A Signal Subspace-Based Subband Approach to Space-Time Adaptive Processing for Mobile Communications

Kehu Yang, *Member, IEEE*, Yimin Zhang, *Member, IEEE*, and Yoshihiko Mizuguchi

***Abstract*—In this paper, we present a novel space-time signal subspace-based subband approach to space-time adaptive processing (STAP) that has been shown to be an effective method to suppress both the intersymbol interference (ISI) and the cochannel interference (CCI) in mobile communications. We first study the performance of STAP and make clear the conditions of perfect processing (i.e., perfect equalization of the desired user signal and perfect suppression of CCI signals). Based on the polyphase representation and the subspace analysis of the signal channels, we propose a space-time signal subspace-based subband approach to STAP, namely, the subband STAP, which highly improves the convergence rate without loss of the steady-state performance. Simulation results show its effectiveness under the procedure of signal subspace estimation and detection.**

***Index Terms*—Mobile communications, multichannel modeling, space-time adaptive processing (STAP), space-time signal subspace, subspace decomposition, subband filtering.**

I. Introduction

It is known that in land mobile communications, the transmitting signals suffer from reflection and scattering by surroundings, and the receiving signals suffer from fading [1], [2] by multipath propagation. As the mobile communications are developing toward the higher speed digital networks [7], the associated communication channels become severely frequency selective, which makes the intersymbol interference (ISI) highly pronounced. Additionally, due to frequency reuse and multiuser access, cochannel user interference (CCI) signals or multiple user access interference (MUAI) signals are present against the desired user signal. Therefore, the system capacity and the communication quality are greatly affected by both the ISI and CCI (or MUAI) problems.

Adaptive arrays, particularly under space-time adaptive processing (STAP), provide effective ways to suppress both the ISI and the CCI, subsequently improving the system capacity and the communication quality [9]–[13]. A STAP system is usually composed of an antenna array and a set of FIR filters after the array elements to perform joint spatial and temporal adaptive processing. Despite its excellent steady-state performance, a STAP system faces the problems of a high computational burden and a low convergence rate. These problems become particularly serious when it operates in the severe fading environments where longer FIR filters are needed. For example, when LMS-type algorithms are used, the convergence rate will become extremely slow, and subsequently, a long training sequence will be required. Furthermore, when the batch processing-based algorithms [e.g., the sample matrix inversion (SMI) method [14]] are employed, the long training sequence required to estimate the correlation vector and the computation burden of the sample matrix inversion will leave real-time adaptation difficult or even impossible.

To ease these problems of the STAP, the authors have proposed the subband adaptive array scheme [24], [26], which, in essence, is an equivalent space-frequency domain approach to STAP. The subband signal processing converts a wideband signal processing problem into a set of parallel narrowband problems; hence, the equivalent time delay spread between multipath rays at each subband becomes much smaller. As a result, the user signals are approximately equalized, and the computational burden at each subband is greatly reduced.

By contrast, the applications of subband filtering in the temporal domain, such as acoustic echo canceling (AEC) [19]–[21], have demonstrated that decorrelating received signals by subband decomposition improves the convergence rate of LMS algorithm under weighted criteria.

In STAP cases, however, the signals received at different array elements are highly correlated in both space and time. To improve the convergence rate of LMS-type algorithms for the underlying STAP systems, the decorrelation of the data with the use of conventional time-domain filter banks is not efficient. Therefore, decorrelating the signals in the joint spatial and temporal domain simultaneously becomes an important issue.

In the last two decades, signal subspace-based processing has been proposed and applied to various fields [16]. By taking advantage of the orthogonality between the signal subspace and its complement, i.e., the noise subspace, the projection of the signals onto its associated signal subspace decorrelates the channel signals and simultaneously compresses the computational dimension by eliminating the unnecessary components that belong only to the noise subspace. In the underlying STAP systems, the signal subspace-based approaches enable us to decorrelate the received signals in both the space and time domains.

Manuscript received July 7, 1999; revised October 24, 2000. The associate editor coordinating the review of this paper and approving it for publication was Dr. Vikram Krishnamurthy.

K. Yang is with ATR Adaptive Communications Research Laboratories, Kyoto, Japan (e-mail: yang@acr.atr.co.jp).

Y. Zhang is with Department of Electrical and Computer Engineering, Villanova University, Villanova, PA 19085 USA (e-mail: zhang@ieee.org).

Y. Mizuguchi was with ATR Adaptive Communications Research Laboratories, Kyoto, Japan. He is now with KDD Research Laboratories Inc., Saitama, Japan (e-mail: mizu@kddlabs.co.jp).

Publisher Item Identifier S 1053-587X(01)00618-3.

Such subspace projection can be realized by employing subband filtering, yielding the space-time signal subspace-based subband STAP (SSTAP).

The contributions of this paper are twofold. First, some additional new results of STAP are presented, in which the conditions of perfect processing (i.e., perfect equalization of the desired user signal and perfect suppression of CCI signals) are derived, and the effect of the delay of the reference signal on the output performance is investigated. Second, a space-time signal subspace-based subband approach to STAP is proposed, where the polyphase representation is employed, and the convergence rate is highly improved without loss of the steady-state performance.

This paper is organized as follows. In Section II, after introducing the signal model, we derive the conditions of perfect processing and the residual error power of STAP. Section III establishes the relationship between the space-time signal subspace and STAP, as well as the detection of signals' components. In Section IV, a space-time signal subspace-based subband approach to STAP is proposed, and the convergence improvement of LMS algorithm under the proposed approach is confirmed by theoretical analysis. Several simulation results are presented in Section V. Section VI concludes the paper.

II. Performance Analysis of STAP

A. Signal Model

Consider a base station using an antenna array of N ($N \geq 1$) elements with P ($P \geq 1$) users. The signal of the desired user is denoted as $s_1(t)$, whereas signals from other users are denoted as $s_p(t), p = 2, \ldots, P$. The array output vector $\mathbf{x}(t)$ is expressed as

$$\begin{aligned}\mathbf{x}(t) &= \sum_{p=1}^{P}\sum_{l=1}^{L_p}\mathbf{a}\left(\theta_l^p\right)\xi_l^p s_p\left(t-\tau_l^p\right)+\mathbf{n}(t)\\ &= \sum_{p=1}^{P}\sum_{m=-\infty}^{+\infty}\bar{s}_p(m)\mathbf{h}_p(t-mT)+\mathbf{n}(t) \qquad (1)\end{aligned}$$

where

$$s_p(t) = \sum_{m=-\infty}^{+\infty}\bar{s}_p(m)\rho_p(t-mT) \qquad (2)$$

$$\mathbf{h}_p(t) = \sum_{l=1}^{L_p}\mathbf{a}\left(\theta_l^p\right)\xi_l^p\rho\left(t-\tau_l^p\right). \qquad (3)$$

The following notations are used in (1)–(3):

$\{\theta_l^p, \tau_l^p, \xi_l^p\}$	Angle-of-arrival (AOA), time delay, and propagation loss corresponding to the lth path of the pth user.
$\mathbf{a}(\theta)$	Array steering vector corresponding to θ.
$\bar{s}_p(m)$	mth information symbol of the pth user.
$\rho_p(t)$	Pulse shaping function of the pth user.
L_p	Total number of multipath rays of the pth user.
T	Symbol duration.
$\mathbf{n}(t)$	Array noise vector.

We make the following assumptions.

A1) The user signals are wide-sense cyclostationary when they are sampled at fractionally spaced symbol cycle, and are wide-sense stationary when they are sampled at the symbol rate. A wide-sense cyclostationary signal vector is defined by $E[\mathbf{x}(t_1)\mathbf{x}^H(t_2)] = E[\mathbf{x}(t_1+T)\mathbf{x}^H(t_2+T)]$ [5], [15], where $(\cdot)^H$ denotes conjugate transpose, and $E[\cdot]$ denotes statistical expectation.

A2) The information symbols $\bar{s}_p(m),\ p = 1, \ldots, P$ are independent and identically distributed (i.i.d.) with $E\{\bar{s}_p(m)\bar{s}_p^*(m)\} = 1$ and $E\{\bar{s}_p(m)\bar{s}_p^*(m-k)\} = 0$ for $\forall k \neq 0$ and uncorrelated with the channel noise vector, where $(\cdot)^*$ denotes complex conjugation.

A3) All channels $\{\mathbf{h}_p(t), p = 1, \ldots, P\}$ are linearly time-invariant, and each of them is of a finite duration within $[0, D_pT]$, where D_p is called the channel order of the pth user.

A4) The noise vector is zero-mean and temporally and spatially white with

$$E\{\mathbf{n}(t)\mathbf{n}^T(t)\} = 0, \quad E\{\mathbf{n}(t)\mathbf{n}^H(t)\} = \sigma^2\mathbf{I}$$

where $(\cdot)^T$ denotes transpose.

Denote Δ as the sampling cycle, and let $J = T/\Delta$ ($J \geq 1$) be the factor of oversampling. Sampling $\mathbf{x}(t)$ at $t = i\Delta + nT$, $n \in (-\infty, +\infty)$, (1) becomes

$$\begin{aligned}&\mathbf{x}(i\Delta+nT)\\ &\quad= \sum_{p=1}^{P}\sum_{d=0}^{D_p}\bar{s}_p(n-d)\mathbf{h}_p(i\Delta+dT)+\mathbf{n}(i\Delta+nT)\\ &\qquad\qquad i = 0, \ldots, J-1. \qquad (4)\end{aligned}$$

With the exploitation of the cyclostationarity of user signals [15], [16], [18] described in Assumption A1, the scheme of the extended multichannel model of fractionally spaced STAP, as illustrated in Fig. 1, can be easily established as

$$\underline{\mathbf{x}}(n) = \sum_{p=1}^{P}\sum_{d=0}^{D_p}\bar{s}_p(n-d)\underline{\mathbf{h}}_p(d)+\underline{\mathbf{n}}(n) \qquad (5)$$

where

$$\underline{\mathbf{x}}(n) = \begin{bmatrix}\mathbf{x}[nT]\\ \mathbf{x}[nT-\Delta]\\ \vdots\\ \mathbf{x}[nT-(J-1)\Delta]\end{bmatrix}$$

$$\underline{\mathbf{h}}(d) = \begin{bmatrix}\mathbf{h}[dT]\\ \mathbf{h}[dT-\Delta]\\ \vdots\\ \mathbf{h}[dT-(J-1)\Delta]\end{bmatrix}$$

$$\underline{\mathbf{n}}(n) = \begin{bmatrix}\mathbf{n}[nT]\\ \mathbf{n}[nT-\Delta]\\ \vdots\\ \mathbf{n}[nT-(J-1)\Delta]\end{bmatrix}.$$

For each n, the dimension of $\underline{\mathbf{x}}(n)$ is $\underline{N} = NJ$, which is called the *number of the extended channels*. The limit of the number of the extended channels by oversampling is discussed in [8].

