

**Simple approach to approximate quantum error correction based on the transpose channel**

Hui Khoon Ng\* and Prabha Mandayam

*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA*

(Received 12 April 2010; published 28 June 2010)

We demonstrate that there exists a universal, near-optimal recovery map—the transpose channel—for approximate quantum error-correcting codes, where optimality is defined using the worst-case fidelity. Using the transpose channel, we provide an alternative interpretation of the standard quantum error correction (QEC) conditions and generalize them to a set of conditions for approximate QEC (AQEC) codes. This forms the basis of a simple algorithm for finding AQEC codes. Our analytical approach is a departure from earlier work relying on exhaustive numerical search for the optimal recovery map, with optimality defined based on entanglement fidelity. For the practically useful case of codes encoding a single qubit of information, our algorithm is particularly easy to implement.

DOI: [10.1103/PhysRevA.81.062342](https://doi.org/10.1103/PhysRevA.81.062342)

PACS number(s): 03.67.Pp

**I. INTRODUCTION**

Quantum error correction (QEC) is one of the cornerstones of quantum information and quantum computing. Since quantum effects are extremely fragile and susceptible to damage by environmental noise, many quantum communication or computational tasks will be impossible without the use of QEC to protect the information from noise. QEC is thus critical for the success of quantum technologies. The idea behind QEC is a simple one—information is stored in a particular part of the system Hilbert space, cleverly chosen based on the noise process, so a recovery operation can be applied to retrieve the information affected by the noise.

Much of the discussion in the past on error correction focuses on *perfect* QEC, where the recovery operation either perfectly corrects the full completely positive trace-preserving (CPTP) noise channel or perfectly corrects the errors conditioned on the fact that fewer than some  $t$  errors occurred. However, an example of a code designed for correcting errors affected by weak amplitude damping noise presented in Ref. [1] suggests that the requirement for perfect recovery may be too stringent for certain tasks. While the smallest known perfect QEC code requires at least five qubits to encode a single qubit, the code in Ref. [1] uses only four qubits to achieve comparable fidelity. This illustrates a key advantage of relaxing the requirement for perfect QEC—one might be able to encode the same amount of information into fewer qubits while retaining a nearly identical level of protection from the noise process. The four-qubit code is also specially designed for the channel in question, a departure from standard QEC codes that seek to perfectly correct up to  $t$  arbitrary errors on the system. This adaptation of the code to the noise channel, an idea emphasized later in Ref. [2], is a crucial factor behind the success of the four-qubit code. Such *approximate* QEC (AQEC) codes reveal the possibility of designing codes that are better tailored to the particular information processing task at hand.

The analysis in Ref. [1] was based on small perturbations of the perfect QEC conditions central to the standard theory

of error correction. Subsequent work on AQEC adopted an alternate approach by recasting it as an optimization problem. One can formulate AQEC as the problem of finding the optimal encoding and recovery maps, given a noise channel and the information we want to encode (qubit or higher-dimensional object), with optimality defined in terms of a chosen measure of fidelity. In this article, optimality is measured in terms of the *worst-case fidelity*, i.e., the fidelity between the input state and the state after noise and recovery, minimized over all possible input states, for given encoding and recovery maps. This is a triple-optimization problem since one needs to optimize over all possible encodings, recovery maps, and input states.

The simplest approach to solving this optimization problem is to hold either the encoding or the recovery map fixed and then perform the optimization over the remaining two variables—the recovery or the encoding map, and the input state. The problem can be further simplified by looking instead at measures based on *entanglement fidelity* [3] and characterizing the performance of the code averaged over some input ensemble. This eliminates the minimization over all input states required for the worst-case fidelity measure. The task of finding the optimal encoding or recovery map is then numerically tractable via convex-optimization methods [4–8], but the resulting recovery is now optimal for an averaged measure of fidelity. Recovery maps which are near-optimal for the average entanglement fidelity have also been constructed analytically, first in Ref. [9] and more recently in Ref. [10].

For many communication or computational tasks, however, one would prefer an assurance that *all* the information stored in the code is well protected. In such cases, the worst-case fidelity is the appropriate measure for determining the optimality of encoding and recovery maps. The resulting double-optimization problem for a given encoding map was examined using semidefinite programming in Ref. [11]. This method, however, requires a relaxation of one of the constraints in the problem, so the recovery map found is typically suboptimal. Furthermore, the numerically computed recovery map is difficult to describe and understand analytically.

In this article, using the worst-case fidelity measure to define optimality, and assuming a fixed encoding, we construct a universal recovery map that is very easy to define analytically. This universal recovery map—the transpose channel [9,12]—gives a worst-case fidelity that cannot be too far from that

---

\*Current address: DSO National Laboratories, and the Centre for Quantum Technologies, National University of Singapore, Singapore.

of the optimal recovery. Using the fact that the transpose channel is the optimal recovery map for perfect QEC codes, we rewrite the error-correction conditions [13–15] for perfect QEC in such a way that the role of the transpose channel is apparent. From this, we derive necessary and sufficient conditions for AQEC founded on the transpose channel, as a natural generalization of the perfect QEC conditions. While AQEC conditions have been derived in the past from an information-theoretic perspective [16–20], our conditions are algebraic and lead to a simple and universal algorithm to find AQEC codes that does not require optimizing over all recovery maps for each encoding map. Furthermore, we demonstrate that the worst-case fidelity for the transpose channel is an easily computable quantity for the most practically useful case of codes encoding a single qubit. Note that AQEC based on the *worst-case* entanglement fidelity was also discussed recently in Ref. [21].

## II. AQEC AS AN OPTIMIZATION PROBLEM

Consider a physical system, with Hilbert space denoted by  $\mathcal{H}$ . In this system, we seek to encode a qudit of information—information carried by a  $d$ -dimensional Hilbert space  $\mathcal{H}_0$ , with  $d \leq \dim(\mathcal{H})$ . In particular, we focus on the case of a *subspace code*, where the qudit is encoded into a  $d$ -dimensional subspace  $\mathcal{C}$ , of  $\mathcal{H}$ . Formally, the information is encoded into  $\mathcal{C}$  via a linear, invertible encoding map  $\mathcal{W}$ . The action of noise on the system is described by a completely positive (CP), trace-preserving (TP) map  $\mathcal{E} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{H})$ .  $\mathcal{E}$  can describe, for example, the Markovian noise acting on the system over some time step, or the effects of a single use of a noisy channel for communication. Complete positivity of  $\mathcal{E}$  entails that its action can be described by a (nonunique) set of Kraus operators  $\{E_i\}_{i=1}^N$  such that  $\mathcal{E}$  acts as  $\mathcal{E}(\rho) = \sum_{i=1}^N E_i \rho E_i^\dagger$ . To denote the noise channel in terms of its Kraus elements, we write  $\mathcal{E} \sim \{E_i\}$ . The fact that  $\mathcal{E}$  is TP is enforced by the condition  $\sum_i E_i^\dagger E_i = \mathbb{I}$ , where  $\mathbb{I}$  is the identity operator for the domain of  $\mathcal{E}$ . After the action of  $\mathcal{E}$ , we perform a CPTP recovery map  $\mathcal{R} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{C})$  to undo the effects of the noise and then decode using  $\mathcal{W}^{-1}$ .

How well the information is protected from the noise can be quantified by the *fidelity* between the input qudit state and the decoded state after noise and recovery. The fidelity between any two states  $\rho$  and  $\sigma$  is given by  $F(\rho, \sigma) \equiv \text{tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}}$ . For a pure state  $\rho \equiv |\psi\rangle\langle\psi|$ , this can be written as  $F(|\psi\rangle, \sigma) \equiv F(|\psi\rangle\langle\psi|, \sigma) = \sqrt{\langle\psi|\sigma|\psi\rangle}$ . Note that,  $0 \leq F(\rho, \sigma) \leq 1$ , with  $F = 0$  if and only if  $\rho$  and  $\sigma$  have orthogonal support, and  $F = 1$  if and only if  $\rho = \sigma$ . The fidelity is thus a measure of how close two states are. Since we will often discuss fidelity for a state before and after the action of a map  $\Phi$ , we use the shorthand  $F(|\psi\rangle, \Phi) \equiv F[|\psi\rangle, \Phi(|\psi\rangle\langle\psi|)]$ .

