Available online at www.sciencedirect.com
ScienceDirect

Journal of the Franklin Institute 347 (2010) 818–835

Journal of The Franklin Institute

www.elsevier.com/locate/jfranklin

The geometry of linearly and quadratically constrained optimization problems for signal processing and communications ☆

Ali Pezeshki[a], Louis L. Scharf[a,b], Edwin K.P. Chong[a,c,*]

[a]Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA
[b]Department of Statistics, Colorado State University, Fort Collins, CO 80523, USA
[c]Department of Mathematics, Colorado State University, Fort Collins, CO 80523, USA

Received 17 December 2009; accepted 15 March 2010

Abstract

Constrained minimization problems considered in this paper arise in the design of beamformers for radar, sonar, 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 present solutions to these problems and establish a connection between them. A majorization result for matrix trace and Poincare's separation theorem play key roles in establishing the connection. We show that our solutions can be formulated as generalized sidelobe cancellers (GSCs), which tie our constrained minimizations to linear minimum mean-squared error (LMMSE) estimations. We then express our solutions in terms of oblique projection matrices and establish the geometry of our constrained minimizations.

Keywords: Constrained minimization; Generalized sidelobe canceller; Linear and quadratic constraints; Majorization; Multi-rank beamforming; Oblique projections; Poincare's separation theorem; Precoder and equalizer design; Quadratic forms

☆This work was supported in part by the Office of Naval Research (ONR) under Contract N00014-04-1-0084 and by NSF under Grants ECCS-0700559 and CCF-0916314.

*Corresponding author at: Department of Electrical and Computer Engineering, Colorado State University, Fort Collins, CO 80523, USA. Tel.: +1 970 491 7858; fax: +1 970 491 2249.

E-mail addresses: Ali.Pezeshki@colostate.edu (A. Pezeshki), Louis.Scharf@colostate.edu (L.L. Scharf), Edwin.Chong@colostate.edu (E.K.P. Chong).

0016-0032/$32.00
doi:10.1016/j.jfranklin.2010.03.005

1. Introduction

1.1. Constrained quadratic minimizations

The minimization of quadratic forms under linear or quadratic constraints is frequently encountered in signal processing and wireless communications. The quadratic form typically measures the power at the output of a filter. We wish to minimize this power under a constraint that a signal of interest is processed by the filter according to a design principle. For example, in minimum variance distortionless response (MVDR) beamforming [1–3] the power $J(\mathbf{w}) = \mathbf{w}^H\mathbf{R}\mathbf{w}$ at the beamformer output is minimized subject to the linear constraint $\mathbf{w}^H\boldsymbol{\psi} = 1$. Here $\mathbf{w}$ is the MVDR beamformer vector, $\mathbf{R}$ is the covariance matrix of the array measurement vector, $\boldsymbol{\psi}$ is the signal signature vector that carries the relative phases and amplitudes induced on the array elements by a propagating waveform, and superscript H denotes Hermitian transpose. The constraint guarantees that a source with known signature vector $\boldsymbol{\psi}$ is passed by the beamformer undistorted.

In this paper, we consider minimizing a quadratic function $J(\mathbf{W})$,

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

with respect to $\mathbf{W} \in \mathbb{W}$, subject to a set of linear or quadratic constraints. Here, the set $\mathbb{W} = \{\mathbf{W} | \mathbf{W} \in \mathbb{C}^{n\times r}; \mathrm{rank}\{\mathbf{W}\} = r<n\}$ is the set of all $n \times r$ complex matrices of rank $r<n$, $\mathbf{R} \in \mathbb{C}^{n\times n}$ is a positive definite (PD) complex matrix, and $\mathrm{tr}\{\cdot\}$ is the trace operator. The quadratic form $\mathbf{W}^H\mathbf{R}\mathbf{W}$ is the covariance matrix of $\mathbf{y} = \mathbf{W}^H\mathbf{x}$, where $\mathbf{x} \in \mathbb{C}^n$ is a zero-mean random vector with covariance $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$, the quadratic function $J(\mathbf{W}) = \mathrm{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} = E[\mathbf{y}^H\mathbf{y}]$ is total variance, and $E[\cdot]$ is expectation. The constraints are as follows.

Linear constraint: The constraint is

$$\mathbf{W}^H\boldsymbol{\Psi} = \mathbf{C}^H, \tag{2}$$

where $\boldsymbol{\Psi} \in \mathbb{C}^{n\times p}$ and $\mathbf{C} \in \mathbb{C}^{p\times r}$ $(r\leq p)$ are full-rank matrices. Thus, our constrained minimization problem is

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

The linear constraint (2) ensures that the action of $\mathbf{W}^H$ on $\boldsymbol{\Psi}$ meets the design constraint $\mathbf{C}^H$. We shall explain the relevance of Eq. (3) in signal processing and wireless communication in Section 1.2. From here on, without loss of generality, we take the constraint matrix $\mathbf{C}^H$ to be of the form $\mathbf{C}^H = \mathbf{L}^H\mathbf{V}^H$, where $\mathbf{L} \in \mathbb{C}^{r\times r}$ is nonsingular and $\mathbf{V} \in \mathbb{C}^{p\times r}$ is left-orthogonal. That is, we assume $\mathbf{V}^H\mathbf{V} = \mathbf{I}$ but $\mathbf{V}\mathbf{V}^H = \mathbf{P}_\mathbf{V}$, where $\mathbf{P}_\mathbf{V}$ is the orthogonal projection onto the r-dimensional subspace $\langle \mathbf{V} \rangle$ spanned by the columns of $\mathbf{V}$.

Quadratic constraint: The constraint is

$$\mathbf{W}^H\mathbf{S}\mathbf{W} = \mathbf{D}, \tag{4}$$

where $\mathbf{S} \in \mathbb{C}^{n\times n}$ is a positive semi-definite (PSD) matrix of rank p and $\mathbf{D} \in \mathbb{C}^{r\times r}$ is a PD matrix, with $r\leq p<n$. Thus, the constrained minimization problem is

$$\min_{\mathbf{W}\in\mathbb{W}} \quad J(\mathbf{W}) = \mathrm{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\}$$
$$\text{subject to} \quad \mathbf{W}^H\mathbf{S}\mathbf{W} = \mathbf{D}. \tag{5}$$

The quadratic constraint in Eq. (5) 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 covariance $\mathbf{D}$.

1.2. Relevance in signal processing and wireless communication

The constrained minimization problems (3) and (5) are relevant in the design of *multi-rank MVDR beamformers* for resolving nonplanar wavefronts with unknown or unpredictable signatures [4–10]. They are also relevant for the design of precoders and equalizers for digital communications [11–16].

In beamforming, $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H]$ is the nonsingular covariance matrix of a complex random measurement vector $\mathbf{x}$, observed by an n-element sensor array, and $\mathbf{y} = \mathbf{W}^H\mathbf{x}$ is the output of a matrix beamformer $\mathbf{W}$ (see Fig. 1(a)). Consequently, the quadratic form $\mathrm{tr}\{E[\mathbf{y}\mathbf{y}^H]\} = \mathrm{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\} = J(\mathbf{W})$ is the power at the output of the beamformer. The matrices $\mathbf{\Psi}$ and $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H]$ are, respectively, the signal subspace matrix and the signal covariance matrix, which characterize a complex wavefront $\mathbf{s}$ that is incident on the array from an angle θ. That is, $\mathbf{s} = \mathbf{\Psi}\mathbf{a}$ and $\mathbf{S} = \mathbf{\Psi}\mathbf{\Gamma}\mathbf{\Psi}^H$, where $\mathbf{a} \in \mathbb{C}^p$ is a complex random vector with covariance matrix $\mathbf{\Gamma} = E[\mathbf{a}\mathbf{a}^H]$. Typically, the columns of $\mathbf{\Psi}$ and $\mathbf{S}$ lie in the subspace spanned by the first $p < n$ Slepian basis vectors, where p depends on the product of the angular spread of the wavefront and the number of sensors [17,18,9].

The problem in Eq. (3) is to minimize the output power $J(\mathbf{W})$ under the linear constraint that the signal subspace matrix $\mathbf{\Psi}$ is transformed to $\mathbf{C}^H$. The constraint matrix $\mathbf{C}^H$ may be *data-independent*, in which case Eq. (3) corresponds to the linearly constrained MVDR beamformers of [4–6], or it may be *data-dependent*, in which case Eq. (3) corresponds to the matched direction beamformers and matched subspace beamformers of [8,9]. In matched direction beamforming the columns of $\mathbf{C}$ are selected as the dominant eigenvectors of

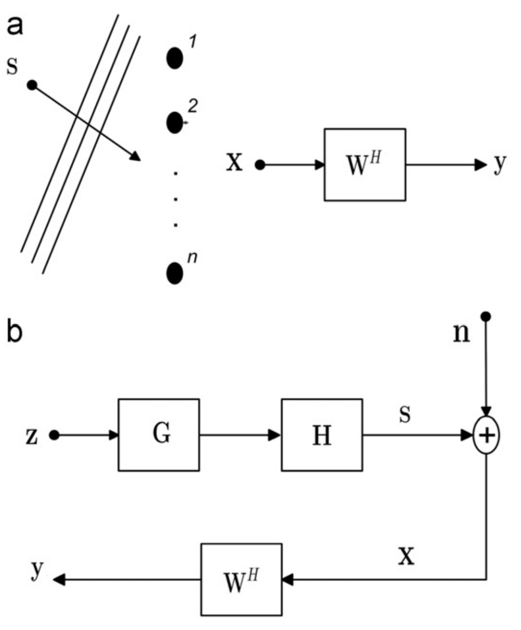

Fig. 1. (a) Multi-rank beamforming and (b) precoder and equalizer design.

$(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}$, while in matched subspace beamforming they are selected as the subdominant eigenvectors of $(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}$.

In the quadratically constrained problem (5), the objective is to minimize the output power $J(\mathbf{W})$ under the quadratic constraint that a wavefront with covariance matrix $\mathbf{S}$ is imaged into an output vector with covariance $\mathbf{D}$. This is the quadratically constrained multi-rank Capon beamforming problem considered in [7].

We note that when $\mathbf{W}$ is a vector and $\mathbf{s}$ is a wavefront known to within a complex constant, the beamforming problems of Eqs. (3) and (5) are equivalent, as they yield the well known (rank-1) MVDR (or Capon) beamformer [1–3].

