An Information Theoretic Approach to Robust Constrained Code Design for MIMO Radars

Mohammad Mahdi Naghsh*, Mahmoud Modarres-Hashemi, Mohammad Alaee, and Ehsan Haj Mirza Alian

Abstract

The problem of transmit code design in multiple-input multiple-output (MIMO) radar systems is addressed in this paper. The problem is considered in both colocated and widely separated antenna radars in the presence of signal-dependent interference and target mobility. Due to complexity of exact expressions of the optimal detector in the MIMO radars, information theoretic criteria are employed as design metrics which results in a general form of non-convex optimization problems. In order to tackle the problem, a novel technique based on the minorization-maximization (MaMi) algorithm is proposed and solutions are presented under practical constraints on the transmit code, namely energy constraint, peak-to-average-power ratio (PAR) constraint, and similarity constraint. Furthermore, the devised method is extended to be robust against uncertainties of the clutter and interference statistics. Finally, numerical examples are used to show the performance of the proposed technique in different situations.

Index Terms

Code design, information theoretic criteria, MIMO radar, minorization-maximization, peak-to-average power ratio, robust

I. Introduction

Transmit code design is an important design challenge in both single-antenna and multiple-antenna radar systems, as it is shown to have a significant impact on the performance of such systems [1]. In particular, multiple-input multiple-output (MIMO) radars, that employ multiple antennas at both the transmitter

The authors are with the Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan 84156-83111, Iran. *Please address all the correspondence to M. M. Naghsh, Email: mm_naghsh@cc.iut.ac.ir

and receiver side, have attracted much interest due to their flexibility of using different waveforms and adaptivity for adjusting them to optimize desired performance criteria.

In general, MIMO radars have shown a better detection performance, more accurate estimation of target parameters, and better resolutions compared to single-antenna radars [2]. There are two well-known structures for such systems, namely *colocated* and *widely separated* antenna MIMO radars. In the latter, the distances between transmit/receive antennas are much larger than the wavelength of transmitted signals. This leads to statistical independence of the reflected signals from the target which provides angular diversity and improves detection performance [3]. On the other hand, in colocated radars, the distances between the transmit/receive antennas are about half of the wavelength of the transmitted signals. In such systems, in contrast to phased-array radars, different waveforms can be transmitted simultaneously from different transmit antennas. This property is referred to as waveform diversity and provides better detection performance, better interference rejection, and more flexibility in generating radiation patterns compared to phased-array radars [2].

Based on the above-mentioned properties, transmit waveform (code) design plays an important role in determining the performance of both colocated and widely separated MIMO radars, and thus, much work has been done in this area (see e.g., [2] and references therein). It is worth noting that the waveform design in such systems depends on several parameters such as mobility of targets, the effect of the signal-dependent interference (clutter), and practical limitations (e.g., peak-to-average power ratio (PAR) considerations). Moreover, depending on the desired system, different design criteria might be used, e.g., criteria related to detection, estimation, classification, etc. [2].

Several works in the last decades have considered the waveform design for improvement of the detection performance in single-antenna radars. However, most of the works are based on simplifying assumptions such as neglecting the effect of the target mobility (i.e. Doppler shift), treating the clutter as a signal-independent interference, and/or assuming perfect a priori knowledge of the statistics of interference [4, 5]. In some other works, clutter is modeled as a wide sense stationary Gaussian process or as a response to a linear time-invariant system for the sake of simplicity [6]. In [7], the problem of code design for single-antenna radars in the presence of clutter and target Doppler shift is addressed; then, signal-to-interference-plus-noise ratio (SINR) at the output of the receiver is maximized as a design criterion (see also [8]).

In multiple-antenna radars, the problem of code design has been addressed for both colocated [1, 9] and widely separated radars [10–13]. However, simplifying assumptions have been made especially on the target Doppler shift, clutter nature and a priori knowledge about its statistics, as well as practical limitations (see e.g., [2, 14, 15] and references therein). Due to complex expressions of exact detection performance metrics in MIMO radars (if any, [16]), information theoretic criteria [13, 17] or SINR [9, 18] have been employed as design metrics. Note that in several situations, maximizing information theoretic criteria can be theoretically justified while the maximization of SINR can not necessarily [15, 17, 19].

Another important issue in code design problems is practical limitations. As an important example, in many applications, the sought code is supposed to be constant-amplitude/PAR-limited. However, this fact is usually not considered [18] or is partially cared about. In other words, in several works, the code design problem is dealt with no PAR constraint and then, a PAR-constrained code is synthesized from the solution of the unconstrained problem. This procedure is associated with a significant performance loss (see e.g., [6, 18, 20]).

In this paper, we consider the problem of transmit code design for both colocated and widely separated MIMO radar systems in the presence of clutter. The aim is to improve the detection performance of a moving target while dealing with practical/implementation limitations as well as uncertainties in a priori knowledge of the interference. To this end, we employ an information-theoretic approach and cast the problem of code design via maximization of information-theoretic criteria; namely, *mutual information* and *J-divergence*. To account for the mostly common used practical/implementation limitations in the radar signal design literature, energy, PAR, and similarity constraints are imposed to the design problems. The constrained design problems are non-convex but all have a structure with respect to (w.r.t.) positive semidefinite (psd) matrices (which can be assumed to be related to the SINR of the optimal detector) and can be exploited to tackle the problems. Therefore, we devise a general method to obtain quality solutions to the design problems associated with the design metrics. The method is based on applying *minorization-maximization* technique to the functions of scalars/psd matrices and leads to obtain stationary points of the problems under some mild conditions. We further robustify the design method to handle uncertainties w.r.t. a priori knowledge of the clutter and (signal-independent) interference, which leads to multi-objective optimization problems. To the best of our knowledge, no information-theoretic code design methodology is addressed in the literature for robust constrained code design. The previous methods mostly design the code via a relaxation of a constraint and then, synthesize the code to satisfy the desired constraint. Moreover, they usually assume a perfect knowledge of the interference [9, 20].

The rest of this paper is organized as follows. In Section II, the signal and system model for the moving target in the presence of clutter is introduced which incorporates both colocated and widely separated MIMO systems. This section also presents the optimal detector associated with the model. We cast the constrained code design problems for the design metrics in Section III. This section also includes derivations of the proposed method to deal with the design problems via an iterative solving of quadratically constrained quadratic programs (QCQPs). Solving the QCQPs subject to the desired constraints is discussed in Section IV. In Section V, the design methodology is extended to be robust against interference uncertainties. Numerical results are provided in Section VI to illustrate the performance of the proposed method. Finally, Section VII concludes the paper.

Notation: Bold lowercase letters and bold uppercase letters are used for vectors and matrices respectively. $\boldsymbol{I}_N$ represents the identity matrix in $\mathbb{C}^{N\times N}$. $\mathbf{1}$ and $\mathbf{0}$ are the all-one and the all-zero vectors/matrices. The Frobenius norm of a matrix $\boldsymbol{X}$ is denoted by $\|\boldsymbol{X}\|_F^2$ whereas the spectral norm of $\boldsymbol{X}$ is denoted by $\|\boldsymbol{X}\|_2$. The l_2-norm of a vector $\boldsymbol{x}$ is denoted by $\|\boldsymbol{x}\|_2^2$. We show vector/matrix transpose by $(\cdot)^T$, the complex conjugate by $(\cdot)^*$, and the Hermitian by $(\cdot)^H$. The symbol $\odot$ stands for element-wise Hadamard product of matrices/vectors. tr$(\cdot)$ is the trace of a square matrix. The notations $\lambda_{max}(\cdot)$ and $\lambda_{min}(\cdot)$ indicate the principal and the minor eigenvalues of a Hermitian matrix, respectively. blkDiag(.) denotes the block diagonal matrix formed by its arguments. vec$(\boldsymbol{X})$ denotes the column-wise stacking of the elements of matrix $\boldsymbol{X}$. $\mathbb{E}\{\cdot\}$ and $\Re(\cdot)$ stand for the statistical expectation and real-part operators respectively. The symbol $\mathcal{CN}(\omega, \boldsymbol{\Sigma})$ denotes the circularly symmetric complex Gaussian distribution with mean ω and covariance $\boldsymbol{\Sigma}$. $\mathbb{R}_+$ represents non-negative real numbers. The notation $\boldsymbol{A} \succ \boldsymbol{B}$ ($\boldsymbol{A} \succeq \boldsymbol{B}$) implies $\boldsymbol{A} - \boldsymbol{B}$ is positive (semi)definite. Finally, $\mathcal{S}_+^N$ denotes the psd cone in $\mathbb{C}^{N\times N}$.

II. Preliminaries

A. Signal and System Model

We consider a MIMO radar system with N_T transmitter and N_R receiver antennas. Let $\boldsymbol{a}_m = [a_m(1), \ldots, a_m(N)]^T \in \mathbb{C}^N$ denote the transmit signal (code) from the mth transmit antenna with N being the length of the code.

1) Colocated MIMO Radars: In the considered colocated MIMO radar, the distance among transmitter antennas is assumed to be sufficiently small such that the reflected signals from a target are correlated across the array and hence the whole array see the target at the direction θ_0 with the reflection coefficient α_0. Moreover, it is supposed that L clutter patches located at $\theta_1, \cdots, \theta_L$ interfere with the target detection at the cell under test. Therefore, $\widetilde{\boldsymbol{r}}(n)$, the $N_R \times 1$ vector of the received samples for the stationary target at the cell under test at the time index n is given by [9][1]:

$$\widetilde{\boldsymbol{r}}(n) = \ \alpha_0 \boldsymbol{s}_r(\theta_0) \boldsymbol{s}_t^T(\theta_0) \boldsymbol{s}_{co}(n)$$

[1]A single target is assumed at the cell under test [2, 10] but the radar can deal with multi-target in the detection range.

$$+\underbrace{\sum_{l=1}^{L}\alpha_l \boldsymbol{s}_r(\theta_l)\boldsymbol{s}_t^T(\theta_l)\boldsymbol{s}_{co}(n)}_{\widetilde{\boldsymbol{c}}(n)}+\widetilde{\boldsymbol{v}}(n),\ n=1,\cdots,N, \tag{1}$$

where $\boldsymbol{s}_r(\theta)$,$\boldsymbol{s}_t(\theta)$, $\widetilde{\boldsymbol{c}}(n)$, and $\widetilde{\boldsymbol{v}}(n)$ are the receive array steering vector[2], transmit array steering vector, clutter, and (signal-independent) interference, respectively. Herein, $\boldsymbol{s}_{co}(n) = [a_1(n),\cdots,a_{N_T}(n)]^T \in \mathbb{C}^{N_T}$ denotes the transmit signal vector of the system at the time index n. Moreover, α_0 and α_l denote the reflected coefficients associated with the radar cross section (RCS) as well as the propagation effects of the target and the lth interference source, respectively. Then, considering target Doppler shift and rearranging $\boldsymbol{X} = [\widetilde{\boldsymbol{r}}(1),\cdots,\widetilde{\boldsymbol{r}}(N)]^T$ and $\boldsymbol{V} = [\widetilde{\boldsymbol{v}}(1),\cdots,\widetilde{\boldsymbol{v}}(N)]^T$, the model in (1) is modified as

$$\boldsymbol{X} = \underbrace{\alpha_0(\boldsymbol{A}\odot\boldsymbol{P})\boldsymbol{\Psi}(\theta_0)}_{\boldsymbol{S}}+\boldsymbol{A}\sum_{l=1}^{L}\alpha_l\boldsymbol{\Psi}(\theta_l)+\boldsymbol{V}, \tag{2}$$

where $\boldsymbol{X}\in\mathbb{C}^{N\times N_R}$ is the received signal matrix from the target at the cell under test contaminated with the clutter component $\boldsymbol{C} = \boldsymbol{A}\sum_{l=1}^{L}\alpha_l\boldsymbol{\Psi}(\theta_l)$ as well as the interference component $\boldsymbol{V}$. Herein, $\boldsymbol{\Psi}(\theta)$ is the steering matrix corresponding to the look angle θ given by $\boldsymbol{\Psi}(\theta) = \boldsymbol{s}_t(\theta)\boldsymbol{s}_r^T(\theta)\in\mathbb{C}^{N_T\times N_R}$ and $\boldsymbol{A}\in\mathbb{C}^{N\times N_T}$ is the transmit code matrix (to be designed) with $\boldsymbol{A}_{n,m} = a_m(n)$. Note that according to the definitions of $\boldsymbol{a}_m$ and $\boldsymbol{s}_{co}(n)$, we have $\boldsymbol{A} = \Big[\boldsymbol{a}_1,\boldsymbol{a}_2,\boldsymbol{a}_3,\ldots,\boldsymbol{a}_{N_t}\Big] = \Big[\boldsymbol{s}_{co}(1),\boldsymbol{s}_{co}(2),\boldsymbol{s}_{co}(3),\ldots,\boldsymbol{s}_{co}(N)\Big]^T$. Also, the matrix $\boldsymbol{P} = [\boldsymbol{p}_d,...,\boldsymbol{p}_d]\in\mathbb{C}^{N\times N_T}$ is the temporal steering matrix associated with target Doppler shift with $\boldsymbol{p}_d = [1,\exp(j2\pi f_d),...,\exp(j2\pi f_d(N-1))]^T$, where f_d is the normalized Doppler shift associated with the moving target.

The covariance matrix of the signal component $\boldsymbol{S}$ can be found using Kronecker properties [21] as

$$\begin{aligned}\boldsymbol{R}_s =& \mathbb{E}\Big\{\text{vec}\Big(\alpha_0(\boldsymbol{A}\odot\boldsymbol{P})\boldsymbol{\Psi}(\theta_0)\Big)\text{vec}\Big(\alpha_0(\boldsymbol{A}\odot\boldsymbol{P})\boldsymbol{\Psi}(\theta_0)\Big)^H\Big\}\\ =&\big(\boldsymbol{I}_{N_R}\otimes[\boldsymbol{A}\odot\boldsymbol{P}]\big)\boldsymbol{T}\left(\boldsymbol{I}_{N_R}\otimes[\boldsymbol{A}\odot\boldsymbol{P}]\right)^H\in\mathcal{S}_+^{NN_R}\end{aligned} \tag{3}$$

where $\boldsymbol{T} = \sigma_s^2\text{vec}\left(\boldsymbol{\Psi}(\theta_0)\right)\text{vec}\left(\boldsymbol{\Psi}(\theta_0)\right)^H \triangleq \boldsymbol{b}\boldsymbol{b}^H$ and $\sigma_s^2 = \mathbb{E}\{\alpha_0\alpha_0^*\}$. Following the same procedure, for the covariance matrix of the clutter component $\boldsymbol{C}$ we have

$$\boldsymbol{R}_c = \left(\boldsymbol{I}_{N_R}\otimes\boldsymbol{A}\right)\boldsymbol{Q}\left(\boldsymbol{I}_{N_R}\otimes\boldsymbol{A}\right)^H, \tag{4}$$

[2]Assuming uniformly spaced linear arrays, the vector $\boldsymbol{s}_t(\theta)$ is given by $\boldsymbol{s}_t(\theta) = [1,\exp(j\frac{2\pi f_c}{c}d_t\sin(\theta)),\cdots,\exp(j\frac{2\pi f_c}{c}d_t\sin(\theta)(N_T-1))]^T$, where c is the wave (light) speed, f_c is the carrier frequency and d_t denotes the inter-element spacing between transmit antennas. Note that a similar equation holds for the receive steering vector $\boldsymbol{s}_r(\theta)$.

where $\boldsymbol{Q} = \sum_{l=1}^{L} \sigma_{c,l}^2 \text{vec}(\boldsymbol{\Psi}(\theta_l))\text{vec}(\boldsymbol{\Psi}(\theta_l))^H$ and $\sigma_{c,l}^2 = \mathbb{E}\{\alpha_l \alpha_l^*\}$, and finally,

$$\boldsymbol{R}_{\boldsymbol{v}} = \mathbb{E}\{\text{vec}(\boldsymbol{V})\text{vec}(\boldsymbol{V})^H\}. \tag{5}$$

It is worth noting that with the assumption of zero Doppler shift, the model is compatible to that of [14], and by neglecting the clutter, the model is simplified to the one in [22] (see also [23]).

