A Novel Wavelet-Based Generalized Sidelobe Canceller

Yi Chu, *Member, IEEE,* and Wen-Hsien Fang, *Member, IEEE*

***Abstract*— This paper presents a novel narrow-band adaptive beamformer with the generalized sidelobe canceller (GSC) as the underlying structure. The new beamformer employs a wavelet-based approach for the design of the blocking matrix of the GSC, which is now constituted by a set of regular M-band wavelet filters. Such a construction of the blocking matrix can not only block the desired signals from the lower path as required provided the wavelet filters have sufficiently high regularity, but it also encompasses the widely used one with ones and minus ones along the diagonals as a special case. In addition, it possesses two advantageous features. First, the eigenvalue spreads of the covariance matrices of the blocking matrix outputs, as demonstrated in various scenarios, are decreased as compared with those of previous approaches. Since the popular least-mean squares (LMS) algorithm has been notorious for its slow convergence rate, the reduction of the eigenvalue spreads can, in general, accelerate the convergence speed of the succeeding LMS algorithm. Second, the new beamformer belongs to a specific type of partially adaptive beamformers, wherein only a portion of the available degree of freedom is utilized in the adaptive processing. As such, the overall computational complexity is substantially reduced when compared to previous works. The issues of choosing the parameters involved for superior performance are also addressed. Simulation results are furnished as well to justify this new approach.**

***Index Terms*— Adaptive beamforming, generalized sidelobe canceller, wavelet filters.**

I. Introduction

THE design of adaptive beamformers is of importance in various disciplines of signal processing applications such as radar, sonar, and geophysical explorations [1]. In many applications, it is not uncommon to use lots of sensors to achieve better interference rejection as well as resolution. The enormous amount of computations involved in this type of arrays, however, may hinder them from practical implementations. To alleviate the computational overhead, two approaches have been advocated in the literature.

The first approach is based on the technique of partial adaptivity, in which only a fraction of the adjusting weights is employed, thus leading to lower computational complexity per iteration in adaptive processing. Several methods have been addressed for designing an optimal partially adaptive beamformer by minimizing the performance degradation (see, e.g., [2], [3]).

Manuscript received August 4, 1997; revised October 29, 1998. This work was supported by National Science Council of R.O.C. under Contracts 87-2213-E-011-004 and 88-2218-E011-007.

The authors are with the Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

Publisher Item Identifier S 0018-926X(99)07965-X.

The second approach concerns the adaptive schemes, among which the least-mean squares (LMS) algorithm, due to its simplicity and robustness, has been widely used as the adaptive algorithm. However, it is well known that the LMS algorithm is notorious for its slow convergence rate, especially for signals whose covariance matrices have widely diverse eigenvalues. To overcome this difficulty, several cascade preprocessors such as the Gram–Schmidt orthogonalization [4] and unitary transforms [5], [6] have been suggested. These preprocessors try to decorrelate the input signal before the adaptive processing, resulting in faster convergence speed.

The Gram–Schmidt orthogonalization preprocessor is less appealing since it requires more adaptive weights to achieve the orthogonalization process. On the other hand, the unitary transforms, which include several Karhunen–Loève-like transforms such as the discrete Fourier transform, the discrete cosine transform, and the wavelet transform comprise a set of *fixed* orthonormal vectors. The transformed signals based on these transforms are roughly uncorrelated and can be further appropriately processed (such as normalization) to reduce the sensitivity to the eigenvalue spreads of the covariance matrices of the input signals. The recently introduced wavelet transform has in particular received a great amount of attention [7], [8]. This new transform works effectively in analyzing nonstationary signals, as it forms a frequency adaptive window on the time-scale plane. Several variants of the wavelet transform LMS algorithms have been addressed and demonstrated to yield faster convergence behavior as compared with other transform-based approaches [9]–[11].

The wavelet transform has also been incorporated in the adaptive beamformer (WASPAB) [11]. The WASPB, which employs the generalized sidelobe canceller (GSC) [12] as the underlying structure, puts a wavelet transform processor preceding the LMS algorithm and indeed exhibits faster null-steering process. In this paper, we further our previous work of [11] and propose a more succinct approach by ingeniously combining the blocking matrix and the wavelet transform process into a single unit.

This new unit is constituted by a set of regular M-band wavelet filters [13], [14]. It is shown that this new unit, as the traditional blocking matrix in the GSC, can also block the desired signals from the lower path as long as the wavelet filters have sufficiently high regularity. This new unit, also being referred to as a blocking matrix, encompasses the widely used one with ones and minus ones along the diagonals as a special case (for the look direction gain constraint). This novel construction of the blocking matrix possesses

two advantageous features. First, the eigenvalue spreads of the covariance matrices of the blocking matrix outputs, as observed in various scenarios, are decreased as compared with those of previous approaches. This in general leads to faster adaptive response of the succeeding LMS algorithm. Second, the new beamformer belongs to a specific type of partially adaptive beamformers, wherein higher dimensional adaptive weights are mapped into lower dimensional ones, thus further reducing the computational overhead. As a consequence of these two characteristics, the computational complexity called for is substantially reduced when compared to previous works. The eigenvalue spreads and the dimension reduced are determined by the parameters of the wavelet filters as well as the corresponding matrix structure of the blocking matrix. To facilitate the choices of these parameters, some suggestive guidelines are also provided, aiming at superior performance.

This paper is organized as follows. Section II provides an overview of the regular M-band wavelets with an emphasis on the properties relevant to the following derivations. In Section III, we then develop our new wavelet-based GSC. The relationship with the corresponding linearly constrained minimum variance (LCMV) adaptive beamformer is addressed. The issues of choices of parameters involved are also treated. Some simulation results are furnished in Section IV to confirm the proposed approach. Section V concludes the whole paper.

II. Regular M-Band Wavelets

Since the proposed approach relies on the regular M-band wavelet filters, in this section, we briefly review the corresponding basic principles with an emphasis on some of their properties which are relevant to the following development. The rationale behind the wavelet decomposition is to represent a signal by a superposition of basis functions called wavelets. However, in contrast to the traditional Fourier transform, the wavelets possess a constant relative bandwidth on the logarithmic frequency axis, thus providing a more effective tool for nonstationary signal analysis.

The M-band orthonormal wavelet basis functions of $L^2(\mathcal{R})$ can be generated by a dilation and translation of a set of prototype wavelets $\{\psi_m(t)\}_{m=1}^{M-1}$ as $\{M^{\frac{i}{2}}\psi_m(M^it-k), m=1,2,\cdots,M-1; i,k\in\mathcal{Z}\}$, where i and k denote the scaling and translation parameters, respectively. In other words, any square integrable function $f(t)$ of $L^2(\mathcal{R})$ can be expressed as

$$f(t)=\sum_{m=1}^{M-1}\sum_{i,k}M^{\frac{i}{2}}f_{m,i,k}\psi_m(M^it-k) \tag{1}$$

where $\{f_{m,i,k}\}$ are the corresponding expansion coefficients. Under the framework of multiresolution analysis, the construction of the M-band wavelets can be carried out by first determining the scaling function $\phi(t)$ which satisfies the following two-scale difference equation [13]:

$$\phi(t)=\sqrt{M}\sum_{k=0}^{L-1}h_0(k)\phi(Mt-k) \tag{2}$$

where $\mathbf{h}_0=[h_0(0),h_0(1),\cdots,h_0(L-1)]^T$ denotes the corresponding (unitary) scaling filter with L being the length of the scaling filter and the superscript ^{T} standing for matrix transpose. $\mathbf{h}_0$ satisfies the constraints of $\sum_k h_0(k)h_0(k+Ml)=\delta(l)$ and $\sum_k h_0(k)=\sqrt{M}$, where $\delta(l)$ is the Kronecker delta function with $\delta(l)=1$ if $l=0$ and $\delta(l)=0$, otherwise. The prototype wavelets $\psi_m(t), m=1,2,\cdots,M-1$, can accordingly be determined by the following two-scale difference equations:

$$\psi_m(t)=\sqrt{M}\sum_{k=0}^{L-1}h_m(k)\phi(Mt-k)\qquad m=1,2,\cdots,M-1 \tag{3}$$

where $\mathbf{h}_m=[h_m(0),h_m(1),\cdots,h_m(L-1)]^T$, $m=1,2,\cdots,M-1$, denote the corresponding (unitary) wavelet filters and satisfy

$$\sum_k h_m(k)h_{m'}(k+Ml)=\delta(m-m')\delta(l). \tag{4}$$

Note that the unitary scaling filter $\mathbf{h}_0$ together with the wavelet filters $\mathbf{h}_m, m=1,2,\cdots,M-1$ form a paraunitary filter bank in which $\mathbf{h}_0$ corresponds to a low-pass filter and $\mathbf{h}_m, m=1,2,\cdots,M-1$, correspond to band/high-pass filters [13]. In addition, as in the two-band case [7], some regularity conditions need to be imposed on the scaling function, leading to the regular M-band wavelets [13], [14]. Several approaches have been addressed in [13], [14] for the construction of P-regular M-band wavelet filters of minimal length $L=MP$. The regularity condition also guarantees that the P-regular M-band wavelet filters have P vanishing moments, i.e.,

$$\sum_k k^r h_m(k)=0$$
$$m=1,2,\cdots,M-1,\quad r=0,1,\cdots,P-1. \tag{5}$$

From the vanishing moment property, we can deduce the following useful property:

Lemma 1: For any integer k_0, we have

$$\sum_k (k_0+k)^r h_m(k)=0$$
$$m=1,2,\cdots,M-1\quad\text{and}\quad r=0,1,\cdots,P-1. \tag{6}$$

Proof: Using the binomial expansion of $(k_0+k)^r=\sum_{i=0}^{r}\frac{r!}{(r-i)!i!}k_0^{(r-i)}k^i$ leads to

$$\begin{aligned}\sum_k (k_0+k)^r h_m(k)&=\sum_k\left[\sum_{i=0}^{r}\frac{r!}{(r-i)!i!}k_0^{(r-i)}k^i\right]h_m(k)\\&=\sum_{i=0}^{r}\frac{r!}{(r-i)!i!}k_0^{(r-i)}\left[\sum_k k^i h_m(k)\right]=0\end{aligned}$$

where we have used the vanishing moments property (5) of the regular wavelet filters. □

III. The Wavelet-Based Generalized Sidelobe Canceller

In this section, we shall first review the GSC which underlines the structure of our proposed adaptive beamformer. A novel wavelet-based blocking matrix is then introduced.