Based on the fidelity measure, we say that a code  $\mathcal{C}$ , together with  $\mathcal{W}$  and  $\mathcal{R}$ , is effective at protecting the information from the noise  $\mathcal{E}$  if the *worst-case fidelity*  $\min_{\rho \in \mathcal{S}(\mathcal{H}_0)} F[\rho, \mathcal{W}^{-1} \circ \mathcal{R} \circ \mathcal{E} \circ \mathcal{W}]$  is close to 1. Here,  $\mathcal{S}(\mathcal{H}_0)$  denotes the set of all states, pure or mixed, in the codespace. In fact, since the fidelity  $F(\rho, \sigma)$  is jointly concave in its arguments (see, for example, Ref. [22]), it suffices to minimize only over pure states in  $\mathcal{S}(\mathcal{H}_0)$ .

Above, we considered a given encoding map  $\mathcal{W}$  and a given recovery map  $\mathcal{R}$ . In reality, one wants to maximize the error correction capability provided by the system by choosing  $\mathcal{W}$  and  $\mathcal{R}$  such that the worst-case fidelity is as close to 1 as possible. The problem of AQEC can thus be phrased as

$$\max_{\mathcal{W}} \max_{\mathcal{R}} \min_{|\psi\rangle \in \mathcal{H}_0} F(|\psi\rangle, \mathcal{W}^{-1} \circ \mathcal{R} \circ \mathcal{E} \circ \mathcal{W}). \quad (1)$$

If the quantity in Eq. (1) attains the maximum possible value of 1, i.e., there exist  $\mathcal{W}$  and  $\mathcal{R}$  such that the worst-case fidelity is 1, then we have perfect QEC.

The simplest approach to solving this optimization problem is to do an exhaustive search over all possible encodings. This amounts to randomly choosing  $d$ -dimensional subspaces  $\mathcal{C} \subset \mathcal{H}$ . For each  $\mathcal{C}$ , we still need to optimize over  $\mathcal{R}$  to maximize worst-case fidelity. For a given  $\mathcal{C}$ , the optimization problem can be written as

$$\max_{\mathcal{R}} \min_{|\psi\rangle \in \mathcal{C}} F(|\psi\rangle, \mathcal{R} \circ \mathcal{E}), \quad (2)$$

where the worst-case fidelity is computed over all pure states in  $\mathcal{C}$  only.

Before proceeding further, let us define some terminology. We will often make use of the square of the fidelity, which we denote as  $F^2(\cdot, \cdot) \equiv [F(\cdot, \cdot)]^2$ . Whenever unambiguous, we will also refer to  $F^2$  as the fidelity. The recovery  $\mathcal{R}$  with the largest worst-case fidelity for a given  $\mathcal{C}$  is the *optimal recovery* and is denoted by  $\mathcal{R}_{\text{op}}$ . The *fidelity loss*  $\eta_{\mathcal{R}}$ , for a given code  $\mathcal{C}$  and a recovery  $\mathcal{R}$ , is defined as

$$\eta_{\mathcal{R}} \equiv 1 - \min_{|\psi\rangle \in \mathcal{C}} F^2(|\psi\rangle, \mathcal{R} \circ \mathcal{E}). \quad (3)$$

The fidelity loss for  $\mathcal{R}_{\text{op}}$ , denoted  $\eta_{\text{op}}$ , is  $\eta_{\text{op}} = \min_{\mathcal{R}} \eta_{\mathcal{R}}$ . We refer to  $\eta_{\text{op}}$  as the *optimal fidelity loss*. A code  $\mathcal{C}$  is said to be  $\epsilon$ -correctable if it has  $\eta_{\text{op}} \leq \epsilon$  for some  $\epsilon \in [0, 1]$ .  $\epsilon$ -correctable codes with  $\epsilon \ll 1$  are approximately correctable in the sense that code states have fidelity at least  $\sqrt{1 - \epsilon} \approx 1 - \epsilon/2$  after the action of the noise and (optimal) recovery.

## III. TRANSPOSE CHANNEL AS UNIVERSAL, NEAR-OPTIMAL RECOVERY

Here, we describe the transpose channel and demonstrate that it is indeed the standard recovery map for perfect QEC codes characterized by the well-known QEC conditions. We then proceed to show that the transpose channel is nearly optimal even in the case of AQEC codes.

### A. Transpose channel

Consider a  $d$ -dimensional code  $\mathcal{C}$  and a CPTP noise channel  $\mathcal{E} \sim \{E_i\}_{i=1}^N$ . Let  $P$  be the projector onto  $\mathcal{C}$  and  $P_{\mathcal{E}}$  be the projector onto  $\mathcal{P}_{\mathcal{E}} \equiv$  the support of  $\mathcal{E}(\mathcal{C})$ . The transpose channel is the CPTP map  $\mathcal{R}_P : \mathcal{B}(\mathcal{P}_{\mathcal{E}}) \rightarrow \mathcal{B}(\mathcal{C})$  such that

$$\mathcal{R}_P(\cdot) \equiv \sum_{i=1}^N P E_i^\dagger \mathcal{E}(P)^{-1/2}(\cdot) \mathcal{E}(P)^{-1/2} E_i P, \quad (4)$$

where the inverse of  $\mathcal{E}(P)$  is taken on its support. The transpose channel can be understood as being composed of three CP maps:  $\mathcal{R}_P = \mathcal{P} \circ \mathcal{E}^\dagger \circ \mathcal{N}$ , where  $\mathcal{P}$  is the projection  $P(\cdot)P$  onto  $\mathcal{C}$ ,  $\mathcal{E}^\dagger$  is the adjoint of  $\mathcal{E}$ , and  $\mathcal{N}$  is the normalization map

$\mathcal{N}(\cdot) = \mathcal{E}(P)^{-1/2}(\cdot)\mathcal{E}(P)^{-1/2}$ . In this form,  $\mathcal{R}_P$  is manifestly independent of the choice of Kraus representation for  $\mathcal{E}$ .

$\mathcal{R}_P$  is a special case of a recovery map introduced in Ref. [9] for reversing the effects of a quantum channel on a given initial state.  $\mathcal{R}_P$  defined here is exactly the case for the initial state  $P/d$ , where  $d$  is the dimension of  $\mathcal{C}$ . In Ref. [23],  $\mathcal{R}_P$  was shown to be useful for correcting information carried by codes preserved according to an operationally motivated notion. The term *transpose channel* owes its origin to Ref. [12], where this channel was first defined in an information-theoretic context. It was shown [24] that the transpose channel has the property of being the unique noise channel that saturates Uhlmann's theorem, i.e., the monotonicity of relative entropy—a fact that was later used to characterize states that saturate the strong subadditivity of quantum entropy [25].

While our focus is on AQEC, understanding the relevance of  $\mathcal{R}_P$  to perfect QEC provides the intuition behind the AQEC conditions presented later. An important characterization of perfect QEC codes is the set of QEC conditions [13–15], which we briefly review here (see, for example, Ref. [22]):

*Theorem 1 (Perfect QEC conditions).* A CPTP recovery  $\mathcal{R}$  that perfectly corrects a CP map  $\mathcal{E}$  on a subspace code  $\mathcal{C}$  exists if and only if

$$\forall i, j, \quad P E_i^\dagger E_j P = \alpha_{ij} P, \quad (5)$$

for some complex matrix  $\alpha$ .

It is useful to rewrite Eq. (5) in a “diagonal” form.  $\alpha$  is clearly Hermitian and can be diagonalized with a unitary  $u$  such that  $\alpha = u D u^\dagger$ , where  $D$  is the diagonal matrix of eigenvalues. The set  $\{F_k \equiv \sum_i u_{ik} E_i\}$  constitutes a different Kraus representation for  $\mathcal{E}$ . With this choice of Kraus representation, the perfect QEC conditions take the diagonal form

$$\forall k, l, \quad P F_k^\dagger F_l P = \delta_{kl} d_{kk} P, \quad (6)$$

where  $d_{kk}$  are the diagonal entries of  $D$ .

