of the SQKF to function reliably at the expense of increased computational complexity. In the final analysis, there is "no free lunch": for every gain we make in system performance, there is a price to be paid.

VII. Concluding Remarks

In this correspondence, we presented the square-root extension of the quadrature Kalman filter. The SQKF exhibits several attractive numerical properties over the conventional QKF at the expense of increased computational complexity. However, a similar square-root formulation for the unscented Kalman filter may not be meaningful, because the covariance estimated by the sigma-point filter itself is not guaranteed to be positive definite, provided that no computational errors are made. We computed the complexity of the SQKF for a generic nonlinear state-space model. The overall complexity for a given problem is essentially not worse than the quantity we computed. For example, reduction in computational complexity is likely to arise in the following contexts: i) When the state, or measurement model is linear, the time-update, or measurement-update of the square-root Kalman filter can be simplified; ii) a careful arrangement of matrix multiplications exploiting their sparseness and replication; iii) use of a multidimensional processor array for fast computation and efficient exploitation of the system memory, e.g., when a matrix $A \in \mathbb{R}^{m \times n}$ is decomposed by the Givens rotation-based QR algorithm in a multidimensional processor array, the computational complexity reduces from $\mathbf{O}(2mn^2)$ to $\mathbf{O}(m+n)$ time units [4].

The QKF is likely to crash, when i) it is allowed to operate for a substantially long period of time without manual intervention in a system with limited word length, in which case, loss of precision due to the accumulation of round off errors, divergence, and other numerical ills are likely to occur, or ii) the numerical aspects of the given application are suspicious, e.g., offline training of neural networks, or iii) the problem under consideration is high dimensional, where error covariance matrices are more likely to lose their properties due to the curse of dimensionality. In all these situations, the use of the SQKF may be considered as a systematic solution.

Acknowledgment

The authors would like to thank the anonymous reviewers for helpful comments.

References

[1] I. Arasaratnam, S. Haykin, and R. J. Elliott, "Discrete-time nonlinear filtering algorithms using Gauss-Hermite quadrature," *Proc. IEEE*, vol. 95, no. 5, pp. 953–977, May 2007.

[2] Y. Bar Shalom, X.-R. Li, and T. Kirubarajan, *Estimation With Applications to Tracking and Navigation*. New York: Wiley, 2001.

[3] M. Grewal and A. Andrews, *Kalman Filtering: Theory and Practice Using Matlab*, 2nd ed. New York: Wiley, 2001.

[4] S. Haykin, *Adaptive Filter Theory*, 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, ch. 10.

[5] P. Maybeck, *Stochastic Models, Estimation and Control*. New York: Academic Press, 1979, vol. I.

[6] P. Kaminski, A. Bryson, and S. Schmidt, "Discrete square root filtering: A survey of current techniques," *IEEE Trans. Autom. Control*, vol. AC-16, no. 6, pp. 727–736, Dec. 1971.

[7] J. Porrill, "Optimal combination and constraints for geometrical senor data," *Int. J. Robot. Res.*, vol. 7, no. 6, pp. 66–77, Dec. 1988.

[8] R. van der Merwe and E. Wan, "The square-root unscented Kalman filter for state and parameter estimation," in *Proc. Int. Conf. Acous. Speech, Signal Process. (ICASSP)*, May 2001, vol. 6, pp. 3461–3464.

[9] M. Varhaegen and P. Van Dooren, "Numerical aspects of different Kalman filter implementations," *IEEE Trans. Autom. Control*, vol. AC-31, no. 10, pp. 907–917, Oct. 1986.

Waveform Synthesis for Diversity-Based Transmit Beampattern Design

Petre Stoica, Jian Li, and Xumin Zhu

***Abstract*—Transmit beampattern design is a critically important task in many fields including defense and homeland security as well as biomedical applications. Flexible transmit beampattern designs can be achieved by exploiting the waveform diversity offered by an array of sensors that transmit probing signals chosen at will. Unlike a standard phased-array, which transmits scaled versions of a single waveform, a waveform diversity-based system offers the flexibility of choosing how the different probing signals are correlated with one another. Recently proposed techniques for waveform diversity-based transmit beampattern design have focused on the optimization of the covariance matrix $\mathbf{R}$ of the waveforms, as optimizing a performance metric directly with respect to the waveform matrix is a more complicated operation. Given an $\mathbf{R}$, obtained in a previous optimization stage or simply pre-specified, the problem becomes that of determining a signal waveform matrix $\mathbf{X}$ whose covariance matrix is equal or close to $\mathbf{R}$, *and* which also satisfies some practically motivated constraints (such as constant-modulus or low peak-to-average-power ratio constraints). We propose a cyclic optimization algorithm for the synthesis of such an $\mathbf{X}$, which (approximately) realizes a given optimal covariance matrix $\mathbf{R}$ under various practical constraints. A numerical example is presented to demonstrate the effectiveness of the proposed algorithm.**

***Index Terms*—Beampattern design, constant modulus, cyclic optimization, low PAR, MIMO, waveform diversity, waveform synthesis.**

I. Introduction

Waveform diversity has been utilized both in multiple-input multiple-output (MIMO) communications and in MIMO radar. In the past decade, communications systems using multiple transmit and receive antennas have attracted significant attention from government agencies, academic institutions and research laboratories, because of their potential for dramatically enhanced throughput and significantly reduced error rate without spectrum expansion. Similarly, MIMO radar systems have recently received the attention of researchers and practitioners alike due to their improved capabilities compared with a standard phased-array radar. A MIMO radar system, unlike a standard phased-array radar, can transmit multiple probing signals that may be chosen at will. This waveform diversity offered by MIMO radar is the main reason for its superiority over standard phased-array radar; see, e.g., [1]–[19]. For colocated transmit and receive antennas, for example, MIMO radar has been shown to have the following appealing features: higher resolution (see, e.g., [1] and [2]), superior moving target detection capability [3], better parameter identifiability [7], [14], and direct applicability of adaptive array techniques [7], [8],

Manuscript received May 23, 2007; revised November 28, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Zhengyuan (Daniel) Xu. This work was supported in part by the Swedish Research Council (VR), by the Army Research Office under Grant W911NF-07-1-0450, by the National Science Foundation under Grant CCF-0634786, and by the Office of Naval Research under Grant N000140710293. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the United States Government.

