Optimum quantum error recovery using semidefinite programming

Andrew S. Fletcher,[1,2,*] Peter W. Shor,[3,†] and Moe Z. Win[1,‡]
[1]*Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
[2]*Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, Massachusetts 02420*
[3]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*
(Received 7 June 2006; published 31 January 2007)

Quantum error correction (QEC) is an essential element of physical quantum information processing systems. Most QEC efforts focus on extending classical error correction schemes to the quantum regime. The input to a noisy system is embedded in a coded subspace, and error recovery is performed via an operation designed to perfectly correct for a set of errors, presumably a large subset of the physical noise process. In this paper, we examine the choice of recovery operation. Rather than seeking perfect correction on a subset of errors, we seek a recovery operation to maximize the entanglement fidelity for a given input state and noise model. In this way, the recovery operation is optimal for the given encoding and noise process. This optimization is shown to be calculable via a semidefinite program, a well-established form of convex optimization with efficient algorithms for its solution. The error recovery operation may also be interpreted as a combining operation following a quantum spreading channel, thus providing a quantum analogy to the classical diversity combining operation.

DOI: 10.1103/PhysRevA.75.012338 PACS number(s): 03.67.Pp, 02.60.Pn

I. INTRODUCTION

Any implementation of quantum computing or communications requires a strategy for error mitigation. Indeed, the development of quantum error correction (QEC) schemes was an important early step in moving quantum computing from an interesting theoretical idea to an exciting field with potential for ground-breaking technological implementations [1]. The importance of efficient and optimum error mitigation only increases as the field advances.

The earliest efforts in QEC used encoding techniques modified from classical error correction schemes [1–5]. Further analysis [6,7] laid the foundation for QEC theory, noting that the important metric is how faithful the statistics of the corrected state remain to the ideal behavior. While that observation suggests quantum error mitigation is thus an optimization problem, most of the subsequent work in the field has appropriately focused on perfect recovery from a set of errors. This emphasis has allowed many techniques to be borrowed from classical error correction and enabled important feasibility studies in quantum computing. It is not, however, the only way to consider controlling for quantum errors [8].

Recently, some authors [9–11] have returned to examining quantum error mitigation as an optimization problem. The essential properties of a quantum state are the statistics of any observable outcome; these are completely encapsulated in the density operator ρ of the state. Noise is introduced by the operation $\mathcal{E}$, which can be thought of as a noisy quantum communications channel. Thus, the goal of any error correction scheme is to design a recovery operation $\mathcal{R}$ such that the recovered state is as faithful a representation as possible of the input, judged by how well the statistics of observables are preserved. The optimum recovery minimizes the "distance" between an input density ρ and the output $\mathcal{R}(\mathcal{E}(\rho))$. This operation may differ from the more traditional QEC recovery operation; such differences illustrate further the contrast between quantum and classical error correction. To distinguish this approach from QEC, we use the term quantum error recovery (QER). It should be emphasized that the optimum QER recovery operation is dependent on a given input density, encoding operation, and noise model.

The paper is organized as follows. In Sec. II, we define the parameters for optimum QER. In Sec. III, the optimum recovery operation is cast as a semidefinite program. Section IV interprets the recovery operation as an optimal combining problem and illustrates the computational benefit of such an interpretation. In Sec. V, QER operations are derived for the amplitude damping channel using codes encoding one qubit into four and five qubits.

II. OPTIMUM QER

Most QEC procedures are designed on the principle of "perfect" correction of arbitrary single qubit errors. Such a design postulates that single qubit errors are the dominant terms in the noise process; thus a scheme that corrects arbitrary single qubit errors and ignores higher order terms will sufficiently mitigate the noise. Pursuit of this approach has led to important results on the feasibility of quantum error correction. However, one may reasonably ask how well this generic approach succeeds in specific cases.

With every quantum code, in the current paradigm, there is an associated recovery operation designed to perfectly correct the dominant errors. This "traditional" recovery (referred hereafter as the QEC recovery operation) applies a syndrome