A. Background

Let us consider an equispaced linear array, which is composed of N-omnidirectional sensor elements. The narrow-band beamformer output at time instant k, $y(k)$, can be expressed as $y(k) = \mathbf{w}^H\mathbf{x}(k)$, where $\mathbf{w}$ and $\mathbf{x}(k)$ denote the weight vector and the array received vector, respectively, and the superscript ^{H} denotes the Hermitian transpose. The received vector $\mathbf{x}(k)$, which is assumed to consist of a single signal under the interference environment of J jammers, can then be represented by

$$\mathbf{x}(k) = s_s(k)\mathbf{a}_S(\theta_s) + \sum_{i=1}^{J} s_i(k)\mathbf{a}_J(\theta_i) + \mathbf{n} \tag{7}$$

where $s_s(k)$ and $s_i(k), i = 1, 2, \cdots, J$ are the waveforms of the desired signal and jammers, respectively

$$\mathbf{a}_S(\theta_s) = \left[e^{j(1-n_0)\mu_s}, e^{j(2-n_0)\mu_s}, \cdots, e^{j(N-n_0)\mu_s}\right]^T \tag{8}$$

denotes the signal vector with an arrival angle θ_s, n_0 is the reference point of the linear array, $\mu_s = \frac{2\pi}{\lambda_w} d_w \sin\theta_s$ with λ_w and d_w being, respectively, the wavelength and sensor distance $\mathbf{a}_J(\theta_i) = [e^{j(1-n_0)\mu_i}, e^{j(2-n_0)\mu_i}, \cdots, e^{j(N-n_0)\mu_i}]^T$, $i = 1, 2, \cdots, J$ corresponds to the jammer arriving from direction θ_i, $\mu_i = \frac{2\pi}{\lambda_w} d_w \sin\theta_i$, and $\mathbf{n}$ is the additive receiver (white) noise.

The LCMV beamformer considered by Frost [15] determines the weight $\mathbf{w}$ by minimizing the output power under some appropriate linear weight constraints. This approach can be expressed by the following linear constrained optimization problem:

$$\min_{\mathbf{w}} \mathbf{w}^H\mathbf{R}_x\mathbf{w} \quad \text{subject to } \mathbf{C}^T\mathbf{w} = \mathbf{f} \tag{9}$$

where $\mathbf{R}_x \triangleq \mathcal{E}\{\mathbf{x}(k)\mathbf{x}(k)^H\}$ is the data covariance matrix with $\mathcal{E}\{\cdot\}$ denoting the expectation operator, $\mathbf{C}$ and $\mathbf{f}$ are an $N \times S$ (full column rank) constraint matrix, and an $S \times 1$ response vector, respectively. In particular, if we consider the mainbeam derivative constraints, which have been utilized to achieve a flatter mainbeam response so that the array is less sensitive to the steering errors [16], [17], [18], the $(S-1)$ order derivative constraints then require that

$$\mathbf{C}_S^T\mathbf{w} = [1, 0, \cdots, 0]^T \tag{10}$$

where the $N \times S$ matrix $\mathbf{C}_S$ corresponds to the $(S-1)$ order derivative constraint matrix as

$$\mathbf{C}_S = [\mathbf{c}_0, \mathbf{c}_1, \cdots, \mathbf{c}_{S-1}] \tag{11}$$

with

$$\mathbf{c}_i = [(1-n_0)^i, (2-n_0)^i, \cdots, (N-n_0)^i]^T, \quad i = 0, 1, \cdots, S-1 \tag{12}$$

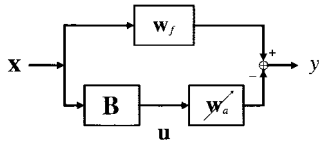

Fig. 1. The structure of the GSC.

in which we have used the expression of $\mathbf{a}_S(\theta_s)$ as given in (8). The GSC reformulates the LCMV to facilitate more efficient implementations as well as performance analysis. The basic principle of the GSC is to decompose the weight vector $\mathbf{w}$ into two components as shown in Fig. 1. The first component $\mathbf{w}_f$, which stands for the fixed target signal filter of the GSC, accounts for the constrained part of the LCMV. The second component, denoting the unconstrained and adaptive part, can be represented as $-\mathbf{B}\mathbf{w}_a$, where $\mathbf{B}$ is constituted by a basis of the null space of $\mathbf{C}^T$. As a consequence, (9) can be rewritten as the following unconstrained problem:

$$\min_{\mathbf{w}_a} (\mathbf{w}_f - \mathbf{B}\mathbf{w}_a)^H\mathbf{R}_x(\mathbf{w}_f - \mathbf{B}\mathbf{w}_a) \tag{13}$$

where $\mathbf{C}^T\mathbf{B} = \mathbf{0}$ and $\mathbf{C}^T\mathbf{w}_f = \mathbf{f}$. The solution of (13) can be easily determined as

$$\mathbf{w}_f = \mathbf{C}(\mathbf{C}^T\mathbf{C})^{-1}\mathbf{f} \quad \text{and} \quad \mathbf{w}_a = (\mathbf{B}^T\mathbf{R}_x\mathbf{B})^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{w}_f. \tag{14}$$

B. The Proposed Wavelet-Based GSC

Let us now consider an $N \times (\lfloor\frac{N-MP}{d}\rfloor + 1)(M-1)$ ($\lfloor\alpha\rfloor$ denotes the largest integer smaller than or equal to α) matrix $\mathbf{B}$ which is constituted by a set of P-regular M-band wavelet filters $\mathbf{h}_m$, $m = 1, 2, \cdots, M-1$, as [19]

$$\mathbf{B}^T = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \vdots \\ \mathbf{H}_{M-1} \end{bmatrix} \tag{15}$$

where $\mathbf{H}_m$, $m = 1, 2, \cdots, M-1$, is a $(\lfloor\frac{N-MP}{d}\rfloor + 1) \times N$ matrix as (16), shown at the bottom of the next page, in which $\mathbf{o}_d$ is a $d \times 1$ zero vector and d is a prespecified integer. It can be readily shown that all of the $(\lfloor\frac{N-MP}{d}\rfloor + 1)(M-1)$ columns of $\mathbf{B}$ form a linearly independent set by using the orthogonality of the unitary wavelet filters as given in (4). In particular, if d is a multiple of M, then all of the column of $\mathbf{B}$ are orthonormal and $\mathbf{B}^T\mathbf{B} = \mathbf{I}$, where $\mathbf{I}$ is an identity matrix. Note that $\mathbf{B}$ does not correspond to a wavelet transform matrix since it does not contain the scaling filter. Such a choice of $\mathbf{B}$ possesses a distinctive feature of "nulling" out the first few terms of the Taylor's series expansion of the desired signal, as justified by the following theorem:

Theorem 1: The matrix $\mathbf{B}$ blocks the first $(P-1)$ order Taylor's series expansion of the desired signal components, i.e.,

$$\mathbf{B}^T\mathbf{a}_{S(P-1)}(\theta_s) = \mathbf{0} \tag{17}$$

where $\mathbf{a}_{S(P-1)}(\theta_s)$ denotes the first $(P-1)$-order Taylor's series expansion of $\mathbf{a}_S$ with respect to the look direction θ_0

and is equal to

$$\mathbf{a}_{S(P-1)}(\theta_s) = \mathbf{a}_S(\theta_0) + \sum_{i=1}^{P-1} \frac{1}{i!} \frac{\partial^i \mathbf{a}_S(\theta_s)}{\partial \theta_s^i}\bigg|_{\theta_s=\theta_0} (\theta_s - \theta_0)^i. \tag{18}$$

