MAJORIZATION-MINIMIZATION TECHNIQUE FOR MULTI-STATIC RADAR CODE DESIGN

Mohammad Mahdi Naghsh[*+], *Mahmoud Modarres-Hashemi*[+], *Shahram ShahbazPanahi*[×], *Mojtaba Soltanalian*[†], *Petre Stoica*[†]

[+]ECE Dept., Isfahan University of Technology, Isfahan, Iran
[×] Faculty of Eng. and App. Science, University of Ontario Institute of Technology, ON, Canada
[†]Dept. of Information Technology, Uppsala University, Uppsala, Sweden

ABSTRACT

In this paper, we study the problem of code design to improve the detection performance of multi-static radar in the presence of signal-dependent clutter. Due to the lack of analytical expression for receiver operation characteristic (ROC), an information-theoretic criterion, namely the KL-divergence, is considered as the design metric. The code design problem is cast as a non-convex optimization problem with a peak-to-average-power ratio (PAR) constraint. We devise a novel method based on Majorization-Minimization to tackle the arising optimization problem. Via numerical investigations, a general analysis of the coded system performance, as well as the behavior of the KL-divergence versus transmit energy is provided.

***Index Terms*-** KL-divergence, Majorization-Minimization

1. INTRODUCTION

Signal design for detection performance improvement has been a long-term research topic in the radar literature. At the receive side, the signals backscattered from undesired obstacles (known as clutter) depend on the transmit signal, whereas noise, unwanted emissions, and jammer emissions do not depend on the transmit signal. The effect of the clutter has been considered in early studies for single-input single-output (SISO) systems [1, 2]. The aim of these studies is to maximize the signal-to-interference-plus-noise-ratio (SINR) by means of joint optimization of the transmit signal and the receive filter. Recently, a related problem has been considered in [3] with a peak-to-average-power ratio (PAR) constraint.

In multi-static scenarios, the expressions for detection performance are too complicated to be amenable to utilization as design metrics (see e.g. [4][5]). In such circumstances, information-theoretic criteria (e.g. the Kullback-Leibler (KL) divergence) can be considered as design metrics to guarantee some types of optimality for the obtained signals [5][6]. Multiple-input multiple-output (MIMO) radar signal design has been studied in [6] using KL-divergence as design metric in the absence of clutter. In [7], KL-divergence has been taken into account for MIMO radar signal design in the absence of clutter. Information-theoretic criteria have also been used in research subjects related to the detection problem [8].

In this paper, we consider the problem of multi-static radar code design in the presence of clutter. We assume the case of one transmit antenna, however the extension of the derivations in this work to the case of multiple transmit antennas (with orthogonal transmission) is straightforward (see e.g. [5]). Using the KL-divergence metric, the code design problem is formulated as a non-convex optimization problem with PAR constraint. To the best of our knowledge, no study of radar code design with PAR constraints using information-theoretic criteria was conducted prior to this work. We devise a novel method based on Majorization-Minimization (MaMi) technique to tackle the problem. MaMi employs a successive usage of majorization functions to obtain simpler "nearest-vector" problems throughout its iterations.

The rest of this paper is organized as follows. Section 2 presents the data modeling and the design problem formulation. The proposed code design scheme is discussed in Section 3. Numerical examples are provided in Section 4. Finally, Section 5 concludes the paper.

2. PROBLEM FORMULATION

2.1. Data Modeling

We consider a multi-static pulsed-radar with one transmitter and N_r widely separated receive antennas. The baseband transmit signal can be formulated as

$$s(t) = \sum_{n=1}^{N} a_n \phi(t - [n-1]T_P) \qquad (1)$$

This work was supported in part by the European Research Council (ERC) under Grant #228044, and the Swedish Research Council.
* Please address all the correspondence to M. M. Naghsh, Phone: (+98) 311-3912450; Fax: (+98) 311-3912451; Email: mm_naghsh@ec.iut.ac.ir

