Available online at www.sciencedirect.com

SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER Linear Algebra and its Applications 412 (2006) 453–470

www.elsevier.com/locate/laa

Least-squares orthogonalization using semidefinite programming

Yonina C. Eldar *

Department of Electrical Engineering, Technion—Israel Institute of Technology, Technion city, Haifa 32000, Israel

Received 22 October 2002; accepted 17 July 2005
Available online 16 September 2005
Submitted by H. Schneider

Abstract

We consider the problem of constructing an optimal set of orthogonal vectors from a given set of vectors in a real Hilbert space. The vectors are chosen to minimize the sum of the squared norms of the errors between the constructed vectors and the given vectors. We show that the design of the optimal vectors, referred to as the least-squares (LS) orthogonal vectors, can be formulated as a semidefinite programming (SDP) problem. Using the many well-known algorithms for solving SDPs, which are guaranteed to converge to the global optimum, the LS vectors can be computed very efficiently in polynomial time within any desired accuracy.

By exploiting the connection between our problem and a quantum detection problem we derive a closed form analytical expression for the LS orthogonal vectors, for vector sets with a broad class of symmetry properties. Specifically, we consider geometrically uniform (GU) sets with a possibly non-abelian generating group, and compound GU sets which consist of subsets that are GU.

AMS classification: 65F25; 15A04; 15A23; 15A99

Keywords: Orthogonalization; Least-squares; Semidefinite programming; Quantum detection; Geometric uniformity

* Tel.: +972 4 829 3256; fax: +972 4 829 5757.
E-mail address: yonina@ee.technion.ac.il

doi:10.1016/j.laa.2005.07.010

1. Introduction

Constructing a set of orthogonal vectors from a given set of vectors is a well-known problem. Specifically, let $\{s_k, 1 \leqslant k \leqslant N\}$ denote a set of N vectors in a real Hilbert space $\mathscr{H}$. Then the problem is to construct a set of orthogonal vectors $\{h_k, 1 \leqslant k \leqslant N\}$ from the vectors $\{s_k, 1 \leqslant k \leqslant N\}$.

One of the most common procedures to orthogonalize a set of vectors is the Gram–Schmidt (GS) [23] method. The GS orthogonal vectors h_k are chosen such that h_1 is in the direction of s_1, h_2 is in the direction of s_2 perpendicular to s_1, and so forth. Thus, h_1 is perfectly aligned with s_1, however h_N may be relatively "far" from s_N. It is immediately obvious that the GS vectors depend on the order in which the vectors $\{s_k\}$ are arranged. Furthermore, the GS approach is not known to be optimal in any sense.

Recently a new method of orthogonalization, or more generally, inner product shaping, has been proposed [10,11] which does not depend on the order of the vectors $\{s_k\}$, and is optimal in a least-squares (LS) sense. Specifically, the constructed vectors h_k are chosen to have some desired inner product structure, so that $\langle h_m, h_k\rangle = r_{mk}$ for a set of numbers r_{mk}, and at the same time to minimize the sum of the squared norms of the error vectors

$$E = \sum_{k=1}^{N}\langle s_k - h_k, s_k - h_k\rangle. \tag{1}$$

In the case in which the desired inner products r_{mk} are specified, the form of the optimal vectors was derived in [10,11]. When the eigenvectors of the inner product matrix with elements r_{mk} are known, and the eigenvalues are chosen to minimize the LS error (1), an analytical solution is hard to obtain. Finding the optimal vectors in this case involves solving a *LS orthogonalization problem*, which is the problem of constructing an orthogonal set of vectors h_k satisfying

$$\langle h_m, h_k\rangle = 0 \quad \text{for } m \neq k \tag{2}$$

that minimize (1), so that they are closest in an LS sense to the given vectors s_k. Note, that in this problem, the norms of the constructed vectors $\langle h_k, h_k\rangle$ are not specified but rather chosen to minimize the LS error (1).

The LS orthogonalization problem of (1) and (2) was first treated in [10], in which it was shown that in general obtaining a closed form analytical expression for the optimal vectors is a difficult problem. An iterative method that is guaranteed to converge to a *local* optimum was proposed. For the special case in which the vectors s_k are defined over a finite *abelian* group of unitary matrices and generated by a single generating vector, a closed form analytical solution for the LS vectors was derived.

In this paper we show that the orthogonal LS vectors can be found by solving a semidefinite programming (SDP) problem, which is a convex optimization problem. By exploiting the many well-known algorithms for solving SDPs [29,2,3,26], the optimal vectors can be computed very efficiently in polynomial time within any desired accuracy. Furthermore, in contrast to the iterative algorithm proposed in [10] which is only guaranteed to converge to a local optimum, algorithms based on SDP are guaranteed to converge to the *global* optimum.

To derive the LS orthogonal vectors as a solution to an SDP, we first formulate (1) and (2) as a non-convex problem in a higher dimension. We then obtain an SDP relaxation of the problem, and establish that the relaxed and original problem have the same solution. Using the SDP formulation, we develop necessary and sufficient optimality conditions on the orthogonal vectors. These conditions are shown to be equivalent to optimality conditions that arise in a certain detection problem in quantum mechanics. We then rely on results obtained in that context to derive simple closed form expressions for the LS orthogonal vectors for vector sets with various symmetry properties, generalizing the result in [10]. Specifically, we consider the case in which the original vector set is defined over a finite group of unitary matrices and generated by a single generating vector, where the group is not constrained to be abelian as in [10]. Such vector sets are referred to as *geometrically uniform* (GU). We then treat *compound GU* (CGU) [12] vector sets in which the vectors are generated by a group of unitary matrices using *multiple* generators. In this case we show that the LS orthogonal vectors are also CGU with generators that can be computed by solving a reduced-size SDP. When the generators of the CGU vector set satisfy a certain constraint, a closed form analytical expression for the optimal generators is given.

After a statement of the problem in Section 2, we develop, in Sections 3 and 4, an SDP representation of our problem. Using this formulation, in Section 5 we consider efficient iterative algorithms that are guaranteed to converge to the globally optimum vectors. We then illustrate the computational steps in the context of a concrete example in Section 6. In Section 7 we rely on results obtained in the context of quantum detection to derive explicit expressions for the LS vectors for vector sets with a broad class of symmetry properties.

2. Least-squares orthogonalization

We denote vectors in a real Hilbert space $\mathscr{H}$ by lowercase letters. General linear transformations are denoted by uppercase letters. $P_{\mathscr{U}}$ denotes the orthogonal projection onto the subspace $\mathscr{U}$, $(\cdot)^*$ and $\mathrm{Tr}(\cdot)$ denote the adjoint and the trace, respectively.