Proof: Since the array has been assumed to be presteered, we can, without loss of generality, take $\theta_0 = 0°$ in the following derivations. From the definition of $\mathbf{a}_S(\theta_s)$ as given in (8), we can readily verify that

$$\frac{\partial^i \mathbf{a}_S(\theta_s)}{\partial \theta_s^i}\bigg|_{\theta_s=\theta_0} = \begin{cases} \left(j\frac{2\pi}{\lambda_w} d_w\right)\mathbf{c}_1, & i = 1 \\ \left(j\frac{2\pi}{\lambda_w} d_w\right)^i \mathbf{c}_i + \sum_{r=1}^{i-1} d_r \mathbf{c}_r, & i = 2, \cdots, P-1 \end{cases} \tag{19}$$

where $\mathbf{c}_i$ is as defined in (12), and d_r are some constants (for example, if $i = 2, d_1 = 0$ and if $i = 3, d_1 = -j2\pi d_w/\lambda_w$ and $d_2 = 0$). Let $\mathbf{b}_n$ be the nth column of $\mathbf{B}$, $n = 1, \cdots, (\lfloor \frac{N-MP}{d} \rfloor + 1)(M-1)$, then

$$\begin{aligned} \mathbf{b}_n^T \mathbf{c}_i &= \sum_{k=0}^{N-1} h_m(k-(t-1)d) c_i(k) \\ &= \sum_{k=0}^{MP-1} h_m(k)(k+(1-n_0+(t-1)d))^i = 0 \\ & \qquad i = 0, 1, \cdots, P-1 \end{aligned} \tag{20}$$

where we have used Lemma 1 and the fact that every n can be expressed as $n = (\lfloor \frac{N-MP}{d} \rfloor + 1)(m-1) + t$, $m = 1, 2, \cdots, M-1$, and $t = 1, 2, \cdots, \lfloor \frac{N-MP}{d} \rfloor + 1$. Since every column of $\mathbf{B}$ is orthogonal to $\mathbf{c}_i$, $i = 0, 1 \cdots, P-1$, we can then obtain $\mathbf{B}^T \mathbf{c}_i = \mathbf{0}$, $i = 0, 1, \cdots, P-1$. It therefore follows that

$$\begin{aligned} &\mathbf{B}^T \mathbf{a}_S(\theta_s) \\ &\quad = \mathbf{B}^T \left(\mathbf{c}_0 + \sum_{i=1}^{P-1} \frac{1}{i!} \left[\left(j\frac{2\pi}{\lambda_w} d_w \right)^i \mathbf{c}_i + \sum_{r=1}^{i-1} d_r \mathbf{c}_r \right] \theta_s^i \right) = 0 \end{aligned}$$

where we have used the fact that $\mathbf{a}_S(\theta_0) = \mathbf{c}_0$. It thus completes the proof. □

Theorem 1 implies that if the desired signal is well approximated by the first P terms of the Taylor's series expansion (or the wavelet filters in $\mathbf{B}$ have sufficiently high regularity), then the desired signal will be "blocked" by the $\mathbf{B}$ matrix as required by the blocking matrix of the GSC structure. From a signal processing point of view, this may be explained by the fact that the proposed wavelet-based blocking matrix amounts to a high-pass spatial filter and thus can block the desired signal, which only contains a low spatial frequency component. To follow, we will then employ such $\mathbf{B}$ as the blocking matrix in our proposed GSC. It is noteworthy that matrix $\mathbf{B}$ is sparse and Toeplitz-like (the diagonal elements of each $\mathbf{H}_m$ are the same) and, thus, does not call for too many computations when the data pass through the blocking matrix (a detailed discussion of computational complexity is provided in Section IV).

We can also note that when $M = 2$ and $P = 1$, there is only one wavelet filter, namely $\mathbf{h}_1 = \frac{1}{\sqrt{2}}[1, -1]^T$, which is just the discrete Haar wavelet. If we further choose $d = 1$, then $\mathbf{B}$ becomes an $N \times (N-1)$ matrix $\mathbf{B}_0$ as

$$\mathbf{B}_0^T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}. \tag{21}$$

The above $\mathbf{B}_0$ corresponds to a normalized version of the widely used blocking matrix (for the look direction gain constraint), which contains ones and minus ones along the diagonals [12], [20].

C. Relationship with the LCMV Beamformer

To follow, we consider the relationship between the wavelet-based GSC described above and the associated LCMV beamformer with the derivative constraints. First, we show that in this case, the columns of $\mathbf{B}$ are orthogonal to those of $\mathbf{C}_S$.

Theorem 2: Assume that $\mathbf{C}_S$ is an $N \times S$ derivative constraint matrix of (11), then the wavelet-based matrix $\mathbf{B}$ as that of (15) (assume that $(\lfloor \frac{N-MP}{d} \rfloor + 1)(M-1) \le N - S$) satisfies

$$\mathbf{B}^T \mathbf{C}_S = \mathbf{0} \qquad \text{if } S \le P. \tag{22}$$

Proof: Since $\mathbf{B}$ is constituted by the P-regular wavelet filters, the proof then follows by noting that every column of $\mathbf{B}$ is orthogonal to that of $\mathbf{C}_S$, as shown in (20). □

As a consequence of Theorem 2, if $\mathrm{rank}(\mathbf{C}_S) + \mathrm{rank}(\mathbf{B}) = N$, then the columns of $\mathbf{C}_S$ and $\mathbf{B}$ all together orthogonally decompose the vector space $\mathcal{R}^N$, in which the weight vector lies. Therefore, if $\mathbf{w}_{\mathrm{opt}}$ denotes the weight vector of the corresponding LCMV, then $\mathbf{w}_{\mathrm{opt}}$ can be uniquely decomposed as $\mathbf{w}_{\mathrm{opt}} = \mathbf{w}_f - \mathbf{B}\mathbf{w}_a^o$. Let the weight vector of the proposed GSC be expressed as $\mathbf{w} = \mathbf{w}_f - \mathbf{B}\mathbf{w}_a$, where $\mathbf{w}_f$ and $-\mathbf{B}\mathbf{w}_a$ correspond to the weight vectors in the upper and lower paths of the GSC, respectively. It has been shown in [20] and [21] that $\mathbf{w}_a = \mathbf{w}_a^o$, viz. the GSC is equivalent to the LCMV in this case.

Next, consider the case when $\mathrm{rank}(\mathbf{C}_S) + \mathrm{rank}(\mathbf{B}) < N$, in which the columns of $\mathbf{C}_S$ together with those of $\mathbf{B}$ do not span $\mathcal{R}^N$. Let $\mathbf{A}$ be any $N \times (N - S - (\lfloor \frac{N-MP}{d} \rfloor + 1)(M-1))$ matrix, which together with $\mathbf{B}$ and $\mathbf{C}_S$, form an orthogonal

$$\mathbf{H}_m = \begin{bmatrix} h_m(0) & h_m(1) & \cdots & h_m(MP-1) & 0 & 0 & \cdots & \cdots \\ \mathbf{o}_d^T & & h_m(0) & h_m(1) & \cdots & h_m(MP-1) & 0 & \cdots \\ \vdots & & \vdots & & \ddots & \ddots & \ddots & \vdots \\ \mathbf{o}_d^T & & \mathbf{o}_d^T & \cdots & h_m(0) & h_m(1) & \cdots & h_m(MP-1) \end{bmatrix} \tag{16}$$

decomposition of $\mathcal{R}^N$, then $\mathbf{w}_{\rm opt}$ of the LCMV can now be expressed as $\mathbf{w}_{\rm opt} = \mathbf{w}_f - \tilde{\mathbf{B}}\mathbf{w}_a^o$, where $\tilde{\mathbf{B}} = [\mathbf{B}, \mathbf{A}]$ and $\tilde{\mathbf{B}}$ and $\mathbf{C}_S$ orthogonally decompose $\mathcal{R}^N$. The adaptive term now becomes

$$\begin{aligned} -\tilde{\mathbf{B}}\mathbf{w}_a^o &= -\tilde{\mathbf{B}}(\tilde{\mathbf{B}}^T\mathbf{R}_x\tilde{\mathbf{B}})^{-1}\tilde{\mathbf{B}}^T\mathbf{R}_x\mathbf{w}_f \\ &= -\mathbf{B}(\mathbf{B}^T\mathbf{R}_x\mathbf{B})^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{w}_f \\ &\quad - \mathbf{Y}\mathbf{A}(\mathbf{A}^T\mathbf{R}_x\mathbf{Y}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{Y}^H\mathbf{R}_x\mathbf{w}_f \end{aligned} \tag{23}$$

where $\mathbf{Y} = \mathbf{I} - \mathbf{B}(\mathbf{B}^T\mathbf{R}_x\mathbf{B})^{-1}\mathbf{B}^T\mathbf{R}_x$, $\mathbf{R}_x$ is assumed to be nonsingular, and we have used the matrix inversion formula in [22]. We can note that the first term of (23) is the same as $-\mathbf{B}\mathbf{w}_a$, where $\mathbf{w}_a = (\mathbf{B}^T\mathbf{R}_x\mathbf{B})^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{w}_f$, whereas the second term accounts for the role of the column space of $\mathbf{A}$. In this case, the proposed GSC belongs to a partially adaptive beamformer since only a portion of the $(N - S)$ adaptive dimension is utilized. The computational complexity called for is reduced as those of previously addressed partially adaptive beamformers. Such an offset of the weight vector, however, inevitably causes some degradation of the resulting performance. More specifically, it can be shown that the array output power of the proposed GSC $\sigma_y^2 (\triangleq \mathbf{w}^H\mathbf{R}_x\mathbf{w})$ can be expressed as [23]

$$\begin{aligned} \sigma_y^2 &= \mathbf{w}_f^H\mathbf{R}_x\mathbf{w}_f - \mathbf{w}_f^H\mathbf{R}_x\mathbf{B}(\mathbf{B}^T\mathbf{R}_x\mathbf{B})^{-1}\mathbf{B}^T\mathbf{R}_x\mathbf{w}_f \\ &= \left(\sigma_y^o\right)^2 + \left|(\mathbf{A}^T\mathbf{R}_x\mathbf{Y}\mathbf{A})^{-\frac{1}{2}}\mathbf{A}^T\mathbf{Y}^H\mathbf{R}_x\mathbf{w}_f\right|^2 \end{aligned} \tag{24}$$

where $(\sigma_y^o)^2 \triangleq \mathbf{w}_{\rm opt}^H\mathbf{R}_x\mathbf{w}_{\rm opt} = \mathbf{w}_f^H\mathbf{R}_x\mathbf{w}_f - \mathbf{w}_f^H\mathbf{R}_x\tilde{\mathbf{B}}(\tilde{\mathbf{B}}^T\mathbf{R}_x\tilde{\mathbf{B}})^{-1}\tilde{\mathbf{B}}^T\mathbf{R}_x\mathbf{w}_f$ denotes the output power of the LCMV, $\|\cdot\|$ is the Euclidean norm, and we have used (23). Therefore, the output power of the proposed GSC σ_y^2 is greater than that of the corresponding LCMV, $(\sigma_y^o)^2$. The results of (23) and (24) can also be verified to be independent of the chosen matrix $\mathbf{A}$ by using the fact that $\tilde{\mathbf{B}}\mathbf{w}_a^o = \mathbf{w}_f - \mathbf{R}_x^{-1}\mathbf{C}_S(\mathbf{C}_S^T\mathbf{R}_x^{-1}\mathbf{C}_S)^{-1}\mathbf{C}_S^T\mathbf{w}_f$ [20].