where $\phi(t)$ is the basic unit-energy transmit pulse, T_P is the pulse repetition period, N is the number of transmitted pulses, and $\{a_n\}_{n=1}^N$ are the deterministic coefficients that are to be "optimally" determined. The vector $\mathbf{a} \triangleq [a_1 \;\; a_2 \;\; \dots \;\; a_N]^T$ is referred to as the code vector of the radar system. The baseband signal received at the k^{th} antenna backscattered from a stationary target can be written as

$$r_k(t) = \underbrace{\alpha_k s(t-\tau_k)}_{\text{target}} + \underbrace{c_k(t)}_{\text{clutter}} + \underbrace{w_k(t)}_{\text{interference}} \tag{2}$$

where α_k is the amplitude of the target return (including the channel effects), $c_k(t)$ is the clutter component, $w_k(t)$ is a Gaussian random process representing the signal-independent interference component (including various types of noise, interference, and jamming), and τ_k is the time corresponding to propagation delay for the path from the transmitter to the target and thereafter to the k^{th} receiver. We assume that the clutter component at the k^{th} receiver is composed of signal echoes produced by many stationary point scatterers located within unambiguous-range with respect to the k^{th} receiver. Accordingly, the clutter component can be expressed as

$$c_k(t) = \sum_{v=1}^{N_c} \rho_{k,v} s(t-\tau_{k,v}) \tag{3}$$

where N_c is the number of point scatterers, $\rho_{k,v}$ is the "amplitude" of the v^{th} scatterer observed by the k^{th} receive antenna, and $\tau_{k,v}$ is the propagation delay at the k^{th} receiver corresponding to the v^{th} scatterer for which we have $\tau_{k,v} \le T_p$.

At the k^{th} receiver, the received signal is matched filtered by $\phi^*(-t)$. Then range-gating is performed by sampling the output of the matched filter at time slots corresponding to a specific radar cell. Let $r_{k,n}$ denote the sample of the filtered $r_k(t)$ at $t = (n-1)T_p + \tau_k$. As $\{\phi(t-[n-1]T_p)\}_{n=1}^N$ are non-overlapping and have unit energy, the effect of the target signature appears as $a_n\alpha_k$ at the $r_{k,n}$. Furthermore, the clutter effect can be expressed as

$$\sum_{v=1}^{N_c} \rho_{k,v} \sum_{i=1}^{N} a_i \int_{-\infty}^{+\infty} \phi(\tau-[i-1]T_p-t)\phi^*(\tau-[n-1]T_p)d\tau. \tag{4}$$

Note that for unambiguous-range clutter scatterers (i.e. scatterers with $\tau_{k,v} \le T_p$), the above integral is zero for $i \ne n$ because $\phi(t-[i-1]T_p-\tau_{k,v})$ and $\phi(t-[n-1]T_p-\tau_k)$ are non-overlapping. Therefore, (4) can be rewritten as

$$a_n \sum_{v=1}^{N_c} \rho_{k,v} \Psi_{n,n}(\tau_k - \tau_{k,v}) \triangleq a_n \widetilde{\rho}_k \tag{5}$$

where $\Psi_{n,n}(.)$ denotes the autocorrelation function of the n^{th} pulse. Therefore, the discrete-time signal corresponding to a certain radar cell for the k^{th} receiver can be described as:

$$\mathbf{r}_k \triangleq \mathbf{s}_k + \mathbf{c}_k + \mathbf{w}_k = \alpha_k \mathbf{a} + \widetilde{\rho}_k \mathbf{a} + \mathbf{w}_k \tag{6}$$

where $\mathbf{r}_k \triangleq [r_{k,1} \;\; r_{k,2} \;\; \cdots \;\; r_{k,N}]^T$, $\mathbf{w}_k \triangleq [w_{k,1} \;\; w_{k,2} \;\; \cdots \;\; w_{k,N}]^T$, $\mathbf{s}_k \triangleq \alpha_k\mathbf{a}$, and $\mathbf{c}_k \triangleq \widetilde{\rho}_k\mathbf{a}$. Herein, $w_{k,n}$ denotes the n^{th} sample of $w_k(t)$ and $\widetilde{\rho}_k$ is a zero-mean complex Gaussian random variable with variance $\sigma^2_{c,k}$ associated with the clutter scatterers.

