Optimization of the Receive Filter and Transmit Sequence for Active Sensing

Petre Stoica, *Fellow, IEEE*, Hao He, *Member, IEEE*, and Jian Li, *Fellow, IEEE*

***Abstract*—This paper discusses the joint design of receive filters and transmit signals for active sensing applications such as radar and active sonar. The goal is to minimize the mean-square error (MSE) of target's scattering coefficient estimate in the presence of clutter and interference, which is equivalent to maximizing the signal-to-clutter-plus-interference ratio. A discrete-time signal model is assumed and practical constant-modulus or low peak-to-average-power ratio (PAR) constraints are imposed on the transmit signal. Several optimization methods are proposed for this joint design. Furthermore, the MSE criterion is expressed in the frequency domain and a corresponding MSE lower bound is derived. Numerical examples for different types of interferences are included to demonstrate the effectiveness of the proposed designs.**

***Index Terms*—Cognitive radar, joint design, probing signal, receive filter.**

I. Introduction

THERE is a long-standing interest in designing radar transmit signals and receive filters for clutter/interference rejection, e.g., [1]–[10]. Clutter (or reverberation in sonar terminology) refers to unwanted echoes that are usually correlated with the transmitted signal, while interference is a term used for noise as well as (adverse) jamming signals. Since the negative impacts from clutter and interference should be minimized at the receiver side, a natural criterion for designing transmit signals and receive filters would be to maximize the signal-to-clutter-plus-interference ratio (SCIR) of the receiver output at the time of target detection.

Manuscript received May 20, 2011; revised September 21, 2011 and November 19, 2011; accepted December 01, 2011. Date of publication December 13, 2011; date of current version March 06, 2012. The associate editor coordinating the review of this manuscript and approving it for publication was . This work was supported in part by the Office of Naval Research (ONR) under Grants N00014-09-1-0211 and N00014-10-1-0054, the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant W911NF-07-1-0450, the Swedish Research Council (VR), and the European Research Council (ERC). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon.

P. Stoica is with the Department of Information Technology, Uppsala University, SE-75105 Uppsala, Sweden (e-mail: ps@it.uu.se).

H. He is with Amazon.com, Seattle, WA 98109-1226 USA (e-mail: haohe@amazon.com.

J. Li is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA, and also with IAA, Inc., Gainesville, FL (e-mail: li@dsp.ufl.edu).

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

Digital Object Identifier 10.1109/TSP.2011.2179652

It is well known that the matched filter (MF) maximizes the signal-to-noise ratio (SNR) in the presence of additive white noise. The matched filter can be implemented as a correlator that multiplies the received signal with a replica of the transmitted signal. The peak of the receiver output indicates the time delay of the target signal. If the received signal also has a Doppler frequency shift due to the relative movement between the target and the platform, a bank of filters is needed, each of which is matched to a specific Doppler frequency.

The matched filter is not able to take care of clutter or jamming suppression, a feature that is left to the transmit signal design. The jamming signals can come from an adversary or interfering radio applications and usually operate in certain frequency bands. Their negative effects can be largely avoided if we put little energy of the transmit signal into those frequency bands. As for clutter that appears as signal-like returns but with different delays or frequency shifts from the signal of interest, its effects can be analyzed using the ambiguity function (AF) [11], [12]. In order to minimize the clutter effects, the AF sidelobes need to be minimized [13]–[15]. When the intrapulse Doppler shifts are negligible, clutter suppression can be related to a more thoroughly studied problem, namely the minimization of autocorrelation sidelobes of the transmit signal [16], [17]. Regarding the dynamic range of the transmit signals, owing to hardware constraints such as the maximum clipping of power amplifiers and A/D converters [18], [19], it is desirable that these signals have constant modulus or low peak-to-average power ratios (PAR). Such constraints are difficult to take into account in the design and are not always addressed properly in the literature.

Alternatively, we can use a "mismatched" filter (MMF) at the receiver side (see, e.g., [9] and the references therein). Using such a filter instead of the matched filter means trading off SNR for SCIR. This tradeoff is legitimate if the radar detection performance is actually clutter or jamming limited. One benefit of using an MMF filter is that it is not subject to the constant-modulus or low PAR constraint, which allows for more degrees of freedom (DOF) in the design. On the other hand, a joint design of the MMF filter and the transmit signal leads to a more complex optimization problem that involves cross ambiguity functions or cross correlations in the negligible Doppler case [3], [20].

In this paper, we consider the joint design of the receive filter and transmit signal. The goal is to minimize the mean-square error (MSE) of a target's scattering coefficient estimate in the presence of clutter and interference. A particular feature of our approach, which distinguishes it from most of the previous design methods, is that we consider a discrete-time signal model which yields designs that can be directly implemented in a digital system (unlike the continuous-time designs of most earlier

approaches). Another distinctive feature of the proposed approach is that it imposes constant-modulus or low PAR constraint on the transmit signal, which is always required in practice and yet has been rarely taken into account in the previous literature.

Section II presents the data model and problem formulation. Sections III, IV and V propose three approaches for the joint design under consideration. More specifically, the algorithm discussed in Section III uses a direct gradient approach to minimize the MSE of target's scattering coefficient estimate, whereas the algorithms in Sections IV and V first transform the MSE metric to the frequency domain and then make use of the Lagrange multiplier approach to solve the minimization problem. An MSE lower bound is also obtained in Section IV. Numerical examples are provided in Section VI and conclusions are drawn in Section VII. Throughout the paper we use bold lowercase/uppercase letters to denote vectors/matrices, respectively. $(\cdot)^H$ indicates vector/matrix conjugate transpose, $(\cdot)^T$ indicates vector/matrix transpose and $(\cdot)^*$ indicates complex conjugate for scalars. $\arg(x)$ denotes the phase of x, $\lfloor x \rfloor$ the largest integer equal to or smaller than x, $\|\cdot\|$ the vector Euclidean norm and $\mathbf{I}_N$ the $N \times N$ identity matrix.

II. Data Model and Problem Formulation

Let

$$\mathbf{s} = [s_1 \quad s_2 \quad \cdots \quad s_N]^T \tag{1}$$

denote the probing sequence that modulates the transmitted train of subpulses [1], [12]. We assume a digital system so that only the discrete-time signal is of concern. The energy of $\{s_k\}_{k=1}^N$ is constrained to be N:

$$\|\mathbf{s}\|^2 = N \tag{2}$$

without any loss of generality. Such an energy constraint must always be imposed in the design.

We constrain $\{s_k\}_{k=1}^N$ to satisfy either

a) a unimodularity constraint

$$\begin{aligned} &|s_k| = 1 \Longleftrightarrow s_k = e^{j\phi_k} \\ &\phi_k \in [0, 2\pi), \quad k = 1, \ldots, N \end{aligned} \tag{3}$$

or

b) a PAR constraint

$$\mathrm{PAR}(\mathbf{s}) \le \mu, \quad \mu \in [1, N] \tag{4}$$

where

$$\mathrm{PAR}(\mathbf{s}) \triangleq \frac{\max_k |s_k|^2}{\frac{1}{N}\sum_{p=1}^N |s_p|^2} = \max_k |s_k|^2 \tag{5}$$

and μ is a predefined parameter that specifies the maximum allowed PAR. Note that when $\mu = N$, (4) imposes no constraint at all; and that when $\mu = 1$, (4) coincides with the unimodularity constraint in (3).

Under the assumption that the intrapulse Doppler is negligible and that the sampling is synchronized to the pulse rate, the received discrete-time baseband data vector after alignment to the range cell of interest, can be written as [7]–[9]

$$\begin{aligned} \mathbf{y} = \alpha_0 \begin{bmatrix} s_1 \\ \vdots \\ s_{N-1} \\ s_N \end{bmatrix} + \alpha_1 \begin{bmatrix} 0 \\ s_1 \\ \vdots \\ s_{N-1} \end{bmatrix} + \cdots + \alpha_{N-1} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ s_1 \end{bmatrix} \\ + \alpha_{-1} \begin{bmatrix} s_2 \\ \vdots \\ s_N \\ 0 \end{bmatrix} + \cdots + \alpha_{-N+1} \begin{bmatrix} s_N \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \boldsymbol{\epsilon} \end{aligned} \tag{6}$$

where α_0 is the scattering coefficient of the current range cell, $\{\alpha_k\}_{k\neq 0}$ are the scattering coefficients of the adjacent range cells that contribute clutter or reverberation components to $\mathbf{y}$, and $\boldsymbol{\epsilon}$ is an interference term that comprises measurement noise as well as other disturbances such as jamming. Note that if the Doppler effects are not negligible (owing to a fast relative motion between the platform and the target), the data model in (6) needs to be modified to accommodate frequency spreading [9] and the analysis should involve two-dimensional ambiguity functions [3], [20]. However such an analysis is not undertaken in this paper where we focus on the simpler data model in (6).

We assume that the covariance matrix of $\boldsymbol{\epsilon}$, viz.,

$$\mathrm{E}\{\boldsymbol{\epsilon}\boldsymbol{\epsilon}^H\} = \boldsymbol{\Gamma} \tag{7}$$

is Toeplitz (which is a weak assumption given the uniform sampling that led to (6)). We also assume that the clutter coefficients $\{\alpha_k\}_{k\neq 0}$ in (6) are independent of one another and of $\boldsymbol{\epsilon}$, and that

$$\mathrm{E}\left\{|\alpha_k|^2\right\} = \beta, \quad k \neq 0. \tag{8}$$

Both $\boldsymbol{\Gamma}$ and β are assumed to be known. In active sensing applications, information on $\boldsymbol{\Gamma}$ and β can be obtained by some form of preprocessing and is usually assumed available in the so-called cognitive systems; see e.g. [21].