P. Stoica is with the Department of Information Technology, Uppsala University, SE-751 05, Uppsala, Sweden (e-mail: ps@it.uu.se).

J. Li and X. Zhu are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: li@dsp.ufl.edu; zhuxm@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.2007.916139

[16]; in addition, the covariance matrix of the probing signal vector transmitted by a MIMO radar system can be designed to approximate a desired transmit beampattern—an operation that, once again, would be hardly possible for conventional phased-array radar [7], [10], [13].

Transmit beampattern design is critically important not only in defense applications, but also in many other fields including homeland security and biomedical applications. In all these applications, flexible transmit beampattern designs can be achieved by exploiting the waveform diversity offered by the possibility of choosing how the different probing signals are correlated with one another.

An interesting current research topic in this area of MIMO radar systems is the optimal synthesis of the transmitted waveforms. Waveform designs have been considered for MIMO radar with widely separated antennas in [11], albeit without imposing any practical constraints (such as the constant-modulus constraint) on the waveforms. Recently proposed techniques for transmit beampattern design or for enhanced radar imaging and parameter estimation using colocated antennas have focused on the optimization of the covariance matrix $\mathbf{R}$ of the waveforms [4], [5], [7], [10], [12], [13], [15], [17]. For example, in a waveform diversity-based ultrasound system, $\mathbf{R}$ can be designed to achieve a beampattern that is suitable for the hyperthermia treatment of breast cancer [20]. Now, instead of designing $\mathbf{R}$, as in the cited references, we might think of designing directly the probing signals by optimizing a given performance measure with respect to the matrix $\mathbf{X}$ of the signal waveforms. However, compared with optimizing the same performance measure with respect to the covariance matrix $\mathbf{R}$ of the transmitted waveforms, optimizing directly with respect to $\mathbf{X}$ is a more complicated problem. This is so because $\mathbf{X}$ has more unknowns than $\mathbf{R}$ and the dependence of various performance measures on $\mathbf{X}$ is more intricate than the dependence on $\mathbf{R}$ (as $\mathbf{R}$ is a quadratic function of $\mathbf{X}$). In effect, there are several recent methods, as mentioned above, that can be used to efficiently compute an optimal covariance matrix $\mathbf{R}$, with respect to several performance metrics; yet the same cannot be said about determining an optimal signal waveform matrix $\mathbf{X}$, which is the *ultimate goal* of the designing exercise.

In this paper, we consider the synthesis of the signal waveform matrix $\mathbf{X}$ for diversity-based transmit beampattern design. With $\mathbf{R}$ obtained in a previous (optimization) stage, our problem is to determine a signal waveform matrix $\mathbf{X}$ whose covariance matrix is equal or close to $\mathbf{R}$, *and* which also satisfies some practically motivated constraints (such as constant-modulus or low peak-to-average-power ratio (PAR) constraints). We present a cyclic optimization algorithm for the synthesis of such an $\mathbf{X}$. We also investigate how the synthesized waveforms and the corresponding transmit beampattern design depend on the enforced practical constraints. A numerical example is provided to demonstrate the effectiveness of the proposed methodology.

Notation: Vectors are denoted by boldface lowercase letters and matrices by boldface uppercase letters. The nth component of a vector $\mathbf{x}$ is written as $x(n)$. The nth diagonal element of a matrix $\mathbf{R}$ is written as R_{nn}. A Hermitian square root of a matrix $\mathbf{R}$ is denoted as $\mathbf{R}^{1/2}$. We use $(\cdot)^T$ to denote the transpose, and $(\cdot)^*$ for the conjugate transpose. The Frobenius norm is denoted as $\|\cdot\|$. The real part of a complex-valued vector or matrix is denoted as $\mathrm{Re}(\cdot)$.

II. Formulation of the Signal Synthesis Problem

Let the columns of $\mathbf{X} \in C^{L\times N}$ be the transmitted waveforms, where N is the number of the transmitters, and L denotes the number of samples in each waveform. Let

$$\mathbf{R} \triangleq \frac{1}{L}\mathbf{X}^*\mathbf{X} \tag{1}$$

be the (sample) covariance matrix of the transmitted waveforms. We assume that $L > N$ (typically $L \gg N$). Note that $\mathbf{X}$ has $2NL$ real-valued unknowns, which is usually a much larger number than the number of unknowns in $\mathbf{R}$, viz. N^2.

The class of unconstrained signal waveform matrices $\mathbf{X}$ that realize a given covariance matrix $\mathbf{R}$ is given by

$$\frac{1}{\sqrt{L}}\mathbf{X}^* = \mathbf{R}^{1/2}\mathbf{U}^* \tag{2}$$

where $\mathbf{U}^*$ is an arbitrary semi-unitary $N \times L$ matrix ($\mathbf{U}^*\mathbf{U} = \mathbf{I}$). Besides realizing (at least approximately) $\mathbf{R}$, the signal waveform matrix must also satisfy a number of practical constraints. Let $\mathcal{C}$ denote the set of signal matrices $\mathbf{X}$ that satisfy these constraints. Then a possible mathematical formulation of the problem of *synthesizing the probing signal matrix* $\mathbf{X}$ is as follows:

$$\min_{\mathbf{X}\in\mathcal{C};\mathbf{U}} \|\mathbf{X} - \sqrt{L}\mathbf{U}\mathbf{R}^{1/2}\|^2. \tag{3}$$