2.2. Design Problem

Using all the received signals, the target detection leads to the following binary hypothesis problem

$$\begin{cases} \mathrm{H}_0 : & \mathbf{r} = \mathbf{c} + \mathbf{w} \\ \mathrm{H}_1 : & \mathbf{r} = \mathbf{s} + \mathbf{c} + \mathbf{w} \end{cases} \tag{7}$$

where $\mathbf{r}$, $\mathbf{s}$, $\mathbf{c}$, and $\mathbf{w}$ are defined by column-wise stacking of $\mathbf{r}_k, \mathbf{s}_k, \mathbf{c}_k$, and $\mathbf{w}_k$ for $k = 1, 2, ..., N_r$. We assume the received signals at various receivers are statistically independent as receivers are widely separated. Furthermore, we consider the Swerling-I model for target return amplitude, i.e. $\alpha_k \sim \mathcal{CN}(0, \sigma_k^2)$ [4]. Let $\{\mathbf{M}_k\}$ denote the covariance matrices of Gaussian random vectors $\{\mathbf{w}_k\}$. Moreover, let us define $\mathbf{D}_k \triangleq (\sigma^2_{c,k}\mathbf{a}\mathbf{a}^H + \mathbf{M}_k)^{-\frac{1}{2}}$ and $\mathbf{x}_k = \mathbf{D}_k\mathbf{r}_k$. The optimal detector referring to the above detection problem can be expressed as [9]

$$T = \sum_{k=1}^{N_r} \frac{\lambda_k |\theta_k|^2}{1+\lambda_k} \underset{\mathrm{H}_1}{\overset{\mathrm{H}_0}{\lessgtr}} \eta \tag{8}$$

where $\theta_k \triangleq \mathbf{a}^H\mathbf{D}_k\mathbf{x}_k / \|\mathbf{a}^H\mathbf{D}_k\|_2$, and

$$\lambda_k = \sigma_k^2 \mathbf{a}^H (\sigma^2_{c,k}\mathbf{a}\mathbf{a}^H + \mathbf{M}_k)^{-1}\mathbf{a}. \tag{9}$$

Although closed-form expressions for probability of detection P_d and probability of false alarm P_{fa} of the optimal detector can be obtained by applying the results of [4], derivation of the analytical ROC is not possible. In such cases, one can resort to information-theoretic criteria. In this paper, we consider the KL-divergence as the design metric to improve the detection performance. The KL-divergence $\mathcal{D}(f_0\|f_1)$ is a metric to measure the "distance" between two pdfs f_0 and f_1. For a binary hypothesis testing problem the Stein Lemma states that [8]

$$\mathcal{D}\left(f(\mathbf{r}|\mathrm{H}_0)\|f(\mathbf{r}|\mathrm{H}_1)\right) = \lim_{N\to\infty} (-1/N)\log(1-P_d)$$

which implies that the maximization of the KL-divergence metric leads to an asymptotical maximization of P_d. In addition, we have that [8]

$$\mathcal{D}(f(\mathbf{r}|\mathrm{H}_0)\|f(\mathbf{r}|\mathrm{H}_1)) = -\,\mathrm{E}\left\{\log\left(\mathcal{L}\right)|\mathrm{H}_0\right\} \tag{10}$$

where $\mathcal{L}$ is the likelihood ratio defined as $\mathcal{L}(\mathbf{r}) \triangleq \frac{f(\mathbf{r}|\mathrm{H}_1)}{f(\mathbf{r}|\mathrm{H}_0)}$. Using (10) and the identity $\log\left(\mathcal{L}\right) = T - \sum_k \log(1+\lambda_k)$ [9], the KL-divergence for (7) can be obtained as

$$\mathcal{D}\left(f(\mathbf{r}|\mathrm{H}_0)\|f(\mathbf{r}|\mathrm{H}_1)\right) = \sum_{k=1}^{N_r} \{\log(1+\lambda_k) - \lambda_k/(1+\lambda_k)\}.$$