Now consider the communication system in Fig. 1(b), where $\mathbf{z} \in \mathbb{C}^p$ is a p-dimensional message vector with covariance matrix $E[\mathbf{z}\mathbf{z}^H] = \mathbf{I}$, $\mathbf{G} \in \mathbb{C}^{m\times p}$ is a precoder matrix, $\mathbf{H} \in \mathbb{C}^{n\times m}$ is a channel matrix, and $\mathbf{n} \in \mathbb{C}^n$ is a proper complex noise vector with covariance matrix $E[\mathbf{n}\mathbf{n}^H] = \mathbf{N} \in \mathbb{C}^{n\times n}$. The matrix $\mathbf{W} \in \mathbb{C}^{n\times r}$ is an equalizer matrix that estimates the message vector $\mathbf{z}$ from the noisy channel output $\mathbf{x} = \mathbf{HGz} + \mathbf{n}$. The matrices $\mathbf{R} = E[\mathbf{x}\mathbf{x}^H] = \mathbf{S} + \mathbf{N}$ and $\mathbf{S} = E[\mathbf{s}\mathbf{s}^H] = \mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{HGG}^H\mathbf{H}^H$ are, respectively, the covariance matrices of the channel output in the presence and absence of noise. The matrix $\mathbf{\Psi} = \mathbf{HG}$ characterizes the combined action of the channel matrix $\mathbf{H}$ and the precoder matrix $\mathbf{G}$. With these interpretations, the linearly constrained problem of Eq. (3) is to minimize the power $J(\mathbf{W}) = \mathrm{tr}\{\mathbf{W}^H\mathbf{RW}\}$ at the receiver, under the constraint that the direct path (noise free) $\mathbf{W}^H\mathbf{\Psi} = \mathbf{W}^H\mathbf{HG}$ from the precoder to the equalizer is equal to $\mathbf{C}^H = \mathbf{L}^H\mathbf{V}^H$. When $\mathbf{L} = \mathbf{I}$, this constraint may be viewed as a *first-order zero-forcing equalization*. In the quadratically constrained problem of Eq. (5), the objective is to minimize the power $J(\mathbf{W})$ at the receiver, under the constraint that the message covariance matrix is transformed to the constraint covariance matrix $\mathbf{D}$. When $\mathbf{D} = \mathbf{I}$, this may be viewed as a *second-order zero-forcing equalization*. When solving Eqs. (3) and (5) for a joint precoder and equalizer design, a constraint on the transmit power, e.g. $\mathrm{tr}\{\mathbf{GG}^H\} \leq \varepsilon^2$, is also imposed. In such a case, the minimization problems of Eqs. (3) and (5) are solved for the equalizer filter $\mathbf{W}$, as if $\mathbf{G}$ were known. Then, the precoder filter $\mathbf{G}$ is designed under the transmit power constraint. The solution to the constrained minimization over $\mathbf{G}$ is different for different precoder designs [12–16] and will not be further elaborated here.

1.3. Main results

We present solutions for Eqs. (3) and (5) and 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 form (2), with $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S}$, $\mathbf{C}^H\mathbf{C} = \mathbf{L}^H\mathbf{V}^H\mathbf{VL} = \mathbf{D}$, $\mathbf{V} \in \mathbb{V}$, and $\mathbb{V}$ the set of all $p \times r$ left-orthogonal complex matrices, is the minimum value of J under the quadratic constraint of Eq. (4). 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 2, and $\mathbf{S}^{H/2} = (\mathbf{S}^{1/2})^H$. The key to this connection is a majorization result for matrix trace and Poincare's separation theorem. We note that majorization theory [19] has been used before for solving constrained minimization problems in signal processing and communications. Examples of such work are [14,16].

We present illuminating circuit diagrams, called generalized sidelobe canceller (GSC) diagrams [20], for the solutions to Eqs. (3) and (5). The GSC diagrams allow for establishing connections between our constrained minimizations and linear minimum mean-squared error (LMMSE) estimations. We show that the minimum value of the quadratic form J in Eqs. (3) and (5) may in fact be interpreted as a weighted mean-squared error (MSE) in an LMMSE estimation. Moreover, we establish the geometry of our constrained minimizations and show that the solutions to Eqs. (3) and (5) can be expressed in terms of oblique projections [21,22].

We note that our aim in this paper is to provide geometrical insights and to highlight the relevance of GSC diagrams for analyzing the solutions to Eqs. (3) and (5). The reader is referred to [6–9] for numerical examples concerning the use of Eqs. (3) and (5) for constructing multi-rank MVDR beamformers. Similarly, numerical results concerning the use of Eqs. (3) and (5) for precoder and equalizer design can be found in [12–16].

2. Preliminaries

Notations and typographic conventions: Scalars are represented by lowercase symbols in Italic fonts, e.g. n. Vectors are represented by lowercase boldface symbols, e.g. $\mathbf{x}$. All matrices are represented by uppercase boldface symbols, e.g. $\mathbf{A}$. Given an $n \times m$ complex matrix $\mathbf{A} \in \mathbb{C}^{n\times m}$, we denote the ith column of $\mathbf{A}$ by $\mathbf{a}_i$, the matrix consisting of the first p columns of $\mathbf{A}$ by $\mathbf{A}_p$, and the matrix consisting of the remaining columns of $\mathbf{A}$ by $\mathbf{A}_\star$. Naturally, $\mathbf{A} = [\mathbf{A}_p \mathbf{A}_\star]$. When $\mathbf{A} \in \mathbb{C}^{n\times n}$ is diagonal, we use $\mathbf{A}_p$ and $\mathbf{A}_\star$ to denote the upper left $p \times p$ and the lower right $(n-p) \times (n-p)$ diagonal blocks of $\mathbf{A}$. We use $\langle \mathbf{A} \rangle$ to denote the linear subspace spanned by the columns of $\mathbf{A}$, and $\mathbf{P_A}$ to denote the orthogonal projection matrix onto $\langle \mathbf{A} \rangle$. Correspondingly, $\mathbf{P}_\mathbf{A}^\perp = \mathbf{I} - \mathbf{P_A}$ denotes the orthogonal projection onto $\langle \mathbf{A} \rangle^\perp$, the linear subspace orthogonal to $\langle \mathbf{A} \rangle$. The matrix $\mathbf{A} \in \mathbb{C}^{n\times m}$ $(n > m)$ is called left-orthogonal when $\mathbf{A}^H\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\mathbf{A}^H = \mathbf{P_A}$. We use $\mathbf{A}^\dagger = (\mathbf{A}^H\mathbf{A})^{-1}\mathbf{A}^H$ to denote the Moore–Penrose pseudo inverse of a tall matrix $\mathbf{A} \in \mathbb{C}^{n\times m}$ $(n > m)$ of rank m, having $\mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$ and $\mathbf{A}\mathbf{A}^\dagger = \mathbf{P_A}$ [23].

We construct the SVD of a rank $p \le \min(n,m)$ rectangular matrix $\mathbf{A} \in \mathbb{C}^{n\times m}$ as

$$\mathbf{A} = \mathbf{U_A}\boldsymbol{\Sigma}_\mathbf{A}\mathbf{V}_\mathbf{A}^H = [\mathbf{U}_{\mathbf{A},p}\mathbf{U}_{\mathbf{A},\star}]\begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{A},p} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{V}_{\mathbf{A},p}^H \\ \mathbf{V}_{\mathbf{A},\star}^H \end{bmatrix} = \mathbf{U}_{\mathbf{A},p}\boldsymbol{\Sigma}_{\mathbf{A},p}\mathbf{V}_{\mathbf{A},p}^H, \tag{6}$$

where $\boldsymbol{\Sigma}_{\mathbf{A},p} = \mathrm{diag}(\sigma_{\mathbf{A},1}, \ldots, \sigma_{\mathbf{A},p})$ carries the nonzero singular values of $\mathbf{A}$, and $\mathbf{U}_{\mathbf{A},p} \in \mathbb{C}^{n\times p}$, $\mathbf{U}_{\mathbf{A},\star} \in \mathbb{C}^{n\times(n-p)}$, $\mathbf{V}_{\mathbf{A},p} \in \mathbb{C}^{m\times p}$, and $\mathbf{V}_{\mathbf{A},p} \in \mathbb{C}^{m\times(m-p)}$ are all left-orthogonal. Clearly, $\langle \mathbf{A} \rangle = \langle \mathbf{U}_{\mathbf{A},p} \rangle$, $\langle \mathbf{A} \rangle^\perp = \langle \mathbf{U}_{\mathbf{A},\star} \rangle$, $\mathbf{P_A} = \mathbf{P}_{\mathbf{U}_{\mathbf{A},p}}$, and $\mathbf{P}_\mathbf{A}^\perp = \mathbf{P}_{\mathbf{U}_{\mathbf{A},\star}}$.

We construct the EVD of a rank $p \le n$ PSD matrix $\mathbf{A} \in \mathbb{C}^{n\times n}$, with nonzero eigenvalues $\lambda_{\mathbf{A},1}^2, \ldots, \lambda_{\mathbf{A},p}^2$, as in Eq. (6), by replacing $\mathbf{V_A}$ and $\boldsymbol{\Sigma}_\mathbf{A}$ with $\mathbf{U_A}$ and $\boldsymbol{\Lambda}_\mathbf{A}^2$, where

$$\boldsymbol{\Lambda}_\mathbf{A}^2 = \begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{A},p}^2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \in \mathbb{C}^{n\times n}; \quad \boldsymbol{\Lambda}_{\mathbf{A},p}^2 = \mathrm{diag}(\lambda_{\mathbf{A},1}^2, \ldots, \lambda_{\mathbf{A},p}^2). \tag{7}$$

Correspondingly, $\mathbf{A}^\dagger = \mathbf{U}_{\mathbf{A},p}\boldsymbol{\Lambda}_{\mathbf{A},p}^{-2}\mathbf{U}_{\mathbf{A},p}^H$, with $\boldsymbol{\Lambda}_{\mathbf{A},p}^{-2} = \mathrm{diag}(1/\lambda_{\mathbf{A},1}^2, \ldots, 1/\lambda_{\mathbf{A},p}^2)$, will be an EVD of $\mathbf{A}^\dagger$.