Suppose we are given a set of N vectors $\{s_k, 1 \leqslant k \leqslant N\}$ in a real Hilbert space $\mathscr{H}$, with inner product $\langle x, y\rangle$ for any $x, y \in \mathscr{H}$. The vectors $\{s_k\}$ span an M-dimensional subspace $\mathscr{U} \subseteq \mathscr{H}$. If the vectors are linearly independent, then $M = N$; otherwise $M < N$. Our objective is to construct a set of orthogonal vectors $\{h_k, 1 \leqslant k \leqslant N\}$ from the given vectors $\{s_k, 1 \leqslant k \leqslant N\}$ that are "closest" to the vectors s_k in the LS sense. Thus we seek the vectors h_k that minimize the LS error (1) subject to (2).

Several potential applications of LS orthonormalization, in which the vectors h_k are constrained to have unit norm, are discussed in [10]. Each of these problems can be generalized to the case in which we allow for additional freedom by also optimizing the norms of the vectors h_k, and not constraining them a priori. For example, a potential application of LS orthonormalization, mentioned in [10], is to a generic classical detection problem in which one of a set of known signals is transmitted, and the objective is to detect the transmitted signal from the signal which has been received over an additive noise channel. The typical receiver used in such problems is the well-known matched filter receiver which consists of correlating the received signal with the possible transmitted signals. We can improve the performance over the matched filter receiver in many cases by correlating the received signal with a set of signals with a specified inner product structure, tailored to the particular problem, that are closest in an LS sense to the transmitted signals [13,11]. Similar applications have also been explored for suppressing interference in multiuser wireless communication systems [14–16]. In such detection problems, we may be able to further improve the performance, by also optimizing the inner products, rather than choosing them a priori. As we noted in the introduction, this involves solving a LS orthogonalization problem.

Expressing the error E of (1) as

$$E = \sum_{k=1}^{N} (\langle s_k, s_k\rangle + \langle h_k, h_k\rangle - 2\langle h_k, s_k\rangle)$$

and defining a set of orthonormal vectors y_k such that $b_k y_k = h_k$ where $b_k^2 = \langle h_k, h_k\rangle$, it follows that minimization of E is equivalent to minimization of

$$E' = \sum_{k=1}^{N} (\langle h_k, h_k\rangle - 2\langle h_k, s_k\rangle) = \sum_{k=1}^{N} \left(b_k^2 - 2b_k\langle y_k, s_k\rangle\right). \tag{3}$$

To determine the optimal vectors h_k we have to minimize E' with respect to b_k and y_k. Fixing y_k and minimizing with respect to b_k, the optimal value of b_k, denoted $\hat{b}_k$, is given by

$$\hat{b}_k = \langle y_k, s_k\rangle.$$

Substituting $\hat{b}_k$ back into (3), the vectors y_k are chosen to maximize

$$R_{ys} = \sum_{k=1}^{N} \langle y_k, s_k\rangle^2 \tag{4}$$

subject to the constraint

$$\langle y_m, y_k\rangle = \delta_{mk}, \tag{5}$$

where $\delta_{mk} = 1$ if $m = k$, and 0 otherwise.

Obtaining a closed form analytical expression for the orthonormal vectors y_k that maximize (4) is in general a difficult problem. An iterative algorithm for computing the optimal vectors that is guaranteed to converge to a *local* optimum was proposed in [10].

In the next section we show that the solution to the problem of (4) and (5) can be obtained by solving an SDP, which is a convex optimization problem. By exploiting the many well-known algorithms for solving SDPs [26,2], the optimal vectors can be computing very efficiently in polynomial time within any desired accuracy. Since an SDP is convex, it does not suffer from local optima, so that SDP-based algorithms are guaranteed to converge to the *global* optimum.

3. Semidefinite relaxation

To develop an efficient algorithm for computing the LS orthogonal vectors, we begin by expressing our problem as a constrained optimization problem in a higher dimension.

3.1. Alternative formulation

We first note that the vectors $\{y_k, 1 \leqslant k \leqslant N\}$ are orthonormal if and only if $\sum_{k=1}^{N} y_k y_k^* = P_{\mathscr{V}}$ where $P_{\mathscr{V}}$ is the orthogonal projection onto the N-dimensional space $\mathscr{V}$, spanned by the vectors y_k. Since the vectors y_k are chosen to maximize R_{ys}, it is obvious that $\mathscr{U} \subseteq \mathscr{V}$ where $\mathscr{U}$ is the M-dimensional space spanned by the vectors s_k, and $M \leqslant N$. Now because $s_k \in \mathscr{U}$, we can write $s_k = P_{\mathscr{U}} s_k$, so that R_{ys} of (4) can be expressed as

$$R_{ys} = \sum_{k=1}^{N} \langle y_k, P_{\mathscr{U}} s_k \rangle^2 = \sum_{k=1}^{N} \langle P_{\mathscr{U}} y_k, s_k \rangle^2 = \sum_{k=1}^{N} \mathrm{Tr}(\Sigma_k \Pi_k),$$

where $\Pi_k = P_{\mathscr{U}} y_k y_k^* P_{\mathscr{U}}$ and $\Sigma_k = s_k s_k^*$. Since for any choice of orthonormal vectors y_k we must have

$$\sum_{k=1}^{N} \Pi_k = \sum_{k=1}^{N} P_{\mathscr{U}} y_k y_k^* P_{\mathscr{U}} = P_{\mathscr{U}} \left(\sum_{k=1}^{N} y_k y_k^* \right) P_{\mathscr{U}} = P_{\mathscr{U}} P_{\mathscr{V}} P_{\mathscr{U}} = P_{\mathscr{U}}$$

the problem of (4) and (5) reduces to finding a set of operators Π_k that maximize

$$J(\{\Pi_k\}) = \sum_{k=1}^{N} \mathrm{Tr}(\Sigma_k \Pi_k) \tag{6}$$

subject to

$$\sum_{k=1}^{N} \Pi_k = P_{\mathscr{U}}; \tag{7}$$

$$\Pi_k = q_k q_k^*, \quad 1 \leqslant k \leqslant N. \tag{8}$$

The conditions (7) and (8) together imply that the vectors $\{q_k, 1 \leqslant k \leqslant N\}$ form a *normalized tight frame*[1] [9,19] for the M-dimensional space $\mathscr{U}$.

