A Polynomial Eigenproblem Approach for General Joint Block Diagonalization

Decai Shi[*]   Yunfeng Cai[†]   Shufang Xu[‡]

LMAM & School of Mathematical Sciences
Peking University, Beijing, 100871, China

August 30, 2013

Abstract

Joint Block Diagonalization (JBD) of a given Hermitian matrix set $\mathcal{A} = \{A_i\}_{i=0}^p$ is to find a nonsingular matrix W such that $W^H A_i W$ for $i = 0, 1, \ldots, p$ are all block diagonal matrices with the same prescribed block diagonal structure. General JBD (GJBD) attempts to solve JBD without knowing the block diagonal structure. GJBD arises in Independent Subspace Analysis (ISA)/Blind Source Separation (BSS), and is more difficult than JBD. In this paper, we show that GJBD of $\{A_i\}_{i=0}^p$ is strongly connected with the eigeninformation of the associated matrix polynomial $P_{\mathcal{A}}(\lambda) = \sum_{i=0}^p \lambda^i A_i$. Under the assumption that $P_{\mathcal{A}}(\lambda)$ has only simple eigenvalues, a solvability theory for GJBD is established and the solutions of GJBD are characterized by the eigeninformation. Based on the established theory, a numerical method is proposed to solve GJBD. Numerical tests show that this method is not only feasible for exact GJBD, but also able to handle approximate GJBD to certain extend.

Key words. general joint block diagonalization, polynomial eigenproblem, tensor decomposition

AMS subject classification. 15A69, 65F18.

1   Introduction

In recent years, the problem of estimating the joint eigen-structure of matrix sets (or tensors) has found a variety of applications, especially in Blind Source Separation (BSS), in which one aims to recover source signals from the observed mixtures, without knowing either the distribution of the sources or the mixing process[5][6]. Joint (Block) Diagonalization has become an important tool in Independent Component/Subspace Analysis based BSS. If the source signals are mutual statistical independent, the mixing or demixing system can be estimated by Joint Diagonalization (JD) of a matrix set, which is used in the well-known JADE [8], eJADE [16], SOBI [2] and Hessian ICA [23][17]. Moreover, if there exist groups of signals of interest, where components from different groups are mutually statistically independent and statistical dependence occurs between components

---

[*]decaishi@gmail.com

[†]yfcai@math.pku.edu.cn

[‡]xsf@math.pku.edu.cn

in the same group, known as multidimensional BSS (MBSS) or group BSS, the mixing or demixing system can be solved by Joint Block Diagonalization (JBD) [18, 4].

Theoretically, it is well know that JD for a single Hermitian matrix A_0 can be solved by an unitary matrix; J(B)D for two Hermitian matrices $\{A_1, A_0\}$ can be solved by the eigenvector matrix of the corresponding Generalized Hermitian Eigenvalue Problem (GHEP) $A_0 x = \lambda A_1 x$; J(B)D for more than two matrices is in general impossible, except two special cases: all of the matrices are commutable; three Hermitian matrices satisfying $A_1 A_2^{-1} A_0 = A_0 A_2^{-1} A_1$ with A_2 positive definite [7]. However, in practical applications, solutions to J(B)D do exist, but in "approximate" sense. Hereafter, we call J(B)D with exact and approximate solutions as *exact* J(B)D and *approximate* J(B)D, respectively.

Great efforts have been devoted to solving J(B)D, and there is a long list of studies on this subject, from unitary JD to non-unitary JD, then to unitary and non-unitary JBD. Overall speaking, current numerical methods can be classified into two categories: Jacobi-like methods and gradient-based methods. Jacobi-like methods for unitary J(B)D are achieved by successive Givens rotations, see JADE[8, 3], eJADE[16], SOBI[2] for unitary JD and [1] for unitary JBD. Following the idea of Jacobi rotations, Jacobi-like methods for non-unitary JBD can be achieved via elementary matrix rotations [21]. A number of gradient-based methods have been proposed for J(B)D. To name a few, for JD, the alternate column diagonal center algorithm (ACDC) by Yeredor [24], the gradient descent algorithm by Joho and Mathis [14]; for JBD, several gradient descent methods by Ghennioui *et. al* [12, 13], and by Zhang *et. al* [25]. For more information, we refer interested readers to [20, 25] and reference therein.

Most of the JBD algorithms for MBSS in literature assume that the block diagonal structure is known, which requires that one needs to known the number of groups and the size of each group beforehand. In practice, however, such an assumption does not often hold. Hereafter we will refer JBD with unknown block structure as *general JBD* (GJBD). Due to its difficulties, GJBD is not well studied, both in theory and algorithm. So far, GJBD is solved simply by JD algorithms followed by a permutation recovering algorithm. However, this method is based on a conjecture[1], and is only partially proved[19]. In [9], GJBD for three Hermitian matrices with one of them positive definite, is solved by a quadratic eigenproblem approach, which is actually a byproduct of an inverse eigenvalue problem. Inspired by this idea, we will study GJBD by a Polynomial Eigen-Problem (PEP) approach.

What follows we first give a definition, then formulate Unitary/Non-Unitary JBD and GJBD mathematically.

Definition 1.1. *We call $\tau_n = (n_1, n_2, \ldots, n_t)$ a partition of positive integer n if $n_1, n_2, \ldots, n_t$ are all positive integers and the sum of them is n, i.e.,*

$$n_1 + n_2 + \cdots + n_t = n.$$

The integer t is called the cardinality *of the partition τ_n, denoted by* $\mathrm{card}(\tau_n)$.

Non-Unitary JBD (NU-JBD). Given a matrix set $\mathcal{A} = \{A_i\}_{i=0}^p$ with A_i Hermitian matrix of order n, and a partition $\tau_n = (n_1, n_2, \ldots, n_t)$. Find a nonsingular matrix $W = W(\tau_n)$ such that

$$W^H A_i W = \mathrm{diag}(A_{i1}, A_{i2}, \ldots, A_{it}), \quad \text{for} \quad i = 0, 1, 2, \ldots, p, \tag{1.1}$$

where $A_{ij} \in \mathbb{C}^{n_j \times n_j}$ for $j = 1, 2, \ldots, t$.

If $W = W(\tau_n)$ in (1.1) is required to be unitary, then the problem is referred to as U-JBD. Obviously, NU(U)-JD can be deemed as a special case of NU(U)-JBD by taking $\tau_n = (1, 1, \ldots, 1)$.

2

Non-Unitary GJBD (NU-GJBD). Given a matrix set $\mathcal{A} = \{A_i\}_{i=0}^p$ with A_i Hermitian matrix of order n. Find a partition $\tau_n' = (n_1', n_2', \ldots, n_t')$ and a nonsingular matrix $W = W(\tau_n')$ such that

$$\operatorname{card}(\tau_n') = \max\{\operatorname{card}(\tau_n) \,|\, \text{there exists a nonsingular matrix } W = W(\tau_n) \text{ solves NU-JBD.}\}$$

Notice that if (τ_n, W) is a solution to NU-GJBD for matrix set $\{A_i\}_{i=0}^p$ with $A_p = I_n$, and $W^H A_p W = W^H W = \operatorname{diag}(A_{p1}, A_{p2}, \ldots, A_{pt})$ is block diagonal, then (τ_n, WD) is a solution to U-JBD for matrix set $\{A_i\}_{i=0}^{p-1}$, where $D = \operatorname{diag}(D_1, D_2, \ldots, D_t)$ with $D_j^H A_{pj} D_j = I_{n_j}$ for all $j = 1, 2, \ldots, t$. This provide us a way to solve U-GJBD by NU-GJBD, so *throughout the rest of this paper, we will only need to concentrate on* NU-GJBD.

Our contributions in this paper are three folds: first, we extend the idea in [9] to NU-GJBD, a solvability theory for NU-GJBD is established; second, the solutions to NU-GJBD are also characterized by the eigeninfomation of the corresponding matrix polynomial; third, a numerical method is given, numerical tests show that this method is not only feasible to solve exact NU-GJBD, but also able to handle approximate NU-GJBD to certain extend.

This paper is organized as follows. In section 2, we give some preliminary results on polynomial eigenvalue problem (PEP) and inverse polynomial eigenvalue problem (IPEP). In section 3, a solvability theory for NU-GJBD is established and the solutions are also characterized. The numerical method and numerical examples are given in section 4. Finally, we present some conclusion remarks in section 5.

2 Preliminary

First, we present some results on PEP and Inverse PEP, which will be used in subsequent sections.

We call the matrix polynomial of the form

$$P(\lambda) = \lambda^p A_p + \lambda^{p-1} A_{p-1} + \cdots + A_0 \tag{2.1}$$

regular self-adjoint matrix polynomial, if A_i is Hermitian matrix of order n for $i = 0, 1, \ldots, p$ and A_p is nonsingular. We will show in the next section that GJBD of a Hermitian matrix set $\mathcal{A} = \{A_i\}_{i=0}^p$ is strongly connected with the self-adjoint polynomial $P(\lambda)$ (2.1), denoted by $P_{\mathcal{A}}(\lambda)$.

Given the matrix coefficients $A_0, \ldots, A_p$, the task of finding scalars $\lambda \in \mathbb{C}$ and nonzero vectors $x, y \in \mathbb{C}^n$ satisfying

$$P_{\mathcal{A}}(\lambda)x = 0, \qquad y^H P_{\mathcal{A}}(\lambda) = 0 \tag{2.2}$$

is known as the *polynomial eigenvalue problem* (PEP). The scalars λ and the corresponding nonzero vectors x, y are called, respectively, the eigenvalues, the right eigenvectors and the left eigenvectors of the matrix polynomial $P_{\mathcal{A}}(\lambda)$. It is easy to see that for self-adjoint matrix polynomial, if x is a right eigenvector corresponding to λ, then x is also a left eigenvector corresponding to $\bar{\lambda}$. Similarly, if y is a left eigenvector corresponding to λ, then y is also a right eigenvector corresponding to $\bar{\lambda}$.