It is noteworthy that the array output power addressed above are based on the Wiener solution. In practice, the resulting output power is the convergent mean squares error (MSE) of the LMS algorithm, which includes not only the MSE based on the optimum Wiener solution but also the excess one due to the employed noisy gradient [1]. As such, it is also crucial to reduce the excess MSE of the GSC in order to attain deep nulls in the resulting beampattern. For the proposed GSC, the excess MSE depends on the choices of parameters M, P, and d, which will be discussed in more detail in the next subsection.

D. Choices of Parameters

In this subsection, we treat the issues of determining the parameters M, P, and d. Our consideration will be based on the excess MSE, the adaptive response of the LMS algorithm, and the output performance. As discussed above, to achieve deep nulls in the beampatterns, we would like to choose these parameters to render a small excess MSE or, alternatively, misadjustment, which provides a measure of the excess MSE as compared with that obtained by the optimum Wiener solution. Also, since the convergence behavior of the LMS algorithm is highly dependent on the eigenvalue spreads of the covariance matrix $\mathbf{R}_u$ of the blocking matrix output $\mathbf{u}$, where $\mathbf{R}_u \triangleq \mathcal{E}\{\mathbf{u}\mathbf{u}^H\}$ and $\mathbf{u} = \mathbf{B}^T\mathbf{x}$ with $\mathbf{B}$ being defined by (15). We would pick these parameters to yield a smaller eigenvalue spread of $\mathbf{R}_u$, which in general leads to faster response. Finally, the parameters chosen should optimize the output performance as well, which is measured in terms of maximum output signal-to-interference-plus-noise-ratio (SINR).

It has been shown in [9] that when data pass through a partial transform the eigenvalue spread and the misadjustment of the succeeding LMS algorithm reduce. It is, however, difficult to derive exact expressions for these two values. In [5], they provide an alternative approach by deriving an upper bound for the eigenvalue spreads of the transformed data in the analysis of transform domain LMS algorithm and then claim that the optimal transform is the one which can attain the lowest upperbound. Likewise, to follow, we provide upper bounds for these two values and then select parameters which attain smaller upper bounds. As [5], although it is not necessarily true that a smaller upper bound would imply a smaller true value, it still provides some guidelines for the choices of these parameters. Additionally, it has been observed in various scenarios [23] (see also the provided simulations in the next section) that choices of parameters based on these suggestive rules are generally in agreement with the true values.

If the desired signal, the jammers, and the contaminated noise are assumed to be uncorrelated, then after some manipulations (detailed derivations are provided in Appendix A), we can obtain the following inequality for the misadjustment $\mathcal{M}$:

$$\mathcal{M} \le \frac{\eta}{2}\left(\left\lfloor\frac{N-MP}{d}\right\rfloor + 1\right)\left(M\sum_{i=1}^{J}\sigma_{Ji}^2 + (M-1)\sigma_n^2\right) \tag{25}$$

where η is the step-size used in the LMS algorithm, σ_{Ji}^2, $i = 1, 2, \cdots, J$, and σ_n^2 stand for the power of the ith jammer and contaminated white noise measured at each element of the array, respectively. Similarly, we can also obtain the following inequality for the eigenvalue spread of $\mathbf{R}_u$ (detailed derivations are provided in Appendix B):

$$\begin{aligned} &\frac{\lambda_{\max}(\mathbf{R}_u)}{\lambda_{\min}(\mathbf{R}_u)} \\ &\quad \le \frac{M\left(\left\lfloor\frac{N-MP}{d}\right\rfloor + 1\right)\sum_{i=1}^{J}\sigma_{Ji}^2 + \sigma_n^2\lambda_{\max}(\mathbf{B}^T\mathbf{B})}{\sigma_n^2\lambda_{\min}(\mathbf{B}^T\mathbf{B})} \end{aligned} \tag{26}$$

where $\lambda_{\max}(\mathcal{A})$ and $\lambda_{\min}(\mathcal{A})$ denote the maximum and minimum eigenvalues of the matrix inside the bracket, respectively, and σ_n^2 is assumed not to equal to zero.

Now, we consider how to choose the parameters M, P, and d to minimize the upper bounds of the inequalities of (25) and (26). We can observe that both the upper bounds of (25) and (26) decrease as P increases (for fixed d and M). It follows, therefore, that the choice of wavelet filters with high regularity is preferred for smaller misadjustments and eigenvalue spreads of $\mathbf{R}_u$.

As for the choice of d, note that since $\mathrm{tr}(\mathbf{B}^T\mathbf{B}) = (\lfloor\frac{N-MP}{d}\rfloor + 1)(M-1) =$ summation of all eigenvalues of $(\mathbf{B}^T\mathbf{B})$, where $\mathrm{tr}(\cdot)$ is the matrix trace operator, we can readily

deduce that $\lambda_{\max}(\mathbf{B}^T\mathbf{B}) \leq \mathrm{tr}(\mathbf{B}^T\mathbf{B}) = (\lfloor \frac{N-MP}{d} \rfloor + 1)(M-1)$ and $\lambda_{\min}(\mathbf{B}^T\mathbf{B}) \leq 1$ (otherwise, it will be in contradiction). Also, since the upper bound of (26) is dictated mainly by the eigenvalue spread of $(\mathbf{B}^T\mathbf{B})$, we can then choose d to minimize this value. Recall that when d is a multiple of M, $\mathbf{B}^T\mathbf{B} = \mathbf{I}$ for which $\lambda_{\max}(\mathbf{B}^T\mathbf{B}) = \lambda_{\min}(\mathbf{B}^T\mathbf{B}) = 1$. As a result, we can choose d as a multiple of M, which then yields a smaller upper bound of (26) (for fixed P and M). In particular, when $M = 2$, choosing $d = 2$ would yield a smaller upper bound of (26) as compared with that by using $d = 1$, which was employed by $\mathbf{B}_0$ of (21). This explains why the proposed approach in general converges more rapidly than that of the previous work which is based on $\mathbf{B}_0$.

The chosen parameters should also maximize the output SINR. First, note that a wavelet filter with high regularity exhibits a fast decaying response. Since the wavelet filters stand for the high-pass spatial filtering (while the scaling filter corresponds to the low-pass spatial filtering), high regularity of the wavelet filters will then form a sharper and wider null in the low spatial frequency part of the spatial response of the blocking matrix. As such, the blocking matrix will block not only the desired signal but also the interfering signals which are supposed to pass through. Therefore, the SINR will somehow degrade if the regularity P of the wavelet filters is chosen widely large, in particular for jammers in the vicinity of the look direction. Additionally, from [13], we know that $(M-1)$ wavelet filters with a larger M will provide better energy compaction, leading to a narrow null in the low spatial frequency part of the blocking matrix spatial response. Along the same line, a larger M is, therefore, preferred as it will cause less degradation of the resulting SINR. The detailed performance analysis can be found in [23] and [24].

IV. Experimental Results and Discussion

Some simulations are carried out in this section to assess the proposed wavelet-based approach. We demonstrate that the proposed one can perform as well as the previous ones but with lower computational complexity. Also, the misadjustments of the LMS algorithm and the eigenvalue spreads of $\mathbf{R}_u$ are investigated with different choices of parameters in various scenarios. To follow, we consider two examples, both of which are based on linear equispaced arrays consisting of N omnidirectional sensors spaced one-half wavelength apart. The GSC, with various blocking matrices along with the derivative constraints, is utilized for the determination of beamformer weights. The received data $\mathbf{x}$ is as that of (7). The employed P-regular M-band wavelet filters are as those addressed in [13].

Example 1: The array considered in Example 1 is composed of 16 sensors. The interference environment consists of one jammer. The gain constraint, i.e., $\mathbf{C}_1$, is used for the GSC. The arrival angle of the desired signal is 0°. The interference-to-noise ratio (INR) is 25 dB and the contaminated white Gaussian noise (WGN) is 10 dB.

