Optimizing completely positive maps using semidefinite programming

Koenraad Audenaert[1,2] and Bart De Moor[1]

[1]*Katholieke Universiteit Leuven, Department of Electrical Engineering (ESAT-SISTA), Kasteelpark Arenberg 10, B-3001 Leuven-Heverlee, Belgium*

[2]*QOLS, Blackett Laboratory, Imperial College of Science, Technology and Medicine, London, SW7 2BW, United Kingdom*

Recently, a lot of attention has been devoted to finding physically realizable operations that realize as closely as possible certain desired transformations between quantum states, e.g., quantum cloning, teleportation, quantum gates, etc. Mathematically, this problem boils down to finding a completely positive trace-preserving (CPTP) linear map that maximizes the (mean) fidelity between the map itself and the desired transformation. In this communication, we want to draw attention to the fact that this problem belongs to the class of so-called semidefinite programming (SDP) problems. As SDP problems are convex, it immediately follows that they do not suffer from local optima. Furthermore, this implies that the numerical optimization of the CPTP map can, and should, be done using methods from the well-established SDP field, as these methods exploit convexity and are guaranteed to converge to the real solution. Finally, we show how the duality inherent to convex and SDP problems can be exploited to prove analytically the optimality of a proposed solution. We give an example of how to apply this proof method by proving the optimality of Hardy and Song's optimal qubit θ shifter (e-print quant-ph/0102100).

The basic problem considered by a number of authors [1–3] is: what physically realizable quantum operation comes closest to a given, but potentially unphysical, transformation between quantum states? The operation is most generally described by a linear map $\mathcal{S}$; the physical realizability requires that the map is completely positive and trace preserving (CPTP). The desired transformation can be specified in a number of ways, for example, by enumerating all possible input-output pairs of pure states $\{|\text{in},k\rangle,|\text{out},k\rangle\}$. The dimensions of the input and output Hilbert spaces, $\mathcal{H}_{\text{in}}$ and $\mathcal{H}_{\text{out}}$, denoted d_1 and d_2, respectively, can in general be different. The symbol k labels the different pairs and can either be discrete or continuous.

In the most commonly used formalism [1,4,5], the CPTP map $\mathcal{S}$ that is to implement the transformation is represented by an operator X acting on the Hilbert space $\mathcal{H}_{\text{in}} \otimes \mathcal{H}_{\text{out}}$. Given any orthonormal basis $\{V_i\}$ for the space of all such operators, X is expressed as $X = \Sigma_i \mathcal{S}(V_i) \otimes V_i^*$, where the asterisk denotes complex (not Hermitian) conjugation. The requirements of complete positivity and trace preservation result in the constraints $X \geq 0$, $\text{Tr}_{\text{out}} X = \mathbb{1}_{\text{in}}$.

The requirement that the map must implement the transformation as closely as possible can be quantified by the mean fidelity F: $F = \Sigma_k \langle \text{out},k | \mathcal{S}(|\text{in},k\rangle\langle\text{in},k|) | \text{out},k \rangle$. The sum in this equation must be an integral with an appropriate measure for k if k is continuous. In terms of the operator X, the fidelity is given by $F = \text{Tr}\, XR$, with $R = \Sigma_k (|\text{in},k\rangle\langle\text{in},k|)^T \otimes |\text{out},k\rangle\langle\text{out},k|$. The great virtue of this measure-of-goodness of the map is that the fidelity is linear in the operator X. In this way the problem has been formulated as an optimization problem

$$P:\begin{cases} \text{maximize} \ \ \text{Tr}\, XR \\ X \geq 0 \\ \text{Tr}_{\text{out}} X = \mathbb{1}_{\text{in}}. \end{cases}$$

In general, optimization problem P cannot be solved analytically and one must resort to numerical methods. Most authors try to solve P using *ad hoc* iteration schemes involving Lagrange multipliers. Using these schemes, various useful results have been obtained. However, in our view, the convergence properties of these schemes are questionable, as it has not been proved that the solution obtained is actually the global optimum. In fact, these methods reportedly get stuck now and then in suboptimal local optima [6].

In this communication, we wish to draw attention to the fact that problem P belongs to a well-studied class of optimization problems called semidefinite programs (SDP). The importance of this fact cannot be overestimated. First of all, semidefinite programs are a subclass of the class of convex optimization problems, and convex problems have the very desirable property that a local optimum is automatically a global optimum. Keeping this in mind we see that the reported presence of local optima in the above iteration schemes is due to the scheme itself, and not to the problem being solved.

