IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 20, 2020, accepted May 3, 2020, date of publication May 6, 2020, date of current version May 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2992810

Waveform Design for MIMO Radar Detection in a Heterogeneous Clutter Environment

BO ZHANG, FENGZHOU DAI, AND NAN SU
National Laboratory of Radar Signal Processing, Xidian University, Xi'an 710071, China

Corresponding author: Fengzhou Dai (fzdai@xidian.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61671359, in part by the Program for the National Science Fund for Distinguished Young Scholars under Grant 61525105, in part by the Shaanxi Innovation Team Project, and in part by the 111 Project under Grant B18039.

ABSTRACT In this study, the problem of transmitted waveform design for a multiple-input-multiple-output (MIMO) radar in a heterogeneous clutter environment is considered. Under a Bayesian framework, the waveform design problem is used to maximize the generalized likelihood ratio for a given false alarm probability. In addition, the proposed problem is nonconvex under the constant modulus constraint. Hence, the alternating direction method of multipliers-based and the Broyden–Fletcher–Goldfarb–Shanno method are adopted to solve this nonconvex optimization problem. Finally, the clutter suppression performance of the proposed algorithms is evaluated in real-world data by numerical simulations. The simulation results show that the proposed methods have better performance on clutter suppression and target detection.

INDEX TERMS MIMO radar, constant modulus constraint, waveform design, heterogeneous clutter, Bayesian framework, nonconvex optimization.

I. INTRODUCTION

In comparison to a phase-array radar, a multiple-input-multiple-output (MIMO) radar manifests significantly better performances in target detection and parameter identification [1]–[4]. Depending on different antenna configurations, MIMO radars can be divided into two classes—distributed MIMO radars that possess widely separated antennas [2] and colocated MIMO radars that possess closely spaced antennas [4].

Studies on the waveform design for colocated MIMO radars can be divided into two main classes: In the first category, MIMO radar waveforms are designed to control the transmitted power distribution in the spatial domain by setting the desired beam pattern. In order to address this problem, two main strategies are usually adopted. In the first strategy, the transmitted beam pattern is synthesized through the optimization of a waveform covariance matrix [4]–[8], and then transmitted waveforms are acquired from the obtained covariance matrix under practical constraints [5], [7], [9]. Lipor *et al.* [8] utilized discrete-Fourier-transform (DFT) coefficients and Toeplitz matrices to form a covariance matrix, which matched the desired beam pattern at low complexity. Finally, with the obtained covariance matrix, practical constraints, such as constant modulus (CM) [9] and the peak to average power ratio [7], were taken into consideration to synthesize these transmitted waveforms. In the second strategy, transmitted waveforms are directly designed to realize the desired beam pattern [10]–[14]. For example, Cheng *et al.* [13] tried to minimize the mean squared error between the synthesized and desired beam patterns based on the alternating direction method of multipliers (ADMM) algorithm. Fan *et al.* [14] also applied the ADMM algorithm; however, paid more attention to the restrictions of the sidelobe and the main lobe of the transmitted beam pattern.

In the second category, transmitted waveforms and receiver filters are designed jointly. MIMO waveforms are generally designed to maximize the output signal to interference plus noise ratio (SINR), which enhances the target detection performance of a MIMO radar and also suppresses signal-dependent interferences (clutter) [15]—[17]. In the literature [15], two sequential optimization procedures were adopted to design the desired waveform by maximizing the SINR value under a CM constraint and a similarity constraint. Sadjad *et al.* [16], with the given prior information of the target and interference angle locations, presented an iterative approach to jointly optimize transmitted waveforms and receiver filters. Tang and Tang [17] applied a cyclic algorithm

The associate editor coordinating the review of this manuscript and approving it for publication was Filbert Juwono.

to achieve an improved SINR under practical constraints. Some researchers [18]–[20] also presented the principle of MIMO waveform design based on the information theory. In the literature [18], [19], the minimum mean-square error (MMSE) was used to optimize MIMO waveforms in colored noise and clutter. In the study [20], instead of solving the intractable problem of maximizing the detection probability for a given false alarm, the relative entropy was employed to improve the detection performance of a MIMO radar.

Instead of maximizing the SINR value, in this study, the waveform design problem is performed by maximizing the likelihood ratio test or generalized likelihood ratio test (GLRT) for a given false alarm probability. In the literature [21]–[25], the GLRT expressions of different environments with prior information (clutter map, statistical information of clutters) were discussed. In some earlier studies [26]–[30], a Bayesian framework was adopted to develop the GLRT expression without secondary data. Therefore, we also consider the GLRT with the Bayesian framework.

In this study, considering the CM constraint, the waveform design problem is nonconvex, and two different approaches are proposed to accomplish the maximization of GLRT. In the first approach, the nonconvex problem is solved by the ADMM method [32]–[34]. In the second approach, the nonconvex problem is first relaxed as an unconstrained optimization problem and then solved by the Broyden–Fletcher–Goldfarb–Shanno (BFGS) method [38]. Finally, the detection probabilities of our designed waveforms the existing orthogonal linear frequency modulation (LFM) waveform are compared through numerical results. In addition, the CM waveform designed by [17] algorithm 2 is also provided to compare.

The remaining paper has been organized as follows: The data model for the proposed problem is formulated in Section II. The first and second algorithms are introduced in Sections III and IV, respectively. Numerical results are presented in Section V, and finally, conclusions are drawn in Section VI.

Notations: The transpose and conjugate transpose operators are denoted by $(\cdot)^T$ and $(\cdot)^H$, respectively, whereas $\odot$ and $\otimes$ represent the Hadamard product and the Kronecker product, respectively. The letter j represents the imaginary unit ($j = \sqrt{-1}$). $\|\cdot\|_2$ denotes the Euclidean norm of a vector, and $|\cdot|$ signifies the absolute value. The lower case letter '$\mathbf{a}$' and the upper-case letter '$\mathbf{A}$' denote vectors and matrices, respectively. vec $(\mathbf{A})$ represents the column vector obtained by stacking its columns, and $\mathbf{I}_N$ stands for the $N \times N$ dimensional identity matrix. $\Re\{\cdot\}$ and Im $\{\cdot\}$ respectively imply the real and imaginary parts of a complex value or a matrix. *etr* $(\cdot)$ stands for the exponential of the trace of the matrix argument, and det $(\cdot)$ denotes the determinant of a matrix.

II. DATA MODELING AND PROBLEM FORMULATION

The system under consideration is a colocated MIMO radar placed on an airborne platform. The radar is equipped with N_T transmitting antennas and N_R receiving antennas. Now, assume the radar transmits a burst of M pulses at a constant pulse repetition frequency (PRF) f_r and $s_n(l), l = 1, 2, \cdots, L; n = 1, 2, \cdots, N_T$ denote the discrete time radar waveform radiated by the n-th antenna (L is the number of time samples of each radar waveform). Furthermore, $\mathbf{s}_l = [s_1(l), \ldots, s_{N_T}(l)]^T$ indicates the N_T dimensional vector that collects transmitted waveforms at l th time; thus $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_L]^T$ can be defined as the transmitted waveform matrix.

Let $\mathbf{x}$ be the received signal in the cell under test (CUT). The problem of detecting targets in the presence of a heterogeneous clutter can be formulated through the following binary hypothesis testing:

$$\begin{cases} H_0: & \mathbf{x} = \mathbf{x}_c \\ H_1: & \mathbf{x} = \mathbf{x}_t + \mathbf{x}_c, \end{cases} \tag{1}$$

where $\mathbf{x}_t$ and $\mathbf{x}_c$ are the vectors that contain target signals and the clutter, respectively.

A. TARGETS MODEL

We assume the propagation is nondispersive, and the transmitted probing signals are narrowband. Considering there are I targets in the space, the i-th target is moving with a speed v_{t_i} relative to the radar platform, then received target signals at the direction of arrival (DOA) θ_{t_i} of the m-th pulse can be expressed as [6],

$$\mathbf{X}_{t_i,m} = \alpha_i e^{j2\pi(m-1)f_{d_i}T_r} \mathbf{b}(\theta_{t_i}) \mathbf{a}^T(\theta_{t_i}) \mathbf{S}^T, \tag{2}$$

where α_i is the target amplitude, $\mathbf{b}(\theta_{t_i})$ is the received array steering vector at θ_{t_i}, $\mathbf{a}(\theta_{t_i})$ is the transmit array steering vector, $f_{d_i} = 2v_{t_i}/\lambda$ is the target Doppler frequency (λ denotes the wavelength), and $T_r = 1/f_r$ is the pulse repetition interval (PRI).

Let $\mathbf{x}_{t_i} = \left[\text{vec}^T(\mathbf{X}_{t,1}), \ldots, \text{vec}^T(\mathbf{X}_{t,M})\right]^T$ be the i-th target response, and

$$\text{vec}(\mathbf{X}_{t_i,m}) = \alpha_i e^{j2\pi(m-1)f_{d_i}T_r} (\mathbf{S} \otimes \mathbf{I}_{N_R}) (\mathbf{a}(\theta_{t_i}) \otimes \mathbf{b}(\theta_{t_i})), \tag{3}$$

then we have

$$\mathbf{x}_{t_i} = \alpha_i \tilde{\mathbf{S}} \mathbf{h}_{t_i}, \tag{4}$$

where $\tilde{\mathbf{S}} = \mathbf{I}_M \otimes \mathbf{S} \otimes \mathbf{I}_{N_R}$, $\mathbf{h}_{t_i} = \mathbf{u}(f_{d_i}) \otimes \mathbf{a}(\theta_{t_i}) \otimes \mathbf{b}(\theta_{t_i})$ and $\mathbf{u}(f_{d_i}) = \left[1, \ldots, e^{j2\pi(M-1)f_{d_i}T_r}\right]^T$ is the temporal steering vector at f_{d_i}.

Let $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \cdots, \alpha_I]^T$, $\mathbf{h}_t = \left[\mathbf{h}_{t_1}, \mathbf{h}_{t_2}, \cdots, \mathbf{h}_{t_I}\right]$ and $\mathbf{x}_t = \left[\mathbf{x}_{t_1}^T, \mathbf{x}_{t_2}^T, \cdots, \mathbf{x}_{t_I}^T\right]^T$, then the targets responses can be given as

$$\mathbf{x}_t = \tilde{\mathbf{S}} \mathbf{h}_t \boldsymbol{\alpha}. \tag{5}$$

B. CLUTTER MODEL

Clutter can be described as unwanted echoes from scatters; hence, it can be regarded as a signal-dependent interference. In the present study, the measured clutter is modelled as the

superposition of signals associated with a large number of patches distributed in the azimuthal direction in the same range cell of the target. The measured signal–dependent clutter associated with the m-th pulse and the κ-th patch in the azimuth can be modeled as

$$\mathbf{X}_{c,m,\kappa} = \delta_{c,m,\kappa} e^{j2\pi(m-1)f_{c,\kappa}T_r} \mathbf{b}\left(\theta_{c,\kappa}\right) \mathbf{a}^T\left(\theta_{c,\kappa}\right) \mathbf{S}^T, \quad (6)$$

where $\delta_{c,m,\kappa}$ is the amplitude associated with the m-th pulse and the κ-th patch, and the parameters $f_{c,\kappa}$ and $\theta_{c,\kappa}$, respectively, denote the Doppler frequency and the DOA of the κ-th patch in the range cell.

