Entanglement Wedge Reconstruction using the Petz Map

Chi Fang Chen[1] and Geoffrey Penington[1]

[1] *Stanford Institute for Theoretical Physics, Stanford University, Stanford CA 94305 USA*

February 11, 2019

Abstract

It has previously been shown that entanglement wedge reconstruction can be achieved in AdS/CFT using a universal recovery channel known as the twirled Petz map. However, this map involves a complicated averaging procedure over bulk and boundary modular time and hence has proved somewhat intractable to evaluate in practice. We show that a much simpler channel, the Petz map, is sufficient for entanglement wedge reconstruction for any code space of fixed finite dimension – no twirling is required. Moreover, the error in the reconstruction will always be non-perturbatively small. From a quantum information perspective, our results extend the use of the Petz map as a general-purpose recovery channel to subsystem and operator algebra quantum error correcting codes.

chifangc@stanford.edu, geoffp@stanford.edu

Contents

1 Introduction **2**

2 Entanglement Wedge Reconstruction and the Petz Map **4**

3 Proof of Theorem 1 **5**

4 Discussion **8**

5 Acknowledgements **8**

1 Introduction

Despite extraordinary progress in the twenty years since it was first proposed, many aspects of the AdS/CFT correspondence remain deeply mysterious. If we want to understand how a bulk gravitational spacetime can be encoded in a non-gravitational boundary field theory, we first need to understand the region of the bulk spacetime that is encoded in (and hence can be 'reconstructed' from) a given boundary subregion A.

Over the course of the last five years, this question has been answered: any boundary region A encodes its 'entanglement wedge' a. This is the bulk region bounded by the region A itself and the Ryu-Takayanagi surface χ_A [1], the minimal area bulk surface anchored to the boundary of A; see Figure 1.[1]

More specifically, given any bulk operator ϕ_a lying within the entanglement wedge a, there exists a boundary operator ϕ_A acting only on the boundary region A which approximately reproduces the action of the bulk operator ϕ_a. The task of finding such an operator ϕ_A is known as entanglement wedge reconstruction.

The conjecture of entanglement wedge reconstruction was developed in [5, 6, 7] and established with increasing levels of rigour in [8, 9, 10]. It was shown in [4, 11], that the error in the reconstruction can be made non-perturbatively small at large gauge group rank N, or equivalently small gravitational coupling G_N.

All this progress was made possible by the realisation in [12] that the task of bulk reconstruction can be rephrased in the language of quantum error correction. Bulk operators in AdS/CFT are only well defined for the "code subspace" $\mathcal{H}_{\text{code}}$ of states with the correct smooth bulk geometry. If we let $J : \mathcal{H}_{\text{code}} \to \mathcal{H}_{CFT}$ be the embedding isometry from this code subspace to the larger CFT Hilbert space, entanglement wedge reconstruction can be rephrased as the task of finding a decoding channel $\mathcal{D}$ that error corrects the channel $\mathcal{N} = [J(\cdot)J^\dagger]_A$, where ρ_A is the restriction of the boundary state ρ to region A. More specifically, entanglement wedge reconstruction is equivalent to the existence of a decoding channel $\mathcal{D}$ such that for all states ρ in the bulk code space,

$$\mathcal{D} \circ \mathcal{N}(\rho) \approx \rho_a, \tag{1}$$

where ρ_a is the restriction of the bulk state ρ to the entanglement wedge a. If such a decoding channel exists, then we can use the adjoint channel $\mathcal{D}^\dagger$, defined by

$$\mathrm{Tr}[\mathcal{D}^\dagger(\phi)\,\sigma] = \mathrm{Tr}[\phi\,\mathcal{D}(\sigma)], \tag{2}$$

for all operators ϕ and states σ, to map bulk operators ϕ_a to boundary reconstructions $\phi_A = \mathcal{D}^\dagger(\phi_a)$ that act only in region A. Since

$$\mathrm{Tr}(\phi_A\,\rho) = \mathrm{Tr}[\phi_a\,\mathcal{D} \circ \mathcal{N}(\rho)] \approx \mathrm{Tr}(\phi_a\,\rho), \tag{3}$$

the expectation values of ϕ_a and ϕ_A approximately agree for all states $\rho \in S(\mathcal{H}_{\text{code}})$. It can be shown that this is also true for higher point correlators [10].

Interestingly, the entanglement wedge a may contain regions outside of the 'causal wedge' of A (the intersection of the past and future of the boundary domain of dependence of A). Given a bulk operator ϕ in the causal wedge of a region A, it is well-understood how to reconstruct the operator ϕ within the boundary region A, given only the bulk and boundary equations of motion, using the so-called HKLL procedure [13]. It was only by introducing the tool of quantum error correction from quantum information that we have begun to understand that the entire entanglement wedge (rather than just the causal wedge) can be reconstructed from region A.

The first clue that more than just the causal wedge was encoded in a boundary region A was given by the Ryu-Takayanagi formula [1, 14]. Including the leading quantum correction [15], the Ryu-Takayanagi formula states that the

[1]This definition is valid within a single static timeslice of a static bulk spacetime. More generally, and more formally, the Ryu-Takayanagi surface χ_A is defined the smallest surface of *extremal* area homologous to A [2]. The entanglement wedge is then the domain of dependence of *any* achronal bulk surface bounded by A and χ_A. At higher orders in perturbation theory, one should use the quantum extremal surface, which extremises the Ryu-Takayanagi formula $\mathcal{A}/4G_N + S_{\text{bulk}}$, where S_{bulk} is the bulk entanglement entropy, rather than simply the classical area $\mathcal{A}$ [3, 4].

