Quantum Chaos and Unitary Black Hole Evaporation

David A. Lowe

Physics Department, Brown University, Providence, RI 02912, USA.

Lárus Thorlacius

Science Institute, University of Iceland,
Dunhaga 3, 107 Reykjavík, Iceland and
Nordita, Stockholm University and KTH Royal Institute of Technology,
Hannes Alfvéns väg 12, SE-106 91 Stockholm, Sweden

Abstract

The formation and evaporation of small AdS black holes in a theory with a holographic dual is governed by the usual rules of quantum mechanics. The eigenstate thermalization hypothesis explains the validity of semiclassical gravity for local bulk observables and can be used to quantify the magnitude of quantum corrections to the semiclassical approximation. The holographic dual produces a basis of black hole states with finite energy width, and observables that are smooth functions on the classical phase space will self-average over a large number of energy eigenstates, exponential in the Bekenstein-Hawking entropy $S$, leading to results that are consistent with semiclassical gravity up to small corrections of order $e^{-S/2}$. As expected, the semiclassical description breaks down for transition amplitudes that contribute to the unitary S matrix of the holographic theory.
I. INTRODUCTION

We revisit the black hole information problem [1] in the context of the AdS/CFT correspondence [2–4]. In this setting, the dual gauge theory provides a unitary time evolution by construction but the implementation of unitarity on the bulk gravitational side is less clear. Much of the work involving black holes in AdS/CFT has focused on ones that are large compared to the AdS length scale and well described using finite temperature quantum field theory [5]. However, the standard formulation of the information problem involves the formation and subsequent evaporation of a macroscopic black hole, i.e. large compared to the Planck scale, in asymptotically flat spacetime. In an asymptotically AdS background this process is most closely modeled by considering a macroscopic black hole that is small compared to the characteristic AdS length scale [6]. The black hole is taken to be macroscopic so that it carries a significant amount of information and evaporates slowly, but it is taken to be small enough for the black hole lifetime to be short compared to the AdS light crossing time to avoid issues arising from the global causal structure of AdS spacetime. This requires a large but finite separation between the AdS scale and the Planck scale on the bulk side of the duality.

One goal of the present work is to extend holographic methods to the study of small AdS black holes by combining Mukhanov’s approach to counting black hole states [7] and the HKLL construction [8, 9] for mapping small perturbations in the asymptotic region of the spacetime into the dual gauge theory. A similar construction for asymptotically flat spacetime was recently given in [10, 11] based on the conformal symmetry of the celestial sphere [12]. Further progress relating asymptotically flat holography to the vanishing cosmological constant limit of AdS/CFT can be found in [13] and it is straightforward to extend our results to that setting. With a holographic map for small AdS black holes in place, we can apply the rules of quantum mechanics and methods from quantum chaos theory to estimate the size of corrections to the semiclassical approximation for observables outside the black hole as well as for observables measured by infalling observers entering the black hole.

In order to keep the presentation as simple as possible, we will mainly focus on small AdS black holes in 3+1 dimensional spacetime but our arguments can be extended to general dimensions. In fact, depending on the precise holographic setup, the small black hole Schwarzschild radius may also be small compared to the characteristic size of the geometry.
transverse to AdS in a consistent string or M theory background and in this case the favored configuration is a higher-dimensional black hole that is localized on the transverse compact space \[14\]. Fortunately, the two main ingredients in our construction, i.e. Mukhanov’s black hole counting and the HKLL formalism, are both easily adapted to general dimensions and our arguments go through in this case as well.

The paper is organized as follows. In Section II we briefly review Mukhanov’s approach to counting black hole states and then outline the construction of finite-width wave packets formed out of quasi-local bulk operators that satisfy the spatial and momentum constraints of Mukhanov states. In Section III we turn our attention to the formation and evaporation of small AdS black holes from a holographic perspective. A key observation is that due to the finite black hole lifetime the gauge theory state that corresponds to an isolated black hole has a natural energy width and necessarily involves a superposition of a large number of energy eigenstates. The energy eigenstates themselves do not correspond to smooth semiclassical geometries but a semiclassical description for appropriately defined black hole observables emerges via eigenstate thermalization. In Section IV we focus on observables that can be measured by infalling observers in a laboratory that enters the black hole. Local bulk operators can be constructed and evolved across the horizon along timelike geodesics and we argue that their expectation values can be computed to a good approximation by semiclassical methods until they approach the region of strong curvature near the singularity. Section V contains some concluding remarks.

II. HOLOGRAPHIC BLACK HOLE EVOLUTION

Consider the formation of an isolated small AdS black hole in a theory with a holographic dual. We would like to identify a set of semiclassical bulk initial states that can account for the black hole entropy and map them to states of the dual field theory. Here we have in mind realizing these states within a single CFT dual to a single asymptotically AdS region. The bulk states will evolve forward in a semiclassical geometric theory while the corresponding field theory states undergo unitary time evolution. By evaluating expectation values of observables we can estimate the size of the corrections to the semiclassical approximation that are imposed by unitarity.