We assume that in all SVD's and EVD's, the singular values and eigenvalues are arranged in descending order, i.e. $\sigma_{\mathbf{A},1} \ge \cdots \ge \sigma_{\mathbf{A},p} > 0$ and $\lambda_{\mathbf{A},1}^2 \ge \cdots \ge \lambda_{\mathbf{A},p}^2 > 0$, unless otherwise stated or implied. Naturally, when the eigenvalues of $\mathbf{A} = \mathbf{U}_{\mathbf{A},p}\boldsymbol{\Lambda}_{\mathbf{A},p}^2\mathbf{U}_{\mathbf{A},p}^H$ are arranged in descending order, the eigenvalues of $\mathbf{A}^\dagger = \mathbf{U}_{\mathbf{A},p}\boldsymbol{\Lambda}_{\mathbf{A},p}^{-2}\mathbf{U}_{\mathbf{A},p}^H$ are arranged in ascending order.

We define one particular square-root of a rank $p \leq n$ PSD matrix $\mathbf{A} \in \mathbb{C}^{n \times n}$ as

$$\mathbf{A}^{1/2} = \mathbf{U}_{\mathbf{A},p} \mathbf{\Lambda}_{\mathbf{A},p}, \tag{8}$$

where $\mathbf{\Lambda}_{\mathbf{A},p} = \mathrm{diag}(\lambda_{\mathbf{A},1}, \ldots, \lambda_{\mathbf{A},p})$, $\lambda_{\mathbf{A},1}, \ldots, \lambda_{\mathbf{A},p} > 0$. It is clear that this square-root of $\mathbf{A}$ satisfies $\mathbf{A}^{1/2}\mathbf{A}^{H/2} = \mathbf{A}$ and $\mathbf{A}^{-H/2}\mathbf{A}^{-1/2} = \mathbf{A}^{\dagger}$, where $\mathbf{A}^{H/2} = (\mathbf{A}^{1/2})^H$ is the Hermitian transpose of $\mathbf{A}^{1/2}$, $\mathbf{A}^{-1/2} = \mathbf{\Lambda}_{\mathbf{A},p}^{-1}\mathbf{U}_{\mathbf{A},p}^H$ is the left-inverse of $\mathbf{A}^{1/2}$ (i.e. $\mathbf{A}^{-1/2}\mathbf{A}^{1/2} = \mathbf{I}$), $\mathbf{A}^{-H/2} = (\mathbf{A}^{-1/2})^H$, and $\mathbf{\Lambda}_{\mathbf{A},p}^{-1} = \mathrm{diag}(1/\lambda_{\mathbf{A},1}, \ldots, 1/\lambda_{\mathbf{A},p})$. When $\mathbf{A}$ is PD, $\mathbf{A}^{\dagger}$ is replaced by $\mathbf{A}^{-1}$.

Letting $\mathbf{A} \in \mathbf{C}^{n \times n}$ and $\mathbf{B} \in \mathbb{C}^{n \times n}$ be PD matrices, we say $\mathbf{A}$ is greater than $\mathbf{B}$, and denote it by $\mathbf{A} \succ \mathbf{B}$, when $\mathbf{A}-\mathbf{B}$ is PD. When $\mathbf{A} \succ \mathbf{B}$, $\mathbf{c}^H\mathbf{A}\mathbf{c} > \mathbf{c}^H\mathbf{B}\mathbf{c}$ for any nonzero complex vector $\mathbf{c} \in \mathbb{C}^n$.

Finally, three theorems play key roles in our developments: a majorization result for matrix trace, Poincare's separation theorem, and a few subspace identities.

Theorem 1 (*A majorization result for matrix trace [19, Chapter 9, H.1.h]*). *Let* $\mathbf{A} \in \mathbb{C}^{n \times n}$ *and* $\mathbf{B} \in \mathbb{C}^{n \times n}$ *be PSD matrices, with eigenvalues* $0 \leq \lambda_{\mathbf{A},1}^2 \leq \cdots \leq \lambda_{\mathbf{A},n}^2$ *and* $\lambda_{\mathbf{B},1}^2 \geq \cdots \geq \lambda_{\mathbf{B},n}^2 \geq 0$. *Then,*

$$\mathrm{tr}\{\mathbf{AB}\} \geq \sum_{i=1}^{n} \lambda_{\mathbf{A},i}^2 \lambda_{\mathbf{B},i}^2. \tag{9}$$

The equality holds when the eigenvectors of $\mathbf{A}$ *and* $\mathbf{B}$ *are equal, i.e.* $\mathbf{U}_{\mathbf{A}} = \mathbf{U}_{\mathbf{B}}$.[1]

Theorem 2 (*Poincare's separation theorem [24, Chapter 11, Theorem 10]*). *Let* $\mathbf{A} \in \mathbb{C}^{n \times n}$ *be a PSD matrix with eigenvalues* $0 \leq \lambda_{\mathbf{A},1}^2 \leq \cdots \leq \lambda_{\mathbf{A},n}^2$ *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* $0 \leq \lambda_{\mathbf{B},1}^2 \leq \cdots \leq \lambda_{\mathbf{B},r}^2$ *be the eigenvalues of* $\mathbf{B} = \mathbf{X}^H\mathbf{A}\mathbf{X} \in \mathbb{C}^{r \times r}$. *Then,*

$$\lambda_{\mathbf{A},i}^2 \leq \lambda_{\mathbf{B},i}^2 \leq \lambda_{\mathbf{A},n-r+i}^2; \quad i = 1, \ldots, r. \tag{10}$$

Theorem 3. *Let* $\mathbf{A} \in \mathbb{C}^{n \times m}$ $(n \geq m)$ *be a rank* $p < m \leq n$ *matrix and* $\mathbf{B} \in \mathbb{C}^{n \times n}$ *a PD matrix. Then,*

$$\langle \mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star} \rangle^{\perp} = \langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p} \rangle \tag{11}$$

and

$$\mathbf{P}_{\mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star}}^{\perp} = \mathbf{P}_{\mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}}. \tag{12}$$

In the special case where $p = m < n$ ($\mathbf{A}$ *is tall and full-rank*) *we also have*

$$\langle \mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star} \rangle^{\perp} = \langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p} \rangle = \langle \mathbf{B}^{-1/2}\mathbf{A} \rangle \tag{13}$$

and

$$\mathbf{P}_{\mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star}}^{\perp} = \mathbf{P}_{\mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}} = \mathbf{P}_{\mathbf{B}^{-1/2}\mathbf{A}}. \tag{14}$$

[1]Letting $\tilde{\lambda}_{\mathbf{A},i}^2$ be $\tilde{\lambda}_{\mathbf{A},i}^2 = \lambda_{\mathbf{A},n-i+1}^2$, we may write the trace inequality in Eq. (9) as $\mathrm{tr}\{\mathbf{AB}\} \geq \sum_{i=1}^{n} \tilde{\lambda}_{\mathbf{A},n-i+1}^2 \lambda_{\mathbf{B},i}^2$, where the $\tilde{\lambda}_{\mathbf{A},i}^2$ are the eigenvalues of $\mathbf{A}$, arranged in descending order. Then, it is interesting to note that the term on the right-hand side (RHS) of this inequality is a discrete-time convolution between $\{0, \{\tilde{\lambda}_{\mathbf{A},i}^2\}_{i=1}^n\}$ and $\{0, \{\lambda_{\mathbf{B},i}^2\}_{i=1}^n\}$, computed at index $n+1$.

Proof. See Appendix A.

3. Quadratic minimizations under linear and quadratic constraints

Theorem 4. *The minimum value* J_0 *of the quadratic form* J *in the linearly constrained minimization problem* (3) *and the minimizer* $\mathbf{W}=\mathbf{W}_0$ *are given by*

$$J_0 = \mathrm{tr}\{\mathbf{L}^H\mathbf{V}^H(\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi})^{-1}\mathbf{V}\mathbf{L}\}, \tag{15}$$

$$\mathbf{W}_0 = \mathbf{R}^{-1}\boldsymbol{\Psi}(\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi})^{-1}\mathbf{V}\mathbf{L}. \tag{16}$$

Proof. This is a simple linearly constrained minimization problem, which can be solved using the method of Lagrange multipliers [25, Chapter 19] and completing the square. □

The linear constraint $\mathbf{W}^H\boldsymbol{\Psi}=\mathbf{L}^H\mathbf{V}^H$ implies the modified linear constraint

$$\mathbf{W}^H\boldsymbol{\Psi}\mathbf{V} = \mathbf{L}^H, \tag{17}$$

which is a constraint that $\mathbf{W}^H$ images the r linear combinations $\boldsymbol{\Psi}\mathbf{V}$ as $\mathbf{L}^H$. However, the modified constraint does not imply the original one. The reason is that the original constraint in Eq. (3) enforces constraints on $p\geq r$ linear combinations of the columns of $\boldsymbol{\Psi}$, not just r of them. It is easy to show that the constraint in Eq. (3) is equivalent to

$$\mathbf{W}^H\boldsymbol{\Psi}[\mathbf{V}\ \mathbf{V}_\star] = \mathbf{L}^H[\mathbf{I}\ \mathbf{0}], \tag{18}$$

where $[\mathbf{V}\ \mathbf{V}_\star]\in\mathbb{C}^{p\times p}$ is an orthogonal matrix.[2] This means that under the original constraint in Eq. (3) $\mathbf{W}^H$ images the r linear combinations $\boldsymbol{\Psi}\mathbf{V}$ as $\mathbf{L}^H$ and the $p-r$ linear combinations $\boldsymbol{\Psi}\mathbf{V}_\star$ as zero. These may be called *zero-forcing* constraints.