For the consecutive samples during the period of M symbols, we form the following vectors:

$$X(n) = \begin{bmatrix} \underline{\mathbf{x}}(n) \\ \underline{\mathbf{x}}(n-1) \\ \vdots \\ \underline{\mathbf{x}}(n-M+1) \end{bmatrix}$$

$$S_p(n) = \begin{bmatrix} \bar{s}_p(n) \\ \bar{s}_p(n-1) \\ \vdots \\ \bar{s}_p(n-M-D_p+1) \end{bmatrix}$$

$$N(n) = \begin{bmatrix} \underline{\mathbf{n}}(n) \\ \underline{\mathbf{n}}(n-1) \\ \vdots \\ \underline{\mathbf{n}}(n-M+1) \end{bmatrix}.$$

Defining the following Sylvester convolution matrix of user p in terms of the $(D_p+1)NJ$-length impulse response of its channel $[\underline{\mathbf{h}}_p^T(0), \underline{\mathbf{h}}_p^T(1), \ldots, \underline{\mathbf{h}}_p^T(D_p)]^T$, as shown in (6) at the bottom of the page. Then, (5) can be extended to

$$X(n) = \sum_{p=1}^{P} \mathbf{H}_p^{(M)} S_p(n) + N(n). \tag{7}$$

Furthermore, let $w_{i,1}, w_{i,2}, \ldots, w_{i,M}$ represent the M weights at the ith extended channel, and denote

$$\underline{\mathbf{w}}_m \triangleq [w_{11,m}, \ldots, w_{NJ,m}]^T, \quad m = 1, 2, \ldots, M \tag{8}$$

$$W \triangleq \left[\underline{\mathbf{w}}_1^T, \ldots, \underline{\mathbf{w}}_M^T\right]^T. \tag{9}$$

Then, the output of the STAP is given by

$$y(n) = \sum_{m=1}^{M} \underline{\mathbf{w}}_m^T \underline{\mathbf{x}}(n-m+1) = W^T X(n). \tag{10}$$

B. Residual Error Power

Under the minimum mean square error (MMSE) criterion, the optimum weights of the STAP are solved from

$$\min_W E|\bar{s}_1(n-v) - y(n)|^2 \tag{11}$$

and are given by the well-known Wiener–Hopf equation

$$W_{\text{MMSE}}^* = \mathbf{R}_X^{-1}\mathbf{r}(v) \tag{12}$$

where $\bar{s}_1(n)$ is the training sequence of the desired user signal, $v \geq 0$ is a delay of the training signal required for the realization of causal filtering

$$\mathbf{R}_X = E[X(n)X^H(n)] \tag{13}$$

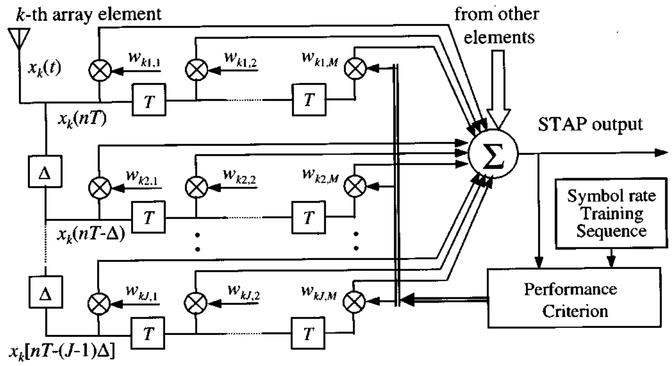

Fig. 1. Scheme of the fractionally spaced STAP.

is the space-time correlation matrix of the signal vector, and

$$\mathbf{r}(v) = E[\bar{s}_1^*(n-v)X(n)] \tag{14}$$

is the correlation vector between the training signal and the signal vector. From (6), (7), (14), and A2, it is evident that $\mathbf{r}(v)$ is the $(v+1)$th column of the signature matrix $\mathbf{H}_1^{(M)}$ when $0 \leq v \leq D_1 + M - 1$.

From (10)–(14), the residual error power of STAP under the MMSE criterion is obtained as

$$\begin{aligned} \sigma_{\text{MMSE}}^2(v) &= E|\bar{s}_1(n-v) - y(n)|^2 \\ &= E|\bar{s}_1(n-v)|^2 - W_{\text{MMSE}}^T \mathrm{R}_X W_{\text{MMSE}}^* \\ &= E|\bar{s}_1(n)|^2 - \mathbf{r}^H(v)\mathbf{R}_X^{-1}\mathbf{r}(v) \end{aligned} \tag{15}$$

where the Hermitian property of $\mathbf{R}_X$ is used.

The residual error power $\sigma_{\text{MMSE}}^2(v)$ is affected by several parameters, such as v and M, etc. In Theorem 1, we prove the relationship between $\sigma_{\text{MMSE}}^2(v)$ and v, and in Theorem 2, we investigate the selection of M and other conditions for perfect processing of a STAP system.

Theorem 1: For a given v, the residual error power is given as

$$\sigma_{\text{MMSE}}^2(v) = E|s_1(n)|^2 - \mathbf{r}_0^H(v)\mathbf{R}_A\mathbf{r}_0(v) \tag{16}$$

where

$$\mathbf{P}_v\mathbf{R}_X^{-1}\mathbf{P}_v^T = \begin{bmatrix} \mathbf{R}_A & \mathbf{C} \\ \mathbf{C} & \mathbf{B} \end{bmatrix} \tag{17}$$

$$\mathbf{P}_v\mathbf{r}(v) = \begin{bmatrix} \mathbf{r}_0(v) \\ \mathbf{0} \end{bmatrix} \tag{18}$$

where $\mathbf{r}_0(v)$ is a nonzero element vector composed of $\underline{\mathbf{h}}_1(m)$, $m = 0, \ldots, D_1$, $\mathbf{R}_A$ is a square matrix with the same dimension of $\mathbf{r}_0(v)$, and $\mathbf{P}_v$ is a permutation matrix [6] that depends on the zero-element structure of $\mathbf{r}(v)$, which is denoted as (19), shown

$$\mathbf{H}_p^{(M)} = \begin{bmatrix} \underline{\mathbf{h}}_p(0) & \cdots & \underline{\mathbf{h}}_p(D_p) & \mathbf{0} & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{h}}_p(0) & \cdots & \underline{\mathbf{h}}_p(D_p) & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \ddots & \cdots & \ddots & \cdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \cdots & \mathbf{0} & \underline{\mathbf{h}}_p(0) & \cdots & \underline{\mathbf{h}}_p(D_p) \end{bmatrix}_{MNJ\times(M+D_p)} \tag{6}$$

at the bottom of the page, where $\mathbf{I}$ and $\mathbf{I}_m$ denote the identity matrices of dimension MNJ and m, respectively.

Proof: Theorem 1 can be proved directly from (15), that is

$$\begin{aligned}\sigma^2_{\mathrm{MMSE}}(v) &= E|s_1(n)|^2 - \mathbf{r}^H(v)\mathbf{R}_X^{-1}\mathbf{r}(v)\\ &= E|s_1(n)|^2 - \mathbf{r}^H(v)\mathbf{P}_v^T\mathbf{P}_v\mathbf{R}_X^{-1}\mathbf{P}_v^T\mathbf{P}_v\mathbf{r}(v)\\ &= E|s_1(n)|^2 - (\mathbf{P}_v\mathbf{r}(v))^H\left(\mathbf{P}_v\mathbf{R}_X^{-1}\mathbf{P}_v^T\right)(\mathbf{P}_v\mathbf{r}(v))\\ &= E|s_1(n)|^2 - \mathbf{r}_0^H(v)\mathbf{R}_A\mathbf{r}_0(v) \qquad (20)\end{aligned}$$

■

It is clear from Theorem 1 that only a part of $\mathbf{R}_X$, i.e., $\mathbf{R}_A$, contributes to the residual error power of STAP. The dimension of $\mathbf{R}_A$ depends on the length of nonzero elements of r(v). This implies that v plays an important role in the output of the residual error power and, therefore, should be properly determined based on the channel characteristics. For example, in the cases of $v = 0$ or $v = D_1 + M - 1$, only one weight is active at each extended channel, and the total degrees-of-freedom (DOF's) of the STAP system is as little as $NJ - 1$. Therefore, in those two cases, the output performances will degrade.

From (6), we understand that in the case of $v = D_p$, the vector $\mathbf{r}(v)$ has longest nonzero length. That is to say, when the channel length is estimated as M_t, we can choose v as

$$v = M_t - 1. \qquad (21)$$

The estimation of channel length M_t is described in the next section.

C. Conditions of Perfect Processing

Theorem 2: In the noise-free case, the perfect processing, i.e., perfect equalization of the desired user and perfect suppression of undesired user signals, can be realized by STAP provided the following.