Depending on the constraint set $\mathcal{C}$, the solution $\mathbf{X}$ to (3) may realize $\mathbf{R}$ exactly or only approximately. Evidently as $\mathcal{C}$ is expanded (i.e., the constraints are relaxed), the matching error in (3) decreases. Whenever the matching error is different from zero, we can use *either* the solution $\hat{\mathbf{X}}$ to (3) as the signal waveform matrix, in which case it will satisfy the constraints but it will only approximately realize $\mathbf{R}$, *or* $\sqrt{L}\hat{\mathbf{U}}\mathbf{R}^{1/2}$, where $\hat{\mathbf{U}}$ is the $\mathbf{U}$-solution of (3), which realizes $\mathbf{R}$ exactly but satisfies only approximately the constraints—the choice between these two signal waveform matrices may be dictated by the application at hand.

The minimization problem in (3) is *nonconvex* due to the nonconvexity of the constraint $\mathbf{U}^*\mathbf{U} = \mathbf{I}$ and possibly of the set $\mathcal{C}$, as well. The constraint $\mathbf{U}^*\mathbf{U} = \mathbf{I}$ generates the so-called Stiefel manifold, and there are algorithms that can be used to minimize a function over the said manifold (see, e.g., [21]). However, these algorithms are somewhat intricate both conceptually and computationally, and their convergence properties are not completely known; additionally, in (3) we also have the problem of minimizing with respect to $\mathbf{X} \in \mathcal{C}$, which may also be nonconvex.

With the above facts in mind, we prefer a cyclic (or alternating) minimization algorithm for solving (3), as suggested in a related context in [22], [23]. We refer to the cited papers for more details on this type of algorithm and its properties.

III. Cyclic Algorithm (CA) for Signal Synthesis

We first summarize the steps of the cyclic minimization algorithm and then describe each step in detail.

Step 0:	Set $\mathbf{U}$ to an initial value (e.g., the elements of $\mathbf{U}$ can be independently drawn from a complex Gaussian distribution with mean 0 and standard deviation 1); alternatively we can start with an initial value for $\mathbf{X}$, in which case the sequence of the next steps should be inverted (note that the initial value of either $\mathbf{U}$ or $\mathbf{X}$ does not necessarily have to satisfy the constraints imposed on these variables in the next steps of the cyclic algorithm).
Step 1:	Obtain the matrix $\mathbf{X} \in \mathcal{C}$ that minimizes (3) for $\mathbf{U}$ fixed at its most recent value.
Step 2:	Determine the matrix $\mathbf{U}$ ($\mathbf{U}^*\mathbf{U} = \mathbf{I}$) that minimizes (3) for $\mathbf{X}$ fixed at its most recent value.
Iteration:	Iterate Steps 1 and 2 until a given stop criterion is satisfied. In the numerical example presented later, we terminate the iteration when the Frobenius norm of the difference between the $\mathbf{U}$ matrices at two consecutive iterations is less than or equal to 10^{-4}.

An important advantage of the above algorithm is that Step 2 has a *closed-form solution*. This solution can be derived in a number of ways

(see, e.g., [24] and [25]). A simple derivation of it runs as follows. For given $\mathbf{X}$, we have that

$$\|\mathbf{X} - \sqrt{L}\mathbf{U}\mathbf{R}^{1/2}\|^2 = \text{const} - 2\text{Re}\left\{\text{tr}[\sqrt{L}\mathbf{R}^{1/2}\mathbf{X}^*\mathbf{U}]\right\}. \quad (4)$$

Let

$$\sqrt{L}\mathbf{R}^{1/2}\mathbf{X}^* = \bar{\mathbf{U}}\mathbf{\Sigma}\hat{\mathbf{U}}^* \quad (5)$$

denote the singular value decomposition (SVD) of $\sqrt{L}\mathbf{R}^{1/2}\mathbf{X}^*$, where $\bar{\mathbf{U}}$ is $N \times N$, $\mathbf{\Sigma}$ is $N \times N$, and $\hat{\mathbf{U}}$ is $L \times N$. Then

$$\text{Re}\left\{\text{tr}[\sqrt{L}\mathbf{R}^{1/2}\mathbf{X}^*\mathbf{U}]\right\} = \text{Re}\left\{\text{tr}[\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}}\mathbf{\Sigma}]\right\} \quad (6)$$

$$= \sum_{n=1}^{N} \text{Re}\left\{[\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}}]_{nn}\right\}\Sigma_{nn}. \quad (7)$$

Because

$$(\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}})(\bar{\mathbf{U}}^*\mathbf{U}^*\hat{\mathbf{U}}) = \hat{\mathbf{U}}^*\mathbf{U}\mathbf{U}^*\hat{\mathbf{U}} \leq \hat{\mathbf{U}}^*\hat{\mathbf{U}} = \mathbf{I} \quad (8)$$

it follows that

$$\text{Re}^2\left\{[\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}}]_{nn}\right\} \leq \left|[\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}}]_{nn}\right|^2 \leq \left[(\hat{\mathbf{U}}^*\mathbf{U}\bar{\mathbf{U}})(\bar{\mathbf{U}}^*\mathbf{U}^*\hat{\mathbf{U}})\right]_{nn} \leq 1 \quad (9)$$

and therefore that

$$\|\mathbf{X} - \sqrt{L}\mathbf{U}\mathbf{R}^{1/2}\|^2 \geq \text{const} - 2\sum_{n=1}^{N}\Sigma_{nn}. \quad (10)$$

The lower bound in (10) is achieved at

$$\hat{\mathbf{U}} = \hat{\mathbf{U}}\bar{\mathbf{U}}^* \quad (11)$$

which is thus the solution to the minimization problem in Step 2 of the cyclic algorithm.

