BOUNDS ON PERTURBATIONS
OF GENERALIZED SINGULAR VALUES AND
OF ASSOCIATED SUBSPACES*

REN-CANG LI†

Abstract. The sensitivity of the generalized singular value decomposition of a matrix pair to
perturbations in the matrix elements is analysed. It is shown how the chordal distances between
the singular values and the angular distance between the generalized singular spaces can be bounded in
terms of the angular distances between the matrices. The main results are generalizations of several
results on the standard singular value problem.

Key words. Grassmann matrix pair, generalized singular value decomposition, generalized
singular subspace, the chordal metric, metrics on Grassmann manifold, unitarily invariant norm,
perturbation bound

AMS(MOS) subject classifications. 15A18, 15A22, 15A23, 15A42

1. Introduction. The generalized singular value decomposition (GSVD) for a
matrix pair of two matrices with the same number of columns was proposed by Van
Loan in 1976. This article addresses the following question: when a matrix pair is
perturbed, by how much can its generalized singular values (GSVs) and subspaces
associated with its GSVD change? This problem was first analysed by Sun in 1983.
(Paige also gave a bound for GSV variations in 1984.) Our main results that are
different from those of Sun and Paige are generalizations of several results on the
standard singular value problem.

Throughout the paper, capital letters are for matrices, lowercase Latin letters are
for column vectors or scalars, and lowercase Greek letters are for scalars; \( \mathbb{C}^{m \times n} \) is
the set of \( m \times n \) complex matrices; \( \mathcal{U}_n \subset \mathbb{C}^{n \times n} \) is the set of \( n \times n \) unitary matrices,
\( \mathbb{C}^m = \mathbb{C}^{m \times 1}, \quad \mathbb{C} = \mathbb{C}^1; \quad \mathbb{R} \) is the real number set. The symbol \( I^{(n)} \) stands for
the \( n \times n \) unit matrix, and \( 0_{m,n} \) for the \( m \times n \) null matrix (also we just write \( I \) and \( 0 \) for
convenience when no confusion arises). \( A^T, \quad A^H, \quad \text{and} \quad A^+ \) denote the transpose, conju
gate transpose, and Moore–Penrose inverse of \( A \), respectively. \( \mathcal{R}(X) \) is the column
space, the subspace spanned by the column vectors of \( X \), and \( P_X \) is the orthogonal
projection onto the column space \( \mathcal{R}(X) \). It is easy to verify that

\[
P_X = XX^+, \quad P_{XX} = X^+X.
\]

We will consider unitarily invariant norms \( \| \cdot \| \) of matrices. In this we follow [34,
pp. 74–87]. To say that the norm is unitarily invariant on \( \mathbb{C}^{m \times n} \) means it satisfies,
besides the usual properties of any norm,

(1) \( \| UAV \| = \| A \| \), for any \( U \in \mathcal{U}_m \), and \( V \in \mathcal{U}_n \).

(2) \( \| A \| = \| A \|_2 \), for any \( A \in \mathbb{C}^{m \times n} \), rank \( A = 1 \).

Two unitarily invariant norms used frequently are the spectral norm \( \| \cdot \|_2 \) and the
Frobenius norm \( \| \cdot \|_F \).

*Received by the editors October 30, 1989; accepted for publication (in revised form) July 31,
†Institute of Applied Physics and Computational Mathematics, P. O. Box 8009, Beijing 100088,
People’s Republic of China. Present address, Department of Mathematics, University of California,
Berkeley, California 94720 (li@math.berkeley.edu).

195
In this paper, very often matrices with different dimensions enter our arguments together, so we adhere to the agreement made on [34, p. 79]. In this way, we have

$$\|CD\| \leq \begin{cases} \|C\|_2 \|D\|_2 \\ \|C\| \|D\|_2 \end{cases},$$

for any $C \in \mathbb{C}^{m \times n}$, $D \in \mathbb{C}^{n \times t}$.

This inequality will be used extensively below. Henceforth, the symbol $\| \cdot \|$ without any subscripts is reserved for a unitarily invariant norm.

For any $A \in \mathbb{C}^{m \times n}$, its singular value decomposition (SVD) may be written as (see, e.g., [34] and [7])

$$A = U\Sigma V^H, \quad \Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots),$$

where $U \in \mathbb{U}_m$, $V \in \mathbb{U}_n$, and $\sigma_1 \geq \cdots \geq 0$ are singular values (SVs) of $A$. With the help of SVD, any unitarily invariant norm can be written as $\|A\| = \Phi(\sigma_1, \sigma_2, \ldots)$ where $\Phi$ is a symmetric gauge function (see, e.g., [23] and [34]). The set of SVs of $A$ is denoted by $\sigma(A)$. Due to the numerical stability of SVD, it has been applied to a variety of practical problems with moderate dimensions. It is an especially major tool for solving ill-conditioned problems in linear algebra. There are quite a number of excellent papers written for various topics relating to SVD and its applications in the literature; see, e.g., [7], [8], [9], [22], [23], [27], [34], [36], [40], and other references therein.

Motivated by SVD, Van Loan [38] and Paige and Saunders [25] suggested several forms of the GSVD of two matrices having the same number of columns. GSVD immediately attracted the attention of a number of numerical analysts. Now, a few algorithms for its numerical computation are available, and applications to practical problems are also being made; see, e.g., [10], [12], [13], [14], and [38].

As to the perturbation theory for the standard singular value problem, the following result has been known for nearly thirty years. It was proved by Mirsky [23] with the help of a powerful theorem for Hermitian matrices of Lidskii [21] and Wielandt [42]. Let $A, \bar{A} \in \mathbb{C}^{m \times n}$, and let $\alpha_1 \geq \cdots \geq \alpha_q$ and $\bar{\alpha}_1 \geq \cdots \geq \bar{\alpha}_q (q = \min\{m, n\})$ be their SVs, respectively. Then for any unitarily invariant norm $\| \cdot \|

$$\|\text{diag}(\alpha_1 - \bar{\alpha}_1, \ldots, \alpha_q - \bar{\alpha}_q)\| \leq \|A - \bar{A}\|.$$ 

Also, Wedin [40] has generalized the celebrated Davis–Kahan sin $\theta$ theorems [4] relating to the invariant subspaces of two self-adjoint operators to cover the SVD. A problem which naturally arises is how the generalized singular values (GSVs) and associated subspaces behave under a perturbation. In 1983, Sun [30] gave a detailed analysis of this problem. In that paper, he generalized several noted results for the standard singular value problem. Although [30] is an excellent paper, there are still a few questions left to be answered. For example, the generalization of (1.3) to cover a general unitarily invariant norm is still open. Part I of this paper is written for this purpose.

The paper is divided into two parts. In the first part, we concentrate our attention on the perturbations of GSVs, and in the second part we focus on the perturbations of subspaces associated with GSVD. The main idea comes from Li’s recent papers [17], [18]. Our results are completely different from those in Sun [30].

Now, we outline definitions relating to GSVD. Readers are referred to Sun [30] for the motivations of these definitions.
Definition 1.1. Let $A, B \in \mathbb{C}^{n \times n}$. $A - \lambda B$ is called a regular matrix pencil of order $n$ if det$(A - \lambda B) \neq 0$ for $\lambda \in \mathbb{C}$. A complex number pair $(\alpha, \beta) \neq (0, 0)$ is termed a generalized eigenvalue of a regular matrix pencil $A - \lambda B$ if det$(\beta A - \alpha B) = 0$. Denote by $\lambda(A, B)$ the set of the generalized eigenvalues (counted according to their algebraic multiplicities) of $A - \lambda B$.

Definition 1.2. Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times n}$. A matrix pair $\{A, B\}$ is an $(m, p, n)$-Grassmann matrix pair (GMP) if rank$(A_B) = n$.

Definition 1.3. Let $\{A, B\}$ be an $(m, n, p)$-GMP. A nonnegative number-pair $(\alpha, \beta)$ is a GSV of the GMP $\{A, B\}$ if

\begin{equation}
(\alpha, \beta) \neq (0, 0), \quad \det(\beta^2 A^H A - \alpha^2 B^H B) = 0, \quad \alpha, \beta \geq 0,
\end{equation}

i.e.,

\begin{equation}
(\alpha, \beta) = (\sqrt{\lambda}, \sqrt{\mu}), \quad \text{where} \quad (\lambda, \mu) \in \lambda(A^H A, B^H B) \quad \text{and} \quad \lambda, \mu \geq 0.
\end{equation}

The set of GSV of $\{A, B\}$ is denoted by $\sigma\{A, B\}$.

The fact we should bear in mind in giving (1.4b) is that for $H, K \in \mathbb{C}^{n \times n}$, $H, K \succeq 0$, if $(\lambda, \mu) \in \lambda(H, K)$, then $(|\lambda|, |\mu|) \in \lambda(H, K)$.

Clearly, if $\{A, B\}$ is an $(m, p, n)$-GMP, then $A^H A - \lambda B^H B$ is a definite matrix pencil, i.e., $x^H A^H A x + x^H B^H B x > 0$ for all $0 \neq x \in \mathbb{C}^n$, and vice versa. The definite pencil $A^H A - \lambda B^H B$ has $n$ generalized eigenvalues, hence GMP $\{A, B\}$ has $n$ GSVs. A well-developed perturbation theory for the generalized eigenvalue problem of definite pencils is available; see, e.g., Sun [31], [32]; Stewart [28]; and Li [17]. Perturbation bounds for the generalized singular value problem can be obtained with the help of the close relation between the two problems. However, the bounds obtained in this way are often unsatisfactory, just like the perturbation bounds for the SVs of a single matrix obtained through the perturbation bounds for the eigenvalues of the Hermitian matrix $A^H A$. So special attention deserves to be paid to perturbations for the generalized singular value problem.

GSVD has several forms in the literature. In this paper, we adopt the following form (see Van Loan [38], Paige and Saunders [25], and Sun [30]).

Theorem 1.1 (GSVD). Let $\{A, B\}$ be an $(m, p, n)$-GMP. Then there exist matrices $U \in \mathbb{U}_m$, $V \in \mathbb{U}_p$, and $Q \in \mathbb{C}^{n \times n}$ nonsingular such that

\begin{equation}
U^H A Q = \Sigma_A, \quad V^H B Q = \Sigma_B,
\end{equation}

where

\begin{equation}
\Sigma_A = \text{diag} \{\alpha_1, \alpha_2, \ldots\},
\end{equation}

\begin{equation}
\Sigma_B = \text{diag} \{\beta_1, \beta_2, \ldots, \beta_{n-1}, \beta_n\},
\end{equation}

meaning that $\Sigma_A$ is zero except for the diagonal starting in the top left corner (leading diagonal), and $\Sigma_B$ is zero except for the diagonal finishing in the bottom right corner (trailing diagonal). $\alpha_{m+1} = \cdots = \alpha_n = 0$, if $m \leq n$; $\beta_1 = \cdots = \beta_{n-p} = 0$, if $p \leq n$; and $\alpha_i, \beta_i \geq 0, \alpha_i^2 + \beta_i^2 = 1$, $i = 1, 2, \ldots, n$.

By Definition 1.3, we have $\sigma\{A, B\} = \{(\alpha_i, \beta_i), \ldots, \} \in (m, n, p)$-GMP $\{A, B\}$ in Theorem 1.1.
Lemma 1.1. If \( \{A, B\} \) is as described in Theorem 1.1, then for any unitarily invariant norm \( \| \cdot \| \),
\[
\| Q \| \leq \| Z^+ \| \quad \text{and} \quad \| Q^{-1} \| \leq \| Z \|, 
\]
where \( Z = (A \quad B) \).

Proof. Bearing rank \( Z = n \) and \( Z^+ Z = I \) in mind, we have
\[
\begin{pmatrix}
U^H \\
V^H
\end{pmatrix} Z Q = \begin{pmatrix}
\Sigma_A \\
\Sigma_B
\end{pmatrix} \Rightarrow \begin{cases}
Q = Z^+ \begin{pmatrix}
U \\
V
\end{pmatrix} \begin{pmatrix}
\Sigma_A \\
\Sigma_B
\end{pmatrix}, \\
Q^{-1} = (\Sigma_A^H, \Sigma_B^H) \begin{pmatrix}
U^H \\
V^H
\end{pmatrix} Z.
\end{cases}
\]
Now, from the above equations, (1.1), and
\[
\left\| \begin{pmatrix}
\Sigma_A \\
\Sigma_B
\end{pmatrix} \right\|_2 = \| (\Sigma_A^H, \Sigma_B^H) \|_2 = 1,
\]
(1.6) follows.