Let $\mathbf{x}_{c,\kappa} = \left[\mathrm{vec}^T\left(\mathbf{X}_{c,1,\kappa}\right), \ldots, \mathrm{vec}^T\left(\mathbf{X}_{c,M,\kappa}\right)\right]^T$; hence, the measured clutter at the κ-th patch in the azimuth can be expressed as

$$\begin{aligned}\mathbf{x}_{c,\kappa} &= \left(\mathbf{I}_M \otimes \mathbf{S} \otimes \mathbf{I}_{N_R}\right)\left(\tilde{\mathbf{u}}\left(f_{c,\kappa}\right) \otimes \mathbf{a}\left(\theta_{c,\kappa}\right) \otimes \mathbf{b}\left(\theta_{c,\kappa}\right)\right) \\ &= \tilde{\mathbf{S}}\mathbf{h}_{c,\kappa}\end{aligned} \quad (7)$$

with $\mathbf{h}_{c,\kappa} = \tilde{\mathbf{u}}\left(f_{c,\kappa}\right) \otimes \mathbf{a}\left(\theta_{c,\kappa}\right) \otimes \mathbf{b}\left(\theta_{c,\kappa}\right)$ and $\tilde{\mathbf{u}}\left(f_{c,\kappa}\right) = diag\left\{\delta_{c,1,\kappa}, \ldots, \delta_{c,M,\kappa}\right\} \mathbf{u}\left(f_{c,\kappa}\right)$.

Now, by taking the clutter from N_c clutter patches, the measured clutter model can be formulated as

$$\mathbf{x}_c = \sum_{\kappa=1}^{N_c} \mathbf{x}_{c,\kappa}. \quad (8)$$

Furthermore, the clutter covariance matrix $\mathbf{R}_{SCM}$ can be expressed as

$$\begin{aligned}\mathbf{R}_{SCM} &= E\left(\mathbf{x}_c \mathbf{x}_c^H\right) = \sum_{\kappa=1}^{N_c} \tilde{\mathbf{S}} h_{c,\kappa} \mathbf{h}_{c,\kappa}^H \tilde{\mathbf{S}}^H \\ &= \tilde{\mathbf{S}}\left(\sum_{\kappa=1}^{N_c} \mathbf{h}_{c,\kappa}\mathbf{h}_{c,\kappa}^H\right)\tilde{\mathbf{S}}^H = \tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H,\end{aligned} \quad (9)$$

where $\Xi = \sum_{\kappa=1}^{N_c} \mathbf{h}_{c,\kappa}\mathbf{h}_{c,\kappa}^H$.

C. PROBLEM FORMULATION

Generally, a heterogeneous clutter can be characterized by the well-known complex compound-Gaussian model:

$$\mathbf{x}_c \sim \mathbb{C}N\left(\mathbf{0}, \mathbf{R}\right), \quad (10)$$

where $\mathbf{R}$ is the covariance matrix. In the field of the multivariate statistical analysis, the inverse Wishart distribution is used as the prior distribution of $\mathbf{R}$ [31], and in practical, the real radar data of SAR and polarimetric SAR (Pol-SAR) show that the covariance of the radar clutter fits the inverse Wishart distribution well [26]–[30]. Therefore, it is assumed that $\mathbf{R}$ is a random matrix drawn from a complex inverse Wishart distribution here; namely, $\mathbf{R} \sim \mathbb{C}W^{-1}\left(\mathbf{\Sigma}, \nu\right)$, where $\mathbb{C}W^{-1}\left(\mathbf{\Sigma}, \nu\right)$ signifies the complex inverse Wishart distribution with the concentration matrix $\mathbf{\Sigma}$ ($\mathbf{\Sigma} > 0$) and the freedom ν ($\nu > MN_RL$). Moreover, the probability density function (p.d.f) of the covariance matrix $\mathbf{R}$ can be expressed as

$$f\left(\mathbf{R}\right) = \frac{\left|\left(\nu - MN_RL\right)\mathbf{\Sigma}\right|^{\nu}}{\tilde{\Gamma}_{N_RL}\left(\nu\right)\left|\mathbf{R}\right|^{\nu+N_RL}} etr\left\{-\left(\nu - MN_RL\right)\mathbf{R}^{-1}\mathbf{\Sigma}\right\}, \quad (11)$$

$\tilde{\Gamma}_{N_RL}\left(\nu\right)$ can be defined as

$$\tilde{\Gamma}_{N_RL}\left(\nu\right) = \pi^{\frac{N_RL\left(N_RL-1\right)}{2}} \prod_{n=1}^{N_RL} \Gamma\left(\nu - n + 1\right), \quad (12)$$

where $\Gamma\left(\cdot\right)$ is the Eulerian Gamma function.

Usually, the target amplitude $\boldsymbol{\alpha}$ is unknown in detection problem. Thus, based on the Neyman − Pearson Theorem [35], the Bayesian GLRT expression for a given false alarm probability can be written as

$$L\left(\mathbf{x}\right) = \frac{\max\limits_{\boldsymbol{\alpha}} f\left(\mathbf{x}\left|H_1\right.\right)}{f\left(\mathbf{x}\left|H_0\right.\right)} = \frac{\max\limits_{\boldsymbol{\alpha}} \int f\left(\mathbf{x};\mathbf{R}\left|H_1\right.\right) f\left(\mathbf{R}\right) d\mathbf{R}}{\int f\left(\mathbf{x};\mathbf{R}\left|H_0\right.\right) f\left(\mathbf{R}\right) d\mathbf{R}}. \quad (13)$$

In addition,

$$\begin{aligned} f\left(\mathbf{x}\left|H_q\right.\right) &= \int f\left(\mathbf{x};\mathbf{R}\left|H_q\right.\right) f\left(\mathbf{R}\right) d\mathbf{R} \\ &\propto \det\left[\left(\mathbf{x} - q\mathbf{x}_t\right)\left(\mathbf{x} - q\mathbf{x}_t\right)^H + \left(\nu - N_RL\right)\mathbf{\Sigma}\right]^{-(\nu+1)}, \\ q &= 0, 1. \end{aligned} \quad (14)$$

Let $\hat{\boldsymbol{\alpha}}$ denote the maximum likelihood estimate (MLE) of $\boldsymbol{\alpha}$ under H_1, and substitute Eqs. (5) and (14) into Eq. (13), yields

$$\begin{aligned} &L\left(\mathbf{x}\right) \\ &= \frac{\det\left[\left(\mathbf{x} - \mathbf{x}_t\right)\left(\mathbf{x} - \mathbf{x}_t\right)^H + \left(\nu - MN_RL\right)\mathbf{\Sigma}\right]^{-(\nu+1)}}{\det\left[\mathbf{x}\mathbf{x}^H + \left(\nu - MN_RL\right)\mathbf{\Sigma}\right]^{-(\nu+1)}} \\ &\propto \frac{\det\left[\left(\mathbf{x} - \tilde{\mathbf{S}}h_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H + \left(\nu - MN_RL\right)\mathbf{\Sigma}\right]^{-1}}{\det\left[\mathbf{x}\mathbf{x}^H + \left(\nu - MN_RL\right)\mathbf{\Sigma}\right]^{-1}}, \end{aligned} \quad (15)$$

where $\propto$ implies "proportional to", and according to [30], the MLE expression of $\boldsymbol{\alpha}$ is given by

$$\hat{\boldsymbol{\alpha}} = \left[\left(\tilde{\mathbf{S}h}_t\right)^H \mathbf{\Sigma}^{-1}\tilde{\mathbf{S}}\mathbf{h}_t\right]^{-1}\left(\tilde{\mathbf{S}h}_t\right)^H \mathbf{\Sigma}^{-1}\mathbf{x}. \quad (16)$$

Furthermore, let us consider the concentration matrix $\mathbf{\Sigma}$ is unknown, then an alternative model of the concentration matrix can be modelled as

$$\mathbf{\Sigma} = \mathbf{R}_{SCM} + \sigma\mathbf{I}_{LN_T} = \tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H + \sigma\mathbf{I}, \quad (17)$$

where σ is the positive diagonal loading coefficient.

According to the Neyman − Pearson theorem, for a given false alarm probability, GLRT expresses the maximized detection probability. Therefore, by substituting (17) into Eq. (15), it can be noticed that the GLRT value is dependent on the transmitted waveform $\mathbf{S}$. Hence, the waveform design problem that maximizes GLRT can be formulated under the CM constraint as follows:

$$\begin{aligned} &\max_{\mathbf{S}} \frac{\det\left[\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\alpha}\right)^H + \tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]^{-1}}{\det\left[\mathbf{x}\mathbf{x}^H + \tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]^{-1}} \\ &s.t.\ \left|\mathbf{S}_{li}\right| = 1 \quad l = 1, \cdots, L;\ i = 1, \cdots, N_T \end{aligned} \quad (18)$$

where $\tilde{\Xi} = (v - N_R L)\,\Xi$, $\tilde{\mathbf{I}} = (v - N_R L)\,\mathbf{I}$, and $\mathbf{S}_{li}$ is the element of l-th column and i-th row of $\mathbf{S}$.

Now, by taking logarithm and negative operations to the cost function of Eq. (18), the equivalent form of Eq. (18) can be expressed as

$$\begin{aligned}&\min_{\mathbf{S}} \ln\left\{\det\left[\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&\quad-\ln\left\{\det\left[\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&s.t.\ |\mathbf{S}_{li}|=1\quad l=1,\cdots,L;\ i=1,\cdots,N_T.\end{aligned}\tag{19}$$

The optimization model in Eq.(19) involves the minimization of a nonconvex cost function and a nonconvex CM constraint, so this is an NP-hard problem. In general, iteration system is used to solve the NP-hard problem. However, the target amplitude vector is unknown and needs to be estimated first at each iteration. Therefore, the specific iteration system is shown in Fig.1, where $\hat{\boldsymbol{\alpha}}_k$ is the MLE of the target amplitude vector with k denoting the iteration time. Then two approaches in Sections III and IV, are proposed to optimize the waveforms respectively.

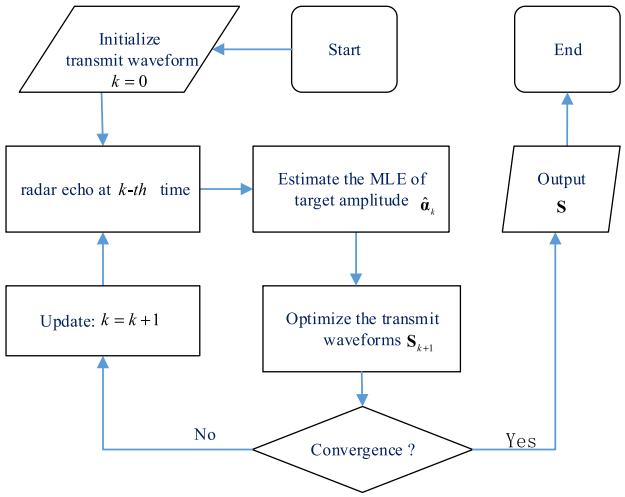


FIGURE 1. Flowchart of the Waveform design system.

III. ADMM-BASED WAVEFORM DESIGN

In the present section, the alternation direction method of multipliers (ADMM) algorithm (a distributed optimization approach) [32]–[34] is presented to solve Eq. (19).

A. ADMM-BASED ALGORITHM

Let $\mathbf{S}_R = \Re\{\mathbf{S}\}, \mathbf{S}_I = \mathrm{Im}\{\mathbf{S}\}$ and $\mathbf{s} = \mathrm{vec}(\mathbf{S})$, then $\mathbf{s}_R = \mathrm{vec}(\mathbf{S}_R)$, and $\mathbf{s}_I = \mathrm{vec}(\mathbf{S}_I)$; thus $\mathbf{s} = \mathbf{s}_R + j\mathbf{s}_I$. Moreover, two auxiliary vectors $\mathbf{v} = \left[v_1,\cdots,v_{LN_T}\right]^T$ and $\mathbf{w} = \left[w_1,\cdots,w_{LN_T}\right]^T$ are also introduced. v_p is the p -th element in $\mathbf{v}$ and w_p is the p-th element in $\mathbf{w}$; therefore,

$$\begin{aligned}v_p&=\left\{(\mathbf{S}_R)_{li}\right\}^2\\&=\left\{(\mathbf{s}_R)_p\right\}^2,\ \begin{matrix}p=(i-1)L+l;\\ l=1,\cdots,L;\ i=1,\cdots,N_T\end{matrix}\end{aligned}\tag{20a}$$

$$\begin{aligned}w_p&=\left\{(\mathbf{S}_I)_{li}\right\}^2\\&=\left\{(\mathbf{s}_I)_p\right\}^2,\ \begin{matrix}p=(i-1)L+l;\\ l=1,\cdots,L;\ i=1,\cdots,N_T.\end{matrix}\end{aligned}\tag{20b}$$

Hence, the CM constraint in Eq. (19) can be reformulated as

$$v_p+w_p=1,\ \begin{matrix}p=(i-1)L+l;\\ l=1,\cdots,L;\ i=1,\cdots,N_T.\end{matrix}\tag{21}$$

Subsequently, Eq. (19) can be transformed into the following equivalent problem:

$$\begin{aligned}&\min_{\mathbf{S}}\ \ln\left\{\det\left[\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&\quad-\ln\left\{\det\left[\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&s.t.\ \mathbf{v}+\mathbf{w}=\mathbf{1},\end{aligned}\tag{22}$$

where $\mathbf{1}$ is a LN_T dimensional vector in which all elements are equal to 1.