The solution to the problem in Step 1 naturally depends on the constraint set $\mathcal{C}$. For example, in radar systems the need to avoid expensive amplifiers and A/D converters has led to the requirement that the transmitted signals have constant modulus (see, e.g., [26]). Let $\{x_n(l)\}_{l=1}^{L}$ denote the elements in the nth column $\mathbf{x}_n$ of the signal waveform matrix $\mathbf{X}$. Then the constant-modulus requirement means that

$$|x_n(l)| = c, \text{ for some given constant } c \text{ and for } l = 1, \cdots, L. \quad (12)$$

(For example, we can choose $c = R_{nn}^{1/2}$; we omit the dependence of c on n for notational simplicity.) Under the constraint in (12), Step 1 of the algorithm has also a *closed-form* solution. Indeed, the generic problem to be solved in such a case is

$$\min_{\psi} |ce^{j\psi} - z|^2 \quad (13)$$

where $c > 0$ and $z \in C$ are given numbers. Because

$$|ce^{j\psi} - z|^2 = \text{const} - 2c|z|\cos[\psi - \arg(z)] \quad (14)$$

the minimizing ψ is evidently given by

$$\psi = \arg(z). \quad (15)$$

Therefore, under the constant-modulus constraint, both steps of the cyclic algorithm have solutions that can be readily computed. However, (12) may be too hard a requirement on the signal matrix in the sense that the corresponding minimum value of the matching criterion in (3) may not be as small as desired. In particular, this means that $(1/L)\hat{\mathbf{X}}^*\hat{\mathbf{X}}$ may not be a good approximation of $\mathbf{R}$ (see, e.g., [15], where it was shown that signals that have constant modulus and take on values in a finite alphabet may fail to realize well a given covariance matrix).

With the above facts in mind, we may be willing to compromise and therefore relax the requirement that the signals have constant modulus. In effect, in some modern radar systems this requirement can be replaced by the condition that the transmitted signals have a *low peak-to-average-power ratio* (PAR). Mathematically, the low PAR requirement can be formulated as follows:

$$\text{PAR}(\mathbf{x}_n) \triangleq \frac{\max_l |x_n(l)|^2}{\frac{1}{L}\sum_{l=1}^{L}|x_n(l)|^2} \leq \rho, \quad \text{for a given } \rho \in [1, L] \quad (16)$$

(where, once again, we omit the dependence of ρ on n for notational simplicity). If we add to (16) a power constraint, viz.

$$\frac{1}{L}\sum_{l=1}^{L}|x_n(l)|^2 = \gamma, \quad (\text{e.g., } \gamma = R_{nn}) \quad (17)$$

then the set $\mathcal{C}$ is described by the equations:

$$\begin{cases} \frac{1}{L}\sum_{l=1}^{L}|x_n(l)|^2 = \gamma, \\ |x_n(l)|^2 \leq \rho\gamma, \qquad l = 1, \cdots, L. \end{cases} \quad (18)$$

While the above constraint set is not convex, an *efficient algorithm* for solving the corresponding problem in Step 1 of the cyclic algorithm has been proposed in [22] and [23]. Note that the constraints in (18) are imposed on $\mathbf{X}$ in a column-wise manner. Consequently, the solution to Step 1 is obtained by dealing with the columns of $\mathbf{X}$ in a one by one fashion.

With the power constraint in (17) enforced, the diagonal elements of $\mathbf{R}$ can be synthesized exactly. If the exact matching of R_{nn} is not deemed necessary, we can relax the optimization by omitting (17). In the Appendix, we show how to modify the algorithm of [22] and [23] in the case where only (16) is enforced. (In the numerical example presented in the following section, (17) will be enforced.)

IV. Numerical Illustration

We present a numerical example to demonstrate the effectiveness of CA for signal synthesis in diversity-based transmit beampattern design applications. Several other examples as well as an application to waveform-diversity based ultrasound hyperthermia can be found in [27].

In a MIMO radar scenario, the power of the probing signal at a generic focal point with coordinates θ, which is the so-called transmitted beampattern, can be shown to be (see, e.g., [4], [8], and [10]):

$$\mathbf{a}^*(\theta)\mathbf{R}\mathbf{a}(\theta) \quad (19)$$

where $\mathbf{R}$ is as defined before

$$\mathbf{a}(\theta) = \left[e^{j2\pi f_0\tau_1(\theta)} \quad e^{j2\pi f_0\tau_2(\theta)} \quad \cdots \quad e^{j2\pi f_0\tau_N(\theta)}\right]^T \quad (20)$$

f_0 is the carrier frequency of the transmitted signal, and $\tau_n(\theta)$ is the time needed by the signal emitted via the nth transmit antenna to arrive at the focal point; in what follows we assume that θ is a one-dimensional angle variable (expressed in degrees).

Assume that we are given a desired transmit beampattern $\phi(\theta)$ defined over a region of interest. First, we use the optimization technique introduced in [13] to choose $\mathbf{R}$ such that the transmit beampattern, $\mathbf{a}^*(\theta)\mathbf{R}\mathbf{a}(\theta)$, approximates (in a least squares sense) the desired transmit beampattern, $\phi(\theta)$, over the region of interest, under the constraint that the elemental transmit power is not larger than a prespecified value (see [13] for details). Then, once the optimal $\mathbf{R}$ has been determined, we can use CA to synthesize the waveform matrix $\mathbf{X}$.