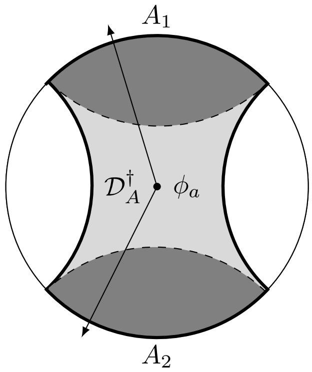

Figure 1: An operator ϕ_a, acting on the entanglement wedge a of $A = A_1 \cup A_2$, can be reconstructed on the boundary region A by the map $\mathcal{D}_A^\dagger : M_a \to M_A$. The solid interior curves represent the RT surface of A and the entire shaded region forms the entanglement wedge a (restricted to a single timeslice). The darker gray areas are the entanglement wedges of A_1 and A_2 individually, and also together form the causal wedge of A. Since the operator ϕ_a is not in the causal wedge of A, we cannot reconstruct it simply by using the bulk and boundary equations of motion. The more sophisticated machinery of quantum error correction is required. Moreover, ϕ_a can only be reconstructed on $A = A_1 \cup A_2$; neither A_1 nor A_2 alone contains any information about ϕ_a.

entanglement entropy S_A of any boundary region A is given by

$$S_A = \frac{\mathcal{A}(\chi_A)}{4G_N} + S_{\text{bulk}}, \tag{4}$$

where $\mathcal{A}(\chi_A)$ is the area of the RT surface χ_A and S_{bulk} is the bulk entanglement entropy associated to the entanglement wedge of A. A quantity that depends only on the reduced density matrix of the state on region A is therefore encoding information that depends on the entire entanglement wedge. Of course, the entanglement entropy is not itself an observable. Somewhat remarkably, however, (4) alone is sufficient to imply the existence of decoding channels $\mathcal{D}$ that can be used for entanglement wedge reconstruction [9, 10]. The key intermediate step, which was shown in [8], is that (4) implies an approximate equality between bulk and boundary relative entropies.

Unfortunately, even though it is, at this point, very well established that entanglement wedge reconstruction is possible in principle (and hence that decoding channels $\mathcal{D}$ must exist), it has proved somewhat challenging to find explicit and practical constructions that work for bulk operators lying outside the causal wedge (and hence for which we cannot use the HKLL prescription). An explicit, if somewhat impractical, general construction was given in [9, 12]. However, this construction relies on the unphysical assumption of exact quantum error correction, which does not exist at finite N.

It was shown in [16] that the evolution of bulk operators in bulk modular time is related via the extrapolate dictionary to the evolution of boundary operators in boundary modular time. Since bulk modular evolution should be linear in the free field approximation $N \to \infty$, one might hope to expand a bulk operator at any point in the entanglement wedge in terms of the modular evolution of operators at the boundary of the wedge – and thus derive an explicit formula for the boundary representation of the bulk operator. However, as yet, the details of this expansion remain unknown, and it is not even clear how to show rigorously that one should exist at all.

Finally, it was demonstrated in [10], using the tools of approximate operator algebra quantum error correction, that entanglement wedge reconstruction can be achieved using the *twirled Petz map*. Rather than being designed with holography in mind, the twirled Petz map is a "universal recovery map", a general purpose decoding channel that is defined for arbitrary quantum error correcting codes. Given an encoding channel $\mathcal{N}$ a fixed state $\sigma \in S(\mathcal{H}_{\text{code}})$, the twirled Petz map has the somewhat complicated form

$$\mathcal{R}_{\sigma,\mathcal{N}}(\rho) = \int dt\, \frac{\pi}{2}\,(\cosh(\pi t) + 1)^{-1}\, \sigma^{\frac{1-it}{2}} \mathcal{N}^\dagger \left([\mathcal{N}(\sigma)]^{-\frac{1-it}{2}} \rho\, [\mathcal{N}(\sigma)]^{-\frac{1+it}{2}} \right) \sigma^{\frac{1+it}{2}}, \tag{5}$$

If we replace σ by the maximally mixed state ω and use the channel $\mathcal{N} = [J(\cdot)J^\dagger]_A$, this leads to the boundary reconstruction ϕ_A of a bulk operator ϕ_a being defined as

$$\phi_A = \mathcal{R}^\dagger_{\omega,\mathcal{N}}(\phi_a) = \frac{1}{d_{\text{code}}} \int dt\, \frac{\pi}{2}\,(\cosh(\pi t) + 1)^{-1}\, \omega_A^{-\frac{1-it}{2}} [J\phi_a J^\dagger]_A \omega_A^{-\frac{1+it}{2}}, \tag{6}$$

where $\omega_A = \mathcal{N}(\omega)$. By using the maximally mixed state, the expression has simplified somewhat. However, it still involves an averaging over boundary modular time with the precisely chosen weighting $\pi/2\,(\cosh(\pi t) + 1)^{-1}$.

In this paper, we will show that such averaging is unnecessary for code spaces of any fixed finite dimension in the semiclassical limit $N \to \infty$ and $G_N \to 0$. Instead it is sufficient to use the much simpler *Petz map* reconstruction

$$\phi_A = \frac{1}{d_{\text{code}}} \omega_A^{-1/2} [J\phi_a J^\dagger]_A\, \omega_A^{-1/2}. \tag{7}$$

We are hopeful that this simpler recovery map should prove significantly more practical to evaluate explicitly; we discuss the challenges and prospects of doing so in Section 4.