One of the principal goals of data processing for (6) is the estimation of α_0. The MMF estimate of α_0 is the following linear function of $\mathbf{y}$:

$$\hat{\alpha}_0 = \frac{\mathbf{w}^H\mathbf{y}}{\mathbf{w}^H\mathbf{s}} \tag{9}$$

where $\mathbf{w}$ is the $N \times 1$ MMF vector. An important special case of (9) is that of the MF estimate

$$\hat{\alpha}_0 = \frac{\mathbf{s}^H\mathbf{y}}{\|\mathbf{s}\|^2} \tag{10}$$

which corresponds to using the sequence itself as the MMF vector

$$\mathbf{w} = \mathbf{s}. \tag{11}$$

Under the stated assumptions [see (7) and (8)], the MSE of (9) can be readily derived:

$$\mathrm{MSE}(\hat{\alpha}_0) = \mathrm{E}\left\{\left|\frac{\mathbf{w}^H\mathbf{y}}{\mathbf{w}^H\mathbf{s}} - \alpha_0\right|^2\right\} = \frac{\mathbf{w}^H\mathbf{R}\mathbf{w}}{|\mathbf{w}^H\mathbf{s}|^2} \tag{12}$$

where

$$\mathbf{R} = \beta \sum_{\substack{k=-N+1 \\ k\neq 0}}^{N-1} \mathbf{J}_k \mathbf{s}\mathbf{s}^H \mathbf{J}_k^H + \mathbf{\Gamma} \tag{13}$$

and $\mathbf{J}_k$ denotes the following shifting matrix:

$$\mathbf{J}_k = \mathbf{J}_{-k}^H = \begin{bmatrix} \overbrace{\begin{matrix} & & \\ & & 1 \end{matrix}}^{k+1} & & 0 \\ & \ddots & \\ & & 1 \\ 0 & & \end{bmatrix}^H_{N\times N}, \quad k = 0, \ldots, N-1. \tag{14}$$

The main goal of this paper is to jointly design the receive filter $\mathbf{w}$ and the transmit sequence $\mathbf{s}$ so as to minimize the MSE in (12). Observe that the denominator in (12) is the power of the signal component in the receiver output, and the numerator is the power of clutter and interference. Hence, minimizing the MSE in (12) is equivalent to maximizing the SCIR. Note also that the MMF estimate has $2N$ additional (real-valued) DOFs as compared to MF; these extra DOFs of (9) should allow a more accurate estimation of α_0 than what is possible using (10).

In what follows, we present three algorithms for the design of $\mathbf{w}$ and $\mathbf{s}$. These algorithms are referred to using the acronym CREW (Cognitive REceiver and Waveform design).

III. CREW(GRA)-A GRADIENT APPROACH

The minimization of (12) with respect to (wrt) $\mathbf{w}$, for fixed $\mathbf{s}$, yields the following well-known closed-form expression for the minimizing vector:

$$\mathbf{w} = \mathbf{R}^{-1}\mathbf{s} \tag{15}$$

to within a multiplicative constant. The proof of (15) is immediate. Let $\mathbf{R}^{\frac{1}{2}}$ denote a Hermitian square root of $\mathbf{R}$. By the Cauchy-Schwartz inequality, we have that

$$\left|\mathbf{w}^H\mathbf{s}\right|^2 = \left|\mathbf{w}^H\mathbf{R}^{1/2}\mathbf{R}^{-1/2}\mathbf{s}\right|^2 \leq (\mathbf{w}^H\mathbf{R}\mathbf{w})(\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}). \tag{16}$$

It follows from (12) and (16) that

$$\mathrm{MSE}(\hat{\alpha}_0) = \frac{\mathbf{w}^H\mathbf{R}\mathbf{w}}{\left|\mathbf{w}^H\mathbf{s}\right|^2} \geq \frac{1}{\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s}} \tag{17}$$

where the lower bound is obtained at (15).

Next we aim to minimize $\frac{1}{(\mathbf{s}^H\mathbf{R}^{-1}\mathbf{s})}$ wrt $\mathbf{s}$ under the unimodularity constraint (3), or equivalently

$$\min_{\boldsymbol{\phi}} -\mathbf{s}^H(\boldsymbol{\phi})\mathbf{R}^{-1}(\boldsymbol{\phi})\mathbf{s}(\boldsymbol{\phi}) \stackrel{\Delta}{=} f(\boldsymbol{\phi}) \tag{18}$$

where

$$\boldsymbol{\phi} = [\phi_1 \quad \cdots \quad \phi_N]^T \tag{19}$$

comprises the phases of $\mathbf{s}$. Gradient-based methods can be used to solve (18), such as the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [22]. In contrast to the classical Newton method, in BFGS the Hessian matrix of the second-order derivatives does not need to be evaluated directly; only the first-order derivative vector $\nabla f(\boldsymbol{\phi})$ is needed. The elements of $\nabla f(\boldsymbol{\phi})$ can be computed as

$$\frac{\partial f(\boldsymbol{\phi})}{\partial \phi_k} = -\left[\frac{\partial \mathbf{s}^H}{\partial \phi_k}\mathbf{R}^{-1}\mathbf{s} + \mathbf{s}^H\frac{\partial \mathbf{R}^{-1}}{\partial \phi_k}\mathbf{s} + \mathbf{s}^H\mathbf{R}^{-1}\frac{\partial \mathbf{s}}{\partial \phi_k}\right], \quad k = 1, \ldots, N \tag{20}$$

where

$$\frac{\partial \mathbf{s}}{\partial \phi_k} = \left[\underbrace{0 \cdots 0}_{k-1} \quad je^{j\phi_k} \quad \underbrace{0 \cdots 0}_{N-k}\right]^T$$
$$\frac{\partial \mathbf{R}^{-1}}{\partial \phi_k} = -\mathbf{R}^{-1}\frac{\partial \mathbf{R}}{\partial \phi_k}\mathbf{R}^{-1} \tag{21}$$

and

$$\frac{\partial \mathbf{R}}{\partial \phi_k} = \beta \sum_{\substack{l=-(N-1) \\ l\neq 0}}^{N-1} \mathbf{J}_l \left(\frac{\partial \mathbf{s}}{\partial \phi_k}\mathbf{s}^H + \mathbf{s}\frac{\partial \mathbf{s}^H}{\partial \phi_k}\right)\mathbf{J}_l^H. \tag{22}$$

The above expression for $\nabla f(\boldsymbol{\phi})$ is all that is needed by a number of available numerical solvers that implement BFGS, such as the "fminunc" function in MATLAB. The so-obtained *gra*dient-based algorithm is called CREW(gra) and is summarized in Table I. Note that CREW(gra) can only deal with the unimodularity constraint. Also note that Step 1 of Table I may require many iterations, during each of which all N elements of $\nabla f(\boldsymbol{\phi})$ need to be recomputed. This makes CREW(gra) computationally expensive in the case of large N (e.g., larger than 400 on an ordinary PC).

TABLE I
THE CREW(GRA) ALGORITHM WITH UNIMODULARITY CONSTRAINT

Step 0: Use an available sequence or a randomly generated one to initialize $\mathbf{s}$.
Step 1: Solve (18) for $\mathbf{s}$ using any gradient-based solver (the expression of the first-order derivative is given in (20)), such as the "fminunc" function in MATLAB that utilizes the BFGS algorithm.
Step 2: Compute $\mathbf{w}$ using (15) with $\mathbf{s}$ provided by Step 1.

IV. CREW(FRE)-A FREQUENCY-DOMAIN APPROACH

Let γ_{i-j} denote the $(i,j)^{\mathrm{th}}$ element of $\mathbf{\Gamma}$, which is an $N \times N$ Toeplitz matrix. Also let $\mathbf{F}$ denote the unitary DFT (discrete Fourier transform) matrix with elements given by

$$F_{kp} = \frac{1}{\sqrt{2N-1}} e^{j\frac{2\pi}{2N-1}(k-1)(p-1)}, \; k, \; p = 1, \ldots, 2N-1. \tag{23}$$

Finally, define the (normalized) DFT of the zero-padded filter

$$\mathbf{h} = [h_1 \quad \cdots \quad h_{2N-1}]^T = \mathbf{F}^H\begin{bmatrix}\mathbf{w} \\ 0\end{bmatrix} = \mathbf{F}^H\tilde{\mathbf{w}} \tag{24}$$

and that of the zero-padded probing sequence

$$\mathbf{x} = [x_1 \quad \cdots \quad x_{2N-1}]^T = \mathbf{F}^H\begin{bmatrix}\mathbf{s} \\ 0\end{bmatrix} = \mathbf{F}^H\tilde{\mathbf{s}} \tag{25}$$

as well as the (normalized) power spectrum of the interference term $\boldsymbol{\epsilon}$ in (6) at the frequencies in $\mathbf{F}$ (see, e.g., [23]):

$$\Phi_p = \frac{1}{2N-1} \sum_{k=-N+1}^{N-1} \gamma_k e^{-j\frac{2\pi}{2N-1}k(p-1)},\ p = 1, \ldots, 2N-1. \tag{26}$$

