Optimized entanglement-assisted quantum error correction

Soraya Taghavi,[1] Todd A. Brun,[1] and Daniel A. Lidar[2]

[1]*Department of Electrical Engineering, Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089, USA*

[2]*Departments of Electrical Engineering, Chemistry and Physics, Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, California 90089, USA*

Using convex optimization, we propose entanglement-assisted quantum error-correction procedures that are optimized for given noise channels. We demonstrate through numerical examples that such an optimized error-correction method achieves higher channel fidelities than existing methods. This improved performance, which leads to perfect error correction for a larger class of error channels, is interpreted in at least some cases by quantum teleportation, but for general channels this interpretation does not hold.

I. INTRODUCTION

Noise is an important obstacle for the scale-up of quantum information devices—both decoherence due to interactions with an external environment and the internal control errors inevitable for any information-processing. The theory of quantum error correction was developed by analogy to classical error-correcting codes to overcome this obstacle [1–4]. Quantum error-correcting codes (QECCs) store quantum information redundantly in an entangled state of multiple quantum systems (usually qubits) in such a way that errors can be detected and corrected without directly measuring (and hence disturbing) the quantum information to be protected. The fact that this can be done at all is quite remarkable and has opened up the prospect that quantum information processing may become a viable technology in the foreseeable future.

Broadly speaking, quantum error correction is useful in two main contexts. The first is quantum computation, where errors that occur during a calculation must be constantly found and corrected to keep them from accumulating. The QECCs used for this purpose must fit into an overall fault-tolerant scheme [5–8], which puts important constraints on the properties of the codes used.

The other main context is quantum communication. Here it is assumed that the steps of encoding and decoding are largely error free, and all errors occur during the transmission of quantum information through a noisy quantum channel. This is the context we consider in this article. It should be noted, though, that this division is not really sharp: Quantum computers may well need to transmit information internally, and quantum communication may be part of a distributed information-processing scheme.

In the context of quantum communications, one can explore the effect of different shared resources on the ability to detect and correct errors. Recently, a great deal of work has been done on entanglement-assisted quantum error correction (EAQEC) and entanglement-assisted quantum error-correcting codes (EAQECCs). In these schemes, the sender (usually called Alice) and receiver (usually called Bob) share some number of pure maximally entangled states, or ebits. Alice can use her halves of these ebits in the encoding procedure, and Bob can use his halves in error correction and decoding. The idea of using entanglement to improve error correction was proposed early in the history of quantum information theory [9], but it took a long time for the first steps toward designing practical codes [10]. This work has led to a generalized theory of QECCs, the entanglement-assisted stabilizer formalism [11–17].

The large majority of work on QECCs has concerned stabilizer codes [3], which can be derived from classical linear codes. In this algebraic approach, one generally tries to design codes that are able to detect and correct an abstract set of errors acting on single qubits, often based on the Pauli operators. QECCs have the property that any linear combination of correctable errors is also a correctable error. Since the Pauli operators form a basis for all single-qubit operators, the ability to correct Pauli-based errors implies the ability to correct more realistic models of noise, provided that errors are not too highly correlated between qubits. Most EAQECCs that have been derived so far extend this stabilizer formalism to include a broader class of codes that utilize shared entanglement.

The standard approach to QEC finds encoding and recovery procedures that ensure perfect recovery of quantum states passing through sufficiently weak noisy channels. The disadvantage of this approach is that it may fail when the noise strength, or error probability, is too high. Another approach to QEC is to tailor QECCs to a particular model of realistic noise, derived from a detailed model or from experimental measurements [18–20], for example, via quantum process tomography [21–27]. The tailored approach to QEC attempts to find encoding and recovery operations which are *optimized* relative to the particular experimentally measured noise or assumed noise model, in the sense that they maximize the fidelity (or minimize the distance) of the encoded output state relative to the input state [28–39]. Both worst-case and average-case performance have been considered in this setting, which usually involves numerical optimization. The advantage of this approach is that it tends to be more robust to variations in the noise strength than the standard approach [35]. Here we further develop the optimized QEC approach by considering EAQEC as an optimization problem.

To formulate the problem for optimization, we break down quantum error correction into three stages. First, the quantum state to be protected is encoded. This procedure appends some number of ancilla qubits in a fixed initial state and then

applies an encoding unitary to the information and ancilla qubits together. These qubits then pass through a noisy channel (assumed to be known). At the receiver's side, more ancillas can be appended, and then a decoding unitary is applied. All the ancillas (encoding or recovery) are discarded, and the state of the information qubits is the output.

In standard quantum error correction, the encoding and recovery ancillas always start in a standard state (usually $|0\rangle$) and are not entangled with each other. Previous work has shown that regular recovery ancillas are redundant in increasing the error-correction fidelity [35]. That is, given the same number of information qubits and encoding ancillas, the presence or absence of recovery ancillas makes no difference to the channel fidelity. This matches the properties of stabilizer codes, where decoding can be done unitarily, and there is no benefit in including recovery ancillas.

To formulate the problem for entanglement-assisted codes, we change the initial state of the ancillas so that noiseless entanglement is shared between the encoding and recovery ancilla qubits. We show that this additional information can enhance the performance of the recovery ancillas and increase the error-correction fidelity in most cases. This initial shared entanglement can in principle improve performance in at least two different ways. It can allow the code to transfer quantum information about the initial state to the recovery block, improving the channel fidelity. This fidelity increase can, in fact, lead to perfect error correction for an important class of error channels (the random unitary channels). The error-correction procedure for this class of channels can be interpreted as teleportation, with only classical information passing through the noisy channel.

However, even in cases where the error-correction protocol is not a form of teleportation, we can still get an improvement in performance. One interpretation of this is that entangled ancillas allow the recovery procedure to extract more information about the errors. This is analogous to superdense coding, where the use of entanglement boosts the rate of classical communication. The fidelity increase here shows the importance of relevant information stored in the ancillas in quantum error correction. The entangled recovery ancillas are fundamentally different than the regular recovery ancillas considered in previous optimization problems that gave no benefit [35].

The optimization procedure here, similar to previous work in this area [28–39], assumes that we know the noise channel. We assume that a channel identification procedure, such as quantum process tomography [18–27], has been performed prior to the error correction. It is also possible to optimize performance in cases where the exact channel model is not known, though we do not consider that here.

The structure of the article is as follows. In Sec. II we formulate the optimization problem in terms of a distance between states that needs to be minimized. The problem is bi-convex in the encoding and recovery operations, and in Secs. III and IV we consider how to optimize the encoding for a given recovery and recovery for a given encoding, respectively. In Sec. V we consider the case of random unitary channels and analyze a simple example for which EAQEC has perfect fidelity, via teleportation. We present numerical examples in Sec. VI, where we contrast different correction

and optimization scenarios. These examples illustrate the improved performance of optimized EAQECCs in a setting where perfect fidelity cannot be achieved. We conclude in Sec. VII. Certain technical details are presented in the Appendix.

