Research Article

Robust Adaptive Modified Newton Algorithm for Generalized Eigendecomposition and Its Application

Jian Yang, Feng Yang, Hong-Sheng Xi, Wei Guo, and Yanmin Sheng

Laboratory of Network Communication System and Control, Department of Automation, University of Science and Technology of China, Hefei, Anhui 230027, China

We propose a robust adaptive algorithm for generalized eigendecomposition problems that arise in modern signal processing applications. To that extent, the generalized eigendecomposition problem is reinterpreted as an unconstrained nonlinear optimization problem. Starting from the proposed cost function and making use of an approximation of the Hessian matrix, a robust modified Newton algorithm is derived. A rigorous analysis of its convergence properties is presented by using stochastic approximation theory. We also apply this theory to solve the signal reception problem of multicarrier DS-CDMA to illustrate its practical application. The simulation results show that the proposed algorithm has fast convergence and excellent tracking capability, which are important in a practical time-varying communication environment.

1. INTRODUCTION

Generalized eigendecomposition has extensive applications in modern signal processing areas, for example, pattern recognition [1, 2], and signal processing for wireless communications [3, 4]. Many efficient adaptive algorithms have been proposed for principal component analysis [5–7], which is a special case of generalized eigendecomposition. However, developing efficient adaptive algorithms for generalized eigendecomposition has not been addressed so far. This paper aims to propose a novel adaptive online algorithm for generalized eigendecomposition.

Consider the matrix pencil $(\mathbf{R}_y, \mathbf{R}_x)$, where $\mathbf{R}_y$ and $\mathbf{R}_x$ are $M \times M$ Hermitian and positive-definite matrices. A scalar λ and $M \times 1$ vector $\mathbf{w}$ that satisfy [8, 9]

$$\mathbf{R}_y \mathbf{w} = \lambda \mathbf{R}_x \mathbf{w} \qquad (1)$$

are called the generalized eigenvalue and corresponding generalized eigenvector of matrix pencil $(\mathbf{R}_y, \mathbf{R}_x)$, respectively. In this paper, we are interested in finding the generalized eigenvector corresponding to the largest eigenvalue.

Many numerical methods have been presented for the generalized eigendecomposition problem [8]. However, these methods are inefficient in a nonstationary signal environment, since they are computationally intensive and belong to the class of batch processing methods. For practical signal processing applications, an adaptive online algorithm is preferred, especially in a nonstationary signal environment. Chatterjee et al. have presented an online generalized eigendecomposition algorithm for linear discriminant analysis (LDA) [10]. However, this algorithm as well as those in [11, 12] are based on the gradient method, and their performance is largely determined by the step size, which is difficult to select in a practical application. To overcome these difficulties, Rao et al. apply a fixed-point algorithm to solve the generalized eigendecomposition problem [13]. The resulting RLS-like algorithm is proven to be more computationally feasible and faster than most of the gradient methods. Recently, by using the recursive least-square learning rule, Yang et al. develop fast adaptive algorithms for the generalized eigendecomposition problem [14]. Besides RLS techniques, the Newton method is also a well-known powerful technique in the area of optimization. By constructing a cost function based on the penalty function method, Mathew and Reddy develop a quasi-Newton adaptive algorithm for estimating the generalized eigenvector corresponding to the smallest generalized eigenvalue [9]. However, this method suffers from the difficulty of selecting an appropriate penalty factor, which requires its priori information of the covariance matrices, which is unavailable in most applications. As a result, this will affect the learning performance. In addition, for

many applications, the generalized eigenvector corresponding to the largest eigenvalue is desired.

In this paper, motivated by the work of Mathew and Reddy [9], we develop an efficient adaptive modified Newton algorithm to track the adaptive principal generalized eigenvector. The basic idea is that we reformulate the generalized eigendecomposition problem as minimizing an unconstrained nonquadratic cost function that has a unique global minimum and no other local minima, and then apply an appropriate Hessian matrix approximation to derive an adaptive modified Newton algorithm. The resulting algorithm is numerically robust no matter whether it is implemented with infinite or finite precision. We also illustrate its application by using it to solve an adaptive signal reception problem in a multicarrier DS-CDMA (MC-DS-CDMA) system [15].

The rest of the paper is organized as follows. In Section 2, we formulate the adaptive signal reception problem in an MC-DS-CDMA system as the principal generalized eigenvector estimation problem, to show the importance of the generalized eigendecomposition technique. In Section 3, the generalized eigendecomposition problem is reinterpreted as a nonlinear optimization problem, and a robust adaptive modified Newton algorithm is developed to estimate the principal generalized eigenvector. The convergence property of the proposed algorithm is also discussed. In Section 4, we present numerical simulation results to show the performance of the proposed algorithm. Conclusions are drawn in Section 5.

2. GENERALIZED EIGENDECOMPOSITION APPLICATION

In this section, we show that it is possible to formulate the signal reception problem in a multicarrier DS-CDMA system [16] as a generalized eigendecomposition problem.

2.1. Signal model of MC-DS-CDMA system

Consider an MC-DS-CDMA system with K simultaneous users. Each one uses the same M carriers. The kth user, for $1 \leq k \leq K$, generates a data sequence:

$$b^{(k)} = \{\ldots, b_0^{(k)}, b_1^{(k)}, b_2^{(k)}, \ldots\} \tag{2}$$

with a symbol interval of T seconds. We assume the data symbols $b_j^{(k)}$ are independent variables with $E[b_j^{(k)}] = 0$ and $E[|b_j^{(k)}|] = 1$.

The kth user is provided a randomly generated signature sequence:

$$a^{(k)} = \{\ldots, a_0^{(k)}, a_1^{(k)}, \ldots, a_{G-1}^{(k)}, \ldots\}, \tag{3}$$

where G is the spreading gain and the elements $a_i^{(k)}$ are modelled as independent and identically distributed (i.i.d.) random variables such that $\Pr(a_i^{(k)} = -1) = \Pr(a_i^{(k)} = 1) = 1/2$. The sequence $a^{(k)}$ is used to spectrally spread the data symbols to form the signal [15]

$$a_k(t) = \sum_{i=-\infty}^{\infty} b_{\lfloor i/G \rfloor}^{(k)} a_i^{(k)} \psi(t - iT_c), \tag{4}$$

where $\lfloor x \rfloor$ denotes the largest integer less than or equal to x, the chip interval T_c is given by $T_c = T/G$, G is the number of chips per symbol interval, and $\psi(t)$ is the common chip waveform for all signals. We assume that the chip waveform $\psi(t)$ is bandlimited, such as the square-root raised-cosine pulse [17], and normalized so that $\int_{-\infty}^{\infty} \|\psi(t)\|^2 dt = T_c$.

Assume a slowly time-varying frequency-selective Rayleigh fading channel. Following the approach [16], by suitably choosing M and the bandwidth of $\psi(t)$, we can assume that each carrier experiences slowly varying flat fading. Then, the received signal in complex form is given by [18]

$$
\begin{aligned}
r(t) = &\sum_{k=1}^{K} \sum_{m=1}^{M} \sqrt{2P_k} \alpha_{k,m} e^{j\omega_m t} \\
&\cdot \left[\sum_{i=-\infty}^{\infty} b_{\lfloor i/G \rfloor}^{(k)} a_i^{(k)} \psi(t - iT_c - \tau_k) \right] + n(t),
\end{aligned} \tag{5}
$$

where ω_m is the frequency of the mth carrier, $\alpha_{k,m}$ accounts for the overall effects of phase shifts and fading for the mth carrier of the kth user, P_k and $\tau_k \in [0, T)$ represent the power for each carrier of the transmitted signal and the delay of the kth user signal, respectively, and $n(t)$ denotes additive white Gaussian noise.