Under the modified linear constraint (17), the minimum value J_0 of J and the minimizer $\mathbf{W}=\mathbf{W}_0$ are

$$J_0 = \mathrm{tr}\{\mathbf{L}^H(\mathbf{V}^H\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi}\mathbf{V})^{-1}\mathbf{L}\}, \tag{19}$$

$$\mathbf{W}_0 = \mathbf{R}^{-1}\boldsymbol{\Psi}\mathbf{V}(\mathbf{V}^H\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi}\mathbf{V})^{-1}\mathbf{L}. \tag{20}$$

It is easy to verify that Eqs. (19) and (20) reduce to Eqs. (15) and (16) when the left-orthogonal matrix $\mathbf{V}\in\mathbb{C}^{p\times r}$ carries r of the eigenvectors of $\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi}$. However, for any other choice of $\mathbf{V}$ we can show that

$$\mathrm{tr}\{\mathbf{L}^H(\mathbf{V}^H\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi}\mathbf{V})^{-1}\mathbf{L}\} < \mathrm{tr}\{\mathbf{L}^H\mathbf{V}^H(\boldsymbol{\Psi}^H\mathbf{R}^{-1}\boldsymbol{\Psi})^{-1}\mathbf{V}\mathbf{L}\}. \tag{21}$$

Therefore, the minimum value of the quadratic form J under the modified linear constraint $\mathbf{W}^H\boldsymbol{\Psi}\mathbf{V}=\mathbf{L}^H$ is always less than or equal to its minimum value under the original linear constraint $\mathbf{W}^H\boldsymbol{\Psi}=\mathbf{L}^H\mathbf{V}^H$. We defer the derivation of Eq. (21) to the end of Section 4 and instead present an intuitive explanation here. As mentioned earlier in this section, the modified constraint (17) does not enforce the $p-r$ zero-forcing constraints that are enforced by the original constraint $\mathbf{W}^H\boldsymbol{\Psi}\mathbf{V}=\mathbf{L}^H$ (or equivalently (18)). Consequently,

[2] Post-multiplying $\mathbf{W}^H\boldsymbol{\Psi}=\mathbf{L}^H\mathbf{V}^H$ by $[\mathbf{V}\ \mathbf{V}_\star]$ yields the constraint in Eq. (18). Conversely, post-multiplying Eq. (18) by $[\mathbf{V}\ \mathbf{V}_\star]^H$ yields the original constraint in Eq. (3).

under the modified constraint there are more degrees of freedom available for minimizing the quadratic form J.

We now consider minimizing $J(\mathbf{W})$ under the quadratic constraint (4).

Theorem 5. *The minimum value J_0 of the quadratic form J in the quadratically constrained minimization problem* (5) *and the minimizer $\mathbf{W}=\mathbf{W}_0$ are given by*

$$J_0 = \mathrm{tr}\{\mathbf{U}_{\mathbf{Q},r}^H \mathbf{Q}^{-1} \mathbf{U}_{\mathbf{Q},r} \mathbf{\Lambda}_{\mathbf{D}}^2\} = \mathrm{tr}\{\mathbf{\Lambda}_{\mathbf{Q},r}^{-2} \mathbf{\Lambda}_{\mathbf{D}}^2\} = \sum_{i=1}^{r} \frac{\lambda_{\mathbf{D},i}^2}{\lambda_{\mathbf{Q},i}^2}, \tag{22}$$

and

$$\mathbf{W}_0 = \mathbf{R}^{-1}\mathbf{S}^{1/2}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}. \tag{23}$$

The matrix $\mathbf{Q} \in \mathbb{C}^{p\times p}$ is the signal-to-signal-plus-noise ratio matrix[3] *and is defined as*

$$\mathbf{Q} = \mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2} = \mathbf{\Lambda}_{\mathbf{S},p}^H \mathbf{U}_{\mathbf{S},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{S},p} \mathbf{\Lambda}_{\mathbf{S},p}. \tag{24}$$

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

3.1. Connections between linearly and quadratically constrained problems

Given $\mathbf{S}$ and $\mathbf{D}$ in the quadratically constrained problem (5), consider a corresponding class of linearly constrained problems of form (3), in which

$$\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S} \quad \text{and} \quad \mathbf{L}^H\mathbf{V}^H\mathbf{V}\mathbf{L} = \mathbf{L}^H\mathbf{L} = \mathbf{D}, \tag{25}$$

and $\mathbf{V} \in \mathbb{V}$ is any $p \times r$ left-orthogonal complex matrix. This equation characterizes a class of linear constraints of form (2) that are quadratically equivalent to the quadratic constraint in Eq. (4).

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. $\mathbf{VT}$, 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 Eq. (25) $\mathbf{L}^H = \mathbf{D}^{1/2} = \mathbf{U}_{\mathbf{D}}\mathbf{\Lambda}_{\mathbf{D}}$.

We also assume that the rank-p matrix $\mathbf{\Psi} \in \mathbb{C}^{n\times p}$ is of the form

$$\mathbf{\Psi} = \mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}_{\mathbf{\Psi},p}\mathbf{I}, \tag{26}$$

where $\mathbf{U}_{\mathbf{\Psi},p} \in \mathbb{C}^{n\times p}$ is left-orthogonal and $\mathbf{\Sigma}_{\mathbf{\Psi},p}$ is diagonal: $\mathbf{\Sigma}_{\mathbf{\Psi},p} = \mathrm{diag}(\sigma_{\mathbf{\Psi},1}, \ldots, \sigma_{\mathbf{\Psi},p})$, $\sigma_{\mathbf{\Psi},1} \geq \cdots \geq \sigma_{\mathbf{\Psi},p} > 0$. One would expect to see an orthogonal matrix $\mathbf{V}_{\mathbf{\Psi}}^H \in \mathbb{C}^{p\times p}$ in place of identity on the right-hand side of Eq. (26). But $\mathbf{V}_{\mathbf{\Psi}}^H$ may always be absorbed in the left-orthogonal matrix $\mathbf{V} \in \mathbb{C}^{p\times r}$ in the constraint matrix $\mathbf{C}^H = \mathbf{L}^H\mathbf{V}^H$, as $(\mathbf{V}^H\mathbf{V}_{\mathbf{\Psi}})^H \in \mathbb{C}^{p\times r}$ will

[3] 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 $\mathbf{N}$. Since 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 $\mathbf{R}$, the matrix $\mathbf{Q}$ may be viewed as the signal-to-signal-plus-noise ratio matrix. When $\mathbf{S}$ and $\mathbf{R}$ are circulant matrices the eigenvalues of $\mathbf{Q}$ are $\lambda_{\mathbf{Q},i}^2 = \mathrm{SNR}_i/(1+\mathrm{SNR}_i)$, where SNR_i is the signal-to-noise ratio (SNR) at the i th discrete Fourier transform (DFT) mode [26].

be left-orthogonal. Considering Eq. (26), it is easy to see that $\mathbf{\Psi}\mathbf{\Psi}^H = \mathbf{S} = \mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}^2_{\mathbf{S},p}\mathbf{U}^H_{\mathbf{S},p}$ implies $\mathbf{\Psi} = \mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p} = \mathbf{S}^{1/2}$.

Theorem 6. *The smallest achievable value for the quadratic form J in the class of linearly constrained problems characterized by Eq.* (25) *is*

$$J_0 = \sum_{i=1}^{r} \frac{\lambda^2_{\mathbf{D},i}}{\lambda^2_{\mathbf{Q},i}}, \tag{27}$$

which is equal to the minimum value of J in the quadratically constrained problem (5). *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* $\mathbf{W} = \mathbf{W}_0$ *is equal to* $\mathbf{W}_0$ *in Eq.* (23).

Proof. See Appendix C.

This theorem shows that the quadratically constrained problem (5) is equivalent to the following linearly constrained problem:

$$\min_{\mathbf{W}\in\mathbb{W},\mathbf{V}\in\mathbb{V}} \quad J = \mathrm{tr}\{\mathbf{W}^H\mathbf{R}\mathbf{W}\}$$
$$\text{subject to} \quad \mathbf{W}^H\mathbf{S}^{1/2} = \mathbf{D}^{1/2}\mathbf{V}^H. \tag{28}$$

4. Geometry and generalized sidelobe canceller diagrams

We now show that the solutions to Eqs. (3) and (5) can be expressed as GSCs [20]. GSCs are particularly illuminating, as they tie the constrained minimizations of Eqs. (3) and (5) to LMMSE estimations and oblique projection operators.

Inserting $\mathbf{\Psi} = \mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}_{\mathbf{\Psi},p}$ in Eq. (16) yields

$$\begin{aligned}\mathbf{W}_0 &= \mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}_{\mathbf{\Psi},p}(\mathbf{\Sigma}^H_{\mathbf{\Psi},p}\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}_{\mathbf{\Psi},p})^{-1}\mathbf{V}\mathbf{L}\\ &= \mathbf{R}^{-H/2}\mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}}\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L},\end{aligned} \tag{29}$$

where

$$\mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}} = \mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}(\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-H/2} \tag{30}$$

is the orthogonal projection matrix onto $\langle \mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}\rangle$. From Theorem 3, we have

$$\mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}} = \mathbf{I}-\mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},\star}} = \mathbf{I}-\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},\star}(\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star})^{-1}\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}^{1/2}. \tag{31}$$

Using Eq. (31), we may rewrite Eq. (29) as

$$\begin{aligned}\mathbf{W}_0 &= \mathbf{R}^{-H/2}[\mathbf{I}-\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},\star}(\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star})^{-1}\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}^{1/2}]\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L}\\ &= [\mathbf{U}_{\mathbf{\Psi},p}-\mathbf{U}_{\mathbf{\Psi},\star}(\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star})^{-1}\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}]\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L} = [\mathbf{U}_{\mathbf{\Psi},p}-\mathbf{U}_{\mathbf{\Psi},\star}\mathbf{F}]\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L},\end{aligned} \tag{32}$$

where

$$\mathbf{F} = (\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star})^{-1}\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}. \tag{33}$$

The last expression in Eq. (32) is the GSC representation of $\mathbf{W}_0$. The solution $\mathbf{W}_0$ is a filter (beamformer or equalizer) that takes the measurement vector $\mathbf{x}$ to the output vector $\mathbf{y} = \mathbf{W}_0^H \mathbf{x}$:

$$\mathbf{y} = \mathbf{W}_0^H \mathbf{x} = \mathbf{L}^H \mathbf{V}^H \mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1} [\mathbf{u} - \mathbf{F}^H \mathbf{v}] = \mathbf{L}^H \mathbf{V}^H \mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1} \mathbf{e}, \tag{34}$$

where $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{e}$ are the following vectors, as illustrated in the *GSC diagram* of Fig. 2:

$$\mathbf{u} = \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{x}, \quad \mathbf{v} = \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{x}, \quad \text{and} \quad \mathbf{e} = \mathbf{u} - \mathbf{F}^H \mathbf{v}. \tag{35}$$