Second, very efficient numerical methods have been devised to solve SDPs, as these problems occur over and over again in various engineering disciplines, operations research, etc. These methods have very good convergence properties, and, moreover, their output includes a *certificate* of convergence consisting of a numerical interval within which the actual solution must lie. Using a sufficient number of iterations, the width of this interval can be made arbitrarily small (apart from numerical errors and given the validity of some technical requirements). In other words: convergence to the real solution is almost always guaranteed and certified.

Third, the way in which these numerical methods work can be exploited to prove analytically that a given proposed solution, e.g., an analytical Ansatz based on an educated guess and on the outcome of numerical experiments, is actually the correct solution.

In the rest of this section we will first discuss the basic mathematical facts of semidefinite programming and then apply them to the problem at hand. For a short introduction to the subject, we refer to [7], and for an in-depth treatment to [8]. Note that [9] presents another application of SDP to quantum mechanics, namely, to finding bounds on the distillable entanglement of mixed bipartite quantum states.

The basic SDP problem is the minimization of a linear function of a real variable $x \in R^m$, subject to a matrix inequality minimize $c^T x$, $F(x) = F_0 + \Sigma_{i=1}^m x_i F_i \geqslant 0$, where the $\geqslant$ sign means that $F(x)$ is positive semidefinite (hence, the term SDP). The problem data are the vector $c \in R^m$ and the $m+1$ real symmetric matrices F_i. Alternatively, the F_i can also be complex Hermitian but this is an atypical formulation within the SDP community (in engineering one typically deals with real quantities).

This problem is called the *primal* problem. Vectors x that satisfy the constraint $F(x) \geqslant 0$ are called *primal feasible points*, and if they satisfy $F(x) > 0$ they are called *strictly feasible points*. The minimal objective value $c^T x$ is by convention denoted as p^* (no complex conjugation) and is called the *primal optimal value*.

Of paramount importance is the corresponding *dual* problem, associated with the primal one, maximize $-\operatorname{Tr} F_0 Z$, $Z \geqslant 0$, $\operatorname{Tr} F_i Z = c_i$, $i = 1, \ldots, m$. Here the variable is the real symmetric (or Hermitian) matrix Z, and the data c, F_i are the same as in the primal problem. Correspondingly, matrices Z satisfying the constraints are called *dual feasible* (or *strictly dual feasible* if $Z > 0$). The maximal objective value $-\operatorname{Tr} F_0 Z$, the *dual optimal value*, is denoted as d^*.

The objective value of a primal feasible point is an upper bound on p^*, and the objective value of a dual feasible point is a lower bound on d^*. The main reason why one is interested in the dual problem is that one can prove that, under relatively mild assumptions, $p^* = d^*$. This holds, for example, if either the primal problem or the dual problem are strictly feasible, i.e., there either exist strictly primal feasible points or strictly dual feasible points. If this or other conditions are not fulfilled, we still have that $d^* \leqslant p^*$. Furthermore, when both the primal and dual problems are strictly feasible, one proves the following optimality condition on x: x is optimal if and only if x is primal feasible and there is a dual feasible Z such that $ZF(x) = 0$. This latter condition is called the *complementary slackness* condition.

In one way or another, numerical methods for solving SDP problems always exploit the inequality $d \leqslant d^* \leqslant p^* \leqslant p$, where d and p are the objective values for any dual feasible point and primal feasible point, respectively. The difference $p - d$ is called the duality gap, and the optimal value p^* is always "bracketed" inside the interval $[d, p]$. These numerical methods try to minimize the duality gap by subsequently choosing better feasible points. Under the requirements of the above-mentioned theorem, the duality gap can be made arbitrarily small (as far as numerical precision allows). This is precisely the reason why one should be happy when an optimization problem turns out to be an SDP problem.

We now apply these generalities to our problem at hand. Problem P can immediately be rewritten as a (primal) SDP problem by noting that the set of Hermitian matrices form a