The recovery map  $\mathcal{R}$  when Eq. (5) is satisfied—which we denote as  $\mathcal{R}_{\text{perf}}$ —is constructed as follows [22]: using the polar decomposition  $F_k P = \sqrt{d_{kk}} U_k P$ ,  $\mathcal{R}_{\text{perf}} : \mathcal{B}(\mathcal{P}_{\mathcal{E}}) \rightarrow \mathcal{B}(\mathcal{C})$  is given by  $\mathcal{R}_{\text{perf}} \sim \{P U_k^\dagger\}$ . One can check that  $\mathcal{R}_{\text{perf}}$  is TP on its domain  $\mathcal{B}(\mathcal{P}_{\mathcal{E}})$  and that for any  $\rho \in \mathcal{B}(\mathcal{C})$ ,  $(\mathcal{R}_{\text{perf}} \circ \mathcal{E})(\rho) = (\sum_k d_{kk}) \rho$ . From the QEC conditions [Eq. (6)], we see that  $\sum_k d_{kk} = \text{tr}[\mathcal{E}(\rho)]$  is independent of  $\rho$  and is exactly equal to 1 if and only if  $\mathcal{E}$  is TP on  $\mathcal{C}$ .  $\mathcal{R}_{\text{perf}}$  thus recovers the original code state, up to any reduction in trace due to the possible non-TP nature of  $\mathcal{E}$ .

A natural question to ask here is how the transpose channel  $\mathcal{R}_P$  relates to the recovery  $\mathcal{R}_{\text{perf}}$  for a given  $\mathcal{E}$  and  $\mathcal{C}$  that satisfy the QEC conditions. Here, we show that they are exactly the same map, as previously noted in Ref. [9]:

*Lemma 2.*  $\mathcal{R}_P = \mathcal{R}_{\text{perf}}$ .

*Proof.* Observe that  $\mathcal{E}(P) = \sum_k (F_k P)(P F_k^\dagger) = \sum_k d_{kk} P_k$ , where  $P_k \equiv U_k P U_k^\dagger$ . Equation (6) gives  $P U_k^\dagger U_l P = \delta_{kl} P$ , so  $P_k$ 's are orthogonal projectors with  $P_k P_l = \delta_{kl} P_k$ . Hence,  $\mathcal{E}(P)^{-1/2} = \sum_k P_k / \sqrt{d_{kk}}$ . The Kraus operators  $\{P F_k^\dagger \mathcal{E}(P)^{-1/2}\}$  of  $\mathcal{R}_P$  can hence be written as

$$P F_k^\dagger \mathcal{E}(P)^{-1/2} = \sum_l \sqrt{d_{kk}/d_{ll}} P U_k^\dagger U_l P U_l^\dagger = P U_k^\dagger, \quad (7)$$

which are exactly the Kraus operators of  $\mathcal{R}_{\text{perf}}$ . ■

Perfect QEC is often discussed for a noise channel that is CP but not necessarily TP. In fact, Theorem 1 and Lemma 2 remain true even for a non-TP  $\mathcal{E}$ . The non-TP scenario is particularly relevant when we deal with a system of  $n$  quantum registers, where each register is independently affected by some CPTP noise  $\mathcal{E}_1$ . One often looks for codes that perfectly correct the noise up to some maximum number  $t$  of quantum registers with errors. Then, instead of having  $\mathcal{E} \equiv \mathcal{E}_1^{\otimes n}$ , the relevant noise channel for perfect QEC describes noise where at most  $t$  registers have errors. Such an  $\mathcal{E}$  is not TP, since we have discarded the part of  $\mathcal{E}_1^{\otimes n}$  that corresponds to having errors in more than  $t$  registers.

Actually, a perfect QEC code for such a non-TP noise channel can be viewed as an AQEC code for the original  $n$ -register noise channel  $\mathcal{E}_1^{\otimes n}$ , which is TP. In our AQEC discussion, the code we look for is approximately correctable on the channel anyway, so  $\mathcal{E}$  is always assumed to be TP, which is often the physically relevant scenario. The TP requirement is also important for fidelity to be a good measure of the efficacy of the recovery operation. Note that the analysis in the remainder of the article applies to a special type of non-TP maps— $\mathcal{E} \sim \{E_i\}$  such that  $\sum_i P E_i^\dagger E_i P = a P$  for  $0 \leq a \leq 1$ , giving an additional proportionality factor  $a$  in our expressions.

## B. Near-optimality of the transpose channel

In general,  $\mathcal{R}_P$  need not be the optimal recovery map  $\mathcal{R}_{\text{op}}$  for a given  $\mathcal{C}$  and  $\mathcal{E}$ . However, in the following theorem and the subsequent corollary, which form the core results of our article, we show that it does not do much worse than  $\mathcal{R}_{\text{op}}$ .

*Theorem 3.* Consider a  $d$ -dimensional code  $\mathcal{C}$  with optimal fidelity loss  $\eta_{\text{op}}$  under a CPTP noise channel  $\mathcal{E}$ . For any  $|\psi\rangle \in \mathcal{C}$ ,

$$F^2(|\psi\rangle, \mathcal{R}_{\text{op}} \circ \mathcal{E}) \leq \sqrt{1 + (d-1)\eta_{\text{op}}} F(|\psi\rangle, \mathcal{R}_P \circ \mathcal{E}). \quad (8)$$

*Proof.* Let  $\{R_j\}$  be a set of Kraus operators of  $\mathcal{R}_{\text{op}} : \mathcal{B}(\mathcal{P}_{\mathcal{E}}) \rightarrow \mathcal{B}(\mathcal{C})$ . For any  $|\psi\rangle \in \mathcal{C}$ , following [9], we have

$$\begin{aligned} F^2(|\psi\rangle, \mathcal{R}_{\text{op}} \circ \mathcal{E}) &\leq \sqrt{\left[ \sum_i |\langle E_i^\dagger \mathcal{E}(P)^{-1/2} E_i \rangle|^2 \right] \left[ \sum_j |\langle R_j \mathcal{E}(P)^{1/2} R_j^\dagger \rangle|^2 \right]}, \\ &\leq \sqrt{\left[ \sum_i |\langle E_i^\dagger \mathcal{E}(P)^{-1/2} E_i \rangle|^2 \right] \left[ \sum_j |\langle R_j \mathcal{E}(P)^{1/2} R_j^\dagger \rangle|^2 \right]}, \end{aligned} \quad (9)$$

where  $\langle \cdot \rangle$  denotes expectation value with respect to  $|\psi\rangle$ . Since  $\mathcal{R}_{\text{op}}$  is TP, we have that  $\sum_j |\langle R_j \mathcal{E}(P)^{1/2} R_j^\dagger \rangle|^2 \leq \langle \sum_j R_j \mathcal{E}(P) R_j^\dagger \rangle = \langle (\mathcal{R}_{\text{op}} \circ \mathcal{E})(P) \rangle$ .

