Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

Constrained Quadratic Minimizations for Signal Processing and Communications

Ali Pezeshki, Louis L. Scharf, Magnus Lundberg, and Edwin K. P. Chong

***Abstract*— Constrained minimization problems considered here arise in the design of multi-dimensional subspace beamformers for radar, sonar, seismology, and wireless communications, and in the design of precoders and equalizers for digital communications. The problem is to minimize a quadratic form, under a set of linear or quadratic constraints. We derive the solutions to these problems and establish connections between them. We show that the quadratically-constrained problem can be solved by solving a set of linearly-constrained problems and then using a majorization argument and Poincare's separation theorem to determine which linearly-constrained problem solves the quadratically-constrained one. In addition, we present illuminating circuit diagrams for our solutions, called generalized sidelobe canceller (GSC) diagrams, which allow us to tie our constrained minimizations to linear minimum mean-squared error (LMMSE) estimations.**

I. Introduction

Let $\mathbf{R} \in \mathbb{C}^{n\times n}$ be a positive definite (PD) complex matrix and $\mathbf{W} \in \mathbb{W} = \{\mathbf{W} \mid \mathbf{W} \in \mathbb{C}^{n\times r};\ \text{rank}\{\mathbf{W}\} = r < n\}$ be the set of all $n \times r$ complex matrices of rank $r < n$. We are interested in minimizing the quadratic function $J(\mathbf{W})$,

$$J(\mathbf{W}) = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\}, \tag{1}$$

with respect to $\mathbf{W}$, subject to a set of linear or quadratic constraints. $\mathbf{W}^H\mathbf{R}\mathbf{W}$ is the covariance matrix for a random vector $\mathbf{y} = \mathbf{W}^H\mathbf{x}$, where $\mathbf{x} \in \mathbb{C}^n$ is a zero-mean random vector with covariance $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$, and $J(\mathbf{W}) = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} = E[\mathbf{y}^H\mathbf{y}]$ is total variance.

The constraints are as follows.

Case 1: Linear Constraint. Let $\mathbf{\Psi} \in \mathbb{C}^{n\times p}$ and $\mathbf{L} \in \mathbb{C}^{r\times r}$ be full-rank matrices, with $r \le p < n$, and $\mathbf{V} \in \mathbb{C}^{p\times r}$ be a left-orthogonal matrix, i.e. $\mathbf{V}^H\mathbf{V} = \mathbf{I}$ but $\mathbf{V}\mathbf{V}^H = \mathbf{P}_{\mathbf{V}} = \mathbf{P}_{\mathbf{V}}^2$ is an orthogonal projection onto the linear subspace spanned by columns of $\mathbf{V}$. Then, the constraint is

$$\mathbf{W}^H\mathbf{\Psi} = \mathbf{L}^H\mathbf{V}^H, \tag{2}$$

which is linear in $\mathbf{W}$. Thus, our constrained minimization problem is

$$\min_{\mathbf{W}\in\mathbb{W}} J(\mathbf{W}) = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} \quad \text{subject to } \mathbf{W}^H\mathbf{\Psi} = \mathbf{L}^H\mathbf{V}^H. \tag{3}$$

The linear constraint in (3) ensures that the action of $\mathbf{W}^H$ on $\mathbf{\Psi}$ meets the design constraint $\mathbf{L}^H\mathbf{V}^H$. It is easy to show that the constraint in (3) is equivalent to

$$\mathbf{W}^H\mathbf{\Psi}\,[\mathbf{V}\ \ \mathbf{V}_\star] = \mathbf{L}^H\,[\mathbf{I}\ \ \mathbf{0}], \tag{4}$$

where $[\mathbf{V}\ \ \mathbf{V}_\star] \in \mathbb{C}^{p\times p}$ is an orthogonal matrix. This means that under the original constraint in (3) $\mathbf{W}^H$ images the r linear combinations $\mathbf{\Psi}\mathbf{V}$ as $\mathbf{L}^H$ and the $p - r$ linear combinations $\mathbf{\Psi}\mathbf{V}_\star$ as zero. These may be called *zero-forcing* constraints, which is a commonly-used term in communications and signal processing.

