IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 7, 2018, accepted June 12, 2018, date of publication July 2, 2018, date of current version July 25, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2852060

A Generalized Minor Component Extraction Algorithm and Its Analysis

HONGZENG LI, BOYANG DU, XIANGYU KONG, YINGBIN GAO, CHANGHUA HU, AND XUHAO BIAN
Xi'an Research Institute of High Technology, Xi'an 710025, China

Corresponding author: Xiangyu Kong (xiangyukong01@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grants 61673387 and 61374120 and in part by the National Science Fund for Distinguished Youth Scholars of China under Grant 61025014.

ABSTRACT Generalized minor component analysis (GMCA) is of great use in modern signal processing. The GMCA algorithms can be simplified to extract the minor generalized eigenvector of the autocorrelation input matrices pencil. In contrast to batching methods, the Hebbian-rule-based algorithm can extract the minor generalized eigenvector online. Few Hebbian-rule-based GMCA algorithms have been reported in the literature, and most of them are not self-stabilizing. Thus, a novel algorithm for GMCA, which is advantageous in terms of good convergence speed, self-stabilizing property, and multiple generalized minor component extraction in sequence, is proposed in this paper. A theoretical analysis verifies these properties via matrix theory and the deterministic discrete-time method. Numerical simulations are conducted to further demonstrate the advantages of the proposed algorithm.

INDEX TERMS Generalized minor component analysis, deterministic discrete time method, Hebbian-rule-based algorithms, self-stabilizing property.

I. INTRODUCTION

IN signal processing, among the eigenvectors of the autocorrelation matrix distilled from the source signal, the minor component (MC) is the one that corresponds to the smallest eigenvalue of the matrix. Similarly, among the eigenvectors of the autocorrelation matrix pencil, which is the autocorrelation matrix of two signals, the generalized minor component (GMC) [1] is the one that corresponds to the smallest generalized eigenvalue of the matrix pencil. The GMC contains the common information of the noise subspace in the two source signals. The techniques to extract the MCs and GMCs are called minor component analysis (MCA) and generalized minor component analysis (GMCA), respectively. These powerful techniques have been widely applied in many areas, including total least squares (TLS) [2], moving target indication [3], clutter cancellation [4], and frequency estimation [5] for MCA and dimension reduction [6], machine learning [7], spectral estimation [8], and adaptive beam forming [9] for GMCA. These applications have common features. First, the number of signals is large. Second, in some applications, the MC must be estimated in real time because the noise subspace is time variant [10]. Third, in some applications, the calculation method must be simple due to limited computational resources. Therefore, due to these requirements, some former techniques, such as the batching method [11], [12], are not suitable.

Therefore, fast and real-time MCA/GMCA approaches based on Hebbian rules have been proposed [13]–[19]. Among MCA algorithms, the Oja-Xu algorithm [13] and AMEX algorithm [14] can track the non-stationary distributed minor component of the signals. Moreover, the OJAm algorithm [15], Kong algorithm [16], and Peng algorithm can not only achieve the above functions, but also are self-stabilizing. In a self-stabilizing algorithm, the state vector can be guaranteed to converge to a normalized MC.

To the best of our knowledge, few studies of GMCA algorithms exist. An online GMCA algorithm was proposed by Ye *et al.* [17] from the perspective of linear discriminant analysis, but the convergence analysis is overly dependent on ordinary differential equation to consistently maintain good speed. Because many MCA algorithms derived from different perspectives have been reported, it is valuable to determine whether the current MCA algorithms can be extended to solve GMCA problems. Two algorithms for GMCA based on the power method [18] and the modified Oja-Xu MCA algorithm [19] have been proposed by Nguyen *et al.* [20].

What is more, frontier research associated with the parallel extraction of GMCA has been explored in terms of the algorithm and its convergence analysis properly [21]. Although these algorithms have good convergence properties, most of them cannot extract multiple GMCs in sequence and are not self-stabilizing. In addition, the lower the computational complexity is, the better the applicability of the algorithm to real applications.

Therefore, we research these issues in this paper. A novel learning algorithm is proposed to solve GMCA problems. Furthermore, the algorithm is extended to extract multiple minor components in sequence. Then, a theoretical analysis illustrates that the algorithm is self-stabilizing. The convergence conditions of the proposed algorithm for both GMC and multiple GMCs are explored through the DDT method. The simulation results illustrate that the proposed algorithm is advantageous in terms of estimation accuracy and convergence speed.

The rest of this paper is organized as follows. Section II provides some notation and preliminary knowledge. Section III presents the proposed algorithm. Then, we analyze the self-stability and the convergence for both single GMC and multiple GMCs in Sections IV, V and VI. Numerical examples in Section VII demonstrate the performance of the proposed algorithms. Section VIII concludes the work.

II. GUIDELINES FOR MANUSCRIPT PREPARATION

A. NOTATION

Bold-face capital and lowercase letters represent matrices and vectors, respectively. The transpose of a matrix or a column vector is noted by superscript $(\bullet)^T$, and the inverse of a non-singular matrix is donated by superscript $(\bullet)^{-1}$. Some notational symbols are listed as follows.

$\boldsymbol{x}$	An M-dimensional sequence
$\boldsymbol{y}$	An N-dimensional sequence
$\boldsymbol{x}(k)$	The k^{th} column vector sampled in $\boldsymbol{x}$
$\boldsymbol{y}(k)$	The k^{th} column vector sampled in $\boldsymbol{y}$
$\boldsymbol{R}$	Autocorrelation matrix
$\boldsymbol{I}$	Identity matrix
η	Learning rate
α	Forgetting factor
E	Expected operator
N	The dimensionality of the input sequence

B. FORMULATIONS RELATED TO GMCA

The goal of generalized eigen component decomposition is to search the vector $\boldsymbol{v}$ and an invariant λ as follows,

$$\boldsymbol{R}_y\boldsymbol{v} = \lambda\boldsymbol{R}_x\boldsymbol{v} \tag{1}$$

where $\boldsymbol{R}_x = E[\boldsymbol{x}\boldsymbol{x}^T]$ and $\boldsymbol{R}_y = E[\boldsymbol{y}\boldsymbol{y}^T]$ are the autocorrelation matrices of two input sources $\boldsymbol{x} \in \mathbb{R}^{n\times M}$ and $\boldsymbol{y} \in \mathbb{R}^{n\times M}$, respectively. In addition, $\boldsymbol{x} \in \mathbb{R}^{n\times M}$ and $\boldsymbol{y} \in \mathbb{R}^{n\times M}$ are stochastic progress vectors with means of zero. Without loss of generality, the vector $\boldsymbol{v}$ and invariant λ in (1) illustrate the generalized eigenvectors and eigenvalues of the matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$.

In addition to the basic concept, a definition related to the algorithm is listed as follows.

Definition 1 [21]: Given an arbitrary vector $\boldsymbol{w}$ with dimensions $n \times 1$ and an arbitrary matrix $\boldsymbol{M}$ with dimensions $n \times n$, the norm of the vector on matrix $\boldsymbol{M}$ can be defined as $\|\boldsymbol{w}\|_{\boldsymbol{M}} = \sqrt{\boldsymbol{w}^T\boldsymbol{M}\boldsymbol{w}}$.