Now, choose a basis  $\{|\psi_i\rangle\}_{i=1}^d$  for  $\mathcal{C}$  with  $|\psi_1\rangle \equiv |\psi\rangle$ . Let  $\rho_i \equiv (\mathcal{R}_{\text{op}} \circ \mathcal{E})(|\psi_i\rangle\langle\psi_i|) = \sum_{kl} \alpha_{kl}^{(i)} |\psi_k\rangle\langle\psi_l|$ , for coefficients satisfying  $\sum_k \alpha_{kk}^{(i)} = 1$  and  $\alpha_{kk}^{(i)} \geq 0 \forall k$ . From the definition of  $\eta_{\text{op}}$ ,  $\alpha_{ii}^{(i)} = \langle \psi_i | \rho_i | \psi_i \rangle \geq 1 - \eta_{\text{op}}$ . This implies  $\sum_{k \neq i} \alpha_{kk}^{(i)} \leq \eta_{\text{op}}$ , which in turn gives  $\alpha_{kk}^{(i)} \leq \eta_{\text{op}} \forall k \neq i$ . Since  $|\psi\rangle = |\psi_1\rangle$  by construction, we get  $\sum_j |\langle R_j \mathcal{E}(P)^{1/2} R_j^\dagger \rangle|^2 \leq \langle (\mathcal{R}_{\text{op}} \circ \mathcal{E})(P) \rangle \leq 1 + (d-1)\eta_{\text{op}}$ . Putting this into Eq. (9), and noting that  $(\sum_i |\langle E_i^\dagger \mathcal{E}(P)^{-1/2} E_i \rangle|^2)^{1/2} \leq F(|\psi\rangle, \mathcal{R}_P \circ \mathcal{E})$ , gives Eq. (8). ■

Let  $\eta_P$  be the fidelity loss for code  $\mathcal{C}$  with  $\mathcal{R}_P$  as the recovery. Then, Theorem 3 implies

*Corollary 4.*  $\eta_P$  satisfies  $\eta_{\text{op}} \leq \eta_P \leq \eta_{\text{op}} f(\eta_{\text{op}}; d)$ , where  $f(\eta; d)$  is the function

$$f(\eta; d) \equiv \frac{(d+1) - \eta}{1 + (d-1)\eta} = (d+1) + O(\eta). \quad (10)$$

*Proof.*  $\eta_P \geq \eta_{\text{op}}$  is true by definition of  $\eta_{\text{op}}$ . For any  $|\psi\rangle \in \mathcal{C}$ , let  $F^2(|\psi\rangle, \mathcal{R}_P \circ \mathcal{E}) \equiv 1 - \eta_{P, \psi}$ . Then, by definition, the fidelity loss is  $\eta_P = \max_{\psi} (\eta_{P, \psi})$ . From Theorem 3,  $1 - \eta_{\text{op}} \leq F^2(|\psi\rangle, \mathcal{R}_{\text{op}} \circ \mathcal{E}) \leq \sqrt{[1 + (d-1)\eta_{\text{op}}](1 - \eta_{P, \psi})}$ . Rearranging gives  $\eta_{P, \psi} \leq \eta_{\text{op}} f(\eta_{\text{op}}; d)$ . Since this holds for all  $\eta_{P, \psi}$ , it also holds for  $\eta_P$ . ■

The inequality  $\eta_P \leq \eta_{\text{op}} f(\eta_{\text{op}}; d)$  makes precise our statement that  $\mathcal{R}_P$  is near-optimal, with the additional factor of  $(d+1)$ . For the most practically relevant case of a code encoding a single qubit, this is only a factor of 3. Note that, for  $\eta_{\text{op}} = 0$ , the inequality in Corollary 4 collapses to  $\eta_P = \eta_{\text{op}}$ , as expected from Lemma 2. Corollary 4 provides necessary and sufficient conditions for  $\mathcal{C}$  to be approximately correctable— $\mathcal{C}$  is approximately correctable if and only if  $\eta_P$  is small.

We do not know if the upper bound on  $\eta_P$  in Corollary 4 is tight. However, the appearance of the dimension  $d$  of the code in the bound is unavoidable, as can be seen from the following example:

*Example 1.* Consider a noise channel  $\mathcal{E}$ , whose action on a code  $\mathcal{C}$  is given by the set of Kraus operators  $\{E_i P\} = \{\sqrt{1-p}P, \sqrt{p}|0\rangle\langle 0|, \sqrt{p}|0\rangle\langle 1|, \dots, \sqrt{p}|0\rangle\langle d-1|\}$ , for  $0 \leq p \ll 1$ .  $\mathcal{E}$  acts like the identity channel on  $\mathcal{C}$ , except for a small damaging component that maps a small part of every code state onto  $|0\rangle$ . For  $d \geq 3$ , one can show that the worst-case fidelity, when using  $\mathcal{R}_P$  as the recovery, occurs for state  $|0\rangle$ . The corresponding fidelity loss is  $\eta_P = (d-1)p/[1 + (d-1)p]$ . Since  $\mathcal{E}$  is nearly the identity channel, we might instead do nothing (identity channel as the recovery), for which the fidelity loss is  $\eta_0 = p$ .  $\eta_0$  is always smaller than  $\eta_P$  for small  $p$ . Since  $\eta_{\text{op}} \leq \eta_0$ , we see that  $\eta_P/\eta_{\text{op}} \geq \eta_P/\eta_0 = (d-1)/[1 + (d-1)p]$ , which grows as  $d$  increases, for fixed  $p$ . Hence, for this noise channel and code, the separation between  $\eta_P$  and  $\eta_{\text{op}}$  grows as  $d$  increases.

That the dimension of the code space appears here is perhaps not surprising. In the next section, we will see that this approach to AQEC using the transpose channel can be viewed as a perturbation from the perfect QEC case. The factor of  $d$  appearing in our bounds can hence be understood as quantifying the number of degrees of freedom in which the approximate case can deviate from the perfect case.

Note, however, that as  $d$  gets large,  $f(\eta; d)$  approaches  $1/\eta$ . In this case, the inequality in Corollary 4 simply becomes the trivial statement  $\eta_{\text{op}} \leq \eta_P \leq 1$ . While we will often only be interested in codes with small values of  $d$ , this demonstrates the weakness in the bounds derived here for large  $d$  values.

#### IV. THE TRANSPOSE CHANNEL AND QEC CONDITIONS

One of the key tools in perfect QEC are the perfect QEC conditions (Theorem 1). Conditions characterizing AQEC codes would likewise be useful. A natural approach is to perturb the perfect QEC conditions to allow for small deviations. For example, the four-qubit code for the amplitude

damping channel in Ref. [1] was shown to obey a set of perturbed QEC conditions. More recently, [26] examined small perturbations of the perfect QEC conditions for general CPTP channels. However, the analysis in Ref. [26] is often complicated, and one wonders if there is a simpler approach using the transpose channel. In this section, we discuss such a set of AQEC conditions built on Corollary 4. We begin by first writing down an alternate but equivalent set of perfect QEC conditions which highlights the role of the transpose channel:

*Theorem 5 (Alternate form of perfect QEC conditions).* A code  $\mathcal{C}$  satisfies the perfect QEC conditions (Theorem 1) if and only if it satisfies

$$\forall i, j, \quad P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = \beta_{ij} P, \quad (11)$$

where  $\beta \equiv \sqrt{\alpha}$ , for  $\alpha$  is defined in Eq. (5).

*Proof.* For a code  $\mathcal{C}$  that satisfies the perfect QEC conditions (Theorem 1), using Eq. (7) and  $P U_k^\dagger U_l P = \delta_{kl} P$ , we have

$$P F_k^\dagger \mathcal{E}(P)^{-1/2} F_l P = \delta_{kl} \sqrt{d_{kk}} P. \quad (12)$$

This diagonal form can be rotated to any other Kraus representation using a unitary  $u$  so that  $F_k = \sum_i u_{ik} E_i$  and  $\alpha = u D u^\dagger$ . Defining  $\beta \equiv \sqrt{\alpha}$  gives Eq. (11), thus showing that if a code  $\mathcal{C}$  satisfies the perfect QEC conditions, it also satisfies Eq. (11).

