Neurocomputing 317 (2018) 79–87

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Adaptive algorithms for computing the principal Takagi vector of a complex symmetric matrix

Maolin Che[a], Sanzheng Qiao[b], Yimin Wei[c,*]

[a] School of Economic Mathematics, Southwest University of Finance and Economics, Chengdu 611130, PR China
[b] Department of Computing and Software, McMaster University, Hamilton, ON L8S 4K1, Canada
[c] School of Mathematical Sciences and Key Laboratory of Mathematics for Nonlinear Sciences, Fudan University, Shanghai, 200433, PR China

ARTICLE INFO

Article history:
Received 3 January 2018
Revised 7 June 2018
Accepted 15 July 2018
Available online 13 August 2018

Communicated by Prof. Liu Xiwei

MSC:
15A18
15A69
65F15
65F10

Keywords:
Complex symmetric matrix
Takagi factorization
Principal Takagi vector
Adaptive algorithm
Asymptotic stability

ABSTRACT

In this paper, we present a unified framework for deriving and analyzing adaptive algorithms for computing the principal Takagi vector of a complex symmetric matrix. Eight systems of complex-valued ordinary differential equations (complex-valued ODEs) are derived and their convergence behavior is analyzed. We prove that the solutions of the complex-valued ODEs are asymptotically stable. The systems can be implemented on neural networks. Finally, we show experimental results to support our analyses.

1. Introduction

A complex symmetric matrix can be diagonalized by a unitary matrix. Specifically, let $\mathbf{A} = \mathbf{B} + \iota\mathbf{C}$ be a complex symmetric matrix, where $\mathbf{B}, \mathbf{C} \in \mathbb{R}^{I\times I}$ are symmetric, then there exist a unitary matrix $\mathbf{Q} \in \mathbb{C}^{I\times I}$ and a positive semidefinite diagonal matrix $\mathbf{\Sigma} \in \mathbb{R}^{I\times I}$ such that (see, e.g., [36])

$$\mathbf{A} = \mathbf{Q}\mathbf{\Sigma}\mathbf{Q}^\top, \quad \mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_I). \tag{1.1}$$

This factorization of a complex symmetric matrix is called the *Autonne-Takagi* factorization, or the *Takagi* factorization in short, originally proposed by Autonne [2] and Takagi [29]. The columns of $\mathbf{Q}$ are called the Takagi vectors of $\mathbf{A}$ and the diagonal elements of $\mathbf{\Sigma}$ are its Takagi values. We denote $\mathbf{\Sigma}$ and $\mathbf{Q}$ as the Takagi value matrix and the associated Takagi vector matrix of $\mathbf{A}$, respectively. In particular, if σ_1 is the largest Takagi value, then we call $\mathbf{q}_1$ the principal Takagi vector associated with σ_1.

The Takagi factorization reveals the symmetry of a complex symmetric matrix. One advantage of the factorization is that it can save storage and computation by about half. The Takagi factorization of a complex symmetric matrix has applications in the Grunsky inequalities [27], computation of the near-best uniform polynomial or rational approximation of a high degree polynomial on a disk [30], the complex independent component analysis [11], and nuclear magnetic resonance [3].

Throughout this paper, unless stated otherwise, I, J, and N denote the index upper bounds, lower case letters $x, u, \dots$ for scalars, bold lower case letters $\mathbf{x}, \mathbf{u}, \dots$ for vectors, bold capital letters $\mathbf{A}, \mathbf{B}, \dots$ for matrices of order I, and calligraphic letters $\mathcal{A}, \mathcal{B}, \dots$ for matrices of order $2I$. This notation is consistently used for entries. For example, the entry with row index i and column index j of a matrix $\mathbf{A}$, i.e., $(\mathbf{A})_{ij}$, is denoted by a_{ij} (also $(\mathbf{x})_i = x_i$). We denote $\overline{\mathbf{A}}$, $\mathbf{A}^\top$ and $\mathbf{A}^*$ the complex conjugate, the transpose and the complex conjugated transpose of $\mathbf{A}$, respectively. We use $|a|$ to denote the modulus of a complex number a. The real and imaginary parts of $\mathbf{z} \in \mathbb{C}^I$ are denoted by $\Re(\mathbf{z})$ and $\Im(\mathbf{z})$, respectively.

There are several ways of computing the Takagi factorization. For example, the complex problem can be transformed into a real

* Corresponding author.
E-mail addresses: cheml@swufe.edu.cn (M. Che), qiao@cas.mcmaster.ca (S. Qiao), ymwei@fudan.edu.cn (Y. Wei).

https://doi.org/10.1016/j.neucom.2018.07.064
0925-2312/

problem of double size. Since $\mathbf{Q}$ is unitary and $\mathbf{Q}^* = (\bar{\mathbf{Q}})^\top$, thus $\mathbf{A}\bar{\mathbf{Q}} = \mathbf{Q}\boldsymbol{\Sigma}$. Let $(\sigma, \mathbf{q})$ be a Takagi value-vector pair and $\mathbf{q} = \mathbf{x} + \iota\mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I$, then $\mathbf{A}\bar{\mathbf{q}} = \sigma\mathbf{q}$, that is,

$$\begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C} & -\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \sigma \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C} & -\mathbf{B} \end{bmatrix} \begin{bmatrix} -\mathbf{y} \\ \mathbf{x} \end{bmatrix} = -\sigma \begin{bmatrix} -\mathbf{y} \\ \mathbf{x} \end{bmatrix}.$$

Denote

$$\mathcal{A} = \begin{bmatrix} \mathbf{B} & \mathbf{C} \\ \mathbf{C} & -\mathbf{B} \end{bmatrix}$$

and note that $\mathcal{A} \in \mathbb{R}^{2I\times 2I}$ is symmetric. Thus, the Takagi factorization of $\mathbf{A}$ can be obtained from the eigenvalue decomposition of the real symmetric $\mathcal{A}$ of double size.

Suppose that $(\sigma, \mathbf{q})$ is a Takagi pair of $\mathbf{A}$. Let $\mathbf{v} = \bar{\mathbf{q}}$ and $\mathbf{u} = \mathbf{q}$, then

$$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \bar{\mathbf{A}}\mathbf{u} = \mathbf{A}^*\mathbf{u} = \sigma\mathbf{v},$$

which implies that $(\sigma, \mathbf{u}, \mathbf{v})$ is a singular pair of $\mathbf{A}$. Hence we can also compute the Takagi factorization of a complex symmetric matrix to obtain its singular value decomposition. The interested readers can be referred to [12,14,15,28] and their references for numerical computation of the singular value decomposition of matrices. However, as shown in [36], for a complex symmetric matrix $\mathbf{A}$, we have

(a) if $(\sigma, \mathbf{q})$ is a Takagi pair of $\mathbf{A}$, then $(\sigma, \mathbf{q}, \bar{\mathbf{q}})$ is its singular pair;
(b) if $(\sigma, \mathbf{u}, \mathbf{v})$ is a singular pair of $\mathbf{A}$, we can not directly obtain the Takagi vector associated with σ from $\mathbf{u}$ and $\mathbf{v}$.

Finally, similar to the computation of the SVD (see, e.g., [12]), Bunse-Gerstner and Gragg [7] stated that a standard algorithm for computing the Takagi factorization of a complex symmetric matrix consists of two stages: First, a complex symmetric matrix is reduced to a complex symmetric tridiagonal matrix; Second, the Takagi factorization of the complex symmetric tridiagonal matrix from the first stage is computed. For the first stage, Qiao et al. [25] derived a block Lanczos method for tridiagonalizing complex symmetric matrices. There are two methods for implementing the second stage: the divide-and-conquer method [36] and a twisted factorization method [37]. There are other numerical algorithms for computing the Takagi factorization of a complex symmetric matrix [7,36,37]. The existing algorithms are designed for the complete Takagi factorization.

In some applications, however, only the principal Takagi vector is required. Especially in the situation where we want to investigate the behavior of each component of the principal Takagi vector of a parameterized complex symmetric matrix. Moreover, we will shortly show that an algorithm for computing the principal Takagi vector of a complex symmetric matrix can be used to compute the complete Takagi factorization. So, in this paper, we focus on the computation of the principal Takagi vector of a complex symmetric matrix.

Let $\mathbf{A} \in \mathbb{C}^{I\times I}$ be symmetric, whose Takagi values are positive and distinct, that is, $\sigma_1 > \sigma_2 > \cdots > \sigma_N > 0$, and $\mathbf{q}_1$ the principal Takagi vector. Rewriting (1.1) as $\mathbf{A} = \sum_{i=1}^I \sigma_i\mathbf{q}_i\mathbf{q}_i^\top$, we get $\mathbf{A} - \sigma_1\mathbf{q}_1\mathbf{q}_1^\top = \sum_{i=2}^I \sigma_i\mathbf{q}_i\mathbf{q}_i^\top$, a symmetric matrix whose principal Takagi vector is $\mathbf{q}_2$. Repeating the process, we can successively compute $\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N$. From (1.1), the Takagi values σ_i can be obtained by $\sigma_i = \mathbf{q}_i^*\mathbf{A}\bar{\mathbf{q}}_i$, $i = 1, 2, \ldots, I$. The following algorithm computes $\mathbf{Q} = [\mathbf{q}_1 \ldots \mathbf{q}_N]$ and $\boldsymbol{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ in the complete Takagi factorization (1.1) by calling a procedure of computing the principal Takagi vector of a complex symmetric matrix.

Input: Symmetric $\mathbf{A} \in \mathbb{C}^{I\times I}$ with positiveand distinct Takagi values.
Output: The Takagi vectors $\mathbf{q}_i$ and Takagi values σ_i with $i = 1, 2, \ldots, I$.
$\mathbf{A}_1 = \mathbf{A}$.
for $i = 1, 2, \ldots, I$ **do**
  Compute $\mathbf{q}_i$: the principal Takagi vector of $\mathbf{A}_i$.
  $\sigma_i = \mathbf{q}_i^*\mathbf{A}_i\bar{\mathbf{q}}_i$.
  $\mathbf{A}_{i+1} = \mathbf{A}_i - \sigma_i\mathbf{q}_i\mathbf{q}_i^\top$.
end for

This paper presents a unified approach to complex-valued neural networks for computing the principal Takagi vector of a complex symmetric matrix with positive and distinct Takagi values.