The augmented Lagrangian (formulated using scaled dual variables) [32] of Eq. (22) can be defined as

$$\begin{aligned}&L_\rho(\mathbf{s}_R,\mathbf{s}_I,\mathbf{u})\\&=\ln\left\{\det\left[\left(\mathbf{x}-\tilde{\mathbf{S}}h_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&\quad-\ln\left\{\det\left[\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&\quad+\frac{\rho}{2}\left(\|\mathbf{v}+\mathbf{w}-\mathbf{1}+\mathbf{u}\|_2^2-\|\mathbf{u}\|_2^2\right),\end{aligned}\tag{23}$$

where $\rho>0$ is the step size and $\mathbf{u}\in\mathbb{R}^{LN_T\times 1}$ is the scaled dual variable. Finally, ADMM is applied to Eq. (23) to determine $\{\mathbf{s}_R,\mathbf{s}_I,\mathbf{u}\}$:

$$\mathbf{s}_R^{k+1}:=\arg\min_{\mathbf{S}_R}L_\rho\left(\mathbf{S}_R,\mathbf{S}_I^k,\mathbf{u}^k\right);\tag{24a}$$

$$\begin{aligned}v_p^{k+1}&=\left\{\left(\mathbf{s}_R^{k+1}\right)_p\right\}^2,\quad p=(i-1)L+l;\\&\quad l=1,\cdots,L;\ i=1,\cdots,N_T;\end{aligned}\tag{24b}$$

$$\mathbf{s}_I^{k+1}:=\arg\min_{\mathbf{S}_I}L_\rho\left(\mathbf{S}_R^{k+1},\mathbf{S}_I,\mathbf{u}^k\right);\tag{24c}$$

$$\begin{aligned}w_p^{k+1}&=\left\{\left(\mathbf{s}_I^{k+1}\right)_p\right\}^2,\quad p=(i-1)L+l;\\&\quad l=1,\cdots,L;\ i=1,\cdots,N_T;\end{aligned}\tag{24d}$$

$$\mathbf{u}^{k+1}:=\mathbf{v}^{k+1}+\mathbf{w}^{k+1}-\mathbf{1}+\mathbf{u}^k;\tag{24e}$$

where k is the time mentioned before in Eq.(19).

Clearly, the constrained optimization problem in Eq. (19) can be divided into several subproblems (Eqs.24(a −(e)). The first subproblem of updating $\mathbf{s}_R$ in Eq. (24a) and the third subproblem of updating $\mathbf{s}_I$ in Eq. (24c) are both unconstrained. The final step in Eq. (24e) defines common update rules of scaled dual variables [32].

Let

$$\begin{aligned}&g_A(\mathbf{s}_R,\mathbf{s}_I)\\&=\ln\left\{\det\left[\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x}-\tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\\&\quad-\ln\left\{\det\left[\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H+\sigma\tilde{\mathbf{I}}\right]\right\}\end{aligned}$$

$g_B(\mathbf{s}_R, \mathbf{s}_I, \mathbf{u}) = \|\mathbf{v}+\mathbf{w}-\mathbf{1}+\mathbf{u}\|_2^2$ and both $g_A(\mathbf{s}_R, \mathbf{s}_I)$ and $g_B(\mathbf{s}_R, \mathbf{s}_I)$ are implicit functions of $\mathbf{s}_R$ and $\mathbf{s}_I$. The solutions of the subproblems in Eqs. (24a) and (24c) can be derived as follows:

Update of $\mathbf{s}_R$: If $\mathbf{b}_R^k = \mathbf{w}^k - \mathbf{1} + \mathbf{u}^k$, then $g_C\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right) = \left\|\mathbf{v}+\mathbf{b}_R^k\right\|_2^2$. The optimization problem in Eq. (24a) can be written as

$$\min_{\mathbf{S}_R} L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right) = g_A\left(\mathbf{s}_R, \mathbf{s}_I^k\right) + \frac{\rho}{2} g_B\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right). \tag{25}$$

Therefore, Eq. (25) is unconstrained. However, it is hard to obtain $\mathbf{s}_R^{k+1}$ by directly setting $\partial L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)/\partial \mathbf{s}_R = \mathbf{0}$. Fortunately, the Quasi-Newton method is one of the most effective methods to solve unconstrained nonlinear optimization problems. Therefore, the BFGS method [38], a type of quasi-Newton method, is adopted to solve Eq. (25). Furthermore, the key is transformed to calculate the gradient of $L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)$ at $\mathbf{s}_R$.

Based on the derivative of a real function with respect to a complex variable [36], the derivative of $g_A\left(\mathbf{s}_R, \mathbf{s}_I^t\right)$ with respect to $\mathbf{s}_R$ can be given as

$$\frac{\partial g_A\left(\mathbf{s}_R, \mathbf{s}_I^k\right)}{\partial \mathbf{s}_R} = \frac{\partial g_A\left(\mathbf{s}_R, \mathbf{s}_I^k\right)}{\partial \mathrm{vec}\left(\mathbf{S}_R\right)} = 2\Re\left\{ \left.\frac{\partial g_A}{\partial \mathbf{s}^*}\right|_{\substack{\mathbf{s}_I=\mathbf{s}_I^k,\\ \hat{\boldsymbol{\alpha}}=\hat{\boldsymbol{\alpha}}_k}} \right\}, \tag{26}$$

which can be computed from Proposition 1.

Proposition 1: The derivative of g_A with respect to vec $(\mathbf{S}^*)$ can be calculated as

$$\begin{aligned}\frac{\partial g_A}{\partial\left(\mathbf{s}^*\right)^T} &= \frac{\partial g_A}{\partial vec^T\left(\mathbf{S}^*\right)} = \frac{\partial g_A(\mathbf{S}, \mathbf{S}^*)}{\partial \mathrm{vec}^T\left(\tilde{\mathbf{S}}^*\right)} \frac{\partial vec\left(\tilde{\mathbf{S}}^*\right)}{\partial vec^T\left(\mathbf{S}^*\right)} \\ &= vec^T\left\{ -\left(\mathbf{x}\mathbf{x}^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H\right)^{-1}\tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}} \right. \\ &\quad \left. + \mathbf{D}^{-1}\tilde{\mathbf{S}}\left(\mathbf{H}_t + \tilde{\Xi}\right) - \mathbf{D}^{-1}\mathbf{x}\left(\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H \right\} \\ &\quad \times \left(\mathbf{I}_M \otimes \mathbf{K}_{N_T N_R, M} \otimes \mathbf{I}_{LN_R}\right) \\ &\quad \times \left(\mathrm{vec}\left(\mathbf{I}_M\right) \otimes \mathbf{P}\right), \end{aligned} \tag{27}$$

where the matrix $\mathbf{D} = \left(\mathbf{x} - \tilde{\mathbf{S}}h_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}$, $\mathbf{H}_t = \mathbf{h}_t\hat{\boldsymbol{\alpha}}\left(\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H$, $\mathbf{P} = \left(\mathbf{I}_{N_t} \otimes \mathbf{K}_{N_R,L} \otimes \mathbf{I}_{N_R}\right)\left(\mathbf{I}_{N_t L} \otimes \mathrm{vec}\left(\mathbf{I}_{N_R}\right)\right)$, $\mathbf{K}_{N_R,L}$ is a $N_R L \times N_R L$ commutation matrix, and $\mathbf{K}_{N_T N_R, M}$ is a $MN_R N_T \times MN_R N_T$ commutation matrix [37].

Proof: See Appendix A.

In addition, with the given $\left\{\mathbf{s}_I^k, \mathbf{u}^k\right\}$, the derivative of $g_B\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_R$ can be written as

$$\frac{\partial g_B\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)}{\partial \mathbf{s}_R} = 4\mathbf{s}_R^3 + 4\mathbf{b}_R^k \odot \mathbf{s}_R, \tag{28}$$

where $\mathbf{s}_R^3$ is defined by taking a cube of each element in $\mathbf{s}_R$.

Proof: See Appendix B.

By combing Eqs.(26) and (28), the gradient of $L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_R$ can be expressed as

$$\begin{aligned}&\nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right) \\ &= \frac{\partial L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)}{\partial \mathbf{s}_R} \\ &= 2\Re\left\{ \left.\frac{\partial g_A}{\partial \mathbf{s}^*}\right|_{\substack{\mathbf{s}_I=\mathbf{s}_I^k,\\ \hat{\boldsymbol{\alpha}}=\hat{\boldsymbol{\alpha}}_k}} \right\} + 2\rho\left(\mathbf{s}_R^3 + \mathbf{b}_R^k \odot \mathbf{s}_R\right). \end{aligned} \tag{29}$$

Herein, let τ be the iteration number and $\mathbf{A}_\tau$ be the approximation to the Hessian matrix of $L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_R$. Then the BFGS method [38] to update $\mathbf{s}_R$ with a fixed step length ur is summarized in algorithm 1.

Algorithm 1 Update of $\mathbf{s}_R$

Inputs: $\left\{\mathbf{s}_R^k, \mathbf{s}_I^k, \mathbf{u}^k\right\}$, the MLE $\hat{\boldsymbol{\alpha}}_k$ of k-th time, step size ur, ρ, iteration stop tolerance εr, and the maximum iteration number $\tau_{\max}$.

Initialize: $\mathbf{A}_0 = \mathbf{I}_{LN_T}$, $\tau = 0$;

1: **while** $\left\|\nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R, \mathbf{s}_I^k, \mathbf{u}^k\right)\right\|_2 > \varepsilon r$ and $\tau < \tau_{\max}$ **do**

2: Compute $\nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R^\tau, \mathbf{s}_I^k, \mathbf{u}^k\right)$ using Eq. (29);

3: Set $\lambda_\tau^R = -ur\mathbf{B}_\tau^{-1}\nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{S}_R^\tau, \mathbf{S}_I^k, \mathbf{u}^k\right)$, update $(\mathbf{s}_R)_{\tau+1} = (\mathbf{s}_R)_\tau + \lambda_\tau^R$;

4: Calculate $\nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R^{\tau+1}, \mathbf{s}_I^k, \mathbf{u}^k\right)$ using Eq. (29), and determine:

$$\mathbf{z}_\tau = \nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R^{\tau+1}, \mathbf{s}_I^k, \mathbf{u}^k\right) - \nabla_{\mathbf{s}_R} L_\rho\left(\mathbf{s}_R^\tau, \mathbf{s}_I^k, \mathbf{u}^k\right);$$

5: if $\mathbf{z}_\tau^T\lambda_\tau^R > 0$, update $\mathbf{A}_{\tau+1}$ with

$$\mathbf{A}_{\tau+1} = \mathbf{A}_\tau + \frac{\mathbf{z}_\tau\mathbf{z}_\tau^T}{\mathbf{z}_\tau^T\lambda_\tau^R} - \frac{\mathbf{A}_\tau\lambda_\tau^R\left(\lambda_\tau^R\right)^T\mathbf{A}_\tau^T}{\left(\lambda_\tau^R\right)^T\mathbf{A}_\tau\lambda_\tau^R};$$ else $\mathbf{A}_{\tau+1} = \mathbf{A}_\tau$;

6: $\tau = \tau + 1$;

7: **end while**

8: Update $\mathbf{v}$ using Eq. (24b)

Outputs: Real part of the MIMO radar waveform $\mathbf{s}_R^{k+1}$ and $\mathbf{v}^{k+1}$

Update of $\mathbf{s}_I$: Due to the symmetry of $\mathbf{s}_R$ and $\mathbf{s}_I$, the update method of $\mathbf{s}_I$(Eq. 24c) is akin to that of $\mathbf{s}_R$ (Eq. 24a).