It is well known that PEP can be transformed to Generalized Eigenvalue Problems(GEP) via *linearization*. Linearization is not unique, what follows we transform PEP $P_{\mathcal{A}}(\lambda)x = 0$ to a GEP via a special linearization.

For any nonzero vector $x \in \mathbb{C}^n$ and scaler $\lambda \in \mathbb{C}$, define

$$u = u(x, \lambda) := \begin{bmatrix} x^\top & \lambda x^\top & \ldots & \lambda^{p-1} x^\top \end{bmatrix}^\top. \tag{2.3}$$

3

Let

$$
L = \begin{bmatrix} -A_0 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & A_p \\ \vdots & \vdots & \cdot^{\cdot^{\cdot}} & \\ 0 & A_p & & \end{bmatrix}, \quad M = \begin{bmatrix} A_1 & A_2 & \cdots & A_p \\ A_2 & A_3 & \cdot^{\cdot^{\cdot}} & \\ \vdots & \cdot^{\cdot^{\cdot}} & & \\ A_p & & & \end{bmatrix}. \tag{2.4}
$$

Then it follows that

$$
P_{\mathcal{A}}(\lambda)x = 0 \implies Lu = \lambda Mu. \tag{2.5}
$$

Conversely, for any nonzero $u = \begin{bmatrix} u_1^\top & u_2^\top & \dots & u_p^\top \end{bmatrix}^\top \in \mathbb{C}^{pn}$, if $Lu = \lambda Mu$, then u is of the form (2.3) and $P_{\mathcal{A}}(\lambda)x = 0$ with $x = u_1$, i.e.,

$$
Lu = \lambda Mu \implies P_{\mathcal{A}}(\lambda)x = 0. \tag{2.6}
$$

Definition 2.1. [11] *A pair of matrices* $(X, \Lambda) \in \mathbb{C}^{n \times pn} \times \mathbb{C}^{pn \times pn}$ *is called a* standard pair *of* $P_{\mathcal{A}}(\lambda)$ *if and only if*

$$
U = U(X, \Lambda) := \begin{bmatrix} X^\top & (X\Lambda)^\top & \dots & (X\Lambda^{p-1})^\top \end{bmatrix}^\top \tag{2.7}
$$

is nonsingular and it holds

$$
A_p X \Lambda^p + A_{p-1} X \Lambda^{p-1} + \cdots + A_0 X = 0, \tag{2.8}
$$

or equivalently,

$$
LU = MU\Lambda. \tag{2.9}
$$

It follows from (2.5),(2.6) and (2.9) that

$$
P_{\mathcal{A}}(\lambda)x = 0 \iff Lu = \lambda Mu \iff (\lambda I - U \Lambda U^{-1})u = 0 \iff (\lambda I - \Lambda)(U^{-1}u) = 0. \tag{2.10}
$$

Therefore, all eigen-information of $P_{\mathcal{A}}(\lambda)$ can be obtained from a standard pair (X, Λ), vice versa.

We now state the spectral decomposition theorem.

Theorem 2.1. [11](Spectral Decomposition) *Given a standard pair* (X, Λ) *of a regular self-adjoint polynomial* $P_{\mathcal{A}}(\lambda)$. *Let* $U = U(X, \Lambda)$,

$$
S_0 = (U^H M U)^{-1}, \tag{2.11}
$$

where M *is defined in (2.4). Then the coefficient matrices* $\{A_i\}_{i=0}^p$ *of* $P_{\mathcal{A}}(\lambda)$ *can be represented in terms of* X, Λ *and* S_0:

$$
A_i = \begin{cases} (X\Lambda^{p-1}S_0 X^H)^{-1}, & i = p; \\ -\sum_{k=i+1}^p A_k X \Lambda^{k+p-i-1} S_0 X^H A_p, & i = p-1, p-2, \dots, 0. \end{cases} \tag{2.12}
$$

IPEP, in contrast, is to reconstruct the coefficient matrices of a matrix polynomial from some known information of its eigenvalues and eigenvectors, which can be stated as follows when all eigen-information is given.

IPEP: Given $(X, \Lambda) \in \mathbb{C}^{n \times pn} \times \mathbb{C}^{pn \times pn}$ with $U = U(X, \Lambda)$ nonsingular, construct the coefficient matrices $A_0, A_1, \dots, A_p$ of a regular self-adjoint matrix polynomial $P_{\mathcal{A}}(\lambda)$ such that

$$
A_p X \Lambda^p + A_{p-1} X \Lambda^{p-1} + \cdots + A_1 X \Lambda + A_0 X = 0.
$$

In order to state the IPEP theorem, we introduce the following notations.

4

Definition 2.2. *Given* $(X, \Lambda) \in \mathbb{C}^{n \times pn} \times \mathbb{C}^{pn \times pn}$, *define*

1. $\mathbb{S}_\Lambda := \{S \in \mathbb{C}^{pn \times pn} \mid S^H = S, \Lambda S = (\Lambda S)^H.\}$

2. $\mathbb{S}_{(X, \Lambda)} := \{S \in \mathbb{S}_\Lambda \mid X \Lambda^k S X^H = 0, k = 0, 1, \ldots, p - 2.\}$

Note here that $\mathbb{S}_{(X, \Lambda)}$ is a linear subspace of $(np)^2 \mathbb{R}$, [1] and not null since S_0 defined in (2.11) belongs to $\mathbb{S}_{(X, \Lambda)}$.

By applying the well-known GLR theory [10][11][15], it is easy to derive the following result.

Theorem 2.2. (IPEP) *Given* $(X, \Lambda) \in \mathbb{C}^{n \times pn} \times \mathbb{C}^{pn \times pn}$ *with* $U = U(X, \Lambda)$ *nonsingular, if there exists a nonsingular* $S \in \mathbb{S}_{(X, \Lambda)}$, *then* (X, Λ) *is a standard pair of* $P_{\mathcal{A}}(\lambda)$, *whose coefficient matrices* $\{A_i\}_{i=0}^p$ *are given by* (2.12).

For NU-GJBD of a Hermitian matrix set $\mathcal{A} = \{A_i\}_{i=0}^p$, there corresponds a self-adjoint matrix polynomial $P_{\mathcal{A}}(\lambda)$, vice versa. For the ease of our following discussions, *we will make the following two assumptions throughout the rest of the paper.*

    A1. A_p is nonsingular;

    A2. The self-adjoint matrix polynomial $P_{\mathcal{A}}(\lambda)$ has only simple eigenvalues.

Assumption A1 is only a technique requirement, assumption A2 is critical, which leads to a simple representation of the element in $\mathbb{S}_\Lambda$. Direct calculation gives rise to the following lemma.

Lemma 2.1. *For a given matrix* Λ *in the form*

$$\Lambda = \mathrm{diag}(\lambda_1, \bar{\lambda}_1, \ldots, \lambda_\ell, \bar{\lambda}_\ell, \lambda_{2\ell+1}, \ldots, \lambda_{pn}), \tag{2.13}$$

where $Im(\lambda_j) \neq 0$ *for* $j = 1, 2, \ldots, \ell$, $\lambda_j \in \mathbb{R}$ *for* $j = 2\ell + 1, \ldots, pn$ *and* $(\lambda_i - \lambda_j)(\lambda_i - \bar{\lambda}_j) \neq 0$ *for* $i \neq j$, *then* $S \in \mathbb{S}_\Lambda$ *if and only if* S *is in the form*

$$S = \mathrm{diag}\left(\begin{bmatrix} 0 & s_1 \\ \bar{s}_1 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & s_\ell \\ \bar{s}_\ell & 0 \end{bmatrix}, s_{2\ell+1}, \ldots, s_{pn}\right), \tag{2.14}$$

where $s_1, \ldots, s_\ell \in \mathbb{C}$, $s_{2\ell+1}, \ldots, s_{pn} \in \mathbb{R}$.

We will use the following definition to partition eigenvalue matrix Λ and $S \in \mathbb{S}_\Lambda$.

Definition 2.3. *For a given matrix* Λ *in the form* (2.13), *a permutation matrix* P *is said to be* Λ-structure preserving *if*
$$P^\top \Lambda P = \mathrm{diag}(\Lambda_{11}, \Lambda_{22}, \ldots, \Lambda_{kk})$$
and Λ_{jj} *is also in the form* (2.13) *for* $j = 1, \ldots, k$.

Notice that if P is Λ-structure preserving, then it also preserves the structure of $S \in \mathbb{S}_\Lambda$, i.e., $P^\top S P = \mathrm{diag}(S_{11}, S_{22}, \ldots, S_{kk})$ with S_{jj} also in the form (2.14).

The following lemma will be used in the next section, and the proof of it is similar to that of Lemma 3.1 in [9].

---

[1] The Hermitian complex matrices of n-by-n do not form a subspace over $\mathbb{C}$, but a subspace of $n^2 \mathbb{R}$ over $\mathbb{R}$.

Lemma 2.2. *Given $(X, \Lambda) \in \mathbb{C}^{n \times pn} \times \mathbb{C}^{pn \times pn}$ with $U = U(X, \Lambda)$ nonsingular. For any $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ nonsingular, we have*

$$U \widetilde{S} S^{-1} U^{-1} = I_p \otimes (\widetilde{A}_p^{-1} A_p), \tag{2.15}$$

where A_p and $\widetilde{A}_p$ are given by (2.12) with $S_0 = S$ and $S_0 = \widetilde{S}$, respectively, and the symbol $\otimes$ denotes the Kronecker product.

We will see later that the number of distinct eigenvalues of $\widetilde{S} S^{-1}$ and their corresponding multiplicities play an important role in NU-GJBD. Now we summarize the properties of the eigenvalus of $\widetilde{S} S^{-1}$, and then give some definitions for the ease of the following discussions.

