Cognitive Radar-Based Sequence Design via SINR Maximization

Linlong Wu, Prabhu Babu, and Daniel P. Palomar, *Fellow, IEEE*

***Abstract*—The ambiguity function plays an important role in radar systems. In fact, many radar design problems can be interpreted from the perspective of persuing desired ambiguity functions to adapt to various application scenes. In this paper, we consider designing a radar sequence, subject to a peak-to-average power ratio (PAR) constraint, to maximize the signal-to-interference plus noise ratio, which can also be interpreted as designing a sequence with a desired ambiguity function. From an optimization point of view, this is equivalent to optimizing a complex quartic function with the PAR constraint. An efficient algorithm based on the general majorization–minimization (MM) method is developed to solve this problem with guaranteed convergence to a stationary point under some mild conditions. In addition, the unit-modulus constraint, as a special case, is considered and another algorithm is proposed, which is the combination of the general MM and the coordinate descent method. Numerical experiments show that the proposed algorithms can shape a desired ambiguity function based on the prior information, and the performance is much better compared with the existing methods.**

***Index Terms*—Radar sequence design, SINR, ambiguity function, majorization-minimization, coordinate descent.**

I. Introduction

In RADAR signal processing, the range-Doppler response of a matched filter to a given finite energy signal with a delay τ and normalized Doppler frequency f, is referred to as the ambiguity function [1] and is defined as

$$\mathcal{R}(\tau,\ f) = \int_{-\infty}^{\infty} u(t)\, u^*(t-\tau)\, e^{j2\pi f t} dt, \tag{1}$$

where $u(t)$ is the radar signal emitted by the transmitter, and τ and f denote the range and the normalized Doppler shift, respectively.

If a baseband signal is modulated as a pulse-coded signal

$$u(t) = \sum_{n=1}^{N} s(n)\, p_n(t), \tag{2}$$

where $\{s(n)\}_{n=1}^{N}$ is the code sequence to be designed and $p_n(t)$ is the ideal rectangular function, the discrete ambiguity function becomes [2]

$$\mathcal{R}(k,\ p) = \sum_{n=1}^{N} s(n)\, s^*(n-k)\, e^{j2\pi \frac{(n-k)}{N} p}, \quad k = -N+1, \ldots,\ N-1,\ p = -\frac{N}{2}, \ldots, \frac{N}{2}. \tag{3}$$

If N is odd, then $p = -\frac{N-1}{2}, \ldots, \frac{N-1}{2}$.

The ambiguity function plays an important role in the design and study of radar systems. For a radar system, a good ambiguity function usually means a good ability to sense and detect. The ideal ambiguity function is thumbtack-like, with the peak corresponding to the range-Doppler bin of the target of interest. However, it is impossible to obtain such an ambiguity function due to the energy limit of the transmitted sequence and the constant volume property of the ambiguity function [1]. Besides this, which ambiguity function should be preferred also depends on the location of the clutters or interference of the surrounding environment. Therefore, in common cases, we are in such a dilemma because although we understand the significance of the ambiguity function, we do not know what its desired shape is, except for the lower sidelobes of the autocorrelation. However, a cognitive approach to radar systems is discussed in [3], which is accessible given the development in signal processing, antennas and computers today and is becoming the leading approach for radar signal processing. Such cognitive radar systems capitalize on the information obtained from the surrounding environment or the prior knowledge stored in the platform, and they tell us the desired shape of the ambiguity function to some extent, at least from the perspective of reducing the interference. To be more specific, the response at some range-Doppler bins corresponding to the known or predicted scatters reduces to be as small as possible, while the response at the corresponding range-Doppler bin of the target of interest is maintained at a relatively high level.

The significance of the ambiguity function and the innovative scheme of cognitive radar systems has motivated active research in developing methods to design sequences with a desired ambiguity function. Some papers focus on communications applications and consider the shaping of the autocorrelation function only (ignoring the Doppler effect). Among them, [4] proposed an algorithm called CAN to design a unimodular sequence with low autocorrelation sidelobes by minimizing an "almost equivalent" problem instead of the original integrated sidelobe level (ISL) one, and then incorporated the spectral constraint in [5]. Later, [6] considered the ISL minimization problem directly, and then [7] extended the solution to the weighted case and

Manuscript received November 24, 2015; revised July 5, 2016 and October 13, 2016; accepted October 18, 2016. Date of publication October 26, 2016; date of current version November 28, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Marco Lops. This work was supported in part by the Hong Kong RGC 16206315 research Grant and in part by the Hong Kong RGC Theme-based Research Scheme under Grant T21-602/15R.

L. Wu and D. P. Palomar are with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: lwuag@ust.hk; palomar@ust.hk).

P. Babu was with the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong, and now with the Centre for Applied Research in Electronics, Indian Institute of Technology Delhi, New Delhi 110016, India (e-mail: prabhubabu@ust.hk).

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/TSP.2016.2621723

the ℓ_p-norm case. From the perspective of the ambiguity function, this kind of problems correspond to the design of a sequence with a desired zero-Doppler frequency ambiguity cut. Further, [8] considered the sequence design for multiple-input-multiple-output (MIMO) radar and shaped the autocorrelation and the crosscorrelation simultaneously. Besides suppressing auto/cross-correlation sidelobe levels, [9] also incorporated the match of the desired beam pattern.

For radar systems, however, the Doppler effect cannot be neglected and the complete ambiguity function needs to be considered, rather than just the autocorrelation. In [10], [11] and [12], the authors considered the radar code design to achieve the best detection performance in the presence of colored Gaussian disturbance, and then formulated a nonconvex quadratic problem, which was solved through semidefinite relaxation (SDR) and then rank-one randomization. Note that the authors introduced a similarity constraint, which was tantamount to imposing some restrictions on the ambiguity function indirectly. In [13] and [14], the authors extended the above work to the robust case with an unknown Doppler frequency of the target. However, the authors treated the disturbance as a whole and did not meticulously consider the characteristics of interference, which in fact could be obtained to some degree according to the cognitive approach we mentioned above. In [15], the authors classified the disturbance into signal-dependent interference and white noise, and under the assumption of the matched filter, the authors arrived at a quartic problem, which was solved based on the maximum-block-improvement (MBI) method. Readers interested in the MBI method may refer to [16] and the references therein. However, we point out that the variant MBI approach deployed in [15] is highly-computational and time-comsuming, which is undesired for real-time applications like radar.

Most of the previous papers consider the design of unimodular sequences. The peak-to-average power ratio (PAR) constraint is a relaxed constraint in the practical sense, although mathematically a more general constraint as the unimodular case is just a particular case. In this paper, with the PAR and finite energy constraints, we focus on designing radar sequences to maximize the SINR based on cognitive radar systems. After some approximations, this problem can also be interpreted as designing sequences with desired ambiguity functions to adapt to specific application scenarios. We develop an efficient algorithm, MIAFIS, based on the general majorization-minimization (MM) method, with a guaranteed convergence to a stationary point under some mild conditions. We also propose some acceleration schemes with faster convergence. Note that the PAR constraint is more general than the unit-modulus one, and MIAFIS can still be deployed for the unimodular case. For the unimodular case, we also propose a new algorithm, CIAFIS, which is the combination of the MM method and coordinate descent method.

The rest of this paper is organized as follows. In Section II, we introduce the system model and formulate the sequence design problem of interest. In Section III, we first introduce the general MM method briefly and then derive the MAIFIS algorithm within the MM framework, followed by convergence analysis and acceleration schemes. In Section IV, we consider a special case of the problem in which the constraints are reduced to the unit-modulus one, and the new CIAFIS algorithm is proposed. In Section V, we compare MIAFIS with CIAFIS and reveal the potential connections between MIAFIS and the traditional gradient projection method. In Section VI, we analyze the performance of the proposed algorithms and compare them with the existing methods. Finally, the conclusions are given in Section VII.

Notation: $\mathbb{R}^n$ and $\mathbb{C}^n$ denote the n-dimensional real and complex vector space, respectively. $\mathbb{R}^{n\times n}$ and $\mathbb{C}^{n\times n}$ denote the $n\times n$ real and complex matrix space, respectively. Boldface upper case letters stand for matrices. Boldface lower case letters stand for column vectors. Standard lower case letters stand for scalars. $\mathrm{Re}(\mathbf{x})$ and $\arg(\mathbf{x})$ denotes the element-wise real part and the phase of a complex vector $\mathbf{x}$, respectively. $(\mathbf{x})^T$, $(\mathbf{x})^*$, $(\mathbf{x})^H$, $\mathrm{tr}(\mathbf{x})$, $\mathrm{vec}(\mathbf{x})$, $\lambda_{\max}(\mathbf{x})$ and $\lambda_{\min}(\mathbf{x})$ denote the transpose, complex conjugate, conjugate transpose, trace, vectorization, largest eigenvalue, and smallest eigenvalue of a matrix $\mathbf{X}$, respectively. $\mathrm{Diag}(\mathbf{x})$ stands for a diagonal matrix with its principal diagonal filled with $\mathbf{x}$. $\mathbf{I}$ denotes the identity matrix. x_i denotes the i-th element of $\mathbf{x}$. $|\cdot|$ denotes the modulus of a complex scalar. $\|\cdot\|$ denotes the ℓ_2-norm of a vector. $\odot$ denotes the Hadamard product. $\nabla(\cdot)$ denotes the gradient of a vector function, and $\mathbb{E}(\cdot)$ denotes the statistic expectation.

II. System Model & Problem Formulation

Consider a monostatic radar system transmitting a coherent burst of coded pulses, with the $N-$dimensional vector of observations modelled as [15]:

$$\mathbf{v} = \alpha\mathbf{s}\odot\mathbf{p}\left(\nu_d\right) + \mathbf{d}\left(\mathbf{s}\right) + \mathbf{n}, \tag{4}$$

where α is a complex parameter accounting for channel propagation and backscattering effects, $\mathbf{s}$ is the vector of coded elements, $\mathbf{p}(\nu_d) = [1, e^{j2\pi\nu_d}, \ldots, e^{j2\pi(N-1)\nu_d}]^T$ is the temporal steering vector, ν_d is the normalized Doppler frequency of the target of interest, $\mathbf{d}(\mathbf{s})$ is the vector of interfering samples, and $\mathbf{n}$ is the vector of the noise samples following the normal distribution $N(\mathbf{0},\, \sigma_n^2\mathbf{I})$ uncorrelated with $\mathbf{d}(\mathbf{s})$.

Note that the interfering vector $\mathbf{d}(\mathbf{s})$ accounts for the clutter returns, which can be expressed as [15]:

$$\mathbf{d}\left(\mathbf{s}\right) = \sum_{i=1}^{N_t}\rho_i\mathbf{J}^{r_i}\left(\mathbf{s}\odot\mathbf{p}\left(\nu_i\right)\right), \tag{5}$$

where N_t is the total number of interfering scatterers, $r_k \in \{0, 1, \ldots N-1\}$, ρ_i and ν_i are, respectively, the range position, the echo complex amplitude, and the normalized Doppler frequency of the i-th scatterer, and $\mathbf{J}^{r_i}$, $r_i \in \{-N+1, \ldots, 0, \ldots, N-1\}$ is the $N\times N$ shift matrix given by

$$\mathbf{J}^{r_i}\left(m,\, n\right) = \begin{cases} 1, & m-n = r_i \\ 0, & m-n \neq r_i. \end{cases} \tag{6}$$

In fact, once the target is defined to be threatening, a track file in the search-and-track modality is opened and continuously updated [15]. This track file usually contains several information, including Doppler velocity measurements [17]. For the details of how the Doppler shift is measured, please refer to Chapter 17

in [18] for details. Thus, we reasonably assume that the Doppler frequency of the target of interest ν_d is known. The output of the matched filter to the echo is given by

$$(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{v} = \alpha\|\mathbf{s}\|^2 + (\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{d}(\mathbf{s}) + (\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{n}, \quad (7)$$

where the last two terms are the disturbance to the target detection.

Consequently, the disturbance power after matched filtering is

$$\begin{aligned}&\mathbb{E}\left[\left|(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{d}(\mathbf{s}) + (\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{n}\right|^2\right]\\ &= \mathbb{E}\left[\left|(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{d}(\mathbf{s})\right|^2\right] + \mathbb{E}\left[\left|(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbf{n}\right|^2\right]\\ &= (\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbb{E}\left[\mathbf{d}(\mathbf{s})\mathbf{d}(\mathbf{s})^H\right](\mathbf{s}\odot\mathbf{p}(\nu_d)) + \sigma_n^2\|\mathbf{s}\|^2.\end{aligned} \quad (8)$$

Thus, the signal-to-interference plus noise ratio (SINR) is

$$\text{SINR} = \frac{|\alpha|^2\|\mathbf{s}\|^4}{(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbb{E}\left[\mathbf{d}(\mathbf{s})\mathbf{d}(\mathbf{s})^H\right](\mathbf{s}\odot\mathbf{p}(\nu_d)) + \sigma_n^2\|\mathbf{s}\|^2}. \quad (9)$$

The problem we will solve is formulated as

$$\begin{aligned}\underset{\mathbf{s}}{\text{maximize}}\quad & \text{SINR}\\ \text{subject to}\quad & \text{PAR}(\mathbf{s}) \le \gamma\\ & \|\mathbf{s}\|^2 = N,\end{aligned} \quad (10)$$

where $\|\mathbf{s}\|^2 = N$ denotes the energy constraint and

$$\text{PAR}(\mathbf{s}) = \frac{\max_{n=1,\ldots,N}\left\{|s_n|^2\right\}}{\sum_{n=1}^N|s_n|^2/N}, \quad (11)$$

and γ is the parameter controlling the acceptable level of PAR with $1 \le \gamma \le N$.

