Recent progress on the black hole information problem

Daniel Harlow

MIT

October 28, 2022

So far I have written three papers with Ooguri-sama:

Constraints on symmetry from holography

Daniel Harlow[1] and Hirosi Ooguri[2,3]
[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*
[3] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*

In this letter we show that a set of old conjectures about symmetries in quantum gravity hold within the Anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence. These conjectures are that no global symmetries are possible, that internal gauge symmetries must come with dynamical objects that transform in all irreducible representations, and that internal gauge groups must be compact. These conjectures are not obviously true from a bulk perspective, they are nontrivial consequences of the non-perturbative consistency of the correspondence. More details of and background for these arguments are presented in an accompanying paper.

Symmetries in Quantum Field Theory and Quantum Gravity

Daniel Harlow[a] and Hirosi Ooguri[b,c]
[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*
[c] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*
E-mail: harlow@mit.edu, ooguri@caltech.edu

ABSTRACT: In this paper we use the AdS/CFT correspondence to refine and then establish a set of old conjectures about symmetries in quantum gravity. We first show that any global symmetry, discrete or continuous, in a bulk quantum gravity theory with a CFT dual would lead to an inconsistency in that CFT, and thus that there are no bulk global symmetries in AdS/CFT. We then argue that any "long-range" bulk gauge symmetry leads to a global symmetry in the boundary CFT, whose consistency requires the existence of bulk dynamical objects which transform in all finite-dimensional irreducible representations of the bulk gauge group. We mostly assume that all internal symmetry groups are compact, but we also give a general condition on CFTs, which we expect to be true quite broadly, which implies this. We extend all of these results to the case of higher-form symmetries. Finally we extend a recently proposed new motivation for the weak gravity conjecture to more general gauge groups, reproducing the "convex hull condition" of Cheung and Remmen.

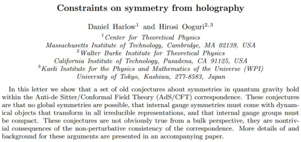
An essential point, which we dwell on at length, is precisely defining what we mean by gauge and global symmetries in the bulk and boundary. Quantum field theory results we meet while assembling the necessary tools include continuous global symmetries without Noether currents, new perspectives on spontaneous symmetry-breaking and 't Hooft anomalies, a new order parameter for confinement which works in the presence of fundamental quarks, a Hamiltonian lattice formulation of gauge theories with arbitrary discrete gauge groups, an extension of the Coleman-Mandula theorem to discrete symmetries, and an improved explanation of the decay $\pi^0 \to \gamma\gamma$ in the standard model of particle physics. We also describe new black hole solutions of the Einstein equation in $d+1$ dimensions with horizon topology $\mathbb{T}^p \times \mathbb{S}^{d-p-1}$.

A universal formula for the density of states in theories with finite-group symmetry

Daniel Harlow[a] and Hirosi Ooguri[b,c]
[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[b] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*
[c] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*
E-mail: harlow@mit.edu, ooguri@caltech.edu

ABSTRACT: In this paper we use Euclidean gravity to derive a simple formula for the density of black hole microstates which transform in each irreducible representation of any finite gauge group. Since each representation appears with nonzero density, this gives a new proof of the completeness hypothesis for finite gauge fields. Inspired by the generality of the argument we further propose that the formula applies at high energy in any quantum field theory with a finite-group global symmetry, and give some evidence for this conjecture.

So far I have written three papers with Ooguri-sama:



Constraints on symmetry from holography

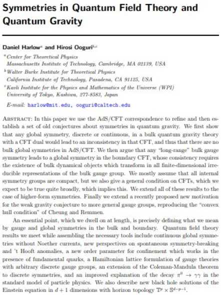
Daniel Harlow[1] and Hirosi Ooguri[2,3]

[1] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[2] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*

[3] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*

In this letter we show that a set of old conjectures about symmetries in quantum gravity hold within the Anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence. These conjectures are that no global symmetries are possible, that internal gauge symmetries must come with dynamical objects that transform in all irreducible representations, and that internal gauge groups must be compact. These conjectures are not obviously true from a bulk perspective, they are nontrivial consequences of the non-perturbative consistency of the correspondence. More details of and background for these arguments are presented in an accompanying paper.



Symmetries in Quantum Field Theory and Quantum Gravity

Daniel Harlow[a] and Hirosi Ooguri[b,c]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

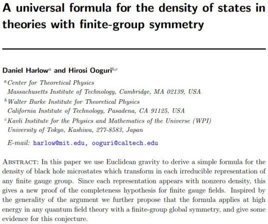
[b] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*

[c] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*

E-mail: harlow@mit.edu, ooguri@caltech.edu

ABSTRACT: In this paper we use the AdS/CFT correspondence to refine and then establish a set of old conjectures about symmetries in quantum gravity. We first show that any global symmetry, discrete or continuous, in a bulk quantum gravity theory with a CFT dual would lead to an inconsistency in that CFT, and thus that there are no bulk global symmetries in AdS/CFT. We then argue that any "long-range" bulk gauge symmetry leads to a global symmetry in the boundary CFT, whose consistency requires the existence of bulk dynamical objects which transform in all finite-dimensional irreducible representations of the bulk gauge group. We mostly assume that all internal symmetry groups are compact, but we also give a general condition on CFTs, which we expect to be true quite broadly, which implies this. We extend all of these results to the case of higher-form symmetries. Finally we extend a recently proposed new motivation for the weak gravity conjecture to more general gauge groups, reproducing the "convex hull condition" of Cheung and Remmen.

