangle$ (common versions of semiclassical gravity including Møller-Rosenfeld), then the test particle will always have the same final particle position, which implies that the test particle is not entangled with the source particle. Furthermore, the source particle notices the backreaction recoil due to the test particle, but now with the gravitational force treated semiclassically.

Each classical outcome $|d_0\rangle, |d_1\rangle$ notices that the classical gravitational equations of motion have broken down for its own branch and can infer the other outcome in the superposition, which leads to a violation of the uncertainty principle. We can deny backreaction recoil to save linearity and unitarity, but this results in the momentum non-conservation issue. Therefore, it initially appears that gravity has to be quantized.

2. Feynman argument

The Feynman argument [13], also an early no-go thought experiment, asks whether semiclassical gravity can survive the double-slit experiment. The logic of the argument is similar to the BR argument - if there is recoil due to the source particle, then we can infer other test particle outcomes even from one test particle outcome branch, leading to a violation of the uncertainty principle, and if recoil is assumed to be non-existent, then we have momentum non-conservation.

The test particle goes through two slits in a superposition. Right around the slits, there is a source particle of mass M. If we assume deterministic semiclassical gravity, then this source particle moves in the same way regardless of the test particle slit outcome, and the test particle recoil due to the source particle reveals the information about the other slit outcome, violating the uncertainty principle.

3. Implicit measurement assumption

In the FBR argument, the trouble arises because we look at what happens in an individual classical outcome. Counterarguments can therefore be formulated: as long as there is no collapse to each classical outcome, there is no actual uncertainty principle violation or superluminal signaling. If there is collapse, then the whole arguments are susceptible to specific collapse mechanisms and interpretations of quantum mechanics.

4. Page-Geilker (PG) experiment

The Page-Geilker (PG) experiment [14] provides an empirical implementation (or variant) of the FBR thought experiment such that we do not need extremely large value of mass to distinguish semiclassical gravity predictions from when gravity is quantized. The cost of the PG experiment relative to FBR is that it largely only works against Møller-Rosenfeld semiclassical gravity in the Copenhagen interpretation. For other collapse mechanisms, the experiment does not even rule out Møller-Rosenfeld.

The experiment starts with a standard Cavendish-style torsion balance with two small test particles (with non-zero mass) on the (left and right) ends. PG modifies it to a single particle setup such that some kind of quantum switch determines at which end the particle would be. Therefore, the particle position is in superposition, and Møller-Rosenfeld predicts that there would be zero net torque, since the torques at both ends cancel each other. But this is not what we get experimentally when the torque is measured.

However, the experiment relies on measuring net torque, and a counterargument may state that wavefunction collapse disturbs classical gravitational fields such that the initial canceling balance breaks down - Møller-Rosenfeld reacts to the collapse-induced change in stress-energy tensor expectation.

C. Eppley-Hannah (EH) argument

The FBR argument can be refined and made more precise as the Eppley-Hannah argument [15]. Suppose that the quantum state is initially a one-particle momentum eigenstate $|p\rangle$. We measure its positional trajectory using a small-momentum small-wavelength classical gravitational wave, which is only possible when gravity is not quantized. Eppley-Hannah then states that there are only three possible scenarios:

1. Gravitational waves collapse the initial quantum state to some position state $|x\rangle$, but since gravitational waves used carry close to zero momentum, the particle also approximately has definite momentum p. That is, we know that the final state is $|x\rangle$ but also know its momentum p without being $|p\rangle$ - a violation of the uncertainty principle.

2. Gravitational waves collapse the initial quantum state in a way consistent with the uncertainty principle. This implies non-conservation of momentum and energy.

3. Gravitational waves alone do not collapse the initial quantum state. But suppose that we have two entangled particles L, R that are spatially separated by distance d, with particle L then collapsed with a quantum measurement while R remaining uncollapsed. Then superluminal signaling is possible, in case gravitational waves respond to the change in L. This is because observers close to R can continuously monitor

gravitational waves without disturbing R. (If we continuously measure R instead, there is necessarily a disturbance to L as well due to entanglement, which defeats the purpose of signal detection.)

As with FBR and PG, the conclusion that semiclassical gravity is untenable depends on the assumed collapse mechanism. In fact, even within the Copenhagen interpretation, the conclusion follows only under instantaneous collapse, as we see in Section II C 1.