Now, let ζ denote the allowed PAR level of the code, viz. $\mathrm{PAR}(\mathbf{a}) = \max_n \{|a_n|^2\}/(\frac{1}{N}\|\mathbf{a}\|_2^2) \leq \zeta$. The problem of PAR-constrained code design by maximizing the KL-divergence metric can be cast as

$$\begin{aligned} \max_{\mathbf{a},\lambda_k} \quad & \sum_{k=1}^{N_r} \{\log(1+\lambda_k) - \lambda_k/(1+\lambda_k)\} \\ \text{subject to} \quad & \lambda_k = \sigma_k^2 \mathbf{a}^H(\sigma_{c,k}^2 \mathbf{a}\mathbf{a}^H + \mathbf{M}_k)^{-1}\mathbf{a} \qquad (11) \\ & \max_{n=1,\ldots,N} \{|a_n|^2\} \leq \zeta\,(e/N) \\ & \|\mathbf{a}\|_2^2 = e, \end{aligned}$$

where e denotes the total transmit energy.

3. THE PROPOSED METHOD

We use the Majorization-Minimization (or Minorization-Maximization) techniques to tackle the non-convex problem in (11). Majorization-Minimization (MaMi) is an iterative technique that can be used for obtaining a locally optimal solution to the general minimization problem

$$\min_{\mathbf{z}} \ \widetilde{f}(\mathbf{z}) \quad \text{subject to} \quad c(\mathbf{z}) \leq 0 \qquad (12)$$

where $\widetilde{f}(.)$ and $c(.)$ are non-convex functions. Each iteration (say the l^{th} iteration) of MaMi consists of two steps:

• Majorization Step: Finding $p^{(l)}(\mathbf{z})$ such that its minimization is simpler than that of $\widetilde{f}(\mathbf{z})$, and that $p^{(l)}(\mathbf{z})$ majorizes $\widetilde{f}(\mathbf{z})$, i.e.

$$p^{(l)}(\mathbf{z}) \geq \widetilde{f}(\mathbf{z}), \ \ \forall \mathbf{z} \ \text{ and } \ p^{(l)}(\mathbf{z}^{(l-1)}) = \widetilde{f}(\mathbf{z}^{(l-1)}) \qquad (13)$$

with $\mathbf{z}^{(l-1)}$ being the value of $\mathbf{z}$ at the $(l-1)^{th}$ iteration.

• Minimization Step: Solving the optimization problem,

$$\min_{\mathbf{z}} \ p^{(l)}(\mathbf{z}) \quad \text{subject to} \quad c(\mathbf{z}) \leq 0 \qquad (14)$$

to obtain $\mathbf{z}^{(l)}$.

We begin by noting that the convex term $g(\lambda_k) = -\lambda_k/(1+\lambda_k)$ can be minorized using its supporting hyperplane at any given $\lambda_k = \lambda_k^{(l)}$ which implies that

$$\sum_{k=1}^{N_r} g(\lambda_k) \geq \sum_{k=1}^{N_r} g(\lambda_k^{(l)}) + \sum_{k=1}^{N_r} g'(\lambda_k^{(l)})(\lambda_k - \lambda_k^{(l)}). \qquad (15)$$

Herein $\lambda_k^{(l)}$ denotes the λ_k obtained at the l^{th} iteration, and $g'(.)$ denotes the first-order derivative of $g(.)$. Furthermore, λ_k in (9) can be simplified using matrix inversion lemma:

$$\begin{aligned} \lambda_k &= \sigma_k^2 \mathbf{a}^H \left(\mathbf{M}_k^{-1} - \sigma_{c,k}^2 \frac{\mathbf{M}_k^{-1}\mathbf{a}\mathbf{a}^H\mathbf{M}_k^{-1}}{1+\sigma_{c,k}^2 \mathbf{a}^H \mathbf{M}_k^{-1}\mathbf{a}} \right) \mathbf{a} \\ &= \sigma_k^2 (\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a})/(1+\sigma_{c,k}^2 \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}). \qquad (16) \end{aligned}$$