Problem (10) is equivalent to

$$\begin{aligned}\underset{\mathbf{s}}{\text{minimize}}\quad & (\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbb{E}\left[\mathbf{d}(\mathbf{s})\mathbf{d}(\mathbf{s})^H\right](\mathbf{s}\odot\mathbf{p}(\nu_d))\\ \text{subject to}\quad & \text{PAR}(\mathbf{s}) \le \gamma\\ & \|\mathbf{s}\|^2 = N.\end{aligned} \quad (12)$$

In [15], the normalized Doppler frequency ν_i of the i-th clutter is modelled as a uniformly distributed random variable. After discretizing the normalized Doppler interval $[-\frac{1}{2},\frac{1}{2}]$ into N_ν bins and approximating the expectation with the sample mean, the objective of problem (12) can be approximately expressed as

$$\begin{aligned}&(\mathbf{s}\odot\mathbf{p}(\nu_d))^H\mathbb{E}\left[\mathbf{d}(\mathbf{s})\mathbf{d}(\mathbf{s})^H\right](\mathbf{s}\odot\mathbf{p}(\nu_d))\\ &\approx \sum_{r=0}^{N-1}\sum_{h=0}^{N_\nu-1}p(r,h)\left|\mathbf{s}^H\mathbf{J}^r\text{Diag}(\mathbf{p}(\nu_h))\mathbf{s}\right|^2,\end{aligned} \quad (13)$$

where $\nu_h = -\frac{1}{2}+\frac{h}{N_v}$, $h = 0,1,\ldots,N_v$, is the discrete normalized Doppler frequency and the target Doppler frequency is set as $\nu_h = 0$ without loss of generality; $p(r,h)$ is the interference power for the range-Doppler bin (r,v_h).

For the objective function of problem (12), there always exists a one-to-one mapping $k \in \{1,2,\ldots,NN_v\} \to (r,\ h) \in \{0,1,\ldots,N-1\}\times\{0,1,\ldots,N_v-1\}$. In the rest of the paper, k is used to represent the corresponding $(r,\ h)$ unless otherwise specified. Let $\mathbf{A}_k = \mathbf{J}^r\text{Diag}(\mathbf{p}(\nu_h))$. Then problem (12) can be written as

$$\begin{aligned}\underset{\mathbf{s}}{\text{minimize}}\quad & \sum_{k=1}^{NN_\nu}p_k\left|\mathbf{s}^H\mathbf{A}_k\mathbf{s}\right|^2\\ \text{subject to}\quad & \text{PAR}(\mathbf{s}) \le \gamma\\ & \|\mathbf{s}\|^2 = N,\end{aligned} \quad (14)$$

which we will solve in the following.

Before proceeding with the design of the solution to problem (14), we address something of our problem formulation:

- $|\mathbf{s}^H\mathbf{A}_k\mathbf{s}|$ and $\{p(r,h)\}$ are the modulus of the ambiguity function of $\mathbf{s}$ and the clutter information at the range-Doppler bin (r,h), respectively. Problem (14) can be interpreted as follows: After perseiving the enviroment by cognitive approaches, the clutter information incorporates into p_ks. By multipling $p(r,h)$ with the square modulus of the ambiguity function at the corresponding range-Doppler bin (r,h) and then minimizing the sum, the designed sequence will has a desired ambiguity function with low values in the range-Doppler bins where $p(r,h)$s have high values.
- Our problem becomes the common ISL problem [4], [6] if we let $\nu_h = 0$ for all h and $p_k = 1$ for all k, which means all the scatters have the same Doppler frequency or we only focus on a specific Doppler frequency of interest. Besides this, if we let $\nu_h = 0$ and choose different values of p_k's, the problem becomes a weighted ISL problem [5], [7].
- Note that γ is in the range $[1,N]$. When $\gamma = 1$, the PAR constraint, together with the constant energy constraint, becomes the unit-modulus constraint, which is widely considered in the literature. By increasing the value of γ, we are relaxing the problem and the value of the optimal objective should be nonincreasing (probably decrease).
- The problem is hard to solve, due to the quartic objective and the nonconvex constraints.

III. Sequence Design via Majorization-Minimization

In this section, we first introduce the general majorization-minimization (MM) and then derive a simple algorithm based on it. Last, we will analyze the convergence of the proposed algorithm and provide two acceleration schemes.

A. Majorization-Minimization Method

The MM method is a powerful optimization scheme, especially when the problem is hard to tackle directly [19]. The idea behind the MM algorithm is to convert the original problem into a sequence of simpler problems to be solved until convergence. The key to using the MM method is to construct a simple majorized problem that can be solved efficiently [20].

Consider a general optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{x} \in \mathcal{X}. \end{aligned} \tag{15}$$

Suppose the problem is hard to directly minimize. Following the general MM idea, we first find $u(\mathbf{x},\ \mathbf{x}^{(\ell)})$, the surrogate function of $f(\mathbf{x})$, which should satisfy the following two requirements at the point $\mathbf{x}^{(\ell)}$:

$$u\big(\mathbf{x},\, \mathbf{x}^{(\ell)}\big) \geq\ f(\mathbf{x}),\ \text{for all}\ \mathbf{x} \in \mathcal{X} \tag{16}$$

$$u\big(\mathbf{x}^{(\ell)},\, \mathbf{x}^{(\ell)}\big) =\ f\big(\mathbf{x}^{(\ell)}\big). \tag{17}$$

Then the MM update is given by

$$\mathbf{x}^{(\ell+1)} = \underset{\mathbf{x}\in\mathcal{X}}{\operatorname{argmin}}\ u\big(\mathbf{x},\, \mathbf{x}^{(\ell)}\big). \tag{18}$$

One interesting and useful property of MM methods is monotonicity:

$$\begin{aligned} f\big(\mathbf{x}^{(\ell+1)}\big) \leq\ & u\big(\mathbf{x}^{(\ell+1)},\, \mathbf{x}^{(\ell)}\big) \leq u\big(\mathbf{x}^{(\ell)},\, \mathbf{x}^{(\ell)}\big) \\ =\ & f\big(\mathbf{x}^{(\ell)}\big), \end{aligned} \tag{19}$$

where the first inequality follows from (16), the second one follows from (18) and the last equality follows from (17).

If the objective function is already convex, then the MM algorithm will converge to the global optimal point, assuming it exists. If convexity or concavity fails, at least one subsequence the MM algorithm generates will converge to one of the stationary points under some conditions. For more details about the convergence, interested readers may refer to [21].

Two things need to be noted here. First, the convergence speed of the MM algorithm is mainly determined by the surrogate function, i.e., how closely it resembles the original function. In some cases, if the surrogate function is ill-constructed, some acceleration techniques have to be adopted. Second, from (19), we see that even when $\mathbf{x}^{(\ell+1)}$ is not the minimizer of $u(\mathbf{x},\, \mathbf{x}^{(\ell)})$, the monotonicity can still be guaranteed as long as it improves the function $u(\mathbf{x}^{(\ell+1)},\, \mathbf{x}^{(\ell)}) \leq u(\mathbf{x}^{(\ell)},\, \mathbf{x}^{(\ell)})$, where the equality means the algorithm has already found a stationary point $\mathbf{x}^{(\ell+1)}$.

B. Majorized Iteration for Ambiguity Function Iterative Shaping

The objective function of problem (14) can be equivalently reformulated as

$$\begin{aligned} \sum_{k=1}^{NN_\nu} p_k \left|\mathbf{s}^H \mathbf{A}_k \mathbf{s}\right|^2 =\ & \sum_{k=1}^{NN_\nu} p_k \left|\mathrm{tr}\left(\mathbf{A}_k \mathbf{S}\right)\right|^2 \\ =\ & \sum_{k=1}^{NN_\nu} p_k \left|\mathrm{vec}\left(\mathbf{S}\right)^H \mathrm{vec}\big(\mathbf{A}_k\big)\right|^2 \\ =\ & \sum_{k=1}^{NN_\nu} p_k \mathrm{vec}\left(\mathbf{S}\right)^H \mathbf{B} \mathrm{vec}\left(\mathbf{S}\right), \end{aligned} \tag{20}$$

where $\mathbf{B} = \sum_{k=1}^{NN_\nu} p_k \mathrm{vec}(\mathbf{A}_k)\mathrm{vec}(\mathbf{A}_k)^H$.

Note that $\mathbf{0} \preceq \mathbf{B} \preceq \mathrm{tr}\left(\mathbf{B}\right)\mathbf{I}$ with $\mathrm{tr}\left(\mathbf{B}\right)$ given by

$$\begin{aligned} \mathrm{tr}\left(\mathbf{B}\right) =\ & \mathrm{tr}\left(\sum_{k=1}^{NN_\nu} p_k \mathrm{vec}\big(\mathbf{A}_k\big)\mathrm{vec}\big(\mathbf{A}_k\big)^H\right) \\ =\ & \sum_{k=1}^{NN_\nu} p_k \mathrm{vec}\big(\mathbf{A}_k\big)^H \mathrm{vec}\big(\mathbf{A}_k\big) \\ =\ & \sum_{k=1}^{NN_\nu} p_k \mathrm{tr}\left(\mathbf{A}_k^H \mathbf{A}_k\right) \\ =\ & \sum_{k=1}^{NN_\nu} p_k \left(N - r\right). \end{aligned} \tag{21}$$

Since $\mathrm{vec}\left(\mathbf{S}\right)^H \mathrm{vec}\left(\mathbf{S}\right) = \mathrm{tr}\left(\mathbf{S}\mathbf{S}\right) = \mathrm{tr}\left(\mathbf{s}\mathbf{s}^H \mathbf{s}\mathbf{s}^H\right) = N^2$ is a constant, we further have the following equivalent problem:

$$\begin{aligned} \underset{\mathbf{S},\mathbf{s}}{\text{minimize}} \quad & \mathrm{vec}\left(\mathbf{S}\right)^H \left(\mathbf{B} - \lambda_u\left(\mathbf{B}\right)\mathbf{I}\right)\mathrm{vec}\left(\mathbf{S}\right) \\ \text{subject to} \quad & |s_n| \leq \sqrt{\gamma},\ n = 1, 2, \ldots, N \\ & \|\mathbf{s}\|^2 = N \\ & \mathbf{S} = \mathbf{s}\mathbf{s}^H, \end{aligned} \tag{22}$$

where $\lambda_u\left(\mathbf{B}\right) = \sum_{k=1}^{NN_\nu} p_k \left(N - r\right)$ is an upper bound of the eigenvalues of $\mathbf{B}$.

Note that the objective function of problem (22) is concave now. We can construct the surrogate function of the objective function of (22) by first-order approximation. Given $\mathbf{S}^{(\ell)} = \mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H$ at the ℓ-th iteration, the first-order approximation is

$$\begin{aligned} u_1\big(\mathbf{S},\ \mathbf{S}^{(\ell)}\big) =\ & 2\mathrm{Re}\left(\mathrm{vec}\left(\mathbf{S}\right)^H \left(\mathbf{B} - \lambda_u\left(\mathbf{B}\right)\mathbf{I}\right)\mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)\right) \\ & + \mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)^H \left(\lambda_u\left(\mathbf{B}\right)\mathbf{I} - \mathbf{B}\right)\mathrm{vec}\big(\mathbf{S}^{(\ell)}\big). \end{aligned} \tag{23}$$

Ignoring the constant terms of (23), the majorized problem of (22) at the point $\mathbf{s}^{(\ell)}$ is given by

$$\begin{aligned} \underset{\mathbf{S},\mathbf{s}}{\text{minimize}} \quad & \mathrm{Re}\big(\mathrm{vec}\left(\mathbf{S}\right)^H \left(\mathbf{B} - \lambda_u\left(\mathbf{B}\right)\mathbf{I}\right)\mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)\big) \\ \text{subject to} \quad & |s_n| \leq \sqrt{\gamma},\ n = 1, 2, \ldots, N \\ & \|\mathbf{s}\|^2 = N \\ & \mathbf{S} = \mathbf{s}\mathbf{s}^H. \end{aligned} \tag{24}$$

We can now undo the change of variable $\mathbf{S} = \mathbf{s}\mathbf{s}^H$ in the objective function of (24):

$$\begin{aligned} & \mathrm{Re}\big(\mathrm{vec}\left(\mathbf{S}\right)^H \left(\mathbf{B} - \lambda_u\left(\mathbf{B}\right)\mathbf{I}\right)\mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)\big) \\ & = \mathrm{Re}\left(\mathrm{vec}\left(\mathbf{S}\right)^H \left(\sum_{k=1}^{NN_v} p_k \mathrm{vec}\big(\mathbf{A}_k\big)\mathrm{vec}\big(\mathbf{A}_k\big)^H\right)\mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)\right) \\ & \quad - \mathrm{Re}\big(\lambda_u\left(\mathbf{B}\right)\mathrm{vec}\left(\mathbf{S}\right)^H \mathrm{vec}\big(\mathbf{S}^{(\ell)}\big)\big) \\ & = \mathrm{Re}\left(\mathrm{tr}\left(\left(\sum_{k=1}^{NN_v} p_k \mathrm{tr}(\mathbf{A}_k^H \mathbf{S}^{(\ell)})\mathbf{A}_k - \lambda_u\left(\mathbf{B}\right)\mathbf{S}^{(\ell)}\right)\mathbf{S}\right)\right) \\ & = \mathrm{Re}\Bigg(\mathbf{s}^H \Bigg(\sum_{k=1}^{NN_v} p_k \big(\mathbf{s}^{(\ell)}\big)^H \mathbf{A}_k^H \mathbf{s}^{(\ell)} \mathbf{A}_k \\ & \quad - \lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}\big(\mathbf{s}^{(\ell)}\big)^H\Bigg)\mathbf{s}\Bigg), \end{aligned} \tag{25}$$

and then problem (24) becomes

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \text{Re}\left(\mathbf{s}^H\left(\mathbf{R}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\right)\mathbf{s}\right) \\ \text{subject to} \quad & |s_n|\le\sqrt{\gamma},\ n=1,2,\ldots,N \\ & \|\mathbf{s}\|^2=N, \end{aligned} \tag{26}$$

where

$$\mathbf{R}=\sum_{k=1}^{NN_v} p_k\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k^H\mathbf{s}^{(\ell)}\mathbf{A}_k. \tag{27}$$