Then it is possible to show (see Appendix A) that:

$$\mathrm{MSE}(\hat{\alpha}_0) + \beta = \frac{2N-1}{\left|\sum_{p=1}^{2N-1} h_p^* x_p\right|^2} \sum_{p=1}^{2N-1} |h_p|^2 \left(\beta |x_p|^2 + \Phi_p\right). \tag{27}$$

The energy constraint can also be written in the frequency domain (cf. Parseval equality):

$$\|\mathbf{s}\|^2 = \|\mathbf{F}\tilde{\mathbf{s}}\|^2 = \|\mathbf{x}\|^2 = N. \tag{28}$$

We can therefore think of formulating the design problem as follows:

$$\min_{\{h_p\},\{x_p\}} \frac{1}{\left|\sum_{p=1}^{2N-1} h_p^* x_p\right|^2} \sum_{p=1}^{2N-1} |h_p|^2 \left(\beta |x_p|^2 + \Phi_p\right)$$
$$\text{s.t.} \sum_{p=1}^{2N-1} |x_p|^2 = N. \tag{29}$$

Note that this is a relaxed version of the original design problem. Indeed, while for given $\mathbf{w}$ and $\mathbf{s}$ we can uniquely determine $\{|h_p|\}$ and $\{|x_p|\}$ from (24) and (25), the converse is not necessarily true. In other words, there might exist no $\mathbf{w}$ and $\mathbf{s}$ that can synthesize exactly a set of given $\{|h_p|\}$ and $\{|x_p|\}$. Some form of approximation will therefore be necessary when converting the result of the optimization problem in (29) to the original design variables $\mathbf{w}$ and $\mathbf{s}$, as detailed later in this section.

Continuing with (29), we note that the minimization of the objective function of this problem wrt $\{h_p\}$, for fixed $\{x_p\}$, can be simply done as follows. It is a consequence of the Cauchy-Schwartz inequality that

$$\begin{aligned} &\frac{\sum_{p=1}^{2N-1} |h_p|^2 \left(\beta |x_p|^2 + \Phi_p\right)}{\left|\sum_{p=1}^{2N-1} h_p^* x_p\right|^2} \\ &= \frac{\sum_{p=1}^{2N-1} |h_p|^2 \left(\beta |x_p|^2 + \Phi_p\right)}{\left|\sum_{p=1}^{2N-1} h_p^* \left(\beta |x_p|^2 + \Phi_p\right)^{1/2} \frac{x_p}{\left(\beta |x_p|^2 + \Phi_p\right)^{1/2}}\right|^2} \\ &\geq \frac{1}{\sum_{p=1}^{2N-1} \frac{|x_p|^2}{\beta |x_p|^2 + \Phi_p}} \end{aligned} \tag{30}$$

where the lower bound is attained for

$$h_p = \frac{x_p}{\beta |x_p|^2 + \Phi_p}. \tag{31}$$

Remark: It follows from (31) that at any frequency where the clutter dominates the interference, i.e., at the values of p where Φ_p in (31) is negligible wrt $\beta |x_p|^2$, we have

$$|h_p| = \frac{1}{\beta |x_p|}. \tag{32}$$

TABLE II
THE BISECTION METHOD FOR (37)

Step 0: Let $f(\lambda) = \frac{1}{\beta} \sum_{p=1}^{2N-1} \max\{\lambda\rho\Phi_p^{1/2} - \Phi_p, 0\}$. Let $f_{\text{obj}} = N$, $\lambda_{\text{left}} = 0$ and $\lambda_{\text{right}} = 1$.
Step 1: Set $\lambda = (\lambda_{\text{left}} + \lambda_{\text{right}})/2$ and compute $f(\lambda)$.
Step 2: If $f(\lambda) < f_{\text{obj}}$, then $\lambda_{\text{left}} \leftarrow \lambda$; else $\lambda_{\text{right}} \leftarrow \lambda$.
Iteration: Repeat Steps 1 and 2 until $|f(\lambda) - f_{\text{obj}}| \leq \epsilon$ where ϵ is a predefined threshold (e.g., 10^{-2}).

In words, if the noise is negligible then the squared magnitude of the frequency characteristic of the optimum receive filter is inversely proportional to the transmit signal spectrum. ■

Making use of (30) and (31) reduces (29) to the following maximization problem wrt $\{z_p = |x_p|^2\}$:

$$\max_{\{z_p\}} \sum_{p=1}^{2N-1} \frac{z_p}{\beta z_p + \Phi_p} \tag{33}$$
$$\text{s.t.} \quad \sum_{p=1}^{2N-1} z_p = N \tag{34}$$
$$z_p \geq 0 \quad \left(z_p = |x_p|^2\right). \tag{35}$$

Because the objective function in (33) can be easily shown to be concave and as the constraints are linear, the above problem is convex and can therefore be globally solved. A continuous version of this problem has appeared in [8] (see also [24]) which studied a related design problem. Like the continuous problem in [8], [24], the discrete one above can be solved using the Lagrange approach. Concretely, the solution to (33) can be shown to be (see Appendix C):

$$|x_p|^2 = z_p = \frac{1}{\beta} \max\left\{\lambda\rho\Phi_p^{1/2} - \Phi_p, 0\right\} \quad (p = 1, \ldots, 2N-1) \tag{36}$$

where λ is determined by the condition that $\{z_p\}$ satisfy the energy constraint

$$\frac{1}{\beta} \sum_{p=1}^{2N-1} \max\left\{\lambda\rho\Phi_p^{1/2} - \Phi_p, 0\right\} = N \tag{37}$$

and where

$$\rho = \frac{\beta N + \sum_{p=1}^{2N-1} \Phi_p}{\sum_{p=1}^{2N-1} \Phi_p^{1/2}}. \tag{38}$$

It can be shown that the solution of (37) cannot be larger than one, that is

$$\lambda \in [0, 1]. \tag{39}$$

Using this fact, one can easily solve (37), for example by means of the bisection method as outlined in Table II.

Remark: The expression in (36) for the optimum $\{|x_p|\}$ has the flavor of water-filling results commonly encountered in the communication literature (as also noted in [5] and [8] for related design problems). More specifically, one can see from (36) that the optimum probing sequence does not contain any power at those frequencies where the interference is strong, i.e., it satisfies $\Phi_p^{\frac{1}{2}} \geq \lambda\rho$. ■

It follows from (27) and (30) that the following *lower bound* on MSE $(\hat{\alpha}_0)$ holds

$$B_{\mathrm{MSE}} = \frac{2N-1}{\sum_{p=1}^{2N-1} \frac{\hat{z}_p}{\beta \hat{z}_p + \Phi_p}} - \beta \tag{40}$$

where $\{\hat{z}_p\}_{p=1}^{2N-1}$ denotes the optimum power spectrum obtained from (36) and (37). This lower bound can be attained only if $\{\hat{z}_p\}_{p=1}^{2N-1}$ can be exactly synthesized by selecting $\{s_k\}_{k=1}^{N}$; see below for details on this synthesis problem.

Once λ is found, $\{|x_p|\}$ is given by the closed-form expression in (36). Note that the phases of $\{x_p\}$, say $\{\psi_p\}$, are not variables of the design problem in (33) and therefore they can be chosen freely. The original design variable $\mathbf{s}$ can be obtained by solving (25) in a least-squares sense:

$$\min_{\mathbf{s},\{\psi_p\}} \|\mathbf{x} - \mathbf{F}^H \tilde{\mathbf{s}}\|^2. \tag{41}$$

Solving (41) wrt $\mathbf{s}$ and $\{\psi_p\}$ is a task that in general requires an iterative solver. We use a cyclic algorithm for this task, see below for details, which is reminiscent of the Sussman–Gerchberg–Saxton procedure [13], [25] and of the CAN algorithm [17], [26], [27].

Consider first the minimization of the objective function in (41) wrt $\mathbf{s}$, for fixed $\{\psi_p\}$ (i.e., $\mathbf{x}$ is fixed). Under the energy constraint, this minimization can be done in closed form:

$$\mathbf{s} = \sqrt{N} \frac{\boldsymbol{\nu}}{\|\boldsymbol{\nu}\|}; \quad \boldsymbol{\nu} = \tilde{\mathbf{F}}\mathbf{x} \tag{42}$$

where $\tilde{\mathbf{F}}$ is the $N \times (2N-1)$ matrix made from the first N rows of $\mathbf{F}$. To prove (42) note that (recall that $\|\mathbf{x}\|^2 = \|\mathbf{s}\|^2 = \|\tilde{\mathbf{s}}\|^2 = N$)

$$\|\mathbf{x} - \mathbf{F}^H \tilde{\mathbf{s}}\|^2 = 2N - 2\mathrm{Re}\{\mathbf{x}^H \mathbf{F}^H \tilde{\mathbf{s}}\} \tag{43}$$

where

$$\begin{aligned} \mathrm{Re}\{\mathbf{x}^H \mathbf{F}^H \tilde{\mathbf{s}}\} &\le |\mathbf{x}^H \mathbf{F}^H \tilde{\mathbf{s}}| = |\mathbf{x}^H \tilde{\mathbf{F}}^H \mathbf{s}| \\ &\le \|\mathbf{s}\| \cdot \|\tilde{\mathbf{F}}\mathbf{x}\| = \sqrt{N}\|\tilde{\mathbf{F}}\mathbf{x}\|. \end{aligned} \tag{44}$$

The upper bound in (44), and hence the minimum value of (43), is achieved for the $\mathbf{s}$ in (42), which also satisfies the energy constraint and is therefore the sought solution.