1. *Kent's counterargument: instantaneous collapse is the problem, not Copenhagen interpretation*

[16] argues that even within the Copenhagen interpretation, the third scenario of the EH argument only works if collapse is instantaneous. Suppose L and R are entangled as follows:

$$\frac{1}{\sqrt{2}}(|0_L, 0_R\rangle + |1_L, 1_R\rangle) \tag{3}$$

Now L is measured at $t = T$ to either of $|0_L\rangle, |1_L\rangle$. Then EH assumes that R immediately collapses to one of $|0_R\rangle, |1_R\rangle$ at $t = T$. Without directly measuring R, one can instead measure the classical gravitational wave around R to infer the measurement outcome of L, which in turn informs the collapsed state of R. But EH argues that this is superluminal signaling.

Suppose instead that collapse is not instantaneous - at $t = T$, even when L is measured to one of the outcomes, R remains in the following mixed state:

$$\frac{1}{2}(|0_R\rangle\langle 0_R| + |1_R\rangle\langle 1_R|) \tag{4}$$

Then the information update via the gravitational wave propagates from L starting at $t = T$, which is first received by R at $t = T + d/c$, with c the speed of light - no superluminal signaling. Right at $t = T + d/c$, the state of R is updated to one of $|0_R\rangle, |1_R\rangle$.

See also [17] for an additional and similar analysis, in particular with regards to the PG experiment.

D. Peres-Terno (PT) argument

The PT argument [18] provides an important example on why mixing a classical system with a quantum system is difficult due to a breakdown in the correspondence principle.

This is not completely a no-go experiment, as some complex interactions may still save the correspondence principle. Nevertheless, generally within quantum mechanics, classical-quantum interactions violate the correspondence principle.

We first have to ask how a classical system may be modeled in quantum mechanics, and this itself can be a contentious question which is not to be addressed in this paper. PT simply invokes the Koopman-von Neumann mechanicms formalism such that the Schrödinger equation for classical systems is obtained. In this formalism, classical systems become probabilistic in sense of classical statistical mechanics.

Mathematically, we define Liouvillian L_v (v has no meaning other than to symbolize Liouvillan, in contrast to Lagrangian),

$$L_v = \left(\frac{\partial H}{\partial y}\right)\left(-i\frac{\partial}{\partial x}\right) - \left(\frac{\partial H}{\partial x}\right)\left(-i\frac{\partial}{\partial y}\right) \tag{5}$$

where H is classical Hamiltonian, y is classical momentum as opposed to the quantum momentum operator $p_x = -i\frac{d}{dx}$ and x is one-dimensional position. L_v is unobservable, though x and y are observable. Then the Schrödinger equation for classical systems is given as:

$$i\frac{\partial \psi}{\partial t} = L_v \psi(x, y, t) \tag{6}$$

where $\psi(x, y, t) = \sqrt{f(x, y, t)}$, with f being the Liouville probability density of classical statistical mechanics.

PT then defines Koopmanian K that serves as a Hamiltonian for classical-quantum systems:

$$K = L_v(x, y, p_x, p_y) + H(q, p) + K_I \tag{7}$$

where $p_y = -i\frac{d}{dy}$, $H(q, p)$ is the Hamiltonian for the purely quantum system with position q and momentum p and K_I is the interaction term. This gives the Schrödinger equation for the hybrid classical-quantum system:

$$i\frac{\partial \Psi}{\partial t} = K\Psi(x, y, q, p, t) \tag{8}$$

where Ψ is the full wavefunction of the hybrid system. The question then is how we define K_I so that the correspondence principle is respected.

PT considers the example of two harmonic oscillators with bilinear coupling - one harmonic oscillator is now classical and the other is quantum. The correspondence principle implies that we should recover the same pre-quantization equations of motion regardless of whether we consider the hybrid system or the purely classical system. However. this is not possible unles we break the commutation relation of the quantum harmonic oscillator. In case of bilinear coupling, we want $[p.K_i] = -kx$ and $[y, K_i] = -kq$ with $[y, p] = 0$ (y is classical momentum), but this is prohibited by Jacobi's identity on y, K_i, p. PT then considers treating the interaction initially as classical Hamiltonian $H_{int} = kqx$ which is then converted to Liouvillan $K_I = L_{v,int}$ by Equation (5):

$$K_I = L_{v,int} = -kqp_y \tag{9}$$

However, even this approach leads to significant deviations from the correspondence principle, so the hybrid approach is difficult to accept.

The PT approach does not invoke measurements and is initially free of counterarguments based on collapse mechanisms. However, it is exactly this problem of classical-quantum interactions that collapse mechanisms deal with, so we have not really escaped collapse mechanisms.