Tables I and II list the misadjustments of the LMS algorithm and the eigenvalue spreads of $\mathbf{R}_u$, respectively, for jammers arriving from several different directions based on: 1) $\mathbf{B}_0$ of (21); 2) $\breve{\mathbf{B}}_0$ as suggested in [20], where

TABLE I
The Misadjustments of the LMS Algorithm inVarious Scenarios of Example 1, Where the Previous Approaches Use the Blocking Matrices $\mathbf{B}_0$ and $\breve{\mathbf{B}}_0$

Jamming	previous approaches		$M=2$			
Angles	$\mathbf{B}_0$	$\breve{\mathbf{B}}_0$	$P=1$	$P=3$	$P=4$	$P=5$
$\theta=\pm15^o$	2.997×10^{-3}	7.639×10^{-3}	1.598×10^{-3}	2.384×10^{-4}	9.887×10^{-5}	4.224×10^{-5}
$\theta=\pm35^o$	1.169×10^{-2}	1.242×10^{-2}	6.235×10^{-3}	5.381×10^{-3}	4.676×10^{-3}	3.869×10^{-3}
$\theta=\pm55^o$	1.751×10^{-2}	1.197×10^{-2}	9.340×10^{-3}	7.569×10^{-3}	6.328×10^{-3}	5.066×10^{-3}
$\theta=\pm75^o$	1.895×10^{-2}	1.264×10^{-2}	1.011×10^{-2}	7.601×10^{-3}	6.335×10^{-3}	5.068×10^{-3}
Jamming	$M=3$			$M=4$		$M=5$
Angles	$P=3$	$P=4$	$P=5$	$P=2$	$P=3$	$P=3$
$\theta=\pm15^o$	1.276×10^{-3}	6.928×10^{-4}	2.863×10^{-4}	4.518×10^{-3}	3.015×10^{-3}	2.830×10^{-3}
$\theta=\pm35^o$	5.700×10^{-3}	3.802×10^{-3}	1.901×10^{-3}	7.585×10^{-3}	5.069×10^{-3}	3.067×10^{-3}
$\theta=\pm55^o$	5.672×10^{-3}	3.795×10^{-3}	1.900×10^{-3}	7.366×10^{-3}	4.965×10^{-3}	3.170×10^{-3}
$\theta=\pm75^o$	5.368×10^{-3}	3.614×10^{-3}	1.819×10^{-3}	7.607×10^{-3}	5.072×10^{-3}	3.114×10^{-3}

TABLE II
The Eigenvalue Spreads of $\mathbf{R}_u$ in Various Scenarios of Example 1, Where the Previous Approaches Use the Blocking Matrices $\mathbf{B}_0$ and $\breve{\mathbf{B}}_0$

Jamming	previous approaches		$M=2$			
Angles	$\mathbf{B}_0$	$\breve{\mathbf{B}}_0$	$P=1$	$P=3$	$P=4$	$P=5$
$\theta=\pm15^o$	7.339×10^4	6.526×10^4	7.922×10^2	1.142×10^2	4.544×10^1	1.812×10^1
$\theta=\pm35^o$	3.032×10^5	1.117×10^5	3.110×10^3	2.685×10^3	2.334×10^3	1.931×10^3
$\theta=\pm55^o$	4.549×10^5	1.077×10^5	4.663×10^3	3.779×10^3	3.160×10^3	2.530×10^3
$\theta=\pm75^o$	4.908×10^5	1.134×10^5	5.046×10^3	3.796×10^3	3.163×10^3	2.531×10^3
Jamming	$M=3$			$M=4$		$M=5$
Angles	$P=3$	$P=4$	$P=5$	$P=2$	$P=3$	$P=3$
$\theta=\pm15^o$	6.331×10^2	3.434×10^2	1.422×10^2	2.251×10^3	1.502×10^3	1.412×10^3
$\theta=\pm35^o$	2.845×10^3	1.898×10^3	9.497×10^2	3.785×10^3	2.530×10^3	1.530×10^3
$\theta=\pm55^o$	2.831×10^3	1.894×10^3	9.489×10^2	3.675×10^3	2.477×10^3	1.582×10^3
$\theta=\pm75^o$	2.679×10^3	1.804×10^3	9.087×10^2	3.795×10^3	2.531×10^3	1.554×10^3

$$\breve{\mathbf{B}}_0^T = \begin{bmatrix} \breve{\mathbf{B}} & \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \breve{\mathbf{B}} & \mathbf{0}_3 & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \breve{\mathbf{B}} & \mathbf{0}_3 \\ \mathbf{0}_3 & \mathbf{0}_3 & \mathbf{0}_3 & \breve{\mathbf{B}} \end{bmatrix} \tag{27}$$

in which $\breve{\mathbf{B}} = [\breve{\mathbf{b}}_1,\ \breve{\mathbf{b}}_2,\ \breve{\mathbf{b}}_3]^T$ with $\breve{\mathbf{b}}_1 = \frac{1}{|\tilde{\mathbf{b}}_1|}\tilde{\mathbf{b}}_1$, $\breve{\mathbf{b}}_2 = \frac{1}{|\tilde{\mathbf{b}}_2|}\tilde{\mathbf{b}}_2$, $\breve{\mathbf{b}}_3 = \frac{1}{|\tilde{\mathbf{b}}_3|}\tilde{\mathbf{b}}_3$, $\tilde{\mathbf{b}}_1 = [0.87, -0.29, -0.29, -0.29]^T$, $\tilde{\mathbf{b}}_2 = [0, 0.82, -0.41, -0.41]^T$, $\tilde{\mathbf{b}}_3 = [0, 0, 0.71, -0.71]^T$, and $\mathbf{0}_3$ is a 3×3 zero matrix; and 3) the proposed wavelet-based blocking matrix as given in (15) with various choices of M and P (d is chosen to be the same as M, as suggested in the previous section). For a more revealing comparison, we take $M = 4$ ($d = 4$) and $P = 3$, resulting in a 16×6 blocking matrix, for the jammer with an arrival direction of 55°. The resulting array output beampattern based on an ensemble average of 50 independent trials by using the proposed blocking matrix is shown in Fig. 2. The difference patterns (gain difference between the proposed one and that of the previous scheme) based on $\mathbf{B}_0$ and $\breve{\mathbf{B}}_0$ are shown in Figs. 3 and 4, respectively. From the figures, we can find that the beampattern based on the proposed approach can attain a deep null at the jammer direction more rapidly than those based on $\mathbf{B}_0$ and $\breve{\mathbf{B}}_0$.

Example 2: The array considered in Example 2 is composed of 48 sensors. The interference environment now consists of two jammers. The second-order derivative constraints, i.e., $\mathbf{C}_3$ (with $n_0 = 24$), is used for the GSC. The incident angle of the desired signal is still 0°.

In Tables III and IV, we list the misadjustments and the eigenvalue spreads of $\mathbf{R}_u$ for jammers arriving from several different directions based on: 1) the (normalized) $N \times (N-3)$

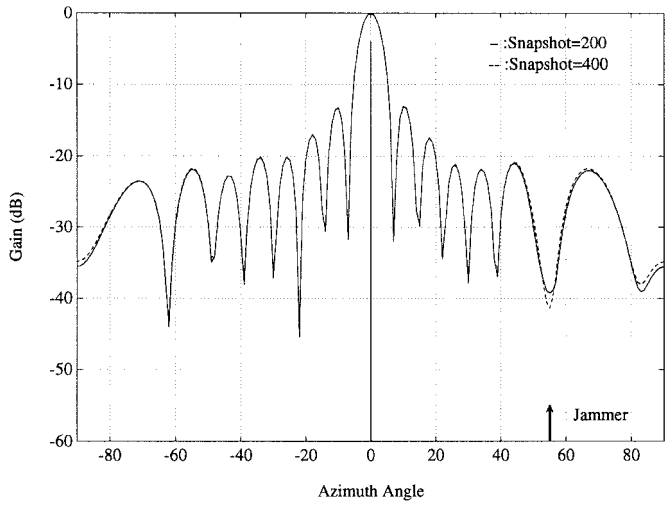

Fig. 2. The output beampattern of Example 1 by using the proposed blocking matrix.

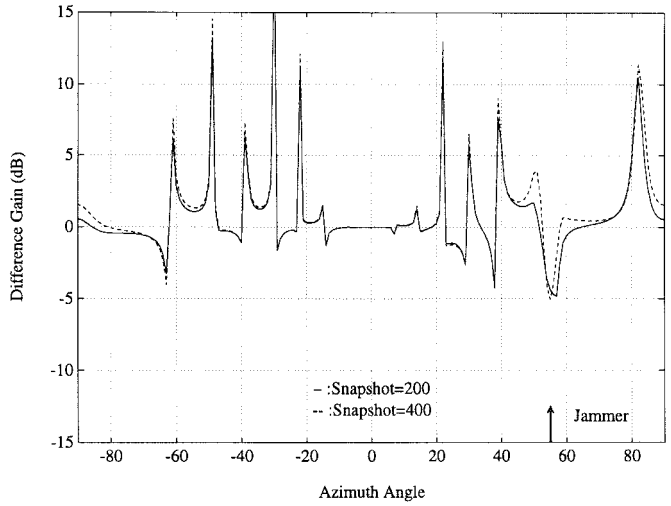

Fig. 3. The difference pattern of Example 1 by using the blocking matrix $\mathbf{B}_0$.

blocking matrix $\mathbf{B}_2$ [25] as

$$\mathbf{B}_2^T = \frac{1}{\sqrt{20}} \begin{bmatrix} 1 & -3 & 3 & -1 & \cdots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -3 & 3 & -1 \end{bmatrix} \tag{28}$$

and 2) the proposed wavelet-based blocking matrix as given in (15) with various choices of M and P (d is chosen to be the same as M). For a revealing comparison, we take $M = 6$ ($d = 6$) and $P = 7$, resulting in a 48×10 blocking matrix, for the jammers arriving from directions ($-45°$, $50°$). The INR's are equal to 20 dB and 30 dB for jammers with directions of arrival $-45°$ and $50°$, respectively, whereas the WGN has a 10-dB SNR. The resulting array output beampattern based on an ensemble average of 50 independent trials by using the proposed blocking matrix is shown in Fig. 5. The difference pattern based on $\mathbf{B}_2$ is shown in Fig. 6. From the figures, again we can see that the beampattern based on the proposed approach can attain deep nulls at the jammer directions faster than that based on $\mathbf{B}_2$.