II. PROBLEM FORMULATION

A quantum dynamical process is a map from an initial state ρ to a final state $\mathcal{E}(\rho)$. In general, the map $\mathcal{E}$ does not represent unitary evolution. However, if we assume a larger closed quantum system that includes the environment (or bath) B of the system, the total evolution becomes unitary. If the initial system-bath state is ρ_{SB}, a quantum channel can be written as

$$\mathcal{E}(\rho) = \text{Tr}_B[U_{SB}\rho_{SB}U_{SB}^\dagger], \qquad (1)$$

where $\rho = \text{Tr}_B[\rho_{SB}]$ is the initial system density matrix, U_{SB} is a unitary operator acting on the joint system-bath Hilbert space $\mathcal{H} = \mathcal{H}_S \otimes \mathcal{H}_B$, $\chi_B = \text{Tr}_S[\rho_{SB}]$ is the initial state of the environment, and Tr_B (Tr_S) denotes the partial trace over the environment (system). In general, a quantum dynamical process described by Eq. (1) is a Hermitian map (i.e., it maps Hermitian operators to Hermitian operators) [40]. When the initial system-bath state is purely classically correlated (formally, has vanishing quantum discord [41]), $\mathcal{E}$ becomes a completely positive (CP), trace-preserving map [42]. This condition is also necessary [40]. In this case the evolution can be represented using the Kraus operator sum representation (OSR) [43–45],

$$\mathcal{E}(\rho) = \sum_i K_i \rho K_i^\dagger, \qquad (2)$$

where the operators K_i, known as Kraus operators, satisfy the normalization condition $\sum_i K_i^\dagger K_i = I$ (identity). Such maps are powerful tools which can represent different types of noise processes that could be acting on the system.

We assume that the error-correction procedures of encoding and recovery, as well as the noisy channel, can all be described using Kraus operators. The entire procedure on the system can then be represented as

$$\rho_0 \xrightarrow{\mathcal{C}} \rho_C \xrightarrow{\mathcal{E}} \rho_E \xrightarrow{\mathcal{R}} \rho_R, \qquad (3)$$

where $\mathcal{C}$, $\mathcal{E}$, and $\mathcal{R}$ are the encoding, error, and recovery operations, respectively. The density matrices ρ_0 through ρ_R are the states of the data qubits and the ancillas, that is, represent the state of the entire system at different points in time. The initial system state $\rho_0 : \mathcal{H}_S \mapsto \mathcal{H}_S$, where $\mathcal{H}_S$ is the total system Hilbert space, is taken to be a product state between the data qubits, represented by ρ_{dat}, and the ancillas, represented by ρ_{anc}; that is, $\rho_0 = \rho_{\text{dat}} \otimes \rho_{\text{anc}}$. Thus, $\mathcal{H}_S = \mathcal{H}_{\text{dat}} \otimes \mathcal{H}_{\text{anc}}$, where $\mathcal{H}_{\text{dat}}$ and $\mathcal{H}_{\text{anc}}$ are, respectively, the Hilbert space of the data qubits and of the ancillas. The ancilla Hilbert space further decomposes into an encoding subspace $\mathcal{H}_{\text{enc}}$ and a recovery subspace $\mathcal{H}_{\text{rec}} : \mathcal{H}_{\text{anc}} = \mathcal{H}_{\text{enc}} \otimes \mathcal{H}_{\text{rec}}$. We write $d = \dim(\mathcal{H}_S)$ and otherwise denote the respective dimensions of the various subspaces by

$$d_X = \dim(\mathcal{H}_X), \quad X \in \{\text{dat,anc,enc,rec}\}. \qquad (4)$$

Note that

$$d = d_{\text{dat}} d_{\text{anc}}, \quad d_{\text{anc}} = d_{\text{enc}} d_{\text{rec}}. \qquad (5)$$

For most of our discussion we assume that all encoding ancillas are pairwise maximally entangled with recovery ancillas prior to the encoding operation, whence $d_{\text{enc}} = d_{\text{rec}}$, and $\rho_{\text{anc}} = (|B\rangle \langle B|)^{\otimes d_{\text{enc}}}$, where $|B\rangle$ is a maximally entangled state between encoding and recovery ancilla pairs. Nevertheless, for generality we leave d_{enc} and d_{rec} unless it is necessary to be more specific. Later, in our numerical examples, we also consider regular (unentangled) encoding and recovery ancillas.

Using the OSR for each of these operations, the overall evolution is

$$\rho_R = \mathcal{R} \circ \mathcal{E} \circ \mathcal{C}(\rho_0) = \sum_{c,e,r}(R_r E_e C_c)\rho_0(R_r E_e C_c)^\dagger, \quad (6)$$

where $\{C_c\}_{c=1}^{m_C}$, $\{E_e\}_{e=1}^{m_E}$, and $\{R_r\}_{r=1}^{m_R}$ are the Kraus operators of the encoding, error, and recovery channels, respectively, all satisfying the normalization condition

$$\sum_c C_c^\dagger C_c = \sum_e E_e^\dagger E_e = \sum_r R_r^\dagger R_r = I. \qquad (7)$$

All the Kraus operators are here represented by square $d \times d$ matrices, that is, are operators on $\mathcal{H}_S$. The number of Kraus operators for each map depends on the implementation and the basis representation, but the number m_E of error operators is bounded above by d^2 [44].

We assume that the encoding operation $\mathcal{C}$ is unitary, represented by a single unitary (Kraus) operator C. The encoded d-dimensional state is therefore

$$\rho_C = C(\rho_{\text{dat}} \otimes \rho_{\text{anc}})C^\dagger. \qquad (8)$$

Since, as illustrated in Fig. 1, the encoding does not act on the second half of the entangled pairs (the recovery ancillas), the encoding unitary C should be represented as $C' \otimes I_{\text{rec}}$, where C' is a $d_{C'} \times d_{C'}$ matrix and where $d_{C'} = d_{\text{dat}} d_{\text{enc}} = d_{\text{dat}}\sqrt{d_{\text{anc}}}$. We write I_X to denote the identity operator acting on a subspace
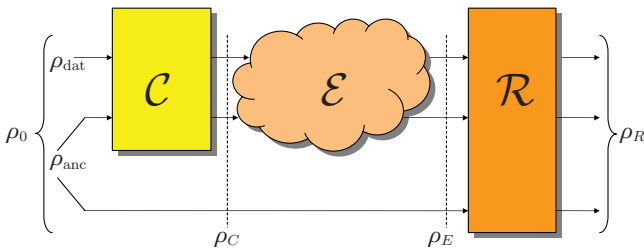

FIG. 1. (Color online) Scheme of entanglement-assisted error correction. The data qubits are represented by ρ_{dat}. The encoding and recovery ancillas are entangled before the encoding procedure $\mathcal{C}$ begins, and are represented by ρ_{anc}. The joint initial system state is $\rho_0 = \rho_{\text{dat}} \otimes \rho_{\text{anc}}$. The recovery ancillas pass through a noiseless channel to the recovery operation $\mathcal{R}$. The information qubits and encoding ancillas pass through a noisy channel $\mathcal{E}$. The output of the recovery procedure is ρ_R. The intermediate states ρ_C (output of the encoding) and ρ_E (output of the noise channel) are also indicated.