According to the matrix analysis [11],

$$\boldsymbol{R}_y\boldsymbol{v}_i = \lambda_i\boldsymbol{R}_x\boldsymbol{v}_i \tag{2}$$

$$\boldsymbol{v}_i^T\boldsymbol{R}_x\boldsymbol{v}_j = \delta_{ij}(i, j = 1, 2, \cdots, n) \tag{3}$$

where λ_i denotes the i^{th} generalized eigenvalue that satisfies $\lambda_1 > \lambda_2 > \cdots > \lambda_n$, corresponding to the generalized eigenvector $\boldsymbol{v}_i$. In addition, δ_{ij} is the Kronecker delta. Generalized eigen component decomposition can be simplified into eigen component decomposition when $\boldsymbol{R}_x = I_n$, $\boldsymbol{R}_y\boldsymbol{v}_i = \lambda_i\boldsymbol{v}_i$.

Because it is impossible to obtain the autocorrelation matrices $\boldsymbol{R}_x$ and $\boldsymbol{R}_y$ beforehand in real time, they are estimated according to the input signals through a weighted window function [22] given as

$$\hat{\boldsymbol{R}}_y(k+1) = \alpha_1\hat{\boldsymbol{R}}_y(k) + \boldsymbol{y}(k+1)\boldsymbol{y}^T(k+1) \tag{4}$$

$$\hat{\boldsymbol{R}}_x(k+1) = \alpha_2\hat{\boldsymbol{R}}_x(k) + \boldsymbol{x}(k+1)\boldsymbol{x}^T(k+1) \tag{5}$$

where $\alpha_1, \alpha_2 \in (0, 1)$ are the forgetting factors, $\hat{\boldsymbol{R}}_x$ and $\hat{\boldsymbol{R}}_y$ denote the estimates of the matrix pencil. Since the restriction of the condition number to a certain range cannot be guaranteed, which means that the inversion of $\hat{\boldsymbol{R}}_x$ is not stable, the inverse matrix must be estimated according to lemma [12]. Let $\boldsymbol{Q}_x$ represent $\boldsymbol{R}_x^{-1}$, so

$$\boldsymbol{Q}_x(k+1) = \frac{1}{\alpha_2}\left[\boldsymbol{Q}_x(k) - \frac{\boldsymbol{Q}_x(k)\boldsymbol{x}(k+1)\boldsymbol{x}^T(k+1)\boldsymbol{Q}_x(k)}{\alpha_2 + \boldsymbol{x}^T(k+1)\boldsymbol{Q}_x(k)\boldsymbol{x}(k+1)}\right] \tag{6}$$

III. NOVEL GENERALIZED MINOR COMPONENT EXTRACTION ALGORITHM

The minor subspace learning rule presented by Douglas to extract the minor subspace [23] is self-stabilizing in that the vectors do not need to be periodically normalized to unit modulus. Based on the Douglas algorithm, we propose a novel GMCA algorithm called generalized Douglas minor component analysis (GDM),

$$\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + \eta(-(\boldsymbol{w}_k^T R_x\boldsymbol{w}_k)^2 R_x^{-1}R_y\boldsymbol{w}_k + \boldsymbol{w}_k^T R_y\boldsymbol{w}_k\boldsymbol{w}_k) \tag{7}$$

where $\boldsymbol{w}_k \in \mathbb{R}^{n\times 1}$ is the weight vector, and $\eta \in (0, 1)$.

Let $\boldsymbol{R}_x$, $\boldsymbol{R}_y$ and $\boldsymbol{R}_x^{-1}$ be replaced by $\hat{\boldsymbol{R}}_x$, $\hat{\boldsymbol{R}}_y$ and $\boldsymbol{Q}_x$ in (7), respectively. The algorithm steps are summarized in Table 1.

After initializing $\hat{\boldsymbol{R}}_x$, $\hat{\boldsymbol{R}}_y$, $\boldsymbol{Q}_x$ and $\boldsymbol{w}_0$, the algorithm regulates the weight vector $\boldsymbol{w}_k$ by repeating the same computations in (8) while $\hat{\boldsymbol{R}}_x$, $\hat{\boldsymbol{R}}_y$, and $\boldsymbol{Q}_x$ are updated until the difference of adjacent weight vectors $\boldsymbol{w}_k$ approaches zero. Then, $\boldsymbol{w}_k$ converges to the GMC of the matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$.

IV. SELF-STABILIZING ANALYSIS

Self-stability reflects the stability of the GMCA algorithm against fluctuations in the weight matrix. The self-stabilizing performance can be revealed through various vector norms. Without loss of generality, the norm of the vector $\boldsymbol{w}_k$ on matrix $\boldsymbol{R}_x$, $\|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x}$ is selected. The Self-stability means that the norm of $\boldsymbol{w}_k$ is global stable with any initial value $\boldsymbol{w}_0$.

Theorem 1: It holds that

$$\lim_{k\to\infty} \|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x} / \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} = 1$$

if the learning factor $\eta << 1$.

Proof: As shown in (7), the norm of the vector $\boldsymbol{w}_{k+1}$ on matrix $\boldsymbol{R}_x$ can be expressed by $\boldsymbol{w}_k$

$$\begin{aligned}
&\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x} \\
&= \boldsymbol{w}^T_{k+1}\boldsymbol{R}_x\boldsymbol{w}_{k+1} \\
&= \left[\boldsymbol{w}_k + \eta(-(\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k)^2\boldsymbol{R}^1_x\boldsymbol{R}_y\boldsymbol{w}_k + \boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k\boldsymbol{w}_k)\right]^T\boldsymbol{R}_x \\
&\quad \times \left[\boldsymbol{w}_k + \eta(-(\boldsymbol{w}(t)^T\boldsymbol{R}_x\boldsymbol{w}_k)^2\boldsymbol{R}^{-1}_x\boldsymbol{R}_y\boldsymbol{w}_k + \boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}(t)\boldsymbol{w}_k)\right] \\
&= \boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k + 2\eta\boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k(1-\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k) + o(\eta^2)
\end{aligned} \tag{8}$$

The ratio of two adjacent norms is

$$\begin{aligned}
&\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x}/\|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} \\
&= \frac{\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k + 2\eta\boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k(1-\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k) + o(\eta^2)}{\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k} \\
&= 1 + 2\eta\boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k(1-\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k) + o(\eta^2)/\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k
\end{aligned} \tag{9}$$

The sign of the function $2\eta\boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k(1-\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k) + o(\eta^2)/\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k$ will not change if $\eta << 1$. Therefore,

$$\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x}/\|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} \begin{cases} > 1 & \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} < 1 \\ = 1 & \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} = 1 \\ < 1 & \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} > 1 \end{cases} \tag{10}$$

(11) shows that $\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x}/\|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x}$ gradually converges to one as k increases. Eventually, the weight vector $\boldsymbol{w}_k$ is independent of the initial value $\boldsymbol{w}_0$.

V. CONVERGENCE ANALYSIS

In this section, we analyze the convergence of the GDM algorithm via DDT. According to previous research [24], the basic task of DDT is to analyze the projection of weight vectors on the GMC. To obtain the DDT system, the conditional expectation $E\left\{\boldsymbol{w}_k/\boldsymbol{w}_0, \boldsymbol{x}(i), i<k\right\}$ of both sides is simultaneously sought.

According to matrix theory [22], both $\boldsymbol{R}_x$ and $\boldsymbol{R}_y$ are positive definite symmetric matrices. The generalized eigenvectors of the matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$ are composed of a set of orthogonal bases in the space $\mathbb{R}^{n\times n}$. Without loss of generality, we choose a specific set as the default set in the following argument of this paper, where the eigenvalues of the matrix pencil are arranged in descending order as $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. Therefore, the weight vector is a linear combination in the set of orthogonal bases

TABLE 1. Adaptive generalized minor component analysis algorithm.