Case 2: Quadratic Constraint. Let $\mathbf{S} \in \mathbb{C}^{n\times n}$ be a positive semi-definite (PSD) matrix of rank p and $\mathbf{D} \in \mathbb{C}^{r\times r}$ be a PD matrix, with $r \le p < n$. Then, the constraint is

$$\mathbf{W}^H\mathbf{S}\mathbf{W} = \mathbf{D}, \tag{5}$$

which is quadratic in $\mathbf{W}$. Thus, the constrained minimization problem in this case is

$$\min_{\mathbf{W}\in\mathbb{W}} J(\mathbf{W}) = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} \quad \text{subject to } \mathbf{W}^H\mathbf{S}\mathbf{W} = \mathbf{D}. \tag{6}$$

The quadratic constraint in (6) ensures that the action of $\mathbf{W}^H$ on a random vector $\mathbf{s} \in \mathbb{C}^n$, with covariance $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H]$, produces a random vector with designed covariance $\mathbf{D}$.

In this paper (also see [1]), we derive closed-form solutions for (3) and (6). The linearly-constrained problem of (3) is convex and may be solved using the method of Lagrange multipliers. The quadratically-constrained problem of (6) on the other hand is nonconvex and deriving a closed form solution for it requires a majorization argument.[1] Further, we establish connections between the linearly- and quadratically-constrained problems. Given $\mathbf{S}$ and $\mathbf{D}$, we show that the minimum value for the quadratic form J under the family of linear constraints of the form (2), with $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S}$, $\mathbf{L}^H\mathbf{V}^H\mathbf{V}\mathbf{L} = \mathbf{L}^H\mathbf{L} = \mathbf{D}$, $\mathbf{V} \in \mathbb{V}$, and $\mathbb{V}$ the set of all $p \times r$ left-orthogonal complex matrices, is the

This work is supported by the DARPA Integrated Sensing and Processing Program, under AFOSR contract FA9550-04-1-0371, and by the Office of Naval Research (ONR), under contract N00014-04-1-0084.

A. Pezeshki is with the Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA, ali@engr.colostate.edu

L. L. Scharf is with the Departments of Electrical and Computer Engineering,and Statistics, Colorado State University, Fort Collins, CO 80523, USA, scharf@engr.colostate.edu

M. Lundberg is with the Department of Electrical Engineering, Luleä University of Technology, Luleä, Sweden, mlg@sm.lthu.se

E. K. P. Chong is with the Departments of Electrical and Computer Engineering, and Mathematics, Colorado State University, Fort Collins, CO 80523, USA, chong@engr.colostate.edu

[1] If $\mathbf{S}$ were nonsingular and $\mathbf{D} = \mathbf{I}$, we could easily solve (6) by solving the generalized eigenvalue problem $\mathbf{R}\mathbf{W} = \mathbf{S}\mathbf{W}\mathbf{\Omega}$, with $\mathbf{\Omega}$ the diagonal eigenvalue matrix, for the r largest eigenvalues and the corresponding eigenvectors. However, for the singular $\mathbf{S}$ this is not possible and a majorization argument is required to determine the closed form solution.

0-7803-9568-9/05/$20.00 ©2005 IEEE

minimum value of J under the quadratic constraint of (5). This minimum is obtained when the left-orthogonal matrix $\mathbf{V}$ carries the r principal eigenvectors of $\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2}$, where $\mathbf{S}^{1/2} \in \mathbb{C}^{n\times p}$ is one particular rectangular square-root of $\mathbf{S} = \mathbf{S}^{1/2}\mathbf{S}^{H/2}$, which we shall define in Section II-B. The implication of this result is that the nonconvex quadratically-constrained problem may be solved by solving a set of convex linearly-constrained problems to determine a set of candidate solutions, which all satisfy the quadratic constraint, and then finding the best candidate in the set. The set of candidate solutions is nonconvex, and hence the search for the best candidate is carried out inside a nonconvex set. It turns out that a majorization result and Poincare's separation theorem are the right tools for finding the best candidate.

In addition, we present illuminating circuit diagrams for the solutions to (3) and (6), called generalized sidelobe canceller (GSC) diagrams, which allow us to establish connections between our constrained minimizations and linear minimum mean-squared error (LMMSE) estimations.

Two theorems play key roles in our developments: a majorization result for matrix trace and Poincare's separation theorem.

Theorem 1. (A majorization result for matrix trace). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ and $\mathbf{B} \in \mathbb{C}^{n\times n}$ be PSD matrices, with eigenvalues $0 \le \lambda^2_{\mathbf{A},1} \le \cdots \le \lambda^2_{\mathbf{A},n}$ and $\lambda^2_{\mathbf{B},1} \ge \cdots \ge \lambda^2_{\mathbf{B},n} \ge 0$, respectively. Then,

$$\mathrm{tr}\{\mathbf{AB}\} \ge \sum_{i=1}^{n} \lambda^2_{\mathbf{A},i}\lambda^2_{\mathbf{B},i}. \tag{7}$$

The equality holds when the eigenvectors of $\mathbf{A}$ and $\mathbf{B}$ are equal.