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

$$E \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} [\mathbf{u}^H \;\; \mathbf{v}^H] = \begin{bmatrix} (\mathbf{R}_{uu} = \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p}) & (\mathbf{R}_{uv} = \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star}) \\ (\mathbf{R}_{vu} = \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p}) & (\mathbf{R}_{vv} = \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star}) \end{bmatrix}. \tag{36}$$

It is easy to recognize that $\mathbf{F}^H = \mathbf{R}_{uv} \mathbf{R}_{vv}^{-1}$ is 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

$$\begin{aligned} \mathbf{R}_{ee} &= E[\mathbf{e}\mathbf{e}^H] = \mathbf{R}_{uu} - \mathbf{R}_{uv} \mathbf{R}_{vv}^{-1} \mathbf{R}_{vu} \\ &= \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p} - \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star} (\mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star})^{-1} \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p} \\ &= \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{1/2} (\mathbf{I} - \mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_\star}) \mathbf{R}^{H/2} \mathbf{U}_{\mathbf{\Psi},p}. \end{aligned} \tag{37}$$

Using Theorem 3, we may write

$$\mathbf{I} - \mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{\Psi},\star}} = \mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{\Psi},p}} = \mathbf{R}^{-1/2} \mathbf{U}_{\mathbf{\Psi},p} (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1} \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-H/2}, \tag{38}$$

and then simplify Eq. (37) to

$$\mathbf{R}_{ee} = (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1}. \tag{39}$$

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

$$\mathrm{MSE} = \mathrm{tr}\{\mathbf{R}_{ee}\} = \mathrm{tr}\{(\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1}\}. \tag{40}$$

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

$$\mathbf{R}_{yy} = E[\mathbf{y}\mathbf{y}^H] = \mathbf{L}^H \mathbf{V}^H \mathbf{\Sigma}_{\mathbf{\Psi},p}^{-1} (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1} \mathbf{\Sigma}_{\mathbf{\Psi},p}^{-H} \mathbf{V} \mathbf{L} = \mathbf{L}^H \mathbf{V}^H (\mathbf{\Psi}^H \mathbf{R}^{-1} \mathbf{\Psi})^{-1} \mathbf{V} \mathbf{L}, \tag{41}$$

where the last equality follows from $\mathbf{\Psi} = \mathbf{U}_{\mathbf{\Psi},p} \mathbf{\Sigma}_{\mathbf{\Psi},p}$. Therefore, $\mathrm{tr}\{\mathbf{R}_{yy}\}$ is a weighted MSE. However, $\mathrm{tr}\{\mathbf{R}_{yy}\}$ is also the minimum value of the quadratic form $J = \mathrm{tr}\{\mathbf{W}^H \mathbf{R} \mathbf{W}\}$ in (15). Thus, in the linearly constrained problem (3) the minimum value of the quadratic form J measures the weighted MSE in the LMMSE estimation problem in Fig. 2. The upper branch of the GSC diagram from $\mathbf{x}$ to $\mathbf{y}$ is the part of $\mathbf{W}_0$ that satisfies the constraint

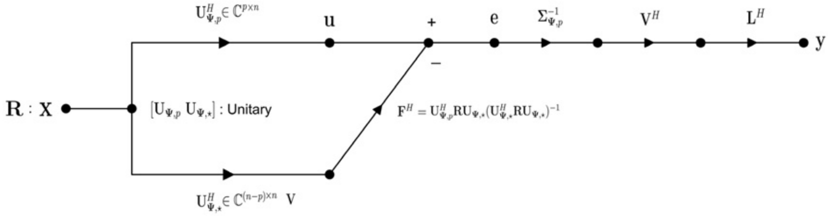

Fig. 2. Generalized sidelobe canceller (GSC) diagram.

$\mathbf{W}_0^H \boldsymbol{\Psi} = \mathbf{L}^H \mathbf{V}^H$, and the lower branch is the part that minimizes the quadratic form J, by minimizing the weighted MSE.

Let us define $\overline{\mathbf{W}}^H = \mathbf{U}_{\boldsymbol{\Psi},p}^H$ and $\underline{\mathbf{W}}^H = \mathbf{F}^H \mathbf{U}_{\boldsymbol{\Psi},\star}^H$ as the upper and lower branches of the GSC diagram from $\mathbf{x}$ to $\mathbf{e}$. Then, we have

$$\underline{\mathbf{W}}^H = \mathbf{F}^H \mathbf{U}_{\boldsymbol{\Psi},\star}^H = \overline{\mathbf{W}}^H \mathbf{K}^H, \tag{42}$$

where

$$\mathbf{K}^H = \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}(\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star})^{-1}\mathbf{U}_{\boldsymbol{\Psi},\star}^H = \mathbf{R}^{1/2}\mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_{\boldsymbol{\Psi},\star}}\mathbf{R}^{-1/2}. \tag{43}$$

It is easy to see that $\mathbf{K}^H$ is idempotent, i.e. $\mathbf{K}^H\mathbf{K}^H = \mathbf{K}^H$. But it is not Hermitian, i.e. $\mathbf{K}^H \neq \mathbf{K}$. Thus, $\mathbf{K}^H$ is an oblique projection matrix [21,22]. By a simple algebraic manipulation, we may rewrite $\mathbf{K}^H$ as

$$\begin{aligned}\mathbf{K}^H &= \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}(\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star})^{-1}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \\ &= \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}(\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star})^{-1}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \\ &= \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}(\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},\star}}\mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star})^{-1}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},\star}}.\end{aligned} \tag{44}$$

Substituting $\mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},\star}} = \mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},p}}^{\perp}$ in Eq. (44) yields

$$\mathbf{K}^H = \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}(\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},p}}^{\perp}\mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star})^{-1}\mathbf{U}_{\boldsymbol{\Psi},\star}^H \mathbf{R}\mathbf{P}_{\mathbf{U}_{\boldsymbol{\Psi},p}}^{\perp}. \tag{45}$$

This shows that $\mathbf{K}^H$ is an oblique projection whose range is $\langle \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star} \rangle$ and whose null space is $\langle \mathbf{U}_{\boldsymbol{\Psi},p} \rangle$ [21]. Similarly, we can show that $(\mathbf{I}-\mathbf{K}^H)$ is an oblique projection whose range is $\langle \mathbf{U}_{\boldsymbol{\Psi},p} \rangle$ and whose null space is $\langle \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star} \rangle$:

$$\mathbf{I}-\mathbf{K}^H = \mathbf{R}^{1/2}\mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\boldsymbol{\Psi},p}}\mathbf{R}^{-1/2} = \mathbf{U}_{\boldsymbol{\Psi},p}(\mathbf{U}_{\boldsymbol{\Psi},p}^H \mathbf{P}_{\mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}}^{\perp}\mathbf{U}_{\boldsymbol{\Psi},p})^{-1}\mathbf{U}_{\boldsymbol{\Psi},p}^H \mathbf{P}_{\mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star}}^{\perp}. \tag{46}$$

Using (42), we may write the filter $\mathbf{W}_0^H$ in terms of the oblique projection $\mathbf{K}^H$ as

$$\mathbf{W}_0^H = \mathbf{L}^H\mathbf{V}^H\boldsymbol{\Sigma}_{\boldsymbol{\Psi},p}^{-1}(\overline{\mathbf{W}}^H + \underline{\mathbf{W}}^H) = \mathbf{L}^H\mathbf{V}^H\boldsymbol{\Sigma}_{\boldsymbol{\Psi},p}^{-1}\mathbf{U}_{\boldsymbol{\Psi},p}^H(\mathbf{I}-\mathbf{K}^H). \tag{47}$$

Correspondingly, we may redraw the GSC diagram as in Fig. 3. In this new diagram, the error vector $\mathbf{e}$ is computed by subtracting from the measurement vector $\mathbf{x}$ its oblique projection by $\mathbf{K}^H$, and transforming the result by $\overline{\mathbf{W}}^H = \mathbf{U}_{\boldsymbol{\Psi},p}^H$. Alternatively, the error vector $\mathbf{e}$ may be computed by obliquely projecting the measurement vector $\mathbf{x}$ by $(\mathbf{I}-\mathbf{K}^H)$, and then transforming the result by $\overline{\mathbf{W}}^H = \mathbf{U}_{\boldsymbol{\Psi},p}^H$. As illustrated in Fig. 4, $\mathbf{K}^H\mathbf{x}$ is the oblique projection of $\mathbf{x}$ onto $\langle \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star} \rangle$ along $\langle \mathbf{U}_{\boldsymbol{\Psi},p} \rangle$, and $(\mathbf{I}-\mathbf{K}^H)\mathbf{x}$ is the oblique projection of $\mathbf{x}$ onto $\langle \mathbf{U}_{\boldsymbol{\Psi},p} \rangle$ along $\langle \mathbf{R}\mathbf{U}_{\boldsymbol{\Psi},\star} \rangle$.

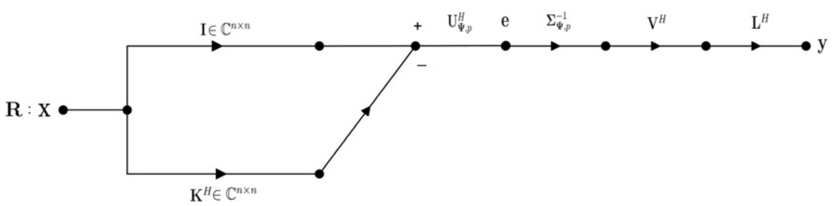

Fig. 3. Oblique projection implementation of the GSC diagram.

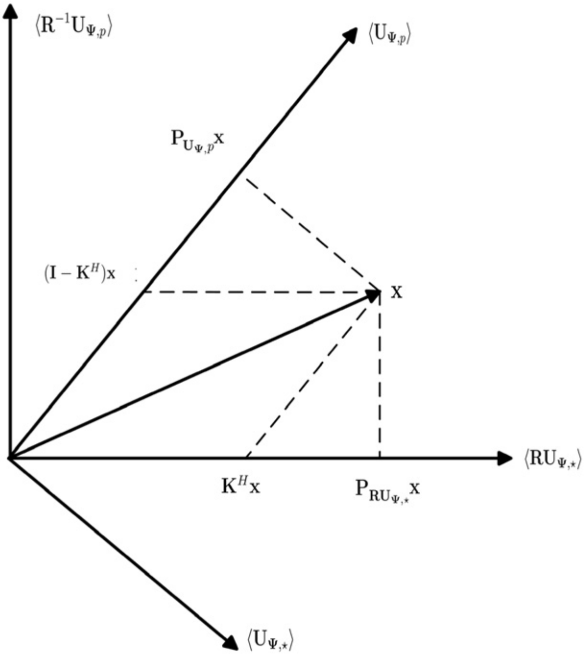

Fig. 4. Oblique projections of $\mathbf{x}$ by $\mathbf{K}^H$ and $(\mathbf{I}-\mathbf{K}^H)$.