Conversely, suppose we start from the diagonal form of Eq. (11) as in Eq. (12), which can be accomplished by choosing  $u$  so that  $\beta$  is diagonal with entries  $\sqrt{d_{kk}}$ . Then taking the square root of Eq. (12) gives  $\mathcal{E}(P)^{-1/4} F_k P = (d_{kk})^{1/4} V_k P$  for some unitary  $V_k$ , so  $F_k P = (d_{kk})^{1/4} \mathcal{E}(P)^{1/4} V_k P$ . Putting this into Eq. (12) gives  $P V_k^\dagger V_l P = \delta_{kl} P$ . Furthermore,  $\mathcal{E}(P)^{1/2} = [\sum_k (F_k P)(P F_k^\dagger)]^{1/2} = \sum_k \sqrt{d_{kk}} V_k P V_k^\dagger$ . Direct computation then gives  $P F_k^\dagger F_l P = \delta_{kl} d_{kk} P$ , which is exactly Eq. (6). Applying an appropriate  $u$  to rotate to the desired Kraus representation gives Eq. (5). ■

Observe that the left-hand side of Eq. (11) is a Kraus operator of  $\mathcal{R}_P \circ \mathcal{E}$ . Thus, the QEC conditions in Theorem 5, and equivalently the original conditions stated in Theorem 1, simply express the fact that  $\mathcal{C}$  is perfectly correctable if and only if  $\mathcal{R}_P \circ \mathcal{E} \propto \mathcal{P}$ . The proportionality factor is  $\sum_{ij} \beta_{ij}^2 = \sum_{ij} \alpha_{ij} = \sum_k d_{kk}$ .

We can now obtain a set of conditions for AQEC by perturbing this alternate form of the QEC conditions.

*Theorem 6 (AQEC conditions).* Consider a CPTP channel  $\mathcal{E} \sim \{E_i\}$ , and a  $d$ -dimensional code  $\mathcal{C}$  with projector  $P$ . Let  $\Delta_{ij} \in \mathcal{B}(\mathcal{C})$  be traceless operators such that

$$P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = \beta_{ij} P + \Delta_{ij}, \quad (13)$$

where  $\beta_{ij} \in \mathbb{C}$ . Then, for  $\epsilon \in [0, 1]$ ,  $\exists \eta \in [0, 1]$  given by

$$\eta = \max_{|\psi\rangle \in \mathcal{C}} \sum_{ij} [|\langle \psi | \Delta_{ij}^\dagger \Delta_{ij} | \psi \rangle - |\langle \psi | \Delta_{ij} | \psi \rangle|^2]. \quad (14)$$

such that

- (i)  $\mathcal{C}$  is  $\epsilon$ -correctable if  $\eta \leq \epsilon$ ;
- (ii)  $\mathcal{C}$  is  $\epsilon$ -correctable only if  $\eta \leq \epsilon f(\epsilon; d)$ , where  $f$  is the function defined in Eq. (10).

*Proof.* The left-hand side of Eq. (13) is a Kraus operator of  $\mathcal{R}_P \circ \mathcal{E}$ . This, along with the TP condition for  $\mathcal{R}_P \circ \mathcal{E}$ , gives

the expression on the right-hand side of Eq. (14) for  $\eta_P$ . Setting  $\eta = \eta_P$ , conditions (i) and (ii) follow from Corollary 4. ■

Equation (14) elucidates how the fidelity loss arises from the presence of the  $\Delta_{ij}$  operators. If  $\Delta_{ij} = 0 \forall i, j$ , we have perfect QEC.

The AQEC conditions, like the perfect QEC conditions, provide a way to check if a code is approximately correctable, without requiring knowledge of the optimal recovery. More precisely, given a maximum tolerable fidelity loss  $\epsilon$  for some information processing task at hand, one can check if a code  $\mathcal{C}$  is  $\epsilon$ -correctable as follows. The AQEC conditions instruct us to compute  $\eta_P$ , which can be done once we know  $\mathcal{C}$  and the noise channel  $\mathcal{E}$ . If  $\eta_P \leq \epsilon$ , then  $\mathcal{C}$  is a good code. If, however,  $\eta_P$  violates the inequality in condition (ii), we know that  $\mathcal{C}$  is not good enough for our purposes. Of course, there is a gap—for  $\eta_P$  taking values  $\epsilon \leq \eta_P \leq \epsilon f(\epsilon; d)$ , we cannot use the conditions to determine if  $\mathcal{C}$  is within our tolerable fidelity loss, but this gap is small for small  $d$ . We do not know if the gap can be shrunk by replacing  $\eta_P$  with the fidelity loss for a different recovery map than the transpose channel, but we believe it is unlikely to vanish completely.

For a general  $\mathcal{C}$ , the fidelity loss  $\eta_P$  may be difficult to compute as it requires a maximization over all states in the code space. However, there is a quick way to check for sufficiency by relaxing condition (i) of Theorem 6:

*Corollary 7.*  $\mathcal{C}$  is  $\epsilon$ -correctable for some  $\epsilon \in [0, 1]$  if  $\|\Delta_{\text{sum}}\| \leq \epsilon$ , where  $\Delta_{\text{sum}} \equiv \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij}$ , and  $\|\cdot\|$  denotes the operator norm.

*Proof.* Observe that the right-hand side of Eq. (14) satisfies  $\sum_{ij} [\langle \psi | \Delta_{ij}^\dagger \Delta_{ij} | \psi \rangle - |\langle \psi | \Delta_{ij}^\dagger | \psi \rangle|^2] \leq \langle \psi | \Delta_{\text{sum}} | \psi \rangle$ . Maximizing this expression over all  $|\psi\rangle \in \mathcal{C}$  gives  $\|\Delta_{\text{sum}}\|$ . Hence,  $\eta_P \leq \|\Delta_{\text{sum}}\|$ , and the sufficiency condition (i) in Theorem 6 is satisfied if  $\|\Delta_{\text{sum}}\| \leq \epsilon$ . ■

Since  $\Delta_{\text{sum}} \geq 0$ , its operator norm is given by its maximum eigenvalue, which is easily computable. In fact, for codes encoding a single qubit, it is easy to show (using the Pauli basis, for example) that  $\|\Delta_{\text{sum}}\| = 1 - \sum_{ij} |\beta_{ij}|^2$ . Note that  $\beta_{ij}$  for any code  $\mathcal{C}$  and noise channel  $\mathcal{E}$  is simply given by  $\beta_{ij} = (1/d) \text{tr}(P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P)$ .

## V. COMPUTING $\eta_P$ FOR QUBIT CODES

Computing  $\eta_P$  for a general code requires an exhaustive optimization over all states in the code. However, for the practically relevant case of codes encoding a single qubit, i.e.,  $\mathcal{C}$  with dimension  $d = 2$ ,  $\eta_P$  turns out to require only simple eigenanalysis to compute.

For a qubit code,  $(\mathcal{R}_P \circ \mathcal{E}) : \mathcal{B}(\mathcal{C}) \rightarrow \mathcal{B}(\mathcal{C})$  is a qubit map. Observe that  $\mathcal{R}_P \circ \mathcal{E}$  is not only CPTP but also unital [i.e.,  $(\mathcal{R}_P \circ \mathcal{E})(P) = P$ ]. Here, we show that the worst-case fidelity for a unital, CPTP qubit map is easy to compute. While our context requires only a unital, CPTP qubit map, we begin with a general CP map  $\Phi \sim \{K_i\}$  on a  $d$ -dimensional Hilbert subspace  $\mathcal{C}$ , so as to highlight why the qubit case is particularly simple.

We begin by choosing a Hermitian basis  $\{O_0, O_1, \dots, O_{d^2-1}\}$  for  $\mathcal{B}(\mathcal{C})$ , where  $O_0 \equiv \mathbb{I}$ ,  $O_\alpha^\dagger = O_\alpha \forall \alpha$ ,  $\text{tr}\{O_\alpha^\dagger O_\beta\} = \delta_{\alpha\beta} d \forall \alpha, \beta$ . The operators  $\{O_\alpha, \alpha = 1, \dots, d^2 - 1\}$  are clearly traceless. Such a basis exists for any

$d$ —for example, one can use the set of standard generators of the  $SU(d)$  group, augmented with the identity operator. The action of  $\Phi$  can be represented as a matrix  $\mathcal{M}$ , acting on vectors [operators in  $\mathcal{B}(\mathcal{C})$ ] in the Hilbert-Schmidt space, with matrix elements

$$\mathcal{M}_{\alpha\beta} \equiv \frac{1}{d} \text{tr}\{O_\alpha \Phi(O_\beta)\}. \quad (15)$$

Since  $\Phi$  is CP and  $O_\alpha$ 's are Hermitian,  $\mathcal{M}_{\alpha\beta}^* = \mathcal{M}_{\alpha\beta}$ , so  $\mathcal{M}$  is a real matrix.