Proof. See [2, ch. 9, H.1.h].

Theorem 2. (Poincare's separation theorem). Let $\mathbf{A} \in \mathbb{C}^{n\times n}$ be an PSD matrix with eigenvalues $0 \le \lambda^2_{\mathbf{A},1} \le \cdots \le \lambda^2_{\mathbf{A},n}$ and $\mathbf{X} \in \mathbb{C}^{n\times r}$ $(n > r)$ a left-orthogonal matrix ($\mathbf{X}^H\mathbf{X} = \mathbf{I}$ and $\mathbf{X}\mathbf{X}^H = \mathbf{P}_\mathbf{X}$). Further, let $0 \le \lambda^2_{\mathbf{B},1} \le \cdots \le \lambda^2_{\mathbf{B},r}$ be the eigenvalues of $\mathbf{B} = \mathbf{X}^H\mathbf{A}\mathbf{X} \in \mathbb{C}^{r\times r}$. Then,

$$\lambda^2_{\mathbf{A},i} \le \lambda^2_{\mathbf{B},i} \le \lambda^2_{\mathbf{A},n-r+i}; \quad i = 1,\ldots,r. \tag{8}$$

Proof. See [3, ch. 11, Thm. 10].

The constrained minimization problems posed in this paper arise in the design of multi-rank adaptive beamformers for radar, sonar, seismology, and wireless communications, where robustness with respect to the non-planar and unpredictable structure of a propagating wavefront is desired [4]-[7]. They also arise in the design of precoders and equalizers for digital communications [8]-[12]. The reader is referred to [1] for discussions on physical interpretations of (3) and (6) in beamforming and diversity combining.

II. Quadratic Minimizations Under Linear and Quadratic Constraints

A. Solution for the Linearly-Constrained Problem

For the linearly-constrained problem of (3), we first establish an appropriate coordinate system for the rank$-p < n$ matrix $\mathbf{\Psi} \in \mathbb{C}^{n\times p}$ by constructing its SVD as

$$\mathbf{\Psi} = \mathbf{U}_\mathbf{\Psi}\mathbf{\Sigma}_\mathbf{\Psi} = \begin{bmatrix} \mathbf{U}_{\mathbf{\Psi},p} & \mathbf{U}_{\mathbf{\Psi},\star} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{\mathbf{\Psi},p} \\ \mathbf{0} \end{bmatrix}. \tag{9}$$

One would expect to see an orthogonal matrix $\mathbf{V}^H_\mathbf{\Psi} \in \mathbb{C}^{p\times p}$ in place of identity on the right-hand-side of (9). But $\mathbf{V}^H_\mathbf{\Psi}$ may always be absorbed in the left-orthogonal matrix $\mathbf{V} \in \mathbb{C}^{p\times r}$ in the constraint $\mathbf{W}^H\mathbf{\Psi} = \mathbf{L}^H\mathbf{V}^H$, as $(\mathbf{V}^H\mathbf{V}_\mathbf{\Psi})^H \in \mathbb{C}^{p\times r}$ will be left-orthogonal. Thus, without loss of generality, we assume that the rank$-p$ matrix $\mathbf{\Psi} \in \mathbb{C}^{n\times p}$ is of the form (9).

The solution to (3) may now be stated as the following theorem.

Theorem 3. The minimum value of the quadratic form J in the linearly-constrained minimization problem of (3) is

$$J_o = \mathrm{tr}\{\mathbf{L}^H\mathbf{V}^H(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}\mathbf{VL}\}. \tag{10}$$

Further, the solution $\mathbf{W} = \mathbf{W}_o$ has two equivalent forms:

1. Standard form

$$\mathbf{W}_o = \mathbf{R}^{-1}\mathbf{\Psi}(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}\mathbf{VL} \tag{11}$$

2. GSC form

$$\mathbf{W}_o = [\mathbf{U}_{\mathbf{\Psi},p} - \mathbf{U}_{\mathbf{\Psi},\star}\mathbf{F}]\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{VL}, \tag{12}$$

where

$$\mathbf{F} = (\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star})^{-1}\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}. \tag{13}$$

Proof. This is a simple linearly-constrained minimization problem, which may be solved using the method of Lagrange multipliers and completing the square, as shown in [1].