1) $\mathbf{H}_1^{(M)}$ is full column rank.
2) The columns of $\mathbf{H}_1^{(M)}$ are linearly independent to the other columns of $\mathbf{H}^{(M)}$, where

$$\mathbf{H}^{(M)} = \left[\mathbf{H}_1^{(M)}, \ldots, \mathbf{H}_P^{(M)}\right]. \qquad (22)$$

3)

$$MNJ \geq \text{column rank}\left\{\mathbf{H}^{(M)}\right\}. \qquad (23)$$

Proof: The first condition is common for single user cases and the proof is given in [16] and [17]. The second condition can be proved by counterevidence. Denoting

$$S(n) = \left[S_1^T(n), \ldots, S_P^T(n)\right]^T \qquad (24)$$

in the noise-free case, (7) can be written as

$$X(n) = \mathbf{H}^{(M)}S(n) \qquad (25)$$

and the output of STAP is given by

$$y(n) = W^TX(n) = W^T\mathbf{H}^{(M)}S(n). \qquad (26)$$

Perfect processing implies that

$$y(n) = \bar{s}_1(n-v), \quad 0 \leq v \leq M + D_1 - 1 \qquad (27)$$

or equivalently

$$W^T\mathbf{H}^{(M)} = \left[\mathbf{0}_{1\times v}^T, \overset{\bar{s}_1(n-v)}{\overset{\downarrow}{1}}, \mathbf{0}_{1\times M_e}^T\right] \qquad (28)$$

where $M_e = \sum_{p=1}^P (D_p + M) - v - 1$, $0 \leq v \leq D_1 + M - 1$.

If the columns of $\mathbf{H}_1^{(M)}$ are not linearly independent to the other columns of $\mathbf{H}^{(M)}$, for example, the v_0th column of $\mathbf{H}_1^{(M)}$ can be expressed as a linear combination of the other columns of $\mathbf{H}^{(M)}$, then (28) cannot stand for $v = v_0 - 1$. That is to say, when the columns of $\mathbf{H}_1^{(M)}$ are not linearly independent to the other columns of $\mathbf{H}^{(M)}$, the conditions of the perfect processing cannot be always satisfied for $0 \leq v \leq D_1 + M - 1$.

To prove the third condition, we compare the number of the weights with the number of equations. It is seen from (28) that there are $M_T = \sum_{p=1}^P (D_p + M)$ equations, whereas the number of the independent equations is specified by column rank$\{\mathbf{H}^{(M)}\}$. The condition for (28) to have a unique solution is that the number of the adjustable weights must be greater than or equal to the number of the independent equations, which goes to (23). ■

It is noteworthy that the presence of common roots in the channels of the undesired users does not bring to STAP any difficulties in performing perfect processing. Common roots in the channels of the undesired signals imply that the number of DOFs required to suppress these undesired signals is reduced. That is to say, common roots in the channels of the undesired users do ease the STAP processing.

$$\mathbf{P}_v^T = \begin{cases}\mathbf{I}, & v \in [0, \ldots, D_1]\\ \begin{bmatrix}\mathbf{0} & \mathbf{I}_{NJ(v-D_1)} & \mathbf{0}\\ \mathbf{I}_{NJ(D_1+1)} & \mathbf{0} & \mathbf{0}\\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{NJ(M-1-v)}\end{bmatrix}, & v \in [D_1 + 1, \ldots, M - 1]\\ \begin{bmatrix}\mathbf{0} & \mathbf{I}_{NJ(v-D_1)}\\ \mathbf{I}_{NJ(M+D_1-v)} & \mathbf{0}\end{bmatrix}, & v \in [M - 1, \ldots, M + D_1 - 1]\end{cases} \qquad (19)$$

Corollary 1: If $\mathbf{H}^{(M)}$ is full column rank, then $\mathrm{rank}\{\mathbf{H}^{(M)}\} = M_T$, and the condition of perfect processing becomes

$$M \geq \frac{1}{NJ-P}\sum_{i=1}^{P} D_i. \tag{29}$$

III. Subspace Analysis of STAP

In the previous section, the conditions of perfect processing of a STAP system are derived, which specify the lower bound of the number of taps in noise-free cases. Since such a bound depends on the length of the channels, it is thus necessary to estimate the channel length. In this section, we consider the estimation from the point of view of the space-time signal subspace of the associated space-time correlation matrix and establish the relationship between the space-time signal subspace and STAP.

A. Subspace Decomposition

Under Assumptions A2–A4, the space-time correlation matrix defined by (13) can be rewritten as

$$\begin{aligned}\mathbf{R}_X &= E[X(n)X^H(n)] \\ &= \sum_{p=1}^{P} \mathbf{H}_p^{(M)}\mathbf{R}_p\mathbf{H}_p^{(M)H} + \mathbf{R}_N \\ &= \mathbf{H}^{(M)}\mathbf{R}_S\mathbf{H}^{(M)H} + \mathbf{R}_N \end{aligned}\tag{30}$$

where

$$\mathbf{R}_p = E\left[S_p(n)S_p^H(n)\right] = \mathbf{I}_{M+D_p} \tag{31}$$

$$\mathbf{R}_S = \mathrm{Diag}\{\mathbf{R}_p, p=1,\ldots,P\} = \mathbf{I}_{M_T} \tag{32}$$

and

$$\mathbf{R}_N = E[N(n)N^H(n)] = \sigma^2\mathbf{I}. \tag{33}$$

Equation (30) shows that the column space of $\mathbf{H}^{(M)}$ is the space-time signal subspace. By using eigendecomposition, the space-time correlation matrix $\mathbf{R}_X$ can be expressed as

$$\mathbf{R}_X = \mathbf{S}\Sigma_s\mathbf{S}^H + \mathbf{G}\Sigma_n\mathbf{G}^H \tag{34}$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{M_D} \geq \lambda_{M_D+1} = \cdots = \lambda_{MNJ} = \sigma^2$ are the eigenvalues of $\mathbf{R}_X$, the columns of $\mathbf{S} = [\mathbf{S}_1,\ldots,\mathbf{S}_{M_D}]$ are the orthonormal eigenvectors associated with eigenvalues $\lambda_1,\lambda_2,\ldots,\lambda_{M_D}$, the columns of $\mathbf{G} = [\mathbf{G}_1,\ldots,\mathbf{G}_{MNJ-M_D}]$ are the orthonormal eigenvectors associated with eigenvalues $\lambda_{M_D+1},\ldots,\lambda_{MNJ}$, $\Sigma_s = \mathrm{diag}\{\lambda_i, i=1,\ldots,M_D\}$, and $\Sigma_n = \mathrm{diag}\{\lambda_i, i=1+M_D,\ldots,MNJ\}$. It is seen that M_D is the rank of $\mathbf{R}_X$ in the absence of noise. In the cases where $\mathbf{H}^{(M)}$ is of full column rank, then $M_D = M_T = \sum_{p=1}^{P}(M+D_p)$, the columns of $\mathbf{S}$ span the space-time signal subspace, and the columns of $\mathbf{G}$ span the space-time noise subspace. We assume in the sequel that $\mathbf{H}^{(M)}$ is of full column rank and that the space-time signal subspace can be estimated by the eigendecomposition of the estimated space-time correlation matrix. In the following, the space-time signal subspace is abbreviated as the signal subspace, and the space-time noise subspace is abbreviated as the noise subspace.

B. Space-Time Signal Subspace and STAP

As we discussed in Section II, $\mathbf{r}(v)$ is the $(v+1)$th column of $\mathbf{H}_1^{(M)}$, which means that $\mathbf{r}(v)$ belongs to the signal subspace. Using the relationship

$$\mathbf{R}_X^{-1} = \mathbf{S}\Sigma_s^{-1}\mathbf{S}^H + \mathbf{G}\Sigma_n^{-1}\mathbf{G}^H \tag{35}$$

we have

$$W_{\mathrm{MMSE}}^* = \mathbf{R}_X^{-1}\mathbf{r}(v) = \mathbf{S}\Sigma_s^{-1}\mathbf{S}^H\mathbf{r}(v) \tag{36}$$

which shows that the optimum weight vector is a linear combination of only the eigenvectors that span the signal subspace. In other words, the optimum weight vector belongs to the signal subspace.

We know that the residual error power of STAP is defined by (15) as

$$\sigma_{\mathrm{MMSE}}^2(v) = E|\bar{s}_1(n)|^2 - W_{\mathrm{MMSE}}^T\mathbf{R}_X W_{\mathrm{MMSE}}^*. \tag{37}$$

Substituting (35) and (36) into (37), we have

$$\sigma_{\mathrm{MMSE}}^2(v) = E|\bar{s}_1(n)|^2 - \mathbf{r}^H(v)\mathbf{S}\Sigma_s^{-1}\mathbf{S}^H\mathbf{r}(v). \tag{38}$$

The above equation clearly shows that only the components belonging to the signal subspace contribute to the residual error power, which implies that using the projection of the received signal vector onto the signal subspace instead of the received signal vector itself does not degrade the output performance. This observation result constitutes the basis of the space-time signal subspace-based approach in the paper, which enables us to reduce the complexity of STAP without reducing the steady-state performance.

C. Detection of Signals' Components in Practice

In this subsection, we investigate the channel length by detecting the dimension of the signal subspace. The problem of the signal detection based on certain theoretic information criteria has been well explored in array signal processing. The Akaike information criterion (AIC) is one of the commonly used criteria that offers the estimated number of signals with the use of all the eigenvalues [25]. Denoting k as an estimate of the number of signals estimated from the correlation matrix $\mathbf{R}_X$, the AIC is given by

$$\begin{aligned}\mathrm{AIC}(k) = &-2(MNJ-k)N_t\log\left(\frac{\prod_{i=k+1}^{MNJ}\lambda_i^{\frac{1}{MNJ-k}}}{\frac{1}{MNJ-k}\sum_{i=k+1}^{MNJ}\lambda_i}\right) \\ &+ 2k(2MNJ-k)\end{aligned}\tag{39}$$

where N_t is the number of the samples used to estimate the space-time correlation matrix $\mathbf{R}_X$, i.e.,

$$\hat{\mathbf{R}}_X = \frac{1}{N_t}\sum_{n=1}^{N_t} X(n)X^H(n). \tag{40}$$

The estimation of the dimension of the signal subspace $\hat{M}_D$ is determined as the value of $k \in \{0,1,\ldots,MNJ-1\}$ that minimizes the AIC. Based on the estimated $\hat{M}_D$, we obtain the lower bound for the required number of the taps of STAP as $M_t = \lfloor \hat{M}_D/NJ \rfloor$, where $\lfloor x \rfloor$ denotes the smallest integer not less than x.

