Wang et al. EURASIP Journal on Advances in Signal Processing 2011, **2011**:15
http://asp.eurasipjournals.com/content/2011/1/15

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

RESEARCH Open Access

Joint optimization of MIMO radar waveform and biased estimator with prior information in the presence of clutter

Hongyan Wang[1*], Guisheng Liao[1], Hongwei Liu[2], Jun Li[1] and Hui Lv[1]

Abstract

In this article, we consider the problem of joint optimization of multi-input multi-output (MIMO) radar waveform and biased estimator with prior information on targets of interest in the presence of signal-dependent noise. A novel *constrained biased* Cramer-Rao bound (CRB) based method is proposed to optimize the waveform covariance matrix (WCM) and biased estimator such that the performance of parameter estimation can be improved. Under a simplifying assumption, the resultant nonlinear optimization problem is solved resorting to a convex relaxation that belongs to the semidefinite programming (SDP) class. An optimal solution of the initial problem is then constructed through a suitable approximation to an optimal solution of the relaxed one (in a least squares (LS) sense). Numerical results show that the performance of parameter estimation can be improved considerably by the proposed method compared to uncorrelated waveforms.

Keywords: Multi-input multi-output (MIMO) radar, waveform optimization, clutter, *constrained biased* Cramer-Rao bound (CRB), Semidefinite programming (SDP)

1 Introduction

Multi-input multi-output (MIMO) radar has attracted more and more attention recently [1-19]. Unlike the traditional phased-array radar which can only transmit scaled versions of a single waveform, MIMO radar can use multiple transmitting elements to transmit arbitrary waveforms. Two categories of MIMO radar systems can be classified by the configuration of the transmitting and receiving antennas: (1) MIMO radar with widely separated antennas (see, e.g., [1,2]), and (2) MIMO radar with colocated antennas (see, e.g., [3]). For MIMO radar with widely separated antennas, the transmitting and receiving elements are widely spaced such that each views a different aspect of the target. This type of MIMO radar can exploit the spatial diversity to overcome performance degradations caused by target scintillations [2]. In contrast, MIMO radar with colocated antennas, the elements of which in transmitting and receiving arrays are close enough such that the target radar cross sections (RCS) observed by MIMO radar are identical, can be used to increase the spatial resolution. Accordingly, it has several advantages over its phased array counterpart, including improved parameter identifiability [4,5], and more flexibility for transmit beampattern design [6-19]. In this article, we focus on MIMO radar with colocated antennas.

One of the most interesting research topics on both types of MIMO radar is the waveform optimization, which has been studied in [6-19]. According to the target model used in the problem of waveform design, the current design methods can be divided into two categories: (1) point target-based design [6-12], and (2) extended target-based design [13-19]. In the case of point targets, the corresponding methods optimize the waveform covariance matrix (WCM) [6-8] or the radar ambiguity function [9-12]. The methods of optimizing the WCM only consider the spatial domain characteristics of the transmitted signals, while the one of optimizing the radar ambiguity function treat the spatial, range, and Doppler domain characteristics jointly. In the case of extended targets, some prior information on the target and noise are used to design the transmitted waveforms.

* Correspondence: gglongs@163.com
[1]National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China
Full list of author information is available at the end of the article

In [7], based on the Cramer-Rao bound (CRB), the problem of MIMO radar waveform design for parameter estimation of point targets has been investigated under the assumption that the received signals do not include the clutter which depends on the transmitted waveforms. However, it is known that the received data is generally contaminated by the clutter in many applications (see, e.g., [13,14]). It is noted that the CRB provides a lower bound on the variance when any *unbiased estimator* is used without employing any prior information. In fact, some prior information may be available in many array signal processing fields (see, e.g., [20-22]), which can be regarded as a constraint on the estimated parameter space. A variant of the CRB for this kind of the constrained estimation problem was developed in [20,22], which is called the *constrained* CRB. Moreover, a *biased estimator* can lower the resulting variance obtained by any *unbiased estimator* generally [23-28]. The variant of the CRB for this case is named as the *biased* CRB. Furthermore, the variance produced by any *unbiased estimator* can be lowered obviously while both *biased estimator* and prior information are used. A variant of the CRB for this case was studied in [29], which can be referred to as the *constrained biased* CRB. Consequently, from the parameter estimation point of view, it is worth studying the waveform optimization problem in the presence of clutter by employing both the *biased estimator* and prior information.

In this article, we consider the problem of joint optimization of the WCM and *biased estimator* with prior information on targets of interest in the presence of clutter. Under the weighted or spectral norm constraint on the bias gradient matrix of the biased estimator, a novel *constrained biased* CRB-based method is proposed to optimize the WCM and biased estimator such that the performance of parameter estimation can be improved. The joint WCM and biased estimator design is formulated in terms of a rather complicated nonlinear optimization problem, which cannot be easily solved by convex optimization methods [30-32]. Under a simplifying assumption, this problem is solved resorting to a convex relaxation that belongs to the semidefinite programming (SDP) class [31]. An optimal solution of the initial joint optimization problem is then constructed through a suitable approximation to an optimal solution of the relaxed one (in a least squares (LS) sense).

The rest of this article is organized as follows. In Section 2, we present MIMO radar model, and formulate the joint optimization of the WCM and *biased estimator*. In Section 3, under the weighted or spectral norm constraint on the gradient matrix, we solve the joint optimization problem resorting to the SDP relaxation, and provide a solution to the problem. In Section IV, we assess the effectiveness of the proposed method via some numerical examples. Finally, in Section V, we draw conclusions and outline possible for future research tracks.

Throughout the article, matrices and vectors are denoted by boldface uppercase and lowercase letters, respectively. We use $\{\cdot\}^T$, $\{\cdot\}^*$, and $\{\cdot\}^H$ to denote the transpose, conjugate, and conjugate transpose, respectively. vec{·} is the vectorization operator stacking the columns of a matrix on top of each other, $\mathbf{I}$ denotes the identity matrix, and $\otimes$ indicates the Kronecker product. The trace, real, and imaginary parts of a matrix are denoted by tr{·}, Re{·}, and Im{·}, respectively. The symbol $\{\cdot\}^\dagger$ denotes Moore-Penrose inverse of a matrix, and $\{\cdot\}^+$ indicates the positive part of a real number. The notation $E\{\cdot\}$ stands for the expectation operator, diag{a} for a diagonal matrix with its diagonal given by the vector $\mathbf{a}$, and $\|\mathbf{A}\|_\mathrm{F}$ for the Frobenius norm of the matrix $\mathbf{A}$. Given a vector function $\mathbf{f}:\mathbb{R}^n \to \mathbb{R}^k$, we denote by $\frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}$ the k × n matrix the *ij*th element of which is $\frac{\partial \mathbf{f}_i}{\partial \boldsymbol{\theta}_j}$. $\Re(\mathbf{A})$ is the range space of a matrix $\mathbf{A}$. Finally, the notation $\mathbf{A} \preccurlyeq \mathbf{B}$ means that $\mathbf{B}$-$\mathbf{A}$ is positive semidefinite.

2 System model and problem formulation

Consider a MIMO radar system with M_t transmitting elements and M_r receiving elements. Let $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_{M_\mathrm{t}}]^T \in \mathbb{C}^{M_\mathrm{t} \times L}$ be the transmitted waveform matrix, where $\mathbf{s}_i \in \mathbb{C}^{L\times 1}, i = 1,2,\ldots,M_\mathrm{t}$ denotes the discrete-time baseband signal of the *i*th transmit element with L being the number of snapshots. Under the assumption that the transmitted signals are narrowband and the propagation is non-dispersive, the received signals by MIMO radar can be expressed as

$$\mathbf{Y} = \sum_{k=1}^{K} \beta_k \mathbf{a}(\theta_k)\mathbf{v}^T(\theta_k)\mathbf{S} + \sum_{i=1}^{N_\mathrm{C}} \rho(\theta_i)\mathbf{a}_\mathrm{c}(\theta_i)\mathbf{v}_\mathrm{c}^T(\theta_i)\mathbf{S} + \mathbf{W}, \quad (1)$$

where the columns of $\mathbf{Y} \in \mathbb{C}^{M_\mathrm{r} \times L}$ are the collected data snapshots, $\{\beta_k\}_{k=1}^K$ are the complex amplitudes proportional to the RCSs of the targets with K being the number of targets at the considered range bin, and $\{\theta_k\}_{k=1}^K$ denote the locations of these targets. The parameters $\{\beta_k\}_{k=1}^K$ and $\{\theta_k\}_{k=1}^K$ need to be estimated from the received signal $\mathbf{Y}$. The second term in the right hand of (1) indicates the clutter data collected by the receiver, $\rho(\theta_i)$ is the reflect coefficient of the clutter patch at θ_i, and N_C ($N_\mathrm{C} \gg M_\mathrm{t}M_\mathrm{r}$ the number of spatial samples of the clutter. The term $\mathbf{W}$ denotes the interference plus noise, which is independent of the clutter. Similar to [7], the columns of $\mathbf{W}$ can be assumed to be independent and identically distributed circularly symmetric complex Gaussian random vectors with mean zero and an unknown covariance $\mathbf{B}$. $\mathbf{a}(\theta_k)$ and $\mathbf{v}(\theta_k)$ denote, respectively, the receiving and transmitting steering vectors for the target located at θ_k, which can be expressed as

$$\begin{aligned}\mathbf{a}(\theta_k) &= \left[e^{j2\pi f_0 \tau_1(\theta_k)}, e^{j2\pi f_0 \tau_2(\theta_k)}, \ldots, e^{j2\pi f_0 \tau_{M_r}(\theta_k)}\right]^T \\ \mathbf{v}(\theta_k) &= \left[e^{j2\pi f_0 \tilde{\tau}_1(\theta_k)}, e^{j2\pi f_0 \tilde{\tau}_2(\theta_k)}, \ldots, e^{j2\pi f_0 \tilde{\tau}_{M_t}(\theta_k)}\right]^T,\end{aligned} \tag{2}$$

where f_0 represents the carrier frequency, $\tau_m(\theta_k)$, $m = 1,2,...M_r$ is the propagation time from the target located at θ_k to the mth receiving element, and $\tilde{\tau}_n(\theta_k)$, $n = 1, 2, \ldots M_t$ is the propagation time from the nth transmitting element to the target. Also, $\mathbf{a}_c(\theta_i)$ and $\mathbf{v}_c(\theta_i)$ denote the receiving and transmitting steering vectors for the clutter patch at θ_i, respectively.