Let $\mathbf{b}_I^k = \mathbf{v}^{k+1} - \mathbf{1} + \mathbf{u}^k$, then $g_B\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right) = \left\|\mathbf{w}+\mathbf{b}_I^k\right\|_2^2$. Hence, the derivative of $g_B\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_I$ can be expressed as

$$\frac{\partial g_B\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right)}{\partial \mathbf{s}_I} = 4\mathbf{s}_I^3 + 4\mathbf{b}_I^k \odot \mathbf{s}_I, \tag{30}$$

where $\mathbf{s}_I^3$ is defined by taking a cube of each element in $\mathbf{s}_I$.

Furthermore, using the expression of Eqs. (27) and the results of algorithm 1, the derivative of $g_A\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_I$ can be written as

$$\begin{aligned}\frac{\partial g_A\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I\right)}{\partial \mathbf{s}_I} &= \frac{\partial g_A\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I\right)}{\partial \mathrm{vec}\left(\mathbf{S}_I\right)} \\ &= 2\mathrm{Im}\left\{ \left.\frac{\partial g_A}{\partial \mathbf{s}^*}\right|_{\substack{\mathbf{s}_R=\mathbf{s}_R^{k+1},\\ \hat{\boldsymbol{\alpha}}=\hat{\boldsymbol{\alpha}}_k}} \right\}, \end{aligned} \tag{31}$$

Therefore, the gradient of $L_\rho\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right)$ with respect to $\mathbf{s}_I$ can be expressed as

$$\nabla_{\mathbf{s}_I} L_\rho\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I, \mathbf{u}^k\right) = 2\rho\left(\mathbf{s}_I^3 + \mathbf{b}_I^k \odot \mathbf{s}_I\right) + 2\mathrm{Im}\left\{\left.\frac{\partial g_A}{\partial \mathbf{s}^*}\right|_{\substack{\mathbf{s}_R=\mathbf{s}_R^{k+1},\\ \hat{\boldsymbol{\alpha}}=\hat{\boldsymbol{\alpha}}_k}}\right\}. \quad (32)$$

In summary, the design process of the ADMM-based waveform (ADMM-W) is summarized in algorithm 2.

Algorithm 2 ADMM-Based Waveform (ADMM-W) Design

Inputs: step size ρ, iteration stop tolerance ς, and the maximum iteration time $K_{\max}$

Initialize: $\mathbf{s}_R, \mathbf{s}_I, \mathbf{s}^0 = \mathbf{s}_R^0 + j\mathbf{s}_I^0$ and $k = 0$;

1: **while** max $\left\{\left|L_\rho\left(\mathbf{s}_R^{k+1}, \mathbf{s}_I^{k+1}, \mathbf{u}^k\right) - L_\rho\left(\mathbf{s}_R^k, \mathbf{s}_I^k, \mathbf{u}^k\right)\right|, \left|\left\|\mathbf{u}^{k+1}\right\|_2 - \left\|\mathbf{u}^k\right\|_2\right|\right\} < \varsigma$ and $k < K_{\max}$ **do**

2: Compute the MLE of target amplitude $\hat{\boldsymbol{\alpha}}_k$ using (16);

3: Update $\mathbf{s}_R^{k+1}$, $\mathbf{v}^{k+1}$ of Eq. (24a) and (24b) using **Algorithm 1**;

4: Update $\mathbf{s}_I^{k+1}$, $\mathbf{w}^{k+1}$ of Eq. (24c) and (24d) using the method akin to **Algorithm 2**;

5: Update $\mathbf{u}^{k+1}$ using Eq. (24e);

6: Compute $\mathbf{s}^{k+1} = \mathbf{s}_R^{k+1} + j\mathbf{s}_I^{k+1}$

7: $k = k + 1$;

8: **end while**

Output: MIMO radar waveform $\mathbf{S}$

B. CONVERGENCE PERFORMANCE AND COMPUTATIONAL COMPLEXITY ANALYSIS

First, the convergence performance of the ADMM-based algorithm is discussed. ADMM can guarantee the global convergence for a convex optimization problem [32]. The theoretical proof of convergence for a nonconvex problem is proposed in Theorem 1 [14], [33], [34]. The following theorem reveals that sequences generated by the ADMM-W algorithm are convergent under some mild conditions.

Theorem 1: Let $\left\{\mathbf{s}_R^k, \mathbf{s}_I^k, \mathbf{u}^k\right\}$ be a sequence generated by ADMM-W with $\rho > 0$ (Eqs.24(a)−(e)). Now, if $\lim_{t\to\infty} \mathbf{u}^{k+1} - \mathbf{u}^k = \mathbf{0}$, then the limit point $\left\{\left(\mathbf{s}_R^\star, \mathbf{s}_I^\star, \mathbf{u}^\star\right)\right\}$ is an optimal solution of Eq. (22).

Proof: Since $\lim_{k\to\infty} \mathbf{u}^{k+1} - \mathbf{u}^k = \mathbf{0}$ and $\rho > 0$, it can obtained from Eq. (24e) that

$$\lim_{k\to\infty} \mathbf{v}^{k+1} + \mathbf{w}^{k+1} = \mathbf{1}. \quad (33)$$

Moreover, from Eqs. (24b), (24d), and (33), the following expressions can be achieved.

$$\begin{aligned} \mathbf{0} \le \mathbf{v}^k \le \mathbf{1}, \\ \mathbf{0} \le \mathbf{w}^k \le \mathbf{1}. \end{aligned} \quad (34)$$

Hence, the sequences $\left\{\mathbf{s}_R^k\right\}$ and $\left\{\mathbf{s}_I^k\right\}$ are bounded, and a stationary point $\left\{\left(\mathbf{s}_R^\star, \mathbf{s}_I^\star, \mathbf{u}^\star\right)\right\}$ exists between them.

$$\lim_{t\to\infty} \left\{\left(\mathbf{s}_R^k, \mathbf{s}_I^k, \mathbf{u}^k\right)\right\} = \left\{\left(\mathbf{s}_R^\star, \mathbf{s}_I^\star, \mathbf{u}^\star\right)\right\}. \quad (35)$$

Therefore, Theorem 1 is proved.

Next, we analyze the computational complexity of ADMM-W algorithm.

As to the computational complexity of ADMM-W, it mainly depends on the number of (outer) iterations and the number of operations at each iteration. Herein, we focus on the analysis of the complexity involved in each (outer) iteration. For Algorithm ADMM-W, the update of $\hat{\boldsymbol{\alpha}}$(with (16)) requires $O\left((LMN_R)^3\right)$ operations, and the optimization of $\mathbf{s}_R$ requires $O\left(\tau_R\left(LN_T\right)^3\right)$ operations with τ_R denoting the number of (inner) iterations of Algorithm 1. Similarly, the optimization of $\mathbf{s}_I$ requires $O\left(\tau_I\left(LN_T\right)^3\right)$ operations with τ_I denoting the number of (inner) iterations. Last, the update of $\mathbf{s}$ requires $O\left(LN_T\right)$ operations.

IV. PHASE-ONLY WAVEFORM DESIGN

Since the ADMM-W algorithm is sensitive to the parameter ρ and the algorithm is quite complicated, it is necessary to find a simpler method to solve the optimization problem (19). In the present section, the phase-only waveform design algorithm, also based on the BFGS method [38], is employed to solve Eq. (19).

A. PHASE-ONLY ALGORITHM

In order to solve this nonconvex problem, CM constraints can be simplified by introducing auxiliary phase variables.

$$\boldsymbol{\Phi} = \begin{bmatrix} \varphi_{11} & \varphi_{12} & \cdots & \varphi_{1N_T} \\ \varphi_{21} & \varphi_{22} & \cdots & \varphi_{2N_T} \\ \vdots & \vdots & \vdots & \vdots \\ \varphi_{L1} & \varphi_{L2} & \cdots & \varphi_{LN_T} \end{bmatrix}, \quad (36)$$

Hence, Eq. (19) is converted into the following optimization problem without performing any modulus operation:

$$\begin{aligned} \min_{\mathbf{S}} \ & \ln\left\{\det\left[\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]\right\} \\ & - \ln\left\{\det\left[\mathbf{x}\mathbf{x}^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]\right\} \\ s.t. \ & \mathbf{S} = e^{j\boldsymbol{\Phi}}. \end{aligned} \quad (37)$$

Now, by substituting $\mathbf{S} = e^{j\boldsymbol{\Phi}}$ into the cost function of Eq. (37), the optimization problem can be relaxed as an unconstrained nonlinear optimization problem $g_A(\varphi)$ associated with $\boldsymbol{\Phi}$:

$$\begin{aligned} \min_{\boldsymbol{\Phi}} & \ g_A(\boldsymbol{\Phi}) \\ & = \ln\left\{\det\left[\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)\left(\mathbf{x} - \tilde{\mathbf{S}}\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]\right\} \\ & \quad - \ln\left\{\det\left[\mathbf{x}\mathbf{x}^H + \tilde{\mathbf{S}}\tilde{\boldsymbol{\Xi}}\tilde{\mathbf{S}}^H + \sigma\tilde{\mathbf{I}}\right]\right\}. \end{aligned} \quad (38)$$

which is an unconstrained problem. Similar to algorithm 1, the BFGS method is also adopted to solve Eq. (38).

According to the derivative theorem of a scalar real-valued function with respect to the complex matrix variables $\mathbf{Z}$ and

$\mathbf{Z}^* \in \mathbb{C}^{N\times Q}$ [36], the gradient of $g_A(\mathbf{\Phi})$ with respect to $\mathbf{\Phi}$ can be given as

$$\begin{aligned}\nabla_{\mathbf{\Phi}} g_A(\mathbf{\Phi}) &= \frac{\partial g_A(\mathbf{\Phi})}{\partial \mathbf{\Phi}} = \frac{\partial g_A}{\partial \mathbf{S}}\frac{\partial \mathbf{S}}{\partial \Phi} + \frac{\partial g_A}{\partial \mathbf{S}^*}\frac{\partial \mathbf{S}^*}{\partial \mathbf{\Phi}} \\ &= \frac{\partial g_A}{\partial \mathbf{S}} \odot (j\mathbf{S}) + \frac{\partial g_A}{\partial \mathbf{S}^*} \odot \left(-j\mathbf{S}^*\right) \\ &= 2\mathrm{Im}\left\{\frac{\partial g_A}{\partial \mathbf{S}^*} \odot \mathbf{S}^*\right\}. \end{aligned} \tag{39}$$

And $\partial g_A \big/ \partial \mathbf{S}^*$ can be calculated from Proposition 1.