2) Widely Separated MIMO Radars: In widely separated MIMO radar systems, the distance among array elements is assumed to be sufficiently large such that the (possible) target or clutter provides uncorrelated reflection coefficients between each pair of radar transmitter and receiver [3]. Therefore, $\boldsymbol{r}_k$, the reflected signal (vector) at the kth receiver can be modeled in the discrete-time domain at the cell under test as [2, 15]

$$\boldsymbol{r}_k = \boldsymbol{s}_{wi,k} + \boldsymbol{c}_k + \boldsymbol{v}_k, \quad k = 1, 2, \cdots, N_R, \tag{6}$$

where $\boldsymbol{r}_k \in \mathbb{C}^N$, $\boldsymbol{v}_k$ is the signal-independent interference at the kth receiver, and $\boldsymbol{s}_{wi,k}$ is the received signal from the target which can be written as

$$\boldsymbol{s}_{wi,k} = \sum_{m=1}^{N_T} (\alpha_{m,k} \boldsymbol{a}_m \odot \boldsymbol{p}_{m,k}), \quad k = 1, \cdots, N_R, \tag{7}$$

with $\alpha_{m,k}$ being the reflection coefficient accounting for the target RCS and the propagation effects associated with the mth transmitter and the kth receiver. In (7), $\boldsymbol{p}_{m,k}$ is the temporal steering vector from the mth transmitter to the kth receiver with $\boldsymbol{p}_{m,k} = [1, \exp(j2\pi f_{d_{m,k}}), \cdots, \exp(j2\pi f_{d_{m,k}}(N-1))]^T$, in which $f_{d_{m,k}}$ is the normalized Doppler shift associated with emitted signal from the mth transmit antenna and captured at the kth receive antenna (which also depends on the angle between the target and the antennas). In widely separated antenna radar systems, there are different reflection coefficients for the target at the kth receiver denoted by $\boldsymbol{\alpha}_k = [\alpha_{1,k}, \cdots, \alpha_{N_T,k}]^T \in \mathbb{C}^{N_T}$. Therefore, the echo of the moving target at the kth receiver can be written as

$$\boldsymbol{s}_{wi,k} = (\boldsymbol{A} \odot \boldsymbol{P}_k)\boldsymbol{\alpha}_k, \quad k = 1, \cdots, N_R, \tag{8}$$

with $\boldsymbol{P}_k = [\boldsymbol{p}_{1,k}, \cdots, \boldsymbol{p}_{N_T,k}] \in \mathbb{C}^{N \times N_T}$ and $\boldsymbol{A}$ is the matrix code as in the colocated MIMO case. Furthermore, in (6), $\boldsymbol{c}_k$ can be expressed as $\boldsymbol{c}_k = \boldsymbol{A}\boldsymbol{\beta}_k$, where $\boldsymbol{\beta}_k \in \mathbb{C}^{N_T}$ is the vector associated with clutter component at the kth receiver antenna [14, 15]. Thus, we can reformulate (6) as[3]

$$\boldsymbol{r}_k = (\boldsymbol{A} \odot \boldsymbol{P}_k)\boldsymbol{\alpha}_k + \boldsymbol{A}\boldsymbol{\beta}_k + \boldsymbol{v}_k, \quad k = 1, 2, \cdots, N_R. \tag{9}$$

[3]The presented model for the colocated and widely separated MIMO radars addresses an ideal beam pattern for the antennas; however, the effects of the imperfect beampatters can be incorporated into the model (see e.g., [7] and references therein for details).

Similar to the colocated case, we can define $\boldsymbol{X} = [\boldsymbol{r}_1, \boldsymbol{r}_2, ..., \boldsymbol{r}_{N_R}] \in \mathbb{C}^{N \times N_R}$ and rewrite (9) as

$$\boldsymbol{X} = \boldsymbol{S} + \boldsymbol{C} + \boldsymbol{V}, \tag{10}$$

where $\boldsymbol{S} = [(\boldsymbol{A} \odot \boldsymbol{P}_1)\boldsymbol{\alpha}_1, ..., (\boldsymbol{A} \odot \boldsymbol{P}_{N_R})\boldsymbol{\alpha}_{N_R}]$, $\boldsymbol{C} = \boldsymbol{A}[\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, ..., \boldsymbol{\beta}_{N_R}]$, and $\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_{N_R}]$.

It should be noted that the reflection coefficients at various receivers are uncorrelated in widely separated MIMO radars and hence, the covariance matrices can be obtained as[4]

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{s}} =& \text{blkDiag}\big[\mathbb{E}\{[(\boldsymbol{A} \odot \boldsymbol{P}_1)\boldsymbol{\alpha}_1][(\boldsymbol{A} \odot \boldsymbol{P}_1)\boldsymbol{\alpha}_1]^H\}, \cdots, \\
& \qquad \mathbb{E}\{[(\boldsymbol{A} \odot \boldsymbol{P}_{N_R})\boldsymbol{\alpha}_{N_R}][(\boldsymbol{A} \odot \boldsymbol{P}_{N_R})\boldsymbol{\alpha}_{N_R}]^H\}\big] \\
=& \text{blkDiag}\big[(\boldsymbol{A} \odot \boldsymbol{P}_1)\boldsymbol{R}_{\boldsymbol{s}_1}(\boldsymbol{A} \odot \boldsymbol{P}_1)^H, \cdots, \\
& \qquad (\boldsymbol{A} \odot \boldsymbol{P}_{N_R})\boldsymbol{R}_{\boldsymbol{s}_{N_R}}(\boldsymbol{A} \odot \boldsymbol{P}_{N_R})^H\big],
\end{aligned} \tag{11}$$

in which $\boldsymbol{R}_{\boldsymbol{s}_k} = \mathbb{E}[\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k^H] \in \mathcal{S}_+^{N_T}$. Similarly, we have

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{c}} =& \text{blkDiag}(\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{c}_1}\boldsymbol{A}^H, \ldots, \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{c}_{N_R}}\boldsymbol{A}^H), \\
\boldsymbol{R}_{\boldsymbol{v}} =& \text{blkDiag}(\boldsymbol{R}_{\boldsymbol{v}_1}, \ldots, \boldsymbol{R}_{\boldsymbol{v}_{N_R}}),
\end{aligned} \tag{12}$$

where $\boldsymbol{R}_{\boldsymbol{c}_k} = \mathbb{E}[\boldsymbol{\beta}_k \boldsymbol{\beta}_k^H]$ and $\boldsymbol{R}_{\boldsymbol{v}_k} = \mathbb{E}[\boldsymbol{v}_k \boldsymbol{v}_k^H]$.

B. Optimal Detector

In both cases of the colocated and the widely separated MIMO radars, the target detection problem leads to the following binary hypothesis test

$$\begin{cases} \mathrm{H}_0: & \boldsymbol{x} = \boldsymbol{c} + \boldsymbol{v} \\ \mathrm{H}_1: & \boldsymbol{x} = \boldsymbol{s} + \boldsymbol{c} + \boldsymbol{v} \end{cases} \tag{13}$$

where $\boldsymbol{x} = \text{vec}(\boldsymbol{X}) \in \mathbb{C}^{NN_R}$, $\boldsymbol{s} = \text{vec}(\boldsymbol{S})$, $\boldsymbol{c} = \text{vec}(\boldsymbol{C})$, and $\boldsymbol{v} = \text{vec}(\boldsymbol{V})$ are assumed to be Gaussian random vectors [2, 6]. Defining $\boldsymbol{y} \triangleq \boldsymbol{D}^{-1/2}\boldsymbol{x}$ with $\boldsymbol{D} = \boldsymbol{R}_c + \boldsymbol{R}_{\boldsymbol{v}}$, the underlying detection problem is equivalently expressed as [24]

$$\begin{cases} \mathrm{H}_0: \boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{NN_R}) \\ \mathrm{H}_1: \boldsymbol{y} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_{NN_R} + \boldsymbol{F}) \end{cases} \tag{14}$$

[4]Note that the devised methodology in this paper can also be applied to the cases for which the aforementioned assumption does not hold.

where $\boldsymbol{F} = \boldsymbol{D}^{-\frac{1}{2}} \boldsymbol{R}_s \boldsymbol{D}^{-\frac{1}{2}} = \widetilde{\boldsymbol{U}} \boldsymbol{\Lambda} \widetilde{\boldsymbol{U}}^H$ with $\boldsymbol{\Lambda}$ being a diagonal matrix containing eigenvalues of $\boldsymbol{F}$ and columns of $\widetilde{\boldsymbol{U}}$ are the associated eigenvectors. Therefore, the canonical form of the optimal Neyman-Pearson (NP) detector is given by [24]

$$\sum_{n=1}^{NN_R} \frac{\lambda_n |z_n|^2}{1+\lambda_n} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta, \tag{15}$$

where λ_n are eigenvalues of $\boldsymbol{F}$, η is the detection threshold, and $\boldsymbol{z} = [z_1, \cdots, z_{NN_R}]^T = \widetilde{\boldsymbol{U}}^H \boldsymbol{y}$.

In the case of widely separated MIMO radars, the matrix $\boldsymbol{F}$ is block-diagonal with blocks

$$\{\widetilde{\boldsymbol{F}}_k\}_{k=1}^{N_R} = \{\boldsymbol{D}_k^{-1/2} (\boldsymbol{A} \odot \boldsymbol{P}_k) \boldsymbol{R}_{\boldsymbol{s}_k} (\boldsymbol{A} \odot \boldsymbol{P}_k)^H \boldsymbol{D}_k^{-1/2}\}_{k=1}^{N_R},$$

where $\boldsymbol{D}_k = \boldsymbol{A} \boldsymbol{R}_{\boldsymbol{c}_k} \boldsymbol{A}^H + \boldsymbol{R}_{\boldsymbol{v}_k}$ (see (11) and (12)). Now, letting $\overline{\boldsymbol{U}}_k \widetilde{\boldsymbol{\Lambda}}_k \overline{\boldsymbol{U}}_k^H$ denote the eigen-decomposition of $\widetilde{\boldsymbol{F}}_k$, the canonical form of the optimal NP detector becomes

$$\sum_{k=1}^{N_R} \sum_{n=1}^{N} \frac{\widetilde{\lambda}_{n,k} |\widetilde{z}_{n,k}|^2}{1+\widetilde{\lambda}_{n,k}} \underset{\mathrm{H}_0}{\overset{\mathrm{H}_1}{\gtrless}} \eta, \tag{16}$$

where $\widetilde{\boldsymbol{z}}_k = [\widetilde{z}_{1,k}, \cdots, \widetilde{z}_{N,k}]^T = \overline{\boldsymbol{U}}_k^H \boldsymbol{D}_k^{-1/2} \boldsymbol{r}_k$, $\boldsymbol{r}_k$ is given in (6), and $\widetilde{\lambda}_{n,k}$ is the nth eigenvalue of the matrix $\widetilde{\boldsymbol{F}}_k$.

III. The Proposed Method

The aim is to design the code matrix $\boldsymbol{A}$ to improve the detection performance of the presented optimal detector. In this section, we consider a knowledge-based (cognitive) system such that the second-order statistics of the target, clutter, and interference are known at the design stage. This can be fulfilled using geological and meteorological data as well as data of previous scans in a cognitive setup. However, in Section V, we extend the proposed method to deal with uncertainties in prior knowledge of interference via employing a robust approach.

The detector presented in Section II-B is optimal in the sense that it maximizes probability of detection while keeping probability of false-alarm below a certain threshold. However, in most cases, finding the best transmission code that maximizes the performance of the optimal detector is not analytically tractable. In such cases, other design criteria referred to as *information-theoretic* criteria can be employed for transmission code design [15, 19]. Herein, we employ mutual information and J-divergence as the information theoretic criteria for the code design. These metrics give some bounds on the performance of the optimal detector, i.e., maximizing each of them enhances the performance of the detector (see [15, 17, 25, 26] and references therein for detailed discussions and the justification of using such metrics).

A. Colocated MIMO radars

It is shown in [17] that maximizing the mutual information ($\mathcal{M}$) between the received signal and the target response leads to a better detection performance in the case of Gaussian distributions. Using the results of [25], the $\mathcal{M}$ metric associated with (14) is given by

$$\begin{aligned} \mathcal{M} &= \log[(\pi e)^N \det(\boldsymbol{I}_{NN_R} + \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{R}_s\boldsymbol{D}^{-\frac{1}{2}})] - \log[(\pi e)^N \det(\boldsymbol{I}_{NN_R})] \\ &= \log\det\left(\boldsymbol{I}_{NN_R} + \boldsymbol{F}\right) \end{aligned} \tag{17}$$

with e being the Napier constant ($e \approx 2.71$). Another information theoretic metric is the J-divergence ($\mathcal{J}$) which is used to measure the similarity between two distributions $f_0 = f(\boldsymbol{y}|\mathrm{H}_0)$ and $f_1 = f(\boldsymbol{y}|\mathrm{H}_1)$. Employing the metric for signal design is based on the Stein Lemma [15]. The $\mathcal{J}$ metric can be computed as [26]

$$\mathbb{E}\{(\mathcal{L}-1)\log(\mathcal{L})|\mathrm{H}_0\} \ = \mathbb{E}\{\log(\mathcal{L})|\mathrm{H}_1\} - \mathbb{E}\{\log(\mathcal{L})|\mathrm{H}_0\},$$

where $\mathcal{L}$ is the likelihood ratio $\mathcal{L} \triangleq f(\boldsymbol{y}|\mathrm{H}_1)/f(\boldsymbol{y}|\mathrm{H}_0)$. Applying the above equation to (14) results in

$$\mathcal{J} = \mathrm{tr}\left\{(\boldsymbol{F} + \boldsymbol{I}_{NN_R})^{-1} + \boldsymbol{F} - \boldsymbol{I}_{NN_R}\right\}.$$

Thus, the code design problem for colocated MIMO radars using the aforementioned criteria can be cast as

$$\begin{aligned} \max_{\boldsymbol{A},\boldsymbol{F},\boldsymbol{D},\boldsymbol{R}_s} \quad & f_{\mathcal{I}}(\boldsymbol{F}) \\ \text{subject to} \quad & \boldsymbol{F} = \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{R}_s\boldsymbol{D}^{-\frac{1}{2}} \\ & \boldsymbol{A} \in \mathcal{C}, \end{aligned} \tag{18}$$

where $\mathcal{I} \in \{\mathcal{M},\ \mathcal{J}\}$, $f_{\mathcal{M}}(\boldsymbol{F}) = \log\det\left(\boldsymbol{I}_{NN_R} + \boldsymbol{F}\right)$, and $f_{\mathcal{J}}(\boldsymbol{F}) = \mathrm{tr}\{(\boldsymbol{F} + \boldsymbol{I}_{NN_R})^{-1} + \boldsymbol{F} - \boldsymbol{I}_{NN_R}\}$. In (18), $\mathcal{C}$ represents a typical constraint set on the code matrix $\boldsymbol{A}$. In this paper, the following three types of constraints on the code matrix $\boldsymbol{A}$ are considered; indeed, the set $\mathcal{C}$ represents one of these sets

1) Note that the transmit energy of the system is given by $\sum_{m=1}^{N_t} \|\boldsymbol{a}_m\|_2^2 = \sum_{n=1}^{N} \|\boldsymbol{s}_{co}(n)\|_2^2 = \|\boldsymbol{A}\|_F^2$. Therefore, to impose the energy constraint, we let $\|\boldsymbol{A}\|_F^2 \leq e_s$, where e_s is the maximum available transmit energy.
2) PAR constraint, i.e., $\|\boldsymbol{A}\|_F^2 = NN_T$ and $\max_{n=1,\cdots,NN_T} |\widetilde{a}(n)|^2 \leq \gamma$, where $\widetilde{a}(n)$ denotes the nth element of the vector $\widetilde{\boldsymbol{a}} = \mathrm{vec}(\boldsymbol{A})$, and γ is the maximum PAR level. The PAR constraint can also be applied column-wise to the matrix $\boldsymbol{A}$ in widely separated MIMO radars (see Section IV-B).

