Quantum Inf Process
DOI 10.1007/s11128-009-0120-2

Quantum error correction via convex optimization

Robert L. Kosut · Daniel A. Lidar

Abstract We show that the problem of designing a quantum information processing error correcting procedure can be cast as a bi-convex optimization problem, iterating between encoding and recovery, each being a semidefinite program. For a given encoding operator the problem is convex in the recovery operator. For a given method of recovery, the problem is convex in the encoding scheme. This allows us to derive new codes that are locally optimal. We present examples of such codes that can handle errors which are too strong for codes derived by analogy to classical error correction techniques.

Keywords Quantum error correction · Convex optimization

1 Introduction

Quantum error correction is essential for the scale-up of quantum information processing devices. A theory of quantum error correcting codes has been developed, in analogy to classical coding for noisy channels, e.g., [7,8,15,16]. Recently error correction design has been cast as an optimization problem with the design variables being the process matrices associated with the encoding and/or recovery channels [4,10,14,17].[1]

R. L. Kosut (✉)
SC Solutions, Sunnyvale, CA 94085, USA
e-mail: kosut@scsolutions.com

D. A. Lidar
Departments of Chemistry, Electrical Engineering, and Physics, Center for Quantum Information Science & Technology, University of Southern California, Los Angeles, CA 90089, USA
e-mail: lidar@usc.edu

Published online: 18 July 2009

Using fidelity measures leads naturally to a convex optimization problem, specifically a semidefinite program (SDP) [2]. The advantage of this approach is that noisy channels which do not satisfy the standard assumptions for perfect correction can be optimized for the best possible encoding and/or recovery.

Here we apply convex optimization via SDP, and iterate between encoding and recovery. For a given encoding operator the problem is convex in the recovery. For a given method of recovery, the problem is convex in the encoding. We further make use of Lagrange Duality to alleviate some of the computational burden associated with solving the SDP for the process matrices. The SDP formalism also allows for a robust design by enumerating constraints associated with different error models. We illustrate the approach with an example where the error system does not assume independent channels.

An intriguing prospect is to integrate the results found here within a complete "black-box" error correction scheme that takes quantum state (or process) tomography as input and iterates until it finds an optimal error correcting encoding and recovery. A similar idea was proposed in [3] for determining bang-bang control pulses.

2 Quantum error correction

2.1 Standard model

A *standard model* [13, §10.3] of an error correction system as shown in the block diagram of Fig. 1 is composed of three *quantum operations*: encoding $\mathcal{C}$, error $\mathcal{E}$, and recovery $\mathcal{R}$.

The input, ρ_S, is the $n_S \times n_S$ dimensional density matrix which contains the quantum information of interest and which is to be processed. We will refer to ρ_S as the *system state* or the *unencoded state*. The output of the encoding operation is ρ_C, the $n_C \times n_C$ dimensional *encoded state*. The error operator, which is also the source of decoherence, corrupts the encoded state and returns σ_C, the $n_C \times n_C$ "noisy" encoded state. Finally, ρ_R is the $n_R \times n_R$ dimensional *recovered state*. The objective considered here is to design $(\mathcal{C}, \mathcal{R})$ so that the map $\rho_S \to \rho_R$ is as close as possible to a desired $n_S \times n_S$ unitary L_S. Hence, ρ_R has the same dimension as ρ_S, that is, $n_R = n_S$. For emphasis we will replace ρ_R with $\hat{\rho}_S$.

Although it is possible for $\mathcal{E}$ to be non-trace preserving, in the model considered here, all three quantum operations are each characterized by a trace-preserving operator-sum-representation (OSR):

$$
\begin{aligned}
\rho_C &= \mathcal{C}(\rho_S) = \sum\nolimits_c C_c \rho_S C_c^\dagger, \ \sum\nolimits_c C_c^\dagger C_c = I_{n_S} \\
\sigma_C &= \mathcal{E}(\rho_C) = \sum\nolimits_e E_e \rho_C E_e^\dagger, \ \sum\nolimits_e E_e^\dagger E_e = I_{n_C} \\
\hat{\rho}_S &= \mathcal{R}(\sigma_C) = \sum\nolimits_r R_r \sigma_C R_r^\dagger, \ \sum\nolimits_r R_r^\dagger R_r = I_{n_C}
\end{aligned} \tag{1}
$$

[1] The current paper appeared as the preprint quant-ph/0606078. It was followed by [10], where we introduced the "indirect fidelity maximization method", which has some advantages over the "direct" method we use here. These advantages are explained in [10].

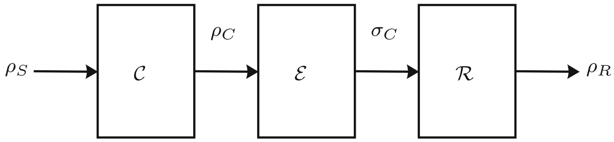

Fig. 1 Standard encoding-error-recovery model of an error correction system

with OSR matrix elements $C_c \in \mathbf{C}^{n_C \times n_S}$, $E_e \in \mathbf{C}^{n_C \times n_C}$, and $R_r \in \mathbf{C}^{n_S \times n_C}$. These engender a single trace-preserving quantum operation, $\mathcal{S}$, mapping ρ_S to $\hat{\rho}_S$,

$$\hat{\rho}_S = \mathcal{S}(\rho_S) = \sum_{r,e,c} S_{rec}\, \rho_S S_{rec}^{\dagger}, \quad S_{rec} = R_r E_e C_c \tag{2}$$

Before we describe our design approach we make a few remarks about the error source and implementation of the encoding and recovery operations.

2.2 Implementation

Any OSR can be equivalently expressed, and consequently physically implemented, as a unitary with ancilla states [13, §8.23]. An equivalent system-ancilla-bath representation of the standard error correction model of Fig. 1 is shown in the block diagram of Fig. 2.

For the encoding operator, $\mathcal{C}$, the *encoding ancilla state*, $|0_{CA}\rangle$, has dimension n_{CA}, and hence, the resulting encoded space has dimension $n_C = n_S\, n_{CA}$. The encoding operation is determined by U_C, the $n_C \times n_C$ unitary encoding operator which produces the encoded state $\sigma_C = U_C(\rho_S \otimes |0_{CA}\rangle\langle 0_{CA}|)U_C^{\dagger}$.

Here we will ignore complications associated with an infinite dimensional bath. The error system is thus equivalent to the $n_E \times n_E$ unitary error operator U_E with uncoupled inputs, the encoded state ρ_C, and ρ_B, the $n_B \times n_B$ bath state. Thus, $n_E = n_S n_{CA} n_B$.