An essential point, which we dwell on at length, is precisely defining what we mean by gauge and global symmetries in the bulk and boundary. Quantum field theory results we meet while assembling the necessary tools include continuous global symmetries without Noether currents, new perspectives on spontaneous symmetry-breaking and 't Hooft anomalies, a new order parameter for confinement which works in the presence of fundamental quarks, a Hamiltonian lattice formulation of gauge theories with arbitrary discrete gauge groups, an extension of the Coleman-Mandula theorem to discrete symmetries, and an improved explanation of the decay $\pi^0 \to \gamma\gamma$ in the standard model of particle physics. We also describe new black hole solutions of the Einstein equation in $d+1$ dimensions with horizon topology $\mathbb{T}^p \times \mathbb{S}^{d-p-1}$.



A universal formula for the density of states in theories with finite-group symmetry

Daniel Harlow[a] and Hirosi Ooguri[b,c]

[a] *Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

[b] *Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, CA 91125, USA*

[c] *Kavli Institute for the Physics and Mathematics of the Universe (WPI), University of Tokyo, Kashiwa, 277-8583, Japan*

E-mail: harlow@mit.edu, ooguri@caltech.edu

ABSTRACT: In this paper we use Euclidean gravity to derive a simple formula for the density of black hole microstates which transform in each irreducible representation of any finite gauge group. Since each representation appears with nonzero density, this gives a new proof of the completeness hypothesis for finite gauge fields. Inspired by the generality of the argument we further propose that the formula applies at high energy in any quantum field theory with a finite-group global symmetry, and give some evidence for this conjecture.

The first two are sort of well-known, but I think the third is also nice: we argue that in any quantum field theory with a finite global symmetry G the high-energy density of states in each representation α obeys $\rho_\alpha(E) = \frac{d_\alpha^2}{|G|}\rho(E)$. Maybe testable in condensed matter systems?

Of course the best part of working with Hirosi is having fun with Hirosi!

Especially the food

Incidentally his impeccable dress extends also to active situations

I thought we could also take a little time for a Japanese lesson.

I thought we could also take a little time for a Japanese lesson.
(I don't speak Japanese)

I thought we could also take a little time for a Japanese lesson.
(I don't speak Japanese)
One interesting aspect of Hirosi is how to spell his name. Many of you presumably know that he is really Hiroshi, but his surname is also interesting.

I thought we could also take a little time for a Japanese lesson.
(I don't speak Japanese)
One interesting aspect of Hirosi is how to spell his name. Many of you presumably know that he is really Hiroshi, but his surname is also interesting.
There is a particularly important distinction you should be aware of:

Ooguri （大栗）

Oguri (小栗)

I thought we could also take a little time for a Japanese lesson.
(I don't speak Japanese)
One interesting aspect of Hirosi is how to spell his name. Many of you presumably know that he is really Hiroshi, but his surname is also interesting.
There is a particularly important distinction you should be aware of:

Ooguri （大栗）

Oguri (小栗)

Please do not diminish our big chestnut by turning him into a little chestnut!

Happy birthday Ooguri-sama!

Introduction

- Hawking's black hole information problem has been one of the driving challenges of theoeretical physics for the last 50 years.

Introduction

- Hawking's black hole information problem has been one of the driving challenges of theoeretical physics for the last 50 years.
- It suggests a tension between gravity and quantum mechanics, neither of which we are willing to give up lightly.

Introduction

- Hawking's black hole information problem has been one of the driving challenges of theoeretical physics for the last 50 years.
- It suggests a tension between gravity and quantum mechanics, neither of which we are willing to give up lightly.
- In the last 10-15 years we have improved substantially our understanding of how spacetime emerges in holography, and in the last few years this has grown into a new understanding of Hawking's paradox.

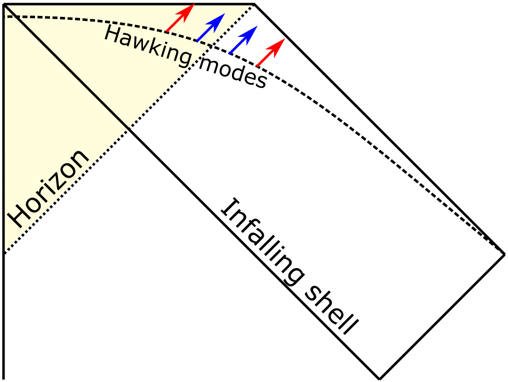
Introduction

- Hawking's black hole information problem has been one of the driving challenges of theoeretical physics for the last 50 years.
- It suggests a tension between gravity and quantum mechanics, neither of which we are willing to give up lightly.
- In the last 10-15 years we have improved substantially our understanding of how spacetime emerges in holography, and in the last few years this has grown into a new understanding of Hawking's paradox.
- It is too soon to say that the problem has been fully resolved, but I think it is fair to say that many of us feel a resolution is in sight.

Introduction

- Hawking's black hole information problem has been one of the driving challenges of theoeretical physics for the last 50 years.
- It suggests a tension between gravity and quantum mechanics, neither of which we are willing to give up lightly.
- In the last 10-15 years we have improved substantially our understanding of how spacetime emerges in holography, and in the last few years this has grown into a new understanding of Hawking's paradox.
- It is too soon to say that the problem has been fully resolved, but I think it is fair to say that many of us feel a resolution is in sight.
- In this talk I will attempt to give a brief overview of what I think is the current status.

You are hopefully familiar with the basic argument:



Entanglement between interior and exterior modes causes the black hole to radiate, losing energy, but this radiation cannot carry information about the infalling shell since this is still deep inside. By the time the black hole reaches Planckian size, it doesn't have enough energy left to return this information to the exterior.

We can describe Hawking's paradox as saying it is impossible to have the following things in one theory:

We can describe Hawking's paradox as saying it is impossible to have the following things in one theory:

(1) A finite black hole entropy

We can describe Hawking's paradox as saying it is impossible to have the following things in one theory:

(1) A finite black hole entropy

(2) A unitary black hole S-matrix

We can describe Hawking's paradox as saying it is impossible to have the following things in one theory:

(1) A finite black hole entropy
(2) A unitary black hole S-matrix
(3) Effective field theory valid away from high energy densities and/or curvatures.