For notational simplicity, (1) can be rewritten as

$$\mathbf{Y} = \sum_{k=1}^{K} \beta_k \mathbf{a}(\theta_k)\mathbf{v}^T(\theta_k)\mathbf{S} + \mathbf{H}_c\mathbf{S} + \mathbf{W}, \tag{3}$$

where $\mathbf{H}_c = \sum_{i=1}^{N} \rho(\theta_i)\mathbf{a}_c(\theta_i)\mathbf{v}_c^T(\theta_i)$, which represents the clutter transfer function similar to the channel matrix in [2]. According to Chen and Vaidyanathan and Wang and Lu [33,34], vec($\mathbf{H}_c$) can be considered as an identically distributed complex Gaussian random vector with mean zero and covariance

$$\mathbf{R}_{\mathbf{H}_c} = E\left[\text{vec}(\mathbf{H}_c)\text{vec}^H(\mathbf{H}_c)\right]. \tag{4}$$

In fact, $\mathbf{R}_{\mathbf{H}_c}$ can be explicitly expressed as (see, e.g., [35]):

$$\mathbf{R}_{\mathbf{H}_c} = \mathbf{V}\Xi\mathbf{V}^H, \tag{5}$$

where
$\mathbf{V} = \left[\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{N_C}\right]$, $\mathbf{v}_i = \mathbf{v}_c(\theta_i) \otimes \mathbf{a}_c(\theta_i)$, $i = 1, 2, \ldots, N_C$, $\Xi = \text{diag}\left\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_{N_C}^2\right\}$, and $\sigma_i^2 = E\left[\rho(\theta_i)\rho^*(\theta_i)\right]$. Note that $\mathbf{R}_{\mathbf{H}_c}$ is a positive semidefinite Hermitian matrix [33].

We now consider the *constrained biased* CRB of the unknown target parameters $\mathbf{x} = \left[\boldsymbol{\theta}^T, \boldsymbol{\beta}_R^T, \boldsymbol{\beta}_I^T\right]^T$, where $\boldsymbol{\beta}_I = \left[\beta_{I,1}, \beta_{I,2}, \cdots, \beta_{I,K}\right]^T$, $\boldsymbol{\beta}_I = \left[\beta_{I,1}, \beta_{I,2}, \cdots, \beta_{I,K}\right]^T$, $\beta_R = \text{Re}(\beta)$, $\beta_R = \text{Re}(\beta)$ and $\beta_I = \text{Im}(\beta)$. According to Zvika and Eldar Yonina [29], if $\Re(\mathbf{UU}^H(\mathbf{I}+\mathbf{D})^H) \subseteq \Re(\mathbf{UU}^H\mathbf{FUU}^H)$, the *constrained biased* CRB can be written as

$$\mathbf{J}_{\text{CBCRB}} = (\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{FU})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H, \tag{6}$$

where

$$\mathbf{D}(\mathbf{x}) = \frac{\partial \mathbf{d}(\mathbf{x})}{\partial \mathbf{x}}, \tag{7}$$

with $\mathbf{d}(\mathbf{x})$ denoting the bias for estimating $\mathbf{x}$. U satisfies:

$$\mathbf{G}(\mathbf{x})\mathbf{U}(\mathbf{x}) = \mathbf{0}, \quad \mathbf{U}^H(\mathbf{x})\mathbf{U}(\mathbf{x}) = \mathbf{I} \tag{8}$$

in which $G(\mathbf{x}) = \frac{\partial \mathbf{g}(\mathbf{x})}{\partial \mathbf{x}}$ is assumed to have full row rank with $\mathbf{g}(\mathbf{x})$ being the equality constraint set on $\mathbf{x}$ and $\mathbf{U}$ is the tangent hyperplane of $\mathbf{g}(\mathbf{x})$ [20].

Following [20,21], some prior information can be available in array signal processing, for example, constant modulus constraint on the transmitted waveform, and the signal subspace constraints in the estimation of the angle-of-arrival. Here, we assume that the complex amplitude matrix $\boldsymbol{\beta} = \text{diag}(\beta_1, \beta_2, ..., \beta_k)$ is known as

$$\begin{aligned} g_i(\mathbf{x}) &= \beta_{R,i} - 1 = 0, \quad i = 1, \ldots, K \\ g_j(\mathbf{x}) &= \beta_{I,j} - 1 = 0, \quad j = K+1, \ldots, 2K \end{aligned} \tag{9}$$

Remark

In practice, the parameters of one target can be estimated roughly from the received data by many methods (see, e.g., [36] for more details). Therefore, we can obtain the imprecise knowledge of one target by transmitting orthogonal (or uncorrelated) waveforms before waveform optimization. In this article, our main interest is only to improve the accuracy of location estimation by optimizing transmitted waveforms. One can see from Section 3 that the waveform optimization is based on the FIM $\mathbf{F}$ that considers the unknown parameters consisting of the location and complex amplitude (see, (11)-(16)). Hence, the estimation of complex amplitude matrix $\boldsymbol{\beta}$ is regarded as prior information for waveform optimization here.

Following (9), we can obtain $\mathbf{G} = [\mathbf{0}_{2K\times K}, \mathbf{I}_{2K\times 2K}]$, where $\mathbf{0}_{2K\times K}$ denotes a zero matrix of size $2K \times K$. Hence, the corresponding null space $\mathbf{U}$ can be expressed as

$$\mathbf{U} = [\mathbf{I}_{K\times K} \ \ \mathbf{0}_{K\times 2K}]^H. \tag{10}$$

Based on the discussion above, the Fisher information matrix (FIM) $\mathbf{F}$ with respect to $\mathbf{x}$ is derived in Appendix A and given by

$$\mathbf{F} = 2\begin{bmatrix} \text{Re}(\mathbf{F}_{11}) & \text{Re}(\mathbf{F}_{12}) & -\text{Im}(\mathbf{F}_{12}) \\ \text{Re}^T(\mathbf{F}_{12}) & \text{Re}(\mathbf{F}_{22}) & -\text{Im}(\mathbf{F}_{22}) \\ -\text{Im}^T(\mathbf{F}_{12}) & -\text{Im}^T(\mathbf{F}_{22}) & \text{Re}(\mathbf{F}_{22}) \end{bmatrix}, \tag{11}$$

where

$$[\mathbf{F}_{11}]_{ij} = \beta_i^* \beta_j \dot{\mathbf{h}}_i^H \left[(\mathbf{I} + (\mathbf{R}_S \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c})^{-1}(\mathbf{R}_S \otimes \mathbf{B}^{-1})\right]\dot{\mathbf{h}}_j, \tag{12}$$

$$[\mathbf{F}_{12}]_{ij} = \beta_i^* \dot{\mathbf{h}}_i^H \left[(\mathbf{I} + (\mathbf{R}_S \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c})^{-1}(\mathbf{R}_S \otimes \mathbf{B}^{-1})\right]\mathbf{h}_j, \tag{13}$$

$$[\mathbf{F}_{22}]_{ij} = \mathbf{h}_i^H \left[\left(\mathbf{I} + (\mathbf{R_S} \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c} \right)^{-1} (\mathbf{R_S} \otimes \mathbf{B}^{-1}) \right] \mathbf{h}_j \quad (14)$$

$$\mathbf{h}_k = \mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k), \quad (15)$$

$$\dot{\mathbf{h}}_k = \frac{\partial(\mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k))}{\partial \theta_k}, \; k = 1, 2, \ldots, K, \quad (16)$$

$$\mathbf{R_S} = \mathbf{S}^* \mathbf{S}^T. \quad (17)$$

The problem of main interest in this study is the joint optimization of the WCM and bias estimator to improve the performance of parameter estimation by minimizing the *constrained biased* CRB of target locations. It can be seen from (6) that the *constrained biased* CRB depends on **U**, **D**, and **F**. In practice, it is not obvious how to choose a particular matrix **D** to minimize the total variance [23]. Even if a bias gradient matrix is given, it may not be suitable because a *biased estimator* reduces the variance obtained by any *unbiased estimator* at the cost of increasing the bias. As a sequence, a tradeoff between the variance and bias should be made, i.e., the biased estimator should be optimized [24]. According to Hero and Cramer-Rao [23], optimizing the bias estimator requires its bias gradient belonging to a suitable class. In this article, two constraints on the bias gradient are considered, i.e., the weighted and spectral norm constraints. In Section 3, with each norm constraint, we treat the joint optimization problem under two design criteria, i. e., minimizing the trace and the largest eigenvalue of the *constrained biased* CRB.

3 Joint optimization

In this section, we demonstrate how the WCM and bias estimator can be jointly optimized by minimizing the *constrained biased* CRB. First of all, this problem is considered under the weighted norm constraint.

A. Joint Optimization With the Weighted Norm Constraint

Similar to [28], the weighted norm constraint can be expressed as

$$\mathrm{tr}(\mathbf{D}^H \mathbf{D} \mathbf{M}) \le \gamma, \quad (18)$$

where **M** is a non-negative definite Hermitian weighted matrix, and γ is a constant which satisfies:

$$\gamma < \mathrm{tr}(\mathbf{M}). \quad (19)$$

First, we consider this problem by minimizing the trace of the *constrained biased* CRB, which is referred to as the *Trace-opt* criterion [7]. Under the weighted norm constraint (18) and the total transmitted power constraint, the optimization problem can be formulated as

$$\begin{array}{ll} \min\limits_{\mathbf{R_S},\mathbf{D}} & \mathrm{tr}(\mathbf{J}_{\mathrm{CBCRB}}) \\ \text{s.t.} & \mathrm{tr}(\mathbf{R_S}) = LP \\ & \mathbf{R_S} \succcurlyeq \mathbf{0} \\ & \mathrm{tr}(\mathbf{D}^H \mathbf{D} \mathbf{M}) \le \gamma \end{array}, \quad (20)$$

where the second constraint holds because the power transmitted by each transmitting element is more than or equal to zero [6], and P is the total transmitted power.