X, and reserve I for the identity operator on the entire system space $\mathcal{H}_S$.

Our goal is to implement a close approximation to a certain desired unitary L on the entire system space. More specifically, the purpose of optimization is to design the encoding $\mathcal{C}$ and recovery $\mathcal{R}$ for a given error channel $\mathcal{E}$ such that the overall map (6) $\rho_0 \mapsto \rho_R$ is as close as possible to some desired unitary $L : \mathcal{H}_S \mapsto \mathcal{H}_S$. This goal is different from that pursued in earlier optimization articles (e.g., [33–35]), where the goal was to obtain a good approximation to a unitary $L_{\text{dat}} : \mathcal{H}_{\text{dat}} \mapsto \mathcal{H}_{\text{dat}}$ acting only on the data qubits. Our present goal is general enough to encapsulate state preservation of the data qubits, for example, by implementing, with as high a fidelity as possible, a unitary which swaps the data qubits into the recovery qubits at the end of the process, while leaving the data and encoding qubits in a new pure state which is tensored with the recovery qubits. Indeed, this is essentially what happens in our teleportation example (Sec. V).

Channel fidelity and distance are typical measures of performance between two quantum channels [28,44,46,47]. The channel fidelity between the error-correction operation $\mathcal{R}\mathcal{E}\mathcal{C}$ and the desired unitary operation L is

$$f = \frac{1}{d_{\text{dat}}^2}\sum_{r,e,}|\text{Tr } L^\dagger R_r E_e C|^2. \qquad (9)$$

We wish to maximize this fidelity, which satisfies $0 \leqslant f \leqslant 1$. From [44, Theorem 8.2], $f = 1$ if and only if there are constants μ_{re} such that

$$R_r E_e C = \mu_{re} L, \quad \sum_{r,e}|\mu_{re}|^2 = 1. \qquad (10)$$

As shown in [35], maximizing the fidelity f is equivalent to minimizing the following distance, which is the approach we focus on henceforth:

$$\delta = \sum_{r,e}\|R_r E_e C - \mu_{re}L\|^2, \qquad (11)$$

where $\|X\| \equiv (\text{Tr}X^\dagger X)^{1/2}$ is the Frobenius norm. The minimization must now include the parameters $\{\mu_{re}\}$ as well. Using straightforward matrix manipulations, we can rewrite this distance as

$$\delta(R,C,\Delta) = \|RE(I_{m_E} \otimes C) - \Delta \otimes L\|^2, \qquad (12)$$

where $\Delta \equiv [\mu_{re}]$ is an $m_R \times m_E$ rectangular matrix, E is the $d \times m_E d$ rectangular error matrix built using the error Kraus operators as horizontally arranged blocks,

$$E = [E_1, \ldots, E_{m_E}], \qquad (13)$$

and R is the $m_R d \times d$ rectangular matrix obtained by vertically stacking the recovery Kraus operators R_r:

$$R = \begin{pmatrix} R_1 \\ \vdots \\ R_{m_R} \end{pmatrix}. \qquad (14)$$

Therefore, we have $\|\Delta\|^2 = \text{Tr } \Delta^\dagger \Delta = \sum_{r,e}|\mu_{re}|^2 = 1$ and $R^\dagger R = \sum_r R_r^\dagger R_r = I$.

Our optimization problem can now be summarized as follows:

minimize $\delta(R,C,\Delta)$ subject to the constraints

$$R^\dagger R = I, \quad C = C' \otimes I_{\text{rec}}, \quad C'^\dagger C' = I_{\text{dat,rec}}, \quad \|\Delta\|^2 = 1,$$
(15)

where $I_{\text{dat,rec}}$ denotes the identity operator on $\mathcal{H}_{\text{dat}} \otimes \mathcal{H}_{\text{rec}}$. There are three matrices of parameters (R,C,Δ) in this expression that are to be optimized. As stated, this optimization is not convex. Therefore, as first suggested by Reimpell and Werner [28], to identify the optimal values of these parameters, we solve this optimization problem iteratively. Starting with an arbitrary initial encoding operation, we first find the optimized recovery operation, which by itself is a convex problem; we then fix the recovery operation obtained in the previous step and find the optimized encoding operation, which is also a convex problem; and so on, alternately optimizing the recovery and encoding operations until the procedure converges. Δ is an intermediary matrix of parameters that must also be recalculated at each step. This iteration continues until the distance (11) stops decreasing. The details of the procedure at each step are provided in the next two sections.

The design of the optimized recovery operation at each step is not affected by the extra constraints added by introducing entanglement in the error-correction procedure. This is so because optimization of the recovery operation does not depend on the details of the encoding operation. However, the extra constraints play an important role in identifying the optimized encoding operation, which must be of the desired form, $C = C' \otimes I_{\text{rec}}$.

The encoding ancillas must be initialized in a maximally entangled state—for example, an Einstein-Podolsky-Rosen (EPR) pair. However, the distance measure defined in (11) is independent of the initial state of the ancilla qubits. To overcome this problem, we add an extra step, an *entangling* operation $U : \mathcal{H}_S \mapsto \mathcal{H}_S$, to the procedure. The entangling operation does not act on the data qubits, so it can be written as $U = I_{\text{dat}} \otimes U_{\text{anc}}$, where $U_{\text{anc}} : \mathcal{H}_{\text{anc}} \mapsto \mathcal{H}_{\text{anc}}$. By including this entangling operation in the evolution, the ancillas can be assumed to have the initial state $|0\rangle$.