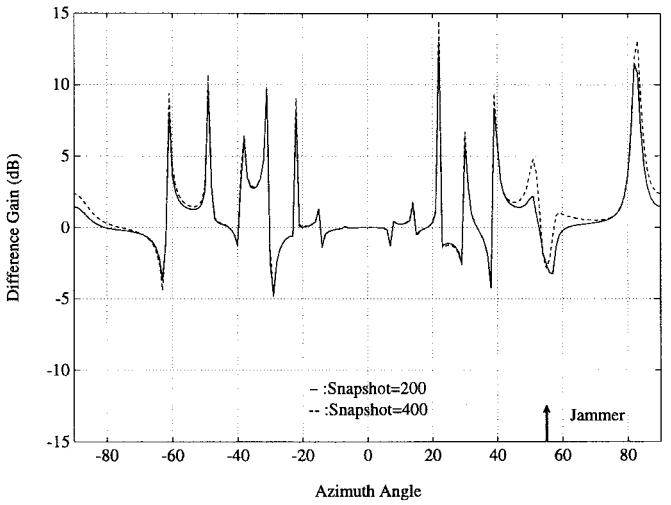

Fig. 4. The difference pattern of Example 1 by using the blocking matrix $\mathbf{B}_0$.

TABLE III
THE MISADJUSTMENTS OF THE LMS ALGORITHM IN VARIOUS SCENARIOS OF EXAMPLE 2, WHERE THE PREVIOUS APPROACH USES THE BLOCKING MATRIX $\mathbf{B}_2$

Jamming Angles(θ_1, θ_2)	previous approach	$M=2$			$M=4$	
		$P=5$	$P=6$	$P=7$	$P=5$	$P=6$
$(-15°, +25°)$	2.220×10^{-4}	2.373×10^{-4}	2.013×10^{-4}	1.721×10^{-4}	1.285×10^{-3}	1.128×10^{-3}
$(-45°, +50°)$	4.252×10^{-3}	1.984×10^{-3}	1.893×10^{-3}	1.798×10^{-3}	1.564×10^{-3}	1.372×10^{-3}
$(+35°, +75°)$	4.416×10^{-3}	1.768×10^{-3}	1.700×10^{-3}	1.627×10^{-3}	1.605×10^{-3}	1.405×10^{-3}
$(-45°, +35°)$	2.710×10^{-3}	1.750×10^{-3}	1.690×10^{-3}	1.622×10^{-3}	1.587×10^{-3}	1.391×10^{-3}
Jamming Angles(θ_1, θ_2)	$M=4$ $P=7$	$M=6$			$M=8$	
		$P=5$	$P=6$	$P=7$	$P=5$	$P=6$
$(-15°, +25°)$	9.689×10^{-4}	1.203×10^{-3}	9.031×10^{-4}	6.023×10^{-4}	8.024×10^{-4}	4.014×10^{-4}
$(-45°, +50°)$	1.179×10^{-3}	1.204×10^{-3}	9.033×10^{-4}	6.023×10^{-4}	8.035×10^{-4}	4.017×10^{-4}
$(+35°, +75°)$	1.204×10^{-3}	1.204×10^{-3}	9.034×10^{-4}	6.024×10^{-4}	8.031×10^{-4}	4.016×10^{-4}
$(-45°, +35°)$	1.194×10^{-3}	1.204×10^{-3}	9.034×10^{-4}	6.024×10^{-4}	8.030×10^{-4}	4.016×10^{-4}

TABLE IV
THE EIGENVALUE SPREADS OF $\mathbf{R}_u$ IN VARIOUS SCENARIOS OF EXAMPLE 2, WHERE THE PREVIOUS APPROACH USES THE BLOCKING MATRIX $\mathbf{B}_2$

Jamming Angles(θ_1, θ_2)	previous approach	$M=2$			$M=4$	
		$P=5$	$P=6$	$P=7$	$P=5$	$P=6$
$(-15°, +25°)$	3.089×10^{9}	9.032×10^{2}	7.745×10^{2}	6.655×10^{2}	3.204×10^{3}	2.832×10^{3}
$(-45°, +50°)$	3.760×10^{10}	4.158×10^{3}	3.798×10^{3}	3.789×10^{3}	3.568×10^{3}	3.253×10^{3}
$(+35°, +75°)$	5.636×10^{10}	4.009×10^{3}	3.841×10^{3}	3.603×10^{3}	3.339×10^{3}	2.938×10^{3}
$(-45°, +35°)$	2.938×10^{10}	3.975×10^{3}	3.807×10^{3}	3.580×10^{3}	3.493×10^{3}	2.830×10^{3}
Jamming Angles(θ_1, θ_2)	$M=4$ $P=7$	$M=6$			$M=8$	
		$P=5$	$P=6$	$P=7$	$P=5$	$P=6$
$(-15°, +25°)$	2.419×10^{3}	2.843×10^{3}	2.239×10^{3}	1.591×10^{3}	1.956×10^{3}	9.032×10^{2}
$(-45°, +50°)$	2.964×10^{3}	2.489×10^{3}	1.861×10^{3}	1.275×10^{3}	1.712×10^{3}	8.584×10^{2}
$(+35°, +75°)$	2.458×10^{3}	2.492×10^{3}	1.947×10^{3}	1.414×10^{3}	1.670×10^{3}	9.783×10^{2}
$(-45°, +35°)$	2.676×10^{3}	2.663×10^{3}	2.036×10^{3}	1.541×10^{3}	1.792×10^{3}	9.089×10^{2}

Remarks:

1) Smaller upper bounds of (25) and (26) (with appropriate choices of parameters) in general imply smaller misadjustments of the succeeding LMS algorithm and the eigenvalue spreads of $\mathbf{R}_u$. These values decrease as P increases for fixed M, which is also in consistence with (25) and (26). In contrast, these values decrease when M decreases only when $(\lfloor \frac{N-MP}{d} \rfloor + 1)$ is fixed. More simulation results to confirm these observations can be found in [23].
2) The misadjustment and the eigenvalue spread of $\mathbf{R}_u$, in general, decrease as the jammer angle is close to the mainlobe. To explain this (for clarity, only one jammer case is considered), suppose that $\mathbf{r} = [1, e^{j\mu_1}, \cdots, e^{j(MP-1)\mu_1}]^T$, which is equal to the

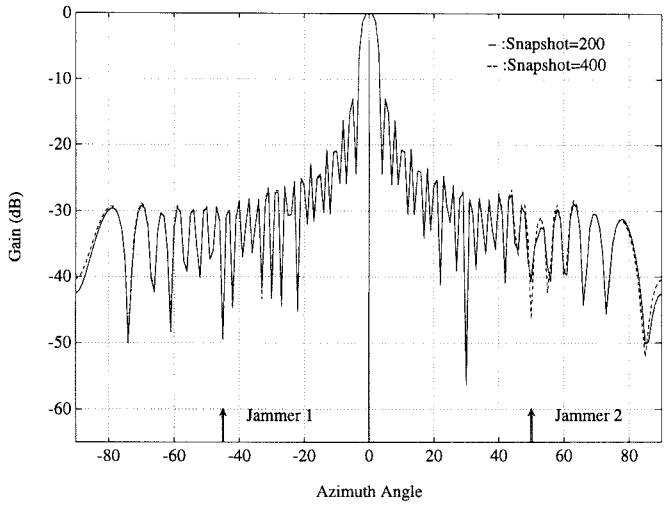

Fig. 5. The output beampattern of Example 2 by using the proposed blocking matrix.

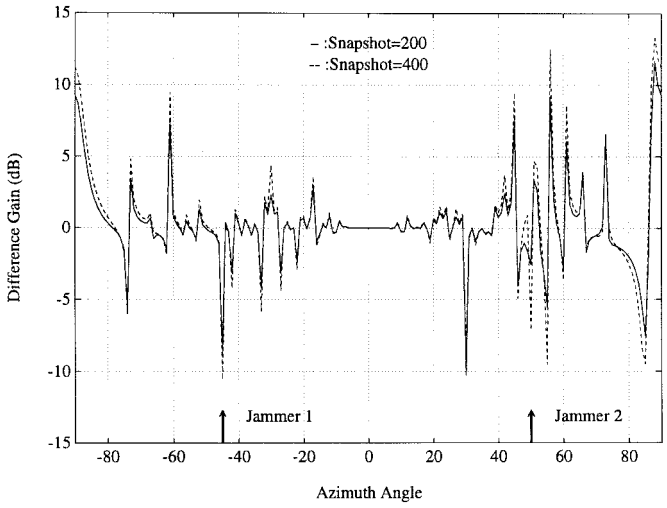

Fig. 6. The difference pattern of Example 2 by using the blocking matrix $\mathbf{B}_2$.

first MP terms of $\mathbf{a}_J(\theta_1)$. $\mathbf{r}_{(P-1)}(\theta_1)$ denotes the first $(P-1)$ order Taylor's series expansion of $\mathbf{a}_J(\theta_1)$ with respect to the look direction θ_0 and is equal to $\mathbf{r}_{(P-1)}(\theta_1) = \mathbf{r}(\theta_0) + \sum_{i=1}^{P-1} \frac{1}{i!} \frac{\partial^i \mathbf{r}(\theta_1)}{\partial \theta_1^i}|_{\theta_1=\theta_0} (\theta_1 - \theta_0)^i$. Following the same steps as those of Theorem 1, it can be shown that $\mathbf{h}_m^T \mathbf{r}_{(P-1)} = 0$. Therefore, λ_J as given in (35) can be reduced to

$$\lambda_J = \mathrm{tr}\,(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = \sigma_J^2 \left(\left\lfloor \frac{N-MP}{d} \right\rfloor + 1 \right) \sum_{m=1}^{M-1} \left| \mathbf{h}_m^T (\mathbf{r} - \mathbf{r}_{(P-1)}) \right|^2 \tag{29}$$

which can be approximated by

$$\lambda_J \approx \sigma_J^2 \left(\left\lfloor \frac{N-MP}{d} \right\rfloor + 1 \right) \left(\frac{2\pi}{\lambda_w} d_w \right)^{2P} (\theta_1 - \theta_0)^{2P} \times \frac{(MP)^{2P+1}}{2P+1} \left[\frac{(P!)}{(2P)!} \right]^2 \tag{30}$$

where we have used [23, eq. (6.57)]. It therefore follows that λ_J and the corresponding misadjustment will decrease by (36). Similarly, we can demonstrate that the eigenvalue spread will also decrease by using (41).