It was established in Theorem 6 that the linearly constrained problem (3) is equivalent to the quadratically constrained problem (5) 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. 2 will be a GSC diagram for (23), 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 $\mathbf{U}_{\mathbf{\Psi},\star} = \mathbf{U}_{\mathbf{S},\star}$. Obviously, the oblique projection formulation for (23) is obtained from (47) after proper replacements. Theorem 6 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 (5), 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 $\mathbf{Q}$.

Combining Eqs. (37) and (39), and noting that $\mathbf{F}$ is of form (33), we may write

$$\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p} = \mathbf{F}^H \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star} \mathbf{F} + (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1}. \tag{48}$$

The term on the LHS of Eq. (48) is the covariance matrix of $\mathbf{u}$. The first term on the RHS of Eq. (48) is the covariance matrix of $\hat{\mathbf{u}} = \mathbf{F}^H \mathbf{v}$, the estimate of $\mathbf{u} = \hat{\mathbf{u}} + \mathbf{e}$ from $\mathbf{v}$, and the second term is the covariance matrix of the error vector $\mathbf{e}$. Therefore, (48) is a decomposition of $\mathbf{R}_{uu} = \mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p}$ into the sum of $\mathbf{R}_{\hat{u}\hat{u}} = \mathbf{F}^H \mathbf{U}_{\mathbf{\Psi},\star}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},\star} \mathbf{F}$ and $\mathbf{R}_{ee} = (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R}^{-1} \mathbf{U}_{\mathbf{\Psi},p})^{-1}$. This decomposition directly follows from the fact that in the LMMSE estimation in Fig. 2 the error vector $\mathbf{e} = \mathbf{u} - \hat{\mathbf{u}}$ is orthogonal to the vector $\mathbf{v}$ and the estimate $\hat{\mathbf{u}} = \mathbf{F}^H \mathbf{v}$. This orthogonality property is illustrated in Fig. 5(a).

In beamforming, the matrix $\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p}$ may be called the *Bartlett matrix*, as $\mathrm{tr}\{\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{R} \mathbf{U}_{\mathbf{\Psi},p}\} = E[(\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{x})^H (\mathbf{U}_{\mathbf{\Psi},p}^H \mathbf{x})]$ is the power at the output of the *multi-rank Bartlett*

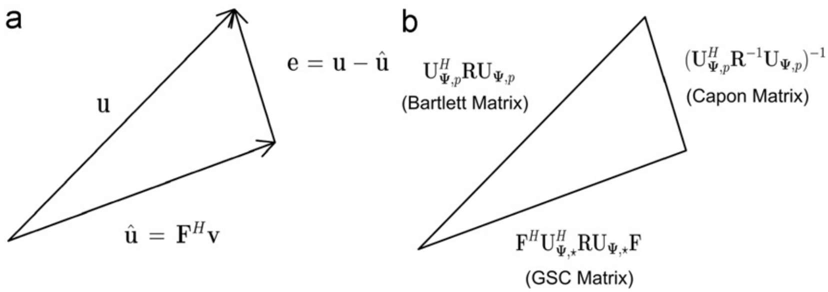

Fig. 5. (a) Orthogonal decomposition of $\mathbf{u}$ into $\hat{\mathbf{u}}$ and $\mathbf{e}$ and (b) Pythagorean decomposition of the Bartlett matrix into the Capon matrix and the GSC matrix.

beamformer $\mathbf{U}^H_{\mathbf{\Psi},p}$ [9,18,27]. Similarly, $\mathbf{F}^H\mathbf{U}^H_{\mathbf{\Psi},\star}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},\star}\mathbf{F}$ and $(\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}$ may be called, respectively, the *GSC matrix* and the *Capon* (*or MVDR*) *matrix*. In light of the orthogonal decomposition of $\mathbf{u}$ in Fig. 5(a), and the fact that the Bartlett, Capon, and GSC matrices are all PD, we may view (48) as a "Pythagorean decomposition" of the Bartlett matrix into the Capon matrix and the GSC matrix, as illustrated in Fig. 5(b).

From Eq. (48), it is obvious that the difference between the Bartlett matrix and the Capon matrix is PD, and therefore we have the matrix inequality

$$(\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}) \succ (\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}. \tag{49}$$

Naturally, Eq. (49) implies the trace inequality

$$\mathrm{tr}\{\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}\} > \mathrm{tr}\{(\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}\}, \tag{50}$$

which shows that the output power of a multi-rank MVDR beamformer[4] is always less than that of a multi-rank Bartlett beamformer. A more general trace inequality that follows from Eq. (49) is

$$\mathrm{tr}\{\mathbf{L}^H\mathbf{V}^H\mathbf{\Sigma}^{-1}_{\mathbf{\Psi},p}\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}\mathbf{U}_{\mathbf{\Psi},p}\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L}\} > \mathrm{tr}\{\mathbf{L}^H\mathbf{V}^H\mathbf{\Sigma}^{-1}_{\mathbf{\Psi},p}(\mathbf{U}^H_{\mathbf{\Psi},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{\Psi},p})^{-1}\mathbf{\Sigma}^{-H}_{\mathbf{\Psi},p}\mathbf{V}\mathbf{L}\}, \tag{51}$$

which shows that in the GSC diagram of Fig. 2 the power of $\mathbf{y}$ is smaller with the lower branch switched in than switched out.

In deriving Eq. (48), we need $\mathbf{R}$ to be PD, and $\mathbf{U}_{\mathbf{\Psi}} = [\mathbf{U}_{\mathbf{\Psi},p}\mathbf{U}_{\mathbf{\Psi},\star}]$ to be an orthogonal matrix. However, no particular relation between $\mathbf{R}$ and $\mathbf{U}_{\mathbf{\Psi}}$ is required. Therefore, in general, the matrix $\mathbf{U}_{\mathbf{\Psi},p}$ can be any $n \times p$ complex left-orthogonal matrix. In other words, considering an arbitrary rank $-p<n$ matrix $\mathbf{M} \in \mathbb{C}^{n\times p}$ with the SVD $\mathbf{M} = \mathbf{U}_{\mathbf{M},p}\mathbf{\Sigma}_{\mathbf{M},p}\mathbf{V}^H_{\mathbf{M}}$, in the GSC diagram of Fig. 2, we can replace $\mathbf{U}_{\mathbf{\Psi},p}$ and $\mathbf{U}_{\mathbf{\Psi},\star}$ by $\mathbf{U}_{\mathbf{M},p}$ and $\mathbf{U}_{\mathbf{M},\star}$, solve for the LMMSE estimator, and then obtain the decomposition

$$\mathbf{U}^H_{\mathbf{M},p}\mathbf{R}\mathbf{U}_{\mathbf{M},p} = \mathbf{F}^H\mathbf{U}^H_{\mathbf{M},\star}\mathbf{R}\mathbf{U}_{\mathbf{M},\star}\mathbf{F} + (\mathbf{U}^H_{\mathbf{M},p}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{M},p})^{-1}, \tag{52}$$

with $\mathbf{F} = (\mathbf{U}^H_{\mathbf{M},\star}\mathbf{R}\mathbf{U}_{\mathbf{M},\star})^{-1}\mathbf{U}^H_{\mathbf{M},\star}\mathbf{R}\mathbf{U}_{\mathbf{M},p}$. Pre-multiplying Eq. (52) by $\mathbf{V}_{\mathbf{M}}\mathbf{\Sigma}^H_{\mathbf{M},p}$ and post-multiplying it by $\mathbf{\Sigma}_{\mathbf{M},p}\mathbf{V}^H_{\mathbf{M}}$, and noting that $\mathbf{\Sigma}_{\mathbf{M},p}$ is a full-rank $p \times p$ diagonal matrix and

[4]Note that in the GSC diagram in Fig. 2, the error vector $\mathbf{e}$ may be viewed as the output of a multi-rank MVDR beamformer that is associated with a linearly constrained minimization of form (3), with $\mathbf{\Psi} = \mathbf{U}_{\mathbf{\Psi},p}$ and $\mathbf{C}^H = \mathbf{L}^H\mathbf{V}^H = \mathbf{I}$.

$\mathbf{V_M} \in \mathbb{C}^{p\times p}$ is an orthogonal matrix, yields

$$\mathbf{M}^H\mathbf{R}\mathbf{M} = \mathbf{V_M}\mathbf{\Sigma}^H_{\mathbf{M},p}\mathbf{F}^H\mathbf{U}^H_{\mathbf{M},\star}\mathbf{R}\mathbf{U}_{\mathbf{M},\star}\mathbf{F}\mathbf{\Sigma}_{\mathbf{M},p}\mathbf{V}^H_{\mathbf{M}} + (\mathbf{N}^H\mathbf{R}^{-1}\mathbf{N})^{-1}, \tag{53}$$

where $\mathbf{N} = \mathbf{U}_{\mathbf{M},p}\mathbf{\Sigma}^{-H}_{\mathbf{M},p}\mathbf{V}^H_{\mathbf{M}}$. Observing that $\mathbf{N}^H\mathbf{M} = \mathbf{I}$, and the first term on the RHS of Eq. (53) is PD, we may state the following theorem.

Theorem 7. *Let* $\mathbf{M} \in \mathbb{C}^{n\times p}$ *and* $\mathbf{N} \in \mathbb{C}^{n\times p}$ $(p<n)$ *be full-rank matrices such that* $\mathbf{N}^H\mathbf{M} = \mathbf{I}$, *i.e.* $\mathbf{N}^H$ *is the left-inverse of* $\mathbf{M}$, *and* $\mathbf{R} \in \mathbb{C}^{n\times n}$ *be a PD matrix. Then,*

$$\mathbf{M}^H\mathbf{R}\mathbf{M} \succ (\mathbf{N}^H\mathbf{R}^{-1}\mathbf{N})^{-1}. \tag{54}$$

A direct consequence of this theorem is the trace inequality $\mathrm{tr}\{\mathbf{M}^H\mathbf{R}\mathbf{M}\} > \mathrm{tr}\{(\mathbf{N}^H\mathbf{R}^{-1}\mathbf{N})^{-1}\}$ *in* [24, Chapter 11].