Remark 1.1. We note that in order to have rank \( n, m + p \geq n \). When \( m + p = n \) we can choose \( Q = Z^{-1} \) and get a trivial GSVD with
\[
\sigma\{A, B\} = \{(1, 0), \ldots, (1, 0), (0, 1), \ldots, (0, 1)\}.
\]

We use the chordal metric on the Riemann sphere to measure the difference between two generalized singular values, and metrics on the Grassmann manifold to measure the difference between two matrix pairs (see Sun [29]-[35] and Li [16]-[18]). For \((\alpha, \beta) \neq (0, 0), (\gamma, \delta) \neq (0, 0)\), the chordal distance between the two points is defined by
\[
\rho((\alpha, \beta), (\gamma, \delta)) \equiv \frac{|\delta \alpha - \gamma \beta|}{\sqrt{|\alpha|^2 + |\beta|^2 \sqrt{|\gamma|^2 + |\delta|^2}}},
\]
Let \( X, Y \in \mathbb{C}^{m \times n} (m > n) \) both have full column rank \( n \), and define the angle \( \Theta(X, Y) \) between \( X \) and \( Y \) as [34]
\[
\Theta(X, Y) \equiv \arccos((X^H X)^{-\frac{1}{2}} X^H Y (Y^H Y)^{-1} Y^H X (X^H X)^{-\frac{1}{2}})^{-\frac{1}{2}} \geq 0.
\]
Here and in the following, \( A > 0 (A \geq 0) \) denotes that \( A \) is a positive definite (nonnegative definite) Hermitian matrix. \( A^{1/2} \) is the unique positive definite (nonnegative definite) square root of \( A \geq 0 \); and \( A^{-1/2} = (A^{1/2})^{-1} \) for \( A > 0 \). The difference between these two points in the Grassmann manifold can be measured by appropriate unitarily invariant norms of \( \sin \Theta(X, Y) \) or of \( F_X - F_Y \). The reader is referred to Lemma 3.1 below for the relations between the unitarily invariant norms of \( \sin \Theta(X, Y) \) or those of \( F_X - F_Y \).
Lemma 1.2. Let $X, Y \in \mathbb{C}^{m \times n}$ ($1 \leq n \leq m - 1$) have full column rank $n$. Suppose that $\tilde{Y} = (Y, Y_1) \in \mathbb{C}^{m \times m}$ is a nonsingular matrix with

$$\tilde{Y}^{-1} = \begin{pmatrix} S_H S_I^H \end{pmatrix}, \quad S \in \mathbb{C}^{m \times n}.$$

Then for any unitarily invariant norm $\| \cdot \|$,

$$\|\sin \Theta(X, Y)\| = \|(S_H S_I)^{-\frac{1}{2}} S_H^H X (X^H X)^{-\frac{1}{2}}\|.$$

Lemma 1.2 is now well known. For a proof of it, the reader is referred to, e.g., Li [17, Lemma 2.1].

Throughout this paper, $\{A, B\}$ and $\{\tilde{A}, \tilde{B}\}$ are always reserved for two $(m, p, n)$-GMPs except when otherwise stated.

Part I. Perturbation bounds for GSV. In §2 we study a geometric representation of a GSV and pairing problems for two sets of GSVs. The former is similar to the statements in [17] for a real generalized eigenvalue, and the latter may be of general interest apart from the role it plays in this paper. We present the main results concerning GSV in §3 and give their proofs in §4.

2. Geometric representation and pairing. By Definition 1.3, it is evident that every GSV can be represented by a pair $(\alpha, \beta)$ of real numbers $\alpha$ and $\beta$ satisfying

$$\alpha, \beta \geq 0, \quad \alpha^2 + \beta^2 = 1. \quad (2.1)$$

Thus a 1 to 1 correspondence between the set of GSVs and a quarter-circular arc $\Gamma$ of the unit circle is established in the following way. $(\alpha, \beta)$ satisfying (2.1) corresponds to a point $z \in \Gamma$, as shown in Fig. 1. Therefore, every $z \in \Gamma$ determines a number pair $(\alpha, \beta)$ satisfying (2.1) by its coordinate. Moreover, $(\alpha, \beta)$ determines a unique angle $\theta = \theta(\alpha, \beta) = \arccos \alpha = \arcsin \beta, (0 \leq \theta \leq \frac{\pi}{2})$. For convenience, in the following text, we treat $z$ and $(\alpha, \beta)$ equally and write $z = (\alpha, \beta)$. The symbol $\Gamma$ is always reserved for the quarter-circular arc.

![Fig. 1. Geometric representation: $z = (\alpha, \beta)$.](image-url)
Given \((\alpha, \beta), (\gamma, \delta)\) corresponding to \(z, w \in \Gamma\), the notation
\[
(\alpha, \beta) \prec (\gamma, \delta) \quad (\langle \alpha, \beta \rangle \preceq (\gamma, \delta))
\]
means
\[
\theta(\alpha, \beta) < \theta(\gamma, \delta) \quad (\theta(\alpha, \beta) \leq \theta(\gamma, \delta));
\]
and the notation \(z \prec w \ (z \preceq w)\) means that (2.2) holds. We will use the notation \((zuv)\) to mean that the points \(z, w, v \in \Gamma\) appear in the order \(z \preceq w \preceq v\). Similarly, we can deal with more than three points. A distance on \(\Gamma\) is defined by (refer to (1.8))
\[
d(z, w) \stackrel{\text{def}}{=} \rho((\alpha, \beta), (\gamma, \delta)).
\]
It is easy to prove that \((\Gamma, d)\) is a complete distance space (see Proposition 2.1 below).

**Proposition 2.1 (Li [17]).** Let \((\alpha, \beta), (\gamma, \delta)\) be two GSVs. Then
\[
\rho((\alpha, \beta), (\gamma, \delta)) = \sin(\theta(\alpha, \beta) - \theta(\gamma, \delta)).
\]

Now, we consider the pairing problem for two sets of GSVs.

**Proposition 2.2.** Let \((\alpha_i, \beta_i)\) and \((\gamma_j, \delta_j)\), \(i, j = 1, 2, \ldots, n\) be 2n GSVs arranged in increasing order, respectively, i.e.,
\[
(\alpha_1, \beta_1) \preceq \cdots \preceq (\alpha_n, \beta_n) \quad \text{and} \quad (\gamma_1, \delta_1) \preceq \cdots \preceq (\gamma_n, \delta_n).
\]
Then
\[
\min_{1 \leq j \leq n} \max_{1 \leq \tau \leq n} \rho((\alpha_j, \beta_j), (\gamma_{\tau(j)}, \delta_{\tau(j)})) = \max_{1 \leq j \leq n} \rho((\alpha_j, \beta_j), (\gamma_j, \delta_j)).
\]
The minimum in (2.5) is taken over all permutations of \(\{1, 2, \ldots, n\}\).

Proposition 2.2 may be of great importance. The reader will find that all our bounds in §3 involve permutations of \(\{1, 2, \ldots, n\}\). Generally, except for their existence, they are unknown. Proposition 2.2 may be the only result now available which tells us what the exact permutations are in a few cases (refer to Theorems 3.1 and 3.2, Corollary 3.1, Remark 3.1, and Lemma 4.2). An insight into the importance of Proposition 2.2 can also be learned by noting some facts arising from the perturbation theories for the standard eigenvalue and singular value problems. Recall that for 2n real numbers \(\alpha_1 \leq \cdots \leq \alpha_n\) and \(\gamma_1 \leq \cdots \leq \gamma_n\), we have
\[
\max_{1 \leq j \leq n} |\alpha_j - \gamma_j| = \min_{\mu} \max_{1 \leq j \leq n} |\alpha_j - \gamma_{\mu(j)}|,
\]
and more generally, for any unitarily invariant norm \(\|\cdot\|\),
\[
\|\text{diag}(\alpha_1 - \gamma_1, \ldots, \alpha_n - \gamma_n)\| = \min_{\mu} \|\text{diag}(\alpha_1 - \gamma_{\mu(1)}, \ldots, \alpha_n - \gamma_{\mu(n)})\|
\]
where \(\min_{\mu}\) is taken over all permutations of \(\{1, 2, \ldots, n\}\). If \(\alpha_i\) and \(\gamma_j\) are no longer real but may be complex, the situation becomes very complicated, and the above two equations are no longer true. The fact that we have not yet completely solved an open problem [23] of extending the well-known Weyl–Lidskii theorem (see, e.g., [34]) to the spectral perturbation of a normal matrix should be partially attributed to the fact that we do not know what a proper pairing of 2n complex numbers is. (Much work on such extensions has been done by several mathematicians; see, e.g., [1] and [2].)

The following proposition will help us to finish our proofs for the nonsquare case in §4.
Proposition 2.3. Let \((\alpha_i, \beta_i)\) and \((\gamma_j, \delta_j)\), \(i, j = 1, 2, \ldots, n\), be 2n GSVs; let 
\(p, q \geq 0\) be two integers; and define 
\[
(\alpha_{n+i}, \beta_{n+i}) = (\gamma_{n+i}, \delta_{n+i}) = (1, 0), \quad i = 1, 2, \ldots, p;
\]
\[
(\alpha_{n+p+j}, \beta_{n+p+j}) = (\gamma_{n+p+j}, \delta_{n+p+j}) = (0, 1), \quad j = 1, 2, \ldots, q.
\]

Also, let \(\mu\) be a permutation of \(\{1, 2, \ldots, n + p + q\}\). Then there exists a permutation \(\tau\) of \(\{1, 2, \ldots, n\}\) such that for any unitarily invariant norm \(\| \cdot \|
\]

\[
\| \text{diag} \left( \rho(\alpha_1, \beta_1), (\gamma_{\mu(1)}, \delta_{\mu(1)}), \ldots, \rho(\alpha_{n+p+q}, \delta_{n+p+q}) \right) \| 
\geq \| \text{diag} \left( \rho(\alpha_{\tau(1)}, \beta_{\tau(1)}), (\gamma_{\tau(1)}, \delta_{\tau(1)}), \ldots, \rho(\alpha_{\tau(n)}, \beta_{\tau(n)}) \right) \|.
\]

In order to keep the paper fairly short, we do not give the details of the proofs of Propositions 2.2 and 2.3, which are based on constructive arguments.

Proposition 2.2 solves the problem: Find a permutation \(\tau\) of \(\{1, 2, \ldots, n\}\) such that

\[
\min_{\mu} \max_{1 \leq j \leq n} \rho(\alpha_j, \beta_j, (\gamma_{\mu(j)}, \delta_{\mu(j)}) )
\]

\[
\equiv \min_{\mu} \| \text{diag} \left( \rho(\alpha_1, \beta_1), (\gamma_{\mu(1)}, \delta_{\mu(1)}), \ldots, \rho(\alpha_n, \beta_n, (\gamma_{\mu(n)}, \delta_{\mu(n)}) ) \right) \|_2
\]

\[
= \| \text{diag} \left( \rho(\alpha_{\tau(1)}, \beta_{\tau(1)}), (\gamma_{\tau(1)}, \delta_{\tau(1)}), \ldots, \rho(\alpha_{\tau(n)}, \beta_{\tau(n)}), (\gamma_{\tau(n)}, \delta_{\tau(n)}) \right) \|_2
\]

\[
= \max_{1 \leq j \leq n} \rho(\alpha_j, \beta_j, (\gamma_{\tau(j)}, \delta_{\tau(j)}) )
\]

Let \(\zeta\) and \(\nu\) be two permutations of \(\{1, 2, \ldots, n\}\) with the property

\[
(\alpha_{\zeta(1)}, \beta_{\zeta(1)}) \preceq \cdots \preceq (\alpha_{\zeta(n)}, \beta_{\zeta(n)}) \quad \text{and} \quad (\gamma_{\nu(1)}, \delta_{\nu(1)}) \preceq \cdots \preceq (\gamma_{\nu(n)}, \delta_{\nu(n)})
\]

then \(\tau = \nu \zeta^{-1}\) is a solution to the problem. In other words, \(\tau\) satisfies (2.7). Generally, there is no global solution (independent of \(\| \cdot \|\)) to the optimization problem

\[
\min_{\mu} \| \text{diag} \left( \rho(\alpha_1, \beta_1, (\gamma_{\mu(1)}, \delta_{\mu(1)})), \ldots, \rho(\alpha_n, \beta_n, (\gamma_{\mu(n)}, \delta_{\mu(n)}) \right) \|
\]

if the minimum is taken over all permutations of \(\{1, 2, \ldots, n\}\). To see this, we give a counterexample: \(n = 2\), \(z_1 = (\alpha_1, \beta_1) = (1, 0), z_2 = (\alpha_2, \beta_2) = (\sqrt{3}/2, 1/2), w_1 = (\gamma_1, \delta_1) = (1/2, \sqrt{3}/2), w_2 = (\gamma_2, \delta_2) = (0, 1)\), and \(\| \cdot \| = \| \cdot \|_F\). It is easy to verify that \(z_1 < z_2, w_1 < w_2\). Therefore \(\zeta, \nu\) and \(\tau = \nu \zeta^{-1}\) are all the identity permutation of \(\{1, 2\}\). Only two permutations exist: one is the identity permutation \(\omega\), i.e., \(\omega(i) = i\) for \(i = 1, 2\); the other is \(\mu\) defined by \(\mu(1) = 2, \mu(2) = 1\). It is easy to verify that

\[
\| \text{diag} \left( \rho(\alpha_1, \beta_1, (\gamma_{\omega(1)}, \delta_{\omega(1)})), \rho(\alpha_2, \beta_2, (\gamma_{\omega(2)}, \delta_{\omega(2)}) \right) \|_2
\]

\[
= \frac{\sqrt{3}}{2} < 1 = \| \text{diag} \left( \rho(\alpha_1, \beta_1, (\gamma_{\mu(1)}, \delta_{\mu(1)})), \rho(\alpha_2, \beta_2, (\gamma_{\mu(2)}, \delta_{\mu(2)}) \right) \|_2,
\]
whereas
\[
\frac{\sqrt{6}}{2} \geq \frac{\sqrt{5}}{2} = \| \text{diag} \left( \rho \left( (\alpha_1, \beta_1), (\gamma_{\omega(1)}, \delta_{\omega(1)}) \right), \rho \left( (\alpha_2, \beta_2), (\gamma_{\omega(2)}, \delta_{\omega(2)}) \right) \right) \|_F.
\]

Remark 2.1. \( \| \cdot \|_2 \) is unitarily invariant, and for \( \| \cdot \|_2 \), Proposition 2.3 is a corollary of Proposition 2.2. To see this, we note that the spectral norm of a diagonal matrix is the maximal modulus of its diagonal elements (refer to (2.7)), and that the values of the qualities before \( \geq \) in the inequality (2.6) are reduced if the permutations \( \mu \) are replaced by those optimal ones obtained analogously to the permutation \( \tau \) of Proposition 2.2 by arranging \( (\alpha_i, \beta_i) \) and \( (\gamma_j, \delta_j) \) in ascending order. Under these optimal permutations, \( (1,0) \) must be paired to \( (1,0) \), and \( (0,1) \) to \( (0,1) \); and thus \( (1,0) \) and \( (0,1) \) can be eliminated as required. We note that for a general unitarily invariant norm, Proposition 2.3 may not be regarded as a corollary of Proposition 2.2 (refer to the counterexample above).