Our strategy for proving the efficacy of the Petz map for entanglement wedge reconstruction builds on work by Barnum and Knill [17], who showed that, for ordinary subspace quantum error correction, the Petz map will always have an *average* decoding error that is almost as small as the average error of the optimal decoding channel. In other words, the Petz map is always 'pretty good'. We extend this work to subsystem and operator algebra quantum error correcting codes and then show that the average error can always be used to bound the worst-case error so long as the dimension of the code space does not grow too quickly in the limit of large N. (We discuss very large code spaces, such as code spaces of black hole microstates, briefly in Section 4.)

In Section 2, we formalise the problem of entanglement wedge reconstruction in the language of quantum error correction and show how to adapt the results of Barnum and Knill to prove that reconstruction is possible using the Petz map. The proof of our main technical result is postponed to Section 3. Section 4 consists of a brief discussion of potential applications and extensions of our work.

2 Entanglement Wedge Reconstruction and the Petz Map

In order to apply information-theoretic techniques to the problem of entanglement wedge reconstruction, we first need to rephrase our task in the language of quantum information. The framework that we use is the same framework used in [10] – finite-dimensional approximate operator algebra quantum error correction.

The AdS/CFT correspondence is a duality between a boundary conformal field theory, with Hilbert space $\mathcal{H}_{CFT}$, and a bulk quantum gravity theory. In principle, if AdS/CFT is supposed to be a true duality between theories, the 'bulk' Hilbert space should be isomorphic to the boundary Hilbert space $\mathcal{H}_{CFT}$. However, a complete, non-perturbative, microscopic description of the entire Hilbert space from a purely bulk perspective, if one exists, remains unknown. Moreover, any such Hilbert space would be dominated by large black holes. Instead, we are normally only interested in a small 'code subspace' $\mathcal{H}_{\text{code}}$ of states with a smooth semiclassical bulk geometry; for example, we might consider small bulk perturbations about the vacuum state. We therefore have an isometry $J : \mathcal{H}_{\text{code}} \to \mathcal{H}_{CFT}$. Equivalently, we can consider the quantum channel $\mathcal{J} = J(\cdot)J^\dagger$ which maps bulk density matrices to boundary density matrices; in fact none of our results rely on $\mathcal{J}$ being an isometry as opposed to a general quantum channel.

We assume that both $\mathcal{H}_{\text{code}}$ and $\mathcal{H}_{CFT}$ are finite-dimensional. In the case of $\mathcal{H}_{\text{code}}$, this is justified by the fact that we cannot include arbitrarily high energy excitations in the bulk without causing significant backreaction and eventually creating a black hole. In the case of $\mathcal{H}_{CFT}$, we should be able to regularise the boundary theory in the UV, while only affecting bulk physics close to the boundary. Of course, the real value of these assumptions for our purposes is that they allow us to apply known results from the large literature on finite-dimensional quantum error correction.

We denote the algebra of observables on the Hilbert space $\mathcal{H}_{\text{code}}$ by $\mathcal{B}(\mathcal{H}_{\text{code}})$ and the algebra of observables on $\mathcal{H}_{CFT}$ by $\mathcal{B}(\mathcal{H}_{CFT})$. The entanglement wedge a has an associated von Neumann subalgebra $\mathcal{M}_a \overset{i}{\hookrightarrow} \mathcal{B}(\mathcal{H}_{\text{code}})$, consisting of bulk observables that act only on a; similarly, the boundary region A is associated with a von Neumann subalgebra $\mathcal{M}_A \overset{i}{\hookrightarrow} \mathcal{B}(\mathcal{H}_{CFT})$. We use the notation from [10], where the space of density matrices on a von Neumann subalgebra $\mathcal{M}$ acting on a Hilbert space $\mathcal{H}$ is denoted by $S(\mathcal{M}) \cong S(\mathcal{H}) \cap \mathcal{M}$. This space is isomorphic to the space of positive normalised linear functionals on the algebra. See the appendix of [10] for more details.

The question of entanglement wedge reconstruction can then be rephrased as the question of whether the channel $\mathcal{N} = \mathcal{J}(\cdot)|_A$ forms an approximate error-correcting code for the algebra $\mathcal{M}_a$. Here, the restriction channel $(\cdot)|_A$ simply projects the density matrix onto the algebra $\mathcal{M}_A$. In other words, entanglement wedge reconstruction is possible if (and only if) there exists a decoding channel $\mathcal{D} : S(\mathcal{M}_A) \to S(\mathcal{M}_a)$ such that

$$\mathcal{D} \circ \mathcal{N}(\rho) \approx \rho|_a, \tag{8}$$

for all states $\rho \in S(\mathcal{H}_{\text{code}})$. The restriction $\rho|_a$ is, of course, the projection of ρ onto $\mathcal{M}_a$. For subsystem algebras, this corresponds to taking a partial trace over the complementary subsystem and hence agrees with the usual notion of a reduced density matrix; operator algebra quantum error correction therefore generalises subsystem quantum error correction.

