Published for SISSA by Springer

Received: June 8, 2016
Accepted: November 23, 2016
Published: December 7, 2016

A holographic model for black hole complementarity

David A. Lowe[a] **and Larus Thorlacius**[b,c]

[a] *Physics Department, Brown University, Providence, RI 02912, U.S.A.*

[b] *University of Iceland, Science Institute, Dunhaga 3, IS-107, Reykjavik, Iceland*

[c] *The Oskar Klein Centre for Cosmoparticle Physics, Department of Physics, Stockholm University, AlbaNova University Centre, 10691 Stockholm, Sweden*

E-mail: lowe@brown.edu, lth@hi.is

Abstract: We explore a version of black hole complementarity, where an approximate semiclassical effective field theory for interior infalling degrees of freedom emerges holographically from an exact evolution of exterior degrees of freedom. The infalling degrees of freedom have a complementary description in terms of outgoing Hawking radiation and must eventually decohere with respect to the exterior Hamiltonian, leading to a breakdown of the semiclassical description for an infaller. Trace distance is used to quantify the difference between the complementary time evolutions, and to define a decoherence time. We propose a dictionary where the evolution with respect to the bulk effective Hamiltonian corresponds to mean field evolution in the holographic theory. In a particular model for the holographic theory, which exhibits fast scrambling, the decoherence time coincides with the scrambling time. The results support the hypothesis that decoherence of the infalling holographic state and disruptive bulk effects near the curvature singularity are complementary descriptions of the same physics, which is an important step toward resolving the black hole information paradox.

Keywords: AdS-CFT Correspondence, Black Holes, Models of Quantum Gravity

ArXiv ePrint: 1605.02061

doi:10.1007/JHEP12(2016)024

JHEP12(2016)024

Contents

1 Introduction

The essence of the black hole information paradox is that the symmetry principle of General Relativity, general covariance, is incompatible with a unitary quantum evolution, where the Hawking radiation [1] carries away the quantum information of the black hole. The internal rigidity of quantum mechanics, and the success of holographic approaches to string theory, such as AdS/CFT [2] and the BFSS matrix model [3], lend weight to the hypothesis that a unitary quantum description should be exact. One is then faced with the problem of how to recover approximate general covariance from such a description.

According to the principle of black hole complementarity, as introduced in [4], physics outside the stretched horizon of a black hole is well described by a local effective field theory but the local description does not extend inside the stretched horizon. As far as outside observers are concerned, the black hole interior is encoded into quantum mechanical degrees of freedom associated with the stretched horizon and residing in a Hilbert space of finite dimension given by the exponential of the Bekenstein-Hawking entropy of the black hole. No information enters the black hole — from the outside point of view infalling matter is absorbed, thermalized and re-emitted by the stretched horizon — and the interior spacetime experienced by a typical observer entering the black hole in free fall is postulated to emerge from the stretched horizon degrees of freedom in a holographic fashion. The interior hologram is constructed from a finite number of degrees of freedom, so the interior bulk theory can at best be approximately local. The question is whether a physical observer inside a black hole, whose measurement resolution is limited both in space and time by the finite size of the black hole, can detect a deviation from local effective field theory in the time allotted before hitting the curvature singularity.

A version of black hole complementarity, which addresses this question, has been proposed by the authors and explored in some recent papers [5, 6]. The construction in [5] applies to an observer falling into a black hole at a prescribed time and is restricted to a limited time period before and after. For the construction to work, the two complementary descriptions of the black hole interior need to satisfy two key requirements. First of all, from the point of view of the outside description, the minimal decoherence time of a black hole must have a lower bound of order the scrambling time $t_{\rm scr} = 4M \log(4M)$. This is the minimum time it takes outside observers to extract quantum information from the black hole after an infalling observer has been absorbed into the stretched horizon. Second, from the viewpoint of the infalling observer, any quantum information that entered the black hole more than a scrambling time earlier must already have been erased from the interior bulk description when the observer enters the black hole.

The second requirement was studied in [6] using a simple model for the bulk physics. It was argued that a sensible holographic description with a finite N (or more precisely the bulk $\hbar$ non-zero), would correspond to a bulk theory with a physical regulator.[1] A minimal requirement is that such a theory should be capable of describing events in a freely-falling frame outside the black hole with a resolution down to a Planck length. Propagating that forward in time, then leads to a lattice black hole model of the type studied in [8] based on a Painlevé-Gullstrand time-slicing. The time coordinate of the infalling frame can be matched to the exterior timelike Killing vector at some finite distance outside the black hole, allowing one to map time evolution in the holographic model to time evolution in the black hole interior. By studying interior propagation of massless fields in this lattice description, it was found that the scrambling time emerges as the maximum coordinate time a signal can propagate before hitting the curvature singularity. This is in sharp contrast to the unregulated description, with exact general covariance, where an infinite coordinate time might pass before collision with the singularity.