2. Basics of real-valued functions of complex variables

For a function $f : \mathbb{C} \to \mathbb{C}$, its complex derivative at $x \in \mathbb{C}$, if it exists, is defined as the limit:

$$f'(x) := \lim_{\Delta x\to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

We know that f is differentiable in the complex sense, if and only if the Cauchy–Riemann conditions hold. However, in many practical applications, functions are not differentiable in the complex sense. In this paper, we consider optimization problems where the objective functions are real valued with complex variables. The objective functions are not complex differentiable, unless they are constant functions [5]. In order to deal with the problem, we provide an alternative formulation which is based on the real derivatives but similar to the complex derivative. The purpose of this section is to introduce some basics of the Wirtinger derivative of a real valued function with complex variables, including the cogradient and the Hessian matrix.

Suppose that $f : \mathbb{C}^I \to \mathbb{R}$ is a real-valued function with complex variables. Let $\mathbf{z} = \mathbf{x} + \iota\mathbf{y} \in \mathbb{C}^I$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I$. We introduce a calculus of the differential operators, developed in principal by Wirtinger, often called the *Wirtinger calculus*. We refer to [6,16,26,31] for the underlying framework of the complex derivatives.

Definition 2.1. Let $\mathbf{z} = \mathbf{x} + \iota\mathbf{y} \in \mathbb{C}^I$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^I$. The cogradient operator $\frac{\partial}{\partial\mathbf{z}}$ and the conjugate cogradient operator $\frac{\partial}{\partial\bar{\mathbf{z}}}$ are defined by

$$\frac{\partial}{\partial\mathbf{z}} = \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial x_1} - \iota\frac{\partial}{\partial y_1} \\ \vdots \\ \frac{\partial}{\partial x_I} - \iota\frac{\partial}{\partial y_I} \end{bmatrix} \quad \text{and} \quad \frac{\partial}{\partial\bar{\mathbf{z}}} = \frac{1}{2}\begin{bmatrix} \frac{\partial}{\partial x_1} + \iota\frac{\partial}{\partial y_1} \\ \vdots \\ \frac{\partial}{\partial x_I} + \iota\frac{\partial}{\partial y_I} \end{bmatrix}.$$

For example, define $f(\mathbf{z}) = \mathbf{z}^*\mathbf{z}$ for all $\mathbf{z} = \mathbf{x} + \iota\mathbf{y} \in \mathbb{C}^I$, then

$$\frac{\partial}{\partial\mathbf{z}} = \mathbf{x} - \iota\mathbf{y}, \quad \frac{\partial}{\partial\bar{\mathbf{z}}} = \mathbf{x} + \iota\mathbf{y}.$$

Note that the Cauchy–Riemann conditions for $f(\mathbf{z})$ to be analytic at $\mathbf{z}$ can be expressed compactly, using the cogradient as $\frac{\partial f}{\partial\bar{\mathbf{z}}} = \mathbf{0}_I$, i.e., f is a function of $\mathbf{z}$ only. Analogously, f is said to be analytic at $\bar{\mathbf{z}}$ if and only if $\frac{\partial f}{\partial\mathbf{z}} = \mathbf{0}_I$.

Let $\frac{\partial}{\partial\mathbf{x}} = [\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_I}]^\top$ and $\frac{\partial}{\partial\mathbf{y}} = [\frac{\partial}{\partial y_1}, \ldots, \frac{\partial}{\partial y_I}]^\top$, it follows immediately from the above definition that

$$\frac{\partial}{\partial\mathbf{x}} = \frac{\partial}{\partial\mathbf{z}} + \frac{\partial}{\partial\bar{\mathbf{z}}}, \quad \frac{\partial}{\partial\mathbf{y}} = \iota\left(\frac{\partial}{\partial\mathbf{z}} - \frac{\partial}{\partial\bar{\mathbf{z}}}\right).$$

The Hessian matrix is defined by

$$\mathcal{H}_1(\mathbf{z}) := \begin{bmatrix} \mathbf{H}_{\mathbf{z}\mathbf{z}}(\mathbf{z}) & \mathbf{H}_{\mathbf{z}\bar{\mathbf{z}}}(\mathbf{z}) \\ \mathbf{H}_{\bar{\mathbf{z}}\mathbf{z}}(\mathbf{z}) & \mathbf{H}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(\mathbf{z}) \end{bmatrix}$$

or

$$\mathcal{H}_2(\mathbf{z}) := \begin{bmatrix} \mathbf{H}_{\mathbf{z}\bar{\mathbf{z}}}(\mathbf{z}) & \mathbf{H}_{\mathbf{z}\mathbf{z}}(\mathbf{z}) \\ \mathbf{H}_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(\mathbf{z}) & \mathbf{H}_{\bar{\mathbf{z}}\mathbf{z}}(\mathbf{z}) \end{bmatrix}$$

where, for $k, l = 1, 2, \ldots, I$,

$$(\mathbf{H_{zz}})_{kl}(\mathbf{z}) = \frac{\partial}{\partial z_l}\left(\frac{\partial f(\mathbf{z})}{\partial z_k}\right), \quad (\mathbf{H_{z\bar{z}}})_{kl}(\mathbf{z}) = \frac{\partial}{\partial \bar{z}_l}\left(\frac{\partial f(\mathbf{z})}{\partial z_k}\right),$$

$$(\mathbf{H_{\bar{z}z}})_{kl}(\mathbf{z}) = \frac{\partial}{\partial z_l}\left(\frac{\partial f(\mathbf{z})}{\partial \bar{z}_k}\right), \quad (\mathbf{H_{\bar{z}\bar{z}}})_{kl}(\mathbf{z}) = \frac{\partial}{\partial \bar{z}_l}\left(\frac{\partial f(\mathbf{z})}{\partial \bar{z}_k}\right).$$

Note that $\mathcal{H}_1(\mathbf{z}) \in \mathbb{C}^{2I\times 2I}$ is symmetric, whereas $\mathcal{H}_2(\mathbf{z}) \in \mathbb{C}^{2I\times 2I}$ is Hermitian.

3. A framework

In this section, we describe a framework of deriving and analyzing neural network systems for computing the principal Takagi vector of a complex symmetric matrix. We refer to [1,4,13,20,33,35,38,39] for solving a complex-valued nonlinear convex programming problem by a complex-valued neural network model.

Wang et al. [34] proposed a dynamics of the complex-valued neural network model for computing the principal Takagi vector of a complex symmetric matrix **A**:

$$\frac{d\mathbf{z}(t)}{dt} = f(\mathbf{z}; \mathbf{A}), \quad \text{where} \quad f(\mathbf{z}; \mathbf{A}) = (\mathbf{z}^*\mathbf{z})\mathbf{A}\bar{\mathbf{z}} - \frac{\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z} + \mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}}{2}\mathbf{z}, \tag{3.1}$$

where $\mathbf{z} = [z_1, z_2, \ldots, z_I]^\top \in \mathbb{C}^I$ represents the state of the network. It is proved in [34] that the equilibrium point[1] of (3.1) corresponds to the principal Takagi vector of **A**. A simple discrete-time iterative algorithm corresponding to the neural network described in (3.1) can be written as

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \eta_k f(\mathbf{z}_k; \mathbf{A}) \tag{3.2}$$

with

$$f(\mathbf{z}_k; \mathbf{A}) = (\mathbf{z}_k^*\mathbf{z}_k)\mathbf{A}\bar{\mathbf{z}}_k - \frac{\mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k + \mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k}{2}\mathbf{z}_k,$$

where $\mathbf{z}_1 \in \mathbb{C}^I$ is a given initial value and $\{\eta_k\}$ is a decreasing gain sequence. Following [17], we make the following assumption on $\{\eta_k\}$.

Assumption 3.1. The gain sequence $\{\eta_k \geq 0\}$ is decreasing such that $\sum_{k=0}^{\infty} \eta_k = \infty$, $\sum_{k=0}^{\infty} \eta_k^r < \infty$ for some $r > 1$ and $\lim_{k\to\infty}(\eta_k^{-1} - \eta_{k-1}^{-1}) < \infty$.

Based on the relationship between the Takagi pairs of **A** and the eigenpairs of $\mathcal{A}$ [17], we can apply the theory of stochastic approximation to prove the convergence of $\mathbf{z}_k$ in (3.2) to the principal Takagi vector of **A**. Note that the stable stationary solution of (3.1) is a convergence point of (3.2). Hence we either solve (3.1) or analyze its stable stationary point.

In the following, from the framework, we derive and analyze eight different adaptive algorithms for computing the principal Takagi vector, as a realization and development of (3.2). The function $f(\mathbf{z}_k; \mathbf{A})$ for various adaptive algorithms is given by:

Type I $\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2}$;

Type II $\frac{1}{\mathbf{z}_k^*\mathbf{z}_k}\left(\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2\mathbf{z}_k^*\mathbf{z}_k}\right)$;

Type III $\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2\mathbf{z}_k^*\mathbf{z}_k} = \mathbf{z}_k^*\mathbf{z}_k \cdot \text{Type II}$;

Type IV $\mathbf{z}_k^*\mathbf{z}_k\left(\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2\mathbf{z}_k^*\mathbf{z}_k}\right) = (\mathbf{z}_k^*\mathbf{z}_k)^2 \cdot \text{Type II}$;

Type V $\frac{2\mathbf{A}\bar{\mathbf{z}}_k}{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k} - \mathbf{z}_k = \frac{2}{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k} \cdot \text{Type I}$;

Type VI $\mathbf{A}\bar{\mathbf{z}}_k - \mu\mathbf{z}_k(\mathbf{z}_k^*\mathbf{z}_k - 1)$;

Type VII $\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2} - \mathbf{z}_k(\mathbf{z}_k^*\mathbf{z}_k - 1)$;

Type VIII $\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2} - \mu\mathbf{z}_k(\mathbf{z}_k^*\mathbf{z}_k - 1)$;

Type I is based on nonlinear programming, Types II, III, and IV are based on the generalized Rayleigh quotient criterion, Type V is based on the information theory criterion, Type VI is based on the penalty function, and Types VII and VIII are based on the the augmented Lagrangian criterion.