Note that in the case in which $M < N$, the optimal orthonormal vectors are not unique. In particular, we have seen that R_{ys} of (4), equivalently J of (6), depend on the vectors y_k only through the projections $q_k = P_{\mathscr{U}} y_k$. Once we obtain the vectors $\hat{q}_k$ that maximize (6), the optimal orthonormal vectors $\hat{y}_k$ can be chosen as any set of orthonormal vectors whose projections onto $\mathscr{U}$ satisfy $P_{\mathscr{U}}\hat{y}_k = \hat{q}_k$. For example, let $\widehat{Q}$ be the matrix with columns $\hat{q}_k$ and let $\widehat{Q}$ have a singular value decomposition $\widehat{Q} = U\Sigma V^*$ where U is a matrix with N orthonormal vectors where the first M vectors span $\mathscr{U}$, Σ is an $N \times N$ diagonal matrix and V is an $N \times N$ unitary matrix. To satisfy (7), the first M diagonal elements of Σ must be equal to 1 and the remaining diagonal elements must be equal 0. Then, we may choose the vectors $\hat{y}_k$ as the columns of

$$\widehat{Y} = UV^*. \tag{9}$$

In the remainder of the paper we focus on finding the optimal projections $\hat{q}_k$. The LS vectors $\hat{h}_k$ are then equal to $\hat{h}_k = \langle \hat{q}_k, s_k\rangle \hat{y}_k$ with the vectors $\hat{y}_k$ given e.g., by (9).

3.2. Convex relaxation

The constraint (8) implies that Π_k is symmetric, positive semidefinite (PSD), and of rank 1. Due to this constraint, the problem of (6)–(8) is a non-convex optimization problem and therefore hard to solve. Removing the rank-1 constraint results in the relaxed *convex* problem of maximizing (6) subject to

$$\begin{aligned} &\sum_{k=1}^{N} \Pi_k = P_{\mathscr{U}};\\ &\Pi_k \geqslant 0, \quad 1 \leqslant k \leqslant N, \end{aligned} \tag{10}$$

where $\Pi_k \geqslant 0$ means that Π_k is symmetric and PSD. We note that a similar approach has been used in the context of Euclidean distance matrix problems; see [1,24].

Let $\widehat{J}$ and $\widehat{J}_{\mathrm{R}}$ denote the optimal values of the original and relaxed problems, respectively. Clearly, $\widehat{J}_{\mathrm{R}} \geqslant \widehat{J}$, so that the optimal value of the relaxed convex problem provides an upper bound on the optimal value of the original non-convex problem. If in addition we can show that the operators $\widehat{\Pi}_k$ that maximize the relaxed problem have rank 1, then it follows that $\widehat{J}_{\mathrm{R}} = \widehat{J}$ and that these operators are also optimal for the original problem.

In the next section we establish the equivalence between the original and relaxed problems by examining the optimality conditions associated with the SDP relaxation.

[1] A set of vectors $\{x_k, 1 \leqslant k \leqslant N\}$ forms a normalized tight frame for an M-dimensional space $\mathscr{V}$ if $\sum_{k=1}^{N} x_k x_k^* = P_{\mathscr{V}}$. If $M = N$, then the vectors x_k are orthonormal; however, when $M < N$, the vectors x_k are linearly dependent, and therefore cannot be orthonormal.

4. Optimality of the semidefinite relaxation

4.1. Optimality conditions

Since the problem of maximizing $J(\{\Pi_k\})$ of (6) subject to (10) is convex and strictly feasible (i.e. there exists operators $\Pi_k > 0$ such that $\sum_k \Pi_k = P_{\mathscr{U}}$), the Karush–Kuhn–Tucker (KKT) conditions [6] are necessary and sufficient for optimality. In our case, these conditions imply that $\{\Pi_k \in \mathscr{U}\}$ are optimal if and only if there exists matrices Δ and $\Psi_k \geqslant 0$, $1 \leqslant k \leqslant N$ such that:

(1) $d\mathscr{L}/d\Pi_k = 0$, $1 \leqslant k \leqslant N$ where the Lagrangian $\mathscr{L}$ is defined by

$$\mathscr{L} = -\sum_{k=1}^{N} \operatorname{Tr}(\Sigma_k \Pi_k) + \operatorname{Tr}\left(\Delta\left(\sum_{k=1}^{N} P_{\mathscr{U}} \Pi_k P_{\mathscr{U}} - P_{\mathscr{U}}\right)\right) - \sum_{k=1}^{N} \operatorname{Tr}(\Psi_k \Pi_k).$$

(2) Feasibility: $\sum_k \Pi_k = P_{\mathscr{U}}$ and $\Pi_k \geqslant 0$, $1 \leqslant k \leqslant N$.

(3) Complementary slackness: $\operatorname{Tr}(\Psi_k \Pi_k) = 0$, $1 \leqslant k \leqslant N$.

In writing the Lagrangian we used the fact that the condition $\sum_{k=1}^{N} \Pi_k = P_{\mathscr{U}}$ can equivalently be written as $\sum_{k=1}^{N} P_{\mathscr{U}} \Pi_k P_{\mathscr{U}} = P_{\mathscr{U}}$.

Differentiating the Lagrangian with respect to Π_k and equating to 0,

$$\Delta_{\mathscr{U}} \triangleq P_{\mathscr{U}}\, \Delta P_{\mathscr{U}} = \Sigma_k + \Psi_k. \tag{11}$$

Since $\Psi_k \geqslant 0$, (11) implies that

$$\Delta_{\mathscr{U}} \geqslant \Sigma_k, \quad 1 \leqslant k \leqslant N. \tag{12}$$

Substituting (11) into the complementary slackness condition,

$$\sum_{k=1}^{N} \operatorname{Tr}\left((\Delta_{\mathscr{U}} - \Sigma_k)\Pi_k\right) = 0. \tag{13}$$

Since $\Delta_{\mathscr{U}} \geqslant \Sigma_k$ and $\Pi_k \geqslant 0$, (13) is satisfied if and only if

$$(\Delta_{\mathscr{U}} - \Sigma_k)\Pi_k = \Pi_k(\Delta_{\mathscr{U}} - \Sigma_k) = 0, \quad 1 \leqslant k \leqslant N. \tag{14}$$

Summing both sides of (14) over k and using (10), we have that at the optimal solution $\widehat{\Pi}_k$,

$$\sum_{k=1}^{N} \Sigma_k \widehat{\Pi}_k = \sum_{k=1}^{N} \widehat{\Pi}_k \Sigma_k. \tag{15}$$

Summing only one side of (14) over k, leads to $\Delta_{\mathscr{U}} = \sum_{k=1}^{N} \Sigma_k \widehat{\Pi}_k$, which together with (12) implies that

$$\sum_{k=1}^{N} \Sigma_k \widehat{\Pi}_k \geqslant \Sigma_j, \quad 1 \leqslant j \leqslant N. \tag{16}$$

We conclude that if $\widehat{\Pi} = \{\widehat{\Pi}_k\}_{k=1}^N$ is optimal, then it must satisfy (15) and (16). The conditions (15) and (16) together with (10) are also sufficient. Indeed, suppose that the operators $\widehat{\Pi}_k$ satisfy (15) and (16). Since $\widehat{\Pi}_k \in \mathscr{U}$, we can define $\Delta_{\mathscr{U}} = \sum_{k=1}^N \widehat{\Pi}_k \Sigma_k$ and $\Psi_k = \Delta_{\mathscr{U}} - \Sigma_k$. It is easy to see that with this choice, the KKT conditions are satisfied.