The GSC form in (12) is interesting from a signal processing and communications point of view. In Section III, we shall show that this form has a circuit diagram implementation, called the GSC diagram, which ties the minimization problem of (3) to LMMSE estimations.

B. Solution for the Quadratically-Constrained Problem

For the quadratically-constrained problem of (6), we first establish the appropriate coordinate system for the rank$-p < n$ PSD matrix $\mathbf{S} \in \mathbb{C}^{n\times n}$ and the PD matrix $\mathbf{D} \in \mathbb{C}^{r\times r}$ by constructing their EVD's as

$$\mathbf{S} = \mathbf{U}_\mathbf{S}\mathbf{\Lambda}^2_\mathbf{S}\mathbf{U}^H_\mathbf{S} = \begin{bmatrix} \mathbf{U}_{\mathbf{S},p} & \mathbf{U}_{\mathbf{S},\star} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}^2_{\mathbf{S},p} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U}^H_{\mathbf{S},p} \\ \mathbf{U}^H_{\mathbf{S},\star} \end{bmatrix}. \tag{14}$$

and

$$\mathbf{D} = \mathbf{U}_\mathbf{D}\mathbf{\Lambda}^2_\mathbf{D}\mathbf{U}^H_\mathbf{D}, \tag{15}$$

where $\mathbf{\Lambda}^2_{\mathbf{S},p} = \mathrm{diag}(\lambda^2_{\mathbf{S},1},\ldots,\lambda^2_{\mathbf{S},p})$, $\lambda^2_{\mathbf{S},1} \ge \cdots \ge \lambda^2_{\mathbf{S},p} > 0$ and $\mathbf{\Lambda}^2_\mathbf{D} = \mathrm{diag}(\lambda^2_{\mathbf{D},1},\ldots,\lambda^2_{\mathbf{D},r})$, $\lambda^2_{\mathbf{D},1} \ge \cdots \ge \lambda^2_{\mathbf{D},r} > 0$. Further, we define the *signal-to-signal-plus-noise ratio matrix* $\mathbf{Q} \in \mathbb{C}^{p\times p}$ as[2]

$$\mathbf{Q} = \mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2} = \mathbf{\Lambda}^H_{\mathbf{S},p}\mathbf{U}^H_{\mathbf{S},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p}, \tag{16}$$

[2]In a signal-plus-noise model, the covariance matrix $\mathbf{R} = \mathbf{S} + \mathbf{N}$ is the sum of the signal covariance matrix $\mathbf{S}$ and the noise covariance matrix $\mathbf{N}$. Due to the fact that in the expression for $\mathbf{Q}$ the square-roots of the signal covariance matrix $\mathbf{S}$ are multiplied by the inverse of the signal-plus-noise covariance matrix $\mathbf{R}$, in signal processing and communications the matrix $\mathbf{Q}$ is usually called the signal-to-signal-plus-noise ratio matrix.

with $\mathbf{S}^{1/2} = \mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p}$, and give it the EVD

$$\mathbf{Q} = \mathbf{U}_{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{Q}}^2\mathbf{U}_{\mathbf{Q}}^H, \tag{17}$$

where $\mathbf{\Lambda}_{\mathbf{Q}}^2 = \text{diag}(\lambda_{\mathbf{Q},1}^2, \ldots, \lambda_{\mathbf{Q},p}^2)$, $\lambda_{\mathbf{Q},1}^2 \geq \cdots \geq \lambda_{\mathbf{Q},p}^2 > 0$. Consequently,

$$\mathbf{Q}^{-1} = \mathbf{U}_{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{Q}}^{-2}\mathbf{U}_{\mathbf{Q}}^H \tag{18}$$

is an EVD of $\mathbf{Q}^{-1} = (\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}$, where $\mathbf{\Lambda}_{\mathbf{Q}}^{-2} = \text{diag}(\frac{1}{\lambda_{\mathbf{Q},1}^2}, \ldots, \frac{1}{\lambda_{\mathbf{Q},p}^2})$ and $0 < \frac{1}{\lambda_{\mathbf{Q},1}^2} \leq \cdots \leq \frac{1}{\lambda_{\mathbf{Q},p}^2}$. We are now ready to solve (6).