Step 1	Let $k=0$, and initialize $\hat{\boldsymbol{R}}_y(0), \hat{\boldsymbol{R}}_x(0), \boldsymbol{Q}_x(0)$. Set the weight matrix $\boldsymbol{w}_0$ randomly.
Step 2	Set $k=k+1$ and update estimates $\hat{\boldsymbol{R}}_y(k), \hat{\boldsymbol{R}}_x(k), \boldsymbol{Q}_x(k)$. Update the $\boldsymbol{w}_k$ as $\boldsymbol{w}_{k+1} = \boldsymbol{w}_k + \eta[-(\boldsymbol{w}^T_k\hat{\boldsymbol{R}}_x(k)\boldsymbol{w}_k)^2\boldsymbol{Q}_x(k)\hat{\boldsymbol{R}}_y(k)\boldsymbol{w}_k + \boldsymbol{w}^T_k\hat{\boldsymbol{R}}_y(k)\boldsymbol{w}_k\boldsymbol{w}_k]$ (11)
Step 3	Return to step 2.

$$\boldsymbol{w}_k = \sum_{i=1}^{n} z_i(k)\boldsymbol{v}_i \tag{12}$$

where $z_i(k) = \boldsymbol{v}^T_i\boldsymbol{R}_x\boldsymbol{w}_k$ is a scalar that represents the projection length of $\boldsymbol{w}(k)$ onto the generalized eigenvector $\boldsymbol{v}_i$. Substituting (12) into (7)

$$z_i(k+1) = \left\{1 + \eta(-\lambda_i(\boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w})^2 + \boldsymbol{w}^T\boldsymbol{R}_y\boldsymbol{w})\right\} z_i(k) \tag{13}$$

where $k \geq 0$. Meanwhile, the Rayleigh quotient should be noted [25],

$$0 < \lambda_n\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k \leq \boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k \leq \lambda_1\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k \tag{14}$$

The details of the analysis are as follows.

Theorem 2: For $k \geq 0$, if $\boldsymbol{w}^T_0\boldsymbol{R}_x\boldsymbol{v}_n \neq 0$ and $0 \leq \|\boldsymbol{w}_0\|_{\boldsymbol{R}_x} \leq 1$, then $\|\boldsymbol{w}_k\|_{\boldsymbol{R}_x} < 1 + \eta\lambda_1$.

Proof: Due to (12) and (13),

$$\begin{aligned}
\|\boldsymbol{w}_{k+1}\|^2_{R_x} &= \sum_{i=1}^{n} z_i^2(k+1) \\
&= \sum_{i=1}^{n}\left[1 + \eta(-\lambda_i(\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k)^2 + \boldsymbol{w}^T_k\boldsymbol{R}_y\boldsymbol{w}_k)\right]^2 z_i^2(k) \\
&< \left[1 + \eta\lambda_1\boldsymbol{w}^T_k\boldsymbol{R}_x\boldsymbol{w}_k\right]^2 \sum_{i=1}^{n} z_i^2(k) \\
&= \left[1 + \eta\lambda_1\|\boldsymbol{w}_k\|^2_{R_x}\right]^2\|\boldsymbol{w}_k\|^2_{R_x}
\end{aligned} \tag{15}$$

For simplicity, assuming $s = \|\boldsymbol{w}_k\|^2_{R_x}$, we define a differentiable function

$$f(s) = (1+\eta\lambda_1 s)^2 s \tag{16}$$

where $s \in [0, 1]$, and the differential form of (16) is

$$\dot{f}(s) = (1+\eta\lambda_1 s)(1+3\eta\lambda_1 s) \tag{17}$$

The zero points of $\dot{f}(s)$ are $s_1 = -1/\eta\lambda_1$ and $s_2 = -3/\eta\lambda_1$. Since $\eta > 0, \lambda_1 > 0$, then $s_1 < s_2 < 0$. As a consequence, $\dot{f}(s) > 0$ for $s \in [0, 1]$, which means that $f(s)$ is a monotonically increasing function. Therefore,

$$f(s) \leq f(1) = (1+\eta\lambda_1)^2 \tag{18}$$

This result illustrates that $\|\boldsymbol{w}_k\|_{R_x} < 1 + \eta\lambda_1$ for $k \in N^*$.

Theorem 3: Suppose that $c = \left[1 - \eta\lambda_1 (1 + \eta\lambda_1)^4\right]$. For $k \geq 0$, if $\|\boldsymbol{w}_k\|_{\boldsymbol{R}_x} < 1 + \eta\lambda_1$ and $\eta\lambda_1 < 0.2$, then $\|\boldsymbol{w}_k\|_{\boldsymbol{R}_x} > c^k \|\boldsymbol{w}_0\|$.

Proof: According to (12) and (13),

$$\begin{aligned}
&\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x} \\
&= \sum_{i=1}^{n} z_i^2(k+1) \\
&= \sum_{i=1}^{n} \left[1 + \eta(-\lambda_i(\boldsymbol{w}_{k+1}^T\boldsymbol{R}_x\boldsymbol{w}_k)^2 + \boldsymbol{w}_k^T\boldsymbol{R}_y\boldsymbol{w}_k)\right]^2 z_i^2(k) \\
&> \left[1 - \eta\lambda_1 \left[\boldsymbol{w}_{k+1}^T\boldsymbol{R}_x\boldsymbol{w}_k\right]^2\right]^2 \sum_{i=1}^{n} z_i^2(k) \\
&= \left[1 - \eta\lambda_1 \|\boldsymbol{w}_k\|^4_{\boldsymbol{R}_x}\right]^2 \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} \qquad (19)
\end{aligned}$$

Since $\|\boldsymbol{w}_k\|_{\boldsymbol{R}_x} < 1 + \eta\lambda_1$, then

$$\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x} \geq \left[1 - \eta\lambda_1(1 + \eta\lambda_1)^4\right]^2 \|\boldsymbol{w}_k\|^2_{\boldsymbol{R}_x} \qquad (20)$$

where $c = \left[1 - \eta\lambda_1 (1 + \eta\lambda_1)^4\right]$, and $c = \left[1 - \eta\lambda_1 (1 + \eta\lambda_1)^4\right] > \left[1 - 0.2 \times (1 + 0.2)^4\right] = 0.58528 > 0$ when $\eta\lambda_1 < 0.2$. As a result,

$$\|\boldsymbol{w}_{k+1}\|^2_{\boldsymbol{R}_x} > c^2 \|\boldsymbol{w}_{k-1}\|^2_{\boldsymbol{R}_x} > \cdots > c^{2k} \|\boldsymbol{w}_0\|^2_{\boldsymbol{R}_x} \qquad (21)$$

which illustrates that $\|\boldsymbol{w}_k\|_{\boldsymbol{R}_x} > c^k \|\boldsymbol{w}_0\|$.

Theorem 4: For $k \geq 0$, if $\boldsymbol{w}_0^T\boldsymbol{R}_x\boldsymbol{v}_n \neq 0$ and $\eta\lambda_1 < 0.2$, then $1 + \eta\boldsymbol{w}^T\boldsymbol{R}_y\boldsymbol{w} - \eta\lambda_i(\boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w})^2 > 0$, which is the proportion of $z_i(k)$ noted in (13).