As was observed by Mukhanov \[7\], the overwhelming majority of initial states in the bulk
that contribute to the black hole entropy will correspond to the time-reverse of outgoing states produced by the evaporation of a black hole of the same mass, i.e. states consisting of a collapsing cloud of out-of-equilibrium radiation directed toward the location where the black hole forms over a time period of order the black hole lifetime. The set of possible initial states of course also includes low-entropy configurations like incoming spherical shells of matter or stellar interiors undergoing gravitational collapse but the dominant contribution to the black hole entropy comes from incoming radiation states of the form considered by Mukhanov.

To enumerate the incoming Mukhanov states, we find it more convenient to consider the equivalent counting problem for outgoing radiation states. The lifetime of a black hole of mass $M$ in 3+1 dimensional spacetime is

$$D \sim M^3.$$  \hspace{1cm} (1)

The entropy of the black hole is correctly computed if we simply calculate the out-of-equilibrium entropy of a quasi-thermal gas of photons, imagining it to be released outward into a region of radius $D$. The energy of the gas must match the energy of the black hole

$$M \sim T^4 V \delta \Omega,$$  \hspace{1cm} (2)

where $T$ is the temperature of the gas, and $\delta \Omega$ is the solid angle subtended by the past projection of the black hole horizon. In this case we have

$$\delta \Omega \sim \frac{M^2}{D^2} \sim \frac{1}{M^4}.$$  \hspace{1cm} (3)

Substituting in we then find that

$$T \sim 1/M,$$  \hspace{1cm} (4)

and the entropy of the gas of radiation

$$S \sim E/T \sim M^2.$$  \hspace{1cm} (5)

A more precise calculation would take into account that the temperature of the gas changes as the black hole evaporates, modulating the energy density of the outgoing radiation.

We note that if the black hole energy changes by $\delta M$, then the entropy of the black hole changes by

$$\delta S = \frac{\delta M}{T_H},$$  \hspace{1cm} (6)
where $T_H$ is the Hawking temperature for a mass $M$ black hole. Thus when emitted, the non-equilibrium outgoing radiation will carry off this entropy together with any extra contribution arising from irreversibility in accord with the generalized second law. In four-dimensional spacetime, the entropy of this quasi-thermal radiation, when treated like black body radiation, becomes

$$\delta S = \frac{4}{3} \frac{\delta M}{T}. \quad (7)$$

The extra entropy production in (7) versus (6) can be viewed as an artifact of the semiclassical approximation, where the emitted radiation contains no mutual correlations [15].

In [7] there is an attempt to match this equilibrium entropy (7) with (6) by adjusting the temperature $T > T_H$. From our perspective, this is unnecessary since the time reverse of the emitted black hole radiation will not obey (7) but rather (6) in the context of unitary evolution. If we consider the full set of ingoing states obeying (7), the important point is that we obtain a overcomplete set of states for forming small black holes. We expect the states accounting for the discrepancy between (6) and (7) will not form black holes, and instead scatter back out on a much shorter timescale when quantum effects are accounted for.

In effect, the Mukhanov conditions amount to requiring the outgoing states are selected from an ensemble of pure states with energy $E = M$ in a region of finite size $D \sim M^3$, with the additional solid angle constraint (8), and the momentum to be predominantly in the radial direction. Such constraints are straightforward to implement using wavepackets of finite width to simultaneously implement the spatial and momentum constraints. These quasi-local operators can be constructed directly in holographic theories (assuming the matching between gravitational and holographic quantities is precise far from the black hole). For example, in the case of AdS/CFT the methods of HKLL [8, 9] can be applied provided $D \gg l_{Pl}$ where $l_{Pl}$ is the Planck length. The Mukhanov conditions will be useful provided the AdS radius of curvature satisfies $R_{AdS} \gg D$.

More specifically, HKLL construct bulk operators as smeared integrals of CFT primaries over compact spacetime regions on the boundary. The resulting expressions are covariant under the conformal group. If one introduces a time-slicing on the boundary, these smeared

---

1 In $d$ spatial dimensions the prefactor is $\frac{(d+1)}{d}$.

2 The region under consideration can be extended to sizes larger than $D$ but in this case the cloud of outgoing Hawking radiation will occupy an outgoing shell of width $D$ with vacuum outside and inside the shell. The radiation is free streaming so the entropy does not change.
integrals can be further localized to a narrow range of time on the boundary, by using free boundary propagation. This is sufficient to build a boundary state, corresponding to the outgoing Mukhanov state. The time reverse of an outgoing Mukhanov state provides an initial state for black hole formation and the procedure may be repeated for this state and arranged so there is no temporal overlap with the final state operator insertions. Our construction thus provides a map for both in- and out-going bulk states, that only have support away from the black hole region and are well described by the semiclassical gravitational theory, into corresponding states in the boundary theory which undergo manifestly unitary time evolution. The bulk states can also be evolved in the semiclassical theory and this bulk evolution will eventually diverge from unitary evolution generated by the boundary theory. In the following, we will use the map from the bulk to the boundary and methods from quantum chaos theory to estimate the size of corrections to the semiclassical approximation for appropriately defined observables.