Theorem 4. The minimum value of the quadratic form J in the quadratically-constrained minimization problem of (6) is

$$J_o = \text{tr}\{\mathbf{U}_{\mathbf{Q},r}^H\mathbf{Q}^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{\Lambda}_{\mathbf{D}}^2\} = \sum_{i=1}^{r}\frac{\lambda_{\mathbf{D},i}^2}{\lambda_{\mathbf{Q},i}^2}, \tag{19}$$

where $\mathbf{U}_{\mathbf{Q},r} \in \mathbb{C}^{p\times r}$ carries the r principal eigenvectors of $\mathbf{Q}$. Further, the solution $\mathbf{W} = \mathbf{W}_o$ has two equivalent forms:

1. Standard form

$$\mathbf{W}_o = \mathbf{R}^{-1}\mathbf{S}^{1/2}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}, \tag{20}$$

2. GSC form

$$\mathbf{W}_o = [\mathbf{U}_{\mathbf{S},p} - \mathbf{U}_{\mathbf{S},\star}\mathbf{F}]\mathbf{\Lambda}_{\mathbf{Q}}^{-H}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}, \tag{21}$$

where $\mathbf{D}^{H/2} = \mathbf{\Lambda}_{\mathbf{D}}^H\mathbf{U}_{\mathbf{D}}^H$ and

$$\mathbf{F} = (\mathbf{U}_{\mathbf{S},\star}^H\mathbf{R}\mathbf{U}_{\mathbf{S},\star})^{-1}\mathbf{U}_{\mathbf{S},\star}^H\mathbf{R}\mathbf{U}_{\mathbf{S},p}. \tag{22}$$

Proof. The proof follows from the majorization result of Theorem 1 and Poincare's separation theorem, as shown in [1].

C. Connections Between Linearly- and Quadratically-Constrained Problems

Our aim in this section is to establish connections between the convex linearly-constrained minimization problem of (3) and the nonconvex quadratically-constrained minimization problem of (6).

Given $\mathbf{S}$ and $\mathbf{D}$ in the quadratically-constrained problem of (6), consider a corresponding class of linearly-constrained problems of the form (3), in which

$$\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S} \quad \text{and} \quad \mathbf{L}^H\mathbf{V}^H\mathbf{V}\mathbf{L} = \mathbf{L}^H\mathbf{L} = \mathbf{D}, \tag{23}$$

and $\mathbf{V} \in \mathbb{V}$ is any $p\times r$ left-orthogonal complex matrix: $\mathbb{V} = \{\mathbf{V} \mid \mathbf{V} \in \mathbb{C}^{p\times r},\ \mathbf{V}^H\mathbf{V} = \mathbf{I},\ \mathbf{V}\mathbf{V}^H = \mathbf{P}_{\mathbf{V}}\}$. This equation characterizes a class of linear constraints of the form (2) that are quadratically equivalent to the quadratic constraint in (5).

Clearly, any linear constraint matrix $\mathbf{L}^H \in \mathbb{C}^{r\times r}$ satisfying $\mathbf{L}^H\mathbf{V}^H\mathbf{V}\mathbf{L} = \mathbf{L}^H\mathbf{L} = \mathbf{D}$ may be expressed as $\mathbf{L}^H = \mathbf{D}^{1/2}\mathbf{T}^H$, where $\mathbf{D}^{1/2} = \mathbf{U}_{\mathbf{D}}\mathbf{\Lambda}_{\mathbf{D}}$ and $\mathbf{T} \in \mathbb{C}^{r\times r}$ is any orthogonal matrix. However, the product of the orthogonal matrix $\mathbf{T}$ and the left-orthogonal matrix $\mathbf{V}$, i.e. $\mathbf{V}\mathbf{T}$, is a $p\times r$ left-orthogonal matrix and hence belongs to $\mathbb{V}$. Therefore from here on, without loss of generality, we assume that $\mathbf{T}$ has been absorbed in $\mathbf{V}$, and hence in the class of linear constraints associated with (23) $\mathbf{L}^H = \mathbf{D}^{1/2} = \mathbf{U}_{\mathbf{D}}\mathbf{\Lambda}_{\mathbf{D}}$. Further, since $\mathbf{\Psi}$ is of the form (9), $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S} = \mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p}^2\mathbf{U}_{\mathbf{S},p}^H$ implies $\mathbf{\Psi} = \mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p} = \mathbf{S}^{1/2}$.

We may now state the following theorem to tie together our linearly- and quadratically-constrained minimizations.