In the present paper, we study a model for the holographic side of this story. The predictions from the bulk side are that one should see free propagation in the interior for a time at least of order the scrambling time [6] followed by the rapid onset of large curvature effects with a timescale of order the Planck time. Suppose we send in a small subsystem in a pure state into the black hole. The subsystem can be viewed as a simple model for a freely falling laboratory, where tests of spacetime locality are carried out. Eventually the quantum information initially contained in the subsystem will come out in the Hawking radiation. Since the Hawking radiation propagates with respect to a local Hamiltonian in the exterior, any interactions there will appear as effects that decohere the state from the interior bulk viewpoint. Moreover, such interactions will appear highly non-local in the infalling frame, and lead to apparent violations of quantum mechanics for the infalling observer. Naively, one might predict that measurements of the Hawking radiation might disrupt such a state in an arbitrarily short time. However, if one insists that the detector itself evolve according to the same Hamiltonian as the black hole degrees of freedom, a finite minimal decoherence time emerges.

[1] The need for a bulk regulator to reconcile a finite black hole entropy with the number of states computed semiclassically was noted already in [7].

JHEP12(2016)024

These statements will be quantified in a spin system we introduce below as a simple model of the holographic stretched horizon. The model we choose exhibits fast scrambling which is conjectured to be a property of the stretched horizon degrees of freedom a wide class of black holes [9]. The infalling Hamiltonian evolution is mapped in the holographic model to a mean field Hamiltonian, dependent on the initial state of the system. We compute the trace distance between the states that are obtained by evolving the initial state with respect to the exact Hamiltonian and with respect to the mean field Hamiltonian. This trace distance provides a measure of the decoherence of the infalling state. We find that the decoherence only becomes significant after at least a scrambling time, matching precisely the expectation from the regulated bulk theory. Moreover the timescale for the rapid onset of decoherence also matches the bulk prediction. The results support the version of black hole complementarity advocated by the authors, where singularity approach is complementary to decoherence of the infalling state, initially outlined in [10]. This represents an important step forward towards solving the black hole information problem.

2 Coherence/decoherence

Let us begin by reviewing the basic ideas of decoherence, which involves some system of interest S, interacting with some much larger system S^C which is often denoted the environment. We suppose the Hilbert space factors as

$$\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_{S^c}$$

Consider a pure state in $\mathcal{H}_S$

$$|\Psi\rangle = (|\psi_1\rangle_S + |\psi_2\rangle_S) \otimes |\chi\rangle_{S^C}$$

Under time evolution this becomes

$$|\Psi'\rangle = e^{-iHt}|\Psi\rangle = \sum_i c_{1i}|e_i\rangle \otimes |f_{1i}\rangle + c_{2i}|e_i\rangle \otimes |f_{2i}\rangle \tag{2.1}$$

where the e_i are some basis of $\mathcal{H}_S$. If there is decoherence, then it is a good approximation to assume $|f_{1i}\rangle$ is orthogonal to $|f_{2j}\rangle$ for *any* i and j. For example, this will typically occur if the Hamiltonian is local in position space and $|f_{1i}\rangle$ and $|f_{2i}\rangle$ are position eigenstates. We will adopt the notation

$$\Psi_S = \mathrm{Tr}_{S^c}|\Psi\rangle\langle\Psi|$$

to denote the partial trace over the complement of S. If there is decoherence, then to a good approximation

$$\Psi'_S \approx \sum_i \left(|c_{1i}|^2 + |c_{2i}|^2\right)|e_i\rangle\langle e_i| \tag{2.2}$$

which means the probabilities add, without cross terms. The end result for Ψ'_S is then the same as if a measurement had collapsed the wavefunction into the states $|e_i\rangle$. Note the probability of each $|e_i\rangle$ is not necessarily equal, so Ψ'_S need not be maximally mixed.