Next observe that using (16), the optimal code $\mathbf{a} = \mathbf{a}_\star$ can be obtained in an iterative manner solving the following maximization at the $(l+1)^{th}$ iteration:

$$\begin{aligned} \max_{\mathbf{a},\lambda_k} \quad & \sum_{k=1}^{N_r} \log(1+\lambda_k) + g\prime\left(\lambda_k^{(l)}\right)\ \lambda_k \qquad (17) \\ \text{subject to} \quad & \lambda_k = \gamma_k - \frac{\gamma_k}{1+\beta_k \mathbf{a}^H \mathbf{M}_k^{-1}\mathbf{a}} \qquad (18) \\ & \max_{n=1,\ldots,N} \{|a_n|^2\} \leq \zeta\,(e/N) \qquad (19) \\ & \|\mathbf{a}\|_2^2 = e, \qquad (20) \end{aligned}$$

where $\gamma_k = \frac{\sigma_k^2}{\sigma_{c,k}^2}$ and $\beta_k = \sigma_{c,k}^2$. Substituting $\{\lambda_k\}$ of (18) into the objective function of (17) leads to the following expression for the objective function:

$$\begin{aligned} \sum_{k=1}^{N_r} \Bigg[& \log\left(1+\gamma_k - \tfrac{\gamma_k}{1+\beta_k \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}\right) \\ & + \left(\tfrac{1}{1+\lambda_k^{(l)}}\right)^2 \left(\tfrac{\gamma_k}{1+\beta_k \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}\right) \Bigg] \qquad (21) \end{aligned}$$

where $\lambda_k^{(l)} = \gamma_k - \gamma_k/(1+\beta_k y_k^{(l)})$ and $y_k^{(l)} = \mathbf{a}^{(l)\,H}\mathbf{M}_k^{-1}\mathbf{a}^{(l)}$. In the following, we exploit theorems 3.1 and 4.2 in [10] to derive a lemma (whose detailed proof is omitted due to the lack of space) that paves the way for obtaining a minorizer of the above logarithmic term.

Lemma 1. *Let* $f(x) = -\log(1+\mu - \frac{\mu}{1+\eta x^2})$ *for some* $\mu, \eta > 0$*. Then for all* $x, \tilde{x} \in \mathbb{R}$ *we have that*

$$\begin{aligned} f(x) \ \leq \ & f(\tilde{x}) + \frac{\eta}{1+\eta\tilde{x}^2}(x^2 - \tilde{x}^2) \\ & - \frac{2\eta\tilde{x}(1+\mu)}{1+\eta(1+\mu)\tilde{x}^2}(x-\tilde{x}) + \eta(1+\mu)(x-\tilde{x})^2. \end{aligned}$$

A minorizer of the logarithmic term can be obtained immediately by employing Lemma 1 with $x_k = \sqrt{\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}$, $\mu = \gamma_k$, and $\eta = \beta_k$. To deal with the expression $\frac{\gamma_k}{1+\beta_k \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}$ in (21) conveniently, we use the convexity of the function $\frac{1}{1+\beta x}$ for $\beta > 0$ which implies

$$\frac{1}{1+\beta x} \geq \frac{1}{1+\beta\tilde{x}} - \frac{\beta}{(1+\beta\tilde{x})^2}(x-\tilde{x}), \quad \forall\, x, \tilde{x}. \qquad (22)$$

As a result, a minorizer of $\frac{\gamma_k}{1+\beta_k \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}$ can be obtained by considering the above inequality for $x_k = \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}$ and $\beta = \beta_k$. Furthermore, by replacing the summation terms in (21) for each k with the obtained minorizers (using Lemma 1 and eq. (22)) and removing the constants, the criterion in (21) turns to:

$$\sum_{k=1}^{N_r} \left[\phi_k^{(l)} \mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a} + \psi_k^{(l)} \sqrt{\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}} \right] \qquad (23)$$

where

$$\phi_k^{(l)} \triangleq \frac{\beta_k}{1+\beta_k y_k^{(l)}} + \beta_k(1+\gamma_k) + \frac{\gamma_k}{(1+\lambda_k^{(l)})^2}\left(\frac{\beta_k}{(1+\beta_k y_k^{(l)})^2}\right)$$

$$\psi_k^{(l)} \triangleq \sqrt{y_k^{(l)}}\left(\frac{2\beta_k(1+\gamma_k)}{1+\beta_k y_k^{(l)}(1+\gamma_k)} + 2\beta_k(1+\gamma_k)\right).$$