By defining $\mathbf{P}=\frac{1}{2}(\mathbf{R}+\mathbf{R}^H)$, we have $\text{Re}(\mathbf{s}^H\mathbf{R}\mathbf{s})=\frac{1}{2}(\mathbf{s}^H\mathbf{R}\mathbf{s}+\mathbf{s}^H\mathbf{R}^H\mathbf{s})=\mathbf{s}^H\mathbf{P}\mathbf{s}$. Then problem (26) can be rewritten as

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \mathbf{s}^H\left(\mathbf{P}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\right)\mathbf{s} \\ \text{subject to} \quad & |s_n|\le\sqrt{\gamma},\ n=1,2,\ldots,N \\ & \|\mathbf{s}\|^2=N. \end{aligned} \tag{28}$$

Now the objective function of problem (28) is quadratic in $\mathbf{s}$, but it is still hard to solve directly because the matrix $\mathbf{P}$ may be indefinite. Thus we propose to majorize the objective function of problem (28) at $\mathbf{s}^{(\ell)}$ again to further simplify the problem we need to solve at every iteration. Similarly to the construction of the first surrogate function $u_1(\mathbf{S},\ \mathbf{S}^{(\ell)})$, we need to find an upper bound of the matrix $(\mathbf{P}-\lambda_u(\mathbf{B})\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H)$.

Before we find the upper bound, let us introduce a useful theorem regarding the bounds of extreme eigenvalues of a Hermitian matrix.

Lemma 1: [22] Let $\mathbf{M}$ be an $n\times n$ complex matrix with real eigenvalues $\lambda(\mathbf{M})$, and let

$$m=\frac{\text{tr}(\mathbf{M})}{n},\ s^2=\frac{\text{tr}(\mathbf{M}^2)}{n-m^2}. \tag{29}$$

Then

$$m-s\sqrt{n-1}\le\ \lambda_{\min}(\mathbf{M})\le m-\frac{s}{\sqrt{n-1}}, \tag{30}$$

$$m+\frac{s}{\sqrt{n-1}}\le\ \lambda_{\max}(\mathbf{M})\le m+s\sqrt{n-1}. \tag{31}$$

We define

$$\mathbf{P}_k=\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k^H\mathbf{s}^{(\ell)}\mathbf{A}_k+\mathbf{A}_k^H\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}, \tag{32}$$

and then $\mathbf{P}=\sum_{k=1}^{NN_v}\frac{p_k}{2}\mathbf{P}_k$. Each $\mathbf{P}_k,\ k=1,2,\ldots,NN_v$, is Hermitian and thus has real eigenvalues. By using Lemma 1 and considering the special structure of $\mathbf{P}_k$, we have the following results.

Lemma 2: Let $\mathbf{P}_k$ be the matrix defined in (32). Then

$$\lambda_{\max}(\mathbf{P}_k)\le\begin{cases}\sqrt{\frac{2(N-r)(N-1)}{N}}\left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|, & \text{for } r\neq 0\\ 2N, & \text{for } r=0,\end{cases}$$

where r represents the range and $r=\lfloor\frac{k}{N_v}\rfloor$.

Proof: See Appendix A. ■

Since $\lambda_{\max}(\mathbf{P})\le\sum_{k=1}^{NN_v}\frac{p_k}{2}\lambda_{\max}(\mathbf{P}_k)$, the upper bound of matrix $\mathbf{P}$ can be expressed as

$$\begin{aligned}\lambda_u(\mathbf{P})=&\sum_{k=N_v+1}^{NN_v}p_k\sqrt{\frac{(N-r)(N-1)}{2N}}\left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|\\&+\sum_{k=1}^{N_v}p_kN,\end{aligned} \tag{33}$$

which is also an upper bound of the eigenvalues of matrix $(\mathbf{P}-\lambda_u(\mathbf{B})\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H)$. Thus, problem (28) is equivalent to

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \mathbf{s}^H\left(\mathbf{P}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H-\lambda_u\left(\mathbf{P}\right)\mathbf{I}\right)\mathbf{s} \\ \text{subject to} \quad & |s_n|\le\sqrt{\gamma},\ n=1,2,\ldots,N \\ & \|\mathbf{s}\|^2=N. \end{aligned} \tag{34}$$

The objective function of (34) can also be majorized by the first-order approximation

$$\begin{aligned} u_2\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)=&\ 2\text{Re}\Big(\mathbf{s}^H\left(\mathbf{P}-\lambda_u(\mathbf{B})\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H-\lambda_u(\mathbf{P})\mathbf{I}\right)\mathbf{s}^{(\ell)}\Big)\\&+\left(\mathbf{s}^{(\ell)}\right)^H\Big(\lambda_u(\mathbf{P})\mathbf{I}-\mathbf{P}+\lambda_u(\mathbf{B})\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\Big)\mathbf{s}^{(\ell)}\\=&\ 2\text{Re}\left(\mathbf{s}^H\left(\mathbf{P}-\left(\lambda_u(\mathbf{B})N+\lambda_u(\mathbf{P})\right)\mathbf{I}\right)\mathbf{s}^{(\ell)}\right)\\&+\left(\mathbf{s}^{(\ell)}\right)^H\left(\lambda_u(\mathbf{P})\mathbf{I}-\mathbf{P}+\lambda_u(\mathbf{B})\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\right)\mathbf{s}^{(\ell)}.\end{aligned} \tag{35}$$

Ignoring the constant terms and the scalar of (35), the majorized problem of (28) is

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \text{Re}\left(\mathbf{s}^H\mathbf{z}\right) \\ \text{subject to} \quad & |s_n|\le\sqrt{\gamma},\ n=1,2,\ldots,N \\ & \|\mathbf{s}\|^2=N, \end{aligned} \tag{36}$$

where

$$\mathbf{z}=\left(\mathbf{P}-\left(\lambda_u\left(\mathbf{B}\right)N+\lambda_u\left(\mathbf{P}\right)\right)\mathbf{I}\right)\mathbf{s}^{(\ell)}. \tag{37}$$

The following lemma gives an optimal solution of problem (36).

Lemma 3: An optimal solution to (36) is given by

$$\mathbf{s}=\mathcal{P}_{\mathcal{S}}\left(\mathbf{z}\right), \tag{38}$$

where

$$\begin{aligned}\mathcal{P}_{\mathcal{S}}(\cdot)=&-\left(\mathbf{1}_{\mathbb{R}_0^+}\left(N-m\gamma\right)\right)\sqrt{\gamma}\mathbf{u}_m\odot e^{j\arg(\cdot)}\\&-\left(\mathbf{1}_{\mathbb{R}^-}\left(N-m\gamma\right)\right)\min\{\beta|\mathbf{z}|,\sqrt{\gamma}\mathbf{1}\}\odot e^{j\arg(\cdot)},\end{aligned} \tag{39}$$

$\min\{\cdot,\cdot\}$, $|\cdot|$ and $e^{j\arg(\cdot)}$ are element-wise operations ,

$$\mathbf{1}_A\left(x\right)=\begin{cases}1, & \text{if } x\in A,\\ 0, & \text{otherwise,}\end{cases} \tag{40}$$

$$\mathbf{u}_m=\left[\underbrace{1,\ldots 1}_{m},\ \underbrace{\sqrt{\frac{N-m\gamma}{N\gamma-m\gamma}},\ldots,\ \sqrt{\frac{N-m\gamma}{N\gamma-m\gamma}}}_{N-m}\right]^T, \tag{41}$$

Algorithm 1: MIAFIS - Majorized Iteration for Ambiguity Function Iterative Shaping.

Input: Initial sequence $\mathbf{s}^{(0)}$
Output: Designed sequence $\mathbf{s}$
$\lambda_u(\mathbf{B}) = \sum_{k=1}^{NN_\nu} p_k(N-r)$
repeat
  $\mathbf{P} = \sum_{k=1}^{NN_\nu} \frac{p_k}{2}(tr(\mathbf{A}_k^H \mathbf{S}^{(\ell)})\mathbf{A}_k + tr(\mathbf{A}_k \mathbf{S}^{(\ell)})\mathbf{A}_k^H)$
  Calculate $\lambda_u(\mathbf{P})$ according to (32)
  $\mathbf{z} = (\mathbf{P} - (\lambda_u(\mathbf{B})N + \lambda_u(\mathbf{P}))\mathbf{I})\mathbf{s}^{(\ell)}$
  $\mathbf{s}^{(\ell+1)} = \mathcal{P}_{\mathcal{S}}(\mathbf{z})$
  $\ell \leftarrow \ell + 1$
until convergence

and

$$\beta \in \left\{ \beta | \sum_{n=1}^{N} \min\left\{ \beta^2 |z_n|^2, \gamma \right\} = N, \beta \in \left[0, \frac{\sqrt{\gamma}}{\min\{|z_n| \,|\, |z_n| \neq 0\}}\right] \right\}. \quad (42)$$

Proof: See Appendix B ■

Note that even though we derive the objective function of (36) through two majorization steps, we can merge the two steps into one and obtain the final surrogate function of the objective function of (14) given by

$$\begin{aligned} u(\mathbf{s}, \mathbf{s}^{(\ell)}) &= 2u_2(\mathbf{s}, \mathbf{s}^{(\ell)}) + 2\lambda_u(\mathbf{P})N + 2\lambda_u(\mathbf{L})N^2 \\ &\quad - \mathrm{vec}(\mathbf{S}^{(\ell)})^H \mathbf{L}\mathrm{vec}(\mathbf{S}^{(\ell)}) \\ &= 4\mathrm{Re}\left(\mathbf{s}^H \mathbf{P}\mathbf{s}^{(\ell)} - (\lambda_u(\mathbf{P}) + \lambda_u(\mathbf{L})N)\mathbf{s}^H \mathbf{s}^{(\ell)}\right) \\ &\quad - \left(2(\mathbf{s}^{(\ell)})^H \mathbf{P}\mathbf{s}^{(\ell)} + \mathrm{vec}(\mathbf{S}^{(\ell)})^H \mathbf{L}\mathrm{vec}(\mathbf{S}^{(\ell)})\right) \\ &\quad + 4N(\lambda_u(\mathbf{P}) + \lambda_u(\mathbf{L})N) \\ &= 4\mathrm{Re}\left(\mathbf{s}^H (\mathbf{P} - (\lambda_u(\mathbf{B})N + \lambda_u(\mathbf{P}))\mathbf{I})\mathbf{s}^{(\ell)}\right) \\ &\quad + \text{constant}. \end{aligned} \quad (43)$$

This surrogate function will enlighten us to an acceleration scheme later.

Now we have finished the derivation of the algorithm for (14), and the complete description of the overall algorithm is given in Algorithm 1. Note that this algorithm we derived is within the general MM framework. Thus the monotonicity can still be guaranteed. As for the convergence, we will prove it in the next subsection.

C. Convergence Analysis

The objective function of problem (14) is bounded by 0, and according to the inequality (19), we have $f(\mathbf{s}^{(0)}) \leq f(\mathbf{s}^{(1)}) \leq \ldots \leq f(\mathbf{s}^{(\ell)})$. Thus, the sequence $\{f(\mathbf{s}^{(\ell)})\}$ generated by MIAFIS is guaranteed to converge to a finite value. In the following, we will analyze the convergence property of the sequence $\{\mathbf{s}^{(\ell)}\}$ generated by MIAFIS.

Algorithm 2: MIAFIS Acceleration via SQUAREM.

Input: Initial sequence $\mathbf{s}^{(0)}$
Output: Designed sequence $\mathbf{s}$
repeat
  $\mathbf{s}_1 = \mathbf{F}_{MM}(\mathbf{s}^{(\ell)})$
  $\mathbf{s}_2 = \mathbf{F}_{MM}(\mathbf{s}_1)$
  $\mathbf{q} = \mathbf{s}_1 - \mathbf{s}^{(\ell)}$
  $\mathbf{v} = \mathbf{s}_2 - \mathbf{s}_1 - \mathbf{q}$
  $\alpha = -\frac{||\mathbf{q}||}{||\mathbf{v}||}$
  $\mathbf{s}^{(\ell+1)} = -\mathcal{P}_{\mathcal{S}}(\mathbf{s}^{(\ell)} - 2\alpha\mathbf{q} + \alpha^2\mathbf{v})$
  while $f(\mathbf{s}^{(\ell+1)}) > f(\mathbf{s}^{(\ell)})$ **do**
    $\alpha = (\alpha - 1)/2$
    $\mathbf{s}^{(\ell+1)} = -\mathcal{P}_{\mathcal{S}}(\mathbf{s}^{(\ell)} - 2\alpha\mathbf{q} + \alpha^2\mathbf{v})$
  end while
  $\ell \leftarrow \ell + 1$
until convergence

Lemma 4: Let $\{\mathbf{s}^{(\ell)}\}$ be the sequence generated by MIAFIS. Then every limit point of $\{\mathbf{s}^{(\ell)}\}$ is a stationary point of problem (14).