JHEP12(2016)024

As it stands, this statement of decoherence is basis dependent. To make a basis independent statement, one can instead quantify the purity of the reduced density matrix Ψ'_S. One way to do this is to compute

$$P = \mathrm{Tr}_S(\Psi'_S)^2$$

which is known as the purity of a density matrix. $P = 1$ for a pure state since a pure state acts as a projector, $(\Psi'_S)^2 = \Psi'_S$, and by normalization $\mathrm{Tr}_s \Psi'_S = 1$. For a mixed state $0 < P < 1$.

Alternatively, one may use the von Neumann entropy

$$\mathcal{S} = -\mathrm{Tr}_s \Psi'_S \log \Psi'_S$$

to quantify the purity of the reduced density matrix. This vanishes for a pure state. For a maximally mixed state $\Psi'_S = \mathbb{1}/n$, on the other hand, $\mathcal{S} = \log n$, with n the dimension of the Hilbert subspace S.

We can then formulate the decoherence time t_d in the following way. Assume at time $t = 0$ Ψ_S is in a pure state. Then define the decoherence time t_d as the time when

$$\mathcal{S}(\Psi_S(t_d)) = \delta \log n \tag{2.3}$$

for some choice of $\delta < 1$. We are not aware of prior appearances of this definition of decoherence time in the literature. This definition should be useful in many other contexts.

In the following we will mostly be interested in studying finite dimensional spin systems. In this class of models, we can reformulate the condition (2.3) as a condition on the trace distance using the results of [11]. We recall the definition

$$\|\Psi_S - \Phi_S\|_1 = \mathrm{Tr}_S \sqrt{(\Psi_S - \Phi_S)^\dagger (\Psi_S - \Phi_S)} \tag{2.4}$$

In [11] it is shown that

$$|\mathcal{S}(\Psi_S) - \mathcal{S}(\Phi_S)| \leq \|\Psi_S - \Phi_S\|_1 \log n \tag{2.5}$$

for two different density matrices in $\mathcal{H}_S$. Therefore the definition of the decoherence time can be reformulated as

$$\|\Psi_S(t_d) - \Phi_S(t_d)\|_1 = \delta \tag{2.6}$$

for some fixed constant $\delta < 1$ and some suitable choice for Φ_S. The state $\Phi_S(t)$ should be chosen to maintain purity under time evolution for the subsystem of interest, but minimize the trace distance as a function of time so the bound (2.5) is as useful as possible. For the models considered here, we will choose Φ_S to evolve according to a local mean field Hamiltonian, as we describe below.

In [12] the statement of fast scrambling was defined in a similar way. The key distinction is that scrambling involves a global mixing of the system, rather than only the mixing of a particular subsystem of interest. The condition for scrambling would then require that, (2.6) should hold for all subsystems, suitably defined, rather than some single small subsystem, as typically considered in the decoherence problem.

JHEP12(2016)024

3 Toy holographic model

While it is interesting to try to derive an effective holographic model for the horizon degrees of freedom of a black hole from some more fundamental description such as AdS/CFT or the Matrix Model, our strategy will be to make some minimal assumptions about such a description, and hope that it carries over to a more precise reconstruction. The key assumption we will make of the model is that it exhibits fast scrambling in the sense of [13], with a scrambling time

$$t \sim \beta \log S_{\rm BH}$$

with β the inverse Hawking temperature of a black hole hole with energy E and $S_{\rm BH}$ the Bekenstein-Hawking entropy of the black hole. Later we will also be interested in carrying out computations in the model for highly entangled states that will model the state of an old black hole entangled with its Hawking radiation. As such, we assume the model contains enough degrees of freedom to model the interior of the black hole and its immediate vicinity. Thus we make the identification that $S \sim N$ the number of sites in the model, and β will be scaled out of the problem. The near-horizon region will not contain all the symmetries of the asymptotic region, so we do not expect conformal symmetry (as in AdS/CFT) or supersymmetry (as in the BFSS model) to be crucial in formulating this effective model. At best the holographic model should contain a version of rotation/translation symmetry, and time translation invariance.

A toy model that exhibits these features is discussed in [12]. This is a spin model with a non-local pairwise interaction. There are N distinct sites with the Hilbert space of tensor product form $\mathcal{H} = \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_N$. The sites interact via a pairwise Hamiltonian $H = \sum_{\langle x,y\rangle} H_{\langle x,y\rangle}$ summing over unordered pairs of sites. The Hamiltonian may therefore be associated with a graph $G = (V, E)$ with N vertices V, and edges E corresponding to the non-zero $H_{\langle x,y\rangle}$. In order to have fast scrambling, the degree of the vertices D should be of order the size of the system. We shall then set $D = N - 1$. To have a sensible limit for large N, we take the pairwise interactions to be bounded $||H_{\langle x,y\rangle}|| < c/D$, for some constant c. Here the operator norm $||\mathcal{O}||$ is defined as the absolute value of the maximum eigenvalue of the operator $\mathcal{O}$.