Without loss of generality, throughout the paper we will consider the signal from the first user as the desired signal and the signals from all other users as interfering signals. Assume that synchronization has been achieved with the transmitted signal of the desired user. Therefore, the delay of the desired signal τ_1 can be taken to be zero. In order to avoid interchip interference for the desired signal when it is chip-synchronous, the waveform is chosen to satisfy the Nyquist criterion. Then the input signal to the first PN correlator (finger) associated with the mth carrier is written as

$$
\begin{aligned}
x_m[g] &= \frac{1}{T_c} \int_{-\infty}^{\infty} r(t) \psi^*(t - gT_c) e^{-j\omega_m t} dt \\
&= \sqrt{2P_1} \alpha_{1,m} b_{\lfloor g/G \rfloor}^{(1)} a_g^{(1)} + \sum_{k=2}^{K} i_{k,m}[g] + n_m[g],
\end{aligned} \tag{6}
$$

where g is the chip index, $n_m[g]$ denotes the component due to AWGN, and

$$i_{k,m}[g] = \sqrt{2P_k} \alpha_{k,m} \sum_{i=-\infty}^{\infty} b_{\lfloor i/G \rfloor}^{(k)} a_i^{(k)} R_\psi[(g-i)T_c - \tau_k] \tag{7}$$

is the component due to the kth user signal, $2 \leq k \leq K$. The function $R_\psi(\cdot)$ is the autocorrelation of the chip waveform defined by

$$R_\psi(t) = \frac{1}{T_c} \int_{-\infty}^{\infty} \psi(s) \psi^*(s - t) ds. \tag{8}$$

The input signal vector can be written as

$$
\begin{aligned}
\mathbf{x}[g] &= [x_1[g], x_2[g], \ldots, x_m[g]]^T \\
&= \mathbf{h}_1 b_{\lfloor g/G \rfloor}^{(1)} a_g^{(1)} + \sum_{k=2}^{K} \mathbf{h}_k \sum_{i=-\infty}^{\infty} b_{\lfloor i/G \rfloor}^{(k)} a_i^{(k)} \\
&\quad \cdot R_\psi[(g-i)T_c - \tau_k] + \mathbf{n}[g],
\end{aligned} \tag{9}
$$

where $\mathbf{h}_k = [\sqrt{2P_k}\alpha_{k,1},\ldots,\sqrt{2P_k}\alpha_{k,m}]^T$, $1 \le k \le K$, and $\mathbf{n}[g] = [n_1[g], n_2[g],\ldots,n_M[g]]^T$ is a zero-mean Gaussian random vector with covariance $\sigma^2\mathbf{I}$.

Then, the output signal of the first PN correlator to extract the signal at the mth carrier can be written as

$$y_m[n] = \frac{1}{\sqrt{G}}\sum_{l=0}^{G-1} a_{l+Gn}^{(1)} x_m[Gn+l] \tag{10}$$

and the output signal vector can be expressed as

$$\begin{aligned}\mathbf{y}[n] &= [y_1[n], y_2[n],\ldots, y_M[n]]^T \\ &= \mathbf{h}_1\sqrt{G}b_n^{(1)} + \sum_{k=2}^{K}\mathbf{h}_k\frac{1}{\sqrt{G}}\sum_{l=0}^{G-1}a_{l+Gn}^{(1)}\sum_{i=-\infty}^{\infty}b_{\lfloor i/G\rfloor}^{(k)} \\ &\quad \cdot a_i^{(k)}R_\psi[(l+Gn-i)T_c - \tau_k] + \mathbf{n}_1[n],\end{aligned} \tag{11}$$

where

$$\mathbf{n}_1(n) = \frac{1}{\sqrt{G}}\sum_{l=0}^{G-1}a_{l+Gn}^{(1)}\mathbf{n}[l+Gn] \tag{12}$$

is the noise component with $E\{\mathbf{n}_1[n]\mathbf{n}_1^H[n]\} = \sigma^2\mathbf{I}$. The received signal vectors $\mathbf{x}[g]$ and $\mathbf{y}[n]$ are referred to as undespreaded and despreaded received signal vectors of the desired user.

2.2. MSINR signal reception problem

From (11), the despreaded signal vector can be rewritten as

$$\mathbf{y}[n] = \mathbf{s}[n] + \mathbf{u}[n], \tag{13}$$

where $\mathbf{s}[n] = \mathbf{h}_1\sqrt{G}b_{[n]}^{(1)}$ denotes the desired signal vector, and $\mathbf{u}[n]$ is the undesired signal vector.

The optimal weight vector under the MSINR performance criterion can be found as [15]

$$\mathbf{w}_{\mathrm{MSINR}} = \arg\max_{\mathbf{w}}\frac{\mathbf{w}^H\mathbf{R}_s\mathbf{w}}{\mathbf{w}^H\mathbf{R}_u\mathbf{w}}, \tag{14}$$

where $\mathbf{R}_s = E\{\mathbf{s}[n]\mathbf{s}^H[n]\}$ and $\mathbf{R}_u = E\{\mathbf{u}[n]\mathbf{u}^H[n]\}$ are the covariance matrices of the desired and undesired signals, respectively. It is obvious that the optimal weight vector $\mathbf{w}_{\mathrm{MSINR}}$ is the generalized eigenvector corresponding to the maximum generalized eigenvalue of the matrix pencil $(\mathbf{R}_s, \mathbf{R}_u)$, that is,

$$\mathbf{R}_s\mathbf{w}_{\mathrm{MSINR}} = \lambda_{\max}\mathbf{R}_u\mathbf{w}_{\mathrm{MSINR}}, \tag{15}$$

where $\lambda_{\max}$ is the maximum generalized eigenvalue. Unfortunately, because $\mathbf{s}[n]$ and $\mathbf{u}[n]$ cannot be separately obtained from the received signal $\mathbf{y}[n]$, it seems difficult to obtain $\mathbf{w}_{\mathrm{MSINR}}$ from (14). In the following, we will propose an improved criterion equivalent to MSINR to overcome the above difficulty.

According to (9) and (11), after some calculations, the autocorrelation matrices $\mathbf{R}_x = E\{\mathbf{x}[g]\mathbf{x}^H[g]\}$ and $\mathbf{R}_y = E\{\mathbf{y}[n]\mathbf{y}^H[n]\}$ are given by, respectively,

$$\begin{aligned}\mathbf{R}_x &= \mathbf{h}_1\mathbf{h}_1^H + \sum_{k=2}^{K}\mathbf{h}_k\mathbf{h}_k^H\sum_{i=-\infty}^{\infty}\|R_\psi(iT_c - \tau_k)\|^2 + \sigma^2\mathbf{I}, \\ \mathbf{R}_y &= G\mathbf{h}_1\mathbf{h}_1^H + \sum_{k=2}^{K}\mathbf{h}_k\mathbf{h}_k^H\sum_{i=-\infty}^{\infty}\|R_\psi(iT_c - \tau_k)\|^2 + \sigma^2\mathbf{I}.\end{aligned} \tag{16}$$

Hence, we have

$$\begin{aligned}\mathbf{R}_x &= \frac{1}{G}\mathbf{R}_s + \mathbf{R}_u, \\ \mathbf{R}_y &= \mathbf{R}_s + \mathbf{R}_u.\end{aligned} \tag{17}$$