real vector space of dimension the square of the matrix dimension. Since we are dealing with matrices over the bipartite Hilbert space $\mathcal{H}_{in} \otimes \mathcal{H}_{out}$ it is convenient to choose the basis vectors of the matrix space accordingly. Let $\{\sigma^j\}$ and $\{\tau^k\}$ be orthogonal bases for Hermitian matrices over $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$, respectively, then $\{\sigma^j \otimes \tau^k\}$ forms an orthogonal basis for $\mathcal{H}_{in} \otimes \mathcal{H}_{out}$. Furthermore, choose the bases so that both σ^0 and τ^0 are the identity matrix (of appropriate dimension) and all other σ^j and τ^k are traceless Hermitian matrices. An obvious choice would be the set of Pauli matrices $\{\sigma^x, \sigma^y, \sigma^z\}$ or generalizations thereof to higher dimensions. We thus have the following parametrization of the matrix X: $X = \Sigma_{j=0}^{d_1^2-1} \Sigma_{k=0}^{d_2^2-1} x_{jk} \sigma^j \otimes \tau^k$. With this parametrization, the TP requirement can be expressed in a straightforward way. The condition $\operatorname{Tr}_{out} X = \mathbb{1}_{in} = \sigma^0$ is fulfilled if and only if $x_{j0} = 0$ for all $j > 0$, and $x_{00} = 1/d_2$. By changing the parametrization of X, this can be taken care of implicitly, $X = \Sigma_{j=1}^{d_1^2-1} \Sigma_{k=1}^{d_2^2-1} x_{jk} \sigma^j \otimes \tau^k + \Sigma_{k=1}^{d_2^2-1} x_{0k} \sigma^0 \otimes \tau^k + \mathbb{1}/d_2$. From this parametrization, and the additional requirement $X \geqslant 0$, it immediately follows that the matrices F_i (in the SDP problem) are given by $F_0 = \mathbb{1}/d_2$, $F_{\cdots i \cdots} = \sigma^j \otimes \tau^k$, with $k \neq 0$. The index "i" in the left-hand side refers to the i of the SDP problem, and corresponds to all possible pairs (j, k) of right-hand side indices with $k \neq 0$. As a shorthand for summation over all these pairs we will use the symbol $\Sigma_{j,k}^*$.

Finally, we can assign values to the vector coefficients c_i as follows. The fidelity F is to be maximized, so we need an additional minus sign; furthermore, in terms of x_{jk}, F equals $F = \Sigma_{j,k}^* x_{jk} \operatorname{Tr}(\sigma^j \otimes \tau^k R) + 1/d_2$, where we have used the fact that $\operatorname{Tr} R = 1$. This yields for the coefficients c_i $c_{\cdots i \cdots} = -\operatorname{Tr}(\sigma^j \otimes \tau^k R)$, and for the optimal fidelity, in terms of the primal optimal value $F_{opt} = -p^* + 1/d_2$.

Using these expressions for the vector c and the matrices F_i (which are only dependent on the dimensions of the problem), one can go about solving the problem P numerically. Again we wish to make a case for the use of freely or commercially available SDP software in contrast to home-brewn solvers. Current SDP solving methods are highly optimized with respect to execution speed and convergence and are easy to use. As some of the F_i are complex, one has to use SDP software that explicitly allows for complex entries (e.g., [11]).

With the above assignments, the dual problem can now be formulated in a rather nice way. The dual objective, to be maximized over all $Z \geqslant 0$, is $d = -\operatorname{Tr} F_0 Z = -\operatorname{Tr} Z/d_2$. The constraint $\operatorname{Tr} F_i Z = c_i$ gets an interesting form $\operatorname{Tr}[\sigma^j \otimes \tau^k (Z + R)] = 0$, with $k \neq 0$. As Z and R are both Hermitian, this means that the matrix $Z + R$ must be of the form $Z + R = a_0 \mathbb{1} + \Sigma_{j \neq 0} a_j \sigma^j \otimes \mathbb{1}$, or, in other words, $Z = a_0 \mathbb{1} + A \otimes \mathbb{1} - R$, with A a traceless Hermitian matrix. With this parametrization for all dual feasible Z, the dual objective becomes $d = -d_1 a_0 + 1/d_2$. Maximizing d thus amounts to minimizing a_0 over all traceless Hermitian matrices A such that the resulting Z is still positive semidefinite. From the parametrization of Z one sees that the smallest feasible value of a_0 for a fixed matrix A is given by $a_0(A) = -\lambda_{min}(A \otimes \mathbb{1} - R)$, where

$\lambda_{\min}$ signifies the minimal eigenvalue of the matrix. The dual problem finally becomes: find the optimal traceless Hermitian matrix A such that this $a_0(A)$ is minimal. The dual optimal value is then $d^* = -d_1 \min_A a_0(A) + 1/d_2$. Note that we have significantly reduced the number of unknown parameters: from $(d_1 d_2)^2$ for Z to $d_1^2 - 1$ for A.