We can describe Hawking's paradox as saying it is impossible to have the following things in one theory:

(1) A finite black hole entropy
(2) A unitary black hole S-matrix
(3) Effective field theory valid away from high energy densities and/or curvatures.

These are all things we really would like to be true, so any resolution of the paradox will teach us something deep!

- In some of the traditional proposals for resolving Hawking's problem, i.e. remnants and/or information loss, one gives up (1) and/or (2) to preserve (3).

- In some of the traditional proposals for resolving Hawking's problem, i.e. remnants and/or information loss, one gives up (1) and/or (2) to preserve (3).
- In particular one gives up the idea that the black hole entropy is actually given by $S = A/4G$.

- In some of the traditional proposals for resolving Hawking's problem, i.e. remnants and/or information loss, one gives up (1) and/or (2) to preserve (3).
- In particular one gives up the idea that the black hole entropy is actually given by $S = A/4G$.
- The string theory counting of black hole microstates, both directly and through the AdS/CFT correspondence, gives strong evidence that indeed we have $S = A/4G$.

- In some of the traditional proposals for resolving Hawking's problem, i.e. remnants and/or information loss, one gives up (1) and/or (2) to preserve (3).
- In particular one gives up the idea that the black hole entropy is actually given by $S = A/4G$.
- The string theory counting of black hole microstates, both directly and through the AdS/CFT correspondence, gives strong evidence that indeed we have $S = A/4G$.
- Moreover AdS/CFT give strong evidence that the S-matrix is unitary.

- In some of the traditional proposals for resolving Hawking's problem, i.e. remnants and/or information loss, one gives up (1) and/or (2) to preserve (3).
- In particular one gives up the idea that the black hole entropy is actually given by $S = A/4G$.
- The string theory counting of black hole microstates, both directly and through the AdS/CFT correspondence, gives strong evidence that indeed we have $S = A/4G$.
- Moreover AdS/CFT give strong evidence that the S-matrix is unitary.

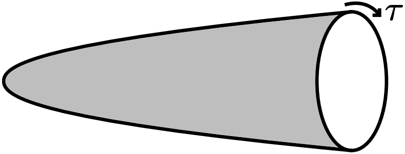
Our challenge is thus to understand what replaces (3).

An aside on Euclidean gravity

There is a somewhat philosophical point which needs to be emphasized here.

An aside on Euclidean gravity

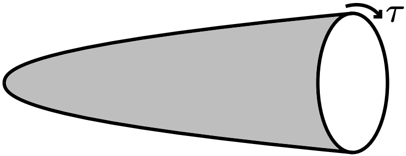
There is a somewhat philosophical point which needs to be emphasized here.



Already in 1976 Gibbons and Hawking used the Euclidean gravity path integral to "derive" the entropy formula:

An aside on Euclidean gravity

There is a somewhat philosophical point which needs to be emphasized here.

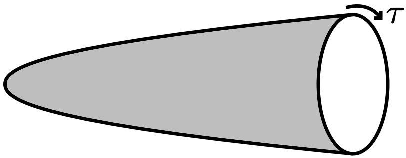


Already in 1976 Gibbons and Hawking used the Euclidean gravity path integral to "derive" the entropy formula:

$$S = (1 - \beta\partial_\beta) \log Z = \frac{A}{4G}.$$

An aside on Euclidean gravity

There is a somewhat philosophical point which needs to be emphasized here.



Already in 1976 Gibbons and Hawking used the Euclidean gravity path integral to "derive" the entropy formula:

$$S = (1 - \beta\partial_\beta)\log Z = \frac{A}{4G}.$$

Why then did I cite string theory and AdS/CFT to justify this formula?

- The reason is that in Euclidean gravity this calculation is *not* manifestly equal to a trace: the thermal circle contracts at the tip of the cigar, so we cannot interpret this "partition function" as a trace in the Hilbert space of canonical gravity.

- The reason is that in Euclidean gravity this calculation is *not* manifestly equal to a trace: the thermal circle contracts at the tip of the cigar, so we cannot interpret this “partition function” as a trace in the Hilbert space of canonical gravity.
- Euclidean gravity allows the semiclassical theory to “know” the number of black hole microstates without knowing what those microstates are.

- The reason is that in Euclidean gravity this calculation is *not* manifestly equal to a trace: the thermal circle contracts at the tip of the cigar, so we cannot interpret this "partition function" as a trace in the Hilbert space of canonical gravity.
- Euclidean gravity allows the semiclassical theory to "know" the number of black hole microstates without knowing what those microstates are.
- It is interesting to think about why (and when) this works, but it isn't my focus today.

- The reason is that in Euclidean gravity this calculation is *not* manifestly equal to a trace: the thermal circle contracts at the tip of the cigar, so we cannot interpret this “partition function” as a trace in the Hilbert space of canonical gravity.
- Euclidean gravity allows the semiclassical theory to “know” the number of black hole microstates without knowing what those microstates are.
- It is interesting to think about why (and when) this works, but it isn’t my focus today.
- The reason why so many people were excited about the work of Strominger and Vafa is that it enabled us to compute black hole entropy *using the definition of entropy*, without relying on the mysterious black box of quantum gravity.

Spacetime emergence

- In understanding what might replace (3) (the validity of EFT away from singularities), it is useful to note that in AdS/CFT it not obvious that *any* version of (3) holds: the fundamental description is the boundary CFT, and the bulk spacetime is at best *emergent*.

Spacetime emergence

- In understanding what might replace (3) (the validity of EFT away from singularities), it is useful to note that in AdS/CFT it not obvious that *any* version of (3) holds: the fundamental description is the boundary CFT, and the bulk spacetime is at best *emergent*.
- On the other hand using SUSY, large N methods, integrability, etc, it is clear that the low-energy states of holographic CFTs indeed match onto the Hilbert space of EFT of AdS.