3) For a partially adaptive GSC with a blocking matrix $\mathbf{B}$ of size $N \times K$, the number of multiplications required in the lower path to attain the desired beampattern is $(N + 2K)T$, where T is the number of iterations required for the LMS to converge. The first term N and the second term $2K$ of the above expression denote the numbers of multiplications required when the data pass through the blocking matrix and for adaptive processing per iteration, respectively. Notice that since the former are fully parallel [26], only one row-column multiplication is counted. If, in particular, each row of $\mathbf{B}^T$ only consists of L nonzero elements, then the number of multiplications required is reduced to $(L+2K)T$. Based on this, the proposed wavelet-based GSC then requires

$$\left(MP + 2\left(\left\lfloor \frac{N-MP}{d} \right\rfloor + 1 \right)(M-1) \right) T \tag{31}$$

multiplications, where we have exploited the sparsity of (15). As a consequence, we can note that the number of multiplications per iteration then decreases from 32 (based on $\mathbf{B}_0$) and 39 (based on $\mathbf{B}_1$) to 24 (based on the proposed $\mathbf{B}$) in Example 1 and from 94 (based on $\mathbf{B}_2$) to 62 (based on the proposed $\mathbf{B}$) in Example 2. Also, as can be noted from Figs. 2–6 that the previous schemes need more iterations and thus call for more computations than the proposed one to attain deep nulls in the jammer directions. These computational savings will become even more substantial for a larger N and an appropriate choice of M and P.

4) The rationale of choosing the blocking matrices $\mathbf{B}_0$, $\check{\mathbf{B}}_0$, and $\mathbf{B}_2$, as suggested in [20] and [25], is due to their sparsity and simplicity. Other approaches such as the singular value decomposition (SVD) and the QR factorization can also be employed to design the blocking matrix. These blocking matrices, however, are generally not sparse and thus induce more computations when data pass through (see the above discussion). Additionally, the determination of these blocking matrices are computationally more demanding. A comparison of these methods can be found in [27].

V. Conclusion

In this paper, we describe a new low complexity wavelet-based GSC in which the blocking matrix is constituted by a set of regular M-band wavelet filters. This new blocking matrix, as justified analytically, can block the desired signal as required, provided the employed wavelet filters are highly regular. Furthermore, the outputs of the blocking matrix not only have reduced dimensions, but their covariance matrices in general also have smaller eigenspreads as compared with those of previous works. As a result, the array can form the desired null-steering beampatterns with substantially reduced computational complexity. Some suggestions regarding the

choices of parameters have also been addressed to achieve smaller eigenvalue spreads, smaller misadjustments of the LMS algorithm and larger output SINR's. Simulation results confirm the proposed wavelet-based approach as well as these suggestive guidelines for the choices of parameters.

APPENDIX A

In this appendix, we derive the inequality for the misadjustment of (25). Let us begin our derivations with $J = 1$ (one jammer case). The jammer covariance matrix $\mathbf{R}_J$ is

$$\begin{aligned}\mathbf{R}_J &= \sigma_J^2 \mathbf{a}_J \mathbf{a}_J^H \\ &= \sigma_J^2 \begin{bmatrix} 1 & e^{-j\mu_1} & \cdots & e^{-j(N-1)\mu_1} \\ e^{j\mu_1} & 1 & \cdots & e^{-j(N-2)\mu_1} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(N-1)\mu_1} & e^{j(N-2)\mu_1} & \cdots & 1 \end{bmatrix}\end{aligned} \tag{32}$$

where σ_J^2 is the jamming power. The covariance matrix of the jammer after passing through the blocking matrix can then be expressed as

$$\mathbf{B}^T \mathbf{R}_J \mathbf{B} = \sigma_J^2 \begin{bmatrix} \mathbf{R}_1 & \times & \cdots & \times \\ \times & \mathbf{R}_2 & \cdots & \times \\ \vdots & \vdots & \ddots & \vdots \\ \times & \times & \cdots & \mathbf{R}_{M-1} \end{bmatrix} \tag{33}$$

where (34), shown at the bottom of the page, $m = 1, 2, \cdots, M-1$, $\mathbf{r} = [1, e^{j\mu_1}, \cdots, e^{j(MP-1)\mu_1}]^T$, and $\times$ are the irrelevant cross terms. From (32), we can note that $\mathbf{R}_J$ is an outer product of $\mathbf{a}_J$, so $\operatorname{rank}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = 1$ and thus has only one nonzero eigenvalue as

$$\lambda_J = \operatorname{tr}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = \sigma_J^2 \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \sum_{m=1}^{M-1} \left| \mathbf{h}_m^T \mathbf{r} \right|^2 \tag{35}$$

where we have used the fact that $\operatorname{tr}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) =$ summation of all eigenvalues of $(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = \lambda_J$. Next, consider the covariance matrix $\mathbf{R}_u$ of the blocking matrix output. Since the jamming and the contaminated noise are assumed to be uncorrelated and the desired signal has been blocked, we can obtain $\mathbf{R}_u = \mathbf{B}^T \mathbf{R}_J \mathbf{B} + \mathbf{B}^T \mathbf{R}_n \mathbf{B}$, where $\mathbf{R}_n \triangleq \mathcal{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma_n^2 \mathbf{I}$ is the noise covariance matrix. The misadjustment $\mathcal{M}$ is therefore given by

$$\mathcal{M} = \frac{\eta}{2} \left(\lambda_J + \sigma_n^2 (M-1) \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \right) \tag{36}$$

where we have used the fact that $\mathcal{M} = \frac{\eta}{2} \operatorname{tr}(\mathbf{R}_u)$ for small η [1]. Also, since $\sum_{m=0}^{M-1} |\mathbf{h}_m^T \mathbf{r}|^2 = M$ [14], which implies that

$$\sum_{m=1}^{M-1} \left| \mathbf{h}_m^T \mathbf{r} \right|^2 \le M \tag{37}$$

we can then arrive at the following inequality:

$$\mathcal{M} \le \frac{\eta}{2} \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \left(\sigma_J^2 M + \sigma_n^2 (M-1) \right) \tag{38}$$

where we have substituted (37) into (35) and used (36). The above derivations can be extended to multiple jammers $(J > 1)$ case straightforwardly. Now, the jammer covariance matrix becomes $\mathbf{R}_J = \sum_{i=1}^{J} \mathbf{R}_{Ji}$, where $\mathbf{R}_{Ji}$, $i = 1, 2, \cdots, J$ correspond to the covariance matrix of the ith jammer and we have used the fact that the jammers are uncorrelated. Similar to (35), $\operatorname{tr}(\mathbf{B}^T \mathbf{R}_J \mathbf{B})$ is now bounded by

$$\begin{aligned}\operatorname{tr}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) &= \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \sum_{i=1}^{J} \sigma_{Ji}^2 \sum_{m=1}^{M-1} \left| \mathbf{h}_m^T \mathbf{r}_i \right|^2 \\ &\le M \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \sum_{i=1}^{J} \sigma_{Ji}^2\end{aligned} \tag{39}$$

where σ_{Ji}^2, denotes the ith jamming power and $\mathbf{r}_i = [1, e^{j\mu_i}, \cdots, e^{j(MP-1)\mu_i}]^T$, $i = 1, 2, \cdots, J$. Following the same approach as the one jammer case, we can get the following inequality for the misadjustment

$$\mathcal{M} \le \frac{\eta}{2} \left(\left\lfloor \frac{N - MP}{d} \right\rfloor + 1 \right) \left(M \sum_{i=1}^{J} \sigma_{Ji}^2 + (M-1) \sigma_n^2 \right).$$

APPENDIX B

Here we derive the upper bound for the eigenvalue spread of (26). As above, we also begin our derivations with the simpler one jammer case. Invoking the inequalities of $\lambda_{\min}(\mathcal{A}) + \lambda_{\min}(\mathcal{B}) \le \lambda_{\min}(\mathcal{A}+\mathcal{B})$ and $\lambda_{\max}(\mathcal{A}) + \lambda_{\max}(\mathcal{B}) \ge \lambda_{\max}(\mathcal{A}+\mathcal{B})$ [28] lead to the following inequality:

$$\frac{\lambda_{\max}(\mathbf{R}_u)}{\lambda_{\min}(\mathbf{R}_u)} \le \frac{\lambda_{\max}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) + \lambda_{\max}(\mathbf{B}^T \mathbf{R}_n \mathbf{B})}{\lambda_{\min}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) + \lambda_{\min}(\mathbf{B}^T \mathbf{R}_n \mathbf{B})}. \tag{40}$$