III. BLACK HOLE EIGENSTATE THERMALIZATION

The eigenstate thermalization hypothesis [16] explains how an isolated quantum system, which is prepared in a far-from-equilibrium initial state, can evolve to a state that is hard to distinguish from thermal equilibrium. The hypothesis, which is expected to hold for quantum systems whose classical limit exhibits dynamical chaos, originates from the work of Berry [17] and was further developed by Deutsch [18] and Srednicki [16, 19]. It has been validated in various model systems but not proven rigorously.

Analogous behavior arises in gravitational physics when non-thermal matter undergoes gravitational collapse to a black hole that exhibits thermal properties. We propose that in a gravitational theory with a holographic dual this is more than an analogy and that eigenstate thermalization can explain the validity of semiclassical gravity for observables that involve a relatively small number of fields. In this case, an isolated black hole formed by gravitational collapse is described by a state in the dual field theory that lives in a Hilbert subspace of dimension $e^{S(E)}$ where $E$ is the energy of the black hole and $S$ is the Bekenstein-Hawking entropy. The eigenstate thermalization hypothesis can be expressed as follows [19]: For any observable $A$ which is a smooth function of the classical phase space coordinates, we assume the matrix elements between energy eigenstates will take the form
$A_{\alpha\beta} = A(E_\alpha)\delta_{\alpha\beta} + e^{-S/2}((E_\alpha+E_\beta)/2)R_{\alpha\beta}, \quad (8)$

where $A(E_\alpha)$ is a smooth function of the energy and $R_{\alpha\beta}$ is a random matrix, whose matrix elements are drawn from a Gaussian distribution with variance of order one. There is a further assumption here that the number of choices for observables $A$ is not exponential in the system size. The $e^{-S/2}$ factor in front of the off-diagonal matrix elements can be understood at a qualitative level by observing that for generic $A$ the matrix elements of $A^2$ should also satisfy (8). To see this, one inserts (8) twice on the right hand side of $(A^2)_{\alpha\beta} = \sum_\gamma A_{\alpha\gamma}A_{\gamma\beta}$ and carries out the sum over $\gamma$. The exponentially large number of terms in the sum combined with elementary properties of random matrices gives an estimate that precisely offsets the suppression from the $e^{-S/2}$ prefactors that accompany the off-diagonal matrix elements [19]. Note that it is only necessary to assume (8) holds for the Hilbert subspace relevant for an isolated black hole, together with its formation and evaporation products. It need not hold for the entire Hilbert space of the complete quantum theory.

Now let us consider eigenstate thermalization in the context of black hole evolution. We have in mind an isolated black hole, formed from an initial pure state $|\psi\rangle$ in gravitational collapse that is reasonably well-localized in time, and then allowed to evaporate down to nothing without further disturbing it. An initial state of this form can always be expanded on the basis of states provided by the holographic construction in Section (II) but we will also assume that $|\psi\rangle$ is well-described by semiclassical evolution in the asymptotic region. We then want to address the question of how long and to what extent the semiclassical evolution remains faithful to the holographic evolution of the state forward in time, both outside and inside the black hole. Of course, it is possible to choose initial states for which the semiclassical approximation fails from the outset but then our geometric description also fails from the outset. A simple example would be a superposition of equal mass states localized in different points in space [20]. The holographic theory nevertheless provides a complete description of the time evolution of such states.

The matrix elements in (8) are between energy eigenstates $|\alpha\rangle$ whose detailed structure depends on the details of the holographic theory. It is worth noting that the energy eigenstates themselves will not have a simple geometric/semiclassical description. Rather they are stationary states formed by superposition of ingoing and outgoing Mukhanov states for
black hole formation at all different times. The initial state $|\psi\rangle = \sum_\alpha C_\alpha |\alpha\rangle$ has an energy width given by

$$\Delta_\psi E = \left( \sum_\alpha |C_\alpha|^2 \left( E_\alpha - \langle E_\psi \rangle \right)^2 \right)^{1/2}, \quad (9)$$

which has a natural value of order $1/E^3$ due to the finite black hole lifetime. This is a narrow resonance but still $\Delta E$ is parametrically larger than $e^{-S(E)}$ and therefore expectation values can self-average over a large number of order $e^{S(E)}$ states. We also require the observable $A$ in (8) to be sufficiently smooth across the range of energies under consideration. The precise criterion is considered in [19] with the conclusion

$$\langle \Delta E \rangle^2 \ll \left| \frac{A(\langle E \rangle)}{A''(\langle E \rangle)} \right|^2. \quad (10)$$