If, instead of the energy constraint, the unimodularity constraint is imposed on $\mathbf{s}$, the minimizing $\mathbf{s}$ will be given by

$$s_k = e^{j \arg(\nu_k)}, \quad k = 1, \ldots, N. \tag{45}$$

This can be seen from the following derivation:

$$\mathrm{Re}\{\mathbf{x}^H \mathbf{F}^H \tilde{\mathbf{s}}\} = \sum_{k=1}^{N} |\nu_k| \cos\left[\arg\{s_k\} - \arg\{\nu_k\}\right] \le \sum_{k=1}^{N} |\nu_k| \tag{46}$$

where the equality is achieved for (45).

If the PAR constraint is imposed, note that $\|\mathbf{x} - \mathbf{F}^H \tilde{\mathbf{s}}\|^2 = \|\mathbf{F}\mathbf{x} - \tilde{\mathbf{s}}\|^2 = \|\boldsymbol{\nu} - \mathbf{s}\|^2$ and thus (47) can be cast as

$$\begin{aligned} &\min_{\mathbf{s}} \ |\mathbf{s} - \boldsymbol{\nu}|^2 \\ &\text{s.t. } \mathrm{PAR}(\mathbf{s}) \le \mu \end{aligned} \tag{47}$$

TABLE III
THE CREW(FRE) ALGORITHM WITH UNIMODULARITY/PAR CONSTRAINT

Step 0: Solve (37) for λ (see Table II). Determine $\{|x_p|\}$ from (36). Use an available sequence or a randomly generated one to initialize $\mathbf{s}$.
Step 1: Compute

$$\psi_p = \arg(\tilde{\mathbf{f}}_p^H \mathbf{s}), \quad p = 1, \ldots, 2N-1. \tag{50}$$

Step 2: Compute

$$\boldsymbol{\nu} = \tilde{\mathbf{F}}\mathbf{x} \tag{51}$$

where $\{x_p = |x_p| e^{j\psi_p}\}$ with $\{|x_p|\}$ provided by Step 0. If only the energy constraint is imposed, use (42) to update $\mathbf{s}$; if the unimodularity constraint is imposed, use (45); and if the PAR constraint is enforced, use (47).
Iteration: Repeat Steps 1 and 2 until convergence.
Step 3: Compute $\mathbf{w}$ using (15) with the vector $\mathbf{s}$ provided by the above iteration.

which can be efficiently solved via a global finite-step "nearest-vector" algorithm introduced in [28].

For given $\mathbf{s}$, the minimization of (41) wrt $\{\psi_p\}$ also has a simple closed-form solution. Let $\mathbf{f}_p^H$ denote the p^{th} row of $\mathbf{F}^H$ and rewrite (41) as

$$\min_{\mathbf{s},\{\psi_p\}} \sum_{p=1}^{2N-1} \left| |x_p| e^{j\psi_p} - \mathbf{F}_p^H \tilde{\mathbf{s}} \right|^2. \tag{48}$$

The minimization of the above function wrt $\{\psi_p\}$, for any fixed $\mathbf{s}$, can be easily done in closed-form [the proof is similar to (46)]:

$$\psi_p = \arg\left\{\mathbf{F}_p^H \tilde{\mathbf{s}}\right\}, \quad p = 1, \ldots, 2N-1. \tag{49}$$

Note that the calculation of $\boldsymbol{\nu}$ in (42) and of $\{\mathbf{f}_p^H \tilde{\mathbf{s}}\}$ in (49) can be done efficiently via FFT (fast Fourier transform). After obtaining $\mathbf{s}$, the receive filter $\mathbf{w}$ is given by (15). The so-obtained *fre*quency domain-based algorithm, which is called CREW(fre) for short, is summarized in Table III. Note that CREW(fre) can handle both unimodularity and PAR constraints and it is computationally efficient due to the leveraged FFT operations.

V. CREW(MAT)-SPECIALIZATION OF CREW(FRE) TO MATCHED FILTERING

In this section we specialize the CREW(fre) algorithm of Section IV to the matched-filter case in which $\mathbf{w} = \mathbf{s}$ and thus $\mathbf{h} = \mathbf{x}$. It follows from (27) that

$$\mathrm{MSE}(\hat{\alpha}_0) = \frac{2N-1}{N^2} \sum_{p=1}^{2N-1} |x_p|^2 \left(\beta |x_p|^2 + \Phi_p\right) - \beta. \tag{52}$$

Therefore, the problem of minimizing the MSE metric becomes [using the variables $z_p = |x_p|^2$, as in (33)]:

$$\begin{aligned} &\min_{\{z_p\}} \sum_{p=1}^{2N-1} \left(\beta z_p^2 + \Phi_p z_p\right) \\ &\text{s.t.} \sum_{p=1}^{2N-1} z_p = N \quad (z_p \ge 0). \end{aligned} \tag{53}$$

TABLE IV
THE CREW(MAT) ALGORITHM WITH UNIMODULARITY/PAR CONSTRAINT

Step 0: Calculate λ from (56) using a bisection method. Then determine $\{|x_p|\}$ from (55). Use an available sequence or a randomly generated one to initialize $\mathbf{s}$.
Step 1, Step 2, Iteration: Same as those in Table III.

The above problem is convex. Its Lagrangian has the following decoupled form:

$$\sum_{p=1}^{2N-1}\left[\beta z_p^2+\Phi_p z_p-\lambda z_p+\frac{N}{2N-1}\lambda\right] \tag{54}$$

and is clearly minimized over $\{z_p \geq 0\}$ at

$$z_p=\frac{1}{2\beta}\max\{\lambda-\Phi_p,0\}. \tag{55}$$

The multiplier λ in (55) is the solution of the constraint equation:

$$\frac{1}{2\beta}\sum_{p=1}^{2N-1}\max\{\lambda-\Phi_p,0\}=N. \tag{56}$$

Let λ_0 satisfy the equation

$$\frac{1}{2\beta}\sum_{p=1}^{2N-1}(\lambda_0-\Phi_p)=N \tag{57}$$

which yields

$$\lambda_0=\frac{2N\beta+\sum_{p=1}^{2N-1}\Phi_p}{2N-1}. \tag{58}$$

If $\lambda_0 \geq \Phi_p$ (for $p=1,\ldots,2N-1$) then the solution of (56) is $\lambda=\lambda_0$. In all other cases λ must satisfy $\lambda<\lambda_0$ (because the left-hand side of (56) is an increasing function of λ that is equal to 0 at $\lambda=0$ and is larger than or equal to N at $\lambda=\lambda_0$). It follows that the solution to (56) must lie in the interval

$$\lambda \in (0,\lambda_0] \tag{59}$$

and it can therefore be conveniently found by the bisection method [similarly to (37)].

Remark: Note again the water-filling character of (55): the optimum probing sequence does not have any power at the frequencies where the interference is strong enough for $\lambda-\Phi_p$ to be negative [see the remark following (39) for a similar comment on the CREW(fre) solution]. ■

Once λ is found, $\{|x_p|=z_p^{\frac{1}{2}}\}$ are determined from (55). The problem that remains is to find $\mathbf{s}$ from the solution of the problem in (41), which can be done as explained in the previous section. The so-obtained *mat*ched filter-based algorithm, named CREW(mat), is summarized in Table IV. Note that CREW(mat) can be interpreted as an extension of the CAN algorithm in [17] to the case in which some $\{\Phi_p\}$ are different from zero (indeed, CREW(mat) reduces to CAN in the no-interference case of $\{\Phi_k \equiv 0\}$).

VI. NUMERICAL EXAMPLES

The covariance matrix of the interference term, viz. $\mathbf{\Gamma}$ [see (7)], is constructed as follows:

$$\mathbf{\Gamma}=\sigma_{\mathrm{J}}^2\mathbf{\Gamma}_{\mathrm{J}}+\sigma^2\mathbf{I} \tag{60}$$

where σ_{J}^2 and σ^2 are the jamming and noise powers, respectively, and

$$\mathbf{\Gamma}_{\mathrm{J}}=\begin{bmatrix} q_0 & q_1^* & \cdots & q_{N-1}^* \\ q_1 & q_0 & & q_{N-2}^* \\ \vdots & & \ddots & \vdots \\ q_{N-1} & q_{N-2} & \cdots & q_0 \end{bmatrix} \quad (q_0=1) \tag{61}$$

is the normalized jamming covariance matrix. $\mathbf{\Gamma}_{\mathrm{J}}$ can be determined directly by the IDFT (inverse DFT) of a desired jamming power spectrum. More concretely, let $\{\eta_p\}_{p=1}^{2N-1}$ denote a given jamming power spectrum at frequencies $\left\{\frac{p-1}{2N-1}\right\}_{p=1}^{2N-1}$. Similarly to (75), the IDFT of $\{\eta_p\}_{p=1}^{2N-1}$ leads to

$$\begin{bmatrix} q_0 & q_1 & \cdots & q_{N-1} & q_{N-1}^* & \cdots & q_1^* \end{bmatrix}^T \tag{62}$$

which needs to be normalized so that $q_0=1$.