Let us consider the following function:

$$f(\mathbf{w}) = \frac{\mathbf{w}^H\mathbf{R}_y\mathbf{w}}{\mathbf{w}^H\mathbf{R}_x\mathbf{w}} = G - \frac{G-1}{\gamma/G+1}, \tag{18}$$

where

$$\gamma = \frac{\mathbf{w}^H\mathbf{R}_s\mathbf{w}}{\mathbf{w}^H\mathbf{R}_u\mathbf{w}} \tag{19}$$

for any $\mathbf{w}$ except for $\mathbf{w}^H\mathbf{R}_u\mathbf{w} = 0$. If $\mathbf{R}_u$ is full rank, this function is valid for any $\mathbf{w} \ne 0$. According to (18), we can see that if $G > 1$, the weight vector $\mathbf{w}$ that maximizes $f(\mathbf{w})$ eventually maximizes γ. Therefore, the optimal weight vector can be found as

$$\mathbf{w}_{\mathrm{MSINR}} = \arg\max_{\mathbf{w}}\frac{\mathbf{w}^H\mathbf{R}_y\mathbf{w}}{\mathbf{w}^H\mathbf{R}_x\mathbf{w}}. \tag{20}$$

Hereby, estimating the MSINR weight vector from (20) instead of (14), we do not need to know or estimate the covariance matrices of $\mathbf{s}[n]$ and $\mathbf{u}[n]$, which are basically not available at the receiving end. Obviously, this is the problem of estimating the principal generalized eigenvector from two observed sample sequences $\mathbf{y}[n]$ and $\mathbf{x}[g]$.

3. ROBUST ADAPTIVE MODIFIED NEWTON ALGORITHM FOR GENERALIZED EIGENDECOMPOSITION

To solve a class of signal processing problems similar to that in Section 2, we construct a novel unconstrained cost function. Then, starting from this cost function, a robust modified Newton algorithm is derived. Its convergence is rigorously analyzed by using stochastic approximation theory.

3.1. Generalized eigendecomposition problem reinterpretation

Let λ_i and $\mathbf{u}_i$ $(1 \le i \le M)$ be the generalized eigenvalue and the corresponding $\mathbf{R}_x$-orthonormalized generalized eigenvector of the matrix pencil $(\mathbf{R}_y, \mathbf{R}_x)$, that is, [9]

$$\begin{aligned}\mathbf{R}_y\mathbf{u}_i &= \lambda_i\mathbf{R}_x\mathbf{u}_i, \\ \mathbf{u}_i^H\mathbf{R}_x\mathbf{u}_j &= \delta_{ij},\end{aligned} \tag{21}$$

where δ_{ij} is the Kronecker delta function.

Consider the following nonlinear scalar cost function:

$$J(\mathbf{w}) = \mathbf{w}^H \mathbf{R}_x \mathbf{w} - \ln(\mathbf{w}^H \mathbf{R}_y \mathbf{w}). \tag{22}$$

As will be shown next, this is a novel criterion for the generalized eigendecomposition problem. In the following theorem, we assume that the maximum generalized eigenvalue of $(\mathbf{R}_y, \mathbf{R}_x)$ has multiplicity 1. The case when the multiplicity of the maximum generalized eigenvalue is larger than 1 will be discussed later.

Theorem 1. *Let* $\lambda_1 > \lambda_2 \geq \cdots \geq \lambda_M > 0$ *be the generalized eigenvalues of the matrix pencil* $(\mathbf{R}_y, \mathbf{R}_x)$. *Then* $\mathbf{w} = \mathbf{u}_1$ *is the unique global minimal point of* $J(\mathbf{w})$ *and the others are saddle points of* $J(\mathbf{w})$.

Proof. See Appendix A.   □

Theorem 1 shows that if the maximum generalized eigenvalue has multiplicity 1, $J(\mathbf{w})$ has a global minimum and no other local minima, and global convergence is guaranteed when one seeks the $\mathbf{R}_x$-orthonormalized generalized eigenvector corresponding to the maximum generalized eigenvalue of $(\mathbf{R}_y, \mathbf{R}_x)$ by iterative methods. When the multiplicity of the maximum generalized eigenvalue is more than 1, there are some local minima. Hence, the iterative algorithm will converge to one of these local minima. Nevertheless, it is not a hindrance for one to seek the principal generalized eigenvector, because these local minima themselves are the $\mathbf{R}_x$-orthonormalized generalized eigenvectors corresponding to the maximum generalized eigenvalue. Therefore, the principal generalized eigenvector estimation problem can be reformulated as the following unconstrained nonlinear optimization problem:

$$\min_{\mathbf{w}} J(\mathbf{w}). \tag{23}$$

3.2. Adaptive modified Newton algorithm derivation

The Hessian matrix of $J(\mathbf{w})$ with respect to $\mathbf{w}$ is derived in Appendix A as

$$\mathbf{H} = \mathbf{R}_x - \mathbf{R}_y (\mathbf{w}^H \mathbf{R}_y \mathbf{w})^{-1} + (\mathbf{w}^H \mathbf{R}_y \mathbf{w})^{-2} \mathbf{R}_y \mathbf{w} \mathbf{w}^H \mathbf{R}_y. \tag{24}$$

In order to simplify the Hessian matrix, we drop the second term on the right-hand side of (24). Therefore, an approximation to the Hessian matrix can be written as:

$$\tilde{\mathbf{H}} = \mathbf{R}_x + (\mathbf{w}^H \mathbf{R}_y \mathbf{w})^{-2} \mathbf{R}_y \mathbf{w} \mathbf{w}^H \mathbf{R}_y. \tag{25}$$

The inverse Hessian matrix is given by

$$\tilde{\mathbf{H}}^{-1} = \mathbf{R}_x^{-1} - \frac{\mathbf{R}_x^{-1} \mathbf{R}_y \mathbf{w} \mathbf{w}^H \mathbf{R}_y \mathbf{R}_x^{-1}}{(\mathbf{w}^H \mathbf{R}_y \mathbf{w})^2 + \mathbf{w}^H \mathbf{R}_y \mathbf{R}_x^{-1} \mathbf{R}_y \mathbf{w}}. \tag{26}$$

Then the modified Newton algorithm for updating the weight vector $\mathbf{w}[n+1]$ can be written as

$$\begin{aligned} \mathbf{w}[n+1] &= \mathbf{w}[n] - [\tilde{\mathbf{H}}^{-1} \nabla J(\mathbf{w})]\big|_{\mathbf{w}=\mathbf{w}[n]} \\ &= \frac{2\mathbf{R}_x^{-1} \mathbf{R}_y \mathbf{w} \mathbf{w}^H \mathbf{R}_y \mathbf{w}}{(\mathbf{w}^H \mathbf{R}_y \mathbf{w})^2 + \mathbf{w}^H \mathbf{R}_y \mathbf{R}_x^{-1} \mathbf{R}_y \mathbf{w}}\bigg|_{\mathbf{w}=\mathbf{w}[n]}. \end{aligned} \tag{27}$$

Remark 2. In the derivation of the updating rule (27), we approximate the Hessian matrix $\mathbf{H}$ by dropping a term so as to make the Hessian matrix $\tilde{\mathbf{H}}$ positive definite, and consequently make the resultant algorithm more robust, since for stabilizing the Newton-type algorithms it is necessary to guarantee that the Hessian matrix is positive definite. Although the approximation causes the resultant Hessian matrix to deviate from the true Hessian matrix, as shown in Section 4, the derived algorithm (27) can asymptotically converge to the principal generalized eigenvector of the matrix pencil $(\mathbf{R}_y, \mathbf{R}_x)$. In addition, the numerical simulation results show that the approximation has little influence on convergence speed and estimation accuracy. Therefore, the approximation is a reasonable step in developing the adaptive modified Newton algorithm.