Substituting (11) back into the Lagrangian, we can immediately determine the dual problem associated with (6) and (10), which becomes

$$\min_{\Delta}\{\mathrm{Tr}(\Delta P_{\mathscr{U}}) : \Delta \geqslant \Sigma_k, 1 \leqslant k \leqslant N\}. \tag{17}$$

As we will see in Section 5, the dual is useful for deriving efficient computational algorithms.

Note that the dual solution Δ can always be chosen to satisfy $\Delta = P_{\mathscr{U}} \Delta P_{\mathscr{U}}$. We therefore denote the solution by $\Delta_{\mathscr{U}}$.

4.2. *Equivalence of relaxed and original problems*

We now use the KKT conditions for optimality to show that if $\Sigma_k = s_k s_k^*$ for some set of vectors s_k spanning $\mathscr{U}$, then $\widehat{\Pi}_k = q_k q_k^*$ for a set of vectors q_k, so that the solution of the relaxed convex problem, and the original non-convex problem, are the same.

From (14) it follows that the operators $\widehat{\Pi}_k$ must lie in the intersection of $\mathscr{U}$ and the null space of $\Delta_{\mathscr{U}} - \Sigma_k$. We denote this intersection by $\mathscr{N}_{\mathscr{U}}(\Delta_{\mathscr{U}} - \Sigma_k)$. Consequently $\mathrm{rank}(\widehat{\Pi}_k) \leqslant \dim(\mathscr{N}_{\mathscr{U}}(\Delta_{\mathscr{U}} - \Sigma_k))$. Because $\Delta_{\mathscr{U}} \geqslant \Sigma_k = s_k s_k^*$, $1 \leqslant k \leqslant N$, it follows that $\Delta_{\mathscr{U}}$ is positive definite on $\mathscr{U}$. Indeed, since the vectors $\{s_k, 1 \leqslant k \leqslant N\}$ span $\mathscr{U}$, for any $u \in \mathscr{U}$ there exists a k such that $|\langle u, s_k\rangle|^2 = \langle u, \Sigma_k u\rangle > 0$, which implies that $\langle u, \Delta_{\mathscr{U}} u\rangle > 0$ for any $u \in \mathscr{U}$. Therefore, $\dim(\mathscr{N}_{\mathscr{U}}(\Delta_{\mathscr{U}} - \Sigma_k)) \leqslant 1$ and

$$\mathrm{rank}(\widehat{\Pi}_k) \leqslant \dim(\mathscr{N}(\Delta_{\mathscr{U}} - \Sigma_k)) \leqslant 1, \quad 1 \leqslant k \leqslant N.$$

We conclude that the operators maximizing the relaxed problem of (6) subject to (10) have the form $\widehat{\Pi}_k = q_k q_k^*$ for a set of vectors q_k. This then implies that the solution to the original problem of (6) subject to (7) and (8) is equal to the solution of the relaxed problem, which is a convex SDP.

We summarize our results in the following theorem:

Theorem 1. *Let $\{s_k, 1 \leqslant k \leqslant N\}$ denote a set of N vectors in a real Hilbert space $\mathscr{H}$, that span an M-dimensional subspace $\mathscr{U} \subseteq \mathscr{H}$. Let $\{\hat{h}_k, 1 \leqslant k \leqslant N\}$ denote N orthogonal vectors that minimize the LS error defined by* (1). *Then $\hat{h}_k$ can be chosen as $\hat{h}_k = \langle \hat{q}_k, s_k\rangle \hat{y}_k$ where*

(1) *the vectors $\hat{q}_k$ form a normalized tight frame for $\mathscr{U}$ and can be obtained as the solution to the semidefinite programming problem of maximizing $J(\{\Pi_k\}) = \sum_{k=1}^N \mathrm{Tr}(\Sigma_k \Pi_k)$ subject to $\Pi_k \geqslant 0$, $1 \leqslant k \leqslant N$ and $\sum_{k=1}^N \Pi_k = P_{\mathscr{U}}$, where $\Sigma_k = s_k s_k^*$. The optimal vectors $\hat{q}_k$ are then given by $\widehat{\Pi}_k = \hat{q}_k \hat{q}_k^*$, where $\{\widehat{\Pi}_k\}$ denote the optimal $\{\Pi_k\}$;*

(2) *the vectors* $\hat{q}_k$ *are optimal if and only if they satisfy* $\sum_{k=1}^{N} \hat{q}_k \hat{q}_k^* = P_{\mathscr{U}}$, *and*

$$\sum_{k=1}^{N} \Sigma_k \widehat{\Pi}_k = \sum_{k=1}^{N} \widehat{\Pi}_k \Sigma_k;$$

$$\sum_{k=1}^{N} \Sigma_k \widehat{\Pi}_k \geqslant \Sigma_j, \quad 1 \leqslant j \leqslant N;$$

(3) *the vectors* $\hat{y}_k$ *are a set of orthonormal vectors such that* $\hat{y}_k = P_{\mathscr{U}} \hat{q}_k$.

Consider the dual problem of maximizing $T(\Delta) = \mathrm{Tr}(P_{\mathscr{U}} \Delta)$ *subject to* $\Delta \geqslant \Sigma_k$, $1 \leqslant k \leqslant N$, *and let* $\widehat{\Delta}$ *denote an optimal* Δ. *Then a necessary and sufficient condition on the optimal vectors* $\hat{q}_k$ *is* $(\Delta_{\mathscr{U}} - \Sigma_k)\hat{q}_k = 0$, $1 \leqslant k \leqslant N$, *where* $\Delta_{\mathscr{U}} = P_{\mathscr{U}} \widehat{\Delta} P_{\mathscr{U}}$.

Except in some particular cases [22,8,27,5,18,17], obtaining a closed-form analytical expression for the optimal vectors $\hat{q}_k$ directly from the necessary and sufficient conditions of Theorem 1 is a difficult and unsolved problem. However, since our problem can be formulated as a (convex) SDP [29,2,26], there are very efficient methods for its solution.