Throughout the following examples we choose $\sigma_{\mathrm{J}}^2=100$, $\sigma^2=0.1$ and $\beta=1$ [see (8)]. The proposed algorithms will be compared with the CAN algorithm (see [17]). The CAN sequence is designed to suppress the correlation sidelobes. In the case of matched filter (i.e., $\mathbf{w}=\mathbf{s}$), low correlation sidelobes lead to small $\mathrm{MSE}(\hat{\alpha}_0)$ [see (12)] if the clutter dominates the interference (i.e., $\mathbf{\Gamma}$ is not considered in the design). The CAN algorithm can handle both unimodularity and PAR constraints. As already mentioned, the proposed CREW(mat) algorithm extends CAN by taking $\mathbf{\Gamma}$ into account (a fact that should lead to smaller MSE than CANs).

Note that, while the CAN sequence is designed under the implicit assumption of using a matched filter, it can also be used with the MMF filter as in (15). The combination of a CAN sequence with the corresponding optimum MMF filter is referred to as CAN-MMF.

In what follows, the jamming type and the sequence length will be varied to examine the effect on algorithm performance in terms of $\mathrm{MSE}(\hat{\alpha}_0)$. The considered algorithms are listed in Table V for easy reference. The Golomb sequence (see, e.g., [12] and [29]) will be used for the initialization of all algorithms. This sequence appears to be a good choice for initialization, although we did observe occasionally that a randomly generated sequence can lead to a lower MSE than the Golomb sequence. The MSE lower bound (see (40)) will be used as a benchmark. Note that CAN, CAN-MMF and the lower bound will be shown in every considered case; however, for CREW(gra), CREW(fre) and CREW(mat) we only show the one that yields the lowest MSE to avoid over-crowding the figures. Also recall that CREW(gra) can only deal with the unimodularity constraint and also that it cannot handle large values of N due to its relatively expensive computation (consequently when we show the MSE of CREW(gra) we limit N to 300).

TABLE V
ALGORITHMS TO BE COMPARED

	transmit sequence $\mathbf{s}$	receive filter $\mathbf{w}$	power constraints
CAN	the CAN sequence (see [17])	the matched filter, i.e., $\mathbf{w} = \mathbf{s}$	unimodularity/PAR
CAN-MMF	the CAN sequence	the optimum MMF filter, i.e., $\mathbf{w} = \mathbf{R}^{-1}\mathbf{s}$ (see (15))	unimodularity/PAR
CREW(gra)	the $\mathbf{s}$ obtained from Table I	the optimum MMF filter	unimodularity
CREW(fre)	the $\mathbf{s}$ obtained from Table III	the optimum MMF filter	unimodularity/PAR
CREW(mat)	the $\mathbf{s}$ obtained from Table IV	the matched filter	unimodularity/PAR

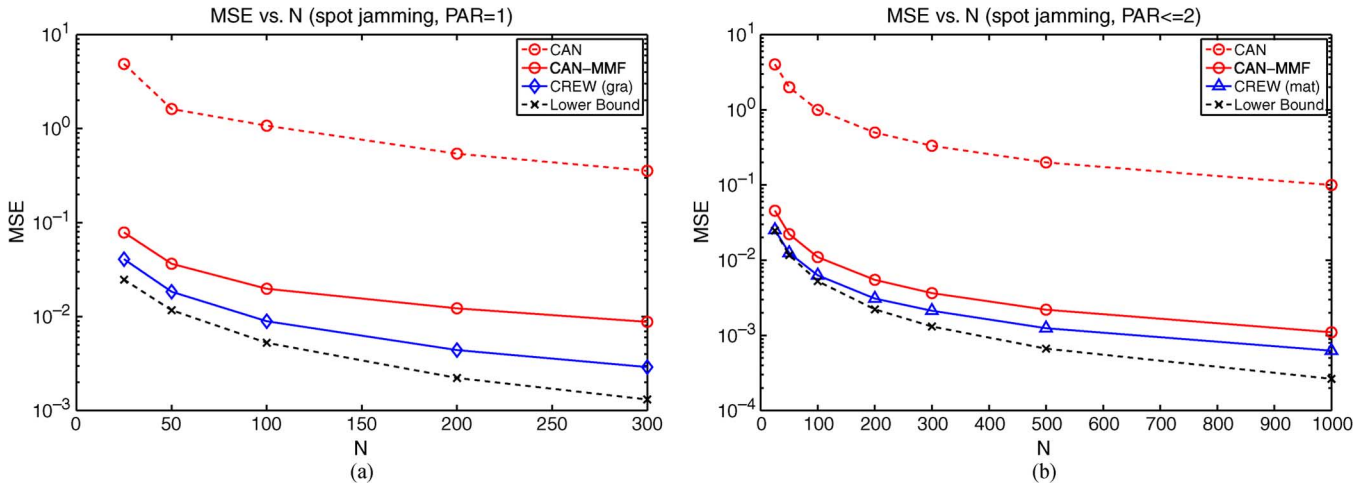

Fig. 1. MSE comparison in the case of spot jamming. (a) MSE of CAN, CAN-MMF and CREW(gra) under the unimodularity constraint. (b) MSE of CAN, CAN-MMF and CREW(mat) under the constraint of $\mathrm{PAR} \leq 2$.

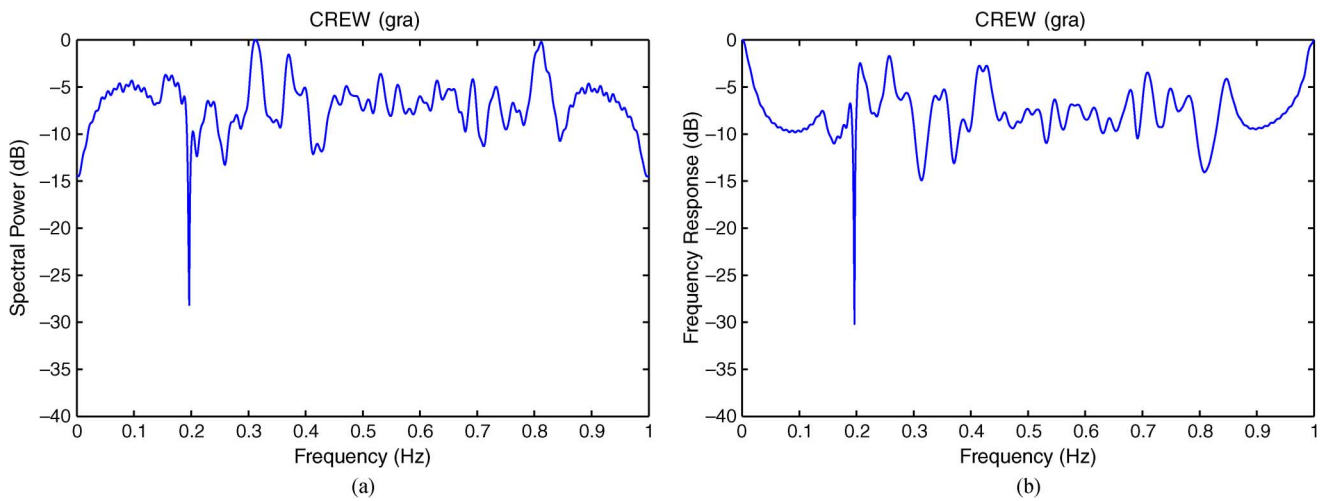

Fig. 2. (a) The spectral power of the length-100 CREW(gra) sequence associated with Fig. 1(a). (b) The frequency response of the corresponding receive filter.

A. Spot Jamming

Assume a spot jamming located at frequency f_0 whose power spectrum is given by

$$\eta_p = \begin{cases} 1, & p = \lfloor (2N-1)f_0 \rfloor \\ 0, & \text{elsewhere.} \end{cases} \qquad p = 1, \ldots, 2N-1. \quad (63)$$

Let $f_0 = 0.2$ Hz.

Fig. 1(a) shows the MSE of CAN, CAN-MMF and CREW(gra) for $N = 25, 50, 100, 200, 300$ under the unimodularity constraint. It can be observed that CAN-MMF provides a significantly smaller MSE than CAN and that CREW(gra) further improves over CAN-MMF. Fig. 1(b) shows the MSE of CAN, CAN-MMF and CREW(mat) for $N = 25, 50, 100, 200, 300, 500, 1000$ under the constraint of $\mathrm{PAR} \leq 2$. It can be observed that CREW(mat) gives the smallest MSE for all N's.

Fig. 2 shows the spectral power of the length-100 CREW(gra) sequence associated with Fig. 1(a) and the frequency response of the corresponding receive filter. Observe the spectral notch at the jamming frequency $f_0 = 0.2$ Hz for both the probing sequence and the receive filter. Except at frequency f_0, the spectrum of the probing sequence and the frequency response of the receive filter are approximately the reciprocal of each other, as explained in the Remark following (31).