Remark 2.2. Proposition 2.3 plays an important role in our proofs in §4. In fact, in order to complete our proof for the cases \( m \neq n \) and/or \( p \neq n \), we augment \( A \) and \( B \) of an \((m,p,n)\)-GMP \( \{A, B\} \) suitably to a bigger GMP \( \{A, B, \tilde{A}, \tilde{B}\} \), with \( \tilde{A}, \tilde{B} \) being square, and \( \sigma \{A, B, \tilde{A}, \tilde{B}\} \) being a union of \( \sigma \{A, B\} \) with a few \((1,0)\) and/or \((0,1)\). Proposition 2.3 guarantees the possibility of removing the added \((1,0)\) and/or \((0,1)\) from the bounds via those for the case \( m = p = n \) (refer to §4).

3. Main perturbation theorems for GSV.

Theorem 3.1. Let \( \{A, B\} \) and \( \{A, B, \tilde{A}, \tilde{B}\} \) be two \((m,p,n)\)-GMPs, and let \( \sigma \{A, B\} = \{ (\alpha_i, \beta_i), i = 1, 2, \ldots, n \} \) and \( \sigma \{A, B, \tilde{A}, \tilde{B}\} = \{ (\tilde{\alpha}_j, \tilde{\beta}_j), j = 1, 2, \ldots, n \} \). Then there exists a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) such that
\[
\max_{1 \leq j \leq n} \rho \left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}) \right) \leq \| \Theta(Z, \tilde{Z}) \|_2,
\]
where
\[
Z = \begin{pmatrix} A \\ B \end{pmatrix}, \quad \tilde{Z} = \begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix} \in \mathbb{C}^{(m+p) \times n}.
\]

Theorem 3.2. The conditions and notations are as described in Theorem 3.1. Then there exists a permutation \( \mu \) of \( \{1, 2, \ldots, n\} \) such that
\[
\left( \sum_{j=1}^{n} \left[ \rho \left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)}) \right) \right]^2 \right)^{1/2} \leq \| \Theta(Z, \tilde{Z}) \|_F.
\]

For two square matrix pairs, we have Theorem 3.3.

Theorem 3.3. To the assumptions of Theorem 3.1, we add that \( m = p = n \). Then there exists a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) such that
\[
\max_{1 \leq j \leq n} \rho \left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}) \right) = \min_{U, V \in \mathfrak{L}_n} \left\| (Z^H V Z)^{-\frac{1}{2}} (A^H U \tilde{A} V Z)^{-\frac{1}{2}} \right\|_F.
\]

We present two corollaries obtained from Theorems 3.1 and 3.2. But first we need the following lemma.
Lemma 3.1. If \( X, Y \in \mathbb{C}^{q \times t} \) (\( q > t \)) have full column rank \( t \), then \( P_X^\dagger (Y - X)(Y^H Y)^{-\frac{1}{2}} \), \( P_Y^\dagger (Y - X)(X^H X)^{-\frac{1}{2}} \), and \( \sin \Theta(X, Y) \) have the same nonzero singular values, where \( P_X^\dagger = I - P_X \), and \( P_Y^\dagger = I - P_Y \). Moreover if their nonzero singular values are \( \sigma_1, \sigma_2, \ldots \), the nonzero singular values of \( P_Y - P_X \) are \( \sigma_1, \sigma_1, \sigma_2, \sigma_2, \ldots \). Therefore,

\[
\begin{align*}
(3.5a) \quad \| P_Y - P_X \|_2 &= \| P_X^\dagger (Y - X)(Y^H Y)^{-\frac{1}{2}} \|_2 = \| P_Y^\dagger (Y - X)(X^H X)^{-\frac{1}{2}} \|_2 \\
&= \| \sin \Theta(X, Y) \|_2;
(3.5b) \quad \frac{1}{\sqrt{2}} \| P_Y - P_X \|_F &= \| P_X^\dagger (Y - X)(Y^H Y)^{-\frac{1}{2}} \|_F = \| P_Y^\dagger (Y - X)(X^H X)^{-\frac{1}{2}} \|_F \\
&= \| \sin \Theta(X, Y) \|_F.
\end{align*}
\]

**Proof.** Choose \( U \in \mathcal{U}_q \) such that

\[
U^H X = \begin{pmatrix} X_1 \\ 0 \end{pmatrix}, \quad U^H Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}, \quad X_1, Y_1 \in \mathbb{C}^{t \times t}.
\]

Obviously, \( X_1 \) is nonsingular. Thus we have

\[
P_Y - P_X = Y(Y^H Y)^{-1} Y^H - X(X^H X)^{-1} X^H
= U \begin{pmatrix} Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_1^H - I^{(t)} \\ Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_1^H \\ Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_2^H \\ Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_2^H \end{pmatrix} U^H
\]

and

\[
(P_Y - P_X)^H (P_Y - P_X) = (P_Y - P_X)^2
= U \begin{pmatrix} I^{(t)} - Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_1^H \\ 0 \\ Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_2^H \\ 0 \end{pmatrix} U^H.
\]

On the other hand, from

\[
P_X^\dagger = U \begin{pmatrix} 0 & 0 \\ 0 & I^{(q-t)} \end{pmatrix} U^H
\]

follows

\[
(3.7) \quad (P_X^\dagger (Y - X))(Y^H Y)^{-1}(P_X^\dagger (Y - X))^H
= U \begin{pmatrix} 0 & 0 \\ 0 & Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_2^H \end{pmatrix} U^H.
\]

Note also that

\[
\Delta(Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_2^H) = \Delta(Y_2^H Y_2(Y_1^H Y_1 + Y_2^H Y_2)^{-1})
= \Delta(I - Y_1^H Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1})
= \Delta(I - Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_1^H),
\]

and

\[
\sin^2 \Theta(X, Y) = I - (Y_1^H Y_1 + Y_2^H Y_2)^{-\frac{1}{2}} Y_1^H Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-\frac{1}{2}}
= \Delta(\sin^2 \Theta(X, Y)) = \Delta(I - Y_1^H Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1})
= \Delta(I - Y_1(Y_1^H Y_1 + Y_2^H Y_2)^{-1} Y_1^H),
\]
where $\Delta(\cdot)$ denotes the set of nonzero eigenvalues of a matrix. From (3.6) and (3.7), we deduce that if the nonzero singular values of $P_X^\perp(Y - X)(Y^HY)^{-\frac{1}{2}}$ are $\sigma_1, \sigma_2, \ldots$, then so are those of $\sin \Theta(X,Y)$, and moreover, those of $P_Y - P_X$ are $\sigma_1, \sigma_2, \sigma_2, \ldots$, so

$$\|P_Y - P_X\|_2 = \|P_X^\perp(Y - X)(Y^HY)^{-\frac{1}{2}}\|_2 = \|\sin \Theta(X,Y)\|_2,$$

$$\frac{1}{\sqrt{2}}\|P_Y - P_X\|_F = \|P_X^\perp(Y - X)(Y^HY)^{-\frac{1}{2}}\|_F = \|\sin \Theta(X,Y)\|_F.$$

These prove the first equations in (3.5a) and (3.5b). By a symmetry argument, other equations in (3.5) also hold. \hfill $\Box$

**Corollary 3.1.** The conditions and notations are as described in Theorem 3.1. Then there exists a permutation $\tau$ of $\{1, 2, \ldots, n\}$ such that

$$\max_{1 \leq j \leq n} \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}) \right) \leq \min \left\{ \| (\tilde{Z} - Z)(\tilde{Z}^HZ)^{-\frac{1}{2}} \|_2, \| (\tilde{Z} - Z)(Z^HZ)^{-\frac{1}{2}} \|_2 \right\}. $$

**Corollary 3.2.** The conditions and notations are as described in Theorem 3.2. Then there exists a permutation $\mu$ of $\{1, 2, \ldots, n\}$ such that

$$\sum_{j=1}^n \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)}) \right)^2 \leq \min \left\{ \| (\tilde{Z} - Z)(\tilde{Z}^HZ)^{-\frac{1}{2}} \|_F, \| (\tilde{Z} - Z)(Z^HZ)^{-\frac{1}{2}} \|_F \right\}. $$

The inequalities (3.1) and (3.3) correspond to two inequalities for perturbations of the ordinary SVD. Let $A, \tilde{A} \in \mathbb{C}^{m \times n}$, with $\sigma(A) = \{\alpha_1, \ldots, \alpha_q\}$ and $\sigma(\tilde{A}) = \{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_q\}$, $q = \min\{m, n\}$. We also assume $\alpha_1 \geq \cdots \geq \alpha_q \geq 0$ and $\tilde{\alpha}_1 \geq \cdots \geq \tilde{\alpha}_q \geq 0$. Then

$$\max_{1 \leq j \leq q} |\alpha_j - \tilde{\alpha}_j| \leq \|A - \tilde{A}\|_2,$$

$$\sqrt{\sum_{j=1}^q |\alpha_j - \tilde{\alpha}_j|^2} \leq \|A - \tilde{A}\|_F,$$

which are nothing but the inequalities (1.3) for the special unitarily invariant norms $\|\cdot\|_2$ and $\|\cdot\|_F$. The generalizations (3.1) and (3.3) of (3.10) and (3.11) are new. A few other generalizations have appeared in the literature. Sun [33] proved the following generalization of (3.10):

$$\max_{1 \leq j \leq n} \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}) \right) \leq \max_{\|x\|_2 = 1} \rho\left( (\|Ax\|_2, \|Bx\|_2), (\|\tilde{A}x\|_2, \|\tilde{B}x\|_2) \right) \leq \|Z^+\|_2 \|Z - \tilde{Z}\|_2,$$

(3.12b)
under the condition that

\begin{equation}
\max_{x \in \mathbb{C}^n, \|x\| = 1} \sqrt{\frac{\| (\tilde{A} - A)x \|_2^2 + \| (\tilde{B} - B)x \|_2^2}{\| Ax \|_2^2 + \| Bx \|_2^2}} < 1,
\end{equation}

and for (3.11), Sun [30, Thm. 2.1] gave the following result:

\begin{equation}
\sqrt{1 - \prod_{j=1}^{n} \left( 1 - \frac{\rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}))}{1 - \frac{\det Z^H \tilde{Z}}{\det Z^H Z \det \tilde{Z}^H \tilde{Z}}} \right)^2} \leq \left\{ 1 - \frac{|\det Z^H \tilde{Z}|^2}{\det Z^H \tilde{Z}^H \tilde{Z}} \right\}^{\frac{1}{2}},
\end{equation}

from which follows [30, Corollary 2.1]

\begin{equation}
\sqrt{\sum_{j=1}^{n} \left[ \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)})) \right]^2} \leq \left\{ n \left[ 1 - \sqrt{\frac{|\det Z^H \tilde{Z}|^2}{\det Z^H \tilde{Z}^H \tilde{Z}}} \right] \right\}^{\frac{1}{2}}.
\end{equation}

Also, Paige [24], along the lines of [30], proved that if \( \alpha_i^2 + \beta_i^2 = \tilde{\alpha}_i^2 + \tilde{\beta}_i^2 = 1 \), \( i, j = 1, 2, \ldots, n \), then there exists a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) such that

\begin{equation}
\sqrt{\sum_{j=1}^{n} (\alpha_j - \tilde{\alpha}_{\tau(j)})^2 + (\beta_j - \tilde{\beta}_{\tau(j)})^2} \leq \min_{Q \in \mathcal{U}_n} \| Z_0 - \tilde{Z}_0 Q \|_F,
\end{equation}

where \( Z_0 = Z(Z^H Z)^{-\frac{1}{2}} \) and \( \tilde{Z}_0 = \tilde{Z}(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \). As can be easily seen, the right-hand side of (3.15) involves a minimization over all \( Q \in \mathcal{U}_n \). Thus it would be very difficult to compute, although the minimization may always provide a good estimation, so Paige [24] also gave (in our notations): if \( \tilde{Z}_1 \in \mathbb{C}^{(m+p) \times (m+p-n)} \) such that \( (\tilde{Z}_0, \tilde{Z}_1) \in \mathcal{U}_{m+p} \), then

\begin{equation}
\| \tilde{Z}_1^H Z_0 \|_F \leq \min_{Q \in \mathcal{U}_n} \| Z_0 - \tilde{Z}_0 Q \|_F \leq \sqrt{2} \| \tilde{Z}_1^H Z_0 \|_F \leq \sqrt{2} \| Z_0 - \tilde{Z}_0 \|_F.
\end{equation}

Now, we are in a position to compare those bounds with ours. First we consider the case for \( \| \cdot \|_2 \). When \( m = p = n \), it is easy to see that (3.4) is the best one among all bounds. It is extremely difficult for us to relate the right-hand side of (3.12a) to that of (3.8) and to \( \| \sin \Theta(Z, \tilde{Z}) \|_2 \) as well. Therefore, we are unable here to say exactly which one of (3.12a) and (3.8) is better than the other and which one of (3.12a) and (3.1) is better as well. As to (3.12b), from

\begin{align*}
\| (Z^H Z)^{-\frac{1}{2}} \|_2 &= \text{the smallest singular value of } Z = \| Z^+ \|_2, \\
\| (\tilde{Z} - Z)(Z^H Z)^{-\frac{1}{2}} \|_2 &\leq \| \tilde{Z} - Z \|_2 \| (Z^H Z)^{-\frac{1}{2}} \|_2 = \| Z^+ \|_2 \| \tilde{Z} - Z \|_2,
\end{align*}

it follows that (3.8) improves (3.12b), and so does (3.1).