Proof: As shown in (13),

$$\begin{aligned}
1 &+ \eta\boldsymbol{w}^T\boldsymbol{R}_y\boldsymbol{w} - \eta\lambda_i(\boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w})^2 \\
&> 1 - \eta\lambda_1(\boldsymbol{w}^T\boldsymbol{R}_x\boldsymbol{w})^2 \\
&\geq 1 - \eta\lambda_1(1 + \eta\lambda_1)^4 \\
&\geq 1 - 0.2 \times (1 + 0.2)^4 \\
&= 0.58528 > 0 \qquad (22)
\end{aligned}$$

where $\eta\lambda_1 \leq 0.2$.

From (12), each $\boldsymbol{w}_k$ can be expressed by a linear combination of the eigenvectors, which is rewritten as

$$\boldsymbol{w}_k = \sum_{i=1}^{n-1} z_i(k)\boldsymbol{v}_i + z_n(k)\boldsymbol{v}_n \qquad (23)$$

where in the k^{th} iteration, $z_i(k)$ is a combination factor corresponding to its eigenvector. Then, the convergence analysis of the GDM algorithm is transferred from vectors $\boldsymbol{w}_k$ to scalars $z_i(k)$. Clearly, if each sequence $\{z_i(k)\}$, $i = 1, 2, \ldots, N-1$ converges to a constant, $\boldsymbol{w}_k$ will approach the MC. This assumption is proved by Theorems 5 and 6.

Theorem 5: For $i = 1, 2, \cdots, n-1$, $\lim_{k\to\infty} z_i(k) = 0$ if $\left[1 - \eta\lambda_1 (1 + \eta\lambda_1)^4\right] < \|\boldsymbol{w}(k)\|_{\boldsymbol{R}_x} < 1 + \eta\lambda_1$ and $1 + \eta\boldsymbol{w}^T\boldsymbol{R}_y\boldsymbol{w} - \eta\lambda_i(\boldsymbol{w}^T\boldsymbol{R}_x\,\boldsymbol{w})^2 > 0$.

Proof: For $k \geq 0$,

$$\begin{aligned}
&\left[\frac{z_i(k+1)}{z_n(k+1)}\right]^2 \\
&= \left[\frac{1 + \eta\boldsymbol{w}_k^T\boldsymbol{R}_y\boldsymbol{w}_k - \eta\lambda_i(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\boldsymbol{w}_k^T\boldsymbol{R}_y\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}\right]^2 \frac{z_i^2(k)}{z_n^2(k)} \\
&= \left[1 - \frac{\eta(\lambda_i - \lambda_n)(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\boldsymbol{w}_k^T\boldsymbol{R}_y\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}\right]^2 \frac{z_i^2(k)}{z_n^2(k)} \\
&\leq \left[1 - \frac{\eta(\lambda_i - \lambda_n)(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}(k))^2}{1 + \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}\right]^2 \frac{z_i^2(k)}{z_n^2(k)} \qquad (24)
\end{aligned}$$

Let

$$\delta_k = \frac{\eta(\lambda_i - \lambda_n)(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}$$

and

$$\begin{aligned}
\theta_k &= \left[1 - \frac{\eta(\lambda_i - \lambda_n)(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}\right]^2 \\
&= (1 - \delta_k^2)^2.
\end{aligned}$$

Then, rewrite (25) as

$$\begin{aligned}
\left[\frac{z_i(k+1)}{z_n(k+1)}\right]^2 &\leq \left[\frac{z_i(k)}{z_n(k)}\right]^2 \theta_k \leq \left[\frac{z_i(k-1)}{z_n(k-1)}\right]^2 \theta_k\theta_{k-1} \\
&\leq \cdots \leq \frac{z_i^2(0)}{z_n^2(0)} \prod_{j=0}^{k} \theta_k \leq \frac{z_i^2(0)}{z_n^2(0)} \theta^{k+1} \qquad (25)
\end{aligned}$$

where $\theta = \max(\theta_0, \theta_1, \cdots \theta_k, \cdots)$. As long as $z_n(k)$ is bounded and $0 < \theta < 1$, then

$$\lim_{k\to\infty} z_i(k) = 0, i = 1, 2, \cdots, n-1 \qquad (26)$$

Because $\{z_i(k)|i = 1, 2, \ldots, n\}$ is bounded, $z_n(k)$ is bounded. If $0 < \theta_k < 1$, then $0 < \theta < 1$. According to the definition of θ_k, if $0 < \delta_k < 1$, then $0 < \theta_k < 1$. Given that

$$\begin{aligned}
1 &+ \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2 \\
&> 1 - \eta\lambda_1(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2 \\
&> 1 - \eta\lambda_1(1 + \eta\lambda_1)^4 \\
&\geq 1 - 0.2 \times (1 + 0.2)^4 \\
&= 0.5853 > 0 \qquad (27)
\end{aligned}$$

and $\lambda_1 > \lambda_n$. Moreover, $(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2 > 0$ and $\delta_k > 0$. Given that

$$\begin{aligned}
\delta_k &= \frac{\eta(\lambda_i - \lambda_n)(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_n(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2} \\
&< \frac{\eta\lambda_1(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2}{1 + \eta\lambda_1\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k - \eta\lambda_1(\boldsymbol{w}_k^T\boldsymbol{R}_x\boldsymbol{w}_k)^2} \\
&< \frac{\eta\lambda_1(1 + \eta\lambda_1)^4}{1 - \eta\lambda_1(1 + \eta\lambda_1)^4} \\
&\leq \frac{0.2 \times (1 + 0.2)^4}{1 - 0.2 \times (1 + 0.2)^4} \\
&= 0.7086 < 1 \qquad (28)
\end{aligned}$$

then $\delta_k < 1$. Therefore, $\lim_{k\to\infty} z_i(k) = 0, i = 1, 2, \cdots, n-1$ is true.

Theorem 6: If $\eta\lambda_1 \leq 0.2$, $\boldsymbol{w}_0^T \boldsymbol{R}_x \boldsymbol{v}_n \neq 0$ and $0 \leq \|\boldsymbol{w}_0\|_{\boldsymbol{R}_x} \leq 1$, it holds that $\lim_{k\to\infty} z_n(k) = \pm 1$.

Proof: According to Theorem 4, there exist k_0, and $\boldsymbol{w}_{k_0}$ converges to the direction of the MC $\boldsymbol{v}_n$, namely, $\boldsymbol{w}_{k_0} = z_n(k_0)\boldsymbol{v}_n$. Substituting $\boldsymbol{w}_{k_0} = z_n(k_0)\boldsymbol{v}_n$into (13) gives

$$z_n(k+1) = \left[1 + \eta\lambda_n z_n^2(k) - \eta\lambda_n z_n^4(k)\right] z_n(k) \tag{29}$$

One is subtracted from both sides of (30) to obtain

$$\begin{aligned} z_n(k+1) - 1 &= (1 + \eta\lambda_n z_n^2(k) - \eta\lambda_n z_n^4(k)) z_n(k) - 1 \\ &= \left[1 - \eta\lambda_n z_n^2(k)(z_n(k)-1)(z_n(k)+1)\right] z_n(k) - 1 \\ &= (z_n(k)-1)\left[1 - \eta\lambda_n z_n^3(k)(z_n(k)+1)\right] \end{aligned} \tag{30}$$