*Electronic address: fletcher@mit.edu
†Electronic address: shor@math.mit.edu
‡Electronic address: moewin@mit.edu

measurement to determine which error occurred and a correction operation dependent on the observed syndrome. For a given code and error process, the QEC recovery operation may not provide the most effective safeguard from error. Depending upon the form of the error process, an alternate recovery operation may be designed that better preserves the input state, based on some measure of statistical "closeness" between the input density ρ and the output density $\mathcal{R}(\mathcal{E}(\rho))$. Commonly used metrics for quantum information arise from the fidelity, defined to be

$$F(\rho,\sigma) = \mathrm{tr}\sqrt{\rho^{1/2}\sigma\rho^{1/2}}, \tag{1}$$

where ρ and σ are density operators. The fidelity F takes a value between 0 and 1, where 1 indicates that the states are identical. Using F_o to represent any fidelity-based measure of channel performance, the error recovery optimization problem is to find

$$\mathcal{R}^{\star} = \arg\max_{\{\mathcal{R}\}} F_o(\mathcal{R}\circ\mathcal{E}), \tag{2}$$

where arg refers to the element of $\{\mathcal{R}\}$ that achieves the maximum. To understand this optimization problem, the remainder of this section will discuss the choice of F_o and the properties of the set $\{\mathcal{R}\}$ of recovery operations.

The fidelity, as defined in (1), measures the closeness between two quantum states; F_o in (2) quantifies the information-preserving behavior of the channel $\mathcal{R}\circ\mathcal{E}$. There are three standard choices of fidelity-based measures of channel performance. In [10], the minimum fidelity was chosen for F_o, and a sub-optimal solution to (2) was obtained. In [12], the ensemble average fidelity was applied to a similar optimization problem. In this paper, we follow the approach of [9] and seek to maximize the *entanglement fidelity*.[1] For a more complete discussion of these channel measures, we refer the reader to [14].

Entanglement fidelity [15] arises from the mathematical concept of mixed state purification. Any mixed quantum state can be represented as a subsystem of a pure state in a larger Hilbert space. The subsystem is mixed due to the entangled nature of the pure state. Consider a mixed state $\rho \in \mathcal{L}(\mathcal{H})$. By defining a reference space $\mathcal{B}$, we may denote ρ as a subsystem of a pure state,

$$\rho = \mathrm{tr}_{\mathcal{B}}|BH\rangle\langle BH|, \tag{3}$$

where $|BH\rangle$ is a pure state in the space $\mathcal{B}\otimes\mathcal{H}$. The entanglement fidelity then measures how faithfully a channel $\mathcal{A}$ maintains the purification (or, equivalently, how well it preserves the entanglement). It is given by

$$F_{\mathrm{ent}}(\rho,\mathcal{A}) = \langle BH|\mathcal{I}_{\mathcal{B}}\otimes\mathcal{A}(|BH\rangle\langle BH|)|BH\rangle, \tag{4}$$

the squared fidelity of the input $|BH\rangle$ with the output of the channel $\mathcal{I}_{\mathcal{B}}\otimes\mathcal{A}$.

With entanglement fidelity as F_o in (2), we define the optimum recovery operation for the error process $\mathcal{E}$ and input distribution ρ as

$$\mathcal{R}_{\rho}^{\star} = \arg\max_{\{\mathcal{R}\}} F_{\mathrm{ent}}(\rho,\mathcal{R}\circ\mathcal{E}). \tag{5}$$

We now turn our attention to the set of recovery operations over which we are optimizing in (5). A quantum operation must be completely positive (CP) and trace-preserving (TP)[16]. This requirement follows directly from the postulates of quantum mechanics wherein the evolution of a closed quantum system is unitary. Let $\mathcal{A}$ be a CPTP map from $\mathcal{L}(\mathcal{H})\mapsto\mathcal{L}(\mathcal{K})$. The most familiar representation of the CPTP map is the Kraus (or operator sum) form, where the mapping is specified by a set of operators $\{A_k\}$ known as the operator elements [16]. The channel output is given by

$$\mathcal{E}(\rho) = \sum_k A_k\rho A_k^{\dagger}, \tag{6}$$

and the CPTP constraint is met when

$$\sum_k A_k^{\dagger}A_k = I, \tag{7}$$

where I is the identity operator in $\mathcal{L}(\mathcal{H})$.

While a properly constrained set of operators fully specifies a quantum channel, the Kraus form is inconvenient for optimization. The most obvious inconvenience is the many-to-one correspondence between sets of operator elements and channel mappings. For this reason, we will utilize the representation of a channel mapping $\mathcal{A}$ by a positive semidefinite operator $X_{\mathcal{A}}$ on $\mathcal{L}(\mathcal{H}\otimes\mathcal{K})$ [17], for which the correspondence is one-to-one [18]. We will refer to the operator $X_{\mathcal{A}}$ as the Choi matrix [19–21].