Spacetime emergence

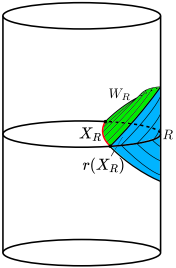
- In understanding what might replace (3) (the validity of EFT away from singularities), it is useful to note that in AdS/CFT it not obvious that *any* version of (3) holds: the fundamental description is the boundary CFT, and the bulk spacetime is at best *emergent*.
- On the other hand using SUSY, large N methods, integrability, etc, it is clear that the low-energy states of holographic CFTs indeed match onto the Hilbert space of EFT of AdS.
- Moreover the high-energy states qualitatively behave as one would expect for states with a big black hole in the bulk, and in some cases we can even confirm the entropy formula (e.g. using the Cardy formula).

Spacetime emergence

- In understanding what might replace (3) (the validity of EFT away from singularities), it is useful to note that in AdS/CFT it not obvious that *any* version of (3) holds: the fundamental description is the boundary CFT, and the bulk spacetime is at best *emergent*.
- On the other hand using SUSY, large N methods, integrability, etc, it is clear that the low-energy states of holographic CFTs indeed match onto the Hilbert space of EFT of AdS.
- Moreover the high-energy states qualitatively behave as one would expect for states with a big black hole in the bulk, and in some cases we can even confirm the entropy formula (e.g. using the Cardy formula).
- What is less clear is how to think about the emergence of spacetime *inside* black holes, and this has been the main topic of recent work.

Quantum extremal surfaces

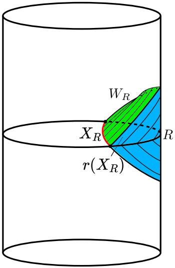
A key insight into the nature of emergent spacetime in AdS/CFT is the *quantum extremal surface formula*: Ryu, Takayanagi, Hubeny, Rangamani, Faulkner, Maldacena, Lewkowycz, Engelhardt, Wall



$$S(\rho_R) = \min_{X_R}\left[\operatorname*{ext}_{X_R}\left(\frac{\mathrm{Area}(X_R)+\dots}{4G} + S_{bulk}(r(X_R))\right)\right].$$

Quantum extremal surfaces

A key insight into the nature of emergent spacetime in AdS/CFT is the *quantum extremal surface formula*: Ryu, Takayanagi, Hubeny, Rangamani, Faulkner, Maldacena, Lewkowycz, Engelhardt, Wall

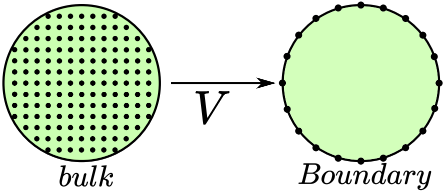


$$S(\rho_R) = \min_{X_R}\left[\operatorname*{ext}_{X_R}\left(\frac{\mathrm{Area}(X_R)+\dots}{4G}+S_{bulk}(r(X_R))\right)\right].$$

Here X_R is required to be *homologous* to R, meaning there exists a bulk region $r(X_R)$ such that $\partial r(X_R) = R - X_R$. The domain of dependence of $r(X_R)$ is called the *entanglement wedge* of R, denoted W_R.

- As with the black hole entropy formula, of which it is a generalization, there is a Euclidean gravity "derivation" of the QES formula. Lewkowycz/Maldacena
- One can also look for a more direct Hilbert space explanation a la Strominger/Vafa, and currently our best direct understanding is based on the idea that the QES formula is a generic feature of quantum error-correcting codes. Almheiri/Harlow/Dong, Pastawski/Yoshida/Harlow/Preskill, Harlow, Hayden/Nezami/Qi/Thomas/Walter

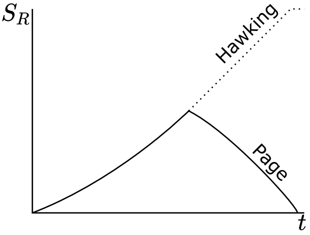
- As with the black hole entropy formula, of which it is a generalization, there is a Euclidean gravity "derivation" of the QES formula. Lewkowycz/Maldacena
- One can also look for a more direct Hilbert space explanation a la Strominger/Vafa, and currently our best direct understanding is based on the idea that the QES formula is a generic feature of quantum error-correcting codes. Almheiri/Harlow/Dong, Pastawski/Yoshida/Harlow/Preskill, Harlow, Hayden/Nezami/Qi/Thomas/Walter
- For the latter the basic idea is that there is an approximate isometry V mapping the bulk effective field theory Hilbert space into the CFT Hilbert space, whose error-correcting properties require the QES formula to hold.



In 2019 these ideas led to the remarkable discovery that the QES formula is smart enough to compute the "Page curve" of an evaporating black hole. Penington, Almheiri/Engelhardt/Marolf/Maxfield

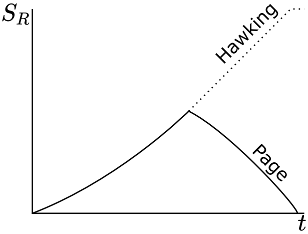
In 2019 these ideas led to the remarkable discovery that the QES formula is smart enough to compute the "Page curve" of an evaporating black hole. Penington, Almheiri/Engelhardt/Marolf/Maxfield

The Page curve is the plot of the von Neumann entropy of the Hawking radiation as a function of time:



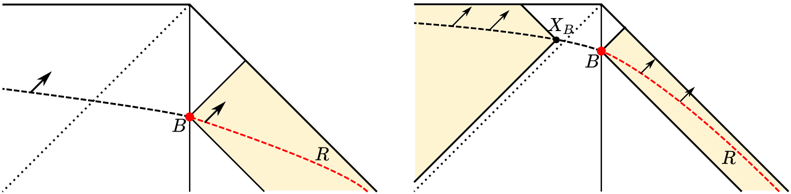
In 2019 these ideas led to the remarkable discovery that the QES formula is smart enough to compute the "Page curve" of an evaporating black hole. Penington, Almheiri/Engelhardt/Marolf/Maxfield

The Page curve is the plot of the von Neumann entropy of the Hawking radiation as a function of time:

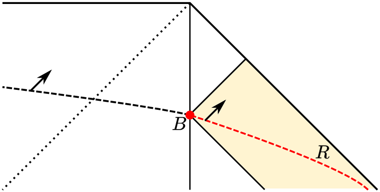


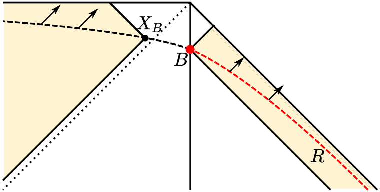
It gives a sharp test of whether or not the black hole S-matrix is unitary.