Define $\beta = 1 - \eta\lambda_n z_n^3(k)(z_n(k)+1)$. For $k > k_0$,

$$\begin{aligned} \beta &= (z_n(k+1)-1)/(z_n(k)-1) \\ &= 1 - \eta\lambda_n z_n^3(k)[z_n(k)+1] \\ &> 1 - \eta\lambda_1(1+\eta\lambda_1)^3(1+\eta\lambda_1+1) \\ &\geq 1 - 0.2 \times (1+0.2)^3(2+0.2) \\ &= 0.2397 > 0 \end{aligned} \tag{31}$$

where $z_n(k) \leq \|\boldsymbol{w_k}\|_B \leq 1 + \eta\lambda_1$.

$$\begin{aligned} |z_n(k+1)-1| &\leq |z_n(k)-1|\,\beta \leq \cdots \leq |z_n(0)-1|\,\beta^{k+1} \\ &\leq |z_n(0)-1|\,(k+1)e^{-\alpha(k+1)} \end{aligned} \tag{32}$$

where $\alpha = -\ln\beta$. Therefore, for any $\varepsilon > 0$, there exists $K > 1$ in

$$\frac{\Pi K e^{-\alpha K}}{(1-e^{-\alpha})^2} \leq \varepsilon \tag{33}$$

where $\Pi = \eta\lambda_1(1+\eta\lambda_1)^2(2+\eta\lambda_1)\,|z_n(0)-1|$.

To analyze the change of $z_n(k)$ when $k > K$, we arbitrarily choose k_1 and k_2, where $k_1 > k_2 > K$. Then,

$$\begin{aligned} |z_n(k_1) - z_n(k_2)| &= \left|\sum_{r=k_2}^{k_1-1}[z_n(r+1) - z_n(r)]\right| \\ &\leq \left|\sum_{r=k_2}^{k_1-1}\left[\eta\lambda_n z_n^2(r) - \eta\lambda_n z_n^4(r)\right]\right| \\ &\leq \sum_{r=k_2}^{k_1-1}\left|\eta\lambda_n z_n^2(r)\right|\,|z_n(r)+1|\,|z_n(r)-1| \\ &\leq \eta\lambda_1(1+\eta\lambda_1)^2(2+\eta\lambda_1)\sum_{r=k_2}^{k_1}|z_n(r)-1| \\ &\leq \Pi\sum_{r=k_2}^{k_1} r e^{-\alpha r} \\ &\leq \Pi\sum_{r=K}^{+\infty} r e^{-\alpha r} \\ &\leq \Pi K e^{-\alpha K}\sum_{r=0}^{+\infty} r\left(e^{-\alpha}\right)^{r-1} \\ &\leq \frac{\Pi K e^{-\alpha K}}{(1-e^{-\alpha})^2} \\ &\leq \varepsilon \end{aligned} \tag{34}$$

According to the definition of a Cauchy series, the series $\{z_n(k)\}$ is convergent. There must exist a constant z^*, where $\lim_{k\to\infty} z_n(k) = z^*$ and $\lim_{k\to\infty}\boldsymbol{w}_k = z^*\boldsymbol{v}_n$. Furthermore, according to Theorem 1, it holds that $\lim_{k\to\infty}\|\boldsymbol{w}_{k+1}\|_{\boldsymbol{R_x}}/\|\boldsymbol{w}_k\|_{\boldsymbol{R_x}} = 1$. Therefore,

$$1 + \eta\lambda_n(z^*)^2 - \eta\lambda_n(z^*)^4 = 0 \tag{35}$$

where $z^* = \lim_{k\to\infty} z_n(k) = \pm 1$.

Remark 1: Theorems 5 and 6 depict the variation of $z_i(k)$ for $i \in \{1, 2, \cdots, n\}$. As a combination of these two conclusions,

$$\begin{aligned} \lim_{k\to\infty}\boldsymbol{w}_k &= \lim_{k\to\infty}(z_n(k)\boldsymbol{v}_n + \sum\nolimits_{i=1}^{n-1} z_i(k)\boldsymbol{v}_i) \\ &= \lim_{k\to\infty} z_n(k)\boldsymbol{v}_n + \sum\nolimits_{i=1}^{n-1}(\boldsymbol{v}_i \lim_{k\to\infty} z_i(k)). \\ &= \pm\boldsymbol{v}_n \end{aligned} \tag{36}$$

Remark 2: According to the complete convergence analysis from Theorems 2 to 6, the GDM algorithm is convergent, and the basic conditions are $\boldsymbol{w}_0^T\boldsymbol{R}_x\boldsymbol{v}_n \neq 0$, $0 \leq \|\boldsymbol{w}_0\|_{\boldsymbol{R}_x} \leq 1$, $\eta\lambda_1 \leq 0.2$. For $\eta\lambda_1 \leq 0.2$, the learning factor is associated with the greatest eigenvalue. The upper limit of λ_1 can be estimated in real applications [24]. In addition, the initial weight vector $\boldsymbol{w}_0$, which satisfies $\boldsymbol{w}_0^T\boldsymbol{R}_x\boldsymbol{v}_n \neq 0$ and $0 \leq \|\boldsymbol{w}_0\|_{\boldsymbol{R}_x} \leq 1$, is generated randomly. In fact, the alternative range of $\|\boldsymbol{w}_0\|_{\boldsymbol{R}_x}$ can be larger when $\eta\lambda_1$ is sufficiently small.

VI. MULTIPLE GENERALIZED MINOR COMPONENT EXTRACTION

In this section, we extend the GDM algorithm to extract multiple GMCs. GMCs can be extracted in sequence according to the analysis of matrix disturbance [12]. In each extraction process, the extracted GMC corresponds to the smallest generalized eigenvalue of the autocorrelation matrix. Meanwhile, before the extraction, the autocorrelation matrix is modified by a designed matrix disturbance. The convergence of the extended algorithm is also analyzed via DDT.

We can obtain a DDT system by using the conditional expectation operator $E\{w_{k+1}/w_0, x(i), i < k\}$ on both sides of the Eq. (7) as

$$\begin{aligned} \boldsymbol{w}_{j,k+1} = \boldsymbol{w}_{j,k} + \eta\Big[&\boldsymbol{w}_{j,k}^T(\boldsymbol{R}_y + \boldsymbol{M}_{j,k})\boldsymbol{w}_{j,k}\boldsymbol{w}_{j,k} \\ &- (\boldsymbol{w}_{j,k}^T\boldsymbol{R}_x\boldsymbol{w}_{j,k})^2\boldsymbol{R}_x^{-1}(\boldsymbol{R}_y + \boldsymbol{M}_{j,k})\boldsymbol{w}_{j,k}\Big] \end{aligned} \tag{37}$$

where n_{mc} is the minimum of the extracted MCs and $j = 1, 2, \cdots, n_{mc}$. $\boldsymbol{M}_{j,k}$ is the matrix disturbance produced by the result of the $(j-1)^{th}$ extraction, and $\boldsymbol{M}_{1,k}=0$.

$$\boldsymbol{M}_{j,k}=\boldsymbol{M}_{j-1,k}+\tau\frac{\boldsymbol{R}_x\boldsymbol{w}_{j-1,k}\boldsymbol{w}_{j-1,k}^T\boldsymbol{R}_x}{\boldsymbol{w}_{j-1,k}^T\boldsymbol{R}_x\boldsymbol{w}_{j-1,k}}=\tau\sum_{i=1}^{j-1}\frac{\boldsymbol{R}_x\boldsymbol{w}_{i,k}\boldsymbol{w}_{i,k}^T\boldsymbol{R}_x}{\boldsymbol{w}_{i,k}^T\boldsymbol{R}_x\boldsymbol{w}_{i,k}}.$$

τ is a scalar that is larger than any other eigenvalue, namely, $\tau>\lambda_1$. The convergence is analyzed through the following theorem.