Note that we cannot ensure that $\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}$ is always positive, where $\mathbf{z}$ is the limit of the sequence $\{\mathbf{z}_k, k = 1, 2, \ldots\}$, which is derived from the above eight adaptive algorithms. If $\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z} < 0$, when considering Type I and setting $\mathbf{u} = e^{\iota\pi/2}\mathbf{z}$, we have

$$\mathbf{A}\bar{\mathbf{u}} = \mathbf{u}\frac{\mathbf{u}^*\mathbf{A}\bar{\mathbf{u}} + \mathbf{u}^\top\bar{\mathbf{A}}\mathbf{u}}{2}, \quad \mathbf{u}^*\mathbf{A}\bar{\mathbf{u}} + \mathbf{u}^\top\bar{\mathbf{A}}\mathbf{u} > 0.$$

Although there exist many numerical algorithms to computing the Takagi factorization of complex symmetric matrices (see, e.g., [7,36,37]), it is more effective to use eight adaptive algorithms for computing the Takagi vectors of complex symmetric matrices in following situations:

(a) when we only want to derive the Takagi vector complex symmetric matrices, associated to the largest Takagi value;
(b) when we want to draw the graph about each component of the Takagi vector of parameterized complex symmetric matrices, associated to the largest Takagi value.

3.1. Type I

For Type I, the adaptive algorithm is derived from the following nonlinear programming

$$\min_{\mathbf{z}\in\mathbb{C}^I}\left\{-\frac{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}{2}\right\}, \quad \text{subject to } \mathbf{z}^*\mathbf{z} = 1.$$

When $f(\mathbf{z}_k; \mathbf{A})$ is Type I, we have

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \eta_k\left(\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2}\right), \tag{3.3}$$

where the sequence $\{\eta_k\}$ satisfies Assumption 3.1 in Section 3. Our model is a generalization of the previous work [21,23], where real **A** and $\mathbf{z}_k$ in $f(\mathbf{z}_k; \mathbf{A})$ are considered. The complex-valued ODE associated with (3.3) is given by

$$\frac{d\mathbf{z}(t)}{dt} = \mathbf{A}\bar{\mathbf{z}} - \mathbf{z}\frac{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}{2}, \tag{3.4}$$

where $\mathbf{z}(t) \in \mathbb{C}^I$ is the continuous time counterpart of $\mathbf{z}_k$ with $t \geq 0$ representing time. The following theorem gives the stable equilibrium points and the convergence properties of (3.4).

Theorem 3.1. *Under the conditions of Assumption 3.1, given a symmetric matrix $\mathbf{A} \in \mathbb{C}^{I\times I}$ with $\sigma_1 > \sigma_2 \geq \cdots \geq \sigma_I \geq 0$, let $\mathbf{z}(t) = \sum_{i=1}^{I} a_i(t)\mathbf{q}_i$ be a solution of the complex-valued ODE (3.4) expressed in terms of the set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_I\}$ of all the unitary Takagi vectors of **A**, then for any initial condition $\mathbf{z}(0) = \mathbf{z}_0 \in \mathbb{C}^I$ with $\mathbf{z}_0^*\mathbf{q}_1 \neq 0$ and the real and imaginary parts of $a_i(0)$ being nonzero, the coefficients $a_i(t)$ $(i = 2, 3, \ldots, I)$ of the solution $\mathbf{z}(t)$ are given by*

$$\frac{\Re(a_i(t))}{\Re(a_1(t))} = \frac{\Re(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1-\sigma_i)t} \quad \text{and}$$
$$\frac{\Im(a_i(t))}{\Re(a_1(t))} = \frac{\Im(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1+\sigma_i)t}. \tag{3.5}$$

[1] The point $\tilde{x} \in \mathbb{R}$ is an equilibrium point for the differential equation

$$\frac{dx}{dt} = f(t, x)$$

if $f(t, \tilde{x}) = 0$ for all t.

Furthermore, we also have $\Re(a_1(t)) \to \pm 1$ *and* $\Im(a_1(t)) \to 0$ *as* $t \to \infty$.

The points $\pm \mathbf{q}_1$ *are (uniformly) asymptotically stable. The domain of contraction of* $\mathbf{q}_1$ *is* $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^* \mathbf{q}_1) > 0\}$ *and that of* $-\mathbf{q}_1$ *is* $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^* \mathbf{q}_1) < 0\}$.

Proof. Substituting $\mathbf{z}(t)$ in (3.4) with $\mathbf{z}(t) = \sum_{i=1}^I a_i(t)\mathbf{q}_i$ and premultiplying the both side with $\mathbf{q}_i^*$, we obtain

$$\frac{da_i(t)}{dt} = \sigma_i \bar{a}_i(t) - a_i(t) \sum_{j=1}^I \sigma_j (\Re(a_j(t))^2 - \Im(a_j(t))^2), \quad i = 1, 2, \dots, I. \tag{3.6}$$

Define $\mathbf{b} \in \mathbb{C}^{I-1}$ such that

$$\Re(b_i) = \frac{\Re(a_i)}{\Re(a_1)} \quad \text{and} \quad \Im(b_i) = \frac{\Im(a_i)}{\Re(a_1)}, \quad i = 2, \dots, I.$$

It then follows from (3.6) that

$$\frac{d\Re(b_i(t))}{dt} = -\Re(b_i)(\sigma_1 - \sigma_i) \quad \text{and}$$
$$\frac{d\Im(b_i(t))}{dt} = -\Im(b_i)(\sigma_1 + \sigma_i), \quad i = 2, \dots, I,$$

which implies (3.5). From (3.5), we have $\Re(a_i) \to 0$ and $\Im(a_i) \to 0$ as $t \to \infty$, for $i = 2, \dots, I$. Thus, we have $\Re(a_i(t)) \to x(t)$ and $\Im(a_i(t)) \to y(t)$ as $t \to \infty$, where $x(t)$ and $y(t)$ satisfies the following ordinary differential equations:

$$\begin{aligned} \frac{dx(t)}{dt} &= \sigma_1 x(t) - \sigma_1 (x(t)^2 - y(t)^2) x(t), \\ \frac{dy(t)}{dt} &= -\sigma_1 y(t) - \sigma_1 (x(t)^2 - y(t)^2) y(t). \end{aligned} \tag{3.7}$$

When letting $u(t) = x(t)^2$ and $v(t) = y(t)^2$, we have

$$\frac{du(t)}{dt} = \sigma_1 u(t) - \sigma_1 (u(t) - v(t)) u(t)^2,$$
$$\frac{dv(t)}{dt} = -\sigma_1 v(t) - \sigma_1 (u(t) - v(t)) v(t)^2,$$

which implies that

$$u(t) = \frac{e^{2\sigma_1 t} u(0)}{e^{2\sigma_1 t} u(0) - 2\sigma_1 v(0) t + 1},$$
$$v(t) = \frac{v(0)}{e^{2\sigma_1 t} u(0) - 2\sigma_1 v(0) t + 1}.$$

with $u(0) = x(0)^2 > 0$ and $v(0) = y(0)^2 > 0$, where $x(0)$ and $y(0)$ are the initial values of (3.7). Thus, $u(t) \to 1$ and $v(t) \to 0$ as $t \to \infty$. which lead to $\Re(a_1(t)) \to \pm 1$ and $\Im(a_1(t)) \to 0$ as $t \to \infty$. Thus the sign of $\Re(a_i(t))$, $i = 2, \dots, I$, is determined by the sign of the initial $\Re(a_1(0)) = \Re(\mathbf{z}(0)^* \mathbf{q}_1)$. From (3.5), $\Re(a_i(t)) \to 0$ and $\Im(a_i(t)) \to 0$, as $t \to \infty$, for $i = 2, 3, \dots, I$. Thus $\mathbf{z}(t) \to \pm \mathbf{q}_1$ as $t \to \infty$. □

Remark 3.1. Similar to the proof of Theorem 3.1, we also have

$$\frac{\Re(a_i(t))}{\Im(a_1(t))} = \frac{\Re(a_i(0))}{\Im(a_1(0))} e^{(\sigma_1 - \sigma_i)t}, \quad \frac{\Im(a_i(t))}{\Im(a_1(t))} = \frac{\Im(a_i(0))}{\Im(a_1(0))} e^{(\sigma_1 + \sigma_i)t}, \quad i = 2, 3, \dots, I.$$

It is easy to verify that both $\Re(a_i(t))$ and $\Im(a_i(0))$ are the high order infinitesimal of $\Im(a_i(0))$, which also hold for the other adaptive algorithms in this paper.

The convergence of (3.3) is now a direct consequences of the above theorem and Theorem 1 of Ljung [17].