The Choi matrix may be defined in several ways. We will employ the common practice of "stacking" columns of a matrix to form a vector. Let C be a bounded linear operator from $\mathcal{H}_1$ to $\mathcal{H}_2$: $C\in\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)$. We define a ket in the Hilbert space $\mathcal{H}_1\otimes\mathcal{H}_2$ associated with C as[2]

$$|C\rangle\rangle = \sum_{ij} c_{ij}|j\rangle_1|i\rangle_2, \tag{8}$$

where $\{|j\rangle_1\}$ and $\{|i\rangle_2\}$ are orthonormal bases for $\mathcal{H}_1$ and $\mathcal{H}_2$, respectively, and $c_{ij}={}_2\langle i|C|j\rangle_1$ is the matrix element of C on these bases. Two useful relations follow directly from this definition. The first one,

$$A\otimes B|C\rangle\rangle = |BCA^T\rangle\rangle, \tag{9}$$

applies whenever the dimensions of A, B, and C indicate that BCA^T is a valid operator.[3] The second relation applies for C_1, $C_2\in\mathcal{L}(\mathcal{H}_1,\mathcal{H}_2)$:

$$\mathrm{tr}_{\mathcal{H}_1}[|C_1\rangle\rangle\langle\langle C_2|] = C_1C_2^{\dagger}\in\mathcal{L}(\mathcal{H}_2). \tag{10}$$

[1] Use of the entanglement fidelity requires an assumption for the input density ρ, leading one to argue that it is less general. The impact of this assumption, however, is relatively benign. For example, the choice of entanglement fidelity has no impact on channel capacities [13].

[2] The notation $|\cdot\rangle\rangle$ is used to emphasize that these kets represents operators, not quantum states.

[3] Note the superscript T indicates the transpose with respect to the specified basis without conjugation.

The Choi matrix for the quantum operation $\mathcal{A}$ is given by $X_{\mathcal{A}} \equiv \Sigma_k |A_k\rangle\rangle\langle\langle A_k|$, and the channel mapping $\mathcal{A}: \mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{K})$ is defined by

$$\mathcal{A}(\rho) = \mathrm{tr}_{\mathcal{H}}[(\rho^T \otimes I) X_{\mathcal{A}}]. \tag{11}$$

The required trace-preserving property can be stated as

$$\mathrm{tr}_{\mathcal{K}} X_{\mathcal{A}} = I \in \mathcal{L}(\mathcal{H}). \tag{12}$$

III. OPTIMUM RECOVERY VIA SEMIDEFINITE PROGRAMMING

The problem given by (5) is a convex optimization problem, and we may approach it with sophisticated tools. Particularly powerful is the semidefinite program (SDP) [22], where the objective function is linear in an input constrained to a semidefinite cone. Indeed, the power of the SDP was a primary motivation in choosing to maximize the entanglement fidelity, which is linear in the quantum operation $\mathcal{R}$.

The definition of entanglement fidelity given in (4) is intuitively useful, but awkward for calculations. An easier form arises when operator elements $\{A_i\}$ for $\mathcal{A}$ are given. The entanglement fidelity is then

$$F_{\mathrm{ent}}(\rho, \mathcal{A}) = \sum_i |\mathrm{tr}(\rho A_i)|^2. \tag{13}$$

From (13), we may derive a calculation rule for the entanglement fidelity when the channel $\mathcal{A}$ is expressed via the Choi matrix. Using (10), we see that $\mathrm{tr}\, A_i \rho = \mathrm{tr} |A_i\rangle\rangle\langle\langle \rho| = \langle\langle \rho | A_i \rangle\rangle$. Inserting this into (13), we obtain the entanglement fidelity in terms of $X_{\mathcal{A}}$:

$$F_{\mathrm{ent}}(\rho, \mathcal{A}) = \sum_i \langle\langle \rho | A_i \rangle\rangle \langle\langle A_i | \rho \rangle\rangle = \langle\langle \rho | X_{\mathcal{A}} | \rho \rangle\rangle. \tag{14}$$

Armed with this expression for the entanglement fidelity, we may now express (5) in a form readily seen to be a semidefinite program. To do this, we must consider the form of the composite operation $\mathcal{R} \circ \mathcal{E}: \mathcal{L}(\mathcal{H}) \mapsto \mathcal{L}(\mathcal{H})$ expressed as a positive operator on $\mathcal{H} \otimes \mathcal{H}$. If the operator elements for the recovery and channel are $\{R_i\}$ and $\{E_j\}$, respectively, then the operator $X_{\mathcal{R}\circ\mathcal{E}}$ is given by

$$X_{\mathcal{R}\circ\mathcal{E}} = \sum_{ij} |R_i E_j\rangle\rangle\langle\langle R_i E_j|. \tag{15}$$

