Waveform Design for MIMO Radars With Matrix Completion

Shunqiao Sun, *Student Member, IEEE*, and Athina P. Petropulu, *Fellow, IEEE*

***Abstract*—It was recently shown that MIMO radars with sparse sensing and matrix completion (MC) can significantly reduce the volume of data required by MIMO radars for accurate target detection and estimation. In MIMO-MC radars, the subsampled target returns are forwarded by the receive antennas to a fusion center, partially filling a matrix, referred to as the data matrix. The data matrix is first completed via MC techniques and then used to estimate the target parameters via standard array processing methods. This paper studies the applicability of MC theory on the data matrix arising in colocated MIMO radars using uniform linear arrays. It is shown that the data matrix coherence, and consequently the performance of MC, is directly related to the transmit waveforms. Among orthogonal waveforms, the optimum choices are those for which, any snapshot across the transmit array has a flat spectrum. The problem of waveform design is formulated as an optimization problem on the complex Stiefel manifold, and is solved via the modified steepest descent method, or the modified Newton algorithm with nonmonotone line search. Although the optimal waveforms are designed for the case of targets falling in the same range bin, sensitivity analysis is conducted to assess the performance degradation when those waveforms are used in scenarios in which the targets fall in different range bins.**

***Index Terms*—Complex Stiefel manifold, matrix completion, multiple-input multiple-output (MIMO) radar, waveform design.**

I. Introduction

UNLIKE traditional phased-array radars which transmit fully correlated signals through their transmit antennas, multiple-input multiple-output (MIMO) radars [3] transmit mutually orthogonal signals. The orthogonality allows the receive antennas to separate the transmitted signals via matched filtering. The target parameters are obtained by processing the phase shifts of the received signals. MIMO radars offer a high degree of freedom [4] and thus enable improved resolution. Since the transmitted signals are orthogonal, the transmit beam is not focused on a particular direction [5], which results in decrease of illumination power. Pulse compression techniques are typically used to improve the range resolution as well as strengthen the receive signal power. The transmit pulse can be coded or modulated, e.g., phase-coded pulse or linear frequency modulated (LFM) pulse [6].

The idea of using low-rank matrix completion (MC) techniques in MIMO radars (termed as MIMO-MC radars) was first proposed in [7], [8] as means of reducing the volume of data required by MIMO radars for accurate target detection and estimation. In particular, [8] considers a colocated pulse MIMO radar scenario with uniform linear arrays (ULAs) at the transmitter and the receiver. The transmit waveforms are pulse coded. Each receive antenna samples the target returns and forwards the obtained samples to a fusion center. Based on the data received, the fusion center formulates a matrix, referred to as "data matrix", which can then be used in standard array processing methods for target detection and estimation. If the data samples are obtained in a Nyquist fashion, and the number of targets is small relative to the number of transmit and receive antennas, the data matrix is low-rank [7]. Thus, it can be recovered from a small subset of its uniformly spaced elements via matrix completion (MC) techniques. By exploiting the latter fact, the MIMO-MC radar receive antennas obtain a small number of samples at uniformly random sampling times. Based on knowledge of the sampling instances, the fusion center populates the data matrix corresponding to Nyquist sampling in a uniform sparse fashion, and subsequently recovers the full matrix via MC techniques. This is referred to as Scheme II in [8]. Alternatively, the receive antennas perform matched filtering with a randomly selected set of transmit waveforms, and forward the results to the fusion center, partially populating the data matrix corresponding to matched filtering with all transmit waveforms. The data matrix is subsequently completed via MC. This is referred to as Scheme I in [8]. As compared to MIMO radars, both Schemes I and II reduce the number of samples that need to be forwarded to the fusion center. If the transmission occurs in a wireless fashion, this translates to savings in power and bandwidth. As compared to MIMO radars based on sparse signal recovery [9]–[11], it has been shown in [8] that MIMO-MC radars achieve similar performance but without requiring a target space grid. This is because MC techniques do not require building a basis matrix based on the discretized target space. However, this advantage is achieved at a cost of higher computation complexity, arising from the need of both matrix recovery and subspace estimation methods.

Details on the general topic of matrix completion and the conditions for matrix recovery can be found in [12]–[14]. The conditions for the applicability of MC on Scheme I of [8] can be found in [15]. In that case, and under ideal conditions, the transmit waveforms do not affect the MC performance. On the

Manuscript received February 02, 2015; revised May 11, 2015; accepted August 10, 2015. Date of publication August 20, 2015; date of current version November 17, 2015. This work was supported by the National Science Foundation under Grant ECCS-1408437. Partial results of this paper have been published in 48th Annual Asilomar Conference on Signals, Systems, and Computers (Asilomar'14) [1] and IEEE Global Conference on Signal and Information Processing (GlobalSIP'14) [2]. The guest editor coordinating the review of this manuscript and approving it for publication was Dr. Yuri Abramovich.

The authors are with the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854 USA (e-mail: shunq.sun@rutgers.edu; athinap@rutgers.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTSP.2015.2469641

other hand, for Scheme II of [8], it was shown in [8], [16] that the transmit waveforms do affect the matrix completion performance, as they directly affect the data matrix coherence [13]; a larger coherence implies that more samples need to be collected for reliable target estimation. That observation motivates the work in this paper, where we explore the relationship between matrix coherence and transmit waveforms for Scheme II of [8], and design waveforms that result in the lowest possible coherence. We should note here that waveform design for MIMO radars has been extensively studied [17]–[25] but not from the point of view of coherence minimization.

Since the waveforms are constrained to be orthogonal, we formulate the design problem as an optimization problem on the complex Stiefel manifold [26], and for its solution employ the modified steepest descent algorithm [27] and the modified Newton algorithm. The local gradient and Hessian of the cost functions are derived in closed forms. We also provide sensitivity analysis of the optimized waveforms for the scenario of targets falling in different range bins. In particular, we show that for relatively small delays, i.e., of the order of the radar pulse duration, the matrix coherence increases only slightly as the maximal range delay increases. Numerical results show that the optimized waveforms result in improved angle estimation performance.

The rest of this paper is organized as follows. Some background on noisy matrix completion and MIMO-MC radars is provided in Section II. The matrix coherence analysis results under the scenario of targets falling in the same range bin are presented in Section III, while the waveform design is addressed in Section IV. Sensitivity analysis of optimized waveforms in the scenario of targets falling in different range bins is studied in Section V. Simulations results are given in Section VI. Finally, Section VII provides some concluding remarks.

Notation

See Table I for the notation used in the paper.

TABLE I
NOTATIONS.

$\Re\{\cdot\}$:	the real part of $\{\cdot\}$
$\Im\{\cdot\}$:	the imaginary part of $\{\cdot\}$
$\mathbf{1}_L$:	the vector of length L with each element as 1
$\|\mathbf{a}\|_2$:	the Euclidean norm of a vector $\mathbf{a}$
$\mathbf{A}^*$:	the complex conjugate of a matrix $\mathbf{A}$
$\mathbf{A}^{(i)}$:	the i-th row of a matrix $\mathbf{A}$
$[\mathbf{A}]_{nm}$:	the nm-th element of a matrix $\mathbf{A}$
$\mathbf{A}^T$:	the transpose of a matrix $\mathbf{A}$
$\mathbf{A}^H$:	the conjugate transpose of a matrix $\mathbf{A}$
$\mathrm{tr}(\mathbf{A})$:	the trace of a matrix $\mathbf{A}$
$\lambda_{\min}(\mathbf{A})$:	the minimal singular value of a matrix $\mathbf{A}$
$\|\mathbf{A}\|_F$:	the Frobenius norm of a matrix $\mathbf{A}$
$\|\mathbf{A}\|_*$:	the nuclear norm of a matrix $\mathbf{A}$
$\mathrm{vec}(\mathbf{A})$:	the vectorization of a matrix $\mathbf{A}$
$\mathbf{A}\otimes\mathbf{B}$:	the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{A}\odot\mathbf{B}$:	the Hadamard product of two matrices $\mathbf{A}$ and $\mathbf{B}$
$\mathbf{I}_M$:	the identity matrix of dimension $M\times M$
$\mathbb{N}_K^+$:	the positive natural number set with cardinality of K

II. BACKGROUND

A. Matrix Completion

The recovery of a rank-r matrix $\mathbf{M}\in\mathbb{C}^{n_1\times n_2}$ based on partial knowledge of its entries depends on the coherence of its left and right singular spaces [12]–[14]. Let the compact singular value decomposition (SVD) of $\mathbf{M}$ be $\mathbf{M}=\sum_{k=1}^{r}\rho_k\mathbf{u}_k\mathbf{v}_k^H$, where $\rho_k, k=1,\ldots,r$ are the singular values, and $\mathbf{u}_k$ and $\mathbf{v}_k$ the corresponding left and right singular vectors, respectively, and let U,V be the subspaces spanned by $\mathbf{u}_k$ and $\mathbf{v}_k$, respectively. The coherence of U (and similarly of V) quantifies the similarity of the singular vectors to the standard basis and is defined as [12]

$$\mu(U)=\frac{n_1}{r}\max_{1\le i\le n_1}\left\|\mathbf{U}^{(i)}\right\|^2\in\left[1,\frac{n_1}{r}\right], \tag{1}$$

where $\mathbf{U}^{(i)}$ denotes the i-th row of matrix $\mathbf{U}=[\mathbf{u}_1,\ldots,\mathbf{u}_r]$. Matrix $\mathbf{M}$ has coherence with parameters μ_0 and μ_1 if

(A0) $\max(\mu(U),\mu(V))\le\mu_0$ for some positive μ_0.

(A1) The maximum element of the $n_1\times n_2$ matrix $\sum_{1\le i\le r}\mathbf{u}_i\mathbf{v}_i^H$ is bounded by $\mu_1\sqrt{r/(n_1n_2)}$ in absolute value, for some positive μ_1.

In fact, it was shown in [12] that if (**A0**) holds, then (**A1**) also holds with $\mu_1\le\mu_0\sqrt{r}$. If matrix $\mathbf{M}\in\mathbb{C}^{n_1\times n_2}$ satisfies (**A0**) and (**A1**), then there exist constants C and c such that if

$$m\ge C\max\left\{\mu_1^2,\mu_0^{1/2}\mu_1,\mu_0n^{1/4}\right\}nr\beta\log n$$

for some $\beta>2$, the minimizer of the optimization program

$$\min\ \|\mathbf{X}\|_*\quad\text{s.t.}\ \ \mathcal{P}_\Omega(\mathbf{X})=\mathcal{P}_\Omega(\mathbf{M}), \tag{2}$$

with $\mathcal{P}_\Omega(\mathbf{M})$ being the observation operator and Ω the set of indices of observed entries with cardinality m, is unique and equal to $\mathbf{M}$ with probability at least $1-cn^{-\beta}$, where $n=\max\{n_1,n_2\}$. For $r\le\mu_0^{-1}n^{1/5}$ the bound can be improved to $m\ge C\mu_0n^{6/5}r\beta\log n$, without affecting the probability of success [12].

Thus, the lower the coherence parameter μ_0, the fewer entries of $\mathbf{M}$ are required to estimate $\mathbf{M}$. Consequently, matrix recovery with as few entries as possible requires the matrix singular vectors to be uncorrelated to the standard basis, or equivalently, sufficiently spread [12]. The coherence lower bound is one. In this paper, we say the coherence of a given matrix is *optimal* as long as it achieves its lowest bound.

In practice, the observations are typically corrupted by noise, i.e., $[\mathbf{Y}]_{ij}=[\mathbf{M}]_{ij}+[\mathbf{E}]_{ij},(i,j)\in\Omega$, where, $[\mathbf{E}]_{ij}$ represents noise. The completion of $\mathbf{M}$ is done by solving a nuclear norm optimization problem subject to the constraint $\|\mathcal{P}_\Omega(\mathbf{X}-\mathbf{Y})\|_F\le\delta$. Assuming that the noise is zero-mean with variance σ^2, $\delta>0$ is a parameter defined as $\delta^2=(m+\sqrt{8m})\sigma^2$ [14]. Let $\hat{\mathbf{M}}$ be the solution of the nuclear norm minimization problem. It can be shown that the error norm $\left\|\mathbf{M}-\hat{\mathbf{M}}\right\|_F$ is bounded as [14]

$$\left\|\mathbf{M}-\hat{\mathbf{M}}\right\|_F\le4\sqrt{\frac{(2n_1n_2+m)\min(n_1,n_2)}{m}}\delta+2\delta \tag{3}$$

with high probability. If $\mathbf{M}$ has favorable coherence properties, the matrix completion for both noiseless and noisy cases will be stable.

B. MIMO-MC Radar

We consider the problem formulation proposed in [8] for scheme II. The scenario involves narrowband orthogonal transmit waveforms, transmitted in pulses with pulse repetition interval T_{PRI} and carrier wavelength λ, K far-filed targets at angles θ_k, and ULAs for transmission and reception, equipped with M_t transmit and M_r receive antennas, respectively, and inter-element spacing d_t and d_r, respectively (see [8, Subsection B of Section II]).

During each pulse, the m-th, $m \in \mathbb{N}^+_{M_t}$ antenna transmits a coded waveform containing N symbols $\{s_m(n)\}, n = 1, \ldots, N$ of duration T_b each. The baseband representation of the waveform is

$$\phi_m(t) = \sum_{n=1}^{N} s_m(n) \Lambda\left[\frac{t-(n-1)T_b}{T_b}\right], t \in [0, T_\phi], \quad (4)$$

where $s_m(n) = a_m(n) e^{j\varphi_m(n)}$, with $\{\varphi_m(n)\}$ uniformly distributed in $[-\pi, \pi]$, and $\{a_m(n)\}$ taking arbitrary positive values; $\Lambda(t)$ is a rectangular pulse with duration $T_\phi = NT_b$. In this paper, we relax the constant amplitude requirement to exploit more degrees of freedom for waveform design. A similar relaxation was also exploited in [24], [25]. We admit however, that this relaxation would require special attention at the radar transmitters. We will assume that the waveforms are sufficiently narrowband, i.e., $1/T_\phi \ll c/\lambda$ and targets are slow moving, i.e., $2\vartheta/\lambda \ll 1/T_\phi$, where c is the speed of light and ϑ is target speed.