The Lieb-Robinson result [14] places bounds on the norm of the commutator of operators localized at different sites, as a function of time. For local interactions, this is to be interpreted as a proof of finite group velocity in nonrelativistic spin systems. In the case at hand, where interactions are non-local, the same method still yields a bound on the norm of the commutator for operators. In particular, in [12] it is shown that

$$\|[O_A(t), O_B]\| \leq \frac{4}{D} \|O_A\| \, \|O_B\| \, |A| e^{8ct} \tag{3.1}$$

Here O_X is a bounded norm operator acting in the Hilbert subspace of the sites in the set X, and B is chosen to be a single site.

The condition for scrambling is set up in [12] as follows. Consider some Hilbert subspace $\mathcal{H}_1$ with dimension of order 1, maximally entangled with some reference system $\mathcal{P}$, which experiences no interactions. Here we set the system $S = \mathcal{H}_1$. Under time evolution, the entanglement between $\mathcal{H}_1$ and $\mathcal{P}$ will decay, which may be quantified by the trace distance

$$\|\Psi_{\mathcal{P}S}(t_*) - \Psi_{\mathcal{P}}(0) \otimes \Psi_S(t_*)\|_1 < \epsilon \,\mathrm{rank}\Psi_{\mathcal{P}}(0) \tag{3.2}$$

for some constant $\epsilon \ll 1$. This may in principle then be used as a definition of scrambling time. A bound on the time t_* can then be obtained by noting that it is bounded by the signaling time from the space S to its complement S^c, which may be bounded using (3.1).

First apply this to an initial state where S is a single site, and the complement subspace S^c has dimension of order N. We assume the initial state is of product form $|\Psi(0)\rangle = |\psi_1\rangle_{\mathcal{P}\mathcal{H}_1} \otimes |\psi_2\rangle_{\mathcal{H}_2} \otimes \cdots \otimes |\psi_N\rangle_{\mathcal{H}_N}$. Applying (3.1) with $B = S$ and $A = S^c$ one finds the timescale t_* is of order a constant.

For the black hole problem, the natural initial state to choose is instead one where the black hole degrees of freedom are maximally entangled with the exterior Hawking radiation. Now essentially the roles of S and S^c are reversed. One takes the system S to be of size of order N, with some small subsystem in a factor pure state. The complement is then of size of order 1. To satisfy the bound (3.2) one again requires signaling between S and S^c, and this timescale is bounded by the Lieb-Robinson result. This yields a timescale of order $\log N$ as expected for a fast scrambling system.

4 Mean field and bulk evolution

At first sight, the results of the previous section are not encouraging for the black hole complementarity scenario. While one can build holographic models that exhibit fast scrambling with t_* proportional to $\log N$, it seems the decoherence time for some small Hilbert subspace in such models will be very short. This is, however, not the right question to ask in the black hole problem. Instead, what one should do is build a model for a laboratory that one sends into the black hole, and then ask whether that laboratory will have a decoherence time sufficiently long that they will not be able to distinguish quantum mechanics failing from their classical demise due to singularity approach.

The eventual failure of quantum mechanics in the infalling laboratory can be traced to the existence of two distinct time evolutions for the state in the lab subspace. One of these will be the exact Hamiltonian evolution according to the holographic Hamiltonian H. The other will be defined according to a mean field Hamiltonian H_{MF}, that we describe in more detail shortly, and corresponds to the usual notion of time evolution in the bulk spacetime. It is important to note that not all states yield sensible mean field evolutions. Moreover, as will be clear, the mean field Hamiltonian depends on the state. We conjecture that states close to smooth bulk spacetimes do have useful mean field descriptions, and that the mean field evolution is dual to the usual time evolution with respect to the bulk Hamiltonian. In the remainder of this section, we explore these issues in the context of our holographic toy model.

The mean field approximation to the time evolution of a density matrix is considered in some generality in [15]. We begin by briefly reviewing the standard mean field construction

JHEP12(2016)024

based on an initial pure state of product form

$$|\Psi(0)\rangle = |\psi_1\rangle_{\mathcal{H}_1} \otimes \cdots \otimes |\psi_N\rangle_{\mathcal{H}_N} \tag{4.1}$$

and later on we adapt it to the case of a highly entangled state corresponding to an old black hole.