In the Heisenberg (adjoint) picture, this condition becomes

$$\mathcal{N}^\dagger \circ \mathcal{D}^\dagger(\phi_a) = \mathcal{J}^\dagger \circ \mathcal{D}^\dagger(\phi_a) \approx \phi_a. \tag{9}$$

Note that, since the adjoint of the restriction channel is simply the embedding of the subalgebra in the larger algebra of observables, $\mathcal{N}^\dagger(O_A) = \mathcal{J}^\dagger(O_A)$ for all operators $O_A \in \mathcal{M}_A$. In other words, $\phi_A = \mathcal{D}^\dagger(\phi_a)$ acts in approximately the same way as ϕ_a

$$\mathrm{Tr}(\phi_A \mathcal{J}(\rho)) \approx \mathrm{Tr}(\phi_a \rho). \tag{10}$$

The complete setup, in both the Schrödinger and Heisenberg pictures, is shown in Figure 2.

$$
\begin{tikzcd}
\mathcal{M}_a \arrow[r, hook, "i_a"] \arrow[d, dashed, "\mathcal{D}^\dagger"'] & \mathcal{B}(\mathcal{H}_{\text{code}}) \\
\mathcal{M}_A \arrow[r, hook, "i_A"'] \arrow[ur, "\mathcal{N}^\dagger"] & \mathcal{B}(\mathcal{H}_{\text{CFT}}) \arrow[u, "\mathcal{J}^\dagger"']
\end{tikzcd}
\qquad
\begin{tikzcd}
S(\mathcal{M}_a) & S(\mathcal{H}_{\text{code}}) \arrow[l, "res_a"'] \arrow[dl, "\mathcal{N}"] \arrow[d, "\mathcal{J}"] \\
S(\mathcal{M}_A) \arrow[u, dashed, "\mathcal{D}"] & S(\mathcal{H}_{\text{CFT}}) \arrow[l, "res_A"]
\end{tikzcd}
$$

Figure 2: In the Heisenberg picture, $\mathcal{M}_a \overset{i}{\hookrightarrow} \mathcal{B}(\mathcal{H}_{\text{code}})$ and $\mathcal{M}_A \overset{i}{\hookrightarrow} \mathcal{B}(\mathcal{H}_{CFT})$ are von Neumann subalgebras acting on the code space $\mathcal{H}_{\text{code}}$ and CFT Hilbert space $\mathcal{H}_{CFT}$ respectively. The Heisenberg channel $\mathcal{J}^\dagger = J^\dagger(\cdot)J$ maps boundary observables to their projection in the code space. The task of entanglement wedge reconstruction is to find a Heisenberg decoding channel $\mathcal{D}^\dagger : \mathcal{M}_a \to \mathcal{M}_A$ that maps bulk observables ϕ_a in $\mathcal{M}_a$ to boundary observables ϕ_A in $\mathcal{M}_A$. When projected into the code space using $\mathcal{J}^\dagger$, the boundary observable ϕ_A should reproduce the original bulk observable ϕ_a. In the Schrödinger picture, the directions of all channels are reversed. The channel $\mathcal{J}$ now maps bulk states to the corresponding boundary states. The Heisenberg channel i_a and i_A embedding the von Neumann subalgebras $\mathcal{M}_a$ and $\mathcal{M}_A$ into the larger algebras of observables $\mathcal{B}(\mathcal{H}_{\text{code}})$ and $\mathcal{B}(\mathcal{H}_{CFT})$ are adjoints of the restriction maps onto $S(\mathcal{M}_a)$ and $S(\mathcal{M}_A)$ respectively. Finally, the decoding channel $\mathcal{D} : S(\mathcal{M}_A) \to S(\mathcal{M}_a)$ now needs to satisfy $\mathcal{D}[\mathcal{J}(\cdot)_A] \approx (\cdot)_a$.

It was argued in [10] that the twirled Petz map provides an example of such a decoding map, with an error that is perturbatively suppressed in $1/N$. It was then shown in [11] that there must exist some decoding channel $\mathcal{D}$ with a non-perturbatively small error; however, this argument was non-constructive. Both results relied heavily on the approximate equality between bulk and boundary relative entropies found in [8]. To show the existence of a decoding channel that is accurate to all orders in perturbation theory requires the refined variant of this approximate equality that was derived in [4]. Here, we shall simply take as our starting assumption the existence of *some* good decoding channel $\mathcal{D}'$. We shall not need to know anything about the details of the channel. We can therefore use the result of [11] to assume that the error for this channel is non-perturbatively small. The following theorem, which we prove in Section 3, then implies that the Petz map is also a good decoding channel:

Theorem 1. *Let $\mathcal{M}_a \overset{i}{\hookrightarrow} \mathcal{B}(\mathcal{H}_{code})$ be a von Neumann subalgebra acting on the code space $\mathcal{H}_{code}$ with dimension d_{code}, let $\mathcal{N}$ be a quantum channel, and suppose that there exists a channel $\mathcal{D}'$ such that $\|\mathcal{D}' \circ \mathcal{N}(\rho) - \rho_a\|_1 < \delta$. Let*

$$\mathcal{P}_{\omega,\mathcal{N}} := \frac{1}{d_{code}} \mathcal{N}^\dagger \left[\mathcal{N}(\omega)^{-1/2}(\cdot)\mathcal{N}(\omega)^{-1/2}\right]$$

be the Petz map with maximally mixed reference state ω. Then

$$\|\mathcal{P}_{\omega,\mathcal{N}} \circ \mathcal{N}(\rho)|_a - \rho_a\|_1 \leq d_{code}\sqrt{2\delta}. \tag{11}$$

Note that our bound on the error when reconstructing the reduced state using the Petz map $\mathcal{P}_{\omega,\mathcal{N}}$ is significantly higher than the original error δ. Not only is the our bound proportional to $\sqrt{\delta}$, but it is also proportional to the dimension d_{code} of the code space. As we shall see in Section 3, the square root appears because of inefficiencies in converting between trace distances and fidelities using the Fuchs-van de Graaf inequalities [18], while the factor of d_{code} appears in order to convert a bound on the average-case error into a bound on the worst-case error.