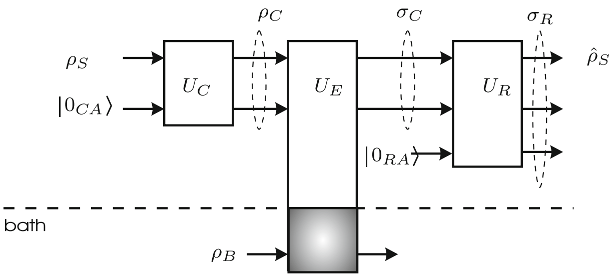

Fig. 2 System-ancilla-bath representation of standard encoding-error-recovery model of error correction system

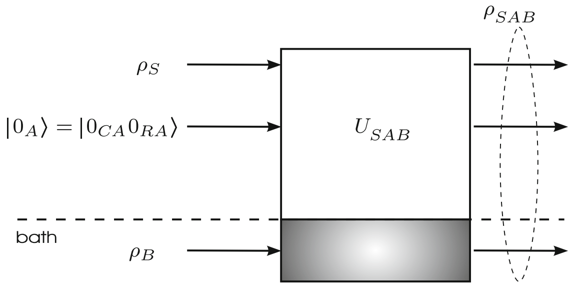

Fig. 3 System-ancilla-bath representation of error correction system

The noisy encoded state σ_C, is the $n_C \times n_C$ reduced state obtained by tracing out the bath from the output of U_E, that is, $\sigma_C = \mathbf{Tr}_B\ U_E(\rho_C \otimes \rho_B)U_E^\dagger$.

The recovery system $\mathcal{R}$ has an additional ancilla $|0_{RA}\rangle$ of dimension n_{RA}. U_R is the $n_R \times n_R$ unitary recovery operator with $n_R = n_S n_{CA} n_{RA}$ and with σ_R the $n_R \times n_R$ full output state $\sigma_R = U_R(\sigma_C \otimes |0_{RA}\rangle\langle 0_{RA}|)U_R^\dagger$. The $n_S \times n_S$ reduced output state, $\hat{\rho}_S$, is given by the partial trace over *all* the ancillas, the bath having been traced out in the previous step. Specifically, $\hat{\rho}_S = \mathbf{Tr}_A\ \sigma_R$.

Caveat emptor The "real" error correction system is unlikely to be perfectly represented by the system shown in Fig. 2, but rather by a full *system-ancilla-bath* interaction [1]. As shown in the block diagram in Fig. 3, U_{QAB} is the $n_{QAB} \times n_{QAB}$ unitary system-ancilla-bath operator, $|0_{CA}0_{RA}\rangle$ is the total ancilla state of dimension $n_{CA}n_{RA}$ and ρ_B is the bath state. The reduced system output state, $\hat{\rho}_S$, is obtained from the full output state ρ_{QAB} by tracing simultaneously over all the ancilla and the bath, $\hat{\rho}_S = \mathbf{Tr}_{AB}\ \rho_{QAB}$. At this level of representation, there is no distinction between the n_{CA} ancilla states used for encoding and the n_{RA} ancilla states used for recovery. The internal design, however, may be constructed with such a distinction.

2.3 Optimal error correction: maximizing fidelity

As stated, the goal is to make the operation $\rho_S \to \hat{\rho}_S$ be as close as possible to a desired unitary operation L_S. Measures to compare two quantum channels are typically based on fidelity or distance, e.g., [5,9]. Let $\mathcal{S}$ denote a trace-preserving quantum channel mapping n-dimensional states to n-dimensional states,

$$\mathcal{S}(\rho) = \sum_k S_k \rho S_k^\dagger, \quad \sum_k S_k^\dagger S_k = I_n \tag{3}$$

The fidelity of the channel compared to identity can be measured in a number of ways, for example:

$$f_{\text{mixed}} = \min_\rho \sum\nolimits_k |\mathbf{Tr}\ S_k \rho|^2$$

$$f_{\text{pure}} = \min_{|\psi\rangle} \sum_k |\langle\psi|S_k|\psi\rangle|^2$$
$$f_{\text{avg}} = \frac{1}{n^2} \sum_k |\mathbf{Tr} S_k|^2 \tag{4}$$

All are in [0,1] and equal to one if and only if $\mathcal{S}(\rho) = \rho$. From (2), $\mathcal{S} = L_S^\dagger \mathcal{R}\mathcal{E}\mathcal{C}$ with OSR elements $S_k = S_{rec} = L_S^\dagger R_r E_e C_c \in \mathbf{C}^{n_S \times n_S}$. Thus $f_{\text{avg}} = 1$ if and only if $\mathcal{R}\mathcal{E}\mathcal{C} = L_S$, i.e., the channel reproduces the desired unitary.

Given $\mathcal{S}$, not all these fidelity measures are easy to calculate. Specifically, f_{mixed} is a convex optimization over all density matrices, that is, over all $\rho \in \mathbf{C}^{n \times n}$, $\rho \geq 0$, $\mathbf{Tr}\ \rho = 1$, and hence, can be numerically obtained. Calculation of f_{avg} is direct. Calculating f_{pure} is, unfortunately, not a convex optimization over all pure states $|\psi\rangle$. If, however, the density matrix associated with f_{mixed} is nearly rank one, then $f_{\text{mixed}} \approx f_{\text{pure}}$.

As a practical matter, when dealing with small channel errors, it does not matter which fidelity measure is used. Therefore it is convenient to use f_{avg}, as it is already in a form explicitly dependent only on the OSR elements. In [14] f_{avg} was also used as the design measure, but specific convex optimization algorithms were not proposed. In [17] a similar optimization was proposed using a distance measure to obtain the recovery given the encoding.

We now focus on the optimization problem,

$$\begin{array}{l} \text{maximize } f_{\text{avg}}(\mathcal{R}, \mathcal{C}) = \frac{1}{n_S^2} \sum_{r,e,c} \left| \mathbf{Tr} L_S^\dagger R_r E_e C_c \right|^2 \\ \text{subject to} \sum_r R_r^\dagger R_r = I_{n_C}, \sum_r C_c^\dagger C_c = I_{n_S} \end{array} \tag{5}$$

The optimization variables are the OSR elements $\{C_c\}$ and $\{R_r\}$. As posed this is a difficult optimization problem. The objective function is not a convex function of either of the design variables and the equality constraints are quadratic, hence, not convex sets. The problem, however, can be approximated using *convex relaxation*, where each non-convex constraint is replaced with a less restrictive convex constraint [2]. This finally results in a bi-convex optimization problem in the encoding and recovery operator elements which can be iterated to yield a local optimum. As we show next, iterating between the two problems is guaranteed to increase fidelity of each of the relaxed problems. When the iterations converge, all that can be said is that a local solution has been found. However, each relaxed solution is in fact optimal, i.e., given an encoding the relaxed recovery is optimal, and conversely, given a recovery the relaxed encoding is optimal.