E. Additional papers, in addition to no-go thought experiments

We have so far focused on established and well-known no-go thought experiments against semiclassical gravity. There are some reviews and additional thought experiments that are worth noting though this cannot be a complete list.

For a philosophy of physics review of quantum gravity and semiclassical gravity, see [4]. For older philosophy of physics reviews, see [1, 2], with some of the points discussed in this paper shared. See [3] as well that notes how the uncertainty principle may be read epistemically in different interpretations of quantum mechanics.

The Schrödinger-Newton equation, which works as the non-relativistic limit of semiclassical gravity, has also served an important role in semiclassical gravity discussions, and [19] suggests that it is practically infeasible to test how the reality difffers from Schrödinger-Newton.

[20] argues that the spatial resolution limit is enough to protect against semiclassical gravity paradoxes. See [21] for a non-relativistic defense of semiclassical gravity. See [22] for a method for distinguishing a quantum system from a classical system, not restricted to

the question of quantum gravity against semiclassical gravity, for which efficiency may be limited.

Recently, [9] argues that semiclassical gravity is very defensible if we take matter-gravity interactions to be classically stochastic and proposes a consistent way of doing this - postquantum theory of gravity - avoiding problems noted in the PT argument. This approach suggests that information is lost whenever black holes evaporate, and [23] suggests information loss as a feature, not a bug.

There have also been very popular and powerful arguments that quantum gravity has the gravitational Gauss law such that information is always available at asymptotic infinity, and the black hole exterior is already purified such that the black hole information problem is non-existent from the start [24–29]. While acknowledging holography of information, it still remains how the information problem is resolved around the black hole horizon where we still non-gravitational quantum field theory to provide accurate approximations. It is in this sense that this paper probes the Page curve.

III. A JAYNES-CUMMINGS (JCM) TOY MODEL OF QUANTUM GRAVITY WITHIN BLACK HOLE SPACETIMES AND A NO-GO AGAINST SEMICLASSICAL GRAVITY

A. The black hole information problem and the small corrections theorem re-cast as no-go thought experiments against semiclassical gravity

The very problem with the black hole information problem [30] is that if we assume the semiclassical gravity picture of black hole evaporation, then the evolution of the black hole exterior becomes non-unitary after complete evaporation. In this case, it is not the assumption of semiclassical gravity that drives the paradox but the expectation that semiclassical approximations should be correct enough [31]. However, we can obviously treat this as a no-go experiment against semiclassical gravity as well.

The small corrections theorem [32] suggests that small corrections to the semiclassical effective field theory picture cannot solve the non-unitarity problem, and we briefly describe its toy qubit model logic.

At each step with interval Δt, we consider a single qubit pair to be produced around the black hole horizon, with one qubit falling into the black hole and the other radiated away to the exterior. At step n, counted from step 1, this can be described as:

$$|\psi^{eff}_{pair,n}\rangle = \cos(\phi)|0^{eff}_{R,n}\rangle|0^{eff}_{BH,b}\rangle + \sin(\phi)|1^{eff}_{R,n}\rangle|1^{eff}_{BH,n}\rangle + O(\epsilon), \quad \epsilon \ll 1 \tag{10}$$

where R represents the black hole exterior radiation(s) and BH represents the black hole interior, with eff representing effective degrees of freedom. von Neumann entropy-wise, we have:

$$\begin{gathered} S(R^{eff}_{n+1} + BH^{eff}_{n+1}) = 0 \\ S(BH^{eff}_{n+1}) = -[\cos^2(\phi)\ln(\cos^2(\phi)) + \sin^2(\phi)\ln(\sin^2(\phi))] \\ S(R_{n+1} + BH_{n+1}) = \epsilon_1 \\ S(BH_{n+1}) > -[\cos^2(\phi)\ln(\cos^2(\phi)) + \sin^2(\phi)\ln(\sin^2(\phi))] - \epsilon_2, \quad \epsilon_1, \epsilon_2 \ll 1 \end{gathered} \tag{11}$$

where R_{n+1} represents the radiation qubit in the exact theory at step $n+1$ and so forth. Now invoke the strong subadditivity relation (for subspaces with independent degrees of freedom) to prove the small corrections theorem:

$$S(A+B) + S(B+C) \geq S(A) + S(C) \tag{12}$$

Let $S(R_{\leq n})$ refer to the entropy of all radiations to step n in the exact theory. Then:

$$\begin{gathered} S(R_{\leq n} + R_{n+1}) + S(R_{n+1} + BH_{n+1}) \geq S(R_{\leq n}) + S(BH_{n+1}) \\ \Rightarrow S(R_{\leq n} + R_{n+1}) \geq S(R_{\leq n}) + S(BH_{n+1}) - S(R_{n+1} + BH_{n+1}) \\ \Rightarrow S(R_{\leq n} + R_{n+1}) \geq S(R_{\leq n}) - [\cos^2(\phi)\ln(\cos^2(\phi)) + \sin^2(\phi)\ln(\sin^2(\phi))] - (\epsilon_1 + \epsilon_2) \end{gathered} \tag{13}$$