This criterion is easily satisfied for a large class of observables in a black hole background where $\Delta E \sim 1/E^3$ and does not represent a severe restriction on the local operators available for semiclassical physics. In fact, eigenstate thermalization applies for states that are as broad as $\Delta E \lesssim O(1)$ but such a broad state would be short lived compared to the macroscopic black holes that are of primary interest here.

Now let us consider two different normalized pure states, $|\psi\rangle$ and $|\psi'\rangle$ satisfying the above criteria,

$$|\psi\rangle = \sum_\alpha C_\alpha |\alpha\rangle, \quad |\psi'\rangle = \sum_\alpha C'_\alpha |\alpha\rangle, \quad (11)$$

and compute the following quantity, representing the fluctuation of the expectation value of $A$ between the two states,

$$\delta A = \langle \psi | A | \psi \rangle - \langle \psi' | A | \psi' \rangle = \sum_{\alpha\beta} D_{\alpha\beta} e^{i(E_\alpha - E_\beta)t/\hbar} A_{\alpha\beta}, \quad (12)$$

where $D_{\alpha\beta} = C_\alpha^* C_\beta - C_\alpha^* C'_\beta$ and normalization implies $\sum_\alpha D_{\alpha\alpha} = 0$. We can estimate $\delta A$ by using (8) for the matrix elements in the energy eigenstate basis,

$$\delta A = \sum_{\alpha\beta} D_{\alpha\beta} \left( A(E_\alpha) \delta_{\alpha\beta} + e^{-S((E_\alpha + E_\beta)/2)/2} R_{\alpha\beta} \right) e^{i(E_\alpha - E_\beta)t/\hbar}. \quad (13)$$

---

3 Even with finite $\Delta E$ one can build states that do not self-average over many energy eigenstates. However such states will not satisfy the semiclassical approximation in the asymptotic region.
Consider the first term, arising from the diagonal contributions, which will be time independent. For simplicity, let us define \( \langle E \rangle_{\psi} = \langle E \rangle_{\psi'} \) but we will allow that \( \Delta_{\psi'} E = c \Delta_{\psi} E \) with \( c = O(1) \). Then the leading portion of the first term will be

\[
\delta A_{\text{diag}} = \frac{1}{2} \left( 1 - c^2 \right) (\Delta_{\psi} E)^2 A'' \left( \langle E \rangle_{\psi} \right),
\]

and then it directly follows from the criteria (10) constraining the width of the band of energy eigenstates that \( |\delta A_{\text{diag}}| \ll |A(\langle E \rangle)| \). The diagonal corrections depend on the energy profiles of the states \( |\psi\rangle \) and \( |\psi'\rangle \) through the slowly varying function \( A(\langle E \rangle_{\psi}) \) but are otherwise insensitive to the detailed properties of the microstates. It is natural to assume that the semiclassical approximation is capable of correctly computing this contribution, since it should correctly compute quantities, including finite size corrections, that depend on the expectation value of the energy. The RST model in two dimensional spacetime provides a detailed example where such finite size corrections are explicitly computable \[21\]. We also note that the finite size corrections emerging from 2d CFT show a similar behavior \[22\].

The off-diagonal corrections, however, depend on the random matrix \( R_{\alpha\beta} \) which depends sensitively on the choice of state. The semiclassical approximation will not capture those effects correctly. The size of the off-diagonal contribution to \( \delta A \) in (13) is governed by the random fluctuations of \( R_{\alpha\beta} \) and for typical microstates \( |\psi\rangle \) and \( |\psi'\rangle \) one obtains

\[
|\delta A_{\text{off-diag}}| \sim e^{-S(\langle E \rangle)/2}.
\]

The answer for the off-diagonal terms in (13) is time dependent but at \( t = 0 \) all the phase factors will be equal to one. On timescales of order the Heisenberg time \( t \sim e^S \) we expect this expectation value to recur.