In the following, we do the analysis for a scenario in which all targets are in the same range bin [3], [4]; in Section V, we will study sensitivity issues arising from targets appearing in different range bins. Suppose that the receive antennas sample the target echoes with sampling interval T_b and forward their samples to a fusion center. Let $\mathbf{X}$ be the matrix formulated at the fusion center, with each antenna contributing a row to $\mathbf{X}$. It holds that

$$\mathbf{X} = \mathbf{W} + \mathbf{J}, \quad (5)$$

where $\mathbf{J}$ is an interference/noise matrix and

$$\mathbf{W} = \mathbf{B}\mathbf{D}\mathbf{A}^T\mathbf{S}^T. \quad (6)$$

In the equation above, $\mathbf{A} \in \mathbb{C}^{M_t \times K}$ is the transmit steering matrix (respectively defined is $\mathbf{B} \in \mathbb{C}^{M_r \times K}$) with $[\mathbf{A}]_{mk} = e^{-j2\pi(m-1)\alpha_k^t}$ and $\alpha_k^t \triangleq d_t \sin(\theta_k)/\lambda, (m,k) \in \mathbb{N}^+_{M_t} \times \mathbb{N}^+_K$, where θ_k is the angle of the k-th target, or equivalently, α_k^t is the spatial frequency corresponding to the k-th target. Matrix $\mathbf{D}$ is defined as

$$\mathbf{D} \triangleq \mathrm{diag}\left(\left[\beta_1\zeta_{q1} \quad \beta_2\zeta_{q2} \quad \cdots \quad \beta_K\zeta_{qK}\right]\right), \quad (7)$$

where $\zeta_{qk} = e^{j2\pi\nu_k(q-1)T_{PRI}}$, with q denoting the pulse index, $\nu_k = 2\vartheta_k/\lambda$ denoting the Doppler shift of the k-th target, and $\{\beta_k\}_{k\in\mathbb{N}^+_K}, \{\vartheta_k\}_{k\in\mathbb{N}^+_K}$ denoting target reflection coefficients and speeds, respectively. $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(N)]^T \in \mathbb{C}^{N\times M_t}$ with $\mathbf{s}(i) = [s_1(i), \ldots, s_{M_t}(i)]^T$, is the sampled waveform matrix, with its vertical dimension corresponding to sampling along time and its horizontal dimension corresponding to sampling across the array (sampling in space). The i-th row of $\mathbf{S}$ can be thought of as the snapshot of the waveforms across the transmit antennas at sampling time i. Due to the assumed orthogonality of the waveforms, it holds that $\mathbf{S}^H\mathbf{S} = \mathbf{I}_{M_t}$ when $N \geq M_t$ [28].

When both M_r, M_t as well as N are larger than K, the noise free data matrix $\mathbf{W}$ is rank-K and can be recovered from a small number of its entries via matrix completion. This fact motivated the approach of [7], [8], which calls for subsampling the target echoes at the receive antennas in a uniformly pseudo-random fashion, partially filling the data matrix at the fusion center, and then completing the data matrix via MC techniques. This approach reduces the number of samples that need to be forwarded to the fusion center. It turns out that the applicability of MC depends on the transmit waveforms. In the next section, we derive the necessary and sufficient conditions which the transmit waveforms must satisfy so that the coherence of $\mathbf{W}$ is asymptotically optimal, i.e., it approaches 1 as M_t increases.

III. Coherence of $\mathbf{W}$ and Optimal Waveform Conditions

In this section, we analyze the coherence of $\mathbf{W}$, for the MIMO radar system defined in the previous section. In particular, (1) we provide sufficient and necessary conditions for the optimal transmit waveforms under which the coherence of $\mathbf{W}$ attains its lowest possible value; (2) under those conditions, we show asymptotic optimality of the coherence of $\mathbf{W}$ w.r.t. the number of transmit/receive antennas; (3) we show that the coherence of $\mathbf{W}$ does not depend on the Doppler shift.

A. The Coherence of $\mathbf{W}$

Let $S_i(\alpha_k^t)$ denote the discrete-time Fourier transform (DTFT) of the i-th snapshot of the transmit waveforms evaluated at spatial frequency α_k^t, i.e.,

$$S_i(\alpha_k^t) = \sum_{m=1}^{M_t} s_m(i) e^{-j2\pi(m-1)\alpha_k^t}, \quad (8)$$

where $\{\{s_m(i)\}_{m\in\mathbb{N}^+_{M_t}}\}$ are the elements in the i-th row of $\mathbf{S}$.

Before we proceed we provide a lemma (whose proof is fairly simple and can also be found in [29]) that will be useful in the subsequent theorems.

Lemma 1: For the MIMO radar system described in Section II-B with $d_t = \lambda/2$, and K targets randomly located at angles $\{\theta_k \in [-\pi/2, \pi/2]\}_{k\in\mathbb{N}^+_K}$, or equivalently, at spatial frequencies $\{\alpha_k^t \in [-1/2, 1/2]\}_{k\in\mathbb{N}^+_K}$, it holds that

$$\sum_{i=1}^{N}\sum_{k=1}^{K} \left|S_i(\alpha_k^t)\right|^2 = KM_t. \quad (9)$$

The waveforms conditions under which the coherence of $\mathbf{W}$ attains its lowest value are summarized in the following theorem.

Theorem 1: (Optimal Waveform Conditions): Consider the MIMO radar systems as defined in Section II-B, with $d_t = \lambda/2$.

The necessary condition under which the coherence of $\mathbf{W}$ *attains its lowest possible value is*

$$\sum_{k=1}^{K}\left|S_i\left(\alpha_k^t\right)\right|^2=\frac{KM_t}{N},\ \forall\, i\in\mathbb{N}_N^+. \tag{10}$$

A sufficient condition for the coherence of $\mathbf{W}$ *to attain its lowest possible value, independent of the target angles is*

$$\left|S_i\left(\alpha_l^t\right)\right|^2=\frac{M_t}{N},\ \forall\, i\in\mathbb{N}_N^+ \text{ and } \forall\,\alpha_l^t\in\left[-\frac{1}{2},\frac{1}{2}\right]. \tag{11}$$

Proof of Theorem 1: To make the proof more tractable, we break it into two parts. In the first part, we characterize the SVD of the matrix $\mathbf{W}$ in order to identify the actions that are needed in order to bound its coherence. In the second part, we derive the optimal conditions of the coded orthogonal waveforms.

1) Characterization of the SVD of $\mathbf{W}$*:* The compact SVD of $\mathbf{W}$ can be expressed as

$$\mathbf{W}=\mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \tag{12}$$

where $\mathbf{U}\in\mathbb{C}^{M_r\times K}$, $\mathbf{V}\in\mathbb{C}^{N\times K}$ such that $\mathbf{U}^H\mathbf{U}=\mathbf{I}_K$, $\mathbf{V}^H\mathbf{V}=\mathbf{I}_K$, and $\mathbf{\Lambda}\in\mathbb{R}^{K\times K}$ is a diagonal matrix containing the singular values of $\mathbf{W}$.

Consider the QR decomposition of $\mathbf{B}$, i.e., $\mathbf{B}=\mathbf{Q}_r\mathbf{R}_r$, with $\mathbf{Q}_r\in\mathbb{C}^{M_r\times K}$, such that $\mathbf{Q}_r^H\mathbf{Q}_r=\mathbf{I}_K$ and $\mathbf{R}_r\in\mathbb{C}^{K\times K}$ an upper triangular matrix. Similarly, consider the QR decomposition of $\mathbf{SA}$, i.e., $\mathbf{SA}=\mathbf{Q}_s\mathbf{R}_s$, with $\mathbf{Q}_s\in\mathbb{C}^{N\times K}$, such that $\mathbf{Q}_s^H\mathbf{Q}_s=\mathbf{I}_K$ and $\mathbf{R}_s\in\mathbb{C}^{K\times K}$ an upper triangular matrix. The matrix $\mathbf{R}_r\mathbf{D}\mathbf{R}_s^T\in\mathbb{C}^{K\times K}$ is rank-K and its SVD can be expressed as $\mathbf{R}_r\mathbf{D}\mathbf{R}_s^T=\mathbf{Q}_1\mathbf{\Delta}\mathbf{Q}_2^H$. Here, $\mathbf{Q}_1\in\mathbb{C}^{K\times K}$ is such that $\mathbf{Q}_1\mathbf{Q}_1^H=\mathbf{Q}_1^H\mathbf{Q}_1=\mathbf{I}_K$ (the same holds for $\mathbf{Q}_2$) and $\mathbf{\Delta}\in\mathbb{R}^{K\times K}$ is non-zero diagonal, containing the singular values of $\mathbf{R}_r\mathbf{D}\mathbf{R}_s^T$. Therefore, it holds that

$$\mathbf{W}=\mathbf{Q}_r\mathbf{Q}_1\mathbf{\Delta}\mathbf{Q}_2^H\mathbf{Q}_s^T=\mathbf{Q}_r\mathbf{Q}_1\mathbf{\Delta}(\mathbf{Q}_s^*\mathbf{Q}_2)^H, \tag{13}$$

which is a valid SVD of $\mathbf{W}$ since $(\mathbf{Q}_r\mathbf{Q}_1)^H\mathbf{Q}_r\mathbf{Q}_1=\mathbf{I}_K$ and $(\mathbf{Q}_s^*\mathbf{Q}_2)^H\mathbf{Q}_s^*\mathbf{Q}_2=\mathbf{I}_K$. Via the uniqueness of the singular values of a matrix, it holds that $\mathbf{\Lambda}=\mathbf{\Delta}$, thus $\mathbf{U}=\mathbf{Q}_r\mathbf{Q}_1$ and $\mathbf{V}=\mathbf{Q}_s^*\mathbf{Q}_2$.

Let $\mathbf{Q}_r^{(i)}$ denote the i-th column of $\mathbf{Q}_r$. The coherence of the row space of $\mathbf{W}$ is

$$\mu\left(U\right)=\frac{M_r}{K}\sup_{i\in\mathbb{N}_{M_r}^+}\left\|\mathbf{Q}_r^{(i)}\mathbf{Q}_1\right\|_2^2=\frac{M_r}{K}\sup_{i\in\mathbb{N}_{M_r}^+}\left\|\mathbf{Q}_r^{(i)}\right\|_2^2 \tag{14}$$

It can be seen from (14) that $\mu\left(U\right)$ is determined by $\mathbf{Q}_r$, which is only related to the receive steering matrix $\mathbf{B}$ and is independent of the transmit waveform $\mathbf{S}$. In the MIMO radar systems under ULA configuration with $d_r=\lambda/2$, under the assumption that the target angles set $\{\theta_k\}_{k\in\mathbb{N}_N^+}$ are distinct with minimal spatial frequency separation ξ_r, it was shown in [15] that

$$\mu\left(U\right)\le\frac{\sqrt{M_r}}{\sqrt{M_r}-\left(K-1\right)\sqrt{\beta_{M_r}\left(\xi_r\right)}}, \tag{15}$$

where $\beta_{M_r}\left(\xi_r\right)$ is the Fejér kernel (see (26)).

Let $\mathbf{Q}_s^{*(i)}$ and $\mathbf{S}^{*(i)}$, $i\in\mathbb{N}_N^+$ denote the i-th row of $\mathbf{Q}_s^*$ and $\mathbf{S}^*$, respectively. For the coherence of the row space of $\mathbf{W}$ we have

$$\begin{aligned}\mu\left(V\right)&=\frac{N}{K}\sup_{i\in\mathbb{N}_N^+}\left\|\mathbf{Q}_s^{*(i)}\mathbf{Q}_2\right\|_2^2=\frac{N}{K}\sup_{i\in\mathbb{N}_N^+}\left\|\mathbf{Q}_s^{*(i)}\right\|_2^2\\&=\frac{N}{K}\sup_{i\in\mathbb{N}_N^+}\left\|\mathbf{S}^{*(i)}\mathbf{A}^*(\mathbf{R}_s^*)^{-1}\right\|_2^2\\&\le\frac{N}{K}\sup_{i\in\mathbb{N}_N^+}\frac{\left\|\mathbf{S}^{*(i)}\mathbf{A}^*\right\|_2^2}{\sigma_{\min}^2\left(\mathbf{R}_s^*\right)},\end{aligned} \tag{16}$$

where

$$\begin{aligned}\sigma_{\min}^2\left(\mathbf{R}_s^*\right)&=\lambda_{\min}\left(\left(\mathbf{R}_s^*\right)^H\mathbf{R}_s^*\right)=\lambda_{\min}\left(\mathbf{R}_s^H\mathbf{R}_s\right)\\&=\lambda_{\min}\left(\mathbf{R}_s^H\mathbf{Q}_s^H\mathbf{Q}_s\mathbf{R}_s\right)=\lambda_{\min}\left(\left(\mathbf{SA}\right)^H\mathbf{SA}\right)\\&=\lambda_{\min}\left(\mathbf{A}^H\mathbf{A}\right).\end{aligned} \tag{17}$$

Here, we use the symbol $\lambda_{\min}\left(\cdot\right)$ to denote the minimal eigenvalue of a matrix. In addition, we apply the fact that the eigenvalues of a Hermitian matrix are real, and the eigenvalues of $\mathbf{X}^*$ are the complex conjugate of the eigenvalues of $\mathbf{X}$. Thus, if $\mathbf{X}$ is Hermitian, its eigenvalues are equal to eigenvalues of $\mathbf{X}^*$.

In the MIMO radar systems under ULA configuration with $d_t=\lambda/2$, under the assumption that the target angles are distinct with minimal spatial frequency separation ξ_t, it was shown in [15] that

$$\lambda_{\min}\left(\mathbf{A}^H\mathbf{A}\right)\ge M_t-\left(K-1\right)\sqrt{M_t\beta_{M_t}\left(\xi_t\right)} \tag{18}$$

where $\beta_{M_t}\left(\xi_t\right)$ is the kernel defined in (26). Therefore, regarding the coherence of the row space of $\mathbf{W}$, we have

$$\mu\left(V\right)\le\frac{N}{K}\sup_{i\in\mathbb{N}_N^+}\frac{\left\|\mathbf{S}^{*(i)}\mathbf{A}^*\right\|_2^2}{M_t-\left(K-1\right)\sqrt{M_t\beta_{M_t}\left(\xi_t\right)}}. \tag{19}$$

Next, we focus on finding the minimum of the supremum of $\left\|\mathbf{S}^{*(i)}\mathbf{A}^*\right\|_2^2$ over $i\in\mathbb{N}_N^+$, which results in the optimal waveform conditions.