Therefore, the entropy of all radiations must keep increasing monotonically, as long as $-[\cos^2(\phi)\ln(\cos^2(\phi)) + \sin^2(\phi)\ln(\sin^2(\phi))] > \epsilon_1 + \epsilon_2$. The original small corrections theorem takes $\phi = \pi/4$, but this is largely about analytic simplicity.

So far, we cannot compare against when gravity is quantized. Indeed, there is a difficulty in constructing even toy models of when gravity is genuinely quantized, and this paper intends to provide a good toy model capturing important details of quantum gravity. And when the comparison is made, we find that quantum gravity restores unitarity for the black hole exterior, while semiclassical gravity cannot.

In the toy model, we show that without violating EFT, the below can be violated at late time evaporation, expressed as an inequality:

$$S(BH^{eff}_{n+1}) \neq -[\cos^2(\phi)\ln(\cos^2(\phi)) + \sin^2(\phi)\ln(\sin^2(\phi))] \tag{14}$$

ϵ_1, ϵ_2 are irrelevant for rest of the paper - we set $\epsilon_1, \epsilon_2 = 0$.

B. JCM toy model intuitions

1. *On correcting for the empty black hole*

The intuition in the JCM toy model is simple. Effective field theory (EFT) on black hole spacetimes predicts long-lasting entanglement (and radiations) at asymptotic infinity whenever a black hole is non-empty, and its radiation temperature actually increases as the black hole becomes smaller. In contrast, EFT predicts no long-lasting entanglement for an empty black hole.

Slightly more technically, the temperature limit of Hawking radiations (in Schwarzschild spacetimes) as black holes become empty is infinity, but the metric-wise limit of Schwarzschild spacetimes as black holes become empty is Minkowski spacetime. These limits produce different entanglement stories, and since the latter holds whenever a black hole is actually empty, we have to provide corrections whenever semiclassical approximations are used. And this correction serves as a truly quantum gravity correction.

If there is a unique empty black hole state, as stated in the no-hair principle, then unitarity is demonstrably restored for the black hole exterior in (the toy model of) quantum gravity. If there are multiple black hole (interior) ground states, then the JCM toy model of this paper is inapplicable.

It is incredibly difficult to calculate the corrections within EFT, and approximate classicality of the underlying spacetimes has served to justify semiclasssical approximations [31]. The toy model of this paper shows that in spite of approximate classicality of background spacetimes, the empty black hole corrections affect calculations for late-time evaporation.

2. *Unique ground state energy logic*

Suppose we divide the universe Hamiltonian H as $H = H_{ext} + H_{BH} + H_I$, where H_{ext} stands for the free Hamiltonian of the black hole exterior, H_{BH} is the free Hamiltonian of the black hole interior and H_I is the interaction term. Now suppose that the exterior remains in the ground state around the black hole horizon - any excitation is washed away as radiation. Then because there is no energy to take away from the exterior, only the black hole (interior) can emit energy. With sufficiently general decoherence induced by H_I, the black hole has to continuously emit energy until it becomes empty.

When the black hole does become completely empty (in any quantum branch), then if the no-hair principle of a unique ground state holds for H_{BH}, the black hole interior carries a pure state, which implies that the black hole exterior is purified.

For semiclassical gravity, we cannot correct for the empty black hole quantum branch and the small corrections theorem kicks in such that the black hole entanglement entropy increases monotonically up to complete evaporation. Semiclassical H_I cannot account for non-generation of entanglement in case of the empty black hole outcome.

3. On quasi-classicality of background spacetime

It is easy to verify that quasi-classicality of background spacetime is satisfied in the toy model. We consider each small time sub-interval for which interaction behavior remains almost identical. This allows us to use the central limit theorem to treat the energy outflow of a sub-interval (and the union of sub-intervals as well) in terms of a normal distribution for early times.