3) similarity constraint, i.e., $\|\boldsymbol{A}-\boldsymbol{A}_0\|_F^2 \leq \epsilon$, where $\boldsymbol{A}_0$ is a desired code matrix and ϵ is the similarity threshold.

Imposing these constraints leads to three different optimization problems with the general form of (18). Note that in general, the stated design problem is non-convex and belongs to a class of NP-hard problems [27]. However, it can be further viewed that $f_{\mathcal{M}}(\boldsymbol{F})$ is a concave function of $\boldsymbol{F} \in \mathcal{S}_+^{NN_R}$ whereas $f_{\mathcal{J}}(\boldsymbol{F})$ is a convex one.

In order to obtain solutions to the optimization problem in (18), we use minorization-maximization (MaMi)[5] technique, an iterative optimization technique, which maximizes a proper upper bound of the objective function to find their maxima [28]. In summary, MaMi can be used for obtaining stationary points of the general maximization problem

$$\max_{\boldsymbol{\omega}} h(\boldsymbol{\omega}), \text{ subject to } \bar{c}(\boldsymbol{\omega}) \leq 0, \tag{19}$$

in an iterative way where $h(.)$ and $\bar{c}(.)$ can be neither concave nor convex. In the ith iteration of MaMi, a so-called *minorizing* function $p^{(i)}(\boldsymbol{\omega})$ is found such that

$$p^{(i)}(\boldsymbol{\omega}) \leq h(\boldsymbol{\omega}), \ \forall \boldsymbol{\omega} \quad \text{and} \quad p^{(i)}(\boldsymbol{\omega}^{(i-1)}) = h(\boldsymbol{\omega}^{(i-1)}),$$

with $\boldsymbol{\omega}^{(i-1)}$ being the value of $\boldsymbol{\omega}$ in the $(i-1)$th iteration. Then, the minorizing function $p^{(i)}(\boldsymbol{\omega})$ is maximized in the ith iteration rather than the function $h(\boldsymbol{\omega})$, i.e.,

$$\max_{\boldsymbol{\omega}} p^{(i)}(\boldsymbol{\omega}) \quad \text{subject to} \quad \bar{c}(\boldsymbol{\omega}) \leq 0 \tag{20}$$

to obtain $\boldsymbol{\omega}^{(i)}$. Logically, $p^{(i)}(\boldsymbol{\omega})$ is chosen such that (20) is easier to solve than (19).

The MaMi method is employed in the following theorem to find a solution to the constrained design problem in (18). Indeed, Theorem 1 deals with the problem in (18) by iteratively solving a proper QCQP. The methods for solving the aforementioned QCQP stated in Theorem 1 are presented in Section IV.

Theorem 1: The design problems in (18) for $\mathcal{M}$ and $\mathcal{J}$ are equivalent. Furthermore, the solution $\boldsymbol{A} = \boldsymbol{A}_\star$ to this problem can be obtained iteratively by solving the following QCQP in the ith iteration

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{H}^{(i)} \widetilde{\boldsymbol{a}} + 2\Re\{(\boldsymbol{g}^{(i)})^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \widetilde{\boldsymbol{a}} \in \widetilde{\mathcal{C}}, \end{aligned} \tag{21}$$

where $\widetilde{\boldsymbol{a}} = \mathrm{vec}(\boldsymbol{A})$ and $\widetilde{\mathcal{C}}$ denotes the constraint set (associated with $\mathcal{C}$) imposed on $\widetilde{\boldsymbol{a}}$. The matrices $\{\boldsymbol{H}^{(i)}\}_i \succeq \boldsymbol{0}$ and the vectors $\{\boldsymbol{g}^{(i)}\}_i$ are given below. ■

[5] Also known as majorization-maximization or MM algorithm in the literature.

Proof. We begin by noting that the covariance of the signal component $\boldsymbol{R}_s$ is rank-1; namely, $\boldsymbol{R}_s = \boldsymbol{k}\boldsymbol{k}^H$ with $\boldsymbol{k} = [\boldsymbol{I}_{N_R} \otimes (\boldsymbol{A} \odot \boldsymbol{P})]\,\boldsymbol{b}$ and $\boldsymbol{b} = \sigma_s \text{vec}\,(\boldsymbol{\Psi}(\theta_0))$. Now, using the determinant property [21], we obtain

$$\begin{aligned}\log\det(\boldsymbol{I}_{NN_R} + \boldsymbol{F}) =& \log\det\left(\boldsymbol{I}_{NN_R} + \boldsymbol{D}^{-\frac{1}{2}}\boldsymbol{k}\boldsymbol{k}^H\boldsymbol{D}^{-\frac{1}{2}}\right) \\ =& \log\left(1 + \boldsymbol{k}^H(\boldsymbol{R}_c + \boldsymbol{R}_{\boldsymbol{v}})^{-1}\boldsymbol{k}\right) = \log\left(1 + J\right)\end{aligned} \tag{22}$$

where $J = \boldsymbol{k}^H(\boldsymbol{R}_c + \boldsymbol{R}_{\boldsymbol{v}})^{-1}\boldsymbol{k} \in \mathbb{R}_+$. Similarly, by employing the trace inequality [21], we have

$$\begin{aligned}\text{tr}\left\{(\boldsymbol{I}_{NN_R} + \boldsymbol{F})^{-1} + \boldsymbol{F}\right\} =& \text{tr}\left\{(1 + J)^{-1} + J\right\} + \text{const.} \\ =& J + \frac{1}{1+J} + \text{const.}\end{aligned} \tag{23}$$

Note that the objective functions in (18) are monotonically increasing functions of the scalar J. Therefore, the design problems for $\mathcal{M}$ and $\mathcal{J}$ are equivalent. Consequently, in the sequel, we derive the QCQP associated with the mutual information[6]. The $\mathcal{J}$ metric can be dealt with similarly.

Remark 1: considering the above observation, we note that the objective functions in (18) can be replaced by the scalar J. Also, one may consider J as a type of SINR at the output of the detector and observe that for the employed model of colocated MIMO radars, code design via maximizing SINR is equivalent to that of maximizing the information theoretic criteria. We herein remark on the fact that SINR maximization has been addressed in the literature for MIMO radar code design but to the best of our knowledge, they are all based on a synthesis stage/suboptimal procedure to deal with e.g., PAR constraint or assume a perfect prior knowledge of interference.

The following lemma is the key to employ MaMi technique for the non-convex design problem and deriving the QCQP.

Lemma 1: For $f_{\mathcal{M}}(\boldsymbol{F}) = \log\det(\boldsymbol{I}_{NN_R} + \boldsymbol{F}) : \mathcal{S}_+^{NN_R} \to \mathbb{R}_+$, we have

$$\log\det(\boldsymbol{I}_{NN_R} + \boldsymbol{F}) = \log(\boldsymbol{u}^H\boldsymbol{B}_{\mathcal{M}}^{-1}\boldsymbol{u}), \tag{24}$$

where $\boldsymbol{u} = [1, \boldsymbol{0}^T_{NN_R\times 1}]^T$ and

$$\boldsymbol{B}_{\mathcal{M}} = \begin{bmatrix} 1 & \boldsymbol{k}^H \\ \boldsymbol{k} & \boldsymbol{R}_s + \boldsymbol{R}_c + \boldsymbol{R}_{\boldsymbol{v}} \end{bmatrix} \in \mathcal{S}_+^{NN_R+1}. \tag{25}$$

Moreover, for any full-column rank matrix $\boldsymbol{U}$, $\log\det(\boldsymbol{U}^H\boldsymbol{B}_{\mathcal{M}}^{-1}\boldsymbol{U})$ is convex w.r.t. $\boldsymbol{B}_{\mathcal{M}} \succ \boldsymbol{0}$. ■

[6]We can equivalently consider the maximization of J; however, we deal with $\log(1 + J)$ to use the results for the widely separated case shortly.

Proof: See Appendix A. $\square$

Using Lemma 1, the solution to (18) for the case of $\mathcal{I} = \mathcal{M}$ can be found by considering the equivalent optimization

$$\begin{aligned} \max_{\boldsymbol{A}\in\mathcal{C},\boldsymbol{k},\boldsymbol{B}_{\mathcal{M}}} \quad & \log(\boldsymbol{u}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{u}) \\ \text{subject to} \quad & \boldsymbol{k} = \left[\boldsymbol{I}_{N_R} \otimes (\boldsymbol{A} \odot \boldsymbol{P})\right] \boldsymbol{b}. \end{aligned} \tag{26}$$

According to Lemma 1, the objective function in (26) is convex w.r.t. $\boldsymbol{B}_{\mathcal{M}}^{-1}$. More concretely, here, the matrix $\boldsymbol{U}$ is given by the vector $\boldsymbol{u}$ and the determinant of the scalar $\boldsymbol{u}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{u}$ is given by $\boldsymbol{u}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{u}$. Therefore, this term can be minorized using its tangent plane at a given $\widetilde{\boldsymbol{B}}_{\mathcal{M}}$ as [29]

$$\log(\boldsymbol{u}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{u}) \geq \log(\boldsymbol{u}^H \widetilde{\boldsymbol{B}}_{\mathcal{M}}^{-1} \boldsymbol{u}) + \operatorname{tr}\{\widetilde{\boldsymbol{\Upsilon}}(\boldsymbol{B}_{\mathcal{M}} - \widetilde{\boldsymbol{B}}_{\mathcal{M}})\},$$

where $\widetilde{\boldsymbol{\Upsilon}} = -\widetilde{\boldsymbol{B}}_{\mathcal{M}}^{-1} \boldsymbol{u} \big(\boldsymbol{u}^H \widetilde{\boldsymbol{B}}_{\mathcal{M}}^{-1} \boldsymbol{u}\big)^{-1} \boldsymbol{u}^H \widetilde{\boldsymbol{B}}_{\mathcal{M}}^{-1}$. Moreover, the first term on the right hand side of the above inequality is constant for a given $\widetilde{\boldsymbol{B}}_{\mathcal{M}}$. Thus, by defining $\boldsymbol{\Upsilon} = -\widetilde{\boldsymbol{\Upsilon}} \succeq \boldsymbol{0}$, the optimization problem for the ith iteration of MaMi can be handled via solving

$$\begin{aligned} \min_{\boldsymbol{A}\in\mathcal{C},\boldsymbol{k},\boldsymbol{B}_{\mathcal{M}}} \quad & \operatorname{tr}(\boldsymbol{\Upsilon}^{(i)} \boldsymbol{B}_{\mathcal{M}}) \\ \text{subject to} \quad & \boldsymbol{k} = \left[\boldsymbol{I}_{N_R} \otimes (\boldsymbol{A} \odot \boldsymbol{P})\right] \boldsymbol{b}. \end{aligned} \tag{27}$$

where $\boldsymbol{\Upsilon}^{(i)}$ denotes the matrix $\boldsymbol{\Upsilon}$ at the ith iteration. Now, letting $\boldsymbol{\Upsilon} = \begin{bmatrix} \upsilon_{11} & \boldsymbol{\upsilon}_{12} \\ \boldsymbol{\upsilon}_{21} & \boldsymbol{\Upsilon}_{22} \end{bmatrix}$ with the same partitioning as that of $\boldsymbol{B}_{\mathcal{M}}$ in (25), neglecting the constant terms, and using hermitian property of $\boldsymbol{\Upsilon}$ and $\boldsymbol{R}_s$, the following equivalent problem is obtained for (27)

$$\begin{aligned} \min_{\boldsymbol{A}\in\mathcal{C},\boldsymbol{R}_s,\boldsymbol{R}_c,\boldsymbol{k}} \quad & 2\Re\big\{\boldsymbol{k}^H \boldsymbol{\upsilon}_{21}^{(i)}\big\} + \operatorname{tr}\left\{\boldsymbol{\Upsilon}_{22}^{(i)}(\boldsymbol{R}_{\boldsymbol{s}} + \boldsymbol{R}_{\boldsymbol{c}})\right\} \\ \text{subject to} \quad & \boldsymbol{R}_{\boldsymbol{s}} = \boldsymbol{k}\boldsymbol{k}^H \\ & \boldsymbol{R}_{\boldsymbol{c}} = (\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})\, \boldsymbol{Q}\, (\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})^H \\ & \boldsymbol{k} = \left[\boldsymbol{I}_{N_R} \otimes (\boldsymbol{A} \odot \boldsymbol{P})\right] \boldsymbol{b}. \end{aligned} \tag{28}$$

The optimization problem above is still implicit in terms of the code matrix $\boldsymbol{A}$. Lemma 2 is the key to obtain explicit expressions in terms of $\boldsymbol{A}$.