Proof: See Appendix C. ■

D. Acceleration Schemes

For the MM algorithm, the convergence speed is mainly determined by the majorized function. In our case, since the surrogate function is constructed through two majorization steps, it might be relatively loose as an upper bound of the original function. Consequently, the convergence seems likely to be relatively slow. In order to accelerate the convergence speed, we need to adopt some acceleration techniques which will be elaborated in the following.

1) Acceleration via SQUAREM: SQUAREM refers to the squared iterative method. It was proposed by Varadhan and Roland [23] and and can be easily implemented as an "off-the-shelf" accelerator for the MM algorithm. Compared with other acceleration schemes, SQUAREM has two important advantages: i) it only requires an MM variable updating scheme and ii) it is guaranteed to converge.

Let $\mathbf{F}_{MM}(\cdot)$ denote the nonlinear fixed-point iteration map of the MIAFIS algorithm. The MM updating scheme can be expressed as $\mathbf{s}^{(\ell+1)} = \mathbf{F}_{MM}(\mathbf{s}^{(\ell)})$. The detailed implementation of the proposed algorithm accelerated via SQUAREM is shown in Algorithm 2. Note that applying SQUAREM may cause two potential problems. First, SQUAREM may violate the PAR and constant energy constraints. Second, it may violate the monotonicity of the proposed MM algorithm. For the first problem, we project the infeasible point back to the constraint set by $-\mathcal{P}_{\mathcal{S}}(\cdot)$. For the second problem, a strategy based on backtracking is adopted to preserve to the monotonicity, which repeatedly halves the distance between -1 and α: $\alpha = (\alpha - 1)/2$ until the monotonicity is achieved. Note that when $\alpha = -1$, $\mathbf{s}^{(\ell)} - 2\alpha\mathbf{q} + \alpha^2\mathbf{v} = \mathbf{s}_2$. According to the monotonicity of the MM algorithm, $f(\mathbf{s}_1) \leq f(\mathbf{s}_2)$. Thus, it is clear that the monotonicity will finally be achieved when the value of α is approaching -1.

The complete description of MIAFIS acceleration via SQUAREM is given in Algorithm 2.

2) Acceleration via Local Majorization: As mentioned above, the potential slowness of the convergence is mainly caused by the double majorization, which may lead to a loose approximation of the original objective function. Besides this, in these two majorization steps, we use the upper bound of $\mathbf{B}$ and and that of $\left(\mathbf{P}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H\right)$, which could make the approximation even looser. Apart from the SQUAREM scheme, which still uses the same surrogate function, a natural idea to accelerate the MM algorithm is to find a better surrogate of the the original quartic objective function at every iteration. Note that the monotonicity of the MM algorithm only requires that $u\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)\geq f\left(\mathbf{s}\right)$ at $\mathbf{s}=\mathbf{s}^{(\ell+1)}$. In other words, $u\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)$ does not have to be a global upper bound of $f\left(\mathbf{s}\right)$ on the whole domain.

Recall the surrogate function of the original objective at the point $\mathbf{s}^{(\ell)}$ in (43). The term $\left(\lambda_u(\mathbf{P})+\lambda_u\left(\mathbf{L}\right)N\right)$ makes the bound globally loose and will influence the convergence speed. By tuning this term, we can achieve a tighter local upper bound of the original objective function at $\mathbf{s}^{(\ell)}$, although it may not be a global upper bound. Thus, we consider the following local upper bound of $f\left(\mathbf{s}\right)$:

$$\begin{aligned}u_t\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right) =\ & 4\mathrm{Re}\left(\mathbf{s}^H\mathbf{P}\mathbf{s}^{(\ell)}-t\mathbf{s}^H\mathbf{s}^{(\ell)}\right)+4Nt\\ &-\left(2\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{P}\mathbf{s}^{(\ell)}+\mathrm{vec}\left(\mathbf{S}^{(\ell)}\right)^H\mathbf{B}\mathrm{vec}\left(\mathbf{S}^{(\ell)}\right)\right)\\ =\ & 4\mathrm{Re}\left(\mathbf{s}^H\mathbf{P}\mathbf{s}^{(\ell)}-t\mathbf{s}^H\mathbf{s}^{(\ell)}\right)+4Nt\\ &-2\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{P}\mathbf{s}^{(\ell)}-f\left(\mathbf{s}^{(\ell)}\right),\end{aligned}\tag{44}$$

where t needs to be chosen such that $u_t\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)\geq f\left(\mathbf{s}\right)$ at the minimizer of $u_t\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)$ over the constraint set, which is

$$\mathbf{s}_t^\star=\mathcal{P}_{\mathcal{S}}\left(\left(\mathbf{P}-t\mathbf{I}\right)\mathbf{s}^{(\ell)}\right).\tag{45}$$

In order to understand the meaning of t, we can interpret our acceleration scheme from another perspective. In fact, the surrogate function (44) can be expressed in the form of

$$\begin{aligned}u\left(\mathbf{s},\ \mathbf{s}^{(\ell)}\right)=&f\left(\mathbf{s}^{(\ell)}\right)+\nabla f\left(\mathbf{s}^{(\ell)}\right)^T\begin{bmatrix}\mathbf{s}-\mathbf{s}^{(\ell)}\\ \left(\mathbf{s}-\mathbf{s}^{(\ell)}\right)^*\end{bmatrix}\\ &+\frac{L}{2}\begin{bmatrix}\mathbf{s}-\mathbf{s}^{(\ell)}\\ \left(\mathbf{s}-\mathbf{s}^{(\ell)}\right)^*\end{bmatrix}^H\begin{bmatrix}\mathbf{s}-\mathbf{s}^{(\ell)}\\ \left(\mathbf{s}-\mathbf{s}^{(\ell)}\right)^*\end{bmatrix},\end{aligned}\tag{46}$$

where $\nabla f(\mathbf{s}^{(\ell)})=[\,\frac{\partial f}{\partial\mathbf{s}}\ \frac{\partial f}{\partial\mathbf{s}^*}\,]_{\mathbf{s}=\mathbf{z}}$, L is an upper bound of L_f, the Lipschitz constant of the function $f\left(\mathbf{s}\right)$, which is hard to obtain in most situations. Compared with (44), it is clear to see that $t=\frac{L}{2}$. In order to make L close to L_f, we adopt such a scheme: starting from a small value of t, we keep increasing its value until monotonicity is achieved. We know that when $t=\lambda_u(\mathbf{P})+\lambda_u\left(\mathbf{B}\right)N$, the monotonicity must be satisfied.

The complete description of MIAFIS acceleration via local majorization is given in Algorithm 3.

Algorithm 3: MIAFIS Acceleration via Local Majorization.

Input: Initial sequence $\mathbf{s}^{(0)}$
Output: Designed sequence $\mathbf{s}$

$\lambda_u(\mathbf{B})=\sum_{k=1}^{NN_\nu}p_k(N-r)$
repeat
  $\mathbf{P}=\sum_{k=1}^{NN_\nu}\frac{p_k}{2}\left(tr(\mathbf{A}_k^H\mathbf{S}^{(\ell)})\mathbf{A}_k+tr(\mathbf{A}_k\mathbf{S}^{(\ell)})\mathbf{A}_k^H\right)$
  Calculate $\lambda_u\left(\mathbf{P}\right)$ according to (32)
  repeat for m in $\{0,1,\ldots,N-1\}$
    $t=\frac{\lambda_u(\mathbf{P})+\lambda_u(\mathbf{B})N}{2^{(N-m)}}$
    $\mathbf{s}_t^\star=\mathcal{P}_{\mathcal{S}}((\mathbf{P}-t\mathbf{I})\mathbf{s}^{(\ell)})$
    $m\leftarrow m+1$
  until $u_t(\mathbf{s}_t^\star,\mathbf{s}^{(\ell)})>f(\mathbf{s}_t^\star)$
  $\mathbf{s}_n^{(\ell+1)}=\mathbf{s}_t^\star$
  $\ell\leftarrow\ell+1$
until convergence

Algorithm 4: MIAFIS - Majorized Iteration for Ambiguity Function Iterative Shaping.

Input: Initial sequence $\mathbf{s}^{(0)}$
Output: Designed sequence $\mathbf{s}$

$\lambda_u(\mathbf{B})=\sum_{k=1}^{NN_\nu}p_k(N-r)$
repeat
  $\mathbf{P}=\sum_{k=1}^{NN_\nu}\frac{p_k}{2}\left(tr(\mathbf{A}_k^H\mathbf{S}^{(\ell)})\mathbf{A}_k+tr(\mathbf{A}_k\mathbf{S}^{(\ell)})\mathbf{A}_k^H\right)$
  Calculate $\lambda_u\left(\mathbf{P}\right)$ according to (32)
  $\mathbf{z}=\left(\mathbf{P}-\left(\lambda_u\left(\mathbf{B}\right)N+\lambda_u\left(\mathbf{P}\right)\right)\mathbf{I}\right)\mathbf{s}^{(\ell)}$
  $\mathbf{s}^{(\ell+1)}=-e^{j\arg(\mathbf{z})}$
  $\ell\leftarrow\ell+1$
until convergence

IV. A Special Case: Unit-modulus Constraint

As mentioned at the end of Section II, when $\gamma=1$, the PAR constraint and constant energy constraint together become the unit-modulus constraint, $|s_n|=1,\ n=1,2,\ldots,N$. The original problem (10) becomes

$$\begin{aligned}\underset{\mathbf{s}}{\text{maximize}}\quad&\text{SINR}\\ \text{subject to}\quad&|s_n|=1,\ n=1,2,\ldots,N.\end{aligned}\tag{47}$$

The MM updating scheme we have derived above becomes much easier for the unit-modulus case, and $\mathcal{P}_{\mathcal{S}}\left(\cdot\right)$ can be simplified to $-e^{j\arg(\cdot)}$. The description of MIAFIS for the unit-modulus case is given in Algorithm 4. Also, all the acceleration schemes can still be adopted here.

However, in the derivation of MIAFIS, except for the conjugate symmetry, the matrix $\left(\mathbf{P}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H\right)$ does not provide any useful properties for us to take advantage of, so obtaining a nice upper bound of its eigenvalues is a nontrivial issue. Thus it is natural to ask whether we can avoid such an issue or not. In the following, we will give a positive answer.

With the unit-modulus constraint, problem (28) becomes

$$\begin{aligned}\underset{\mathbf{s}}{\text{minimize}}\quad&\mathbf{s}^H\mathbf{Q}\mathbf{s}\\ \text{subject to}\quad&|s_n|=1,\ n=1,2,\ldots,N,\end{aligned}\tag{48}$$

where $\mathbf{Q}^{(\ell)}=\mathbf{P}-\lambda_u\left(\mathbf{B}\right)\mathbf{s}^{(\ell)}(\mathbf{s}^{(\ell)})^H$.

Two features of problem (48) should be noted. First, we obtain problem (48) within the MM framework, which means that the monotonicity can still be guaranteed as long as $\left(\mathbf{s}^{(\ell+1)}\right)^H\mathbf{Q}^{(\ell)}\mathbf{s}^{(\ell+1)}\leq(\mathbf{s}^{(\ell)})^H\mathbf{Q}^{(\ell-1)}\mathbf{s}^{(\ell)}$. Second, the feasi-

Algorithm 5: CIAFIS - Coordinate Iteration for Ambiguity Function Iterative Shaping.

Input: Initial sequence $\mathbf{s}^{(0)}$
Output: Designed sequence $\mathbf{s}$
1: $\lambda_u(\mathbf{B}) = \sum_{k=1}^{NN_\nu} p_k(N-r)$
2: **repeat**
3: $\quad \mathbf{P} = \sum_{k=1}^{NN_\nu} \frac{p_k}{2}\left(tr(\mathbf{A}_k^H \mathbf{S}^{(\ell)})\mathbf{A}_k + tr(\mathbf{A}_k \mathbf{S}^{(\ell)})\mathbf{A}_k^H\right)$
4: $\quad$ **for** n=1:N **do**
5: $\quad\quad s_n^{(\ell+1)} = -e^{j\arg(\mathbf{q}_{-n}^H \mathbf{s}_{-n})}$
6: $\quad$ **end for**
7: $\quad \ell \leftarrow \ell + 1$
8: **until** convergence

ble set of problem (48) is a Cartesian product for $s_n,\ n = 1, 2, \ldots, N$.

We define $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N]$, $\mathbf{q}_{-n} = [q_{1,n}, q_{2,n}, \ldots, q_{n-1,n}, q_{n+1,n}, \ldots q_{N,n}]^T$ and $\mathbf{s}_{-n} = [s_1^{(\ell+1)}, s_2^{(\ell+1)}, \ldots, s_{n-1}^{(\ell+1)}, s_{n+1}^{(\ell)}, \ldots s_n^{(\ell)}]^T$. For the coordinate $s_n,\ n = 1, 2, \ldots, N$, we have the following problem:

$$\begin{aligned} \underset{s_n}{\text{minimize}} \quad & 2\text{Re}\left(\mathbf{q}_{-n}^H \mathbf{s}_{-n} s_n^*\right) \\ \text{subject to} \quad & |s_n| = 1,\ n = 1, 2, \ldots, N, \end{aligned} \tag{49}$$

which has a closed-form solution

$$s_n^{(\ell+1)} = -e^{j\arg\left(\mathbf{q}_{-n}^H \mathbf{s}_{-n}\right)}. \tag{50}$$

The complete description of this algorithm is given in Algorithm 5.