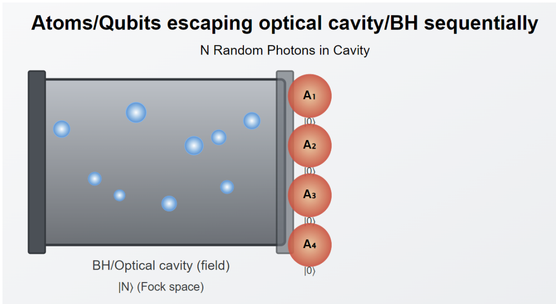
Around the entropy turnover time ('Page time'), entanglement entropy decreases, and since entropy serves as a proxy for variance of the total energy emitted as well, variance is bounded, and quasi-classicality holds down to complete evaporation. Note that by the no-hair principle, without charge and angular momentum, stable metric depends only on the mass-energy of the black hole.

We can further reduce variance of each sub-interval by tuning ϕ in Equation (10) such that we may practically assert complete classicality within each sub-interval.

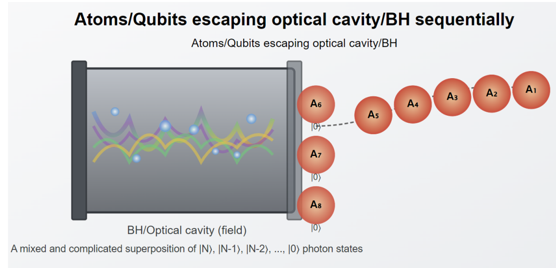
4. Additional remarks

Even before stating the full JCM toy model, these theoretical intuitions are sufficient to provide the conclusion that gravity needs to be quantized if the black hole exterior is to be purified at the end of black hole evaporation. We now proceed to the specification of the JCM toy model.

We also note that while the toy model of this paper is significantly different, the Jaynes-



(a) Initial evaporation configuration



(b) Late-time evaporation configuration

FIG. 1: JCM black hole model configurations at initial and late times

Cummings (JCM) model had been used as a toy model of black holes in [33].

C. JCM toy model setup

Each kth qubit (atom in original JCM), initially in $|0\rangle$ interacts sequentially with the black hole (optical cavity in original JCM) 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 representing the number of particles in the black hole. There are M qubits.

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

$$H_{rad} = \omega_a \sigma_+ \sigma_- = \omega_a |1\rangle\langle 1| \tag{15}$$

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:

$$H_{BH} = \omega_a a^\dagger a \tag{16}$$

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 $\hat{n}|N\rangle = N|N\rangle$.

Now most importantly, JCM interaction Hamiltonian H_I that couples a single qubit and the black hole (when under interaction) goes as:

$$H_I = g(\sigma_+ a + \sigma_- a^\dagger) \tag{17}$$

Since we have same ω_a in the free Hamiltonians, this JCM 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 rest of the discussions.

D. Variable interaction time (Δt_k) case

The final outgoing state of the kth qubit is determined with unitary U_k

$$U_k|0,n\rangle = \cos(g\sqrt{n}\Delta t_k)|0,n\rangle - i\sin(g\sqrt{n}\Delta t_k)|1,n-1\rangle \tag{18}$$

where $|q_1,q_2\rangle$ refers to $q_1 \in \{0,1\}$ (kth qubit), $q_2 \in \{0,1,..,N\}$ (black hole). For variable interaction time Δt_k, we choose

$$\Delta t_k = \frac{\pi}{4g\sqrt{\langle n_k\rangle}} \tag{19}$$

where $\langle n_k\rangle$ is average excitation numbers in the black hole at the start of the kth interaction. This matches the semiclassical requirement of the small corrections theorem that is reproduced up to phase for $\langle n_k\rangle \in \mathbb{N}$ (with $\phi = \pi/4$):

$$U_k|0,\langle n_k\rangle\rangle \approx \frac{1}{\sqrt{2}}\left(|0,\langle n_k\rangle\rangle + |1,\langle n_k\rangle - 1\rangle\right) \tag{20}$$

E. Fixed interaction time $\Delta t_k = \Delta t$ case

We can instead fix interaction time as constant $\Delta t_k = \Delta t$ ($U_k = U$) such that:

$$\Delta t = \frac{\pi}{4g\sqrt{N}} \tag{21}$$

reproducing only for $n = N$ up to phase,

$$U|0,N\rangle \approx \frac{1}{\sqrt{2}}\left(|0,N\rangle + |1,N-1\rangle\right) \tag{22}$$

F. Simplified model

We can modify interaction unitary $U_k = U$ (fixed interaction time case) to be slightly different from the JCM case:

$$\begin{aligned} U|0,n\rangle &= \frac{1}{\sqrt{2}}\left(|0,n\rangle + |1,n-1\rangle\right) \quad (0 < n) \\ U|0,0\rangle &= |0,0\rangle \end{aligned} \tag{23}$$

This simplified model continues to respect conservation of non-interaction energy. The model is analytically simpler and can be beneficial for those insisting on the same interaction behavior regardless of n. We arrive at the same conclusions qualitatively regardless of the models used.