We apply the following equations to recursively estimate $\mathbf{R}_x$ and $\mathbf{R}_y$:

$$\mathbf{R}_x[n+1] = \mu \mathbf{R}_x[n] + \mathbf{x}[n+1] \mathbf{x}^H[n+1], \tag{28}$$

$$\mathbf{R}_y[n+1] = \beta \mathbf{R}_y[n] + (1 - \beta)\mathbf{y}[n+1]\mathbf{y}^H[n+1], \tag{29}$$

where $0 < \mu, \beta < 1$ are the forgetting factors.

Let $\mathbf{P}[n+1] = \mathbf{R}_x^{-1}[n+1]$. Then we get

$$\mathbf{P}[n+1] = \frac{1}{\mu} \mathbf{P}[n] \left(\mathbf{I} - \frac{\mathbf{x}[n+1]\mathbf{x}^H[n+1]\mathbf{P}[n]}{\mu + \mathbf{x}^H[n+1]\mathbf{P}[n]\mathbf{x}[n+1]} \right). \tag{30}$$

Postmultiplying both sides of (29) with $\mathbf{w}[n]$, we have

$$\begin{aligned} \mathbf{R}_y[n+1]\mathbf{w}[n] &= \beta \mathbf{R}_y[n]\mathbf{w}[n] \\ &\quad + (1-\beta)\mathbf{y}[n+1]\mathbf{y}^H[n+1]\mathbf{w}[n]. \end{aligned} \tag{31}$$

Applying the projection approximation [5] yields

$$\mathbf{r}[n+1] = \mathbf{R}_y[n+1]\mathbf{w}[n+1] \approx \mathbf{R}_y[n+1]\mathbf{w}[n]. \tag{32}$$

Then (31) can be rewritten as

$$\mathbf{r}[n+1] = \beta \mathbf{r}[n] + (1-\beta)\mathbf{y}[n+1]c^*[n+1], \tag{33}$$

where $c[n+1] = \mathbf{w}^H[n]\mathbf{y}[n+1]$. In addition, we define $d[n+1] = \mathbf{w}^H[n]\mathbf{R}_y[n+1]\mathbf{w}[n]$. Then according to (29) we obtain

$$d[n+1] = \beta d[n] + (1-\beta)c^*[n+1]c[n+1]. \tag{34}$$

Let

$$\tilde{\mathbf{w}}[n+1] = \frac{\mathbf{R}_y[n+1]\mathbf{w}[n]}{\mathbf{w}^H[n]\mathbf{R}_y[n+1]\mathbf{w}[n]} \tag{35}$$

so that the update rule of $\mathbf{w}[n+1]$ can be rewritten as

$$\mathbf{w}[n+1] = \frac{2\mathbf{P}[n+1]\tilde{\mathbf{w}}[n+1]}{1 + \tilde{\mathbf{w}}^H[n+1]\mathbf{P}[n+1]\tilde{\mathbf{w}}[n+1]}. \tag{36}$$

Thus, the adaptive modified Newton algorithm can be summarized as

$$\mathbf{P}[n+1] = \frac{1}{\mu}\mathbf{P}[n]\left(\mathbf{I} - \frac{\mathbf{x}[n+1]\mathbf{x}^H[n+1]\mathbf{P}[n]}{\mu + \mathbf{x}^H[n+1]\mathbf{P}[n]\mathbf{x}[n+1]}\right),$$

$$c[n+1] = \mathbf{w}^H[n]\mathbf{y}[n+1],$$

$$\mathbf{r}[n+1] = \beta\mathbf{r}[n] + (1-\beta)\mathbf{y}[n+1]c^*[n+1],$$

$$d[n+1] = \beta d[n] + (1-\beta)c[n+1]c^*[n+1],$$

$$\widetilde{\mathbf{w}}[n+1] = \frac{\mathbf{r}[n+1]}{d[n+1]},$$

$$\mathbf{w}[n+1] = \frac{2\mathbf{P}[n+1]\widetilde{\mathbf{w}}[n+1]}{1 + \widetilde{\mathbf{w}}^H[n+1]\mathbf{P}[n+1]\widetilde{\mathbf{w}}[n+1]}. \tag{37}$$

The simplest way to choose the initial values is to set $\mathbf{P}[0] = \eta_1\mathbf{I}$, $\mathbf{w}[0] = \mathbf{r}[0] = \eta_2[1 \ 0 \ \cdots \ 0]^T$, and $d[0] = \eta_3$, where η_i $(i = 1, 2, 3)$ are appropriate positive values. During deriving the algorithm (37), we have adopted the projection approximation approach [5]. The rationality of using projection approximation has been concretely explained in [5]. In this paper, the numerical results show that using the projection approximation has little impact on the performance of the proposed algorithm.

Note that the update step for $\mathbf{P}[n]$ involves subtraction. Hence, the numerical error may cause $\mathbf{P}[n]$ to lose the Hermitian positive definiteness, while $\mathbf{P}[n]$ is theoretically Hermitian positive definite. An efficient and robust way is to apply the QR-update method to calculate the square root matrices $\mathbf{P}^{1/2}[n]$ [19]. Because $\mathbf{P}[n] = \mathbf{P}^{1/2}[n]\mathbf{P}^{H/2}[n]$, the Hermitian positive definiteness remains regardless of any numerical error.

3.3. Convergence analysis

In this section, we apply the stochastic approximation method, which is developed by Ljung [20], and Kushner and Clark [21], to analyze the convergence property of the proposed algorithm based on updating rule (27). According to the stochastic approximation theory, a deterministic ordinary differential equation (ODE) can be associated with the recursive stochastic approximation algorithm, and the convergence of the algorithm can be studied in terms of this differential equation.

The ordinary differential equation corresponding to the proposed algorithm based on updating rule (27) can be written as

$$\frac{d\mathbf{w}(t)}{dt} = \frac{2\mathbf{R}_x^{-1}\mathbf{R}_y\mathbf{w}(t)\mathbf{w}^H(t)\mathbf{R}_y\mathbf{w}(t)}{\left(\mathbf{w}^H(t)\mathbf{R}_y\mathbf{w}(t)\right)^2 + \mathbf{w}^H(t)\mathbf{R}_y\mathbf{R}_x^{-1}\mathbf{R}_y\mathbf{w}(t)} - \mathbf{w}(t). \tag{38}$$

We have the following theorem to demonstrate the convergence of $\mathbf{w}(t)$.

Theorem 3. *Given the matrix pencil* $(\mathbf{R}_y, \mathbf{R}_x)$, *whose largest generalized eigenvalue* λ_1 *has multiplicity* 1, *and assuming that*

$\mathbf{u}_1^H\mathbf{R}_x\mathbf{w}(0) \neq 0$, *then the ODE* (38) *has a global asymptotically stable equilibrium state at* $(\lambda_1, \gamma\mathbf{u}_1)$, *where* γ *is a constant complex number with norm* $\|\gamma\| = 1$.

Proof. See Appendix B. □

Note that if $\|\gamma\| = 1$, $\gamma\mathbf{u}_1$ is also the $\mathbf{R}_x$-orthornormalized generalized eigenvector corresponding to the maximum generalized eigenvalue of $(\mathbf{R}_y, \mathbf{R}_x)$. Theorem 3 also shows that although we approximate the Hessian matrix when deriving the updating rule (27), the resultant algorithm can asymptotically converge to the principal generalized eigenvector.

4. SIMULATIONS

In this section, we apply the proposed algorithm to the signal reception problem in multicarrier DS-CDMA, and perform numerical simulation to investigate its performance. For each run, the proposed algorithm in this paper, the direct eigendecomposition method, the TTJ algorithm [15], and sample matrix/iterative (SMIT) [12] are implemented simultaneously in the simulations. The data in each plot is the average over 100 independent runs.