Theorem 7: If $\eta[\lambda_1+(n_{mc}-1)\tau]\le 0.2$ and $\boldsymbol{w}_{j,0}^T\boldsymbol{R}_x\boldsymbol{v}_{n-j+1}\neq 0$, it holds that for $j=1,2,\cdots,n_{mc}$, $\lim_{k\to\infty}\boldsymbol{w}_{j,k}=\pm\boldsymbol{v}_{n-j+1}$.

Proof: For $k>0$, let $\lambda_{i,j}$ denote the eigenvalue of the symmetric non-positive $\boldsymbol{R}_y+\boldsymbol{M}_{j,k}$ with

$$\lambda_{1,j}>\lambda_{2,j}>\cdots\lambda_{n,j}\ge 0 \tag{38}$$

Given that $M_j(k)=0$, $\lambda_{1,1}=\lambda_1$, where λ_1 is the largest eigenvalue of the autocorrelation matrix $\boldsymbol{R}_x$. According to the update equation of $\boldsymbol{M}_{j,k}$,

$$\boldsymbol{R}_y+\boldsymbol{M}_{j+1,k}=\boldsymbol{R}_y+\boldsymbol{M}_{j,k}+\tau\frac{\boldsymbol{R}_x\boldsymbol{w}_{j,k}\boldsymbol{w}_{j,k}^T\boldsymbol{R}_x}{\boldsymbol{w}_{j,k}^T\boldsymbol{R}_x\boldsymbol{w}_{j,k}} \tag{39}$$

According to [11],

$$\lambda_{1,j+1}\ge\lambda_{1,j}\ge\lambda_{2,j+1}\ge\cdots\lambda_{n,j+1}\ge\lambda_{n,j} \tag{40}$$

and

$$\sum_{i=1}^{n}\lambda_{i,j+1}-\sum_{i=1}^{n}\lambda_{i,j}=\tau \tag{41}$$

Therefore, $\lambda_{1,j+1}\ge\lambda_{1,j}$ and $\lambda_{1,j+1}\le\lambda_{1,j}+\tau$. If $j=n_{mc}$,

$$\lambda_{1,q}\ge\lambda_{1,n_{mc}-1}\ge\cdots\lambda_{n,2}\ge\lambda_{n,1}=\lambda_1 \tag{42}$$

and

$$\lambda_{1,n_{mc}}\le\lambda_{1,n_{mc}-1}+\tau\le\cdots\le\lambda_{n,1}+(n_{mc}-1)\tau=\lambda_1+(n_{mc}-1)\tau \tag{43}$$

Then, it holds that

$$\eta<1/\{5[\lambda_1+(n_{mc}-1)\tau]\}\le 1/(5\lambda_{1,n_{mc}})\le\cdots\le 1/(5\lambda_{1,1})=1/(5\lambda_1) \tag{44}$$

if $\eta[\lambda_1+(n_{mc}-1)\tau]\le 0.2$.

According to Theorems 2-6, $\boldsymbol{w}_{j,k}$ converges to eigenvectors corresponding to the smallest eigenvalues for $j=1,2,\cdots,n_{mc}$ conditional on (44). If so, $\lim_{k\to\infty}\boldsymbol{w}_{1,k}=\pm v_n$ when $j=1$.

Assuming that $j=2$,

$$\lim_{k\to\infty}\boldsymbol{R}_y+\boldsymbol{M}_{2,k}=\sum_{i=1}^{n}\lambda_i\boldsymbol{R}_x\boldsymbol{v}_i\boldsymbol{v}_i^T+\tau\boldsymbol{R}_x\boldsymbol{v}_n\boldsymbol{v}_n^T=\sum_{i=1}^{n-1}\lambda_i\boldsymbol{R}_x\boldsymbol{v}_i\boldsymbol{v}_i^T+(\tau+\lambda_n)\boldsymbol{R}_x\boldsymbol{v}_n\boldsymbol{v}_n^T. \tag{45}$$

Since $\tau+\lambda_n\ge\lambda_1$, repeat steps (39) - (44); then, $\lim_{k\to\infty}\boldsymbol{w}_{2,k}=\pm\boldsymbol{v}_{n-1}$. Therefore, for $j=1,2,\cdots,n_{mc}$, it holds that

$$\lim_{k\to\infty}\boldsymbol{w}_{j,k}=\pm v_{n+1-j} \tag{46}$$

VII. SIMULATION EXPERIMENTS

In this section, we use three examples to evaluate the effectiveness of the GDM algorithm. The first example illustrates the convergence of the algorithm. The second experiment demonstrates the self-stability of the GDM algorithm. The third experiment verifies the convergence of the sequential GMCA algorithm. In all the examples, the results are the average of 100 independent runs in MATLAB.

A. EXAMPLE 1: CONVERGENCE

In this section, an example is designed to depict the conclusions drawn from Theorems 2-6. We examine the varying curve of the projection length $z_i(k)$, which means that if $z_i(k)$ converges to zero for $i=1,2,\cdots,n-1$ and $z_n(k)$ converges to unity modulus, then the GDM algorithm is convergent. The convergence speeds of three algorithms, i.e., a the GDM algorithm , GMOX algorithm [19] proposed by Nguyen in 2013, and Ye algorithm [17] proposed by M. Ye in 2008, are compared.

A randomly generated matrix pencil is given in (47) and (48), as shown at the bottom of the next page. The generalized eigenvalues of the matrix pencil $(\boldsymbol{R}_y,\boldsymbol{R}_x)$ are $\lambda_1=2.7015$, $\lambda_2=2.2032$, $\lambda_3=0.6181$, $\lambda_4=0.5673$, $\lambda_5=0.4561$, and $\lambda_6=0.2028$. We use the three algorithms to extract the GMC of the matrix pencil.

Without loss of generality, the learning factor $\eta=0.020<0.2/\lambda_1$. According to Theorem 6, the learning factor is in the domain of the convergence condition. In addition, the initial weight vectors are as follows, (49) and (50), as shown at the bottom of the next page.

In Fig. 1 and 2, the convergence curves of the GDM algorithm are solid lines, those of the GMOX algorithm are dotted lines and the Ye algorithm are short-dashed lines. z_i, $i=1,2,\cdots,5$ converges to zero rapidly while z_6 converges to

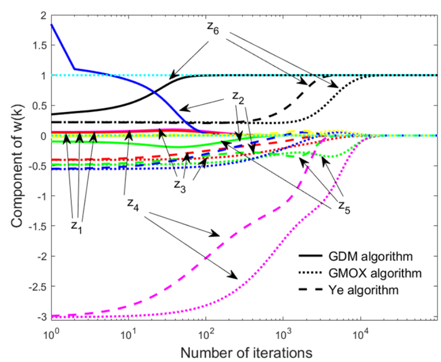


FIGURE 1. Convergence curves of $z_i, i=1,2,\cdots,6$ for $w_0=w_a$.