As mentioned in Section II, the CA solution to (3) may be chosen to realize $\mathbf{R}$ exactly or only approximately. When the signal waveforms are synthesized as $\sqrt{L}\hat{\mathbf{U}}\mathbf{R}^{1/2}$, where $\hat{\mathbf{U}}$ is the solution to (3)

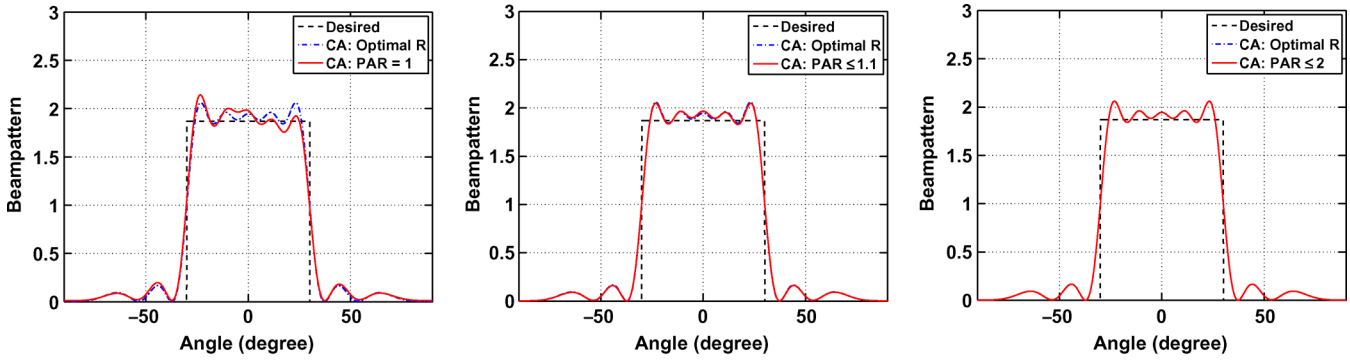

Fig. 1. Beampattern matching design with the desired main-beam width of 60° and under the uniform elemental power constraint. The probing signals are synthesized for $N = 10$ and $L = 256$ by using CA under (a) PAR $= 1$ (resulting in $\delta = -29.7891$ dB), (b) PAR ≤ 1.1 (resulting in $\delta = -41.7237$ dB), and (c) PAR ≤ 2 (resulting in $\delta = -119.5251$ dB).

obtained via CA, then they realize $\mathbf{R}$ exactly, but satisfy the PAR constraints only approximately. We refer to the so-synthesized waveforms as *the CA synthesized waveforms with optimal* $\mathbf{R}$ (abbreviated as *optimal* $\mathbf{R}$). When we use the solution $\hat{\mathbf{X}}$ to (3) obtained via CA as the transmitted signal waveform matrix, then $\hat{\mathbf{X}}$ will satisfy the PAR constraints, but will realize $\mathbf{R}$ only approximately. We refer to the so-synthesized waveforms as *the CA synthesized waveforms with PAR* $\leq \rho$ (abbreviated as *PAR* $\leq \rho$).

In the present example, the transmit array is assumed to be a uniform linear array (ULA) comprising $N = 10$ sensors with half-wavelength inter-element spacing. The sample number L is set equal to 256, and the CA algorithm is initialized using the $\mathbf{U}$ described in Step 0.

The desired beampattern has one wide main-beam centered at 0° with a width of 60°. Fig. 1(a)–(c) shows the beampatterns corresponding to the CA synthesized waveforms under the constraints of PAR $= 1$ (constant-modulus), PAR ≤ 1.1, and PAR ≤ 2, respectively. For comparison purposes, we also show the desired beampattern $\phi(\theta)$. Note that the beampattern obtained using the CA synthesized waveforms is close to the desired one even under the constant-modulus constraint.

We also note from Fig. 1 that the beampatterns obtained using *the CA synthesized waveforms with optimal* $\mathbf{R}$ are slightly different from those obtained using *the CA synthesized waveforms with PAR* $\leq \rho$. Let

$$\hat{\mathbf{R}} = \frac{1}{L}\hat{\mathbf{X}}^*\hat{\mathbf{X}} \tag{21}$$

be the sample covariance matrix corresponding to *the CA synthesized waveforms with* $PAR \leq \rho$, and let

$$\delta = \|\hat{\mathbf{R}} - \mathbf{R}\| \tag{22}$$

denote the norm of the difference between $\hat{\mathbf{R}}$ and $\mathbf{R}$. Then we have $\delta = -29.7891$ dB, -41.7237 dB, and -119.5251 dB for Fig. 1(a), (b), and (c), respectively. As expected, the difference decreases as the PAR value increases. For the case of PAR $= 2$, the difference is essentially zero. The mean-squared error (MSE) of $\hat{\mathbf{R}}$ (i.e., the average value of δ^2), obtained under PAR $= 1$ and estimated via 100 Monte Carlo simulation runs, is shown in Fig. 2 as a function of the sample number L. Note that, as also expected, the MSE decreases as L increases.

Fig. 3 shows the actual PAR values of *the CA synthesized waveforms with optimal* $\mathbf{R}$ corresponding to Fig. 1. These PAR values are also compared to those associated with the waveform matrix obtained by pre-multiplying $\mathbf{R}^{1/2}$ with a 256×10 matrix whose columns contain orthogonal Hadamard code sequences of length 256. The sample covariance matrix of the colored Hadamard sequences is also equivalent

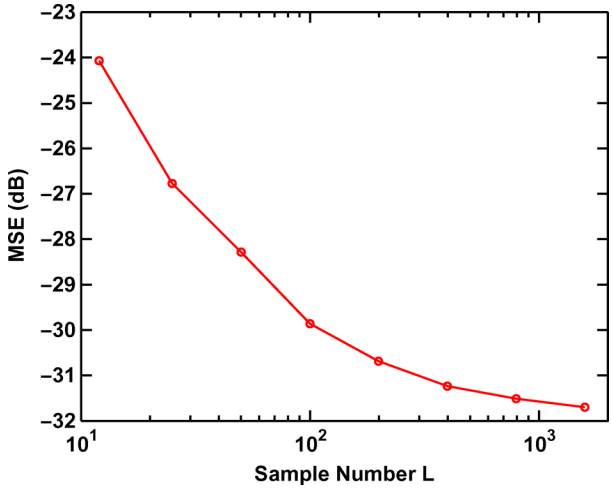

Fig. 2. MSE of $\hat{\mathbf{R}}$ for *the CA synthesized waveforms with* PAR $= 1$, as a function of the sample number L.