III. OPTIMIZED ENCODING OPERATION

The derivation of the optimized encoding operator is similar to what was done previously in [35], but requires including the entangling operator U and a careful consideration of the subspaces involved. Including the entangling operator, the distance measure (11) can be rewritten as

$$
\begin{aligned}
\delta(R,C,\Delta) &= \|RE(I_{m_E} \otimes CU) - \Delta \otimes L\|^2 \\
&= \sum_{r,e} \|R_r E_e CU - \mu_{re} L\|^2 \\
&= \text{Tr}\left[I + C^\dagger C - UL^\dagger \sum_{r,e} \mu_{re}^* R_r E_e C - \text{H.c.} \right],
\end{aligned}
$$
(16)

where in the second line we used the relation between Eqs. (11) and (12) and in the third line we used the definition of

the Frobenius norm and the normalization conditions (7) and (10).

As a first step we need to find the unconstrained minimum of $\delta(R,C,\Delta)$ with respect to C; that is, we need to find the solution to $\partial\delta/\partial C = 0$ without introducing the condition $C^\dagger C = I$. To this end we note the following matrix differentiation identities, valid for any pair of matrices A and Z with compatible dimensions, with A being independent of the elements of Z:

$$\frac{\partial}{\partial Z}\text{Tr}[AZ] = A^t, \quad \frac{\partial}{\partial Z}\text{Tr}[AZ^\dagger] = 0, \quad (17)$$

$$\frac{\partial}{\partial Z}\text{Tr}[ZZ^\dagger] = Z^*, \quad \frac{\partial}{\partial Z}\text{Tr}[AZ^\dagger Z] = Z^* A^t, \quad (18)$$

$$\frac{\partial}{\partial Z}\text{Tr}[A(Z \otimes I)] = (\text{Tr}_2 A)^t, \quad (19)$$

where in the last identity the partial trace is over the subspace acted on by the identity operator. We prove all these identities in the Appendix.

Therefore, assuming that the recovery operators R_r are fixed, we have

$$\frac{\partial\delta}{\partial C} = C^* - \left(UL^\dagger \sum_{r,e} \mu_{re}^* R_r E_e \right)^t, \quad (20)$$

so that the solution to the unconstrained optimization problem $\frac{\partial\delta}{\partial C} = 0$ is

$$\bar{C} = \sum_{r,e} \mu_{re}(R_r E_e)^\dagger L U^\dagger. \quad (21)$$

We wish to impose the additional structure $C = C' \otimes I_{\text{rec}}$, that is, $C' = \frac{1}{d_{\text{rec}}}\text{Tr}_{\text{rec}}C$, where the partial trace is over the recovery ancillas. Thus, the unconstrained solution over $\mathcal{H}_{\text{dat}} \otimes \mathcal{H}_{\text{enc}}$ becomes

$$
\begin{aligned}
\bar{C}' &= \arg\min_{C'} \delta(R, C' \otimes I_{\text{rec}}, \Delta) \\
&= \frac{1}{d_{\text{rec}}}\text{Tr}_{\text{rec}}\left[\sum_{r,e} \mu_{re}(R_r E_e)^\dagger L \left(I_{\text{dat}} \otimes U_a^\dagger\right) \right].
\end{aligned}
$$
(22)

The reader can check that this result can also be derived by introducing the decomposition $C = C' \otimes I_{\text{rec}}$ directly into Eq. (16) and using identity (19), that is,

$$\frac{\partial\delta}{\partial C'} = d_{\text{rec}}C'^* - \left(\text{Tr}_{\text{rec}}\left[UL^\dagger \sum_{r,e} \mu_{re}^* R_r E_e \right] \right)^t. \quad (23)$$

This is important for the next step.

The constrained problem (15) requires that $C'^\dagger C' = I_{\text{dat,rec}}$. To solve this problem, we form the Lagrangian

$$\mathcal{L} = \delta(R,C',\Delta) + \text{Tr}[P(C'^\dagger C' - I_{\text{dat,rec}})], \quad (24)$$

where P is a Hermitian Lagrange multiplier matrix (because the constraint $C'^\dagger C' = I_{\text{dat,rec}}$ is Hermitian). Setting $\partial_{C'}\mathcal{L} = 0$, we find, using Eqs. (18), (22), and (23): $C'^* - \bar{C}'^* + C'^* P^t = 0$; that is, the solution to the constrained problem is

$$C' = \bar{C}'(I_{\text{dat,rec}} + P)^{-1}. \quad (25)$$

To eliminate P, we note that the constraint $C'^\dagger C' = I_{\text{dat,rec}}$ now implies $\bar{C}'^\dagger \bar{C}' = (I_{\text{dat,rec}} + P)^2$, from which we have

$(I_{\text{dat,rec}} + P)^{-1} = (\bar{C}'^{\dagger}\bar{C}')^{-1/2}$. Thus, finally,

$$C' = \bar{C}'(\bar{C}'^{\dagger}\bar{C}')^{-1/2}, \qquad (26)$$

where $\bar{C}'$ is given in Eq. (22). This encoding operation is optimized for the given recovery operation R.

IV. OPTIMIZED RECOVERY OPERATION

Unlike the case of the optimized encoding operation, we do not impose any specific tensor product structure on the recovery, so that derivation of the optimized recovery operator is essentially the same as what was done previously in [35], except that again we must include the entangling operator. Here we briefly outline the steps required to optimize the recovery operator. Using the constraints in (15) and fixing the encoding operation C, we express the distance (11) as

$$\delta = \|RE(I_{m_E} \otimes CU) - \Delta \otimes L\|^2$$
$$= 2d - 2\text{Re}\,\text{Tr}\,RE(\Delta^{\dagger} \otimes CUL^{\dagger}). \qquad (27)$$

Minimizing (27) with respect to R is the same as maximizing the last term. As shown in [35, Appendix], this maximization yields

$$\max_{R^{\dagger}R=I} \text{Re}\,\text{Tr}\,RE(\Delta^{\dagger} \otimes CUL^{\dagger})$$
$$= \text{Tr}\sqrt{E(\Gamma \otimes I)E^{\dagger}}, \qquad (28)$$

where the matrix Γ is defined as

$$\Gamma = \Delta^{\dagger}\Delta. \qquad (29)$$

The constraint $\|\Delta\|^2 = 1$ from (15) is equivalent to $\text{Tr}\,\Gamma = 1$, and $\Gamma \geqslant 0$ by definition. Therefore, the optimization problem for Γ is

$$\text{maximize} \quad \text{Tr}\sqrt{E(\Gamma \otimes I)E^{\dagger}}$$
$$\text{subject to} \quad \Gamma \geqslant 0, \quad \text{Tr}\,\Gamma = 1. \qquad (30)$$

The optimized Γ can be obtained by solving an equivalent semidefinite programming (SDP) problem, or by performing a constrained least-squares optimization [35]. The matrix Δ can be obtained from Γ by the definition (29) [see Eq. (21) of Ref. [35] for details]. Once the optimized Δ is known, the optimized recovery matrix can be found from

$$R = [v_1, \ldots, v_d][u_1, \ldots, u_d]^{\dagger}, \qquad (31)$$

where $\{v_i, u_i\}_{i=1}^{d}$ are, respectively, the right and left singular vectors in the singular value decomposition of the matrix $E(\Delta^{\dagger} \otimes CUL^{\dagger})$, with the singular values in descending order.