Now, the density operator corresponding to any pure state  $|\psi\rangle$  in  $\mathcal{C}$  can be expanded in terms of the Hermitian basis as

$$|\psi\rangle\langle\psi| = \frac{1}{d} (\mathbb{I} + \mathbf{s} \cdot \mathbf{O}) = \frac{1}{d} \vec{s} \cdot \vec{O}, \quad (16)$$

where  $\mathbf{s}$  is a real  $(d^2 - 1)$ -element vector,  $\vec{s} \equiv (1, \mathbf{s})$ ,  $\mathbf{O} \equiv (O_1, O_2, \dots, O_{d^2-1})$ , and  $\vec{O} \equiv (\mathbb{I}, \mathbf{O})$ .  $\mathbf{s}$  is not an arbitrary vector, but in general has to obey some constraints in order for it to correspond to a pure state.

Using Eqs. (15) and (16), the fidelity for a state  $|\psi\rangle \in \mathcal{C}$  under the map  $\Phi$  can be written as  $F^2(|\psi\rangle, \Phi) = \frac{1}{d} s^T \mathcal{M} s$ , where  $s$  is  $\vec{s}$  written as a column vector, and the superscript  $T$  denotes the transpose. We can rewrite the expression for the fidelity using the symmetrized version of  $\mathcal{M}$ :  $\mathcal{M}_{\text{sym}} \equiv \frac{1}{2} (\mathcal{M} + \mathcal{M}^T)$ . Observe that  $s^T \mathcal{M}_{\text{sym}} s = s^T \mathcal{M} s$ . Therefore,

$$F^2(|\psi\rangle, \Phi) = s^T \mathcal{M}_{\text{sym}} s. \quad (17)$$

Finding the worst-case fidelity is hence equivalent to the following minimization problem for a real, symmetric matrix  $\mathcal{M}_{\text{sym}}$ :

$$\text{minimize: } s^T \mathcal{M}_{\text{sym}} s, \quad (18a)$$

$$\text{constraint: } s \text{ corresponds to a pure state.} \quad (18b)$$

For  $d > 2$ , the constraint Eq. (18b) is difficult to write down. Even if we relax the constraint to include mixed states, it is not known in general what  $s$  corresponding to a (positive, trace-1) density operator looks like. This constrained minimization problem is hence not simple for a general  $d$ .

For qubits ( $d = 2$ ), however, the constraint equation is simple to write down. We choose the operator basis to be the Pauli basis. Given an orthonormal basis  $\{|v_1\rangle, |v_2\rangle\}$  for the qubit code space, the Pauli basis  $\{\sigma_0 \equiv \mathbb{I}_2, \sigma_x, \sigma_y, \sigma_z\}$  can be constructed as

$$\begin{aligned} \sigma_0 &= |v_1\rangle\langle v_1| + |v_2\rangle\langle v_2| \equiv \mathbb{I}_2, \\ \sigma_x &= |v_1\rangle\langle v_2| + |v_2\rangle\langle v_1|, \end{aligned} \quad (19)$$

$$\sigma_y = -i(|v_1\rangle\langle v_2| - |v_2\rangle\langle v_1|),$$

$$\text{and } \sigma_z = |v_1\rangle\langle v_1| - |v_2\rangle\langle v_2|.$$

Equation (16) then corresponds to the Bloch sphere representation of a pure state, with the Bloch vector  $\mathbf{s} \equiv (s_x, s_y, s_z)$  satisfying  $\|\mathbf{s}\| = (s_x^2 + s_y^2 + s_z^2)^{1/2} = 1$ . The constraint Eq. (18b) becomes

$$\text{constraint: } s = (1, \mathbf{s}), \text{ with } \|\mathbf{s}\| = 1. \quad (18b')$$

The constrained minimization problem can then be solved using the Lagrange multiplier method.

For the case of a CPTP qubit map that is also unital, the minimization problem can be further simplified. For any CPTP, unital  $\Phi$  (arbitrary  $d$ ),  $\mathcal{M}$  takes the form

$$\mathcal{M} = \begin{pmatrix} 1 & | & 0 & \dots & 0 \\ \hline 0 & & & & \\ \vdots & & \mathcal{T} & & \\ 0 & & & & \end{pmatrix}. \quad (20)$$

The first row comes from the fact that  $\Phi$  is TP, and the first column from the fact that  $\Phi$  is unital.  $\mathcal{T}$  is a  $(d-1) \times (d-1)$  real matrix. Defining  $\mathcal{T}_{\text{sym}} \equiv \frac{1}{2}(\mathcal{T} + \mathcal{T}^T)$ , Eq. (17) can be written as  $F^2(|\psi\rangle, \Phi) = \frac{1}{d}(1 + \mathbf{s}^T \mathcal{T}_{\text{sym}} \mathbf{s})$ . This means that we can equivalently minimize  $\mathbf{s}^T \mathcal{T}_{\text{sym}} \mathbf{s}$  instead of the original  $s^T \mathcal{M}_{\text{sym}} s$  in Eq. (18a). For a qubit CPTP, unital  $\Phi$  then, the constrained minimization problem becomes

$$\text{minimize: } \mathbf{s}^T \mathcal{T}_{\text{sym}} \mathbf{s}, \quad (21a)$$

$$\text{constraint: } \|\mathbf{s}\| = \sqrt{s_x^2 + s_y^2 + s_z^2} = 1. \quad (21b)$$

This simply tells us to minimize the expectation value of  $\mathcal{T}_{\text{sym}}$  with respect to all real *unit* vectors  $\mathbf{s}$ .

Since  $\mathcal{T}_{\text{sym}}$  is real and symmetric, it can be diagonalized with an orthogonal matrix  $Q$  so that  $\mathcal{T}_{\text{sym}} = Q^T \mathcal{T}_D Q$ , where  $\mathcal{T}_D$  is a real, diagonal matrix of eigenvalues of  $\mathcal{T}_{\text{sym}}$ . Then  $\mathbf{s}^T \mathcal{T}_{\text{sym}} \mathbf{s} = (Q\mathbf{s})^T \mathcal{T}_D (Q\mathbf{s})$ .  $Q$ , being orthogonal, preserves the length of the vector it acts on. The minimization problem Eq. (21) thus corresponds to minimizing the expectation value of  $\mathcal{T}_D$  over all real unit vectors. As  $\mathcal{T}_D$  is real and diagonal, this minimum expectation value is exactly the smallest eigenvalue of  $\mathcal{T}_D$  (and hence of  $\mathcal{T}_{\text{sym}}$ ), attained by the corresponding eigenvector normalized to unit length. Therefore, we see that the fidelity loss for a CPTP, unital qubit map  $\Phi$  is given by

$$\eta_\Phi = 1 - \min_{|\psi\rangle \in \mathcal{C}} F^2(|\psi\rangle, \Phi) = \frac{1}{2}(1 - t_{\min}), \quad (22)$$

where  $t_{\min}$  is the smallest eigenvalue of  $\mathcal{T}_{\text{sym}}$  corresponding to the map  $\Phi$ . Setting  $\Phi = \mathcal{R}_P \circ \mathcal{E} \circ \mathcal{P}$  gives  $\eta_P$ . Note that, for  $\Phi$  with a Hermitian-closed set<sup>1</sup> of Kraus operators, as is the case for  $\Phi \equiv \mathcal{R}_P \circ \mathcal{E} \circ \mathcal{P} \sim \{PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j^\dagger P\}$ ,  $\mathcal{T}$  is symmetric so that  $\mathcal{T}_{\text{sym}} = \mathcal{T}$ .