By now this calculation is well-known, so I'll just flash the picture:



By now this calculation is well-known, so I'll just flash the picture:





At early times the entanglement wedge of the Hawking radiation R just contains the Hawking radiation, so according to Hawking's picture of the bulk its entropy should grow with time. Eventually however it is advantageous to include an "island" in the black hole interior, which contains modes which purify the Hawking radiation entropy and whose generalized entropy also has a term proportional to the horizon area. This gives the decreasing part of the curve!

Although this calculation is beautiful, there are several immediate questions:

Although this calculation is beautiful, there are several immediate questions:

- If Hawking's picture of the bulk leads to information loss, why is it correct to use it in the QES calculation?

Although this calculation is beautiful, there are several immediate questions:

- If Hawking's picture of the bulk leads to information loss, why is it correct to use it in the QES calculation?
- What does this calculation say about the validity of assumption (3) in Hawking's paradox? In other words, what is the breakdown of EFT which avoids Hawking's contradiction?
- If quantum error correction is the fundamental justification of the QES formula, what is the nature of the code which is at work here?

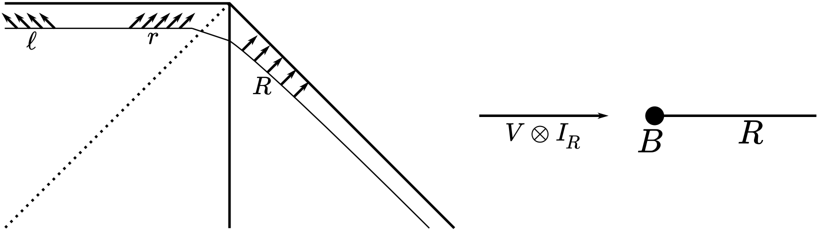
Although this calculation is beautiful, there are several immediate questions:

- If Hawking's picture of the bulk leads to information loss, why is it correct to use it in the QES calculation?
- What does this calculation say about the validity of assumption (3) in Hawking's paradox? In other words, what is the breakdown of EFT which avoids Hawking's contradiction?
- If quantum error correction is the fundamental justification of the QES formula, what is the nature of the code which is at work here?

In the last few minutes I'll briefly describe some recent work which we are optimistic gives answers to these questions. Akers/Engelhardt/Harlow/Penington/Vardhan

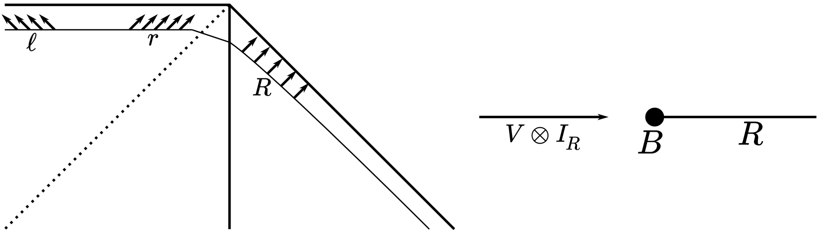
An encoding map for the BH interior

The idea is to look for a holographic encoding map $V : \mathcal{H}_\ell \otimes \mathcal{H}_r \to \mathcal{H}_B$ on a nice slice:



An encoding map for the BH interior

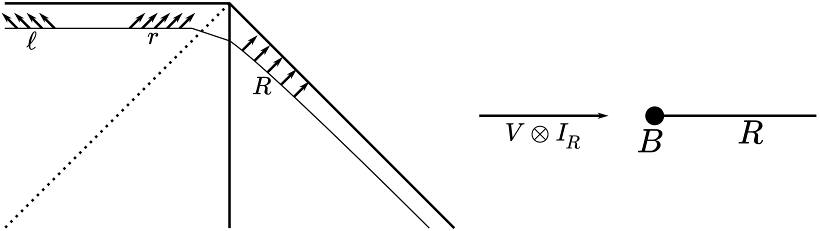
The idea is to look for a holographic encoding map $V : \mathcal{H}_\ell \otimes \mathcal{H}_r \to \mathcal{H}_B$ on a nice slice:



The basic problem, which is in some sense the essence of the information problem, is that at late times we have $|\ell||r| \gg |B|$, so the map V cannot be an isometry.