The estimate in (15) will hold for most microstates but there are special states for which the off-diagonal contribution is considerably larger. The maximal possible value will be realized when either \( |\psi\rangle \) or \( |\psi'\rangle \) happens to be close to the eigenstate of the matrix \( R_{\alpha\beta} \) that belongs to its largest eigenvalue. It then follows from the Wigner semi-circle law \[23\] that \( |\delta A_{\text{off-diag}}|_{\text{max}} \sim O(1) \) but the probability for a randomly chosen microstate to be close to such an eigenstate is extremely small based on a Haar measure for \( U(e^S) \). Let us constrain \( |C_{\alpha} - \tilde{C}_{\alpha}| < \epsilon \ll 1 \), where \( \tilde{C}_{\alpha} \) specifies the eigenstate of \( R_{\alpha\beta} \) in question. Then the likelihood of picking the state would be of order \( e^{-S} \), which is negligible. We will therefore use (15) as an estimate of the off-diagonal term for the states of interest.
The exponentially small off-diagonal corrections to observables computed using a semi-classical approximation are impossible to detect unless one projects onto states that are very close to energy eigenstates. To probe how the semiclassical approximation would break down outside the black hole in an operationally well-defined way, we are led to consider initial and final states for which the off-diagonal corrections in (15) are comparable to the diagonal corrections coming from (14). To measure violations of the semiclassical approximation outside the black hole, one must in effect tune $\Delta E \lesssim e^{-S(E)/4}$ which requires enormously long timescales in the asymptotic region and presumably requires extremely large measuring apparatus and repeated measurements on identically prepared systems [24]. As noted above, an isolated black hole formed in gravitational collapse does not have such sharply tuned energy width and we can therefore expect the semiclassical approximation to be valid for the computation of expectation values of local operators throughout the geometry outside the black hole up to corrections of order $e^{-S(E)/2}$.

There are, however, key aspects of the physics of black hole evolution where the semiclassical approximation fails badly. This includes, for instance, transition amplitudes between in- and out-states. Consider an ingoing state $|\psi\rangle$ that is well-described by the semiclassical theory in the asymptotic region, along with a different outgoing state $|\psi'\rangle$ that also has a good semiclassical description and compute a matrix element of local operators,

$$\langle O(x)O(y) \rangle_{\psi\psi'} = \langle \psi'|O(x)O(y)|\psi\rangle.$$  

From the point of view of the semiclassical approximation, we could compute this correlator in two ways. First, begin with the initial state $|\psi\rangle$, evolve it forward in time, producing an outgoing thermal flux of radiation in accord with Hawking’s original computation [25]. The correlator can then be evaluated as an expectation value in this semiclassical background.

On the other hand, we can also begin with the bra-state $\langle \psi'|$ and evolve that backward in time according to the same set of rules, producing a thermal incoming flux of radiation in accord with the time reverse of Hawking’s calculation. The states $|\psi\rangle$ and $|\psi'\rangle$ can be any states satisfying the Mukhanov conditions, which allows a wide variation in the energy fluxes, subject to the average null energy condition. In general, the answers obtained by the two semiclassical computations will disagree by relative errors of $\mathcal{O}(1)$ throughout the

---

4 We will turn our attention to observables for infalling observers in Section IV below.

5 The RST model [21] provides an example where back-reaction can be fully incorporated at the semiclassical level to justify this assumption.
geometry. In the asymptotic region, the predicted answers will disagree by relative errors of $O(1)$ simply because the expectation value of the incoming flux associated with the state $|\psi\rangle$, which is the input to the semiclassical solution of the gravitational equations, will differ by corrections of relative $O(1)$ with the back-evolved prediction for the expectation value of the incoming flux associated with the final state $|\psi'\rangle$, and vice versa. Likewise, near the apparent horizon the answers will disagree by errors of $O(1)$. We conclude that even if $|\psi\rangle$ and $|\psi'\rangle$ lead to well-behaved semiclassical geometries in the past/future asymptotic region respectively, the semiclassical approximation fails when computing the correlator (16). This is of course just restating the information problem, which amounts in this context to the failure of semiclassical physics to reproduce the unitary $S$ matrix of the holographic dual theory.

Another example where semiclassical physics might be expected to fail is in computing non-local quantities such as entanglement entropy between the black hole and the outgoing radiation. Since we assume the black hole involves a finite dimensional Hilbert subspace under unitary evolution, and follows the rules of quantum mechanics, Page’s calculation of entanglement entropy immediately follows [26]. In particular, the entanglement entropy will follow a so-called Page curve and fall to zero at late times for any incoming pure state. In an interesting development [27, 28], a Page curve was obtained using semiclassical methods for a large AdS black hole coupled to an external conformal field theory reservoir by adapting a geometric prescription for calculating generalized entropy [29–32]. The fact that a fine-grained quantity like entanglement entropy is correctly reproduced by a semiclassical prescription is highly non-trivial but does not on its own provide much insight into the underlying quantum physics.