Lemma 2: Let $\boldsymbol{G} = \boldsymbol{b}\boldsymbol{\upsilon}_{21}^H$, $\widetilde{\boldsymbol{g}} = \sum_{k=1}^{N_R} \operatorname{vec}\big(\boldsymbol{G}_{kk}{}^H\big)$, $\boldsymbol{g} = (\widetilde{\boldsymbol{g}} \odot \widetilde{\boldsymbol{p}}^*)$, $\widetilde{\boldsymbol{p}} = \operatorname{vec}(\boldsymbol{P})$, and $\widetilde{\boldsymbol{a}} = \operatorname{vec}(\boldsymbol{A})$, where $\boldsymbol{G}_{kk} \in \mathbb{C}^{N \times N_T}$ is a submatrix of $\boldsymbol{G}$ containing rows $(k-1)N_T + 1$ to kN_T and columns $(k-1)N + 1$ to kN. Then we have

$$\Re\{\boldsymbol{k}^H \boldsymbol{\upsilon}_{21}\} = \Re\{\boldsymbol{g}^H \widetilde{\boldsymbol{a}}\}, \tag{29}$$

$$\mathrm{tr}\{\boldsymbol{\Upsilon}_{22}(\boldsymbol{R}_s+\boldsymbol{R}_c)\} = \widetilde{\boldsymbol{a}}^H \boldsymbol{H}\, \widetilde{\boldsymbol{a}},$$

where

$$\boldsymbol{H} = \sum_{l=1}^{N_R}\sum_{k=1}^{N_R}\left[(\boldsymbol{T}_{kl}^H \otimes \boldsymbol{\Upsilon}_{22,lk}) \odot (\widetilde{\boldsymbol{p}}\widetilde{\boldsymbol{p}}^H)^T + \boldsymbol{Q}_{kl}^H \otimes \boldsymbol{\Upsilon}_{22,lk}\right],$$

in which $\boldsymbol{T}_{kl}$ and $\boldsymbol{Q}_{kl}$ are submatrices of $\boldsymbol{T}$ and $\boldsymbol{Q}$ (see (3) and (4)), respectively, containing rows $(k-1)N_T+1$ to kN_T and columns $(l-1)N_T+1$ to lN_T, and $\boldsymbol{\Upsilon}_{22,lk}$ is a submatrix of $\boldsymbol{\Upsilon}_{22}$ containing rows $(l-1)N+1$ to lN and columns $(k-1)N+1$ to kN. ■

Proof: See Appendix B. □

Now, by applying Lemma 2 to (28), we have the following equivalent optimization problem in the ith iteration

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{H}^{(i)}\widetilde{\boldsymbol{a}} + 2\Re\{(\boldsymbol{g}^{(i)})^H\widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \widetilde{\boldsymbol{a}} \in \tilde{\mathcal{C}}, \end{aligned} \tag{30}$$

in which $\boldsymbol{H}^{(i)}$ and $\boldsymbol{g}^{(i)}$ are available using eq. (29) and it completes the proof of Theorem 1. □

B. Widely separated MIMO radars

For widely separated MIMO radars, the information theoretic criteria $\mathcal{M}$ and $\mathcal{J}$ are given by

$$\mathcal{M} = \sum_{k=1}^{N_R} \log\det(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_k), \tag{31}$$

$$\mathcal{J} = \sum_{k=1}^{N_R} \mathrm{tr}\{(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_k)^{-1} + \widetilde{\boldsymbol{F}}_k - \boldsymbol{I}_N\},$$

and therefore, the corresponding constrained design problems can be cast as

$$\begin{aligned} \max_{\boldsymbol{A}\in\mathcal{C},\{\widetilde{\boldsymbol{F}}_k\}} \quad & \sum_{k=1}^{N_R} f_{W,\mathcal{I}}(\widetilde{\boldsymbol{F}}_k) \\ \text{subject to} \quad & \widetilde{\boldsymbol{F}}_k = \boldsymbol{D}_k^{-\frac{1}{2}}\boldsymbol{R}_{\boldsymbol{s}_k}\boldsymbol{D}_k^{-\frac{1}{2}}, \end{aligned} \tag{32}$$

where $\mathcal{I} \in \{\mathcal{M}, \mathcal{J}\}$, $f_{W,\mathcal{M}}(\widetilde{\boldsymbol{F}}_k) = \log\det(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_k)$, and $f_{W,\mathcal{J}}(\widetilde{\boldsymbol{F}}_k) = \mathrm{tr}\{(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_k)^{-1} + \widetilde{\boldsymbol{F}}_k - \boldsymbol{I}_N\}$. It should be noted that the above constrained design problem is non-convex. Furthermore in this case, as opposed to the colocated case, the results of employing $\mathcal{M}$ and $\mathcal{J}$ metrics are not generally equivalent (see Section VI-B). This is due to the fact that the covariance matrix of the target component $\boldsymbol{R}_s$ is no longer rank-1. Therefore, the following theorem that presents the results for the widely separated antenna case is proved for each metric separately.

Theorem 2: The solution $\boldsymbol{A} = \boldsymbol{A}_\star$ to the optimization problem (32) can be obtained iteratively by solving the following QCQP in the ith iteration

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{H}_{\mathcal{W},\mathcal{I}}^{(i)} \widetilde{\boldsymbol{a}} + 2\,\Re\{(\boldsymbol{g}_{\mathcal{W},\mathcal{I}}^{(i)})^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \widetilde{\boldsymbol{a}} \in \widetilde{\mathcal{C}}, \end{aligned} \tag{33}$$

where the matrices $\{\boldsymbol{H}_{\mathcal{W},\mathcal{I}}^{(i)}\}_i \succeq \boldsymbol{0}$ and the vectors $\{\boldsymbol{g}_{\mathcal{W},\mathcal{I}}^{(i)}\}_i$ are defined for each of the information theoretic metrics $\mathcal{I} \in \{\mathcal{M}, \mathcal{J}\}$ separately in the sequel. ■

1) Mutual Information: The case of mutual information ($\mathcal{M}$) is similar to that of the colocated case; however, here the covariance matrices $\{\boldsymbol{R}_{s_k}\}_k$ can be full-rank. Therefore, we apply MaMi technique to the functions of psd matrices. Letting $\boldsymbol{R}_{s_k} = \boldsymbol{Y}_k \boldsymbol{Y}_k^H \in \mathcal{S}_+^{N_T}$ and by defining

$$\boldsymbol{B}_{k,\mathcal{M}} = \begin{bmatrix} \boldsymbol{B}_{k,\mathcal{M}_{11}} & \boldsymbol{B}_{k,\mathcal{M}_{12}} \\ \boldsymbol{B}_{k,\mathcal{M}_{12}}^H & \boldsymbol{B}_{k,\mathcal{M}_{22}} \end{bmatrix} \in \mathcal{S}_+^{N+N_T}, \tag{34}$$

in which $\boldsymbol{B}_{k,\mathcal{M}_{11}} = \boldsymbol{I}_{N_T}$, $\boldsymbol{B}_{k,\mathcal{M}_{12}} = \boldsymbol{Y}_k^H(\boldsymbol{A} \odot \boldsymbol{P}_k)^H$, and $\boldsymbol{B}_{k,\mathcal{M}_{22}} = \boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + (\boldsymbol{A} \odot \boldsymbol{P}_k)\boldsymbol{R}_{s_k}(\boldsymbol{A} \odot \boldsymbol{P}_k)^H + \boldsymbol{R}_{v_k}$, we have[7]

$$\begin{aligned} \sum_{k=1}^{N_R} \log\det(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_k) &= \sum_{k=1}^{N_R} \log\det(\boldsymbol{I}_{N_T} + \boldsymbol{J}_k) \\ &= \sum_{k=1}^{N_R} \log\det(\boldsymbol{U}^H \boldsymbol{B}_{k,\mathcal{M}}^{-1} \boldsymbol{U}) \end{aligned} \tag{35}$$

with $\boldsymbol{U} = [\boldsymbol{I}_{N_T} \quad \boldsymbol{0}_{N_T \times N}]^T$ and

$$\boldsymbol{J}_k = \boldsymbol{Y}_k^H (\boldsymbol{A} \odot \boldsymbol{P}_k)^H (\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + \boldsymbol{R}_{v_k})^{-1} (\boldsymbol{A} \odot \boldsymbol{P}_k) \boldsymbol{Y}_k. \tag{36}$$

Next, using the results of Lemma 1, (35) can be minorized by its supporting hyperplane at given $\{\widetilde{\boldsymbol{B}}_{k,\mathcal{M}}\}_k$ as

$$\begin{aligned} \sum_{k=1}^{N_R} \log\det\left(\boldsymbol{U}^H \boldsymbol{B}_{k,\mathcal{M}}^{-1} \boldsymbol{U}\right) \geq & \sum_{k=1}^{N_R} \log\det\left(\boldsymbol{U}^H \widetilde{\boldsymbol{B}}_{k,\mathcal{M}}^{-1} \boldsymbol{U}\right) \\ & + \sum_{k=1}^{N_R} \operatorname{tr}\left\{\widetilde{\boldsymbol{\Phi}}_k \big(\boldsymbol{B}_{k,\mathcal{M}} - \widetilde{\boldsymbol{B}}_{k,\mathcal{M}}\big)\right\}, \end{aligned}$$

where $\widetilde{\boldsymbol{\Phi}}_k = -\widetilde{\boldsymbol{B}}_{k,\mathcal{M}}^{-1} \boldsymbol{U} \big(\boldsymbol{U}^H \widetilde{\boldsymbol{B}}_{k,\mathcal{M}}^{-1} \boldsymbol{U}\big)^{-1} \boldsymbol{U}^H \widetilde{\boldsymbol{B}}_{k,\mathcal{M}}^{-1}$. Let the psd matrix $-\widetilde{\boldsymbol{\Phi}}_k$ be partitioned in accordance with $\boldsymbol{B}_{k,\mathcal{M}}$ as $\begin{bmatrix} \boldsymbol{\Phi}_{11,k} & \boldsymbol{\Phi}_{12,k} \\ \boldsymbol{\Phi}_{21,k} & \boldsymbol{\Phi}_{22,k} \end{bmatrix}$. By neglecting the constants, it can be shown that the following problem

[7] See (11) for the definition of $\boldsymbol{R}_s$ and $\boldsymbol{R}_{s_k}$.

is solved in the ith iteration of the MaMi procedure

$$\min_{\boldsymbol{A}\in\mathcal{C}} \quad \sum_{k=1}^{N_R}[\mathrm{tr}\{(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{R}_{s_k}(\boldsymbol{A}\odot\boldsymbol{P}_k)^H)\boldsymbol{\Phi}_{22,k}^{(i)}\} \\ +2\Re\,\mathrm{tr}\{\{\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H\boldsymbol{\Phi}_{21,k}^{(i)}\}\}] \tag{37}$$

Next, by employing standard Kronecker product properties and some algebraic manipulations, we obtain the following explicit form w.r.t. $\widetilde{\boldsymbol{a}}$ in the ith iteration

$$\min_{\widetilde{\boldsymbol{a}}} \quad \widetilde{\boldsymbol{a}}^H\boldsymbol{H}_{\mathcal{M},\mathcal{W}}^{(i)}\widetilde{\boldsymbol{a}} + 2\,\Re\{(\boldsymbol{g}_{\mathcal{M},\mathcal{W}}^{(i)})^H\widetilde{\boldsymbol{a}}\} \tag{38}$$
$$\text{subject to} \quad \widetilde{\boldsymbol{a}}\in\widetilde{\mathcal{C}},$$

where by defining $\widetilde{\boldsymbol{p}}_k = \mathrm{vec}(\boldsymbol{P}_k)$ we let

$$\boldsymbol{H}_{\mathcal{M},\mathcal{W}}^{(i)} = \sum_{k=1}^{N_R}[(\boldsymbol{R}_{s_k}^H\otimes\boldsymbol{\Phi}_{22,k}^{(i)})\odot(\widetilde{\boldsymbol{p}}_k\widetilde{\boldsymbol{p}}_k^H)^T + \boldsymbol{R}_{c_k}^H\otimes\boldsymbol{\Phi}_{22,k}^{(i)}],$$
$$\boldsymbol{g}_{\mathcal{M},\mathcal{W}}^{(i)} = \sum_{k=1}^{N_R}\mathrm{vec}(\boldsymbol{\Phi}_{21,k}^{(i)}\boldsymbol{Y}_k^H)\odot\widetilde{\boldsymbol{p}}_k^*.$$

2) J-divergence: For this case, the objective function is $f_{\mathcal{J}} = \sum_{k=1}^{N_R}\mathrm{tr}\{(\boldsymbol{I}_N+\widetilde{\boldsymbol{F}}_k)^{-1}+\widetilde{\boldsymbol{F}}_k-\boldsymbol{I}_N\}$. Using the trace identity and the fact that $\boldsymbol{R}_{s_k}=\boldsymbol{Y}_k\boldsymbol{Y}_k^H$, we have

$$\mathrm{tr}\{(\boldsymbol{I}_N+\widetilde{\boldsymbol{F}}_k)^{-1}+\widetilde{\boldsymbol{F}}_k-\boldsymbol{I}_N\} = \mathrm{tr}\{(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)^{-1}+\boldsymbol{J}_k\}+\text{const.},$$

where $\boldsymbol{J}_k$ is defined in (36). Now, we first observe that the function $\mathrm{tr}\{(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)^{-1}\}$ is convex w.r.t. the psd matrix $\boldsymbol{J}_k$. Therefore, a linear minorizer is available considering the tangent plane at a given $\widetilde{\boldsymbol{J}}_k$

$$\mathrm{tr}\{(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)^{-1}\} \geq \mathrm{tr}\{(\boldsymbol{I}_{N_T}+\widetilde{\boldsymbol{J}}_k)^{-1}\} \\ -\mathrm{tr}\{(\boldsymbol{I}_{N_T}+\widetilde{\boldsymbol{J}}_k)^{-2}(\boldsymbol{J}_k-\widetilde{\boldsymbol{J}}_k)\}. \tag{39}$$

The above minorizer leads to the following optimization in the ith iteration

$$\max_{\boldsymbol{A}\in\mathcal{C},\{\boldsymbol{J}_k\}} \quad \sum_{k=1}^{N_R}\mathrm{tr}\{\boldsymbol{L}_k^{(i)}\boldsymbol{J}_k\} \tag{40}$$
$$\text{subject to} \quad \boldsymbol{J}_k = \boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H+\boldsymbol{R}_{v_k})^{-1} \\ (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k.$$

with $\boldsymbol{L}_k^{(i)} \triangleq \boldsymbol{I}_{N_T}-(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k{}^{(i)})^{-2}$. It should be noted that the matrices $\{\boldsymbol{L}_k^{(i)}\}_k$ are psd $\forall\, k, i$, since $\lambda_{min}(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)\geq 1$. Therefore, there exists a matrix $\boldsymbol{E}_k^{(i)}$ such that $\boldsymbol{L}^{(i)} = \boldsymbol{E}_k^{(i)}\boldsymbol{E}_k^{(i)H}$ and thus, we have the following equivalent form for (40)

$$\max_{\boldsymbol{A}\in\mathcal{C}} \quad \sum_{k=1}^{N_R} \mathrm{tr}\{\boldsymbol{I}_{N_T} + \boldsymbol{E}_k^{(i)H} \boldsymbol{Y}_k^H (\boldsymbol{A}\odot\boldsymbol{P}_k)^H \\ (\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + \boldsymbol{R}_{v_k})^{-1}(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k\boldsymbol{E}_k^{(i)}\}. \tag{41}$$

Using tricks similar to those of the mutual information case, the objective function in (41) can be rewritten as a convex function of auxiliary matrices which lays the ground for applying the MaMi technique. More precisely, by defining $\boldsymbol{\Gamma} \triangleq \boldsymbol{I}_{N_T} + \boldsymbol{E}_k^{(i)H}\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H+\boldsymbol{R}_{v_k})^{-1}(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k\boldsymbol{E}_k^{(i)}$ and using the matrix inversion Lemma, it is observed that $\boldsymbol{\Gamma} = \boldsymbol{U}^H\boldsymbol{B}_{\mathcal{J}}^{-1}\boldsymbol{U}$ where $\boldsymbol{U} = [\boldsymbol{I}_{N_T} \;\; \boldsymbol{0}_{N_T\times N}]^T$ and

$$\boldsymbol{B}_{k,\mathcal{J}} = \begin{bmatrix} \boldsymbol{B}_{k,\mathcal{J}_{11}} & \boldsymbol{B}_{k,\mathcal{J}_{12}} \\ \boldsymbol{B}_{k,\mathcal{J}_{12}}^H & \boldsymbol{B}_{k,\mathcal{J}_{22}} \end{bmatrix} \tag{42}$$

with $\boldsymbol{B}_{k,\mathcal{J}_{11}} = \boldsymbol{I}_{N_T}$, $\boldsymbol{B}_{k,\mathcal{J}_{12}} = \boldsymbol{E}_k^{(i)H}\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H$, and $\boldsymbol{B}_{k,\mathcal{J}_{22}} = \boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k\boldsymbol{L}_k^{(i)}\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H + \boldsymbol{R}_{v_k}$. Furthermore, using a proof similar to that of Lemma 1, it can be shown that $\mathrm{tr}(\boldsymbol{U}^H\boldsymbol{B}_{k,\mathcal{J}}^{-1}\boldsymbol{U})$ is convex w.r.t. $\boldsymbol{B}_{k,\mathcal{J}} \succ \boldsymbol{0}$. Therefore, the solution to (41) in the ith iteration can be obtained via solving

$$\max_{\boldsymbol{A}\in\mathcal{C},\boldsymbol{B}_{k,\mathcal{J}}} \quad \sum_{k=1}^{N_R} \mathrm{tr}\{\boldsymbol{U}^H\boldsymbol{B}_{k,\mathcal{J}}^{-1(i)}\boldsymbol{U}\}, \tag{43}$$

which is non-convex. To tackle (43), employing the MaMi technique, we minorize $\mathrm{tr}\{\boldsymbol{U}^H\boldsymbol{B}_{k,\mathcal{J}}^{-1(i)}\boldsymbol{U}\}$ using its supporting hyperplane at a given $\widetilde{\boldsymbol{B}}_{k,\mathcal{J}}$ as

$$\mathrm{tr}\{\boldsymbol{U}^H\boldsymbol{B}_{k,\mathcal{J}}^{-1}\boldsymbol{U}\} \geq \mathrm{tr}\{\boldsymbol{U}^H\widetilde{\boldsymbol{B}}_{k,\mathcal{J}}^{-1}\boldsymbol{U}\} \\ - \mathrm{tr}\{\widetilde{\boldsymbol{B}}_{k,\mathcal{J}}^{-1}\boldsymbol{U}\boldsymbol{U}^H\widetilde{\boldsymbol{B}}_{k,\mathcal{J}}^{-1}(\boldsymbol{B}_{k,\mathcal{J}} - \widetilde{\boldsymbol{B}}_{k,\mathcal{J}})\}.$$