Remark. The trace inequality in Eq. (21) follows from Theorem 7, by replacing $\mathbf{R}$ and $\mathbf{M} = \mathbf{N}$ with $(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}$ and $\mathbf{V}$, and noting that the traces on the LHS and RHS of Eq. (21) are, respectively, sums of the quadratic forms of $(\mathbf{V}^H\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi}\mathbf{V})^{-1}$ and $\mathbf{V}^H(\mathbf{\Psi}^H\mathbf{R}^{-1}\mathbf{\Psi})^{-1}\mathbf{V}$ that are constructed by pre-multiplying these matrices by the rows of $\mathbf{L}^H$ and post-multiplying them by the columns of $\mathbf{L}$.

5. Conclusions

The constrained minimization problems considered here arise in the design of multi-rank beamformers for radar, sonar, 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 or quadratic constraints. Solutions to these problems were derived and connections between the linearly and quadratically constrained minimizations were established. The majorization result of Theorem 1 and Poincare's separation theorem played key roles in establishing the connection. We presented GSC diagrams for our solutions, clarifying in the process the connections between the linearly and quadratically constrained minimizations and LMMSE estimations. We then showed that our solutions can be cast in terms of oblique projections and established the geometry of our constrained minimizations.

Appendix A

A.1. Proof of Theorem 3

The subspace identity in Eq. (11) follows by observing that $\langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}\rangle$ and $\langle \mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star}\rangle$ are p- and $(n–p)$-dimensional subspaces of $\mathbb{C}^n$, and that Euclidean inner products between $\mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}$ and $\mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star}$ are zero:

$$\begin{aligned}(\mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p})^H\mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star} &= \mathbf{U}^H_{\mathbf{A},p}\mathbf{U}_{\mathbf{A},\star} = \mathbf{0},\\ (\mathbf{B}^{H/2}\mathbf{U}_{\mathbf{A},\star})^H\mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p} &= \mathbf{U}^H_{\mathbf{A},\star}\mathbf{U}_{\mathbf{A},p} = \mathbf{0}.\end{aligned} \tag{55}$$

The projection identity in Eq. (12) is a direct consequence of Eq. (11). In the special case where $p = m<n$, in the SVD of $\mathbf{A} = \mathbf{U}_{\mathbf{A},p}\mathbf{\Sigma}_{\mathbf{A},p}\mathbf{V}^H_{\mathbf{A},p}$, $\mathbf{V}_{\mathbf{A},p} \in \mathbb{C}^{p\times p}$ is an orthogonal matrix and

hence $\mathbf{\Sigma}_{\mathbf{A},p}\mathbf{V}_{\mathbf{A},p}^H \in \mathbb{C}^{p\times p}$ is nonsingular. Consequently, $\langle \mathbf{B}^{-1/2}\mathbf{A}\rangle = \langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}\mathbf{\Sigma}_{\mathbf{A},p}\mathbf{V}_{\mathbf{A},p}^H\rangle = \langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}\rangle$, which yields the additional identities in Eqs. (13) and (14). When $p<m$, however, $\langle \mathbf{B}^{-1/2}\mathbf{A}\rangle$ is a subspace of $\langle \mathbf{B}^{-1/2}\mathbf{U}_{\mathbf{A},p}\rangle$.

A.2. Proof of Theorem 5

Consider the EVD of $\mathbf{S} = \mathbf{U_S\Lambda_S^2U_S^H}$. Since $\mathbf{U_S} = [\mathbf{U}_{\mathbf{S},p}\ \mathbf{U}_{\mathbf{S},\star}] \in \mathbb{C}^{n\times n}$ is an orthogonal basis for $\mathbb{C}^n$, without loss of generality, we may express $\mathbf{W} \in \mathbb{C}^{n\times r}$ as

$$\mathbf{W} = [\mathbf{U}_{\mathbf{S},p}\ \mathbf{U}_{\mathbf{S},\star}]\begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{U}_{\mathbf{S},\star}\mathbf{B}, \tag{56}$$

where $\mathbf{A} \in \mathbb{C}^{p\times r}$ and $\mathbf{B} \in \mathbb{C}^{(n-p)\times r}$.

We start by satisfying the quadratic constraint $\mathbf{W}^H\mathbf{SW} = \mathbf{D}$. Substituting for $\mathbf{W}$ from Eq. (56), we may simplify the constraint to

$$\mathbf{W}^H\mathbf{SW} = [\mathbf{A}^H\ \ \mathbf{B}^H]\begin{bmatrix}\mathbf{\Lambda}_{\mathbf{S},p}^2 & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\mathbf{A}\\ \mathbf{B}\end{bmatrix} = \mathbf{D}, \tag{57}$$

or equivalently

$$\mathbf{A}^H\mathbf{\Lambda}_{\mathbf{S},p}^2\mathbf{A} = \mathbf{D}. \tag{58}$$

Therefore, $\mathbf{A}$ must be of the form

$$\mathbf{A} = \mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{E}\mathbf{D}^{H/2} = \mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^H\mathbf{U}_{\mathbf{D}}^H, \tag{59}$$

where $\mathbf{E} \in \mathbb{C}^{p\times r}$ $(p\geq r)$ is a left-orthogonal matrix, yet to be determined.

Inserting $\mathbf{W} = \mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{U}_{\mathbf{S},\star}\mathbf{B}$ in $J(\mathbf{W}) = \mathrm{tr}\{\mathbf{W}^H\mathbf{RW}\}$ yields

$$\begin{aligned} J(\mathbf{W}) &= \mathrm{tr}\{(\mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{U}_{\mathbf{S},\star}\mathbf{B})^H\mathbf{R}(\mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{U}_{\mathbf{S},\star}\mathbf{B})\}\\ &= \mathrm{tr}\{(\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}\mathbf{B})^H(\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A} + \mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}\mathbf{B})\}. \end{aligned} \tag{60}$$

Assuming $\mathbf{A}$ (or equivalently $\mathbf{E}$) is fixed, we minimize J with respect to $\mathbf{B}$. Since columns of $\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}$ and $\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}$ do not lie in the same subspace this minimization results in a least-squares solution for $\mathbf{B}$ of the form

$$\mathbf{B} = -(\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star})^\dagger\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A}. \tag{61}$$

Plugging $\mathbf{B}$ in Eq. (60) yields

$$\begin{aligned} J &= \mathrm{tr}\{\mathbf{A}^H\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{1/2}[\mathbf{I}-\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}(\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star})^\dagger]\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A}\}\\ &= \mathrm{tr}\{\mathbf{A}^H\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{1/2}[\mathbf{I}-\mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}}]\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A}\}\\ &= \mathrm{tr}\{\mathbf{A}^H\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{1/2}\mathbf{P}^\perp_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}}\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{A}\}. \end{aligned} \tag{62}$$

Using Theorem 3, we may write

$$\mathbf{P}^\perp_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}} = \mathbf{P}_{\mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{S},p}} = \mathbf{R}^{-1/2}\mathbf{U}_{\mathbf{S},p}(\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p})^{-1}\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{-H/2}, \tag{63}$$

and then use Eq. (63) to simplify Eq. (62) as

$$J = \mathrm{tr}\{\mathbf{A}^H(\mathbf{U}_{\mathbf{S},p}^H\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p})^{-1}\mathbf{A}\}. \tag{64}$$

Plugging $\mathbf{A}=\mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{H}\mathbf{U}_{\mathbf{D}}^{H}$ in Eq. (64) yields

$$\begin{aligned}J&=\mathrm{tr}\{\mathbf{U}_{\mathbf{D}}\mathbf{\Lambda}_{\mathbf{D}}\mathbf{E}^{H}\mathbf{\Lambda}_{\mathbf{S},p}^{-1}(\mathbf{U}_{\mathbf{S},p}^{H}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p})^{-1}\mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{H}\mathbf{U}_{\mathbf{D}}^{H}\}\\&=\mathrm{tr}\{\mathbf{E}^{H}(\mathbf{\Lambda}_{\mathbf{S},p}^{H}\mathbf{U}_{\mathbf{S},p}^{H}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p})^{-1}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{2}\},\end{aligned}\tag{65}$$

where the second equality follows from the cyclic property of trace and $\mathbf{U}_{\mathbf{D}}^{H}\mathbf{U}_{\mathbf{D}}=\mathbf{I}$. Using Eq. (24), we may rewrite Eq. (65) as

$$J=\mathrm{tr}\{\mathbf{E}^{H}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{2}\}=\mathrm{tr}\{\mathbf{E}^{H}\mathbf{Q}^{-1}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{2}\}.\tag{66}$$

We may now minimize J with respect to the left-orthogonal matrix $\mathbf{E}\in\mathbb{C}^{p\times r}$. Let $0<\lambda_{\tilde{\mathbf{Q}},1}^{2}\leq\cdots\leq\lambda_{\tilde{\mathbf{Q}},r}^{2}$ be the eigenvalues of $\tilde{\mathbf{Q}}=\mathbf{E}^{H}\mathbf{Q}^{-1}\mathbf{E}$. Then from Theorem 1, we have

$$J=\mathrm{tr}\{\tilde{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{D}}^{2}\}\geq\sum_{i=1}^{r}\lambda_{\tilde{\mathbf{Q}},i}^{2}\lambda_{\mathbf{D},i}^{2}.\tag{67}$$

Since $\mathbf{E}$ is left-orthogonal, from Theorem 2, the eigenvalues of $\mathbf{Q}^{-1}$ and $\tilde{\mathbf{Q}}=\mathbf{E}^{H}\mathbf{Q}^{-1}\mathbf{E}$, i.e. $1/\lambda_{\mathbf{Q},i}^{2}$ and $\lambda_{\tilde{\mathbf{Q}},i}^{2}$, satisfy the inequalities

$$\frac{1}{\lambda_{\mathbf{Q},i}^{2}}\leq\lambda_{\tilde{\mathbf{Q}},i}^{2}\leq\frac{1}{\lambda_{\mathbf{Q},p-r+i}^{2}},\quad i=1,\ldots,r.\tag{68}$$