It can be seen from (6) that $\mathrm{J}_{\mathrm{CBCRB}}$ is a linear function of F^{-1}, and a quadratic one of **D**. Moreover, **F** is a nonlinear function of $\mathbf{R_S}$, which can be seen from (11)-(14). As a sequence, this problem is a rather complicated nonlinear optimization one, and hence it is difficult to be treated by convex optimization methods [30-32]. In order to solve it, we make a simplifying assumption that $\mathbf{R_S} \otimes \mathbf{B}^{-1}$ spans the same subspace as $\mathbf{R}_{\mathbf{H}_c}$, i.e.,

$$\Re(\mathbf{R_S} \otimes \mathbf{B}^{-1}) = \Re(\mathbf{R}_{\mathbf{H}_c}), \quad (21)$$

the rationality of which is proved under a certain condition in Appendix B. Under this assumption, according to Horn and Johnson [37], the product of $\mathbf{R_S} \otimes \mathbf{B}^{-1}$ and $\mathbf{R}_{\mathbf{H}_c}$, denoted by $\mathbf{R_{SC}}$, is positive semidefinite, i.e.,

$$\mathbf{R_{SC}} \succcurlyeq \mathbf{0} \quad (22)$$

With (22), the problem in (20) can be solved by SDP relying on the following lemma [38, pp. 472]:

Lemma 1

(*Schur's Complement*) *Let* $\mathbf{Z} = \begin{bmatrix} \mathbf{A} & \mathbf{B}^H \\ \mathbf{B} & \mathbf{C} \end{bmatrix}$ *be a Hermitian matrix with* $\mathbf{C} \succ \mathbf{0}$, *then* $\mathbf{Z} \succcurlyeq \mathbf{0}$ *if and only if* $\Delta\mathbf{C} \succcurlyeq \mathbf{0}$, *where* $\Delta\mathbf{C}$ *is the Schur complement of* **C** *in* **Z** *and is given by* $\Delta\mathbf{C} = \mathbf{A}\text{-}\mathbf{B}^H\mathbf{C}^{-1}\mathbf{B}$.

Using Lemma 1, the proposition 1 below can reformulate the nonlinear objective in (20) as a linear one, and give the corresponding linear matrix inequality (LMI) formulations of the first two constraints, which is proved in Appendix C.

Proposition 1

Using matrix manipulations, the first two constraints in (20) can be converted into the following LMIs:

$$\begin{bmatrix} \tau & \mathrm{vec}(\mathbf{I}_{M_t M_r})^H \\ \mathrm{vec}(\mathbf{I}_{M_t M_r}) & \mathbf{I}_{M_t M_r} \otimes (\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c}) \end{bmatrix} \succcurlyeq \mathbf{0} \quad (23)$$

$$\mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta \mathbf{I}, \quad (24)$$

where

$$\mathbf{E} = (\mathbf{I} + (\mathbf{R}_S \otimes \mathbf{B}^{-1})\mathbf{R}_{H_c})^{-1} (\mathbf{R_S} \otimes \mathbf{B}^{-1}). \quad (25)$$

and τ, β *are given in* (75) *and* (87), *respectively.* According to Lemma 1, the matrix $\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c}$ must be positive definite, which can be guaranteed by (72). From

(11)-(14) and (25), it is known that the nonlinear objective in (20) can be converted into a linear one with respect to **E**.

With (6), (23) and (24), the problem (20) can be equivalently represented as

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad \operatorname{tr}((\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) \le t\\ &\qquad \operatorname{tr}(\mathbf{D}^H\mathbf{D}\mathbf{M}) \le \gamma\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{\mathbf{H}_c})\end{bmatrix} \succcurlyeq 0\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}\end{aligned} \tag{26}$$

where t is an auxiliary variable.

It is noted that the terms in the left hand of the first two constraint inequalities in (26) are quadratic functions of **D**, and hence these inequalities are not LMIs. The Proposition 2 below can give the LMI formulations of these inequalities, which is proved in Appendix D.

Proposition 2

Using Lemma 1 and some matrix lemmas, the first two constraint inequalities in (26) *can be, respectively, expressed as*

$$\begin{bmatrix} t & (\operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H))^H\\ \operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) & (\mathbf{I}\otimes(\mathbf{U}^H\mathbf{F}\mathbf{U}))\end{bmatrix} \succcurlyeq \mathbf{0}, \tag{27}$$

$$\begin{bmatrix} \gamma & \operatorname{vec}(\mathbf{M}^{1/2}\mathbf{D}^H)^H\\ \operatorname{vec}(\mathbf{D}\mathbf{M}^{1/2}) & \mathbf{I}\end{bmatrix} \succcurlyeq \mathbf{0} \tag{28}$$

Now, the joint optimization problem (20) can be readily cast as an SDP

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad \begin{bmatrix} t & (\operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H))^H\\ \operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) & (\mathbf{I}\otimes(\mathbf{U}^H\mathbf{F}\mathbf{U}))\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix} \gamma & \operatorname{vec}(\mathbf{M}^{1/2}\mathbf{D}^H)^H\\ \operatorname{vec}(\mathbf{D}\mathbf{M}^{1/2}) & \mathbf{I}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{H_c})\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}\end{aligned} \tag{29}$$

Next, the joint optimization problem is treated by minimizing the largest eigenvalue of the constrained biased CRB, which is referred to as the Eigen-opt criterion [7]. Similar to the case of the Trace-opt criterion, the problem can be expressed as

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad (\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H \preccurlyeq t\mathbf{I}\\ &\qquad \operatorname{tr}(\mathbf{D}^H\mathbf{D}\mathbf{M}) \le \gamma\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{\mathbf{H}_c})\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}\end{aligned} . \tag{30}$$

Using Lemma 1 and the results above, this problem is equivalent to SDP as

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad \begin{bmatrix} t\mathbf{I} & (\mathbf{I}+\mathbf{D})\mathbf{U}\\ ((\mathbf{I}+\mathbf{D})\mathbf{U})^H & \mathbf{U}^H\mathbf{F}\mathbf{U}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix} \gamma & \operatorname{vec}(\mathbf{M}^{1/2}\mathbf{D}^H)^H\\ \operatorname{vec}(\mathbf{D}\mathbf{M}^{1/2}) & \mathbf{I}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{\mathbf{H}_c})\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}\end{aligned} \tag{31}$$

B Joint Optimization With the Spectral Norm Constraint

The spectral norm constraint, similar to [28], can be written as

$$\mathbf{T}^H\mathbf{D}\mathbf{D}^H\mathbf{T} \preccurlyeq \gamma\mathbf{I}, \tag{32}$$

where **T** is a non-negative definite Hermitian matrix, and γ is a constant satisfying:

$$\gamma < \lambda_{\max}^2(\mathbf{T}), \tag{33}$$

with $\lambda_{\max}(\mathbf{T})$ denoting the largest eigenvalue of **T**.

First, we consider the *trace-opt* criterion. Under the spectral norm constraint (32), the problem can be similarly written as

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{R}_\mathbf{S}} t\\ &\text{s.t.}\quad \operatorname{tr}((\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) \le t\\ &\qquad \mathbf{T}^H\mathbf{D}\mathbf{D}^H\mathbf{T} \preccurlyeq \gamma\mathbf{I}\\ &\qquad \operatorname{tr}(\mathbf{R}_\mathbf{S}) = LP\\ &\qquad \mathbf{R}_\mathbf{S} \succcurlyeq \mathbf{0}\end{aligned} . \tag{34}$$

Following *Lemma 1* and the propositions above, (34) can be recast as SDP

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad \begin{bmatrix} t & (\operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H))^H\\ \operatorname{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) & (\mathbf{I}\otimes(\mathbf{U}^H\mathbf{F}\mathbf{U}))\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix} \gamma\mathbf{I} & \mathbf{T}^H\mathbf{D}\\ \mathbf{D}^H\mathbf{T} & \mathbf{I}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{\mathbf{H}_c})\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}\end{aligned} . \tag{35}$$

Second, similar to the discussion above, the optimization problem under the *Eigen-opt* criterion can be represented as SDP

$$\begin{aligned}&\min_{t,\mathbf{D},\mathbf{E}} t\\ &\text{s.t.}\quad \begin{bmatrix} t\mathbf{I} & (\mathbf{I}+\mathbf{D})\mathbf{U}\\ ((\mathbf{I}+\mathbf{D})\mathbf{U})^H & \mathbf{U}^H\mathbf{F}\mathbf{U}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix} \gamma\mathbf{I} & \mathbf{T}^H\mathbf{D}\\ \mathbf{D}^H\mathbf{T} & \mathbf{I}\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \begin{bmatrix}\tau & \operatorname{vec}(\mathbf{I}_{M_tM_r})^H\\ \operatorname{vec}(\mathbf{I}_{M_tM_r}) & \mathbf{I}_{M_tM_r}\otimes(\mathbf{I}-\mathbf{E}\mathbf{R}_{\mathbf{H}_c})\end{bmatrix} \succcurlyeq \mathbf{0}\\ &\qquad \mathbf{0} \preccurlyeq \mathbf{E}\mathbf{R}_{H_c} \preccurlyeq \beta\mathbf{I}\end{aligned} . \tag{36}$$

After obtaining the optimum $\mathbf{E}$ from (29), (31), (35), and (36), the term $\mathbf{R_{SB}} = \mathbf{R_S} \otimes \mathbf{B}^{-1}$ can be solved via (25), which can be reshaped as

$$(\mathbf{I}_{M_t M_r} + \mathbf{R_{SB}}\mathbf{R_{H_c}})\mathbf{E} = \mathbf{R_{SB}}. \quad (37)$$

From (37), we have

$$\mathbf{R_{SB}} = \mathbf{E}(\mathbf{I}_{M_t M_r} - \mathbf{R_{H_c}}\mathbf{E})^{-1}. \quad (38)$$

Scale $\mathbf{R_{SB}}$ such that

$$\mathrm{tr}(\alpha\mathbf{R_{SB}}) = LP\mathrm{tr}(\mathbf{B}^{-1}), \quad (39)$$

where α is a scalar which satisfies the equality constraint.