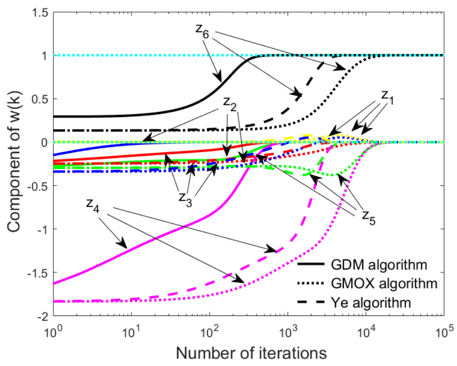


FIGURE 2. Convergence curves of $z_i, i = 1, 2, \cdots, 6$ for $w_0 = w_b$.

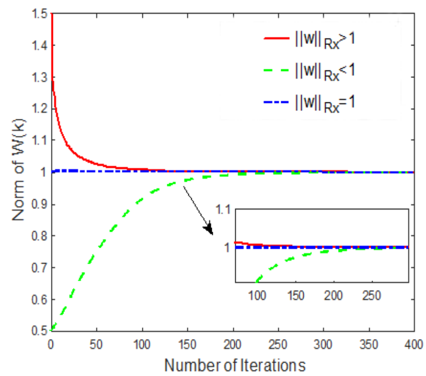


FIGURE 3. Norm curves of the GDM algorithm in three cases.

unity modulus. In addition, iteration number of the GDM algorithm when z_6 converges to unity modulus is less than that of the GMOX and Ye algorithms. The convergence of z_6 is faster in Fig. 1 than that in Fig. 2. These results illustrate three facts. First, the GDM algorithm is convergent if the learning rate satisfies the conditions in Theorems 2-5. Second, regardless of the initial weight, the GDM algorithm performs better than the GMOX and Ye algorithms in terms of convergence speed. Third, for a larger modulus of the initial weight vector $\boldsymbol{w}_a$, the GDM algorithm has better convergence speed than that for $\boldsymbol{w}_b$.

B. EXAMPLE 2: SELF-STABILITY

In this section, an example is evaluated to verify whether the $\lim_{k\to\infty} \|\boldsymbol{w}_k\|_{R_x} = 1$ holds for different moduli of the initial weight vector. With the same matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$ and learning rate as in Example 1, three randomly generated vectors, whose moduli are less then, equal to and greater than one, are selected.

Fig. 3 shows that the norm of $\boldsymbol{w}_k$ converges to a constant regardless of the initial weight vector. According to this simulation experiment, we can conclude that the GDM algorithm is self-stabilizing.

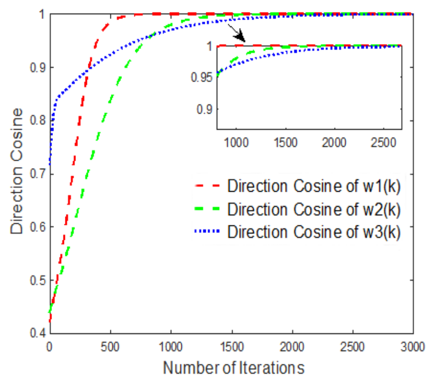


FIGURE 4. DC curves of the sequential GDM algorithm.

C. EXAMPLE 3: MULTIPLE EXTRACTION IN SEQUENCE

In this section, we use the matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$ in (47) and (48) to extract the three MCs in sequence. According to algorithm (37), let $\tau = 3 > \lambda_1$ and let the learning factor $\eta = 0.02 < 0.2/\lambda_1 - (n_{mc} - 1)\tau$, where $n_{mc} = 3$. The initial weight vector $\boldsymbol{w}_0 = \boldsymbol{w}_a$. To measure the convergence speed and accuracy of these algorithms, we calculate the direction

$$\boldsymbol{R}_y = \begin{bmatrix} 0.0614 & 0.0174 & -0.0127 & -0.0049 & 0.0110 & -0.0294 \\ 0.0174 & 0.0661 & -0.0189 & 0.0052 & 0.0131 & -0.0226 \\ -0.0127 & -0.0189 & 0.0578 & 0.0008 & -0.0169 & 0.0317 \\ -0.0049 & 0.0052 & 0.0008 & 0.0473 & 0.0024 & -0.0023 \\ 0.0110 & 0.0131 & -0.0169 & 0.0024 & 0.0506 & -0.0082 \\ -0.0294 & -0.0226 & 0.0317 & -0.0023 & -0.0082 & 0.0471 \end{bmatrix} \tag{47}$$

$$\boldsymbol{R}_x = \begin{bmatrix} 0.0644 & 0.0092 & 0.0027 & 0.0244 & -0.0128 & -0.0126 \\ 0.0092 & 0.0681 & -0.0160 & -0.0004 & -0.0013 & 0.0239 \\ 0.0027 & -0.0160 & 0.0803 & -0.0192 & -0.0034 & 0.0161 \\ 0.0244 & -0.0004 & -0.0192 & 0.0675 & 0.0087 & 0.0082 \\ -0.0128 & -0.0013 & -0.0034 & 0.0087 & 0.0690 & 0.0287 \\ -0.0126 & 0.0239 & 0.0161 & 0.0082 & 0.0287 & 0.0715 \end{bmatrix} \tag{48}$$

$$\boldsymbol{w}_a = \begin{bmatrix} -5.8744 & 4.6453 & -3.8261 & -2.0116 & -8.0562 & -2.3446 \end{bmatrix}^T \tag{49}$$

$$\boldsymbol{w}_b = \begin{bmatrix} -3.5757 & 2.8276 & -2.3289 & -1.2244 & -4.9038 & -1.4272 \end{bmatrix}^T \tag{50}$$

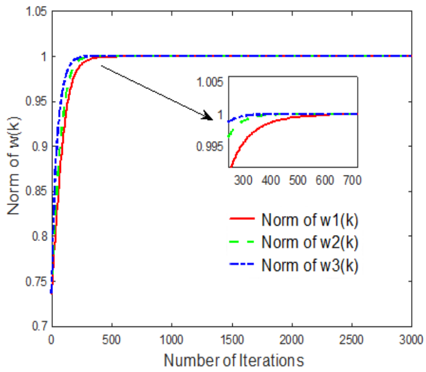


FIGURE 5. Norm curves of the sequential GDM algorithm.

cosine for each iteration, which is given by

$$Direction\ Cosine(k) = \left| \boldsymbol{w}_{i,k}^{T} v_{6-i} \right| / \left\| \boldsymbol{w}_{i,k}^{T} \right\| \left\| v_{6-i} \right\| \ (i = 1, 2, 3) \tag{51}$$

where v_{6-i} is the true eigenvector associated with the eigenvalue of the autocorrelation matrix pencil $(\boldsymbol{R}_y, \boldsymbol{R}_x)$. Fig. 4 shows that all the direction cosines of $\boldsymbol{w}_{i,k}^{T}$ converge to one, which means $\boldsymbol{w}_{i,k}^{T}$ converges to the corresponding MC. Fig. 5 shows that each norm converges to one, which is consistent with Theorem 7, so all three MCs are self-stabilizing.

VIII. CONCLUSION

In this paper, a novel algorithm was proposed for GMCA. The convergence analysis was accomplished through DDT method. Then the algorithm was proved to be self-stabilizing. A sequential GMCs extraction algorithm was also derived from the algorithm. Simulation results illustrate that the GDM algorithm was advantageous in estimation accuracy and convergence speed compared with some other algorithms.

REFERENCES

[1] X. Y. Kong, C. H. Hu, Z. S. Duan, *Principal Component Analysis Networks and Algorithms*. Singapore: Springer, 2017, pp. 185–231.