2) Optimal Waveform Conditions: The transmit steering matrix $\mathbf{A}$ under a ULA configuration has a Vandermonde structure. Consequently, it holds that

$$\begin{aligned}&\left\|\mathbf{S}^{*(i)}\mathbf{A}^*\right\|_2^2\\&=\mathbf{S}^{*(i)}\mathbf{A}^*\left(\mathbf{S}^{*(i)}\mathbf{A}^*\right)^H\\&=\sum_{k=1}^{K}\sum_{m=1}^{M_t}s_m\left(i\right)e^{-j2\pi(m-1)\alpha_k^t}\sum_{m'=1}^{M_t}s_{m'}^*\left(i\right)e^{j2\pi\left(m'-1\right)\alpha_k^t}\\&=\sum_{k=1}^{K}\left|\sum_{m=1}^{M_t}s_m\left(i\right)e^{-j2\pi(m-1)\alpha_k^t}\right|^2=\sum_{k=1}^{K}\left|S_i\left(\alpha_k^t\right)\right|^2,\end{aligned} \tag{20}$$

Therefore, the coherence bound of the row space of $\mathbf{W}$ is

$$\mu\left(V\right)\le\frac{N}{K}\frac{\sup\limits_{i\in\mathbb{N}_N^+}\sum\limits_{k=1}^{K}\left|S_i\left(\alpha_k^t\right)\right|^2}{M_t-\left(K-1\right)\sqrt{M_t\beta_{M_t}\left(\xi_t\right)}}. \tag{21}$$

The lowest possible coherence bound of $\mu(V)$ can be achieved by finding waveforms that minimize $\sup_{i\in\mathbb{N}_N^+}\sum_{k=1}^K |S_i(\alpha_k^t)|^2$. This can be formulated as a min-max optimization problem subject to the constraint given in Lemma 1, i.e.,

$$\min_{\mathbf{S}} \left(\max_{i\in\mathbb{N}_N^+} \sum_{k=1}^K \left|S_i\left(\alpha_k^t\right)\right|^2 \right)$$
$$\text{s.t.} \sum_{i=1}^N \sum_{k=1}^K \left|S_i\left(\alpha_k^t\right)\right|^2 = KM_t \tag{22}$$

Since $\sum_{k=1}^K |S_i(\alpha_k^t)|^2 \geq 0$, for $i\in\mathbb{N}_N^+$, the optimal solution of the min-max optimization problem is

$$\sum_{k=1}^K \left|S_i\left(\alpha_k^t\right)\right|^2 = \frac{KM_t}{N}, \quad \forall\ i\in\mathbb{N}_N^+. \tag{23}$$

The solution, as shown in (23) depends on the specific target spatial angles $\{\alpha_k^t\}_{k\in\mathbb{N}_K^+}$. Since these angles are not known, we need to consider every possible angle θ_l in the angle space $[-\pi/2,\pi/2]$, or every $\alpha_l^t \in [-1/2,1/2]$. Thus, the optimal waveforms should sufficiently satisfy:

$$\left|S_i\left(\alpha_l^t\right)\right|^2 = \frac{M_t}{N}, \ \forall\, i\in\mathbb{N}_N^+ \ \text{ and } \ \forall\, \alpha_l^t \in \left[-\frac{1}{2},\frac{1}{2}\right]. \tag{24}$$

The condition of (24) indicates that the power spectrum of each snapshot should be flat in the spatial frequency range $\alpha_l^t\in[-1/2,1/2]$, which would be satisfied if each waveform snapshot was white noise type sequence with variance M_t/N. This completes the proof of Theorem 1. □

Under the optimal waveforms conditions stated in the Theorem 1, the coherence of matrix $\mathbf{W}$ is asymptotical optimal, as stated in the following theorem.

Theorem 2: (Coherence of $\mathbf{W}$*): Consider the MIMO radar system as presented in Section II-B and K distinct targets. Let the minimum spatial frequency separation of the targets be*

$$x = \min_{(i,j)\in\mathbb{N}_K^+\times\mathbb{N}_K^+, i\neq j} \frac{d_h}{\lambda}\left(\sin\theta_i - \sin\theta_j\right), h\in\{t,r\}. \tag{25}$$

and assume that $|x|\geq\xi_h\neq 0, h\in\{t,r\}$. Let us also define

$$\beta_{M_h}(x) = \frac{1}{M_h}\frac{\sin^2(\pi M_h x)}{\sin^2(\pi x)}, \tag{26}$$

For $d_t = d_r = \lambda/2$ and under the optimal waveform conditions stated in Theorem 1, as long as

$$K \leq \min_{h\in\{t,r\}} \left\{ \sqrt{\frac{M_h}{\beta_{M_h}(\xi_h)}} \right\}, \tag{27}$$

*the matrix $\mathbf{W}$ obeys the conditions (**A0**) and (**A1**) with*

$$\mu_0 \triangleq \max_{h\in\{t,r\}} \left\{ \frac{\sqrt{M_h}}{\sqrt{M_h} - (K-1)\sqrt{\beta_{M_h}(\xi_h)}} \right\}, \tag{28}$$

and $\mu_1 \triangleq \sqrt{K}\mu_0$ with probability 1.

Proof of Theorem 2: Following Theorem 1, for waveforms that satisfy the necessary condition (23), it holds that

$$\mu(V) \leq \inf_{\mathbf{S}} \left(\frac{N}{K} \sup_{i\in\mathbb{N}_N^+} \frac{\sum_{k=1}^K |S_i(\alpha_k^t)|^2}{M_t - (K-1)\sqrt{M_t\beta_{M_t}(\xi_t)}} \right)$$
$$= \frac{\sqrt{M_t}}{\sqrt{M_t} - (K-1)\sqrt{\beta_{M_t}(\xi_t)}}. \tag{29}$$

Consequently, under the optimal waveforms conditions, and via inequality (15), we get (28). It was shown in [12] that in the general case, $\mu_1 = \mu_0\sqrt{K}$ always holds true. Thus both (**A0**) and (**A1**) hold. □

Remarks

1) Asymptotic Optimal Coherence of $\mathbf{W}$*:* It should be noted that kernel $\beta_{M_h}(x), h\in\{t,r\}$ is a periodic function of x.

For $d_t = d_r = \lambda/2$, the spatial frequency separation corresponding to both transmit and receive arrays, satisfy $|x|\in(0,1/2]$. If $\xi \triangleq \max\{\xi_r,\xi_t\}\neq 0$, we can find a small constant ξ and $0 < \xi < 1/\min\{M_t,M_r\}$ such that the Dirichlet kernel $\sin(\pi M_h\xi)/\sin(\pi\xi) = \mathcal{O}(1)$ and the kernel $\beta_{M_h}(\xi)$ satisfy $\sqrt{\beta_{M_h}(\xi)} = \sin(\pi M_h\xi)/(\sqrt{M_h}\sin(\pi\xi)) = \mathcal{O}\left(1/\sqrt{M_h}\right)$. Consequently, the values of $\beta_{M_h}(\xi)$ decrease as $M_h, h\in\{t,r\}$ increase. Then, for any fixed K, if $\sqrt{M_h}\geq K\sqrt{\beta_{M_h}(\xi)}, h\in\{t,r\}$, or equivalently, if

$$M_h \geq K\frac{\sin(\pi M_h\xi)}{\sin(\pi\xi)} = \mathcal{O}(K), \tag{30}$$

both (15) and (29) hold. Consequently, under the optimal waveform conditions, it holds that

$$\lim_{M_t\to\infty}\mu(V) \leq \lim_{M_t\to\infty}\frac{\sqrt{M_t}}{\sqrt{M_t} - (K-1)\sqrt{\beta_{M_t}(\xi)}} = 1.$$

Since $\mu(V)\geq 1$, via the coherence definition, it must hold that under the optimal waveform conditions $\mu(V) = 1$ in the limit w.r.t. M_t. Similarly, it must hold that $\mu(U) = 1$ in the limit w.r.t. M_r. As a result, the coherence of $\mathbf{W}$ is asymptotically optimal.

It should be noted that the spatial frequency separation requirement is not restrictive. For example, in a ULA with M antennas, the spatial frequency separation of targets should be larger than the resolution of the array, i.e., $1/M$. As it can be seen in the proof of Theorem 2, Theorem 2 holds even when the spatial frequency separation of the targets is less than the resolution of the array.

2) Coherence and Doppler Shift: It can be easily seen from Theorem 2 that the coherence of $\mathbf{W}$ does not depend on the Doppler shift $\{\nu_k\}_{k\in\mathbb{N}_K^+}$ under the assumption that targets are moving slowly.

IV. Waveform Design Under Spatial Power Spectra Constraints

Theorem 1 states that, among the class of orthogonal waveforms, and for MIMO radars using ULAs, the optimal wave-

form matrix should have rows that are white-noise type functions, i.e., the waveform snapshots across the transmit antennas should be white. In this section, we propose a scheme to optimally design the transmit waveform matrix for MIMO-MC radars. In particular, the design problem is formulated as optimization on matrix manifolds [30]. Due to the orthogonality constraint on the transmit waveforms, which are the columns of matrix $\mathbf{S}$, the matrix manifold is the complex Stiefel manifold, which is non-convex. The solution can be obtained via the modified steepest descent algorithm [27], or the modified Newton algorithm with a nonmonotone line search method [31]. In the following, the derivative and Hessian of the objective function w.r.t. the waveform matrix are obtained in a closed form.

A. Problem Formulation

Let us discretize the angle space $[-\pi/2, \pi/2]$ into L phases $\{\theta_l\}_{l\in\mathbb{N}_L^+}$, corresponding to the spatial frequencies $\{\alpha_l^t\}_{l\in\mathbb{N}_L^+}$. Let $c_{il} = \mathbf{S}^{*(i)}\mathbf{A}^*(\theta_l)$ for $i \in \mathbb{N}_N^+$. According to the optimal condition (24), it holds that

$$|c_{il}|^2 = \left|S_i\left(\alpha_l^t\right)\right|^2 = \frac{M_t}{N},\ i \in \mathbb{N}_N^+,\ l \in \mathbb{N}_L^+. \tag{31}$$

Define $\mathbf{A}^* = [\mathbf{A}^*(\theta_1), \ldots, \mathbf{A}^*(\theta_L)]$ and $\mathbf{F} = \mathbf{S}^*\mathbf{A}^*$. It holds that $[\mathbf{F}\odot\mathbf{F}^*]_{il} = |c_{il}|^2$. Based on (24), let us define the objective function

$$f(\mathbf{S}) = \left\|\mathbf{F}\odot\mathbf{F}^* - \frac{M_t}{N}\mathbf{1}_N\mathbf{1}_L^T\right\|_F^2. \tag{32}$$

The waveform design problem is formulated as

$$\begin{aligned} \min\quad & f(\mathbf{S}) \\ \text{s.t.}\quad & \mathbf{S}^H\mathbf{S} = \mathbf{I}_{M_t}. \end{aligned} \tag{33}$$

Due to the orthogonal constraint, $\mathbf{S}$ belongs to the complex Stiefel manifold $\mathcal{S}(N, M_t)$, defined as

$$\mathcal{S}(N, M_t) = \left\{\mathbf{S} \in \mathbb{C}^{N\times M_t} : \mathbf{S}^H\mathbf{S} = \mathbf{I}_{M_t}\right\}. \tag{34}$$

The nonconvexity of the orthogonal constraint on the complex Stiefel manifold makes the waveform design problem challenging. In the following we adopt the modified steepest descent algorithm [27], or the modified Newton algorithm on the Stiefel manifold to solve the problem of (33).

B. Derivative and Hessian of Cost Function $f(\mathbf{S})$

In this subsection, we will address the derivative and Hessian of the cost function $f(\mathbf{S})$, defined in (32), w.r.t. the variable $\mathbf{S}$. First, based on the second order Taylor series approximation (see [27]), the cost function $f : \mathbb{C}^{N\times M_t} \to \mathbb{R}$ can be written as

$$\begin{aligned} f(\mathbf{S}+\delta\mathbf{Z}) = f(\mathbf{S}) &+ \delta\Re\left\{\mathrm{tr}\left(\mathbf{Z}^H\mathrm{D}_\mathbf{S}\right)\right\} \\ &+ \frac{\delta^2}{2}\mathrm{vec}(\mathbf{Z})^H\mathrm{H}_\mathbf{S}\mathrm{vec}(\mathbf{Z}) \\ &+ \frac{\delta^2}{2}\Re\left\{\mathrm{vec}(\mathbf{Z})^T\mathrm{C}_\mathbf{S}\mathrm{vec}(\mathbf{Z})\right\} + \mathcal{O}\left(\delta^3\right), \end{aligned} \tag{35}$$

where $\mathrm{D}_\mathbf{S} \in \mathbb{C}^{N\times M_t}$ is the derivative of f evaluated at $\mathbf{S}$, and the matrix pair $\mathrm{H}_\mathbf{S}, \mathrm{C}_\mathbf{S} \in \mathbb{C}^{NM_t\times NM_t}$ are the Hessian of f evaluated at $\mathbf{S}$. To ensure uniqueness, we require that $\mathrm{H}_\mathbf{S} = \mathrm{H}_\mathbf{S}^H, \mathrm{C}_\mathbf{S} = \mathrm{C}_\mathbf{S}^T$.