Yet, due to the non-concavity of the terms $\left\{\sqrt{\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}\right\}$, dealing with the maximization of the criterion in (23) appears to be complicated. However, $\sqrt{\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}}$ can be minorized using its supporting hyperplane at any given $\tilde{\mathbf{a}}$; more precisely,

$$\sqrt{\mathbf{a}^H\mathbf{M}_k^{-1}\mathbf{a}} \geq \sqrt{\tilde{\mathbf{a}}^H\mathbf{M}_k^{-1}\tilde{\mathbf{a}}} + \text{Real}\left(\frac{\tilde{\mathbf{a}}^H\mathbf{M}_k^{-1}}{\sqrt{\tilde{\mathbf{a}}^H\mathbf{M}_k^{-1}\tilde{\mathbf{a}}}}(\mathbf{a}-\tilde{\mathbf{a}})\right). \tag{24}$$

Ultimately, using eq. (23) and (24) as well as removing the constants, the optimization problem associated with the $(l+1)^{th}$ iteration will become as follows:

$$\begin{array}{lll} \min\limits_{\mathbf{a}} & \mathbf{a}^H\left(\sum\limits_{k=1}^{N_r}\phi_k^{(l)}\mathbf{M}_k^{-1}\right)\mathbf{a} - \text{Real}\left(\sum\limits_{k=1}^{N_r}\mathbf{a}^H\mathbf{d}_k^{(l)}\right) & \\ \text{subject to} & \max\limits_{n=1,\ldots,N}\{|a_n|^2\} \leq \zeta\,(e/N) & (25) \\ & \|\mathbf{a}\|_2^2 = e & \end{array}$$

where $\mathbf{d}_k^{(l)} \triangleq (\psi_k^l/\sqrt{y_k^{(l)}})\mathbf{M}_k^{-1}\mathbf{a}^{(l)}$. The problem in (25) can be recast equivalently as

$$\begin{array}{lll} \max\limits_{\mathbf{a}} & \widehat{\mathbf{a}}^H\mathbf{K}\,\widehat{\mathbf{a}} & (26) \\ \text{subject to} & \max\limits_{n=1,\cdots,N}\{|a_n|^2\} \leq \zeta\,(e/N) & \\ & \|\mathbf{a}\|_2^2 = e & \end{array}$$

where $\widehat{\mathbf{a}} = [\mathbf{a} \quad 1]^T$, $\mathbf{K} = \mu\mathbf{I}_{N+1} - \mathbf{J}$, and

$$\mathbf{J} = \begin{bmatrix} \left(\sum_{k=1}^{N_r}\phi_k^{(l)}\mathbf{M}_k^{-1}\right) & -0.5\left(\sum_{k=1}^{N_r}\mathbf{d}_k^{(l)}\right) \\ -0.5\left(\sum_{k=1}^{N_r}\mathbf{d}_k^{(l)}\right)^H & 0 \end{bmatrix}.$$

Herein, $\mu > \mu_{max}$ with μ_{max} being the maximum eigenvalue of the matrix $\mathbf{J}$. The code vector $\mathbf{a}$ at the $(l+1)^{th}$ iteration of MaMi can be obtained from $\mathbf{a}^{(p)}$ (at convergence), using the power method-like iterations [11]:

$$\begin{array}{lll} \max\limits_{\mathbf{a}^{(p+1)}} & \|\mathbf{a}^{(p+1)} - \breve{\mathbf{a}}^{(p)}\| & (27) \\ \text{subject to} & \max\limits_{n=1,\cdots,N}\{|a_n^{(p+1)}|^2\} \leq \zeta(e/N) & \\ & \|\mathbf{a}^{(p+1)}\|_2^2 = e & \end{array}$$