We consider a K-user asynchronous MC-DS-CDMA system of $M = 12$ carriers with processing gain $G = 32$. The system uses a square-root raised-cosine chip pulse with roll-off factor of 0.8 [17]. It is customary to truncate $\psi(t)$ such that it spans only several chips [18], and we assume that the duration of the pulse is $4T_c$. Throughout this section, the signal-to-noise ratio (SNR) of the desired user is fixed at 20 dB.

To evaluate the convergence speed and the estimate accuracy, the direction cosine and the normalized projection error (NPE) [22] are defined, respectively, as

$$\text{direction cosine} = \frac{\|\mathbf{w}^H(k)\mathbf{w}_{\text{MSINR}}\|}{\|\mathbf{w}(k)\| \|\mathbf{w}_{\text{MSINR}}\|},$$

$$\text{NPE} = 1 - \left(\frac{\|\mathbf{w}^H(k)\mathbf{w}_{\text{MSINR}}\|}{\|\mathbf{w}(k)\| \|\mathbf{w}_{\text{MSINR}}\|}\right)^2, \tag{39}$$

where $\mathbf{w}_{\text{MSINR}}$ is the theoretically optimal combining weight vector and can be computed by [23]

$$\mathbf{w}_{\text{MSINR}} = \mathbf{R}_u^{-1}\mathbf{h}_1. \tag{40}$$

We use the MSINR performance to assess the MAI suppression capability of the proposed algorithm. The expression for calculating the SINR at the nth iteration is given by

$$\text{SINR}(n) = 10\log\frac{\mathbf{w}^H[n]\mathbf{R}_s\mathbf{w}[n]}{\mathbf{w}^H[n]\mathbf{R}_u\mathbf{w}[n]}. \tag{41}$$

The proposed algorithm starts with initial values $\mathbf{r}[0] = \mathbf{w}[0] = [1 \ 0 \ \cdots \ 0]^T$, $d[0] = 1$, $\mathbf{P}[0] = 0.01\mathbf{I}$, $\mu = 0.995$, and $\beta = 0.8$. For the direct eigendecomposition method, we use the same method as (28) and (29) to estimate the $\mathbf{R}_x$ and $\mathbf{R}_y$ at the nth iteration. The initial values $\mathbf{R}_x[0] = 0.1\mathbf{I}$, $\mathbf{R}_y[0] = 0.1\mathbf{I}$, and a forgetting factor of 0.9 are set. We also start the TTJ algorithm with $\mathbf{w}[0] = [1 \ 0 \ \cdots \ 0]^T$. But its step size should be regulated according to different simulation environments.
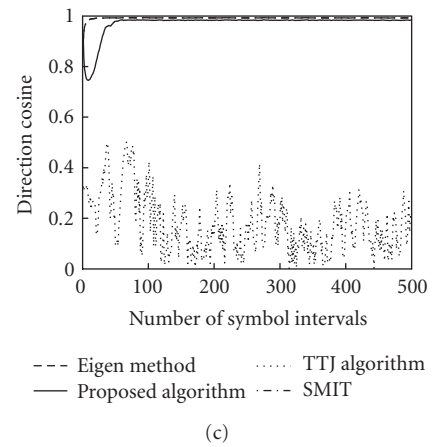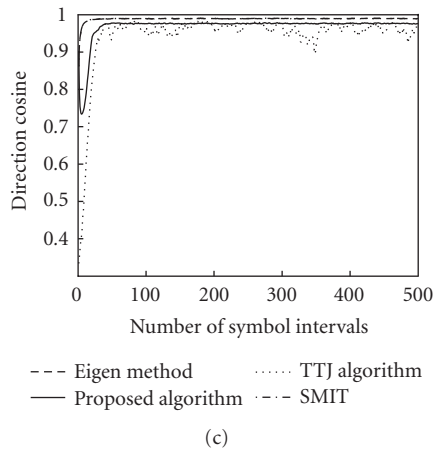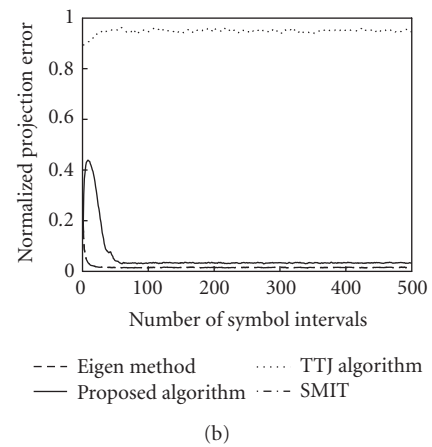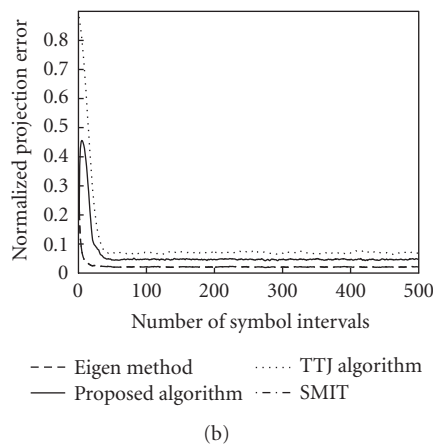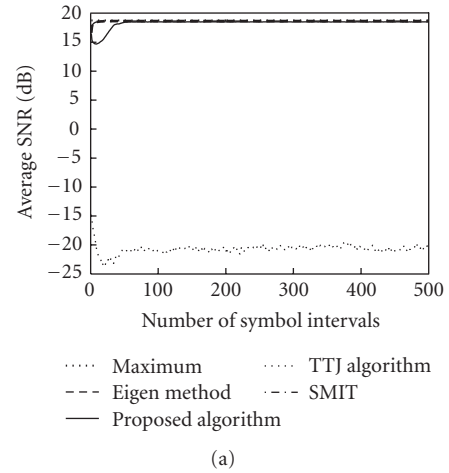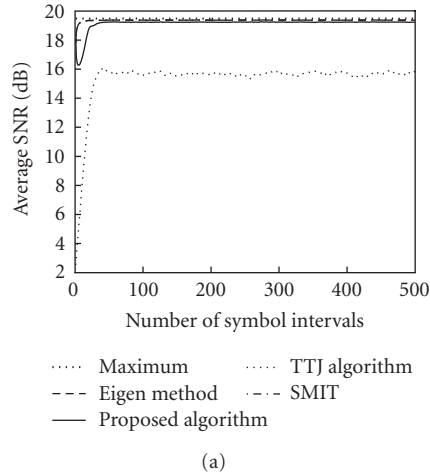
FIGURE 1: (a) SINR performance in the case of two interferers. (b) Normalized projection error in the case of two interferers. (c) Direction cosine performance in the case of two interferers.

FIGURE 2: (a) SINR performance in the case of five interferers. (b) Normalized projection error in the case of five interferers. (c) Direction cosine performance in the case of five interferers.