For the case of \( \| \cdot \|_F \), we could not compare (3.14a) to ours and to (3.15), but comparisons can be done if (3.14b) is used instead of (3.14a). A rather detailed
comparison was conducted by Paige [24], whose conclusion is that the bound (3.15) (cf. (3.16)) is always just about as effective as the bound (3.14b). However, for large \(n\) the reverse cannot be said, i.e., for large \(n\) there is a possibility that the right-hand side of (3.14b) is quite large, while that of (3.15) remains small. Now we show that (3.3) and (3.15) are almost equivalent, but independent. To this end, we note [24, (2.11)]

\[
\frac{1}{\sqrt{2}} \sum_{j=1}^{n} \left[ (\alpha_j - \tilde{\alpha}_r(j))^2 + (\beta_j - \tilde{\beta}_r(j))^2 \right] \\
\leq \sqrt{\sum_{j=1}^{n} \rho^2 \left( (\alpha_j, \beta_j), (\tilde{\alpha}_r(j), \tilde{\beta}_r(j)) \right)^2} \\
\leq \sqrt{\sum_{j=1}^{n} \left[ (\alpha_j - \tilde{\alpha}_r(j))^2 + (\beta_j - \tilde{\beta}_r(j))^2 \right]}.
\]

By Lemma 1.2, it follows that for any unitarily invariant norm \(\| \cdot \|\)

\[
\| \sin \Theta(Z, \tilde{Z}) \| = \| \tilde{Z}_0^H \tilde{Z}_0 \|.
\]

Now, it will not be difficult for the reader, with (3.16)–(3.18) in mind, to derive a bound of the kind (3.15) via (3.3) and vice versa. The derived bound of one kind via the other is slightly weaker than its original by a factor \(\sqrt{2}\). Comparing (3.9) with (3.15) seems to be somewhat troublesome. We end our comparisons with the following typical example [17], [31]: let \(\{A, B\}\) be an \((m, p, n)\)-GMP, and \(\{\tilde{A}, \tilde{B}\} = \{(1 + r)A, (1 + r)B\}\) its perturbed GMP. In this example, \(\sigma(A, B) = \sigma(\tilde{A}, \tilde{B})\) and the right-hand sides of (3.1), (3.3), (3.4), (3.12a), (3.14), and (3.15) are all zero, thus those inequalities produce the best estimates. Those of (3.8), (3.9), and (3.12b) may be very large if \(r\) is chosen badly. This says that metrics on the Grassmann manifold have their own advantages in measuring the differences between two GMPs.

In short, we can say that our bounds are completely different from and at least as effective as, if not better than, those of Sun and Paige. Also, ours are formulated in an elegant way in very simple and intuitive terms.

Paige [24] and Sun [35] also included discussions on the case when \(\{A, B\}\) is not an \((m, p, n)\)-GMP, i.e., \(\text{rank } A < n\). In my opinion, the GSVD problem for such a case should be regarded as an ill-conditioned one, as any small arbitrary perturbation to \(A\) and \(B\) may change the rank of \(A\). The situation is quite similar to the generalized eigenvalue problem for singular matrix pencils [12], [13], [36]. Thus, without imposing any additional assumptions such as

\[
\text{rank } \begin{pmatrix} A \\ B \end{pmatrix} = \text{rank } \begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix},
\]

it is almost impossible for us to develop a perfect perturbation theory for GSVD of \(\{A, B\}\) with \(\begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix}\) having no full column rank. The work done by Paige [24] and Sun [35] is indeed valuable for understanding the behavior of GSVD of a non-GMP in the presence of perturbations.
Remark 3.1. The reader may notice that we have used the same symbol \( \tau \) in (3.1), (3.4), (3.8), (3.12), (3.14), and (3.15) for several permutations. This is not a result of our carelessness, but is intentional. Recall that ([24], [35]) all these \( \tau \)'s can be given explicitly, as we noted after (2.7). This is not the case for \( \mu \) in (3.3) and (3.9). Generally, \( \tau \) and \( \mu \) are different (refer to the counterexample at the end of \( \S 2 \)).

Remark 3.2. The bounds (3.10) and (3.11) can be obtained from (3.1) and (3.3) using a special limiting procedure. Such a procedure has also been used to derive the perturbation bounds for the standard eigenvalue problem and for the standard singular value problem via bounds for generalized problems; cf., e.g., Stewart [28], Sun [29], and Li [17]. We omit the derivations here.

We have given perturbation bounds in the two special unitarily invariant norms \( \| \cdot \|_2 \) and \( \| \cdot \|_F \). Now, we consider the case when a general unitarily invariant norm is employed.

**Theorem 3.4.** The conditions and notations are as described in Theorem 3.1. Let

\[
(\alpha_{n+i}, \beta_{n+i}) = (-\alpha_i, \beta_i), \quad (\tilde{\alpha}_{n+j}, \tilde{\beta}_{n+j}) = (-\tilde{\alpha}_j, \tilde{\beta}_j), \quad i, j = 1, 2, \ldots, n.
\]

Then there exists a permutation \( \nu \) of \( \{1, 2, \ldots, q\} \) such that for any unitarily invariant norm \( \| \cdot \| \)

\[
\begin{align*}
\| \text{diag} \left( \rho \left( (\alpha_1, \beta_1), (\tilde{\alpha}_{\nu(1)}, \tilde{\beta}_{\nu(1)}) \right), \ldots, \rho \left( (\alpha_q, \beta_q), (\tilde{\alpha}_{\nu(q)}, \tilde{\beta}_{\nu(q)}) \right) \right) \| & \leq \frac{\pi}{2} \| P_2 - P_2 \|,
\end{align*}
\]

where

\[
q = 2n, \quad \text{if } m \leq n \text{ and } p \leq n;
\]

\[
\begin{cases}
q = 2n + 2(p - n), \\
(\alpha_{2n+i}, \beta_{2n+i}) = (\tilde{\alpha}_{2n+j}, \tilde{\beta}_{2n+j}) = (1, 0), \quad 1 \leq i, j \leq 2(p - n), \quad \text{if } m \leq n < p;
\end{cases}
\]

\[
\begin{cases}
q = 2n + 2(m - n), \\
(\alpha_{2n+i}, \beta_{2n+i}) = (\tilde{\alpha}_{2n+j}, \tilde{\beta}_{2n+j}) = (0, 1), \quad 1 \leq i, j \leq 2(m - n), \quad \text{if } m > n \geq p;
\end{cases}
\]

\[
\begin{cases}
q = 2n + 2(p - n) + 2(m - n), \\
(\alpha_{2n+i}, \beta_{2n+i}) = (\tilde{\alpha}_{2n+j}, \tilde{\beta}_{2n+j}) = (1, 0), \quad 1 \leq i, j \leq 2(p - n), \quad \text{if } m > n \quad \text{and } p > n.
\end{cases}
\]

The appearance of \((\alpha_i, \beta_i)\) and \((\tilde{\alpha}_j, \tilde{\beta}_j)\) \((i, j > n)\) in this theorem makes it unsatisfactory, since the permutation \( \nu \) may pair some of the \((\alpha_i, \beta_i)\) \((i \leq n)\) to some of the \((\tilde{\alpha}_j, \tilde{\beta}_j)\) \((j > n)\). Therefore, it will be important to find a permutation \( \nu \) with the property \( \nu(i) \leq n, \quad i = 1, 2, \ldots, n \), such that (3.20) holds. On the other hand, when \( \max\{m, p\} > n \) quite a few \((1, 0)\) and/or \((0, 1)\) that are not GSVs of the pairs considered intrude themselves into (3.20). This also makes (3.20) unsatisfactory. The reader will find in \( \S 4 \) that in the derivation of Theorems 3.1 and 3.3 there are also the same intruders \((1, 0)\) and/or \((0, 1)\). Fortunately, Propositions 2.2 and 2.3 help us to eliminate them. So, we intend to eliminate \((1, 0)\) and/or \((0, 1)\) in the right-hand side of (3.20) by choosing \( \nu \) suitably. If we, indeed, could do all these, then Theorem 3.4 would become definitive. To this end, any answer (positive or negative) to the following conjecture would be helpful.
Conjecture. Let \((\alpha_i, \beta_i) \) and \((\tilde{\alpha}_j, \tilde{\beta}_j)\), \(i, j = 1, 2, \ldots, n\), be 2n GSV, and let \((\alpha_{n+i}, \beta_{n+i})\) and \((\tilde{\alpha}_{n+j}, \tilde{\beta}_{n+j})\) \((i, j = 1, 2, \ldots, n)\) be defined by (3.19), and \(\nu\) be a permutation of \(\{1, 2, \ldots, 2n\}\). Does there exist a permutation \(\mu\) of \(\{1, 2, \ldots, n\}\) such that for any unitarily invariant norm \(\|\cdot\|\),

\[
\left\| \sum_i \rho(\alpha_i, \beta_i, \tilde{\alpha}_i, \tilde{\beta}_i) \right\| 
\leq \left\| \text{diag} \left( \rho(\alpha_1, \beta_1, \tilde{\alpha}_1, \tilde{\beta}_1), \ldots, \rho(\alpha_{n}, \beta_{n}, \tilde{\alpha}_n, \tilde{\beta}_n) \right) \right\|
\]

where \(\Sigma = \text{diag} \left( \rho(\alpha_1, \beta_1, \tilde{\alpha}_1, \tilde{\beta}_1), \ldots, \rho(\alpha_{n}, \beta_{n}, \tilde{\alpha}_n, \tilde{\beta}_n) \right)\).

If there is a positive answer to this conjecture, elimination of \((1, 0)\) (or \((0, 1)\)) from (3.20) becomes possible. This can be seen from (and by the proof of) Proposition 2.3.

We finish with a corollary of Theorem 3.4, but first we need the following lemma.

Lemma 3.2. The conditions and notations are as described in Lemma 3.1. Then for any unitarily invariant norm \(\|\cdot\|\),

\[
(3.21a) \quad \|P_Y - P_X\| \leq 2 \left\| P_X^\frac{1}{4} (Y - X)(Y^H Y)^{-\frac{1}{2}} \right\|

(3.21b) \quad = 2 \left\| P_Y^\frac{1}{4} (Y - X)(X^H X)^{-\frac{1}{2}} \right\|
\]

Equation (3.21a) becomes an identity if \(\|\cdot\|\) is the trace class norm, i.e., the sum of all the singular values of a matrix.

Proof. Let \(\Phi(\cdot)\) be the symmetric gauge function associated with the \(\|\cdot\|\). Then by Lemma 3.1, we have

\[
\|P_Y - P_X\| = \Phi(\sigma_1, \sigma_1, \sigma_2, \sigma_2, \ldots)
\leq \Phi(\sigma_1, 0, \sigma_2, 0, \ldots) + \Phi(0, \sigma_1, 0, \sigma_2, \ldots)
= 2 \left\| P_X^\frac{1}{4} (Y - X)(Y^H Y)^{-\frac{1}{2}} \right\|
= 2 \left\| P_Y^\frac{1}{4} (Y - X)(X^H X)^{-\frac{1}{2}} \right\|
\]

This establishes (3.21) and means that (3.21) is an equality if \(\|\cdot\|\) is the trace class norm. \(\square\)

Corollary 3.3. The conditions and notations are as described in Theorem 3.4. Then

\[
(3.22) \quad \left\| \text{diag} \left( \rho(\alpha_1, \beta_1, \tilde{\alpha}_1, \tilde{\beta}_1), \ldots, \rho(\alpha_n, \beta_n, \tilde{\alpha}_n, \tilde{\beta}_n) \right) \right\|
\leq \left\{ \begin{array}{l}
\pi \left\| (\tilde{Z} - Z)(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \right\|,
\pi \left\| (\tilde{Z} - Z)(Z^H Z)^{-\frac{1}{2}} \right\|.
\end{array} \right.
\]

4. Proofs of Theorems 3.1–3.4. Before proving Theorems 3.1–3.3, we give five lemmas, the first four of which may also be of independent interest.

Lemma 4.1. Partition \(X \in U_n\) as

\[
X = \begin{pmatrix}
X_{11} & X_{12} \\
X_{21} & X_{22}
\end{pmatrix}, \quad X_{11} \in \mathbb{C}^{k \times \ell}.
\]
If \( k + \ell > n \), then \( \|X_{11}\|_2 = 1 \).

Proof. It follows from \( k + \ell > n \) that
\[
\mathcal{R} \left( \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} \right) \cap \mathcal{R} \left( \begin{pmatrix} I^{(k)} \\ 0 \end{pmatrix} \right) \neq \emptyset,
\]
in other words, there exist \( g \in \mathbb{C}^\ell, \ h \in \mathbb{C}^k \) with \( \|g\|_2 = \|h\|_2 = 1 \) such that
\[
\begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} g = \begin{pmatrix} I^{(k)} \\ 0 \end{pmatrix} h \overset{\text{def}}{=} x.
\]
Obviously, \( \|x\|_2 = 1 \). On the other hand
\[
X_{11} = \begin{pmatrix} I^{(k)} \\ 0 \end{pmatrix} \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} \Rightarrow \|X_{11}\|_2 \geq \|h^H \begin{pmatrix} I^{(k)} \\ 0 \end{pmatrix} \begin{pmatrix} X_{11} \\ X_{21} \end{pmatrix} g = x^H x = 1.
\]
This, together with the fact \( \|X_{11}\|_2 \leq \|X\|_2 = 1 \), leads to \( \|X_{11}\|_2 = 1 \).