where $\breve{\mathbf{a}}^{(p)}$ represents the vector containing the first N entries of $\mathbf{K}\ \widehat{\mathbf{a}}^{(p)}$. The optimization problem (27) is a "nearest-vector" problem with PAR constraint and can be solved efficiently using a recursive algorithm proposed in [12]: If the magnitudes of the entries of $\sqrt{e}\,\breve{\mathbf{a}}^{(p)}/\|\breve{\mathbf{a}}^{(p)}\|_2$ are below $\sqrt{\zeta(e/N)}$ then $\mathbf{a}^{(p+1)} = \sqrt{e}\,\breve{\mathbf{a}}^{(p)}/\|\breve{\mathbf{a}}^{(p)}\|_2$ is the solution. Otherwise, the entry of $\mathbf{a}^{(p+1)}$ corresponding to the entry of $\breve{\mathbf{a}}^{(p)}$ (say a_{max}) with maximal magnitude is given by $\sqrt{\zeta(e/N)}\,e^{j\arg(a_{max})}$; and the other entries of $\mathbf{a}^{(p+1)}$ can be obtained by solving the same type of "nearest-vector" problem but with the remaining energy i.e. $e - \zeta(e/N)$.

Finally, the steps of MaMi algorithm are summarized in Table 1.

Table 1. The MaMi Algorithm for maximizing the KL-divergence with a PAR constraint

Step 0: Initialize $\mathbf{a}$ with a random vector in $\mathbb{C}^N$ and set the iteration number l to 0.
Step 1: Solve the problem in (25) iteratively considering the nearest-vector problem in (27); set $l \leftarrow l+1$.
Step 2: Compute $\phi_k^{(l)}$ and $\mathbf{d}_k^{(l)}$.
Step 3: Repeat steps 1 and 2 until a pre-defined stop criterion is satisfied, e.g. $\|\mathbf{a}^{(l+1)} - \mathbf{a}^{(l)}\|_2 \leq \xi$ for some $\xi > 0$.

4. SIMULATION RESULTS

In this section, we present numerical examples to examine the performance of the proposed algorithm. In particular, we compare the system performance for coded and uncoded (employing the code vector $\mathbf{a} = \sqrt{e/N}\,\mathbf{1}$) scenarios.

In this section, we assume the code length $N = 10$, the number of receivers $N_r = 4$, variances of the target components given by $\sigma_k^2 = 1$ (for $1 \leq k \leq 4$), and variances of the clutter components given by $(\sigma_{c,1}^2, \sigma_{c,2}^2, \sigma_{c,3}^2, \sigma_{c,4}^2) = (0.125, 0.25, .5, 1)$. Furthermore, we assume that the k^{th} interference covariance matrix $\mathbf{M}_k$ is given by $[\mathbf{M}_k]_{m,n} = (1-0.15k)^{|m-n|}$. The ROC is used to evaluate the detection performance of the system using analytical expressions for P_d and P_{fa} (see eqs. (32)-(34) in [4]).

Fig. 1 show the ROCs associated with the coded system (employing the optimized codes) with no PAR constraint and with PAR $= 1$ as well as the uncoded system for $e = 10$. It can be observed that the performance of the coded system outperforms that of the uncoded system significantly. Furthermore, a performance degradation is observed for constant modulus code design as compared to the unconstrained design. This can be explained using the fact that the feasibility set of the unconstraint design problem is larger that that of the constrained design.

The behavior of the KL-divergence criterion versus the transmit energy e is investigated in Fig. 2 for the coded system (with no PAR constraint, and PAR $= 1$) and the uncoded system. It is observed that for sufficiently large values of transmit energy, the performance improvement obtained by increasing e is negligible (saturation phenomenon). Moreover, a saturation of the coded system always occurs before that in the un-

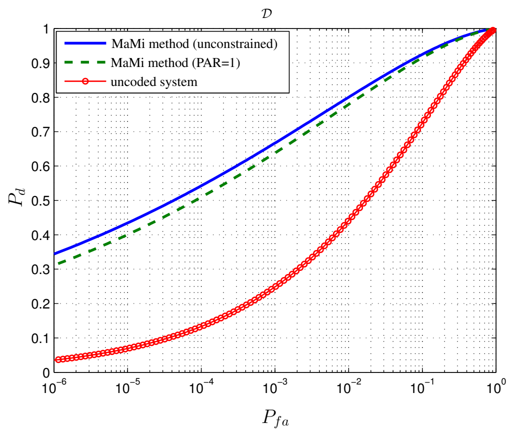

Fig. 1. ROCs of optimally coded and the uncoded systems.