Consequently, as $\widetilde{\boldsymbol{B}}_{\mathcal{J}}$ is fixed, the problem in the ith iteration turns into

$$\min_{\boldsymbol{A}\in\mathcal{C}} \quad \sum_{k=1}^{N_R} \mathrm{tr}\{\boldsymbol{W}_k^{(i)}\boldsymbol{B}_{k,\mathcal{J}}\}, \tag{44}$$

with $\boldsymbol{W}_k^{(i)} = (\boldsymbol{B}_{k,\mathcal{J}}^{(i)})^{-1}\boldsymbol{U}\boldsymbol{U}^H(\boldsymbol{B}_{k,\mathcal{J}}^{(i)})^{-1} \succeq \boldsymbol{0}$. By partitioning $\boldsymbol{W}_k^{(i)}$ similar to that of $\boldsymbol{B}_{k,\mathcal{J}}$ and neglecting the constants, we come up with the equivalent problem

$$\min_{\boldsymbol{A}\in\mathcal{C}} \quad \sum_{k=1}^{N_R}[\mathrm{tr}\{(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H \\ +(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k\boldsymbol{L}_k^{(i)}\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H)\boldsymbol{W}_{22,k}^{(i)}\} \\ +2\Re\,\mathrm{tr}\{\boldsymbol{E}_k^{(i)H}\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H\boldsymbol{W}_{21,k}^{(i)}\}], \tag{45}$$

which can be recast in terms of $\widetilde{\boldsymbol{a}}$ by employing techniques similar to those of $\mathcal{M}$ as

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{H}^{(i)}_{\mathcal{J},\mathcal{W}} \widetilde{\boldsymbol{a}} + 2\Re\{(\boldsymbol{g}^{(i)}_{\mathcal{J},\mathcal{W}})^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \widetilde{\boldsymbol{a}} \in \widetilde{\mathcal{C}}, \end{aligned} \tag{46}$$

where

$$\begin{aligned} \boldsymbol{H}^{(i)}_{\mathcal{J},\mathcal{W}} = \sum_{k=1}^{N_R} [((\boldsymbol{Y}_k \boldsymbol{L}^{(i)}_k \boldsymbol{Y}^H_k) \otimes \boldsymbol{W}^{(i)}_{22,k}) \\ \odot (\widetilde{\boldsymbol{p}}_k \widetilde{\boldsymbol{p}}^H_k)^T + \boldsymbol{R}^H_{c_k} \otimes \boldsymbol{W}^{(i)}_{22,k}], \\ \boldsymbol{g}^{(i)}_{\mathcal{J},\mathcal{W}} = \sum_{k=1}^{N_R} \operatorname{vec}(\boldsymbol{W}^{(i)}_{21,k} {\boldsymbol{E}^{(i)}_k}^H \boldsymbol{Y}^H_k) \odot \widetilde{\boldsymbol{p}}^*_k. \end{aligned}$$

IV. Solving the QCQP in Each Iteration

For both colocated and widely separated MIMO systems, the QCQP in each iteration of the devised method is in the form of

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{\mathcal{H}} \widetilde{\boldsymbol{a}} + 2\,\Re\{\boldsymbol{g}^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \widetilde{\boldsymbol{a}} \in \widetilde{\mathcal{C}}, \end{aligned} \tag{47}$$

where $\boldsymbol{\mathcal{H}}$ and $\boldsymbol{g}$ depend on the scenario (see Theorem 1 and 2). The superscript (i) has been dropped for the simplicity of the notation.

A. Energy Constraint

In this case, the constraint set is given by $\|\widetilde{\boldsymbol{a}}\|_2^2 \leq e_s$. Therefore, the resulting QCQP is a convex optimization problem. This problem can be solved using the results of the analytical solution provided in [30]. It should be noted that applying the energy constraint at each transmit antenna separately in widely separated MIMO radars does not affect the convexity of the QCQP in (47).

B. Peak-to-Average Power Ratio (PAR) Constraint

For PAR constraint, we consider the problem for the two cases of the colocated and the widely separated MIMO radar systems.

a) Colocated MIMO Radars: In this case, transmit antennas are close to each other and hence the same PAR constraint is applied to the signal transmitted from all antennas. Thus, the optimization problem in the ith iteration can be cast as

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{\mathcal{H}} \widetilde{\boldsymbol{a}} + 2\,\Re\{\boldsymbol{g}^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \max_{n=1,\cdots,NN_T} |\widetilde{a}(n)|^2 \leq \gamma \\ & \|\widetilde{\boldsymbol{a}}\|_2^2 = NN_T, \end{aligned} \tag{48}$$

with $\widetilde{a}(n)$ denoting the nth element of the vector $\widetilde{\boldsymbol{a}}$ and γ being the maximum peak power. The QCQP in (48) is NP-hard in general [27]. However, it can be equivalently rewritten in the form of

$$\begin{aligned} \max_{\hat{\boldsymbol{a}}} \quad & \hat{\boldsymbol{a}}^H \widetilde{\boldsymbol{\mathcal{H}}}\, \hat{\boldsymbol{a}} \\ \text{subject to} \quad & \max_{n=1,\cdots,NN_T} |\widetilde{a}(n)|^2 \leq \gamma \\ & \|\widetilde{\boldsymbol{a}}\|_2^2 = NN_T, \end{aligned} \tag{49}$$

with $\hat{\boldsymbol{a}} = [\widetilde{\boldsymbol{a}}^T,\ 1]^T$, $\widetilde{\boldsymbol{\mathcal{H}}} = \mu \boldsymbol{I}_{NN_T+1} - \boldsymbol{\mathcal{K}}$, and $\boldsymbol{\mathcal{K}} = \begin{bmatrix} \boldsymbol{\mathcal{H}} & \boldsymbol{g} \\ \boldsymbol{g}^H & 0 \end{bmatrix}$, for any $\mu > \lambda_{\max}(\boldsymbol{\mathcal{K}})$. Now, according to [15, 27], the problem in (49) can be solved iteratively via solving the "*nearest-vector*" problem at the $(l+1)$th iteration

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}^{(l+1)}} \quad & \|\widetilde{\boldsymbol{a}}^{(l+1)} - \boldsymbol{a}^{(l)}\|_2^2 \\ \text{subject to} \quad & \max_{n=1,\cdots,NN_T} |\widetilde{a}^{(l+1)}(n)|^2 \leq \gamma \\ & \|\widetilde{\boldsymbol{a}}^{(l+1)}\|_2^2 = NN_T, \end{aligned} \tag{50}$$

where $\boldsymbol{a}^{(l)}$ represents the vector containing the first NN_T entries of $\widetilde{\boldsymbol{\mathcal{H}}}\hat{\boldsymbol{a}}^{(l)}$. A recursive algorithm is proposed in [31] for solving (50). It is worth noting that the total transmit energy in (50) is considered to be NN_T. However, if another energy level is required, the final solution can be scaled to meet the desired energy level [15].

b) Widely Separated MIMO Radars: Different PAR constraints (with parameters γ_k) might be applied to the code transmitted by each antenna in widely separated systems. Letting $\widetilde{\boldsymbol{a}} = [\widetilde{\boldsymbol{a}}_1^T,\ \ \widetilde{\boldsymbol{a}}_2^T,\ \cdots\ ,\widetilde{\boldsymbol{a}}_{N_T}^T]^T$, (48) is modified to

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{\mathcal{H}} \widetilde{\boldsymbol{a}} + 2\,\Re\{\boldsymbol{g}^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \max_{n=1,\cdots,N} |\widetilde{a}_k(n)|^2 \leq \gamma_k, \forall k \\ & \|\widetilde{\boldsymbol{a}}_k\|_2^2 = N, \quad k = 1, \cdots, N_T. \end{aligned} \tag{51}$$

where $\widetilde{a}_k(n)$ is the $n$the element of the vector $\widetilde{\boldsymbol{a}}_k$. Similar to the colocated case, (51) can be written in terms of $\hat{\boldsymbol{a}} = [\widetilde{\boldsymbol{a}}^T,\ 1]^T$ as

$$\begin{aligned} \max_{\widetilde{\boldsymbol{a}}} \quad & \hat{\boldsymbol{a}}^H \widetilde{\boldsymbol{\mathcal{H}}}\ \hat{\boldsymbol{a}} \\ \text{subject to} \quad & \max_{n=1,\cdots,N} |\widetilde{a}_k(n)|^2 \leq \gamma_k, \forall k \\ & \|\widetilde{\boldsymbol{a}}_k\|_2^2 = N, \quad k = 1, \cdots, N_T. \end{aligned} \tag{52}$$

Iterative tackling of the above problem leads to the following optimization in the $(l+1)$th iteration

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}^{(l+1)}} \quad & \|\widetilde{\boldsymbol{a}}^{(l+1)} - \boldsymbol{a}^{(l)}\|_2^2 \\ \text{subject to} \quad & \max_{n=1,\cdots,N} |\widetilde{a}_k^{(l+1)}(n)|^2 \leq \gamma_k, \forall k \\ & \|\widetilde{\boldsymbol{a}}_k^{(l+1)}\|_2^2 = N, \quad k = 1, \cdots, N_T. \end{aligned} \tag{53}$$

with $\boldsymbol{a}^{(l)}$ being the first NN_T elements of $\widetilde{\boldsymbol{\mathcal{H}}}\hat{\boldsymbol{a}}^{(l)}$ (and being partitioned similar to $\widetilde{\boldsymbol{a}}$-see above). This problem is equivalent to

$$\begin{aligned} \min_{\{\widetilde{\boldsymbol{a}}_k^{(l+1)}\}_k} \quad & \sum_{k=1}^{N_T} \|\widetilde{\boldsymbol{a}}_k^{(l+1)} - \boldsymbol{a}_k^{(l)}\|_2^2 \\ \text{subject to} \quad & \max_{n=1,\cdots,N} |\widetilde{a}_k^{(l+1)}(n)|^2 \leq \gamma_k, \forall k \\ & \|\widetilde{\boldsymbol{a}}_k^{(l+1)}\|_2^2 = N, \quad k = 1, \cdots, N_T. \end{aligned} \tag{54}$$

Interestingly, the optimization problem in (54) is separable w.r.t. k and hence, can be solved via solving the following set of N_T problems

$$\begin{aligned} \min_{\{\widetilde{\boldsymbol{a}}^{(l+1)}\}_k} \quad & \|\widetilde{\boldsymbol{a}}_k^{(l+1)} - \boldsymbol{a}_k^{(l)}\|_2^2 \\ \text{subject to} \quad & \max_{n=1,\cdots,N} |\widetilde{a}_k^{(l+1)}(n)|^2 \leq \gamma_k \\ & \|\widetilde{\boldsymbol{a}}_k^{(l+1)}\|_2^2 = N, \quad k = 1, \cdots, N_T, \end{aligned} \tag{55}$$

whose solution is similar to that of (50).

C. Similarity Constraint

In this case, the problem of code design at the ith iteration can be cast as [20]

$$\begin{aligned} \min_{\widetilde{\boldsymbol{a}}} \quad & \widetilde{\boldsymbol{a}}^H \boldsymbol{\mathcal{H}} \widetilde{\boldsymbol{a}} + 2\,\Re\{\boldsymbol{g}^H \widetilde{\boldsymbol{a}}\} \\ \text{subject to} \quad & \|\widetilde{\boldsymbol{a}} - \widetilde{\boldsymbol{a}}_0\|_2^2 \leq \epsilon, \ \ \|\widetilde{\boldsymbol{a}}\|_2^2 = e_s, \end{aligned} \tag{56}$$

where $\widetilde{\boldsymbol{a}}_0$ is a code with "*good*" properties and ϵ is the similarity threshold [5, 32]. The QCQP in (56) has a hidden convexity. Indeed, the semidefinite relaxation of (56) is tight and the resulting SDP can be written as

$$\begin{aligned} \min_{\boldsymbol{Z}} \quad & \operatorname{tr}(\boldsymbol{A}_1 \boldsymbol{Z}) \\ \text{subject to} \quad & \operatorname{tr}(\boldsymbol{A}_2 \boldsymbol{Z}) = e_s, \ \operatorname{tr}(\boldsymbol{A}_3 \boldsymbol{Z}) \leq 0, \ \operatorname{tr}(\boldsymbol{A}_4 \boldsymbol{Z}) = 1, \end{aligned} \tag{57}$$

with $\boldsymbol{Z} = \begin{bmatrix} \widetilde{\boldsymbol{a}}\widetilde{\boldsymbol{a}}^H & \widetilde{\boldsymbol{a}}t^* \\ \widetilde{\boldsymbol{a}}^H t & |t|^2 \end{bmatrix}$ where t is an auxiliary variable and

$$\boldsymbol{A}_1 = \boldsymbol{\mathcal{K}}, \qquad \boldsymbol{A}_2 = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & 0 \end{bmatrix},$$

$$\boldsymbol{A}_3 = \begin{bmatrix} \boldsymbol{I} & -\widetilde{\boldsymbol{a}}_0 \\ -\widetilde{\boldsymbol{a}}_0^H & \|\widetilde{\boldsymbol{a}}_0\|_2^2 - \epsilon \end{bmatrix}, \quad \boldsymbol{A}_4 = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & 1 \end{bmatrix}.$$

Now, if $[(\widetilde{\boldsymbol{a}}_\star)^T \quad t_\star]^T$ is a solution to (57), then $\widetilde{\boldsymbol{a}}^{(k)} = \widetilde{\boldsymbol{a}}_\star / t_\star$ is an optimal solution to (56) (see [17, 33]).

The proposed method for dealing with the constrained code design problem is summarized in Algorithm 1. Note that for both colocated and widely separated systems, a QCQP is solved in each iteration of the proposed method. The parameters of the objective function in the QCQP depends on the system type (colocated or widely separated) and the employed design metric. We herein remark on the fact that applying the devised method to the constrained problems leads to a monotonically increasing sequence of the objective values of the design problem which guarantees the convergence of the sequence (see [15, 28] for details).

Remark 2 (computational complexity): the computational complexity of the proposed method per iteration can be taken into account[8] by noting the fact that according to Algorithm 1, each iteration of the method consists of first updating the parameters of the associated QCQP and second, solving the QCQP. We consider the order of the computational complexity (per iteration) assuming that the matrix multiplications (and hence e.g. the inverse) can be performed with complexity of $\mathcal{O}(n^{2.3})$ for a matrix in $\mathbb{C}^{n \times n}$ [34]. Also, without loss of generality, we consider typical cases in which the complexity for updating the aforementioned parameters is mainly determined by the required matrix inversions (when compared to other matrix multipications/manipulations). For the colocated MIMO system, at each iteration, first,

[8] The computational complexity is linear w.r.t. the number of iterations.