In the first simulation experiment, we consider the case when there are two interferers whose received powers are 10 dB stronger than the desired user. Figure 1 shows the simulation results. It can be observed that the eigenmethod and the proposed algorithm outperform the TTJ algorithm. The reason is that the TTJ algorithm belongs to the stochastic gra-

dient algorithm class and its fixed step size is chosen based on some tradeoff between tracking capability and accuracy; too small a value will bring on slow convergence and too large a value will lead to overshoot and instability [19]. The eigen method and SMIT have the best performance. However, their computational complexity is very high. Compared to these
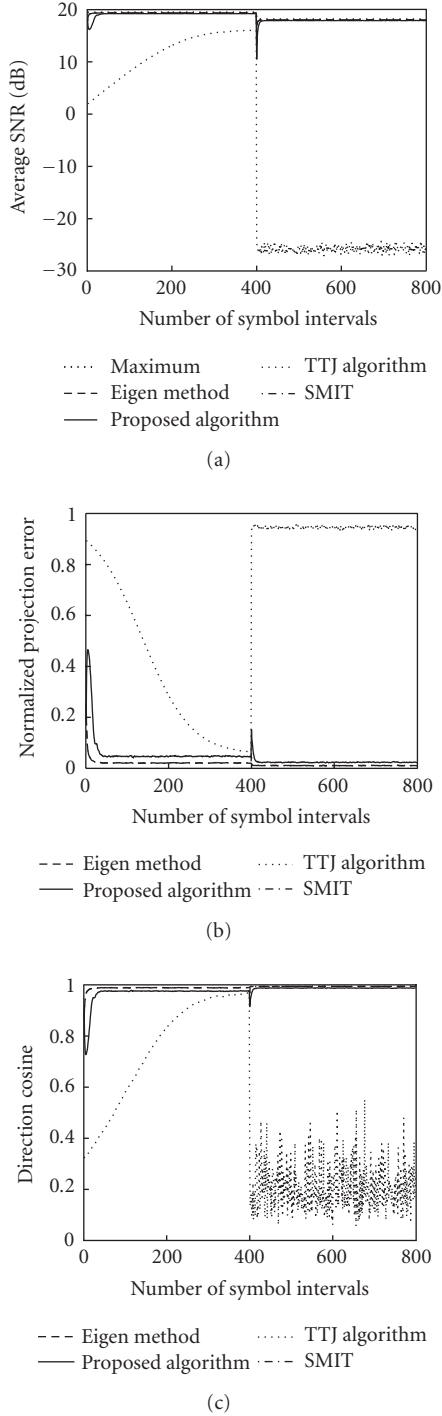
Figure 3: (a) SINR performance in the dynamical signal environment. (b) Normalized projection error in the dynamical signal environment. (c) Direction cosine performance in the dynamical signal environment.

methods, the complexity of the proposed algorithm has been greatly reduced, while its performance degrades only slightly. The simulation results also show that the approximation of the Hessian matrix and the projection approximation have little influence on the performance of the proposed algo-

rithm, since its performance approaches that of the eigen method, which uses neither of these approximation techniques.

In the next simulation experiment, we investigate the performance of the proposed algorithm in a signal environment with strong interference. We assume that there are two 10 dB, two 20 dB, and one 30 dB interferers. The simulation results in Figure 2 show that the performance of the eigen method and the proposed algorithm hardly changes, whereas the performance of the TTJ algorithm degrades rapidly. This is not surprising because at each step the TTJ algorithm uses a single instantaneous sample to update the weight vector, and as a result, the estimated weight vector oscillates around the MSINR combining weight vector. As the number and powers of the interferers increase, the oscillation becomes more dramatic and the amplitude increases. Consequently, the averaged performance degrades greatly in this scenario. In contrast, the proposed algorithm uses all of the data samples available up to the time instant $n + 1$ to estimate the optimal weight vector, and as a result, it performs well in a signal environment with strong interference. This experiment also shows that in the case with strong interferers, using the Hessian matrix approximation and the projection approximation has only a slight impact on the performance of the proposed algorithm.

In the final experiment, we study the tracking capability of the proposed algorithm in a dynamic environment. At the beginning, there are two 10 dB interferers, and at symbol interval 400, three 20 dB, one 30 dB, and one 40 dB interferers are added. Figure 3 shows the simulation results. Because there are few interferers and their powers are not very strong in the first phase, the TTJ algorithm performs very well. But in the second phase, too much interference and unregulated fixed step size cause the performance to degrade greatly. It can be observed that the eigen method, SMIT, and the proposed algorithm can rapidly adapt to the suddenly changed signal environment. This is because of using the forgetting factor in the recursive covariance matrix estimator. The simulation results also show that in time-varying environment the influence of the Hessian matrix approximation and the projection approximation is small.

Therefore, from the above simulation results in various signal environments, we conclude that the proposed algorithm has rapid convergence, sufficient estimation accuracy, and good tracking capability. These properties make it very useful in a practical signal environment, especially when the interfering power increases due to many practical reasons, such as too many interferers, incorrect power control, time-varying channel.

5. CONCLUSIONS

In this paper, we have studied the principal generalized eigenvector estimation problem. We proposed a new unconstrained cost function for the generalized eigendecomposition problem. Then, based on the proposed cost function, we have derived a robust adaptive modified Newton algorithm. The convergence of the proposed algorithm has been

rigorously analyzed. In addition, we applied the proposed algorithm to the adaptive signal reception problem in multi-carrier DS-CDMA systems, and the numerical simulation results show that the proposed algorithm has fast convergence and excellent tracking capability, which are very useful for a practical communication environment.

APPENDICES

A. PROOF OF THEOREM 1

Proof. Let ∇_R and ∇_I be the gradient operators with respect to the real and imaginary parts of $\mathbf{w}$. According to [19], the complex gradient operator is defined as $\nabla = (1/2)[\nabla_R + j\nabla_I]$. After some calculation, we can derive the gradient of $J(\mathbf{w})$ as

$$\nabla J(\mathbf{w}) = \mathbf{R}_x\mathbf{w} - \mathbf{R}_y\mathbf{w}(\mathbf{w}^H\mathbf{R}_y\mathbf{w})^{-1}. \tag{A.1}$$

When $\mathbf{w} = \mathbf{u}_i$, it is easy to show that $\nabla J(\mathbf{u}_i) = 0$. This implies that any $\mathbf{R}_x$-orthonormalized generalized eigenvector, $\mathbf{u}_i$, of $(\mathbf{R}_y, \mathbf{R}_x)$ is the stationary point of $J(\mathbf{w})$.

Conversely, $\nabla J(\mathbf{w}) = 0$ means

$$\mathbf{R}_y\mathbf{w} = (\mathbf{w}^H\mathbf{R}_y\mathbf{w})\mathbf{R}_x\mathbf{w}. \tag{A.2}$$

Hence, $\mathbf{w}$ is the generalized eigenvector of $(\mathbf{R}_y, \mathbf{R}_x)$, and the corresponding generalized eigenvalue is $(\mathbf{w}^H\mathbf{R}_y\mathbf{w})$. Premultiplying the both sides of (A.2) with $\mathbf{w}^H$ we have

$$\mathbf{w}^H\mathbf{R}_y\mathbf{w} = (\mathbf{w}^H\mathbf{R}_y\mathbf{w})(\mathbf{w}^H\mathbf{R}_x\mathbf{w}). \tag{A.3}$$

Since $\mathbf{R}_y$ is positive definite, $\mathbf{w}^H\mathbf{R}_y\mathbf{w} > 0$ for $\mathbf{w} \neq \mathbf{0}$. Therefore, we get $\mathbf{w}^H\mathbf{R}_x\mathbf{w} = 1$. This shows that stationary point, $\mathbf{w}$, of $J(\mathbf{w})$ is the $\mathbf{R}_x$-orthonormalized generalized eigenvector of $(\mathbf{R}_y, \mathbf{R}_x)$.

From above analysis, we conclude that $\mathbf{w}$ is a stationary point of $J(\mathbf{w})$ if and only if $\mathbf{w}$ is the $\mathbf{R}_x$-orthonormalized generalized eigenvector of $(\mathbf{R}_y, \mathbf{R}_x)$.