Starting from (4.1) one builds a state dependent mean field Hamiltonian

$$\begin{aligned} H^{\mathrm{MF}} &= \sum_x H_x^{\mathrm{MF}}(t) \\ H_x^{\mathrm{MF}} &= \sum_y \mathrm{tr}_y \left(H_{\langle x,y \rangle} \Psi_y^{\mathrm{MF}}(t) \right) \end{aligned} \tag{4.2}$$

where Ψ^{MF} evolves according to H^{MF} from the same initial state $|\Psi(0)\rangle$. A key point is that with these definitions, and choice of initial state, the mean field Hamiltonian never generates entanglement between different sites, and remains in the same product form as the initial state. This mean field Hamiltonian then has the expected properties of the holographic dual of the bulk gravity Hamiltonian. As is well known, the bulk Hamiltonian generates smooth time evolution all the way to the curvature singularity, with minimal entanglement being generated. This feature is in fact the origin of the information problem.[2]

One then wishes to calculate the timescale for which the trace norm distance between $\Psi_x(t)$ and $\Psi_x^{\mathrm{MF}}(t)$ remains small. This maps onto a problem solved in [12] for the spin model considered above, via careful application of Lieb-Robinson bounds applied to an expansion of the matrix element

$$\left\langle \Psi(t) | \mathbb{1} - \Psi_x^{\mathrm{MF}}(t) | \Psi(t) \right\rangle = 1 - \left\langle \Psi_x^{\mathrm{MF}}(t) | \Psi_x(t) | \Psi_x^{\mathrm{MF}}(t) \right\rangle \tag{4.3}$$

by making a Dyson series expansion in $H - H_x^{\mathrm{MF}}$. This matrix element in turn places a bound on the trace distance between the states (2.4). Using the result of [11], this then places a bound on the von Neumann entropy $H(\Psi_x(t))$. One finds

$$\left\langle \Psi_x^{\mathrm{MF}}(t) | \Psi_x(t) | \Psi_x^{\mathrm{MF}}(t) \right\rangle \leq \frac{c'}{D} e^{c'' t} \tag{4.4}$$

where $c'$and c'' are constants independent of N. For $D = N - 1$ these quantities become of order 1 when $t \sim \log N$.

Making contact with black hole physics, the holographic description should be useful both inside and outside the black hole horizon. An initial state of the form (4.1) is relevant outside the black hole horizon, or for a recently formed black hole prior to scrambling. To make further progress we need to generalize the mean field results of [12] to highly entangled states.

Suppose we choose a maximally entangled initial state where we have a pairwise entanglement between $\mathcal{H}_{2k}$ and $\mathcal{H}_{2k+1}$ for all $k \geq 1$. Then we can almost map the problem

[2]It should be noted the mean field Hamiltonian depends on the choice of initial state via (4.2). The state dependence of the boundary to bulk map is emphasized in [16]. However in the present construction it is also important that the boundary to bulk mapping is time dependent, which follows from $[H, H_x^{\mathrm{MF}}] \neq 0$.

JHEP12(2016)024

into the one just considered by coarse graining, and viewing $\mathcal{H}_{2k} \otimes \mathcal{H}_{2k+1}$ as a pure state on a single coarse grained site.[3] An important new feature is that the coarse grained Hamiltonian now has a self-interaction term. Such a term must be treated exactly in the mean field approximation. For this initial state, we therefore define

$$
\begin{aligned}
H^{\mathrm{MF}} &= \sum_x H_x^{\mathrm{MF}}(t) \\
H_x^{\mathrm{MF}} &= H_{\langle x,x\rangle} + \sum_{y\backslash x} \mathrm{tr}_y \left(H_{\langle x,y\rangle} \Psi_y^{\mathrm{MF}}(t) \right)
\end{aligned}
$$

where the sums are over coarse grained sites $x = 1, \cdots, N/2$. With this Hamiltonian, we may then proceed as above to compute the trace distance between the mean field state and the exact evolution, or equivalently the von Neumann entropy of the exact reduced density matrix, obtaining the same scaling with N (though different constants c' and c'') via (4.4).