Algorithm 1 The Proposed Method for the Constrained Code Design Using Information Theoretic Criteria $\mathcal{I} \in \{\mathcal{J}, \mathcal{M}\}$

1: Initialize $\widetilde{\boldsymbol{a}}$ with a random vector in $\mathbb{C}^{NN_T}$ and set the iteration number i to 0.

2: Solve the QCQP of Theorem 1 or Theorem 2 (depending on the desired system and metric $\mathcal{I}$) via

- the results of Section IV-A for the energy constraint,
- considering (50) for the PAR constraint in colocated systems and considering (55) for the PAR constraint in widely separated systems,
- the results of Section IV-C for the similarity constraint,

to obtain $\widetilde{\boldsymbol{a}}^{(i+1)}$; set $l \leftarrow i + 1$.

3: Update the parameters of the objective function in the QCQP of Step 2 according to the desired system and metric $\mathcal{I}$.

4: Repeat steps 2 and 3 until a pre-defined stop criterion is satisfied, e.g., the change in successive values of the metric is less than a predefined threshold ξ.

the inverse of the matrix $\boldsymbol{B}_{\mathcal{M}}$ must be computed possessing a complexity of $\mathcal{O}((NN_R)^{2.3})$. Second, a QCQP must be solve that its computational complexity depends on the constraint set: i) The QCQP with energy constraint can be solved via the semi-closed-form solution in [30] possessing a complexity of $\mathcal{O}((NN_T)^{2.3})$. ii) The QCQP for the PAR constraint is NP-hard and is tackled via an iterative algorithm. This algorithm is initialized by the computation of the principal eigenvalue of the matrix $\mathcal{K}$ with complexity of $\mathcal{O}((NN_T)^2)$; then, each iteration of the aforementioned algorithm is performed efficiently, e.g., with a complexity of $\mathcal{O}(NN_T)$ for the unimodular case. iii) For the similarity constraint, the QCQP is recast as an SDP with $\mathcal{O}((NN_T)^{4.5})$ complexity [33]. Using above discussion, the case of widely separated systems can be dealt with straightforwardly.

V. Extensions of the Proposed Method

In practice, there might exist uncertainties w.r.t. a priori knowledge of the interference statistics which should reasonably be considered in the code design stage. In this section, we model the uncertainties of clutter and noise covariances using their spectral norms. That is, for colocated systems, it is assumed that for covariance matrices of noise and clutter we have

$$\|\boldsymbol{R}_v - \boldsymbol{R}^0\|_2 \leq \zeta_n, \ \ \|\boldsymbol{Q} - \boldsymbol{Q}^0\|_2 \leq \zeta_c,$$

respectively. For the widely separated case, we suppose different uncertainties at various receive antennas as

$$\|\boldsymbol{R}_{v_k} - \boldsymbol{R}_{v,k}^0\|_2 \leq \zeta_{n,k}, \ \ \|\boldsymbol{R}_{c_k} - \boldsymbol{R}_{c,k}^0\|_2 \leq \zeta_{c,k}, \tag{58}$$

where $\boldsymbol{R}^0, \boldsymbol{Q}^0$, $\boldsymbol{R}_{v,k}^0$, and $\boldsymbol{R}_{c,k}^0$ are given psd matrices. Note that the scalar parameters ζ_n, ζ_c, $\zeta_{n,k}$, and $\zeta_{c,k}$ rule the corresponding uncertainty regions. In such cases, we can robustify the design via improving the system performance in the worst case scenario [7, 8, 29]. Hence, a robust constrained code design problem can be cast as

$$\begin{aligned}
\max_{\boldsymbol{A}\in\mathcal{C}} \ \min_{\boldsymbol{R}_v,\boldsymbol{Q}} \quad & f_{\mathcal{I}} \\
\text{subject to} \quad & \|\boldsymbol{R}_v - \boldsymbol{R}^0\|_2 \leq \zeta_n \\
& \|\boldsymbol{Q} - \boldsymbol{Q}^0\|_2 \leq \zeta_c \\
& J = \boldsymbol{k}^H[(\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})\boldsymbol{Q}(\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})^H + \boldsymbol{R}_{\boldsymbol{v}}]^{-1}\boldsymbol{k} \\
& \boldsymbol{k} = [\boldsymbol{I}_{N_R} \otimes (\boldsymbol{A} \odot \boldsymbol{P})]\boldsymbol{b} \\
& \boldsymbol{R}_v \succeq \boldsymbol{0}, \ \boldsymbol{Q} \succeq \boldsymbol{0},
\end{aligned} \tag{59}$$

for colocated MIMO radars with $\mathcal{I} \in \{\mathcal{M}, \mathcal{J}\}$ and $f_{\mathcal{M}} = \log(1+J)$, $f_{\mathcal{J}} = J + 1/(J+1)$; and

$$\begin{aligned}
\max_{\boldsymbol{A}\in\mathcal{C}} \ \min_{\{\boldsymbol{R}_{v_k},\boldsymbol{R}_{c_k}\}} \quad & f_{W,\mathcal{I}} \\
\text{subject to} \quad & \|\boldsymbol{R}_{v_k} - \boldsymbol{R}_{v,k}^0\|_2 \leq \zeta_{n,k}, \forall k \\
& \|\boldsymbol{R}_{c_k} - \boldsymbol{R}_{c,k}^0\|_2 \leq \zeta_{c,k}, \forall k \\
& \boldsymbol{J}_k = \boldsymbol{Y}_k^H(\boldsymbol{A} \odot \boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + \boldsymbol{R}_{v_k})^{-1} \\
& \qquad (\boldsymbol{A} \odot \boldsymbol{P}_k)\boldsymbol{Y}_k \\
& \boldsymbol{R}_{v_k} \succeq \boldsymbol{0}, \ \boldsymbol{R}_{c_k} \succeq \boldsymbol{0},
\end{aligned} \tag{60}$$

for widely separated MIMO radars with $f_{W,\mathcal{M}} = \sum_{k=1}^{N_R} \log\det(\boldsymbol{I}_{N_T} + \boldsymbol{J}_k)$ and $f_{W,\mathcal{J}} = \sum_{k=1}^{N_R} \mathrm{tr}\{(\boldsymbol{I}_{N_T} + \boldsymbol{J}_k)^{-1} + \boldsymbol{J}_k\}$. Interestingly, the robust problems above can be dealt with via the devised methods in Theorem 1 and 2 considering the results of the following theorem.

Theorem 3: The solutions to the robust design (i.e., max-min) problems in (59) and (60), can be obtained by considering the following optimizations, respectively

$$\begin{aligned}
\max_{\boldsymbol{A}\in\mathcal{C}} \quad & f_{\mathcal{I}} \\
\text{subject to} \quad & J = \boldsymbol{k}^H[(\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})(\boldsymbol{Q}^0 + \zeta_c\boldsymbol{I}_{N_T N_R})(\boldsymbol{I}_{N_R} \otimes \boldsymbol{A})^H
\end{aligned}$$

$$+(\boldsymbol{R}^0+\zeta_n\boldsymbol{I}_{NN_R})]^{-1}\boldsymbol{k} \tag{61}$$

$$\boldsymbol{k}=[\boldsymbol{I}_{N_R}\otimes(\boldsymbol{A}\odot\boldsymbol{P})]\,\boldsymbol{b},$$

with $\mathcal{I}\in\{\mathcal{M},\mathcal{J}\}$ and $f_{\mathcal{M}}=\log(1+J)$, $f_{\mathcal{J}}=J+1/(J+1)$; and

$$\begin{aligned}\max_{\boldsymbol{A}\in\mathcal{C}} \quad & f_{W,\mathcal{I}} \\ \text{subject to} \quad & \boldsymbol{J}_k=\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H[\boldsymbol{A}(\boldsymbol{R}_{c,k}^0+\zeta_{c,k}\boldsymbol{I}_{N_T})\boldsymbol{A}^H \\ & \qquad +(\boldsymbol{R}_{v,k}^0+\zeta_{n,k}\boldsymbol{I}_N)]^{-1}(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k,\end{aligned} \tag{62}$$

with $f_{W,\mathcal{M}}=\sum_{k=1}^{N_R}\log\det(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)$ and $f_{W,\mathcal{J}}=\sum_{k=1}^{N_R}\mathrm{tr}\{(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k)^{-1}+\boldsymbol{J}_k\}$. ■

The problems in (61) and (62) are in the form of the problems in Theorem 1 and 2, respectively. Therefore, we have the following corollary.

Corollary 1: Solutions to the robust constrained design problems can be obtained by iteratively solving the stated QCQPs in Theorem 1 and 2 when the matrices $\{\boldsymbol{H}_{\mathcal{I}}\}$ and vectors $\{\boldsymbol{g}_{\mathcal{I}}\}$ are updated according to the results of Theorem 3 (for $\mathcal{I}\in\{\mathcal{M},\mathcal{J}\}$).

Proof. We prove this theorem for the case of widely separated MIMO radars as the colocated one is straightforward. We consider the inner minimization in (60) and observe that it is separable w.r.t. k. Furthermore, for every k, we have

$$\begin{aligned}\min_{\boldsymbol{R}_{v_k},\boldsymbol{R}_{c_k}\succeq\boldsymbol{0}} \quad & \log\det\left(\boldsymbol{I}_{N_T}+\boldsymbol{J}_k\right) \\ \text{subject to} \quad & \|\boldsymbol{R}_{v_k}-\boldsymbol{R}_{v,k}^0\|_2\le\zeta_{n,k} \\ & \|\boldsymbol{R}_{c_k}-\boldsymbol{R}_{c,k}^0\|_2\le\zeta_{c,k} \\ & \boldsymbol{J}_k=\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H+\boldsymbol{R}_{v_k})^{-1} \\ & \qquad (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k.\end{aligned} \tag{63}$$

Note also that the function $f(\boldsymbol{X})=\log\det(\boldsymbol{I}_{N_1}+\boldsymbol{X}):\mathcal{S}_+^{N_1}\to\mathbb{R}_+$ is monotonic w.r.t. $\boldsymbol{X}\succeq\boldsymbol{0}$ [20]. Hence, to obtain the solution to the above problem, we consider the following multi-objective optimization problem

$$\begin{aligned}\min_{\boldsymbol{R}_{v_k},\boldsymbol{R}_{c_k}\succeq\boldsymbol{0}} \quad & \boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H+\boldsymbol{R}_{v_k})^{-1} \\ & \qquad (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k \\ \text{subject to} \quad & \|\boldsymbol{R}_{v_k}-\boldsymbol{R}_{v,k}^0\|_2\le\zeta_{n,k} \\ & \|\boldsymbol{R}_{c_k}-\boldsymbol{R}_{c,k}^0\|_2\le\zeta_{c,k}.\end{aligned} \tag{64}$$

The constraint in (64) on $\boldsymbol{R}_{v_k}$ is indeed

$$\boldsymbol{R}_{v,k}^0 - \zeta_{n,k}\boldsymbol{I}_N \preceq \boldsymbol{R}_{v_k} \preceq \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N. \tag{65}$$

Similarly, for $\boldsymbol{R}_{c_k}$ we have $\boldsymbol{R}_{c,k}^0 - \zeta_{c,k}\boldsymbol{I}_{N_T} \preceq \boldsymbol{R}_{c_k} \preceq \boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T}$. Accordingly, the problem in (64) boils down to the problem of the following form

$$\begin{aligned}
\min_{\{\boldsymbol{R}_{v_k},\boldsymbol{R}_{c_k}\}\succeq\boldsymbol{0}} \quad & \boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H+\boldsymbol{R}_{v_k})^{-1} \\
& \qquad\qquad (\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k \\
\text{subject to} \quad & \boldsymbol{R}_{v,k}^0 - \zeta_{n,k}\boldsymbol{I}_N \preceq \boldsymbol{R}_{v_k} \preceq \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N \\
& \boldsymbol{R}_{c,k}^0 - \zeta_{c,k}\boldsymbol{I}_{N_T} \preceq \boldsymbol{R}_{c_k} \preceq \boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T}.
\end{aligned} \tag{66}$$

Therefore, the minimizer to the problem in (64) is given by (see Appendix C)

$$\boldsymbol{R}_{v_k}^\star = \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N, \quad \boldsymbol{R}_{c_k}^\star = \boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T}, \tag{67}$$

which completes the proof. $\square$

VI. Numerical Results

In this section, we provide several numerical simulations to evaluate the performance of the proposed code design algorithm for both colocated and widely separated MIMO radar systems. In the simulation setup, unless otherwise explicitly stated, a MIMO radar with $N_T = 5$ transmit antennas, $N_R = 3$ receive antennas, and the code length $N = 11$ is considered. As for the constraints, the transmit energy e_s is supposed to be equal to 1. Moreover, the phase code design is addressed meaning that the PAR constraint with $\gamma = 1$ is imposed to the code design (without loss of generality we assume that the same PAR constraint is applied to the whole transmitters, i.e., assuming the colcated scenario-see Section IV-B). For the similarity constraint, the Barker code of length $N = 11$ is considered as the reference code for each transmitter. Also, a similarity threshold $\epsilon = 0.75$ is taken into account. The proposed method is initialized by the reference code Barker at each transmitter. The threshold ξ as the stoping criterion of Algorithm 1 is set to 10^{-4}. Moreover, the `CVX` toolbox [35] is used for solving the convex optimization problems.

A. Colocated MIMO Radar Systems

For the collocated system we let the target at the angle $\theta_0 = 25$ degrees with normalized Doppler shift $f_d = 0.25$; also, the number of interfering clutter patches is supposed to be $L = 7$ around θ_0, viz. $\{\theta_l\} = \{22, 23, 24, 25, 26, 27, 28\}$. A homogenous clutter environment is dealt with herein implying that $\sigma_{c,1} = \sigma_{c,2} \cdots = \sigma_{c,L}$. Also, a colored noise with the exponential correlation shape is considered for which the (m, n)th entry of the covariance matrix is $\sigma_n^2 \rho^{|m-n|}$ with $\sigma_n^2 = 1, \rho = 0.5$. For this case, signal to noise ratio and clutter to noise ratio are defined as SNR $= \frac{\sigma_s^2}{\sigma_n^2} e_s$ and CNR $= \frac{\sigma_{c,l}^2}{\sigma_n^2} e_s$. Note that according to Theorem 1, employing mutual information is equivalent to that of J-divergence; therefore, without loss of generality, in this subsection, we include the results for the mutual information. We begin by numerically addressing the convergence of the devised method. Fig. 1 shows the values of the mutual information vs. the number of iterations for SNR $= 0$ dB and CNR $= 0$ dB. The values of the criterion increase by increasing the number of iterations such that the improvement in the successive iterations reaches the predefined threshold (see Algorithm 1). In the constrained designs, smaller values are observed for the metric compared to the case of just energy constraint due to the smaller feasibility regions. However, the performance degradation of the phase-code design is negligible which highlights the effectiveness of the design methodology. This is attributed to the fact that the proposed method directly deals with the constraint instead of employing a suboptimal synthesis procedure (see e.g., [9, 20]). Note that the size of the feasibility region and hence the performance loss for the similarity-constrained case depends on the value of similarity threshold ϵ.