With the above algorithm, we observe that we avoid the upper bound at the cost of updating $\mathbf{P}$ at every iteration. However, since the majority of the parameters p_k are zeros, the computation cost is not too high. Similarly, we can also apply the acceleration scheme SQUAREM to accelerate it.

V. Comparison and Connection

In this section, we will review and gain extra insight into the proposed two algorithms. We first compare MIAFIS and CIAFIS to reveal their advantages and disadvantages. Then we reveal the connections between MIAFIS and the traditional gradient projection method.

A. Comparison Between MIAFIS and CIAFIS

First, both of the algorithms are based on the general MM method, so the monotonicity can be guaranteed and some acceleration techniques can be adopted. The difference lies after the first majorization: for the majorized problem, at every iteration, MIAFIS chooses to majorize again, while CIAFIS turns to the coordinate descent method. As mentioned above, the original intent of CIAFIS is to avoid the upper bound issue of the eigenvalues of an complicated matrix. In addition, the feasible set is Cartesian and a closed-form solution can be found for every coordinate. This makes CIAFIS a suitable solution for the sequence design with the unit-modulus constraint.

In Algorithm 4, the computation of the algorithm per iteration is dominated by $\left((\mathbf{s}^{(\ell)})^H \mathbf{A}_k^H \mathbf{s}^{(\ell)} \mathbf{A}_k \mathbf{s}^{(\ell)}\right)$, while in Algorithm 5, it is dominated by $\left((\mathbf{s}^{(\ell)})^H \mathbf{A}_k^H \mathbf{s}^{(\ell)} \mathbf{A}_k\right)$. However, CIAFIS also needs to update all the N coordinates sequentially, which is dominated by $\left(\hat{\mathbf{q}}_{-n}^H \hat{\mathbf{s}}_{-n}\right)$ per update. Therefore, the time complexities of both MIAFIS and CIAFIS are $\mathcal{O}\left(N^2\right)$.

CIAFIS not only avoids the upper bound issue, but also has the same time complexity as MIAFIS. However, the convergence is the cost CIAFIS has to pay. Notice that in the general MM algorithm, we need to obtain the global optimal solution of the majorized function over the feasible set in order to guarantee the convergence to stationary points, theoretically. In CIAFIS, we solve the majorized problem through the coordinate descent method, which at best can provide a local optimal solution. Nevertheless, the monotonicity, as mentioned above, can still be guaranteed, and our numerical experiments indicate CIAFIS usually reaches a stationary point.

B. Connections Between MIAFIS and the Traditional Gradient Projection Method

Note that in MIAFIS with the PAR constraint, the update rule for $\mathbf{s}$ is given by $\mathbf{s}^{(\ell+1)} = \mathcal{P}_{\mathcal{S}}(\mathbf{z})$, where $\mathbf{z} = (\mathbf{P} - (\lambda_u(\mathbf{L})N + \lambda_u(\mathbf{P}))) \mathbf{s}^{(\ell)}$. $\mathcal{P}_{\mathcal{S}}(\cdot)$ could be considered a projection operation. $\mathbf{z}$ includes the current iteration point $\mathbf{s}^{(\ell)}$ and the scaled derivative, $\mathbf{P}\mathbf{s}^{(\ell)}$, of the original objective function with respect to $\mathbf{s}^*$. This reminds us of the traditional gradient projection method. In order to reveal the potential connection, we will review the derivation of MIAFIS from the perspective of the gradient projection method in the following.

Recall that with the PAR constraint, after the second majorization, we obtain problem (36):

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \text{Re}\left(\mathbf{s}^H \mathbf{z}\right) \\ \text{subject to} \quad & |s_n| \le \sqrt{\gamma},\ n = 1, 2, \ldots, N \\ & \|\mathbf{s}\|^2 = N, \end{aligned} \tag{51}$$

which is equivalent to

$$\begin{aligned} \underset{\mathbf{s}}{\text{minimize}} \quad & \|\mathbf{s} - \hat{\mathbf{z}}\|^2 \\ \text{subject to} \quad & \mathbf{s} \in \mathcal{S}, \end{aligned} \tag{52}$$

where

$$\hat{\mathbf{z}} = -\mathbf{z} = ((\lambda_u(\mathbf{L})N + \lambda_u(\mathbf{P})) - \mathbf{P})\mathbf{s}^{(\ell)}, \tag{53}$$

$$\mathcal{S} = \left\{\mathbf{s} \in \mathbb{C}^N \mid \|\mathbf{s}\|^2 = N, |s_n| \le \sqrt{\gamma},\ n = 1, 2, \ldots, N\right\}. \tag{54}$$

Reviewing the update rule we have derived in MIAFIS, we have

$$\begin{aligned} \mathbf{s}^{(\ell+1)} &= \mathcal{P}_{\mathcal{S}}(\mathbf{z}) \\ &= -\mathcal{P}_{\mathcal{S}}(\hat{\mathbf{z}}) \\ &= -\mathcal{P}_{\mathcal{S}}\left(((\lambda_u(\mathbf{L})N + \lambda_u(\mathbf{P}))\mathbf{I} - \mathbf{P})\mathbf{s}^{(\ell)}\right) \\ &= -\mathcal{P}_{\mathcal{S}}\left(\left(\mathbf{I} - \frac{\mathbf{P}}{\lambda_u(\mathbf{L})N + \lambda_u(\mathbf{P})}\right)\mathbf{s}^{(\ell)}\right) \\ &= -\mathcal{P}_{\mathcal{S}}\left(\mathbf{s}^{(\ell)} - \frac{1}{\lambda_u(\mathbf{L})N + \lambda_u(\mathbf{P})}\mathcal{D}_{\mathbf{s}^*} f\left(\mathbf{s}^{(\ell)}\right)\right), \end{aligned} \tag{55}$$

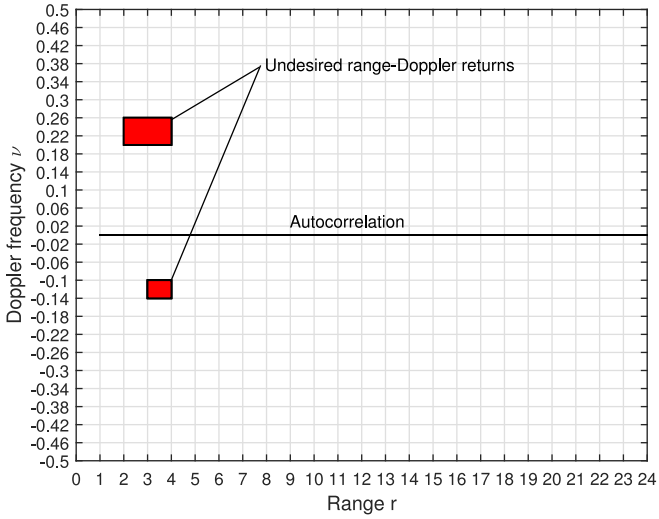

Fig. 1. Range-Doppler interference scenario for test.

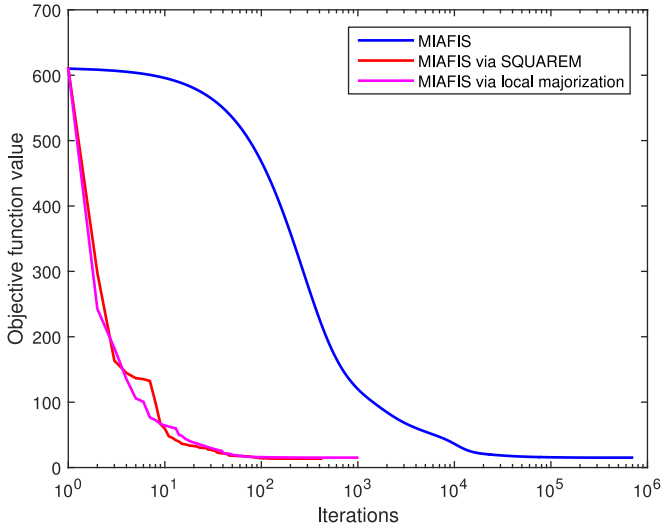

Fig. 2. Convergence of MIAFIS algorithms for $N = 25$.

where $\mathcal{D}_{\mathbf{s}^*} f(\mathbf{s}^{(\ell)})$ is the derivative of the original objective function of (14) with respect to $\mathbf{s}^*$ at the point $\mathbf{s} = \mathbf{s}^{(\ell)}$. Therefore, it is very clear that the update rule of MIAFIS has the same structure as the traditional gradient projection method. However, our update rule works on the complex variables directly. In addition to this, the traditional gradient projection method only works over a convex set, which is essential to guarantee its convergence [24]. Since our algorithm is within the MM framework, the convergence can still be guaranteed even over a nonconvex set. For radar applications, due to the finite energy limit in reality, the feasible set is usually nonconvex. In addition, the step size of our update rule is adaptive, which provides more flexibility and avoids the trouble of using the Armijo rule.

VI. Numerical Experiments

In this section, we first present the testing scenario and give a brief explanation. Second, we present some experimental results to illustrate the proposed algorithm with the PAR constraint. Last, for the unit-modulus case, we compare the performance of the proposed algorithms with the existing algorithms to show the improved performance. All experiments were implemented in MATLAB R2014b and performed on a PC with a 3.30 GHz i5-4950 CPU and 8 GB RAM.

A. Testing Scenario

We consider the range-Doppler interference scenario shown in Figure 1, where we take the sequence length $N = 25$ as an example to illustrate our performance analysis. In this interference map, the red blocks correspond to the regions of unwanted range-Doppler returns (e.g., interference, fake target, etc.). The normalized Doppler frequency axis is discretized into $N_v = 50$ bins, and consequently the discrete Doppler frequency is given by $v_h = -\frac{1}{2} + \frac{h}{50},\ h = 0, 1, \ldots, 49$. In addition, a uniform interference power is assumed among the interference bins. To be more specific, we take

$$p(r,h) = \begin{cases} 1 & (r,h) \in \{2,3,4\} \times \{35,36,37,38\} \\ 1 & (r,h) \in \{3,4\} \times \{18,19,20\} \\ 1 & (r,h) \in \{1,2,\ldots,\ N-1\} \times \{25\} \\ 0 & \text{otherwise.} \end{cases} \tag{56}$$

Note that in this interference map, we not only suppress the unwanted range-Doppler returns, but also control the ISL over all the lags of the autocorrelation of the transmitted sequence. Also, weighted ISL control can be readily incorporated by letting the $p(r,h)$ corresponding to some particular sidelobes be zero. Furthermore, by assuming a uniform interference power, we don't give any priority or preference to the interference or ISL. In the following, all the simulations are based on the above scenario (56) unless otherwise specified. Also note that all the simulations in the following two sections are based on the particular system model described above. To the best of our knowledge, there is no benchmark considering both the same system model and the PAR constraint.

B. Sequence Design to Improve Ambiguity Function with the PAR Constraint

In Figure 2, we plot the convergence curves of the objective value with respect to the number of iterations for the above scenario. A randomly generated sequence is used as the initial one, and the PAR parameter here is $\gamma = 4$. From this figure, we can see clearly that the two accelerated algorithms require far fewer iterations (around 2-3 orders of magnitude less).

Recall that the squared magnitude of the ambiguity function of the matched filter of the radar sequence $\mathbf{s}$ after normalization is given by

$$g_{\mathbf{s}}(r,\ v) = \frac{1}{\|\mathbf{s}\|^2} \left| \mathbf{s}^H \mathbf{J}^r \operatorname{diag}(\mathbf{p}(v))\, \mathbf{s} \right|^2. \tag{57}$$

In Figure 3, we compare the ambiguity function of the initial random sequence and that of the designed sequence for the

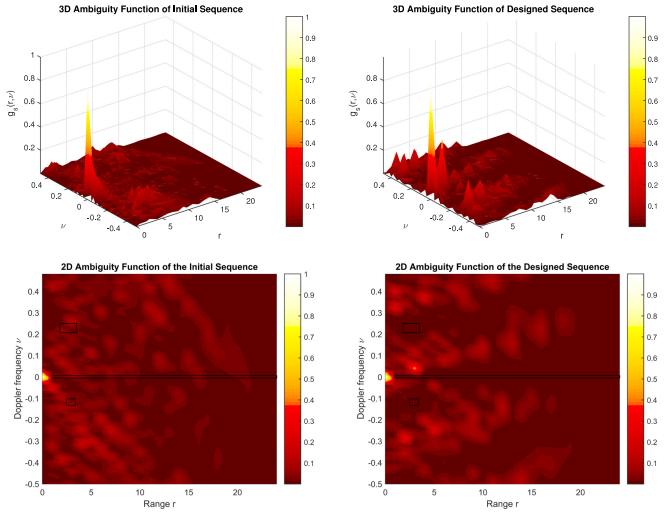

Fig. 3. Ambiguity function. Left: Initial sequence; Right: Designed sequence.