Combining Eqs. (67) and (68), and plugging in $\mathbf{Q}^{-1}=\mathbf{U}_{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{Q}}^{-2}\mathbf{U}_{\mathbf{Q}}^{H}$ yields

$$J=\mathrm{tr}\{\mathbf{E}^{H}\mathbf{U}_{\mathbf{Q}}\mathbf{\Lambda}_{\mathbf{Q}}^{-2}\mathbf{U}_{\mathbf{Q}}^{H}\mathbf{E}\mathbf{\Lambda}_{\mathbf{D}}^{2}\}\geq\sum_{i=1}^{r}\frac{\lambda_{\mathbf{D},i}^{2}}{\lambda_{\mathbf{Q},i}^{2}}.\tag{69}$$

The equality holds when $\mathbf{E}^{H}\mathbf{U}_{\mathbf{Q}}=[\mathbf{I}_{r}\ \mathbf{0}]$ or equivalently $\mathbf{E}=\mathbf{U}_{\mathbf{Q},r}$. Thus, the minimum value of the quadratic form J is

$$J_{0}=\sum_{i=1}^{r}\frac{\lambda_{\mathbf{D},i}^{2}}{\lambda_{\mathbf{Q},i}^{2}}.\tag{70}$$

Correspondingly, the solution $\mathbf{W}=\mathbf{W}_{0}$ is

$$\begin{aligned}\mathbf{W}_{0}&=\mathbf{U}_{\mathbf{S},p}\mathbf{A}+\mathbf{U}_{\mathbf{S},\star}\mathbf{B}\\&=[\mathbf{U}_{\mathbf{S},p}-\mathbf{R}^{-H/2}\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}(\mathbf{U}_{\mathbf{S},\star}^{H}\mathbf{R}\mathbf{U}_{\mathbf{S},\star})^{-1}\mathbf{U}_{\mathbf{S},\star}^{H}\mathbf{R}^{1/2}\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}]\mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}\\&=\mathbf{R}^{-H/2}(\mathbf{I}-\mathbf{P}_{\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},\star}})\mathbf{R}^{H/2}\mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}\\&=\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p}(\mathbf{U}_{\mathbf{S},p}^{H}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p})^{-1}\mathbf{\Lambda}_{\mathbf{S},p}^{-H}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2},\end{aligned}\tag{71}$$

where the last equality follows from Eq. (63). Finally, a simple algebraic manipulation on Eq. (71) yields

$$\begin{aligned}\mathbf{W}_{0}&=\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p}(\mathbf{\Lambda}_{\mathbf{S},p}^{H}\mathbf{U}_{\mathbf{S},p}^{H}\mathbf{R}^{-1}\mathbf{U}_{\mathbf{S},p}\mathbf{\Lambda}_{\mathbf{S},p})^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}\\&=\mathbf{R}^{-1}\mathbf{S}^{1/2}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}.\end{aligned}\tag{72}$$

A.3. Proof of Theorem 6

Inserting $\mathbf{L}^H = \mathbf{D}^{1/2} = \mathbf{U_D \Lambda_D}$ and $\mathbf{\Psi} = \mathbf{S}^{1/2}$ in Eq. (15) yields

$$J_0 = \mathrm{tr}\{\mathbf{U_D \Lambda_D V}^H (\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{V \Lambda_D^H U_D^H}\} = \mathrm{tr}\{\mathbf{V}^H \mathbf{Q}^{-1}\mathbf{V \Lambda_D^2}\}. \tag{73}$$

Noting that $\mathbf{V}$ is left-orthogonal, using Theorems 1 and 2, and an argument similar to that made in deriving Eqs. (67) to (69), we can derive the inequality

$$J_0 = \mathrm{tr}\{\mathbf{V}^H \mathbf{Q}^{-1}\mathbf{V \Lambda_D^2}\} \geq \sum_{i=1}^{r} \frac{\lambda_{\mathbf{D},i}^2}{\lambda_{\mathbf{Q},i}^2}. \tag{74}$$

The RHS of this inequality is equal to the minimum value of the quadratic form J in Eq. (22). The equality in Eq. (74) holds when $\mathbf{V}^H \mathbf{U}_{\mathbf{Q},r} = [\mathbf{I}_r \ \mathbf{0}]$ or equivalently $\mathbf{V} = \mathbf{U}_{\mathbf{Q},r}$. Finally, inserting $\mathbf{\Psi} = \mathbf{S}^{1/2}$, $\mathbf{L} = \mathbf{D}^{H/2}$, and $\mathbf{V} = \mathbf{U}_{\mathbf{Q},r}$ in Eq. (16) yields

$$\mathbf{W}_0 = \mathbf{R}^{-1}\mathbf{S}^{1/2}(\mathbf{S}^{H/2}\mathbf{R}^{-1}\mathbf{S}^{1/2})^{-1}\mathbf{U}_{\mathbf{Q},r}\mathbf{D}^{H/2}, \tag{75}$$

which is equal to $\mathbf{W}_0$ in Eq. (23).

References

[1] J. Capon, R.J. Greenfield, R.J. Kolker, Multidimensional maximum-likelihood processing for a large aperture seismic array, Proc. IEEE 55 (1967) 192–211.
[2] J. Capon, High-resolution frequency-wavenumber spectrum analysis, Proc. IEEE 57 (1969) 1408–1418.
[3] H.L. Van Trees, Optimum Array Processing, Wiley Interscience, New York, NY, 2002.
[4] H. Cox, Sensitivity considerations in adaptive beamforming, in: J.W.R. Griffiths, P.L. Stocklin, C. Van Schooneveld (Eds.), Signal Processing (Proceedings of the NATO Advanced Study Institute Signal Processing with Particular Reference to Underwater Acoustics, Loughborough, UK, August 1972), Academic, New York, London, 1973, pp. 619–645.
[5] B.D. Van Veen, W. van Drongelen, M. Yuchtman, A. Suzuki, Localization of brain electrical activity via linearly constrained minimum variance spatial filtering, IEEE Trans. Biomed. Eng. 44 (1997) 867–880.
[6] H. Cox, A. Pezeshki, L.L. Scharf, M. Lundberg, H. Lai, Multi-rank adaptive beamforming with linear and quadratic constraints, in: Conference Record, 39th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, 2005, pp. 1692–1697.
[7] L.L. Scharf, A. Pezeshki, M. Lundberg, Multi-rank adaptive beamforming, in: Proceedings of the 15th IEEE Signal Processing Workshop, Bordeaux, France, 2005.
[8] L.L. Scharf, A. Pezeshki, B.D. Van Veen, H. Cox, O. Besson, Eigenvalue beamforming using a multi-rank MVDR beamformer, in: Proceedings of the Fifth Workshop on Defence Applications of Signal Processing, Queensland, Australia, 2006.
[9] A. Pezeshki, B.D. Van Veen, L.L. Scharf, H. Cox, M. Lundberg, Eigenvalue beamforming using a multi-rank MVDR beamformer and subspace selection, IEEE Trans. Signal Process. 56 (2008) 1954–1967.
[10] S. Shahbazpanahi, A.B. Gershman, Z. Luo, K.M. Wong, Robust adaptive beamforming for general-rank signal models, IEEE Trans. Signal Process. 51 (2003) 2257–2269.
[11] M.K. Tsatsanis, Z. Xu, Performance analysis of minimum variance CDMA receivers, IEEE Trans. Signal Process. 46 (1998) 3014–3022.
[12] A. Scaglione, P. Stoica, S. Barbarossa, G.B. Giannakis, H. Sampath, Optimal designs for space–time linear precoders and decoders, IEEE Trans. Signal Process. 50 (2002) 1051–1064.
[13] Y. Ding, T.N. Davidson, Z. Luo, K.M. Wong, Minimum BER block precoders for zero-forcing equalization, IEEE Trans. Signal Process. 51 (2003) 2410–2423.
[14] D.P. Palomar, J.M. Cioffi, M.A. Lagunas, Joint Tx–Rx beamforming design for multicarrier MIMO channels: a unified framework for convex optimization, IEEE Trans. Signal Process. 51 (2003) 2381–2401.
[15] S. Shahbazpanahi, M. Beheshti, A.B. Gershman, M. Gharavi-Alkhansari, K.M. Wong, Minimum variance linear receivers for multiaccess MIMO wireless systems with space-time block coding, IEEE Trans. Signal Process. 52 (2004) 3306–3313.

[16] D.P. Palomar, Y. Jiang, MIMO Transceiver Design via Majorization Theory, Foundations and Trends in Communications and Information Theory, vol. 3, Now Publishers, 2006, pp. 331–551.
[17] D. Slepian, Prolate spheroidal wave functions, Fourier analysis and uncertainty—V: the discrete case, Bell Syst. Technol. J. (1978) 1371–1430.
[18] C.T. Mullis, L.L. Scharf, Quadratic estimators of the power spectrum, in: S. Haykin (Ed.), Advances in Spectrum Estimation, vol. 1, Prentice Hall, Englewood Cliffs, NJ, 1990, pp. 1–57 (Chapter 1).
[19] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York, 1979.
[20] L.J. Griffiths, C.W. Jim, An alternative approach to linearly constrained adaptive beamforming, IEEE Trans. Antennas Propag. AP 30 (1982) 27–34.
[21] R.T. Behrens, L.L. Scharf, Signal processing applications of oblique projection operators, IEEE Trans. Signal Process. 42 (1994) 1413–1424.
[22] S. Kayalar, H.L. Weinert, Oblique projections: formulas, algorithms and error bounds, Math. Control Signals Systems 2 (1989) 33–45.
[23] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., Johns Hopkins University Press, Baltimore, MD, 1996.
[24] J.R. Magnus, H. Neudecker, Matrix Differential Calculus with Applications in Statistics and Econometrics, Wiley, New York, NY, 1988.
[25] E.K.P. Chong, S.H. Żak, An Introduction to Optimization, third ed., Wiley, New York, NY, 2008.
[26] L.L. Scharf, C.T. Mullis, Canonical coordinates and the geometry of inference, rate and capacity, IEEE Trans. Signal Process. 48 (2000) 824–831.
[27] D.J. Thomson, Spectrum estimation and harmonic analysis, Proc. IEEE 70 (1982) 1055–1096.