P1. For any $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ nonsingular, it follows from Lemma 2.1 that

$$\widetilde{S} S^{-1} = \text{diag}(\theta_1, \bar{\theta}_1, \ldots, \theta_\ell, \bar{\theta}_\ell, \theta_{2\ell+1}, \ldots, \theta_{pn}), \tag{2.16}$$

where $\theta_j = \frac{\tilde{s}_j}{s_j}$ for all j, and $\theta_j \in \mathbb{C}$ for $j = 1, 2, \ldots, \ell$, $\theta_j \in \mathbb{R}$ for $j = 2\ell + 1, \ldots, pn$.

P2. For any $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ nonsingular, it follows from Lemma 2.2 that each eigenvalue of $\widetilde{S} S^{-1}$ repeats p times.

Definition 2.4. *For any $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ nonsingular, let $\lambda_d(\widetilde{S} S^{-1}) = \{\mu_1, \mu_2, \ldots, \mu_t\}$ be the distinct eigenvalues of $\widetilde{S} S^{-1}$ within the closed upper complex plane $i\mathbb{C}^+ = \{a + bi \mid a \in \mathbb{R}, b \geq 0\}$, $Im(\mu_j) > 0$ for $j = 1, \ldots, \hat{\ell}$, $\mu_j \in \mathbb{R}$ for $j = \hat{\ell} + 1, \ldots, t$. [2] Assume the multiplicity of μ_j is pm_j, then define eigenvalue multiplicity vector ζ_n for $\widetilde{S} S^{-1}$ as*

$$\zeta_n = \zeta_n(\widetilde{S} S^{-1}) = (2m_1, \ldots, 2m_{\hat{\ell}}, m_{\hat{\ell}+1}, \ldots, m_t). \tag{2.17}$$

and the eigenvalue multiplicity vector ζ_n^{opt} for $\mathbb{S}_{(X, \Lambda)}$ as

$$\zeta_n^{\text{opt}} = \zeta_n^{\text{opt}}(\mathbb{S}_{(X, \Lambda)}) = \text{argmax}\{\text{card}(\zeta_n) \mid \zeta_n = \zeta_n(\widetilde{S} S^{-1}), S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)} \text{ nonsingular.}\} \tag{2.18}$$

REMARK **2.1.**

1. Notice here that ζ_n is a partition of n, and $t = \text{card}(\zeta_n)$.

2. $\text{card}(\zeta_n(\widetilde{S} S^{-1}))$ is the number of distinct eigenvalues of $\widetilde{S} S^{-1}$ within the upper closed complex plane, and $\text{card}(\zeta_n^{\text{opt}})$ is the maximal.

3. In the definition of ζ_n^{opt}, S can be fixed as S_0, i.e.,

$$\zeta_n^{\text{opt}} = \zeta_n^{\text{opt}}(\mathbb{S}_{(X, \Lambda)}) = \text{argmax}\{\text{card}(\zeta_n) \mid \zeta_n = \zeta_n(\widetilde{S} S_0^{-1}), \widetilde{S} \in \mathbb{S}_{(X, \Lambda)} \text{ nonsingular.}\} \tag{2.19}$$

As a matter of fact, if $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ are chosen such that $\zeta_n^{\text{opt}}(\mathbb{S}_{(X, \Lambda)}) = \zeta_n(\widetilde{S} S^{-1})$, then let $\widetilde{S} S^{-1} = \text{diag}(\theta_j)$, $S S_0^{-1} = \text{diag}(\theta_j^{(0)})$, for any $\alpha \in \mathbb{R}$ and

$$\alpha \notin \left\{\alpha \mid \alpha = \frac{\theta_j \theta_j^{(0)} - \theta_i \theta_i^{(0)}}{\theta_i^{(0)} - \theta_j^{(0)}}, \theta_i^{(0)} - \theta_j^{(0)} \neq 0\right\},$$

it holds $\zeta_n^{\text{opt}}(\mathbb{S}_{(X, \Lambda)}) = \zeta_n((\widetilde{S} + \alpha S) S_0^{-1})$.

---

[2]Note that $\hat{\ell} \leq \ell$ since for $j = 1, \ldots, \ell$, $\theta_j \in \mathbb{C}$ in (2.16) can be real.

3 NU-GJBD Theory

In this section, we will establish a theory on NU-GJBD. First, we give the following theorem, which gives necessary and sufficient conditions for NU-JBD.

Theorem 3.1. *Let (X, Λ) be a standard pair of $P_{\mathcal{A}}(\lambda)$, $U = U(X, \Lambda)$ and $S_0 = \left(U^H M U\right)^{-1}$. Given a partition $\tau_n = (n_1, n_2, \ldots, n_t)$, then the following statements are equivalent:*

(1). *There exists a nonsingular matrix $W = W(\tau_n)$ which solves NU-JBD;*

(2). *There exists a nonsingular matrix $W = W(\tau_n)$ and a Λ-structure preserving matrix P such that*

$$W^{-1}XP = \operatorname{diag}(X_{11}, X_{22}, \ldots, X_{tt}), \tag{3.1}$$

where $X_{jj} \in \mathbb{C}^{n_j \times p n_j}$ for $j = 1, 2, \ldots, t$;

(3). *There exists a nonsingular matrix $\widetilde{S} \in \mathbb{S}_{(X,\Lambda)}$ and a Λ-structure preserving matrix P such that*

$$P^\top \widetilde{S} S_0^{-1} P = \operatorname{diag}(D_{11}, D_{22}, \ldots, D_{tt}), \tag{3.2}$$

where $D_{jj} \in \mathbb{C}^{p n_j \times p n_j}$ for $j = 1, 2, \ldots, t$ with $\lambda(D_{ii}) \cap \lambda(D_{jj}) = \emptyset$ for $i \neq j$.

Proof. We proceed by showing that $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (1)$.

$(1) \Rightarrow (2)$. Given τ_n, suppose $W = W(\tau_n)$ solves NU-JBD. Then there exists a Λ-structure preserving matrix P such that $P^\top \Lambda P = \operatorname{diag}(\Lambda_{11}, \Lambda_{22}, \ldots, \Lambda_{tt})$ with the spectrum of Λ_{jj} equals the spectrum of $P_j(\lambda) = \sum_{i=0}^{p} \lambda^i A_{ij}$ for $j = 1, \ldots, t$. Now we can rewrite $\sum_{i=0}^{p} A_i X \Lambda^i = 0$ as

$$\sum_{i=0}^{p} (W^H A_i W)(W^{-1}XP)(P^\top \Lambda P)^i = 0.$$

That is

$$\sum_{i=0}^{p} \operatorname{diag}\left(A_{i1}, A_{i2}, \ldots, A_{it}\right)(W^{-1}XP) \operatorname{diag}(\Lambda_{11}, \Lambda_{22}, \ldots, \Lambda_{tt})^i = 0.$$

Denote

$$W^{-1}XP = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1t} \\ X_{21} & X_{22} & \cdots & X_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t1} & X_{t2} & \cdots & X_{tt} \end{bmatrix}, \tag{3.3}$$

then we have

$$\sum_{i=0}^{p} A_{il} X_{lj} \Lambda_{jj}^i = 0, \quad l, j = 1, \ldots, t.$$

Since every eigenvalue of Λ is simple, it follows $X_{lj} = 0$ for $l \neq j$. Therefore, (3.1) holds.

$(2) \Rightarrow (3)$. Since P is a Λ-structure preserving matrix, we can denote

$$P^\top \Lambda P = \operatorname{diag}(\Lambda_{11}, \Lambda_{22}, \ldots, \Lambda_{tt}), \quad P^\top S_0 P = \operatorname{diag}(S_{11}, S_{22}, \ldots, S_{tt}),$$

where $\Lambda_{jj}, S_{jj} \in \mathbb{C}^{pn_j \times pn_j}$ for $j = 1, \ldots, t$. Then it follows $S_{jj} \in \mathbb{S}_{(X_{jj}, \Lambda_{jj})}$. Let

$$\widetilde{S} = P \operatorname{diag}(S_{11}, 2S_{22}, \ldots, tS_{tt}) P^\top,$$

then $\widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ is nonsingular and

$$P^\top \widetilde{S} S_0^{-1} P = \operatorname{diag}(I_{n_1}, 2I_{n_2}, \ldots, tI_{n_t}),$$

which means (3) holds.

(3) $\Rightarrow$ (1). According to Lemma 2.2, we have

$$U \widetilde{S} S_0^{-1} U^{-1} = I_p \otimes (\widetilde{A}_p^{-1} A_p), \tag{3.4}$$

where A_p and $\widetilde{A}_p$ are given by (2.12) with S_0 and $\widetilde{S}$, respectively. Since

$$P^\top \widetilde{S} S_0^{-1} P = \operatorname{diag}(D_{11}, D_{22}, \ldots, D_{tt}), \tag{3.5}$$

with $\lambda(D_{ii}) \cap \lambda(D_{jj}) = \emptyset$ for $i \neq j$, it follows that there exists a nonsingular matrix W such that

$$W^{-1} \widetilde{A}_p^{-1} A_p W = \operatorname{diag}(F_{11}, F_{22}, \ldots, F_{tt}), \tag{3.6}$$

where $F_{jj} \in \mathbb{C}^{n_j \times n_j}$ and $\lambda(F_{jj}) \subset \lambda(D_{jj})$ for $j = 1, 2, \ldots, t$.