Given $\mathbf{R_{SB}}$, $\mathbf{R_S}$ can be constructed via a suitable approximation to it (in a LS sense), which is formulated as

$$\begin{aligned} &\mathbf{R_S} = \arg\min_{\mathbf{R_S}} \left\| \mathbf{R_{SB}} - \mathbf{R_S} \otimes \mathbf{B}^{-1} \right\|_F \\ &s.t. \quad \mathrm{tr}(\mathbf{R_S}) = LP \\ &\qquad \mathbf{R_S} \succeq \mathbf{0} \end{aligned} \quad (40)$$

The problem above can be equivalently represented as

$$\begin{aligned} &\min_{\mathbf{R_S},t} t \\ &s.t. \quad \left\| \mathbf{R_{SB}} - \mathbf{R}_S \otimes \mathbf{B}^{-1} \right\|_F \le t \\ &\qquad \mathrm{tr}(\mathbf{R_S}) = LP \\ &\qquad \mathbf{R_S} \succeq \mathbf{0} \end{aligned}. \quad (41)$$

Using Lemma 1, (41) can be equivalently represented as an SDP

$$\begin{aligned} &\min_{\mathbf{R_S},t} t \\ &s.t. \quad \begin{bmatrix} t & \mathrm{vec}^H(\mathbf{R_{SB}} - \mathbf{R_S} \otimes \mathbf{B}^{-1}) \\ \mathrm{vec}(\mathbf{R_{SB}} - \mathbf{R_S} \otimes \mathbf{B}^{-1}) & \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \\ &\qquad \mathrm{tr}(\mathbf{R_S}) = LP \\ &\qquad \mathbf{R_S} \succeq \mathbf{0} \end{aligned}. \quad (42)$$

Using many well-known algorithms (see, e.g., [30-32]) for solving SDP problems, the problems in (29), (31), (35), (36), and (42) can be solved very efficiently. In the following examples, the optimization toolbox in [32] is used for these problems. It is noted that we only obtain the WCM other than the ultimate transmitted waveforms in this article. In practice, the ultimate waveforms can be asymptotically synthesized by using the method in [39].

4 Numerical examples

In this section, some examples are provided to illustrate the effectiveness of the proposed method as compared with the uncorrelated transmitted waveforms (i.e., $\mathbf{R_S} = (P / M_t)\mathbf{I}$).

Consider a MIMO radar system with $M_t = 5$ transmitting elements and $M_r = 5$ receiving elements. We use the following two MIMO radar systems with various antenna configurations: MIMO radar (0.5, 0.5), and MIMO radar (2.5, 0.5), where the parameters specifying each radar system are the inter-element spacing of the transmitter and receiver (in units of wavelengths), respectively. Let the weighted matrix $\mathbf{M} = \mathbf{I}$ and $\gamma = 1$ in the case of the weighted norm constraint, and $\mathbf{T} = \mathbf{I}$ and $\gamma = 0.5$ in the other case. In the following examples, two targets with unit amplitudes are considered, which are located, respectively, at $\theta_1 = 0^o$ and $\theta_2 = 13^o$ for MIMO radar (0.5, 0.5), and $\theta_1 = 0^o$and $\theta_2 = 7^o$ for MIMO radar (2.5, 0.5). The number of snapshots is $L = 256$. The array signal-to-noise ratio (ASNR) in the following examples varying from -10 to 50 dB is defined as PM_tM_r/σ_W^2, where σ_W^2 denotes the variance of the additive white thermal noise. The clutter is modelled as $N_c = 10000$ discrete patches equally spaced on the range bin of interest. The RCSs of these clutter patches are modelled as independent and identically distributed zero mean Gaussian random variables, which are assumed to be fixed in the coherent processing interval (CPI). The clutter-to-noise ratio (CNR) is defined as $\mathrm{tr}(R_{H_c})/\sigma_W^2$, which ranges from 10 to 50 dB. There is a strong jammer at -11° with an array interference-to-noise ratio (AINR) equal to 60 dB, defined as the product of the incident interference power and M_r divided by σ_W^2. The jammer is modeled as point source which transmits white Gaussian signal uncorrelated with the signals transmitted by MIMO radar.

From Section 3, it is known that the joint optimization problem is based on the CRB that requires the specification of some parameters, e.g., the target location and clutter covariance matrix. In practice, the target parameters and clutter covariance can be estimated by using the method in [36,35], respectively.

In order to examine the effectiveness of the proposed method, we will focus on the following three cases: the CRB of two angles with exactly known initial parameters, the effect of the optimal *biased estimator* or prior information on the CRB, and the effect of the initial parameter estimation errors on the CRB.

A.The CRB Without Initial Estimation Errors

Figure 1 shows the optimal transmit beampatterns under the *Trace-opt* criterion in the case of ASNR = 50 dB and CNR = 10 dB. It can be seen that a notch is placed almost at the jammer location. Moreover, the difference between the powers obtained by two targets is large because only the total CRB is minimized here excluding the CRB of every parameter. As a sequence, for a certain parameter, the CRB obtained by the optimal waveforms may be larger than that of uncorrelated waveforms.

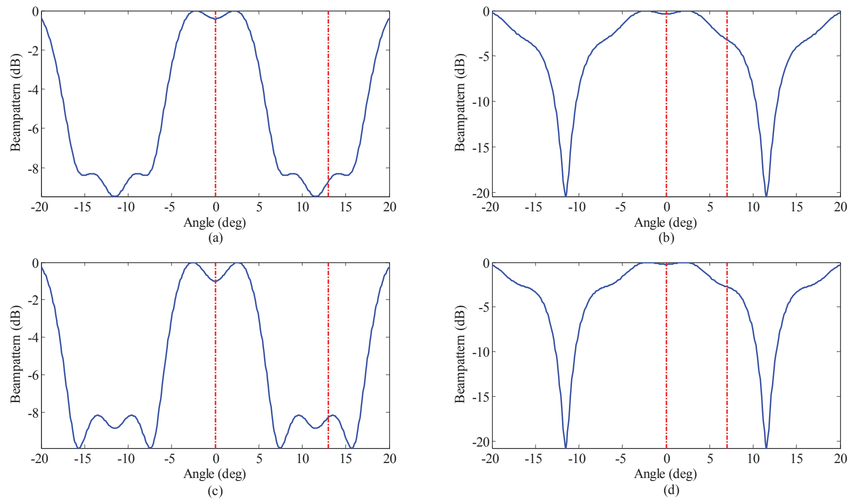

Figure 1 Optimal transmit beam patterns under the *Trace-opt* criterion with ASNR = 50 dB and CNR = 10 dB. **(a)** With the weighted norm constraint for MIMO radar (0.5, 0.5). **(b)** With the weighted norm constraint for MIMO radar (2.5, 0.5). **(c)** With the spectral norm constraint for MIMO radar (0.5, 0.5). **(d)** With the spectral norm constraint for MIMO radar (2.5, 0.5).

Figure 2 shows the CRB of two angles as a function of ASNR or CNR. One can see that the CRB obtained by our method or uncorrelated waveforms decreases as the increasing of ASNR, while increases as the decreasing of CNR. Moreover, the CRB under the *Trace-opt* or *Eigen-opt* criterion is much lower than that of uncorrelated waveforms, regardless of ASNR or CNR. Furthermore, under the same norm constraint, the *Trace-opt* criterion leads to a lower total CRB than the *Eigen-opt* criterion. Besides, by comparing Figure 2a with 2c or Figure 2b with 2d, it follows that the total CRB for MIMO radar (2.5, 0.5) is lower than that for MIMO radar (0.5, 0.5). This is because the virtual receiving array aperture for the former radar is much larger than that for the latter [3].

B.Effect of the Optimal Biased Estimator or Prior Information on the CRB

In this subsection, we will study the CRB obtained by only using the optimal *biased estimator* or prior information.

First, only the optimal *biased estimator* is employed. In this case, let the matrix **u** in (6) be equal to **I** (All other parameters are the same as the previous

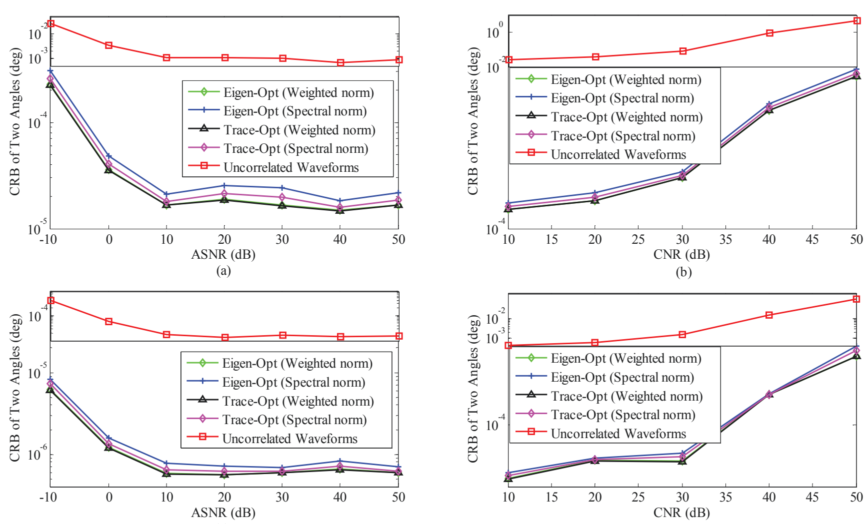

Figure 2 CRB of two angles versus ASNR or CNR. **(a)** CRB versus ASNR with CNR = 10 dB for MIMO radar (0.5, 0.5). **(b)** CRB versus CNR with ASNR = -10 dB for MIMO radar (0.5, 0.5). **(c)** CRB versus ASNR with CNR = 10 dB for MIMO radar (2.5, 0.5). **(d)** CRB versus CNR with ASNR = -10 dB for MIMO radar (2.5, 0.5).