Theorem 5. The smallest achievable value for the quadratic form J in the class of linearly-constrained problems characterized by (23) is

$$J_o = \sum_{i=1}^{r}\frac{\lambda_{\mathbf{D},i}^2}{\lambda_{\mathbf{Q},i}^2}, \tag{24}$$

which is equal to the minimum value of J in the quadratically-constrained problem of (6). This is obtained when the left-orthogonal matrix $\mathbf{V}$ carries the r principal eigenvectors of $\mathbf{Q} = \mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2}$, i.e. when $\mathbf{V} = \mathbf{U}_{\mathbf{Q},r}$. In such a case, the solution $\mathbf{W} = \mathbf{W}_o$ is equal to $\mathbf{W}_o$ in (20).

Proof. See [1].

This theorem shows that the quadratically-constrained problem of (6) may be solved by solving the following linearly-constrained problem:

$$\begin{aligned} &\min_{\mathbf{W}\in\mathbb{W},\mathbf{V}\in\mathbb{V}} \quad J = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} \\ &\text{subject to } \ \mathbf{W}^H\mathbf{S}^{1/2} = \mathbf{D}^{1/2}\mathbf{V}^H. \end{aligned} \tag{25}$$

By fixing $\mathbf{V} \in \mathbb{V}$, the above minimization problem is reduced to the convex linearly-constrained minimization problem of (3), with $\mathbf{\Psi} = \mathbf{S}^{1/2}$ and $\mathbf{L}^H = \mathbf{D}^{1/2}$, the solution to which is

$$\mathbf{W}_o = \mathbf{R}^{-1}\mathbf{S}^{1/2}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{V}\mathbf{D}^{H/2}. \tag{26}$$

The matrix $\mathbf{W}_o$ is a candidate solution for (25), or equivalently (6), as it satisfies the quadratic constraint of (6). The candidate solutions $\mathbf{W}_o$, obtained by considering every $\mathbf{V} \in \mathbb{V}$, form a nonconvex set. Consequently, the solution to (6) may be found by searching this nonconvex set for the *best candidate*, i.e. the one that yields the smallest value of the quadratic form $J = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\}$. The proof in [1] suggests that the majorization result of Theorem 1 and Poincare's separation theorem are the right tools for performing such a search.

III. Generalized Sidelobe Canceller Diagrams

In Section II, we derived alternative forms for the solutions to (3) and (6) and named them the GSC forms. In this Section, we clarify our terminology by presenting circuit diagrams for (12) and (21), which in signal processing and communications are called generalized sidelobe canceller (GSC) diagrams [13]. Here the GSC diagrams are particularly illuminating, as they tie the constrained minimizations of (3) and (6) to LMMSE estimations.

Consider the solution $\mathbf{W}_o$ in (12) for the linearly-constrained problem of (3). This solution is a matrix (filter) that takes the (input) random vector $\mathbf{x}$ to the (output) random vector $\mathbf{y} = \mathbf{W}_o^H\mathbf{x}$,

$$\begin{aligned} \mathbf{y} = \mathbf{W}_o^H\mathbf{x} &= \mathbf{L}^H\mathbf{V}^H\mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1}[\mathbf{U}_{\mathbf{\Psi},p}^H - \mathbf{F}^H\mathbf{U}_{\mathbf{\Psi},\star}^H]\mathbf{x} \\ &= \mathbf{L}^H\mathbf{V}^H\mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1}[\mathbf{u} - \mathbf{F}^H\mathbf{v}] \\ &= \mathbf{L}^H\mathbf{V}^H\mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1}\mathbf{e}, \end{aligned} \tag{27}$$

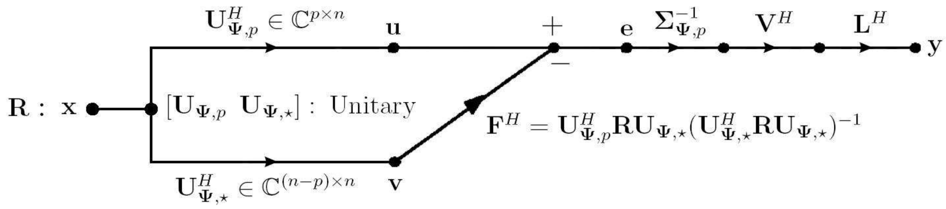

Fig. 1. Generalized sidelobe canceller (GSC) diagram.