An encoding map for the BH interior

The idea is to look for a holographic encoding map $V : \mathcal{H}_\ell \otimes \mathcal{H}_r \to \mathcal{H}_B$ on a nice slice:

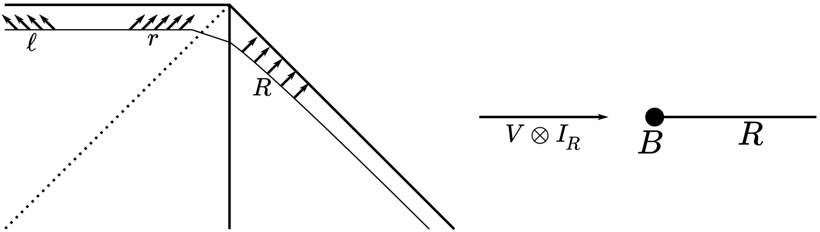


The basic problem, which is in some sense the essence of the information problem, is that at late times we have $|\ell||r| \gg |B|$, so the map V cannot be an isometry.
This sounds scary, but it also has a silver lining: since V is not an isometry we have

$$\mathrm{tr}_B\left((V \otimes I_R)\rho_{\ell r R}(V^\dagger \otimes I_R)\right) \neq \mathrm{tr}_{\ell r}\left(\rho_{\ell r R}\right).$$

An encoding map for the BH interior

The idea is to look for a holographic encoding map $V : \mathcal{H}_\ell \otimes \mathcal{H}_r \to \mathcal{H}_B$ on a nice slice:



The basic problem, which is in some sense the essence of the information problem, is that at late times we have $|\ell||r| \gg |B|$, so the map V cannot be an isometry.

This sounds scary, but it also has a silver lining: since V is not an isometry we have

$$\mathrm{tr}_B \Big((V \otimes I_R)\rho_{\ell r R}(V^\dagger \otimes I_R) \Big) \neq \mathrm{tr}_{\ell r}\left(\rho_{\ell r R}\right).$$

We argue this is the key distinction between Page and Hawking!

In more detail the essence of our proposal is the following:

There is a large set of "null states" in the Hilbert space of effective field theory inside a black hole, each of which is annihilated by the holographic map to the fundamental degrees of freedom. This however cannot be detected by any observer who does not perform an operation of exponential complexity.

In more detail the essence of our proposal is the following:

There is a large set of "null states" in the Hilbert space of effective field theory inside a black hole, each of which is annihilated by the holographic map to the fundamental degrees of freedom. This however cannot be detected by any observer who does not perform an operation of exponential complexity.

In the previous language we advocate the following replacement:

~~(3) EFT valid wherever there is not a large energy density/curvature.~~

(3*) EFT valid *for sub-exponential states/observables* wherever there is not large energy density/curvature.

In more detail the essence of our proposal is the following:

There is a large set of "null states" in the Hilbert space of effective field theory inside a black hole, each of which is annihilated by the holographic map to the fundamental degrees of freedom. This however cannot be detected by any observer who does not perform an operation of exponential complexity.

In the previous language we advocate the following replacement:

~~(3) EFT valid wherever there is not a large energy density/curvature.~~

(3*) EFT valid *for sub-exponential states/observables* wherever there is not large energy density/curvature.

Indeed we can construct models where appropriate analogous of (1), (2), and (3*) are all manifestly true. They are thus compatible, and so if we are willing to accept (3*) then the information problem is resolved in these models.

More concretely these models have the following features:

More concretely these models have the following features:

- $V \otimes I_R$ preserves the overlaps of all states of sub-exponential complexity

More concretely these models have the following features:

- $V \otimes I_R$ preserves the overlaps of all states of sub-exponential complexity
- In sub-exponential states the entropy of R can either be computed directly in the fundamental description or in the effective description using the QES formula Engelhardt/Wall 2014, with the same results up to $O(e^{-\gamma S_{BH}})$ for some $\gamma > 0$.

More concretely these models have the following features:

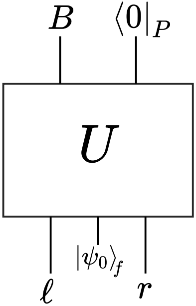
- $V \otimes I_R$ preserves the overlaps of all states of sub-exponential complexity
- In sub-exponential states the entropy of R can either be computed directly in the fundamental description or in the effective description using the QES formula Engelhardt/Wall 2014, with the same results up to $O(e^{-\gamma S_{BH}})$ for some $\gamma > 0$.
- Sub-exponential interior observables can be (non-linearly but unambiguously) reconstructed in the fundamental description, in agreement with the effective description up to $O(e^{-\gamma S_{BH}})$.

A static model

I do not have time to show any calculations, but I'll give you two examples.

A static model

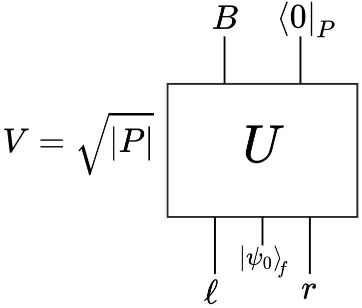
I do not have time to show any calculations, but I'll give you two examples. Our first model works likes this:

$V = \sqrt{|P|}$ 

Here U is a unitary drawn from the Haar ensemble.

A static model

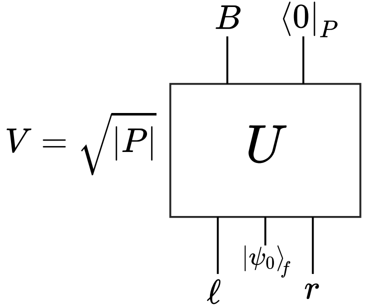
I do not have time to show any calculations, but I'll give you two examples.
Our first model works likes this:



Here U is a unitary drawn from the Haar ensemble.
V is not necessarily an isometry due to the post-selection onto $\langle 0|_P$.

A static model

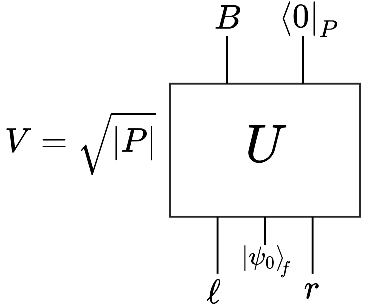
I do not have time to show any calculations, but I'll give you two examples. Our first model works likes this:



Here U is a unitary drawn from the Haar ensemble.
V is not necessarily an isometry due to the post-selection onto $\langle 0|_P$.
I emphasize that U is drawn *only once*; we are not doing "fundamental averaging".