Let $\boldsymbol{\varphi} = \mathrm{vec}^T(\mathbf{\Phi})$; thus the gradient of $g_A(\mathbf{\Phi})$ with respect to $\mathbf{\Phi}$ can be expressed as

$$\nabla_{\boldsymbol{\varphi}} g_A(\boldsymbol{\varphi}) = \mathrm{vec}\left[\nabla_{\mathbf{\Phi}} g_A(\mathbf{\Phi})\right]. \tag{40}$$

Based on the BFGS method, the phase vector $\boldsymbol{\varphi}_{k+1}$ at iteration $k+1$ can be updated as

$$\boldsymbol{\varphi}_{k+1} = \boldsymbol{\varphi}_k - u_k \mathbf{B}_k^{-1} \nabla_{\boldsymbol{\varphi}_k} g_A(\varphi_k), \tag{41}$$

where u_k is the selected step size in the current search direction and $\mathbf{B}_k$ is the approximation to the Hessian matrix of $g_A(\boldsymbol{\varphi}_k)$ with respect to $\boldsymbol{\varphi}_k$. Algorithm 3 summarizes the phase-only waveform design (POW) process using the BFGS method.

Algorithm 3 Phase-Only Waveform (POW) Design

Inputs: step size $u, \beta, u \in (0,1), \beta \in (0, 0.5)$, iteration stop tolerance ε, and the maximum iteration time $K_{\max}$;
Initialize: $\mathbf{B}_0 = \mathbf{I}_{LN_T}$, $\boldsymbol{\varphi}_0$, $k = 0$
1: while $\left\|\nabla_{\boldsymbol{\varphi}_k} g_A(\boldsymbol{\varphi}_k)\right\|_2 > \varepsilon$ and $k < K_{\max}$ do
2: Compute the MLE of target amplitude $\hat{\boldsymbol{\alpha}}_k$ using Eq.(16);
3: Compute $\nabla_{\boldsymbol{\varphi}_k} g_A(\boldsymbol{\varphi}_k)$ using Eqs. (39)–(40), and determine the search direction $\mathbf{p}_k = -\mathbf{B}_k^{-1}\nabla_{\boldsymbol{\varphi}_k} g_A(\boldsymbol{\varphi}_k)$;
4: Perform a line search with the Armijo rule [38] to find the smallest positive integer that holds the following formula.
$g_A(\boldsymbol{\varphi}_k + u^{mk}\mathbf{p}_k) \le g_A\left(\boldsymbol{\varphi}_k\right) + \beta u^{mk} \nabla_{\boldsymbol{\varphi}} g_A(\boldsymbol{\varphi}_k)^T \mathbf{p}_k$;
5: Set $\mathbf{r}_k = u^{mk}\mathbf{p}_k$ to update $\boldsymbol{\varphi}_{k+1} = \boldsymbol{\varphi}_k + \mathbf{r}_k$;
6: Calculate $\nabla_{\boldsymbol{\varphi}_{k+1}} g_A(\boldsymbol{\varphi}_{k+1})$ with Eqs. (39)–(40), and determine
$\mathbf{y}_k = \nabla_{\boldsymbol{\varphi}_{k+1}} g_A(\boldsymbol{\varphi}_{k+1}) - \nabla_{\boldsymbol{\varphi}_k} g_A(\boldsymbol{\varphi}_k)$;
7: if $\mathbf{y}_k^T \mathbf{r}_k > 0$, update $\mathbf{B}_{k+1} = \mathbf{B}_k + \frac{\mathbf{y}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{r}_k} - \frac{\mathbf{B}_k \mathbf{r}_k \mathbf{r}_k^T \mathbf{B}_k^T}{\mathbf{r}_k^T \mathbf{B}_k \mathbf{r}_k}$; else $\mathbf{B}_{k+1} = \mathbf{B}_k$;
8: $k = k + 1$;
9: end while
Output: MIMO radar waveform $\mathbf{S}$

B. CONVERGENCE PERFORMANCE AND COMPUTATIONAL COMPLEXITY ANALYSIS

POW is based on the BFGS method, and the convergence process of the BFGS can be found in literature [38], thus the convergence analysis of POW is omitted here.

The computational complexity of POW mainly depends on the number of (outer) iterations and the number of operations at each iteration. In precise, at each iteration, the update of $\hat{\boldsymbol{\alpha}}$(with (16)), same with ADMM-W, requires $O\left((LMN_R)^3\right)$ operations. The optimization of $\boldsymbol{\varphi}$ requires $O\left((LN_T)^3\right)$. Compared with ADMM-W, the computational complexity of POW per iteration is lower.

V. SIMULATION

In the current section, a MIMO radar system consisting of 4 transmit antennas (N_T) and 4 receiving antennas (N_R) is considered. The proposed waveforms are tested using MCARM data (obtained from the MCARM file re050575) [39]. In this file, datacubes are comprised of 22 channels, 128 pulses, and 630 unambiguous range cells. The element 1:4, the channel 1:4:13, and the pulse 1:16 are, respectively, assumed as the receiver array, the transmit array, and the number of pulses in a CPI. The diagonal loading coefficient in Eq. (18) is set as $\sigma = 0.4$.

Some numerical examples are also provided to demonstrate the performances of the waveforms designed by algorithms 1, 2 and 3. In both algorithms 1 and 3, the phase vector $\boldsymbol{\varphi}$, $\mathbf{s}_R$, and $\mathbf{s}_I$ are initialized to be random.

A. MCARM FILE re050575 DATA PREPROCESSING

In the re050575 file of range bin 290, only the clutter data of 22 channels are available; however, it is indispensable to obtain the clutter patch in the azimuth dimension. Therefore, an iterative adaptive approach (IAA) algorithm [40] is employed to estimate the clutter patch amplitude $\delta_{c,m,k}$ described in Eq. (6). Figure 2a displays the origin clutter distribution at range bin 290, and the estimated clutter patch amplitude $\delta_{c,m,k}$ is presented in Fig. 1b. Furthermore, the clutter sample covariance matrix $\mathbf{R}_{SCM}$ in Eq. (9) is obtained to measure the concentration matrix $\mathbf{\Sigma}$ in Eq. (17).

In this paper, we consider I point targets model with fixed target amplitudes from trial to trial. In this case, the signal-to-clutter ratio (SCR) is defined as

$$\mathrm{SCR} = \frac{\sum_{i=1}^{I} |\alpha_i|^2}{\sum_{m=1}^{M}\sum_{\kappa=1}^{N_c} \left|\delta_{c,m,\kappa}\right|^2}. \tag{42}$$

Besides, the space-time cross-ambiguity can evaluate the clutter suppression performance of the waveform. And according to the space-time cross-ambiguity expression for a target in [17], we can extend that the space-time cross-ambiguity for multiple targets as

$$P(\theta, f_d) = \left|\sum_{i=1}^{I} \mathbf{w}_i^H \tilde{\mathbf{S}}\mathbf{h}_{t_i}\right|^2, \tag{43}$$

where $\mathbf{w}_i = \left(\mathbf{R}_{SCM} + \sigma\mathbf{I}\right)^{-1}\tilde{\mathbf{S}}\mathbf{h}_{t_i}$ is the optimal receiver filter for the i-th target.

B. PERFORMANCES OF DIFFERENT ALGORITHMS FOR A SINGLE TARGET

In the present subsection, the performances of the phase-only algorithm and the ADMM-based algorithm are evaluated

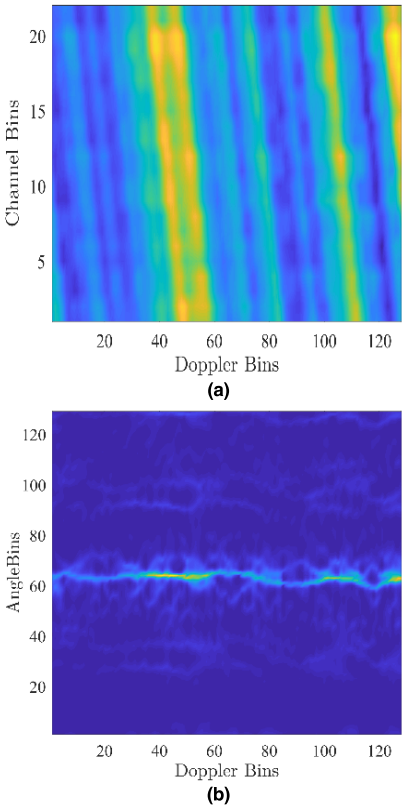


FIGURE 2. (a) Origin clutter amplitude distribution at range bin 290 and (b) estimated clutter patch amplitude.

with a low SCR = 5dB of a target, i.e. $I = 1$. The target is artificially injected at Doppler bin 3, range bin 290, and angle bin 65. Parameters for MCARM data are presented in Table 1. The number of time samples and the freedom are considered as $L = 16$ and $v = MN_RL + 2$, respectively.

TABLE 1. Parameters for the MCARM file re050575.

Parameters	Value
Pulse number (M)	128
Element spacing (d)	0.1092 m
Carrier frequency	1240 MHz
Bandwidth	800 KHz
PRF	1984 Hz
λ	0.2419 m

1) CONVERGENCE PERFORMANCE AND CM PROPERTY OF ADMM-W

In the current experiment, the convergence performance and the CM property of the ADMM waveform (ADMM-W)-based design algorithm are explored. The step sizes of algorithm 2 for each parameter ρ are set $ur = 0.3$, respectively.

The convergence performance of the ADMM-W-based design algorithm under different step sizes (ρ) is investigated.

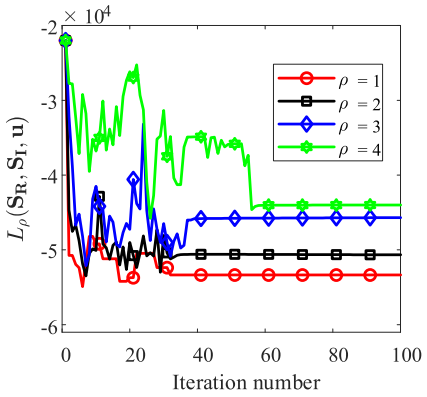


FIGURE 3. Influences of step size (ρ) on the convergence of ADMM-W at SCR = 5dB for a single target.

Figure 3 presents a comparison between the obtained objective values L_ρ ($\mathbf{s}_R$, $\mathbf{s}_I$, $\mathbf{u}$) at different step sizes. It is noticeable that the ADMM-based algorithm with a smaller step size (for example, $\rho = 0.5$) converged faster than that with a larger step size (for example, $\rho = 4$); however, a too small step size (for example, $\rho = 0.5$) causes a big oscillation in the obtained objective function values. In addition, the objective function gradually increases with the increasing step size. It should be noted that the choice of parameter ρ has a great influence on the performance of the ADMM-based algorithm. As to the computational complexity of ADMM-W, the run-time in each (outer) iteration is about 278s for each parameter ρ.

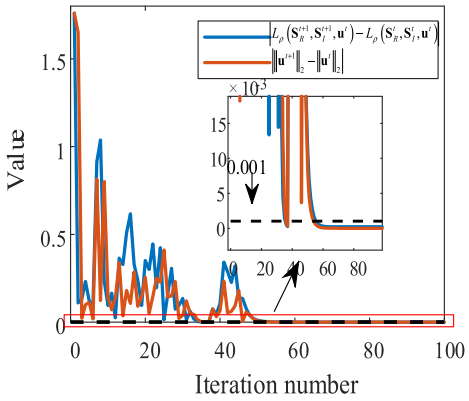


FIGURE 4. ADMM-W with SCR = 5dB, $\rho = 1$ for a single target.

Furthermore, the CM property of the obtained waveforms is analyzed by considering $\rho = 1$ and the iteration stop tolerance $\varsigma = 0.001$. It is evident from Fig.4 that the values start to converge after 45 iterations. Let the symbol 'Tx' denote the transmit antenna. Figure 5 presents the modules, and the phase values of ADMM-W. We can see that both the real and imaginary parts exist between -1 and 1. Therefore, from Fig.5a, it can be inferred that the ADMM-W fulfilled the constant modulus constraint of the MIMO radar.

2) CONVERGENCE PERFORMANCE AND CM PROPERTY OF POW

In the current experiment, the convergence performance and the CM property of the POW algorithm are explored.