where $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{e}$ are the following vectors, as illustrated in the GSC diagram of Fig. 1:

$$\mathbf{u} = \mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{x}, \quad \mathbf{v} = \mathbf{U}_{\mathbf{\Psi},\star}^H\mathbf{x}, \quad \text{and} \quad \mathbf{e} = \mathbf{u} - \mathbf{F}^H\mathbf{v}. \tag{28}$$

In this diagram, the vector $\mathbf{x}$ is decomposed into two sets of coordinates $\mathbf{u} = \mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{x}$ and $\mathbf{v} = \mathbf{U}_{\mathbf{\Psi},\star}^H\mathbf{x}$, with composite covariance matrix

$$\begin{aligned} & E\begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}\begin{bmatrix} \mathbf{u}^H & \mathbf{v}^H \end{bmatrix} \\ &= \begin{bmatrix} (\mathbf{R}_{uu} = \mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}) & (\mathbf{R}_{uv} = \mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star}) \\ (\mathbf{R}_{vu} = \mathbf{U}_{\mathbf{\Psi},\star}^H\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}) & (\mathbf{R}_{vv} = \mathbf{U}_{\mathbf{\Psi},\star}^H\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star}) \end{bmatrix}. \end{aligned} \tag{29}$$

From (13), it is easy to recognize that $\mathbf{R}_{uv}\mathbf{R}_{vv}^{-1} = \mathbf{F}^H$, making $\mathbf{F}^H$ the LMMSE filter in estimating $\mathbf{u}$ from $\mathbf{v}$. Correspondingly, $\mathbf{e} = \mathbf{u} - \mathbf{F}^H\mathbf{v}$ is the error in such an estimation, with covariance [1]

$$\begin{aligned} \mathbf{R}_{ee} = E[\mathbf{e}\mathbf{e}^H] &= \mathbf{R}_{uu} - \mathbf{R}_{uv}\mathbf{R}_{vv}^{-1}\mathbf{R}_{vu} \\ &= (\mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}. \end{aligned} \tag{30}$$

The trace of $\mathbf{R}_{ee}$ measures the mean-squared error (MSE) in estimating $\mathbf{u}$ from $\mathbf{v}$:

$$\text{MSE} = \text{tr}\{\mathbf{R}_{ee}\} = \text{tr}\{(\mathbf{U}_{\mathbf{\Psi},p}^H\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}\}. \tag{31}$$

With this interpretation, the output vector $\mathbf{y}$ may be viewed as a weighted (colored) error vector, with covariance

$$\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{L}^H\mathbf{V}^H(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}\mathbf{V}\mathbf{L}, \tag{32}$$

Therefore, $\text{tr}\{\mathbf{R}_{yy}\}$ is a weighted MSE. However, $\text{tr}\{\mathbf{R}_{yy}\}$ is also the minimum value of the quadratic form $J = \text{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\}$ in (10). Thus, in the linearly-constrained problem of (3) the minimum value of the quadratic form J measures the weighted MSE in the LMMSE estimation problem of Fig. 1. The upper branch of the GSC diagram from $\mathbf{x}$ to $\mathbf{y}$ is the part of the filter $\mathbf{W}_o$ that satisfies the constraint $\mathbf{W}_o^H\mathbf{\Psi} = \mathbf{L}^H\mathbf{V}^H$, and the lower branch is the part that minimizes the quadratic form J, by minimizing the weighted MSE.

It was established in Theorem 5 that the linearly-constrained problem of (3) solves the quadratically-constrained problem of (6) when $\mathbf{\Psi} = \mathbf{S}^{1/2}$, $\mathbf{L}^H = \mathbf{D}^{1/2}$, and $\mathbf{V} = \mathbf{U}_{\mathbf{Q},r}$. Therefore, the GSC diagram in Fig. 1 will be a GSC diagram for (21), if $\mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1}$, $\mathbf{V}^H$, and $\mathbf{L}^H$ are replaced by $\mathbf{\Lambda}_{\mathbf{S},p}^{-1}$, $\mathbf{U}_{\mathbf{Q},r}^H$, and $\mathbf{D}^{1/2}$. The rest of the diagram remains the same, as $\mathbf{\Psi} = \mathbf{S}^{1/2}$ implies $\mathbf{U}_{\mathbf{\Psi},p} = \mathbf{U}_{\mathbf{S},p}$ and $\mathbf{U}_{\mathbf{\Psi},\star} = \mathbf{U}_{\mathbf{S},\star}$. Theorem 5 also shows that when $\mathbf{\Psi} = \mathbf{S}^{1/2}$ and $\mathbf{L}^H = \mathbf{D}^{1/2}$ the smallest achievable weighted MSE is equal to the minimum value of J in the quadratically-constrained problem of (6), and is obtained when the left-orthogonal matrix $\mathbf{V}$ consists of the r principal eigenvectors of the signal-to-signal-plus-noise ratio matrix $\mathbf{Q}$.