Lemma 4.2. Suppose that \( \alpha_i, \beta_i, \tilde{\alpha}_j, \tilde{\beta}_j \geq 0, \alpha_i^2 + \beta_i^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1, \ i, j = 1, 2, \ldots, n \), and set
\[
\begin{cases}
\Lambda = \text{diag}(\alpha_1, \ldots, \alpha_n), \\
\Omega = \text{diag}(\beta_1, \ldots, \beta_n), \\
\tilde{\Lambda} = \text{diag}(\tilde{\alpha}_1, \ldots, \tilde{\alpha}_n), \\
\tilde{\Omega} = \text{diag}(\tilde{\beta}_1, \ldots, \tilde{\beta}_n).
\end{cases}
\]
Let \( U, V \) be two \( n \times n \) unitary matrices. Then there exists a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) such that
\[
\max_{1 \leq j \leq n} \rho( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)})) \leq \| \Lambda U \tilde{\Omega} - \Omega \tilde{\Lambda} \|_2.
\]

Remark 4.1. An alternative formulation of (4.2) is
\[
\min_{U, V} \max_{1 \leq j \leq n} \rho( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)})) = \min_{U, V} \| \Lambda U \tilde{\Omega} - \Omega \tilde{\Lambda} \|_2.
\]
This is because permutation matrices are also unitary.

A proof of Lemma 4.2 using Lemma 4.1 was implied in the proofs of Proposition 3.4 in Li [17]. But in the special case of Lemma 4.2, for convenience, we give an easy proof here.

Proof of Lemma 4.2. Without loss of generality, we assume that
\[
(\alpha_1, \beta_1) \preceq \cdots \preceq (\alpha_n, \beta_n) \quad \text{and} \quad (\tilde{\alpha}_1, \tilde{\beta}_1) \preceq \cdots \preceq (\tilde{\alpha}_n, \tilde{\beta}_n).
\]
Choose an index \( t \) satisfying
\[
\eta \overset{\text{def}}{=} \max_{1 \leq j \leq n} \rho( (\alpha_j, \beta_j), (\tilde{\alpha}_j, \tilde{\beta}_j)) = \rho((\alpha_t, \beta_t), (\tilde{\alpha}_t, \tilde{\beta}_t)) > 0.
\]
If \( \eta = 0 \), then (4.2) becomes trivial. So we assume also \( \eta > 0 \). The possible relative positions of \( (\alpha_t, \beta_t) \) and \( (\tilde{\alpha}_t, \tilde{\beta}_t) \) on \( \Gamma \) is one of
\[
(\alpha_t, \beta_t) \prec (\tilde{\alpha}_t, \tilde{\beta}_t), \quad (\tilde{\alpha}_t, \tilde{\beta}_t) \prec (\alpha_t, \beta_t).
\]
By a symmetry argument, it suffices to consider the first case, i.e., \((\alpha_i, \beta_i) \prec (\bar{\alpha}_i, \bar{\beta}_i)\). Denote by

\[
\theta_i = \theta(\alpha_i, \beta_i), \quad \bar{\theta}_j = \theta(\bar{\alpha}_j, \bar{\beta}_j), \quad i, j = 1, 2, \ldots, n.
\]

It is easy to verify that

\[
(4.4) \quad \eta = \sin(\bar{\theta}_t - \theta_t)
\]

and

\[
(4.5) \quad \begin{cases}
\min_{1 \leq i \leq t} \alpha_i = \alpha_t = \cos \theta_t, \\
\max_{1 \leq i \leq t} \beta_i = \beta_t = \sin \theta_t,
\end{cases} \quad \begin{cases}
\max_{t \leq j \leq n} \bar{\alpha}_j = \bar{\alpha}_t = \cos \bar{\theta}_t, \\
\min_{t \leq j \leq n} \bar{\beta}_j = \bar{\beta}_t = \sin \bar{\theta}_t.
\end{cases}
\]

Next, we partition \(\Lambda, \Omega, \bar{\Lambda}, \bar{\Omega}, U,\) and \(V\) as

\[
\Lambda = \text{diag}(\Lambda_1, \Lambda_2), \quad \Omega = \text{diag}(\Omega_1, \Omega_2), \quad \Lambda_1, \Omega_1 \in \mathbb{C}^{t \times t},
\]

\[
\bar{\Lambda} = \text{diag}(\bar{\Lambda}_1, \bar{\Lambda}_2), \quad \bar{\Omega} = \text{diag}(\bar{\Omega}_1, \bar{\Omega}_2), \quad \bar{\Lambda}_1, \bar{\Omega}_1 \in \mathbb{C}^{(t-1) \times (t-1)},
\]

and

\[
U = \begin{pmatrix}
U_{11} & U_{12} \\
U_{21} & U_{22}
\end{pmatrix}, \quad V = \begin{pmatrix}
V_{11} & V_{12} \\
V_{21} & V_{22}
\end{pmatrix}.
\]

Therefore,

\[
E \overset{\text{def}}{=} \Lambda U \bar{\Omega} - \Omega V \bar{\Lambda} = \begin{pmatrix}
\Lambda_1 U_{11} \bar{\Omega}_1 - \Omega_1 V_{11} \bar{\Lambda}_1 & \Lambda_1 U_{12} \bar{\Omega}_2 - \Omega_1 V_{12} \bar{\Lambda}_2 \\
\Lambda_2 U_{21} \bar{\Omega}_1 - \Omega_2 V_{21} \bar{\Lambda}_1 & \Lambda_2 U_{22} \bar{\Omega}_2 - \Omega_2 V_{22} \bar{\Lambda}_2
\end{pmatrix},
\]

which yields

\[
(4.6) \quad \|E\|_2 \geq \|\Lambda_1 U_{12} \bar{\Omega}_2 - \Omega_1 V_{12} \bar{\Lambda}_2\|_2.
\]

We will now prove that the right-hand side of (4.6) is greater than or equal to \(\eta\). From (4.5) it follows that

\[
(4.7) \quad \begin{cases}
\|\Lambda_1^{-1}\|_2^{-1} = \min_{1 \leq i \leq t} \alpha_i = \cos \theta_t, \\
\|\Lambda_2\|_2 = \max_{t \leq j \leq n} \bar{\alpha}_j = \cos \bar{\theta}_t,
\end{cases} \quad \begin{cases}
\|\Omega_1\|_2 = \max_{1 \leq i \leq t} \beta_i = \sin \theta_t, \\
\|\Omega_2^{-1}\|_2^{-1} = \min_{t \leq j \leq n} \bar{\beta}_j = \sin \bar{\theta}_t,
\end{cases}
\]

and by Lemma 4.1, it follows that

\[
\begin{cases}
U_{12}, V_{12} \in \mathbb{C}^{t \times (n-t+1)} \\
t + (n - t + 1) = n + 1 > n
\end{cases} \quad \Rightarrow \|U_{12}\|_2 = \|V_{12}\|_2 = 1.
\]

Combining (4.6) with (4.7) and the above equations produces

\[
\|E\|_2 \geq \|\Lambda_1 U_{12} \bar{\Omega}_2\|_2 - \|\Omega_1 V_{12} \bar{\Lambda}_2\|_2 \\
\geq \|\Lambda_1^{-1}\|_2^{-1} \|U_{12}\|_2 \|\bar{\Omega}_2^{-1}\|_2^{-1} - \|\Omega_1\|_2 \|V_{12}\|_2 \|\bar{\Lambda}_2\|_2 \\
= \sin \bar{\theta}_t \cos \theta_t - \cos \bar{\theta}_t \sin \theta_t \\
= \sin(\bar{\theta}_t - \theta_t).
\]

By (4.4) and (4.3), this is exactly what we need. \(\square\)
Lemma 4.3. The conditions and notations are as described in Lemma 4.2. Then there exists a permutation $\mu$ of $\{1, 2, \ldots, n\}$ such that

\[
\left(\sum_{j=1}^{n} \left[ \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)}) \right) \right]^2 \right)^{1/2} \leq \frac{1}{\sqrt{2}} \left( \| \Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda} \|_F^2 + \| \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda} \|_F^2 \right)^{1/2}.
\]

Proof. Denote $U = (u_{ij})$, $V = (v_{ij})$. Then ($\Re \lambda$ denotes the real part of the complex number $\lambda$)

\[
\| \Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda} \|_F^2 + \| \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda} \|_F^2 = \sum_{i,j} \left[ |\alpha_i \tilde{\beta}_j u_{ij} - \beta_i \tilde{\alpha}_j v_{ij}|^2 + |\alpha_i \tilde{\beta}_j v_{ij} - \beta_i \tilde{\alpha}_j u_{ij}|^2 \right]
\]

\[
= \sum_{i,j} \left[ (\alpha_i^2 \tilde{\beta}_j^2 + \beta_i^2 \tilde{\alpha}_j^2)(|u_{ij}|^2 + |v_{ij}|^2) - 4\Re \alpha_i \beta_j \tilde{\beta}_j \tilde{\alpha}_j u_{ij} v_{ij} \right]
\]

\[
\geq \sum_{i,j} \left[ (\alpha_i^2 \tilde{\beta}_j^2 + \beta_i^2 \tilde{\alpha}_j^2)(|u_{ij}|^2 + |v_{ij}|^2) - 2\alpha_i \beta_i \tilde{\alpha}_j \tilde{\beta}_j (|u_{ij}|^2 + |v_{ij}|^2) \right]
\]

\[
= \sum_{i,j} (\alpha_i \tilde{\beta}_j - \beta_i \tilde{\alpha}_j)^2 (|u_{ij}|^2 + |v_{ij}|^2)
\]

\[
de \equiv 2 \sum_{i,j} (\alpha_i \tilde{\beta}_j - \beta_i \tilde{\alpha}_j)^2 h_{ij},
\]

where

\[
H = (h_{ij}) \equiv \left( \frac{|u_{ij}|^2 + |v_{ij}|^2}{2} \right)
\]

is, evidently, a doubly stochastic matrix [22]. Thanks to the celebrated Birkhoff theorem: every doubly stochastic matrix is a convex combination of permutation matrices (see, e.g., Marshall and Olkin [22], [34, pp. 85–86]), and by a technique used frequently (see Hoffman and Wielandt [11] and Sun [29]), we can prove that there is a permutation $\mu$ of $\{1, 2, \ldots, n\}$ such that

\[
\| \Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda} \|_F^2 + \| \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda} \|_F^2 \geq 2 \sum_{j=1}^{n} (\alpha_j \tilde{\beta}_{\mu(j)} - \beta_j \tilde{\alpha}_{\mu(j)})^2
\]

\[= 2 \sum_{j=1}^{n} \left[ \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)}) \right) \right]^2,
\]

which yields (4.8) in a straightforward way. \qed

Remark 4.2. It can be seen from the above proof that (4.8) can be reversed in the sense that there is a permutation $\nu$ of $\{1, 2, \ldots, n\}$ such that

\[
\frac{1}{\sqrt{2}} \left( \| \Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda} \|_F^2 + \| \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda} \|_F^2 \right)^{1/2} \leq \frac{1}{\sqrt{2}} \left( \sum_{j=1}^{n} \left[ \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\nu(j)}, \tilde{\beta}_{\nu(j)}) \right) \right]^2.
\]
Readers may also find that the assumption $\alpha_j^2 + \beta_j^2 = \tilde{\alpha}_j^2 + \tilde{\beta}_j^2 = 1$ is not essential. If this assumption, indeed, does not hold, it is sufficient to replace the left-hand side of (4.8) by

$$\sqrt{\sum_{j=1}^{n} |\alpha_j \tilde{\beta}_{\mu(j)} - \beta_j \tilde{\alpha}_{\mu(j)}|^2}$$

and the right-hand side of (4.9) by

$$\sqrt{\sum_{j=1}^{n} |\alpha_j \tilde{\beta}_{\mu(j)} + \beta_j \tilde{\alpha}_{\mu(j)}|^2}.$$

Then (4.8) and (4.9) with the modified left-hand side and right-hand side, respectively, remain valid.