## VI. EXAMPLE: AMPLITUDE DAMPING CHANNEL

As an example to illustrate our discussion so far, let us look at the noise channel considered in Ref. [1]—the amplitude damping channel. The single-qubit amplitude damping channel  $\mathcal{E}_{\text{AD}}$  is the CPTP channel with Kraus operators

$$E_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix} \quad \text{and} \quad E_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}, \quad (23)$$

written in some qubit basis  $\{|0\rangle, |1\rangle\}$ .  $\mathcal{E}_{\text{AD}}$  can be thought of as describing energy dissipation for a system where  $|0\rangle$  is the ground state and  $|1\rangle$  is some excited state.  $\gamma$  is then the probability of a transition from the excited state to the ground

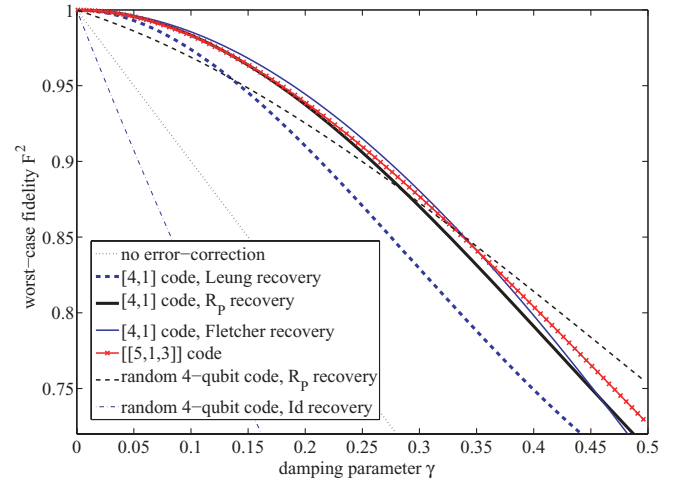


FIG. 1. (Color online) Codes for the amplitude damping channel, for  $0 \leq \gamma \leq 0.5$ .

state. In the absence of any encoding or recovery, the worst-case fidelity for a single qubit undergoing  $\mathcal{E}_{\text{AD}}$  falls off as  $1 - \gamma$ , as  $\gamma$  increases (see Fig. 1, line labeled “no error correction”).

A code that uses four physical qubits to protect a single qubit of information against amplitude damping noise was constructed by Leung *et al.* [1]. Assuming that the noise acts independently on the qubits, the four-qubit noise channel is just four copies of  $\mathcal{E}_{\text{AD}}$ , i.e.,  $\mathcal{E}_{\text{AD}}^{\otimes 4}$ . The four-qubit subspace code constructed in Ref. [1] is the span of the following two states:

$$|0_L\rangle \equiv \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle), \quad (24)$$

$$\text{and } |1_L\rangle \equiv \frac{1}{\sqrt{2}}(|0011\rangle + |1100\rangle).$$

$|0_L\rangle$  and  $|1_L\rangle$ , respectively, represent the  $|0\rangle$  and  $|1\rangle$  states of the single qubit of information we want to encode in the four-qubit Hilbert space. We denote this code as the [4,1] code, where the first entry in the brackets corresponds to the number of qubits in the system, and the second entry is the number of qubits of information encoded in the system. It was shown in Ref. [1] that this code satisfies the perfect QEC conditions for  $\mathcal{E}_{\text{AD}}^{\otimes 4}$ , except for small corrections of order  $\gamma^2$  and hence a recovery operation similar to  $\mathcal{R}_{\text{perf}}$  can be constructed. We refer to this recovery map as the *Leung recovery*. The worst-case fidelity for this code and recovery is plotted as a function of  $\gamma$  in Fig. 1. Clearly, the [4,1] code is able to significantly raise the worst-case fidelity for the encoded qubit of information, as compared to the no-error-correction case.

In the same figure, we also plot the worst-case fidelity using the transpose channel  $\mathcal{R}_P$  as the recovery operation, instead of the Leung recovery, for the same [4,1] code. We see that using the transpose channel as the recovery map gives a higher fidelity than the original Leung recovery.

For comparison, we also look at a recovery map for the [4,1] code constructed by Fletcher *et al.* in Ref. [27]. This recovery map, which we refer to as the *Fletcher recovery*, was originally optimized for an averaged measure of fidelity. We instead

<sup>1</sup>A set  $\mathcal{K} \equiv \{K_i\}$  is Hermitian closed if  $K_i \in \mathcal{K}$  if and only if  $K_i^\dagger \in \mathcal{K}$ .

compute the worst-case fidelity for this recovery,<sup>2</sup> also plotted in Fig. 1. For small values of  $\gamma$ , the Fletcher recovery gives the best performance compared to the other recovery maps, despite being optimized for an averaged measure of fidelity. However, it only does marginally better than the transpose channel recovery.

We also compare the performance of the  $[[4,1]]$  code under these different recovery maps with that of a code that satisfies the perfect QEC conditions. The smallest code capable of perfectly correcting an arbitrary error on any single qubit requires five qubits. The relevant noise channel now is  $\mathcal{E}_{AD}^{\otimes 5}$ . The five-qubit code  $[[14,28]]$ , usually referred to as the  $[[5,1,3]]$  code,<sup>3</sup> satisfies the perfect QEC conditions for the CP channel comprising only the single-qubit (Pauli) errors in  $\mathcal{E}_{AD}^{\otimes 5}$ . Using the corresponding  $\mathcal{R}_{perf}$  as the recovery for the  $[[5,1,3]]$  code, we compute the worst-case fidelity for the noise channel  $\mathcal{E}_{AD}^{\otimes 5}$ , for different values of  $\gamma$ . As the plot in Fig. 1 shows, the  $[[5,1,3]]$  code performs better than the  $[[4,1]]$  code with Leung recovery, but the  $[[4,1]]$  code uses one qubit less to encode the same amount of information. The  $[[4,1]]$  code with the transpose channel as recovery has nearly identical worst-case fidelity as the  $[[5,1,3]]$  code, while the one with Fletcher recovery does slightly better for small values of  $\gamma$ .

These observations clearly demonstrate the benefit of going beyond codes described by the perfect QEC conditions. Furthermore, while the  $[[5,1,3]]$  code is capable of perfectly correcting an arbitrary single-qubit error in a system subjected to *any* noise channel, the comparison with the  $[[4,1]]$  code with various recovery maps clearly show the gain that one might achieve by adapting the codes and recovery to the noise channel in question.

Last, we also compute the worst-case fidelity for randomly generated four-qubit codes, using the transpose channel as the recovery map. Computing  $F^2$  for about 500 randomly selected codes took less than half an hour on a typical laptop computer. We plot the worst-case fidelity for the best code in Fig. 1 (line marked “random 4-qubit code,  $\mathcal{R}_P$  recovery”). For small values of  $\gamma$ , this random code does not do as well as the other codes discussed so far for the amplitude damping channel, but it still does significantly better than the case without error correction. Furthermore, for  $\gamma \gtrsim 0.35$ , our randomly generated code actually outperforms all the other codes. For comparison, we have also plotted the worst-case fidelity for this randomly generated code in the absence of the transpose channel recovery, i.e., with the identity channel as the recovery map (line marked “random 4-qubit code, Id recovery”). One should keep in mind the ease with which the

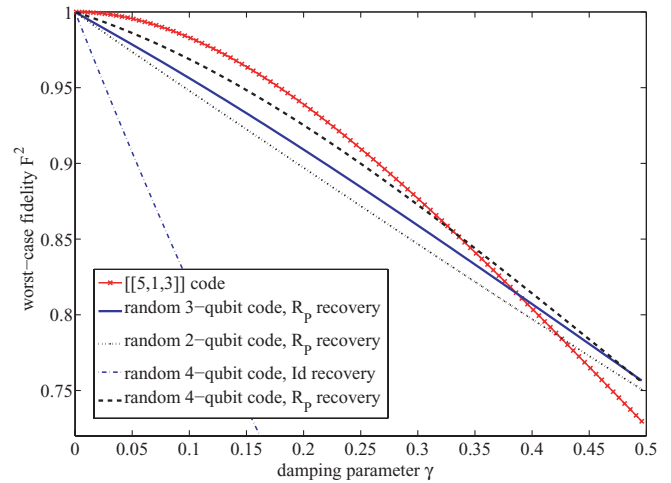


FIG. 2. (Color online) Randomly generated two-, three-, and four-qubit codes using the transpose channel as the recovery map. For comparison, we have also plotted the worst-case fidelity for the  $[[5,1,3]]$  code and that of the randomly generated four-qubit code with no recovery (i.e., identity channel as recovery).

performance of the randomly generated code was achieved, due to the fact that the transpose channel is a near-optimal recovery map for *any* code.