IV. Conclusions

The constrained minimization problems considered here arise in the design of multi-rank beamformers for radar, sonar, seismology, and wireless communications, and in the design of precoders and equalizers for digital communications. The aim is to minimize variance (or power) under a constraint that certain subspace signals are passed through the matrix filter undistorted. This leads to quadratic minimizations under a set of linear and quadratic constraints. Closed-form solutions to these problems have been derived and connections between the linearly- and quadratically-constrained minimizations have been established.

Evidently, deriving the closed-form solution for the nonconvex quadratically-constrained problem requires the majorization argument of Theorem 1 and Poincare's separation theorem, as shown in [1]. Alternatively, one may solve such a nonconvex problem by solving a class of convex linearly-constrained problems to determine a set of candidate solutions (which all satisfy the quadratic constraint), and then finding the best candidate in this set. The set of candidate solutions is nonconvex, and hence the search for the best candidate is performed inside a nonconvex set. Interestingly, the majorization result of Theorem 1 and Poincare's separation theorem appear to be the right tools to perform such a search.

In addition, we have presented generalized sidelobe canceller diagrams for our solutions, clarifying in the process the connections between the linearly- and quadratically-constrained minimizations and LMMSE estimations.

References

[1] A. Pezeshki, L. L. Scharf, and H. Cox, "On linearly- and quadratically-constrained quadratic minimization problems for beamforming and diversity combining," submitted to *SIAM Journal on Matrix Analysis and Applications*, 2005.

[2] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*. New York: Academic Press, 1979.

[3] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Applications in Statistics and Econometrics*. Wiley, 1988.

[4] H. Cox, "Sensitivity considerations in adaptive beamforming," in *Signal Processing (Proc. NATO Advanced Study Inst. Signal Processing with Particular Reference to Underwater Acoust., Loughborough, U.K., Aug. 1972)*, J. W. R. Griffiths, P. L. Stocklin, and C. Van Schooneveld, Eds. New York and London: Academic, 1973, pp. 619–645.

[5] B. D. Van Veen, W. van Drongelen, M. Yuchtman, and A. Suzuki, "Localization of brain electrical activity via linearly constrained minimum variance spatial filtering," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 867–880, Sept. 1997.

[6] M. Lundberg, L. L. Scharf, and A. Pezeshki, "Multi-rank Capon beamforming," in *Conf. Rec. Thirty-eighth Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 2004, pp. 2335–2339.

[7] L. L. Scharf, A. Pezeshki, and M. Lundberg, "Multi-rank adaptive beamforming," in *Proc. 15th IEEE Signal Processing Workshop*, Bordeaux, France, July 17-20 2005.

[8] Y. Ding, T. N. Davidson, Z. Luo, and K. M. Wong, "Minimum BER block precoders for zero-forcing equalization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2410–2423, Sept. 2003.

[9] D. P. Palomar, J. M. Cioffi, and M. A. Lagunas, "Joint Tx-Rx beamforming design for multicarrier MIMO channels: A unified framework for convex optimization," *IEEE Trans. Signal Process.*, vol. 51, pp. 2381–2401, Sept. 2003.

[10] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath, "Optimal designs for space-time linear precoders and decoders," *IEEE Trans. Signal Process.*, vol. 50, pp. 1051–1064, May 2002.

[11] S. Shahbazpanahi, M. Beheshti, A. B. Gershman, M. Gharavi-Alkhansari, and K. M. Wong, "Minimum variance linear receivers for multiaccess MIMO wireless systems with space-time block coding," *IEEE Trans. Signal Process.*, vol. 52, pp. 3306–3313, Dec. 2004.

[12] M. K. Tsatsanis and Z. Xu, "Performance analysis of minimum variance CDMA receivers," *IEEE Trans. Signal Process.*, vol. 46, pp. 3014–3022, Nov. 1998.

[13] L. J. Griffiths and C. W. Jim, "An alternative approach to linearly constrained adaptive beamforming," *IEEE Trans. Antennas Propagation*, vol. AP-30, pp. 27–34, Jan. 1982.