Next, we illustrate the detection performance associated with the designed code via the proposed method (shown in Fig. 1). To this end, the receiver operating characteristic (ROC) of the optimal detector (15) is considered. Fig. 2 shows the ROC for the detector when the system employs the designed code as per mutual information criterion. The ROC is obtained numerically using the results of [16]. The figure also includes the performance of the systems employing random coding (with i.i.d. Gaussian elements for the code), the system with all-one codes (i.e., an uncoded system with a scaled version of $\boldsymbol{A} = \mathbf{1}$), the system with the code matrix obtained by the method in [36], and the system with quasi-orthogonal waveforms [22][20] at the transmit side. A significant performance improvement can be observed for the system that employs the designed code when compared to other systems. Such an improvement over quasi-orthogonal waveforms is also reported in [20]—however, the method in [20] is only applicable for stationary targets and when the location of the target does not overlap with that of the clutter patches. Regarding the method of [36], the improvement can be explained using the fact that this method implicitly assumes receivers

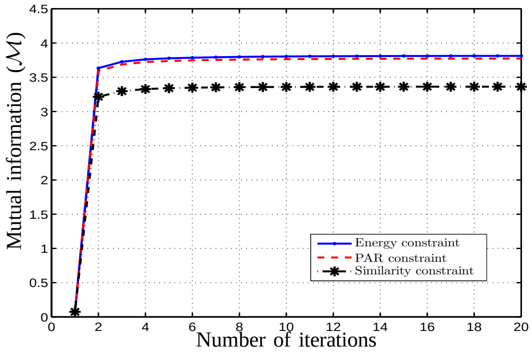

Fig. 1. The values of the mutual information versus the number of iterations for the proposed constrained code design method.

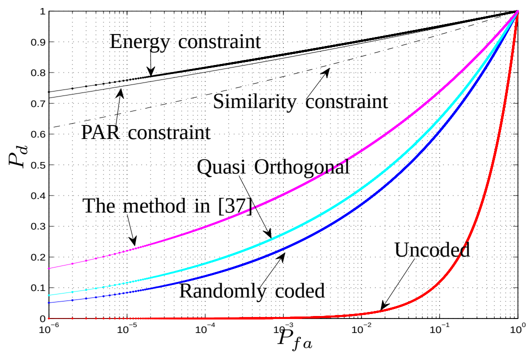

Fig. 2. ROC of the optimal detector for the designed codes via mutual information criterion and several other methods.

with matched filters not optimal filters. Similar to Fig. 1, phase-code design has a negligible loss in the detection probability compared to the energy constraint case which highlights the effectiveness of the devised method.

The effect of the target Doppler shift on the detection performance in the presence of clutter is illustrated in Fig. 3. The figure shows the probability of detection P_d vs. normalized target Doppler shift for the proposed design methodology. In the figure, we set the probability of false alarm to $P_{fa} = 10^{-4}$, $\text{SNR} = 0$ dB, and $\text{CNR} = 10$ dB. As expected, when the Doppler shift f_d is near zero, the probability of detection P_d is small. On the other hand, for a wide-range of the values of f_d, this figure shows relatively high detection probabilities. Indeed, in presence of signal-dependent clutter, the target Doppler shift helps to discriminate the target from the stationary clutter. Similar to Fig. 1 and Fig. 2, the phase-code design shows a slightly lower P_d compared to the energy constraint case.

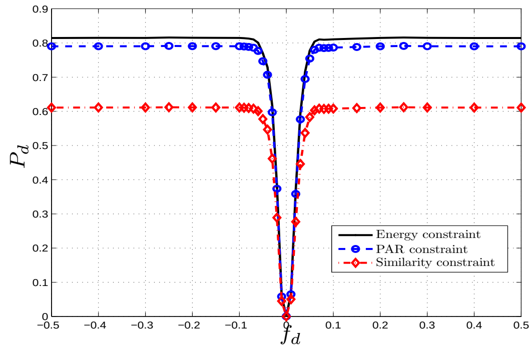

Fig. 3. Probability of detection vs. normalized Doppler shift f_d for the proposed design method.

We examine the the robustness of the system performance against angle dis-adjustments of the target for the proposed method. In Fig. 4 the values of the mutual information $\mathcal{M}$ is plotted versus angle mismatches. More precisely, we design the code assuming $\theta_0 = 25$ but compute the values of the metric for the case that the target is actually located at some other angle $\theta \neq 25$. Expectedly, such an angle dis-adjustment for the target results in performance degradation. However, the metric remains at the 95th percentile line for a relatively wide-range of the angle error (around ± 5 degrees); indeed, the values of $\mathcal{M}$ obtained by the devised algorithm show a robustness w.r.t. angle dis-adjustments of the target. Interestingly, this is observed for the phase-code design as well. In order to illustrate how degradation of the value of the mutual information affects the detection performance, we report the corresponding probability of detection. The detection probability P_d with energy constraint at the peak (no angle mismatch), 95th, and 90th percentile of the curves in Fig. 4 are respectively equal to 0.8160, 0.7829, and 0.7412. Here, a fixed probability of false alarm $P_{fa} = 10^{-4}$ is assumed. It can be seen that the detection performance degrades slightly in the considered scenario.

B. Widely separated MIMO radar systems

In this case, we define

$$\text{SNR}_k = \frac{\sigma_{s,k}^2}{\sigma_{n,k}^2} e_s, \quad \text{CNR}_k = \frac{\sigma_{c,k}^2}{\sigma_{n,k}^2} e_s \qquad k = 1, \cdots, N_R,$$

where $\sigma_{n,k}^2 = 1$ is the power of the noise, and $\sigma_{s,k}^2$ as well as $\sigma_{c,k}^2$ are associated with the target and clutter power at the kth receiver, respectively. We consider examples in which $\{\text{SNR}_k\}_k = \text{SNR}$ and $\{\text{CNR}_k\}_k = \text{CNR}$. Furthermore, for this case, the target, clutter, and interference are characterized by the covariance matrices $\{\boldsymbol{R}_{s_k}\}_{k=1}^{N_R}$, $\{\boldsymbol{R}_{c_k}\}_k$, and $\{\boldsymbol{R}_{v_k}\}_k$, respectively. We consider the exponential correlation shape

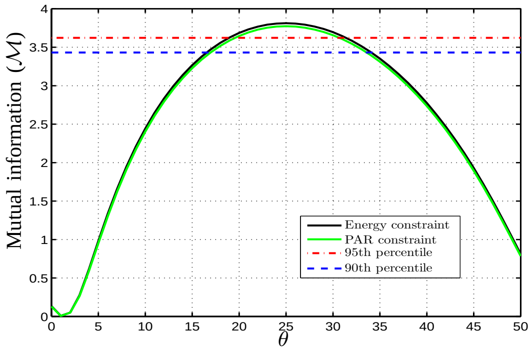

Fig. 4. An illustration of the robustness of the values of the mutual information against angle dis-adjustments for the target.

for the matrices (see above) with parameters $\{\rho_{s_k}\}_{k=1}^3 = 0.5$, $\{\rho_{c_k}\} = 0.5$, and $\{\rho_{v_k}\} = (1-1/(2k))0.5$, respectively. Also, we let $\{f_{d_{m,k}}\}_{m,k} = 0.15 f_{m,k} + 0.25$ with $f_{m,k}$ being a random variable uniformly distributed in $[-0.5, 0.5]$. Note that the results of the monotonic behavior of the proposed method and the improvement in the detection performance are similar to those of the colocated case and we do not report them herein. As opposed to the colocated case, employing mutual information and J-divergence leads to different solutions, i.e., the criteria are not generally equivalent. To address this point, we include a typical behavior of the rank of the designed code matrix $\boldsymbol{A}$. In Fig. 5, the rank of the designed code matrix $\boldsymbol{A}$ is depicted versus SNR for the information theoretic criteria $\mathcal{M}, \mathcal{J}$ for CNR $= 0$ dB with the energy constraint. As expected, various metrics lead to different solutions: at low SNR regimes, both the criteria results in rank-one solutions; by increasing the SNR, $\mathcal{M}$ results in the maximum possible rank for $\boldsymbol{A} \in \mathbb{C}^{11\times 5}$ (i.e., 5) but the achieved rank associated with $\mathcal{J}$ will be equal to 2. The figure also plots the rank corresponding to the uncoded system and the random coded system. The former is equal to 1 and the latter is equal to 5 (a behavior similar to that of the randomly coded system was numerically observed for the code matrix of [36] as well as quasi-orthogonal waveforms–see the discussions of Fig. 2). We herein remark on the fact that the solutions (and the rank behavior) depend on the considered scenario; e.g., $\{f_{d_{m,k}}\}$, the covariance matrices, etc. However, we numerically observed that $\mathcal{M}$ tends to increase the rank of the designed code matrix $\boldsymbol{A}$ at lower SNRs when compared to the $\mathcal{J}$. As to the corresponding detection performance, we observed cases in which the designed $\boldsymbol{A}$ according to $\mathcal{J}$ leads to better detection performance than that of $\mathcal{M}$ and also vice versa. Indeed, no metric is universally better than another. Note that the comparison of the aforementioned design criteria for detection performance improvement is out of the scope of this paper (the interested reader can refer to [26] and references

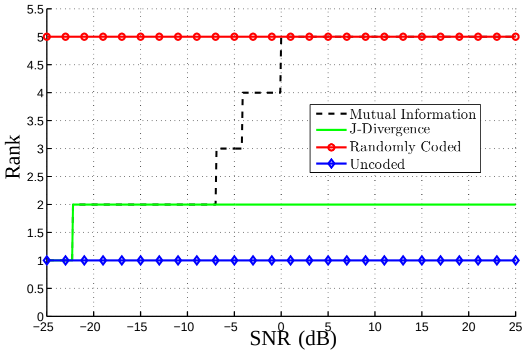

Fig. 5. The rank of the code matrix $\boldsymbol{A}$ designed based on mutual information and J-divergence metrics (under energy constraint) versus SNR.

therein for discussions on this point).

Now, we illustrate the effectiveness of the robust code design approach presented in Theorem 3 in Section V. To this end, we first assume a perfect a priori knowledge for the clutter and interference and obtain the sought code matrix $\boldsymbol{A}$. Next, we design the robust code matrix $\boldsymbol{A}_{robust}$ according to the results of the Theorem 3. The robust code design problem (see (62)) is cast with $\{\zeta_{n,k}\}_{k=1}^{3} = \{0.1\lambda_{max}(\boldsymbol{R}^{0}_{v_k})\}_k$, $\{\zeta_{c,k}\}_{k=1}^{3} = \{0.3\lambda_{max}(\boldsymbol{R}^{0}_{c_k})\}_k$, and $\{\boldsymbol{R}^{0}_{v,k}, \boldsymbol{R}^{0}_{c,k}\}_k$ equal to those employed when we had a perfect a priori knowledge (see above). Then, to examine the robustness of the code matrices $\boldsymbol{A}$ and $\boldsymbol{A}_{robust}$ w.r.t. uncertainties of the interference, we consider the uncertainty sets

$$\|\boldsymbol{R}_{v_k} - \boldsymbol{R}^{0}_{v,k}\|_2 \leq \widetilde{\zeta}_{n,k}, \quad \|\boldsymbol{R}_{c_k} - \boldsymbol{R}^{0}_{c,k}\|_2 \leq \widetilde{\zeta}_{c,k}, k = 1, 2, 3$$

and report the minimum values of the criterion[9] over the intersection of the above sets versus various values of $\widetilde{\zeta}_{n,k}, \widetilde{\zeta}_{c,k}$. Without loss of generality, in Fig. 6 we show the values of the mutual information for this setup for $\widetilde{\zeta}_{n,k} = \{0.09\lambda_{max}(\boldsymbol{R}_{v_k})\}_k$ versus $\zeta_{c,1} \triangleq \widetilde{\zeta}_{c,k}$. Note that in this example, the matrices $\{\boldsymbol{R}^{0}_{c,k}\}_{k=1}^{3}$ and the corresponding uncertainty sets for the clutter statistics at various receivers are the same; hence, we use the parameter $\zeta_{c,1}$ to rule the sets. It can be seen that the system with code matrix $\boldsymbol{A}_{robust}$ is more robust w.r.t. the uncertainties when compared to that with $\boldsymbol{A}$, i.e., the system employing the non-robust design. Note that in this figure, at $\zeta_{c,1} = 0$ the value of the robust design is reasonably larger than the non-robust design because there are uncertainties w.r.t. noise covariances $\{\boldsymbol{R}_{v_k}\}_k$.

[9]The worst (minimum) value of the criterion can be found using the results of Appendix C.

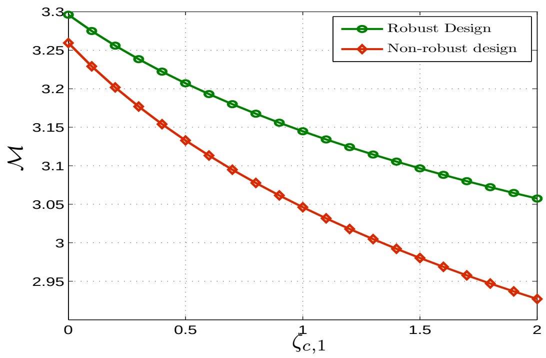

Fig. 6. An illustration of the robustness of the devised robust design method w.r.t. uncertainties of clutter and interference. Herein, we set $\widetilde{\zeta}_{n,k} = \{0.09\lambda_{max}(\boldsymbol{R}_{v_k})\}_k$ and show values of $\mathcal{M}$ versus $\zeta_{c,1}$.

C. Computational complexity

Now we consider the computational complexity of the proposed method. To this end we report the computational time of the proposed method on a standard PC with CPU CoRe i5 3GHz and 4GB RAM. Due to the fact that the computational times depend on the employed starting point, the results have been averaged on 20 random starting points. Table I shows the computational times of the proposed method for the colocated system and the widely separated one assuming the parameters of subsections VI-A and VI-B, respectively. For the colocated system, mutual information and J-divergence lead to the same performance and hence we report the case of mutual information (see Remark 1). It can be seen that the computational times for the similarity constraint are much higher than other constraints in both colocated and widely separated systems due to the heavy burden of solving an SDP at each iteration. We also numerically observed that in the widely separated case, more iterations are required to satisfy the stopping criterion $\xi = 10^{-4}$ of Algorithm 1 when compared to the colocated one (specially for energy and PAR constraints). Such an observation along with the fact that updating the parameters of QCQP in Theorem 2 is more involved can explain the higher computational times for the widely separated case comparing to those of the colocated case in the table.

VII. Conclusion

In this paper, we proposed an information theoretic design methodology for the constrained code design in colocated and widely separated MIMO radar systems. We employed mutual information and J-divergence criteria to improve detection performance of a moving target in the presence of clutter. The design problem with various constraints including energy, PAR, and similarity constraint were cast and

TABLE I

THE AVERAGE COMPUTATIONAL TIME (IN SEC.) OF THE PROPOSED METHOD IN VARIOUS CASES ON A STANDARD PC

Constraint / System type (Criterion)	Energy	PAR	Similarity
Colocated ($\mathcal{M}$)	0.11	0.09	7.87
Widely separated ($\mathcal{M}$)	1.10	0.80	11.14
Widely separated ($\mathcal{J}$)	1.15	0.89	13.46

then a method was devised based on applying minorization-maximization (MaMi) technique to obtain solutions to the design problem for the colocated and the widely separated systems. Moreover, we extended the proposed method to be robust w.r.t. uncertainties of a priori knowledge of the clutter/interference. The effectiveness of the proposed method was illustrated in numerical examples and it was shown that the system employing the proposed method outperforms other methods. Possible future research topics include robust code design w.r.t. target Doppler shift.