In the next section we develop fast computational methods for finding the optimal normalized tight frame vectors $\hat{q}_k$. In Section 7 we consider some special cases in which these vectors have a closed form expression, by applying results developed in the context of quantum detection.

5. Computational aspects

To develop efficient computational methods, we note that the dual problem (17) involves fewer decision variables than the primal maximization problem (6). Therefore, it is advantageous to solve the dual problem and then use (14) with $\widehat{\Pi}_k = \hat{q}_k \hat{q}_k^*$ to determine the optimal vectors $\hat{q}_k$, rather than solving the primal problem directly.

The operator $\Delta_{\mathscr{U}}$ that minimizes $\mathrm{Tr}(P_{\mathscr{U}} \Delta)$ subject to $\Delta \geqslant \Sigma_k$, $1 \leqslant k \leqslant N$ can be computed in Matlab using the linear matrix inequality (LMI) Toolbox. Convenient interfaces for using the LMI toolbox are the Matlab packages IQCβ [25] and self-dual-minimization (SeDuMi) [28] (together with the Yalmip interface).

Once we determine $\Delta_{\mathscr{U}}$, the optimal operators $\widehat{\Pi}_k = \hat{q}_k \hat{q}_k^*$ can be computed using (14) and (7). Specifically, from (14) it follows that $\widehat{\Pi}_k$ can be expressed as

$$\widehat{\Pi}_k = a_k x_k x_k^*, \tag{18}$$

where $a_k \geqslant 0$, and x_k is a normalized vector that spans the null space of $\Delta_{\mathscr{U}} - \Sigma_k$ which can be determined using the eigendecomposition of $\Delta_{\mathscr{U}} - \Sigma_k$. To satisfy (7) we must have

$$\sum_{k=1}^{N} a_k x_k x_k^* = P_{\mathscr{U}}. \tag{19}$$

Let $e = \text{vec}(P_{\mathscr{U}})$ and $z_k = \text{vec}(x_k x_k^*)$, where $v = \text{vec}(V)$ denotes the vector obtained by stacking the columns of V. Then we can express (19) as

$$Za = e, \tag{20}$$

where Z is the matrix of columns z_k and a is the vector with components a_k. If the matrix Z has full column rank, then the unique solution to (20) is

$$a = (Z^*Z)^{-1}Z^*e.$$

In the general case, Z will not have full column rank and there will be many solutions a to (20). Each such vector defines a corresponding set of optimal operators $\widehat{\Pi}_k$ via (18). To find a unique solution we may seek the vector[2] $a \geqslant 0$ that satisfies (20), and such that $\sum_{k=1}^{N} \text{Tr}(\widehat{\Pi}_k) = \sum_{k=1}^{N} a_k$ is minimized. Our problem therefore reduces to

$$\min\langle 1, a\rangle, \tag{21}$$

where 1 denotes the vector with components that are all equal 1, subject to

$$\begin{aligned} Za &= e; \\ a &\geqslant 0. \end{aligned} \tag{22}$$

The problem of (21) and (22) is a linear programming problem that can be solved very efficiently using standard linear programming tools [7].

6. Example of the least-squares orthogonal vectors

We now consider an example illustrating the computational steps involved in determining the LS orthogonal vectors.

Consider the case in which we are given two vectors s_k, $1 \leqslant k \leqslant 2$ in $\mathscr{R}^2$ where

$$s_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{23}$$

In this example the vectors s_k are linearly independent so that $\mathscr{U} = \mathscr{R}^2$ and $P_{\mathscr{U}} = I$.

To compute the LS orthogonal vectors $\hat{h}_k$, we first find the orthonormal vectors $\hat{q}_k$ that maximize (6), using the procedure developed in Section 5. Specifically, we define $\Sigma_k = s_k s_k^*$ and determine the matrix $\Delta_{\mathscr{U}}$ that minimizes $\text{Tr}(\Delta)$ subject to $\Delta \geqslant \Sigma_k$ by using any SDP software package, which results in

$$\Delta_{\mathscr{U}} = \begin{bmatrix} 1.448 & -0.724 \\ -0.724 & 1.171 \end{bmatrix}. \tag{24}$$

[2] The inequality is to be understood as a component-wise inequality.

Using the eigendecomposition of $\Delta_{\mathscr{U}} - \Sigma_k$ we conclude that the null space of $\Delta_{\mathscr{U}} - \Sigma_k$ has dimension 1 for $k = 1, 2$ and is spanned by the vector x_k where

$$x_1 = \begin{bmatrix} 0.851 \\ -0.526 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -0.526 \\ 0.851 \end{bmatrix}.$$

The optimal orthonormal vectors are therefore given by $\hat{q}_k = \sqrt{a}_k x_k$ with a_k denoting the kth component of a. From (20) a must satisfy

$$\begin{bmatrix} 0.724 & 0.276 \\ 0.447 & -0.447 \\ 0.447 & -0.447 \\ 0.276 & 0.724 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \tag{25}$$

Since the matrix in (25) has full column rank, there is a unique solution $a_1 = a_2 = 1$ so that

$$\hat{q}_1 = \begin{bmatrix} 0.851 \\ 0.526 \end{bmatrix}, \quad \hat{q}_2 = \begin{bmatrix} -0.526 \\ 0.851 \end{bmatrix}. \tag{26}$$

We can immediately verify that the vectors $\hat{q}_k$ of (26) together with $\Delta_{\mathscr{U}}$ of (24) satisfy the necessary and sufficient conditions of Theorem 1. Finally, the LS orthogonal vectors are given by $\hat{h}_k = b_k \hat{q}_k$ with $b_k = \langle \hat{q}_k, s_k \rangle$ which yields

$$\hat{h}_1 = \begin{bmatrix} 0.724 \\ 0.447 \end{bmatrix}, \quad \hat{h}_2 = \begin{bmatrix} -0.724 \\ 1.171 \end{bmatrix}. \tag{27}$$

In Fig. 1 we plot the original vectors s_k given by (23), together with the LS orthogonal vectors of (27).

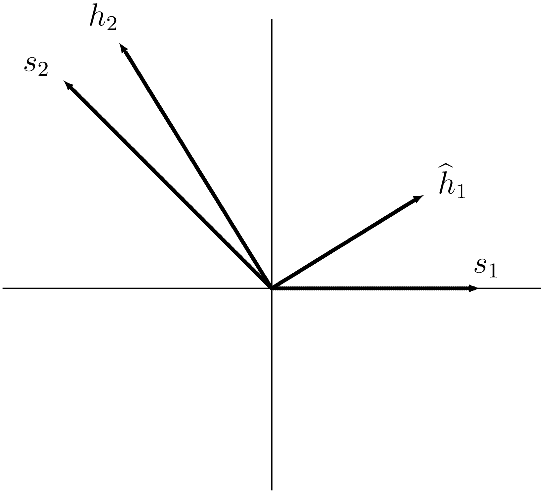

Fig. 1. Illustration of the least-squares orthogonal vectors. The original vectors s_k are given by (23), and the optimal orthogonal vectors $\hat{h}_k$ are given by (27).