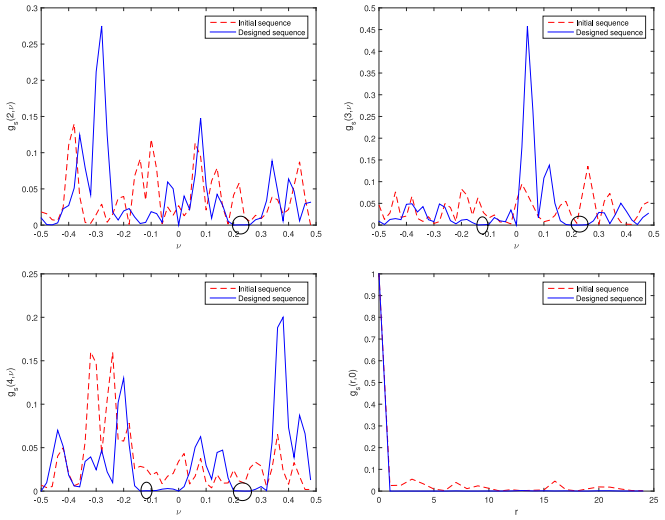

Fig. 4. Ambiguity function cut at different $r = 2, 3, 4$; Ambiguity function cut at $\nu = 0$.

above scenario. From this comparison, we can see clearly that the unwanted range-Doppler responses in the two red blocks are suppressed to a very low level and the ISL is significantly improved. In Figure 4, we provide the range-cuts for different values of r, and also the autocorrelation. From both Figure 3 and Figure 4, it is clear to see that the shaped ambiguity function experiences deep nulls where interference is foreseen, which indicates that the proposed algorithms shape the range-Doppler response very well.

In Figure 5, we show the achieved performance with respect to the PAR parameter γ. $\mathbf{s}_{ini}$ and $\mathbf{s}_{new}$ are the initial sequence and the corresponding designed sequence, respectively, for a particular γ. Specifically, except the case of $\gamma = 1$, where $\mathbf{s}_{ini}$ is generated randomly, $\mathbf{s}_{ini}$ for every γ is $\mathbf{s}_{new}$ for the previous γ. The objective value is the converged value for every γ. The left-hand plot is the curve of the objective value with respect to the PAR parameter γ, while the right-hand plot shows the difference between $\mathbf{s}_{ini}$ and $\mathbf{s}_{new}$ for every value of γ. From

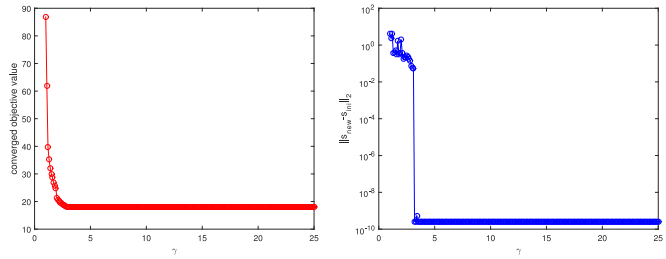

Fig. 5. Left: Objective value v.s. PAR parameter γ (the stepsize of γ is 0.1); Right: $||\mathbf{s}_{new} - \mathbf{s}_{ini}||_2$ v.s. PAR parameter γ.

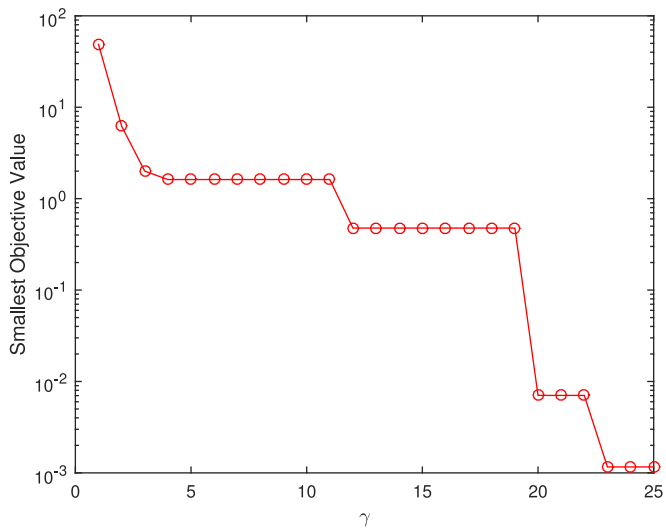

Fig. 6. Smallest objective value v.s. γ. Each point is averaged over 1000 random trials.

Figure 5, we can clearly see that the objective value is nonincreasing with γ, which is consistent with our intuition that the feasible set becomes larger when γ keeps increasing. In particular, when the value of γ changes from 1 to 2, the objective value decreases significantly. But the objective value barely changed for $\gamma > 5$, which can be verified by the right-hand plot, where each $\mathbf{s}_{new}$ is almost the same for $\gamma > 5$. From the perspective of optimization, it is advantageous to choose a large γ to achieve a good performance. However, in practice, a large γ will result in a designed sequence with a high PAR, which is usually undesirable, for it is difficult to implement and usually strains the analog circuit [25].

In Figure 6, different from the plot scheme above, we run MIAFIS 1000 times with different initial sequences for every value of γ, and we plot the smallest objective value among them. Note that there are $2N = 50$ degrees of freedom (the free phases and magnitudes of sequence $\{s_n\}_{n=1}^{N}$), and our goal is to match 42 tags of the range-Doppler response (the interference and ISL regions of the whole range-Doppler plan). Thus, there are enough degrees of freedom, and, consequently, the global optimal value can be driven close to 0, which can be observed clearly in this figure, especially when $\gamma > 22$. Besides, the above simulation scheme reminds us that, in practice, we can run our algorithm with different initial points independently and then

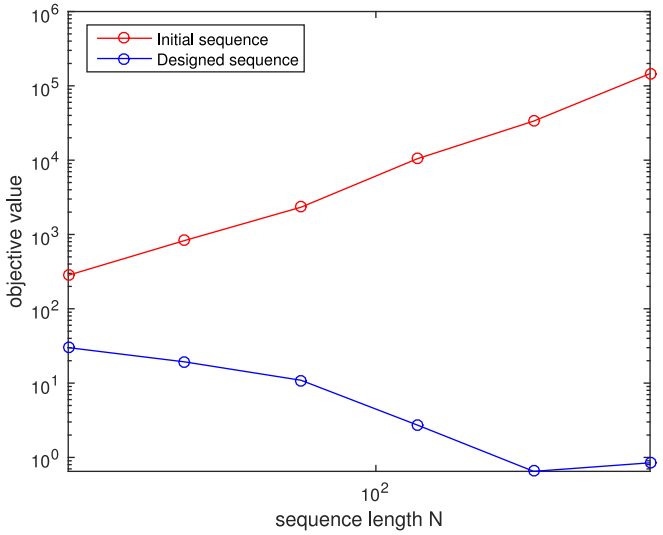

Fig. 7. Objective value versus sequence length N.

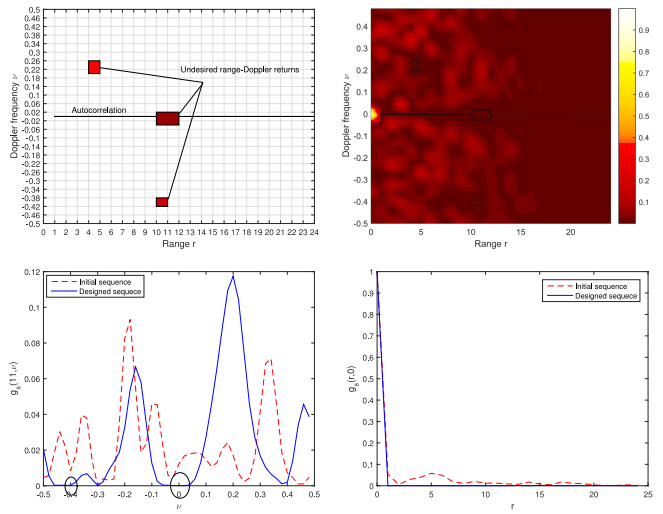

Fig. 8. Numerical resluts on the testing scenario given by (58). Top left: Testing scenario; Top right : Ambiguity Function; Bottom left: Ambiguity function cut at $r = 11$; Bottom right: Ambiguity function cut at $\nu = 0$.

select the best solution among the outputs, which is probably the optimal one.

Also, in Figure 7, we plot the curves of the objective value with respect to different sequence lengths N for different sequences. Note that we can see clearly that our designed sequences still achieve good performance even for a long sequence. In addition, if the scenario does not change, we can shape the ambiguity function much better by using a longer sequence, which means we can achieve a better SINR.

Since our algorithms are locally convergent, we test our our algorithm on two more testing scenarios given by, respectively,

$$p(r,h)=\begin{cases}0.5 & (r,h)\in\{4,5\}\times\{35,36,37,38\}\\ 2 & (r,h)\in\{10,11,12\}\times\{23,24,25,26\}\\ 1 & (r,h)\in\{10,11\}\times\{4,5,6\}\\ 1 & (r,h)\in\{1,\ldots,9,13,\ldots,N-1\}\times\{25\}\\ 0 & \text{otherwise.}\end{cases} \tag{58}$$

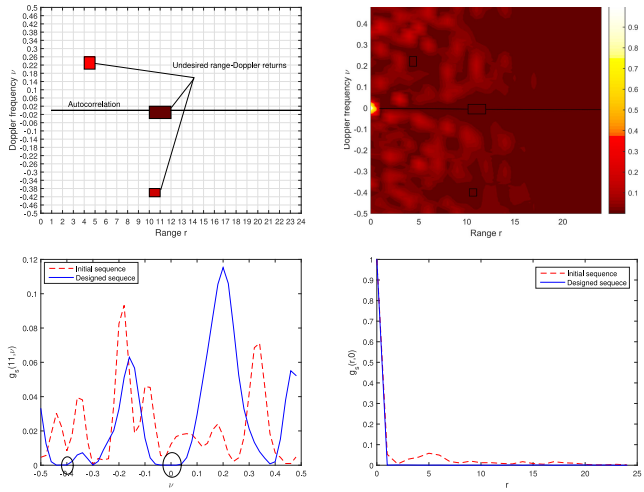

Fig. 9. Numerical resluts on the testing scenario given by (59). Top left: Testing scenario; Top right : Ambiguity Function; Bottom left: Ambiguity function cut at $r = 11$; Bottom right: Ambiguity function cut at $\nu = 0$.

and

$$p(r,h)=\begin{cases}0.5 & (r,h)\in\{4,5\}\times\{35,36,37,38\}\\ 4 & (r,h)\in\{10,11,12\}\times\{23,24,25,26\}\\ 1 & (r,h)\in\{10,11\}\times\{4,5,6\}\\ 1 & (r,h)\in\{1,\ldots,9,13,\ldots,N-1\}\times\{25\}\\ 0 & \text{otherwise.}\end{cases} \tag{59}$$

Compared with the testing scenario shown in 1, these two tesing scenarios have different locations and values of nonzero $p\,(r,h)$s. The simulation resluts are shown in Figure 8 and Figure 9. In both figures, the ambiguity functions shaped by our algorithm match the corresponding range-Doppler interference scenarios very well. This indicates that our algorithm adapts to various situations.

C. Sequence Design to Improve Ambiguity Function with Unit-modulus Constraint

In Figure 10 and Figure 11, we plot the convergence curves for the unit-modulus case using MIAFIS and CIAFIS. From this figure, we can see that CIAFIS also works very well and achieves almost the same performance as MIAFIS.

Note that the performances of both MIAFIS and CIAFIS are related to the initial sequence. In order to show the extent of the influence of the initial sequence to our algorithms, we plot the curve of the averaged SINR with sequence length N in Figure 12. The curve is averaged over 100 trials, which have different randomly-generated initial sequence. The bar for each point indicates the range of the acieved SINR among the 100 trials for the corresponding N. From Figure 12, the range is usually less than 3 dB.

In Figure 13 and Figure 14, we compare, respectively, the ambiguity functions and the ambiguity function cuts of the designed sequences by our methods and the benchmark, MBI-type methods. The testing scenario is shown in Figure 1 and the sequence length is $N = 25$. Notice that for the MBI-type methods,

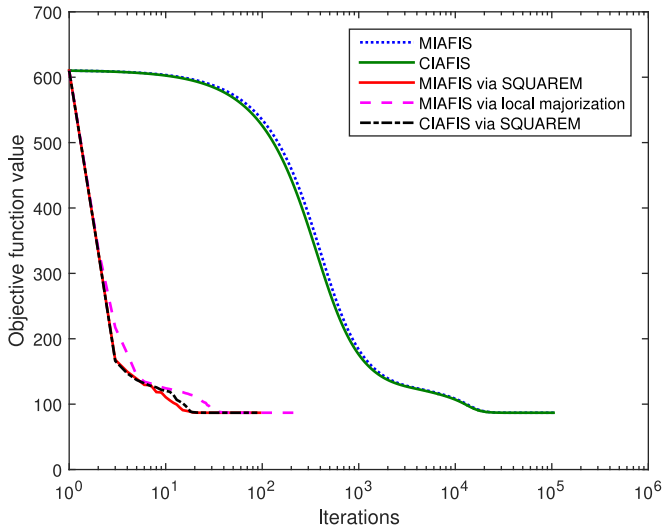

Fig. 10. Objective value versus the number of iterations. $N = 25$.

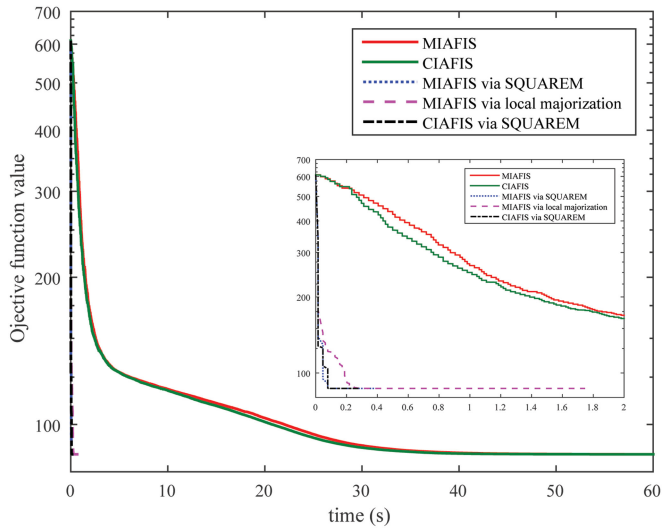

Fig. 11. Objective value versus time (in seconds). $N = 25$.

we choose the best one (MBIQ with 0.5λ) for the comparison. From Figure 13, it is clear to see that all of the methods can shape a desired ambiguity function although the designed sequences are different. Figure 14 shows that all the shaped ambiguity functions have deep nulls where the interference is forseen.