APPENDIX A
PROOF OF LEMMA 1

Without loss of generality, we prove the Lemma for the fixed $k = 1$ in widely separated MIMO systems. The case of the colocated systems is straightforward. According to the matrix inversion lemma [21], for the block diagonal matrix $\boldsymbol{B}_{1,\mathcal{M}} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{12}^H & \boldsymbol{B}_{22} \end{bmatrix}$ and $\boldsymbol{U} = [\boldsymbol{I}, \boldsymbol{0}]^T$ of a proper dimension, we have

$$\boldsymbol{U}^H \boldsymbol{B}_{1,\mathcal{M}}^{-1} \boldsymbol{U} = (\boldsymbol{B}_{11} - \boldsymbol{B}_{12} \boldsymbol{B}_{22}^{-1} \boldsymbol{B}_{12}^H)^{-1} \triangleq \boldsymbol{C}_1^{-1}. \tag{68}$$

On the other hand, $\log\det(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_1) = \log\det(\boldsymbol{I}_{N_T} + \boldsymbol{J}_1)$. Now, letting $\hat{\boldsymbol{A}}_k = \boldsymbol{A} \odot \boldsymbol{P}_k$ and applying the matrix inversion lemma to $(\boldsymbol{I}_{N_T} + \boldsymbol{J}_1)^{-1}$ we have

$$\begin{aligned}(\boldsymbol{I}_{N_T} + \boldsymbol{J}_1)^{-1} = \boldsymbol{I}_{N_T} - \boldsymbol{Y}_k^H \hat{\boldsymbol{A}}_k^H (\hat{\boldsymbol{A}}_k \boldsymbol{R}_{s_k} \hat{\boldsymbol{A}}_k^H \\ + \boldsymbol{A} \boldsymbol{R}_{c_k} \boldsymbol{A}^H + \boldsymbol{R}_{\boldsymbol{v}})^{-1} \hat{\boldsymbol{A}}_k \boldsymbol{Y}_k = \boldsymbol{C}_1.\end{aligned} \tag{69}$$

Hence, by using (68), (69), and (34) it is concluded that $\log\det(\boldsymbol{I}_N + \widetilde{\boldsymbol{F}}_1) = \log\det(\boldsymbol{U}^H \boldsymbol{B}_{1,\mathcal{M}}^{-1} \boldsymbol{U})$.

In order to prove the convexity of the function $\log\det(\boldsymbol{U}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{U})$ w.r.t. $\boldsymbol{B}_{\mathcal{M}}$ we resort to the minimization property in [29]. More precisely, for a fixed $\boldsymbol{B}_{\mathcal{M}}$, the matrix $\Psi_\star = \boldsymbol{B}_{11} - \boldsymbol{B}_{12} \boldsymbol{B}_{22}^{-1} \boldsymbol{B}_{12}^H (= \boldsymbol{C}_1)$

is a solution to

$$\begin{aligned} \min_{\boldsymbol{B}_{\mathcal{M}}, \Psi \succ \mathbf{0}} \quad & f(\Psi, \boldsymbol{B}_{\mathcal{M}}) = \log\det\left(\Psi^{-1}\right) \\ \text{subject to} \quad & \begin{bmatrix} \Psi & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \preceq \boldsymbol{B}_{\mathcal{M}} = \begin{bmatrix} \boldsymbol{B}_{11} & \boldsymbol{B}_{12} \\ \boldsymbol{B}_{12}^H & \boldsymbol{B}_{22} \end{bmatrix}. \end{aligned} \tag{70}$$

Hence, $\Psi_\star^{-1} = \boldsymbol{U}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{U}$ and we have $\min_\Psi f(\Psi, \boldsymbol{B}_{\mathcal{M}}) = \log\det(\boldsymbol{U}^H \boldsymbol{B}_{\mathcal{M}}^{-1} \boldsymbol{U})$ that completes the proof.

APPENDIX B
PROOF OF LEMMA 2

Using the trace properties, we can write

$$\begin{aligned} \Re\{\boldsymbol{k}^H \boldsymbol{\upsilon}_{21}\} =& \Re\{\boldsymbol{\upsilon}_{21}((\boldsymbol{I}_{N_R} \otimes [\boldsymbol{A} \odot \boldsymbol{P}])\boldsymbol{b})\} \\ =& \operatorname{tr}\{(\boldsymbol{I}_{N_R} \otimes [\boldsymbol{A} \odot \boldsymbol{P}])\, \boldsymbol{b}\boldsymbol{\upsilon}_{21}\} \\ =& \sum_{k=1}^{N_R} \operatorname{tr}\{\boldsymbol{G}_{kk}[\boldsymbol{A} \odot \boldsymbol{P}]\} \qquad (\text{with } \boldsymbol{G} \triangleq \boldsymbol{b}\boldsymbol{\upsilon}_{21}) \\ =& \Big[\sum_{k=1}^{N_R} \operatorname{vec}(\boldsymbol{G}_{kk}{}^H)\Big]^H \operatorname{vec}(\boldsymbol{A} \odot \boldsymbol{P}) \\ =& \widetilde{\boldsymbol{g}}^H(\widetilde{\boldsymbol{p}} \odot \widetilde{\boldsymbol{a}}) \qquad (\text{with } \widetilde{\boldsymbol{g}} \triangleq \sum_{k=1}^{N_R} \operatorname{vec}(\boldsymbol{G}_{kk}{}^H)) \\ =& \boldsymbol{g}^H \widetilde{\boldsymbol{a}}. \end{aligned}$$

Furthermore, we have

$$\begin{aligned} \operatorname{tr}\{\boldsymbol{\Upsilon}_{22}\boldsymbol{R}_s\} &= \operatorname{tr}\left\{\boldsymbol{\Upsilon}_{22}(\boldsymbol{I}_{N_R} \otimes [\boldsymbol{A} \odot \boldsymbol{P}])\boldsymbol{T}(\boldsymbol{I}_{N_R} \otimes [\boldsymbol{A} \odot \boldsymbol{P}])^H\right\} \\ &= \operatorname{tr}\Big\{\sum_{l=1}^{N_R}\sum_{k=1}^{N_R} [\boldsymbol{A} \odot \boldsymbol{P}]^H \boldsymbol{\Upsilon}_{22,lk} [\boldsymbol{A} \odot \boldsymbol{P}] \boldsymbol{T}_{kl}\Big\} \\ &= \sum_{l=1}^{N_R}\sum_{k=1}^{N_R} (\operatorname{vec}[\boldsymbol{A} \odot \boldsymbol{P}])^H (\boldsymbol{T}_{kl}^H \otimes \boldsymbol{\Upsilon}_{22,lk}) \operatorname{vec}[\boldsymbol{A} \odot \boldsymbol{P}] \\ &= \sum_{l=1}^{N_R}\sum_{k=1}^{N_R} (\widetilde{\boldsymbol{a}}^H \odot \widetilde{\boldsymbol{p}}^H)[\boldsymbol{T}_{kl}^H \otimes \boldsymbol{\Upsilon}_{22,lk}](\widetilde{\boldsymbol{p}} \odot \widetilde{\boldsymbol{a}}) \\ &= \widetilde{\boldsymbol{a}}^H \Big[\sum_{l=1}^{N_R}\sum_{k=1}^{N_R} \left(\boldsymbol{T}_{kl}^H \otimes \boldsymbol{\Upsilon}_{22,lk}\right) \odot \left(\widetilde{\boldsymbol{p}}\widetilde{\boldsymbol{p}}^H\right)^T\Big] \widetilde{\boldsymbol{a}}. \end{aligned}$$

Similarly, for the term associated with clutter we can write $\operatorname{tr}\{\boldsymbol{\Upsilon}_{22}\boldsymbol{R}_c\} = \widetilde{\boldsymbol{a}}^H[\sum_{l=1}^{N_R}\sum_{k=1}^{N_R}\boldsymbol{Q}_{kl}^H\otimes\boldsymbol{\Upsilon}_{22,lk}]\widetilde{\boldsymbol{a}}$. Therefore, $\operatorname{tr}\{\boldsymbol{\Upsilon}_{22}(\boldsymbol{R}_s+\boldsymbol{R}_c)\} = \widetilde{\boldsymbol{a}}^H\boldsymbol{H}\widetilde{\boldsymbol{a}}$.

APPENDIX C
THE SOLUTION TO THE PROBLEM IN (66)

According to (65),

$$\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + \boldsymbol{R}_{v_k} \preceq \boldsymbol{A}(\boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T})\boldsymbol{A}^H + \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N.$$

Therefore,

$$\begin{aligned}&\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H(\boldsymbol{A}\boldsymbol{R}_{c_k}\boldsymbol{A}^H + \boldsymbol{R}_{v_k})^{-1}(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k \succeq\\ &\boldsymbol{Y}_k^H(\boldsymbol{A}\odot\boldsymbol{P}_k)^H[\boldsymbol{A}(\boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T})\boldsymbol{A}^H + \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N]^{-1}\\ &(\boldsymbol{A}\odot\boldsymbol{P}_k)\boldsymbol{Y}_k,\end{aligned}$$

with the equality when

$$\boldsymbol{R}_{v_k}^\star = \boldsymbol{R}_{v,k}^0 + \zeta_{n,k}\boldsymbol{I}_N, \quad \boldsymbol{R}_{c_k}^\star = \boldsymbol{R}_{c,k}^0 + \zeta_{c,k}\boldsymbol{I}_{N_T}, \tag{71}$$

REFERENCES

[1] B. Friedlander, "On Signal Models for MIMO Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3655–3660, October 2012.

[2] J. Li and P. Stoica, *MIMO Radar Signal Processing*. Wiley, 2008.

[3] A. Haimovich, R. Blum, and L. Cimini, "MIMO Radar with Widely Separated Antennas," *IEEE Signal Processing Magazine*, vol. 25, no. 1, pp. 116–129, 2008.

[4] D. F. Delong and E. M. Hofstteter, "On the design of optimum radar waveforms for clutter rejection," *IEEE Trans. Inf. Theory*, vol. 13, no. 3, pp. 454–463, Jul. 1967.

[5] A. De Maio, S. De Nicola, Y. Huang, Z.-Q. Luo, and S. Zhang, "Design of phase codes for radar performance optimization with a similarity constraint," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 610–621, Feb 2009.

[6] S. Kay, "Optimal Signal Design for Detection of Gaussian Point Targets in Stationary Gaussian Clutter/Reverberation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 1, pp. 31–41, June 2007.

[7] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. De Maio, and A. Aubry, "A Doppler Robust Design of Transmit Sequence and Receive Filter in the Presence of Signal-Dependent Interference," *IEEE Transactions on Signal Processing*, vol. 62, no. 4, pp. 772–785, Feb 2014.

[8] A. Aubry, A. D. Maio, and M. M. Naghsh, "Optimizing Radar Waveform and Doppler Filter Bank via Generalized Fractional Programming," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 8, pp. 1387–1399, Dec 2015.

[9] G. Cui, H. Li, and M. Rangaswamy, "MIMO Radar Waveform Design With Constant Modulus and Similarity Constraints," *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 343–353, Jan 2014.

[10] A. De Maio and M. Lops, "Design Principles of MIMO Radar Detectors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 3, pp. 886–898, July 2007.

[11] Q. He and R. Blum, "Diversity Gain for MIMO Neyman Pearson Signal Detection," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 869–881, March 2011.

[12] A. De Maio, M. Lops, and L. Venturino, "Diversity-Integration Tradeoffs in MIMO Detection," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 5051–5061, Oct 2008.

[13] Y. Yang and R. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 1, pp. 330–343, January 2007.

[14] T. Naghibi and F. Behnia, "MIMO Radar Waveform Design in the Presence of Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 47, no. 2, pp. 770–781, April 2011.

[15] M. M. Naghsh, M. Modarres-Hashemi, S. Shahbazpanahi, M. Soltanalian, and P. Stoica, "Unified Optimization Framework for Multi-Static Radar Code Design Using Information-Theoretic Criteria," *IEEE Transactions on Signal Processing*, vol. 61, no. 21, pp. 5401–5416, Nov 2013.

[16] M. M. Naghsh and M. Modarres-Hashemi, "Exact theoretical performance analysis of optimum detector in statistical multi-input multi-output radars," *IET Radar, Sonar & Navigation*, vol. 6, no. 2, pp. 99–111, 2012.

[17] F. Gini, A. De Maio, and L. Patton, *Waveform Design and Diversity for Advanced Radar Systems*. Institution of Engineering and Technology, 2012.

[18] S. M. Karbasi, A. Aubry, V. Carotenuto, M. M. Naghsh, and M. H. Bastani, "Knowledge-based design of space-time transmit code and receive filter for a multiple-input multiple-output radar in signal-dependent interference," *Radar, Sonar & Navigation, IET*, vol. 9, no. 8, pp. 1124–1135, 2015.

[19] S. Kay, "Waveform design for multistatic radar detection," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 45, no. 3, pp. 1153–1166, 2009.

[20] B. Tang, M. M. Naghsh, and J. Tang, "Relative Entropy-Based Waveform Design for MIMO Radar Detection in the Presence of Clutter and Interference," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3783–3796, July 2015.

[21] K. B. Petersen, M. S. Pedersen *et al.*, "The matrix cookbook," *Technical University of Denmark*, vol. 7, p. 15, 2008.

[22] B. Tang, J. Tang, and Y. Peng, "MIMO radar waveform design in colored noise based on information theory," *Trans. Sig. Proc.*, vol. 58, no. 9, pp. 4684–4697, Sep. 2010.

[23] T. Strohmer and B. Friedlander, "Analysis of sparse MIMO radar," *Applied and Computational Harmonic Analysis*, vol. 37, no. 3, pp. 361 – 388, 2014.

[24] S. Kay, *Fundamentals of Statistical Signal Processing: Detection theory*. Prentice-Hall PTR, 1998.

[25] T. Cover and J. Thomas, *Elements of Information Theory*. Wiley, 2006.

[26] T. Kailath, "The Divergence and Bhattacharyya Distance Measures in Signal Selection," *IEEE Transactions on Communication Technology*, vol. 15, no. 1, pp. 52–60, February 1967.

[27] M. Soltanalian and P. Stoica, "Designing Unimodular Codes Via Quadratic Optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 5, pp. 1221–1234, March 2014.

[28] P. Stoica and Y. Selén, "Cyclic minimizers, majorization techniques, and the expectation-maximization algorithm: a refresher," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 112–114, 2004.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[30] M. M. Naghsh, M. Soltanalian, P. Stoica, and M. Modarres-Hashemi, "Radar code design for detection of moving targets," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 4, pp. 2762–2778, October 2014.

[31] J. Tropp, I. Dhillon, R. Heath, and T. Strohmer, "Designing structured tight frames via an alternating projection method," *IEEE Transactions on Information Theory*, vol. 51, no. 1, pp. 188–209, Jan 2005.

[32] J. Li, J. Guerci, and L. Xu, "Signal waveform's optimal under restriction design for active sensing," in *Fourth IEEE Workshop on Sensor Array and Multichannel Processing*, July 2006, pp. 382–386.

[33] A. Aubry, A. De Maio, A. Farina, and M. Wicks, "Knowledge-aided (potentially cognitive) transmit signal and receive filter design in signal-dependent clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 49, no. 1, pp. 93–117, Jan. 2013.

[34] A. M. Davie and A. J. Stothers, "Improved bound for complexity of matrix multiplication," *Proceedings of the Royal Society of Edinburgh: Section A Mathematics*, vol. 143, no. 02, pp. 351–369, 2013.

[35] M. Grant and S. Boyd, "CVX: Matlab Software for Disciplined Convex Programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[36] H. He, P. Stoica, and J. Li, "Designing unimodular sequence sets with good correlation-including an application to MIMO radar," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, pp. 4391–4405, 2009.