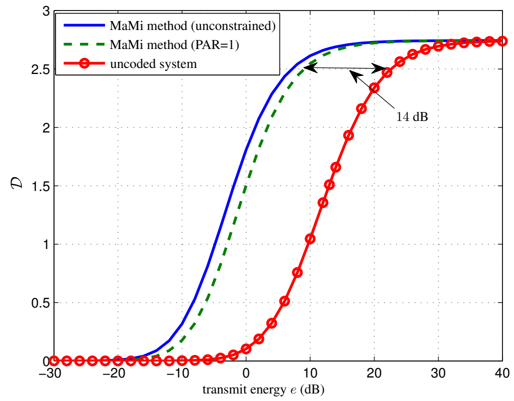

Fig. 2. Behavior of KL-divergence versus transmit energy e for the coded and uncoded systems.

coded system, which was expected: employing an optimized code enables the system to perform closer to the best possible performance at lower values of e. An approximate decrease of $16\ dB$ and $14\ dB$ in the required transmit energy of the coded systems with no PAR constraint and with PAR=1 is observed, respectively, compared to the uncoded system for $\mathcal{D} = 2.5$.

5. CONCLUSIONS

A multi-static radar code design scheme based on KL-divergence was considered in the presence of clutter. A novel method was devised to tackle the highly non-linear and non-convex design problem using the Majorization-Minimization (MaMi) technique. MaMi relies on successive (linear as well as quadratic) majorizations such that each iteration of the algorithm can be handled using nearest-vector optimizations. Numerical examples were provided to examine the effectiveness of the proposed method. The metric's saturation phenomenon, as the transmit energy increases, was also investigated.

A. REFERENCES

[1] D. F. Delong and E. M. Hofstteter, "On the design of optimum radar waveforms for clutter rejection," *IEEE Trans. Inf. Theory*, vol. 13, pp. 454–463, Jul. 1967.

[2] L. J. Spafford, "Optimum radar signal processing in clutter," *IEEE Trans. Inf. Theory*, vol. 14, pp. 734–743, Sep. 1968.

[3] P. Stoica, H. He, and J. Li, "Optimization of the receive filter and transmit sequence for active sensing," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1730 –1740, Apr. 2012.

[4] M. M. Naghsh and M. Modarres-Hashemi, "Exact theoretical performance analysis of optimum detector for statistical MIMO radars," *IET Journal on Radar, Sonar, and Navigation*, vol. 6, pp. 99–111, 2012.

[5] S. M. Kay, "Waveform design for multistatic radar detection," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 45, pp. 1153–1165, Jul. 2009.

[6] E. Grossi and M. Lops, "Space-time code design for MIMO detection based on Kullback-Leibler divergence," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3989–4004, 2012.

[7] B. Tang, J. Tang, and Y. Peng, "MIMO radar waveform design in colored noise based on information theory," *IEEE Trans. Signal Process.*, vol. 58, pp. 4684–4697, Sep. 2010.

[8] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun.*, vol. 15, pp. 52–60, Feb. 1967.

[9] S. M. Kay, *Fundamentals of Statistical Signal Processing-Volume II: Detection Theory*, 1st ed. New Jersey: Prentice Hall, 1998.

[10] J. D. Leeuw and K. Lange, "Sharp quadratic majorization in one dimension," *Computational Statistics and Data Analysis*, vol. 53, pp. 2471–2484, May 2009.

[11] M. Soltanalian, B. Tang, J. Li, and P. Stoica, "Joint design of the receive filter and transmit sequence for active sensing," *IEEE Signal Processing Letters*, vol. 20, no. 5, pp. 423–426, May.

[12] J. A. Tropp, I. S. Dhillon, R. W. Heath, and T. Strohmer, "Designing structured tight frames via an alternating projection method," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 188–208, Jan. 2005.