7. Least-squares orthogonalization for symmetric vector sets

We now consider several special cases under which the optimality conditions of Theorem 1 can be solved to yield closed form solutions for the vectors $\hat{q}_k$. Instead of deriving these results from scratch, we rely on the connection between our problem and a detection problem that arises in quantum mechanics.

7.1. Connection with quantum detection

In a quantum detection problem, a system is prepared in one of N known states, where each state is represented by an Hermitian PSD operator drawn from a collection of known operators $\{\rho_k, 1 \leqslant k \leqslant N\}$ with prior probabilities $\{p_k > 0, 1 \leqslant k \leqslant N\}$. The system is then subjected to a measurement comprising N PSD Hermitian measurement operators $\{\Pi_k, 1 \leqslant k \leqslant N\}$, in order to determine the state prepared. The problem is to choose these operators to maximize the probability of correct detection which is given by [22]

$$P_d = \sum_{k=1}^{N} \mathrm{Tr}(\rho_k' \Pi_k), \tag{28}$$

where $\rho_k' = p_k \rho_k$. With $\mathscr{U}$ denoting the space spanned by the eigenvectors of the operators $\{\rho_k', 1 \leqslant k \leqslant N\}$, to constitute a measurement, the operators Π_k must satisfy

$$\begin{aligned} &\sum_{k=1}^{N} \Pi_k = P_{\mathscr{U}}; \\ &\Pi_k \geqslant 0, \quad 1 \leqslant k \leqslant N. \end{aligned} \tag{29}$$

Comparing (6) and (10) with (28) and (29) we see that our problem is equivalent to choosing a set of quantum measurement operators to maximize the probability of correct detection in a quantum detection problem, with $\Sigma_k = \rho_k'$. In this context, it has been shown in [20] that necessary and sufficient optimality conditions on the operators Π_k are exactly the conditions (15) and (16). From these conditions, closed form solutions where derived for some special cases [18,17]. In this section we summarize results from [17] relevant to LS orthogonalization.

7.2. Geometrically uniform vector sets

We first treat the case in which the vectors $\{s_k, 1 \leqslant k \leqslant N\}$ are defined over a (not necessarily abelian) group of unitary matrices and are generated by a single generating vector. Such a vector set is called *geometrically uniform* (GU) [21].

Let $\mathscr{G} = \{U_k, 1 \leqslant k \leqslant N\}$ be a finite group of N unitary matrices U_k. That is, $\mathscr{G}$ contains the identity matrix I; if $\mathscr{G}$ contains U_k, then it also contains its inverse $U_k^{-1} = U_k^*$; and the product $U_k U_j$ of any two elements of $\mathscr{G}$ is in $\mathscr{G}$ [4]. A vector set generated by $\mathscr{G}$ using a single generating vector s is a set $\mathscr{S} = \{s_k = U_k s, U_k \in \mathscr{G}\}$. Such a vector set has strong symmetry properties and is called GU.

We note that in [10] a GU state set was defined over an *abelian* group of unitary matrices; here we are not requiring the group $\mathscr{G}$ to be abelian.

For a GU vector set the optimal vectors $\hat{q}_k$ are also GU with the same generating group $\mathscr{G}$ as the original vector set. Thus, $\{\hat{q}_k = U_k q, U_k \in \mathscr{G}\}$ where the generator q is given by

$$q = ((SS^*)^{1/2})^{\dagger} s.$$

Here $(\cdot)^{1/2}$ denotes the unique symmetric PSD square root, and $(\cdot)^{\dagger}$ denotes the Moore-Penrose pseudo inverse.

7.3. Compound geometrically uniform vector sets

We now consider vector sets which consist of subsets that are GU, and are therefore referred to as *compound geometrically uniform* (CGU) [12].

A CGU vector set is defined as a set of vectors $\mathscr{S} = \{s_{km} = U_k s_m, 1 \leqslant k \leqslant L, 1 \leqslant m \leqslant R\}$ where the matrices $\{U_k, 1 \leqslant k \leqslant L\}$ are unitary and form a group $\mathscr{G}$, and the vectors $\{s_m, 1 \leqslant m \leqslant R\}$ are the generators. For concreteness we assume that $U_1 = I$ so that $s_{1m} = s_m$. A CGU vector set is in general not GU. However, for every m, the vectors $\{s_{km}, 1 \leqslant k \leqslant L\}$ are GU with generating group $\mathscr{G}$.

An example of a CGU vector set is illustrated in Fig. 2. In this example the vector set is $\{s_{km} = U_k s_m, U_k \in \mathscr{G}\}$ where $\mathscr{G} = \{I_2, U\}$ with

$$U = \frac{1}{2} \begin{bmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{bmatrix} \tag{30}$$

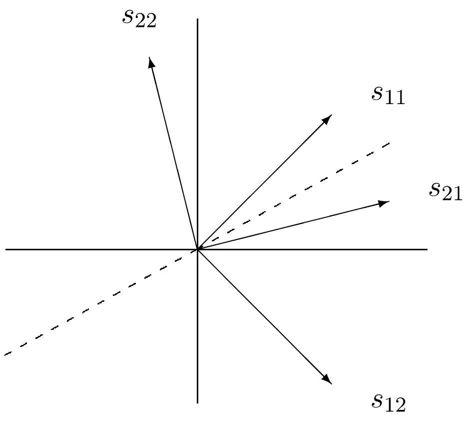

Fig. 2. A compound geometrically uniform vector set. The vector sets $\mathscr{S}_1 = \{s_{11}, s_{21}\}$ and $\mathscr{S}_2 = \{s_{12}, s_{22}\}$ are both geometrically uniform (GU) with the same generating group; both sets are invariant under a reflection about the dashed line. However, the combined set $\mathscr{S} = \{s_{11}, s_{21}, s_{12}, s_{22}\}$ is no longer GU.

and the generating vectors are

$$s_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}, \quad s_2 = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}.$$

The matrix U represents a reflection about the dashed line in Fig. 2. Thus, the vector s_{21} is obtained by reflecting the generator s_1 about this line, and similarly the vector s_{22} is obtained by reflecting the generator s_2 about this line.