3.2. Types II, III and IV

All the three types are derived from the objective function:

$$J(\mathbf{z}_k; \mathbf{A}) = -\frac{\mathbf{z}_k^* \mathbf{A} \bar{\mathbf{z}}_k + \mathbf{z}_k^\top \bar{\mathbf{A}} \mathbf{z}_k}{2 \mathbf{z}_k^* \mathbf{z}_k}, \tag{3.8}$$

called the generalized Rayleigh quotient of $\mathbf{A}$. Our models are generalizations of the previous work by Cirrincione et al. [10], Luo et al. [18] and Oja [22], where the real case is considered. The gradient of (3.8) with respect to $\bar{\mathbf{z}}_k$ is

$$\nabla_{\bar{\mathbf{z}}_k} J(\mathbf{z}_k; \mathbf{A}) = \frac{-1}{\mathbf{z}_k^* \mathbf{z}_k} \left(\mathbf{A} \bar{\mathbf{z}}_k - \frac{\mathbf{z}_k^* \mathbf{A} \bar{\mathbf{z}}_k + \mathbf{z}_k^\top \bar{\mathbf{A}} \mathbf{z}_k}{2 \mathbf{z}_k^* \mathbf{z}_k} \mathbf{z}_k \right).$$

Hence, the adaptive gradient descent algorithm for Type II is

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \eta_k \nabla_{\bar{\mathbf{z}}_k} J(\mathbf{z}_k; \mathbf{A}) = \mathbf{z}_k + \eta_k \frac{1}{\mathbf{z}_k^* \mathbf{z}_k} \left(\mathbf{A} \bar{\mathbf{z}}_k - \frac{\mathbf{z}_k^* \mathbf{A} \bar{\mathbf{z}}_k + \mathbf{z}_k^\top \bar{\mathbf{A}} \mathbf{z}_k}{2 \mathbf{z}_k^* \mathbf{z}_k} \mathbf{z}_k \right), \tag{3.9}$$

and the complex-valued ODE associated with (3.9) is

$$\frac{d\mathbf{z}(t)}{dt} = \frac{1}{\mathbf{z}^* \mathbf{z}} \left(\mathbf{A} \bar{\mathbf{z}} - \mathbf{z} \frac{\mathbf{z}^* \mathbf{A} \bar{\mathbf{z}} + \mathbf{z}^\top \bar{\mathbf{A}} \mathbf{z}}{2 \mathbf{z}^* \mathbf{z}} \right), \tag{3.10}$$

where $\mathbf{z}(t) \in \mathbb{C}^I$ is the continuous time counterpart of $\mathbf{z}_k$. For Type III, the adaptive gradient descent algorithm is

$$\mathbf{z}_{k+1} = \mathbf{z}_k - \eta_k (\mathbf{z}_k^* \mathbf{z}_k) \nabla_{\bar{\mathbf{z}}_k} J(\mathbf{z}_k; \mathbf{A}) = \mathbf{z}_k + \eta_k \left(\mathbf{A} \bar{\mathbf{z}}_k - \frac{\mathbf{z}_k^* \mathbf{A} \bar{\mathbf{z}}_k + \mathbf{z}_k^\top \bar{\mathbf{A}} \mathbf{z}_k}{2 \mathbf{z}_k^* \mathbf{z}_k} \mathbf{z}_k \right), \tag{3.11}$$

and the complex-valued ODE associated with (3.11) is

$$\frac{d\mathbf{z}(t)}{dt} = \mathbf{A} \bar{\mathbf{z}} - \mathbf{z} \frac{\mathbf{z}^* \mathbf{A} \bar{\mathbf{z}} + \mathbf{z}^\top \bar{\mathbf{A}} \mathbf{z}}{2 \mathbf{z}^* \mathbf{z}}. \tag{3.12}$$

For Type IV, the adaptive gradient descent algorithm is

$$\begin{aligned} \mathbf{z}_{k+1} &= \mathbf{z}_k - \eta_k (\mathbf{z}_k^* \mathbf{z}_k)^2 \nabla_{\bar{\mathbf{z}}_k} J(\mathbf{z}_k; \mathbf{A}) \\ &= \mathbf{z}_k + \eta_k (\mathbf{z}_k^* \mathbf{z}_k) \left(\mathbf{A} \bar{\mathbf{z}}_k - \frac{\mathbf{z}_k^* \mathbf{A} \bar{\mathbf{z}}_k + \mathbf{z}_k^\top \bar{\mathbf{A}} \mathbf{z}_k}{2 \mathbf{z}_k^* \mathbf{z}_k} \mathbf{z}_k \right), \end{aligned} \tag{3.13}$$

and the complex-valued ODE associated with (3.13) is

$$\frac{d\mathbf{z}(t)}{dt} = \mathbf{z}^* \mathbf{z} \left(\mathbf{A} \bar{\mathbf{z}} - \mathbf{z} \frac{\mathbf{z}^* \mathbf{A} \bar{\mathbf{z}} + \mathbf{z}^\top \bar{\mathbf{A}} \mathbf{z}}{2 \mathbf{z}^* \mathbf{z}} \right). \tag{3.14}$$

The solutions of the above three complex-valued ODEs and their properties are summarized in the following theorem.

Theorem 3.2. *Under the conditions of Assumption (3.1), given a symmetric matrix* $\mathbf{A} \in \mathbb{C}^{I \times I}$ *with* $\sigma_1 > \sigma_2 \geq \dots \geq \sigma_I > 0$, *let* $\mathbf{z}(t) = \sum_{i=1}^I a_i(t) \mathbf{q}_i$ *be a solution of the complex-valued ODEs (3.10) or (3.12) or (3.14) expressed in terms of the set* $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_I\}$ *of all the unitary Takagi vectors of* $\mathbf{A}$, *then for any initial condition* $\mathbf{z}(0) = \mathbf{z}_0 \in \mathbb{R}^I$ *with* $\Re(\mathbf{z}_0^* \mathbf{q}_1) \neq 0$ *and the real and imaginary parts of* $a_i(0)$ *being nonzero, we have the following statements:*

(a) *for all* $t \geq 0$, *the coefficients* $a_i(t)$ $(i = 2, 3, \dots, I)$ *of the solution for (3.10) are given by*

$$\begin{aligned} \frac{\Re(a_i(t))}{\Re(a_1(t))} &= \frac{\Re(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 - \sigma_i)t / \|\mathbf{z}_0\|_2^2}, \\ \frac{\Im(a_i(t))}{\Re(a_1(t))} &= \frac{\Im(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 + \sigma_i)t / \|\mathbf{z}_0\|_2^2}; \end{aligned} \tag{3.15}$$

(b) *for all* $t \geq 0$, *the coefficients* $a_i(t)$ $(i = 2, 3, \dots, I)$ *of the solution for (3.12) are given by*

$$\frac{\Re(a_i(t))}{\Re(a_1(t))} = \frac{\Re(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 - \sigma_i)t}, \quad \frac{\Im(a_i(t))}{\Re(a_1(t))} = \frac{\Im(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 + \sigma_i)t}; \tag{3.16}$$

(c) *for all* $t \geq 0$, *the coefficients* $a_i(t)$ $(i = 2, 3, \dots, I)$ *of the solution for (3.14) are given by*

$$\begin{aligned} \frac{\Re(a_i(t))}{\Re(a_1(t))} &= \frac{\Re(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 - \sigma_i)t \|\mathbf{z}_0\|_2^2}, \\ \frac{\Im(a_i(t))}{\Re(a_1(t))} &= \frac{\Im(a_i(0))}{\Re(a_1(0))} e^{-(\sigma_1 + \sigma_i)t \|\mathbf{z}_0\|_2^2}; \end{aligned} \tag{3.17}$$

(d) *furthermore, for (3.10), (3.12) and (3.14), we also have* $\Re(a_1(t))\to\pm\|\mathbf{z}_0\|_2$ *and* $\Im(a_1(t))\to 0$ *as* $t\to\infty$.

The points $\pm\mathbf{q}_1$ *are (uniformly) asymptotically stable. The domain of contraction of* $\mathbf{q}_1$ *is* $\mathbb{D}(\mathbf{q}_1)=\{\mathbf{z}\in\mathbb{C}^I:\Re(\mathbf{z}^*\mathbf{q}_1)>0\}$ *and that of* $-\mathbf{q}_1$ *is* $\mathbb{D}(\mathbf{q}_1)=\{\mathbf{z}\in\mathbb{C}^I:\Re(\mathbf{z}^*\mathbf{q}_1)<0\}$.

Proof. Analogous to the proof of Theorem 3.1. □

For Type II (3.10), the equations in (3.15) imply $\Re(a_i(t))\to 0$ and $\Im(a_i(t))\to 0$, as $t\to\infty$, for $i=2,3,\ldots,I$. For Type III (3.12), the equations in (3.16) imply $\Re(a_i(t))\to 0$ and $\Im(a_i(t))\to 0$, as $t\to\infty$, for $i=2,3,\ldots,I$. For Type IV (3.14), the equations in (3.17) imply $\Re(a_i(t))\to 0$ and $\Im(a_i(t))\to 0$, as $t\to\infty$, for $i=2,3,\ldots,I$. Thus, if $\|\mathbf{z}_0\|_2=1$ for Types II, III and IV, then $\mathbf{z}(t)\to\pm\mathbf{q}_1$ as $t\to\infty$.

Note that if $\mathbf{z}$ satisfies (3.10) or (3.12) or (3.14), then

$$\frac{d\|\mathbf{z}(t)\|_2^2}{dt}=\mathbf{z}(t)^*\frac{d\mathbf{z}(t)}{dt}+\frac{d\mathbf{z}(t)}{dt}^*\mathbf{z}(t)=0,$$

implying that $\|\mathbf{z}(t)\|_2=\|\mathbf{z}(0)\|_2$, $t>0$, for all the three types. Applying the approximation $\|\mathbf{z}_k\|_2\approx\|\mathbf{z}_0\|_2=\|\mathbf{z}(0)\|_2$ to Type II (3.9) and Type IV (3.13), we have the following simplified Types II and IV algorithms

$$(3.9)':\quad \mathbf{z}_{k+1}=\mathbf{z}_k+\eta_k\|\mathbf{z}_0\|_2^{-2}\left(\mathbf{A}\bar{\mathbf{z}}_k-\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k+\mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2\mathbf{z}_k^*\mathbf{z}_k}\mathbf{z}_k\right),$$

$$(3.13)':\quad \mathbf{z}_{k+1}=\mathbf{z}_k+\eta_k\|\mathbf{z}_0\|_2^{2}\left(\mathbf{A}\bar{\mathbf{z}}_k-\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k+\mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2\mathbf{z}_k^*\mathbf{z}_k}\mathbf{z}_k\right).$$

Although, the three adaptive algorithms (3.9)′, (3.11) and (3.13)′ differ only in their gain constants, our experiments show that they perform differently.