G. Reduced interaction time case: semiclassical spacetime upheld

Interaction time Δt_k or coupling constant g can be reduced by factor k_r such that more M is allowed for the given time. As M becomes larger, the total energy emitted becomes more semiclassical due to the central limit theorem (CLT). This is simplest to explore in the simplified model, though qualitatively conclusions remain the same. U now instead generates in the simplified model of Section III F:

$$\begin{gathered} U|0,n\rangle = \cos(\pi/(4k_r))|0,n\rangle + \sin(\pi/(4k_r))|1,n-1\rangle \\ U|0,0\rangle = |0,0\rangle \end{gathered} \tag{24}$$

If we can ignore the empty black hole case, variance in energy emitted per each qubit $Var[E_{k,rad}]$ goes as:

$$Var[E_{k,rad}] = \langle (E_{k,rad})^2 \rangle - (\langle E_{k,rad} \rangle)^2 = \sin^2\left(\frac{\pi}{4k_r}\right) - \sin^4\left(\frac{\pi}{4k_r}\right) \tag{25}$$

CLT says that for K such identical qubits, with $\sigma^2 = Var[E_{k,rad}]$:

$$E_{rad,total,k} - K\langle E_{k,rad} \rangle \xrightarrow{d} \mathcal{N}\left(0, K\sigma^2\right) \tag{26}$$

It can be noted that variance is lowered approximately by factor of $(1/k_r)^2$ as k_r is varied, while interaction time is reduced by $1/k_r$, achieving the goal of enlarging the number of interacting qubits within the same time duration and lowering total variance. The point is that with large k_r, we see practically zero variance in total energy emitted from the semiclassical prediction.

In the models, σ eventually changes over time. However, we can slice time into different small sub-intervals of sufficiently large qubits being emitted and apply CLT to each sub-interval. The CLT conclusion therefore remains.

H. Simulation results and analysis

Model simulation results are provided in Figure 2, confirming unitarity even after complete black hole evaporation, regardless of the full JCM or the simplified model, reduction in interaction time or not. Since the interaction time reduction does not change unitarity, semiclassical spacetime is maintained by the discussion in Section III G.

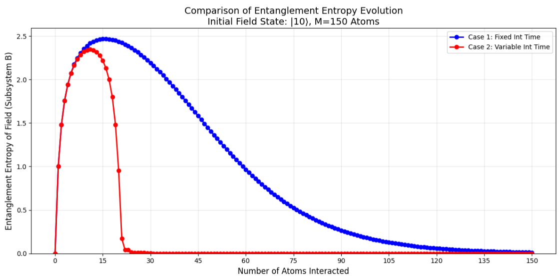
I. Is a collapse mechanism really irrelevant in the JCM toy model?

If we believe that a black hole is a classical system such that we may need to attempt modeling it in the fashion of the PG thought experiment, then a collapse mechanism can spoil this no-go thought experiment.

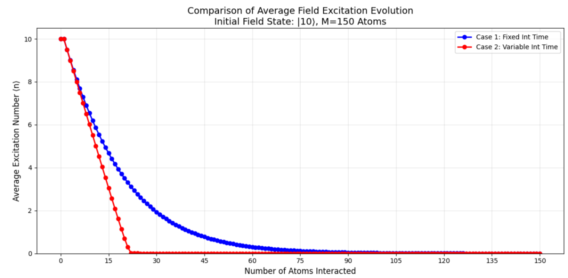
More generally, if our expectation about effective field theory interaction turns out to be wrong, then we cannot clearly separate the semiclassical gravity case from the quantum gravity case, and the no-go thought experiment does not work. In such a case, collapse mechanisms may be relevant. Otherwise, they are irrelevant in this thought experiment.