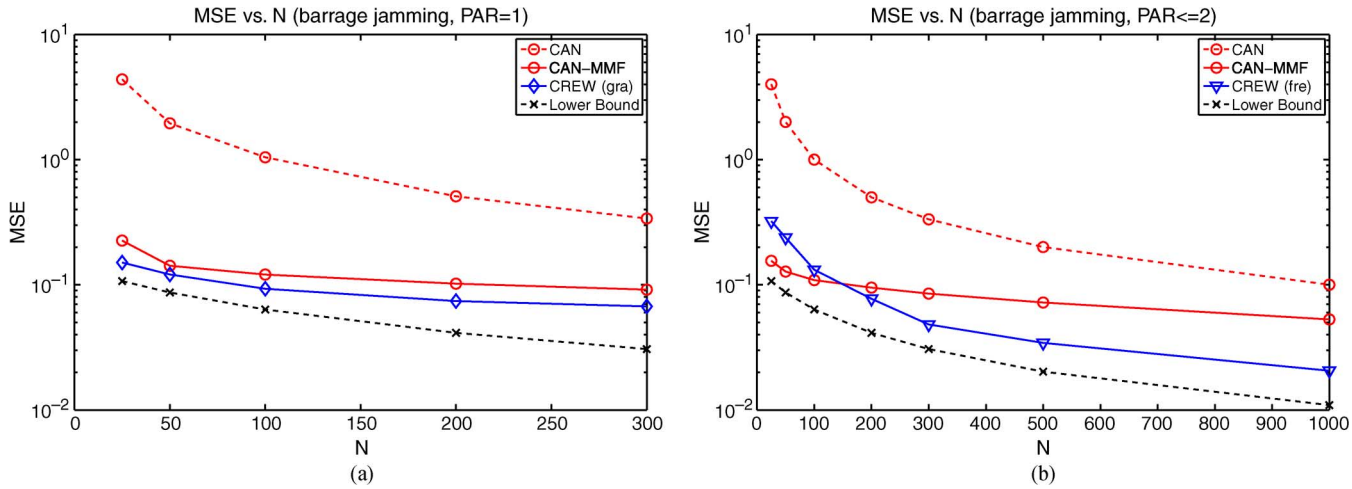

Fig. 3. MSE comparison in the case of barrage jamming. (a) MSE of CAN, CAN-MMF and CREW(gra) under the unimodularity constraint. (b) MSE of CAN, CAN-MMF and CREW(fre) under the constraint of PAR $\leq$ 2.

B. Barrage Jamming

Consider a barrage jamming located in the frequency band $[f_1, f_2]$ whose power spectrum is given by

$$\eta_p = \begin{cases} 1, & \lfloor (2N-1)f_1 \rfloor \leq p \leq \lfloor (2N-1)f_2 \rfloor \\ 0, & \text{elsewhere} \end{cases} \qquad p = 1, \ldots, 2N-1. \quad (64)$$

Let $f_1 = 0.2$ Hz and $f_2 = 0.3$ Hz.

Fig. 3(a) shows the MSE of CAN, CAN-MMF, and CREW(gra) under the unimodularity constraint. Once again CREW(gra) gives a lower MSE than CAN or CAN-MMF does. Fig. 3(b) shows the MSE of CAN, CAN-MMF, and CREW(fre) under the constraint of PAR $\leq$ 2. Except for N's smaller than 100, CREW(fre) gives the lowest MSE among the three algorithms. Moreover, the performance of CREW(fre) tends to follow the MSE lower bound as N increases, in contrast to CAN-MMF's performance that deviates more and more from the lower bound as N increases.

C. Robust Design

Finally we discuss the robustness issue of the aforementioned designs. Consider a strong spot jamming that operates at frequency f_0. In this case the optimum filter in (15) will have a notch at frequency f_0. However, if the information about the jamming frequency is not accurate, the receive filter in (15) will have a notch at the assumed frequency and so it will pass the jamming interference that in reality operates at a different frequency. To combat such a problem, we can assume a barrage jamming located in a frequency band $[f_1, f_2]$ Hz, which is wide enough to include the true frequency of the spot jamming. In this way the receive filter will have a wider stop-band and the jamming interference will be annihilated. Similarly in the case of imprecise information about a barrage jamming, we can consider a frequency band that is wide enough to cover the true jamming band, for the sake of robustness.

As an example, let (64) be the true jamming power spectrum with $f_1 = 0.2$ and $f_2 = 0.3$ Hz. We use CREW(fre) to design the transmit sequence and the receive filter. Three scenarios are considered: the exact values of f_1 and f_2 are known, a wrong jamming band [0.25, 0.35] Hz is used, and an imprecise but sufficiently wide jamming band [0.15, 0.35] Hz is used (a case referred to as robust design). Fig. 4 shows the MSE (calculated using the true jamming spectrum) for N = 25, 50, 100, 200, 300, 500, 1000 under the constraint of PAR $\leq$ 2. It can be observed that the robust design is only weakly affected by the imprecise jamming information.

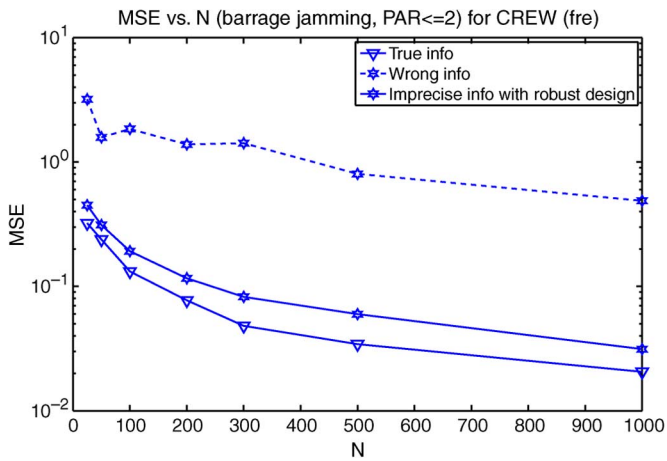

Fig. 4. MSE comparison of CREW(fre) concerning an imprecise knowledge of the jamming frequency.

VII. Concluding Remarks

We have discussed three algorithms, namely CREW(gra), CREW(fre) and CREW(mat), that can be used for the joint optimization of receive filters and transmit signals. The design goal is to minimize the MSE of target's scattering coefficient estimate or, equivalently, to maximize the signal-to-clutter-plus-interference ratio. Practical unimodularity or low-PAR constraints are taken into account in the transmit sequence design. The CREW(gra) algorithm uses a gradient-based approach to minimize the MSE that results from using the optimum MMF filter. Usually it yields the best performance for unimodular sequences but it is relatively computationally expensive and cannot be

used for the more general PAR constraint. The CREW(fre) algorithm operates in the frequency domain and seeks to obtain an optimum power spectrum for the transmit sequence. The said optimum power spectrum leads to an MSE lower bound. CREW(fre) can be specialized to the matched-filter case, which yields the CREW(mat) algorithm that can be considered as an extension of the previously suggested CAN approach to the case of non-white interference. Both CREW(fre) and CREW(mat) are computationally efficient and can easily handle sequence lengths $\sim 10^3$. Numerical examples have been provided to illustrate the effectiveness of the proposed receive filters and transmit sequences.

APPENDIX A
PROOF OF (27)

We begin the proof by including the zero-delay sequence into the matrix $\mathbf{R}$ as defined in (13):

$$\mathbf{R}+\beta\mathbf{s}\mathbf{s}^H=\beta\sum_{k=-N+1}^{N-1}\mathbf{J}_k\mathbf{s}\mathbf{s}^H\mathbf{J}_k^H+\mathbf{\Gamma}=\beta\mathbf{A}^H\mathbf{A}+\mathbf{\Gamma} \quad (65)$$

where

$$\mathbf{A}^H=\begin{bmatrix} s_1 & 0 & \cdots & 0 & s_N & s_{N-1} & \cdots & s_2 \\ s_2 & s_1 & & \vdots & 0 & s_N & & \vdots \\ \vdots & \vdots & \ddots & 0 & \vdots & & & s_N \\ s_N & s_{N-1} & \cdots & s_1 & 0 & 0 & \cdots & 0 \end{bmatrix} \quad (66)$$

is an $N\times(2N-1)$ matrix. Using (65) we rewrite the MSE in (12) as

$$\mathrm{MSE}(\hat{\alpha}_0)=\frac{\mathbf{w}^H(\beta\mathbf{A}^H\mathbf{A}+\mathbf{\Gamma})\mathbf{w}}{|\mathbf{w}^H\mathbf{s}|^2}-\beta. \quad (67)$$

Note that

$$\mathbf{A}^H\mathbf{A}=\begin{bmatrix} c_0 & c_1^* & \cdots & c_{N-1}^* \\ c_1 & c_0 & & c_{N-2}^* \\ \vdots & & \ddots & \vdots \\ c_{N-1} & c_{N-2} & & c_0 \end{bmatrix} \quad (68)$$