From (3.4) and (3.5) we also have

$$\widetilde{A}_p^{-1} A_p X P = X P \operatorname{diag}(D_{11}, D_{22}, \ldots, D_{tt}). \tag{3.7}$$

Partition $W^{-1} X P$ as the form (3.3) and substitute (3.6) into (3.7) yield

$$\begin{bmatrix} F_{11} & & & \\ & F_{22} & & \\ & & \ddots & \\ & & & F_{tt} \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1t} \\ X_{21} & X_{22} & \cdots & X_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t1} & X_{t2} & \cdots & X_{tt} \end{bmatrix} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1t} \\ X_{21} & X_{22} & \cdots & X_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ X_{t1} & X_{t2} & \cdots & X_{tt} \end{bmatrix} \begin{bmatrix} D_{11} & & & \\ & D_{22} & & \\ & & \ddots & \\ & & & D_{tt} \end{bmatrix}.$$

Comparing the (i, j)−block of both sides of the equation, we get

$$F_{ii} X_{ij} = X_{ij} D_{jj}.$$

It follows that $X_{ij} = 0$ whenever $i \neq j$, since $\lambda(F_{ii}) \cap \lambda(D_{jj}) = \emptyset$. Therefore,

$$W^{-1} X P = \operatorname{diag}(X_{11}, X_{22}, \ldots, X_{tt}).$$

Since P is a Λ-structure preserving matrix, we have

$$P^\top \Lambda P = \operatorname{diag}(\Lambda_{11}, \Lambda_{22}, \ldots, \Lambda_{tt}), \qquad P^\top S_0 P = \operatorname{diag}(S_{11}, S_{22}, \ldots, S_{tt}).$$

According to (2.12), we have

$$\begin{aligned}
A_p^{-1} &= X \Lambda^{p-1} S_0 X^H = W(W^{-1} X P)(P^\top \Lambda P)^{p-1}(P^\top S_0 P)(W^{-1} X P)^H W^H \\
&= W \operatorname{diag}(X_{11}, \ldots, X_{tt}) \operatorname{diag}(\Lambda_{11}, \ldots, \Lambda_{tt})^{p-1} \operatorname{diag}(S_{11}, \ldots, S_{tt}) \operatorname{diag}(X_{11}^H, \ldots, X_{tt}^H) W^H \\
&= W \operatorname{diag}(X_{11} \Lambda_{11}^{p-1} S_{11} X_{11}^H, \ldots, X_{tt} \Lambda_{tt}^{p-1} S_{tt} X_{tt}^H) W^H \\
&= W \operatorname{diag}\left((A_{p1})^{-1}, \ldots, (A_{pt})^{-1}\right) W^H,
\end{aligned}$$

8

and hence

$$W^H A_p W = \text{diag}(A_{p1}, A_{p2}, \ldots, A_{pt}).$$

Similarly, from (2.12) we can recursively get

$$W^H A_i W = \text{diag}(A_{i1}, A_{i2}, \ldots, A_{it})$$

for $i = p - 1, \ldots, 0$. $\qquad\qquad\square$

REMARK **3.1.**

1. Lemma 3.2 in [9] actually proves (3) $\Rightarrow$ (1) of Theorem 3.1 for $p = 2$. Here we proves the equivalences of the three statements.

2. The equality (3.1) plays a central role in Theorem 3.1, W in (3.1) solves NU-JBD, P in (3.1) is the same as P in (3.2).

Now we are ready to give the NU-GJBD theorem.

Theorem 3.2. *Let (X, Λ) be a standard pair of $P_{\mathcal{A}}(\lambda)$, $U = U(X, \Lambda)$ and $S_0 = \left(U^H M U\right)^{-1}$. Then there exists a nonsingular matrix W such that $(\zeta_n^{\text{opt}}, W)$ solves NU-GJBD, where ζ_n^{opt} is defined in (2.18), or equivalently (2.19).*

Proof. Assume $\zeta_n^{\text{opt}} = (n_1, n_2, \ldots, n_t)$. Then there exists a nonsingular matrix $\widetilde{S} \in \mathbb{S}_{(X,\Lambda)}$ and a Λ-structure preserving matrix P such that

$$P^\top \widetilde{S} S_0^{-1} P = \text{diag}(D_{11}, D_{22}, \ldots, D_{tt}),$$

where $D_{jj} \in \mathbb{C}^{pn_j \times pn_j}$ with $\lambda(D_{ii}) \cap \lambda(D_{jj}) = \emptyset$ for $i \neq j$. Then according to Theorem 3.1, there exists a nonsingular matrix W which solves NU-JBD. Consequently, it is sufficient if we can prove that if $(\tau_n, \widetilde{W})$ also solves NU-GJBD, then $\text{card}(\tau_n) \leq \text{card}(\zeta_n^{\text{opt}})$. What follows we prove this by contradiction.

Assume there exists $(\tau_n, \widetilde{W})$ solves NU-GJBD with $\text{card}(\tau_n) > \text{card}(\zeta_n^{\text{opt}})$. Therefore, $\widetilde{W}$ solves NU-JBD. Let $\tau_n = (n'_1, n'_2, \ldots, n'_{\tilde{t}})$, then using Theorem 3.1, there exists a nonsingular $\widetilde{S} \in \mathbb{S}_{(X,\Lambda)}$ and a Λ-structure preserving matrix P such that

$$P^\top \widetilde{S} S_0^{-1} P = \text{diag}(D'_{11}, D'_{22}, \ldots, D'_{\tilde{t}\tilde{t}}),$$

where $D'_{jj} \in \mathbb{C}^{pn'_j \times pn'_j}$ with $\lambda(D'_{ii}) \cap \lambda(D'_{jj}) = \emptyset$ for $i \neq j$. Consequently, $\text{card}(\zeta_n(\widetilde{S} S_0^{-1})) = \tilde{t} > \text{card}(\zeta_n^{\text{opt}})$, which contradicts with the definition of ζ_n^{opt}. $\qquad\square$

Theorem 3.2 not only tells the existence of the solutions to NU-GJBD, but also together with Theorem 3.1 characterize the solution by the eigeninformation of the corresponding polynomial eigenvalue problem. See the following example.

Example 3.1. *Consider the Hermitian matrix set $\mathcal{A} = \{A_2, A_1, A_0\}$, where*

$$A_2 = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -16 & 14 \\ 0 & 0 & 14 & -16 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 & -10+3i & 10+3i \\ 0 & 0 & -10-3i & 10-3i \\ -10-3i & -10+3i & 0 & 0 \\ 10-3i & 10+3i & 0 & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} -8 & -4 & 0 & 0 \\ -4 & -8 & 0 & 0 \\ 0 & 0 & 12 & -8 \\ 0 & 0 & -8 & 12 \end{bmatrix}.$$

The corresponding quadratic matrix polynomial is

$$Q_{\mathcal{A}}(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0,$$

and its eigeninformation can be given by (X, Λ), *where*

$$X = \begin{bmatrix} -1 & 1 & -1 & 1 & 2+i & 2-i & 1+2i & 1-2i \\ 1 & -1 & 1 & -1 & 2+i & 2-i & 1+2i & 1-2i \\ 1 & 1 & 1 & 1 & i & -i & -1 & -1 \\ 1 & 1 & 1 & 1 & -i & i & 1 & 1 \end{bmatrix}, \quad \Lambda = \mathrm{diag}(i, -i, 2i, -2i, -1+i, -1-i, 1+i, 1-i).$$

By calculations, we have

$$S_0 = \frac{1}{8} \, \mathrm{diag}\left(\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \frac{1}{30}\begin{bmatrix} 0 & -2+i \\ -2-i & 0 \end{bmatrix}, \frac{1}{30}\begin{bmatrix} 0 & -2-i \\ -2+i & 0 \end{bmatrix} \right),$$

and for any $S \in \mathbb{S}_{(X,\Lambda)}$, *it is in the following form*

$$S = \mathrm{diag}\left(\begin{bmatrix} 0 & s \\ \bar{s} & 0 \end{bmatrix}, \begin{bmatrix} 0 & -s \\ -\bar{s} & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2r - ri \\ 2r + ri & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2r + ri \\ 2r - ri & 0 \end{bmatrix} \right),$$

where $s \in \mathbb{C}$, $r \in \mathbb{R}$ *are arbitrary nonzero numbers. Then we have*

$$S S_0^{-1} = 8 \, \mathrm{diag}(si, -\bar{s}i, si, -\bar{s}i, -30r, -30r, -30r, -30r).$$

Consequently, by choosing $s = \frac{1}{8}$, $r = -\frac{1}{240}$, *we have* $\zeta_n^{\mathrm{opt}} = \frac{1}{p}(2 \times 2, 4) = (2, 2)$ *and* $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = 2$. *Using Theorem 3.2, we know there exists a nonsingular* W *such that* $W^H A_i W = \mathrm{diag}(A_{i1}, A_{i2})$, *where* A_{i1}, A_{i2} *are all 2-by-2 Hermitian matrices for* $i = 0, 1, 2$.

Now given $\tau_n = \zeta_n^{\mathrm{opt}} = (2, 2)$, *using Theorem 3.1, in order to determine* W, *one need to first find a* Λ*-structure preserving matrix* P *satisfying (3.2), then from (3.1) one can obtain* W. *In this example,* $P = I_8$, *and* W *can be obtained by computing a row operation matrix* W^{-1} *such that* $W^{-1}[x_1 \, x_2 \, Re(x_5) \, Im(x_5)] = \mathrm{diag}(X_{11}, X_{22})$, *where* X_{11}, X_{22} *are all 2-by-2 matrices. This is due to the fact that* $\mathrm{span}([x_1 \, x_2 \, x_3 \, x_4]) = \mathrm{span}([x_1 \, x_2])$, $\mathrm{span}([x_5 \, x_6 \, x_7 \, x_8]) = \mathrm{span}([Re(x_5) \, Im(x_5)])$.

REMARK **3.2.** The procedure in the above example actually give us a recipe to compute the solution $(\zeta_n^{\mathrm{opt}}, W)$ of NU-GJBD.

Next, we try to establish the relationship between $\mathrm{card}(\zeta_n^{\mathrm{opt}})$ and the dimension of $\mathbb{S}_{(X,\Lambda)}$. Let us first see the following two examples.