[2] D. Feng, Z. Bao, and L.-C. Jiao, "Total least mean squares algorithm," *IEEE Trans. Signal Process.*, vol. SP-46, no. 8, pp. 2122–2130, Aug. 1998.

[3] R. Klemm, "Adaptive airborne MTI: An auxiliary channel approach," *IEE Proc. F-Commun., Radar Signal Process.*, vol. 134, no. 3, pp. 269–276, Jan. 1987.

[4] S. Barbarossa and E. D'Addio, and G. Galati, "Comparison of optimum and linear prediction techniques for clutter cancellation," *IEE Proc. F-Commun., Radar Signal Process.*, vol. 134, no. 3, pp. 277–282, Jan. 1987.

[5] G. Mathew and V. U. Reddy, "Development and analysis of a neural network approach to Pisarenko's harmonic retrieval method," *IEEE Trans. Signal Process.*, vol. 42, no. 3, pp. 663–667, Mar. 1994.

[6] J. V. Stone, "Blind deconvolution using temporal predictability," *Neurocomputing*, vol. 49, nos. 1–4, pp. 79–86, 2002.

[7] O. Mangasarian and E. Wild, "Multisurface proximal support vector machine classification via generalized eigenvalues," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 1, pp. 69–74, Jan. 2006.

[8] H. Chen, T. Sarkar, S. Dianat, and J. Brule, "Adaptive spectral estimation by the conjugate gradient method," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-34, no. 2, pp. 272–284, Apr. 1986.

[9] S. Shahbazpanahi, A. B. Gershman, Z.-Q. Luo, and K. M. Wong, "Robust adaptive beamforming for general-rank signal models," *IEEE Trans. Signal Process.*, vol. 51, no. 9, pp. 2257–2269, Sep. 2003.

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

[11] W. Shougen and Z. Shuqin, "An algorithm for Ax=λBx with symmetric and positive-definite A and B," *SIAM J. Matrix Anal. Appl.*, vol. 12, no. 4, pp. 654–660, 1991.

[12] G. H. Golub and C. F. Van Loan, *Matrix Computations* 4th ed. Baltimore, MD, USA: Johns Hopkins Univ. Press, 2013.

[13] L. Xu, E. Oja, and C. Y. Suen, "Modified Hebbian learning for curve and surface fitting," *Neural Network.*, vol. 5, no. 3, pp. 441–457, 1992.

[14] S. Ouyang, Z. Bao, G.-S. Liao, and P. C. Ching, "Adaptive minor component extraction with modular structure," *IEEE Trans. Signal Process.*, vol. 49, no. 9, pp. 2127–2137, Sep. 2001.

[15] D.-Z. Feng, W.-X. Zheng, and Y. Jia, "Neural network learning algorithms for tracking minor subspace in high-dimensional data stream," *IEEE Trans. Neural Netw.*, vol. 16, no. 3, pp. 513–521, May 2005.

[16] X. Kong, C. Hu, and C. Han, "A dual purpose principal and minor subspace gradient flow," *IEEE Trans. Signal Process.*, vol. 60, no. 1, pp. 197–210, Jan. 2012.

[17] M. Ye, Y. Liu, H. Wu, and Q. Liu, "A few online algorithms for extracting minor generalized eigenvectors," in *Proc. IEEE Int. Joint Conf. Neural Network.*, Jun. 2008, pp. 1714–1720.

[18] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford, U.K.: Clarendon, 1965.

[19] D. Peng and Y. Zhang, "A modified Oja–Xu MCA learning algorithm and its convergence analysis," *IEEE Trans. Circuits Syst., II, Exp. Briefs*, vol. 54, no. 4, pp. 348–352, Apr. 2007.

[20] T. D. Nguyen, N. Takahashi, and I. Yamada, "An adaptive extraction of generalized eigensubspace by using exact nested orthogonal complement structure," *Multidimensional Syst. Signal Process.*, vol. 24, no. 3, pp. 457–483, 2013.

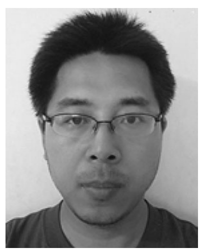
[21] G. Yingbin, K. Xiangyu, H. Changhua, L. Hongzeng, and H. Li'an, "A generalized information criterion for generalized minor component extraction," *IEEE Trans. Signal Process.*, vol. 65, no. 4, pp. 947–959, Feb. 2017.

[22] T. Tanaka, "Fast generalized eigenvector tracking based on the power method," *IEEE Signal Process. Lett.*, vol. 16, no. 11, pp. 969–972, Nov. 2009.

[23] S. C. Douglas, S.-Y. Kong, and S. C. Amari, "A self-stabilized minor subspace rule," *IEEE Signal Process Lett.*, vol. 5, no. 12, pp. 230–328, Dec. 1998.

[24] D. Peng and Z. Yi, "Convergence analysis of a deterministic discrete time system of feng's MCA learning algorithm," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3626–3632, Sep. 2006.

[25] R. A. Horn, C. R. Johnson, *Matrix Analysis*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 2012.

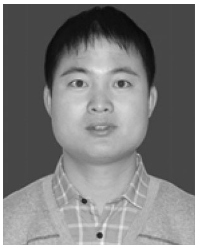


HONGZENG LI was born in 1979. He received the bachelor's, master's, and Ph.D. degrees from Northwestern Polytechnical University, Xi'an, China, in 2002, 2005, and 2008, respectively, all in control science and engineering. He is currently a Lecturer with the Xi'an Research Institute of High Technology, Xi'an, China. His research interests include fault diagnosis and fault tolerant control.

BOYANG DU was born in 1990. He received the B.S. and M.S. degrees from the Xi'an Research Institute of High Technology, Xi'an, China, in 2013 and 2016, respectively, where he is currently pursuing the Ph.D. degree in control science and engineering. His research interests include signal processing.

XIANGYU KONG was born in 1967. He received the bachelor's degree from the Beijing Institute of Technology, Beijing, China, in1990, the master's degree the from Xi'an Research Institute of High Technology, Xi'an, China, in 2000, and the Ph.D. degree in control science and engineering from the Xi'an Jiao tong University, Xi'an, in 2005.

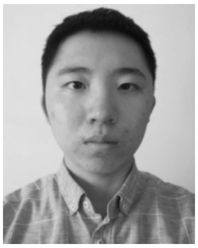
He is currently a Professor with the Xi'an Research Institute of High Technology. His research interests include adaptive signal processing, system modeling, and fault diagnosis.



YINGBIN GAO was born in 1986. He received the bachelor's, master's, and Ph.D. degrees from the Xi'an Research Institute of High Technology, Xi'an, China, in 2009, 2011, and 2016, respectively, all in control engineering. He is currently an Engineer with the China Electronics Technology Group Corp 54th Research Institute, Shijiazhuang, China. His current research interests include adaptive signal processing and feature extraction.

CHANGHUA HU was born in 1979. He received the bachelor's and master's degrees from the Xi'an Research Institute of High Technology, Xi'an, China, in 1987 and 1990, respectively, and the Ph.D. degree from Northwestern Polytechnical University, Xi'an, in 1996.

He is currently a Professor with the Xi'an Research Institute of High Technology. His current research has been supported by the National Science Fund for Distinguished Young Scholars of China. His research interests include fault diagnosis and prediction, life prognosis, and fault tolerant control.



XUHAO BIAN was born in 1997. He received the bachelor's degree from the Xi'an Research Institute of High Technology, Xi'an, China, in 2014. His research interests include signal processing.