3.3. Type V

When $f(\mathbf{z}_k;\ \mathbf{A})$ is defined by Type V, we have the following adaptive method:

$$\mathbf{z}_{k+1}=\mathbf{z}_k+\eta_k\left(\frac{2\mathbf{A}\bar{\mathbf{z}}_k}{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k+\mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}-\mathbf{z}_k\right). \tag{3.18}$$

Our model is a generalization of the previous work by Plumbly [24] and Miao and Hua [19], where real $\mathbf{A}$ and $\mathbf{z}_k$ in (3.18) are considered. The complex-valued ODE associated with the adaptive method (3.18) is

$$\frac{d\mathbf{z}(t)}{dt}=\frac{2\mathbf{A}\bar{\mathbf{z}}}{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}-\mathbf{z}, \tag{3.19}$$

where $\mathbf{z}(t)\in\mathbb{C}^I$ is the continuous time counterpart of $\mathbf{z}_k$ with t representing time. If $\mathbf{z}$ satisfies (3.19), then

$$\frac{d\|\mathbf{z}(t)\|_2^2}{dt}=\mathbf{z}(t)^*\frac{d\mathbf{z}(t)}{dt}+\frac{d\mathbf{z}(t)}{dt}^*\mathbf{z}(t)=2(1-\|\mathbf{z}(t)\|_2^2),$$

from which we obtain $\|\mathbf{z}(t)\|_2^2=1+(\|\mathbf{z}(t)\|_2^2-1)e^{-2t}$, implying that $\|\mathbf{z}(t)\|_2\to 1$ as $t\to\infty$.

We now analyze the stable stationary points of the complex-valued ODE (3.19). An energy function for (3.19) is given by

$$E(\mathbf{z})=\mathbf{z}^*\mathbf{z}-\frac{1}{2}\ln\left(\left(\frac{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}{2}\right)^2\right).$$

Since $\lim_{\|\mathbf{z}\|_2\to 0}E(\mathbf{z})=+\infty$ and $\lim_{\|\mathbf{z}\|_2\to+\infty}E(\mathbf{z})=+\infty$, the function $E(\mathbf{z})$ has global minimum points. The following theorem shows a relation between the principal Takagi vector of $\mathbf{A}$ and the global minima of $E(\mathbf{z})$.

Theorem 3.3. *Under the conditions of Assumption 3.1, let* $\mathbf{A}\in\mathbb{C}^{I\times I}$ *be a symmetric matrix with* $\sigma_1>\sigma_2\geq\cdots\geq\sigma_I>0$, *then the two converging points* $\pm\mathbf{q}_1$ *of the complex-valued ODE (3.19) are the global minimum points of the energy function* $E(\mathbf{z})$, *which has no other local minimum point. In addition,* $\pm\mathbf{q}_i$, $i=2,\ldots,I$ *are the saddle points of* $E(\mathbf{z})$.

Proof. The gradients of $E(\mathbf{z})$ with respect to $\bar{\mathbf{z}}$ and $\mathbf{z}$ are respectively

$$\frac{\partial E(\mathbf{z})}{\partial\bar{\mathbf{z}}}=\mathbf{z}-\frac{2\mathbf{A}\bar{\mathbf{z}}}{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}},$$
$$\frac{\partial E(\mathbf{z})}{\partial\mathbf{z}}=\bar{\mathbf{z}}-\frac{2\bar{\mathbf{A}}\mathbf{z}}{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}.$$

It then follows that the stationary points of $E(\mathbf{z})$ are indeed the Takagi vectors $\pm\mathbf{q}_i$, $i=1,2,\ldots,I$ of $\mathbf{A}$. The Hessian matrix of $E(\mathbf{z})$ is

$$\mathbf{H}(\mathbf{z}):=\begin{bmatrix}E_{\mathbf{z}\bar{\mathbf{z}}}(\mathbf{z}) & E_{\mathbf{z}\mathbf{z}}(\mathbf{z})\\ E_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(\mathbf{z}) & E_{\bar{\mathbf{z}}\mathbf{z}}(\mathbf{z})\end{bmatrix},$$

where

$$E_{\mathbf{z}\mathbf{z}}(\mathbf{z})=\frac{-2\bar{\mathbf{A}}}{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}+\frac{4\bar{\mathbf{A}}\mathbf{z}\mathbf{z}^\top\bar{\mathbf{A}}}{(\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z})^2},$$

$$E_{\mathbf{z}\bar{\mathbf{z}}}(\mathbf{z})=\mathbf{I}_I+\frac{4\bar{\mathbf{A}}\mathbf{z}\mathbf{z}^*\mathbf{A}}{(\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z})^2},$$

$$E_{\bar{\mathbf{z}}\mathbf{z}}(\mathbf{z})=\mathbf{I}_I+\frac{4\mathbf{A}\bar{\mathbf{z}}\mathbf{z}^\top\bar{\mathbf{A}}}{(\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z})^2},$$

$$E_{\bar{\mathbf{z}}\bar{\mathbf{z}}}(\mathbf{z})=\frac{-2\mathbf{A}}{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}+\frac{4\mathbf{A}\bar{\mathbf{z}}\mathbf{z}^*\mathbf{A}}{(\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}}+\mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z})^2}.$$

In particular, for $\mathbf{q}_1$, we have

$$\mathbf{H}(\pm\mathbf{q}_1)\begin{bmatrix}\bar{\mathbf{q}}_1\\ \pm\mathbf{q}_1\end{bmatrix}=2\begin{bmatrix}\bar{\mathbf{q}}_1\\ \pm\mathbf{q}_1\end{bmatrix}\quad\text{and}$$

$$\mathbf{H}(\pm\mathbf{q}_1)\begin{bmatrix}\bar{\mathbf{q}}_i\\ \pm\mathbf{q}_i\end{bmatrix}=\frac{\sigma_1-\sigma_i}{\sigma_1}\begin{bmatrix}\bar{\mathbf{q}}_i\\ \pm\mathbf{q}_i\end{bmatrix},\quad i=2,3,\ldots,I.$$

Since all the eigenvalues of $\mathbf{H}(\pm\mathbf{q}_1)$ are positive, $\mathbf{H}(\pm\mathbf{q}_1)$ is positive definite. Thus the stationary points $\pm\mathbf{q}_1$ are the local minimal points of $E(\mathbf{z})$.

Similarly, it can be verified that

$$\mathbf{H}(\pm\mathbf{q}_i)\begin{bmatrix}\bar{\mathbf{q}}_i\\ \pm\mathbf{q}_i\end{bmatrix}=2\begin{bmatrix}\bar{\mathbf{q}}_i\\ \pm\mathbf{q}_i\end{bmatrix}\quad\text{and}$$

$$\mathbf{H}(\pm\mathbf{q}_i)\begin{bmatrix}\bar{\mathbf{q}}_1\\ \pm\mathbf{q}_1\end{bmatrix}=\frac{\sigma_i-\sigma_1}{\sigma_i}\begin{bmatrix}\bar{\mathbf{q}}_1\\ \pm\mathbf{q}_1\end{bmatrix},\quad i=2,3,\ldots,I.$$

Since $\mathbf{H}(\pm\mathbf{q}_i)$, $2\leq i\leq I$ has two positive eigenvalues and $(2I-2)$ negative eigenvalues, it is indefinite. Hence, $\pm\mathbf{q}_i$ are saddle points of $E(\mathbf{z})$. Since $E(\mathbf{z})$ has only two local minimal points $\pm\mathbf{q}_1$ and $E(\mathbf{q}_1)=E(-\mathbf{q}_1)$, $\pm\mathbf{q}_1$ are the global minima. □

The rate of convergence for (3.18) can be obtained from the equation $\|\mathbf{z}(t)\|_2^2=1+(\|\mathbf{z}(t)\|_2^2-1)e^{-2t}$. A unique feature of this algorithm is that the time constant for $\|\mathbf{z}(t)\|_2$ is 1 and independent of the Takagi structure of $\mathbf{A}$.

3.4. Type VI

When $f(\mathbf{z}_k;\ \mathbf{A})$ is given by Type VI, we have the following adaptive procedure:

$$\mathbf{z}_{k+1}=\mathbf{z}_k+\eta_k\left(\mathbf{A}\bar{\mathbf{z}}_k-\mu\mathbf{z}_k(\mathbf{z}_k^*\mathbf{z}_k-1)\right), \tag{3.20}$$

where $\mu>0$. Our method is a generalization the previous work by Chauvin [8], where both $\mathbf{A}$ and $\mathbf{z}_k$ in $f(\mathbf{z}_k;\ \mathbf{A})$ are assumed to be real. The complex-valued ODE associated with the adaptive method (3.20) is

$$\frac{d\mathbf{z}(t)}{dt}=\mathbf{A}\bar{\mathbf{z}}-\mu\mathbf{z}(\mathbf{z}^*\mathbf{z}-1), \tag{3.21}$$

where $\mu > 0$ and $\mathbf{z}(t) \in \mathbb{C}^I$ is the continuous time counterpart of $\mathbf{z}_k$ with t representing time. The solution of (3.21) and its properties are summarized in the theorem below.

Theorem 3.4. *Under the conditions of Assumption 3.1, for a given symmetric matrix $\mathbf{A} \in \mathbb{C}^{I\times I}$ with $\sigma_1 > \sigma_2 \geq \cdots \geq \sigma_I > 0$, let $\mathbf{z}(t) = \sum_{i=1}^{I} a_i(t)\mathbf{q}_i$ be a solution of the complex-valued ODE (3.21) expressed in terms of the set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_I\}$ of all the unitary Takagi vectors of $\mathbf{A}$, then for any initial condition $\mathbf{z}(0) = \mathbf{z}_0 \in \mathbb{C}^I$ with $\mathbf{z}_0^*\mathbf{q}_1 \neq 0$ and the real and imaginary parts of $a_i(0)$ being nonzero, the coefficients $a_i(t)$ $(i = 2, 3, \ldots, I)$ of the solution are given by*

$$\frac{\Re(a_i(t))}{\Re(a_1(t))} = \frac{\Re(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1-\sigma_i)t}, \quad \frac{\Im(a_i(t))}{\Re(a_1(t))} = \frac{\Im(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1+\sigma_i)t}. \tag{3.22}$$

Furthermore, we also have $\Re(a_1(t)) \to \pm\sqrt{1+\sigma_1/\mu}$ and $\Im(a_1(t)) \to 0$ as $t \to \infty$.

The points $\pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ are (uniformly) asymptotically stable. The domain of contraction of $\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ is $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^*\mathbf{q}_1) > 0\}$ and that of $-\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ is $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^*\mathbf{q}_1) < 0\}$.