Example 3.2. *Consider the Hermitian matrix set* $\mathcal{A} = \{A_2, A_1, A_0\}$, *where*

$$A_2 = \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \quad A_1 = \begin{bmatrix} -2 & -4 \\ -4 & -5 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 2 & 4 \\ 4 & 4 \end{bmatrix}.$$

The corresponding quadratic matrix polynomial is

$$Q_{\mathcal{A}}(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0,$$

whose eigeninformation can by given by (X, Λ):

$$X = \begin{bmatrix} 1 & 1 & -2 & -2 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad \Lambda = \mathrm{diag}(1+i, 1-i, 1, -4).$$

10

By calculations, we have

$$S_0 = \left(U(X,\Lambda)^H M U(X,\Lambda)\right)^{-1} = \mathrm{diag}\left(\frac{1}{2}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \frac{1}{5}, -\frac{1}{5}\right),$$

and $S \in \mathbb{S}_{(X,\Lambda)}$ if and only if S is of the form

$$S = \mathrm{diag}\left(s\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, r, -r\right),$$

where $s, r \in \mathbb{R}$. Therefore, we have

$$\dim(\mathbb{S}_{(X,\Lambda)}) = 2.$$

On the other hand,

$$SS_0^{-1} = \mathrm{diag}(-2s, -2s, 5r, 5r),$$

and hence $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = 2$. So we have $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = \dim(\mathbb{S}_{(X,\Lambda)})$.

Example 3.3. *Consider the Hermitian matrix set $\mathcal{A} = \{A_2, A_1, A_0\}$, where*

$$A_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 3i \\ -3i & 0 \end{bmatrix}, \quad A_0 = \begin{bmatrix} 0 & -2 \\ -2 & 0 \end{bmatrix}.$$

The corresponding quadratic matrix polynomial is

$$Q_{\mathcal{A}}(\lambda) = \lambda^2 A_2 + \lambda A_1 + A_0,$$

whose eigeninformation can be given by (X, Λ):

$$X = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}, \quad \Lambda = \mathrm{diag}(i, -i, 2i, -2i).$$

By calculations, we have

$$S_0 = \left(U(X,\Lambda)^H M U(X,\Lambda)\right)^{-1} = \mathrm{diag}\left(\begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix}, \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\right),$$

and $S \in \mathbb{S}_{(X,\Lambda)}$ if and only if S is of the form

$$S = \mathrm{diag}\left(\begin{bmatrix} 0 & s \\ \bar{s} & 0 \end{bmatrix}, \begin{bmatrix} 0 & -s \\ -\bar{s} & 0 \end{bmatrix}\right),$$

where $s \in \mathbb{C}$. Therefore, we have

$$\dim(\mathbb{S}_{(X,\Lambda)}) = 2.$$

On the other hand,

$$SS_0^{-1} = \mathrm{diag}(-si, \bar{s}i, -si, \bar{s}i).$$

and hence $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = 1$. So we have $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = \frac{1}{2}\dim(\mathbb{S}_{(X,\Lambda)})$.

Example 3.4. *In Example 3.1, $\frac{1}{2}\dim(\mathbb{S}_{(X,\Lambda)}) < \mathrm{card}(\zeta_n^{\mathrm{opt}}) < \dim(\mathbb{S}_{(X,\Lambda)})$.*

11

As a matter of fact, we have the following inequality which reveals the relationship between $\mathrm{card}(\zeta_n^{\mathrm{opt}})$ and $\dim(\mathbb{S}_{(X,\Lambda)})$.

Theorem 3.3.
$$\frac{1}{2}\dim(\mathbb{S}_{(X,\Lambda)}) \le \mathrm{card}(\zeta_n^{\mathrm{opt}}) \le \dim(\mathbb{S}_{(X,\Lambda)}).$$

Proof. Assume $(\zeta_n^{\mathrm{opt}}, W)$ solves NU-GJBD with $\zeta_n^{\mathrm{opt}} = (n_1, \ldots, n_t)$. The eigeninformation of $P_{\mathcal{A}}(\lambda)$ is given by (X, Λ). Then there exists a Λ-structure preserving matrix P such that

$$W^{-1}XP = \mathrm{diag}(X_{11}, \ldots, X_{tt}), \qquad P^{\top}\Lambda P = \mathrm{diag}(\Lambda_{11}, \ldots, \Lambda_{tt}),$$

where $X_{jj} \in \mathbb{C}^{n_j \times pn_j}$, $\Lambda_{jj} \in \mathbb{C}^{pn_j \times pn_j}$. Now define $\mathbb{S}_{(X_{jj}, \Lambda_{jj})}$ and $\zeta_{n_j}^{\mathrm{opt}}(\mathbb{S}_{(X_{jj}, \Lambda_{jj})})$ as $\mathbb{S}_{(X,\Lambda)}$ and $\zeta_n^{\mathrm{opt}}(\mathbb{S}_{(X,\Lambda)})$ with $X = X_{jj}$, $\Lambda = \Lambda_{jj}$, respectively. Then we know that

$$\dim(\mathbb{S}_{(X,\Lambda)}) = \sum_{j=1}^{t} \dim(\mathbb{S}_{(X_{jj}, \Lambda_{jj})}), \qquad \mathrm{card}(\zeta_{n_j}^{\mathrm{opt}}(\mathbb{S}_{(X_{jj}, \Lambda_{jj})})) = 1.$$

Next we show our conclusion by showing that $\dim(\mathbb{S}_{(X_{jj}, \Lambda_{jj})}) = 1$ or 2. Define $U_{jj} = U(X_{jj}, \Lambda_{jj})$ as $U(X, \Lambda)$ with $X = X_{jj}$ and $\Lambda = \Lambda_{jj}$, M_{jj} as M in (2.4) with $A_i = A_{ij}$, $S_{jj}^{(0)} = (U_{jj}^H M_{jj} U_{jj})^{-1}$. Then for any $S_{jj} \in \mathbb{S}_{(X_{jj}, \Lambda_{jj})}$, the eigenvalues of $S_{jj}(S_{jj}^{(0)})^{-1}$ are the same real number or complex conjugates. In the former case, there exists $\alpha \in \mathbb{R}$ such that $S_{jj} = \alpha S_{jj}^{(0)}$, and hence $\dim(\mathbb{S}_{(X_{jj}, \Lambda_{jj})}) = 1$; in the latter case, if $\dim(\mathbb{S}_{(X_{jj}, \Lambda_{jj})}) > 2$, there exist $S_{jj}^{(1)}, S_{jj}^{(2)} \in \mathbb{S}_{(X_{jj}, \Lambda_{jj})}$ such that $\lambda(S_{jj}^{(k)}(S_{jj}^{(0)})^{-1}) = \{\alpha_k \pm i\beta_k\}$ with $\beta_k \ne 0$ for $k = 1, 2$, and

$$\frac{1}{\beta_1}(S_{jj}^{(1)} - \alpha_1 S_{jj}^{(0)})(S_{jj}^{(0)})^{-1} \ne \pm \frac{1}{\beta_2}(S_{jj}^{(2)} - \alpha_2 S_{jj}^{(0)})(S_{jj}^{(0)})^{-1}.$$

Let

$$S_{jj} = \frac{1}{\beta_1}(S_{jj}^{(1)} - \alpha_1 S_{jj}^{(0)}) + \frac{2}{\beta_2}(S_{jj}^{(2)} - \alpha_2 S_{jj}^{(0)}),$$

then $S_{jj} \in \mathbb{S}_{(X_{jj}, \Lambda_{jj})}$, and $\lambda(S_{jj}(S_{jj}^{(0)})^{-1}) = \{\pm 3i, \pm i\}$. Consequently, $\mathrm{card}(S_{jj}(S_{jj}^{(0)})^{-1}) = 2$, which contradicts with $\mathrm{card}(\zeta_{n_j}^{\mathrm{opt}}(\mathbb{S}_{(X_{jj}, \Lambda_{jj})})) = 1$. $\qquad\square$

REMARK **3.3.** We can see from the proof that $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = \dim(\mathbb{S}_{(X,\Lambda)})$ if and only if for any $S \in \mathbb{S}_{(X,\Lambda)}$, the eigenvalues of SS_0^{-1} are all real numbers. If A_p is positive definite, then using Lemma 2.2, all eigenvalues of SS_0^{-1} are real. Therefore, $\mathrm{card}(\zeta_n^{\mathrm{opt}}) = \dim(\mathbb{S}_{(X,\Lambda)})$, which agrees with the conclusion Theorem 3.2 in [9].

Let $\tau_n = (n_1, \ldots, n_t)$, and assume (τ_n, W) solves NU-GJBD, then $(\tau_n\Pi_t, WT\Pi)$ also solves NU-GJBD, where $\Pi_t \in \mathbb{R}^{t \times t}$ is a permutation matrix, $\Pi \in \mathbb{R}^{n \times n}$ is the *corresponding block permutation matrix*, which can be obtained by replacing the j-th column of Π_t by I_{n_j}, T is a *block scaling matrix* of the form $T = \mathrm{diag}(T_{11}, \ldots, T_{tt})$ with $T_{jj} \in \mathbb{C}^{n_j \times n_j}$. The "uniqueness" problem arises: is the solution unique up to such block permutation and block scaling? The answer is yes. In fact, we have the following theorem.

Theorem 3.4. *The solution to* NU-GJBD *is unique up to block permutation and block scaling, i.e., if (τ_n, W) and $(\tilde{\tau}_n, \widetilde{W})$ both solve* NU-GJBD, *then there exists a permutation matrix $\Pi_t \in \mathbb{R}^{t \times t}$ and a block scaling matrix T such that*

$$\tilde{\tau}_n = \tau_n \Pi_t, \qquad \widetilde{W} = WT\Pi, \tag{3.8}$$

where $\Pi \in \mathbb{R}^{n \times n}$ is the block permutation matrix corresponding with Π_t.