A static model

I do not have time to show any calculations, but I'll give you two examples. Our first model works likes this:



Here U is a unitary drawn from the Haar ensemble.
V is not necessarily an isometry due to the post-selection onto $\langle 0|_P$.
I emphasize that U is drawn *only once*; we are not doing "fundamental averaging".
The claimed results follow from standard unitary integration technology.

A dynamical model

The previous model describes the holographic map on a single time-slice. What about dynamics?

A dynamical model

The previous model describes the holographic map on a single time-slice. What about dynamics?

- As time evolves the various Hilbert space dimensions all change, so we need to introduce

$$V_t : \mathcal{H}_{\ell(t)} \otimes \mathcal{H}_{r(t)} \to \mathcal{H}_{B(t)}.$$

A dynamical model

The previous model describes the holographic map on a single time-slice. What about dynamics?

- As time evolves the various Hilbert space dimensions all change, so we need to introduce

$$V_t : \mathcal{H}_{\ell(t)} \otimes \mathcal{H}_{r(t)} \to \mathcal{H}_{B(t)}.$$

- We'd then like to define (discrete) time evolutions U_t in the fundamental description and u_t in the effective description which are "equivariant" in the sense that

$$V_{t+1} u_t = U_{t+1} V_t.$$

A dynamical model

The previous model describes the holographic map on a single time-slice. What about dynamics?

- As time evolves the various Hilbert space dimensions all change, so we need to introduce
$$V_t : \mathcal{H}_{\ell(t)} \otimes \mathcal{H}_{r(t)} \to \mathcal{H}_{B(t)}.$$
- We'd then like to define (discrete) time evolutions U_t in the fundamental description and u_t in the effective description which are "equivariant" in the sense that
$$V_{t+1} u_t = U_{t+1} V_t.$$
("Evolve then encode" equals "encode then evolve")

A dynamical model

The previous model describes the holographic map on a single time-slice. What about dynamics?

- As time evolves the various Hilbert space dimensions all change, so we need to introduce

$$V_t : \mathcal{H}_{\ell(t)} \otimes \mathcal{H}_{r(t)} \to \mathcal{H}_{B(t)}.$$

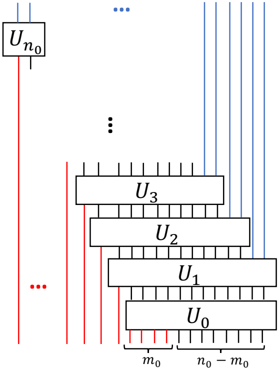
- We'd then like to define (discrete) time evolutions U_t in the fundamental description and u_t in the effective description which are "equivariant" in the sense that

$$V_{t+1} u_t = U_{t+1} V_t.$$

  ("Evolve then encode" equals "encode then evolve")
- We would like to model the full dynamical process proposed by Hawking: collapse some matter to form a black hole, and then watch it evaporate and see if the process is unitary.

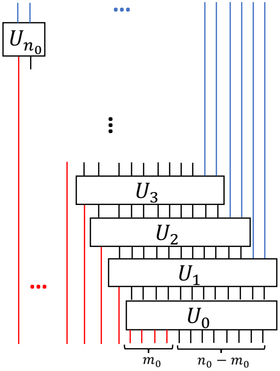
We'll start in the fundamental description.

We'll start in the fundamental description.



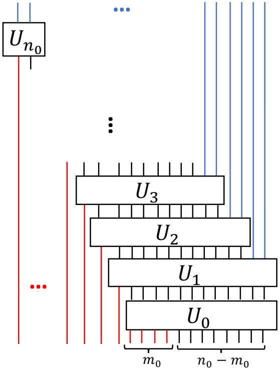
- We model the collapsing shell as m_0 qudits whose state we control and $n_0 - m_0$ qudits in a fixed state $|\psi_0\rangle_f$, with the collapse being implemented by a random unitary U_0.

We'll start in the fundamental description.



- We model the collapsing shell as m_0 qudits whose state we control and $n_0 - m_0$ qudits in a fixed state $|\psi_0\rangle_f$, with the collapse being implemented by a random unitary U_0.
- At each time t we then act with a random unitary U_{t+1} which absorbs one ingoing qudit from the reservoir and radiates two outgoing qudits.

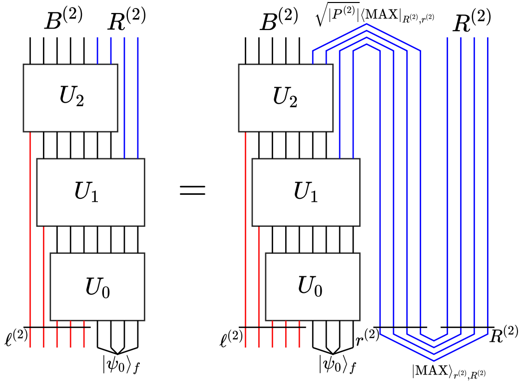
We'll start in the fundamental description.



- We model the collapsing shell as m_0 qudits whose state we control and $n_0 - m_0$ qudits in a fixed state $|\psi_0\rangle_f$, with the collapse being implemented by a random unitary U_0.
- At each time t we then act with a random unitary U_{t+1} which absorbs one ingoing qudit from the reservoir and radiates two outgoing qudits.
- This model clearly has a finite BH entropy and a unitary S-matrix.

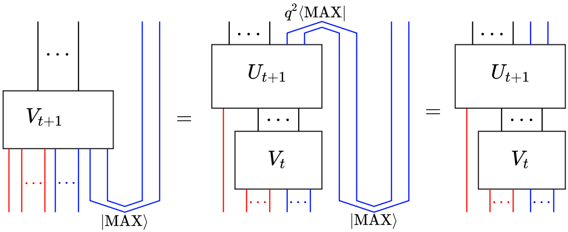
To understand the interior we need to define a holographic map V.