The complex-valued derivative $\mathrm{D}_\mathbf{S}$ is used in the modified steepest descent method. To calculate the Newton direction with the standard Newton method [32], we will use the second order Taylor series expansion of the function $f : \mathbb{R}^{2NM_t} \to \mathbb{R}$ in the well-known vector form with real-valued elements as

$$f(\mathbf{s}+\delta\mathbf{z}) = f(\mathbf{s}) + \delta\mathbf{z}^T\mathbf{d} + \frac{\delta^2}{2}\mathbf{z}^T\mathbf{H}\mathbf{z} + \mathcal{O}\left(\delta^3\right), \tag{36}$$

where the vector $\mathbf{s} \in \mathbb{R}^{2NM_t}$ is defined as

$$\mathbf{s} \triangleq \begin{pmatrix}\mathbf{s}^{\mathrm{re}}\\ \mathbf{s}^{\mathrm{im}}\end{pmatrix} \triangleq \begin{pmatrix}\Re\{\mathrm{vec}(\mathbf{S})\}\\ \Im\{\mathrm{vec}(\mathbf{S})\}\end{pmatrix}. \tag{37}$$

In the above, $\mathbf{d} \in \mathbb{R}^{2NM_t}$ is the derivative of $f(\mathbf{s})$ evaluated at $\mathbf{s}$, similarly defined in terms of its real and imaginary parts as $\mathbf{s}$ is in (37), and $\mathbf{H} \in \mathbb{R}^{2NM_t\times 2NM_t}$ is the Hessian of $f(\mathbf{s})$ evaluated at $\mathbf{s}$ (for definitions, see Section IV-D). The derivatives and Hessians developed in the above two Taylor series expansion forms of (35) and (36) can be transformed into each other [33].

The derivative and Hessian of the cost function $f(\mathbf{S})$, along with the derivation details are given in the Appendix.

C. Modified Steepest Descent on the Complex Stiefel Manifold

Here, we apply the modified steepest descent method of [27] to solve the optimization problem of (33). Let $\mathcal{T}_\mathbf{S}(N, M_t)$ denote the tangent space, i.e., the plane that is tangent to the complex Stiefel manifold at point $\mathbf{S} \in \mathcal{S}(N, M_t)$ [26]. The inner product in the tangent space is defined using the canonical metric [26] in the complex-value case, i.e.,

$$\langle\mathbf{Z}_1, \mathbf{Z}_2\rangle = \Re\left\{\mathrm{tr}\left[\mathbf{Z}_2^H\left(\mathbf{I} - \frac{1}{2}\mathbf{S}\mathbf{S}^H\right)\mathbf{Z}_1\right]\right\}, \tag{38}$$

for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{T}_\mathbf{S}(N, M_t)$.

Let $\mathbf{Z}^k \in \mathcal{T}_\mathbf{S}(N, M_t)$ be the steepest descent at point $\mathbf{S}^k \in \mathcal{S}(N, M_t)$ in the k-th iteration. The steepest descent algorithm starts from $\mathbf{S}^k$ and moves along $\mathbf{Z}^k$ with a step size δ, i.e.,

$$\mathbf{S}^{k+1} = \mathbf{S}^k + \delta\mathbf{Z}^k. \tag{39}$$

To preserve the orthogonality during the update steps, the new point $\mathbf{S}^{k+1}$ is projected back to the complex Stiefel manifold, i.e., $\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta\mathbf{Z}^k\right)$, where Π is the projection operator. For a matrix $\mathbf{S} \in \mathbb{C}^{N\times M_t}$ with $N \geq M_t$ and with SVD $\mathbf{S} = \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^H$, the point in the Stiefel manifold that is nearest to $\mathbf{S}$ in the Frobenius norm sense is given by $\Pi(\mathbf{S}) = \tilde{\mathbf{U}}\mathbf{I}_{N,M_t}\tilde{\mathbf{V}}^H$ [27].

The modified steepest descent is defined as follows [27]. Let $g\left(\mathbf{Z}^k\right) = f\left(\Pi\left(\mathbf{S}^k + \mathbf{Z}^k\right)\right)$ be the local cost function for $\mathbf{S}^k \in \mathcal{S}(N, M_t)$. The gradient of $g\left(\mathbf{Z}^k\right)$ at $\mathbf{Z}^k = \mathbf{0}$ under the canonical inner product (38) is

$$\tilde{\nabla}_\mathbf{S} f\left(\mathbf{S}^k\right) = \nabla_\mathbf{S} f\left(\mathbf{S}^k\right) - \mathbf{S}^k\left(\nabla_\mathbf{S} f\left(\mathbf{S}^k\right)\right)^H\mathbf{S}^k, \tag{40}$$

where $\nabla_\mathbf{S} f(\mathbf{S}) = \mathrm{D}_\mathbf{S}$ denotes the derivative of $f(\mathbf{S})$ (see (63)). Then, the modified steepest descent is $\mathbf{Z}^k = -\tilde{\nabla}_\mathbf{S} f\left(\mathbf{S}^k\right)$.

Algorithm 1 Modified steepest descent algorithm

1: **Initialize**: Choose $\mathbf{S}^0 \in \mathcal{S}(N, M_t)$ and parameters $\alpha, \eta, \epsilon \in (0,1)$, $0 < \beta < \sigma < 1$. Set $\delta = 1, C_0 = f\left(\mathbf{S}^0\right), Q_0 = 1, k = 0$.
2: **Descent direction update**: Compute the descent direction as $\mathbf{Z}^k = -\tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^k\right)$ via (40).
3: **Convergence test**: If $\left\langle \mathbf{Z}^k, \mathbf{Z}^k \right\rangle \leq \epsilon$, then stop.
4: **Line search update**: Compute $\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right)$ and $\tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^{k+1}\right)$.
If $f\left(\mathbf{S}^{k+1}\right) \geq \beta\delta \left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^k\right), \mathbf{Z}^k \right\rangle + C_k$ and $\left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^{k+1}\right), \mathbf{Z}^k \right\rangle \leq \sigma \left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^k\right), \mathbf{Z}^k \right\rangle$, then set $\delta = \alpha\delta$ and repeat Step 4.
5: **Cost update**: $Q_{k+1} = \eta Q_k + 1$, $C_{k+1} = \left[\eta Q_k C_k + f\left(\mathbf{S}^{k+1}\right)\right]/Q_{k+1}$.
6: Perform update $\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right)$, $k = k+1$. Go to Step 2.

The step size δ is chosen using a nonmonotone line search method based on [31], i.e., so that

$$f\left(\Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right)\right) \leq C_k + \beta\delta \left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^k\right), \mathbf{Z}^k \right\rangle, \tag{41}$$

$$\left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right)\right), \mathbf{Z}^k \right\rangle \geq \sigma \left\langle \tilde{\nabla}_{\mathbf{S}} f\left(\mathbf{S}^k\right), \mathbf{Z}^k \right\rangle. \tag{42}$$

Here, C_k is taken to be a convex combination of the function values $f\left(\mathbf{S}^0\right), f\left(\mathbf{S}^1\right), \ldots, f\left(\mathbf{S}^k\right)$, i.e.,

$$C_{k+1} = \frac{\left[\eta Q_k C_k + f\left(\mathbf{S}^{k+1}\right)\right]}{Q_{k+1}}, \tag{43}$$

where $Q_{k+1} = \eta Q_k + 1$, $C_0 = f\left(\mathbf{S}^0\right)$ and $Q_0 = 1$. In the above, the parameter η controls the degree of nonmonotonicity. When $\eta = 0$, the line search is the usual monotone Wolfe or Armijo line search [34]. When $\eta = 1$, then

$$C_k = \frac{1}{k+1} \sum_{i=0}^{k} f\left(\mathbf{S}^{i+1}\right). \tag{44}$$

The modified steepest descent algorithm is summarized in Algorithm 1.

D. Modified Newton Algorithm on the Complex Stiefel Manifold

With expressions for the derivative and Hessian of the cost function given in (63), (64) and (65), we can now formulate the Newton method [32] to solve the waveform design problem of (33).

First, the Newton search direction is calculated as follows. Let $\mathbf{Z}^k \in \mathbb{C}^{N \times M_t}$ denote the Newton search direction in the k-th iteration. Along the lines of [35], we arrange the complex-valued elements of $\mathbf{Z}^k$ into a real-valued vector $\mathbf{z}^k$ of length $2NM_t$, defined as

$$\mathbf{z}^k \triangleq \begin{pmatrix} \mathbf{z}^k_{\mathrm{re}} \\ \mathbf{z}^k_{\mathrm{im}} \end{pmatrix} \triangleq \begin{pmatrix} \Re\left\{\mathrm{vec}\left(\mathbf{Z}^k\right)\right\} \\ \Im\left\{\mathrm{vec}\left(\mathbf{Z}^k\right)\right\} \end{pmatrix}. \tag{45}$$

Let us also define the real-valued vector

$$\mathbf{d}^k \triangleq \begin{pmatrix} \Re\left\{\mathrm{vec}\left(\mathrm{D}_{\mathbf{S}^k}\right)\right\} \\ \Im\left\{\mathrm{vec}\left(\mathrm{D}_{\mathbf{S}^k}\right)\right\} \end{pmatrix}, \tag{46}$$

and the real-valued matrix

$$\mathbf{H}^k \triangleq \begin{bmatrix} \Re\left\{\mathrm{H}_{\mathbf{S}^k} + \mathrm{C}_{\mathbf{S}^k}\right\} & -\Im\left\{\mathrm{H}_{\mathbf{S}^k} + \mathrm{C}_{\mathbf{S}^k}\right\} \\ \Im\left\{\mathrm{H}_{\mathbf{S}^k} - \mathrm{C}_{\mathbf{S}^k}\right\} & \Re\left\{\mathrm{H}_{\mathbf{S}^k} - \mathrm{C}_{\mathbf{S}^k}\right\} \end{bmatrix}. \tag{47}$$

Following the standard Newton method, the vector $\mathbf{z}^k$ is computed as

$$\mathbf{z}^k = -\left[\mathbf{H}^k + \sigma_k \mathbf{I}\right]^{-1} \mathbf{d}^k, \tag{48}$$

where $\sigma_k \geq 0$ is chosen to make the matrix $\mathbf{H}^k + \sigma_k \mathbf{I}$ positive definite. Consequently, the complex-valued Newton search direction can be found as $\mathbf{Z}^k = \mathrm{mat}_{N \times M_t}\left(\mathbf{z}^k_{\mathrm{re}} + j\mathbf{z}^k_{\mathrm{im}}\right)$, which is the inverse vector operation defined in (45).

In the k-th iteration, the standard Newton method performs the update

$$\mathbf{S}^{k+1} = \mathbf{S}^k + \delta \mathbf{Z}^k, \tag{49}$$

where δ is the step size. Since the waveforms are on the complex Stiefel manifold, to preserve the orthogonality in the modified Newton method, the new point $\mathbf{S}^{k+1}$ is projected back to the complex Stiefel manifold, i.e.,

$$\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right). \tag{50}$$

It should be pointed out that the Hessian matrix defined in (47) is not always positive definite. In each step we choose σ_k such that $\mathbf{H}_k + \sigma_k \mathbf{I}$ is positive definite. If matrix $\mathbf{H}_k$ has nonpositive eigenvalues, σ_k should be larger than $-\lambda_{\min}\left(\mathbf{H}_k\right)$, where $\lambda_{\min}$ is the minimal eigenvalue of $\mathbf{H}_k$. In the local area of the minimum, the modified Newton update will approach the pure Newton step. However, if σ_k is chosen very large, the modified Newton search direction will be close to the negative steepest descent. Last, the step size δ could be obtained using a nonmonotone line search method along the lines of [31]. Throughout the modified Newton method, the inner product of two matrices $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathbb{C}^{N \times M_t}$ is defined as $\langle \mathbf{Z}_1, \mathbf{Z}_2 \rangle = \mathrm{tr}\left(\mathbf{Z}_1^H \mathbf{Z}_2\right)$. The modified Newton algorithm is summarized in Algorithm 2.

V. Performance of Proposed Waveforms When Targets Fall in Different Range Bins

In the above derivation of the optimum waveform we used the model of [3], which is valid for targets falling in the same range bin. When the targets fall in different range bins, the model has to be modified to account for different delays in the transmit waveforms, corresponding to the different targets. In that case, the orthogonality of the delayed waveforms cannot be guaranteed.

In this section, we consider the scenario of targets falling in different range bins, and determine the effect on performance when using the waveforms designed in Section IV, i.e., waveforms optimized under the simplified scenario of targets falling

Algorithm 2 Modified Newton algorithm

1: **Initialize**: Choose $\mathbf{S}^0 \in \mathcal{S}(N, M_t)$ and parameters $\alpha, \eta, \epsilon \in (0,1)$, $0 < \beta < \sigma < 1$. Set $\delta = 1, C_0 = f\left(\mathbf{S}^0\right), Q_0 = 1, k = 0$.
2: Compute the derivative $\mathrm{D}_{\mathbf{S}^k}$ with (63) as well as Hessian $\mathrm{H}_{\mathbf{S}^k}, \mathrm{C}_{\mathbf{S}^k}$ with (64) and (65).
3: **Convergence test**: If $\langle \mathrm{D}_{\mathbf{S}^k}, \mathrm{D}_{\mathbf{S}^k} \rangle \leq \epsilon$, then stop.
4: **Newton search direction computation**: Compute the real-valued vector $\mathbf{d}^k$ via (46), as well as the real-valued matrix $\mathbf{H}^k$ via (47). Compute the vector $\mathbf{z}^k$ via (48). The Newton search direction is arranged as $\mathbf{Z}^k = \mathrm{mat}_{N \times M_t}\left(\mathbf{z}^k_{\mathrm{re}} + j\mathbf{z}^k_{\mathrm{im}}\right)$.
5: **Line search update**: Compute $\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta \mathbf{Z}^k\right)$ and $\mathrm{D}_{\mathbf{S}^{k+1}}$. If $f\left(\mathbf{S}^{k+1}\right) \geq \beta\delta\Re\left\{\left\langle \mathrm{D}_{\mathbf{S}^{k+1}}, \mathbf{Z}^k \right\rangle\right\} + C_k$ and $\Re\left\{\left\langle \mathrm{D}_{\mathbf{S}^{k+1}}, \mathbf{Z}^k \right\rangle\right\} \leq \sigma\Re\left\{\left\langle \mathrm{D}_{\mathbf{S}^k}, \mathbf{Z}^k \right\rangle\right\}$, then set $\delta = \alpha\delta$ and repeat Step 5.
6: **Cost update**: $Q_{k+1} = \eta Q_k + 1$, $C_{k+1} = \left[\eta Q_k C_k + f\left(\mathbf{S}^{k+1}\right)\right]/Q_{k+1}$.
7: Perform update $\mathbf{S}^{k+1} = \Pi\left(\mathbf{S}^k + \delta\mathbf{Z}^k\right)$, $k = k+1$. Go to Step 2.

in the same range bin. In the following we will assume that the orthogonal waveforms satisfy (24).

To model this case we will adopt the model of [36], modified for a ULA array configuration and assuming that the Doppler shifts within a pulse can be ignored because the targets are moving slowly. Let $d_{\max}(0)$ and $d_{\min}(0)$ be the maximum and minimum ranges of K far field targets at the initial sampling time. The maximum range delay normalized by T_s is

$$N_1 = \left\lfloor \frac{2\left(d_{\max}(0) - d_{\min}(0)\right)}{cT_s} \right\rfloor. \tag{51}$$

Let us set the length of the sampling window at the receiver end to $\tilde{N} = N + N_1$, where $N_1 < N$ and $\tilde{N}T_s < T_{PRI}$; we will assume that the target reflections fall within this sampling window. During the q-th pulse, the k-th target echo received at the l-th receive antenna and the corresponding demodulated baseband signal are given in [36, equations (3) and (4), respectively]. For our scenario, the noise-free sampled received data at the l-th receive antenna during the q-th pulse can be approximated as

$$\mathbf{r}_l^T \approx \sum_{k=1}^{K} \beta_k \zeta_{qk} e^{-j2\pi 2d_k(0)/\lambda} e^{-j2\pi(l-1)\alpha_k^r} \mathbf{a}^T(\theta_k)\, \mathbf{S}^T \mathbf{C}_{\tau_k}^T, \tag{52}$$

where $\alpha_k^r \triangleq d_r \sin(\theta_k)/\lambda$, $\tau_k = \lfloor (2(d_k(0) - d_{\min}(0)))/cT_s \rfloor$ with $d_k(0)$ denoting the range of the k-th target at the initial sampling time; ζ_{qk} is the Doppler shift defined in (7); $\mathbf{C}_{\tau_k} = \left[\mathbf{0}_{N\times\tau_k} \quad \mathbf{I}_N \quad \mathbf{0}_{N\times(N_1-\tau_k)}\right]^T \in \mathbb{C}^{\tilde{N}\times N}$; $\mathbf{a}(\theta_k) = \left[1, e^{-j2\pi\alpha_k^t}, \ldots, e^{-j2\pi(M_t-1)\alpha_k^t}\right]^T$ is the transmit steering vector. By collecting samples from all receive antennas at the fusion center, the noise-free data matrix can be constructed as

$$\mathbf{W} = [\mathbf{r}_1, \ldots, \mathbf{r}_{M_r}]^T = \mathbf{B}\boldsymbol{\Sigma}\boldsymbol{\Gamma}^T, \tag{53}$$

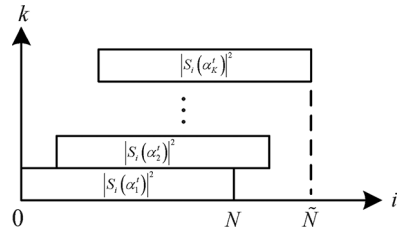

Fig. 1. Illustration of transmit waveform spectra versus the index of transmit waveform arrival at the receiver end.

where $\boldsymbol{\Gamma} = [\mathbf{C}_{\tau_1}\mathbf{S}\mathbf{a}(\theta_1), \ldots, \mathbf{C}_{\tau_K}\mathbf{S}\mathbf{a}(\theta_K)] \in \mathbb{C}^{\tilde{N}\times K}$, $\boldsymbol{\Sigma} = \mathrm{diag}\left\{\left[\beta_1\zeta_{q1}e^{-j2\pi 2d_1(0)/\lambda}, \ldots, \beta_K\zeta_{qK}e^{-j2\pi 2d_K(0)/\lambda}\right]\right\}$. We should note that although $\mathbf{W}$ and $\boldsymbol{\Sigma}$ do depend on the pulse index, for notational convenience, this dependence will not be shown.

To analyze the coherence $\mu(V)$ of $\mathbf{W}$, we do a QR decomposition of $\boldsymbol{\Gamma}$, along the lines of Section III, i.e., $\boldsymbol{\Gamma} = \mathbf{Q}_s\mathbf{R}_s$ with $\mathbf{Q}_s \in \mathbb{C}^{\tilde{N}\times K}$ such that $\mathbf{Q}_s^H\mathbf{Q}_s = \mathbf{I}_K$. Let $\mathbf{Q}_s^{*(i)}$ denote the i-th row of $\mathbf{Q}_s^*$. It holds that

$$\mu(V) = \frac{\tilde{N}}{K}\sup_{i\in\mathbb{N}_{\tilde{N}}^+}\left\|\mathbf{Q}_s^{*(i)}\right\|^2 \leq \frac{\tilde{N}}{K}\sup_{i\in\mathbb{N}_{\tilde{N}}^+}\frac{\left\|\boldsymbol{\Gamma}^{*(i)}\right\|^2}{\sigma_{\min}^2(\mathbf{R}_s^*)}, \tag{54}$$

where $\sigma_{\min}^2(\mathbf{R}_s^*) = \lambda_{\min}\left(\mathbf{R}_s^H\mathbf{Q}_s^H\mathbf{Q}_s\mathbf{R}_s\right) = \lambda_{\min}(\boldsymbol{\Phi})$ and $\boldsymbol{\Phi} = \boldsymbol{\Gamma}^H\boldsymbol{\Gamma}$ with $[\boldsymbol{\Phi}]_{k_1k_2} = \mathbf{a}^H(\theta_{k_1})\mathbf{S}^H\mathbf{C}_{\tau_{k_1}}^T\mathbf{C}_{\tau_{k_2}}\mathbf{S}\mathbf{a}(\theta_{k_2})$. It holds that $\mathbf{S}^H\mathbf{C}_{\tau_{k_1}}^T\mathbf{C}_{\tau_{k_2}}\mathbf{S} = \mathbf{S}^H\mathbf{J}_{\tau_{k_2}-\tau_{k_1}}\mathbf{S} = \mathbf{R}_{\tau_{k_2}-\tau_{k_1}}^T$, where $\mathbf{R}_n$ is the waveform covariance matrix [19] with $[\mathbf{R}_n]_{m_1m_2} = \sum_{k=n+1}^{N} s_{m_1}(k)\, s_{m_2}^*(k-n)$ and $\mathbf{J}_n$ is a shifting matrix defined as [19]

$$\mathbf{J}_n = \begin{bmatrix} \overbrace{0\cdots 0}^{n}1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}_{N\times N} = \mathbf{J}_{-n}^T, \tag{55}$$

for $n = 0, \ldots, N_1$. Since the waveforms are orthogonal it holds that $\mathbf{R}_0 = \mathbf{I}_{M_t}$, and $\boldsymbol{\Phi}$ becomes

$$\boldsymbol{\Phi} = \begin{bmatrix} M_t & \cdots & \mathbf{a}^H(\theta_1)\mathbf{R}_{\tau_K-\tau_1}^*\mathbf{a}(\theta_K) \\ \vdots & \ddots & \vdots \\ \mathbf{a}^H(\theta_K)\mathbf{R}_{\tau_K-\tau_1}^T\mathbf{a}(\theta_1) & \cdots & M_t \end{bmatrix}. \tag{56}$$

In order to compute bounds for $\lambda_{\min}(\boldsymbol{\Phi})$, we need to first compute $\mathrm{tr}(\boldsymbol{\Phi})$ and $\mathrm{tr}(\boldsymbol{\Phi}^2)$ (see [37]). It can be seen than $\mathrm{tr}(\boldsymbol{\Phi}) = KM_t$ and $\mathrm{tr}\left(\boldsymbol{\Phi}^2\right) = KM_t^2 + 2\sum_{k_2=k_1+1}^{K}\sum_{k_1=1}^{K}|[\boldsymbol{\Phi}]_{k_1k_2}|^2$ with

$$\begin{aligned}[\boldsymbol{\Phi}]_{k_1k_2} &= \mathbf{a}^H(\theta_{k_1})\mathbf{R}_{\tau_{k_2}-\tau_{k_1}}^*\mathbf{a}(\theta_{k_2}) \\ &= \sum_{m_2=1}^{M_t}\sum_{m_1=1}^{M_t} e^{j2\pi\left[(m_1-1)\alpha_{k_1}^t - (m_2-1)\alpha_{k_2}^t\right]} \\ &\quad \cdot r_{m_1m_2}^*(\tau_{k_2} - \tau_{k_1}).\end{aligned} \tag{57}$$

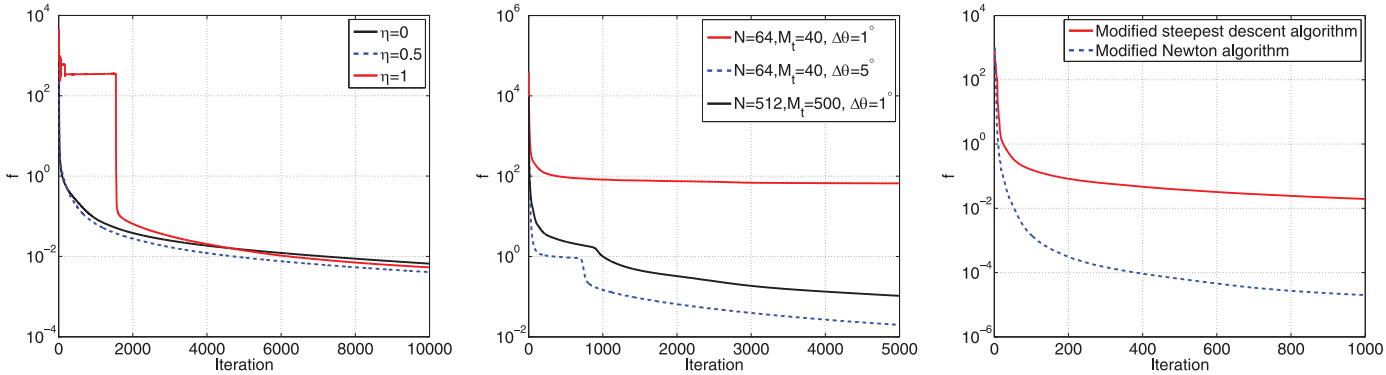

Fig. 2. Objective function of (32) vs. iteration number: (a) DOA space $[-10^\circ : 1^\circ : 10^\circ]$ and using the modified steepest descent algorithm; (b) DOA space $[-90^\circ : \Delta\theta : 90^\circ]$, $\Delta\theta = 1^\circ, 5^\circ$ and using the modified steepest descent algorithm; (c) comparison of the modified steepest descent and the modified Newton algorithms for $M_t = 20, N = 32$ and DOA space $[-5^\circ : 1^\circ : 5^\circ]$.

Then, the lower bound of $\lambda_{\min}(\mathbf{\Phi})$, can be found as (see Theorem 2.1 in [37])

$$\begin{aligned}\lambda_{\min}(\mathbf{\Phi}) &\geq \frac{\mathrm{tr}(\mathbf{\Phi})}{K} - \sqrt{K-1}\sqrt{\frac{\mathrm{tr}\left(\mathbf{\Phi}^2\right)}{K} - \left(\frac{\mathrm{tr}(\mathbf{\Phi})}{K}\right)^2} \\ &= M_t - \sqrt{1-\frac{1}{K}}\sqrt{2\sum_{k_2=k_1+1}^{K}\sum_{k_1=1}^{K}\left|[\mathbf{\Phi}]_{k_1k_2}\right|^2}.\end{aligned} \tag{58}$$

Next, we find the minimum supremum of $\left\|\mathbf{\Gamma}^{*(i)}\right\|^2$ over $i \in \mathbb{N}_{\tilde{N}}^+$. Without loss of generality, we assume that the ranges of targets are ordered as $d_1(0) \leq \cdots \leq d_K(0)$. Thus, $\tau_1 = 0, \tau_K = N_1$. To better understand the term $\left\|\mathbf{\Gamma}^{*(i)}\right\|^2$, in Fig. 1 we give an illustration of the power spectra of the transmitted waveforms corresponding to spatial frequency α_k^t, in the order in which the waveforms arrive at the receiver end. It holds that

$$\left\|\mathbf{\Gamma}^{*(i)}\right\|^2 = \begin{cases} \left|S_i(\alpha_1^t)\right|^2, 1 \leq i \leq \tau_2 \\ \sum_{k=1}^{2}\left|S_{i-\tau_k}(\alpha_k^t)\right|^2, \tau_2 < i \leq \tau_3 \\ \vdots \\ \sum_{k=1}^{K}\left|S_{i-\tau_k}(\alpha_k^t)\right|^2, N_1 < i \leq N \\ \sum_{k=2}^{K}\left|S_{i-\tau_k}(\alpha_k^t)\right|^2, N < i \leq N+\tau_2 \\ \vdots \\ \left|S_{i-N_1}(\alpha_K^t)\right|^2, N+\tau_{K-1} < i \leq N+N_1 \end{cases} \tag{59}$$

It is easy to verify that

$$\sum_{i=1}^{\tilde{N}}\left\|\mathbf{\Gamma}^{*(i)}\right\|^2 = \sum_{i=1}^{N}\sum_{k=1}^{K}\left|S_i(\alpha_k^t)\right|^2 = KM_t. \tag{60}$$

It can be seen that the maximum value of $\left\|\mathbf{\Gamma}^{*(i)}\right\|^2$ over $i \in \mathbb{N}_{\tilde{N}}^+$ is determined by the spatial power spectra $S_i(\alpha_k^t), i \in \mathbb{N}_N^+$ as well as the delay corresponding to each target. Based on the conditions of (24), it holds that $\sup_{i\in\mathbb{N}_{\tilde{N}}^+}\left\|\mathbf{\Gamma}^{*(i)}\right\|^2 = \sum_{k=1}^{K}\left|S_{i-\tau_k}(\alpha_k^t)\right|^2 = KM_t/N$.

Combining the above we get

$$\mu(V) \leq \frac{\left(1+\frac{N_1}{N}\right)M_t}{M_t - \sqrt{1-\frac{1}{K}}\sqrt{2\sum_{k_2=k_1+1}^{K}\sum_{k_1=1}^{K}\left|[\mathbf{\Phi}]_{k_1k_2}\right|^2}}. \tag{61}$$

One can see that as M_t increases, the above bound tends to $1 + N_1/N < 2$.

Corollary 1: If the orthogonal waveforms not only satisfy the sufficient condition stated in (24) but are also designed to have zero auto and cross-correlations for maximal normalized delay range N_1, i.e., $\mathbf{R}_n = \mathbf{0}_{M_t}, \quad n = 1, \ldots, N_1$, then it holds that $\mathbf{\Phi} = M_t\mathbf{I}_{M_t}$ and thus $\lambda_{\min}(\mathbf{\Phi}) = M_t$. As a result,

$$\mu(V) \leq \left(1+\frac{N_1}{N}\right) < 2. \tag{62}$$

Proof: The conclusion is straightforward from the above analysis and the proof is omitted here. □

The above analysis shows that, if the proposed waveforms were to be used in a scenario in which the targets fall in different range bins, the coherence $\mu(V)$ would increase slightly as the maximum normalized range delay increases. Fortunately, the coherence $\mu(V)$ is bounded by 2 for most of the possible values of $N_1 < N$ (see Fig. 3(b) for an example). Our analysis also indicates that if the proposed waveforms were designed to additionally have good correlation properties, they would yield the lowest upper bound of coherence $\mu(V)$. Waveforms with good correlation properties result in better range resolution and have been widely investigated [19], [38]. One could combine the condition of (24) with waveform correlation criterion used in [19], [38] to design waveforms for the scenario in which targets have different angles and fall into different range bins. Such waveforms not only result in the lowest matrix coherence but also have better range resolution. This kind of design will be subject of future study.

VI. Numerical Results

In this section, we provide numerical results to demonstrate the performance of the proposed waveform matrix

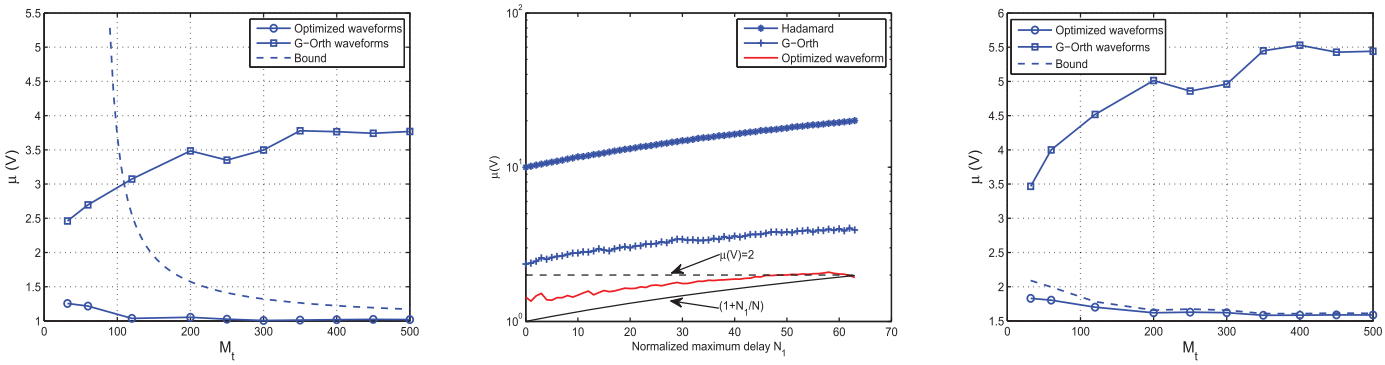

Fig. 3. Coherence $\mu(V)$ for $K = 4$ targets located at $[-10°, -5°, 0°, 1°]$. (a) $\mu(V)$ and its bound, defined in (29), for targets in the same range bin; (b) $\mu(V)$ versus N_1 with $M_t = 40$, $N = 64$ for targets in different range bins; (c) $\mu(V)$ and its bound, defined in (61), with $N_1 = N/2$ for targets in different range bins.

design schemes. Both transmit and receive antennas are configured as ULAs with $d_t = \lambda/2, d_r = M_t\lambda/2$ and carrier frequency $f_c = 5 \times 10^9$ Hz. The pulse repetition interval is $T_{PRI} = 1/4000$ s and the pulse duration is $T_\phi = N \times 8 \times 10^{-7}$ s. The target reflection coefficients $\{\beta_k\}_{k\in\mathbb{N}_K^+}$ remain constant during the processing interval [6].

A. Performance Comparison of Waveform Design Methods

We first design the waveform matrix by applying the modified steepest descent method. We take $N = 64, M_t = 40$ and the direction of arrival (DOA) space $[-10° : 1° : 10°]$. Correspondingly, $\alpha_k^t \in [-0.0868, 0.0868]$, the number of discretized angles is $L = 21$ and the steering matrix $\mathbf{A}$ has dimension 40×21. In the nonmonotone line search, we chose $\beta = 0.01$ and $\sigma = 0.99$. These values are selected by trial and error to satisfy the Wolfe conditions (41) and (42). In addition, we set $\alpha = 0.5$ to adjust the step size, and set $\epsilon = 10^{-5}$ as the stopping check value. The initial step size is set to $\delta = 0.1$. The iteration is initialized with a column-wise Hadamard matrix $\mathbf{S}^0 \in \mathcal{S}(64, 40)$, i.e., a matrix that has Hadamard sequences in its columns. The convergence of the proposed modified steepest descent algorithm for $\eta = 1, 0.5, 0$ is shown in Fig. 2(a). As it can be seen from Fig. 2(a), the objective value f under $\eta = 0.5$ decreases the fastest, while under $\eta = 1$ decreases very slowly for number of iterations less than 2000. The simulation results indicate that the performance of the line search method could be improved if historical objective values are partially utilized in each iteration, as indicated in (43). The simulation results also show that the value of the objective function, f, approaches its global minimal, i.e., 0. The corresponding optimal solution, $\mathbf{S}$, is not unique, and depends on the initial point and the step size. Based on extensive simulations, not shown here due to space constraints, all solutions result in very similar MC recovery performance (see simulations in [2], [29]).

Since the complex Stiefel manifold is not a convex set, there is no guarantee that the algorithms will converge to the global minimum. In the problem of (33), the number of equations is $M_t(M_t+1)/2+NL$, which should be less than the total available combinations, $2M_tN$. Consequently, to make the objective function zero, it must hold that $L < 2M_t - (M_t^2 + M_t)/2N$. Our simulations show that for the entire DOA space $[-90° : 1° : 90°]$, corresponding to $L = 181$, when the dimension of $\mathbf{S}$ is relatively small, e.g., $N = 64, M_t = 40$ and therefore $L > 2M_t - (M_t^2 + M_t)/2N \approx 67$, the objective value gets stuck to local minima; however, if the spacing increases, for example to 5°, corresponding to $L = 37$, the iteration converges to the global minimum. If the dimension is relatively large, e.g., $N = 512, M_t = 500$, even for small spacing, i.e., 1°, the objective value converges to its global minimum (see Fig. 2(b) for $\eta = 0$).

Next we provide a numerical example to compare the performance of the modified steepest descent algorithm and the modified Newton algorithm. Since the complexity of the Newton method increases with the size of the matrix, we conduct the comparison for $N = 32, M_t = 20$ and DOA space $[-5° : 1° : 5°]$. A column-wise Hadamard waveform matrix $\mathbf{S}^0 \in \mathbb{C}^{32\times 20}$ is used as initial search point for both algorithms. The performance comparison is illustrated in Fig. 2(c), where it can be seen that the value of the objective function, $f(\mathbf{S})$, under the modified Newton algorithm decreases much faster than that under the modified steepest descent algorithm.

B. Coherence Properties Under Optimized Waveforms

1) Targets in the Same Range Bin: We first look at the coherence properties under the optimized waveforms for the scenario in which all targets fall in the same range bin. In Fig. 3(a), we plot the coherence $\mu(V)$ of matrix $\mathbf{W}$ and its bound, defined in (29), versus the number of transmit antennas, corresponding to $K = 4$ targets located at angles $[-10°, -5°, 0°, 1°]$. The optimized waveforms for different values of M_t are obtained by solving the problem of (33) via Algorithm 1 focusing on DOA space $[-10° : 1° : 10°]$, i.e., $L = 21$. For comparison, the coherence $\mu(V)$ under the Gaussian orthogonal (G-Orth) waveform matrix is also plotted, where the results are averaged over 100 independent implementations, and in each implementation the waveforms are generated randomly. It can be seen that the average coherence under G-Orth waveforms is higher than the coherence under the optimized waveforms over the entire M_t range. On the other hand, under the optimized waveforms, our simulations show that for different number of targets, the coherence is always bounded by the bound of (29), and approaches its smallest value (not necessarily in a monotone way) when M_t increases. The simulation results in Fig. 3(a) confirm the conclusions in Theorem 2 , i.e., when the waveforms satisfy the optimal waveform conditions stated in Theorem 1 , the matrix coherence $\mu(V)$ is asymptotically optimal w.r.t. M_t. We should

note, however, that the rather big coherence difference between the optimized and the G-Orth waveforms, does not translate to substantial difference in terms of matrix recovery error. Indeed, the G-Orth waveforms perform very closely with the optimized ones when M_t becomes larger.

2) Targets in Different Range Bins: Next, we conduct simulations considering the scenario of targets falling into different range bins, and test the coherence $\mu(V)$ for different maximum normalized delay N_1 and number of transmit antennas M_t.

In Fig. 3(b), we plot the coherence $\mu(V)$ versus N_1 and for $K=4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$ for $M_t = 40, N = 64$. The maximum range $d_{\max}(0)$ is set so that N_1 takes values from 0 to $N-1$. The simulation results are averaged over 100 independent implementations, and in each implementation the ranges of the middle two targets are chosen randomly in $[d_{\min}(0), d_{\max}(0)]$. It can be seen from Fig. 3(b) that $\mu(V)$ increases as N_1 becomes larger. In addition, $\mu(V)$ under the optimized waveforms is the lowest as compared to Hadamard and G-Orth waveforms, which is slightly higher than $1 + N_1/N$ and less than 2 for most N_1 values. In Fig. 3(c), we plot $\mu(V)$ versus M_t for $K = 4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$, where $N_1 = N/2$. The results are averaged over 100 independent implementations; in each implementation the two middle target ranges are chosen randomly in $[d_{\min}(0), d_{\max}(0)]$, and G-Orth waveforms are randomly generated. It can be seen from Fig. 3(c) that the averaged $\mu(V)$ under the optimized waveforms decreases slightly as M_t increases and is bounded by (61).

Our simulation results show that when targets are not all in the same range bin, $\mu(V)$, computed under the optimized waveforms, is close to $1 + N_1/N$ for a wide range of M_t, N_1 values.

C. Matrix Recovery Error Performance

Here, we look at the MC performance as function of the portion of observed entries, p, corresponding to the optimized waveform matrix, for $K = 4$ targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$, respectively. For each configuration, both $M_r = 20$ and $M_r = 40$ antennas are tested. The targets are moving slowly with speeds $[1, 5, 10, 15]$ m/s. The signal-to-noise ratio (SNR), defined as the power of all receive signals at the receiver end over the power of noise, is set to 25 dB. The simulation results are averaged over 50 independent runs, where in each run, the noise is randomly generated. The optimized waveform matrices with $M_t = 20, N = 64$ are obtained via Algorithm 1, focusing on DOA space $[-5^\circ : 0.1^\circ : 5^\circ]$. In the simulations, the data matrix is recovered via the SVT algorithm of [39].

1) Targets in the Same Range Bin: We take all targets to fall in the same range bin. Fig. 4(a) shows the recovery error, suggesting that the optimized waveform matrix results in significantly better performance as compared to the column-wise Hadamard matrix, especially for small values of p. One can see that in order to achieve an error around 5%, MC with the optimized waveforms requires about 50% of the data matrix entries for $M_r = 20$, and 30% for $M_r = 40$. On the other hand, MC with a column-wise Hadamard matrix requires more than 60% of the data matrix entries for both $M_r = 20$ and $M_r = 40$. In the same figure, we also compare the optimized waveforms against

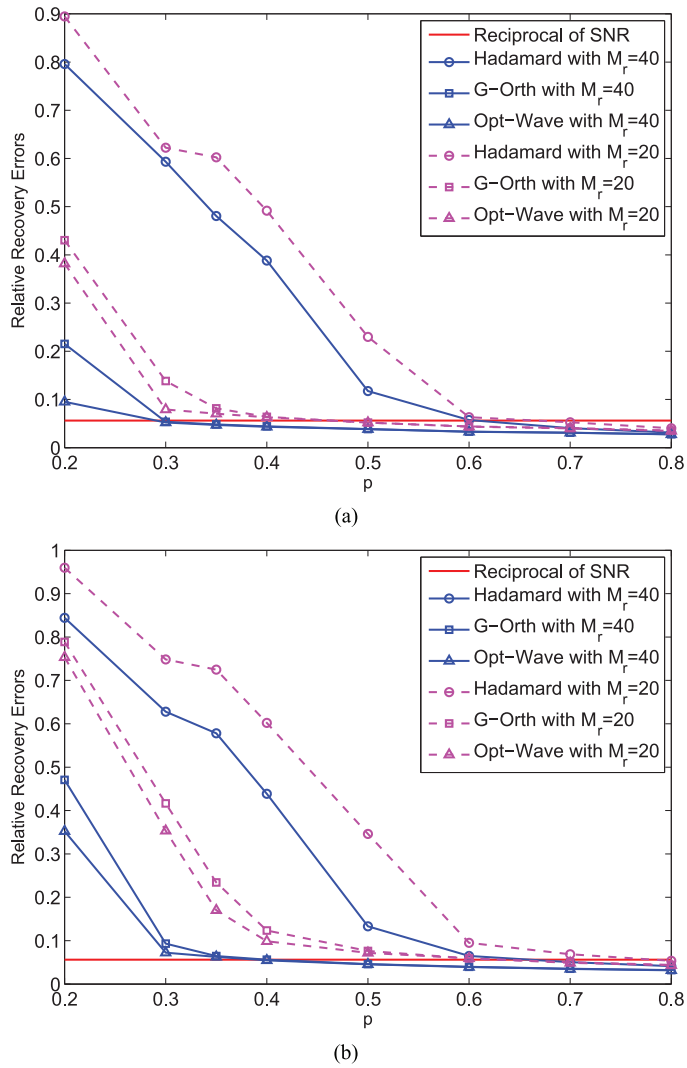

Fig. 4. MC error vs. p for targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$. (a) all targets fall in the same range bin; (b) the ranges are $[2020, 3820, 3820, 6220]$ m.

column-wise G-Orth waveforms. One can see that the former result in lower MC recovery error for smaller p's, while their advantage diminishes for higher p's. Based on our experience with simulations, the range of p over which the optimized waveforms have an advantage over the G-Orth waveforms shrinks as M_t increases. This observation suggests that for large number of transmit antennas, the G-Orth waveforms behave like optimal in the sense that they achieve a comparable MC performance as the optimized waveforms. Due to the substantially lower computational cost involved, in such cases G-Orth waveforms would be preferable in a practical scenario.