It is noted that the estimation of the space-time correlation matrix $\mathbf{R}_X$ and its subspaces does not require any *a priori* information of the signals and can be performed during the period other than that of the training sequence being present. Moreover, the subspace decomposition and detection can be performed by the method of fast subspace decomposition [23] to improve the computation speed.

IV. Signal Subspace-Based Subband Approach

A. Polyphase Representation and Subband Approach

The z transform of the weight vectors of the STAP filter $\underline{\mathbf{w}}_l$, $l = 1, \ldots, M$ is expressed as

$$\mathbf{W}(z) = \sum_{l=0}^{M-1} z^{-l} \underline{\mathbf{w}}_{l+1}. \tag{41}$$

As an implementation method of STAP filtering, by using the generalized polyphase representation [4], [22], $\mathbf{W}(z)$ can be expressed by the following polyphase representation:

$$\mathbf{W}(z) = \sum_{k=1}^{\underline{M}} z^{-(k-1)} \mathbf{W}_k(z^L) \tag{42}$$

with

$$\mathbf{W}_k(z) = \sum_{n=0}^{K-1} z^{-n} \underline{\mathbf{w}}'_{nL+k} \tag{43}$$

as its polyphase components, where $\underline{\mathbf{w}}'_{nL+k}$ can be determined by comparing (43) with (41). In (42) and (43)

- $L \geq 1$ sparsity factor;
- K number of the coefficients in each sparse subband filter (i.e., the order of $\mathbf{W}_k(z^L)$ is $K-1$);
- $\underline{M} \geq L$ transformation size, i.e., the length of the associated filterbank defined in the sequel.

These parameters link with M by $M = (K-1)L + \underline{M}$. Therefore, by increasing the sparsity factor L and the transform size $\underline{M}$, we can reduce K, which is the order of the FIR filter at each subband.

The STAP output can, therefore, be written as

$$Y(z) = \mathbf{W}^T(z)\underline{\mathbf{x}}(z) = \sum_{k=1}^{\underline{M}} z^{-(k-1)} \mathbf{W}_k^T(z^L)\underline{\mathbf{x}}(z) \tag{44}$$

where $\underline{\mathbf{x}}(z)$ and $Y(z)$ express the z-transform of $\underline{\mathbf{x}}(n)$ and $y(n)$, respectively.

In order to establish the relationship between STAP and its subband approaches, we introduce the following full column rank transformation matrix $\mathbf{T}$:

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \cdots & \mathbf{T}_{1\underline{K}} \\ \vdots & \ddots & \vdots \\ \mathbf{T}_{\underline{M}1} & \cdots & \mathbf{T}_{\underline{M}\underline{K}} \end{bmatrix} \tag{45}$$

where $\underline{K} \leq \underline{M}$, and $\mathbf{T}_{ij}$, $i = 1, \ldots, \underline{M}$, $j = 1, \ldots, \underline{K}$ are the submatrices of $\mathbf{T}$. Each submatrix is of dimension $NJ \times NJ$. We define a new set of weights under the transformed scheme $\mathbf{F}_i(z)$, $i = 1, \ldots, \underline{K}$ by

$$\begin{bmatrix} \mathbf{W}_1(z^L) \\ \vdots \\ \mathbf{W}_{\underline{M}}(z^L) \end{bmatrix} = \mathbf{T}^* \times \begin{bmatrix} \mathbf{F}_1(z^L) \\ \vdots \\ \mathbf{F}_{\underline{K}}(z^L) \end{bmatrix}. \tag{46}$$

In the above transform, when $\underline{K} = \underline{M}$, the transform is full rank, and the STAP filtering after the transform of (46) keeps the same performance. On the other hand, if $\underline{K} < \underline{M}$, the transform is rank reduced, and the performance of the STAP filter after the transform may be inferior to that of the original STAP.

However, the reduced-rank transform can be performed without performance loss by using the signal subspace matrix as the transformation matrix. As shown by the results of Section III, the optimum weight vector of STAP belongs to the signal subspace, and only the components belonging to the signal subspace contribute to the output signal power of STAP. Therefore, using the signal subspace matrix as the reduced-rank transformation matrix loses no signal components. To summarize, such a transform based on the signal subspace has threefold advantages.

1) A part of the computations is reduced because of the dimension reduction.
2) There is no performance loss.
3) The convergence rate is improved when LMS-type algorithms are used.

These properties are very important to the practical implementations of STAP systems.

Substituting (46) into (44), we have

$$\begin{aligned} Y(z) &= \sum_{j=1}^{\underline{K}} \sum_{i=1}^{\underline{M}} z^{-(i-1)} \mathbf{F}_j^T(z^L)\mathbf{T}_{ij}^H \underline{\mathbf{x}}(z) \\ &= \sum_{j=1}^{\underline{K}} \mathbf{F}_j^T(z^L) \left(\sum_{i=1}^{\underline{M}} z^{-(i-1)} \mathbf{T}_{ij}^H \right) \underline{\mathbf{x}}(z) \\ &= \sum_{j=1}^{\underline{K}} \mathbf{F}_j^T(z^L) \mathbf{G}_j^T(z) \underline{\mathbf{x}}(z) \end{aligned} \tag{47}$$

where

$$\mathbf{G}_j(z) = \sum_{i=1}^{\underline{M}} z^{-(i-1)} \mathbf{T}_{ij}^*, \quad j = 1, \ldots, \underline{K} \tag{48}$$

are a set of filters and are termed as the generalized space-time filterbank (GSTF), which is an extension of the conventional subband filterbank. Equation (47) shows that space-time adaptive filtering can be equivalently realized via subband processing by using space-time filterbanks. The different implementation schemes of (41), (42), (46), and (47) are plotted in Fig. 2(a)–(d), respectively.

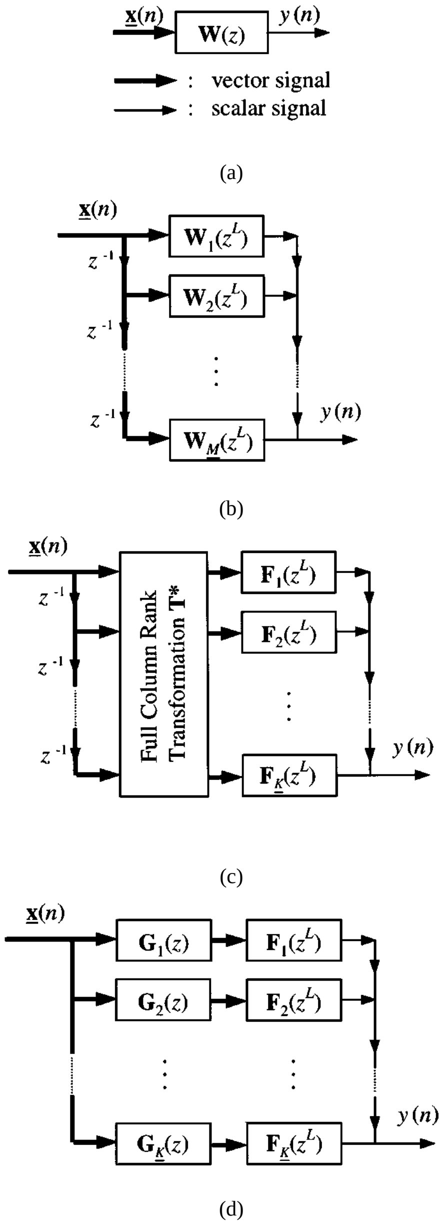

Fig. 2. Subband realization of STAP: (a) Space-time filter. (b) Polyphase implementation of STAP. (c) Full colomn rank transformation. (d) Subband STAP.

B. Signal Subspace-Based Subband Approach

Using the signal subspace matrix $\mathbf{S}$ as the transformation matrix $\mathbf{T}$ yields the signal subspace-based subband approach. To derive the expression of the GSTF $\mathbf{G}_j(z)$, we partition $\mathbf{S}$ into submatrices as follows:

$$\mathbf{S} = [\mathbf{S}_1, \mathbf{S}_2, \ldots, \mathbf{S}_{M_t}] = \begin{bmatrix} \mathbf{S}_{11} & \cdots & \mathbf{S}_{1M_t} \\ \vdots & \ddots & \vdots \\ \mathbf{S}_{M1} & \cdots & \mathbf{S}_{MM_t} \end{bmatrix} \tag{49}$$

where $\mathbf{S}_{ij}$, $i = 1, \ldots, M$, and $j = 1, \ldots, M_t$ are the submatrices of $\mathbf{S}$, where each one is of dimension $NJ \times NJ$. In this analysis, we use two taps at each subband. Similar to (47), the z-domain SSTAP output of signal subspace-based subband approach is given by

$$\begin{aligned} Y(z) &= \sum_{j=1}^{M_t} \sum_{i=1}^{M} z^{-(i-1)} \mathbf{F}_j^T(z^L) \mathbf{S}_{ij}^H \underline{\mathbf{x}}(z) \\ &= \sum_{j=1}^{M_t} \mathbf{F}_j^T(z^L) \left(\sum_{i=1}^{M} z^{-(i-1)} \mathbf{S}_{ij}^H \right) \underline{\mathbf{x}}(z). \end{aligned} \tag{50}$$

In comparison with (47), we have

$$\mathbf{G}_j(z) = \sum_{i=1}^{M} z^{-(i-1)} \mathbf{S}_{ij}^* \tag{51}$$

$$\mathbf{F}_j(z) = \sum_{l=0}^{1} z^{-l} \underline{\mathbf{f}}_{lL+j-1} \tag{52}$$

and the equivalent number of the taps of each extended STAP channel is $(2-1) \times L + M_t$.