**Remark 4.3.** Similarly to Remark 4.1, we see that Lemma 4.3 and Remark 4.2 yield the following results:

$$\min_{\mu} \sqrt{\sum_{j=1}^{n} \left[ \rho((\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)})) \right]^2} = \min_{\nu, \in U_n} \frac{1}{\sqrt{2}} \left[ \|\Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda}\|^2_F + \|\Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda}\|^2_F \right]^{\frac{1}{2}}$$

and

$$\max_{\nu, \in U_n} \frac{1}{\sqrt{2}} \left[ \|\Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda}\|^2_F + \|\Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda}\|^2_F \right]^{\frac{1}{2}} = \max_{\nu} \sqrt{\sum_{j=1}^{n} \left[ \rho((\alpha_j, \beta_j), (\Lambda_\nu(j), \tilde{\beta}_{\nu(j)})) \right]^2}.$$

**Lemma 4.4.** The conditions and notations are as described in Lemma 4.2. Let $(\alpha_{n+1}, \beta_{n+1})$ and $(\tilde{\alpha}_{n+1}, \tilde{\beta}_{n+1})$ $(i, j = 1, 2, \ldots, n)$ be defined by (3.19). Then there exists a permutation $\nu$ of $\{1, 2, \ldots, 2n\}$ such that for any unitarily invariant norm $\| \cdot \|$,

$$\|\text{diag} \left( \rho((\alpha_1, \beta_1), (\tilde{\alpha}_{\nu(1)}, \tilde{\beta}_{\nu(1)})), \ldots, \rho((\alpha_{2n}, \beta_{2n}), (\tilde{\alpha}_{\nu(2n)}, \tilde{\beta}_{\nu(2n)}) \right) \| \leq \frac{\pi}{2} \|\text{diag}(\Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda}, \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda})\|.$$

**Proof.** Set $I \equiv I^{(n)}$ and

$$\tilde{U} \overset{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} V & V \\ U & -U \end{pmatrix}, \quad \tilde{\Gamma} \overset{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} I & I \\ I & -I \end{pmatrix}, \quad \tilde{I} \overset{\text{def}}{=} \frac{1}{\sqrt{2}} \begin{pmatrix} I & -I \\ I & I \end{pmatrix}.$$

An easy verification shows that $\tilde{U}, \tilde{\Gamma}, \tilde{I} \in U_{2n}, \tilde{U}^H = \tilde{\Gamma}^H, \tilde{\Gamma}^2 = I^{(2n)},$ and

$$\tilde{I} \begin{pmatrix} \Lambda & \Lambda \\ \Lambda & -\Lambda \end{pmatrix} \tilde{I} = \begin{pmatrix} \Lambda & \Lambda \\ \Lambda & -\Lambda \end{pmatrix}, \quad \tilde{I} \begin{pmatrix} \Omega & \Omega \\ \Omega & -\Omega \end{pmatrix} \tilde{I} = \begin{pmatrix} \Omega & \Omega \\ \Omega & -\Omega \end{pmatrix}.$$
Note that
\[
\begin{pmatrix}
\Lambda \\
\Lambda
\end{pmatrix}
\tilde{U}
\begin{pmatrix}
\tilde{\Omega} \\
\tilde{\Omega}
\end{pmatrix}
- 
\begin{pmatrix}
\Omega \\
\Omega
\end{pmatrix}
\tilde{U}
\begin{pmatrix}
\tilde{\Lambda} \\
-\tilde{\Lambda}
\end{pmatrix}
= \text{diag}(\Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda}, \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda}) \tilde{I},
\]
so by (4.11),
\[
\begin{pmatrix}
\Lambda \\
-\Lambda
\end{pmatrix}
\tilde{I}
\begin{pmatrix}
\tilde{\Omega} \\
\tilde{\Omega}
\end{pmatrix}
- 
\begin{pmatrix}
\Omega \\
\Omega
\end{pmatrix}
\tilde{I}
\begin{pmatrix}
\tilde{\Lambda} \\
-\tilde{\Lambda}
\end{pmatrix}
= \tilde{I}
\begin{pmatrix}
\Lambda \\
\Lambda
\end{pmatrix}
\tilde{I} \tilde{U}
\begin{pmatrix}
\tilde{\Omega} \\
\tilde{\Omega}
\end{pmatrix}
- 
\tilde{I}
\begin{pmatrix}
\Omega \\
\Omega
\end{pmatrix}
\tilde{I} \tilde{U}
\begin{pmatrix}
\tilde{\Lambda} \\
-\tilde{\Lambda}
\end{pmatrix}
= \tilde{I} \text{diag}(\Lambda U \tilde{\Omega} - \Omega V \tilde{\Lambda}, \Lambda V \tilde{\Omega} - \Omega U \tilde{\Lambda}) \tilde{I}.
\]

Now, a result from Li [19] completes the proof. □

**Lemma 4.5** ([4, Chap. II, Thm. 5.1]). Suppose \(X \in \mathbb{C}^{n \times n}\) is partitioned as
\[
X = \begin{pmatrix}
X_{11} & X_{12} \\
X_{21} & X_{22}
\end{pmatrix},
\]
then for any unitarily invariant norm \(\| \cdot \|\) we have
\[
\|X\| \geq \|\text{diag}(X_{11}, X_{22})\|.
\]

Now, we are ready for the proofs.

Suppose that \((m, p, n)\)-GMP \(\{U, B\}\) and \(\{\tilde{A}, \tilde{B}\}\) have the GSVDs
\[
\begin{align*}
U^H A Q &= \Sigma_A, \\
V^H B Q &= \Sigma_B,
\end{align*}
\]
and
\[
\begin{align*}
\tilde{U}^H \tilde{A} Q &= \Sigma_{\tilde{A}}, \\
\tilde{V}^H \tilde{B} Q &= \Sigma_{\tilde{B}},
\end{align*}
\]
where \(\Sigma_A\) and \(\Sigma_B\) are defined by (1.5b) and (1.5c), \(\Sigma_{\tilde{A}}\) and \(\Sigma_{\tilde{B}}\) are also defined by (1.5b) and (1.5c), but with \(\alpha_i, \beta_i\) replaced by \(\tilde{\alpha}_i, \tilde{\beta}_i\), \(\tilde{U}, \tilde{V} \in \mathcal{U}_m\), and \(\tilde{V}, \tilde{V} \in \mathcal{U}_p\), \(Q, \tilde{Q} \in \mathbb{C}^{n \times n}\) are nonsingular. Suppose also, without loss of generality, that \(\alpha_1^2 + \beta_1^2 = \tilde{\alpha}_1^2 + \tilde{\beta}_1^2 = 1\), \(i, j = 1, 2, \ldots, n\). Define \(\Lambda, \tilde{\Omega}, \tilde{\Lambda}, \tilde{\tilde{\Omega}}\) by (4.1).

We first show our theorems for the square case, i.e., \(m = p = n\).

For this case \(\Sigma_A = \Lambda, \Sigma_B = \tilde{\Omega}, \Sigma_{\tilde{A}} = \tilde{\Lambda}, \Sigma_{\tilde{B}} = \tilde{\tilde{\Omega}}\). It is easy to verify that
\[
B^H V U^H A - A^H U V^H B = Q^{-H} \Lambda Q^{-1} - Q^{-H} \tilde{\Lambda} Q^{-1} = 0
\]
\[
\Rightarrow (B^H V U^H A - A^H U V^H B) Z^+ = 0,
\]
therefore,
\[
-(B^H V U^H \tilde{A} - A^H U V^H \tilde{B}) Z^+
\]
\[
= -(B^H V U^H \tilde{A} - A^H U V^H \tilde{B}) Z^+ + (B^H V U^H A - A^H U V^H B) Z^+
\]
\[
= (B^H, -A^H) \begin{pmatrix} V^H & U^H \end{pmatrix} \begin{pmatrix} A & B \end{pmatrix} Z^+ - (B^H, -A^H) \begin{pmatrix} V^H & U^H \end{pmatrix} \begin{pmatrix} \tilde{A} \\ \tilde{B} \end{pmatrix} Z^+
\]
\[
= (B^H, -A^H) \begin{pmatrix} V^H & U^H \end{pmatrix} (P_Z - P_{\tilde{Z}}).
\]
Because \( \text{rank } \tilde{Z} = n \), \( \tilde{Z} + \tilde{Z} = I^{(n)} \), postmultiplying both sides of the above equation by \( \tilde{Z} \) leads to

\[
(4.13) \quad B^H VU^H \tilde{A} - A^H UV^H \tilde{B} = -(B^H, -A^H) \begin{pmatrix} VU^H & U^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \tilde{Z}.
\]

Inserting (4.12) into (4.13) we obtain

\[
(4.14) \quad \Lambda V^H \tilde{V} \tilde{\Omega} - \Omega U^H \tilde{U} \tilde{\Lambda} = (\Omega, -\Lambda) \begin{pmatrix} U^H & V^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U} & \tilde{V} \end{pmatrix} \begin{pmatrix} \tilde{\Lambda} & \tilde{\Omega} \end{pmatrix} \overset{def}{=} E.
\]

Now, by \( V^H \tilde{V} \in \mathcal{U}_n \), \( U^H \tilde{U} \in \mathcal{U}_n \), and Lemma 4.2, it follows that there is a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) such that

\[
\|P_Z - P_{\tilde{Z}}\|_2 \geq \|E\|_2 \geq \|\Lambda V^H \tilde{V} \tilde{\Omega} - \Omega U^H \tilde{U} \tilde{\Lambda}\|_2 \geq \max_{1 \leq j \leq n} \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)}) \right),
\]

which is (3.1).

To prove (3.3), we have, by interchanging the positions of \( \{A, B\} \) and \( \{\tilde{A}, \tilde{B}\} \) in (4.14), that

\[
\tilde{\Lambda} \tilde{V}^H \Omega - \tilde{\Omega} \tilde{U}^H \Lambda = (\tilde{\Omega}, -\tilde{\Lambda}) \begin{pmatrix} \tilde{U}^H & \tilde{V}^H \end{pmatrix} (P_{\tilde{Z}} - P_Z) \begin{pmatrix} U & V \end{pmatrix} \begin{pmatrix} \Lambda & \Omega \end{pmatrix},
\]

which gives

\[
(4.15) \quad \Lambda U^H \tilde{U} \tilde{\Omega} - \Omega V^H \tilde{V} \Lambda = (\Lambda, \Omega) \begin{pmatrix} U^H & V^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U} & \tilde{V} \end{pmatrix} \begin{pmatrix} \tilde{\Omega} & -\tilde{\Lambda} \end{pmatrix} \overset{def}{=} F.
\]

In the above derivation, the fact that \( P_Z \) and \( P_{\tilde{Z}} \) are Hermitian is employed (see, e.g., [34, p. 106]). Now, by \( V^H \tilde{V} \in \mathcal{U}_n \), \( U^H \tilde{U} \in \mathcal{U}_n \), and Lemma 4.3, it follows that there is a permutation \( \mu \) of \( \{1, 2, \ldots, n\} \) such that

\[
(4.16) \quad \|E\|_F^2 + \|F\|_F^2 \geq 2 \sum_{j=1}^n \rho\left( (\alpha_j, \beta_j), (\tilde{\alpha}_{\mu(j)}, \tilde{\beta}_{\mu(j)}) \right)^2.
\]

We also have

\[
(4.17) \quad \|E\|_F^2 + \|F\|_F^2 \leq \left\| \begin{pmatrix} \Omega & -\Lambda \\ \Lambda & \Omega \end{pmatrix} \begin{pmatrix} U^H & V^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U} & \tilde{V} \end{pmatrix} \begin{pmatrix} \tilde{\Lambda} & -\tilde{\Omega} \\ -\tilde{\Lambda} & \tilde{\Omega} \end{pmatrix} \right\|_F^2 = \|P_Z - P_{\tilde{Z}}\|_F^2,
\]

since

\[
\begin{pmatrix} \Omega & -\Lambda \\ \Lambda & \Omega \end{pmatrix}, \begin{pmatrix} \tilde{\Lambda} & -\tilde{\Omega} \\ -\tilde{\Lambda} & \tilde{\Omega} \end{pmatrix} \in \mathcal{U}_n.
\]

Inequality (3.3) is a consequence of (4.16) and (4.17).
In order to prove (3.4), we note that
\[ Z^H Z = A^H A + B^H B = Q^{-H} Q^{-1}, \quad \tilde{Z}^H \tilde{Z} = \tilde{Q}^{-H} \tilde{Q}^{-1} \]
\[ \Rightarrow Q_1 = Q^{-1}(Z^H Z)^{-\frac{1}{2}} \in \mathcal{U}_n, \quad \tilde{Q}_1 = \tilde{Q}^{-1}(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \in \mathcal{U}_n. \]

Therefore, from Lemma 4.2 and Proposition 2.2 (see Remark 2.1), it follows that there is a permutation \( \tau \) of \( \{1, 2, \ldots, n\} \) independent of \( U_1 \) and \( V_1 \) such that
\[ \|(Z^H Z)^{-\frac{1}{2}} (A^H V_1 \tilde{B} - B^H U_1 \tilde{A}) (\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}}\|_2 = \|Q_1^H (\Lambda U^H V_1 \tilde{\Omega} - \Omega V^H U_1 \tilde{\Lambda}) \tilde{Q}_1\|_2 \]
\[ = \|\Lambda U^H V_1 \tilde{\Omega} - \Omega V^H U_1 \tilde{\Lambda}\|_2 \]
\[ \geq \max_{1 \leq j \leq n} \rho(\alpha_j, \beta_j, (\tilde{\alpha}_{\tau(j)}, \tilde{\beta}_{\tau(j)})). \]

Since \( U_1, V_1 \) are arbitrary unitary matrices, and the equality above holds for some special \( U_1 \) and \( V_1 \), we have proved (3.4).

To prove Theorem 3.4, it suffices to use Lemmas 4.4 and 4.5 with the two matrices \( E \) and \( F \) defined by (4.14) and (4.15), respectively, and noting that \( E \) and \( F \) are the two \( n \times n \) diagonal blocks of the following \( 2n \times 2n \) matrix:
\[ \begin{pmatrix} E & * \\ * & F \end{pmatrix} = \begin{pmatrix} \Omega & -\Lambda \\ \Lambda & \Omega \end{pmatrix} \begin{pmatrix} U^H \\ V^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U} & \Omega \\ \tilde{V} & -\Lambda \end{pmatrix}. \]

So far, we have completed the proof for Theorem 3.3 and proved Theorems 3.1, 3.2, and 3.4 for the case of \( m = p = n \).