However, so long as the error using the original decoding channel $\mathcal{D}'$ was non-perturbatively small, the Petz map error will also be non-perturbatively small so long as the dimension of the code space does not grow superpolynomially in the limit of large N. For most code spaces of interest, such as perturbations about the vacuum, the code space dimension will be $O(1)$ and so this factor is unproblematic. We discuss very large code spaces, such as code spaces containing large numbers of black hole microstates, briefly in Section 4. However, so long as we stick to the confines of perturbative excitations of quantum fields in a fixed gravitational background, the Petz map can always be trusted. No twirling is required.

3 Proof of Theorem 1

Spiritually, Theorem 1 is rooted in the following classic result about ordinary subspace quantum error correction:

Theorem 2 ([17])**.** *Given any pair of quantum channels $\mathcal{D}'$, $\mathcal{N}$, and ensemble of commuting density matrices (p_k, ρ_k) whose sum $\sum_k p_k \rho_k = \rho$, the Petz map*

$$\mathcal{P}_{\rho,\mathcal{N}}[\cdot] := \rho^{1/2}\mathcal{N}^\dagger \left[\mathcal{N}(\rho)^{-1/2}(\cdot)\mathcal{N}(\rho)^{-1/2}\right] \rho^{1/2}$$

with reference state ρ, satisfies

$$\sum_i p_k F(\rho_k, \mathcal{P}_{\mathcal{N},\rho} \circ \mathcal{N}) \geq (\sum_k p_k F(\rho_k, \mathcal{D}' \circ \mathcal{N}))^2. \tag{12}$$

Here, the entanglement fidelity $F(\sigma, \mathcal{Z})$ is defined by

$$F(\sigma, \mathcal{Z}) := \langle\sigma| V_{\mathcal{Z}}^{\dagger} \left(|\sigma\rangle\langle\sigma| \otimes \mathbb{1}_E\right) V_{\mathcal{Z}} |\sigma\rangle ,$$

where $|\sigma\rangle \in \mathcal{H}_{code} \otimes \mathcal{H}_R$ is a purification of $\sigma \in S(\mathcal{H}_{code})$ and $V_{\mathcal{Z}} : \mathcal{H}_{code} \to \mathcal{H}_{code} \otimes \mathcal{H}_E$ is a Stinespring extension of the channel $\mathcal{Z} : S(\mathcal{H}_{code}) \to S(\mathcal{H}_{code})$.

Theorem 2 states that $\mathcal{P}_{\mathcal{N},\rho} \circ \mathcal{N}$ is close to the identity when measured using the average entanglement fidelity of an ensemble $\{\rho_i\}$ with average state ρ. Note that, unlike in Theorem 1, there is no factor of d_{code} in the size of the error for the Petz map $\mathcal{P}_{\mathcal{N},\rho}$ compared to the original decoding channel $\mathcal{D}'$. Instead, (12) implies that the error, measured using the average entanglement fidelity, increases by at most a factor of two.[2] The factor of d_{code} will appear when we convert this average error into a worst-case error.

For concreteness, let us write down an explicit basis for the von Neumann subalgebra $\mathcal{M}_a$. (The exact description of $\mathcal{J}$ and $\mathcal{M}_A$ (and hence $\mathcal{N}$) are unimportant for our purposes.) It is a fact about finite-dimensional von Neumann algebras that we can always find a set of Hilbert spaces $\mathcal{H}_\alpha$ and $\mathcal{H}_{\bar{\alpha}}$, parameterised by α, such that [19]

$$\begin{aligned} \mathcal{M}_a &= \bigoplus_\alpha \mathcal{B}(\mathcal{H}_\alpha) \otimes \mathbb{1}_{\bar{\alpha}}, \\ \mathcal{H}_{code} &= \bigoplus_\alpha \mathcal{H}_\alpha \otimes \mathcal{H}_{\bar{\alpha}} \end{aligned} \tag{13}$$

Note that

$$\sum_\alpha d_\alpha d_{\bar{\alpha}} = d_{\text{code}}, \tag{14}$$

where d_α, $d_{\bar{\alpha}}$ and d_{code} are the dimensions of $\mathcal{H}_\alpha$, $\mathcal{H}_{\bar{\alpha}}$ and $\mathcal{H}_{\text{code}}$ respectively. In this basis, any state $\rho_a \in S(\mathcal{M}_a)$ can be parameterised as

$$\rho_a = \sum_\alpha p_\alpha \rho_\alpha \otimes \omega_{\bar{\alpha}} = \sum_{\alpha, i_\alpha} p_\alpha p_{i_\alpha}^{(\alpha)} |i_\alpha\rangle\langle i_\alpha| \otimes \omega_{\bar{\alpha}}, \tag{15}$$

where the states $\omega_{\bar{\alpha}} \in S(\mathcal{H}_{\bar{\alpha}})$ are maximally mixed, $\rho_\alpha \in S(\mathcal{H}_\alpha)$ are normalised density matrices, p_α and $p_i^{(\alpha)}$ are normalised probability distributions, and $|i_\alpha\rangle$ forms an orthonormal basis for $\mathcal{H}_\alpha$.