IV. INFALLING OBSERVABLES

To build observables for infalling observers we begin with observations made in our previous work [33]. Consider a black hole that has been formed, and then left isolated for a scrambling time, of order $\beta \log S$ where $\beta$ is the inverse Hawking temperature. After this scrambling phase [34], we assume we can apply (8) to the holographic dual. We presume the black hole has been formed from a Mukhanov state that is well-described by the semiclassical approximation (i.e. not, for example, a macroscopic superposition of well-separated black
holes). According to the results of [33] an effective field theory may be set up on a set of
timeslices corresponding to a freely infalling lattice of spatial points moving along timelike
gedesics [35]. Local operators corresponding to infalling “labs” may then be evolved across
the horizon. Moreover the Hamiltonian that evolves these operators along timelike geodesics
may also be constructed in the same manner. We assume that outside the horizon, the lab
states are well-described by a semiclassical approximation, and do not involve strong grav-
itational effects such as the formation of additional black holes inside the lab. Due to the
physical cutoff [33], the time evolution of such operators (and their expectation values) be-
comes independent of the details of what was sent in earlier (i.e. more than $\beta \log S$ earlier).
Thus from the gravity perspective, the details of the black hole microstate do not influence
the evolution of these operators along timelike geodesics. It remains to check then whether
this semiclassical evolution remains accurate when computed using holographic methods.

Outside the horizon of the black hole, in the vicinity of the region where the Mukhanov
states are set up, the lab operators $\phi(x)$ can be mapped precisely into the holographic
description. Likewise the infalling Hamiltonian density acting on this set of operators may be
defined as an operator in the holographic theory. Our philoso phy will be that the expectation
values of this family of operators (i.e. including their fluctuations $|\phi(x) - \langle \phi(x) \rangle|^2$) provides
a complete description of the observables accessible to a local infalling observer. Consider

$$|\text{lab + black hole}\rangle = \{\phi\} |\psi\rangle. \quad (17)$$

This state will typically live in a larger Hilbert subspace than the black hole itself. Here $\{\phi\}$
is some set of operators built of products and sums of the $\phi$’s. The expectation values that
describe the infalling observer will involve matrix elements with states such as

$$\langle \text{lab + black hole}| = \langle \psi| \{\phi\}'.$$

Here $\{\phi\}'$ is some different set, describing the lab at some future time. These matrix elements
can then be computed using the methods described in Section (III) above,

$$\langle \psi| \{\phi\}' \{\phi\} \psi \rangle = \sum_{\alpha \beta} C_\alpha C_\beta^* \langle \beta| \{\phi\}' \{\phi\} |\alpha\rangle = \sum_{\alpha \beta} C_\alpha C_\beta^* A_{\alpha \beta}. \quad (19)$$

These expectation values can be computed to an accuracy that depends on $\Delta E$ for the choice
of black hole states, and the entropy of the black hole. For $\Delta E \sim O(1/E^3)$, we expect they
will be computed to a good approximation by semiclassical methods, as described above.
Now if we have too many independent operators, one can always build linear combinations for which the above statement is no longer true. This bound becomes sharpest when we minimize $\Delta E$ to the point where the corrections are not captured by the semiclassical approximation, i.e. $\Delta E \sim e^{-S/4}$. For an infaller to detect such corrections, they need access to an enormous number of order $N_{lab} \sim e^{S/2}$ operators.

The number of operators an infalling lab has access to can be estimated in a variety of ways. One would be to simply use a holographic bound and assume that if the infalling space of states lives in a Hilbert subspace of dimension $e^{S_{BH}(E_{lab})}$ then the dimension of the space of operators would be of order $e^{2S_{BH}(E_{lab})}$. For $E_{lab} \ll M$ this number is far less than $e^{S(M)/2}$.

An alternative would be to use a spatial lattice cutoff as in [33] and combine that with a lattice cutoff in the time direction. With physically reasonable choices for these cutoffs, it is again easy to obtain an estimate that is far less than $e^{S(M)/2}$ assuming the energy of the lab is much smaller than the energy of the black hole. Again we conclude that the family of matrix elements (19) can be computed to high accuracy by the semiclassical approximation. Once this is established we then have a self-consistent approximation which allows us to push these operators forward along their timelike geodesics into the interior of the black hole, using evolution with respect to the infalling time translation operator and compute their expectation values to the precision stated.

The semiclassical evolution will eventually deviate significantly from the unitary evolution of the corresponding states in the holographic dual theory. The time scale on which this occurs was estimated in a simplified holographic model in [36, 37] and found to match the black hole scrambling time. Interestingly, this is also the maximum time an infalling observers can avoid the black hole singularity in the infalling lattice model of [33, 35]. The holographic model resolves the black hole singularity in the sense that approaching the region of strong curvature coincides with the breakdown of the semiclassical approximation for infalling observables but the holographic evolution remains well defined.