Remark 4.4. We describe how (4.14) and (4.15) relate to the projective matrix spaces [26]. We note that the matrix space considered by [26] is all \( n \times 2n \) complex matrices with full row rank \( n \). Let \( \{A, B\} \) and \( \{\tilde{A}, \tilde{B}\} \) be two \((n, n, n)\)-GMPs having GSVD as described above. Then \( Z \) and \( \tilde{Z} \) defined by (3.2) are two \( 2n \times n \) complex matrices with full column rank \( n \), and \( Z^H \) and \( \tilde{Z}^H \) are two \( n \times 2n \) complex matrices with full row rank \( n \). We now want to show that the chordal distance \( \chi(Z^H, \tilde{Z}^H) \) introduced by Schwarz and Zaks [26] as a generalization of the chordal distance of the scalar case (refer to (1.8)) is nothing but \( \|\sin \Theta(Z, \tilde{Z})\|_2 \); more specifically, we show
\begin{align}
(4.18a) \quad \chi(Z^H, \tilde{Z}^H) &= \|AV^H \tilde{\Omega} - \Omega U^H \tilde{\Lambda}\|_2 = \|\Lambda U^H \tilde{\Omega} - \Omega V^H \tilde{\Lambda}\|_2 \\
(4.18b) &= \|\sin \Theta(Z, \tilde{Z})\|_2 = \|P_Z - P_{\tilde{Z}}\|_2.
\end{align}

To this end, we see by easy verification that
\[ \begin{pmatrix} U \Lambda & -U \Omega \\ V \Omega & V \Lambda \end{pmatrix}, \quad \begin{pmatrix} \Lambda U^H & \Omega V^H \\ -\Omega U^H & \Lambda V^H \end{pmatrix}, \quad \begin{pmatrix} \tilde{U} \Lambda & -\tilde{U} \Omega \\ \tilde{V} \Omega & \tilde{V} \Lambda \end{pmatrix}, \quad \begin{pmatrix} \tilde{\Lambda} \tilde{U}^H & \tilde{\Omega} \tilde{V}^H \\ -\tilde{\Omega} \tilde{U}^H & \tilde{\Lambda} \tilde{V}^H \end{pmatrix} \]
are four \( 2n \times 2n \) unitary matrices. Thus, for any unitarily invariant norm \( \| \cdot \| \), by Lemma 1.2 we have
\[ \|\sin \Theta(Z, \tilde{Z})\| = \|\sin \Theta(ZQ, \tilde{ZQ})\| \quad \text{(refer to (4.12))} \]
\[ = \begin{cases} \|\Lambda U^H \tilde{\Omega} - \Omega U^H \tilde{\Lambda}\|, \\
\|\Lambda U^H \tilde{\Omega} - \Omega V^H \tilde{\Lambda}\|. \end{cases} \]
For \( \| \cdot \| = \| \cdot \|_2 \), this proves the first equation in (4.18b), while the second equation follows from Lemma 3.1. The equation (4.18a) follows immediately from the definition of \( \chi(Z^H, \tilde{Z}^H) \) in [26]. By the way, we point out an interesting result following from (4.18). The reader can easily deduce that

\[(4.20) \quad \|\AVH \tilde{\tilde{W}} - \Omega U^H \tilde{\tilde{U}} \tilde{\tilde{A}}\|_2 \leq \|\PZ - \tilde{\PZ}\|_2, \quad \|\Lambda U^H \tilde{\tilde{W}} - \Omega V^H \tilde{\tilde{V}} \tilde{\tilde{A}}\|_2 \leq \|\PZ - \tilde{\PZ}\|_2. \]

In fact, we have already employed these two inequalities to prove (3.1). Equations (4.18a) and (4.18b) say that both equalities in (4.20) are always attained for some \( Z \) and \( \tilde{Z} \).

The rest of our proof is to reduce the case when \( A \) or \( B \) is not square to the square case that has just been proved above by augmenting the considered matrices suitably with zero block and unit matrices with appropriate dimensions and then applying our theorems for the square case together with Proposition 2.3. For instance, we consider the case: \( m > n \) and \( p > n \). The augmentation is done as follows:

\[(4.21) \quad \tilde{A} = \begin{pmatrix} I^{(p-n)} & A \\ I^{(p-n)} & 0_{m+p-n,m-n} \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} B \\ 0_{m+p-n,p-n} \end{pmatrix}, \quad \tilde{A} = \begin{pmatrix} I^{(p-n)} & \tilde{A} \\ I^{(p-n)} & 0_{m+p-n,m-n} \end{pmatrix}, \quad \tilde{B} = \begin{pmatrix} \tilde{B} \\ 0_{m+p-n,p-n} \end{pmatrix}, \]

from which we have

\[
\sigma(\tilde{A}, \tilde{B}) = \{(1,0), \ldots, (1,0)\} \cup \sigma(A, B) \cup \{(0,1), \ldots, (0,1)\}, \\
\sigma(\tilde{A}, \tilde{B}) = \{(1,0), \ldots, (1,0)\} \cup \sigma(\tilde{A}, \tilde{B}) \cup \{(0,1), \ldots, (0,1)\}.
\]

We leave to the reader to fill in the details.

We have completed the proofs of Theorems 3.1–3.4. \( \square \)

**Part II. Perturbation bounds for generalized singular subspaces.** This part is organized as follows. We outline the background of the problem in \( \S 5 \), based on material from Sun [30, \S 1]. Our main results are discussed in \( \S 6 \) and 8 according to the different features of the results. We will clarify these features in the two sections. The proofs of those results in \( \S 6 \) are given in detail in \( \S 7 \), while the proofs of those in \( \S 8 \) are only outlined in their own section.

**5. Preliminaries to generalized singular subspaces.** Let \( \{A, B\} \) be an \( (m, p, n) \)-GMP, whose GSVD is given by (1.5). We take a natural number \( \ell \) satisfying

\[(5.1) \quad \max\{n - p, 0\} < \ell < \min\{m, n\}. \]

From (1.5) we obtain formally

\[
A \left( \begin{array}{c} Q_1, \\ Q_2 \end{array} \right) = \left( \begin{array}{cc} U_1, & U_2 \end{array} \right) \times_{m-\ell} \left( \begin{array}{cc} A_{11} & 0 \\ 0 & A_{22} \end{array} \right),
\]

\[
B(Q_1, Q_2) = \left( \begin{array}{cc} V_1, & V_2 \end{array} \right) \times_{n-\ell} \left( \begin{array}{cc} B_{11} & 0 \\ 0 & B_{22} \end{array} \right),
\]

\[(5.2a) \quad A \left( \begin{array}{cc} Q_1, & Q_2 \end{array} \right) = \left( \begin{array}{cc} U_1, & U_2 \end{array} \right) \times_{m-\ell} \left( \begin{array}{cc} A_{11} & 0 \\ 0 & A_{22} \end{array} \right),
\]

\[
B(Q_1, Q_2) = \left( \begin{array}{cc} V_1, & V_2 \end{array} \right) \times_{n-\ell} \left( \begin{array}{cc} B_{11} & 0 \\ 0 & B_{22} \end{array} \right).
\]
and
\[
A^H(U_1, U_2) = \begin{pmatrix} \ell & n-\ell \\ n-\ell & \ell \end{pmatrix} \begin{pmatrix} A_{11}^H & 0 \\ 0 & A_{22}^H \end{pmatrix},
\]

\[B^H(V_1, V_2) = \begin{pmatrix} p+\ell-n & n-\ell \\ n-\ell & p+\ell-n \end{pmatrix} \begin{pmatrix} B_{11}^H & 0 \\ 0 & B_{22}^H \end{pmatrix}.
\]

Here \((U_1, U_2)\) and \((V_1, V_2)\) are unitary matrices; \(Q = (Q_1, Q_2)\) and
\[
P = Q^{-H} = \begin{pmatrix} \ell & n-\ell \\ n-\ell & \ell \end{pmatrix},
\]

are nonsingular matrices. Motivated by (5.2) and (5.3), Sun [30] suggested the following definition.

**Definition 5.1.** Suppose that \(\{A, B\}\) is an \((m, p, n)\)-GMP and (5.1) holds for a natural number \(\ell\). Let \(\mathcal{X}_1 \in \mathbb{C}^m\), \(\mathcal{X}_2 \in \mathbb{C}^n\) be subspaces of dimension \(\ell\); let \(\mathcal{X}_2 \in \mathbb{C}^p\) be a subspace of dimension \(p + \ell - n\). Then \(\{\mathcal{X}_j, j = 1, 2, 3, 4\}\) forms a set of generalized singular subspaces (GSSSs) for \(\{A, B\}\) if
\[
A\mathcal{X}_3 \subset \mathcal{X}_1, \quad B\mathcal{X}_3 \subset \mathcal{X}_2, \quad A^H\mathcal{X}_1 \subset \mathcal{X}_4, \quad B^H\mathcal{X}_2 \subset \mathcal{X}_4.
\]

Clearly, according to (5.2) and (5.3), the subspaces \(\mathcal{X}_1 = \mathcal{R}(U_1), \mathcal{X}_2 = \mathcal{R}(V_1), \mathcal{X}_3 = \mathcal{R}(Q_1),\) and \(\mathcal{X}_4 = \mathcal{R}(P_1)\) form a set of GSSSs for \(\{A, B\}\). Further, [30] also proved that if there is a set \(\{\mathcal{X}_j, j = 1, 2, 3, 4\}\) of GSSSs of \(\{A, B\}\) as described in Definition 5.1, then there exist unitary matrices \(U = (U_1, U_2)\) and \(V = (V_1, V_2)\) and a nonsingular matrix \(Q = (Q_1, Q_2)\) with \(\mathcal{X}_1 = \mathcal{R}(U_1), \mathcal{X}_2 = \mathcal{R}(V_1), \mathcal{X}_3 = \mathcal{R}(Q_1),\) and \(\mathcal{X}_4 = \mathcal{R}(P_1)\) such that formally we have decompositions (5.2) and (5.3), where \(\{A_{11}, B_{11}\}\) and \(\{A_{22}, B_{22}\}\) are \((\ell, p + \ell - n, \ell)\)-GMP and \((m - \ell, n - \ell, n - \ell)\)-GMP, respectively.

In the sequel, we assume the \((m, p, n)\)-GMP \(\{A, B\}\) has the decomposition (1.5), and that \((m, p, n)\)-GMP \(\{\tilde{A}, \tilde{B}\}\) has the same decomposition with notation marked with tildes, e.g., corresponding to \(U_1\), we have \(\tilde{U}_1\).

Note that in this paper we do not require the blocks \(A_{11}, B_{11}, \ldots\) in (5.2) to be diagonal as we did with the corresponding blocks in (1.5). Moreover, we do not require \(P_1\) and \(Q_1\) to have orthonormal columns as Sun did [30], we only require the completed matrices \(P\) and \(Q\) to be nonsingular.

Now, we define
\[
\Theta_1 \overset{df}{=} \Theta(U_1, \tilde{U}_1), \quad \Theta_2 \overset{df}{=} \Theta(V_1, \tilde{V}_1), \quad \Theta_3 \overset{df}{=} \Theta(Q_1, \tilde{Q}_1), \quad \Theta_4 \overset{df}{=} \Theta(P_1, \tilde{P}_1).
\]

By Lemma 1.2, we have
\[
\|\sin \Theta_1\| = \|\tilde{U}_2^H U_1\|, \quad \|\sin \Theta_2\| = \|\tilde{V}_2^H V_1\|,
\]
\[
\|\sin \Theta_3\| = \|\tilde{P}_{20}^H Q_{10}\|, \quad \|\sin \Theta_4\| = \|\tilde{Q}_{20}^H P_{10}\|.
\]
Here

\[(5.7) \quad P_{j0} = P_j (P_j^H P_j)^{-\frac{1}{2}}, \quad Q_{j0} = Q_j (Q_j^H Q_j)^{-\frac{1}{2}}, \quad j = 1, 2,\]

and thus \(P_{j0}^H P_{j0} = Q_{j0}^H Q_{j0} = I\), similarly for \(\tilde{P}_{j0}\) and \(\tilde{Q}_{j0}\).

In the theorems of §6, as well as those of §8, there is an essential assumption—an assumption on the separations of two sets of GSVs.

Let \(\Gamma_1\) and \(\Gamma_2\) be two connected arcs on \(\Gamma\) which separate each other as shown in Fig. 2, where \(\Gamma_j\) is the arc corresponding to the angle \(\theta_j, j = 1, 2\). It is easy to verify the following.

**Description I** (Li [17]). In Fig. 2,

\[(5.8) \quad \min_{(\alpha, \beta) \in \Gamma_1} \min_{(\gamma, \delta) \in \Gamma_2} \rho((\alpha, \beta), (\gamma, \delta)) = \sin \psi = \eta.\]

Besides, there are two other descriptions.

**Description II.** There exist \(\alpha \geq 0, \delta > 0, \alpha + \delta \leq 1\), such that

\[(5.9) \quad \max_{w \in \Gamma_2} d(w, (0, 1)) \leq \alpha \quad \text{and} \quad \min_{z \in \Gamma_1} d(z, (0, 1)) \geq \alpha + \delta.\]

For Fig. 2, it suffices to choose \(\alpha = \sin \theta_2, \alpha + \delta = \sin(\theta_2 + \psi)\). Description II was used by Sun [30] for establishing perturbation bounds for GSSSs.