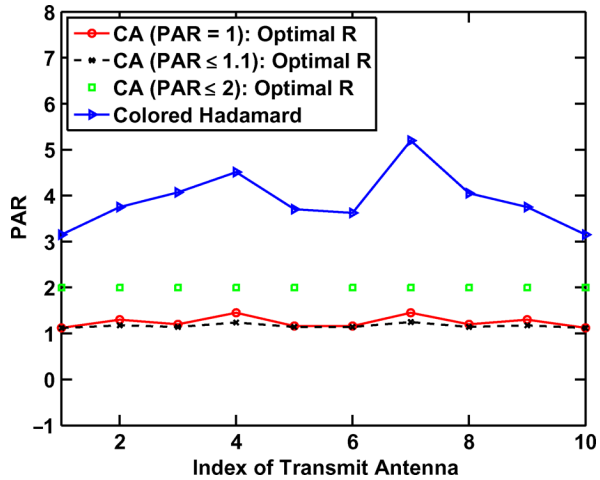

Fig. 3. Actual PAR values for *the CA synthesized waveforms with optimal* $\mathbf{R}$ and for colored Hadamard code sequences.

to $\mathbf{R}$. Note that *the CA synthesized waveforms with optimal* $\mathbf{R}$ have much lower PAR values than the colored Hadamard code sequences. Note also that the actual PAR values of *the CA synthesized waveforms with optimal* $\mathbf{R}$ obtained under PAR ≤ 1.1 are slightly lower than the PAR values obtained under PAR $= 1$.

V. Concluding Remarks

We have considered the problem of waveform synthesis for diversity-based flexible transmit beampattern design. Optimization of a performance metric *directly* with respect to the signal matrix can lead to an intractable problem even under a relatively simple low PAR constraint. For this reason, we proposed the following strategy: *first* optimize the performance metric of interest with respect to the signal covariance matrix $\mathbf{R}$; and *then* synthesize a signal waveform matrix that, under the low PAR constraint, realizes (at least approximately) the optimal covariance matrix derived in the first step. We have presented a cyclic optimization algorithm for the synthesis of a signal waveform matrix to (approximately) realize a given covariance matrix $\mathbf{R}$ under the constant-modulus constraint or the low PAR constraint. The output of the cyclic algorithm can be used to obtain either a waveform matrix whose covariance matrix is exactly equal to $\mathbf{R}$ but whose PAR is slightly larger than the imposed value, or a waveform matrix with the imposed PAR but whose covariance matrix may differ slightly from $\mathbf{R}$—the type of application will dictate which one of these two kinds of waveforms will be more useful. A numerical example has been provided to demonstrate that the proposed algorithm for waveform synthesis is quite effective.

Appendix
On Enforcing Solely the PAR Constraint

Consider the following generic form of the problem

$$\min_{\mathbf{s}} \|\mathbf{s}-\mathbf{z}\|^2 \quad \text{s.t.} \quad \text{PAR}(\mathbf{s}) \le \rho \tag{23}$$

where $\mathbf{z}$ is given and $\text{PAR}(\mathbf{s})$ is as defined in (16). Hence, we have omitted the power constraint (17), which should lead to a smaller matching error. Because $\text{PAR}(\mathbf{s})$ is insensitive to the scaling of $\mathbf{s}$, let us parameterize $\mathbf{s}$ as

$$\mathbf{s} = c\mathbf{x}; \quad \|\mathbf{x}\|^2 = 1; \quad \text{where } c \ge 0 \text{ is a variable.} \tag{24}$$

Using (24) in (23) yields

$$\|\mathbf{s}-\mathbf{z}\|^2 = \|c\mathbf{x}-\mathbf{z}\|^2 = c^2 - 2c\text{Re}(\mathbf{x}^*\mathbf{z}) + \text{const.} \tag{25}$$

If $\text{Re}(\mathbf{x}^*\mathbf{z}) \le 0$, then the minimum value of (25) with respect to $c \ge 0$ occurs at $c = 0$. If $\text{Re}(\mathbf{x}^*\mathbf{z}) > 0$, then the minimization of (25) with respect to $c \ge 0$ gives

$$c = \text{Re}(\mathbf{x}^*\mathbf{z}) \tag{26}$$

and the value of (25) corresponding to (26) is smaller than the value associated with $c = 0$. Because $\text{PAR}(\mathbf{s}) = \text{PAR}(\mathbf{x})$ does not depend on the phases of the elements of $\mathbf{x}$, we can always choose $\mathbf{x}$ such that $\text{Re}(\mathbf{x}^*\mathbf{z}) > 0$—so that we achieve a smaller value of (25). Consequently, the minimizing value $c \ge 0$ of (25) is always given by (26). The remaining problem is

$$\max_{\mathbf{x}} \text{Re}(\mathbf{x}^*\mathbf{z}) \quad \text{s.t.} \quad \|\mathbf{x}\|^2 = 1 \quad \text{and} \quad \text{PAR}(\mathbf{x}) \le \rho \tag{27}$$

or equivalently

$$\min_{\mathbf{x}} \|\mathbf{x}-\mathbf{z}\|^2 \quad \text{s.t.} \quad \|\mathbf{x}\|^2 = 1 \quad \text{and} \quad \text{PAR}(\mathbf{x}) \le \rho \tag{28}$$

which has the form required by the algorithm of [22], [23]. Therefore, we can solve (28) using the said algorithm and then compute $\mathbf{s} = c\mathbf{x}$ with c given by (26).