Equation (50) can be cast into

$$Y(z) = \mathbf{F}^T(z^L) \bar{\mathbf{x}}(z) \tag{53}$$

where

$$\mathbf{F}(z^L) = \begin{bmatrix} \mathbf{F}_1(z^L) \\ \vdots \\ \mathbf{F}_{M_t}(z^L) \end{bmatrix}, \quad \bar{\mathbf{x}}(z) = \begin{bmatrix} \underline{\mathbf{x}}^{(1)}(z) \\ \vdots \\ \underline{x}^{(M_t)}(z) \end{bmatrix}$$

and

$$\underline{\mathbf{x}}^{(j)}(z) = \mathbf{G}_j^T(z) \underline{\mathbf{x}}(z).$$

Here, $\bar{\mathbf{x}}(z)$ is the input signal vector in the subband realization, whereas $\mathbf{F}(z)$ is the weight vector in the subband approach. In the manner of temporal expression, (53) is expressed as

$$y(n) = \sum_{l=0}^{1} \mathbf{f}_l^T(n) \bar{\mathbf{x}}(n - lL) \tag{54}$$

in which $\bar{\mathbf{x}}(n)$ and $\mathbf{f}_l$ are the inverse z-transform of $\bar{\mathbf{x}}(z)$ and $\mathbf{F}(z)$, respectively. From (49)–(54) and the definition of $X(n)$ in (7), we have the following relation as

$$\bar{\mathbf{x}}(n) = \mathbf{S}^H X(n). \tag{55}$$

C. Convergence Rate

In this subsection, the improvement of the convergence rate under (53) over the conventional STAP scheme is investigated. Using the LMS algorithm, the weights of STAP in (10) are updated according to

$$\underline{\mathbf{w}}_m(n+1) = \underline{\mathbf{w}}_m(n) + \mu e^*(n) \underline{\mathbf{x}}(n - m + 1) \qquad m = 1, \ldots, M \tag{56}$$

where

$$e(n) = \bar{s}_1(n - v) - y(n) \tag{57}$$

is the error signal, $y(n)$ is the STAP output defined in (10), and μ is the common step size. The selection of μ will be discussed in Section V. In the subband STAP, because the decorrelation is performed simultaneously in space and time by the space-time subband filtering, the LMS algorithm can be employed with the different step size at each subband, i.e.,

$$\mathbf{f}_l(n+1) = \mathbf{f}_l(n) + \mu\Phi e^*(n)\bar{\mathbf{x}}(n-lL), \quad l=0,1 \tag{58}$$

where $e(n)$ has the same form as (57), whereas $y(n)$ is defined in (54), and the entries of the diagonal matrix Φ represent the different step sizes of the associated subbands. The reasonable choice of Φ is to select

$$\Phi = \Sigma_s^{-1} \tag{59}$$

which equalizes the signal power of different subbands and thus highly reduces the spread of the nonzero eigenvalues of the associated correlation matrix.

It is well known that the convergence rate of the LMS algorithm mainly depends on the eigenvalue spread of the associated correlation matrix. Let

$$\mathbf{f}_\alpha(n) = \left[\mathbf{f}_0^T(n), \mathbf{f}_1^T(n)\right]^T \tag{60}$$

and

$$\mathbf{x}_\alpha(n) = [\bar{\mathbf{x}}^T(n), \bar{\mathbf{x}}^T(n-L)]^T \tag{61}$$

respectively, denote the weight vector and the signal vector after subband decomposition; then, the SSTAP output becomes

$$y(n) = \mathbf{f}_\alpha^T(n)\mathbf{x}_\alpha(n) \tag{62}$$

and the weights are updated according to

$$\begin{aligned}\mathbf{f}_\alpha(n+1) &= \mathbf{f}_\alpha(n) + \mu\Lambda e^*(n)\mathbf{x}_\alpha(n)\\ &= \mathbf{f}_\alpha(n) - \mu\Lambda\mathbf{x}_\alpha(n)\mathbf{x}_\alpha^H(n)\mathbf{f}_\alpha^*(n)\\ &\quad + \mu\Lambda\bar{s}_1^*(n-v)\mathbf{x}_\alpha(n)\end{aligned} \tag{63}$$

where $\Lambda = \mathbf{I}_2 \otimes \Phi$, and $\otimes$ expresses Kronecker product. From the general convergence analysis of the LMS algorithm, the convergence rate of the LMS algorithm based on (63) mainly depends on the eigenvalue spread of the weighted correlation matrix

$$\mathbf{R}_\alpha = E\left[\Lambda\mathbf{x}_\alpha(n)\mathbf{x}_\alpha^H(n)\right]. \tag{64}$$

It is known that the zero eigenvalues of the associated correlation matrix do not influence the convergence rate of the LMS algorithm [22]. Therefore, the spread of the nonzero eigenvalues of $\mathbf{R}_\alpha$ is emphasized. In the following, we present two theorems to show the number of the nonzero eigenvalues of $\mathbf{R}_\alpha$ and their spread.

Theorem 3: The matrix $\mathbf{R}_\alpha$ has $\min\{M_T - PL, 2M_T - N_n\}$ zero eigenvalues, where $N_n = \min\{2M_T, NJ(M+L)\}$.

Proof: In terms of (55) and (61), $\mathbf{R}_\alpha$ is expressed as

$$\begin{aligned}\mathbf{R}_\alpha &= \Lambda E\left[\mathbf{x}_\alpha(n)\mathbf{x}_\alpha^H(n)\right]\\ &= \Lambda\begin{bmatrix}\mathbf{S}^H\mathbf{R}_X\mathbf{S} & \mathbf{S}^H\mathbf{R}_X(L)\mathbf{S}\\ \mathbf{S}^H\mathbf{R}_X^H(L)\mathbf{S} & \mathbf{S}^H\mathbf{R}_X\mathbf{S}\end{bmatrix}\\ &= \Lambda\begin{bmatrix}\mathbf{S}^H & \mathbf{0}\\ \mathbf{0} & \mathbf{S}^H\end{bmatrix}\begin{bmatrix}\mathbf{R}_X & \mathbf{R}_X(L)\\ \mathbf{R}_X^H(L) & \mathbf{R}_X\end{bmatrix}\begin{bmatrix}\mathbf{S} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}\end{bmatrix}\end{aligned} \tag{65}$$

where $\mathbf{R}_X(L) = E[X(n)X^H(n-L)]$ and $\mathbf{R}_X$ is defined in (30). Under Assumptions A2 and A4

$$\begin{aligned}\mathbf{R}_X(L) &= E[X(n)X^H(n-L)]\\ &= \sum_{p=1}^{P}\mathbf{H}_p^{(M)}\mathbf{R}_p(L)\mathbf{H}_p^{(M)H} + \mathbf{R}_N(L)\\ &= \mathbf{H}^{(M)}\mathbf{R}_S(L)\mathbf{H}^{(M)H} + \mathbf{R}_N(L)\end{aligned} \tag{66}$$

in which

$$\mathbf{R}_p(L) = E\left[S_p(n)S_p^H(n-L)\right] = \mathbf{J}_{M+D_p}(L) \tag{67}$$

$$\mathbf{R}_S(L) = \mathrm{Diag}\{\mathbf{R}_p(L), p=1,\ldots,P\} \tag{68}$$

and

$$\mathbf{R}_N(L) = E[N(n)N^H(n-L)] = \sigma^2\mathbf{J}_{NJM}(NJL) \tag{69}$$

where $\mathbf{J}_m(L)$ expresses an $m \times m$ matrix with its Lth down diagonal line of entries of 1s and all the other entries of 0s, i.e.,

$$\mathbf{J}_m(L) = \begin{bmatrix}0 & \cdots & 0 & \cdots & 0\\ \vdots & \cdots & \vdots & \cdots & \vdots\\ 1 & 0 & \vdots & \cdots & \vdots\\ 0 & \ddots & 0 & \cdots & \vdots\\ 0 & \cdots & 1 & 0 & 0\end{bmatrix}. \tag{70}$$

Therefore, substituting (66) and (30) into (65), we obtain

$$\begin{aligned}\mathbf{R}_\alpha = \Lambda&\begin{bmatrix}\mathbf{S}^H\mathbf{H}^{(M)} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}^H\mathbf{H}^{(M)}\end{bmatrix}\begin{bmatrix}\mathbf{R}_S(0) & \mathbf{R}_S(L)\\ \mathbf{R}_S^T(L) & \mathbf{R}_S(0)\end{bmatrix}\\ &\times\begin{bmatrix}\mathbf{H}^{(M)H}\mathbf{S} & \mathbf{0}\\ \mathbf{0} & \mathbf{H}^{(M)H}\mathbf{S}\end{bmatrix} + \Lambda\begin{bmatrix}\mathbf{S}^H & \mathbf{0}\\ \mathbf{0} & \mathbf{S}^H\end{bmatrix}\\ &\times\begin{bmatrix}\sigma^2\mathbf{I}_{NJM} & \mathbf{R}_N(L)\\ \mathbf{R}_N^T(L) & \sigma^2\mathbf{I}_{NJM}\end{bmatrix}\begin{bmatrix}\mathbf{S} & \mathbf{0}\\ \mathbf{0} & \mathbf{S}\end{bmatrix}.\end{aligned} \tag{71}$$

From the basic theory of the linear algebra and (71), the rank of the matrix $\mathbf{R}_\alpha$ is the maximum of the ranks of the two matrix terms in the right side of (71). The ranks of those two matrix terms are $M_T + PL$ and $N_n = \min\{2M_T, NJ(M+L)\}$, respectively. Therefore

$$\mathrm{rank}\{\mathbf{R}_\alpha\} = \max\{M_T + PL, N_n\}. \tag{72}$$

From the linear algebraic theory, (72) tells that the matrix $\mathbf{R}_\alpha$ has $\min\{M_T - PL, 2M_T - N_n\}$ zero eigenvalues. ■

It is noted that the presence of zero eigenvalues brings an infinite number of optimum solutions in terms of the Wiener–Hopf equation (12), which, however, provide the same residual error power.

Next, we present a theorem that determines the spread of the nonzero eigenvalues of the subband STAP system under the assumption of the noise-free environments.