(see Appendix B for a proof) where $\{c_k\}_{k=0}^{N-1}$ are the correlations of $\{s_n\}_{n=1}^{N}$ (with $c_0=N$):

$$c_k=\sum_{n=k+1}^{N}s_ns_{n-k}^*=c_{-k}^*,\quad k=0,1,\ldots,N-1. \quad (69)$$

To express the objective function (67) in the frequency domain, we will make use of some standard properties of Toeplitz and circulant matrices. First note that $\mathbf{\Gamma}$ can be readily embedded in a $(2N-1)\times(2N-1)$ circulant matrix:

$$\mathbf{C}^i=\begin{bmatrix} \gamma_0 & \gamma_1^* & \cdots & \gamma_{N-1}^* & \gamma_{N-1} & \gamma_{N-2} & \cdots & \gamma_1 \\ \gamma_1 & \gamma_0 & & \vdots & \gamma_{N-1}^* & \gamma_{N-1} & & \vdots \\ \vdots & & \ddots & \gamma_1^* & \vdots & & & \gamma_{N-1} \\ \gamma_{N-1} & \cdots & \gamma_1 & \gamma_0 & \gamma_1^* & \cdots & & \gamma_{N-1}^* \\ \gamma_{N-1}^* & \gamma_{N-1} & \cdots & \gamma_1 & \gamma_0 & \gamma_1^* & \cdots & \gamma_{N-2}^* \\ \vdots & & & \vdots & & & & \vdots \\ \gamma_1^* & \cdots & \gamma_{N-1}^* & \gamma_{N-1} & \gamma_{N-2} & \cdots & & \gamma_0 \end{bmatrix} \quad (70)$$

where the superscript i designates "interference". Somewhat similarly, we can also embed $\mathbf{A}^H\mathbf{A}$ in a $(2N-1)\times(2N-1)$ square circulant matrix:

$$\mathbf{C}^s=\begin{bmatrix} c_0 & c_1^* & \cdots & c_{N-1}^* & c_{N-1} & c_{N-2} & \cdots & c_1 \\ c_1 & c_0 & & \vdots & c_{N-1}^* & c_{N-1} & & \vdots \\ \vdots & & \ddots & c_1^* & \vdots & & & c_{N-1} \\ c_{N-1} & \cdots & c_1 & c_0 & c_1^* & \cdots & & c_{N-1}^* \\ c_{N-1}^* & c_{N-1} & \cdots & c_1 & c_0 & c_1^* & \cdots & c_{N-2}^* \\ \vdots & & & \vdots & & & & \vdots \\ c_1^* & \cdots & c_{N-1}^* & c_{N-1} & c_{N-2} & \cdots & & c_0 \end{bmatrix} \quad (71)$$

where the superscript s denotes "signal and signal-dependent clutter."

In addition, let [see (24) and (25)]

$$\tilde{\mathbf{w}}=\begin{bmatrix}\mathbf{w}\\ 0\end{bmatrix}_{(2N-1)\times 1} \quad (72)$$

and

$$\tilde{\mathbf{s}}=\frac{\mathbf{s}}{0}_{(2N-1)\times 1} \quad (73)$$

denote, respectively, the receive filter and probing sequence vectors padded with $N-1$ zeros. Using this notation we can write the objective function in (67) as

$$\mathrm{MSE}(\hat{\alpha}_0)+\beta=\frac{\tilde{\mathbf{w}}^H(\beta\mathbf{C}^s+\mathbf{C}^i)\tilde{\mathbf{w}}}{|\tilde{\mathbf{w}}^H\tilde{\mathbf{s}}|^2}. \quad (74)$$

Note that $\{\Phi_p\}_{p=1}^{2N-1}$ (see (26)) is the (real-valued and nonnegative) DFT of the first column of $\mathbf{C}^i$. Indeed

$$\begin{aligned}&\sum_{k=0}^{N-1}\gamma_k e^{-j\frac{2\pi}{2N-1}kp}+\sum_{k=N}^{2N-2}\gamma_{(2N-1)-k}^* e^{-j\frac{2\pi}{2N-1}kp}\\ &=\sum_{k=0}^{N-1}\gamma_k e^{-j\frac{2\pi}{2N-1}kp}+\sum_{k=N}^{2N-2}\gamma_{k-(2N-1)} e^{-j\frac{2\pi}{2N-1}(k-2N+1)p}\\ &=\sum_{k=0}^{N-1}\gamma_k e^{-j\frac{2\pi}{2N-1}kp}+\sum_{k=-N+1}^{-1}\gamma_k e^{-j\frac{2\pi}{2N-1}kp}\end{aligned} \quad (75)$$

which equals (26) up to the multiplicative constant $\frac{1}{(2N-1)}$. Similarly, the (normalized) power spectrum of the probing sequence $\mathbf{s}$, given by $\{|x_p|^2\}_{p=1}^{2N-1}$ (see, e.g., [23]), is the DFT of the first column of $\mathbf{C}^s$ up to the multiplicative constant $\frac{1}{(2N-1)}$.

Making use of the above notation and observations along with a standard property of circulant matrices (see, e.g. [30]) allows us to rewrite (74) in a frequency-domain form:

$$\begin{aligned}
&\frac{\tilde{\mathbf{w}}^H(\beta\mathbf{C}^s+\mathbf{C}^i)\tilde{\mathbf{w}}}{\left|\tilde{\mathbf{w}}^H\tilde{\mathbf{s}}\right|^2}\\
&\quad=\frac{2N-1}{\left|\tilde{\mathbf{w}}^H\mathbf{F}^H\mathbf{F}\tilde{\mathbf{s}}\right|^2}\tilde{\mathbf{w}}^H\\
&\qquad\cdot\left(\beta\mathbf{F}\begin{bmatrix}|x_1|^2 & & 0\\ & \ddots & \\ 0 & & |x_{2N-1}|^2\end{bmatrix}\mathbf{F}^H\right.\\
&\qquad\qquad\left.+\mathbf{F}\begin{bmatrix}\Phi_1 & & 0\\ & \ddots & \\ 0 & & \Phi_{2N-1}\end{bmatrix}\mathbf{F}^H\right)\tilde{\mathbf{w}}\\
&\quad=\frac{2N-1}{\left|\sum_{p=1}^{2N-1}h_p^*x_p\right|^2}\sum_{p=1}^{2N-1}|h_p|^2\left(\beta|x_p|^2+\Phi_p\right). \quad (76)
\end{aligned}$$

With this observation the proof is concluded.

Appendix B
Proof of (68)

Observe that the product matrix $\mathbf{A}^H\mathbf{A}$ is not affected by a permutation of the columns in $\mathbf{A}^H$. This implies that

$$\mathbf{A}^H\mathbf{A}=\tilde{\mathbf{A}}^H\tilde{\mathbf{A}} \quad (77)$$

where

$$\tilde{\mathbf{A}}^H=\begin{bmatrix}s_N & s_{N-1} & \cdots & s_1 & 0 & \cdots & 0\\ 0 & s_N & & s_2 & s_1 & & \vdots\\ \vdots & & \ddots & \vdots & & \ddots & \vdots\\ 0 & 0 & & s_N & s_{N-1} & \cdots & s_1\end{bmatrix}_{N\times(2N-1)}. \quad (78)$$

Therefore the $(i,j)^{\text{th}}$ element of $\mathbf{A}^H\mathbf{A}$ is given by (assuming that $s_k=0$ for $k\neq[1,N]$)

$$\begin{aligned}
[\mathbf{A}^H\mathbf{A}]_{ij}&=\sum_{p=1}^{2N-1}s_{N+i-p}s_{N+j-p}^*\\
&=\sum_{k=1}^{N}s_ks_{k-(i-j)}^*=c_{i-j} \quad (79)
\end{aligned}$$

and the proof of (68) is concluded.

Appendix C
Lagrange Approach to Solving (33)

The second-order derivative of the objective function in (33) can be shown to be nonpositive:

$$\sum_{p=1}^{2N-1}\frac{\partial^2}{\partial^2 z_p}\left\{\frac{z_p}{\beta z_p+\Phi_p}\right\}=-\sum_{p=1}^{2N-1}\frac{2\beta\Phi_p}{(\beta z_p+\Phi_p)^3}\le 0 \quad (80)$$

since β, $\{z_p\}$ and $\{\Phi_p\}$ are all non-negative. Hence the objective is a concave function. As the constraints in (34) and (35) are linear, it follows that (33) is a convex optimization problem that can be solved by maximizing the following Lagrange function wrt $\{z_p\ge 0\}$:

$$\begin{aligned}
L(\{z_p\}_{p=1}^{2N-1})&=\sum_{p=1}^{2N-1}\frac{z_p}{\beta z_p+\Phi_p}-\lambda\left(\sum_{p=1}^{2N-1}z_p-N\right)\\
&=\sum_{p=1}^{2N-1}\left(\frac{z_p}{\beta z_p+\Phi_p}-\lambda z_p\right)+\lambda N \quad (81)
\end{aligned}$$

where $\lambda>0$ is the Lagrange multiplier. In fact, owing to the decoupled form of (81), we only need to consider maximizing the following function:

$$\tilde{f}(z_p)=\frac{z_p}{\beta z_p+\Phi_p}-\lambda z_p. \quad (82)$$

Similarly to (80), it can be shown that $\tilde{f}(z_p)$ is concave; the maximizer of this function can be obtained by setting its first-order derivate to zero:

$$\frac{\partial\tilde{f}(z_p)}{\partial z_p}=\frac{\Phi_p}{(\beta z_p+\Phi_p)^2}-\lambda=0 \quad (83)$$

which leads to the solution

$$\hat{z}_p=\frac{\sqrt{\Phi_p/\lambda}-\Phi_p}{\beta}. \quad (84)$$

If $\hat{z}_p\ge 0$, then $\hat{z}_p$ is the sought solution; if $\hat{z}_p<0$, considering that $\tilde{f}(z_p)$ (a concave function) is increasing for $z_p\le\hat{z}_p$ and is decreasing for $z_p\ge\hat{z}_p$, $\tilde{f}(z_p)$ must attain its maximum (wrt $z_p\ge 0$) at $z_p=0$. Therefore, the maximizer of $\tilde{f}(z_p)$, as well as the solution to (33), is given by

$$z_p=\max\left\{\frac{\sqrt{\Phi_p/\lambda}-\Phi_p}{\beta},0\right\},\quad p=1,\ldots,2N-1. \quad (85)$$

The solution as shown in (36) is obtained from (85) by replacing $\sqrt{\frac{1}{\lambda}}$ with $\lambda\rho$ so that the parameter λ lies in the range [0, 1] [see (39)].