To understand the interior we need to define a holographic map V. Here is how it works:

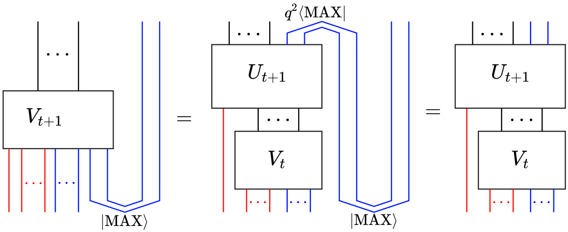


By "bending around" the outgoing modes using post-selection, we can re-interpret that fundamental dynamics as a non-isometric holographic map $V \otimes I_R$ acting on a (maximally-entangled) Hawking state in the effective description!

This map is easily shown to be equivariant:

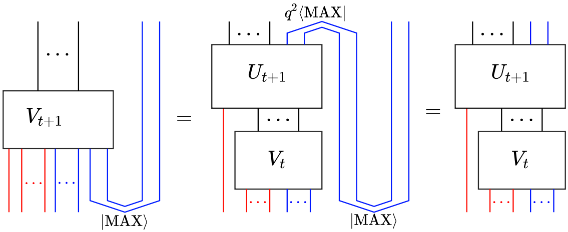


This map is easily shown to be equivariant:



We can also check that our results from the static case carry through:

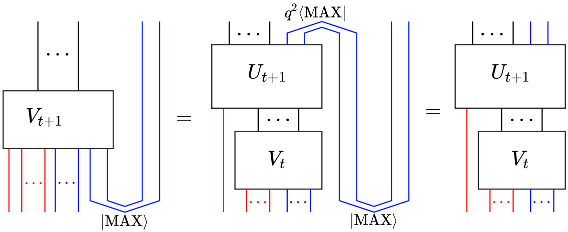
This map is easily shown to be equivariant:



We can also check that our results from the static case carry through:

- Approximate isometry on sub-exponential states

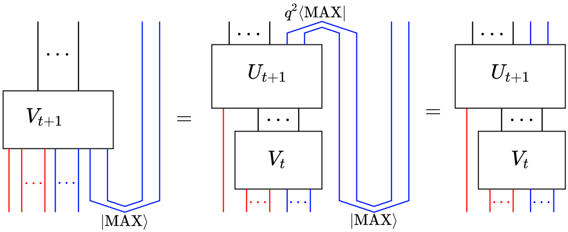
This map is easily shown to be equivariant:



We can also check that our results from the static case carry through:

- Approximate isometry on sub-exponential states
- QES formula valid for computing entropy

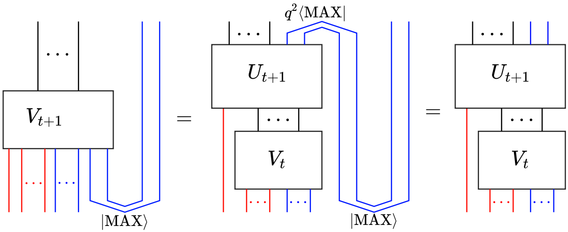
This map is easily shown to be equivariant:



We can also check that our results from the static case carry through:

- Approximate isometry on sub-exponential states
- QES formula valid for computing entropy
- Invertible encoding on sub-exponential states and a reconstructable (but non-linear) measurement theory.

This map is easily shown to be equivariant:



We can also check that our results from the static case carry through:

- Approximate isometry on sub-exponential states
- QES formula valid for computing entropy
- Invertible encoding on sub-exponential states and a reconstructable (but non-linear) measurement theory.

The Hayden-Preskill scrambling argument can also be checked in a more refined version of the model where the U_t are less random.

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013
- non-linearity Papadodimas/Raju 2013-2014

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013
- non-linearity Papadodimas/Raju 2013-2014
- quantum error correction Almheiri/Dong/Harlow 2014

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013
- non-linearity Papadodimas/Raju 2013-2014
- quantum error correction Almheiri/Dong/Harlow 2014
- null states many, recently Marolf/Maxfield 2020

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013
- non-linearity Papadodimas/Raju 2013-2014
- quantum error correction Almheiri/Dong/Harlow 2014
- null states many, recently Marolf/Maxfield 2020
- ...

Conclusion

I've presented what we view as a viable scenario for resolving the information problem.
It incorporates many of the ingredients which have been previously proposed:

- post-selection Horowitz/Maldacena 2003
- a cutoff on complexity Harlow/Hayden 2013
- non-linearity Papadodimas/Raju 2013-2014
- quantum error correction Almheiri/Dong/Harlow 2014
- null states many, recently Marolf/Maxfield 2020
- ...

It does *not* use Euclidean gravity in any way, which I view as a major advantage (no black box), but it does give results which agree with it where appropriate (i.e. the QES formula).

So far we have only realized this scenario in fairly crude toy models.

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance
- Diffeomorphism invariance

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance
- Diffeomorphism invariance
- Propagating gravitons

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance
- Diffeomorphism invariance
- Propagating gravitons
- ...

So far we have only realized this scenario in fairly crude toy models.
In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance
- Diffeomorphism invariance
- Propagating gravitons
- ...

The fact that we can resolve the paradox without including these indicates that they are not essential to it, but on the other hand perhaps something breaks once we try to include them? I do not think so, but we should check.

So far we have only realized this scenario in fairly crude toy models. In particular they do not have

- Locality (i.e. a definite number of spacetime dimensions)
- Continuous time
- Lorentz invariance
- Diffeomorphism invariance
- Propagating gravitons
- ...

The fact that we can resolve the paradox without including these indicates that they are not essential to it, but on the other hand perhaps something breaks once we try to include them? I do not think so, but we should check.

I am optimistic that the encoding map V could be constructed fairly uniquely in AdS/CFT from the "extrapolate" dictionary, but this would also require an assumption that the horizon is smooth for quick-collapse states. Ideally we could instead derive this from some non-perturbative version of string theory.