As can be seen from the figure, the vector set is not GU. In particular, there is no isometry that transforms s_{11} into s_{12} while leaving the set invariant. However, the vector sets $\mathscr{S}_1 = \{s_{11}, s_{21}\}$ and $\mathscr{S}_2 = \{s_{12}, s_{22}\}$ are both GU with generating group $\mathscr{G}$.

For CGU vector sets the optimal vectors $\hat{q}_{km}$ are also CGU with the same generating group $\mathscr{G}$ as the original vector set, so that $\hat{q}_{km} = U_k q_m$ for a set of generators $\{q_m, 1 \leqslant m \leqslant R\}$. These generators can be found by first finding the symmetric operators $\widehat{\Pi}_m$ that maximize

$$\sum_{m=1}^{R} \mathrm{Tr}(\Sigma_m \Pi_m),$$

where $\Sigma_m = s_m s_m^*$ subject to the constraints

$$\Pi_m \geqslant 0, \quad 1 \leqslant m \leqslant R;$$
$$\sum_{k=1}^{L}\sum_{m=1}^{R} U_k \Pi_m U_k^* = P_{\mathscr{U}}.$$

The optimal generators q_m are then given by $\widehat{\Pi}_m = q_m q_m^*$, where the operators $\widehat{\Pi}_m$ are guaranteed to have rank 1. Since this problem is an SDP, the generators can be computed very efficiently.

There are some special cases in which there is a simple closed form solution for the generators q_m. Specifically, if

$$s_m^*((SS^*)^{1/2})^{\dagger} s_m = \alpha, \quad 1 \leqslant m \leqslant R \tag{31}$$

for some constant α, then

$$q_m = ((SS^*)^{1/2})^{\dagger} s_m.$$

A special class of CGU vector sets that satisfy (31) are *CGU vector sets with commuting GU generators*. In this case $\{s_{km} = U_k s_m, U_k \in \mathscr{G}\}$ where the generators s_m are GU, so that $\{s_m = V_m s, V_m \in \mathscr{Q}\}$ for some generator s and unitary matrices V_m that form a group $\mathscr{Q}$. In addition, U_p and V_t commute up to a phase factor for all t and p: $U_p V_t = V_t U_p \mathrm{e}^{\mathrm{j}\theta(p,t)}$ where $\theta(p, t)$ is an arbitrary phase function that may depend on the indices p and t. (In the special case in which $\theta = 0$ the resulting vector set is GU [12]). For such sets the optimal generators q_m are also GU with generating group $\mathscr{Q}$, i.e. $q_m = V_m q$, where

$$q = ((SS^*)^{1/2})^{\dagger} s.$$

As an example of a CGU vector set with commuting GU generators, consider the set $\{s_{km} = U_k V_m s, U_k \in \mathscr{G}, V_m \in \mathscr{Q}\}$, where the generating groups are $\mathscr{G} = \{I_2, U\}$ and $\mathscr{Q} = \{I_2, B\}$, with

$$U = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

and the generating vector is

$$s = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \\ 0 \\ 0 \end{bmatrix}.$$

Since $UB = -BU$, the groups $\mathscr{G}$ and $\mathscr{Q}$ commute up to a phase factor.

The vectors s_{km} span the subspace $\mathscr{R}^2$ of $\mathscr{R}^4$, so that the optimal normalized tight frame vectors $\hat{q}_{km}$ will also lie in $\mathscr{R}^2$. In Fig. 3 we plot the vector set s_{km} in $\mathscr{R}^2$.

From our general results, $\{\hat{q}_{km} = U_k V_m q,\ 1 \leqslant k, m \leqslant 2\}$, where $q = ((SS^*)^{1/2})^\dagger s$. Since

$$SS^* = 2 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

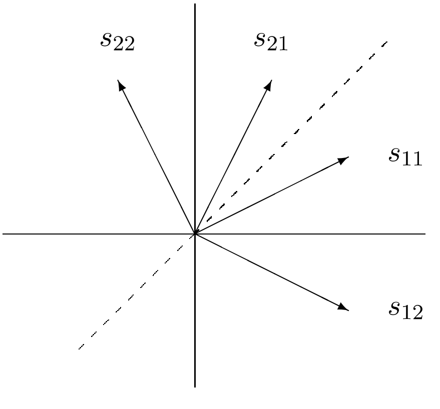

Fig. 3. A compound geometrically uniform (GU) vector set with commuting GU generators. The vector sets $\mathscr{S}_1 = \{s_{11}, s_{21}\}$ and $\mathscr{S}_2 = \{s_{12}, s_{22}\}$ are both GU with the same generating group; both sets are invariant under a reflection about the dashed line. The set of generators $\{s_{11}, s_{12}\}$ is GU and is invariant under a reflection about the x-axis. The combined set $\mathscr{S} = \{s_{11}, s_{21}, s_{12}, s_{22}\}$ is no longer GU. Nonetheless, the optimal normalized tight frame vectors are generated by a single generating vector and are given by $\hat{q}_{km} = (1/\sqrt{2})s_{km}$.

we have

$$((SS^*)^{1/2})^\dagger = (1/\sqrt{2})\begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\end{bmatrix}$$

so that $q = (1/\sqrt{2})s$ and consequently $\hat{q}_{km} = (1/\sqrt{2})s_{km}$.

Note, that in the example of Fig. 2 the CGU vector set also has GU generators. Specifically, the set of generators $\{s_1, s_2\}$ with $s_1 = s_{11}$ and $s_2 = s_{12}$ is invariant under a reflection about the x-axis: $s_2 = Bs_1$ where

$$B = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}.$$

However, the group of generators $\mathscr{Q} = \{I_2, B\}$ does not commute up to a phase with the generating group $\mathscr{G} = \{I_2, U\}$, where U is given by (30) and represents a reflection about the dashed line in Fig. 2. This can be verified graphically from Fig. 2: Applying B to s_{11} followed by U results in the vector s_{22}. On the other hand, if we apply U to s_{11} followed by B, then the resulting vector is the reflection of s_{21} about the x-axis, which is not related to s_{22} by a phase factor.

Next, consider the vector set in Fig. 3. In this case $\mathscr{Q} = \{I_2, B\}$ and $\mathscr{G} = \{I_2, U\}$ where now B represents a reflection about the x-axis and U represents a reflection about the dashed line in Fig. 3. We can immediately verify from the figure that applying U and then B to any vector in the set results in a vector that is equal up to a minus sign to the vector that is obtained from first applying B and then U. For example, applying B to s_{11} followed by U results in s_{22}. If, on the other hand, we first apply U and then B to s_{11}, then we obtain the reflection of s_{21} about the x-axis, which is equal to $-s_{22}$.