We now have all the ingredients we need to begin a proof of Theorem 1. Let $\mathcal{Z} = \mathcal{P}_{\omega,\mathcal{N}} \circ \mathcal{N}$. We first note that $\mathcal{Z}$ is a self-adjoint superoperator. For any operator ϕ,

$$\begin{aligned} \operatorname{Tr}[\phi\, \mathcal{Z}(\rho)] &= \frac{1}{d_{\text{code}}} \operatorname{Tr}\left[\phi \mathcal{N}^\dagger \left(\mathcal{N}(\omega)^{-1/2} \mathcal{N}(\rho) \mathcal{N}(\omega)^{-1/2}\right)\right] && (16) \\ &= \operatorname{Tr}[\mathcal{Z}(\phi)\rho] = \operatorname{Tr}[\phi\, \mathcal{Z}^\dagger(\rho)]. && (17) \end{aligned}$$

Hence we have $\mathcal{Z} = \mathcal{Z}^\dagger$. Note that this argument relied crucially on the fact that the reference state for the Petz map $\mathcal{P}_{\omega,\mathcal{N}}$ is maximally mixed.

Let $\phi_a \in \mathcal{M}_a$ be a Hermitian operator, which we can assume to have eigendecomposition

$$\phi_a = \lambda_{i_\alpha} |i_\alpha\rangle\langle i_\alpha| . \tag{18}$$

Then we can bound the operator norm

$$\begin{aligned} \|\mathcal{Z}^\dagger(\phi_a) - \phi_a\|_\infty &\leq \|\mathcal{Z}^\dagger(\phi_a) - \phi_a\|_1 && (19) \\ &\leq \sum_{\alpha, i_\alpha} |\lambda_{i_\alpha}|\, \|(\mathcal{Z}^\dagger - \mathbb{1})[|i_\alpha\rangle\langle i_\alpha| \otimes \mathbb{1}_{\bar{\alpha}}]\|_1 && (20) \\ &= \sum_{\alpha, i_\alpha} |\lambda_{i_\alpha}|\, d_{\bar{\alpha}} \|\mathcal{Z}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1 && (21) \end{aligned}$$

where the first inequality uses the monotonicity of the Schatten p-norms, the second inequality used the triangle inequality, and, in the final equality, we factored out $d_{\bar{\alpha}}$ so that $\rho_\alpha^{i_\alpha} = |i_\alpha\rangle\langle i_\alpha| \otimes \omega_{\bar{\alpha}}$ are normalised states, and more importantly we used the fact that the channel $\mathcal{Z}$ is self-adjoint. We now simply need to bound the average trace norm error of the channel $\mathcal{Z}$ on states $\rho_a \in S(\mathcal{M}_a)$. This is quadratically controlled by Theorem 2:

[2] An entanglement fidelity $F(\rho, \mathcal{D} \circ \mathcal{N}) = 1$ implies perfect recovery of a purification of ρ. Hence, we can naturally quantify the recovery error when decoding using the channel $\mathcal{D}'$ by

$$\delta = 1 - \sum_k p_k F(\rho_k, \mathcal{D}' \circ \mathcal{N}).$$

The equivalent error measure, when decoding using the Petz map $\mathcal{P}_{\rho,\mathcal{N}}$, will then be bounded by $2\,\delta$.

Proposition 2.1.

$$\sum_{i_\alpha,\alpha} \frac{d_{\bar{\alpha}}}{d_{code}} \|\mathcal{Z}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1^2 \leq 2\delta \tag{22}$$

Proof. We first note that

$$\sum_{i_\alpha,\alpha} \frac{d_{\bar{\alpha}}}{d_{\text{code}}} \rho_\alpha^{i_\alpha} = \omega. \tag{23}$$

Hence

$$\sum_{i_\alpha,\alpha} \frac{d_{\bar{\alpha}}}{d_{code}} \|\mathcal{Z}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1^2 \leq \sum_{i_\alpha,\alpha} \frac{d_{\bar{\alpha}}}{d_{\text{code}}} (1 - F(\rho_\alpha^{i_\alpha}, \mathcal{Z}[\rho_\alpha^{i_\alpha}])) \tag{24}$$

$$\leq 1 - (\sum_\alpha \frac{d_{\bar{\alpha}}}{d_{\text{code}}} F(\rho_\alpha^{i_\alpha}, \mathcal{D}' \circ \mathcal{N}[\rho_\alpha^{i_\alpha}]))^2 \tag{25}$$

$$\leq 1 - (\sum_\alpha \frac{d_{\bar{\alpha}}}{d_{\text{code}}} (1 - \|\mathcal{D}' \circ \mathcal{N}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1))^2 \tag{26}$$

$$\leq 2\delta, \tag{27}$$

where the first inequality uses one of the Fuchs-van de Graaf inequalities [18], the second uses (23) and Theorem 2, the fourth again uses the Fuchs-van de Graaf inequalities, and the fifth uses our assumption $\|\mathcal{D}'\mathcal{N}(\rho) - \rho_a\|_1 < \delta$ and (14). ☐