Theorem 4: In the absence of noise, the eigenvalue system of $\mathbf{R}_\alpha$ is the same as that of the associated signal correlation matrix

$$\mathbf{R}_\beta = \begin{bmatrix}\mathbf{R}_S(0) & \mathbf{R}_S(L)\\ \mathbf{R}_S^T(L) & \mathbf{R}_S(0)\end{bmatrix}. \tag{73}$$

Proof: In the absence of noise, (71) is represented as

$$\mathbf{R}_\alpha = \begin{bmatrix} \Phi & \mathbf{0} \\ \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} \mathbf{S}^H\mathbf{H}^{(M)} & \mathbf{0} \\ \mathbf{0} & \mathbf{S}^H\mathbf{H}^{(M)} \end{bmatrix} \times \begin{bmatrix} \mathbf{R}_S(0) & \mathbf{R}_S(L) \\ \mathbf{R}_S^T(L) & \mathbf{R}_S(0) \end{bmatrix} \begin{bmatrix} \mathbf{H}^{(M)H}\mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}^{(M)H}\mathbf{S} \end{bmatrix}. \tag{74}$$

Define

$$\Lambda_u = \Sigma_s^{1/2} \tag{75}$$

and

$$\mathbf{U} = \mathbf{S}^H\mathbf{H}^{(M)} = \Lambda_u \underline{\mathbf{U}}. \tag{76}$$

By using the results of (30)–(34) in the noise-free case, it is straightforward to show that $\underline{\mathbf{U}}$ is unitary from

$$\begin{aligned}\underline{\mathbf{U}}\underline{\mathbf{U}}^H &= \Lambda_u^{-1}\mathbf{S}^H\mathbf{H}^{(M)}\mathbf{H}^{(M)H}\mathbf{S}\Lambda_u^{-1} \\ &= \Sigma_s^{-1/2}\mathbf{S}^H\mathbf{S}\Sigma_s\mathbf{S}^H\mathbf{S}\Sigma_s^{-1/2} = \mathbf{I}.\end{aligned} \tag{77}$$

Therefore, $\mathbf{R}_\alpha$ can be written as

$$\begin{aligned}\mathbf{R}_\alpha &= \begin{bmatrix} \Phi & \mathbf{0} \\ \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} \Lambda_u\underline{\mathbf{U}}^H & \mathbf{0} \\ \mathbf{0} & \Lambda_u\underline{\mathbf{U}}^H \end{bmatrix} \times \begin{bmatrix} \mathbf{R}_S(0) & \mathbf{R}_S(L) \\ \mathbf{R}_S^T(L) & \mathbf{R}_S(0) \end{bmatrix} \begin{bmatrix} \underline{\mathbf{U}}\Lambda_u & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{U}}\Lambda_u \end{bmatrix} \\ &= \begin{bmatrix} \Phi\Lambda_u & \mathbf{0} \\ \mathbf{0} & \Phi\Lambda_u \end{bmatrix} \begin{bmatrix} \underline{\mathbf{U}}^H & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{U}}^H \end{bmatrix} \times \begin{bmatrix} \mathbf{R}_S(0) & \mathbf{R}_S(L) \\ \mathbf{R}_S^T(L) & \mathbf{R}_S(0) \end{bmatrix} \begin{bmatrix} \underline{\mathbf{U}} & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{U}} \end{bmatrix} \begin{bmatrix} \Lambda_u & \mathbf{0} \\ \mathbf{0} & \Lambda_u \end{bmatrix}.\end{aligned} \tag{78}$$

According to (59) and (75) and the definition of

$$\Phi\Lambda_u = \Sigma_s^{-1/2} \tag{79}$$

let

$$\Lambda_\alpha \triangleq \begin{bmatrix} \Phi\Lambda_u & \mathbf{0} \\ \mathbf{0} & \Phi\Lambda_u \end{bmatrix} = \begin{bmatrix} \Sigma_s^{-1/2} & \mathbf{0} \\ \mathbf{0} & \Sigma_s^{-1/2} \end{bmatrix} \tag{80}$$

$$\mathbf{U}_\alpha \triangleq \begin{bmatrix} \underline{\mathbf{U}}^H & \mathbf{0} \\ \mathbf{0} & \underline{\mathbf{U}}^H \end{bmatrix}. \tag{81}$$

In terms of (73), (79), (80), and (81), (78) is simplified as

$$\mathbf{R}_\alpha = \Lambda_\alpha\mathbf{U}_\alpha\mathbf{R}_\beta\mathbf{U}_\alpha^H\Lambda_\alpha^{-1}. \tag{82}$$

It is seen from (80) that Λ_α is a diagonal matrix with all the diagonal elements greater than zero; therefore, Λ_α is a special orthogonal matrix. Since $\underline{\mathbf{U}}$ is unitary, it is clear that $\mathbf{U}_\alpha$ is also a unitary matrix. Because the unitary or the orthogonal transformation matrix-based similarity transformation does not change the eigenvalues, the eigenvalue system of $\mathbf{R}_\alpha$ is the same as that of $\mathbf{R}_\beta$. ■

From Theorem 4, we see that the spread of the nonzero eigenvalues of $\mathbf{R}_\alpha$ is the same as that of $\mathbf{R}_\beta$ and is no longer dependent on the channels. By solving the characteristic equation of $\mathbf{R}_\beta$, i.e., (det$(\lambda\mathbf{I} - \mathbf{R}_\beta) = 0$ ("det" denotes determinant), we obtained that for the cases of $0 < L < M_t$ (selection of L is restricted by the requirement of polyphase representation), all the nonzero eigenvalues of $\mathbf{R}_\beta$ only take value of either 1 or 2, which means the spread is 2. It can be seen that in the presence of noise, the spread of the nonzero eigenvalues of $\mathbf{R}_\alpha$ will increase. In the case of high input SNR, however, the increase of the spread of the nonzero eigenvalues of $\mathbf{R}_\alpha$ is not significant.

As we show in the sequel by numerical examples, the spread of the nonzero eigenvalues of $\mathbf{R}_\alpha$ is greatly reduced. Hence, the convergence rate of LMS algorithm performed by (63) under SSTAP will be faster than that under the conventional STAP.

V. Simulation Results

In this section, computer simulation results are presented to demonstrate the improvement of the convergence rate of the proposed signal subspace-based subband approach of STAP over that of the conventional STAP.

A uniform circular array with three elements of identical omnidirectional responses is employed. The interelement spacing is $\sqrt{3}\lambda$, where λ is the wavelength of the radio frequency. The oversampling factor is assumed to be $J = 2$. The scenarios of multiple users are considered. All the user signals are modulated by QPSK with raised-cosine pulse shaping filtering, where the roll-off factor is assumed to be 1.0. We assume that six rays are being received at the array for each user with different elevation θ and azimuth φ. Without loss of generality, the propagation loss of the first ray of each signal is assumed to be 1, and the relative time delay of this ray is assumed as 0.

The training sequence is assumed to be the ideal replica of the desired user signal. The taps of the STAP filter at each extended channel is assumed to be $M = 20.2000$ random integers, which are generated as the data source of each user. The input SNR of the first ray (for $\xi = 1$) is 10 dB for each user signal, and the noise power is defined as 0 dB.

We use the output residual error power, i.e.,

$$\varepsilon^2_{\mathrm{STAP}}(l) = \frac{1}{L_t}\sum_{n=1}^{L_t}\left|\bar{s}_1(n-v) - \sum_{m=1}^{M}\underline{\mathbf{w}}_m^T(l)\underline{\mathbf{x}}(n-m+1)\right|^2 \tag{83}$$

and

$$\varepsilon^2_{\mathrm{SSTAP}}(l) = \frac{1}{L_t}\sum_{n=1}^{L_t}\left|\bar{s}_1(n-v) - \mathbf{f}_\alpha^T(l)\mathbf{x}_\alpha(n)\right|^2 \tag{84}$$

to illustrate the output performances of STAP and SSTAP, respectively. In (83) and (84), we use another set of data of the same environment to examine the residual error power of the ongoing updating weight vectors, where l is the iteration number, and L_t is the length of the data. We also take $L_t = 2000$ here. In the simulations, the LMS algorithm is employed for both of the STAP and the SSTAP, and the initial states of the weight vectors are set to zero for all the cases. The maximum step size for LMS algorithm is generally given by [3]

$$0 < \mu < \frac{2}{\text{total input power}}. \tag{85}$$

Here, we choose

$$\mu = \frac{0.4}{\text{total input power}}. \tag{86}$$

TABLE I
(a) PARAMETERS OF THE DESIRED USER.
(b) PARAMETERS OF THE INTERFERENCE USER #1. (c) PARAMETERS OF THE INTERFERENCE USER #2

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-12.3	24.6	0	1.0
2	-28.0	30.7	0.99	0.02-0.84i
3	-13.1	46.7	1.16	0.09+0.80i
4	-0.80	13.0	3.89	-0.75-0.26i
5	-24.0	48.8	5.69	-0.54-0.44i
6	-26.0	25.9	7.41	-0.52-0.29i

(a)

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-8.6	33.6	0	1.0
2	-21.7	46.8	0.65	0.78+0.06i
3	-21.2	77.1	1.09	0.65-0.33i
4	-27.2	67.0	6.43	-0.58-0.17i
5	-10.9	76.8	6.69	0.06+0.54i
6	-26.0	59.0	9.46	-0.39-0.34i

(b)

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-6.6	120.6	0	1.0
2	-3.3	147.3	1.29	0.04+0.86i
3	-8.7	125.2	1.74	0.26+0.79i
4	-9.4	151.9	5.73	0.70+0.29i
5	-14.0	124.8	6.47	0.49+0.06i
6	-0.30	159.3	8.15	-0.37-0.25i

(c)

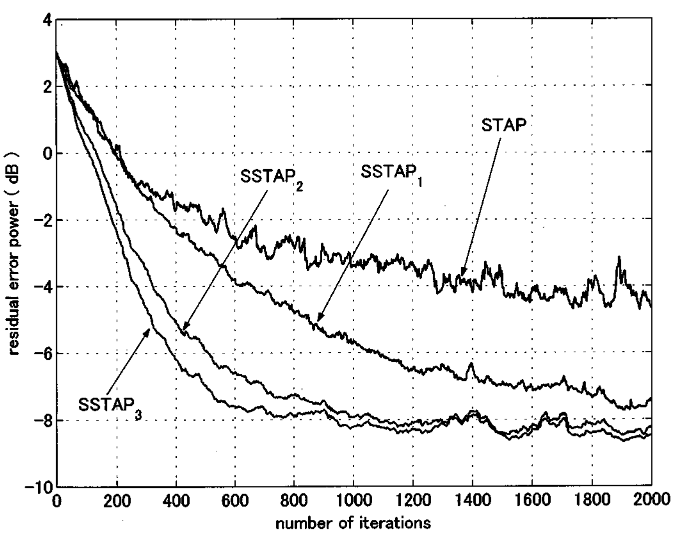

Fig. 3. Simulation results for Case 1. SSTAP1, SSTAP2, and SSTAP3 use 220, 520, and 1020 samples to estimate the signal subspace; STAP: $\mu = 2.0124 \times 10^{-5}$, $v = 12$; SSTAP$_1$: $\mu = 0.000\,861\,8$, $v = 16$; SSTAP$_2$: $\mu = 0.0020$, $v = 13$; SSTAP$_3$: $\mu = 0.0024$, $v = 12$.