This is now a nice model for an old evaporating black hole after the Page time $t \sim \beta S_{\mathrm{BH}}$, where the interior degrees of freedom are maximally entangled with the exterior Hawking radiation. The decoherence time, defined according to the definition (2.3) is now of order $t \sim \log N$ matching the scrambling time. We also note once a time of order the scrambling time has passed, the bound (4.4) increases with a rise time of order 1, matching the bulk expectation of strong curvature effects with an onset of order the Planck time.

The same kind of computations can be carried out for a variety of initial states. For example, to mimic an observer falling in and carrying out quantum experiments, we can choose to separate the infalling site x into two sites x_1 and x_2 with $||\mathcal{H}_1|| \gg ||\mathcal{H}_2||$. Let us assume some strong coupling between these sites, with coupling to the other sites bounded as above. Defining the decoherence time using (2.6) for the choice $S = x_2$ will lead to a decoherence time of order 1, independent of N. This means the decoherence time for measurements internal to the infalling lab can be made rapidly, as expected. However it is still true that the combined state on $\mathcal{H}_1 \otimes \mathcal{H}_2$ will remain pure for a time of order the scrambling time using the above construction. Thus measurements of the Hawking radiation will not lead to rapid decoherence of the state, as naively expected.

To test the idea that the minimal decoherence time due to measurement of Hawking radiation really matches the scrambling time, one can try to generalize the above discussion to any density matrix with a pure state subfactor representing the infalling system. The above derivation will generalize provided the mean field approximation holds for the evolution of the subsystem of interest. It seems natural that states representing smooth spacetime geometries will correspond to good mean field states, however the converse need not be true. It would be very interesting to see more directly how this class of states emerges as a class of attractor states from large N holographic theories.

[3] We note that the Schmidt decomposition [17] implies this special state is unitary equivalent to the generic maximally entangled state. Converting from Schrodinger picture to Heisenberg picture, this may be viewed as conjugation of the Hamiltonian by a constant unitary matrix. For the pairwise interaction considered in this model, a general unitary transformation will induce self-interactions, but preserve the pairwise form of the Hamiltonian. The condition $||H_{(x,y)}|| < c/D$ is preserved by this unitary conjugation, so the above proof goes through.

5 Comments

In the above we have argued the minimal decoherence time of an infalling state is to be identified with the scrambling time, subject to assumptions about the form of the holographic model. In this context, we have proven one of the key assumptions about the approach to black hole complementarity described in [5, 6].

The other key assumption relied on details of the holographic reconstruction of bulk spacetimes, namely that general covariance is softly broken through the introduction of a Planck length regulator. This was explored in a regulated model of the bulk in [6], where it was found that it was sufficient for infalling degrees of freedom to retain coherence for a scrambling time. That work describes the details of the construction, including the important conditions placed on bulk timeslices compatible with the regulator.

Since we are concluding the timescales match, this is an important success for building interior degrees of freedom in a holographic theory. Moreover, if the bound is saturated, one also predicts the timescale of order 1 associated with the rapid rise expected from strong curvature near the singularity. Away from this region the trace distance will be of order $1/N$.

Since these corrections are to be essentially interpreted as violations of quantum mechanics for the infalling observer, it is of great interest to quantify to what extent these are tolerable. The trace distance may be interpreted directly as the probability of an ideal experiment detecting the difference between two states [17]. As an initial crude estimate, if we take $N \sim S_{\text{universe}} \approx 10^{88}$ (assuming domination by CMB photons) and assume the nonlocal effect produces a Planck energy particle with a probability $1/N$ per unit Planck time per degree of freedom, we can apply it to the atoms in the Earth's atmosphere to conclude one Planck energy ultra high energy cosmic ray would appear about every 10^7 years. This is conceivably a detectable effect, but apparently rather harmless.

It is natural to conjecture some version of the same matching of scrambling time with interior decoherence time holds in all holographic theories of quantum gravity. It remains an important open problem to directly derive holographic effective theories of the horizon degrees of freedom from more fundamental descriptions such as AdS/CFT or Matrix Models, and test this conjecture. It will also be very interesting to further explore mean field approximations in such holographic descriptions. Of course one expects the mean field (or master field formulation of a large N theory) will coincide with the bulk gravity description at leading order. However having a formulation directly in the Hilbert space of the underlying holographic description is needed to carry out computations analogous to those of section 4.

We note that generic holographic states in more realistic models may well contain singularities leading to a breakdown of the mean field approach. From the bulk perspective, we would expect for sufficiently large N, a version of cosmic censorship [18] to hold. The additional singularities will then be censored by their own horizons leaving a smooth geometry outside where we expect mean field to remain accurate. It will be interesting to see what extent this may be derived in the holographic model considered here.