IV. CONCLUSION

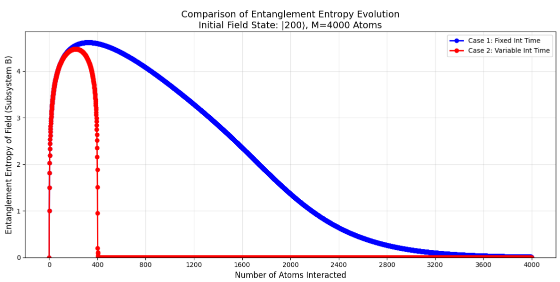
Well-established no-go thought experiments against semiclassical gravity are susceptible to different descriptions of wavefunction collapse mechanisms. This paper utilizes no measurement and is not vulnerable to collapse mechanisms, as long as the usual expectation about effective field theory interactions between the black hole exterior and the interior remains accurate. If the black hole exterior does not need to be purified at the end, then the necessity of quantizing gravity follows, with semiclassical gravity only worsening the matter due to the original Hawking argument [30] and the small corrections theorem. Quantum gravity allows for purification by distinguishing the (unique) empty black hole interior case with other non-empty black hole cases.



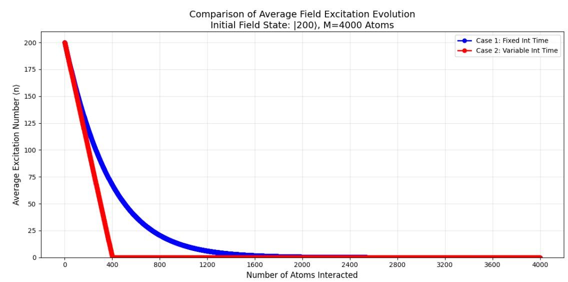
(a) Black hole entanglement entropy evolution. ($N = 10$, $M = 150$, $g = 1$, $k_r = 1$, full JCM)



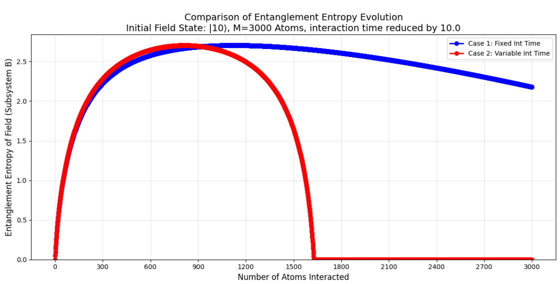
(b) Average black hole excitation number $\langle n \rangle$ plot. ($N = 10$, $M = 150$, $g = 1$, $k_r = 1$, full JCM)



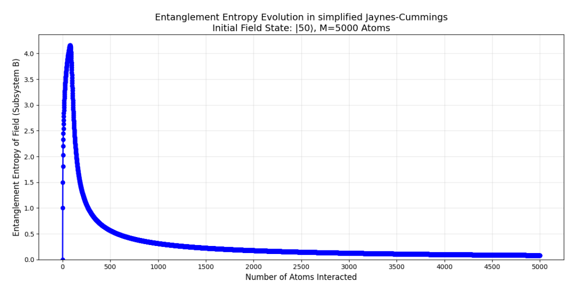
(c) Black hole entanglement entropy evolution. ($N = 200$, $M = 4000$, $g = 1$, $k_r = 1$, full JCM)



(d) Average black hole excitation number $\langle n \rangle$ plot. ($N = 200$, $M = 4000$, $g = 1$, $k_r = 1$, full JCM)



(e) Black hole entanglement entropy evolution. ($N = 10$, $M = 3000$, $g = 1$, $k_r = 10$, full JCM)



(f) Black hole entanglement entropy evolution. ($N = 50$, $M = 5000$, $g = 1$, $k_r = 1$, simplified model)

FIG. 2: Model simulations. For Figure 2a, 2b, 2c, 2d and 2e, the red and symmetric curves represent the variable interaction time case, while the blue curves represent the fixed interaction time case.

DATA AVAILABILITY AND DECLARATION OF INTERESTS

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

[1] Nick Huggett and Craig Callender, "Why quantize gravity (or any other field for that matter)?" Philosophy of Science **68** (2001), 10.1086/392923.

[2] Christian Wüthrich, "To quantize or not to quantize: Fact and folklore in quantum gravity," Philosophy of Science **72**, 777–788 (2005).

[3] James Mattingly, "Is quantum gravity necessary?" in *The Universe of General Relativity*, edited by A. J. Kox and Jean Eisenstaedt (Birkhäuser Boston, Boston, MA, 2005) pp. 327–338.

[4] Karen Crowther, *Why Do We Want a Theory of Quantum Gravity?*, Elements in the Philosophy of Physics (Cambridge University Press, 2025).

[5] David Wallace, "Taking particle physics seriously: A critique of the algebraic approach to quantum field theory," Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics **42**, 116–125 (2011).

[6] Romeo Brunetti, Klaus Fredenhagen, and Rainer Verch, "The generally covariant locality principle – a new paradigm for local quantum field theory," Communications in Mathematical Physics **237**, 31–68 (2003).