Then, we compare the performance of our methods with the MBI methods in terms of the CPU time and the achieved SINR, as shown in Figure 15. The two red blocks of the testing scenario shown in Figure 1 remain unchanged, while the ISL control line continues to increase with the sequence length N. Considering the CPU time, our methods are much better than the MBI-type methods. Note that in each MBI-type method, the parameter λ has to be evaluated appropriately to guarantee the monotonicity. However, evaluating the satisfactory value of λ is not trivial, and in [15] it was in fact obtained by solving an optimization problem. For more details, please refer to [15]. Even if we do not include the time for obtaining λ, our methods are still better that

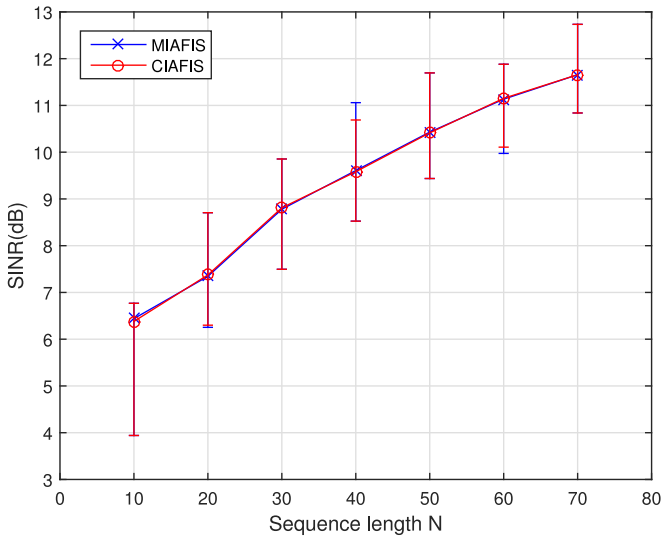

Fig. 12. SINR versus sequence length. Each curve is averaged over 100 random trials

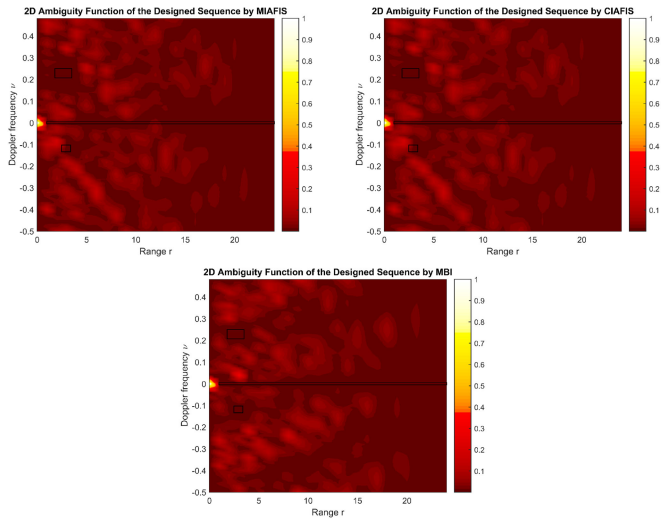

Fig. 13. Comparison of the ambiguity functions. Top left: MIAFIS; Top right: CIAFIS; Bottom: MBI.

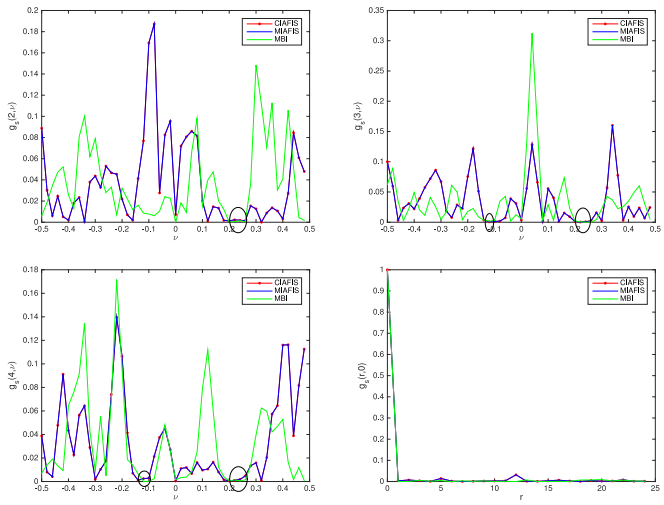

Fig. 14. Comparison of the ambiguity function cuts.

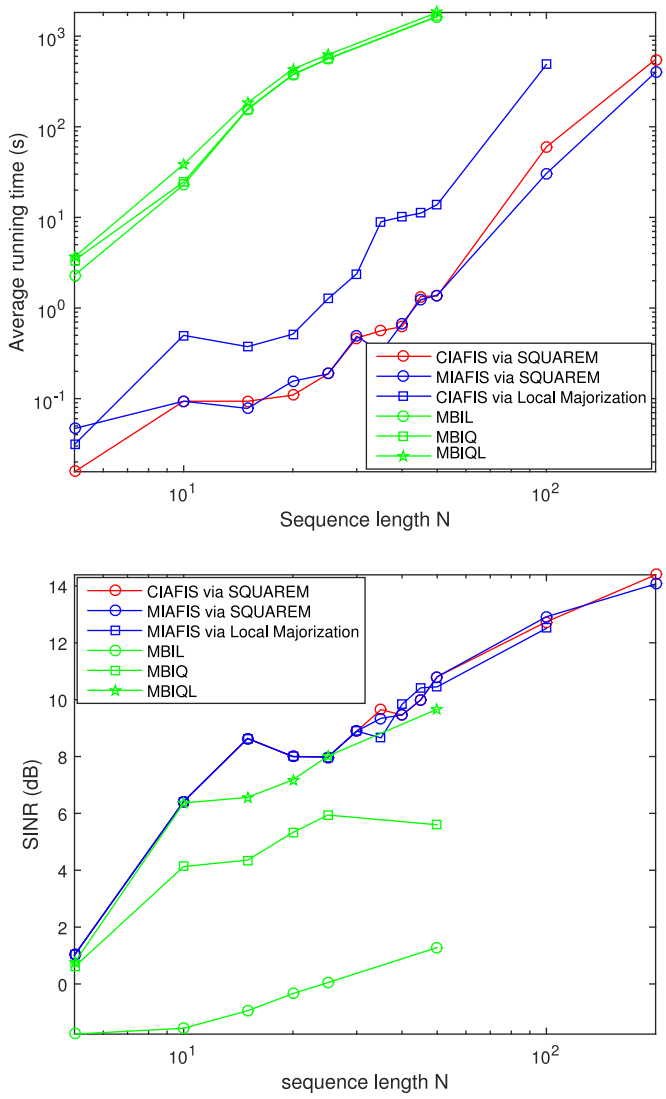

Fig. 15. Top: Average CPU time v.s. N ; Bottom: Average objective value v.s. N, where $|\alpha|^2 = 1$ and $\sigma_n^2 = 0.5$. Each curve is averaged over 100 random trials.

the MBI-type methods, and can outperform them. Considering the SINR, it is also very clear to see that a longer sequence achieves a higher SINR, which can also be interpreted from the perspective of degrees of freedom. Note that for the MBI-type methods, the running time will increase vastly for $N \geq 100$ and the memory of the PC we use is insufficient, thus we only provide the cases for $N \leq 50$. However, this does not affect the comparison and performance evaluation.

VII. Conclusion

Sequence design for radar systems can be interpreted from the perspective of improving the ambiguity function. Starting from raising the SINR, we have derived an efficient algorithm called MIAFIS for sequence design with the PAR and finite energy constraints to improve the ambiguity function. The MIAFIS is derived based on the general MM method and its convergence to a stationary point can be guaranteed. Additionally, in case of the ill-construction of the majorization function, two acceleration schemes have been considered. We also consider the unimodular case, which is a particular case of the general PAR formulation. Another efficient algorithm, CIAFIS, based on the MM method and coordinate descent method, has been proposed. Numerical experiments show the efficiency of the proposed algorithms in designing sequences with desired ambiguity functions and demonstrate their advantages in terms of both the CPU time and SINR compared with the existing methods.

Appendix

A. Proof of Lemma 2

Proof: First, we have

$$\begin{aligned}\mathrm{tr}(\mathbf{P}_k) &= \mathrm{tr}\left(\mathrm{tr}\left(\mathbf{A}_k^H\mathbf{S}^{(\ell)}\right)\mathbf{A}_k + \mathrm{tr}\left(\mathbf{A}_k\mathbf{S}^{(\ell)}\right)\mathbf{A}_k^H\right)\\ &= \mathrm{tr}\left(\mathbf{A}_k^H\mathbf{S}^{(\ell)}\right)\mathrm{tr}\left(\mathbf{A}_k\right) + \mathrm{tr}\left(\mathbf{A}_k\mathbf{S}^{(\ell)}\right)\mathrm{tr}\left(\mathbf{A}_k^H\right)\\ &= 2\mathrm{Re}\left(\mathrm{tr}\left(\mathbf{A}_k^H\mathbf{S}^{(\ell)}\right)\mathrm{tr}\left(\mathbf{A}_k\right)\right) \qquad (60)\end{aligned}$$

and

$$\mathrm{tr}(\mathbf{P}_k^2) = \mathrm{vec}\left(\mathbf{P}_k\right)^H \mathrm{vec}\left(\mathbf{P}_k\right). \qquad (61)$$

If $r \neq 0$, $\mathrm{Tr}(\mathbf{A}_k) = 0$. Thus $\mathrm{Tr}\left(\mathbf{P}_k\right) = 0$. Considering the special structure of $\mathbf{P}_k$, we have

$$\begin{aligned}\mathrm{Tr}(\mathbf{P}_k^2) =& \ \mathrm{vec}\left(\mathbf{P}_k\right)^H \mathrm{vec}\left(\mathbf{P}_k\right)\\ =& \left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|^2 \mathrm{vec}\left(\mathbf{A}_k\right)^H \mathrm{vec}\left(\mathbf{A}_k\right)\\ &+ \left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|^2 \mathrm{vec}\left(\mathbf{A}_k^H\right)^H \mathrm{vec}\left(\mathbf{A}_k^H\right)\\ &+ \left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\mathrm{vec}\left(\mathbf{A}_k\right)^H\mathrm{vec}\left(\mathbf{A}_k^H\right)\\ &+ \left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k^H\mathbf{s}^{(\ell)}\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k^H\mathbf{s}^{(\ell)}\mathrm{vec}\left(\mathbf{A}_k^H\right)^H\mathrm{vec}\left(\mathbf{A}_k\right)\\ =& \ 2\left(N-r\right)\left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|^2, \qquad (62)\end{aligned}$$

where the last equality holds because $\mathrm{vec}(\mathbf{A}_k)^H\mathrm{vec}(\mathbf{A}_k) = \mathrm{vec}(\mathbf{A}_k^H)^H\mathrm{vec}(\mathbf{A}_k^H) = N - r$ and $\mathrm{vec}(\mathbf{A}_k)^H\mathrm{vec}(\mathbf{A}_k^H) = \mathrm{vec}\left(\mathbf{A}_k^H\right)^H\mathrm{vec}(\mathbf{A}_k) = 0$.

Thus according to Lemma 1, we have

$$m = 0,\ s^2 = \frac{2\left(N-r\right)}{N}\left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|^2 \qquad (63)$$

and

$$\lambda_{\max}(\mathbf{P}_k) \leq \sqrt{\frac{2\left(N-r\right)\left(N-1\right)}{N}}\left|\left(\mathbf{s}^{(\ell)}\right)^H\mathbf{A}_k\mathbf{s}^{(\ell)}\right|. \qquad (64)$$

If $r = 0$, then $\mathbf{P}_k$ is a diagonal matrix. So we just need to find an upper bound of the diagonal elements. We have

$$\mathrm{Tr}\left(\mathbf{A}_k\mathbf{S}^{(\ell)}\right) = \sum_{i=0}^{N-1} e^{j2\pi i v_h}. \qquad (65)$$

Define $\mathbf{p}=\left[1,e^{j2\pi\nu_h},\ldots,e^{j2\pi(N-1)\nu_h}\right]^T$, and then we have

$$\begin{aligned}\mathbf{P}_k &= \left(\sum_{i=0}^{N-1}e^{j2\pi i v_h}\right)\left(\mathrm{Diag}\left(\mathbf{p}\right)\right)^H \\ &\quad+\left(\sum_{i=0}^{N-1}e^{-j2\pi i v_h}\right)^H\left(\mathrm{Diag}\left(\mathbf{p}\right)\right) \\ &= \mathrm{Diag}\left(\left\{\sum_{i=0}^{N-1}2\cos\left(2\pi\left(i-d\right)v_h\right)\right\}_{d=0}^{N-1}\right). \qquad (66)\end{aligned}$$

For $d=0,1,\ldots,N-1$, we have $\sum_{i=0}^{N-1}2\cos(2\pi(i-d)v_h)\leq 2N$. Thus,

$$\lambda_{\max}(\mathbf{P}_k)\leq 2N. \qquad (67)$$

The proof is complete. ■

B. Proof of Lemma 3

Proof: The objective of problem (36) can be expressed equivalently as

$$\mathrm{Re}\left(\mathbf{s}^H\mathbf{z}\right)=\sum_{n=1}^{N}|s_n||z_n|\cos\left(\arg\left(z_n\right)-\arg\left(s_n\right)\right). \qquad (68)$$

Since $|s_n||z_n|\geq 0$ and the argument and the magnitude of s_n are independent, $\mathrm{Re}\left(\mathbf{s}^H\mathbf{z}\right)$ get its minimum when $\arg\left(s_n\right)=\arg\left(z_n\right)+(2k+1)\pi$, $\forall k=0,\pm1,\ldots$. Therefore, $s_n=-|s_n|e^{j\arg(z_n)}$ and problem (36) can be further simplified as

$$\begin{aligned}\underset{|s_n|}{\text{maximize}}\quad & \sum_{n=1}^{N}|s_n||z_n| \\ \text{subject to}\quad & |s_n|\leq\sqrt{\gamma} \\ & \sum_{n=1}^{N}|s_n|^2=N,\ n=1,2,\ldots,N, \qquad (69)\end{aligned}$$

which can be rewritten in a vector form:

$$\begin{aligned}\underset{|\mathbf{s}|}{\text{maximize}}\quad & |\mathbf{s}|^T|\mathbf{z}| \\ \text{subject to}\quad & 0\leq|\mathbf{s}|\leq\sqrt{\gamma}\mathbf{1} \\ & |\mathbf{s}|^T|\mathbf{s}|=N, \qquad (70)\end{aligned}$$

where $|\cdot|$ denotes elementwise absolute value.