The following algorithm summarizes the preceding method for encoding and recovery optimization in the presence of entanglement between the encoding and recovery blocks:

Initialize C
Repeat
  (a) Optimal recovery
    maximize (30) to find Γ,
    solve for Δ via (29),
    solve for R via (31),
  (b) Optimal encoding
    solve for C via (26),
Until distance (11) stops decreasing.

Since in each iteration of this algorithm the distance δ only decreases, the converged solution to this optimization is guaranteed to be at least a locally optimal solution to (15). We apply this procedure to a selection of different channels in what follows to assess its performance in practice.

V. RANDOM UNITARY CHANNELS

An error channel $\mathcal{E}$ is called a *random unitary channel* if we can decompose it into the probabilistic application of one of a finite set of unitary operations,

$$\mathcal{E}(\rho) = \sum_{i=1}^{n} p_i V_i \rho V_i^{\dagger}, \qquad (32)$$

where the V_i are unitary operators and $p_i, i = 1, \ldots, n$, is a probability distribution. Random unitary channels describe noise processes that can be corrected using classical information extracted from the environment [48]. It is clear that any random unitary channel should be unital, meaning that $\mathcal{E}(I) = I$. However, the inverse relation holds only for channels on qubits and is not true for higher dimensions [49,50].

There are two operations in (1) that are not random unitary. These operations are the introduction of the ancillas and the partial trace over the environment. Therefore, an error channel need not be random unitary, of form (32), in general. The necessary and sufficient conditions for a channel to be a random unitary are discussed in [51]. In [52], an upper bound was found on the number of unitaries needed in (32). Here we show that if the number of unitary operators required in (32) is at most two, our optimized error-correction method can perfectly correct the error using only one EPR pair as ancilla qubits.

In fact, for this case, the optimized encoding and recovery operations are known and are equivalent to teleportation. To see this, consider a channel that can be decomposed into two unitaries, V_1 and V_2, as follows:

$$\mathcal{E}(\rho) = (1 - p)V_1 \rho V_1^{\dagger} + p V_2 \rho V_2^{\dagger}. \qquad (33)$$

Suppose Alice wants to send Bob one qubit of data through this noisy channel. One pair of entangled qubits enables Alice and Bob to make this communication error free. Assume this pair of qubits was prepared by an entanglement source and sent to Alice and Bob, each taking one of the qubits. Alice can apply an encoding operation on her qubits—the data qubit and her half of the ebit—before sending them to Bob through the noisy channel. Bob applies the recovery operation to all three qubits once he receives the message.

The optimized encoding operation that Alice uses in this scenario is done by first applying $V_1^{\dagger}$ to her qubits (recall that we assume that the channel is known) to make the error channel equivalent to

$$\mathcal{E}(\rho) = (1 - p)I\rho I + p V_2 V_1^{\dagger} \rho V_1 V_2^{\dagger}. \qquad (34)$$

Alice continues the encoding operation by applying

$$C = \tfrac{1}{2}(|v_1 v_1\rangle\langle B_1| + |v_1 v_2\rangle\langle B_2| + |v_2 v_1\rangle\langle B_3| + |v_2 v_2\rangle\langle B_4|), \qquad (35)$$

where v_1 and v_2 are the eigenvectors of the unitary $V_2 V_1^{\dagger}$, and B_{1-4} are an orthonormal basis of maximally entangled
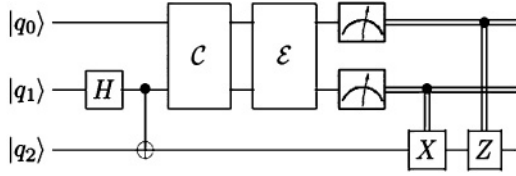
FIG. 2. Quantum error correction using teleportation. The information qubit $|q_0\rangle$ is recovered with unity fidelity in spite of passing through the noise channel $\mathcal{E}$. The first two operations (Hadamard gate on $|q_1\rangle$ and controlled-NOT between $|q_1\rangle$ and $|q_2\rangle$) represent the entangling operation. This is followed by the encoding operation $\mathcal{C}$ described by Eq. (35). The last two operations represent conditional X and Z gates applied to $|q_2\rangle$, dependent on the measurement outcomes of $|q_0\rangle$ and $|q_1\rangle$. See the text for additional details.

states—for example, the Bell states $(|00\rangle \pm |11\rangle)/\sqrt{2}$ and $(|10\rangle \pm |01\rangle)/\sqrt{2}$. This encoding enables us to send two bits of classical information safely through the noisy channel. These two bits of classical information, together with the second half of the ebit, enable us to recover the original state.

To see this in more detail, suppose that the initial state of the data qubit is $|\phi_i\rangle = a|0\rangle + b|1\rangle$. The state of the entire system after applying the entangling unitary is

$$|\psi_i\rangle = \frac{1}{\sqrt{2}}[a(|000\rangle + |011\rangle) + b(|100\rangle + |111\rangle)]. \quad (36)$$

After applying the encoding (35) and error map on the first two qubits, the final state of the entire system is

$$|\psi_{\text{tot}}\rangle = \frac{1}{2\sqrt{2}}[|v_1 v_1\rangle(a|0\rangle + b|1\rangle) + |v_1 v_2\rangle(a|0\rangle - b|1\rangle)$$
$$+ |v_2 v_1\rangle(a|1\rangle + b|0\rangle) + |v_2 v_2\rangle(a|1\rangle - b|0\rangle)]. \quad (37)$$

This is the state of the system once Bob receives it.

At this point, as shown in Fig. 2, he can recover the original state $|\phi_i\rangle$ by measuring the state of the first two qubits. Based on the result of the measurement he applies the appropriate Pauli operator to the last qubit, his half of the ebit, to recover the original state. For example, if he measures the first two qubits to be $|v_2 v_1\rangle$, he should apply σ_x to the last qubit to recover the original state.

VI. EXAMPLES

We present three noise channels as examples: bit flip, bit-phase flip, and depolarizing. We note that in all three cases the noisiest point does not occur at the highest noise rate point ($p = 1$) but rather at some intermediate noise rate value: $p = 1/2$, $p = 2/3$, $p = 3/4$, respectively. Above this value each channel is qualitatively different from the same channel below this value.

A. The bit-flip channel

As an example of a random unitary channel with two unitary operators, we consider the bit-flip-error channel, or binary symmetric channel, in which bit-flip errors occur independently with probability p on each qubit. This channel can be represented as

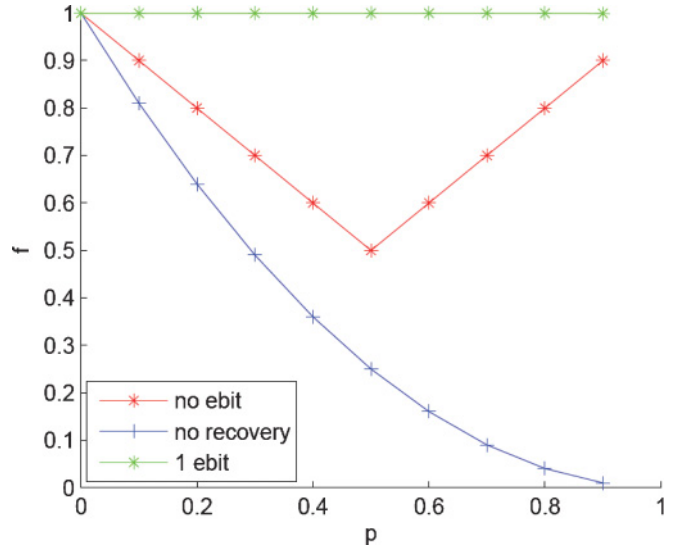$$\mathcal{E}(\rho) = (1 - p)I\rho I + p\sigma_x \rho \sigma_x, \quad (38)$$

FIG. 3. (Color online) Optimized fidelity for the bit-flip channel. Results are shown for the case of one data qubit. The blue (bottom) curve is the fidelity without any recovery. The red (middle) curve is the optimized fidelity obtained for encoding using a single encoding qubit but for recovery without entanglement, that is, with a single recovery qubit in a product state with the encoding qubit. This represents the standard scenario of (channel-optimized) QEC. The green (top) curve is the fidelity obtained in the presence of a single ebit, that is, the EAQEC scheme shown in Fig. 1. The fidelity in this case is unity for all values of the bit-flip probability p, showing that our optimization recovers the perfect encoding and recovery protocol of teleportation discussed in the text. We note that above $p = 1/2$, a bit-flip channel is actually an X gate followed by a bit-flip channel with rate $1 - p$, so that in this sense the monotonically decreasing fidelity in the case without recovery can be viewed as artificial for $p > 1/2$; that is, the noisiest point is at $p = 1/2$.

where henceforth $\sigma_x, \sigma_y, \sigma_z$ denote the standard 2×2 Pauli matrices. As shown in Fig. 3, this error can be corrected using an ebit as the error-correcting resource. The ebit again is shared between the encoding and recovery parts. The encoding acts on the initial data qubit and the first entangled qubit (the encoding qubit). Considering that the error is already in the form (34), we do not need to apply the initial unitary operator here. Therefore, the optimized encoding operator based on (35) is

$$C = |++\rangle\langle B_1| + |+-\rangle\langle B_2| + |-+\rangle\langle B_3| + |--\rangle\langle B_4|, \quad (39)$$

where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, and B_1–B_4 form an orthonormal basis of maximally entangled states. The first entangled qubit together with the initial data qubit are corrupted by the bit-flip error, while the other entangled qubit (the recovery qubit) stays intact. The recovery operation acts on all the qubits and can reproduce the initial state by measuring the first two qubits as discussed earlier, followed by a SWAP between the recovery qubit and the data qubit, which perfectly restores the state of the data qubit.

While the optimization procedure we employ is not guaranteed to find the global maximum of the channel fidelity, we
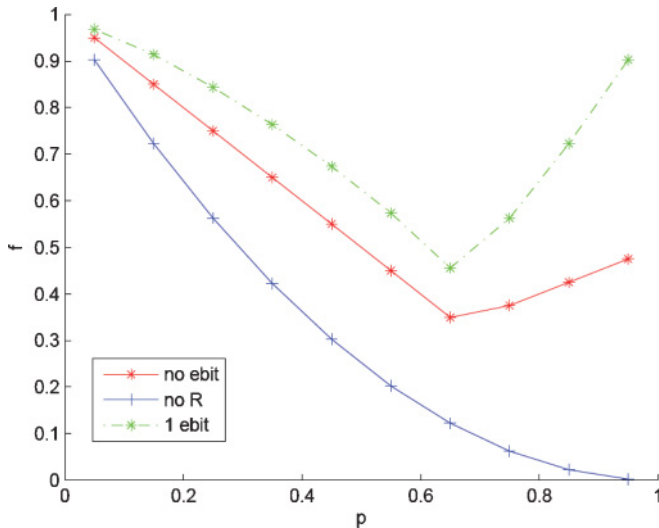
FIG. 4. (Color online) Optimized fidelity for the bit-phase-flip channel. As in Fig. 3 there is a single data qubit going through the noise channel unprotected (blue, bottom curve), optimally protected by a single encoding and a single recovery qubit in a product state (red, middle curve), and optimally protected by an ebit (green, top curve).
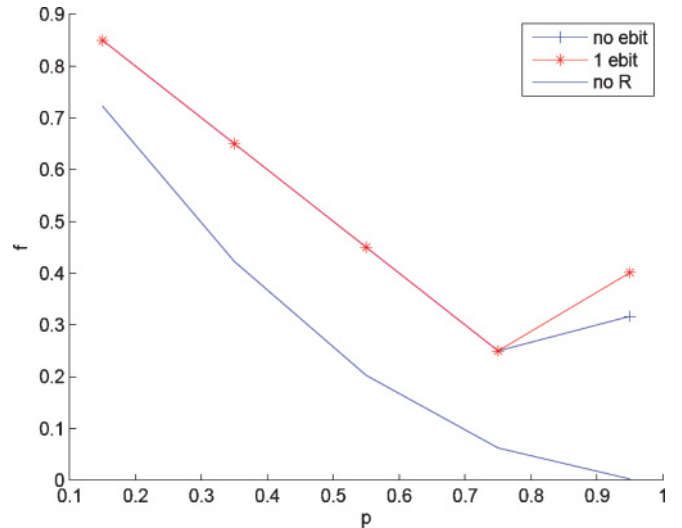

FIG. 5. (Color online) Optimized fidelity for the depolarizing channel. As in Figs. 3 and 4, there is a single data qubit going through the noise channel unprotected (blue, bottom curve), optimally protected by a single encoding and a single recovery qubit in a product state (blue, middle curve, overlapping red for $p \leqslant 3/4$), and optimally protected by an ebit (red, top curve).

see that in this case it does, indeed, find the optimal solution. Convex optimization has rediscovered the protocol of quantum teleportation.

B. The bit-flip–phase-flip channel

In this example both bit-flip and phase-flip errors occur with equal probabilities:

$$\mathcal{E}(\rho) = (1 - p)I\rho I + \frac{p}{2}\sigma_x\rho\sigma_x + \frac{p}{2}\sigma_z\rho\sigma_z. \qquad (40)$$

The error occurs independently on different qubits. This error channel clearly cannot be represented in the form of (33). Hence, our teleportation argument does not apply here. There is no choice of basis for which classical information can be sent through this channel error free.

While entanglement cannot enable perfect error correction in this case, it can still increase the useful information about the errors available to the recovery block and therefore increase the fidelity of the error correction. Figure 4 presents the fidelity of error correction for this channel for different values of p. The only difference between the upper and middle curves is the presence of entanglement. Without entanglement, the extra qubit used in the recovery acts as a regular recovery ancilla and therefore does not increase the fidelity. The minimum fidelity ("noisiest point") in both cases occurs at $p = 2/3$, where the channel is symmetric in I, σ_x, and σ_z : $\mathcal{E}(\rho) = (1/3)(I\rho I + \sigma_x\rho\sigma_x + \sigma_z\rho\sigma_z)$.

C. The depolarizing channel

In this example we consider a particularly simple error model that is widely used to describe noisy quantum systems: the depolarizing channel. This important channel is often considered to be the quantum equivalent of the binary symmetric channel in classical error correction. It occurs when

the state of the system can be completely mixed by the action of the channel. The noise is unbiased in the sense that it generates bit-flip errors, phase-flip errors, or both with equal probability; this is represented by the map

$$\mathcal{E}(\rho) = (1 - p)\rho + \frac{p}{3}(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z). \qquad (41)$$

The depolarizing channel shrinks the radius of the Bloch sphere by a factor of $1 - p$, while preserving its shape [44].

Interestingly, for this error channel the ebit does *not* increase the fidelity for $p < 3/4$. The optimized fidelity of the channel with and without using an ebit is presented in Fig. 5. A break point occurs at $p = 3/4$ (the noisiest point), when the coefficients in (41) become equal. The fidelities of both cases are the same for $p < 3/4$, but the ebit can slightly increase the fidelity for $p > 3/4$.

An observation that may explain this difference is that for $p < 3/4$ we can write $\mathcal{E}(\rho) = (1 - 4p/3)\rho + (4p/3)\mathcal{T}(\rho)$, where $\mathcal{T}$ is the *twirling* operation,

$$\mathcal{T}(\rho) = \frac{1}{4}(\rho + \sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z) = \frac{1}{2}I, \quad (42)$$

for any valid density matrix ρ. In other words, with probability $4p/3$ the channel replaces the qubit with a maximally mixed qubit. For $p > 3/4$ this interpretation is no longer possible, and it is in this regime that shared entanglement seems to give improvement.

VII. CONCLUSION

Previous work on optimized quantum error correction [28–39] considered the "standard," entanglement-unassisted setting. In this work we addressed, for the first time, the problem of optimized EAQEC. Our methodology is a generalization of the optimization approach developed in Ref. [35], which involves a biconvex optimization problem, iterating

between encoding and recovery. Our main technical innovation appears in the encoding step, where we introduced a method to account for the constraint of initial entanglement between the encoding and recovery ancillas.

Our method returns the optimized channel fidelity for a given noise channel in the presence of shared entanglement between the encoding and recovery block. We showed that entanglement can substantially increase the error-correction fidelity. For specific error channels this fidelity increase can lead to perfect error correction. In these cases, the error-correction procedure can be interpreted as quantum teleportation. For cases in which perfect error correction is not possible, this teleportation interpretation does not apply; however, our optimization method returns the optimized fidelity for these channels, in most cases showing improvement with the resource of shared entanglement.

For one channel we found no improvement: the depolarizing channel. For the small codes that we considered in this article, optimized standard quantum error correction performed just as well as optimized EAQEC on this channel. It is possible that this equivalence between entanglement-assisted and standard quantum error correction will not hold for larger code words on the depolarizing channel. Larger code words would use either more ancillas or more ebits in the encoding process. The possibility that larger codes might show a greater difference is suggested by the fact that a three-qubit code word with two ebits exists that can correct an arbitrary single-qubit error [11], while the smallest standard code that corrects an arbitrary error is of size five [53]. However, the performance of *optimized* quantum error correction for larger code words remains a subject for future research. Larger code words make for a more numerically intensive optimization procedure.

Considering that regular recovery ancillas are redundant in the optimized error-correction procedure, the results of this article show that information about errors in the encoded data, transferred to the recovery block through entanglement with the encoded state, can increase the fidelity for many channels. We expect that results for a wider variety of error channels and larger code words will continue to bear this out.

ACKNOWLEDGMENTS

APPENDIX: TRACE DIFFERENTIATION IDENTITIES

We prove Eqs. (17) and (18). These are one-line proofs which follow directly from the definition of differentiation by a matrix [54, Sec. 10.6]: The derivative $\frac{\partial}{\partial Z} f(Z)$ of a differentiable scalar-valued function $f(Z)$ of a matrix argument $Z \equiv [z_{ij}] \in \mathbb{C}^{m \times n}$ is the $m \times n$ matrix whose (i,j) entry is $\frac{\partial}{\partial z_{ij}} f(Z)$:

$$\frac{\partial}{\partial Z} \mathrm{Tr}[AZ] = \left[\frac{\partial}{\partial z_{ij}} \sum_{ij} z_{ij} a_{ji} \right] = [a_{ji}] = A^t, \quad \text{(A1)}$$

$$\frac{\partial}{\partial Z} \mathrm{Tr}[AZ^\dagger] = \left[\frac{\partial}{\partial z_{ij}} \sum_{ij} a_{ij} z_{ij}^* \right] = 0, \quad \text{(A2)}$$

$$\frac{\partial}{\partial Z} \mathrm{Tr}[ZZ^\dagger] = \left[\frac{\partial}{\partial z_{ij}} \sum_{ij} z_{ij} z_{ij}^* \right] = [z_{ij}^*] = Z^*, \quad \text{(A3)}$$

$$\frac{\partial}{\partial Z} \mathrm{Tr}[AZ^\dagger Z] = \left[\frac{\partial}{\partial z_{ij}} \sum_{ij} z_{ij} (AZ^\dagger)_{ji} \right] = [(AZ^\dagger)_{ji}] = Z^* A^t,$$

$$\text{(A4)}$$

where for the second identity we used the fact that for a complex variable $z = x + iy$ $(x, y \in \mathbb{R})$ it holds that $\partial z^* / \partial z = \partial (x - iy)/\partial x - i \partial (x - iy)/\partial y = 0$.