12

Proof. As (τ_n, W) and $(\tilde{\tau}_n, \widetilde{W})$ both solve NU-GJBD, then we know that $\mathrm{card}(\tau_n) = \mathrm{card}(\tilde{\tau}_n) = \mathrm{card}(\zeta_n^{\mathrm{opt}})$, and hence we can assume $\tau_n = (n_1, \ldots, n_t)$, $\tilde{\tau}_n = (\tilde{n}_1, \ldots, \tilde{n}_t)$, where $t = \mathrm{card}(\zeta_n^{\mathrm{opt}})$.

Let (X, Λ) be the eigenformation matrix pair of $P_{\mathcal{A}}(\lambda)$, $S_0 = (U(X, \Lambda)^H M U(X, \Lambda))^{-1}$ with M defined in (2.4). It follows from Theorem 3.1 that there exist nonsingular $S, \widetilde{S} \in \mathbb{S}_{(X, \Lambda)}$ and Λ-structure preserving matrices $P, \widetilde{P}$ such that

$$P^\top S S_0^{-1} P = \mathrm{diag}(D_{11}, \ldots, D_{tt}), \qquad \widetilde{P}^\top \widetilde{S} S_0^{-1} \widetilde{P} = \mathrm{diag}(\widetilde{D}_{11}, \ldots, \widetilde{D}_{tt}), \tag{3.9}$$

where $D_{jj} \in \mathbb{C}^{pn_j \times pn_j}$, $\widetilde{D}_{jj} \in \mathbb{C}^{p\tilde{n}_j \times p\tilde{n}_j}$ for $j = 1, 2, \ldots, t$, and $\lambda(D_{ii}) \cap \lambda(D_{jj}) = \emptyset$, $\lambda(\widetilde{D}_{ii}) \cap \lambda(\widetilde{D}_{jj}) = \emptyset$ for $i \neq j$. And also

$$W^{-1} X P = \mathrm{diag}(X_{11}, \ldots, X_{tt}), \qquad \widetilde{W}^{-1} X \widetilde{P} = \mathrm{diag}(\widetilde{X}_{11}, \ldots, \widetilde{X}_{tt}), \tag{3.10}$$

where $X_{jj} \in \mathbb{C}^{n_j \times pn_j}$, $\widetilde{X}_{jj} \in \mathbb{C}^{\tilde{n}_j \times p\tilde{n}_j}$.

Now let

$$S S_0^{-1} = \mathrm{diag}(\theta_j), \qquad \widetilde{S} S_0^{-1} = \mathrm{diag}(\tilde{\theta}_j),$$

Using the fact that for any $1 \leq i, j \leq pn$, $\theta_i \neq \theta_j$, it holds $\alpha \theta_i + \tilde{\theta}_i \neq \alpha \theta_j + \tilde{\theta}_j$, where $\alpha \in \mathscr{T}$ and

$$\mathscr{T} := \mathbb{R} \setminus \left\{ \alpha \,\middle|\, \alpha = \frac{\tilde{\theta}_j - \tilde{\theta}_i}{\theta_i - \theta_j}, \ \theta_i \neq \theta_j \right\},$$

we know that for any $1 \leq i, j \leq pn$, $\theta_i = \theta_j$, it holds $\tilde{\theta}_i = \tilde{\theta}_j$. Otherwise, for any $\alpha \in \mathscr{T}$, we have

$$\mathrm{card}((\alpha S + \widetilde{S}) S_0^{-1}) > \mathrm{card}(\zeta_n(S S_0^{-1})) = \mathrm{card}(\zeta_n^{\mathrm{opt}}),$$

which contradicts with definition of ζ_n^{opt}. Similarly, we know that for any $1 \leq i, j \leq pn$, $\tilde{\theta}_i = \tilde{\theta}_j$, it holds $\theta_i = \theta_j$. Therefore, there exists a permutation matrix Π_t such that $\tilde{\tau}_n = \tau_n \Pi_t$, and

$$\Pi \widetilde{W}^{-1} X \widetilde{P} (\Pi^\top \otimes I_p) = \Pi \, \mathrm{diag}(\widetilde{X}_{11}, \ldots, \widetilde{X}_{tt})(\Pi^\top \otimes I_p) = \mathrm{diag}(\widehat{X}_{11}, \ldots, \widehat{X}_{tt}), \tag{3.11}$$

where Π is the permutation matrix corresponding with Π_t, $\widehat{X}_{jj} \in \mathbb{C}^{n_j \times pn_j}$. Now let $\widehat{P} = \widetilde{P}(\Pi^\top \otimes I_p)$, then

$$\widehat{P}^\top \widetilde{S} S_0^{-1} \widehat{P} = \mathrm{diag}(\widehat{D}_{11}, \ldots, \widehat{D}_{tt}), \tag{3.12}$$

where $\widehat{D}_{jj} \in \mathbb{C}^{pn_j \times pn_j}$ for $j = 1, 2, \ldots, t$, and $\lambda(\widehat{D}_{ii}) \cap \lambda(\widehat{D}_{jj}) = \emptyset$ for $i \neq j$. Notice that the block structure of $P^\top \widetilde{S} S_0^{-1} P$ is the same as that of $\widehat{P}^\top \widetilde{S} S_0^{-1} \widehat{P}$, so we can select Π_t such that $\widehat{P} = P$. Consequently, (3.11) can be rewritten as

$$\Pi \widetilde{W}^{-1} X P = \mathrm{diag}(\widehat{X}_{11}, \ldots, \widehat{X}_{tt}), \tag{3.13}$$

then using (3.10), we have

$$\mathrm{diag}(\widehat{X}_{11}, \ldots, \widehat{X}_{tt}) = \Pi \widetilde{W}^{-1} X P = \Pi \widetilde{W}^{-1} W (W^{-1} X P) = \Pi \widetilde{W}^{-1} W \, \mathrm{diag}(X_{11}, \ldots, X_{tt}).$$

Let

$$\Pi \widetilde{W}^{-1} W = [T_{ij}]^{-1}, \tag{3.14}$$

where $T_{ij} \in \mathbb{C}^{n_i \times n_j}$ for $i, j = 1, 2, \ldots, t$, then it follows that $T_{ij} = 0$ for $i \neq j$, and $T_{jj} \widehat{X}_{jj} = X_{jj}$ for $j = 1, 2, \ldots, t$. Using the fact that both X_{jj} and $\widehat{X}_{jj}$ are of full row rank, we know that T_{jj} is unique and nonsingular. Thus, T is a block scaling matrix. Our conclusion immediately follows from (3.14).

$\square$

4 Numerical Tests

In this section, we first briefly discuss the numerical methods to solve NU-GJBD, then present some numerical results.

4.1 Numerical Method

The computation of the solutions to the NU-GJBD requires four major steps:

Step 1. Compute all the eigenvalues and eigenvectors of $P_\mathcal{A}(\lambda)$ and formulate the eigeninformation matrix pair (X, Λ);

Step 2. Find a basis of $\mathbb{S}_{(X,\Lambda)}$;

Step 3. Determine partition τ_n and structure preserving matrix P;

Step 4. Compute transformation matrix W.

The first step can be achieved by linearization and computing the general eigenvalue problem $(\lambda M - L)u = 0$, where L, M are from (2.4). The last three steps can be followed from the algorithm in [9], we will omit the details here. As a matter of fact, Example 3.1 is carried out under such framework. Hereafter we will refer to this method as *Polynomial Eigenproblem Approach Solver (PEAS)* for GJBD.

It worth mentioning here that the robustness of PEAS heavily depends on the eigenvalue distribution of $P_\mathcal{A}(\lambda)$. When some of the eigenvalues are close, it would be hard to determine the basis of $\mathbb{S}_{(X,\Lambda)}$, which would make PEAS problematic.

4.2 Numerical Examples

Now we present several numerical examples to illustrate the performance of PEAS. All the numerical examples were carried out using MATLAB 8.0, with machine epsilon $\epsilon = 2.2 \times 10^{-16}$. The quality of joint block diagonalization is measured by the *off-diagonal index*

$$\text{resid}_{\text{offdiag}} = \frac{1}{p+1} \sum_{i=0}^{p} \frac{\| \text{OffBdiag}_{\tau_n}(W^H A_i W)\|_F}{\|W^H A_i W\|_F},$$

where (τ_n, W) solves NU-GJBD, OffBdiag_{τ_n} is defined as

$$\text{OffBdiag}_{\tau_n}(A) = \begin{bmatrix} 0_{n_1 \times n_1} & A_{12} & \cdots & A_{1t} \\ A_{21} & 0_{n_2 \times n_2} & \cdots & A_{2t} \\ \vdots & \vdots & \ddots & \vdots \\ A_{t1} & A_{t2} & \cdots & 0_{n_t \times n_t} \end{bmatrix}.$$

Example 4.1. *Let*

$$A_i = V^H D_i V, \quad i = 0, 1, 2, 3,$$

where

$$D_0 = \text{diag}\left(5, \begin{bmatrix} -5 & 4i \\ -4i & 8 \end{bmatrix}, \begin{bmatrix} 9 & 1-7i \\ 1+7i & -7 \end{bmatrix}, -5\right),$$

$$D_1 = \text{diag}\left(6, \begin{bmatrix} -5 & 9-3i \\ 9+3i & -6 \end{bmatrix}, \begin{bmatrix} -5 & 2-i \\ 2+i & -3 \end{bmatrix}, 7\right),$$

$$D_2 = \text{diag}\left(2, \begin{bmatrix} 1 & 8-4i \\ 8+4i & 5 \end{bmatrix}, \begin{bmatrix} 5 & -2+i \\ -2-i & -8 \end{bmatrix}, -9\right),$$

$$D_3 = \text{diag}\left(1, \begin{bmatrix} 6 & 9-7i \\ 9+7i & 1 \end{bmatrix}, \begin{bmatrix} -1 & -10-3i \\ -10+3i & -7 \end{bmatrix}, 6\right),$$

and

$$V = \begin{bmatrix} 0.31+0.26i & 0.45+0.15i & 0.07-0.32i & 0.08+0.14i & 0.54+0.35i & 0.23+0.41i \\ 0.52-0.65i & 0.08-0.82i & 0.44+0.77i & 0.39-0.13i & 0.14+0.51i & 0.12-0.04i \\ 0.16-0.68i & 0.22+0.53i & 0.10+0.81i & 0.25-0.86i & 0.85-0.40i & 0.18+0.90i \\ 0.60+0.74i & 0.91-0.99i & 0.96-0.86i & 0.80-0.57i & 0.62+0.07i & 0.24-0.94i \\ 0.92+0.57i & 0.23-0.82i & 0.04+0.64i & 0.64+0.54i & 0.74+0.68i & 0.36+0.78i \\ 0.95+0.05i & 0.35-0.01i & 0.16-0.73i & 0.45+0.29i & 0.18+0.18i & 0.62+0.08i \end{bmatrix}.$$