Applying (9), this becomes

$$X_{\mathcal{R}\circ\mathcal{E}} = \sum_{ij} E_j^T \otimes I |R_i\rangle\rangle\langle\langle R_i| E_j^* \otimes I = \sum_j (E_j^T \otimes I) X_{\mathcal{R}} (E_j^* \otimes I), \tag{16}$$

where the $*$ represents complex conjugation, without transposition. The entanglement fidelity is then

$$F_{\mathrm{ent}}(\rho, \mathcal{R} \circ \mathcal{E}) = \sum_j \langle\langle \rho | (E_j^T \otimes I) X_{\mathcal{R}} (E_j^* \otimes I) | \rho \rangle\rangle = \mathrm{tr}\, X_{\mathcal{R}} C_{\rho,\mathcal{E}}, \tag{17}$$

where

$$C_{\rho,\mathcal{E}} = \sum_j E_j^* \otimes I |\rho\rangle\rangle\langle\langle \rho| E_j^T \otimes I = \sum_j |\rho E_j^\dagger\rangle\rangle\langle\langle \rho E_j^\dagger|. \tag{18}$$

We may now express the optimization problem (5) in the simple form

$$X_{\mathcal{R}_\rho}^\star = \arg\max_X \mathrm{tr}(X C_{\rho,\mathcal{E}}), \quad \text{such that } X \geqslant 0,\ \mathrm{tr}_{\mathcal{H}} X = I. \tag{19}$$

This form illustrates plainly the linearity of the objective function and the semidefinite and equality structure of the constraints. Indeed, this is the exact form of the optimization problem in [12], which first pointed out the value of semidefinite programs (SDP) for optimizing quantum channels.

The value of an SDP for optimization is two-fold. First, an SDP is a subclass of convex optimization, and thus a local optimum is guaranteed to be a global optimum. Second, there are efficient and well-understood algorithms for computing the optimum of a semidefinite program. These algorithms are sufficiently mature to be widely available. Thus, by expressing the optimum recovery channel as an SDP, we have explicit means to compute the solution for an arbitrary channel $\mathcal{E}$.

IV. QUANTUM DIVERSITY COMBINING

In the preceding analysis, the method of encoding has been implied by the choice of ρ. Indeed, in most treatments of QEC, the input density is restricted to a subspace called the quantum error correcting code (QECC). If P_C is a projection operator onto the code subspace, then $\rho = P_C \rho P_C$ implies that the state ρ is within the code subspace. The channel is typically defined such that the input and output spaces are identical (i.e., $\mathcal{H} = \mathcal{K}$), and the noise process generally perturbs the system from the code subspace. While this representation is a perfectly legitimate model for the error process (and a convenient form when viewing QEC in a mode comparable to classical error correction) the dimensionality is unnecessarily high for the optimization routine.

Recall that $X_{\mathcal{R}}$ is a Hermitian operator on the space $\mathcal{H} \otimes \mathcal{K}$. The optimization thus has $d_{\mathcal{H}}^2 d_{\mathcal{K}}^2$ real parameters. Even for a [[5, 1, 3]][4] code, the smallest for arbitrary single qubit errors [7,23], this optimization has 2^{20} optimization variables. The high dimensionality can be alleviated somewhat by embedding the encoding into the noise process and redefining $\mathcal{E}$ as a quantum spreading channel where $d_{\mathcal{H}} < d_{\mathcal{K}}$.

The transform to a quantum spreading channel is illustrated in Fig. 1. Consider the noise process $\mathcal{E}: \mathcal{K} \mapsto \mathcal{K}$ with operator elements $\{E_i\}$, input density $\rho \in \mathcal{L}(\mathcal{K})$, and code projector

[4]The notation $[[n,k,d]]$ refers to a quantum code encoding k qubits of information into a n qubit system. The parameter d refers to the *weight* of the code [14].