References

[1] W. D. Rummler, "A technique for improving the clutter performance of coherent pulse train signals," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 3, no. 6, pp. 898–906, Nov. 1967.

[2] D. F. DeLong, Jr. and E. M. Hofstetter, "On the design of optimum radar waveforms for clutter rejection," *IEEE Trans. Inf. Theory*, vol. 13, no. 3, pp. 454–463, Jul. 1967.

[3] L. J. Spafford, "Optimum radar signal processing in clutter," *IEEE Trans. Inf. Theory*, vol. 14, no. 5, pp. 734–743, Sep. 1968.

[4] C. A. Stutt and L. J. Spafford, "A 'best' mismatched filter response for radar clutter discrimination," *IEEE Trans. Inf. Theory*, vol. 14, no. 2, pp. 280–287, Mar. 1968.

[5] M. R. Bell, "Information theory and radar waveform design," *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1578–1597, Sep. 1993.

[6] S. U. Pillai, H. S. Oh, D. C. Youla, and J. R. Guerci, "Optimal transmit-receiver design in the presence of signal-dependent interference and channel noise," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 577–584, Mar. 2000.

[7] S. D. Blunt and K. Gerlach, "Adaptive pulse compression via MMSE estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 42, no. 2, pp. 572–584, Apr. 2006.

[8] S. Kay, "Optimal signal design for detection of Gaussian point targets in stationary Gaussian clutter/reverberation," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 31–41, Jun. 2007.
[9] P. Stoica, J. Li, and M. Xue, "Transmit codes and receive filters for radar," *IEEE Signal Process. Mag.*, vol. 25, no. 6, pp. 94–109, Nov. 2008.
[10] A. Aubry, A. De Maio, A. Farina, and M. Wicks, "Knowledge-aided (potentially cognitive) transmit signal and receive filter design in signal-dependent clutter," *IEEE Trans. Aerosp. Electron. Syst.*, 2011, submitted for publication.
[11] P. M. Woodward, *Probability and Information Theory with Applications to Radar*. New York: Pergamon, 1957.
[12] N. Levanon and E. Mozeson, *Radar Signals*. New York: Wiley, 2004.
[13] S. Sussman, "Least-square synthesis of radar ambiguity functions," *IEEE Trans. Inf. Theory*, vol. 8, no. 3, pp. 246–254, 1962.
[14] J. D. Wolf, G. M. Lee, and C. E. Suyo, "Radar waveform synthesis by mean-square optimization techniques," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 5, no. 4, pp. 611–619, 1968.
[15] I. Gladkova and D. Chebanov, "On the synthesis problem for a waveform having a nearly ideal ambiguity functions," presented at the Int. Conf. Radar Syst., Toulouse, France, Oct. 2004.
[16] D. V. Sarwate, "Meeting the Welch bound with equality," in *Proc. Phin Sequences Their Appl. (SETA)*, S. Springer, Ed., 1999, pp. 79–102.
[17] 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.
[18] M. I. Skolnik, *Radar Handbook*, 3rd ed. New York: McGraw-Hill, 2008.
[19] L. K. Patton and B. D. Rigling, "Autocorrelation and modulus constraints in radar waveform optimization," in *Proc. Int. Waveform Diversity Des. Conf.*, Kissimmee, FL, Feb. 2009, pp. 150–154.
[20] D. F. DeLong, Jr. and E. M. Hofstetter, "The design of clutter-resistant radar waveforms with limited dynamic range," *IEEE Trans. Inf. Theory*, vol. 15, no. 3, pp. 376–385, May 1969.
[21] S. Haykin, "Cognitive radar: A way of the future," *IEEE Signal Process. Mag.*, vol. 23, no. 1, pp. 30–40, Jan. 2006.
[22] R. Fletcher, "A new approach to variable metric algorithms," *Comput. J.*, vol. 13, no. 3, pp. 317–322, Aug. 1970.
[23] P. Stoica and R. L. Moses, *Spectral Analysis of Signals*. Upper Saddle River, NJ: Prentice-Hall, 2005.
[24] R. A. Romero, J. Bae, and N. A. Goodman, "Theory and application of SNR and mutual information matched illumination waveforms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 47, no. 2, pp. 912–927, Apr. 2011.
[25] R. W. Gerchberg and W. O. Saxton, "A practical algorithm for the determination of the phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
[26] 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.
[27] S. U. Pillai, K. Y. Li, R. Zheng, and B. Himed, "Design of unimodular sequences using generalized receivers," in *Proc. IEEE Radar Conf.*, Washington, DC, May 2010, pp. 729–734.
[28] 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. 188–209, Jan. 2005.
[29] S. W. Golomb, *Shift Register Sequences*. San Francisco, CA: Holden-Day, 1967.
[30] G. H. Golub and C. F. Van Loan, *Matrix Computations*. Baltimore, MD: Johns Hopkins Univ. Press, 1984.

Petre Stoica (F'94) received the D.Sc. degree in automatic control from the Polytechnic Institute of Bucharest (BPI), Bucharest, Romania, in 1979 and an honorary Ph.D. degree in science from Uppsala University (UU), Uppsala, Sweden, in 1993.

He was formerly a Professor of system identification and signal processing with the Faculty of Automatic Control and Computers at BPI. He is currently a Professor of systems modeling with the Division of Systems and Control, the Department of Information Technology at UU. He held longer visiting positions with Eindhoven University of Technology, Eindhoven, The Netherlands; Chalmers University of Technology, Gothenburg, Sweden (where he held a Jubilee Visiting Professorship); UU; the University of Florida, Gainesville, FL; and Stanford University, Stanford, CA. His main scientific interests are in the areas of system identification, time series analysis and prediction, statistical signal and array processing, spectral analysis, wireless communications, and radar signal processing. He has published nine books, ten book chapters, and approximately 500 papers in archival journals and conference records. The most recent book he coauthored, with R. Moses, is *Spectral Analysis of Signals* (Prentice-Hall, 2005).

Dr. Stoica is on the editorial boards of six journals: the *Journal of Forecasting*; *Signal Processing*; *Circuits, Signals, and Signal Processing*; *Digital Signal Processing: A Review Journal*; *Signal Processing Magazine*; and *Multidimensional Systems and Signal Processing*. He was a co-guest editor for several special issues on system identification, signal processing, spectral analysis, and radar for some of the aforementioned journals, as well as for the *IEE Proceedings*. He was corecipient of the IEEE ASSP Senior Award for a paper on statistical aspects of array signal processing. He was also recipient of the Technical Achievement Award of the IEEE Signal Processing Society. In 1998, he was the recipient of a Senior Individual Grant Award of the Swedish Foundation for Strategic Research. He was also co-recipient of the 1998 EURASIP Best Paper Award for Signal Processing for a work on parameter estimation of exponential signals with time-varying amplitude, a 1999 IEEE Signal Processing Society Best Paper Award for a paper on parameter and rank estimation of reduced-rank regression, a 2000 IEEE Third Millennium Medal, and the 2000 W. R. G. Baker Prize Paper Award for a paper on maximum likelihood methods for radar. He was a member of the international program committees of many topical conferences. From 1981 to 1986, he was a Director of the International Time-Series Analysis and Forecasting Society, and he was also a member of the IFAC Technical Committee on Modeling, Identification, and Signal Processing. He is also a member of the Royal Swedish Academy of Engineering Sciences, an honorary member of the Romanian Academy, and a fellow of the Royal Statistical Society.

Hao He (S'08–M'11) received the B.Sc. degree from the University of Science and Technology of China (USTC), Hefei, China, in 2007 and the Ph.D. degree from the University of Florida, Gainesville, in 2011, both in electrical engineering.

His research interests were in the areas of radar/sonar waveform design and spectral estimation. He is currently a Software Development Engineer for Amazon.com, Seattle, WA.

Jian Li (S'87–M'91–SM'97–F'05) received the M.Sc. and Ph.D. degrees in electrical engineering from Ohio State University, Columbus, in 1987 and 1991, respectively.

From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, Ohio State University, Columbus. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, where she is currently a Professor. In Fall 2007, she was on sabbatical leave at MIT, Cambridge, Massachusetts. Her current research interests include spectral estimation, statistical and array signal processing, and their applications.

Dr. Li is a Fellow of IET. She is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award and the 1996 Office of Naval Research Young Investigator Award. She was an Executive Committee Member of the 2002 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 2002. She was an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1999 to 2005, an Associate Editor of the *IEEE Signal Processing Magazine* from 2003 to 2005, and a member of the Editorial Board of *Signal Processing*, a publication of the European Association for Signal Processing (EURASIP), from 2005 to 2007. She has been a member of the Editorial Board of *Digital Signal Processing—A Review Journal*, a publication of Elsevier, since 2006. She is presently a member of the Sensor Array and Multichannel (SAM) Technical Committee of the IEEE Signal Processing Society. She is a coauthor of the papers that have received the First and Second Place Best Student Paper Awards, respectively, at the 2005 and 2007 Annual Asilomar Conferences on Signals, Systems, and Computers, Pacific Grove, CA. She is also a coauthor of the paper that received the M. Barry Carlton Award for the best paper published in the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS in 2005.