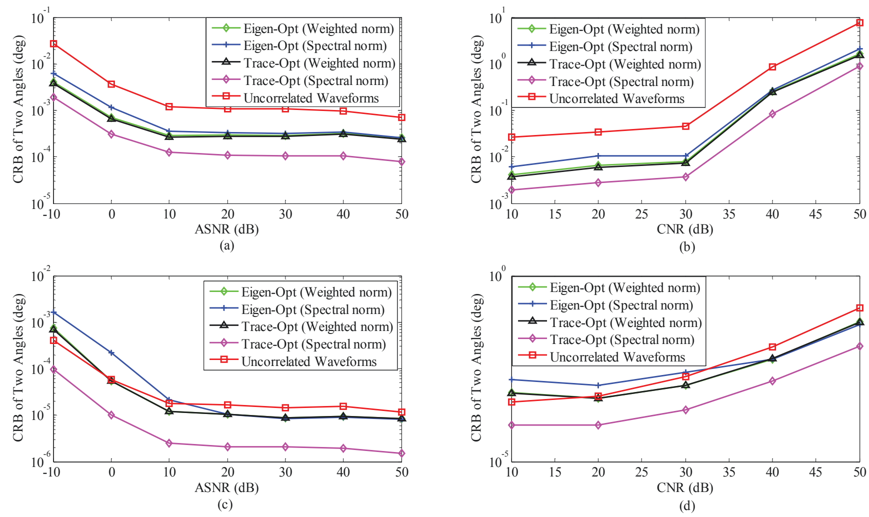

Figure 3 CRB of two angles obtained only by using the optimal *biased estimator*, as well as that of the uncorrelated waveforms, versus ASNR or CNR. **(a)** CRB versus ASNR with CNR = 10 dB for MIMO radar (0.5, 0.5). **(b)** CRB versus CNR with ASNR = -10 dB for MIMO radar (0.5, 0.5). **(c)** CRB versus ASNR with CNR = 10 dB for MIMO radar (2.5, 0.5). **(d)** CRB versus CNR with ASNR = -10 dB for MIMO radar (2.5, 0.5).

examples.). The variant of the CRB for this case is the *biased* CRB as mentioned above. Figure 3 shows the CRB in this case as a function of ASNR or CNR. It can be seen that the optimal *biased estimator* may lead to a little higher CRB than using the uncorrelated waveforms sometimes, which is because the total CRB of the amplitudes of two targets is not taken into account here. Moreover, the *Trace-opt* criterion leads to higher improvement of the CRB than the *Eigen-opt* one under the same norm constraint, which is similar to the results obtained from Figure 2.

Second, we examine the CRB obtained by only using the prior information. In this case, let the matrix $\mathbf{D}$ in (6) be equal to $\mathbf{0}_{3k\times\ 3k}$ and all the other parameters remain the same as the previous examples. The variant of the CRB for this case is the *constrained* CRB as stated above. Figure 4 shows the CRB in the case as a function of ASNR or CNR. One can observe that the contributions of the prior information to two optimization criteria are almost identical, and the prior information can significantly improve the accuracy of parameter estimation with the uncorrelated waveforms.

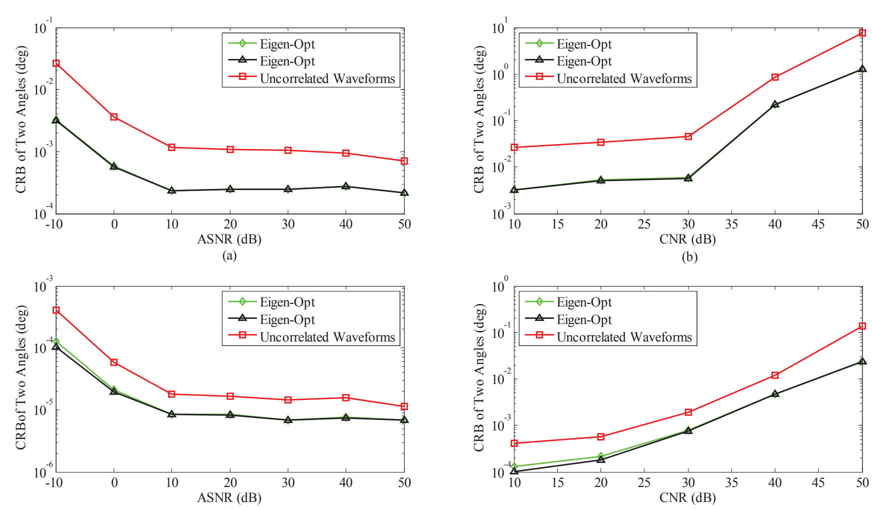

Figure 4 CRB obtained only by using the prior information, along with that of the uncorrelated waveforms, versus ASNR or CNR. **(a)** CRB versus ASNR with CNR = 10 dB for MIMO radar (0.5, 0.5). **(b)** CRB versus CNR with ASNR = -10 dB for MIMO radar (0.5, 0.5). **(c)** CRB versus ASNR with CNR = 10 dB for MIMO radar (2.5, 0.5). **(d)** CRB versus CNR with ASNR = -10 dB for MIMO radar (2.5, 0.5).

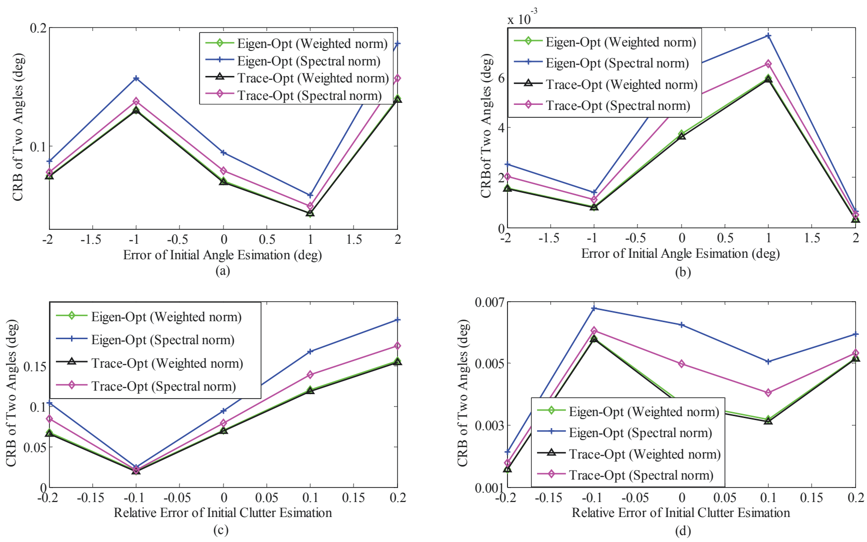

Figure 5 CRB versus angle or clutter estimation error with ASNR = -10 dB and CNR = 50 dB. **(a)** CRB versus initial angle estimation error for MIMO radar (0.5, 0.5). **(b)** CRB versus initial angle estimation error for MIMO radar (2.5, 0.5). **(c)** CRB versus initial clutter estimation error for MIMO radar (0.5, 0.5). **(d)** CRB versus initial clutter estimation error for MIMO radar (2.5, 0.5).

C. Effect of the Initial Parameter Estimation Errors on the CRB

In this subsection, we consider the effect of the initial angle or clutter estimation error on the CRB of two angles. It is noted that the relative error of the clutter estimate is defined as the ratio of the estimation error of the initial total clutter power to the exact one.

Figure 5 shows the CRB versus the estimation error of the initial angle or clutter power with ASNR = -10 dB and CNR = 50 dB under the condition that all the other parameters are exact. We can see that the CRB varies with the estimate error of the angle or clutter very apparently, which indicates that the proposed method is very sensitive to these errors. Hence, the robust method for waveform design is worthy of investigating in the future.

5 Conclusions

In this article, we have proposed a novel *constrained biased* CRB-based method to optimize the WCM and biased estimator to improve the performance of parameter estimation of point targets in MIMO radar in the presence of clutter. The resultant nonlinear optimization problem can be solved resorting to the SDP relaxation under a simplifying assumption. A solution of the initial problem is provided via approximating to an optimal solution of the SDP one (in a LS sense). Numerical examples show that the proposed method can significantly improve the accuracy of parameter estimation in the case of uncorrelated waveforms. Moreover, under the weighted norm constraint, the *Trace-opt* criterion results in a lower CRB than the *Eigen-opt* one. As illustrated by examples in Section IV, the performance of the proposed method may be degraded when the initial parameter estimates are exploited. One way to overcome this performance degradation is to develop a more robust algorithm for joint optimization against the estimation error, which will be investigated in the future.

Appendix A

Fisher information matrix

Consider the signal model in (3), and stack the columns of $\mathbf{Y}$ in a $M_r L \times 1$ vector as

$$\mathbf{y} = (\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \sum_{k=1}^{K} \beta_k (\mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k)) + (\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \mathrm{vec}(\mathbf{H}_c) + \mathrm{vec}(\mathbf{W}). \quad (43)$$

Similar to [7], we calculate the FIM with respect to $\boldsymbol{\theta}$, $\boldsymbol{\beta}_R$, $\boldsymbol{\beta}_I$ (Here we only consider one-dimensional targets.). According to Xu et al. [40], we have

$$F(x_i, x_j) = 2\mathrm{Re}\left\{ \mathrm{tr}\left[\frac{\partial \left((\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \sum_{k=1}^{K} \beta_k (\mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k)) \right)^H}{\partial x_i} \mathbf{Q}^{-1} \frac{\partial \left((\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \sum_{k=1}^{K} \beta_k (\mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k)) \right)}{\partial x_j} \right] \right\}, \quad (44)$$

where $\mathbf{Q}$ denotes the covariance of the clutter plus interference and noise, which can be represented as

$$\mathbf{Q} = E\left\{ \left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \mathrm{vec}(\mathbf{H}_c) + \mathrm{vec}(\mathbf{W}) \right] \left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \mathrm{vec}(\mathbf{H}_c) + \mathrm{vec}(\mathbf{W}) \right]^H \right\} \quad (45)$$

With (4), (45) can be simplified as

$$\mathbf{Q} = (\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \mathbf{R}_{\mathbf{H}_c} (\mathbf{S}^* \otimes \mathbf{I}_{M_r}) + \mathbf{I}_{M_t} \otimes \mathbf{B} \quad (46)$$

Let $\mathbf{h}_k = \mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k)$. Note that

$$F(\theta_i, \theta_j) = 2\mathrm{Re}\left\{\mathrm{tr}\left[\frac{\partial\left((\mathbf{S}^T \otimes \mathbf{I}_{M_r})\sum_{k=1}^{K}\beta_k\mathbf{h}_k\right)^H}{\partial\theta_i}\mathbf{Q}^{-1}\frac{\partial\left((\mathbf{S}^T \otimes \mathbf{I}_{M_r})\sum_{k=1}^{K}\beta_k\mathbf{h}_k\right)}{\partial\theta_j}\right]\right\}. \quad (47)$$