Proof. Analogous to the proof of Theorem 3.1. □

It follows from (3.22) that $\Re(a_i(t)) \to 0$ and $\Im(a_i(t)) \to 0$, as $t \to \infty$, for $i = 2, 3, \ldots, I$. Thus $\mathbf{z}(t)$ converges to $\pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$. In terms of the energy function

$$E(\mathbf{z}) = -\frac{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}{2} + \frac{\mu}{2}(\mathbf{z}^*\mathbf{z} - 1)^2$$

for the adaptive procedure (3.20), the stable stationary points of (3.21) are given in the following theorem.

Theorem 3.5. *Under Assumption 3.1, let $\mathbf{A} \in \mathbb{C}^{I\times I}$ be a symmetric matrix with $\sigma_1 > \sigma_2 \geq \cdots \geq \sigma_I > 0$, then the two converging points $\pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ of the complex-valued ODE (3.21) are the global minimum points of $E(\mathbf{z})$, which has no other local minimum point. In addition, $\pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_i$ $i = 2, \ldots, I$, are saddle points of $E(\mathbf{z})$.*

Proof. Analogous to the proof of Theorem 3.3. □

3.5. Types VII and VIII

When $f(\mathbf{z}_k; \mathbf{A})$ is given by Type VIII, we have the following adaptive procedure:

$$\mathbf{z}_{k+1} = \mathbf{z}_k + \eta_k\left(\mathbf{A}\bar{\mathbf{z}}_k - \mathbf{z}_k\frac{\mathbf{z}_k^*\mathbf{A}\bar{\mathbf{z}}_k + \mathbf{z}_k^\top\bar{\mathbf{A}}\mathbf{z}_k}{2} - \mu\,\mathbf{z}_k(\mathbf{z}_k^*\mathbf{z}_k - 1)\right), \tag{3.23}$$

where $\mu > 0$. Note that when $\mu = 1$, the above adaptive procedure reduces to Type VII. The complex-valued ODE associated with the adaptive method (3.23) is

$$\frac{d\mathbf{z}(t)}{dt} = \mathbf{A}\bar{\mathbf{z}} - \mathbf{z}\frac{\mathbf{z}^*\mathbf{A}\bar{\mathbf{z}} + \mathbf{z}^\top\bar{\mathbf{A}}\mathbf{z}}{2} - \mu\,\mathbf{z}(\mathbf{z}^*\mathbf{z} - 1), \tag{3.24}$$

where $\mu > 0$ and $\mathbf{z}(t) \in \mathbb{C}^I$ is the continuous time counterpart of $\mathbf{z}_k$ with t representing time. The solution of (3.24) and its properties are summarized in the theorem below.

Theorem 3.6. *Under the conditions of Assumption 3.1, given a symmetric matrix $\mathbf{A} \in \mathbb{C}^{I\times I}$ with $\sigma_1 > \sigma_2 \geq \cdots \geq \sigma_I > 0$, let $\mathbf{z}(t) = \sum_{i=1}^{I} a_i(t)\mathbf{q}_i$ be a solution of the complex-valued ODE (3.24) expressed in term of the set $\{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_I\}$ of all the unitary Takagi vectors $\mathbf{A}$, then for any initial condition $\mathbf{z}(0) = \mathbf{z}_0 \in \mathbb{C}^I$ with $\mathbf{z}_0^*\mathbf{q}_1 \neq 0$ and the real and imaginary parts of $a_i(0)$ being nonzero, the coefficients $a_i(t)$ $(i = 1, 2, \ldots, I)$ of the solution $\mathbf{z}(t)$ for $t > 0$ are given by*

$$\frac{\Re(a_i(t))}{\Re(a_1(t))} = \frac{\Re(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1-\sigma_i)t}, \quad \frac{\Im(a_i(t))}{\Re(a_1(t))} = \frac{\Im(a_i(0))}{\Re(a_1(0))}e^{-(\sigma_1+\sigma_i)t}. \tag{3.25}$$

Furthermore, we also have $\Re(a_1(t)) \to \pm\sqrt{1+\sigma_1/\mu}$ and $\Im(a_1(t)) \to 0$ as $t \to \infty$.

The points $\pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ are (uniformly) asymptotically stable. The domain of contraction of $\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ is $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^*\mathbf{q}_1) > 0\}$ and that of $-\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ is $\mathbb{D}(\mathbf{q}_1) = \{\mathbf{z} \in \mathbb{C}^I : \Re(\mathbf{z}^*\mathbf{q}_1) < 0\}$.

Proof. Analogous to the proof of Theorem 3.1. □

It follows from (3.25) that $\Re(a_i(t)) \to 0$ and $\Im(a_i(t)) \to 0$, as $t \to \infty$, for $i = 2, 3, \ldots, I$, and $\Re(a_1(t)) \to \pm\sqrt{1+\sigma_1/\mu}$, $\Im(a_1(t)) \to 0$. Thus $\mathbf{z}(t) \to \pm\sqrt{1+\sigma_1/\mu}\,\mathbf{q}_1$ as $t \to \infty$.

4. Numerical examples

In this section, some computer simulation results are shown to illustrate our models. All computations were carried out in MATLAB Version 2016a, which has a unit of roundoff $2^{-53} \approx 1.1 \times 10^{-16}$, on a laptop with Intel Core i5-3470M CPU (3.20 GHz) and 8.00 GB RAM. All floating-point numbers were rounded to two digits after the decimal point.

Che et al. [9] presented an iterative algorithm for computing the largest Takagi value of complex symmetric matrices, which is similar to the power method [12] for matrix eigenvalue problems. Here we denote this algorithm by "PM". In this section, we also compare the proposed methods with PM for different testing complex symmetric matrices.

First, we ran our adaptive algorithms on a single data set with various starting vectors $\mathbf{z}_0$. Then, we generated several data samples and used the same starting vector $\mathbf{z}_0$.

We refer to the Takagi values and Takagi vectors of a symmetric matrix $\mathbf{A} \in \mathbb{R}^{I\times I}$ computed by the divide-and-conquer method [36] as the actual values. In order to measure the convergence and accuracy of the algorithms, we computed the percentage direction cosine at kth iteration of each adaptive algorithm defined by

$$\text{cosine}(k) = \frac{100|\mathbf{z}_k^*\boldsymbol{\phi}_1|}{\|\mathbf{z}_k\|_2},$$

where $\mathbf{z}_k$ is the approximate principal Takagi vector of $\mathbf{A}$ at kth iteration and $\boldsymbol{\phi}_1$ is the actual principal Takagi vector computed by the divide-and-conquer method. Thus, cosine$(k) \leq 100$ and the larger the more accurate $\mathbf{z}_k$ is.

Example 4.1. A Hankel matrix is a square matrix with constant skew-diagonals (positive sloping diagonals). By definition, an $I \times I$ complex Hankel matrix can be generated by a given sequence $\{x_i \in \mathbb{C} : i = 1, 2, \ldots, 2I-1\}$ which determines its first column and last row.

For a given positive integer I, the first column and the last row of a Hankel matrix $\mathbf{A}$ are given by

$$\begin{cases} a_{k,1} = x_k, & k = 1, 2, \ldots, I, \\ a_{I,k-I+1} = x_k, & k = I, I+1, \ldots, 2I-1, \end{cases}$$

where all x_k are given by [32]:

$$x_k = \exp((-0.01 + 0.04\pi\iota)k) + \exp((-0.02 + 0.44\pi\iota)k).$$

When $I = 1000$, Fig. 1 shows the error

$$\text{ERR}(k) := \|\mathbf{A}\bar{\mathbf{z}}_k - \sigma_k\mathbf{z}_k\|_2,$$

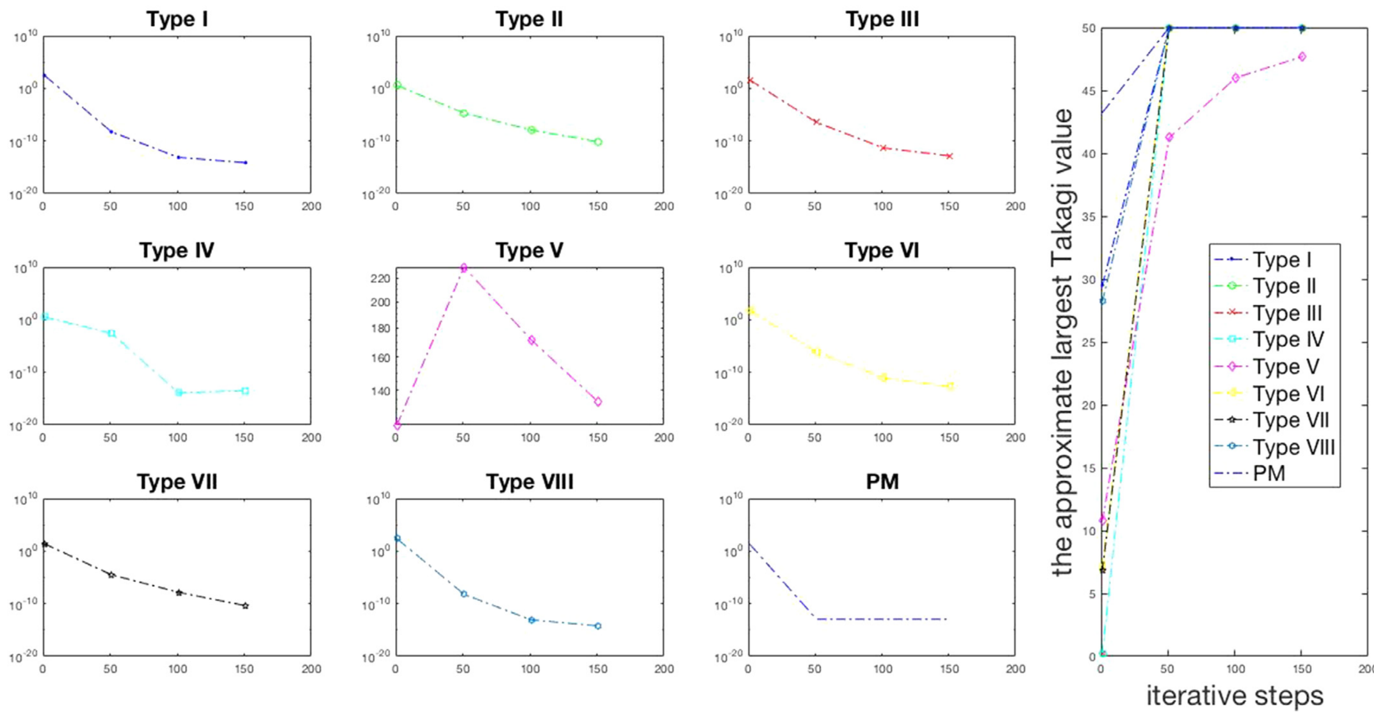


Fig. 1. The error ERR(k) and the principal Takagi value of $\mathbf{A}$ computed at the kth iteration of the proposed algorithms using a given random initial $\mathbf{z}_0$ in Example 4.1 with $\eta_k = 1/(10+k)$.