V. CONCLUSIONS

Even when the semiclassical approximation holds in the asymptotic region, it breaks down with corrections of $O(1)$ for transition amplitudes. To make an analogy with particle physics,
we are pointing out that rare decays of the black hole can only be explained by supplementing
the semiclassical picture of spacetime near the horizon of the black hole by new quantum
effects that are of the same order of magnitude as the classical terms. On the other hand,
for almost all local observables of physical interest, self-averaging over of order $e^S$ energy
eigenstates occurs. This renders the semiclassical approximation (with finite size corrections
included) accurate up to corrections of order $e^{-S/2}$. Until we are able to conduct repeated
experiments involving black hole formation and evaporation in the lab, it seems difficult
to discriminate between different theories of quantum gravity that obey these rules. An
exception to this would be repeated observations of the end point of Hawking evaporation
where the effective $S$ is of $O(1)$ and repeated observations of even local observables will
depend in detail on the structure of the underlying theory of quantum gravity.

The experience of infalling observers can likewise be recast as correlators of local observ-
ables initially located outside the horizon. As explained in [33, 36, 37], these observables
can be evolved along timelike geodesics and their time evolution can be evaluated. Given
the constraints on the types of projection operators that can be built from such observables,
the self-averaging over $e^S$ states will still be in play, and we can in principle evaluate such
observables to an accuracy of order $e^{-S/2}$ using a semiclassical approximation until they
enter the high curvature region well-inside the horizon.

These results offer support to the black hole complementarity paradigm for solving the
black hole information problem [38]. Information propagating across the horizon is well-
described by semiclassical evolution toward the curvature singularity. At the same time,
expectation values of local operators outside the horizon receive corrections of order $e^{-S/2}$
which permits unitary evolution to an outgoing cloud of Hawking radiation. We have iden-
tified criteria which lead to accurate semiclassical evolution of states. Physically interesting
states almost always satisfy these criteria. Conversely precise criteria can be given for states
which violate the semiclassical approximation and which cannot be interpreted geometrically.
ACKNOWLEDGMENTS

Work supported in part by DOE grant de-sc0010010 Task A, Icelandic Research Fund grant 228952-051, and a grant from the University of Iceland Research Fund.

[1] S. W. Hawking, “Breakdown of Predictability in Gravitational Collapse,”
*Phys. Rev. D* **14** (1976) 2460–2473

[2] J. M. Maldacena, “The Large N limit of superconformal field theories and supergravity,”
*Adv. Theor. Math. Phys.* **2** (1998) 231–252, arXiv:hep-th/9711200

[3] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, “Gauge theory correlators from noncritical string theory,”
*Phys. Lett. B* **428** (1998) 105–114, arXiv:hep-th/9802109

[4] E. Witten, “Anti-de Sitter space and holography,”
*Adv. Theor. Math. Phys.* **2** (1998) 253–291, arXiv:hep-th/9802150

[5] E. Witten, “Anti-de Sitter space, thermal phase transition, and confinement in gauge theories,”
*Adv. Theor. Math. Phys.* **2** (1998) 505–532, arXiv:hep-th/9803131

[6] D. A. Lowe and L. Thorlacius, “AdS / CFT and the information paradox,”
*Phys. Rev. D* **60** (1999) 104012, arXiv:hep-th/9903237

[7] V. Mukhanov, “On the origin of black-hole entropy,”
*Foundations of Physics* **33** (02, 2003) 271–277

[8] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, “Local bulk operators in AdS/CFT: A Boundary view of horizons and locality,”
*Phys. Rev. D* **73** (2006) 086003, arXiv:hep-th/0506118 [hep-th]

[9] A. Hamilton, D. N. Kabat, G. Lifschytz, and D. A. Lowe, “Holographic representation of local bulk operators,”
*Phys. Rev. D* **74** (2006) 066009, arXiv:hep-th/0606141 [hep-th]

[10] C. Liu and D. A. Lowe, “Conformal wave expansions for flat space amplitudes,”
*JHEP* **07** (2021) 102, arXiv:2105.01026 [hep-th]

[11] C. Liu and D. A. Lowe, “Conformal wavefunctions for graviton amplitudes,”
*JHEP* **01** (2022) 148, arXiv:2109.00037 [hep-th]

[12] S. Pasterski, S.-H. Shao, and A. Strominger, “Flat Space Amplitudes and Conformal Symmetry of the Celestial Sphere,”
*Phys. Rev. D* **96** no. 6, (2017) 065026
[13] D. A. Lowe and D. M. Ramirez, “BMS Symmetry via AdS/CFT,” *JHEP* **01** (2021) 075
arXiv:2007.02851 [hep-th]

[14] R. Gregory and R. Laflamme, “Black strings and p-branes are unstable,”
*Phys. Rev. Lett.* **70** (1993) 2837–2840 arXiv:hep-th/9301052.

[15] W. H. Zurek, “Entropy Evaporated by a Black Hole,” *Phys. Rev. Lett.* **49** (1982) 1683–1686.