The inequality of (40) can be further simplified as

$$\frac{\lambda_{\max}(\mathbf{R}_u)}{\lambda_{\min}(\mathbf{R}_u)} \le \frac{\lambda_J + \sigma_n^2 \lambda_{\max}(\mathbf{B}^T \mathbf{B})}{\sigma_n^2 \lambda_{\min}(\mathbf{B}^T \mathbf{B})} \tag{41}$$

where we have used the facts that $\lambda_{\max}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = \lambda_J$, $\lambda_{\min}(\mathbf{B}^T \mathbf{R}_J \mathbf{B}) = 0$, $\lambda_{\max}(\mathbf{B}^T \mathbf{R}_n \mathbf{B}) = \sigma_n^2 \lambda_{\max}(\mathbf{B}^T \mathbf{B})$, and

$$\begin{aligned}\mathbf{R}_m &= \sigma_J^2 \left| \mathbf{h}_m^T \mathbf{r} \right|^2 \\ &\times \begin{bmatrix} 1 & e^{-jd\mu_1} & \cdots & e^{-jd\left\lfloor \frac{N-MP}{d} \right\rfloor \mu_1} \\ e^{jd\mu_1} & 1 & \cdots & e^{-jd\left(\left\lfloor \frac{N-MP}{d} \right\rfloor - 1 \right) \mu_1} \\ \vdots & \vdots & \ddots & \vdots \\ e^{jd\left\lfloor \frac{N-MP}{d} \right\rfloor \mu_1} & e^{jd\left(\left\lfloor \frac{N-MP}{d} \right\rfloor - 1 \right) \mu_1} & \cdots & 1 \end{bmatrix}\end{aligned} \tag{34}$$

$\lambda_{\min}(\mathbf{B}^T\mathbf{R}_n\mathbf{B}) = \sigma_n^2\lambda_{\min}(\mathbf{B}^T\mathbf{B})$. Substituting (35) and (37) into (41), we arrive at the following inequality:

$$\frac{\lambda_{\max}(\mathbf{R}_u)}{\lambda_{\min}(\mathbf{R}_u)} \le \frac{M\left(\left\lfloor\frac{N-MP}{d}\right\rfloor+1\right)\sigma_J^2+\sigma_n^2\lambda_{\max}(\mathbf{B}^T\mathbf{B})}{\sigma_n^2\lambda_{\min}(\mathbf{B}^T\mathbf{B})}. \quad (42)$$

The above derivations can also be extended to multiple jammers case straightforwardly (assume that $(\lfloor\frac{N-MP}{d}\rfloor+1)(M-1) > J > 1$). Similar to (35), the maximum eigenvalue of $\mathbf{B}^T\mathbf{R}_J\mathbf{B}$, λ_J, is now bounded by $\lambda_J \le M(\lfloor\frac{N-MP}{d}\rfloor+1)\sum_{i=1}^{J}\sigma_{Ji}^2$. Following the same approach as the one jammer case, we can get the following upper bound for the eigenvalue spread of $\mathbf{R}_u$:

$$\frac{\lambda_{\max}(\mathbf{R}_u)}{\lambda_{\min}(\mathbf{R}_u)} \le \frac{M\left(\left\lfloor\frac{N-MP}{d}\right\rfloor+1\right)\sum_{i=1}^{J}\sigma_{Ji}^2+\sigma_n^2\lambda_{\max}(\mathbf{B}^T\mathbf{B})}{\sigma_n^2\lambda_{\min}(\mathbf{B}^T\mathbf{B})}.$$

ACKNOWLEDGMENT

The authors would like to thank the reviewers for many useful comments and suggestions, which have enhanced the quality and readability of this paper.

REFERENCES

[1] R. A. Monzingo and T. W. Miller, *Introduction to Adaptive Arrays*. New York: Wiley, 1980.
[2] D. J. Chapman, "Partial adaptivity for large array," *IEEE Trans. Antennas Propagat.,*, vol. AP-24, pp. 685–696, Sept. 1976.
[3] B. D. Van Veen and R. A. Roberts, "Partially adaptive beamformer design via output power minimization," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 1524–1532, Nov. 1987.
[4] R. T. Compton Jr., *Adaptive Antennas*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[5] S. S. Narayan, A. M. Peterson, and M. J. Narasimha, "Transform domain LMS algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 609–615, June 1983.
[6] J. C. Lee and C. K. Un, "Performance of transform-domain LMS adaptive digital filters," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 499–510, June 1986.
[7] I. Daubechies, *Ten Lectures on Wavelets*. Philadelphia, PA: SIAM, 1992.
[8] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Wellesley, MA: Wellesley-Cambridge Press, 1996.
[9] N. Erdol and F. Basbug, "Wavelet transform based adaptive filters: Analysis and new results," *IEEE Trans. Signal Processing*, vol. 44, pp. 2163–2171, Sept. 1996.
[10] M. I. Doroslovački and H. Fan, "Wavelet-based linear system modeling and adaptive filtering," *IEEE Trans. Signal Processing*, vol. 44, pp. 1156–1167, May 1996.
[11] S.-H. Chang, C.-C. Chang, C.-L. Chang, and W.-H. Fang, "Spatial processing technique adaptive beamforming (SPTABF) via compactly supported orthogonal wavelets," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, Atlanta, GA, 1996, pp. 3173–3176.
[12] L. J. Griffiths and C. W. Jim, "Alternative approach to linear constrained adaptive beamforming," *IEEE Trans. Antennas Propagat.*, vol. AP-30, pp. 27–34, Jan. 1982.
[13] P. Steffen, P. N. Heller, R. A. Gopinath, and C. S. Burrus, "Theory of regular M-band wavelet bases," *IEEE Trans. Signal Processing*, vol. 41, pp. 1766–1777, Dec. 1993.
[14] P. N. Heller, "rank M wavelets with N vanishing moments," *SIAM J. Matrix Anal. Appl.*, vol. 16, pp. 502–519, Apr. 1995.
[15] O. L. Frost III, "An algorithm for linearly constrained adaptive array processing," *Proc. IEEE*, vol. 60, pp. 926–935, Aug. 1972.
[16] M. H. Er and A. Cantoni, "Derivative constraints for broad-band element space antenna array processors," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1378–1393, Dec. 1983.
[17] K. C. Huarng and C. C. Yeh, "Performance analysis of derivative constraint adaptive arrays with pointing errors," *IEEE Trans. Antennas Propagat.*, vol. 40, pp. 975–981, Aug. 1992.
[18] K. M. Buckley and L. J. Griffiths, "An adaptive sidelobe canceller with derivative constraints," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 311–319, Mar. 1986.
[19] Y. Chu, W.-H. Fang, and S.-H. Chang, "A novel wavelet-based generalized sidelobe canceller," in *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Seattle, WA., 1998, pp. 2497–2500.
[20] C. W. Jim, "A comparison of two LMS constrained optimal array structures," *Proc. IEEE*, vol. 65, pp. 1730–1731, Dec. 1977.
[21] K. M. Buckley, "Broad-band beamforming and the generalized sidelobe canceller," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1322–1323, Oct. 1986.
[22] L. L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*. Reading, MA: Addison-Wesley, 1991.
[23] Y. Chu, "Transform domain statistical signal processing: A wavelet-based approach," Ph.D. dissertation, Nat. Taiwan Univ. Sci. Technol., Taipei, Taiwan, 1998.
[24] Y. Chu, W.-H. Fang, and S.-H. Chang, "Performance analysis of wavelet-based generalized sidelobe cancellers," in *Proc. Eur. Signal Processing Conf.*, Rhode Islands, Greece, 1998, pp. 1981–1984.
[25] N. K. Jablon, "Steady state analysis of the generalized sidelobe canceller by adaptive noise cancelling techniques," *IEEE Trans. Antennas Propagat.*, vol. AP-34, pp. 330–337, Mar. 1986.
[26] F. T. Leighton, *Introduction to Parallel Algorithms and Architectures: Arrays, Trees, and Hypercubes*. San Mateo, CA: Morgan Kaufmann, 1992.
[27] Y.-Y. Wang, W.-H. Fang, and J.-T. Chen, "Improved wavelet-based beamformers with dynamic subband selection," *IEEE AP-S Int. Symp.*, to be published.
[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.

Yi Chu (S'96–M'98) was born in Taipei, Taiwan, R.O.C., in 1966. He received the Dipl.E.E. (electronic engineering) from the National Taipei University of Technology, Taiwan, in 1989, and the M.S. (electronic engineering) and Ph.D. degrees from the National Taiwan University of Science and Technology, Taipei, in 1994 and 1998, respectively.

From 1990 to 1992, he was in the Army where he served as a Jiann-An Project in Hua-Lian, Taiwan. Since 1999 he has been an Assistant Professor of Electronic Engineering at Chien Hsin Institute of Technology, Chungli, Taiwan, R.O.C. His research interests include statistical signal processing, adaptive filtering, array signal processing, and wireless communications.

Wen-Hsien Fang (S'87–M'91) was born in Taipei, Taiwan, R.O.C., in 1961. He received the B.S. degree from the National Taiwan University, Taipei, in 1983, and the M.S.E. (electrical engineering) and Ph.D. (computer science) degrees from the University of Michigan, Ann Arbor, in 1988 and 1991, respectively.

From 1988 to 1991, he was a Research Assistant at the University of Michigan. In 1991 he joined the faculty of National Taiwan University of Science and Technology, where he currently holds the position Associate Professor in the Department of Electronic Engineering. His research interests include array signal processing, wireless communications, fast algorithms for signal processing, and their very large scale integration (VLSI) implementations.