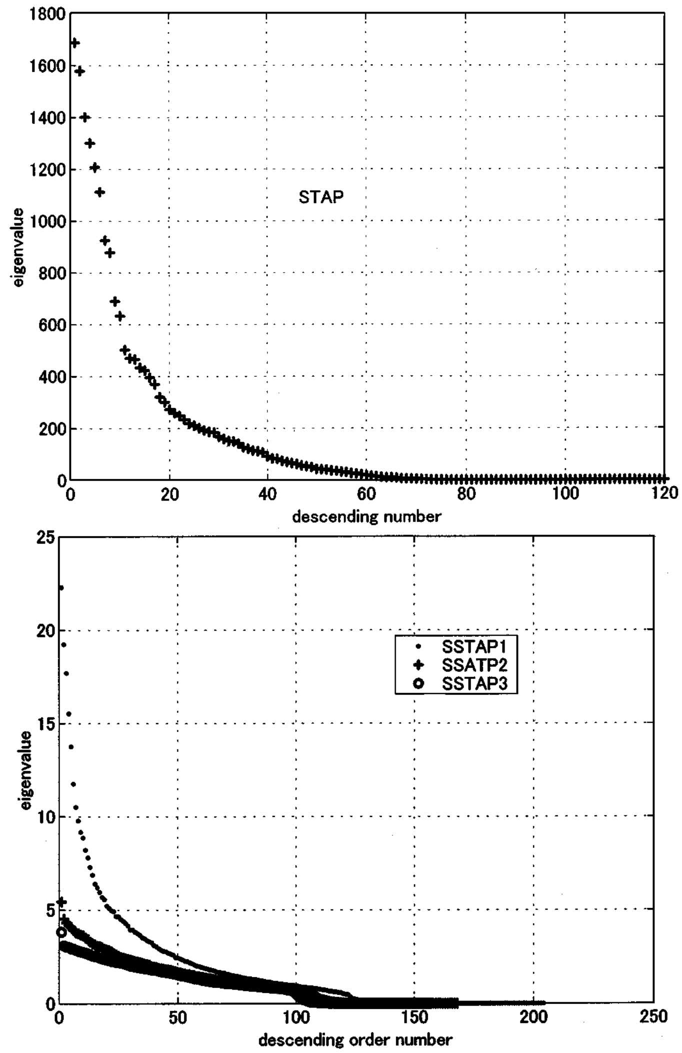

Fig. 4. Eigenvalue distribution of the correlation matrices. (Top) STAP. (Bottom) SSTAPs.

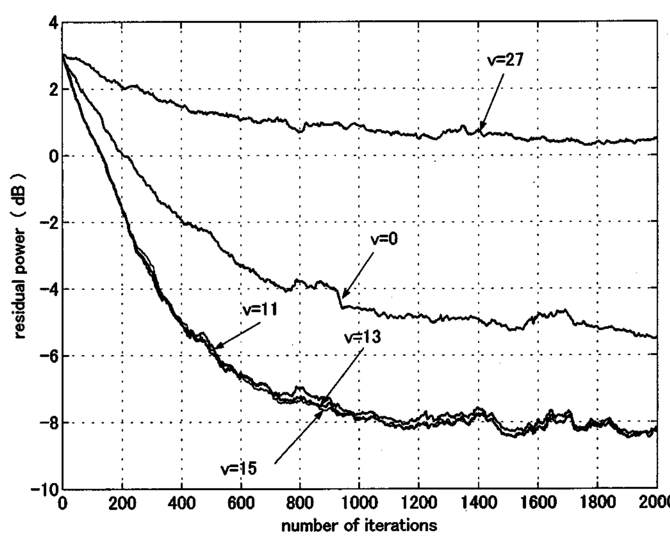

Fig. 5. Comparison of the residual error power under the different delays for Case 1 SSTAP$_2$: $\mu = 0.0020$, and $v = 0, 11, 13, 15, 27$.

For STAP, the total input power of all the taps is equal to the trace of $\mathbf{R}_X$, whereas for SSTAP, the total input power of all the taps equals the trace of $\mathbf{R}_\alpha$.

In order to evaluate the effect of the accuracy of the estimated signal subspace on the output performance of SSTAP, three numbers of the samples are used for the estimation of the signal subspace, namely, $N_t = 220, 520$, and 1020, respectively. Therefore, we have three different estimated signal subspace.

For comparison of the performance, an equivalent number of weights is considered. In addition, at each subband, two taps are taken, and L is selected to meet $M = (2-1) \times L + M_t$.

In the following, simulations are performed in two cases. Case 1 considers a three-user scenario, whereas Case 2 investigates a four-user scenario.

Case 1: Three users are present: One is the desired user, and the other two are interference users. The parameters for those three users are randomly generated and are listed in Table I(a)-(c), respectively. The length of the desired user channels is approximately 8, and the longest channel length of interference users is approximately 10. Before the weight adaptation of SSTAP, subspace decomposition is performed by eigendecomposition, and subsequently, AIC criterion is used to estimate the dimension of the signal subspace. The delay of the training sequence v is chosen based on (21), i.e.,

$$v = \frac{\text{dimension of the estimated signal subspace}}{\text{number of the extended array channels}} - 1. \quad (87)$$

According to the estimated dimension of the signal subspace, the full column rank matrix composed of the eigenvectors of the signal subspace is defined, from which SSTAP is performed. The estimated dimension of the signal subspace based on the AIC is approximately 84, and therefore, $M_t = 84/NJ = 14$. From the parameters listed in Table I, the actual column rank of the associated $\mathbf{H}^{(M)}$ is calculated as $\sum_{p=1}^{3}(D_p + M) \approx 84$. Therefore, the estimation of the dimension of the signal subspace truly describes the actual length of the channels.

Fig. 3 shows the residual error power versus the number of iterations. In this figure, SSTAP_1 denotes the first type of SSTAP in which the signal subspace is estimated by using 200 data samples, i.e., $N_t = 220$. Similarly, SSTAP_2 is for $N_t = 520$, and SSTAP_3 is for $N_t = 1020$. In Fig. 4, the eigenvalue distributions of the correlation matrices of the STAP and the SSTAP's are plotted, which clearly shows that the eigenvalue spread of the associated correlation matrix of SSTAP under the signal subspace-based subband approach is highly reduced.

From Figs. 3 and 4, it is evident that 1) all the three types of SSTAP have smaller eigenvalue spread and provide faster convergence rate than that of the conventional STAP, respectively, and 2) larger number of the samples used for the estimation of the signal subspace yields smaller eigenvalue spread and faster convergence rate, which is due to the fact that the larger the number of the samples to be used, the more accurate the estimated signal subspace obtained will be. In practice, the number N_t should be determined by considering how fast the environment is varying, as well as the computational burden the system can support.

In order to show the efficiency of the selection of the delay v as mentioned in (87), the residual error power performances are shown in Fig. 5 for five different v under the assumption of SSTAP_2, where $v = 13$ is that selected according to (87). The figure clearly shows that SSTAP under the selection of v according to (87) nearly output the best performance. On the other hand, when we take value of $v = 0$ or $v = D_1 + M - 1 = 27$, the output performance greatly degrades.

TABLE II
(a) PARAMETERS OF THE DESIRED USER. (b) PARAMETERS OF THE INTERFERENCE USER #1. (c) PARAMETERS OF THE INTERFERENCE USER #2. (d) PARAMETERS OF THE INTERFERENCE USER #3

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-10.3	22.6	0	1.0
2	-28.8	11.1	1.05	-0.66-0.40i
3	-5.1	3.4	1.33	-0.06+0.73i
4	-3.7	22.3	1.84	0.60+0.30i
5	-0.9	1.1	3.04	-0.46-0.19i
6	-26.6	30.5	3.41	0.24-0.41i

(a)

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-5.6	37.6	0	1.0
2	-4.0	38.8	1.33	-0.82+0.07i
3	-29.1	34.4	1.44	0.37-0.38i
4	-6.1	40.5	3.54	0.09-0.47i
5	-8.4	75.8	6.29	0.05+0.44i
6	-1.7	49.8	6.38	-0.17+0.25i

(b)

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-4.6	130.6	0	1.0
2	-11.4	137.7	2.21	0.94+0.00i
3	-9.0	157.5	2.96	-0.04-0.55i
4	-13.1	135.3	3.75	0.32+0.40i
5	-7.3	133.5	5.97	0.42+0.27i
6	-12.5	158.6	6.25	-0.34+0.20i

(c)

No.	θ (deg.)	φ (deg.)	τ (sym.)	ξ
1	-9.6	108.6	0	1.0
2	-0.2	103.7	0.03	0.30-0.82i
3	-7.0	99.1	2.15	-0.52+0.47i
4	-14.2	94.3	2.57	0.09+0.58i
5	-14.8	102.9	3.63	0.28+0.43i
6	-14.4	102.2	4.75	0.47+0.14i

(d)

Case 2: Four users are present. The parameters for the four users are listed in Table II(a)–(d), respectively, and the length of the four user channels are assumed to be shorter than that in Case 1. The length of the desired user channels is about 4, whereas the longest channel length of interference users is about 7. The detected dimension of the signal subspace based on AIC is approximately 96, and subsequently, M_t is estimated as 16. The result is very close to the actual dimension of $\sum_{p=1}^{3}(D_p + M) \approx 99$, as is calculated from Table II.