These expressions for the primal and dual problems can be used for proving that a certain proposed solution is optimal. To that purpose one needs to propose primal and dual feasible points x and A; if the resulting primal and dual objective values p and d turn out to be equal to each other, then x and A are optimal feasible points and $p = d = p^* = d^*$. Alternatively, any feasible choice for x and A gives upper and lower bounds on the optimal value p^*, resulting in lower and upper bounds, respectively, for the fidelity of problem P. For example, setting $A = 0$ gives $a_0(A) = \lambda_{\max}(R)$ resulting in the upper bound $F \leq d_1 \lambda_{\max}(R)$, which was already derived in [3].

Using the method of the previous paragraph, one can test whether the feasible points are optimal or not, but it does not solve the problem of finding these points. As there is no hope for solving the primal and dual problems analytically for all but the simplest problems, one must resort to numerical methods. Luckily, efficient methods abound and some implementations are freely available on the web. From the numerical results one can then try to guess the analytical form of the solution, or at least try to propose an Ansatz containing a few unknown parameters. If the number of parameters is small they could be found by solving the primal and dual problem using the Ansatz.

Even this could be relatively complicated, especially for the dual problem, as this is an eigenvalue problem. An alternative for solving the dual problem is offered by the complementary slackness (CS) condition, which does not require solving an eigenvalue equation. Supposing that a correct guess has been made for X of the primal problem, one then has to solve the *linear* equation $(a_0 1 + A \otimes 1 - R)X = 0$ in the unknowns a_0 and A. Of course, one then still has to prove that the resulting Z is dual feasible, i.e., is positive semidefinite, and this could still require solving an eigenvalue problem.

As an example of this proof technique, we now consider the problem of constructing an optimal qubit θ shifter, first considered by Hardy and Song [10] and prove that their "quantum scheme" shifter (see also [3]) is optimal.

A qubit θ shifter is a device that transforms a pure state $\psi(\theta, \phi) = \cos(\theta/2)|0\rangle + \exp(i\phi)\sin(\theta/2)|1\rangle$ into another pure state $\psi(\theta + \alpha, \phi)$. This is a nonphysical operation and has, therefore, to be approximated. Hardy and Song consider both a universal approximated shifter, with fidelity independent of θ, and a shifter with θ-dependent fidelity optimizing the *mean* fidelity. The mean fidelity of the nonuniversal shifter is better than for the universal one. Hardy and Song proposed an optimal scheme ("quantum scheme") for θ shifting and proved its optimality using a method of Buzek, Hillery, and Werner [12]. We will now do the same, but using our general approach.
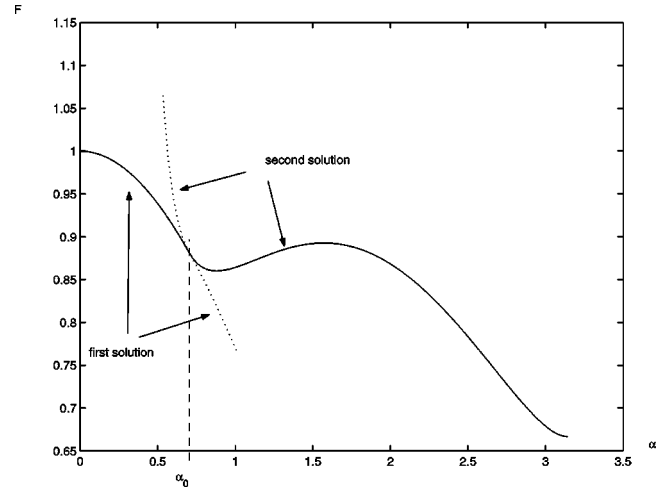
The matrix R for the shifter is given by

FIG. 1. Optimal mean fidelity F of the θ shifter in terms of the shifting angle α, obtained from solving the associated SDP problem. The dotted lines represent the two solutions outside of their respective region of applicability. The full line represents the actual solution.

$$R = \begin{bmatrix} r_1 & 0 & 0 & r_5 \\ 0 & r_2 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ r_5 & 0 & 0 & r_4 \end{bmatrix},$$

with $r_1 = 1/4 + c - s$, $r_2 = 1/4 - c + s$, $r_3 = 1/4 - c - s$, $r_4 = 1/4 + c + s$, $r_5 = 2c$, $c = \frac{1}{12}\cos\alpha$, and $s = (\pi/16)\sin\alpha$.