Applying PEAS on $\mathcal{A} = \{A_i\}_{i=0}^{3}$, we get

$$\tau_n = (2, 2, 1, 1),$$

and

$$W = \begin{bmatrix} 1.00 & -0.10+0.45i & 4.08-0.54i & -0.04-0.06i & -0.14-0.61i & 0.37+0.42i \\ 0.76+1.59i & 0.21+0.17i & -0.90-4.21i & 0.01+0.18i & -0.42-0.10i & 0.48-0.26i \\ 0.10-1.13i & 0.33+0.11i & 1.02-2.71i & -0.03+0.11i & -0.69-0.50i & 0.28-0.25i \\ -4.66+4.08i & 0.10-0.41i & -4.99-0.18i & 0.08+0.16i & -0.77+1.00i & -0.63-1.19i \\ -2.39-2.85i & 0.05+0.14i & 3.35-0.81i & -0.11+0.02i & 0.45-0.95i & 1.02+0.57i \\ 4.42-0.39i & -0.40-0.17i & 0.36+6.44i & -0.04-0.19i & 0.89+0.75i & -0.72+0.66i \end{bmatrix},$$

where we report all numbers in three significant digits only, though all calculations were carried out in full precision. The off-diagonal index is

$$\text{resid}_{\text{offdiag}} = 8.3290e-15,$$

which shows PEAS is feasible.

Example 4.2. *In this example, we investigate how the separability of $\lambda(P_\mathcal{A})$ affects the performance of PEAS.*

Let

$$A_0 = W^H D_0 W, \quad A_1 = W^H D_1 W, \quad A_2 = W^H D_2 W,$$

where

$$D_2 = diag(1, 1, 1), \quad D_1 = \text{diag}(-3, -2, -2), \quad D_0 = \text{diag}(2, 1-\epsilon^2, 2),$$

and

$$W = \begin{bmatrix} 0.5417-0.7983i & 0.1186+0.1690i & 0.5641+0.3498i \\ -0.8669-0.6363i & -0.5002+0.1359i & 0.1803-0.0979i \\ -0.0946-0.0403i & -0.5303+0.4836i & -0.5898-0.9762i \end{bmatrix}.$$

We know the spectrum of $P_{\mathcal{A}}$ is

$$\lambda(P_{\mathcal{A}}) = \{2, 1, 1 - \epsilon, 1 + \epsilon, 1 + i, 1 - i\}.$$

In order to determine the basis of $\mathbb{S}_{(X,\Lambda)}$ in step 2 of PEAS, one need to determine the null space of certain matrix, we denote its singular values by σ_1, σ_2, ..., in ascending order. We will see in the following table that as ϵ deceases, σ_3 increases from a near zero number. So numerically, we have two choices for the dimension of $\mathbb{S}_{(X,\Lambda)}$. In the following table, for different ϵ, we present σ_1, σ_2, σ_3 and $\text{resid}_{\text{offdiag}}$ for different choices of $\dim \mathbb{S}_{(X,\Lambda)}$.

	singular values			$\text{resid}_{\text{offdiag}}$	
ϵ	σ_1	σ_2	σ_3	$\dim \mathbb{S}_{(X,\Lambda)} = 2$	$\dim \mathbb{S}_{(X,\Lambda)} = 3$
10^{-1}	$4.10e-17$	$8.72e-16$	$5.61e-14$	$1.33e-15$	$1.37e-14$
10^{-2}	$3.57e-17$	$7.58e-16$	$6.41e-12$	$1.55e-15$	$9.34e-13$
10^{-3}	$9.48e-17$	$4.93e-16$	$3.72e-10$	$2.09e-14$	$3.12e-10$
10^{-4}	$5.25e-17$	$1.34e-15$	$2.52e-08$	$1.34e-15$	$1.98e-08$
10^{-5}	$5.68e-17$	$1.20e-15$	$4.58e-06$	$1.96e-15$	$1.06e-06$
10^{-6}	$7.23e-17$	$1.04e-15$	$6.43e-04$	$1.80e-16$	$9.89e-05$
10^{-7}	$6.48e-17$	$1.50e-15$	0.0239	$2.59e-15$	0.0084
10^{-8}	$6.90e-17$	$4.91e-16$	0.1288	$9.28e-16$	0.2066

Table 4.1: Three smallest singular values and $\text{resid}_{\text{offdiag}}$ for different $\dim \mathbb{S}_{(X,\Lambda)}$

Table 4.1 shows that the first and second smallest singular values σ_1 and σ_2 are essentially unchanged as ϵ decreases. But the third smallest singular value σ_3 increases greatly as ϵ decreases. Moreover, comparing the third column with the fifth column, fourth with last, respectively, we can see that $\text{resid}_{\text{offdiag}}$ is in the same order of the corresponding singular value.

In specific, A_0, A_1 and A_2 can be joint diagonalized by W^{-1}, that is,

$$W^{-H} P_{\mathcal{A}}(\lambda) W^{-1} = \lambda^2 D_2 + \lambda D_1 + D_0 = P_{\mathcal{D}}(\lambda) = \text{diag}\left(P_{\mathcal{D}_{11}}(\lambda), P_{\mathcal{D}_{22}}(\lambda), P_{\mathcal{D}_{33}}(\lambda)\right),$$

where $P_{\mathcal{D}_{11}}(\lambda) = \lambda^2 - 3\lambda + 2$, $P_{\mathcal{D}_{22}}(\lambda) = \lambda^2 - 2\lambda + 1 - \epsilon^2$, $P_{\mathcal{D}_{33}}(\lambda) = \lambda^2 - 2\lambda + 2$. Apparently,

$$\lambda(P_{\mathcal{D}_{11}}) = \{2, 1\}, \quad \lambda(P_{\mathcal{D}_{22}}) = \{1 + \epsilon, 1 - \epsilon\}, \quad \lambda(P_{\mathcal{D}_{33}}) = \{1 + i, 1 - i\}.$$

Then $\text{dist}\left(\lambda(P_{\mathcal{D}_{33}}), \lambda(P_{\mathcal{D}_{11}}) \cup \lambda(P_{\mathcal{D}_{22}})\right) = 1$ and $\text{dist}\left(\lambda(P_{\mathcal{D}_{11}}), \lambda(P_{\mathcal{D}_{22}})\right) = |\epsilon|$, where $\text{dist}(A, B) = \min\{|a - b| \mid a \in A, b \in B\}$. That is why $\text{resid}_{\text{offdiag}}$ for $\dim(\mathbb{S}_{(X,\Lambda)}) = 2$ maintains a high accuracy as $|\epsilon|$ decrease to zero, but $\text{resid}_{\text{offdiag}}$ for $\dim(\mathbb{S}_{(X,\Lambda)}) = 3$ increases greatly.

Example 4.3. *In this example, we use PEAS to separate artificial mixtures of three natural images, which is a BSS problem. The images were downloaded from "ICA group of Laboratory of Computer and Information Science of Helsinki University of Technology"[26]. The original 256×512 pixel images were vectorised by stacking the columns next to each other. The mixing matrix B is*

$$B = \begin{bmatrix} 0.3 & 0.4 & 0.2 \\ 0.3 & 0.3 & 0.6 \\ 0.4 & 0.3 & 0.2 \end{bmatrix}.$$

The original images are given in Figure 4.1 and mixed images in Figure 4.2.

We use the algorithm introduced in [23] to obtain a set of derivative matrices, then apply PEAS to it, solving a JD problem. Finally, the separated images are given in Figure 4.3. Since the indeterminacy of BSS[22], the order of separated images are different from that of original images, and the second figure in Figure 4.3 is black and white reversed. After reversing, we get Figure 4.4.

Figure 4.1: Original images

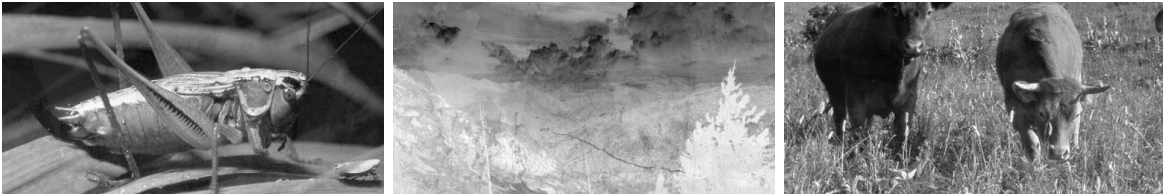Figure 4.2: Mixed images

Figure 4.3: Separated images

Figure 4.4: The second image after reversing

The off-diagonal index of the derivative matrix set corresponding to the original images and separated images are 0.0754, 0.0067, respectively. This means a smaller off-diagonal index does NOT necessarily leads to a better quality of separated images.