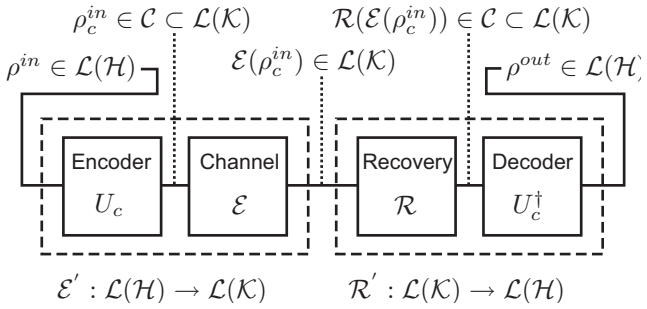

FIG. 1. Transform from standard quantum error correction to quantum spreading channel. By considering the error channel $\mathcal{E}$ together with the encoding U_C and the recovery operation $\mathcal{R}$ with the decoder $U_C^\dagger$, the dimension of the SDP may be reduced by a factor of $(\dim \mathcal{K}/\dim \mathcal{H})^2$. Note that the output of the encoder U_C and the recovery channel $\mathcal{R}$ live on the code subspace $\mathcal{C}$.

$$P_C = \sum_n^{d_\mathcal{H}} |n\rangle_{LL}\langle n|, \tag{20}$$

where $|n\rangle_L \in \mathcal{K}$ are the logical states of the code. Since the input is in the codespace, ρ is preserved by the code projector: $\rho = P_C \rho P_C$. We can reduce the dimensionality of the optimization by transforming this problem. Let $\mathcal{H}$ be a $d_\mathcal{H}$ dimensional Hilbert space with orthonormal basis $\{|n\rangle_\mathcal{H}\}$. $\mathcal{H}$ can be considered the space of the information source. The encoding process is accomplished by the operator $U_C = \Sigma_n^{d_\mathcal{H}} |n\rangle_{L\mathcal{H}}\langle n|$ that maps the basis states of $\mathcal{H}$ to the logical states in $\mathcal{K}$. U_C is an isometry: $U_C^\dagger U_C = I$. Note that

$$U_C U_C^\dagger = \sum_n^{d_\mathcal{H}} |n\rangle_{LL}\langle n| = P_C. \tag{21}$$

If we redefine $\rho' = U_C^\dagger \rho U_C \in \mathcal{L}(\mathcal{H})$ and the operator elements $E_i' = E_i U_C$, then we see the processes $\mathcal{E}(\rho)$ and $\mathcal{E}'(\rho')$ are identical:

$$\mathcal{E}'(\rho') = \sum_i E_i U_C U_C^\dagger \rho U_C U_C^\dagger E_i^\dagger = \sum_i E_i P_C \rho P_C E_i^\dagger = \mathcal{E}(\rho). \tag{22}$$

By enacting such a transformation, the optimization dimension is reduced from $d_\mathcal{K}^4$ to $d_\mathcal{K}^2 d_\mathcal{H}^2$. For the [[5, 1, 3]] code, the reduction is from 2^{20} to 2^{12}.

The above transformation illustrates an alternative interpretation of recovering an encoded quantum state after an error process. We may instead consider an unencoded state input into a *quantum spreading channel*, a channel in which the output dimension is greater than the input dimension (i.e., $\dim \mathcal{K} > \dim \mathcal{H}$). The recovery operation is an attempt to combine the spread output back into the input space, presumably with the intent to minimize information loss. The recovery operation is then the quantum analog to the classical communications concept of diversity combining.

Classical diversity combining describes a broad class of problems in communications and radar systems. In its most general form, we may consider any class of transmission problems in which the receiver observes multiple transmission channels. These channels could arise due to multi-path scattering, frequency diversity (high bandwidth transmissions where channel response varies with frequency), spatial diversity (free-space propagation to multiple physically separated antennas), time diversity, or some combination of the four. Diversity combining is a catch-all phrase for the process of exploiting the multiple channel outputs to improve the quality of transmission (e.g., by reducing error or increasing data throughput).

In a general description of classical diversity, the input signal is coupled through the channel to a receiver system of higher dimensionality. Consider a communication signal with a single transmitter antenna and N receiver antennae. Often, the desired output is a signal of the same dimension as the input, a scalar in this case. Diversity combining is then the process of extracting the maximum information from the N-dimensional received system. In most communications systems, this combining is done at either the analog level (leading to beam-forming or multi-user detection) or digital level (making the diversity system a kind of repeater code). Thus, the natural inclination is to equate diversity combining with either beam-forming or repeater codes. The most general picture of diversity combining, however, is an operation that recombines the channel output into a signal of the same dimension as the input. Thus, it is appropriate to consider a quantum spreading channel to be a quantum diversity channel, and the recovery operation to be a quantum diversity combiner.

V. EXAMPLES

To illustrate the potential benefit of optimum QER, we numerically compare the performance of the optimum QER with QEC recovery for two encoding schemes. First, we examine the [[5, 1, 3]] code in the presence of the amplitude damping channel. Second, we compare the four-bit amplitude damping code from [8] with the optimum QER. The latter is a particularly apt choice for comparison, as the code was designed for a specific channel and sought to only approximately satisfy the quantum error correcting conditions.

In both cases, the noise channel $\mathcal{E}_a$ is the amplitude damping channel. For a single qubit, $\mathcal{E}_a$ has operator elements

$$E_0 = \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{bmatrix} \quad \text{and} \quad E_1 = \begin{bmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{bmatrix}. \tag{23}$$

This channel is a commonly encountered model, where the parameter γ indicates the probability of decaying from state $|1\rangle$ to $|0\rangle$ (i.e., the probability of losing a photon). Amplitude damping is a logical choice to illustrate the benefits of optimum QER, as the operation is not symmetric with respect to $|0\rangle$ and $|1\rangle$.

A. Five-qubit code

The five-qubit code was independently discovered by Refs.[7] and [23]. We will here follow the treatment in [14] and specify the code via the generators $\{g_1, g_2, g_3, g_4\}$ and the logical $\bar{Z}$ and $\bar{X}$ operations given in Table I. The code subspace $\mathcal{C}$ is the two-dimensional subspace that is the +1

TABLE I. Generators and logical operations of the five qubit (i.e., [[5,1,3]]) code.

Name	Operator
g_1	$XZZXI$
g_2	$IXZZX$
g_3	$XIXZZ$
g_4	$ZXIXZ$
$\bar{Z}$	$ZZZZZ$
$\bar{X}$	$XXXXX$

eigenspace of the generators g_i. The logical states $|0\rangle_L$ and $|1\rangle_L$ are the $+1$ and -1 eigenkets of $\bar{Z}$ on $\mathcal{C}$.

To compute the optimum recovery for this code, we assume that the logical states are equally likely; that is, $\rho = \frac{1}{2}|0\rangle_{LL}\langle 0| + \frac{1}{2}|1\rangle_{LL}\langle 1|$. (This choice of ρ in fact assumes nothing about the choice of codewords; rather, it is the maximum entropy distribution constrained to the code space $\mathcal{C}$.) The noise channel is $\mathcal{E}_a^{\otimes 5}$, the amplitude damping channel acting independently on each qubit. For each choice of the parameter γ, the optimum recovery operation $\mathcal{R}_\rho^\star$ is computing according to (5). We compare the entanglement fidelity $F_{\text{ent}}(\rho, \mathcal{R}\circ\mathcal{E}_a^{\otimes 5})$ for both $\mathcal{R}_\rho^\star$ and $\mathcal{R}_{\text{QEC}}$ in Fig. 2.

Figure 2 illustrates clearly the difference between optimum QER and QEC recoveries for large values of γ. It is also instructive to compare the techniques for small values of γ. We do this numerically by calculating the polynomial expansion of $F_{\text{ent}}(\rho, \mathcal{R}\circ\mathcal{E})$ as γ goes to zero. The entanglement fidelity for the optimum QER has the form $F_{\text{ent}}(\rho, \mathcal{R}\circ\mathcal{E}) \approx 1 - 1.166\gamma^2 + \mathcal{O}(\gamma^3)$. In contrast, the QEC recovery is $F_{ent}(\rho, \mathcal{R}\circ\mathcal{E}) \approx 1 - 2.5\gamma^2 + \mathcal{O}(\gamma^3)$.

B. Four-qubit "approximate" code

In [8], Leung *et al.* recognized the advantage that channel-specific error recovery schemes can have over generic QEC routines. To illustrate the advantage, they designed a code specific to the amplitude damping channel with good performance for small γ that only required four qubits of overhead, in contrast to the generic five qubit code. The code only approximately meets the QEC conditions, and as a result, the "corrected" state is somewhat distorted from the input, even when the dominant error term occurs. In this way, the procedure is based upon principles similar to the optimum QER we have developed.