Because

$$\frac{\partial\left((\mathbf{S}^T \otimes \mathbf{I}_{M_r})\sum_{k=1}^{K}\beta_k\mathbf{h}_k\right)}{\partial\theta_i} = (\mathbf{S}^T \otimes \mathbf{I}_{M_r})\beta_i\dot{\mathbf{h}}_i, \quad (48)$$

then

$$\begin{aligned} F(\theta_i, \theta_j) &= 2\mathrm{Re}\left\{\mathrm{tr}\left[\beta_i^*\beta_j\dot{\mathbf{h}}_i^H(\mathbf{S}^* \otimes \mathbf{I}_{M_r})\mathbf{Q}^{-1}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\dot{\mathbf{h}}_j\right]\right\} \\ &= 2\mathrm{Re}\left\{\beta_i^*\beta_j\dot{\mathbf{h}}_i^H(\mathbf{S}^* \otimes \mathbf{I}_{M_r})\left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{R}_{\mathbf{H}_c}(\mathbf{S}^* \otimes \mathbf{I}_{M_r}) + \mathbf{I}_{M_t} \otimes \mathbf{B}\right]^{-1}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\dot{\mathbf{h}}_j\right\} \end{aligned} \quad (49)$$

Let
$\mathbf{A} = (\mathbf{S}^* \otimes \mathbf{I}_{M_r})\left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{R}_{\mathbf{H}_c}(\mathbf{S}^* \otimes \mathbf{I}_{M_r}) + \mathbf{I}_{M_t} \otimes \mathbf{B}\right]^{-1}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})$.
By using matrix inversion lemma, we can get

$$\begin{aligned} \mathbf{A} &= (\mathbf{S}^* \otimes \mathbf{I}_{M_r})\left\{\mathbf{I}_{M_t} \otimes \mathbf{B}^{-1} - (\mathbf{S}^T \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}\left[\mathbf{I}_{M_tM_r} + ((\mathbf{S}^*\mathbf{S}^T) \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}\right]^{-1}(\mathbf{S}^* \otimes \mathbf{B}^{-1})\right\}(\mathbf{S}^T \otimes \mathbf{I}_{M_r}) \\ &= (\mathbf{S}^*\mathbf{S}^T) \otimes \mathbf{B}^{-1} - ((\mathbf{S}^*\mathbf{S}^T) \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}\left[\mathbf{I}_{M_tM_r} + ((\mathbf{S}^*\mathbf{S}^T) \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}\right]^{-1}(\mathbf{S}^*\mathbf{S}^T) \otimes \mathbf{B}^{-1} \\ &= (\mathbf{I}_{M_tM_r} + (\mathbf{R}_S \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c})^{-1}(\mathbf{R}_S \otimes \mathbf{B}^{-1}) \end{aligned} \quad (50)$$

where $\mathbf{R}_\mathbf{S} = \mathbf{S}^*\mathbf{S}^\mathrm{T}$. With (50), (49) can be rewritten as

$$F(\theta_i, \theta_j) = 2\mathrm{Re}\left\{\beta_i^*\beta_j\dot{\mathbf{h}}_i^H\left[\mathbf{I}_{M_tM_r} + (\mathbf{R}_S \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}\right]^{-1}(\mathbf{R}_S \otimes \mathbf{B}^{-1})\dot{\mathbf{h}}_j\right\}, \quad (51)$$

and hence

$$\mathbf{F}(\boldsymbol{\theta}, \boldsymbol{\theta}) = 2\mathrm{Re}(\mathbf{F}_{11}), \quad (52)$$

where $\mathbf{F}_{11}$ is given in (12).
Similarly, we have

$$\frac{\partial\left((\mathbf{S}^T \otimes \mathbf{I}_{M_r})\sum_{k=1}^{K}\beta_k\mathbf{h}_k\right)}{\partial\beta_{R,i}} = (\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{h}_k, \quad (53)$$

and

$$\frac{\partial\left((\mathbf{S}^T \otimes \mathbf{I}_{M_r})\sum_{k=1}^{K}\beta_k\mathbf{h}_k\right)}{\partial\beta_{I,i}} = j(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{h}_k. \quad (54)$$

Hence

$$\mathbf{F}(\boldsymbol{\theta}, \boldsymbol{\beta}_R) = \mathbf{F}^T(\boldsymbol{\theta}, \boldsymbol{\beta}_R) = 2\mathrm{Re}(\mathbf{F}_{12}), \quad (55)$$

and

$$\mathbf{F}(\boldsymbol{\theta}, \boldsymbol{\beta}_I) = \mathbf{F}^T(\boldsymbol{\theta}, \boldsymbol{\beta}_I) = -2\mathrm{Im}(\mathbf{F}_{12}), \quad (56)$$

where $\mathbf{F}_{12}$ is given in (13).
We also have

$$\mathbf{F}(\boldsymbol{\beta}_R, \boldsymbol{\beta}_R) = \mathbf{F}(\boldsymbol{\beta}_I, \boldsymbol{\beta}_I) = 2\mathrm{Re}(\mathbf{F}_{22}), \quad (57)$$

and

$$\mathbf{F}(\boldsymbol{\beta}_I, \boldsymbol{\beta}_R) = \mathbf{F}^T(\boldsymbol{\beta}_R, \boldsymbol{\beta}_I) = -2\mathrm{Im}(\mathbf{F}_{22}) \quad (58)$$

where $\mathbf{F}_{22}$ is given in (14).
From (49) and (55)-(58), we can obtain (11) immediately.

Appendix B
Proof of the rationality of (21)

It is known that the CRB for an unbiased estimator can be achieved by using the minimum mean square error (MMSE) estimator [27]. Therefore, from the parameter estimation perspective, the optimal transmitted waveforms can be obtained through minimizing the MMSE estimation error. For convenience of derivation, we stack the collected data in (3) into a $M_rL \times 1$ vector as

$$\mathbf{y} = (\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{h}_\mathrm{t} + (\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{h}_\mathrm{c} + \mathrm{vec}(\mathbf{W}), \quad (59)$$

where $\mathbf{h}_\mathrm{t}$ = vec ($\mathbf{H}_\mathrm{t}$), $\mathbf{H}_\mathrm{t} = \sum_{k=1}^{K}\beta_k(\mathbf{v}(\theta_k) \otimes \mathbf{a}(\theta_k))$, and $\mathbf{h}_\mathrm{c}$ = vec ($\mathbf{H}_\mathrm{c}$). In order to minimize the MSE, the optimal MMSE estimator, denoting by $\mathbf{G}_\mathrm{op}$, should be firstly obtained. According to Eldar Yonina [28], $\mathbf{G}_\mathrm{op}$ can be obtained by solving the following optimization problem:

$$\mathbf{G}_\mathrm{op} = \arg\min_{\mathbf{G}} E\left\{\left\|\mathbf{h}_\mathrm{t} - \mathbf{G}\mathbf{y}\right\|_F^2\right\}, \quad (60)$$

Differentiating the above function with respect to $\mathbf{G}$ and setting it to zero, we have

$$\mathbf{G}_\mathrm{op} = \mathbf{R}_{\mathbf{H}_\mathrm{t}}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H\left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r})(\mathbf{R}_{\mathbf{H}_\mathrm{t}} + \mathbf{R}_{\mathbf{H}_\mathrm{c}})(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H + \mathbf{I}_{M_t} \otimes \mathbf{B}\right]^{-1}, \quad (61)$$

where $\mathbf{R}_{\mathbf{H}_\mathrm{t}} = E[\mathbf{h}_\mathrm{t}\mathbf{h}_\mathrm{t}^H]$. Hence, the MMSE estimate of $\mathbf{h}_\mathrm{t}$ can be represented as:

$$\hat{\mathbf{h}}_\mathrm{t} = \mathbf{G}_\mathrm{op}\mathbf{y}. \quad (62)$$

Accordingly, the MMSE estimation error can be written as

$$\varepsilon_\mathrm{MMSE} = \mathrm{tr}\left\{(\mathbf{h}_\mathrm{t} - \hat{\mathbf{h}}_\mathrm{t})(\mathbf{h}_\mathrm{t} - \hat{\mathbf{h}}_\mathrm{t})^H\right\}. \quad (63)$$

By substituting (61) and (62) into the equation above and using matrix inversion lemma, (63) can be rewritten as

$$\begin{aligned} \varepsilon_{MMSE} &= \mathrm{tr}\left\{\mathbf{R}_{\mathbf{H}_t} - \mathbf{R}_{\mathbf{H}_t}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H\left[(\mathbf{S}^T \otimes \mathbf{I}_{M_r})(\mathbf{R}_{\mathbf{H}_t} + \mathbf{R}_{\mathbf{H}_c})(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H + \mathbf{I}_{M_t} \otimes \mathbf{B}\right]^{-1}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{R}_{\mathbf{H}_t}\right\} \\ &= \mathrm{tr}\left\{\mathbf{R}_{\mathbf{H}_t} - \mathbf{R}_{\mathbf{H}_t}(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H(\mathbf{I}_{M_t} \otimes \mathbf{B}^{-1/2})\right. \\ &\quad \times\left[(\mathbf{I}_{M_t} \otimes \mathbf{B}^{-1/2})(\mathbf{S}^T \otimes \mathbf{I}_{M_r})(\mathbf{R}_{\mathbf{H}_t} + \mathbf{R}_{\mathbf{H}_c})(\mathbf{S}^T \otimes \mathbf{I}_{M_r})^H(\mathbf{I}_{M_t} \otimes \mathbf{B}^{-1/2}) + \mathbf{I}\right]^{-1} \\ &\quad \left.\times(\mathbf{I}_{M_t} \otimes \mathbf{B}^{-1/2})(\mathbf{S}^T \otimes \mathbf{I}_{M_r})\mathbf{R}_{\mathbf{H}_t}\right\} \\ &= \mathrm{tr}\left\{\mathbf{R}_{\mathbf{H}_t} - \mathbf{R}_{\mathbf{H}_t}(\mathbf{S}^* \otimes \mathbf{B}^{-1/2})\left[(\mathbf{S}^T \otimes \mathbf{B}^{-1/2})(\mathbf{R}_{\mathbf{H}_t} + \mathbf{R}_{\mathbf{H}_c})(\mathbf{S}^* \otimes \mathbf{B}^{-1/2}) + \mathbf{I}\right]^{-1}(\mathbf{S}^T \otimes \mathbf{B}^{-1/2})\mathbf{R}_{\mathbf{H}_t}\right\} \end{aligned} \quad (64)$$