Although the waveform design requires angle space discretization, as suggested by simulations not included here due to space limitations, the performance is not sensitive when the targets fall between grid points when $p \geq 0.4$.

2) Targets in Different Range Bins: Let us continue on the configuration in Fig. 4(a), but set the $K = 4$ target ranges to $[2020, 3820, 3820, 6220]$ m, corresponding to maximal normalized delay $N_1 = 35$. As shown in Fig. 4(b), the recovery error under the optimized waveform matrix has a similar performance trend as the scenario considered in Fig. 4(a), i.e., the optimized waveform matrix results in significant performance

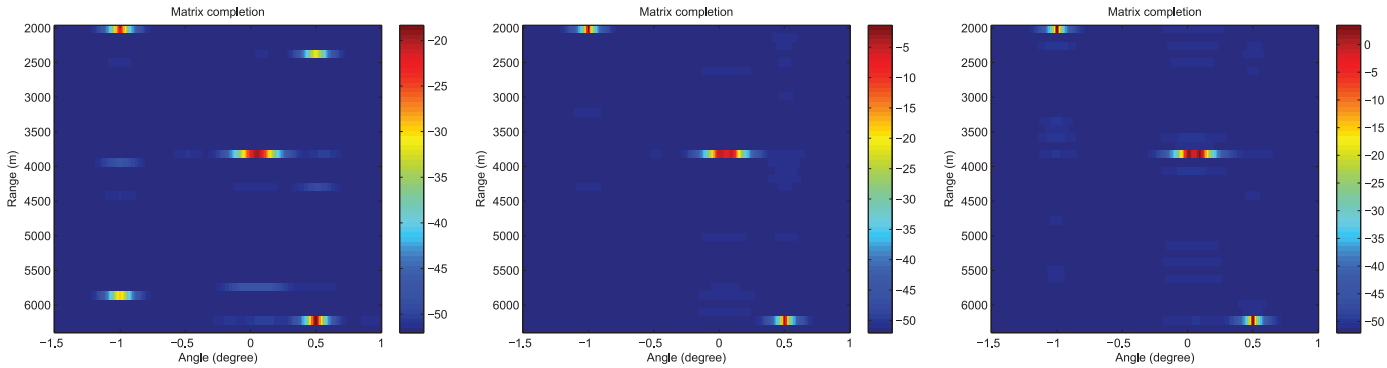

Fig. 5. Angle-Range images using MIMO-MC radars for $K = 4$ targets located at $[-1°, 0°, 0.1°, 0.5°]$ with ranges corresponding to $[2020, 3820, 3820, 6220]$ m: (a) Hadamard waveforms; (b) G-Orth waveforms; (c) Optimized waveforms. The other parameters are $M_t = 20, N = 64, M_r = 20, Q = 40, p = 0.35$.

improvement as compared to the Hadamard matrix. As compared to Fig. 4(a), a slightly larger portion of samples is required for MC under optimized waveforms to achieve a recovery error less than the inverse of SNR, i.e., the ratio of noise power over the targets' power. This is because the coherence in the scenario of targets falling in different range bins is slightly higher, as indicated in (61).

In both Fig. 4(a) and (b), one can see that increasing M_r could improve the matrix completion performance for all tested waveforms..

D. Target Estimation Performance via MC

Continuing on the scenario in Fig. 4(b), we look at the target estimation performance after the data matrix is recovered via MC. Since the targets fall in different range bins, range compression (pulse compression) is first applied to the recovered data matrix. Then, the DOA and speed estimation follow using the subspace methods, such as MUSIC [40]. The details of DOA estimation using MUSIC method are addressed in [8, equation (17)].

Fig. 5 shows the Angle-Range image of the target scene of Fig. 4(b), using the MUSIC method following MC recovery. Total $Q = 40$ pulses are transmitted and $p = 0.35$ portion samples are collected from $M_r = 20$ antennas. One can see that the performance of optimized and G-Orth waveforms is better than that of the Hadamard waveform. False alarms at angles $-1°$ and $0.5°$ in the wrong range bins are triggered under the Hadamard waveform; this is due to the waveform poor correlation properties. Furthermore, the two middle targets at range 3820 m, are unresolvable under the Hadamard waveform. In this case, the performance under the optimized waveform is slightly better than that under G-Orth waveform in terms of magnitude, which would result in better DOA resolution. However, as p increases, the two waveforms perform comparably (see Fig. 4(b)).

Next, we access the capability of the MC based method to resolve two closely located targets in the same range bin. We change the scenario in Fig. 5 by setting $\theta_3 = \theta_2 + \Delta\theta$ and keep the rest parameters unchanged; here θ_k denotes the angle of the k-th target. Two targets are considered to be resolved if $\left|\hat{\theta}_k - \theta_k\right| \leq \Delta\theta/2,\ k = 2, 3$, where $\hat{\theta}_k$ denotes the estimation

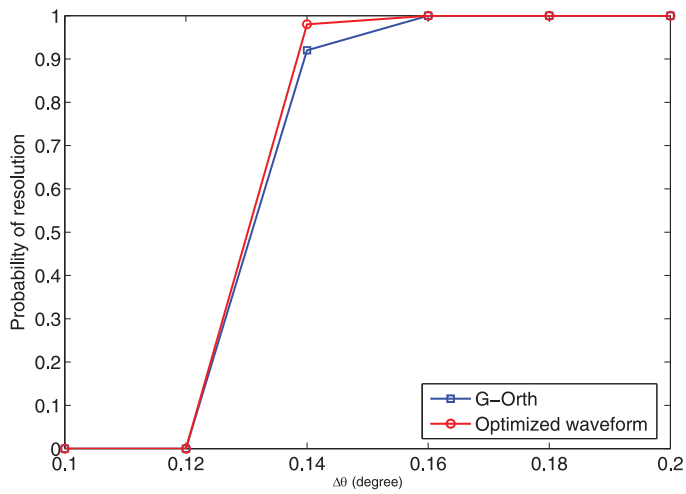

Fig. 6. Probability of resolution comparison for optimized and G-Orth waveforms for $p = 0.35$.

of the k-th target [40]. Fig. 6 shows the probability of resolution comparison between the optimized and G-Orth waveforms for $p = 0.35$. The results are obtained based on 50 independent runs, and in each run, the noise is randomly generated. The probability of resolution is calculated by counting the number of successful resolvable events over the total number of runs. It can be seen that the optimized waveform results in a better probability of resolution as compared to the G-Orth waveform. High resolution could be achieved using MC under both optimized and G-Orth waveforms.

VII. Conclusions

We have presented an analysis of the coherence of the data matrix arising in MIMO-MC radar with ULA configurations and transmitting orthogonal waveforms. We have shown that, the data matrix attains its lowest possible coherence if the waveform snapshots across the transmit array have flat power spectra for all time instances. The waveform design problem has been approached as an optimization problem on the complex Stiefel manifold and has been solved via the modified steepest descent algorithm and the modified Newton algorithm. The numerical results have shown that as the number of antennas increases, the optimized waveforms result in optimal data matrix coherence, i.e., 1, and thus, only a small portion of samples are needed

for the data matrix recovery. For a particular array, the optimal waveforms depend on the target space to be investigated; for different regions of the target space, the corresponding optimal waveforms can be constructed a priori. Since their construction involves high computational complexity, the optimal waveforms can be used as benchmark against easily constructed waveforms. For example, our simulations revealed that as the number of transmit antennas increases, simply transmitting G-Orth waveforms results in comparable matrix recovery performance as transmitting optimized waveforms. Thus, given the cost of computing the optimized waveforms, certain applications and under certain conditions may treat G-Orth waveforms as optimal. Although the optimal waveforms are designed based on the assumption that the targets fall in the same range bin, our analysis and simulations showed that they cause only small amount of performance degradation for relatively small delays, i.e., of the order of the symbol interval when used in a scenario in which the targets appear in different range bins.

Appendix
Derivative and Hessian of $f(\mathbf{S})$

In the following, we list the derivative and Hessian of the cost function $f(\mathbf{S})$.

$$\mathrm{D}_{\mathbf{S}} = 2\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)-\mathbf{N}\right]\odot\mathbf{Y}^T\right\}\mathbf{A}^H, \tag{63}$$

$$\mathrm{H}_{\mathbf{S}} = 4\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}^*\right)\mathrm{diag}\left(\mathrm{vec}\left(2\tilde{\mathbf{H}}^T-\mathbf{N}^T\right)\right)\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}, \tag{64}$$

$$\mathrm{C}_{\mathbf{S}} = 4\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}\right)\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\odot\mathbf{Y}^*\right)\right)\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}, \tag{65}$$

where $\mathbf{N}=(M_t/N)\mathbf{1}_N\mathbf{1}_L^T$, $\tilde{\mathbf{H}}=\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\in\mathbb{R}^{N\times L}$ and $\mathbf{Y}=\mathbf{A}^T\mathbf{S}^T\in\mathbb{C}^{L\times N}$. In the above, $\mathbf{P}_{N\times M_t}$ is a commutation matrix, such that

$$\mathrm{vec}\left(\mathbf{Z}^T\right)=\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right), \tag{66}$$

which can be expressed as [41]

$$\mathbf{P}_{N\times M_t}=\sum_{m=1}^{N}\sum_{n=1}^{M_t}\left(\mathbf{E}_{mn}\otimes\mathbf{E}_{mn}^T\right), \tag{67}$$

where $\mathbf{E}_{mn}$ is a matrix of dimension $N\times M_t$ with 1 at its mn-th position and zeros elsewhere. It holds that $\mathbf{P}_{N\times M_t}=\mathbf{P}_{M_t\times N}^T$. It is easy to verify that $\mathrm{H}_{\mathbf{S}}=\mathrm{H}_{\mathbf{S}}^H, \mathrm{C}_{\mathbf{S}}=\mathrm{C}_{\mathbf{S}}^T$.

To derive the above derivative and Hessian matrices, we give the following two Lemmas (proofs are in [29]):

Lemma 2: Let $\mathbf{A}\in\mathbb{C}^{M_t\times L}, \mathbf{Z}\in\mathbb{C}^{N\times M_t}, \mathbf{Y}\in\mathbb{C}^{L\times N}, \mathbf{H}\in\mathbb{R}^{N\times L}$ *be arbitrary matrices. It can be shown that*

$$\mathrm{tr}\left\{\mathbf{H}\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\mathbf{Y}\right]\right\}=\mathrm{tr}\left\{\mathbf{Z}^H\left[\left(\mathbf{H}\odot\mathbf{Y}^T\right)\mathbf{A}^H\right]\right\}.$$

Lemma 3: Let $\tilde{\mathbf{H}}\in\mathbb{C}^{N\times L}$ *and* $\mathbf{G},\mathbf{M}\in\mathbb{C}^{L\times N}$ *be general matrices with arbitrary elements. It holds that*

$$\mathrm{tr}\left(\left(\mathbf{G}\odot\mathbf{M}\right)\tilde{\mathbf{H}}\right)=\left[\mathrm{vec}\left(\mathbf{G}\right)\right]^T\mathrm{vec}\left(\mathbf{M}\odot\tilde{\mathbf{H}}^T\right).$$

Since $\mathbf{F}\odot\mathbf{F}^*$ and $\mathbf{N}$ are real-valued matrices, the objective function $f(\mathbf{S})$ defined in (32) can be written as

$$f\left(\mathbf{S}\right)=\mathrm{tr}\left\{\left(\mathbf{F}\odot\mathbf{F}^*-\mathbf{N}\right)\left(\mathbf{F}\odot\mathbf{F}^*-\mathbf{N}\right)^T\right\}. \tag{68}$$

In order to find the derivative and Hessian of $f(\mathbf{S})$, we do the following expansion:

$$\begin{aligned}f\left(\mathbf{S}+\delta\mathbf{Z}\right)=&\mathrm{tr}\Big\{\left[\left(\left(\mathbf{S}+\delta\mathbf{Z}\right)^*\mathbf{A}^*\right)\odot\left(\left(\mathbf{S}+\delta\mathbf{Z}\right)\mathbf{A}\right)-\mathbf{N}\right]\\&\times\left[\left(\left(\mathbf{S}+\delta\mathbf{Z}\right)^*\mathbf{A}^*\right)\odot\left(\left(\mathbf{S}+\delta\mathbf{Z}\right)\mathbf{A}\right)-\mathbf{N}\right]^T\Big\}\\=&f\left(\mathbf{S}\right)+\delta\mathrm{tr}\left\{\mathbf{T}\right\}+\delta^2\mathrm{tr}\left\{\mathbf{T}'\right\}+\mathcal{O}\left(\delta^3\right).\end{aligned} \tag{69}$$

Here,

$$\begin{aligned}\mathbf{T}=&\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]-\mathbf{N}\right\}\\&\times\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T+\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^H\right\}\\&+\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]+\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^*\right\}\\&\times\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]^T-\mathbf{N}^T\right\},\end{aligned} \tag{70}$$

$$\begin{aligned}\mathbf{T}'=&\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\\&+\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right\}^T\\&+\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]^T\\&+\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]^T\right\}^T\\&+\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\\&+\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right\}^H\\&-\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\mathbf{N}^T\\&-\left\{\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\mathbf{N}^T\right\}^T.\end{aligned} \tag{71}$$

Thus, it holds that

$$\begin{aligned}\mathrm{tr}\left\{\mathbf{T}\right\}=&\mathrm{tr}\Big(2\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]-\mathbf{N}\right\}\\&\times\left\{\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T+\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^H\right\}\Big)\\=&\Re\left\{\mathrm{tr}\left(\mathbf{H}\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{S}^T\right)\right]\right)\right\},\end{aligned} \tag{72}$$