[16] M. Srednicki, “Chaos and quantum thermalization,” *Phys. Rev. E* **50** (Aug, 1994) 888–901
https://link.aps.org/doi/10.1103/PhysRevE.50.888

[17] M. V. Berry, “Regular and irregular semiclassical wavefunctions,”
*Journal of Physics A: Mathematical and General* **10** no. 12, (Dec, 1977) 2083–2091
https://doi.org/10.1088/0305-4470/10/12/016

[18] J. Deutsch, “Quantum statistical mechanics in a closed system,” *Phys. Rev. A* **43** no. 4, (February, 1991) 2046.

[19] M. Srednicki, “Thermal fluctuations in quantized chaotic systems,”
*Journal of Physics A: Mathematical and General* **29** no. 4, (Feb, 1996) L75–L79
https://doi.org/10.1088/0305-4470/29/4/003

[20] R. M. Wald, *General Relativity* Chicago Univ. Pr., Chicago, USA, 1984.

[21] J. G. Russo, L. Susskind, and L. Thorlacius, “The Endpoint of Hawking radiation,”
*Phys. Rev. D* **46** (1992) 3444–3449 arXiv:hep-th/9206070

[22] S. Datta, P. Kraus, and B. Michel, “Typicality and thermality in 2d CFT,”
*JHEP* **07** (2019) 143 arXiv:1904.00668 [hep-th]

[23] E. P. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions,” *Annals of Mathematics* **62** no. 3, (1955) 548–564. http://www.jstor.org/stable/1970079

[24] D. Harlow and P. Hayden, “Quantum Computation vs. Firewalls,” *JHEP* **06** (2013) 085
arXiv:1301.4504 [hep-th]

[25] S. W. Hawking, “Particle Creation by Black Holes,”
*Commun. Math. Phys.* **43** (1975) 199–220 [Erratum: Commun.Math.Phys. 46, 206 (1976)].

[26] D. N. Page, “Average entropy of a subsystem,” *Phys. Rev. Lett.* **71** (1993) 1291–1294
arXiv:gr-qc/9305007

[27] A. Almheiri, N. Engelhardt, D. Marolf, and H. Maxfield, “The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole,” *JHEP* **12** (2019) 063
[28] G. Penington, “Entanglement Wedge Reconstruction and the Information Paradox,”
\textit{JHEP} \textbf{09} (2020) 002, \texttt{arXiv:1905.08255 [hep-th]}

[29] S. Ryu and T. Takayanagi, “Holographic derivation of entanglement entropy from AdS/CFT,” \textit{Phys. Rev. Lett.} \textbf{96} (2006) 181602, \texttt{arXiv:hep-th/0603001}

[30] V. E. Hubeny, M. Rangamani, and T. Takayanagi, “A Covariant holographic entanglement entropy proposal,” \textit{JHEP} \textbf{07} (2007) 062, \texttt{arXiv:0705.0016 [hep-th]}

[31] T. Faulkner, A. Lewkowycz, and J. Maldacena, “Quantum corrections to holographic entanglement entropy,” \textit{JHEP} \textbf{11} (2013) 074, \texttt{arXiv:1307.2892 [hep-th]}

[32] N. Engelhardt and A. C. Wall, “Quantum Extremal Surfaces: Holographic Entanglement Entropy beyond the Classical Regime,” \textit{JHEP} \textbf{01} (2015) 073, \texttt{arXiv:1408.3203 [hep-th]}

[33] D. A. Lowe and L. Thorlacius, “Quantum information erasure inside black holes,” \textit{JHEP} \textbf{12} (2015) 096, \texttt{arXiv:1508.06572 [hep-th]}

[34] Y. Sekino and L. Susskind, “Fast Scramblers,” \textit{JHEP} \textbf{10} (2008) 065, \texttt{arXiv:0808.2096 [hep-th]}

[35] S. Corley and T. Jacobson, “Lattice black holes,” \textit{Phys. Rev. D} \textbf{57} (1998) 6269–6279, \texttt{arXiv:hep-th/9709166}

[36] D. A. Lowe and L. Thorlacius, “A holographic model for black hole complementarity,” \textit{JHEP} \textbf{12} (2016) 024, \texttt{arXiv:1605.02061 [hep-th]}

[37] D. A. Lowe and L. Thorlacius, “Black hole holography and mean field evolution,” \textit{JHEP} \textbf{01} (2018) 049, \texttt{arXiv:1710.03302 [hep-th]}

[38] L. Susskind, L. Thorlacius, and J. Uglum, “The Stretched horizon and black hole complementarity,” \textit{Phys. Rev. D} \textbf{48} (1993) 3743–3761, \texttt{arXiv:hep-th/9306069}