which has the same form as Equation 3 shown in [19]. Therefore, according to Theorem 4 in [19], if $\mathbf{R}_{\mathbf{H}_t}$ and $\mathbf{R}_{\mathbf{H}_c}$ can be joint diagonalized, we can obtain

$$\mathbf{R}_\mathbf{S} \otimes \mathbf{B}^{-1} = \mathbf{Q}(\mathbf{\Lambda}_\mathrm{t} + \mathbf{\Lambda}_\mathrm{c})^{\dagger}[\mu\mathbf{\Lambda}_\mathrm{t} - \mathbf{I}]^{+}\mathbf{Q}^{H}, \tag{65}$$

where $\mathbf{\Lambda}_\mathrm{t}$ and $\mathbf{\Lambda}_\mathrm{c}$ are, respectively, the diagonal matrices with each diagonal entry given by a real and nonnegative eigenvalue of $\mathbf{R}_{\mathbf{H}_t}$ and $\mathbf{R}_{\mathbf{H}_c}$, $\mathbf{Q}$ is the unitary eigenvector matrix of $\mathbf{R}_{\mathbf{H}_t}$ and $\mathbf{R}_{\mathbf{H}_c}$, and μ is a scalar constant that satisfies the transmitted power constraints. It can be seen from (65) that $\mathbf{R}_\mathbf{S} \otimes \mathbf{B}^{-1}$ spans indeed the same subspace as $\mathbf{R}_{\mathbf{H}_c}$. The proof is completed.

Appendix C

Proof of proposition 1

In order to convert the objective in (20) into a linear function, let

$$\mathbf{E} = (\mathbf{I} + (\mathbf{R}_\mathbf{S} \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c})^{-1}(\mathbf{R}_\mathbf{S} \otimes \mathbf{B}^{-1}), \tag{66}$$

then

$$\mathbf{E}\mathbf{R}_{\mathbf{H}_c} = (\mathbf{I} + \mathbf{R}_\mathrm{SC})^{-1}\mathbf{R}_\mathrm{SC}. \tag{67}$$

It is noted that $\mathbf{E}\mathbf{R}_{\mathbf{H}_c}$ is a Hermitian matrix under the aforementioned assumption. Substituting (66) into (12)-(14), we can see that $\mathbf{F}$ is the linear function with respect to $\mathbf{E}$. Because

$$(\mathbf{I} + \mathbf{M})^{-1}(\mathbf{I} + \mathbf{M}) = \mathbf{I}, \tag{68}$$

we have

$$(\mathbf{I} + \mathbf{M})^{-1}\mathbf{M} = \mathbf{I} - (\mathbf{I} + \mathbf{M})^{-1}. \tag{69}$$

Combining (67) and (69), we can obtain

$$\mathbf{E}\mathbf{R}_{\mathbf{H}_c} = \mathbf{I} - (\mathbf{I} + \mathbf{R}_\mathrm{SC})^{-1}. \tag{70}$$

Hence

$$(\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1} = \mathbf{I} + \mathbf{R}_\mathrm{SC}. \tag{71}$$

Because $\mathbf{R}_\mathrm{SC} \succcurlyeq \mathbf{0}$, we have

$$\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c} \succ 0. \tag{72}$$

From (71), it follows that

$$\mathrm{tr}((\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1}) = \mathrm{tr}(\mathbf{I}) + \mathrm{tr}(\mathbf{R}_\mathrm{SC}), \tag{73}$$

Using a well-known inequality in matrix theory, we have

$$\begin{aligned}\mathrm{tr}((\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1}) &\le M_\mathrm{t}M_\mathrm{r} + \eta_{\max}(\mathbf{R}_{\mathbf{H}_c})tr(\mathbf{R}_\mathbf{S} \otimes \mathbf{B}^{-1}), \\ &= \tau\end{aligned} \tag{74}$$

where

$$\tau = M_\mathrm{t}M_\mathrm{r} + LPtr(\mathbf{B}^{-1})\eta_{\max}(\mathbf{R}_{\mathbf{H}_c}), \tag{75}$$

and $\eta_{\max}(\mathbf{R}_{\mathbf{H}_c})$ is the largest eigenvalue of $\mathbf{R}_{\mathbf{H}_c}$.

With $\mathrm{tr}(\mathbf{ABC}) = \mathrm{vec}(\mathbf{A}^H)^H(\mathbf{I} \otimes \mathbf{B})\mathrm{vec}(\mathbf{C})$, we can obtain

$$\begin{aligned}\mathrm{tr}((\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1}) &= \mathrm{tr}(\mathbf{I}(\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1}I) \\ &= \mathrm{vec}(\mathbf{I}_{M_tM_r})^H(\mathbf{I}_{M_tM_r} \otimes (\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c})^{-1})\mathrm{vec}(\mathbf{I}_{M_tM_r}). \\ &= \mathrm{vec}(\mathbf{I}_{M_tM_r})^H(\mathbf{I}_{M_tM_r} \otimes (\mathbf{I} - \mathbf{E}\mathbf{R}_{\mathbf{H}_c}))^{-1}\mathrm{vec}(\mathbf{I}_{M_tM_r})\end{aligned} \tag{76}$$

Using Lemma 1, and (74)-(76), (23) follows immediately. In order to obtain (24), we rely on the following lemma.

Lemma 2

Let $\mathbf{A}$ *and* $\mathbf{B}$*be positive and non-negative definite Hermitian matrix, respectively. Then,* $\mathbf{AB} \succcurlyeq \mathbf{0}$ *if* $\mathbf{AB}$ *is a Hermitian matrix.*

Proof: According to the similarity property of the matrices [38], $\mathbf{AB}$ is similar to a Hermitian matrix $\mathbf{A}^{-1/2}\mathbf{ABA}^{1/2} = \mathbf{A}^{1/2}\mathbf{BA}^{1/2}$. Hence, if we can obtain $\mathbf{x}^H\mathbf{A}^{1/2}\mathbf{BA}^{1/2}\mathbf{x} \ge 0, \ \forall\mathbf{x}$, then $\mathbf{AB} \succcurlyeq \mathbf{0}$. Let $\mathbf{y} = \mathbf{A}^{1/2}\mathbf{x}$, then

$$\mathbf{x}^H\mathbf{A}^{1/2}\mathbf{BA}^{1/2}\mathbf{x} = \mathbf{y}^H\mathbf{By}. \tag{77}$$

Following the definition of the non-negative matrix [38], we have

$$\mathbf{x}^H\mathbf{A}^{1/2}\mathbf{BA}^{1/2}\mathbf{x} \ge 0. \tag{78}$$

Thus, $\mathbf{AB} \succcurlyeq \mathbf{0}$.

Following Lemma 2, it is obvious that

$$\mathbf{E}\mathbf{R}_{\mathbf{H}_c} = (\mathbf{I}_{M_tM_r} + \mathbf{R}_\mathrm{SC})^{-1}\mathbf{R}_\mathrm{SC} \succcurlyeq \mathbf{0}. \tag{79}$$

It is noted that $\mathbf{R}_\mathbf{sc}$ can be diagonalized by its eigenvalue decomposition, i.e.,

$$\mathbf{R}_\mathrm{SC} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^H, \tag{80}$$

where $\mathbf{u}$ is a unitary matrix and $\Sigma = \mathrm{diag}\{\lambda_1,\lambda_2,..., \lambda MtMr\}$ is a diagonal matrix with each diagonal entry given by a eigenvalue. With $\mathbf{R}_\mathbf{SC} \succcurlyeq \mathbf{0}$, we can obtain $\lambda_i \ge 0$, $i = 1,2,...,M_tM_r$. Then (79) can be rewritten as

$$\mathbf{E}\mathbf{R}_{\mathbf{H}_c} = \mathbf{U}(\mathbf{\Sigma} + \mathbf{I})^{-1}\mathbf{\Sigma}\mathbf{U}^H. \tag{81}$$

Denote the eigenvalue of $\mathbf{E}\mathbf{R}_{\mathbf{H}_c}$ by $\gamma_i(\mathbf{E}\mathbf{R}_{\mathbf{H}_c}), \ i = 1, 2, \cdots, M_\mathrm{t}M_\mathrm{r}$. From (81), $\gamma_i(\mathbf{E}\mathbf{R}_{\mathbf{H}_c})$ can be expressed as

$$\gamma_i(\mathbf{E}\mathbf{R}_{\mathbf{H}_c}) = \frac{\lambda_i}{\lambda_i + 1} \tag{82}$$

From (82), it is known that $\gamma_i(\mathbf{E}\mathbf{R}_{\mathbf{H}_c})$ increases monotonically with λ_i. Hence,

$$\gamma_{\max}(\mathbf{E}\mathbf{R}_{\mathbf{H}_c}) = \frac{\lambda_{\max}}{\lambda_{\max} + 1}, \tag{83}$$

where $\gamma_{\max}(\mathbf{ER}_{\mathbf{H}_c})$ and $\lambda_{\max}$ are the largest eigenvalues of $\mathbf{ER}_{\mathbf{H}_c}$ and $\mathrm{R}_{\mathbf{sc}}$, respectively. As discussion above, we have

$$\mathrm{tr}(\mathbf{R}_{\mathbf{SC}}) = \mathrm{tr}((\mathbf{R}_{\mathbf{S}} \otimes \mathbf{B}^{-1})\mathbf{R}_{\mathbf{H}_c}) \leq LP\eta_{\max}(\mathbf{R}_{\mathbf{H}_c})tr(\mathbf{B}^{-1}). \quad (84)$$

Because $\mathbf{R}_{\mathrm{SC}} \succcurlyeq \mathbf{0}$, then

$$\lambda_{\max} \leq \mathrm{tr}(\mathbf{R}_{\mathbf{SC}}). \quad (85)$$

With (83)-(85), we have

$$\gamma_{\max}(\mathbf{ER}_{\mathbf{H}_c}) \leq \beta, \quad (86)$$

where

$$\beta = \frac{LP\eta_{\max}(\mathbf{R}_{\mathbf{H}_c})tr(\mathbf{B}^{-1})}{LP\eta_{\max}(\mathbf{R}_{\mathbf{H}_c})tr(\mathbf{B}^{-1}) + 1}. \quad (87)$$

Then, we have

$$\mathbf{ER}_{\mathbf{H}_c} \preccurlyeq \beta\mathbf{I}. \quad (88)$$

By combining (79) and (88), (24) follows immediately.