Fig. 6 plots the residual error power versus the number of iterations. Similar to Case 1, the delay in the training sequence is chosen as that mentioned in (87). This figure again shows that SSTAP outperforms the conventional STAP over the convergence rate. In Fig. 6, we see that even when the number of users is larger than that of the array elements, the extended channels achieved from oversampling provide some extra degrees of freedom and enable the STAP and the SSTAP systems output reasonable residual error power performance.

It is noted that in practice, the length of the training sequence depends on the number of iterations required for convergence. The smaller the number of iterations required for the steady

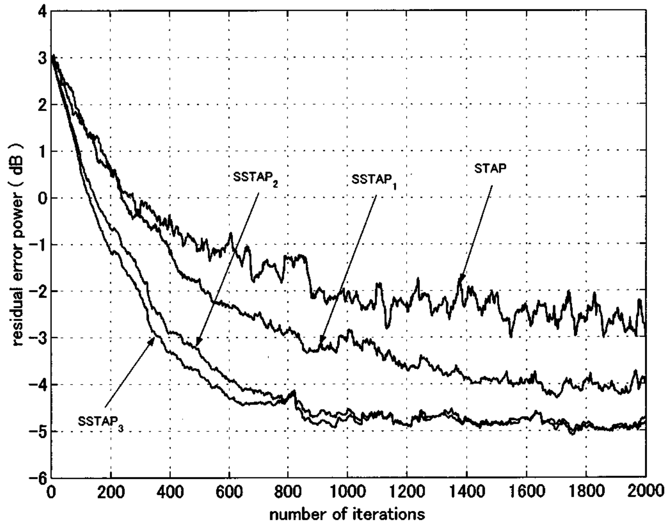

Fig. 6. Simulation results for Case 2. SSTAP1, SSTAP2, and SSTAP3 use 220, 520, and 1020 samples to estimate the signal subspace; STAP: $\mu = 1.9970 \times 10^{-5}$, $v = 15$; SSTAP_1: $\mu = 0.000\,984\,53$, $v = 17$; SSTAP_2: $\mu = 0.0017$, $v = 15$; SSTAP_3: $\mu = 0.0019$, $v = 15$.

state, the shorter the training sequence. From Figs. 3 and 6, 600 symbols of the training sequence are required for the convergence of the two cases of SSTAP.

VI. Conclusion

In this paper, we have investigated the conditions of STAP systems to realize perfect processing in noise-free situations and have proposed the space-time signal subspace-based subband approach to STAP, which is based on the polyphase representation. Because of the decorrelation by the space-time signal subspace-based subband filtering, the subband approach of STAP greatly improves the convergence rate without reducing the residual error power performance. Simulation results have been presented, which well confirmed the theoretical results.

Acknowledgment

The authors would like to give thanks to Dr. B. Komiyama, ATR Adaptive Communications Research Laboratories, for his encouragement and helpful discussions, and to Dr. T. Ohira and Y. Kamiya, ATR Adaptive Communications Research Laboratories, for their discussions.

References

[1] J. G. Proakis, *Digital Communications*, 2nd ed. New York: McGraw-Hill, 1989.

[2] G. L. Stuber, *Principles of Mobile Communications*. Boston, MA: Kluwer, 1996.

[3] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.

[4] P. P. Vaidyanathan, *Multirate Systems and Filter banks*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[5] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill, 1991.

[6] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford, U.K.: Oxford Univ. Press, 1965.

[7] G. B. Giannakis, "Highlights of signal processing for communications," *IEEE Signal Processing Mag.*, vol. 16, pp. 14–50, Mar. 1999.

[8] A. J. van der Veen, "Resolution limits of blind multi-user multichannel identification schemes—The bandlimited case," in *Proc. ICASSP*, Atlanta, GA, May 1996.

[9] A. F. Naguib, A. Paulraj, and T. Kailath, "Capacity improvement with base-station antenna arrays in cellular CDMA," *IEEE Trans. Veh. Technol.*, vol. 43, pp. 691–698, Aug. 1994.

[10] S. Anderson, M. Millnert, M. Viberg, and B. Wahlberg, "An adaptive array for mobile communication systems," *IEEE Trans. Veh. Technol.*, vol. 40, pp. 230–236, 1991.

[11] J.-W. Liang and A. Paulraj, "Two stage CCI/ISI reduction with space-time processing in TDMA cellular networks," in *Proc. 30th Annual Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 1996.

[12] A. J. Paulraj and C. B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 49–83, Nov. 1997.

[13] R. Kohno, "Spatial and temporal communication theory using adaptive antenna array," *IEEE Pers. Commun.*, vol. 5, pp. 28–35, Feb. 1998.

[14] I. S. Reed, J. D. Mallett, and L. E. Brennan, "Rapid convergence rate in adaptive arrays," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, Nov. 1974.

[15] L. Tong, G. Xu, and T. Kailath, "Blind identification and equalization based on second-order statistics: A time domain approach," *IEEE Trans. Inform. Theory*, vol. 40, pp. 340–349, Mar. 1994.

[16] E. Moulines, P. Duhamel, J.-F. Cardoso, and S. Mayrargue, "Subspace methods for the blind identification of multichannel FIR filters," *IEEE Trans. Signal Processing*, vol. 43, pp. 516–525, Feb. 1995.

[17] C. Papadias and D. T. M. Slock, "Fractionally spaced equalization of linear polyphase channels and related blind techniques based on multichannel linear prediction," *IEEE Trans. Signal Processing*, vol. 47, pp. 641–654, Mar. 1999.

[18] J. R. Treichler, I. Fijalkow, and C. R. Johnson, Jr., "Fractional spaced equalizers, how long should they really be?," *IEEE Signal Processing Mag.*, vol. 13, pp. 65–81, May 1996.

[19] A. Gilloire and M. Vetterli, "Adaptive filtering in subbands with critical sampling: Analysis, experiments, and applications to acoustic echo cancellation," *IEEE Trans. Signal Processing*, vol. 40, pp. 1862–1875, Aug. 1992.

[20] M. de Courville and P. Dujamel, "Adaptive filtering in subbands using a weighted criterion," *IEEE Trans. Signal Processing*, vol. 46, pp. 2359–2371, Sept. 1998.

[21] S. Sandeep and V. U. Reddy, "A new approach to subband adaptive filtering," *IEEE Trans. Signal Processing*, vol. 47, pp. 655–664, Mar. 1999.

[22] M. R. Petraglia and S. K. Mitra, "Adaptive FIR filter structure based on the generalized subband decomposition of FIR filters," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 354–362, June 1993.

[23] G. Xu and T. Kailath, "Fast subspace decomposition," *IEEE Trans. Signal Processing*, vol. 42, pp. 539–551, Mar. 1994.

[24] Y. Zhang, K. Yang, and M. G. Amin, "Adaptive array processing for multipath fading mitigation via exploitation of filter banks," *IEEE Trans. Antennas Propagat.*, to be published.

[25] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.

[26] Y. Zhang, K. Yang, and Y. Karasawa, "Subband CMA adaptive arrays in multipath fading environment," *IEICE Trans. Commun.*, vol. J82-B, pp. 97–108, Jan. 1999.

Kehu Yang (M'99) received the B.E., M.S., and Ph.D. degrees from Xidian University (formerly The Northwest Telecommunications Engineering Institute), Xi'an, China, in 1982, 1984, and 1995, respectively.

He joined Xidian University in 1985, where he became an Associate Professor in May 1996. Since December 1998, he has been a Visiting Researcher at ATR Adaptive Communications Research Laboratories, Kyoto, Japan. His research interests include array signal processing, space-time adaptive processing, and space-time multiuser detection for mobile communications.

Yimin Zhang (M'88) received the M.S. and Ph.D. degrees from the University of Tsukuba, Tsukuba, Japan, in 1985 and 1988, respectively.

He joined the Faculty of the Department of Radio Engineering, Southeast University, Nanjing, China, in 1988. He served as a Technical Manager at Communication Laboratory Japan, Kawasaki, from 1995 to 1997 and as a Visiting Researcher at ATR Adaptive Communications Research Laboratories, Kyoto, Japan, from 1997 to 1998. Currently, he is a Research Fellow at the Department of Electrical and Computer Engineering, Villanova University, Villanova, PA. His current research interests are in the areas of array signal processing, space-time adaptive processing, multiuser detection, blind signal processing, and time–frequency analysis.

Yoshihiko Mizuguchi was born in Hiroshima, Japan, in 1946. He received the B.E. and M.S. degrees in electrical engineering from Hiroshima University and the Dr.Eng. degree from Tokyo Institute of Technology, Tokyo, Japan, in 1969, 1971, and 1985, respectively.

He joined the R&D Laboratories of Kokusai Denshin Denwa Co. (KDD), Saitama, Japan, in 1971. From 1986 to 1988, he worked in the R&D Department of INTELSAT. From 1988 to 1997, he was a Senior Research Engineer and a Project Manager of the KDD R&D Labs. From 1997 to 2000, he was a Department Head with ATR Adaptive Communications Research Labs, Kyoto, Japan. Since August 2000, he has been back with the KDD R&D Labs. He had been engaged in research and development of microwave antennas for satellite communications and terrestrial radio relay links. His current research interests are in the area of multimedia mobile communication technology.

Dr. Mizuguchi is a member of the Institute of Electronics, Information, and Communication Engineering (IEICE). He received the Excellent Paper Award of the IEICE in 1979.