Let $\mathbf{H} = \nabla\nabla^H J(\mathbf{w})$ be the $M \times M$ Hessian matrix [7] of $J(\mathbf{w})$ with respect to the vector $\mathbf{w}$. After some calculations, the Hessian matrix $\mathbf{H}$ is given as

$$\mathbf{H} = \mathbf{R}_x - \mathbf{R}_y(\mathbf{w}^H\mathbf{R}_y\mathbf{w})^{-1} + (\mathbf{w}^H\mathbf{R}_y\mathbf{w})^{-2}\mathbf{R}_y\mathbf{w}\mathbf{w}^H\mathbf{R}_y. \tag{A.4}$$

Since $\mathbf{R}_x$ is positive definite, we have $\mathbf{R}_x = \mathbf{V}\mathbf{V}^H$, where $\mathbf{V}$ is an invertible $M \times M$ matrix. Let $\mathbf{e}_i = \mathbf{V}^H\mathbf{u}_i$ and $\mathbf{C} = \mathbf{V}^{-1}\mathbf{R}_y(\mathbf{V}^{-1})^H$. According to (21) we obtain

$$\begin{aligned}\mathbf{C}\mathbf{e}_i &= \lambda_i\mathbf{e}_i, \\ \mathbf{e}_i^H\mathbf{e}_j &= \delta_{ij}.\end{aligned} \tag{A.5}$$

Obviously, λ_i and $\mathbf{e}_i$ are the eigenvalue and the corresponding eigenvector of $\mathbf{C}$.

Let $\mathbf{e} = \mathbf{V}^H\mathbf{w}$. Then we get

$$\mathbf{H} = \mathbf{V}\left[\mathbf{I} - \frac{\mathbf{C}}{\mathbf{e}^H\mathbf{C}\mathbf{e}} + \frac{\mathbf{C}\mathbf{e}\mathbf{e}^H\mathbf{C}}{(\mathbf{e}^H\mathbf{C}\mathbf{e})^2}\right]\mathbf{V}^H = \mathbf{V}\mathbf{F}(\mathbf{e})\mathbf{V}^H, \tag{A.6}$$

where

$$\mathbf{F}(\mathbf{e}) = \mathbf{I} - \frac{\mathbf{C}}{\mathbf{e}^H\mathbf{C}\mathbf{e}} + \frac{\mathbf{C}\mathbf{e}\mathbf{e}^H\mathbf{C}}{(\mathbf{e}^H\mathbf{C}\mathbf{e})^2}. \tag{A.7}$$

From the fact that $\mathbf{e}_1^H\mathbf{C}\mathbf{e}_1 = \lambda_1$ and $\mathbf{C}\mathbf{e}_1\mathbf{e}_1^H = \lambda_1^2\mathbf{e}_1\mathbf{e}_1^H$ we have

$$\begin{aligned}\mathbf{F}(\pm\mathbf{e}_1) &= \mathbf{I} - \frac{\mathbf{C}}{\lambda_1} + \mathbf{e}_1\mathbf{e}_1^H, \\ \mathbf{F}(\pm\mathbf{e}_1)\mathbf{e}_1 &= \mathbf{e}_1, \\ \mathbf{F}(\pm\mathbf{e}_1)\mathbf{e}_i &= \left(1 - \frac{\lambda_i}{\lambda_1}\right)\mathbf{e}_i,\end{aligned} \tag{A.8}$$

where $i = 2, \ldots, M$. Since $(1 - \lambda_i/\lambda_1) > 0$, all the eigenvalues of $\mathbf{F}(\mathbf{e}_1)$ are positive. We can conclude that $\mathbf{F}(\mathbf{e})$ is positive definite at the point $\mathbf{e} = \pm\mathbf{e}_1$. Similarly, we can derive

$$\begin{aligned}\mathbf{F}(\mathbf{e}_i)\mathbf{e}_1 &= \left(1 - \frac{\lambda_1}{\lambda_i}\right)\mathbf{e}_1, \\ \mathbf{F}(\mathbf{e}_i)\mathbf{e}_i &= \mathbf{e}_i,\end{aligned} \tag{A.9}$$

where $i = 2, \ldots, M$. Because $(1 - \lambda_1/\lambda_i) < 0$, $F(\mathbf{e}_i)$ is neither positive definite nor negative definite. According to (A.6), we have

$$\mathbf{H}|_{\mathbf{w}=\pm\mathbf{u}_i} = \mathbf{V}\mathbf{F}(\mathbf{e}_i)\mathbf{V}^H. \tag{A.10}$$

It is clear that $\mathbf{H}$ is positive definite at the stationary point $\mathbf{w} = \mathbf{u}_1$. At any other stationary point $\mathbf{u}_i$ $(i = 2, \ldots, M)$, $\mathbf{H}$ is neither positive definite nor negative definite. This means that $\mathbf{w} = \mathbf{u}_1$ is the unique global minimal point of $J(\mathbf{w})$, and the other stationary points $\mathbf{u}_i$ $(i = 2, \ldots, M)$ are saddle points of $J(\mathbf{w})$. □

B. PROOF OF THEOREM 3

Proof. The vector $\mathbf{w}(t)$ can be expressed as a linear combination of M generalized eigenvectors $\mathbf{u}_i$ of $(\mathbf{R}_y, \mathbf{R}_x)$, which is given by

$$\mathbf{w}(t) = \sum_{i=1}^{M}\alpha_i(t)\mathbf{u}_i, \tag{B.1}$$

where $\alpha_i(t)$ are complex coefficients.

Substituting (B.1) into (38) and premultiplying by $\mathbf{u}_l^H\mathbf{R}_x$ yield

$$\begin{aligned}\frac{d\alpha_l(t)}{dt} &= \left[\left(\sum_{i=1}^{M}\lambda_i\|\alpha_i(t)\|^2\right)^2 + \sum_{i=1}^{M}\lambda_i^2\|\alpha_i(t)\|^2\right]^{-1} \\ &\quad \cdot \left(2\lambda_l\alpha_l(t)\sum_{i=1}^{M}\lambda_i\|\alpha_i(t)\|^2\right) - \alpha_l(t).\end{aligned} \tag{B.2}$$

Under the assumption $\mathbf{u}_1^H\mathbf{R}_x\mathbf{w}(0) \neq 0$ we can define $\theta_l = \alpha_l(t)/\alpha_1(t)$, $l = 2, \ldots, M$. Then we have

$$\frac{d\theta_l}{dt} = \left(\alpha_1(t)\frac{d\alpha_l(t)}{dt} - \alpha_l(t)\frac{d\alpha_1(t)}{dt}\right)\alpha_1^{-2}(t). \tag{B.3}$$

Substituting (B.2) into (B.3) yields

$$\frac{d\theta_l}{dt} = -(\lambda_1 - \lambda_l)\kappa(t)\theta_l(t), \qquad (B.4)$$

where

$$\kappa(t) = \left(2\sum_{i=1}^{M}\lambda_i\|\alpha_i(t)\|^2\right)$$
$$\times \left[\left(\sum_{i=1}^{M}\lambda_i\|\alpha_i(t)\|^2\right)^2 + \sum_{i=1}^{M}\lambda_i^2\|\alpha_i(t)\|^2\right]^{-1}. \qquad (B.5)$$

Since $\kappa(t) > 0$ for all $t > 0$, $\lim_{t\to\infty}\theta_l = 0$, $l = 2,\ldots,M$. It follows that $\lim_{t\to\infty}\alpha_l(t) = 0$, $l = 2,\ldots,M$, and $\mathbf{w}(t) = \alpha_1(t)\mathbf{u}_1$ is an asymptotically stable solution of (38).