Appendix D

Proof of proposition 2

Because $\mathrm{tr}(\mathbf{ABCD}) = \mathrm{vec}(\mathbf{A}^H)^H(\mathbf{D}^H \otimes \mathbf{B})\mathrm{vec}(\mathbf{C})$, then

$$\begin{aligned}
&\mathrm{tr}((\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) \\
&= \mathrm{tr}((\mathbf{I}+\mathbf{D})\mathbf{U}(\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1}\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H\mathbf{I}) \\
&= (\mathrm{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H))^H(\mathbf{I} \otimes (\mathbf{U}^H\mathbf{F}\mathbf{U})^{-1})\mathrm{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H) \\
&= (\mathrm{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H))^H(\mathbf{I} \otimes (\mathbf{U}^H\mathbf{F}\mathbf{U}))^{-1}\mathrm{vec}(\mathbf{U}^H(\mathbf{I}+\mathbf{D})^H)
\end{aligned} \quad (89)$$

Evidently,

$$\mathrm{tr}(\mathbf{D}^H\mathbf{D}\mathbf{M}) = \mathrm{tr}(\mathbf{D}^H\mathbf{D}\mathbf{M}^{1/2}\mathbf{M}^{1/2}), \quad (90)$$

where $\mathbf{M}^{1/2}$ is the square root of $\mathbf{M}$ [38]. With tr $(\mathbf{ABC}) = \mathrm{tr}(\mathbf{CAB})$, (90) can be rewritten as

$$\mathrm{tr}(\mathbf{D}^H\mathbf{D}\mathbf{M}) = \mathrm{tr}(\mathbf{M}^{1/2}\mathbf{D}^H\mathbf{D}\mathbf{M}^{1/2}). \quad (91)$$

With Lemma 1, (89) and (91), we can obtain (27) and (28).

Abbreviations

AINR: array interference-to-noise ratio; ASNR: array signal-to-noise ratio; CNR: clutter-to-noise ratio; CPI: coherent processing interval; CRB: Cramer-Rao bound; FIM: Fisher information matrix; LMI: linear matrix inequality; LS: least squares; MIMO: multi-input multi-output; MMSE: minimum mean square error; RCS: radar cross sections; SDP: semidefinite programming; WCM: waveform covariance matrix

Acknowledgements

The authors would like to thank Dr. Magnus Jansson and the anonymous reviewers for their thoughtful and to-the-point comments and suggestions which greatly improved the manuscript. This study is sponsored in part by NSFC under Grant 60825104, Program for Changjiang Scholars and Innovative Research Team in University under Grant IRT0954, and the Major State Basic Research Development Program of China (973 Program) under Grant 2010CB731903, 2011CB707001.

Author details

[1]National Key Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China [2]School of Science, Xidian University, Xi'an 710071, China

Competing interests

The authors declare that they have no competing interests.

Received: 25 January 2011 Accepted: 29 June 2011
Published: 29 June 2011

References

1. E Fishler, A Haimovich, R Blum, D Chizhik, L Cimini, R Valenzuela, MIMO radar: an idea whose time has come. in *Proceedings of the IEEE Radar Conference* Newark, NJ, USA, 26-29, April 2004, p. 71
2. E Fishler, A Haimovich, R Blum, L Cimini, D Chizhik, R Valenzuela, Spatial diversity in radars-models and detection performance. IEEE Trans Signal Process. **54**(3), 823–838 (2006)
3. J Li, P Stoica, MIMO radar with colocated antennas. IEEE Signal Process Mag. **24**(5), 106–114 (2007)
4. J Li, P Stoica, L Xu, W Roberts, On parameter identifiability of MIMO radar. IEEE Signal Process Lett. **14**(12), 968–971 (2007)
5. H Yan, J Li, G Liao, Multitarget identification and localization using bistatic MIMO radar systems. EURASIP J. Adv Signal Process. **2008**, Article ID 283483 (2008)
6. S Peter, J Li, Y Xie, On probing signal design for MIMO radar. IEEE Trans. Signal Process. **55**(8), 4151–4161 (2007)
7. J Li, L Xu, P Stoica, KW Forsythe, DW Bliss, Range compression and waveform optimization for MIMO radar: a Cramer-Rao bound based study. IEEE Trans Signal Process. **55**(8), 4151–4161 (2007)
8. DR Fuhrmann, GS Antonio, Transmit beamforming for MIMO radar systems using signal cross-correlation. IEEE Trans Aerosp Electron Syst. **44**(1), 171–186 (2008)
9. J Li, L Xu, P Stoica, Z Xiayu, Signal synthesis and receiver design for MIMO radar imaging. IEEE Trans Signal Process. **56**(8), 3959–3968 (2008)
10. CY Chen, PP Vaidyanathan, MIMO radar ambiguity properties and optimization using frequency-hopping waveforms. IEEE Trans Signal Process. **56**(12), 5926–5936 (2008)
11. B Liu, Z He, J Zeng, BY Liu, Polyphase Orthogonal Code Design for MIMO Radar Systems. in *International Conference on Radar* (2006), p. 1
12. B Liu, Z He, Q He, Optimization of Orthogonal Discrete Frequency-Coding Waveform Based on Modified Genetic Algorithm for MIMO Radar. in *International Conference on Communication, Circuits, and Systems* (2007), p. 966
13. B Friedlander, Waveform design for MIMO radars. IEEE Trans Aerosp Electron Syst. **43**(3), 1227–1238 (2007)
14. CY Chen, PP Vaidyanathan, MIMO radar waveform optimization with prior information of the extended target and clutter. IEEE Trans Signal Process. **57**(9), 3533–3544 (2009)
15. MR Bell, Information theory and radar waveform design. IEEE Trans Inf Theory. **39**(5), 1578–1597 (1993)
16. Y Yang, R Blum, MIMO radar waveform design based on mutual information and minimum mean-square error estimation. IEEE Trans Aerosp Electron Syst. **43**(1), 330–343 (2007)
17. Y Yang, R Blum, Minimax robust MIMO radar waveform design. IEEE J Sel Top Signal Process. **1**(1), 147–155 (2007)
18. A Leshem, O Naparstek, A Nehorai, Information theoretic adaptive radar waveform design for multiple extended targets. IEEE J Sel Topics Signal Process. **1**(1), 42–55 (2007)
19. T Naghibi, M Namvar, F Behnia, Optimal and robust waveform design for MIMO radars in the presence of clutter. Signal Process. **90**(4), 1103–1117 (2010)
20. JD Gorman, AO Hero, Lower bounds for parametric estimation with constraints. IEEE Trans Inf Theory. **26**(6), 1285–1301 (1990)
21. M Brian Sadler, J Richard Kozick, T Moore, Bounds on bearing and symbol estimation with side information. IEEE Trans Signal Process. **49**(4), 822–834 (2001)
22. P Stoica, C Ng Boon, On the Cramer-Rao bound under parametric constraints. IEEE Signal Process Lett. **5**(7), 177–179 (1998)
23. AO Hero, A Cramer–Rao, *Type Lower Bound for Essentially Unbiased Parameter Estimation* (Lincoln Lab., Mass. Inst. Technol., Lexington, MA, 1992). Tech. Rep. 890 DTIC AD-A246666
24. AO Hero, JA Fessler, M Usman, Exploring estimator bias-variance tradeoffs using the uniform CR bound. IEEE Trans Signal Process. **44**, 2026–2041 (1996)

25. C Eldar Yonina, BT Aharon, N Arkadi, Robust mean-squared error estimation in the presence of model uncertainties. IEEE Trans Signal Process. **53**(1), 168–181 (2005)
26. C Eldar Yonina, AV Oppenheim, Covariance shaping least-squares estimation. IEEE Trans Signal Process. **51**(3), 686–697 (2003)
27. C Eldar Yonina, Minimax MSE estimation of deterministic parameters with noise covariance uncertainties. IEEE Trans Signal Process. **54**(1), 138–145 (2006)
28. C Eldar Yonina, Minimum variance in biased estimation: bounds and asymptotically optimal estimators. IEEE Trans Signal Process. **52**(7), 1915–1930 (2004)
29. BH Zvika, C Eldar Yonina, On the constrained Cramér-Rao bound with a singular fisher information matrix. IEEE Signal Process Lett. **16**(6), 453–456 (2009)
30. A Ben-Tal, A Nemirovski, *Lectures on Modern Convex Optimization, Series on Optimization* (MPS-SIAM, Philadelphia, 2001)
31. L Vandenberghe, S Boyd, Semidefinite programming. SIAM Rev. **38**(1), 40–95 (1996)
32. J Lofberg, YALMIP, A toolbox for modeling and optimization in MATLAB. in *Proceedings of the CACSD Conference* (Taipei, Taiwan, 2004), p. 284
33. CY Chen, PP Vaidyanathan, MIMO radar space-time adaptive processing using prolate spheroidal wave functions. IEEE Trans Signal Process. **56**(2), 623–635 (2008)
34. G Wang, Y Lu, Clutter rank of STAP in MIMO radar with waveform diversity. IEEE Trans Signal Process. **58**(2), 938–943 (2010)
35. J Ward, *Space-Time Adaptive Processing for Airborne Radar MIT* (Lincoln Lab., Lexington, 1994). Tech. Rep. 1015
36. P Stoica, RL Moses, *Spectral Analysis of Signals* (Prentice-Hall, Englewood Cliffs, 2005)
37. RA Horn, CR Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985)
38. HL ütkepohl, *Handbook of Matrices* (Wiley, New York, 1996)
39. P Stoica, J Li, X Zhu, Waveform synthesis for diversity-based transmit beam pattern design. IEEE Trans Signal Process. **56**(6), 2593–2598 (2008)
40. L Xu, J Li, P Stoica, Target detection and parameter estimation for MIMO radar systems. IEEE Trans Aerosp Electron Syst. **44**(3), 927–939 (2008)

doi:10.1186/1687-6180-2011-15
Cite this article as: Wang *et al.*: **Joint optimization of MIMO radar waveform and biased estimator with prior information in the presence of clutter.** *EURASIP Journal on Advances in Signal Processing* 2011 **2011**:15.