The ansatz for the primal feasible point, which was already given by Fiurasek [3], is

$$X = \begin{bmatrix} \cos^2\beta & 0 & 0 & \cos\beta \\ 0 & \sin^2\beta & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \cos\beta & 0 & 0 & 1 \end{bmatrix}.$$

There are two regimes, depending on the value of α. For $\alpha \leq \alpha_0 = \arctan(8/3\pi)$, put $\cos\beta = 1$, and for $\alpha \geq \alpha_0$, $\cos\beta = c/(s-c)$. This gives as primal objective fidelity $F = (1 + \cos\alpha)/2$ for $\alpha \leq \alpha_0$, $F = 1/2 + 2s + 2c^2/(s-c)$, for $\alpha \geq \alpha_0$. Going over to the dual problem, we now present our own Ansatz for the dual feasible point A, which was inspired by numerical results: simply consider *diagonal* A only. This means that A is parametrized by a single number, say δ, and equals $A = \delta\sigma^z$. This gives for Z

$$Z = \begin{bmatrix} a_0 + \delta - r_1 & 0 & 0 & -r_5 \\ 0 & a_0 + \delta - r_2 & 0 & 0 \\ 0 & 0 & a_0 - \delta - r_3 & 0 \\ -r_5 & 0 & 0 & a_0 - \delta - r_4 \end{bmatrix}.$$

To prove optimality of Fiurasek's primal Ansatz and our dual Ansatz, we use the complementary slackness condition (for finding the optimal value for a_0 and δ). The CS condition $ZX=0$ gives rise to just three independent equations ($(a_0 + \delta - r_1)\cos\beta - r_5 = 0$, $(a_0 + \delta - r_2)\sin^2\beta = 0$, $(a_0 - \delta - r_4) - r_5\cos\beta = 0$. As could be expected, there are two different solutions $a_0 = 1/4 + 3c$, $\delta = -s$, $\cos\beta = 1$, and $a_0 = 1/4 + s + c^2/(s-c)$, $\delta = -s/(s-c)$, $(r_2 - r_1)\cos\beta = r_5$. The third equation of each set shows us that the first solution pertains to the case $\alpha \leq \alpha_0$ and the second solution to the other case. The mean fidelity, resulting from these solutions is then

$$F = \begin{cases} 2(1/4+3c) = (1+\cos\alpha)/2, & \text{for } \alpha \leq \alpha_0, \\ 1/2+2s+2c^2/(s-c) = 1/2 + (\pi/8)\sin\alpha + 2\cos^2\alpha/(9\pi\sin\alpha - 12\cos\alpha), & \text{for } \alpha > \alpha 0, \end{cases} \tag{1}$$

which is depicted in Fig. 1. These values are exactly the ones obtained in the Primal problem, so this proves the optimality of the two Ansatze, provided $Z \geq 0$ in both cases. It is a basic exercise in linear algebra to calculate the eigenvalues of Z in both cases; noting that $0 \leq s \leq 2c$ in the case $\alpha \leq \alpha_0$, and $c \leq s$ in the other case, one can indeed show that Z is always positive semidefinite, proving its feasibility.

To conclude, we have noted that the problem P, which has to be solved for finding CPTP maps that optimally approximate certain desired qubit transformations, is a semidefinite programming (SDP) problem. From this observation, it follows that P can be efficiently solved using standard SDP software, and that there is no need for *ad hoc* solution methods, which could suffer from bad convergence properties. Furthermore, we presented a method for proving analytically that an Ansatz for the solution of P is optimal. We hope that the present communication will be useful for those working in the field of determining optimal CP maps or optimal quantum measurements.

[1] M.F. Sacchi, Phys. Rev. A **63**, 054104 (2001).

[2] S. Massar and S. Popescu, Phys. Rev. A **61**, 062303 (2000).

[3] J. Fiurasek, e-print quant-ph/0105124 (2001).

[4] J. de Pillis, Pac. J. Math. **23**, 129 (1967).

[5] A. Jamiolkowski, Rep. Math. Phys. **3**, 275 (1972).

[6] J. Fiurasek (private communication, 2001).

[7] L. Vandenberghe and S. Boyd, SIAM Rev. **38**, 49 (1996).

[8] Y. Nesterov and A. Nemirovsky, *Interior-Point Polynomial Methods in Convex Programming*, Vol. 13 of *Studies in Applied Mathematics* (SIAM, Philadelphia, 1994).

[9] E. Rains, IEEE Trans. Inf. Theory (to be published).

[10] L. Hardy and D.D. Song, Phys. Rev. A **63**, 032304 (2001); **64**, 032301 (2001).

[11] J.F. Sturm, Optimization Methods and Software **11−12**, 625 (1999). Software available at http://members.tripodnet.nl/SeDuMi

[12] V. Buzek, M. Hillery, and F. Werner, J. Mod. Opt. **47**, 2112 (2001).