[7] Stefan Hollands and Robert M. Wald, "Axiomatic quantum field theory in curved spacetime," Communications in Mathematical Physics **293**, 85–125 (2009).

[8] Lajos Diósi, "A healthier stochastic semiclassical gravity: world without schrödinger cats," General Relativity and Gravitation **57** (2025), 10.1007/s10714-025-03396-z.

[9] Jonathan Oppenheim, "A postquantum theory of classical gravity?" Phys. Rev. X **13** (2023), 10.1103/PhysRevX.13.041040.

[10] Suvrat Raju and Pushkal Shrivastava, "Critique of the fuzzball program," Physical Review D **99** (2019), 10.1103/physrevd.99.066009.

[11] Niels Bohr and Léon Rosenfeld, "On the question of the measurability of electromagnetic field quantities," in *Foundations of Quantum Physics II (1933–1958)*, Niels Bohr Collected Works,

Vol. 7, edited by Jørgen Kalckar (Elsevier, 1996) pp. 123–166.

[12] Léon Rosenfeld, "On quantization of fields," Nuclear Physics **40**, 353–356 (1963).

[13] Richard P. Feynman, "The role of gravitation in physics," in *Magic Without Magic: John Archibald Wheeler*, edited by J. Klauder (Freeman, San Francisco, 1972) pp. 434–467.

[14] Don N. Page and C. D. Geilker, "Indirect evidence for quantum gravity," Phys. Rev. Lett. **47**, 979–982 (1981).

[15] Kenneth Eppley and Eric Hannah, "The necessity of quantizing the gravitational field," Foundations of Physics **7**, 51–68 (1977).

[16] Adrian Kent, "Simple refutation of the Eppley–Hannah argument," Classical and Quantum Gravity **35** (2018), 10.1088/1361-6382/aaea20.

[17] André Großardt, "Three little paradoxes: Making sense of semiclassical gravity," AVS Quantum Science **4** (2022), 10.1116/5.0073509.

[18] Asher Peres and Daniel R. Terno, "Hybrid classical-quantum dynamics," Phys. Rev. A **63** (2001), 10.1103/PhysRevA.63.022101.

[19] S Carlip, "Is quantum gravity necessary?" Classical and Quantum Gravity **25** (2008), 10.1088/0264-9381/25/15/154010.

[20] Erik Rydving, Erik Aurell, and Igor Pikovski, "Do gedanken experiments compel quantization of gravity?" Physical Review D **104** (2021), 10.1103/physrevd.104.086024.

[21] Antoine Tilloy, "Does gravity have to be quantized? lessons from non-relativistic toy models," Journal of Physics: Conference Series **1275** (2019), 10.1088/1742-6596/1275/1/012006.

[22] C. Marletto and V. Vedral, "Witnessing the quantumness of a system by observing only its classical features," npj Quantum Information **3** (2017), 10.1038/s41534-017-0040-4.

[23] Jonathan Oppenheim, "Is it time to rethink quantum gravity?" International Journal of Modern Physics D **32** (2023), 10.1142/s0218271823420245.

[24] Steven B. Giddings, Donald Marolf, and James B. Hartle, "Observables in effective gravity," Physical Review D **74** (2006), 10.1103/physrevd.74.064018.

[25] Tomás Andrade, Donald Marolf, and Cédric Deffayet, "Can hamiltonians be boundary observables in parametrized field theories?" Classical and Quantum Gravity **28** (2011), 10.1088/0264-9381/28/10/105002.

[26] Donald Marolf, "Holography without strings?" Classical and Quantum Gravity **31**, 015008 (2013).

[27] Chandramouli Chowdhury, Olga Papadoulaki, and Suvrat Raju, "A physical protocol for observers near the boundary to obtain bulk information in quantum gravity," SciPost Phys. **10**, 106 (2021).

[28] Alok Laddha, Siddharth Prabhu, Suvrat Raju, and Pushkal Shrivastava, "The holographic nature of null infinity," SciPost Physics **10** (2021), 10.21468/scipostphys.10.2.041.

[29] Suvrat Raju, "Failure of the split property in gravity and the information paradox," Classical and Quantum Gravity **39** (2022), 10.1088/1361-6382/ac482b.

[30] Stephen W. Hawking, "Particle creation by black holes," Communications in Mathematical Physics **43**, 199–220 (1975).

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

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

[33] Wen-Yu Wen, "Hawking radiation as stimulated emission," Physics Letters B **803** (2020), 10.1016/j.physletb.2020.135348.