Without loss of generality, we assume that $|z_1|\geq|z_2|\geq\cdots\geq|z_N|$ and the number of nonzero elements of $\mathbf{z}$ is m. Then the solution to problem (69) is as follows:

if $m\gamma\leq N$, then $|s_n|=\sqrt{\gamma}$, for $n=1,\ldots,m$. For $n=m+1,\ldots,N$, we have $\sum_{n=m+1}^{N}|s_n|^2=N-m\gamma$ and $0\leq|s_n|\leq\sqrt{\gamma}$. Thus, there are multiple solutions to problem (69) for this case and one of them is given by

$$|s_n|=\begin{cases}\sqrt{\gamma}, & n=1,\ldots,m,\\ \sqrt{\frac{N-m\gamma}{N-m}}, & n=m+1,\ldots,N.\end{cases} \qquad (71)$$

If $m\gamma>N$, then the solution to problem (69) is given by

$$|s_n| = \min\left\{\beta|z_n|,\sqrt{\gamma}\right\},n=1,\ldots,N, \qquad (72)$$

where

$$\begin{aligned}\beta\in & \left\{\beta|\sum_{n=1}^{N}\min\left\{\beta^2|z_n|^2,\gamma\right\}\right. \\ & \left.=N,\beta\in\left[0,\frac{\sqrt{\gamma}}{\min\{|z_n|\,|\,|z_n|\neq 0\}}\right]\right\}. \qquad (73)\end{aligned}$$

Since the function $g(\beta)=\sum_{n=1}^{N}\min\{\beta^2|z_n|^2,\gamma\}$ is strictly increasing within $[0,\frac{\sqrt{\gamma}}{\min\{|z_n|\,|\,|z_n|\neq0\}}]$ and $g(0)=0$, only one unique β exists. Numerically, the simple bisection method can be adopted to find the unique β with a high degree of accuracy.

Therefore, the solution to problem (36) is given by

$$\mathbf{s}=\mathcal{P}_{\mathcal{S}}\left(\mathbf{z}\right), \qquad (74)$$

where

$$\begin{aligned}\mathcal{P}_{\mathcal{S}}\left(\cdot\right) = & -\left(\mathbf{1}_{\mathbb{R}_0^+}\left(N-m\gamma\right)\right)\sqrt{\gamma}\mathbf{u}_m\odot e^{j\arg(\cdot)} \\ & -\left(\mathbf{1}_{\mathbb{R}^-}\left(N-m\gamma\right)\right)\min\{\beta|\mathbf{z}|,\sqrt{\gamma}\mathbf{1}\}\odot e^{j\arg(\cdot)}, \qquad (75)\end{aligned}$$

$\min\{\cdot,\cdot\}$, $|\cdot|$ and $e^{j\arg(\cdot)}$ are element-wise operations, and

$$\mathbf{1}_A\left(x\right)=\begin{cases}1, & \text{if } x\in A,\\ 0, & \text{otherwise.}\end{cases} \qquad (76)$$

$$\mathbf{u}_m=\left[\underbrace{1,\ldots 1}_{m},\underbrace{\sqrt{\frac{N-m\gamma}{N\gamma-m\gamma}},\ldots,\sqrt{\frac{N-m\gamma}{N\gamma-m\gamma}}}_{N-m}\right]^T. \qquad (77)$$

■

C. Proof of Lemma 4

Proof: First, every point of the sequence $\{\mathbf{s}^{(\ell)}\}$ is bounded with $0\leq|\mathbf{s}^{(\ell)}|\leq\sqrt{\gamma}$. According to Theorem 2.17 in [26], at least one limit point must exist.

Denote the objective function of problem (14) by $f(\mathbf{s})$ and the feasible set by $\mathcal{S}$. Consider a limit point $\mathbf{z}$ and the corresponding subsequence $\{\mathbf{s}^{(\ell_i)}\}$. We have

$$\begin{aligned}u\left(\mathbf{s}^{(\ell_{i+1})},\,\mathbf{s}^{(\ell_{i+1})}\right) &= f\left(\mathbf{s}^{(\ell_{i+1})}\right)\leq f\left(\mathbf{s}^{(\ell_i+1)}\right) \\ &\leq u\left(\mathbf{s}^{(\ell_i+1)},\,\mathbf{s}^{(\ell_i)}\right)\leq u\left(\mathbf{s},\,\mathbf{s}^{(\ell_i)}\right), \\ &\forall\mathbf{s}\in\mathcal{S}. \qquad (78)\end{aligned}$$

Letting $i\to\infty$, we obtain

$$u\left(\mathbf{z},\,\mathbf{z}\right)\leq u\left(\mathbf{s},\,\mathbf{z}\right),\ \forall\mathbf{s}\in\mathcal{S}, \qquad (79)$$

which implies

$$\nabla u\left(\mathbf{z},\,\mathbf{z}\right)^T\begin{bmatrix}\mathbf{s}-\mathbf{z}\\ \left(\mathbf{s}-\mathbf{z}\right)^*\end{bmatrix}\geq 0,\ \forall\mathbf{s}\in\mathcal{S}, \qquad (80)$$

where

$$\nabla u\left(\mathbf{z},\,\mathbf{z}\right)=\left[\begin{array}{cc}\frac{\partial u}{\partial\mathbf{s}} & \frac{\partial u}{\partial\mathbf{s}^*}\end{array}\right]_{(\mathbf{s},\,\mathbf{s}^*)=(\mathbf{z},\,\mathbf{z}^*)}. \qquad (81)$$

From the deviation of the majorization function (43) of the objective of problem (14), we can see clearly that

$$\nabla f(\mathbf{z}) = \nabla u(\mathbf{z}, \mathbf{z}). \tag{82}$$

Therefore, $\mathbf{z}$ is a stationary point for problem (14). ■

ACKNOWLEDGMENT

The authors are deeply grateful to Dr. A. Aubry and Prof. A. De Maio for the MBI simulation code.

REFERENCES

[1] N. Levanon and E. Mozeson, *Radar Signals*. Hoboken, NJ, USA: Wiley, 2004.
[2] H. He, J. Li, and P. Stoica, *Waveform Design for Active Sensing Systems: A Computational Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2012.
[3] S. Haykin, "Cognitive radar: A way of the future," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 30–40, Jan. 2006.
[4] P. Stoica, H. He, and J. Li, "New algorithms for designing unimodular sequences with good correlation properties," *IEEE Trans. Signal Process.*, vol. 57, no. 4, pp. 1415–1425, Apr. 2009.
[5] H. He, P. Stoica, and J. Li, "Waveform design with stopband and correlation constraints for cognitive radar," in *Proc. 2nd Int. Workshop Cogn. Inf. Process.*, Elba Island, Italy, Jun. 2010, pp. 344–349.
[6] J. Song, P. Babu, and D. P. Palomar, "Optimization methods for designing sequences with low autocorrelation sidelobes," *IEEE Trans. Signal Process.*, vol. 63, no. 15, pp. 3998–4009, Aug. 2015.
[7] J. Song, P. Babu, and D. P. Palomar, "Sequence design to minimize the weighted integrated and peak sidelobe levels," *IEEE Trans. Signal Process.*, vol. 64, no. 8, pp. 2051–2064, Apr. 2016.
[8] 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.
[9] 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.
[10] A. De Maio, S. D. Nicola, Y. Huang, S. Zhang, and A. Farina, "Code design to optimize radar detection performance under accuracy and similarity constraints," *IEEE Trans. Signal Process.*, vol. 56, no. 11, pp. 5618–5629, Nov. 2008.
[11] A. De Maio, S. De Nicola, Y. Huang, Z.-Q. Luo, and S. Zhang, "Design of phase codes for radar performance optimization with a similarity constraint," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 610–621, Feb. 2009.
[12] A. De Maio, S. De Nicola, Y. Huang, D. P. Palomar, S. Zhang, and A. Farina, "Code design for radar STAP via optimization theory," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 679–694, Feb. 2010.
[13] A. De Maio, Y. Huang, and M. Piezzo, "A Doppler robust max-min approach to radar code design," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4943–4947, Sep. 2010.
[14] A. De Maio, Y. Huang, M. Piezzo, S. Zhang, and A. Farina, "Design of optimized radar codes with a peak to average power ratio constraint," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2683–2697, Jun. 2011.
[15] A. Aubry, A. De Maio, B. Jiang, and S. Zhang, "Ambiguity function shaping for cognitive radar via complex quartic optimization," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5603–5619, Nov. 2013.
[16] B. Chen, S. He, Z. Li, and S. Zhang, "Maximum block improvement and polynomial optimization," *SIAM J. Optim.*, vol. 22, no. 1, pp. 87–107, 2012.
[17] B. R. Mahafza, *Introduction to Radar Analysis*. Boca Raton, FL, USA: CRC, 1998.
[18] M. A. Richards, J. A. Scheer, W. A. Holm, and W. L. Melvin, *Principles of Modern Radar: Basic Principles*. Raleigh, NC, USA: SciTech, 2010.
[19] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statist.*, vol. 58, no. 1, pp. 30–37, 2004.
[20] Y. Sun, P. Babu, and D. Palomar, "Majorization-Minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. PP, no. 99, pp. 1–1, 2016.
[21] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, 2013.
[22] H. Wolkowicz and G. P. Styan, "Bounds for eigenvalues using traces," *Linear Algebra Appl.*, vol. 29, pp. 471–506, 1980.
[23] R. Varadhan and C. Roland, "Simple and globally convergent methods for accelerating the convergence of any EM algorithm," *Scandinavian J. Statist.*, vol. 35, no. 2, pp. 335–353, 2008.
[24] D. P. Bertsekas, *Nonlinear Programming*. St. Belmont, MA, USA: Athena Scientific, 1999.
[25] T. Jiang and Y. Wu, "An overview: Peak-to-average power ratio reduction techniques for OFDM signals," *IEEE Trans. Broadcast.*, vol. 54, no. 2, pp. 257–268, Jun. 2008.
[26] S. Ponnusamy and H. Silverman, *Complex Variables with Applications*. Berlin, Germany: Springer, 2007.

Linlong Wu received the B.E. degree in electronic information from the Xi'an Jiaotong University, Xi'an, China, in 2014. He is currently working toward the Ph.D. degree in electronic and computer engineering at the Hong Kong University of Science and Technology, Hong Kong. His research interests include optimization algorithms with applications in signal processing, MIMO radar, and financial engineering.

Prabhu Babu received the Ph.D. degree in electrical engineering from the Uppsala University, Uppsala, Sweden, in 2012. From 2013–2016, he was a Postdoctoral Fellow with the Hong Kong University of Science and Technology. He is currently with the Centre for Applied Research in Electronics, Indian Institute of Technology Delhi, New Delhi, India.

Daniel P. Palomar (S'99–M'03–SM'08–F'12) received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia, Barcelona, Spain, in 1998 and 2003, respectively.

He is currently a Professor in the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong, which he joined in 2006. Since 2013, he is a Fellow of the Institute for Advance Study (IAS), HKUST. He had previously held several research appointments, namely, at King's College London, London, U.K.; Stanford University, Stanford, CA, USA; Telecommunications Technological Center of Catalonia, Barcelona, Spain; Royal Institute of Technology, Stockholm, Sweden; University of Rome "La Sapienza", Rome, Italy; and Princeton University, Princeton, NJ, USA. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to financial systems, big data systems, and communication systems.

Dr. Palomar received the 2004/06 Fulbright Research Fellowship, the 2004 and 2015 (co-author) Young Author Best Paper Awards by the IEEE Signal Processing Society, the 2015–16 HKUST Excellence Research Award, the 2002/03 best Ph.D. prize in Information Technologies and Communications by the Technical University of Catalonia, the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation, and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT. He is a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING 2016 Special Issue on Financial Signal Processing and Machine Learning for Electronic Trading and has been an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE 2010 Special Issue on Convex Optimization for Signal Processing the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special Issue on Game Theory in Communication Systems and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks.