On the Design of Constant Modulus Probing Signals for MIMO Radar

Y.-C. Wang*, *Member, IEEE*,
X. Wang†, Hongwei Liu†, *Member, IEEE*, and
Z.-Q. Luo‡ , *Fellow, IEEE*

Abstract

Probing signal waveforms play a central role in the signal processing performance of a MIMO radar. In practice, for a given desired beam pattern, we need to design a probing signal waveform whose beam pattern closely matches the desired one and whose auto-correlation and cross-correlation sidelobes are kept low. The latter properties are important to mitigate undesirable interference caused by multiple targets or scatterers. In this paper, we present an efficient optimization method to design a constant modulus probing signal which can synthesize a desired beam pattern while maximally suppressing both the auto-correlation and cross-correlation sidelobes at/between given spacial angles. We formulate this problem as an unconstrained minimization of a fourth order trigonometric polynomial and propose an efficient quasi-Newton iterative algorithm to solve it. Besides, we give an analysis of the local minima of the fourth order trigonometric polynomial and prove that any local minima is a 1/2-approximation of its global optimal solution. Numerical examples show that the proposed approach compares favorably with the existing approach.

Index Terms

Multiple-input multiple-output (MIMO) radar, Beam pattern, Constant modulus probing signal, Spacial auto-/cross-correlation.

*State Key Laboratory on Integrated Service Networks, Xidian University, Xi'an, 710071, China. Y.-C Wang started this work while he was a postdoctoral fellow at the University of Minnesota. (e-mail: ychwang@mail.xidian.edu.cn)

† National Lab. of Radar Signal Processing, Xidian University, Xi'an, 710071, China.

‡ Department of Electrical Computer Science Engineering, University of Minnesota, Minneapolis, 55455, USA. (e-mail: luozq@umn.edu)

I. INTRODUCTION

Multi-input and multi-output (MIMO) transmission and reception is a promising paradigm for the next generation radar systems [1] - [4]. Unlike the phased-array radar, a MIMO radar allows independent probing signals at different antennas. Through this additional diversity, a MIMO radar can deliver a higher detection performance and a better spatial resolution. In particular, if the transmitters are widely separated in space (called the uncollocated configuration), then each of them can provide an independent view of the target, which, when appropriately combined, will improve detection performance [5] - [7]. Alternatively, if the antennas are placed in close proximity, different probing signals from various collocated transmitters can generate various desired beam patterns, leading to an improved directional resolution. Besides, the collocated MIMO radar has other advantages such as interference rejection capability [8] - [10]. In this paper, we focus on the collocated system.

A central signal processing challenge in MIMO radar research is to design probing signals that are constant modulus [11] and satisfy certain beam pattern specifications. The existing design approaches can be classified into three categories: (1) maximizing the mutual information between the received signal and the impulse response of the target [12] - [14]; (2) optimizing the range, angular, and doppler resolution based on radar ambiguity function [15] - [17]; (3) matching a desired beam pattern using independent constant modulus signals while suppressing the spacial auto-correlation and cross-correlation sidelobes [18]- [20].

This paper considers the third design approach for MIMO waveform design. In [18], Fuhrmannn showed how to create high directionality or omni-directionality beam pattern through waveform covariance matrix $\mathbf{R}$. In [19], Stoica exploited semidefinite programming technique to design $\mathbf{R}$, where suppressing spacial cross-correlation levels at temporal delay zero was considered. In [20], Ahmed proposed two algorithms to design constant modulus waveforms, which satisfy the specifications of $\mathbf{R}$ in [19]. In [21], Li proposed a cyclic algorithm to synthesize constant modulus signals as well as pursuing the desired auto-/cross-correlation characteristics. In [22], an alternating approach is proposed for jointly optimizing constant modulus probing signal and receiver filter bank.

Different from the existing approaches [18] - [22], we propose to optimize probing signal waveforms to meet the beam pattern specification directly. We formulate this problem as an unconstrained fourth-order trigonometric polynomial minimization model and propose a quasi-Newton iterative algorithm to solve it approximately. Simulation results demonstrate that the resulting design procedure compares favorably

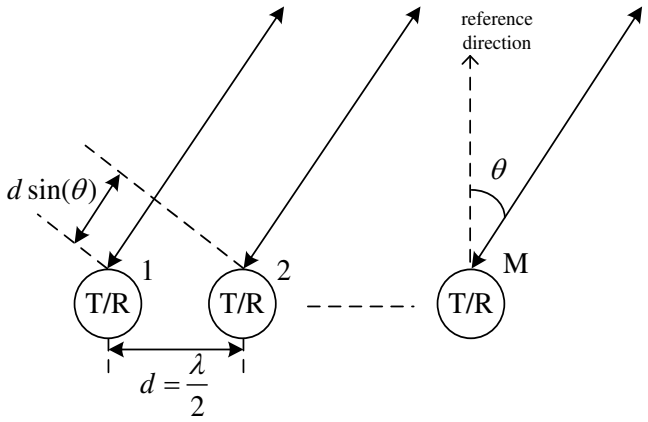


Fig. 1. MIMO radar transceiver equipped with M antennas (uniform linear array and half wavelength inter-element spacing, $\lambda/2$). θ is the spacial direction of interest.

with the existing approach in terms of both the algorithm speed and the quality of the obtained waveforms.

II. Problem Formulation and Analysis

A. System parameters

Consider a MIMO radar equipped with M transmitting antennas as shown in figure 1. The probing signal matrix $\mathbf{X} \in \mathbb{C}^{L\times M}$ and the steering vector $\mathbf{a} \in \mathbb{C}^{M}$ are

$$\mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1M} \\ \vdots & \ddots & \vdots \\ x_{L1} & \cdots & x_{LM} \end{bmatrix}, \tag{1}$$

$$\mathbf{a}_\theta = \left[1 \; e^{j\pi\sin(\theta)} \; \cdots \; e^{j\pi(M-1)\sin(\theta)}\right]^T, \tag{2}$$

where θ belongs to an angle set Θ representing the spacial direction and L is the temporal length of the probing signal. The beam pattern, which describes the power distribution of the probing signals in the spacial domain, is defined as

$$P(\theta) = \mathbf{a}_\theta^H \mathbf{X}^H \mathbf{X} \mathbf{a}_\theta. \tag{3}$$

For any integers d, ℓ with $0 < d < L$, $-d \leq \ell \leq d$, we define an $L \times L$ matrix

$$S_\ell = \begin{bmatrix} & \overbrace{0 \quad \cdots \quad 0 \quad 1}^{\ell \text{ zeros}} & \mathbf{0} & \\ & & \ddots & \\ & & & 1 \\ & \mathbf{0} & & \end{bmatrix}.$$

Then the time delayed probing signal matrix can be described as $\mathbf{S}_\ell^H \mathbf{X}$, and the spacial auto-/cross-correlation functions for the probing signal $\mathbf{X}$ and its delayed version become

$$P_{\mathrm{c}}(\ell, \theta_i, \theta_j) = \mathbf{a}_{\theta_i}^H \mathbf{X}^H \mathbf{S}_\ell \mathbf{X} \mathbf{a}_{\theta_j}, \tag{4}$$

where $\theta_i, \theta_j \in \hat{\Theta} = \{\theta_1, ..., \theta_K\}$ and $\hat{\Theta}$ is an angle set of beam pattern locations. Naturally, $\hat{\Theta} \subset \Theta$. If $\theta_i = \theta_j = \theta_k$, we denote $P_{\mathrm{c}}(\ell, \theta_k, \theta_k)$ by

$$P_{\mathrm{ac}}(\ell, \theta_k) = \mathbf{a}_{\theta_k}^H \mathbf{X}^H \mathbf{S}_\ell \mathbf{X} \mathbf{a}_{\theta_k}, \tag{5}$$

which is the spacial auto-correlation function. If $\theta_i \neq \theta_j$, the $P_{\mathrm{c}}(\ell, \theta_k, \theta_k)$ becomes the spacial cross-correlation function which we denote by

$$P_{\mathrm{cc}}(\ell, \theta_i, \theta_j) = \mathbf{a}_{\theta_i}^H \mathbf{X}^H \mathbf{S}_\ell \mathbf{X} \mathbf{a}_{\theta_j}. \tag{6}$$

B. Model formulation

To maximize the efficiency in a MIMO radar system, power amplifiers typically have to operate in the saturation mode. As a result, it is important that the probing signals for a MIMO radar have constant modulus, which mathematically means

$$|x_{\ell m}| = 1, \ell = 1, \cdots, L; m = 1, \cdots, M.$$

Because the beam pattern (3) describes the spacial power distribution of probing signals, we can specify a desired beam pattern that focuses the signal power along the directions of interest. This can effectively reduce the clutters' impact and extend the detection distance. Moreover, we can use the desired beam

pattern to set an upper bound for the auto-correlation sidelobes ($P_{\rm ac}(\ell,\theta_k),\ \ell\neq 0$) and cross-correlation levels ($P_{\rm cc}(\ell,\theta_i,\theta_j)$). The former aims to minimize the effects of clutters and the latter decrease the interference between signals from different directions. For these reasons, we propose the following optimization model

$$\min_{\alpha,\mathbf{X}} \quad w_{\rm b}^2 e_{\rm b}(\alpha,\mathbf{X}) + w_{\rm ac}^2 e_{\rm ac}(\mathbf{X}) + w_{\rm cc}^2 e_{\rm cc}(\mathbf{X}), \tag{7a}$$

$$s.t. \quad |x_{ij}| = 1, \quad i=1,\cdots,L; \quad j=1,\cdots,M, \tag{7b}$$

where

$$e_{\rm b}(\alpha,\mathbf{X}) = \sum_{\theta\in\Theta} |\alpha\mathbf{p}(\theta) - \mathbf{a}_\theta^H\mathbf{X}^H\mathbf{X}\mathbf{a}_\theta|^2, \tag{8a}$$

$$e_{\rm ac}(\mathbf{X}) = \sum_{\ell=1}^{d}\sum_{\theta_k\in\hat{\Theta}} |\mathbf{a}_{\theta_k}^H\mathbf{X}^H\mathbf{S}_\ell\mathbf{X}\mathbf{a}_{\theta_k}|^2, \tag{8b}$$

$$e_{\rm cc}(\mathbf{X}) = \sum_{\ell=0}^{d}\sum_{\substack{\theta_i\neq\theta_j\\ \theta_i,\theta_j\in\hat{\Theta}}} |\mathbf{a}_{\theta_i}^H\mathbf{X}^H\mathbf{S}_\ell\mathbf{X}\mathbf{a}_{\theta_j}|^2. \tag{8c}$$

In the optimization model (7), the term $e_{\rm b}$ in the objective function captures the beam pattern mismatching error, while the minimization of the other two terms $e_{\rm ac}$, $e_{\rm ac}$ represents the suppression of the auto-/cross-correlation sidelobe levels, respectively. More specifically, in the definition of $e_{\rm b}$ (cf. (8)), α is an unknown scaling factor to be optimized, $\mathbf{p}(\theta)$ is the desired beam pattern. The weights $w_{\rm b}$, $w_{\rm ac}$ and $w_{\rm cc}$ are positive (chosen by the user), which can make trade-off of performances amongst matching desired beam pattern and suppressing auto-/cross- correlation sidelobe levels. Parameter d denotes the considered maximum temporal delay. Because $P_{\rm ac}^*(-\ell,\theta_k) = P_{\rm ac}(\ell,\theta_k)$ and $P_{\rm cc}^*(-\ell,\theta_i,\theta_j) = P_{\rm ac}(\ell,\theta_j,\theta_i)$, correlation measures for $\ell<0$ are not included.

The optimization model (7) involves minimizing a nonconvex fourth-order polynomial with some nonlinear equality constraints, which is numerically difficult handle. Notice that the constant modulus constraints are equivalent to every entry of $\mathbf{X}$ lying on the unit circle, i.e., $x_{\ell m} = e^{j\phi_{\ell m}}$. Using $\phi_{\ell m}$ as optimization variables and writing $\mathbf{X}$ as $\mathbf{X}(\boldsymbol{\phi})$ where $\boldsymbol{\phi}$ is an $L\times M$ real matrix, we can drop the constant modulus constraints and formulate (7) as an unconstrained fourth order trigonometric polynomial minimization problem

$$f(\alpha,\boldsymbol{\phi}) = w_{\rm b}^2 e_{\rm b}(\alpha,\mathbf{X}(\boldsymbol{\phi})) + w_{\rm ac}^2 e_{\rm ac}(\mathbf{X}(\boldsymbol{\phi})) + w_{\rm cc}^2 e_{\rm cc}(\mathbf{X}(\boldsymbol{\phi})). \tag{9}$$

Then we obtain

$$\min_{\alpha,\boldsymbol{\phi}} \; f(\alpha, \boldsymbol{\phi}), \tag{10}$$

where both α and each entry $\phi_{\ell m} \in \boldsymbol{\phi}$ are real valued variables.

Like (7), the unconstrained optimization model (10) is still nonconvex. However, the unconstrained formulation makes the problem amenable to the use of L-BFGS type iterative procedures which can implemented efficiently. Moreover, the unconstrained formulation has a strong property that every local minimum is a 1/2-approximation of the global minimum.

C. Analysis of local minima

Computing the global optimal solution of the minimization problem, either (7) or (10), is difficult due to the nonconvexity of the objective function. Following the analysis of nonconvex quadratic minimization problem in [23], we show below that any local minima of a 4-th order trigonometric polynomial (11) is a $1/2$-approximation of its global minimum. In particular, let us consider the following 4-th order trigonometric polynomial function minimization problem

$$\min_{\boldsymbol{\phi}} \quad f(\boldsymbol{\phi}) = \frac{1}{2}(g(e^{-j\boldsymbol{\phi}}) - c)^2 \tag{11}$$

where $g(\mathbf{x}) = \mathbf{x}^H\mathbf{M}\mathbf{x}$ is a quadratic function of $\mathbf{x}$ and $\mathbf{M}$ is a Hermitian matrix, c is a real-valued number. This optimization problem (11) is clearly in the form of the unconstrained minimization problem (10). So analyzing the local minima of (11) can provide useful insight on the quality of optimal design of constant modulus waveforms based on (10).

Lemma: Let $\hat{\boldsymbol{\phi}}$ be a local minimizer of (11). Then $\hat{\boldsymbol{\phi}}$ is a $\frac{1}{2}$-approximation of the global minimum of (11) in the sense that

$$\frac{f(\hat{\boldsymbol{\phi}}) - f_{\min}}{f_{\max} - f_{\min}} \leq \frac{1}{2}, \tag{12}$$

where $f_{\min}$ and $f_{\max}$ are respectively the global minimum and global maximum value of (11).

Proof: Let $\hat{\boldsymbol{\phi}}$ be a local minimizer of (11). Then

$$0 = \nabla f(\hat{\boldsymbol{\phi}}) = -j(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\nabla g(e^{-j\hat{\boldsymbol{\phi}}}).$$

If $g(e^{-j\hat{\boldsymbol{\phi}}}) - c = 0$, then clearly $\hat{\boldsymbol{\phi}}$ is a global minimizer, so (12) trivially. Now consider the case where $g(e^{-j\hat{\boldsymbol{\phi}}}) - c \neq 0$, in which case we have $\nabla g(e^{-j\hat{\boldsymbol{\phi}}}) = 0$. Then the second order necessary optimality

condition implies that

$$\begin{aligned} 0 &\preceq \nabla^2 f(\hat{\boldsymbol{\phi}}) \\ &= (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\left[\nabla^2 g(e^{-j\hat{\boldsymbol{\phi}}})\right] + \nabla g(e^{-j\hat{\boldsymbol{\phi}}})(\nabla g(e^{-j\hat{\boldsymbol{\phi}}}))^T \\ &= (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\left[\nabla^2 g(e^{-j\hat{\boldsymbol{\phi}}})\right], \end{aligned}$$

where we have used chain rule and the property $\nabla g(e^{-j\hat{\boldsymbol{\phi}}}) = 0$. Let $\mathbf{x} \in \mathbb{C}^n$ be any vector with $|\mathbf{x}_i| = 1$ for all i. The proof of [23, Thm. 3] implies

$$\begin{aligned} (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\left(\mathbf{x}^H\mathbf{M}\mathbf{x} + (\mathbf{x}^*)^H\mathbf{M}(\mathbf{x}^*) - 2g(e^{-j\hat{\boldsymbol{\phi}}})\right) \geq 0, \\ \mathbf{x} \in \mathbb{C}^n \text{ with } |x_i| = 1,\ \forall\ i, \end{aligned}$$

where $\mathbf{x}^*$ denotes the complex conjugation of $\mathbf{x}$. Specializing this inequality to the global minimizer $\tilde{\mathbf{x}} = e^{-j\tilde{\boldsymbol{\phi}}} = \underset{\boldsymbol{\phi}}{\operatorname{argmin}} f(\boldsymbol{\phi})$. Then we obtain

$$(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\left(g(\tilde{\mathbf{x}}) + g(\tilde{\mathbf{x}}^*) - 2g(e^{-j\hat{\boldsymbol{\phi}}})\right) \geq 0.$$

Since the left hand side can be written as

$$(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\,(g(\tilde{\mathbf{x}}) + g(\tilde{\mathbf{x}}^*) - 2c) - 2(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2,$$

it follows that

$$2(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2 \leq (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)\,(g(\tilde{\mathbf{x}}) + g(\tilde{\mathbf{x}}^*) - 2c)\,.$$

This further implies

$$\begin{aligned} &4(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^4 \\ &\leq (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2(g(\tilde{\mathbf{x}}) + g(\tilde{\mathbf{x}}^*) - 2c)^2 \\ &\leq 2(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2\left((g(\tilde{\mathbf{x}}) - c)^2 + (g(\tilde{\mathbf{x}}^*) - c)^2\right) \\ &\leq 2(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2\,(f_{\min} + f_{\max}) \end{aligned}$$

where we have used $(g(\tilde{\mathbf{x}}) - c)^2 = f_{\min}$ and $(g(\tilde{\mathbf{x}}^*) - c)^2 \leq f_{\max}$. Canceling the factor $2(g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2$ yields

$$f(\hat{\boldsymbol{\phi}}) = (g(e^{-j\hat{\boldsymbol{\phi}}}) - c)^2 \leq (f_{\min} + f_{\max})\,/2,$$

which establishes (12).

III. Implementing a Quasi-Newton Method

The unconstrained minimization model (10) can be solved, approximately but effectively, using a quasi Newton algorithm, for example L-BFGS (Limited-Memory Broyden Fletcher Goldfarb and Shanno algorithm, see the details in [25]). The outline of this algorithm is given in Figure 2.

In the L-BFGS algorithm, the computational cost of each L-BFGS iteration is dominated by the evaluation of $f(\alpha, \boldsymbol{\phi})$ and $\nabla f(\alpha, \boldsymbol{\phi})$. Direct computation of f and ∇f according to (8) is inefficient, especially for large d. In what follows, we describe an efficient method to compute $f(\alpha, \boldsymbol{\phi})$ and $\nabla f(\alpha, \boldsymbol{\phi})$.

Define

$$\mathbf{b}_{\theta\theta} = \mathrm{vec}(\mathbf{a}_\theta \mathbf{a}_\theta^H), \quad A_1 = \sum_{\theta\in\Theta} \mathbf{p}(\theta),$$
$$\mathbf{A}_2 = -\sum_{\theta\in\Theta} \mathbf{p}(\theta)\mathbf{b}_{\theta\theta}, \quad \mathbf{A}_3 = \sum_{\theta\in\Theta} \mathbf{b}_{\theta\theta}\mathbf{b}_{\theta\theta}^H.$$

where $\mathrm{vec}(\cdot)$ vectorizes a matrix by stacking its columns on top of one another. We can rewrite the objective function in the matrix form. For $e_{\mathrm{b}}(\alpha, \mathbf{X}(\boldsymbol{\phi}))$, we can derive it as

Initialize the number of L-BFGS updates, m, and the iteration index $k = 1$. Choose the initial point $(\alpha, \boldsymbol{\phi})_1$ and compute the corresponding f_1, ∇f_1. Set the initial search direction $\mathbf{g}_1 = -\nabla f_1$. To be simple, we denote $(\alpha, \boldsymbol{\phi})$ by $\mathbf{x}$.

Repeat

1 Use f_k, ∇f_k, and $\mathbf{g}_k$ to determine the step length μ_k satisfying the specific line search rule.
2 Compute $\mathbf{s}_k = \mu_k \mathbf{g}_k$ and $\mathbf{x}_{k+1} = \mathbf{x}_k + \mu_k \mathbf{g}_k$.
3 Compute ∇f_{k+1} and $\mathbf{y}_k = \nabla f_{k+1} - \nabla f_k$.
4 Let $\mathbf{q} = \nabla f_{k+1}$.
5 **for** $i = k, k-1, \cdots, k-m+1$
$t_i = \frac{\mathbf{s}_i^H \mathbf{q}}{\mathbf{y}_i^H \mathbf{s}_i}$, $\mathbf{q} = \mathbf{q} - t_i \mathbf{y}_i$.
end (for)
6 $\mathbf{r} = \frac{\mathbf{s}_k^H \mathbf{y}_k}{\mathbf{y}_k^H \mathbf{y}_k}\mathbf{q}$.
7 **for** $i = k-m+1, k-m+2, \cdots, k$
$\beta = \frac{\mathbf{y}_i^H \mathbf{r}}{\mathbf{y}_i^H \mathbf{s}_i}$, $\mathbf{r} = \mathbf{r} + (t_i - \beta)\mathbf{s}_i$.
end (for)
8 $\mathbf{g}_{k+1} = -\mathbf{r}$.
9 $k = k + 1$.

until a pre-set termination condition is satisfied.

Fig. 2. Outline of L-BFGS algorithm.

$$w_b^2 e_{\mathrm{b}}(\alpha, \mathbf{X}(\boldsymbol{\phi})) = \mathbf{v}(\alpha, \boldsymbol{\phi})^H \mathbf{Q} \mathbf{v}(\alpha, \boldsymbol{\phi}), \tag{13}$$

where

$$\mathbf{v}(\alpha, \boldsymbol{\phi}) = \begin{bmatrix} \alpha \\ \mathrm{vec}(\mathbf{X}(\boldsymbol{\phi})^H \mathbf{X}(\boldsymbol{\phi})) \end{bmatrix},$$
$$\mathbf{Q} = w_b^2 \begin{bmatrix} A_1 & \mathbf{A}_2^H \\ \mathbf{A}_2 & \mathbf{A}_3 \end{bmatrix},$$

and superscript $(\cdot)^H$ denotes conjugate transpose. Let $\mathbf{c}_{\theta_k}(\boldsymbol{\phi}) = \mathbf{X}(\boldsymbol{\phi})\mathbf{a}_{\theta_k}$ and $\bar{\mathbf{c}}_{\theta_k}$ is obtained by reversing the order of the entries of $\mathbf{c}_{\theta_k}$. We have

$$\begin{aligned} &\sum_{\ell=-L+1}^{L-1} \sum_{\theta_k \in \theta} |\mathbf{a}_{\theta_k}^H \mathbf{X}(\boldsymbol{\phi})^H \mathbf{S}_\ell \mathbf{X}(\boldsymbol{\phi}) \mathbf{a}_{\theta_k}|^2 \\ &= \sum_{\theta_k \in \theta} \sum_{\ell=-L+1}^{L-1} |\mathbf{c}_{\theta_k}(\boldsymbol{\phi})^H \mathbf{S}_\ell \mathbf{c}_{\theta_k}(\boldsymbol{\phi})|^2 \\ &= \sum_{\theta_k \in \theta} \|\mathbf{c}_{\theta_k}^*(\boldsymbol{\phi}) \otimes \bar{\mathbf{c}}_{\theta_k}(\boldsymbol{\phi})\|_2^2, \end{aligned} \tag{14}$$

where the superscript $(\cdot)^*$ denotes complex conjugation, $\otimes$ is the convolution operator, and $\|\cdot\|$ is the 2-norm operator. Truncating $\mathbf{c}_{\theta_k}^*(\boldsymbol{\phi}) \otimes \bar{\mathbf{c}}_{\theta_k}(\boldsymbol{\phi})$ from 1 to d and denoting it as $\mathbf{d}_{\theta_k\theta_k}(\boldsymbol{\phi})$, we obtain

$$e_{\mathrm{ac}}(\mathbf{X}(\boldsymbol{\phi})) = w_{ac}^2 \sum_{\theta_k \in \Theta} \|\mathbf{d}_{\theta_k\theta_k}(\boldsymbol{\phi})\|_2^2. \tag{15}$$

Through similar derivations, we get

$$e_{\mathrm{cc}}(\mathbf{X}(\boldsymbol{\phi})) = w_{cc}^2 \sum_{\theta_i, \theta_j \in \Theta} \|\mathbf{d}_{\theta_i\theta_j}(\boldsymbol{\phi})\|_2^2. \tag{16}$$

Combining (13), (15), and (16), we obtain the following reformulation of $f(\alpha, \boldsymbol{\phi})$ as

$$f(\alpha, \boldsymbol{\phi}) = \mathbf{v}(\alpha, \boldsymbol{\phi})^H \mathbf{Q} \mathbf{v}(\alpha, \boldsymbol{\phi}) + u(\boldsymbol{\phi}), \tag{17}$$

where

$$u(\boldsymbol{\phi}) = w_{ac}^2 \sum_{\theta_k \in \Theta} \|\mathbf{d}_{\theta_k\theta_k}(\boldsymbol{\phi})\|_2^2 + w_{cc}^2 \sum_{\substack{\theta_i \neq \theta_j \\ \theta_i, \theta_j \in \theta}} \|\mathbf{d}_{\theta_i\theta_j}(\boldsymbol{\phi})\|_2^2.$$

According to (17), we can compute $\nabla f(\alpha, \boldsymbol{\phi})$ as follows

$$\frac{\partial f(\alpha,\boldsymbol{\phi})}{\partial \alpha} = 2\mathrm{Re}\left([1 \quad \mathbf{0}^T]\mathbf{Q}\mathbf{v}(\alpha,\boldsymbol{\phi})\right), \tag{18a}$$

$$\frac{\partial f(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}} = 2\mathrm{Re}\left(\left[\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}\right]^H \mathbf{Q}\mathbf{v}(\alpha,\boldsymbol{\phi})\right) + \frac{\partial u(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}, \tag{18b}$$

where the entries in $\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ and $\frac{\partial g(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ are

$$\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \phi_{lm}} = \begin{bmatrix} 0 & \dfrac{\partial \mathrm{vec}(\mathbf{X}(\boldsymbol{\phi})\mathbf{X}(\boldsymbol{\phi})^H)}{\partial \phi_{lm}} \end{bmatrix}^T, \tag{19}$$

$$\begin{aligned} \frac{\partial u(\boldsymbol{\phi})}{\partial \phi_{lm}} =& w_{ac}^2 \sum_{\theta_k \in \theta} 2\mathrm{Re}\left(\mathbf{d}_{\theta_k\theta_k}^H \frac{\partial \mathbf{d}_{\theta_k\theta_k}}{\partial \phi_{lm}}\right) \\ &+ w_{cc}^2 \sum_{\substack{\theta_i \neq \theta_j \\ \theta_i,\theta_j \in \theta}} 2\mathrm{Re}\left(\mathbf{d}_{\theta_i\theta_j}^H \frac{\partial \mathbf{d}_{\theta_i\theta_j}}{\partial \phi_{lm}}\right). \end{aligned} \tag{20}$$

Exploiting special structures such as convolution and sparsity, we can compute f and ∇f efficiently. Let us first consider the computation of $f(\alpha, \boldsymbol{\phi})$. Recall that the entries of $\mathbf{X}(\boldsymbol{\phi})$ can be written as $e^{\phi_{\ell m}}$ for some angle $\phi_{\ell m}$. Exploiting this structure, we can write the entries of $\mathbf{X}(\boldsymbol{\phi})^H\mathbf{X}(\boldsymbol{\phi})$ as $\sum_{\ell=1}^{L} e^{j(\phi_{k\ell}-\phi_{\ell i})}$, where $i, k = 1, \cdots, m$ are the row and column indices respectively. Because the value of $e^{j\phi_{lm}}$ can be obtained by table look-up, computing $\mathbf{X}(\boldsymbol{\phi})^H\mathbf{X}(\boldsymbol{\phi})$ or $\mathbf{v}(\alpha, \boldsymbol{\phi})$ can be free of multiplication operations. Similarly, both $\mathbf{c}_{\theta_i} = \mathbf{X}(\boldsymbol{\phi})\mathbf{a}_{\theta_i}$ and $\mathbf{c}_{\theta_j} = \mathbf{X}(\boldsymbol{\phi})\mathbf{a}_{\theta_i}$ are also free of complex multiplications. Also, since $\mathbf{v}(\alpha, \boldsymbol{\phi}) \in \mathbb{C}^{LM+1}$, $\mathbf{Q} \in \mathbb{C}^{(M^2+1)\times(M^2+1)}$, we can compute $\mathbf{v}(\alpha, \boldsymbol{\phi})^H\mathbf{Q}\mathbf{v}(\alpha, \boldsymbol{\phi})$ using $(M^2+1)^2 + M^2 + 1$ complex multiplications. If we use FFT to evaluate $\mathbf{d}_{\theta_i\theta_j} = \mathbf{c}_{\theta_i} \otimes \bar{\mathbf{c}}_{\theta_j}$ $(i, j = 1, \cdots, K)$ and notice that the size of θ is K, all of the terms $\mathbf{d}_{\theta_i\theta_j}$ $(i, j = 1, \cdots, K)$ can be computed with $3K^2(2L-1)\log_2(2L-1)$ complex multiplications. Since $d < L$, we conclude that the total number of complex multiplications required to compute $f(\alpha, \boldsymbol{\phi})$ is $\mathcal{O}(M^4 + 6K^2L\log_2 2L)$.

As for $\nabla f(\alpha, \boldsymbol{\phi})$, $\frac{\partial f(\alpha,\boldsymbol{\phi})}{\partial \alpha}$ can be obtained for free from the computation of $\frac{\partial f(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$, which consists of two terms, $(\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}})^H\mathbf{Q}\mathbf{v}(\alpha, \boldsymbol{\phi})$ and $\frac{\partial g(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$. For the first term, $\mathbf{Q}\mathbf{v}(\alpha, \boldsymbol{\phi})$ has been obtained during the computation of $f(\alpha, \boldsymbol{\phi})$. Notice that $\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \phi_{lm}}$ has only $2M-1$ none-zero entries, so it takes $LM(2M-1)$ complex multiplications to obtain $(\frac{\partial \mathbf{v}(\alpha,\boldsymbol{\phi})}{\partial \boldsymbol{\phi}})^H\mathbf{Q}\mathbf{v}(\alpha, \boldsymbol{\phi})$. In the second term $\frac{\partial u(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$, $\mathbf{d}_{\theta_i\theta_j}$ has been obtained during the computation of $f(\alpha, \boldsymbol{\phi})$. Its derivative, $\frac{\partial \mathbf{d}_{\theta_i\theta_j}}{\partial \phi_{lm}}$, can be computed by truncating $\mathbf{c}_{\theta_i} \otimes \frac{\partial \bar{\mathbf{c}}_{\theta_j}}{\partial \phi_{lm}} + \bar{\mathbf{c}}_{\theta_j} \otimes \frac{\partial \mathbf{c}_{\theta_i}}{\partial \phi_{lm}}$. Because $\frac{\partial \bar{\mathbf{c}}_{\theta_k}}{\partial \phi_{lm}}$ or $\frac{\partial \mathbf{c}_{\theta_k}}{\partial \phi_{lm}}$ has only one none-zero constant modulus entry, $\frac{\partial \mathbf{d}_{\theta_i\theta_j}}{\partial \phi_{lm}}$ can be implemented using only additions. For $\mathbf{d}_{\theta_i\theta_j}^H \frac{\partial \mathbf{d}_{\theta_i\theta_j}}{\partial \phi_{lm}}$, it takes $\mathcal{O}(d)$ complex multiplications. So $\frac{\partial u(\boldsymbol{\phi})}{\partial \boldsymbol{\phi}}$ requires $\mathcal{O}(K^2LMd)$

complex multiplications. In all, the total number of complex multiplications to compute $\nabla f(\alpha, \boldsymbol{\phi})$ is $\mathcal{O}(2LM^2 + K^2LMd)$.

When compared with the original expression of $f(\alpha, \phi)$ (8)-(9), the new expression (17) offers a significant computational saving. In particular, the matrix $\mathbf{Q}$ in (17) can be obtained in advance and used in all iterations. In this way, we avoid computing $\mathbf{Q}$ for each $\theta \in \Theta$ as would have been the case for the original expression (8)-(9). This is a major advantage because the size of angle set Θ, i.e., $|\Theta|$, is extremely large in practice. Another source of efficiency is that by formulating $e_{\rm ac}(\mathbf{X}(\boldsymbol{\phi}))$ and $e_{\rm cc}(\mathbf{X}(\boldsymbol{\phi}))$ as a convolution, we can use FFT algorithm to compute them efficiently. In this way, the value of the temporal lag d takes no effect on the complexity required to compute $f(\alpha, \phi)$. Finally, when using (9) to compute ∇f directly, we recognize that f is of the form $f(\mathbf{X}) = \|\mathbf{C} - \mathbf{A}\mathbf{X}\mathbf{B}\mathbf{X}^H\mathbf{A}^H\|_F^2$, so that $\nabla f(\mathbf{X}) = -\mathbf{A}^H\mathbf{C}^H\mathbf{A}\mathbf{X}\mathbf{B} - \mathbf{A}^H\mathbf{C}\mathbf{A}\mathbf{X}\mathbf{B}^H + 2\mathbf{A}^H\mathbf{A}\mathbf{X}\mathbf{B}\mathbf{X}^H\mathbf{A}^H\mathbf{A}\mathbf{X}\mathbf{B}^H + 2\mathbf{A}^H\mathbf{A}\mathbf{X}\mathbf{B}^H\mathbf{X}^H\mathbf{A}^H\mathbf{A}\mathbf{X}\mathbf{B}$, where $\|\cdot\|_F$ is the Frobenius norm operator. Table I shows a rough comparison of the complexity to compute f and ∇f using (8)-(9) and (17).

TABLE I
COMPARISON OF THE COMPLEXITY TO COMPUTE f AND ∇f

	Original expression (8)-(9)	New expression (17)
$f(\alpha, \boldsymbol{\phi})$	$\mathcal{O}(\lvert\Theta\rvert L + K^2 dL)$	$\mathcal{O}(M^4 + 6K^2L\log_2 2L)$
$\nabla f(\alpha, \boldsymbol{\phi})$	$\mathcal{O}(3\lvert\Theta\rvert LM + 2K^2LMd)$	$\mathcal{O}(2LM^2 + K^2LMd)$

IV. NUMERICAL RESULTS

We present several numerical examples to illustrate the performance of the proposed, L-BFGS, approach. Consider MIMO radar system equipped with uniform linear antennas and inter-element spacing of half a wavelength. The antenna number, waveform length, and temporal delay are denoted by M, L, and d, respectively. Their pre-set values can be found in the corresponding simulations. The angle set Θ covers $(-90^o,\ 90^o]$ with spacing 0.1^o, and two interested targets are located in the directions of $\theta_1 = -40^o$ and $\theta_2 = 30^0$, i.e., $\hat{\Theta} = \{-40^o, 30^o\}$, and the desired beam pattern is

$$\mathbf{p}(\theta) = \begin{cases} 1, & \theta \in [\theta_k - 10^o, \theta_k + 10^o],\ k = 1, 2, \\ 0, & \text{otherwise.} \end{cases}$$

Weights $w_{\rm b}$, $w_{\rm ac}$, $w_{\rm cc}$ are (1, 1, 1) or (1, 8, 8). We have implemented the L-BFGS algorithm combined with the Armijo-Goldstein stepsize rule to solve (10). The number of L-BFGS updates that we store is 5,

and the termination condition is for the maximum iteration number, 200, to be reached or $\|\Delta f(\alpha, \boldsymbol{\phi})\|_F$ is less than 10^{-4} . We have compared the L-BFGS algorithm with the approach proposed by Li et.al. in [21]. We also set a maximum iteration number of 200 as the stopping criterion for the Li approach.

The L-BFGS approach and Li approach are implemented in the MATLAB2008a/Windows XP environment on a 1.7GHz Athlon XII computer with 3GB of RAM. The Matlab package used in our simulations can be downloaded from the first author's website [24].

A. Synthesized beam pattern

Using the optimized $\mathbf{X}$, we plot the corresponding beam pattern in figures 3(a)-3(c) together with the desired beam pattern. Simulation parameters are shown in the caption of these figures. Table II lists the mean-squared error (MSE) between the desired beam pattern and the synthesized beam pattern. We can see that both the Li approach and the L-BFGS approach do a good job in approximating the desired beam pattern. We also see that increasing antenna number can result in better approximation.

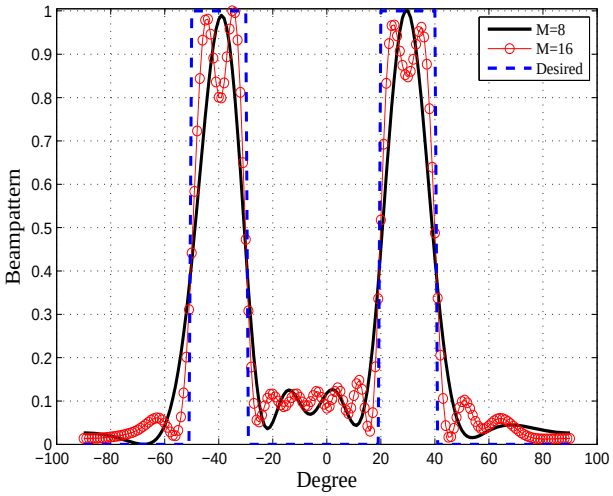
TABLE II
MSE COMPARISON

	L-BFGS Approach				Li approach	
(M, L, d)	(8, 128, 8)		(16, 128, 8)		(8, 128, 8)	(16, 128, 8)
$(w_{\rm b}, w_{\rm ac}, w_{\rm cc})$[1]	(1, 1, 1)	(1, 8, 8)	(1, 1, 1)	(1, 8, 8)	⁄	⁄
MSE	2.94×10^{-2}	3.01×10^{-2}	1.56×10^{-2}	1.61×10^{-2}	2.95×10^{-2}	1.57×10^{-2}

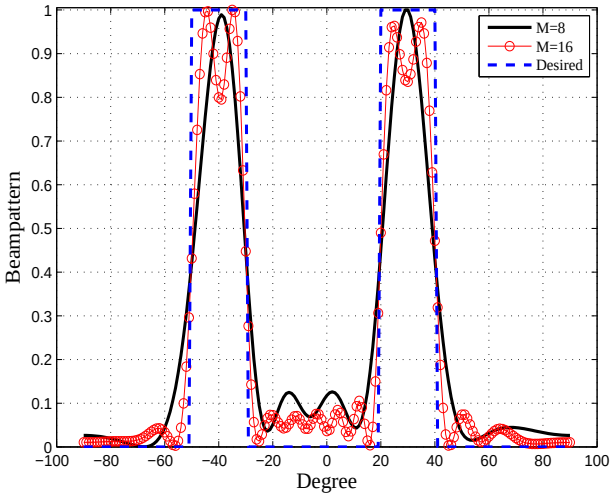
In the Li approach [21], the settings for $w_{\rm b}$, $w_{\rm ac}$, $w_{\rm cc}$ are fixed as (1, 1, 1) and there is no scheme to adjust them.

B. Spacial correlation characteristics

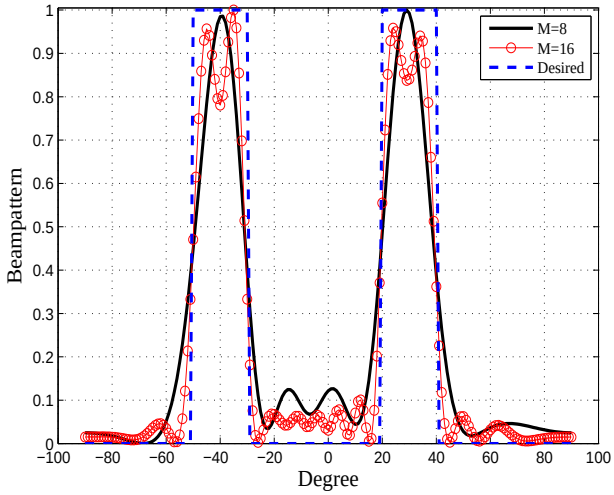
Figure 4 shows the normalized spacial auto-correlation characteristics. From the figure, we see that auto-correlation levels are symmetric with respect to the origin. This means that only taking $\ell > 0$ into account are sufficient for suppressing auto-correlation sidelobes. It can be also seen that, for the L-BFGS approach and the Li approach, increasing L can lower auto-correlation sidelobe levels while increasing delay parameter d tends to drive up the sidelobe levels. This is reasonable because a large L or small d means more degrees of design freedom. Furthermore, we can see that increasing weights $\omega_{\rm ac}$ and $\omega_{\rm cc}$ can also lower auto-correlation sidelobe levels. Comparing figure 4(a) with 4(c), we can see that the auto-correlation sidelobe levels decrease by about -40dB to -50dB for both approaches when L is increased



(a) L-BFGS approach, (1,1,1)



(b) L-BFGS approach, (1,8,8)



(c) Li approach

Fig. 3. Comparison of the synthesized beam pattern. Antenna parameters M, waveform length L, and temporal delay d are (8, 128, 8) and (16, 128, 8). Weights, (ω_{b}, ω_{ac}, ω_{cc}) for the L-BFGS approach are (1, 1, 1) and (1, 8, 8).

from 32 to 128 (d is fixed to 8). Comparing 4(a) with 4(b), we can also determine the effects of parameter d quantitatively. For example, when d is increased from 8 to 16 with a fixed $L = 32$, the sidelobe levels increase by about 10dB. Furthermore, while increasing ω_{ac} and ω_{cc} from 1 to 8, we can see that auto-correlation sidelobe levels decrease by about -30dB to -35dB with a fixed $d = 8$. From these three figures, we also see that the L-BFGS approach has a comparable or better auto-correlation performance than the Li approach.

Figure 5 shows spacial cross-correlation characteristics. We see that cross-correlation levels of $(-40^o, 30^o)$ in the section $\ell \geq 0$ is symmetric to that of $(30^o, -40^o)$ in the section $\ell \leq 0$. This verifies that, in (7), it is reasonable to only consider the cross-correlation levels for $\ell \geq 0$. Furthermore, we find parameters L and d have similar effects in figure 4, i.e., increasing L, decreasing d, or increasing ω_{ac} and ω_{cc} will lead to lower cross-correlation levels. Specifically, comparing figure 5(a) with 5(c), we have the following observations: when we increase L from 32 to 128, the L-BFGS approach and the Li approach can reduce the cross-correlation levels by about 30dB. Comparing 5(a) with 5(b), we can also determine the qualitative effects of delay parameter d and weights ω_{ac} and ω_{cc} on cross-correlation effects in both approaches.

C. Execution time

As mentioned before, the computational complexity of each iteration in the L-BFGS approach is dependent on objective function f and gradient ∇f, which are $\mathcal{O}(M^4 + 3K^2 L \log_2 2L)$ and $\mathcal{O}(2LM^2 + LMd)$, respectively. In the Li approach, the main computation, in each iteration, is to perform the singular value decomposition, the complexity of which is

$$4(d + L - 1)^2(Md) + 8(d + L - 1)(Md)^2 + 9(Md)^3,$$

or rough $\mathcal{O}((Md)^3 + L(Md)^2 + MdL^2)$.

Table II lists the execution time (averaged over 100 runs, 200 iterations) for the proposed L-BFGS approach and the Li approach [21], as a function of the antenna number M, pulse length L, and delay parameter d. The measurements provide a rough estimate of the efficiency of two approaches. While the execution time of both approaches increases with M, L, or d, the rate of increase appears more rapid for the Li approach. For example, when system parameters (M, L, d) change from $(8, 16, 8)$ to $(16, 128, 127)$, the execution time of the Li's approach grows from 2.18s to 22198s. In contrast, the L-BFGS approach executes very efficiently. Even for large parameters, such as $(16, 256, 255)$, execution time is only 162

seconds. This low execution time indicates that the L-BFGS approach is suitable for applications involving large L and d.

TABLE III
SIMULATION TIME (SECOND)

(M,L,d)	L-BFGS approach	Li Approach [21]
(8,16,8)	2.72s	2.18s
(8,32,16)	4.43s	11.16s
(8,128,16)	21.25s	19.46s
(8,128,32)	22.04s	81.65s
(8,128,127)	23.01s	3193s
(16,128,127)	62.6s	22198s
(16,256,255)	162s	/

V. CONCLUDING REMARKS

In this paper, we have considered a direct way to synthesize constant modulus waveforms for a MIMO radar. Compared to the state-of-the-art design procedures (e.g., [21]), the proposed procedure (based on the L-BFGS algorithm) offers a comparable or better performance and yet is substantially more efficient, especially for situations involving a large number of system parameters.

ACKNOWLEDGMENT

The authors appreciate the associate editor and the anonymous reviewers of this paper who help them to elevate the quality of this paper. The first author also wants to thank Dr. M. Razaviyayn in UMN, who gives the insightful discussions in the beginning of this research work.

REFERENCES

[1] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, R. Valenzuela, "MIMO Radar: An Idea Whose Time Has Come," in *Proc. 2004 IEEE Radar Conference*, Philadelphia, Pa, USA, pp. 71–78, Apr. 2004.

[2] K. Forsythe, D. Bliss, G. Fawcett, "Multiple-input multiple-output (MIMO) Radar: Performance issues," in *Prod. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, vol.1, pp. 310–315, Nov. 2004.

[3] D. W. Bliss, K. W. Forsythe, "Multi-input multiple-output (MIMO) radar and imaging: Degrees of freedom and resolution," *in Proc. 37th Asilomar Conf. Signals, Syst. and Comput.*, pp.54–59, Nov. 2003.

[4] F. Robey, S. Coutts, D. Weikle, J. McHarg, K. Cuomo, "MIMO radar theory and experimental results," in *MIMO radar theory and experimental results, in Proc. 38th Asilomar Conf. Signals, Syst. and Comput.*, pp.300–304, Nov. 2004.

[5] A. Haimovich, R. Blum, L. Cimini, "MIMO Radar with Widely Separated Antennas," *IEEE Signal Mag.*, pp. 116–129, Jan. 2008.

[6] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, R. Valenzuela, “Performance of MIMO Radar systems: Advantages of angular diversity,” in *Prod. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, vol.1, pp. 305–309, Nov. 2004.

[7] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, R. Valenzuela, “Spatial diversity in radars - models and detection performance,” *IEEE Trans. on Signal Process.*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[8] J. Li, P. Stoica, “MIMO Radar with colocated antennas,” *IEEE Signal Mag.*, vol.24, no.5, pp. 106–114, Sept. 2007.

[9] C.-Y, Chen, and P.P. Vaidyanathan, “MIMO Radar space-time adaptive processing using prolate spheroidal wave functions,” *IEEE Trans. Signal Process.*, vol.56., no.2, pp. 623–635, Feb. 2008.

[10] H. Lehmann, A. Haimovich, R. Blum, L. Cimini, “High resolution capabilities of MIMO radar,” in *Prod. 40th Asilomar Conf. Signals, Syst. Comput.*, pp.25-30, Nov. 2006.

[11] Y.-C. Wang, L. Dong, X. Xue, K-C, Yi, “On the design of constant modulus sequences with low correlation sidelobes levels,” *IEEE Commun. Letts*, vol.16., no.3, pp. 330–343, Apr. 2012.

[12] Y. Yang, R. Blum, “Radar waveform design based on mutual information and mean-square error estimation,” *IEEE Trans. Aerosp. Electron. Syst.*, vol.43., no.1, pp. 330–343, Jan.2007.

[13] B. Friedlander, “Waveform design for MIMO radars,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, pp.1227–1238, Jul. 2007.

[14] A. Leshem, O. Naparstek, A. Nehorai, “Information theoretic adaptive radar waveform design for multiple extended targets,” *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp.42–55, Jun. 2007.

[15] G. Antonio, D. Fuhrmann, and F. Robey, “MIMO radar ambiguity functions,” *IEEE J. Sel. Topics in Signal Process.*, vol.1, pp.167–177, Jun. 2007.

[16] C.-Y, Chen, P. Vaidyanathan, “MIMO Radar Ambiguity Properties and Optimization Using Frequency-Hopping Waveforms,” *IEEE Trans. Signal Process.*, vol.56., no.12, pp. 5926–5936, Dec. 2008.

[17] C.-Y, Chen, P.P. Vaidyanathan, “MIMO Radar Waveform Optimization With Prior Information of the Extended Target and Clutter,” *IEEE Trans. Signal Process.*, vol.57., no.9, pp. 3533–3544, Sept. 2009.

[18] D. R. Fuhrmann, G. S. 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.

[19] P. Stoica, J. Li, Y. Xie, “On Probing Signal Design For MIMO Radar,” *IEEE Trans. Signal Process.*, vol.55, no.8, pp. 4151–4160, Aug. 2007.

[20] S. Ahmed and J. Thompson, Y. R. Petillot and B. Mulgrew, “Unconstrained synthesis of covariance matrix for MIMO radar transmit beampattern,” *IEEE Trans. Sig. Proc.*, vol.59, no.8, pp. 3837 – 3849, Aug, 2011.

[21] J. Li, P. Stoica, X.-Y, Zheng, “Signal Synthesis and Receiver Design for MIMO Radar Imaging,” *IEEE Trans. Signal Process.*, vol.56., no.8, pp. 3959–3968, Aug. 2008.

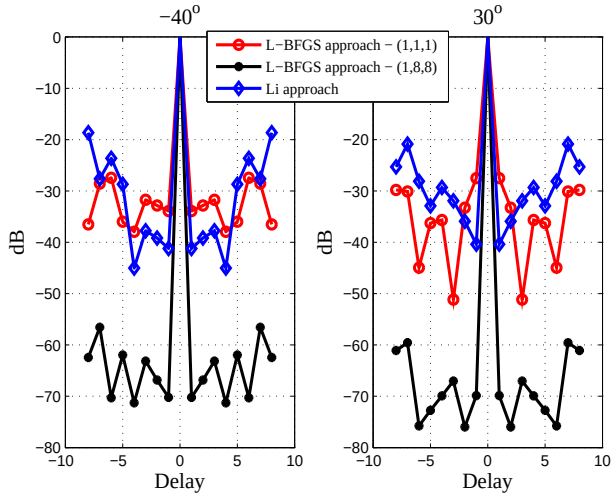
[22] Y.-C. Wang, H.-W. Liu, Z.-Q. Luo, “Iterative Design of MIMO Radar Transmit Waveforms and Receive Filter Bank,” in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal. Process.*,, Dallas, Ti, USA. pp. 2770–2773, Mar. 2010.

[23] M. Kisialiou, X.-D. Luo and Z.-Q. Luo, “Efficient Implementation of Quasi-Maximum-Likelihood Detection Based on Semidefinite Relaxation,” *IEEE Transactions on Signal Processing*, Vol. 57, No. 12, pp. 4811–4822, 2009.

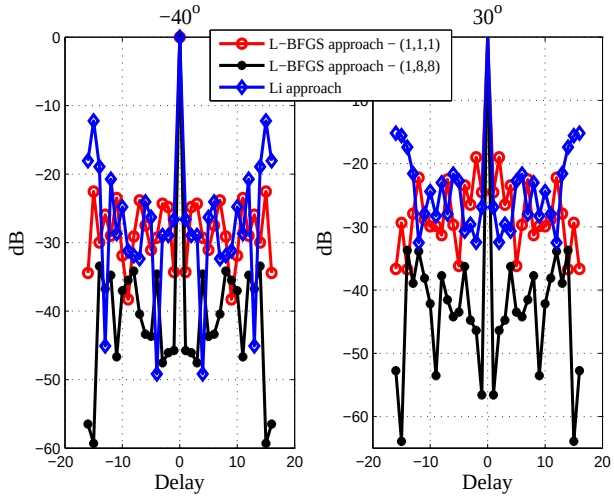
[24] http://web.ee.xidian.edu.cn/ychwang/paper.html

[25] J. Nocedal, S. Wright “Numerical Optimization,” Berlin: Springer-Verlag Press, pp.224–227.

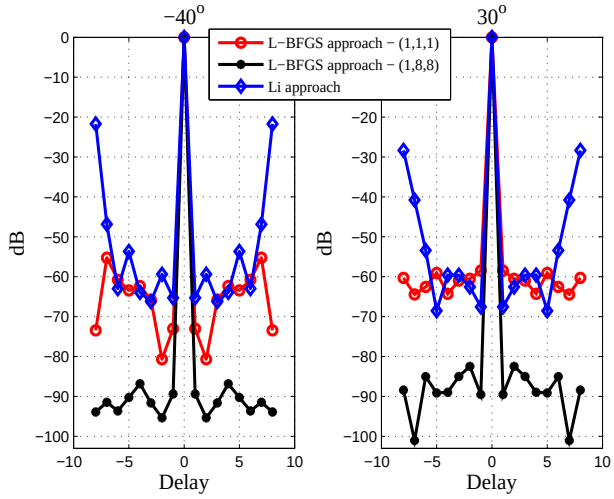
[26] Y.-X Yuan, W.-Y Sun “Optimization Theory and Method” Beijing: Science Press, pp.94–100.



(a) $M = 8$, $L = 32$, and $d = 8$

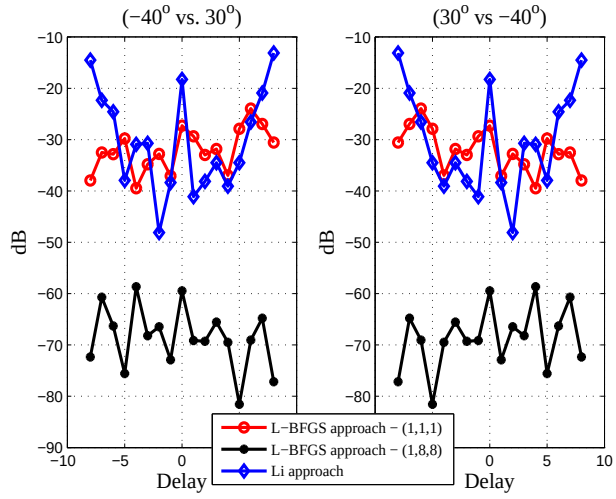


(b) $M = 8$, $L = 32$, and $d = 16$

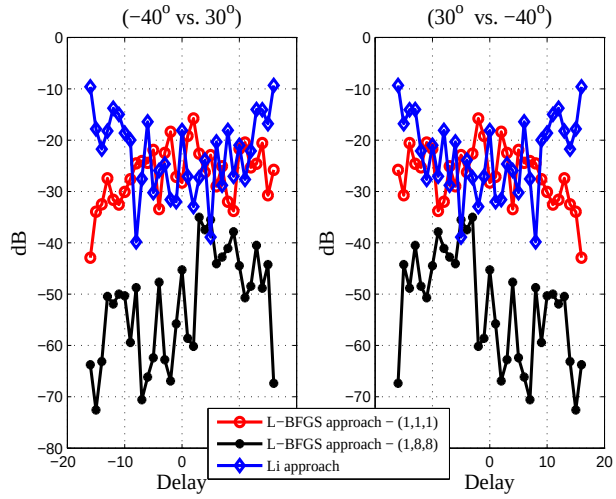


(c) $M = 8$, $L = 128$, and $d = 8$

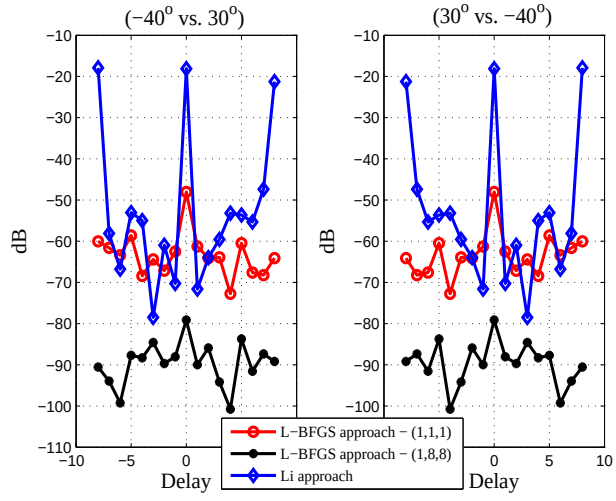
Fig. 4. Comparison of auto-correlation characteristics. Weights, (ω_{b}, ω_{ac}, ω_{cc}) for the L-BFGS approach are (1, 1, 1) and (1, 8, 8), respectively.



(a) $M = 8$, $L = 32$, and $d = 8$



(b) $M = 8$, $L = 32$, and $d = 16$



(c) $M = 8$, $L = 128$, and $d = 8$

Fig. 5. Comparison of cross-correlation characteristics. Weights, (ω_{b}, ω_{ac}, ω_{cc}) for the L-BFGS approach are (1, 1, 1) and (1, 8, 8), respectively.