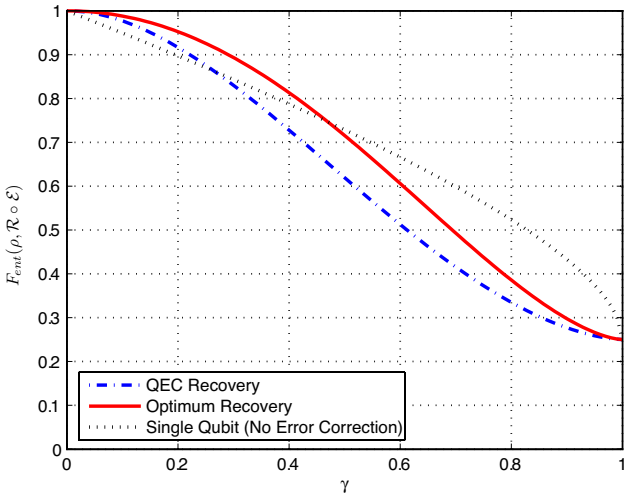

FIG. 2. (Color online) Entanglement fidelity vs γ for the five qubit code and the amplitude damping channel $\mathcal{E}_a^{\otimes 5}$. γ refers to the damping parameter of the channel. Entanglement fidelity for a single qubit and no error correction [i.e., $F_{\text{ent}}(\rho, \mathcal{E}_a)$] is included for comparative purposes.

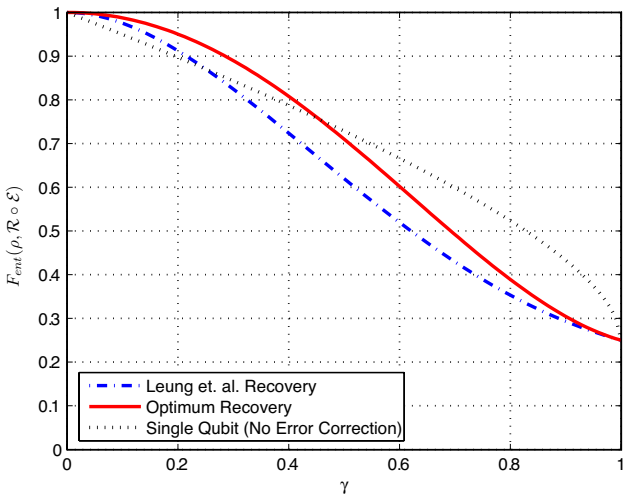

FIG. 3. (Color online) Entanglement fidelity vs γ for the four qubit code of Leung *et al.* [8] and the amplitude damping channel $\mathcal{E}_a^{\otimes 4}$. Entanglement fidelity for a single qubit and no error correction [i.e., $F_{\text{ent}}(\rho, \mathcal{E}_a)$] is included for comparative purposes.

The logical states are given by

$$|0\rangle_L = \frac{1}{\sqrt{2}}(|0000\rangle + |1111\rangle), \tag{24}$$

$$|1\rangle_L = \frac{1}{\sqrt{2}}(|0011\rangle + |1100\rangle), \tag{25}$$

and the recovery operation is specified by the circuits in Fig. 2 of [8]. We note that the recovery operation depends explicitly on the parameter γ. We compare the recovery of Leung *et al.* with the optimum QER computed according to (5), once again assuming the completely mixed input density $\rho = \frac{1}{2}|0\rangle_{LL}\langle 0| + \frac{1}{2}|1\rangle_{LL}\langle 1|$. The numerical comparison for various values of γ is provided in Fig. 3. As γ goes to zero, the entanglement fidelity for the optimum QER is numerically determined to have the form $F_{ent}(\rho, \mathcal{R}\circ\mathcal{E}) \approx 1 - 1.25\gamma^2 + \mathcal{O}(\gamma^3)$. In contrast, the Leung *et al.* recovery is $F_{\text{ent}}(\rho, \mathcal{R}\circ\mathcal{E}) \approx 1 - 2.75\gamma^2 + \mathcal{O}(\gamma^3)$.

C. Commentary on examples

It is not surprising that the optimum QER operation outperforms the five qubit code and the Leung code. It is, however, noteworthy that the difference can be significant, for even relatively small values of γ. We may conclude that channel specific recoveries can significantly improve the performance of error correcting procedures. This conclusion

was shared by Leung *et al.*, who lamented the lack of a general method to design such channel-specific schemes. The SDP formalism outlined in this paper provides such a general method.

Perhaps the most notable improvement in optimum QER can be noted in the "single qubit (no error correction)" comparisons in Figs. 2 and 3. These curves represent the entanglement fidelity of a single qubit transmitted through the noisy channel. For values of γ where the recovered entanglement fidelity lies below the "single qubit," we see that the error mitigation procedure does more harm than good. Performing optimum QER as opposed to QEC significantly extends the values of γ for which error mitigation is valid. This suggests optimum QER will be a particularly valuable technique for noisier systems.