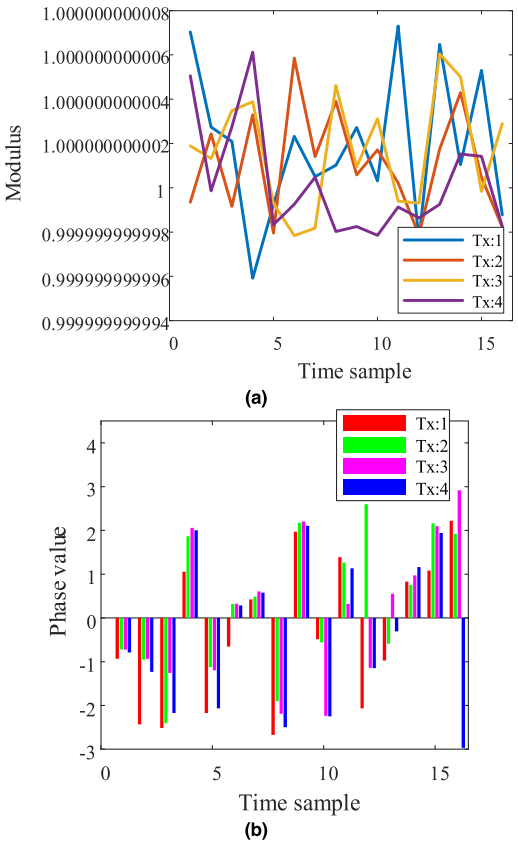


FIGURE 5. (a) modules, and (b) phase values of ADMM-W with $\rho = 1$, SCR $= 5$dB for a single target.

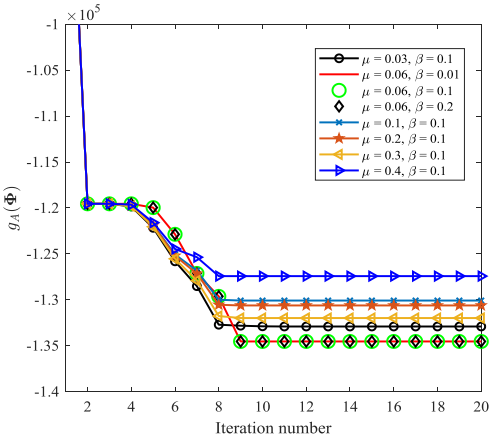


FIGURE 6. Influences of step size (μ, β) on the convergence of POW at SCR $= 5$dB.

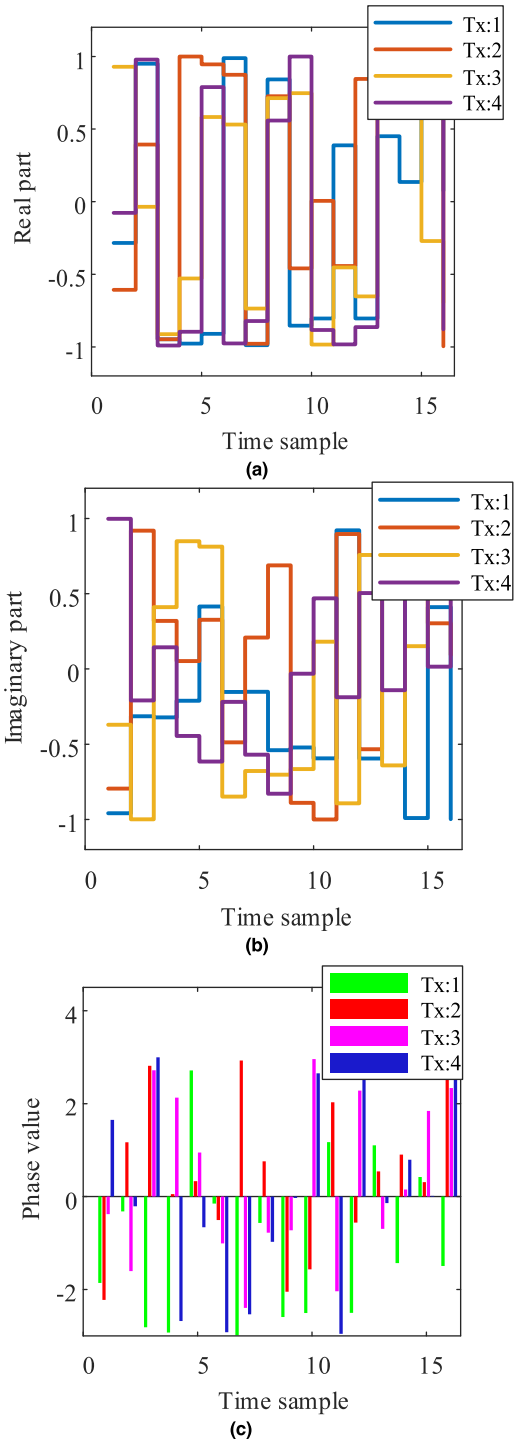


FIGURE 7. (a) Real part, (b) imaginary part, and (c) phase values of POW with $\mu = 0.06$, $\beta = 0.1$ for a single target.

The influences of step size on the convergence of the POW design algorithm are first examined. The curves of the objective value $g_A(\mathbf{\Phi})$ with different step sizes (μ, β) and the same initial phase value are plotted in Fig.6. It is observable that the POW design algorithm has the fastest convergence rate at $\mu = 0.4$, whereas a large step size μ results in a large objective function value. In addition, the step size β has little effect on the objective function. When $\mu = 0.06$, $\beta = 0.01$, $g_A(\mathbf{\Phi})$ obtained the local minimum value.

As to the computational complexity of POW, the run-time in each (outer) iteration is about 31s, which is significantly lower than ADMM-W.

It is evident from Fig.7 that both real (Fig.7a) and imaginary parts (Fig.7b) of each antenna existed between -1 and 1. Moreover, the phase values of the POW algorithm are presented in Fig.7c.

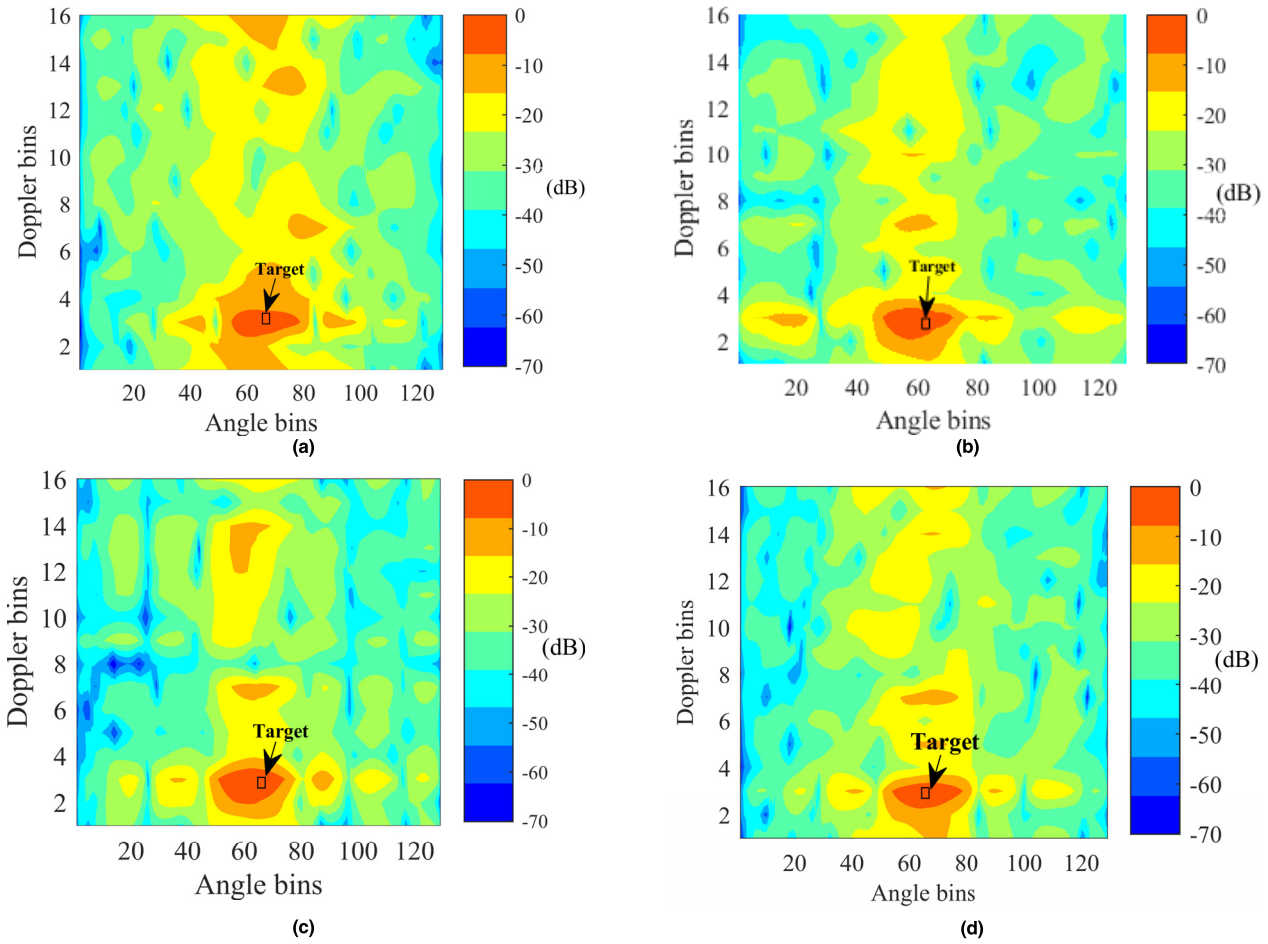


FIGURE 8. Space-time cross-ambiguity of (a) orthogonal LFM, (b) [17]: Algorithm 2 (c) ADMM-W with $\rho = 2$, and (d) POW with $\mu = 0.4$, $\beta = 0.01$ for a single target.

C. CLUTTER SUPPRESSION PERFORMANCES OF THE WAVEFORMS DESIGNED FOR SINGLE TARGET

In order to assess the space-time cross-ambiguity of different waveforms at SCR = 5dB of a target, their clutter suppression performances are evaluated.

In [17], the authors proposed some algorithms to jointly optimize the transmit CM waveform and receive filter for single target MIMO radar systems. Especially, the Algorithm 2 is one of the method to design CM waveform. Thus, the waveform designed by Algorithm 2 is provided here. In addition, the orthogonal LFM waveform [15] is also offered to compare, i.e., the corresponding waveform matrix is given by

$$\mathbf{S}_{LFM}(n_t, l) = \exp\left\{j2\pi n_t (l-1)/L\right\} \exp\left\{j\pi (l-1)^2/L\right\}, \tag{44}$$

where $n_t = 1, 2, \ldots, N_T$ and $l = 1, 2, \ldots, L$.

Fig.8 shows the space-time cross-ambiguity (Eq.43) associated with the different waveforms. It is clear from Fig. 8 that the target is located in the main lobe (Doppler bin 3 and angle bin 65) for all the waveforms. Compared to Fig.8b, Fig.8c and Fig.8d, Fig.8a has the biggest region in yellow. Moreover, as shown in Fig.8b and Fig.8d, the Algorithms 2 of [17] has similar performance with POW. As a consequence, our designed waveforms can suppress the power of clutter to a desired low level of Doppler bins and angle bins.

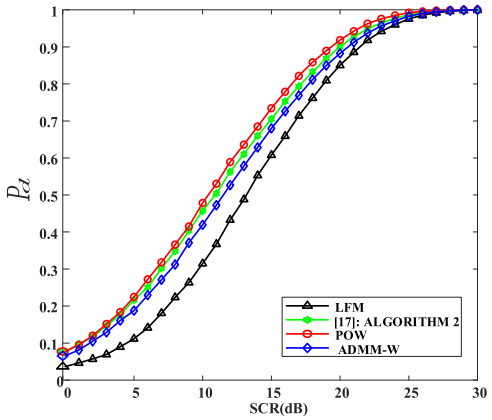


FIGURE 9. P_d of different waveforms versus SCR for $P_{fa} = 10^{-3}$.