3 Optimal error correction via bi-convex relaxation

3.1 Process matrix problem formulation

Following the procedure used in quantum process tomography [13, §8.4.2], [11] we expand each of the OSR elements $R_r \in \mathbf{C}^{n_S \times n_C}$ and $C_c \in \mathbf{C}^{n_C \times n_S}$ in a set of basis matrices, respectively, for $\mathbf{C}^{n_S \times n_C}$ and $\mathbf{C}^{n_C \times n_S}$, that is,

$$\begin{array}{l} R_r = \sum_i x_{ri}\ B_{Ri} \\ C_c = \sum_i x_{ci}\ B_{Ci} \end{array} \tag{6}$$

where $\{B_{Ri} \in \mathbf{C}^{n_S \times n_C},\ B_{Ci} \in \mathbf{C}^{n_C \times n_S} \mid i = 1, \ldots, n_S n_C\}$ and the $\{x_{ri}\}$ and $\{x_{ci}\}$ are complex scalars. Problem (5) can then be equivalently expressed as,

$$\begin{array}{ll} \text{maximize} & f_{\text{avg}}(X_R, X_C) = \sum_{ijk\ell}(X_R)_{ij}(X_C)_{k\ell}F_{ijk\ell} \\ \text{subject to} & \sum_{ij}\ (X_R)_{ij}\ B_{Ri}^\dagger B_{Rj} = I_{n_C} \\ & \sum_{k\ell}\ (X_C)_{k\ell}\ B_{Ck}^\dagger B_{C\ell} = I_{n_S} \\ & (X_R)_{ij} = \sum_r\ x_{ri}x_{rj}^* \\ & (X_C)_{k\ell} = \sum_c\ x_{ck}x_{c\ell}^* \\ & F_{ijk\ell} = \sum_e\ g_{ik}g_{j\ell}^*/n_s^2 \\ & g_{ik} = \mathbf{Tr}\ L_S^\dagger B_{Ri}E_e B_{Ck} \end{array} \tag{7}$$

The optimization variables are the $n_S n_C \times n_S n_C$ *process matrices* X_R and X_C and the scalars $\{x_{ri}\}$ and $\{x_{ci}\}$. The problem data which describes the desired unitary and error system is contained in the $\{F_{ijk\ell}\}$. The equality constraints $(X_R)_{ij} = \sum_r\ x_{ri}x_{rj}^*$ and $(X_C)_{k\ell} = \sum_c\ x_{ck}x_{c\ell}^*$ are both quadratic, exposing again that this is not a convex optimization problem.

3.2 Design of $\mathcal{R}$ given $\mathcal{C}$ and $\mathcal{E}$

In this section and in the remainder of the paper we set the desired logical operation to identity, i.e., $L_S = I_S$; just error correction, not correction and computation. This is without loss of generality as a desired logical operation can be added everywhere.

Suppose the encoding $\mathcal{C}$ is given (and $L_S = I_S$). Then optimizing only over $\mathcal{R}$ in (7) can be equivalently expressed as,

$$\begin{array}{ll} \text{maximize} & f_{\text{avg}}(X_R, \mathcal{C}) = \mathbf{Tr}\ X_R W_R(\mathcal{E}, \mathcal{C}) \\ \text{subject to} & \sum_{i,j}\ (X_R)_{ij}\ B_{Ri}^\dagger B_{Rj} = I_{n_C} \\ & (X_R)_{ij} = \sum_r\ x_{ri}x_{rj}^* \\ & (W_R(\mathcal{E}, \mathcal{C}))_{ij} = \sum_{c,k,\ell}\ x_{ck}x_{c\ell}^*\ F_{ijk\ell} \\ & \qquad\qquad\qquad\quad = \sum_{k,\ell}\ (X_C)_{k\ell}\ F_{ijk\ell} \end{array} \tag{8}$$

The optimization variables are the matrix $X_R \in \mathbf{C}^{n_S n_C \times n_S n_C}$ and the scalars $\{x_{ri}\}$. The problem data is contained in the positive semidefinite matrix $W_R(\mathcal{E}, \mathcal{C}) \in \mathbf{C}^{n_S n_C \times n_S n_C}$. The objective function is now linear in X_R, which is of course a convex function. However, each of the equality constraints, $(X_R)_{ij} = \sum_r\ x_{ri}x_{rj}^*$ is quadratic, and thus does not form a convex set. This set of quadratic equality constraints can be relaxed to the matrix inequality constraint, $X_R \geq 0$, that is, X_R is positive semidefinite, a convex set in the elements of X_R. A *convex relaxation* of (8) is then,

$$\begin{array}{ll} \text{maximize } \mathbf{Tr}\ X_R W_R(\mathcal{E}, \mathcal{C}) \\ \text{subject to } X_R \geq 0, \quad \sum_{i,j} (X_R)_{ij}\ B_{Ri}^\dagger B_{Rj} = I_{n_C} \end{array} \tag{9}$$

This class of convex optimization problems is referred to as an SDP, for *semidefinite program* [2].[2] For a given encoding $\mathcal{C}$ (or X_C), the optimal solution to the relaxed problem (9), $X_R^{\text{rlx\_opt}}$, provides an upper bound on the average fidelity objective in (5) or (7). From the fidelity inequalities, we can derive a lower bound. Specifically, the (unknown, possibly unknowable) solution to the original problem (5), is bounded as follows:

$$\begin{array}{rl} f_{\text{mixed}}(\mathcal{R}^{\text{rlx\_opt}}, \mathcal{C}) & \leq \max_{\mathcal{R}}\ f_{\text{pure}}(\mathcal{R}, \mathcal{C}) \\ & \leq f_{\text{avg}}(\mathcal{R}^{\text{rlx\_opt}}, \mathcal{C}) \end{array} \tag{10}$$

where $\mathcal{R}^{\text{rlx\_opt}}$ is the OSR with elements $\left\{R_r^{\text{rlx\_opt}}\right\}$ obtained from $X_R^{\text{rlx\_opt}}$ via the singular value decomposition,

$$\begin{array}{c} X_R^{\text{rlx\_opt}} = VSV^\dagger \\ \Downarrow \\ R_r^{\text{rlx\_opt}} = \sqrt{s_r} \sum_{i=1}^{n_S n_C} V_{ir} B_{Ri},\ r = 1, \ldots, n_S n_C \end{array} \tag{11}$$

where $V \in \mathbf{C}^{n_S n_C \times n_S n_C}$ is unitary, $S = \text{diag}(s_1\ \cdots\ s_{n_S n_C})$ with singular values in decreasing order, $s_1 \geq s_2 \geq\ \cdots\ \geq s_{n_S n_C} \geq 0$.

3.3 Design of $\mathcal{C}$ given $\mathcal{R}$ and $\mathcal{E}$

Repeating the previous steps, optimizing only over $\mathcal{C}$ in (7) can be equivalently expressed as,

$$\begin{array}{ll} \text{maximize} & f_{\text{avg}}(X_C, \mathcal{R}) = \mathbf{Tr}\ X_C W_C(\mathcal{E}, \mathcal{R}) \\ \text{subject to} & \sum_{k,\ell} (X_C)_{k\ell}\ B_{Ck}^\dagger B_{C\ell} = I_{n_S} \\ & (X_C)_{k\ell} = \sum_c\ x_{ck} x_{c\ell}^* \\ & (W_C(\mathcal{E}, \mathcal{R}))_{k\ell} = \sum_{r,i,j}\ x_{ri} x_{rj}^*\ F_{ijk\ell} \\ & \qquad\qquad\qquad\ \ = \sum_{i,j}\ (X_R)_{ij}\ F_{ijk\ell} \end{array} \tag{12}$$

The optimization variables are the matrix $X_C \in \mathbf{C}^{n_S n_C \times n_S n_C}$ and the scalars $\{x_{ci}\}$ with all the problem data contained in the symmetric positive semidefinite matrix $W_C(\mathcal{E}, \mathcal{R}) \in \mathbf{C}^{n_S n_C \times n_S n_C}$. In this case, however, the basis matrices, $\{B_{Ci}\}$ are $n_C \times n_S$. Repeating the previous procedure of relaxing the quadratic equality constrain to $X_C \geq 0$, we obtain the convex relaxation of (7) as the SDP,

[2] A standard SDP is to minimize a linear objective function subject to convex inequalities and linear equalities. The objective function in (9) is the maximization of a linear function which is equivalent to the minimization of its negative, and hence, is a linear objective function.

$$\begin{array}{l} \text{maximize } \mathbf{Tr}\ X_C W_C(\mathcal{E}, \mathcal{R}) \\ \text{subject to } X_C \geq 0, \quad \sum_{i,j} (X_C)_{ij}\ B_{Ci}^\dagger B_{Cj} = I_{n_S} \end{array} \tag{13}$$

Analogously to (11), for a given recovery $\mathcal{R}$, the (unknown, possibly unknowable) solution to the original problem (5), is bounded as follows:

$$\begin{array}{rl} f_{\text{mixed}}(\mathcal{R}, \mathcal{C}^{\text{rlx\_opt}}) & \leq \max_{\mathcal{C}}\ f_{\text{pure}}(\mathcal{R}, \mathcal{C}) \\ & \leq f_{\text{avg}}(\mathcal{R}, \mathcal{C}^{\text{rlx\_opt}}) \end{array} \tag{14}$$

where $\mathcal{C}^{\text{rlx\_opt}}$ is the OSR with elements $\left\{C_c^{\text{rlx\_opt}}\right\}$ obtained from $X_C^{\text{rlx\_opt}}$ via the singular value decomposition,

$$\begin{array}{c} X_C^{\text{rlx\_opt}} = VSV^\dagger \\ \Downarrow \\ C_c^{\text{rlx\_opt}} = \sqrt{s_c}\ \sum_{i=1}^{n_S n_C}\ V_{ic} B_{Ci},\ c = 1, \ldots, n_S n_C \end{array} \tag{15}$$

where $V \in \mathbf{C}^{n_S n_C \times n_S n_C}$ is unitary, $S = \text{diag}(s_1\ \cdots\ s_{n_S n_C})$ with singular values in decreasing order, $s_1 \geq s_2 \geq\ \cdots\ \geq s_{n_S n_C} \geq 0$.

3.4 Iterative bi-convex algorithm

Proceeding analogously as in [14], the two separate optimizations for $\mathcal{C}$ and $\mathcal{R}$ can be combined into the following iteration.

initialize encoding $\hat{\mathcal{C}}$ and stopping level ϵ

repeat

1. **optimize recovery**
   (a) compute $X_R^\star$ as solution to:

$$\begin{array}{l} \text{maximize } \mathbf{Tr}\ X_R W_R(\mathcal{E}, \hat{\mathcal{C}}) \\ \text{subject to } X_R \geq 0,\ \sum_{i,j} (X_R)_{ij}\ B_{Ri}^\dagger B_{Rj} = I_{n_C} \end{array}$$

   (b) use (11) to compute $\mathcal{R}^\star$ from $X_R^\star$
2. **optimize encoding**
   (a) compute $X_C^\star$ as solution to:

$$\begin{array}{l} \text{maximize } \mathbf{Tr}\ X_C W_C(\mathcal{E}, \mathcal{R}^\star) \\ \text{subject to } X_R \geq 0,\ \sum_{i,j} (X_C)_{ij}\ B_{Ci}^\dagger B_{Cj} = I_{n_S} \end{array}$$

   (b) use (15) to compute $\mathcal{C}^\star$ from $X_C^\star$
3. **compute change in fidelity**

$$\Delta f_{\text{avg}} = f_{\text{avg}}(\mathcal{R}^\star, \mathcal{C}^\star) - f_{\text{avg}}(\mathcal{R}^\star, \hat{\mathcal{C}})$$

4. **reset**

$$\hat{\mathcal{C}} = \mathcal{C}^\star$$

until

$$\Delta f_{\text{avg}} < \epsilon$$

The algorithm returns $(\mathcal{R}^\star, \mathcal{C}^\star)$. The optimization in each of the steps is a convex optimization and hence fidelity will increase in each step, thereby converging to a local solution of the joint relaxed problem. This solution is not necessarily a local solution to the original problem (5) or (7). However, the upper and lower bounds obtained will apply. The optimization steps can be reversed by starting with an initial recovery and then starting the iteration by optimizing over encoding. We remark again, that via (11) each relaxed solution produces an equivalent OSR, and hence the relaxed solution is optimal in each iteration between encoding and recovery.

3.5 Decoherence resistant encoding

If the sole purpose of encoding is to sustain the information state ρ_S, then the desired operation is the identity ($L_S = I_S$). For a given error $\mathcal{E}$, finding an optimal encoding by solving (13) is equivalent to finding a *decoherence-resistant-subspace*. If there is perfect error correction, then we have found a *decoherence-free-subspace* [12]. In [18], this problem was considered using f_{pure}, the pure state fidelity.

Usually in this case the recovery operation in Fig. 1, which we refer to as the *Nominal recovery*, is simply the partial trace over the encoding ancilla, that is,

$$\begin{aligned} \hat{\rho}_S &= \mathcal{R}(\sigma_C) \\ &= \mathbf{Tr}_{CA}\, \sigma_C = \begin{bmatrix} \mathbf{Tr}\ (\sigma_C)_{[1,1]} & \cdots & \mathbf{Tr}\ (\sigma_C)_{[1,n_S]} \\ \vdots & \vdots & \vdots \\ \mathbf{Tr}\ (\sigma_C)_{[n_S,1]} & \cdots & \mathbf{Tr}\ (\sigma_C)_{[n_S,n_S]} \end{bmatrix} \end{aligned} \quad (16)$$

where the $(\sigma_C)_{[i,j]}$ are the n_S^2 sub-block matrices of σ_C, each being $n_{CA} \times n_{CA}$. Hence, the OSR elements of the nominal recovery are given by

$$\begin{aligned} &(R_r)_{ij} = \begin{cases} 1 & j = (i-1)n_{CA} + r \\ 0 & \text{else} \end{cases} \\ &r = 1, \ldots, n_{CA}, \;\; j = 1, \ldots, n_S \end{aligned} \quad (17)$$

4 Robust error correction

The bi-convex optimization can be extended to the case where the error system is one of a number of possible error systems, that is,

$$\mathcal{E} \in \{ \mathcal{E}_\alpha \mid \alpha = 1, \ldots, \ell \} \tag{18}$$

where each $\mathcal{E}_\alpha$ has OSR elements $\{E_{\alpha e}\}$. The worst-case fidelity design problem, by analogy with (7), is then:

$$\begin{array}{ll} \text{maximize } \min_\alpha \; \sum_{ijk\ell} (X_R)_{ij} \; (X_C)_{k\ell} \; F_{\alpha ijk\ell} \\ \text{subject to } \; X_R, \; X_C \text{ constrained as in (7)} \\ \qquad\qquad F_{\alpha ijk\ell} = \sum_e \; g_{\alpha ik} g_{\alpha j\ell}^* / n_s^2 \\ \qquad\qquad g_{\alpha ik} = \mathbf{Tr} \; L_S^\dagger B_{Ri} E_{\alpha e} B_{Ck} \end{array} \tag{19}$$

Iterating as before between $\mathcal{R}$ and $\mathcal{C}$ results again in separate convex optimization problems, each of which is an SDP. Specifically, for a given encoding $\mathcal{C}$, a robust recovery is obtained from,

$$\begin{array}{ll} \text{maximize } \min_\alpha \; \mathbf{Tr} \; X_R W_R(\mathcal{E}_\alpha, \mathcal{C}) \\ \text{subject to } \; X_C \geq 0, \quad \sum_{i,j} \; (X_C)_{ij} \; B_{Ci}^\dagger B_{Cj} = I_{n_S} \end{array} \tag{20}$$

Similarly, for a given recovery $\mathcal{R}$, a robust encoding is obtained from,

$$\begin{array}{ll} \text{maximize } \min_\alpha \; \mathbf{Tr} \; X_C W_C(\mathcal{E}_\alpha, \mathcal{R}) \\ \text{subject to } \; X_R \geq 0, \quad \sum_{i,j} \; (X_R)_{ij} \; B_{Ri}^\dagger B_{Rj} = I_{n_C} \end{array} \tag{21}$$

5 Computing the solution via Lagrange Duality

The main difficulty with embedding the OSR elements into either X_R or X_C is *scaling with qubits*. Specifically, the number of design parameters needed to determine either X_C or X_R scales exponentially with the number of qubits. Although exponential scaling at the moment seems unavoidable, we show in this section that solving the dual SDPs associated with either (9) or (13) requires fewer parameters, and thus engenders a reduced computational burden.

The convex optimization problems (9) and (13) are both SDPs of the form,

$$\begin{array}{ll} \text{maximize } \mathbf{Tr} \; XW \\ \text{subject to } \; X \geq 0, \quad \sum_{ij} X_{ij} B_i^\dagger B_j = I_m \end{array} \tag{22}$$

with optimization variable $X = X^\dagger \in \mathbf{C}^{n \times n}$, $n = rm$ for some integer r, and with each basis matrix $B_i \in \mathbf{C}^{r \times m}$. We will refer to this as the *primal problem*. Accounting for the linear (matrix) equality constraint and the Hermiticity of X, the number of real optimization variables in the primal problem (22) is $p = n^2 - m^2 = (r^2 - 1)m$.

Solving (9) for X_R, gives $n = n_S n_C$, $r = n_S$, $m = n_C = n_S n_{CA}$ for $p_R = (n_S^2 - 1) n_C^2$ primal parameters. Solving (13) for X_C, gives $n = n_S n_C$, $r = n_C$, $m = n_S$ for $p_C = (n_C^2 - 1) n_S^2$ primal paramters. Exponential growth in computation occurs because each of these dimensions are exponential in the number of qubits, i.e., $n_S = 2^{q_S}$, $n_C = 2^{q_S+q_{CA}}$, and so on.

The computational burden can be somewhat alleviated by appealing to *Lagrange Duality Theory* [2, Ch. 5]. In the Appendix we show that the *dual problem* associated with the *primal problem* (22) is also an SDP, specifically,

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\ Y \\ \text{subject to} & K(Y) - W \geq 0, \quad K_{ij}(Y) = \mathbf{Tr}\ B_j^\dagger B_i Y \end{array} \tag{23}$$

with optimization variable $Y = Y^\dagger \in \mathbf{C}^{m\times m}$. The number of (real) optimization variables for the dual problem is then at most $d = m^2$. The reduction of primal to dual parameters is thus $p/dp = r^2 - 1$. Solving for the recovery dual gives the reduction as $n_S^2 - 1$ and for the encoding dual as $n_C^2 - 1$. We show in the Appendix that if $(X^{\rm opt}, Y^{\rm opt})$ solve the primal and dual problems, respectively, then:

$$\begin{aligned} \mathbf{Tr}\ X^{\rm opt} W &= \mathbf{Tr}\ Y^{\rm opt} \\ (K(Y^{\rm opt}) - W)X^{\rm opt} &= 0 \end{aligned} \tag{24}$$

The second equation above together with the linear equality constraint in (22) can be used to obtain the primal solution $X^{\rm opt}$ from the dual solution $Y^{\rm opt}$. That is, solve for $X^{\rm opt}$ from the set of *linear* equations,

$$\begin{aligned} (K(Y^{\rm opt}) - W)X^{\rm opt} &= 0 \\ \sum\nolimits_{ij} X_{ij}^{\rm opt} B_i^\dagger B_j &= I_m \end{aligned} \tag{25}$$

As is well known, very efficient methods exist for solving linear equations [6]. We remark again that the "indirect method" presented in [10] has some inherent computational advantages over the direct method presented here.

6 Example

In this illustrative example, the goal is to preserve a single information qubit using a single ancilla qubit. Thus, the desired logical gate is the identity, that is, $L_S = I_2$, with $n_S = n_{CA} = 2$, and hence, $n_C = 4$. We made two error systems, $\mathcal{E}_a$ and $\mathcal{E}_b$, by randomly selecting the unitary bath representation as shown in Fig. 2 as follows: Each error system has a single qubit bath state, $|0\rangle_B$, thus $n_B = 2$. The Hamiltonian for each system, $H_E = H_E^\dagger \in \mathbf{C}^{n_E\times n_E}$, $n_E = n_C n_B = 8$, was chosen randomly and then adjusted to have the magnitude (maximum singular value) $\|H_E\| = \delta_E = 0.75$. Then the unitary representing the error system was set to $U_E = \exp(-iH_E)$ and from this the corresponding OSR $\mathcal{E}$ was computed. The OSR elements to three decimal points for the two random systems are as follows.

OSR elements of$\mathcal{E}_a$

$$E_{a1} = \begin{bmatrix} 0.9 - 0.049i & 0.193 + 0.194i & -0.161 + 0.039i & -0.135 + 0.156i \\ -0.159 + 0.148i & 0.887 - 0.046i & 0.148 - 0.025i & -0.168 - 0.081i \\ 0.167 + 0.061i & -0.07 + 0.004i & 0.905 + 0.161i & 0.16 + 0.125i \\ 0.124 + 0.137i & 0.167 - 0.155i & -0.203 + 0.118i & 0.844 - 0.26i \end{bmatrix}$$

$$E_{a2} = \begin{bmatrix} 0.053 - 0.063i & -0.034 + 0.082i & 0.148 - 0.085i & 0.13 - 0.076i \\ -0.168 - 0.01i & 0.141 + 0.073i & 0.008 + 0.091i & -0.074 - 0.024i \\ 0.119 + 0.053i & 0.207 - 0.02i & 0.043 + 0.01i & -0.063 - 0.21i \\ 0.123 + 0.066i & 0.027 + 0.008i & 0.07 - 0.058i & 0.098 - 0.11i \end{bmatrix}$$

OSR elements of $\mathcal{E}_b$

$$E_{b1} = \begin{bmatrix} 0.943 + 0.018i & -0.14 - 0.024i & 0.076 - 0.081i & 0.04 - 0.163i \\ 0.107 + 0.062i & 0.876 + 0.068i & -0.06 - 0.021i & -0.127 + 0.06i \\ -0.025 - 0.042i & 0.122 + 0.073i & 0.889 - 0.035i & 0.043 - 0.078i \\ -0.017 - 0.094i & 0.095 + 0.035i & -0.032 - 0.089i & 0.88 + 0.113i \end{bmatrix}$$

$$E_{b2} = \begin{bmatrix} 0.07i & -0.2 - 0.082i & 0.028 - 0.083i & 0.179 + 0.206i \\ -0.003 - 0.147i & 0.138 - 0.155i & 0.202 + 0.306i & 0.045 - 0.134i \\ 0.049 + 0.084i & -0.149 + 0.217i & 0.143 - 0.04i & 0.024 + 0.174i \\ -0.191 + 0.095i & -0.081 - 0.097i & 0.007 - 0.127i & 0.035 - 0.167i \end{bmatrix}$$

Neither of these error systems is of the standard type, e.g., there is no independent channel structure. The choice of $\delta_E = 0.75$ is perhaps extreme, but is motivated here by our desire to demonstrate that the optimization procedure can handle errors that are beyond the range of classically-inspired quantum error correction. For this particular set of error systems, we do not know if there exists an encoding/recovery pair limited to using a single encoding ancilla state which can bring perfect correction. This also motivates the search for the still elusive black-box error correction discussed in the introduction.

For each of the error systems we ran the bi-convex iteration 100 times starting with the initial recovery operator $\mathcal{R}_0$ given by the partial trace operation (17). Denote $(\mathcal{R}_{a1}, \mathcal{C}_{a1})$ and $(\mathcal{R}_{a100}, \mathcal{C}_{a100})$ as the 1st and 100th iteration pairs optimized for $\mathcal{E}_a$, and similarly $(\mathcal{R}_{b1}, \mathcal{C}_{b1})$ and $(\mathcal{R}_{b100}, \mathcal{C}_{b100})$ as the 1st and 100th iteration pairs optimized for $\mathcal{E}_b$. Table 1 shows the average fidelities $f_{\text{avg}}(\mathcal{R}, \mathcal{E}, \mathcal{C})$ for some of the possible combinations.

As Table 1 clearly shows, fidelity tuned for a specific error, either $\mathcal{E}_a$ or $\mathcal{E}_b$ in this example, saturated to the levels shown (0.9997) in about 100 iterations. However, as is clearly indicated in the table, neither of the optimized codes is robust to *both* $\mathcal{E}_a$ and $\mathcal{E}_b$. That is, each does very poorly when the error Hamiltonian H_E changes even though its magnitude δ_E remain unchanged. The codes, however, are not overly sensitive to small changes in the error magnitude δ_E, provided that the structure of the Hamiltonian H_E remains unchanged. However, robust codes can be obtained by optimizing over the set of errors as described in Sect. 4 and illustrated in the example to follow.

Table 1 Average fidelities

Type	$\mathcal{R}, \mathcal{C}$	$\mathcal{E}_a$	$\mathcal{E}_b$
Optimal encoding ($\mathcal{E}_a$) Nominal recovery	$\mathcal{R}_0, \mathcal{C}_{a1}$	0.9686	0.7631
Optimal encoding ($\mathcal{E}_a$) Optimal recovery ($\mathcal{E}_a$) 1 iteration	$\mathcal{R}_{a1}, \mathcal{C}_{a1}$	0.9719	0.7805
Optimal encoding ($\mathcal{E}_a$) Optimal recovery ($\mathcal{E}_a$) 100 iterations	$\mathcal{R}_{a100}, \mathcal{C}_{a100}$	0.9997	0.6261
Optimal encoding ($\mathcal{E}_b$) Nominal recovery	$\mathcal{R}_0, \mathcal{C}_{b1}$	0.7445	0.9091
Optimal encoding ($\mathcal{E}_b$) Optimal recovery ($\mathcal{E}_b$) 1 iteration	$\mathcal{R}_{b1}, \mathcal{C}_{b1}$	0.7843	0.9441
Optimal encoding ($\mathcal{E}_b$) Optimal recovery ($\mathcal{E}_b$) 100 iterations	$\mathcal{R}_{b100}, \mathcal{C}_{b100}$	0.7412	0.9997

The optimal encodings and recoveries are based on using error models either $\mathcal{E}_a$ or $\mathcal{E}_b$ as indicated. When the code is optimized for $\mathcal{E}_a$ (upper three rows) the fidelities are high (column 3) but are considerably lower when the same code is applied to $\mathcal{E}_b$ (column 4). The same applies reversley when the code is optimized for $\mathcal{E}_b$ (bottom three rows)

By raising the number of ancillas it is of course possible to make the system robust. This, however, introduces considerable complexity. What the table suggests is that an alternate route is to tune for maximal fidelity, say, in a particular module. This of course can only be done on the actual system.

For each of the optimizations, the 8×8 process matrices X_C and X_R, associated, respectively with each $\mathcal{C}$ and $\mathcal{R}$ were of reduced rank. For all the optimized $\mathcal{C}$, each process matrix X_C was found to have a single dominant singular value, and hence, there is a single dominant 4×2 OSR element which characterizes $\mathcal{C}$. For the optimized $\mathcal{R}$, each X_R was found to have two dominant singular values, and hence, there are two dominant 2×4 OSR elements which characterize $\mathcal{R}$. For example, the recovery/encoding pair $(\mathcal{R}_{a100}, \mathcal{C}_{a100})$ has the OSR elements:

$$C_1 = \begin{bmatrix} -0.629 & 0.189 - 0.332i \\ 0.455 + 0.378i & 0.207 + 0.24i \\ 0.42 + 0.063i & -0.425 - 0.358i \\ 0.13 + 0.233i & 0.626 + 0.226i \end{bmatrix}$$

$$R_1 = \begin{bmatrix} -0.707 & 0.532 - 0.342i & 0.194 - 0.175i & 0.103 - 0.138i \\ 0.134 + 0.087i & 0.009 - 0.166i & -0.103 + 0.404i & 0.833 - 0.276i \end{bmatrix}$$

$$R_2 = \begin{bmatrix} -0.603 & -0.528 + 0.461i & -0.262 - 0.131i & 0.172 - 0.163i \\ 0.313 - 0.103i & 0.259 + 0.104i & -0.374 - 0.728i & 0.174 - 0.333i \end{bmatrix}$$

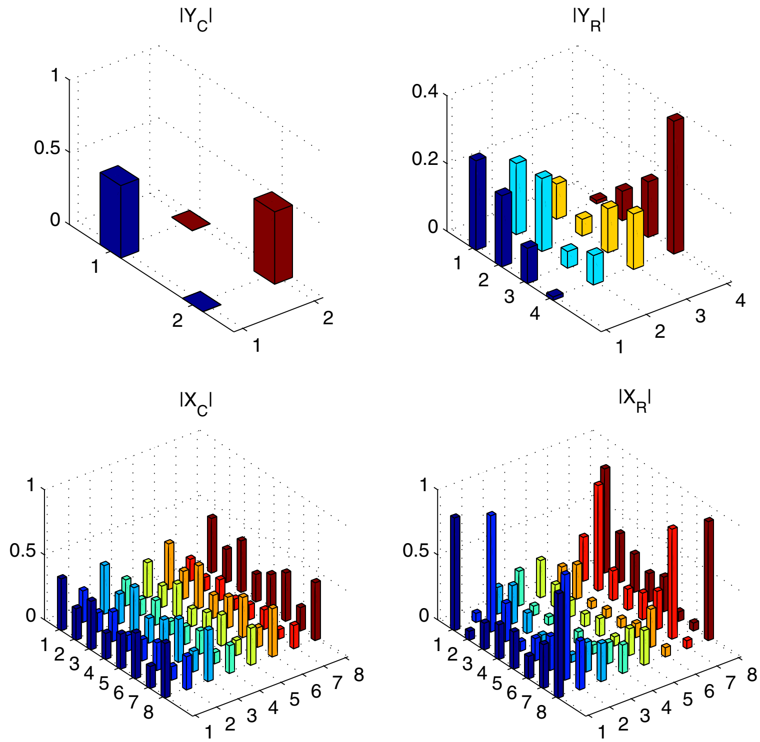

Fig. 4 Magnitudes of primal-dual pairs (X_C, Y_C) and (X_R, Y_R) corresponding to $(\mathcal{R}_{a100}, \mathcal{C}_{a100})$ optimized for $\mathcal{E}_a$

It is not obvious that these correspond to any of the standard codes. Since $C_1^\dagger C_1 = I_2$ and $\sum_{i=1}^{2} R_i^\dagger R_i = I_4$, we can construct the encoding and recovery unitaries in Fig. 2 as,

$$U_C = [C_1 \; C_2], \quad U_R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix}$$

where $C_2 \in \mathbf{C}^{2\times 2}$ is arbitrary as long as U_C is unitary, or equivalently, $C_1^\dagger C_2 = 0$ and $C_2^\dagger C_2 = I_2$. Observe that U_R is already a 4×4 unitary.

Bar plots of the magnitude of the elements in the primal-dual pairs (X_C, Y_C) and (X_R, Y_R) corresponding, respectively to $(\mathcal{R}_{a100}, \mathcal{C}_{a100})$ and $(\mathcal{R}_{b100}, \mathcal{C}_{b100})$ are shown in Figs. 4 and 5. From many of such similar plots we have observed some common structure which may be used to reduce the computational burden. For example, the dual encoding process matrix Y_C is nearly diagonal while the dual recovery process matrix Y_R is block diagonal.

We also computed a *robust* encoding and recovery for the error set $\{\mathcal{E}_a, \mathcal{E}_b\}$ by iterating between (20) and (21). The resultant average fidelities are in Table 2.

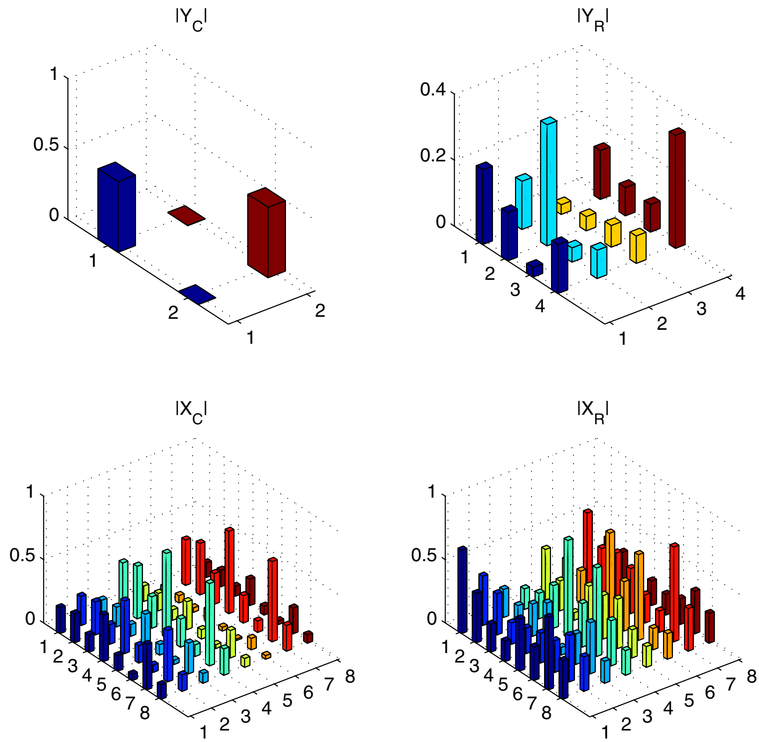

Fig. 5 Magnitudes of primal-dual pairs (X_C, Y_C) and (X_R, Y_R) corresponding to $(\mathcal{R}_{b100}, \mathcal{C}_{b100})$ optimized for $\mathcal{E}_b$

Table 2 Average robust fidelities

Type	$\mathcal{R}$, $\mathcal{C}$	$\mathcal{E}_a$	$\mathcal{E}_b$
Robust encoding Nominal recovery	$\mathcal{R}_0$, $\mathcal{C}_{ab1}$	0.8840	0.8840
Robust encoding Robust recovery 1 iteration	$\mathcal{R}_{ab1}$, $\mathcal{C}_{ab1}$	0.9284	0.9284
Robust encoding Robust recovery 100 Iterations	$\mathcal{R}_{ab100}$, $\mathcal{C}_{ab20}$	0.9576	0.9576

Here the codes are optimized to handle both error systems $\mathcal{E}_a$ and $\mathcal{E}_b$

Comparing Tables 1 and 2 clearly shows that a robust design is possible although at a cost of performance. Also in this case after 100 iterations the robust fidelity did not increase. In addition, the rank of the process matrices X_C and X_R remained as before at 1 and 2, respectively, and the resulting OSR elements do not appear standard.