Example 4.4. *We apply PEAS to the convolutive blind separation of audio signals to illustrate the effectiveness of PEAS in the CBSS context. The original audio signals (three music) are given in Figure 4.5, which are $T = 2^{18} = 262144$ time samples. The convolutive system, which is defined by a set of unknown filters with responses, is the same as that in section 4.3 of [25]. The 6 observed music signals are given in Figure 4.7. We divide the observations into $N_b = 2^6 = 64$ epochs of $T_b = 2^{12} = 4096$ samples and use a rectangular window of width T_b to estimate the correlation matrices R_i, where $i = 1, 2, 3$. By applying PEAS to $\{R_i\}_{i=1}^3$, we get the unmixing matrix. The separated signals are given in Figure 4.6.*

We can see from the above figures that the resulting signals are permuted and their amplitudes are changed, but well separated. In practice, such indeterminacy is acceptable.

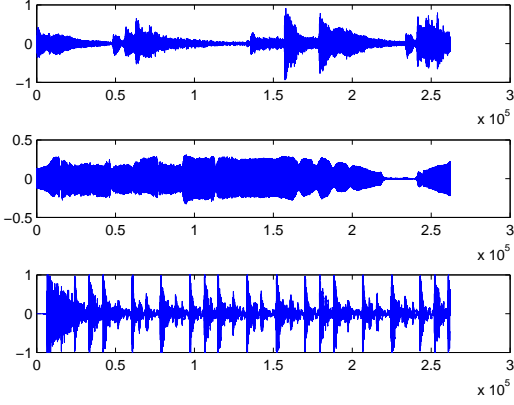*In this example, the off-diagonal index is $\mathrm{resid}_{\mathrm{offdiag}} = 0.0039$.*
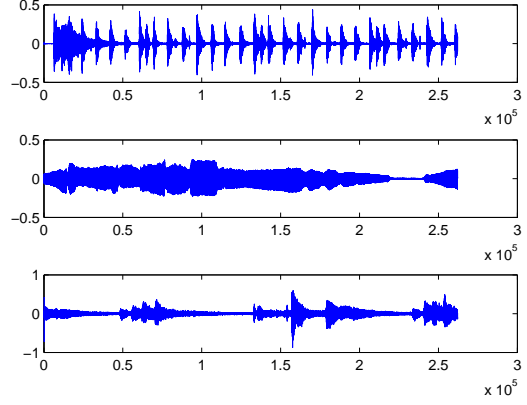
Figure 4.5: Original music signals

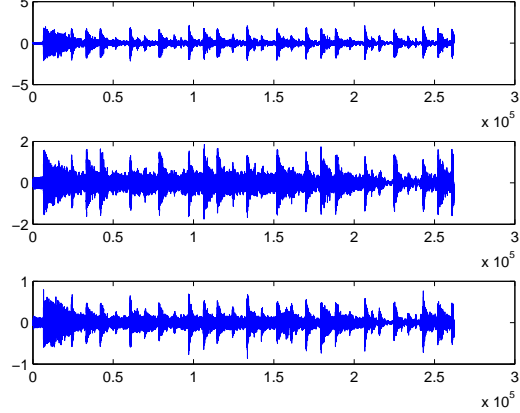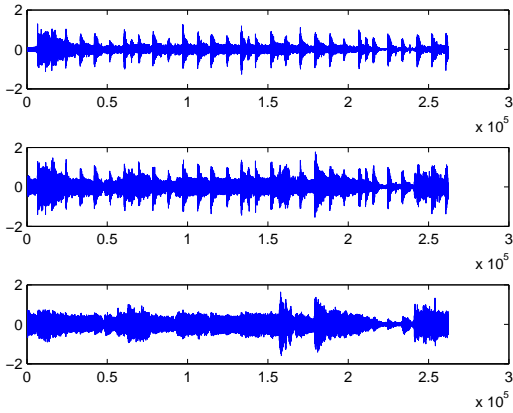Figure 4.6: Separated music signals

Figure 4.7: Observed music signals

REMARK **4.1.** Example 4.3 and 4.4, PEAS are applied to solve approximate JD and GJBD problems, respectively. The results are satisfying for practical applications.

5 Conclusion

In this paper, we show that GJBD of $\mathcal{A} = \{A_i\}_{i=0}^p$ is strongly connected with the eigeninformation of the associated matrix polynomial $P_{\mathcal{A}}(\lambda) = \sum_{i=0}^p \lambda^i A_i$. Under the assumption that $P_{\mathcal{A}}(\lambda)$ has only simple eigenvalues, a solvability theory for GJBD is established and the solutions of GJBD are characterized by the eigeninformation. Furthermore, we prove that the solution to NU-GJBD is unique up to block permutation and block scaling. Based on the established theory, we briefly present PEAS, which solves NU-GJBD. Our limited numerical tests show that PEAS is not only able to handle exact NU-GJBD well when the eigenvalues of $P_{\mathcal{A}}(\lambda)$ are well separated, and also feasible and effective for approximate NU-GJBD problems arising in practical applications, e.g. BSS and CBSS.

Finally, we should point out that there are obviously unfinished tasks in this study, e.g., extension of the theory presented here to the case when $P_{\mathcal{A}}(\lambda)$ has semi-simple eigenvalues, or even deficient eigenvalues; improvement of PEAS, since it is just a rough framework; and stable analysis for PEAS, etc.

References

[1] K. Abed-Meraim and A. Belouchrani, *Algorithms for joint block diagonalization*, in 12th European Signal Processing Conference, 2004 (EUSIPCO 2004), pp. 209-212.

[2] A. Belouchrani, K. Abed-Meraim, J. F. Cardoso and E. Moulines, *A blind source separation technique using second-order statistics*, IEEE Trans. Signal Process. 45(1997), pp. 434-444.

[3] J. F. Cardoso and A. Souloumiac, *Jacobi angles for simultaneous diagonalization*, SIAM J. Matrix Anal. Appl. 17(1996), pp. 161-164.

[4] J. F. Cardoso, *Multidimensional independent component analysis*, in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 1998 (ICASSP 1998), pp. 1941-1944.

[5] S. Choi, A. Cichocki, H. M. Park and S. Y. Lee, *Blind Source Separation and Independent Component Analysis: A Review*, Neural Information Processing-Letters and Reviews, 6(2005), pp. 1-57.

[6] P. Comon and C. Jutten, *Handbook of Blind Source Separation: Independent Component Analysis and Applicitions*, Academic Press, 2010.

[7] T. K. Caughey and M. E. J. O'Kelly, *Classical normal modes in damped linear dynamic systems*, J. Appl. Mech., 32(1965), pp. 583-588.

[8] J. F. Cardoso and A. Souloumiac, *Blind beamforming for non-Gaussian signals*, IEE Proceedings F 140(1993), pp. 362-370.

[9] Y. F. Cai and S. F. Xu, *On a quadratic inverse eigenvalue problem*, Inverse Problems 25(2009), 085004.

[10] I. Gohberg, P. Lancaster and L. Rodman, *Spectral analysis of selfadjoint matrix polynomials*, Ann. of Math. (2) 112(1980), pp. 33-71.

[11] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Computer Science and Applied Mathematics. Academic Press, Inc., New York, London, 1982.

[12] H. Ghennioui, N. Thirion-Moreau, E. Moreau, A. Adib and D. Aboutajdine, *Non unitary joint block diagonalization of complex matrices using a gradient approach*, Independent Component Analysis and Signal Separation, Lecture Notes in Comput. Sci. 4666, Springer, London, UK, 2007, pp. 201-208.

[13] H. Ghennioui, N. Thirion-Moreau, E. Moreau and D. Aboutajdine, *Gradient-based joint block diagonalization algorithms: Application to blind separation of FIR convolutive mixtures*, Signal Processing 90(2010), pp. 1836-1849.

[14] M. Joho and H. Mathis, *Joint diagonalization of correlation matrices by using gradient methods with application to blind signal separation*, in Sensor Array and Multichannel Signal Processing Workshop Proceedings, 2002, pp. 273-277.

[15] P. Lancaster, *Isospectral vibrating systems. Part 1. The spectral method*, Linear Algebra Appl. 409(2005), pp. 51-69.

[16] E. Moreau, *A generalization of joint-diagonalization criteria for source separation*, IEEE Trans. Signal Process. 49(2001), pp. 530-541.

[17] F. J. THEIS, *A new concept for separability problems in blind source separation*, Neural Comput. 16(2004), pp. 1827-1850.

[18] F. J. THEIS, *Blind signal separation into groups of dependent signals using joint block diagonalization*, in the IEEE International Symposium on Circuits and Systems, 2005 (ISCAS 2005), pp. 5878-5881.

[19] F. J. THEIS, *Towards a general independent subspace analysis*, in Proceedings of the Neural Information Processing Systems Conference, 2006, pp. 1361-1368.

[20] F. J. THEIS AND Y. INOUYE, *On the use of joint diagonalization in blind signal processing*, in Proceedings of the IEEE International Symposium on Circuits and Systems, 2006, (ISCAS 2006), 4 pp. -3589.

[21] P. TICHAVSKY AND Z. KOLDOVSKY, *Algorithms for nonorthogonal approximate joint block-diagonalization*, 20th European Singnal Processing Conference, 2012 (EUSIPCO 2012), pp. 2094-2098.

[22] L. TONG, R. W. LIU, V. C. SOON AND Y. F. HUANG, *Indeterminacy and identifiability of blind identification*, IEEE Trans. Circuits Syst. 38(1991), pp. 499-509.

[23] A. YEREDOR, *Blind source separation via the second characteristic function*, Signal Processing 80(2000), pp. 897-902.

[24] A. YEREDOR, *Non-orthogonal joint diagonalization in the least-squares sense with application in blind source separation*, IEEE Trans. Signal Process. 50(2002), pp. 1545-1553.

[25] W. T. ZHANG, S. T. LOU AND H. M. LU, *Fast nonunitary joint block diagonalization with degenerate solution elimination for convolutive blind source separation*, Digital Signal Processing 22(2012), pp. 808-819.

[26] www.cis.hut.fi/projects/ica/data/images/