D. DETECTION PERFORMANCE TEST FOR SINGLE TARGET

In the current subsection, the detection performances of our designed waveforms, the waveform designed by the

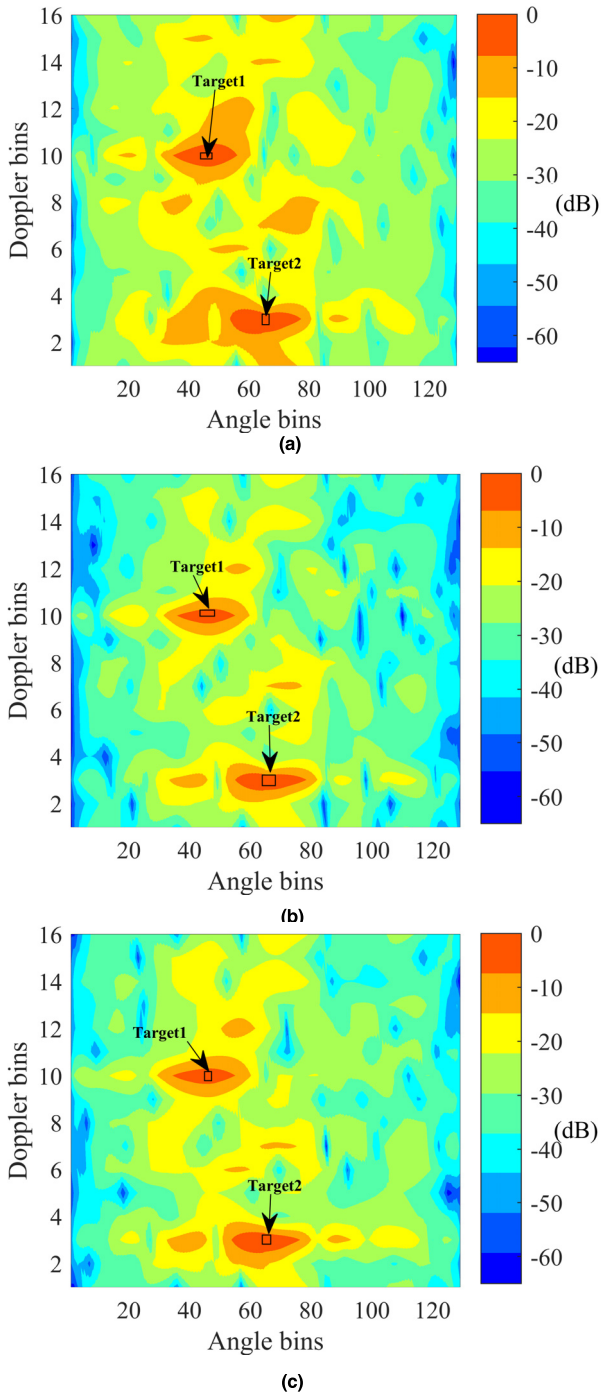


FIGURE 10. Space-time cross-ambiguity of (a) orthogonal LFM, (b) ADMM-W with $\rho = 2$, and (c) POW with $\mu = 0.06$, $\beta = 0.2$ for two targets.

Algorithm 2 of [17] and the orthogonal LFM waveform (Eq. 44) are compared. In order to save simulation time, the value of M is set to 1 and the other parameters are kept unchanged. According to Eq.(15), under the proposed Bayesian framework, the detector associated with a transmitted waveform for a given probability of false alarm P_{fa} (referred as the Bayesian-Optimal-detector $T_{Bayesian}$) can be expressed as

$$T_{Bayesian} = \frac{\det\left[(\mathbf{x}-\mathbf{x}_t)(\mathbf{x}-\mathbf{x}_t)^H + (\nu - MN_RL)\,\boldsymbol{\Sigma}\right]^{-1}}{\det\left[\mathbf{x}\mathbf{x}^H + (\nu - MN_RL)\,\boldsymbol{\Sigma}\right]^{-1}} \underset{H_0}{\overset{H_1}{>}} \gamma, \tag{45}$$

where γ is the threshold (estimated from the results of $100/P_{fa}$ Monte Carlo independent trials). The values of detection probability (P_d) are also obtained from these independent trials.

Figure 9 displays the detection probability P_d versus SCR curves of different waveforms for given false alarm probability $P_{fa} = 10^{-3}$. It is evident that the performance of radar detection is improved obviously by waveforms design. And the $T_{Bayesian}$ detector had the best performance with POW, better performance with ADMM-W, worst with LFM. In addition, the waveform designed by Algorithm 2 of [17] has a close detection performance with POW.

E. CLUTTER SUPPRESSION PERFORMANCES OF THE WAVEFORMS DESIGNED FOR TWO TARGETS

In the present subsection, the space-time cross-ambiguity associated with LFM of (44), ADMM-W and POW are evaluated in the scenario for two targets, i.e., $I = 2$. The target 1 is artificially injected at Doppler bin 10, and angle bin 45. The target 2 is artificially injected at Doppler bin 3, and angle bin 65. And the rest parameters are not changed.

Since the method of [17] is only applicable to a single target scenarios, it is not appropriate to compare Algorithm 2 of [17] with our algorithm here. Behaviors similar to Fig.8, Fig. 10 assessed the space-time cross-ambiguity associated with the designed waveforms by ADMM-W with $\rho = 2$ and POW with $\mu = 0.06$, $\beta = 0.2$. As shown in Fig.10, the space-time cross-ambiguity function has two mainlobes, such as the first target is located in Doppler bin 10, angle bin 45, and the second target is located in Doppler bin 3, and angle bin 65. We also can see that in Fig.10b, the blue area is the largest and the yellow area is the smallest, that is, the clutter suppression effect of ADMM-W is the best. As a result, the waveforms designed by us manifest significantly better performance in suppressing the clutter of Doppler bins and angle bins of two targets.

VI. CONCLUSIONS

In this study, the transmit waveform design for target detection problems in heterogeneous clutter was addressed for an airborne colocated MIMO radar. For the signal-dependent clutter model, a Bayesian framework was considered, and the corresponding GLRT expression for waveform design problem was established. The POW and the ADMM-W methods were proposed to solve this problem. Simulation results revealed that the optimal waveform significantly improved the detection performance and suppressed the strong clutter of MCARM dataset; therefore, it can be inferred that the Bayesian detector manifested a better detection performance. In addition, for a single target, the ADMM-W method is

not recommended because of its high complexity, long time consuming and sometimes not easy to converge. However, for multiple targets, it should be considered a compromise between the two methods.

APPENDIX A
PROOF OF PROPOSITION 1

Let we define that

$$D_{\tilde{\mathbf{S}}^*} \triangleq g_A(\mathbf{S},\mathbf{S}^*)\frac{\partial g_A}{\partial \mathrm{vec}^T\left(\tilde{\mathbf{S}}^*\right)}, \tag{46}$$

$$D_{\mathbf{S}_t^*} \triangleq \left(\tilde{\mathbf{S}}^*\right)\frac{\partial \mathrm{vec}\left(\tilde{\mathbf{S}}^*\right)}{\partial \mathrm{vec}^T\left(\mathbf{S}^*\right)}. \tag{47}$$

and

$$\begin{aligned}\mathrm{vec}\left(\mathbf{S}^*\otimes\mathbf{I}_{N_R}\right) &= \left(\mathbf{I}_{N_t}\otimes\mathbf{K}_{N_R,L}\otimes\mathbf{I}_{N_R}\right)\\ &\quad\times\left[\left(\mathbf{I}_{N_tL}\mathrm{vec}\left(\mathbf{S}^*\right)\right)\otimes\left(\mathrm{vec}\left(\mathbf{I}_{N_R}\right)1\right)\right]\\ &= \left(\mathbf{I}_{N_t}\otimes\mathbf{K}_{N_R,L}\otimes\mathbf{I}_{N_R}\right)\\ &\quad\times\left[\left(\mathbf{I}_{N_tL}\otimes\mathrm{vec}\left(\mathbf{I}_{N_R}\right)\right)\left(\mathrm{vec}\left(\mathbf{S}^*\right)\otimes 1\right)\right]\\ &= \left(\mathbf{I}_{N_t}\otimes\mathbf{K}_{N_R,L}\otimes\mathbf{I}_{N_R}\right)\\ &\quad\times\left(\mathbf{I}_{N_tL}\otimes\mathrm{vec}\left(\mathbf{I}_{N_R}\right)\right)\mathrm{vec}\left(\mathbf{S}^*\right)\\ &= \mathbf{P}\mathrm{vec}\left(\mathbf{S}^*\right)\end{aligned} \tag{48}$$

where $\mathbf{P}=\left(\mathbf{I}_{N_t}\otimes\mathbf{K}_{N_R,L}\otimes\mathbf{I}_{N_R}\right)\left(\mathbf{I}_{N_tL}\otimes\mathrm{vec}\left(\mathbf{I}_{N_R}\right)\right)$.

Hence, it can be expressed that

$$\begin{aligned}&\mathrm{vec}\left(\tilde{\mathbf{S}}^*\right)\\ &= \mathrm{vec}\left(\mathbf{I}_M\otimes\mathbf{S}^*\otimes\mathbf{I}_{N_R}\right)\\ &= \left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\left(\mathrm{vec}\left(\mathbf{I}_M\right)\otimes\mathrm{vec}\left(\mathbf{S}^*\otimes\mathbf{I}_{N_R}\right)\right)\\ &= \left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\left(\mathrm{vec}\left(\mathbf{I}_M\right)\otimes\mathbf{P}\mathrm{vec}\left(\mathbf{S}^*\right)\right)\\ &= \left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\left(\mathrm{vec}\ \left(\mathbf{I}_M\right)\otimes\mathbf{P}\right)\mathrm{vec}\left(\mathbf{S}^*\right)\end{aligned} \tag{49}$$

Furthermore,

$$\begin{aligned}d\mathrm{vec}\left(\tilde{\mathbf{S}}^*\right) &= \left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\\ &\quad\times\left(\mathrm{vec}\left(\mathbf{I}_M\right)\otimes\mathbf{P}\right)\left[d\mathrm{vec}\left(\mathbf{S}^*\right)\right]\end{aligned} \tag{50}$$

Based on Table 3.2 in literature [36], it can be obtained that

$$\begin{aligned}D_{\mathbf{S}}\left(\tilde{\mathbf{S}}^*\right) &= \frac{\partial \mathrm{vec}\left(\tilde{\mathbf{S}}^*\right)}{\partial \mathrm{vec}^T\left(\mathbf{S}^*\right)}\\ &= \left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\left(\mathrm{vec}\left(\mathbf{I}_M\right)\otimes\mathbf{P}\right)\end{aligned} \tag{51}$$

Moreover, the differential of g_A can be described as

$$\begin{aligned}dg_A = &-Tr\left\{\tilde{\Xi}\tilde{\mathbf{S}}^H\left(\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}\right)^{-1}d\tilde{\mathbf{S}}\right\}\\ &-Tr\left\{\tilde{\Xi}^T\tilde{\mathbf{S}}^T\left(\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H\right)^{-T}d\tilde{\mathbf{S}}^*\right\}\\ &+Tr\left\{\hat{\alpha}\mathbf{B}^{-1}\mathbf{x}\mathbf{h}_t^H\left(d\tilde{\mathbf{S}}^H\right)\right\}+Tr\left\{\mathbf{B}^{-1}\left(d\tilde{\mathbf{S}}\right)\left(\mathbf{h}_t\hat{\alpha}\right)\mathbf{x}^H\right\}\\ &-Tr\left\{\mathbf{B}^{-1}\left(d\tilde{\mathbf{S}}\right)\left(\mathbf{H}_t+\tilde{\Xi}\right)\tilde{\mathbf{S}}^H\right\}\\ &-Tr\left\{\mathbf{B}^{-1}\tilde{\mathbf{S}}\left(\mathbf{H}_t+\tilde{\Xi}\right)\left(d\tilde{\mathbf{S}}^H\right)\right\}.\end{aligned} \tag{52}$$