7 Conclusions

We have shown that the design of a quantum error correction system can be cast as a bi-convex iteration between encoding and recovery, each being a semidefinite program (SDP). The SDP formalism also allows for a robust design by enumerating constraints associated with different error models. We illustrated the approach with an example where the error map does not assume independent channels. The example data as seen in Table 1 shows that the codes are not robust to structural changes in the Hamiltonian even though the error magnitude δ_E remains unchanged. Although no measure of robustness is specifically put forward, when we optimize over a set of errors as seen in Table 2 a robust code is obtained for the set of errors. At present, formulating a precise measure of robustness seems to be an open question.

Acknowledgements This work has been funded under the DARPA QuIST Program (Quantum Information Science & Technology) and (to D. A. L.) NSF CCF-0523675 and ARO (Quantum Algorithms, W911NF-05-1-0440). We would like to thank I. Walmsley and D. Browne (Oxford), C. Brif, M. Grace, and H. Rabitz (Princeton), and A. Shabani (USC) for many fruitful discussions.

Appendix: dual problem

We apply Lagrange Duality Theory [2, Ch.5]. Write the primal problem (22) as a minimization,

$$\begin{array}{l} \text{minimize } -\mathbf{Tr}\ XW \\ \text{subject to } X \geq 0, \quad \sum_{ij} X_{ij} C_{ij} = I_m \end{array} \tag{26}$$

with optimization variable $X = X^\dagger \in \mathbf{C}^{n\times n}$. The *Lagrangian* for (26) is,

$$\begin{aligned} L(X, Z, Y) &= -\mathbf{Tr}\ XW = \mathbf{Tr}\ XZ - \mathbf{Tr}\ Y(I_m - \textstyle\sum_{ij} X_{ij} B_i^\dagger B_j) \\ &= \textstyle\sum_{ij} X_{ij}(-W_{ji} - Z_{ji} + \mathbf{Tr}\ YC_{ji}) - \mathbf{Tr}\ Y \end{aligned} \tag{27}$$

where $Z = Z^\dagger \in \mathbf{C}^{n\times n}$ and $Y = Y^\dagger \in \mathbf{C}^{m\times m}$ are Lagrange multipliers associated with the (Hermitian) inequality and equality constraints, respectively. The *Lagrange dual function* is then,

$$\begin{aligned} g(Z, Y) &= \inf_X\ L(X, Z, Y) \\ &= \begin{cases} -\mathbf{Tr}\ Y & Z_{ji} = \mathbf{Tr}\ YC_{ij} - W_{ji} \\ -\infty & \text{otherwise} \end{cases} \end{aligned} \tag{28}$$

For any Y and $Z \geq 0$, $g(Z, Y)$ yields a lower bound on the optimal objective $-\mathbf{Tr}\ X^{\text{opt}} W$. The largest lower bound from this dual function is then $\max\{\, g(Z, Y) \mid Z \geq 0 \,\}$. Eliminating Z, this can be written equivalently as,

$$\begin{array}{l} \text{minimize } \mathbf{Tr}\ Y \\ \text{subject to } K(Y) - W \geq 0, \quad K_{ij}(Y) = \mathbf{Tr}\ YC_{ij} \end{array} \tag{29}$$

with optimization variable $Y = Y^\dagger \in \mathbf{C}^{m\times m}$. This is precisely the result in (23). Because the problem is strictly convex, the dual optimal objective is equal to the primal optimal objective as stated in the first line of (24). The *complementary slackness* condition gives the second line in (24).

References

1. Alicki, R., Lidar, D.A., Zanardi, P.: Internal consistency of fault-tolerant quantum error correction in light of rigorous derivations of the quantum markovian limit. Phys. Rev. A **73**, 052311 (2005)
2. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
3. Byrd, M.S., Lidar, D.A.: Empirical determination of bang-bang operations. Phys. Rev. A **67**, 012324 (2003)
4. Fletcher, A.S., Shor, P.W., Win, M.Z.: Optimum quantum error recovery using semidefinite programming. Phys. Rev. A **75**, 012338 (2007)
5. Gilchrist, A., Langford, N.K., Nielsen, M.A.: Distance measures to compare real and ideal quantum processes. Phys. Rev. A **71**, 062310 (2005)
6. Golub, G.H., Van Loan, C.F.: Matrix Computations. Johns Hopkins University Press, Maryland (1983)
7. Gottesman, D.: Class of quantum error-correcting codes saturating the quantum hamming bound. Phys. Rev. A **54**, 1862 (1996)
8. Knill, E., Laflamme, R.: Theory of quantum error-correcting codes. Phys. Rev. A **55**, 900 (1997)
9. Kosut, R.L., Grace, M., Brif, C., Rabitz, H.: On the distance between unitary propagators of quantum systems of differing dimensions. Eprint. quant-ph/0606064
10. Kosut, R.L., Shabani, A., Lidar, D.A.: Robust quantum error correction via convex optimization. Phys. Rev. Lett. **100**, 020502 (2008)
11. Kosut, R.L., Walmsley, I.A., Rabitz, H.: Optimal experiment design for quantum state and process tomography and hamiltonian parameter estimation. Eprint. quant-ph/0411093. Mohseni, M., Rezakhani, A.T., Lidar, D.A.: Quantum process tomography: resource analysis of different strategies. Phys. Rev. A **77**, 032322 (2008)
12. Lida, D.A., Chuang, I.L., Whaley, K.B.: Decoherence-free subspaces for quantum computation. Phys. Rev. Lett. **81**, 2594 (1998). Shabani, A., Lidar, D.A.: Theory of initialization-free decoherence-free subspaces and subsystems. Phys. Rev. A **72**, 043203 (2005)
13. Nielsen, M.A., Chuang, I.L.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
14. Reimpell, M., Werner, R.F.: Iterative optimization of quantum error correcting codes. Phys. Rev. Lett. **94**, 080501 (2005)
15. Shor, P.W.: Scheme for reducing decoherence in quantum memory. Phys. Rev. A **52**, R2493 (1995)
16. Steane, A.M.: Error correcting codes in quantum theory. Phys. Rev. Lett. **77**, 793 (1996)
17. Yamamoto, N., Hara, S., Tsumara, K.: Suboptimal quantum error correcting procedure based on semidefinite programming. Phys. Rev. A **71**, 022322 (2005)
18. Zanardi, P., Lidar, D.A.: Purity and state fidelity of quantum channels. Phys. Rev. A **70**, 012315 (2004)