Applying Proposition 2.1 to (19), we find

$$\sum_{\alpha,i_\alpha} |\lambda_{i_\alpha}|\, d_{\bar{\alpha}} \|\mathcal{Z}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1 \leq \|\phi_a\|_\infty \sum_{\alpha,i_\alpha} \sqrt{d_{\bar{\alpha}} d_{\text{code}}} \cdot \sqrt{\frac{d_{\bar{\alpha}}}{d_{\text{code}}}} \|\mathcal{Z}[\rho_\alpha^{i_\alpha}] - \rho_\alpha^{i_\alpha}\|_1 \tag{28}$$

$$\leq \|\phi_a\|_\infty \sqrt{\sum_{\alpha,i_\alpha} d_{\bar{\alpha}} d_{\text{code}}} \cdot \sqrt{2\delta} \tag{29}$$

$$= \|\phi_a\|_\infty\, d_{\text{code}} \cdot \sqrt{2\delta}, \tag{30}$$

where, in the first inequality, we used the fact that $\|\phi_a\|_\infty \geq |\lambda_{i_\alpha}|$ for all λ_{i_α} and, in the second inequality, we used the Cauchy-Schwarz inequality. We therefore find that

$$\|\mathcal{Z}^\dagger(\phi_a) - \phi_a\|_\infty \leq \|\phi_a\|_\infty d_{code} \sqrt{2\delta}.$$

Since

$$\|\mathcal{Z}(\rho)_a - \rho_a\|_1 = \sup_{\phi \in \mathcal{B}(\mathcal{H}_{\text{code}})} \frac{1}{\|\phi\|_\infty} \operatorname{Tr}(\phi[\mathcal{Z}(\rho)_a - \rho_a]) \tag{31}$$

$$= \sup_{\phi_a \in \mathcal{M}_a} \frac{1}{\|\phi_a\|_\infty} \operatorname{Tr}([\mathcal{Z}^\dagger(\phi_a) - \phi_a]\rho) \leq \sup_{\phi_a \in \mathcal{M}_a} \frac{1}{\|\phi_a\|_\infty} \|\mathcal{Z}^\dagger(\phi_a) - \phi_a\|_\infty, \tag{32}$$

we immediately find our desired result

$$\|(\mathcal{P}_{\omega,\mathcal{N}} \circ \mathcal{N}[\rho])_a - \rho_a\|_1 = \|(\mathcal{Z}[\rho])_a - \rho_a\|_1 \leq d_{code}\sqrt{2\delta}, \tag{33}$$

for any state $\rho \in S(\mathcal{H}_{\text{code}})$.

Note that we could have directly seen from Proposition 2.1 using the triangle inequality that for any state $\rho_a \in S(\mathcal{M}_a)$ we have

$$\|\mathcal{Z}(\rho_a)_a - \rho_a\|_1 \leq \sqrt{2\delta d_{\text{code}}}. \tag{34}$$

However, although this is a tighter bound than (33), it only applies to states in the code space that are of the form given in (15). In the Heisenberg picture, we want our reconstructed operator to work for all states in the code space – not just states of the form (15).

The same problem of extending reconstruction from states $\rho_a \in S(\mathcal{M}_a)$ to all states $\rho \in S(\mathcal{H}_{\text{code}})$ was previously encountered for the twirled Petz map in [10]. It was shown that the approximate equality between bulk and boundary relative entropies [8] implied that any state $\rho \in S(\mathcal{H}_{\text{code}})$ satisfies

$$\mathcal{N}(\rho) \approx \mathcal{N}(\rho_a). \tag{35}$$

Hence (34) implies that, for all states $\rho \in S(\mathcal{H}_{\text{code}})$, we have

$$\|\mathcal{Z}(\rho)_a - \rho\|_1 \leq \sqrt{2\delta d_{\text{code}}} + \varepsilon. \tag{36}$$

where ε is independent of d_{code} and $\varepsilon \to 0$ in the limit $N \to \infty$.

However, (35) is really only true because of the complementary recovery property of AdS/CFT. Not only does region A learn everything about the entanglement wedge a, it also learns nothing about the complementary bulk region $\bar{a}$, which is the entanglement wedge of region $\bar{A}$. In general, operator algebra quantum error correcting codes will not even approximately satisfy (35); as a simple example, consider the case where $\mathcal{N}$ is the identity channel and $\mathcal{M}_a$ is any proper subalgebra of the algebra of observables $\mathcal{B}(\mathcal{H}_{\text{code}})$. It follows that (36) is specific to holographic codes. In contrast, Theorem 1 is a completely general fact about operator algebra quantum error correction. Theorem 1 is therefore a true extension of the range of validity of the Petz map as a general-purpose recovery channel to operator algebra and subsystem codes.

4 Discussion

We have successfully shown that entanglement wedge reconstruction can be achieved using the Petz map, so long as the dimension of the code space, for which the boundary reconstruction is expected to be valid, is not too large. In particular, the Petz map is valid so long as the code space dimension does not grow superpolynomially in the limit of large N. In practice, this is almost always the case for code spaces of interest.

However, it is worth commenting briefly on the major exception to this rule: namely code spaces containing large numbers of black hole microstates; for a detailed discussion, see, for example, [11]. The entropy of such code spaces may in general be $O(1/G_N)$. Hence the dimension of the code space may be exponential in N. However, as yet, the only black hole microstates that we understand to any degree whatsoever are generic, equilibrium microstates. For code spaces made out of such microstates, we would expect the worst-case and average reconstruction errors to approximately agree, even though, in principle, the large code space dimension means that very large differences between them are possible. It is therefore reasonable to hope that the Petz map will still be valid for entanglement wedge reconstruction. Meanwhile, for non-generic, finely-tuned black hole microstates, it is not clear that entanglement wedge reconstruction is possible at all. Hence there is good reason to expect that, whenever entanglement wedge reconstruction is possible, it can be achieved using the Petz map.