According to the definition of the derivative *df* (Page no. 76 in literature [36]), the gradient of g_A with respect to $\tilde{\mathbf{S}}^*$ can be given as

$$\begin{aligned}\frac{\partial g_A}{\partial \mathrm{vec}^T\left(\tilde{\mathbf{S}}^*\right)} = \mathrm{vec}^T\Big[&-\left(\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H\right)^{-1}\tilde{\mathbf{S}}\tilde{\Xi}\\ &+\mathbf{B}^{-1}\tilde{\mathbf{S}}\left(\mathbf{H}_t+\tilde{\Xi}\right)-\mathbf{B}^{-1}\mathbf{x}\left(\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H\Big].\end{aligned} \tag{53}$$

Now, combing Eqs.(48)−(51), it can be proved that

$$\begin{aligned}\frac{\partial g_A}{\partial \mathrm{vec}^T\left(\mathbf{S}^*\right)} &= D_{\tilde{\mathbf{S}}^*}g_A(\mathbf{S},\mathbf{S}^*)D_{\mathbf{S}^*}\left(\tilde{\mathbf{S}}^*\right)\\ &= \mathrm{vec}^T\Big[-\left(\mathbf{x}\mathbf{x}^H+\tilde{\mathbf{S}}\tilde{\Xi}\tilde{\mathbf{S}}^H\right)^{-1}\tilde{\mathbf{S}}\tilde{\Xi}\\ &\quad+\mathbf{B}^{-1}\tilde{\mathbf{S}}\left(\mathbf{H}_t+\tilde{\Xi}\right)-\mathbf{B}^{-1}\mathbf{x}\left(\mathbf{h}_t\hat{\boldsymbol{\alpha}}\right)^H\Big]\\ &\quad\left(\mathbf{I}_M\otimes\mathbf{K}_{N_TN_R,M}\otimes\mathbf{I}_{LN_R}\right)\left(\mathrm{vec}\left(\mathbf{I}_M\right)\otimes\mathbf{P}\right).\end{aligned} \tag{54}$$

Hence, proposition 1 is proved.

APPENDIX B

With the given function $g_B\left(\mathbf{s}_R,\mathbf{s}_I^k,\mathbf{u}^k\right)$, it can be expressed that

$$\begin{aligned}g_B\left(\mathbf{s}_R,\mathbf{s}_I^k,\mathbf{u}^k\right) &= \left\|\mathbf{v}+\mathbf{b}_R^k\right\|_2^2=\left(\mathbf{v}+\mathbf{b}_R^k\right)^T\left(\mathbf{v}+\mathbf{b}_R^k\right)|\\ &= \mathbf{v}^T\mathbf{v}+2\left(\mathbf{b}_R^k\right)^T\mathbf{v}+C\\ &= \sum_{i=1}^{LN_T}(\mathbf{s}_R)_i^4+2\left(\mathbf{b}_R^k\right)_i(\mathbf{s}_R)_i^2+C,\end{aligned} \tag{55}$$

where C is unrelated to $\mathbf{s}_R$.

Let $(\mathbf{s}_R)_i$ denote the i-th element in $\mathbf{s}_R$, then

$$\frac{\partial g_C\left(\mathbf{s}_R,\mathbf{s}_I^k,\mathbf{u}^k\right)}{\partial\left(\mathbf{s}_R\right)_i} = 4\left(\mathbf{s}_R\right)_i^3+4\left(\mathbf{b}_R^k\right)_i(\mathbf{s}_R)_i. \tag{56}$$

Hence, the proof is completed.

REFERENCES

[1] 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.

[2] H. Li, Z. Wang, J. Liu, and B. Himed, "Moving target detection in distributed MIMO radar on moving platforms," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1524–1535, Dec. 2015.

[3] H. Li and B. Himed, "Transmit subaperturing for MIMO radars with co-located antennas," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 55–65, Feb. 2010.

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

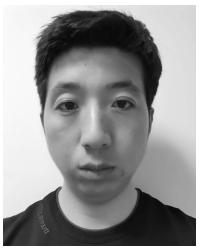
[5] D. R. Fuhrmann and G. San Antonio, "Transmit beamforming for MIMO radar systems using signal cross-correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 1, pp. 171–186, Jan. 2008.

[6] P. Stoica, J. Li, and X. Zhu, "Waveform synthesis for diversity-based transmit beampattern design," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2593–2598, Jun. 2008.

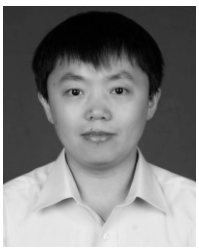
[7] S. Ahmed, J. S. Thompson, Y. R. Petillot, and B. Mulgrew, "Finite alphabet constant-envelope waveform design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5326–5337, Nov. 2011.

[8] J. Lipor, S. Ahmed, and M.-S. Alouini, "Fourier-based transmit beampattern design using MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2226–2235, May 2014.

[9] X. Yu, G. Cui, T. Zhang, and L. Kong, "Constrained transmit beampattern design for colocated MIMO radar," *Signal Process.*, vol. 144, pp. 145–154, Mar. 2018.
[10] Y.-C. Wang, X. Wang, H. Liu, and Z.-Q. Luo, "On the design of constant modulus probing signals for MIMO radar," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4432–4438, Aug. 2012.
[11] X. Zhang, Z. He, L. Rayman-Bacchus, and J. Yan, "MIMO radar transmit beampattern matching design," *IEEE Trans. Signal Process.*, vol. 63, no. 8, pp. 2049–2056, Apr. 2015.
[12] S. Ahmed and M.-S. Alouini, "MIMO radar transmit beampattern design without synthesising the covariance matrix," *IEEE Trans. Signal Process.*, vol. 62, no. 9, pp. 2278–2289, May 2014.
[13] Z. Cheng, Z. He, S. Zhang, and J. Li, "Constant modulus waveform design for MIMO radar transmit beampattern," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4912–4923, Sep. 2017.
[14] W. Fan, J. Liang, and J. Li, "Constant modulus MIMO radar waveform design with minimum peak sidelobe transmit beampattern," *IEEE Trans. Signal Process.*, vol. 66, no. 16, pp. 4207–4222, Aug. 2018.
[15] G. Cui, H. Li, and M. Rangaswamy, "MIMO radar waveform design with constant modulus and similarity constraints," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 343–353, Jan. 2014.
[16] S. Imani, M. M. Nayebi, and S. A. Ghorashi, "Transmit signal design in colocated MIMO radar without covariance matrix optimization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 5, pp. 2178–2186, Oct. 2017.
[17] B. Tang and J. Tang, "Joint design of transmit waveforms and receive filters for MIMO radar space-time adaptive processing," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4707–4722, Sep. 2016.
[18] T. Naghibi and F. Behnia, "MIMO radar waveform design in the presence of clutter," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 770–781, Apr. 2011.
[19] B. Tang, J. Tang, and Y. Peng, "Waveform optimization for MIMO radar in colored noise: Further results for estimation-oriented criteria," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1517–1522, Mar. 2012.
[20] B. Tang, M. M. Naghsh, and J. Tang, "Relative entropy-based waveform design for MIMO radar detection in the presence of clutter and interference," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3783–3796, Jul. 2015.
[21] M. M. Naghsh, E. H. M. Alian, M. M. Hashemi, and M. M. Nayebi, "Cognitive MIMO radars: An information theoretic constrained code design method," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Budapest, U.K., Aug. 2016, pp. 2215–2219.
[22] W. Huleihel, J. Tabrikian, and R. Shavit, "Optimal adaptive waveform design for cognitive MIMO radar," *IEEE Trans. Signal Process.*, vol. 61, no. 20, pp. 5075–5089, Oct. 2013.
[23] N. Sharaga, J. Tabrikian, and H. Messer, "Optimal cognitive beamforming for target tracking in MIMO radar/onar," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 8, pp. 1440–1450, Dec. 2015.
[24] F. Bandiera, O. Besson, and G. Ricci, "Knowledge-aided covariance matrix estimation and adaptive detection in compound-Gaussian noise," *IEEE Trans. Signal Process.*, vol. 58, no. 10, pp. 5390–5396, Oct. 2010.
[25] G. T. Capraro, A. Farina, H. Griffiths, and M. C. Wicks, "Knowledge-based radar signal and data processing: A tutorial review," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 18–29, Jan. 2006.
[26] S. Bidon, O. Besson, and J.-Y. Tourneret, "Knowledge-aided STAP in heterogeneous clutter using a hierarchical Bayesian algorithm," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 3, pp. 1863–1879, Jul. 2011.
[27] F. Bandiera, O. Besson, and G. Ricci, "Adaptive detection of distributed targets in compound-Gaussian noise without secondary data: A Bayesian approach," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 5698–5708, Dec. 2011.
[28] Y. I. Abramovich, O. Besson, and B. A. Johnson, "Conditional expected likelihood technique for compound Gaussian and Gaussian distributed noise mixtures," *IEEE Trans. Signal Process.*, vol. 64, no. 24, pp. 6640–6649, Dec. 2016.
[29] A. Turlapaty and Y. Jin, "Multi-parameter estimation in compound Gaussian clutter by variational Bayesian," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4663–4678, Sep. 2016.
[30] L. Hong, F. Dai, and X. Wang, "Knowledge-based wideband radar target detection in the heterogeneous environment," *Signal Process.*, vol. 144, pp. 169–179, Oct. 2018.
[31] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013.
[32] S. Boyd, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2010.
[33] Z. Wen, C. Yang, X. Liu, and S. Marchesini, "Alternating direction methods for classical and ptychographic phase retrieval," *Inverse Problems*, vol. 28, no. 11, Nov. 2012, Art. no. 115010.
[34] M. Hong, Z.-Q. Luo, and M. Razaviyayn, "Convergence analysis of alternating direction method of multipliers for a family of nonconvex problems," *SIAM J. Optim.*, vol. 26, no. 1, pp. 337–364, Jan. 2016.
[35] M. S. Kay, *Fundamentals of Statistical Signal Processing*, vol. 2. Upper Saddle River, NJ, USA: Prentice-Hall, 1993.
[36] A. Hjrungnes, *Complex-Valued Matrix Derivatives: With Applications in Signal Processing and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
[37] B. N. Datta, *Numerical Linear Algebra and Applications*. Philadelphia, PA, USA: SIAM, 2010.
[38] J. Nocedal and S. Wright, *Numerical optimization*. Springer, 2006.
[39] B. Himed, "MCARM/STAP data analysis," Air Force Res. Lab., New York, NY, USA, Tech. Rep. AFRL-SN-RSTR-1999, 1999, vol. 48.
[40] T. Yardibi, J. Li, P. Stoica, M. Xue, and A. B. Baggeroer, "Source localization and sensing: A nonparametric iterative adaptive approach based on weighted least squares," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 46, no. 1, pp. 425–443, Jan. 2010.



BO ZHANG received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2014, where he is currently pursuing the Ph.D. degree in signal processing with the National Laboratory of Radar Signal Processing. His research interests include MIMO radar waveform design and radar target detection.



FENGZHOU DAI received the B.S., M.S., and Ph.D. degrees in electronic engineering from Xidian University, Xi'an, China, in 2002, 2005, and 2010, respectively. He is currently an Associate Professor with the National Laboratory of Radar Signal Processing, Xidian University. His research interests include radar signal processing and microwave imaging.

NAN SU received the B.S. degree in electronic engineering from Xidian University, Xi'an, China, in 2012, where he is currently pursuing the Ph.D. degree in signal processing with the National Laboratory of Radar Signal Processing. His research interests include micromotion feature extraction and parameter estimation.