Finally, we also consider the possibility of constructing two-qubit and three-qubit codes for the amplitude damping channel. Because the transpose channel is near-optimal for any code, it can be used as a good recovery map for the codes we generate, thus eliminating the need to search for a good recovery for every randomly selected code. The worst-case fidelity for the best codes we found are plotted in Fig. 2. For comparison, we also plot the worst-case fidelities for the randomly generated four-qubit code mentioned in the previous paragraph, with the transpose channel and the identity channel as recovery maps. The corresponding graphs for the two- and three-qubit codes with identity channel as recovery are close to that of the four-qubit code. From the figure, we see that while the worst-case fidelity decreases as the number of physical qubits decreases, the two- and three-qubit codes in fact do not perform too badly compared to the four-qubit code or the  $[[5,1,3]]$  code. Such codes may be of relevance whenever the desire to lower resource requirements trumps the need for the best possible worst-case fidelity.

## VII. CONCLUSIONS AND OPEN PROBLEMS

In this work, we demonstrated the crucial role the transpose channel plays in perfect QEC, and used it to formulate a simple approach to characterizing and finding AQEC codes. Compared to previous work based on numerically generated recovery maps specific to the noise channel in question, the universal and analytically simple form of our transpose channel makes it particularly useful toward developing a better understanding of AQEC. While not being the optimal recovery in the case of AQEC codes, the near-optimality of the transpose channel provides a simple algorithm for identifying codes that satisfy some maximum fidelity loss requirements, without

<sup>2</sup>The recovery map we used is given in Table II of Ref. [27]. The Fletcher recovery map actually depends on two parameters  $\alpha$  and  $\beta$  which can be numerically optimized, for each value of  $\gamma$ , for the best recovery map. For simplicity, we set  $\alpha = \beta = 1/\sqrt{2}$  in our plot, which corresponds to the “code-projected recovery” in Ref. [27] with comparable performance as the fully optimized recovery.

<sup>3</sup>The first two entries in the double brackets mean the same as in the  $[[4,1]]$  code. The third entry is the distance parameter given by  $2t + 1$ , where  $t$  is the number of errors in the system the code can perfectly correct. The five-qubit code is capable of correcting an error on any qubit, so its distance parameter is equal to 3.

having to perform a difficult optimization over all recovery maps for every possible encoding. Furthermore, our approach, founded on the worst-case fidelity rather than an averaged measure of fidelity, provides the often desirable guarantee that the code found is able to protect *all* information that can be stored in the code with some minimum fidelity. We have also shown that the case of qubit codes is particularly easy to handle, and our method of computing the worst-case fidelity for a CPTP qubit map might be useful in contexts beyond our present discussion.

There are many interesting related open problems. An immediate question is whether the gap present in our AQEC conditions between the necessary and sufficient conditions (arising from the inequality in Corollary 4) can be reduced, either by improving the bound in Theorem 3 or by using a different recovery map that might perform better than the transpose channel. It would be very interesting if a simple and universal recovery map could be found, for which the dimension of the code does not appear in the worst-case fidelity. There is also the question of whether it might be possible to extend our efficient method of computing the worst-case fidelity for qubit codes to higher dimensional codes and more general channels. Finally, we expect that the transpose channel can also be used to study approximate codes more general than subspace codes, like, for example, OQEC codes [29].

Another important problem is to figure out whether the transpose channel can be easily implemented using measurements and gates. In the case of perfect QEC, the transpose channel (or equivalently  $\mathcal{R}_{\text{perf}}$ ) can be implemented simply using syndrome measurements and conditional gates (see, for example, Ref. [22]). In order for AQEC codes to be useful for computational or communication tasks, it must be possible to implement the recovery operation using physical operations that are not overly complicated or demanding in resources. This is in fact another advantage of our analytical approach over numerically constructed recovery maps for which no practical implementation structure may be apparent (although, see Ref. [7]).

AQEC provides a new and mostly unexplored arena of possibilities for the design of codes to protect information from noise for use in quantum information processing tasks. Our work provides an analytical characterization of AQEC and further analytical understanding will undoubtedly prove invaluable toward unlocking the full potential of AQEC.

#### ACKNOWLEDGMENTS

We thank David Poulin for introducing us to the problem of approximation quantum error correction and for many insightful discussions. This work is supported by NSF under Grant No. PHY-0803371.

- 
- [1] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, *Phys. Rev. A* **56**, 2567 (1997).
  - [2] A. S. Fletcher, Ph.D. thesis (MIT), e-print [arXiv:0706.3400](https://arxiv.org/abs/0706.3400) [quant-ph] (2007).
  - [3] B. Schumacher, *Phys. Rev. A* **54**, 2614 (1996).
  - [4] M. Reimpell and R. F. Werner, *Phys. Rev. Lett.* **94**, 080501 (2005).
  - [5] R. L. Kosut and D. A. Lidar, *Quant. Inf. Proc.* **8**, 443 (2009).
  - [6] A. S. Fletcher, P. W. Shor, and M. Z. Win, *Phys. Rev. A* **75**, 012338 (2007).
  - [7] A. S. Fletcher, P. W. Shor, and M. Z. Win, *Phys. Rev. A* **77**, 012320 (2008).
  - [8] R. L. Kosut, A. Shabani, and D. A. Lidar, *Phys. Rev. Lett.* **100**, 020502 (2008).
  - [9] H. Barnum and E. Knill, *J. Math. Phys.* **43**, 2097 (2002).
  - [10] J. Tyson, e-print [arXiv:0907.3386v4](https://arxiv.org/abs/0907.3386v4) [quant-ph].
  - [11] N. Yamamoto, S. Hara, and K. Tsumura, *Phys. Rev. A* **71**, 022322 (2005).
  - [12] M. Ohya and D. Petz, *Quantum Entropy and Its Use* (Springer-Verlag, Berlin, Heidelberg, 1993).
  - [13] A. Ekert and C. Macchiavello, *Phys. Rev. Lett.* **77**, 2585 (1996).
  - [14] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Phys. Rev. A* **54**, 3824 (1996).
  - [15] E. Knill and R. Laflamme, *Phys. Rev. A* **55**, 900 (1997).
  - [16] B. Schumacher and M. D. Westmoreland, *Quant. Info. Proc.* **1**, 5 (2002).
  - [17] R. Klesse, *Phys. Rev. A* **75**, 062315 (2007).
  - [18] F. Buscemi, *Phys. Rev. A* **77**, 012309 (2008).
  - [19] C. Bény, TQC 2009 LNCS 5906, p. 6675 (2009), e-print [arXiv:0907.4207](https://arxiv.org/abs/0907.4207).
  - [20] J. M. Renes, e-print [arXiv:1003.1150v2](https://arxiv.org/abs/1003.1150v2) [quant-ph].
  - [21] C. Bény and O. Oreshkov, *Phys. Rev. Lett.* **104**, 120501 (2010).
  - [22] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).
  - [23] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, *Phys. Rev. Lett.* **100**, 030501 (2008).
  - [24] D. Petz, *Rev. Math. Phys.* **15**, 79 (2003).
  - [25] P. Hayden, R. Jozsa, D. Petz, and A. Winter, *Commun. Math. Phys.* **246**, 359 (2004).
  - [26] P. Mandayam and D. Poulin (in preparation, 2009).
  - [27] A. S. Fletcher, P. W. Shor, and M. Z. Win, *IEEE Trans. Info. Theory* **54**, 5705 (2008).
  - [28] R. Laflamme, C. Miquel, J.-P. Paz, and W. H. Zurek, *Phys. Rev. Lett.* **77**, 198 (1996).
  - [29] D. Kribs, R. Laflamme, D. Poulin, and M. Lesosky, *Quant. Inf. Comp.* **6**, 382 (2006).