The alternative discussed in Section III is to constrain $\|\mathbf{s}\|^2 = \|\mathbf{z}\|^2$ (which is the case when we choose $\gamma = R_{nn}$ in (17)). The use of this constraint is logical if we want to match R_{nn} exactly (for strict transmission power control, for example). However, if matching R_{nn} exactly is not a necessary condition, then a smaller matching error between $\mathbf{s}$ and $\mathbf{z}$ is obtained using (26) and (28).

References

[1] D. W. Bliss and K. W. Forsythe, "Multiple-input multiple-output (MIMO) radar and imaging: Degrees of freedom and resolution," in *Proc. 37th Asilomar Confe. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2003, vol. 1, pp. 54–59.

[2] I. Bekkerman and J. Tabrikian, "Spatially coded signal model for active arrays," in *Proc. 2004 IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Montreal, QC, Canada, Mar. 2004, vol. 2, pp. ii/209–ii/212.

[3] K. W. Forsythe, D. W. Bliss, and G. S. Fawcett, "Multiple-input multiple-output (MIMO) radar: Performance issues," in *Proc. 38th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2004, vol. 1, pp. 310–315.

[4] D. R. Fuhrmann and G. San Antonio, "Transmit beamforming for MIMO radar systems using partial signal correlations," in *Proc. 38th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2004, vol. 1, pp. 295–299.

[5] K. W. Forsythe and D. W. Bliss, "Waveform correlation and optimization issues for MIMO radar," in *Proc. 39th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2005, pp. 1306–1310.

[6] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenzuela, "Spatial diversity in radars—Models and detection performance," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 823–838, Mar. 2006.

[7] J. Li and P. Stoica, "MIMO radar—Diversity means superiority," presented at the 14th Annual Workshop on Adaptive Sensor Array Processing (Invited), MIT Lincoln Lab., Lexington, MA, Jun. 2006.

[8] L. Xu, J. Li, and P. Stoica, "Adaptive techniques for MIMO radar," presented at the 4th IEEE Workshop on Sensor Array and Multi-Channel Processing, Waltham, MA, Jul. 2006.

[9] L. Xu, J. Li, and P. Stoica, "Radar imaging via adaptive MIMO techniques," presented at the 14th Eur. Signal Processing Conf. (Invited), Florence, Italy, Sep. 2006.

[10] J. Li, P. Stoica, and Y. Xie, "On probing signal design for MIMO radar," presented at the 40th Asilomar Conf. Signals, Systems, Computers (Invited), Pacific Grove, CA, Oct. 2006.

[11] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, Jan. 2000.

[12] L. Xu, J. Li, P. Stoica, K. W. Forsythe, and D. W. Bliss, "Waveform Optimization for MIMO radar: A Cramer-Rao bound based study," presented at the 2007 IEEE Int. Conf. Acoustics, Speech, Signal Processing, Honolulu, HI, Apr. 2007.

[13] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, Aug. 2007.

[14] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of MIMO radar," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 968–971, Dec. 2007.

[15] D. R. Fuhrmann and G. San Antonio, "Transmit beamforming for MIMO radar systems using signal cross-correlation," *IEEE Trans. Aerosp. Electron. Syst.*, to be published.

[16] L. Xu, J. Li, and P. Stoica, "Adaptive MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.* [Online]. Available: http://www.sal.ufl.edu/sal/MIMORadar_AES.pdf to appear

[17] J. Li, L. Xu, P. Stoica, K. W. Forsythe, and D. W. Bliss, "Range compression and waveform optimization for MIMO radar: A Cramér–Rao bound based study," *IEEE Trans. Signal Process.* vol. 56, no. 1, pp. 218–232, Jan. 2008.

[18] J. Li and P. Stoica, "MIMO radar with colocated antennas: Review of some recent work," *IEEE Signal Process. Mag.* vol. 24, no. 5, pp. 106–114, Sep. 2007.

[19] J. Li and P. Stoica, Eds., *MIMO Radar Signal Processing*. New York: Wiley, 2008.

[20] B. Guo and J. Li, "Waveform diversity based ultrasound system for hyperthermia treatment of breast cancer," *IEEE Trans. Biomed. Eng.* vol. 55, no. 2, pt. 2, pp. 822–826, Feb. 2008.

[21] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.

[22] J. A. Tropp, I. S. Dhillon, R. W. Heath, and T. Strohmer, "CDMA signature sequences with low peak-to-average-power ratio via alternating projection," in *Proc. 37th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Nov. 2003, vol. 1, pp. 475–479.

[23] 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.
[24] C. R. Rao, "Matrix approximations and reduction of dimensionality in multivariate statistical analysis," in *Multivariate Analysis*, P. R. Krishnaiah, Ed. Amsterdam, The Netherlands: North-Holland, 1980, vol. 5, pp. 3–22.
[25] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[26] N. Levanon and E. Mozeson, *Radar Signals*. New York: Wiley, 2004.
[27] P. Stoica, J. Li, X. Zhu, and B. Guo, "Waveform synthesis for diversity-based transmit beampattern design," Univ. of Florida, Gainesville, FL, Internal Tech. Rep. [Online]. Available: http://www.sal.ufl.edu/sal/Txbeampattern.pdf

On Using *a priori* Knowledge in Space-Time Adaptive Processing

Petre Stoica, Jian Li, Xumin Zhu, and Joseph R. Guerci

***Abstract*—In space–time adaptive processing (STAP), the clutter covariance matrix is routinely estimated from secondary "target-free" data. Because this type of data is, more often than not, rather scarce, the so-obtained estimates of the clutter covariance matrix are typically rather poor. In knowledge-aided (KA) STAP, an *a priori* guess of the clutter covariance matrix (e.g., derived from knowledge of the terrain probed by the radar) is available. In this note, we describe a computationally simple and fully automatic method for combining this prior guess with secondary data to obtain a theoretically optimal (in the mean-squared error sense) estimate of the clutter covariance matrix. The authors apply the proposed method to the KASSPER data set to illustrate the type of achievable performance.**