Table 1
The percentage direction cosines of the principal Takagi vector of $\mathbf{A}$ computed by the adaptive algorithms at iterations $k = 250, 300$ for random initial $\mathbf{z}_0$ in Example 4.1.

$\|\mathbf{z}_0\|_2$	k	Type I	Type II	Type III	Type IV	Type VI	Type VII	Type VIII
1.54	250	94.79	69.05	94.76	99.09	94.78	94.79	94.80
	300	97.23	91.12	97.22	99.85	97.22	97.23	97.24
1.63	250	65.30	49.22	65.17	93.35	65.24	65.31	65.32
	300	76.89	56.32	76.78	99.03	76.83	76.89	76.90
1.94	250	83.05	57.45	82.93	99.59	83.02	83.05	83.07
	300	90.13	70.57	90.05	99.98	90.10	90.13	90.14
2.02	250	96.64	78.01	96.60	99.96	96.60	96.59	97.23
	300	98.23	90.47	98.21	100.00	98.23	98.30	98.57
2.17	250	92.43	63.57	92.36	99.96	92.42	92.43	92.45
	300	95.89	79.39	95.85	100.00	95.88	95.89	95.90
2.26	250	92.08	62.27	91.99	99.98	92.06	92.08	92.10
	300	95.69	77.60	95.64	100.00	95.68	95.69	95.70
2.25	250	99.59	79.62	99.58	100.00	99.59	99.62	99.56
	300	99.79	93.27	99.79	100.00	99.79	99.81	99.77
2.48	250	8.90	3.31	8.85	97.04	8.91	8.91	8.93
	300	12.37	10.62	12.30	99.99	12.38	12.37	12.40
2.26	250	70.30	36.85	70.09	99.89	70.27	70.31	70.35
	300	80.95	49.12	80.79	100.00	80.93	80.95	80.99
2.30	250	75.59	47.24	75.39	99.94	75.56	75.59	75.64
	300	84.95	57.14	84.81	100.00	84.93	84.95	84.99

where $\mathbf{z}_k$ is the principal Takage vector and σ_k is the corresponding Takagi value of $\mathbf{A}$ computed at the kth iteration of each of the eight algorithms using any given random initial $\mathbf{z}_0$ in Example 4.1. These algorithms take 3.18s, 2.52s, 2.94s, 2.73s, 2.47s, 1.70s, 2.53s, 2.63s and 2.70s, respectively. As seen in Fig. 1, Type V (3.18) is the worst. Hence, we do not consider the numerical implementation of the adaptive algorithm with Type V.

When $I = 10$, applying the MATLAB function `svd`, the first two Takagi values of $\mathbf{A}$ are 9.06 and 8.24, respectively. The rest of the Takagi values are much smaller than the first two. The initial values of all the adaptive algorithms are $\mathbf{z}_0 = 0.5\mathbf{r}$, where $\mathbf{r} \in \mathbb{C}^{10}$ is a complex random vector such that the real parts and the imaginary parts of its entries obey the standard normal distribution. For all the algorithms, we set gain factor $\eta_k = 1/(100+k)$. For Type VI (3.20) and Type VIII (3.23), we set $\mu = 0.1$.

The results are summarized in Table 1, where we report the percentage direction cosine after $k = 250$ and 300 iterations for each algorithm. Table 1 shows that Type IV (3.13) converged faster than the others.

Example 4.2. Suppose that I is a positive integer. The testing complex symmetric matrices $\mathbf{A} \in \mathbb{C}^{I\times I}$ were generated by the following steps [37]:

1. generate a random unitary matrix $\mathbf{Q} \in \mathbf{C}^{I\times I}$: $[\mathbf{Q}, \mathbf{R}] = \mathrm{qr}(\mathrm{randn}(I) + \iota\,\mathrm{randn}(I))$;
2. generate a diagonal matrix $\mathbf{D} \in \mathbb{R}^{I\times I}$ with diagonal elements

$$d_{11} = 25, \quad d_{22} = 25/c, \quad d_{kk} = \mathrm{rand}(1), \quad k = 3, 4, \ldots, I,$$

Table 2
The percentage direction cosines of the principal Takagi vector of **A** computed by the adaptive algorithms at iterations $k = 250, 300$ for different complex symmetric matrices **A** with varying ratios σ_1/σ_2 in Example 4.2.

σ_1/σ_2	k	Type I	Type II	Type III	Type IV	Type VI	Type VII	Type VIII
1.10	250	76.28	54.89	76.00	94.41	76.07	76.28	76.29
	300	87.27	63.94	87.06	99.54	87.11	87.24	87.25
1.14	250	82.07	55.96	81.75	98.03	81.83	82.07	82.08
	300	93.12	66.55	92.98	99.95	93.01	93.12	93.13
1.20	250	86.51	56.93	86.18	99.29	86.26	86.51	86.52
	300	96.29	68.89	96.18	99.99	96.20	96.29	96.29
1.23	250	89.82	57.81	89.51	99.73	89.59	89.82	89.83
	300	97.95	70.97	97.88	100.00	97.90	97.95	97.95
1.30	250	92.28	58.60	91.99	99.89	92.07	92.28	92.29
	300	98.84	72.84	98.79	100.00	98.80	98.84	98.84
1.32	250	94.09	59.32	93.84	99.95	93.90	94.09	94.10
	300	99.32	74.50	99.29	100.00	99.29	99.32	99.32
1.40	250	95.43	59.98	95.21	99.98	95.26	95.43	95.44
	300	99.59	75.98	99.57	100.00	99.57	99.59	99.59
1.41	250	96.42	60.58	96.23	99.99	96.28	96.42	96.43
	300	99.74	77.31	99.73	100.00	99.73	99.74	99.74
1.47	250	97.17	61.13	97.01	100.00	97.05	97.17	97.18
	300	99.84	78.50	99.83	100.00	99.83	99.84	99.84
1.50	250	97.74	61.64	97.59	100.00	97.63	97.74	97.74
	300	99.89	79.56	99.89	100.00	99.89	99.90	99.89

Table 3
The percentage direction cosines of the principal Takagi vector of **A** computed by the adaptive algorithms at iterations $k = 250, 300$ for different complex symmetric matrices **A** with varying Takagi values σ_1 and σ_2 in Example 4.3.

σ_1, σ_2	k	Type I	Type II	Type III	Type IV	Type VI	Type VII	Type VIII
11.58,6.32	250	88.95	30.10	88.31	99.97	88.85	88.98	89.08
	300	99.23	38.83	99.20	100.00	99.22	99.24	99.24
16.63,6.49	250	88.92	30.13	88.29	99.96	88.82	88.95	89.05
	300	99.18	38.92	99.14	100.00	99.16	99.18	99.18
11.73,6.92	250	88.69	30.19	88.06	99.94	88.58	88.71	88.81
	300	98.98	39.07	98.93	100.00	98.96	98.98	98.98
11.84,7.18	250	88.71	30.26	88.10	99.93	88.60	88.74	88.83
	300	98.87	39.27	98.83	100.00	98.86	98.87	98.88
12.14,7.64	250	89.13	30.46	88.55	99.91	89.01	89.15	89.23
	300	98.76	39.86	98.71	100.00	98.74	98.76	98.76
12.54,8.08	250	89.80	30.72	89.25	99.91	89.67	89.81	89.88
	300	98.72	40.64	98.68	100.00	98.71	98.72	98.73
12.87,8.67	250	89.90	30.96	89.39	99.89	89.77	89.92	89.97
	300	98.48	41.32	98.43	100.00	98.46	98.48	98.49
13.57,9.33	250	90.99	31.48	90.53	99.88	90.85	91.00	91.04
	300	98.55	42.83	98.50	100.00	98.53	98.55	98.55
14.09,9.88	250	91.49	31.89	91.07	99.88	91.35	91.50	91.53
	300	98.52	43.95	98.47	100.00	98.50	98.52	98.53
17.97,11.66	250	96.96	35.45	96.77	100.00	96.87	96.96	96.98
	300	99.73	53.49	99.72	100.00	99.73	99.73	99.32

where $c \in \{1.10, 1.14, 1.20, 1.23, 1.30, 1.32, 1.40, 1.41, 1.47, 1.50\}$;

3. compute $\mathbf{A} = \mathbf{QDQ}^\top$.

The initial value of all the adaptive algorithms was set to $\mathbf{z}_0 = 0.5\mathbf{r}$, where $\mathbf{r} \in \mathbb{C}^I$ is a complex random vector such that the real parts and the imaginary parts of its entries obey the standard normal distribution. For all the algorithms, we set $\eta_k = 1/(100 + k)$. For Type VI (3.20) and Type VII (3.23), we set $\mu = 0.1$.

For the case of $I = 10$, Table 2 lists the percentage direction cosines after $k = 250$ and 300 iterations for each algorithm. Table 2 shows that Type IV (3.13) converges faster than the others, the convergence behavior of Type I (3.3), Type III (3.11), Type VI (3.20), Type VII (3.23) with $\mu = 1$, and Type VIII (3.23) is similar and the convergence behavior of Type II (3.9) is similar. As expected, the algorithm converges faster when the ratio σ_1/σ_2 is larger.