JHEP12(2016)024

In the near term, it would be interesting to generalize the present work beyond spin models to systems with an infinite dimensional Hilbert space at each site. Some recent applications of Lieb-Robinson bounds to such systems appear in [19–22]. It appears promising that these results can be generalized to models that exhibit fast scrambling.

A Comparison with firewall story

It is useful to compare the black hole complementarity picture with the popular firewall story.

A.1 Disagreement in overlap region

The formulation of [5, 6] implies that the infalling state begins to disagree with the exact state outside the black hole a scrambling time prior to horizon crossing. One might attempt to reach a contradiction with this approach by taking an outgoing Hawking particle emitted in this overlap region, performing a quantum computation on this, and the earlier Hawking radiation, and sending the result to the infalling observer. If this could be done with appreciable probability, it would result in a violation of the ordinary rules of quantum mechanics for the infaller, who would notice that upon horizon crossing, the required entangled Hawking partner was not there.

If the unitary transformation is done using the holographic model we have described above, the quantum computation maps to unitary evolution with respect to some dense pairwise interactions in the holographic theory. The above estimates apply to this state as well, and one again concludes that the time necessary for such a computation is of order the scrambling time. Thus by construction, the infaller has already crossed the horizon.

A.2 Precomputation

One might try an extreme version of the above by precomputing the quantum state, and arranging for the infalling state to be entangled with the outgoing Hawking particle in the overlap region. In this case the black hole complementarity picture fails, and is unpredictive for the experience of the infalling observer.

However we should then try to quantify how surprising it is for the approach to fail for special states. To arrange for one Hawking particle to be precisely entangled with a given infaller requires picking a vector lying in an $e^{S_{\rm BH}}$ dimensional Hilbert space, modulo unitary transformations that just act on the other Hawking particles. Standard estimates, matching the trace distance of such states to the required state then give the time for such a transformation to be constructed of order

$$e^{ce^{S_{\rm BH}(M)}}/e^{ce^{S_{\rm BH}(M-1/\beta)}} \approx e^{c'e^{S_{\rm BH}(M)}}$$

where c and c' are positive constants of order 1. This is parametrically the same as upper bounds on the quantum Poincare recurrence time [23, 24].

We conclude that with enough available time, one can reliably create infalling states that do not have sensible bulk evolution in the interior. However unless one completely

JHEP12(2016)024

isolates the state, for an extremely long time, any quantum noise will defeat the coherence required to see these unusual effects. Turning the argument around, a typical infalling observer will see drama with a probability of order $e^{-c' e^{S_{\rm BH}}}$.

A.3 Microcanonical ensemble

In [25] it was suggested the microcanonical ensemble would lead to an infalling number operator that always was of order 1 for any mode, due to entanglement with the exterior operators. Loopholes in this argument were already pointed out in [26, 27] and the present work and its companion paper [6] provide a concrete realization of these ideas. The new ingredients are the time-dependent (with respect to CFT time) bulk regulator [6] and the corresponding time and state dependent boundary to bulk map via the mean field Hamiltonian.

In particular, if one were to replace the mean field evolution by evolution with respect to a state independent operator a decoherence time of order 1 will emerge for typical states according to (2.6). For example if we tried to represent the bulk Hamiltonian by the identity operator, this would correspond to the choice $\Phi_S(t) = \Psi_S(0)$ in (2.6). For the kinds of highly entangled states considered in section 4 this will lead to a timescale of order 1. The mean field approach is essential for correctly matching the holographic dual of the bulk Hamiltonian.

Acknowledgments

D.L. wishes to thank Daniel Harlow for a helpful discussion. The research of D.L. was supported in part by DOE grant de-sc0010010. The research of L.T. was supported in part by Icelandic Research Fund grant 163422-051, the University of Iceland Research Fund, and the Swedish Research Council under contract 621-2014-5838.

References

[1] S.W. Hawking, *Particle Creation by Black Holes*, *Commun. Math. Phys.* **43** (1975) 199 [*Erratum ibid.* **46** (1976) 206] [INSPIRE].

[2] J.M. Maldacena, *The large-N limit of superconformal field theories and supergravity*, *Int. J. Theor. Phys.* **38** (1999) 1113 [hep-th/9711200] [INSPIRE].

[3] T. Banks, W. Fischler, S.H. Shenker and L. Susskind, *M theory as a matrix model: A conjecture*, *Phys. Rev.* **D 55** (1997) 5112 [hep-th/9610043] [INSPIRE].