Next we prove Eq. (19). Let

$$A = \sum a_{mnpq} |m\rangle\langle n| \otimes |p\rangle\langle q|, \quad \text{(A5)}$$

$$\tilde{Z} = Z \otimes I = \sum z_{m'n'} |m'\rangle\langle n'| \otimes |p'\rangle\langle p'|. \quad \text{(A6)}$$

Then

$$\mathrm{Tr}A\tilde{Z} = \mathrm{Tr} \sum a_{mnpp'} z_{nn'} |m\rangle\langle n'| \otimes |p\rangle\langle p'|$$
$$= \sum_{mnp} a_{mnpp} z_{nm}, \quad \text{(A7)}$$

so that

$$\frac{\partial}{\partial Z} \mathrm{Tr}[A(Z \otimes I)] = \left[\frac{\partial}{\partial z_{mn}} \sum_{mnp} a_{mnpp} z_{nm} \right]$$
$$= \left[\sum_p a_{nmpp} \right] = (\mathrm{Tr}_2 A)^t, \quad \text{(A8)}$$

where the partial trace is over the subspace acted on by the identity operator in $Z \otimes I$.

---

[1] P. W. Shor, Phys. Rev. A **52**, R2493 (1995).

[2] A. M. Steane, Phys. Rev. Lett. **77**, 793 (1996).

[3] D. Gottesman, Phys. Rev. A **54**, 1862 (1996).

[4] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).

[5] D. Aharonov and M. Ben-Or, SIAM J. Comput. **38**, 1207 (2008).

[6] A. Yu. Kitaev, Russ. Math. Surv. **52**, 1191 (1996).

[7] D. Gottesman, Phys. Rev. A **57**, 127 (1998).

[8] J. Preskill, Proc. R. Soc. London A **454**, 385 (1998).

[9] C. H. Bennett, D. P. DiVincenzo, J. Smolin, and W. Wootters, Phys. Rev. A **54**, 3824 (1996).

[10] G. Bowen, Phys. Rev. A **66**, 052313 (2002).

[11] T. Brun, I. Devetak, and M.-H. Hsieh, Science **314**, 436 (2006).

[12] T. Brun, I. Devetak, and M.-H. Hsieh, e-print arXiv:quant-ph/0608027.

[13] M.-H. Hsieh, I. Devetak, and T. Brun, Phys. Rev. A **76**, 062313 (2007).

[14] M. M. Wilde and T. A. Brun, Phys. Rev. A **77**, 064302 (2008).

[15] I. Kremsky, M.-H. Hsieh, and T. A. Brun, Phys. Rev. A **78**, 012341 (2008).

[16] M.-H. Hsieh, T. A. Brun, and I. Devetak, Phys. Rev. A **79**, 032340 (2009).

[17] M. M. Wilde and T. A. Brun, Phys. Rev. A **81**, 042333 (2010).

[18] J. B. Altepeter, D. Branning, E. Jeffrey, T. C. Wei, P. G. Kwiat, R. T. Thew, J. L. O'Brien, M. A. Nielsen, and A. G. White, Phys. Rev. Lett. **90**, 193601 (2003).

[19] Y. Weinstein, T. Havel, J. Emerson, N. Boulant, M. Saraceno, S. Lloyd, and D. Cory, J. Chem. Phys. **121**, 6117 (2004).

[20] M. Howard, J. Twamley, C. Wittman, T. Gaebel, F. Jelezko, and J. Wrachtrup, New J. Phys. **8**, 33 (2006).

[21] I. Chuang and M. Nielsen, J. Mod. Opt. **44**, 2455 (1997).

[22] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).

[23] G. M. D'Ariano and P. LoPresti, Phys. Rev. Lett. **86**, 4195 (2001).

[24] M. Mohseni and D. A. Lidar, Phys. Rev. Lett. **97**, 170501 (2006).

[25] J. Emerson, M. Silva, O. Moussa, C. Ryan, M. Laforest, J. Baugh, D. G. Cory, and R. Laflamme, Science **317**, 1893 (2007).

[26] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Phys. Rev. A **77**, 032322 (2008).

[27] M. Branderhorst, J. Nunn, I. Walmsley, and R. Kosut, New J. Phys. **11**, 115010 (2009).

[28] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94**, 080501 (2005).

[29] N. Yamamoto, S. Hara, and K. Tsumura, Phys. Rev. A **71**, 022322 (2005).

[30] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **75**, 012338 (2007).

[31] A. Fletcher, P. Shor, and M. Win, IEEE Trans. Inf. Theory **54**, 5705 (2008).

[32] A. S. Fletcher, P. W. Shor, and M. Z. Win, Phys. Rev. A **77**, 012320 (2008).

[33] R. Kosut and D. A. Lidar, Quantum Inf. Proc. **8**, 443 (2009).

[34] R. L. Kosut, A. Shabani, and D. A. Lidar, Phys. Rev. Lett. **100**, 020502 (2008).

[35] S. Taghavi, R. L. Kosut, and D. A. Lidar, IEEE Trans. Inf. Theory **56**, 1461 (2010).

[36] G. Balló and P. Gurin, Phys. Rev. A **80**, 012326 (2009).

[37] C. Bény and O. Oreshkov, Phys. Rev. Lett. **104**, 120501 (2010).

[38] H. K. Ng and P. Mandayam, Phys. Rev. A **81**, 062342 (2010).

[39] J. Tyson, J. Math. Phys. **51**, 092204 (2010).

[40] A. Shabani and D. A. Lidar, Phys. Rev. Lett. **102**, 100402 (2009).

[41] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2002).

[42] C. Rodríguez-Rosario, K. Modi, A.-M. Kuah, E. Sudarshan, and A. Shaji, J. Phys. A **41**, 205301 (2008).

[43] K. Kraus, *States, Effects and Operations, Fundamental Notions of Quantum Theory* (Academic, Berlin, 1983).

[44] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2000).

[45] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, Oxford, 2002).

[46] A. Gilchrist, N. K. Langford, and M. A. Nielsen, Phys. Rev. A **71**, 062310 (2005).

[47] D. Kretschmann and R. F. Werner, New J. Phys. **6**, 26 (2004).

[48] M. Gregoratti and R. Werner, J. Mod. Opt. **50**, 915 (2003).

[49] S. L. Tregub, Sov. Math. **30**, 105 (1986).

[50] B. Kümmerer and H. Maassen, Commun. Math. Phys. **109**, 1 (1987).

[51] K. M. R. Audenaert and S. Scheel, New J. Phys. **10**, 023011 (2008).

[52] F. Buscemi, Phys. Lett. A **360**, 256 (2006).

[53] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[54] D. S. Bernstein, *Matrix Mathematics* (Princeton University Press, Princeton, 2005).