Let $\mathbf{Y}=\mathbf{A}^T\mathbf{S}^T$. Following Lemma 2, it holds that

$$\mathrm{tr}\left(\mathbf{T}\right)=\Re\left\{\mathrm{tr}\left(\mathbf{Z}^H\left[\left(\mathbf{H}\odot\mathbf{Y}^T\right)\mathbf{A}^H\right]\right)\right\}. \tag{73}$$

In addition, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\mathbf{T}'\right)\\&=2\mathrm{tr}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right)\\&\quad+2\mathrm{tr}\left(\left[\left(\mathbf{S}\mathbf{A}\right)\odot\left(\mathbf{Z}^*\mathbf{A}^*\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right)\\&\quad+2\Re\left\{\mathrm{tr}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right)\right\}\\&\quad-2\mathrm{tr}\left(\mathbf{N}\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\right).\end{aligned} \tag{74}$$

Now, we focus on the first term on the right side of (74). Following Lemma 3, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{S}\mathbf{A}\right)\right]\left[\left(\mathbf{Z}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right)\\&=\mathrm{tr}\left(\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\tilde{\mathbf{H}}\right)\\&=\left[\mathrm{vec}\left(\mathbf{A}^H\mathbf{Z}^H\right)\right]^T\mathrm{vec}\left(\left(\mathbf{A}^T\mathbf{Z}^T\right)\odot\tilde{\mathbf{H}}^T\right),\end{aligned} \tag{75}$$

where $\tilde{\mathbf{H}} = (\mathbf{S}^*\mathbf{A}^*) \odot (\mathbf{S}\mathbf{A}) \in \mathbb{R}^{N\times L}$. Further, via (66), it holds that

$$\begin{aligned}\left[\mathrm{vec}\left(\mathbf{A}^H\mathbf{Z}^H\right)\right]^T &= \left[\left(\mathbf{I}_N \otimes \mathbf{A}^H\right)\mathrm{vec}\left(\mathbf{Z}^H\right)\right]^T \\ &= \left\{\left(\mathbf{I}_N \otimes \mathbf{A}^H\right)\left[\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right)\right]^*\right\}^T \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^H \mathbf{P}_{M_t\times N}\left(\mathbf{I}_N \otimes \mathbf{A}^*\right), \end{aligned} \quad (76)$$

as well as

$$\begin{aligned}&\mathrm{vec}\left(\left(\mathbf{A}^T\mathbf{Z}^T\right)\odot\tilde{\mathbf{H}}^T\right) \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\tilde{\mathbf{H}}^T\right)\right)\mathrm{vec}\left(\mathbf{A}^T\mathbf{Z}^T\right) \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\tilde{\mathbf{H}}^T\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathrm{vec}\left(\mathbf{Z}^T\right) \\ &= \mathrm{diag}\left(\mathbf{P}_{N\times L}\mathrm{vec}\left(\tilde{\mathbf{H}}\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (77)$$

Consequently, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\tilde{\mathbf{H}}\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^H\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}^*\right)\mathrm{diag}\left(\mathbf{P}_{N\times L}\mathrm{vec}\left(\tilde{\mathbf{H}}\right)\right) \\ &\quad\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (78)$$

Let us focus on the second term on the right side of (74). It holds that

$$\begin{aligned}&\mathrm{tr}\left(\left[\left(\mathbf{S}\mathbf{A}\right)\odot\left(\mathbf{Z}^*\mathbf{A}^*\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right) \\ &= \mathrm{tr}\left(\left[\left(\mathbf{A}^H\mathbf{S}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]^H\left[\left(\mathbf{A}^H\mathbf{S}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\right) \\ &= \mathrm{tr}\left(\left[\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]^H\left[\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right)\right]^H\mathrm{vec}\left(\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right), \end{aligned} \quad (79)$$

as well as

$$\begin{aligned}&\mathrm{vec}\left(\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right) \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\mathrm{vec}\left(\mathbf{A}^T\mathbf{Z}^T\right) \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathrm{vec}\left(\mathbf{Z}^T\right) \\ &= \mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (80)$$

Consequently, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\left[\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]^H\left[\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^H\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}^*\right)\left[\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\right]^H \\ &\quad\times\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^H\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}^*\right)\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}\odot\mathbf{Y}^*\right)\right) \\ &\quad\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (81)$$

Next, let us focus on the third term on the right side of (74). With (80), it holds that

$$\begin{aligned}&\mathrm{tr}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right) \\ &= \left[\mathrm{vec}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right)\right]^T\mathrm{vec}\left(\left[\left(\mathbf{S}^*\mathbf{A}^*\right)\odot\left(\mathbf{Z}\mathbf{A}\right)\right]^T\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right)\right]^T\mathrm{vec}\left(\mathbf{Y}^*\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^T\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}\right)\left[\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\right]^T \\ &\quad\times\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\right)\right)\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^T\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}\right)\mathrm{diag}\left(\mathrm{vec}\left(\mathbf{Y}^*\odot\mathbf{Y}^*\right)\right) \\ &\quad\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (82)$$

Finally, let us focus on the forth term on the right side of (74). Via (76) (77) and Lemma 3, it holds that

$$\begin{aligned}&\mathrm{tr}\left(\mathbf{N}\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\right) \\ &= \mathrm{tr}\left(\left[\left(\mathbf{A}^H\mathbf{Z}^H\right)\odot\left(\mathbf{A}^T\mathbf{Z}^T\right)\right]\mathbf{N}\right) \\ &= \left[\mathrm{vec}\left(\mathbf{A}^H\mathbf{Z}^H\right)\right]^T\mathrm{vec}\left(\left(\mathbf{A}^T\mathbf{Z}^T\right)\odot\mathbf{N}^T\right) \\ &= \left[\mathrm{vec}\left(\mathbf{Z}\right)\right]^H\mathbf{P}_{M_t\times N}\left(\mathbf{I}_N\otimes\mathbf{A}^*\right)\mathrm{diag}\left(\mathbf{P}_{N\times L}\mathrm{vec}\left(\mathbf{N}\right)\right) \\ &\quad\times\left(\mathbf{I}_N\otimes\mathbf{A}^T\right)\mathbf{P}_{N\times M_t}\mathrm{vec}\left(\mathbf{Z}\right). \end{aligned} \quad (83)$$

Therefore, from the expansion of $f(\mathbf{S}+\delta\mathbf{Z})$ and the matrix form of the second-order Taylor series (35), we finally obtain (63) (64) (65).

REFERENCES

[1] S. Sun and A. P. Petropulu, "On the applicability of matrix completion on MIMO radars," in *48th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 2–5, 2014, pp. 306–310.

[2] S. Sun and A. P. Petropulu, "On waveform design for MIMO radar with matrix completion," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, Atlanta, GA, USA, Dec. 3–5, 2014, pp. 473–477.

[3] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 106–114, Sep. 2007.

[4] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of MIMO radar," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 968–971, Dec. 2007.

[5] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, Aug. 2007.

[6] N. Levanon and E. Mozeson, *Radar Signals*. Hoboken, NJ, USA: Wiley, 2004.

[7] S. Sun, A. P. Petropulu, and W. U. Bajwa, "Target estimation in colocated MIMO radar via matrix completion," in *Proc. IEEE 38th Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Vancouver, BC, Canada, May 2013, pp. 4144–4148.

[8] S. Sun, W. U. Bajwa, and A. P. Petropulu, "MIMO-MC radar: A MIMO radar approach based on matrix completion," *IEEE Trans. Aerosp. Electron. Syst.*, 2015, to be published.

[9] C. Y. Chen and P. P. Vaidyanathan, "Compressed sensing in MIMO radar," in *Proc. 42th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Oct. 26–29, 2008, pp. 41–44.

[10] T. Strohmer and B. Friedlander, "Compressed sensing for MIMO radar – algorithms and performance," in *Proc. 43th Annu. Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, USA, Nov. 1–4, 2009, pp. 464–468.

[11] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 146–163, Feb. 2010.

[12] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundat. Comput. Math.*, vol. 9, no. 6, pp. 717–772, 2009.

[13] E. J. Candès and T. Tao, "The power of convex relaxation: Near-optimal matrix completion," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, May 2010.

[14] E. J. Candès and Y. Plan, "Matrix completion with noise," *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, Jun. 2010.

[15] D. S. Kalogerias and A. P. Petropulu, "Matrix completion in colocated MIMO radar: Recoverability, bounds & theoretical guarantees," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 309–321, Feb. 2014.

[16] D. S. Kalogerias, S. Sun, and A. P. Petropulu, "Sparse sensing in colocated MIMO radar: A matrix completion approach," in *Proc. IEEE 13th Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Athens, Greece, Dec. 12–15, 2013, pp. 000496–000502.

[17] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, Jan. 2007.

[18] Y. Yang and R. S. Blum, "Minimax robust MIMO radar waveform design," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 147–155, Jun. 2007.
[19] H. He, P. Stoica, and J. Li, "Designing unimodular sequence sets with good correlations-including an application to MIMO radar," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4391–4405, Nov. 2009.
[20] K. W. Forsythe and D. W. Bliss, "MIMO radar waveform constraints for GMTI," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 21–32, Feb. 2010.
[21] H. Deng, "Polyphase code design for orthogonal netted radar systems," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 3126–3135, Oct. 2004.
[22] C. Y. Chen and P. P. Vaidyanathan, "MIMO radar ambiguity properties and optimization using frequency-hopping waveforms," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 5926–5936, Feb. 2008.
[23] G. S. Antonio, D. R. Fuhrmann, and F. C. Robey, "MIMO radar ambiguity functions," *IEEE Trans. Signal Process.*, vol. 1, no. 1, pp. 167–177, Jan. 2007.
[24] S. Gogineni and A. Nehorai, "Frequency-hopping code design for MIMO radar estimation using sparse modeling," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3022–3035, Jun. 2012.
[25] Y. Yu, S. Sun, R. N. Madan, and A. P. Petropulu, "Power allocation and waveform design for the compressive sensing based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 898–909, Apr. 2014.
[26] A. Edelamn, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
[27] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, Mar. 2002.
[28] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multi-dimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, Nov. 2010.
[29] S. Sun and A. P. Petropulu, "Waveform design for MIMO radar with matrix completion," Tech. Rep. Rutgers Univ., New Brunswick, NJ, USA, 2015 [Online]. Available: http://eceweb1.rutgers.edu/~cspl/publications/Waveform-TR.pdf
[30] P. A. Absil, R. Mathony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Univ. Press, 2008.
[31] H. Zhang and W. W. Hager, "A nonmonotone line search technique and its application to unconstrained optimization," *SIAM J. Optim.*, vol. 14, no. 4, pp. 1043–1056, 2004.
[32] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*. New York, NY, USA: Springer, 2007.
[33] M. Joho and K. Rahbar, "Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation," in *Proc. IEEE Sens. Array Multichannel Signal Process. Workshop (SAM)*, Rosslyn, VA, USA, Jun. 2002.
[34] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Nashua, NH, USA: Athena Scientific, 1999.
[35] M. Joho, "Newton method for joint approximate diagonalization of positive definite Hermitian matrices," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1205–1218, 2008.
[36] Y. Yu, A. P. Petropulu, and H. V. Poor, "CSSF MIMO radar: Compressive-sensing and step-frequency based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1490–1504, Apr. 2012.
[37] H. Wolkowicz and G. P. H. Styan, "Bounds for eigenvalues using traces," *Linear Algebra and Its Applicat.*, vol. 29, no. 2, pp. 471–506, 1980.
[38] L. Xu and Q. Liang, "Zero correlation zone sequence pair sets for MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 3, pp. 2100–2113, Jul. 2012.
[39] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, no. 2, pp. 1956–1982, 2010.
[40] H. L. Van Trees, *Optimum Arrary Processing*. New York, NY, USA: Wiley, 2002.
[41] J. R. Magnus and H. Neudecker, "Symmetry, 0–1 matrices and Jacobians: A review," *Econ. Theory*, vol. 2, no. 2, pp. 157–190, 1986.

Shunqiao Sun (S'11) received the B.E. degree from Southern Yangtze University, Wuxi, China, in 2004, and the M.S. degree from Fudan University, Shanghai, China, in 2011, both in electrical engineering. He is currently working towards the Ph.D. degree at the Department of Electrical and Computer Engineering, Rutgers, The State University of New Jersey. He has more than five years industrial experience, including a summer intern at Mitsubishi Electric Research Labs (MERL), Cambridge, MA.

His research interests are in statistical signal processing, MIMO radar with sparse sensing and wireless communications.

Athina P. Petropulu (M'87–SM'01–F'08) received her undergraduate degree from the National Technical University of Athens, Greece, and the M.Sc. and Ph.D. degrees from Northeastern University, Boston MA, all in electrical and computer engineering. Between 1992–2010 she was faculty at the Department of Electrical and Computer Engineering at Drexel University. Since 2010, she is Professor and Chair of the Electrical and Computer Engineering Department at Rutgers, The State University of New Jersey. She has held visiting appointments at SUPELEC in France and at Princeton University. Dr. Petropulu's research interests span the area of statistical signal processing with applications in wireless communications, networking and radar systems. She is recipient of the 1995 Presidential Faculty Fellow Award in Electrical Engineering given by NSF and the White House. She is the co-author (with C.L. Nikias) of the textbook entitled, "Higher-Order Spectra Analysis: A Nonlinear Signal Processing Framework," (Englewood Cliffs, NJ: Prentice-Hall, Inc., 1993).

Dr. Petropulu is Fellow of IEEE. She has served as Editor-in-Chief of the IEEE TRANSACTIONS ON SIGNAL PROCESSING (2009–2011), IEEE Signal Processing Society Vice President-Conferences (2006–2008), and member-at-large of the IEEE Signal Processing Board of Governors. She was the General Chair of the 2005 International Conference on Acoustics Speech and Signal Processing (ICASSP-05), Philadelphia PA. She is co-recipient of the 2005 IEEE Signal Processing Magazine Best Paper Award, and the 2012 Meritorious Service Award from the IEEE Signal Processing Society.