Therefore, when t is large enough and $l = 1$, (B.2) can be simplified as

$$\frac{d\alpha_1(t)}{dt} = \frac{\alpha_1(t)[1 - \|\alpha_1(t)\|^2]}{1 + \|\alpha_1(t)\|^2}. \qquad (B.6)$$

In order to show that $\lim_{t\to\infty}\|\alpha_1(t)\| = 1$ we define $z(t) = \|\alpha_1(t)\|^2$ and $V[z(t)] = [z(t) - 1]^2$. Their time derivatives are

$$\dot{z}(t) = \alpha_1^*(t)\dot{\alpha}_1(t) + \dot{\alpha}_1^*(t)\alpha_1(t)$$
$$= 2\|\alpha_1(t)\|^2\frac{1 - \|\alpha_1(t)\|^2}{1 + \|\alpha_1(t)\|^2},$$
$$\dot{V}[z(t)] = 2[z(t) - 1]\dot{z}(t)$$
$$= -4\frac{[1 - \|\alpha_1(t)\|^2]^2\|\alpha_1(t)\|^2}{1 + \|\alpha_1(t)\|^2}. \qquad (B.7)$$

According to the theory of Lyapunov stability, $V(z)$ is a Lyapunov function, and $z = 1$ is asymptotically stable. Moreover, from (B.6) and $\lim_{t\to\infty}\|\alpha_1(t)\| = 1$, we can conclude $\lim_{t\to\infty}\alpha_1(t) = \gamma$, where $\|\gamma\| = 1$. Hence, $\mathbf{w}(t)$ in (38) will asymptotically converge to the stable solution $\gamma\mathbf{u}_1$. □

ACKNOWLEDGMENT

REFERENCES

[1] J. Lu, K. N. Plataniotis, and A. N. Venetsanopoulos, "Face recognition using LDA-based algorithms," *IEEE Transactions on Neural Networks*, vol. 14, no. 1, pp. 195–200, 2003.

[2] S. Fidler, D. Skočaj, and A. Leonardis, "Combining reconstructive and discriminative subspace methods for robust classification and regression by subsampling," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 3, pp. 337–350, 2006.

[3] T. F. Wong, T. M. Lok, J. S. Lehnert, and M. D. Zoltowski, "A linear receiver for direct-sequence spread-spectrum multiple-access systems with antenna arrays and blind adaptation," *IEEE Transactions on Information Theory*, vol. 44, no. 2, pp. 659–676, 1998.

[4] J. Yang, H. Xi, F. Yang, and Y. Zhao, "Fast adaptive blind beamforming algorithm for antenna array in CDMA systems," *IEEE Transactions on Vehicular Technology*, vol. 55, no. 2, pp. 549–558, 2006.

[5] B. Yang, "Projection approximation subspace tracking," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 95–107, 1995.

[6] S. Ouyang, P. C. Ching, and T. Lee, "Robust adaptive quasi-Newton algorithms for eigensubspace estimation," *IEE Proceedings: Vision, Image and Signal Processing*, vol. 150, no. 5, pp. 321–330, 2003.

[7] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent Component Analysis*, John Wiley & Sons, New York, NY, USA, 2001.

[8] G. H. Golub and C. F. VanLoan, *Matrix Computations*, John Hopkins University Press, Baltimore, Md, USA, 1991.

[9] G. Mathew and V. U. Reddy, "A quasi-Newton adaptive algorithm for generalized symmetric eigenvalue problem," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2413–2422, 1996.

[10] C. Chatterjee, V. P. Roychowdhury, J. Ramos, and M. D. Zoltowski, "Self-organizing algorithms for generalized eigendecomposition," *IEEE Transactions on Neural Networks*, vol. 8, no. 6, pp. 1518–1530, 1997.

[11] D. Xu, J. C. Principe, and H.-C. Wu, "Generalized eigendecomposition with an on-line local algorithm," *IEEE Signal Processing Letters*, vol. 5, no. 11, pp. 298–301, 1998.

[12] D. R. Morgan, "Adaptive algorithms for solving generalized eigenvalue signal enhancement problems," *Signal Processing*, vol. 84, no. 6, pp. 957–968, 2004.

[13] Y. N. Rao, J. C. Principe, and T. F. Wong, "Fast RLS-like algorithm for generalized eigendecomposition and its applications," *The Journal of VLSI Signal Processing*, vol. 37, no. 2-3, pp. 333–344, 2004.

[14] J. Yang, H. Xi, F. Yang, and Y. Zhao, "RLS-based adaptive algorithms for generalized eigen-decomposition," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1177–1188, 2006.

[15] T. M. Lok, T. F. Wong, and J. S. Lehnert, "Blind adaptive signal reception for MC-CDMA systems in Rayleigh fading channels," *IEEE Transactions on Communications*, vol. 47, no. 3, pp. 464–471, 1999.

[16] S. Kondo and L. B. Milstein, "Performance of multicarrier DS CDNA systems," *IEEE Transactions on Communications*, vol. 44, no. 2, pp. 238–246, 1996.

[17] J. G. Proakis, *Digital Communications*, McGraw-Hill, New York, NY, USA, 1995.

[18] J. Namgoong, T. F. Wong, and J. S. Lehnert, "Subspace multiuser detection for multicarrier DS-CDMA," *IEEE Transactions on Communications*, vol. 48, no. 11, pp. 1897–1908, 2000.

[19] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Upper Saddle River, NJ, USA, 2002.

[20] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.

[21] H. J. Kushner and D. S. Clark, *Stochastic Approximation Methods for Constrained and Unconstrained Systems*, Springer, New York, NY, USA, 1978.

[22] D. R. Morgan, J. Benesty, and M. M. Sondhi, "On the evaluation of estimated impulse responses," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 174–176, 1998.

[23] T. M. Lok and T. F. Wong, "Transmitter and receiver optimization in multicarrier CDMA systems," *IEEE Transactions on Communications*, vol. 48, no. 7, pp. 1197–1207, 2000.

Jian Yang received the B.S., M.S., and Ph.D. degrees from the University of Science and Technology of China (USTC), Hefei, China, in 2001, 2003, and 2005, respectively. He is currently with the Laboratory of Network Communication System and Control in USTC. His research area is multimedia communication and signal processing, including adaptive streaming media system design and performance optimization, adaptive load balance, adaptive filtering, antenna array signal processing, and frequency estimation.

Feng Yang received the B.S. degree in electrical engineering from Tongji University, Shanghai, China, in 2001, and the M.S. degree from USTC, Hefei, China, in 2003. He is currently pursuing the Ph.D. degree. His current research interests include adaptive filtering theory, MC-CDMA systems, and MIMO systems.

Hong-Sheng Xi received the B.S. and M.S. degrees in applied mathematics from the University of Science and Technology of China (USTC), Hefei, China, in 1980 and 1985, respectively. He is currently the Dean of the Department of Automation at USTC. He also directs the Laboratory of Network Communication System and Control. His research interests include stochastic control systems, discrete-event dynamic systems, network performance analysis and optimization, and wireless communications.

Wei Guo received his B.S. degree and Ph.D. degree in China University of Science and Technology and Chinese Academy of Sciences in 1983 and 1992, respectively. He worked in Communication Research Laboratory, Japan, and Hong Kong University of Science and Technology, in 1994-1995 and 1998, respectively. Professor Wei is the Member of the Communication Expert Group, State High Technology Project (863 Project), and the core Member of the Technical Group, China 3G Mobile Communication System Project. His current research interests are the concept and key technology for the 4G Mobile Communication system.

Yanmin Sheng received the B.S. degree in automation from University of Science and Technology of China, Hefei, China, in 2002, the Ph.D. degree in control science and engineering from University of Science and Technology of China, Hefei, China, in 2007. He has worked in areas of wireless communication, adaptive theory, and application, and statistical theory. His current research interests include particle filter application in communication, OFDM, and MIMO.