[4] L. Susskind, L. Thorlacius and J. Uglum, *The stretched horizon and black hole complementarity*, *Phys. Rev.* **D 48** (1993) 3743 [hep-th/9306069] [INSPIRE].

[5] D.A. Lowe and L. Thorlacius, *Black hole complementarity: The inside view*, *Phys. Lett.* **B 737** (2014) 320 [arXiv:1402.4545] [INSPIRE].

[6] D.A. Lowe and L. Thorlacius, *Quantum information erasure inside black holes*, *JHEP* **12** (2015) 096 [arXiv:1508.06572] [INSPIRE].

[7] G. 't Hooft, *On the Quantum Structure of a Black Hole*, *Nucl. Phys.* **B 256** (1985) 727 [INSPIRE].

[8] S. Corley and T. Jacobson, *Lattice black holes*, *Phys. Rev.* **D 57** (1998) 6269 [hep-th/9709166] [INSPIRE].

[9] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [arXiv:0808.2096] [INSPIRE].

[10] D.A. Lowe and L. Thorlacius, *Comments on the black hole information problem*, *Phys. Rev.* **D 73** (2006) 104027 [hep-th/0601059] [INSPIRE].

[11] M. Fannes, *A continuity property of the entropy density for spin lattice systems*, *Commun. Math. Phys.* **31** (1973) 291.

[12] N. Lashkari, D. Stanford, M. Hastings, T. Osborne and P. Hayden, *Towards the Fast Scrambling Conjecture*, *JHEP* **04** (2013) 022 [arXiv:1111.6580] [INSPIRE].

[13] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *JHEP* **09** (2007) 120 [arXiv:0708.4025] [INSPIRE].

[14] E.H. Lieb and D.W. Robinson, *The finite group velocity of quantum spin systems*, *Commun. Math. Phys.* **28** (1972) 251 [INSPIRE].

[15] A.Z. Adrian Muntean, J.D.M. Rademacher and A. Zagaris eds., *Macroscopic and Large Scale Phenomena: Coarse Graining, Mean Field Limits and Ergodicity*, in *Lecture Notes in Applied Mathematics and Mechanics*, vol. 3, Springer (2016).

[16] K. Papadodimas and S. Raju, *State-Dependent Bulk-Boundary Maps and Black Hole Complementarity*, *Phys. Rev.* **D 89** (2014) 086010 [arXiv:1310.6335] [INSPIRE].

[17] M.A. Nielsen and I.L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th edition, Cambridge University Press, New York, NY, U.S.A. (2011).

[18] R. Penrose, *Gravitational collapse: The role of general relativity*, *Riv. Nuovo Cim.* **1** (1969) 252 [INSPIRE].

[19] B. Nachtergaele, H. Raz, B. Schlein and R. Sims, *Lieb-robinson bounds for harmonic and anharmonic lattice systems*, *Commun. Math. Phys.* **286** (2009) 1073 [arXiv:0712.3820].

[20] B. Nachtergaele, B. Schlein, R. Sims, S. Starr and V. Zagrebnov, *On the Existence of the Dynamics for Anharmonic Quantum Oscillator Systems*, *Rev. Math. Phys.* **22** (2010) 207 [arXiv:0909.2249] [INSPIRE].

[21] B. Nachtergaele and R. Sims, *On the dynamics of lattice systems with unbounded on-site terms in the Hamiltonian*, arXiv:1410.8174.

[22] H. Grundling and G. Rudolph, *Dynamics for QCD on an infinite lattice*, arXiv:1512.06319 [INSPIRE].

[23] D.N. Page, *Information loss in black holes and/or conscious beings?*, hep-th/9411193 [INSPIRE].

[24] L. Campos Venuti, *The recurrence time in quantum mechanics*, arXiv:1509.04352.

JHEP12(2016)024

[25] D. Marolf and J. Polchinski, *Gauge/Gravity Duality and the Black Hole Interior*, *Phys. Rev. Lett.* **111** (2013) 171301 [arXiv:1307.4706] [INSPIRE].

[26] S.G. Avery and D.A. Lowe, *Event horizons and holography*, arXiv:1310.7999 [INSPIRE].

[27] S.G. Avery and D.A. Lowe, *Typical Event Horizons in AdS/CFT*, *JHEP* **01** (2016) 082 [arXiv:1501.05573] [INSPIRE].