We summarize our results regarding CGU vector sets in the following theorem:

Theorem 2 (CGU vector sets). *Let $\mathscr{S} = \{s_{km} = U_k s_m, U_k \in \mathscr{G}, 1 \leqslant m \leqslant R\}$ be a compound geometrically uniform vector set generated by a finite group $\mathscr{G}$ of unitary matrices and generators $\{s_m, 1 \leqslant m \leqslant R\}$, and let $\mathscr{U}$ denote the space spanned by the vectors s_{km}. Then the normalized tight frame vectors $\hat{q}_{km}$ that maximize (6) subject to (7) and (8) are given by $\{\hat{q}_{km} = U_k q_m, U_k \in \mathscr{G}, 1 \leqslant m \leqslant R\}$, where the generators q_m can be obtained as the solution to the semidefinite programming problem of maximizing $J(\{\Pi_m\}) = \sum_{m=1}^R \mathrm{Tr}(\Sigma_m \Pi_m)$ with $\Sigma_m = s_m s_m^*$, subject to $\Pi_m \geqslant 0, 1 \leqslant m \leqslant R$ and $\sum_{k=1}^L \sum_{m=1}^R U_k \Pi_m U_k^* = P_{\mathscr{U}}$. The generators q_m are then given by $\widehat{\Pi}_m = q_m q_m^*$, where $\{\widehat{\Pi}_m\}$ denote the optimal $\{\Pi_m\}$. In addition,*

(1) *If $s_m^*((SS^*)^{1/2})^\dagger s_m = \alpha$ for $1 \leqslant m \leqslant R$ then $q_m = ((SS^*)^{1/2})^\dagger s_m$.*
(2) *If the generators $\{s_m = V_m s, 1 \leqslant m \leqslant R\}$ are geometrically uniform with $U_k V_m = V_m U_k \mathrm{e}^{\mathrm{j}\theta(k,m)}$ for all k, m, then $q_m = V_m q$ where $q = ((SS^*)^{1/2})^\dagger s$.*

Acknowledgment

The author wishes to acknowledge the reviewer whose comments greatly improved the presentation of the results.

References

[1] A.Y. Alfakih, A. Khandani, H. Wolkowicz, Solving Euclidean distance matrix completion problems via semidefinite programming, Comput. Optim. Appl. 12 (1999) 13–30.
[2] F. Alizadeh, Combinatorial optimization with interior point methods and semi-definite matrices, Ph.D. thesis, University of Minnesota, Minneapolis, MN, 1991.
[3] F. Alizadeh, Optimization over the positive-definite cone: interior point methods and combinatorial applications, in: P. Pardalos (Ed.), Advances in Optimization and Parallel Computing, North-Holland, The Netherlands, 1992.
[4] M.A. Armstrong, Groups and Symmetry, Springer-Verlag, New York, 1988.
[5] M. Ban, K. Kurokawa, R. Momose, O. Hirota, Optimum measurements for discrimination among symmetric quantum states and parameter estimation, Int. J. Theor. Phys. 36 (1997) 1269–1288.
[6] D.P. Bertsekas, Nonlinear Programming, second ed., Athena Scientific, Belmont, MA, 1999.
[7] D. Bertsimas, J. Tsitsiklis, Introduction to Linear Optimization, Athena Scientific, Belmont, MA, 1997.
[8] M. Charbit, C. Bendjaballah, C.W. Helstrom, Cutoff rate for the m-ary PSK modulation channel with optimal quantum detection, IEEE Trans. Inform. Theory 35 (1989) 1131–1133.
[9] I. Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA, 1992.
[10] Y.C. Eldar, Least-squares inner product shaping, Linear Algebra Appl. 348 (2002) 153–174.
[11] Y.C. Eldar, Quantum signal processing, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA, 2001. Available from: <http://www.ee.technion.ac.il/Sites/People/YoninaEldar/>.
[12] Y.C. Eldar, H. Bölcskei, Geometrically uniform frames, IEEE Trans. Inform. Theory 49 (2003) 993–1006.
[13] Y.C. Eldar, A.V. Oppenheim, D. Egnor, Orthogonal and projected orthogonal matched filter detection, Signal Processing 84 (2004) 677–693.
[14] Y.C. Eldar, A.V. Oppenheim, Orthogonal multiuser detection, Signal Processing 82 (2002) 321–325.
[15] Y.C. Eldar, A.M. Chan, An optimal whitening approach to linear multiuser detection, IEEE Trans. Inform. Theory 49 (2003) 2156–2171.
[16] Y.C. Eldar, A.V. Oppenheim, Covariance shaping least-squares estimation, IEEE Trans. Signal Process. 51 (2003) 686–697.
[17] Y.C. Eldar, A. Megretski, G.C. Verghese, Optimal detection of symmetric mixed quantum states, IEEE Trans. Inform. Theory 50 (2004) 1198–1207.
[18] Y.C. Eldar, G.D. Forney Jr., On quantum detection and the square-root measurement, IEEE Trans. Inform. Theory 47 (2001) 858–872.
[19] Y.C. Eldar, G.D. Forney Jr., Optimal tight frames and quantum measurement, IEEE Trans. Inform. Theory 48 (2002) 599–610.
[20] Y.C. Eldar, A. Megretski, G.C. Verghese, Designing optimal quantum detectors via semidefinite programming, IEEE Trans. Inform. Theory 49 (2003) 1012–1017.
[21] G.D. Forney Jr., Geometrically uniform codes, IEEE Trans. Inform. Theory 37 (1991) 1241–1260.
[22] C.W. Helstrom, Quantum Detection and Estimation Theory, Academic Press, New York, 1976.
[23] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, Cambridge, UK, 1985.
[24] M. Laurent, A tour d'horizon on positive semidefinite and Euclidean distance matrix completion problems, in: Topics in Semidefinite and Interior-Point Methods, The Fields Institute for Research in Mathematical Sciences, Communications Series, vol. 18, American Mathematical Society, Providence, RI, 1998.

[25] A. Megretski, C.-Y. Kao, U. Jönsson, A. Rantzer, A guide to IQCβ: Software for robustness analysis. Available from: <http://web.mit.edu/~cykao/www>.
[26] Y. Nesterov, A. Nemirovski, Interior-Point Polynomial Algorithms in Convex Programming, SIAM, Philadelphia, PE, 1994.
[27] M. Osaki, M. Ban, O. Hirota, Derivation and physical interpretation of the optimum detection operators for coherent-state signals, Phys. Rev. A 54 (1996) 1691–1701.
[28] J.F. Sturm, Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones, Optim. Methods and Softw. 11–12 (1999) 625–653.
[29] L. Vandenberghe, S. Boyd, Semidefinite programming, SIAM Rev. 38 (1996) 40–95.