Example 4.3. The testing complex symmetric matrices were generated the same as Example 4.2 except for the second step, we set

$$\sigma_1 \in \{11.58, 11.63, 11.73, 11.84, 12.14, 12.54, 12.87, 13.57, 14.09, 17.97\};$$

$$\sigma_2 \in \{6.32, 6.49, 6.92, 7.18, 7.64, 8.08, 8.67, 9.33, 11.66\}.$$

So, the diagonal entries of **D** are

$$d_{11} = \sigma_1, \quad d_{22} = \sigma_2, \quad d_{kk} = \text{rand}(1), \quad k = 3, 4, \dots, 10.$$

The initial values of all adaptive algorithms were $\mathbf{z}_0 = 0.5\mathbf{r}$, where $\mathbf{r} \in \mathbb{C}^{10}$ is a complex random vector such that the real parts and the imaginary parts of its entries obey the standard normal distribution. For all the algorithms, we set $\eta_k = 1/(100 + k)$. For Type VI (3.20) and Type VIII (3.23), we set $\mu = 10$.

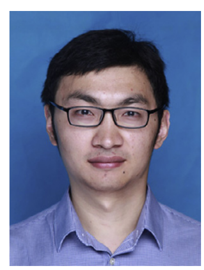
The results are listed in Table 3, showing that the convergence behavior of all adaptive algorithms is similar to the results in Example 4.2.

Acknowledgments

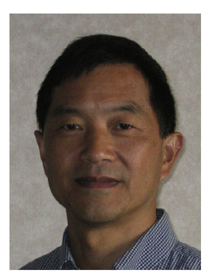
The authors would like to thank the Editor Prof. Zidong Wang and three reviewers for their detailed comments. First author is supported by the Fundamental Research Funds for the Central Universities under grant 221410004005040130, second author is supported by International Cooperation Project of Shanghai Municipal Science and Technology Commission under grant 16510711200 and Natural Science and Engineering Research Council (NSERC) of Canada under grant RGPIN-2014-04252, third author is supported by the National Natural Science Foundation of China under grant 11771099 and International Cooperation Project of Shanghai Municipal Science and Technology Commission under grant 16510711200.

References

[1] I. Aizenberg, Complex-Valued Neural Networks with Multi-Valued Neurons, Studies in Computational Intelligence, 253, Springer-Verlag, New York, 2011.
[2] L. Autonne, Sur les matrices hypohermitiennes et sur les matrices unitaires, Ann. Univ. Lyon 38 (1915) 1–77.
[3] A.D. Bain, Chemical exchange in NMR, Prog. Nuclear Magn. Reson. Spectrosc. 43 (2003) 63–103.
[4] M. Bohner, V.S.H. Rao, S. Sanyal, Global stability of complex-valued neural networks on time scales, Dif. Equ. Dyn. Syst. 19 (2011) 3–11.
[5] P. Bouboulis, Wirtinger's calculus in general Hilbert spaces, arXiv preprint arXiv:1005.5170 (2010).
[6] D. Brandwood, A complex gradient operator and its application in adaptive array theory, IEE Proc. F Commun. Radar Signal Process. 130 (1) (1983) 11–16.
[7] A. Bunse-Gerstner, W.B. Gragg, Singular value decompositions of complex symmetric matrices, J. Comput. Appl. Math. 21 (1) (1988) 41–54.
[8] Y. Chauvin, Principal component analysis by gradient descent on a constrained linear Hebbian cell, Int. Symp. Neural Netw. (1989) 373–380.
[9] M. Che, L. Qi, Y. Wei, Iterative algorithms for computing US- and U-eigenpairs of complex tensors, J. Comput. Appl. Math. 317 (2017) 547–564.
[10] G. Cirrincione, M. Cirrincione, J. Herault, S. Van Huffel, The MCA EXIN neuron for the minor component analysis, IEEE Trans. Neural Netw. 13 (1) (2002) 160–187.
[11] P. Comon, Independent component analysis, a new concept? Signal Proces. 36 (3) (1994) 287–314.
[12] G.H. Golub, C.F. Van Loan, Matrix Computations, fourth ed., Johns Hopkins University Press, Baltimore, MD, 2013.
[13] A. Hirose, Complex-Valued Neural Networks: Theories and Applications, 5, World Scientific, 2003.
[14] Z. Jia, D. Niu, An implicitly restarted refined bidiagonalization lanczos method for computing a partial singular value decomposition, SIAM J. Matrix Anal. Appl. 25 (1) (2003) 246–265.
[15] Z. Jia, D. Niu, A refined harmonic lanczos bidiagonalization method and an implicitly restarted algorithm for computing the smallest singular triplets of large matrices, SIAM J. Sci. Comput. 32 (2) (2010) 714–744.
[16] K. Kreutz-Delgado, The complex gradient operator and the CR-calculus, arXiv preprint arXiv:0906.4835 (2009).
[17] L. Ljung, Analysis of recursive stochastic algorithms, IEEE Trans. Autom. Control 22 (4) (1977) 551–575.
[18] F. Luo, R. Unbehauen, A. Cichocki, A minor component analysis algorithm, Neural Netw. 10 (2) (1997) 291–297.
[19] Y. Miao, Y. Hua, Fast subspace tracking and neural network learning by a novel information criterion, IEEE Trans. Signal Process. 46 (7) (1998) 1967–1979.
[20] T. Nitta, Complex-Valued Neural Networks: Utilizing High-Dimensional Parameters, Hershey, PA, 2009.
[21] E. Oja, Neural networks, principal components, and subspaces, Int. J. Neural Syst. 1 (01) (1989) 61–68.
[22] E. Oja, Principal components, minor components and linear neural networks, Neural Netw. 5 (01) (1992) 927–935.
[23] E. Oja, J. Karhunen, On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix, J. Math. Anal. Appl. 106 (1) (1985) 69–84.
[24] M.D. Plumbley, Lyapunov functions for convergence of principal component algorithms, Neural Netw. 8 (1) (1995) 11–23.
[25] S. Qiao, G. Liu, W. Xu, Block Lanczos tridiagonalization of complex symmetric matrices, in: F.T. Luk (Ed.), Advanced Signal Processing Algorithms, Architectures, and Implementations XV, 5910, Proceedings of SPIE, 2005, pp. 314–324, doi:10.1117/12.615410.
[26] R. Remmert, Theory of Complex Functions, Graduate Texts in Mathematics, 122, Springer-Verlag, New York, 1991, doi:10.1007/978-1-4612-0939-3.
[27] G. Schober, Univalent Functions-Selected Topics, Part of the Lecture Notes in Mathematics book series, 478, Springer-Verlag, New York, 1975.
[28] H.D. Simon, H. Zha, Low-rank matrix approximation using the lanczos bidiagonalization process with applications, SIAM J. Sci. Comput. 21 (6) (1999) 2257–2274.
[29] T. Takagi, On an algebraic problem reluted to an analytic theorem of carathéodory and fejér and on an allied theorem of Landau, Jpn. J. Math. Trans. Abstr. 1 (1924) 83–93.
[30] L.N. Trefethen, Near-circularity of the error curve in complex Chebyshev approximation, J. Approx. Theory 31 (4) (1981) 344–367.
[31] A. Van Den Bos, Complex gradient and Hessian, in: Proceedings of the IEEE Vision, Image and Signal Processing, 141, 1994, pp. 380–383.
[32] S. Van Huffel, Enhanced resolution based on minimum variance estimation and exponential data modeling, Signal Proces. 33 (1993) 333–355.
[33] H. Wang, S. Duan, T. Huang, L. Wang, C. Li, Exponential stability of complex–valued memristive recurrent neural networks, IEEE Trans. Neural Netw. Learn. Syst. 28 (3) (2017) 766–771.
[34] X. Wang, M. Che, Y. Wei, Complex-valued neural networks for the Takagi vector of complex symmetric matrices, Neurocomputing 223 (2017) 77–85.
[35] Z. Wang, Z. Guo, L. Huang, X. Liu, Dynamical behavior of complex-valued Hopfield neural networks with discontinuous activation functions, Neural Proces. Lett. 45 (3) (2017) 1039–1061.
[36] W. Xu, S. Qiao, A divide-and-conquer method for the Takagi factorization, SIAM J. Matrix Anal. Appl. 30 (1) (2008) 142–153.
[37] W. Xu, S. Qiao, A twisted factorization method for symmetric SVD of a complex symmetric tridiagonal matrix, Numer. Linear Algebra Appl. 16 (10) (2009) 801–815.
[38] S. Zhang, Y. Xia, J. Wang, A complex-valued projection neural network for constrained optimization of real functions in complex variables, IEEE Trans. Neural Netw. 26 (12) (2015) 3227–3238.
[39] S. Zhang, Y. Xia, W.X. Zheng, A complex-valued neural dynamical optimization approach and its stability analysis, Neural Netw. 61 (2015) 59–67.



Maolin Che is a Lecturer at School of Economic Mathematics of Southwest University of Finance and Economics in Chengdu. He obtained Ph.D. degree at School of Mathematical Sciences of Fudan University in 2017. He received the B.S. degree in applied mathematics from Neijiang Normal University of Sichuan in China, 2012. His current interests include numerical linear algebra and theory and computation of tensor.



Sanzheng Qiao received his B.S. degree from Shanghai Teacher's University, Shanghai, China in 1966, and the M.S. degree in computer science and Ph.D. degree in applied mathematics from Cornell University,in 1986 and 1987 respectively. He was an assistant professor of computer science at Ithaca College, Ithaca, NY from 1987 to 1988. From 1989 to 1993, he was an assistant professor in the Department of Computer Science and Systems at McMaster University, Hamilton, Ontario, Canada, and an associate professor from 1993 to 1999. He is now a professor emeritus in Department of Computing and Software at McMaster University. His research interests include numerical linear algebra, distributed/parallel scientific computing, numerical methods for signal processing, and numerical software.

Yimin Wei received his B.S. and Ph.D. degrees in computational mathematics from Shanghai Normal University in 1991 and from Fudan University in 1997, respectively. He was a visiting scholar with Division of Engineering and Applied Science of Harvard University from 2000 to 2001. He was a Lecturer of Department of Mathematics at Fudan University from 1997 to 2000. From 2001 to 2005, he was an associate professor in School of Mathematical Science at Fudan University. He is now a Professor with School of Mathematical Science of Fudan University. His research interests include numerical linear and multilinear algebra,sensitivity in computational linear control and perturbation analysis.