We have not made any serious attempt, in this paper, to actually evaluate the Petz map in particular cases. Despite the fact that the Petz map is much simpler to write down, and, in principle, much easier to evaluate, than the twirled Petz map, there remain significant obstacles to doing so. Let us briefly discuss the challenges involved. We wish to explicitly find

$$\phi_A = \frac{1}{d_{code}} \omega_A^{-1/2} [J\phi_a J^\dagger]_A \, \omega_A^{-1/2}. \tag{37}$$

The operator $J\phi_a J^\dagger$ can be found by taking the global HKLL boundary reconstruction ϕ_a^{HKLL} and projecting it into the code space

$$J\phi_a J^\dagger = P_{\text{code}} \phi_a^{HKLL} P_{\text{code}}. \tag{38}$$

The main challenge lies in finding the restriction of this operator to region A. For simplicity, we assume, in accordance with common practice, but not with reality, that the CFT Hilbert space factorises as $\mathcal{H}_{CFT} \cong \mathcal{H}_A \otimes \mathcal{H}_{\bar{A}}$ with $\mathcal{M}_A \cong \mathcal{B}(\mathcal{H}_A)$; the restriction map is then simply a partial trace over $\mathcal{H}_{\bar{A}}$. The difficulty is that HKLL procedure gives an operator ϕ_a that is not localised in time. To take the partial trace over region $\bar{A}$, we need to use the Heisenberg equations of motion to rewrite ϕ_a in terms of operators at time zero.[3] Such operators will in general be very complicated and hard to evaluate. Essentially, the obstruction is simply the usual obstruction to evaluating anything that is not protected by symmetry on the boundary side in AdS/CFT: strongly coupled quantum field theories are hard to deal with – the miracle is the existence of the bulk, which is only weakly coupled.

5 Acknowledgements

We would like to thank Patrick Hayden, Richard Nally, Grant Salton and Michael Walter for valuable discussions. GP is supported by Simons Foundation ("It from Qubit") and AFOSR grant number FA9550-16-1-0082. CFC was supported by the Physics/Applied Physics/SLAC Summer Research Program for undergraduates at Stanford University.

[3] For a more detailed discussion of similar issues, see [10].

References

[1] Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from the anti–de Sitter space/conformal field theory correspondence. *Physical Review Letters*, 96(18):181602, 2006.

[2] Veronika E Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A covariant holographic entanglement entropy proposal. *Journal of High Energy Physics*, 2007(07):062, 2007.

[3] Netta Engelhardt and Aron C Wall. Quantum extremal surfaces: holographic entanglement entropy beyond the classical regime. *Journal of High Energy Physics*, 2015(1):73, 2015.

[4] Xi Dong and Aitor Lewkowycz. Entropy, extremality, euclidean variations, and the equations of motion. *Journal of High Energy Physics*, 2018(1):81, 2018.

[5] Bartłomiej Czech, Joanna L Karczmarek, Fernando Nogueira, and Mark Van Raamsdonk. The gravity dual of a density matrix. *Classical and Quantum Gravity*, 29(15):155009, 2012.

[6] Matthew Headrick, Veronika E Hubeny, Albion Lawrence, and Mukund Rangamani. Causality & holographic entanglement entropy. *Journal of High Energy Physics*, 2014(12):162, 2014.

[7] Aron C Wall. Maximin surfaces, and the strong subadditivity of the covariant holographic entanglement entropy. *Classical and Quantum Gravity*, 31(22):225007, 2014.

[8] Daniel L Jafferis, Aitor Lewkowycz, Juan Maldacena, and S Josephine Suh. Relative entropy equals bulk relative entropy. *Journal of High Energy Physics*, 2016(6):4, 2016.

[9] Xi Dong, Daniel Harlow, and Aron C Wall. Reconstruction of bulk operators within the entanglement wedge in gauge-gravity duality. *Physical Review Letters*, 117(2):021601, 2016.

[10] Jordan Cotler, Patrick Hayden, Geoffrey Penington, Grant Salton, Brian Swingle, and Michael Walter. Entanglement wedge reconstruction via universal recovery channels. *arXiv preprint arXiv:1704.05839*, 2017.

[11] Patrick Hayden and Geoffrey Penington. Learning the alpha-bits of black holes. *arXiv preprint arXiv:1807.06041*, 2018.

[12] Ahmed Almheiri, Xi Dong, and Daniel Harlow. Bulk locality and quantum error correction in AdS/CFT. *Journal of High Energy Physics*, 2015(4):163, 2015.

[13] Alex Hamilton, DN Kabat, Gilad Lifschytz, and David A Lowe. Local bulk operators in ads/cft and the fate of the btz singularity. *AMS/IP Stud. Adv. Math*, 44:85, 2008.

[14] Aitor Lewkowycz and Juan Maldacena. Generalized gravitational entropy. *Journal of High Energy Physics*, 2013(8):90, 2013.

[15] Thomas Faulkner, Aitor Lewkowycz, and Juan Maldacena. Quantum corrections to holographic entanglement entropy. *Journal of High Energy Physics*, 2013(11):74, 2013.

[16] Thomas Faulkner and Aitor Lewkowycz. Bulk locality from modular flow. *Journal of High Energy Physics*, 2017(7):151, 2017.

[17] Howard Barnum and Emanuel Knill. Reversing quantum dynamics with near-optimal quantum and classical fidelity. *Journal of Mathematical Physics*, 43(5):2097–2106, 2002.

[18] Christopher A Fuchs and Jeroen Van De Graaf. Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Transactions on Information Theory*, 45(4):1216–1227, 1999.

[19] Daniel Harlow. The Ryu–Takayanagi formula from quantum error correction. *Communications in Mathematical Physics*, 354(3):865–912, 2017.