Finally, it is worth noting the duality between optimum QER and optimum encoding. We have derived the optimal operation $\mathcal{R}$ for a given encoding and noise process. By the same process, one may derive the optimum encoding given a recovery operation and noise process. This can most easily be seen by noting that the encoding operation is a valid quantum operation, and in fact, a spreading operation; it is thus subject to the same semidefinite cone constraints as the recovery operation. In a manner similar to those suggested by [24,9], one can conceivably obtain a channel-specific error recovery scheme by alternatively holding the recovery fixed and optimizing the encoding, and holding the encoding fixed and optimizing the recovery. Full analysis of such a technique is deferred for future consideration.

VI. CONCLUSION

The structure of quantum operations allows quantum error correction to be approached as an optimization problem. Specifically, optimum recovery of an encoded quantum state from an error process can be solved numerically using semidefinite programming when optimality is interpreted as a maximization of the entanglement fidelity. This analysis suggests the ability to systematically search for recovery operations for complicated error schemes beyond those readily analyzed and corrected through more traditional QEC methodologies. This problem is, in general, the optimum combining operation following a quantum spreading channel and is thus a quantum parallel to the classical problem of diversity combining. We have shown, through the example of the amplitude damping channel, that knowledge of the physical noise process can be exploited to significantly improve the performance of an error recovery operation.

Several questions remain to fully understand the impact of optimum QER on implementing a quantum information processing device. One should like to understand qualitatively the action of the optimum recovery, so as to develop intuition in quantum code and recovery design. It will also be important to establish the ease with which such operations could be implemented in any of the models for quantum computation. Furthermore, we showed the exponential growth of the SDP in the channel dimension, thus making it difficult to analyze high dimensional channels or long codes. It will be valuable to develop computationally feasible techniques to determine optimal or near-optimal recoveries in such cases. Finally, it is of interest to understand how optimum QER will impact fault tolerant quantum computation, particularly the potential propagation of error.

Recently, two different sets of authors, [25,26] presented parallel work which explores the iterative encoding and recovery problem via semidefinite programming.

ACKNOWLEDGMENTS

A.S.F. would like to thank the Department of the Air Force, who sponsored this work under AF Contract No. FA8721-05-C-0002. P.W.S. thanks the National Science Foundation for support through Grant CCF-0431787. Opinions, interpretations, recommendations, and conclusions are those of the authors and are not necessarily endorsed by the United States Government.

[1] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).
[2] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[3] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[4] A. R. Calderbank, E. M. Rains, P. W. Shor, and N. J. A. Sloane, Phys. Rev. Lett. **78**, 405 (1997).
[5] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).
[6] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).
[7] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and K. W. Wootters, Phys. Rev. A **54**, 3824 (1996).
[8] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Phys. Rev. A **56**, 2567 (1997).
[9] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94**, 080501 (2005).
[10] N. Yamamoto, S. Hara, and K. Tsumura, Phys. Rev. A **71**, 022322 (2005).
[11] B. G. Bodmann, D. W. Kribs, and V. I. Paulsen, in *Decoherence-insensitive quantum communication by optimal C^*-encoding* (2006), quant-ph/0605235
[12] K. Audenaert and B. De Moor, Phys. Rev. A **65**, 030302(R) (2002).
[13] H. Barnum, E. Knill, and M. A. Nielsen, IEEE Trans. Inf. Theory **46**, 1317 (2000).
[14] M. A. Nielsen and I. L. Chuang, in *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[15] B. Schumacher, Phys. Rev. A **54**, 2614 (1996).
[16] K. Kraus, in *States, Effects, and Operations: Fundamental Notions of Quantum Theory, Vol. 190 of Lecture Notes in Physics* (Springer-Verlag, Berlin, 1983).
[17] M.-D. Choi, Linear Algebr. Appl. **10**, 285 (1975).
[18] G. M. D'Ariano and P. Lo Presti, Phys. Rev. A **64**, 042308 (2001).
[19] T. F. Havel, J. Math. Phys. **44**, 534 (2003).

[20] C. M. Caves, J. Supercond. **12**, 707 (1999).
[21] J. de Pillis, Pac. J. Math. **23**, 129 (1967).
[22] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).
[23] R. Laflamme, C. Miquel, J.-P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).
[24] P. W. Shor, Math. Program. **97**, 311 (2003).
[25] M. Reimpell, R. F. Werner, and K. Audenaert, in *Comment on "Optimum quantum error recovery using semidefinite programming"* (2006), quant-ph/0606059
[26] R. L. Kosut and D. A. Lidar, in *Quantum error corretion via convex optimization* (2006), quant-ph/0606078