***Index Terms*—Convex combination, general linear combination, knowledge-aided, space–time adaptive processing.**

I. Introduction and Preliminaries

In standard space–time adaptive processing (STAP), the clutter-and-noise covariance matrix, let us call it $\mathbf{R}$, is estimated from secondary data (presumed to be target free), let us say $\{\mathbf{y}(n)\}_{n=1}^{N}$, by means of the well-known formula (see, e.g., [1]–[5])

$$\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N}\mathbf{y}(n)\mathbf{y}^{*}(n) \tag{1}$$

where $(\cdot)^{*}$ denotes the conjugate transpose. However, frequently the dimension of $\mathbf{R}$ (denoted by M in what follows) is larger than, or at best comparable with, N. The result is that $\hat{\mathbf{R}}$ is, more often than not, a poor estimate of $\mathbf{R}$ (particularly so when $M \gg N$; see Section III of this note for such a case).

To estimate $\mathbf{R}$ more accurately, we can try to make use of prior knowledge on the terrain probed by the radar, acquired either from a previous scanning or from a map-based study. In knowledge-aided STAP (KA-STAP), an initial guess of $\mathbf{R}$, let us say $\mathbf{R}_0$, is obtained in this way (see, e.g., [4] and [5]). The problem is then to "combine" $\hat{\mathbf{R}}$ and $\mathbf{R}_0$ into an estimate of $\mathbf{R}$, preferably much more accurate than both $\hat{\mathbf{R}}$ and $\mathbf{R}_0$.

In this note, we will consider a "linear combination" of $\hat{\mathbf{R}}$ and $\mathbf{R}_0$ (see, e.g., [6]–[8]):

$$\tilde{\mathbf{R}} = \alpha\mathbf{R}_0 + \beta\hat{\mathbf{R}}; \quad \alpha > 0 \text{ and } \beta > 0. \tag{2}$$

Because α and β are constrained to be positive, and as typically $\mathbf{R}_0 > 0$ (positive definite), we have that $\tilde{\mathbf{R}} > 0$, which is a desirable feature. We will also consider the following "convex combination" (CC):

$$\tilde{\mathbf{R}} = \alpha\mathbf{R}_0 + (1-\alpha)\hat{\mathbf{R}}; \quad \alpha \in (0,1). \tag{3}$$

The constraint in (3) on α is imposed, once again, to guarantee that $\tilde{\mathbf{R}} > 0$. In general, there is no obvious reason why β in (2) should be constrained to equal $1-\alpha$, like in (3); however, the convex combination, (3), is a more parsimonious description of $\mathbf{R}$ than the general linear combination (GLC) in (2), which is probably the reason why (3) is commonly used in the literature (see, e.g., [4] and the references there).

The first goal of this note is to obtain the α and β that minimize the mean-squared error (MSE) of $\tilde{\mathbf{R}}$:

$$\mathrm{MSE} = E\left\{\|\tilde{\mathbf{R}} - \mathbf{R}\|^2\right\} \tag{4}$$

for both (2) and (3); hereafter, $\|\cdot\|$ denotes the Frobenius matrix norm or the Euclidean vector norm, depending on the context. We stress that this is a constrained estimation problem: the two classes of covariance matrix estimates in (3) and (2) have only one and, respectively, two free parameters. This means that the estimate $\tilde{\mathbf{R}}$ takes values in a restricted set that in general does not contain $\mathbf{R}$; therefore, the trivial but problematic minimizer $\tilde{\mathbf{R}} = \mathbf{R}$ of (4) is in general avoided.

Let α_0 and β_0 denote the optimal values of α and β that minimize (4). The second goal of this note is to discuss how to obtain estimates, $\hat{\alpha}_0$ and $\hat{\beta}_0$, of α_0 and β_0 from the available data (as we will see shortly, and as expected, both α_0 and β_0 depend on the unknown matrix $\mathbf{R}$). Finally, we will explain how to use the proposed estimates of $\mathbf{R}$, viz. $\tilde{\mathbf{R}}$ in (2) or (3) with $\alpha = \hat{\alpha}_0$ and $\beta = \hat{\beta}_0$, in a KA-STAP exercise based on the KASSPER data set [4], [9] (KASSPER stands for knowledge-aided sensor signal processing and expert reasoning).

Estimation of a large-dimension covariance matrix from a limited number of samples in the manner outlined above (and detailed in the next section) has been originally proposed in [6] and its main references (with an emphasis on applications in economics) and later on considered in several other papers, for example in [10] (with a focus on applications in bioinformatics) and in [11] (with an emphasis on array processing applications). However, the cited papers considered only the case of *real-valued data*. Our approach in this note is an extension of that in [6] to the complex-valued data case, as well as to a general $\mathbf{R}_0$ matrix ([6] considers the case of $\mathbf{R}_0 = \mathbf{I}$ only). Also our proofs are more explicit than those in [6], despite the more general, complex-valued data case we consider here. Finally, to conclude this brief review of the relevant literature, we refer the reader interested in KA-STAP and in the problem described in this note, to the more detailed paper [12], also by the present authors, which presents a maximum likelihood based approach to KA-STAP.

Manuscript received January 26, 2007; revised October 25, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Daniel Fuhrmann. This work was sponsored in part by the Defense Advanced Research Projects Agency under Grant HR0011-06-1-0031. Opinions, interpretations, conclusions, and recommendations are those of the authors and are not necessarily endorsed by the U.S. Government.

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

J. Li and X. Zhu are with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL 32611-6130 USA (e-mail: li@dsp.ufl.edu; zhuxm@ufl.edu).

J. R. Guerci is with SAIC, McLean, VA 22102 USA.

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.2007.914347