**Description III.** There exist \(\alpha \geq 0, \delta > 0, \alpha + \delta \leq 1\), such that

\[(5.10) \quad \max_{z \in \Gamma_1} d(z, (1, 0)) \leq \alpha \quad \text{and} \quad \min_{w \in \Gamma_2} d(w, (1, 0)) \geq \alpha + \delta.\]

For Fig. 2, it suffices to choose \(\alpha = \sin \theta_1, \alpha + \delta = \sin(\theta_1 + \psi)\).
PROPOSITION 5.2. The above three descriptions are equivalent. Moreover, if Description II or III holds, then Description I holds with

\[(5.11) \quad \eta = (\alpha + \delta) \sqrt{1 - \alpha^2} - \alpha \sqrt{1 - (\alpha + \delta)^2}.\]

According to the standard sense for a distance between two sets with respect to a metric, \(\eta\), as well as \(\tilde{\eta}\), \(\tilde{\eta}\), and \(\tilde{\eta}\), which will be introduced in \(\S\S 6\) and 8, is more reasonable than \(\delta\) in describing gaps, and it would be better not to refer to \(\alpha\), \(\delta\) in the final bounds. Although Sun’s bounds \([30]\) could be related to \(\eta\) without referring to \(\alpha\), \(\delta\) (see our discussions at the end of \(\S 8\)), Sun \([30]\) indeed used \(\delta\) to describe it and regarded it as the gap between two sets of GSVs. It is easy to verify the following:

\[
(2\alpha + \delta) \sqrt{1 - (\alpha + \delta)^2} < (\alpha + \delta) \sqrt{1 - \alpha^2} + \alpha \sqrt{1 - (\alpha + \delta)^2} < (2\alpha + \delta) \sqrt{1 - \alpha^2}
\]

\[
\Rightarrow \frac{1}{\sqrt{1 - (\alpha + \delta)^2}} > \frac{2\alpha + \delta}{(\alpha + \delta) \sqrt{1 - \alpha^2} + \alpha \sqrt{1 - (\alpha + \delta)^2}} = \frac{\eta}{\delta} > \frac{1}{\sqrt{1 - \alpha^2}}.
\]

Given a set \(\sigma\) consisting of several GSVs, we hereafter prefer to use the notation \(\sigma \subset \Gamma_1\) (or \(\sigma \subset \Gamma_2\)), which means that all points in \(\Gamma\) corresponding to elements of \(\sigma\) belong to \(\Gamma_1\) (or \(\Gamma_2\)).

6. PERTURBATION THEOREMS FOR GSSES. In this section, we are only concerned with establishing bounds for unitarily invariant norms of \(\sin\Theta_1\) and \(\sin\Theta_2\).

THEOREM 6.1. Let \(\{A, B\}\) and \(\{\tilde{A}, \tilde{B}\}\) be two \((m, p, n)\)-GMPs with the decompositions \((5.2)\) and \((5.3)\), and \(U = (U_1, U_2) \in \mathcal{U}_m\) and \(V = (V_1, V_2) \in \mathcal{U}_n\). Suppose that \(\sigma\{A_{11}, B_{11}\} = \{(\alpha_i, \beta_i), i = 1, 2, \ldots, \ell\}\) and \(\sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = \ell + 1, \ldots, n\}\), with \(\Gamma_1\) and \(\Gamma_2\) as shown in Fig. 2 and with \(\eta\) as defined by \((5.8)\). If

\[(6.1) \quad \sigma\{A_{11}, B_{11}\} \subset \Gamma_1, \quad \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \subset \Gamma_2, \quad \eta > 0,
\]

then for any unitarily invariant norm \(\|\cdot\|\)

\[(6.2) \quad \left\| \begin{pmatrix} \sin \Theta_1 \\ \sin \Theta_2 \end{pmatrix} \right\| \leq \frac{1}{\eta} \left\| \begin{pmatrix} U_1^H & V_1^H \\ V_1^H & U_1^H \end{pmatrix} (P_Z - P_\tilde{Z}) \begin{pmatrix} \tilde{U}_2 \\ \tilde{V}_2 \end{pmatrix} \right\|
\]

\[
\max_{i = 1, 2} \|\sin \Theta_i\| \leq \frac{1}{\eta} \max_i \left\{ \left\| \begin{pmatrix} U_i^H & V_i^H \\ V_i^H & U_i^H \end{pmatrix} (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \right\| \right. \]

\[(6.3) \left. \left\| \begin{pmatrix} \tilde{U}_2^H \\ \tilde{V}_2^H \end{pmatrix} (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \right\| \right\}.
\]

Here \(Z\) and \(\tilde{Z}\) are defined by \((3.2)\).

Among all unitarily invariant norms, \(\|\cdot\|_F\) often attracts additional attention due to its own properties. Here this is also the case.

THEOREM 6.2. Let \(\{A, B\}\) and \(\{\tilde{A}, \tilde{B}\}\) be two \((m, p, n)\)-GMPs with the decompositions \((5.2)\) and \((5.3)\), and \(U = (U_1, U_2) \in \mathcal{U}_m\) and \(V = (V_1, V_2) \in \mathcal{U}_n\). Suppose that \(\sigma\{A_{11}, B_{11}\} = \{(\alpha_i, \beta_i), i = 1, 2, \ldots, \ell\}\) and \(\sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} = \{(\tilde{\alpha}_j, \tilde{\beta}_j), j = \ell + 1, \ldots, n\}\). Let

\[(6.4a) \quad \eta \overset{\text{def}}{=} \min_{\ell + 1 \leq j \leq n} \rho((\alpha_i, \beta_i), (\tilde{\alpha}_j, \tilde{\beta}_j)),
\]

\[(6.4b) \quad \eta_1 \overset{\text{def}}{=} \min_{\ell + 1 \leq j \leq n} \rho((1, 0), (\tilde{\alpha}_j, \tilde{\beta}_j)) = \min_{\ell + 1 \leq j \leq n} \tilde{\beta}_j,
\]

\[(6.4c) \quad \eta_2 \overset{\text{def}}{=} \min_{1 \leq i \leq \ell} \rho((\alpha_i, \beta_i), (0, 1)) = \min_{1 \leq i \leq \ell} \alpha_i,
\]
and

\[
\tilde{\eta} \overset{\text{def}}{=} \begin{cases} 
1 & \text{if } m \leq n \text{ and } p \leq n, \\
\min\{\eta, \eta_1\} & \text{if } m \leq n < p, \\
\min\{\eta, \eta_2\} & \text{if } m > n \geq p, \\
\min\{\eta, \eta_1, \eta_2\} & \text{if } m > n \text{ and } p > n.
\end{cases}
\]

Then if \( \tilde{\eta} > 0 \), we have for the Frobenius norm \( \| \cdot \|_F \),

\[
\| \sin \Theta_1 \|_F^2 + \| \sin \Theta_2 \|_F^2 \leq \frac{1}{\tilde{\eta}} \left\| \left( \begin{array}{cc} U_1^H & V_1^H \\
V_1 & Z \end{array} \right) (P_Z - P_{\tilde{Z}}) \left( \begin{array}{c} \tilde{U}_2 \\
\tilde{V}_2 \end{array} \right) \right\|_F^2.
\]

\[
\| \sin \Theta_1 \|_F^2 + \| \sin \Theta_2 \|_F^2 \leq \frac{1}{\tilde{\eta}} \left\| \left( \begin{array}{cc} U_1^H & V_1^H \\
V_1 & Z \end{array} \right) (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \right\|_F^2
\]

\[
+ \left\| \left( \begin{array}{c} \tilde{U}_2^H \\
\tilde{V}_2 \\
\end{array} \right) (Z - \tilde{Z})(Z^H Z)^{-\frac{1}{2}} \right\|_F^2.
\]

Equations (6.4) and (6.5) also provide a description of the separation between two sets \( \sigma\{A_{11}, B_{11}\} \) and \( \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \) of GSVs. This description is much weaker than that employed in Theorem 6.1 (cf. Descriptions I, II, and III in §5). In fact, if (6.1) holds with end points of \( \Gamma_1 \) and \( \Gamma_2 \) being elements of \( \sigma\{A_{11}, B_{11}\} \) or \( \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \), then (6.4a) and (6.5) both hold with \( \eta = \tilde{\eta} = \eta \), but the reverse cannot be said, because generally (6.4) and (6.5) in no way guarantee that there are \( \Gamma_1 \) and \( \Gamma_2 \) as shown in Fig. 2 such that (6.1) holds. The definitions (6.4) and (6.5) have some straightforward interpretations. Equation (6.4a) describes exactly how \( \sigma\{A_{11}, B_{11}\} \) and \( \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \) separate in the sense of the chordal distance. Equations (6.4b), (6.4c), and (6.5) say that in some cases we should regard \((1,0)\) or \((0,1)\) as GSVs of \( \sigma\{A_{11}, B_{11}\} \) or of \( \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \). In one way, the reader might think inclusion of \((1,0)\) or \((0,1)\) unsatisfactory, but we think it is quite reasonable in view of (5.2) as well as the GSVD described in Theorem 1.1. In fact, \((1,0) \in \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \) if \( m < n \), while \((1,0) \in \sigma\{A_{11}, B_{11}\} \) if \( n > p \), and in giving \( \eta \) of Theorem 6.1, we indeed regard \((1,0) \in \sigma\{A_{11}, B_{11}\} \) and \((0,1) \in \sigma\{\tilde{A}_{22}, \tilde{B}_{22}\} \) as \((1,0) \in \Gamma_1 \) and \((0,1) \in \Gamma_2 \).

It can be seen that in the case of \( \| \cdot \| = \| \cdot \|_F \), (6.6) is formally just (6.2), but the former is valid under much weaker conditions. The following theorem gives bounds in a general unitarily invariant norm in terms of \( \tilde{\eta} \). Here a constant enters into our bounds.

**Theorem 6.3.** The conditions and notations are as described in Theorem 6.2. Then there exists a constant \( c \geq 1 \) such that for any unitarily invariant norm \( \| \cdot \| \) we have

\[
\left\| \left( \begin{array}{cc} \sin \Theta_1 \\
\sin \Theta_2 \end{array} \right) \right\| \leq c \cdot \frac{1}{\tilde{\eta}} \left\| \left( \begin{array}{cc} U_1^H & V_1^H \\
V_1 & Z \end{array} \right) (P_Z - P_{\tilde{Z}}) \left( \begin{array}{c} \tilde{U}_2 \\
\tilde{V}_2 \end{array} \right) \right\|_F,
\]

\[
\left\| \left( \begin{array}{cc} \sin \Theta_1 \\
\sin \Theta_2 \end{array} \right) \right\| \leq c \cdot \frac{1}{\tilde{\eta}} \left\{ \left\| \left( \begin{array}{cc} U_1^H & V_1^H \\
V_1 & Z \end{array} \right) (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \right\|_F + \left\| \left( \begin{array}{c} \tilde{U}_2^H \\
\tilde{V}_2 \end{array} \right) (Z - \tilde{Z})(Z^H Z)^{-\frac{1}{2}} \right\|_F \right\}.
\]
We do not know what the smallest constant $c$ that satisfies (6.8) and (6.9) is, but thanks to Bhatia, Davis, and Koosis [3], we shall prove that it can always be chosen so that

\[ c \leq \frac{\pi}{2} \int_0^\pi \frac{\sin t}{t} \, dt < 2.91. \]

Theorem 6.2 tells us that for the Frobenius norm the best constant $c$ is 1.

Remark 6.1. The assumption that there must be positive gaps, say $\eta, \tilde{\eta} > 0$, between $\sigma(A_{11}, B_{11})$ and $\sigma(A_{22}, B_{22})$ is essential. It is not difficult for the reader to find an example with $\eta = 0$ or $\tilde{\eta} = 0$ for which small perturbations in $A$ and/or $B$ lead to big angles $\Theta_1$ and $\Theta_2$ (refer to Davis and Kahan [4] and Stewart [27], [34] for more intuitive insight on this matter).

7. Proofs of Theorems 6.1–6.3. Without loss of generality, let $(m, p, n)$-GMPs $\{A, B\}$ and $\{\widetilde{A}, \widetilde{B}\}$ have decompositions (5.2) and (5.3) with

\[
\begin{pmatrix}
A_{11} & A_{22}
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
B_{11} & B_{22}
\end{pmatrix}
\]

being the matrices $\Sigma_A$ and $\Sigma_B$, respectively, described in Theorem 1.1 and partitioned suitably, and with

\[
\begin{pmatrix}
\widetilde{A}_{11} & \widetilde{A}_{22}
\end{pmatrix}
\quad \text{and} \quad
\begin{pmatrix}
\widetilde{B}_{11} & \widetilde{B}_{22}
\end{pmatrix},
\]

denoted by $\Sigma_\tilde{A}$ and $\Sigma_\tilde{B}$, being the matrices obtained from $\Sigma_A$ and $\Sigma_B$ by replacing $\alpha_j, \beta_j$ by $\tilde{\alpha}_j, \tilde{\beta}_j$, respectively. Otherwise, for example, because $\{A_{jj}, B_{jj}\}$ is GMP, we can find its GSVD, say $M_j^H A_{jj} L_j = \Sigma_{aj}$ and $N_j^H B_{jj} L_j = \Sigma_{bj}$, where $M_j$ and $N_j$ are unitary matrices with appropriate dimensions, and $L_j$ is a nonsingular matrix, also with appropriate dimension. Thus (5.2) becomes

\[
A(Q_1 L_1, Q_2 L_2) = (U_1 M_1, U_2 M_2) \begin{pmatrix}
\Sigma_{a1} & \Sigma_{a2}
\end{pmatrix},
\]

\[
B(Q_1 L_1, Q_2 L_2) = (V_1 N_1, V_2 N_2) \begin{pmatrix}
\Sigma_{b1} & \Sigma_{b2}
\end{pmatrix},
\]

as required.

Set

\[ \Sigma_{aj} = A_{jj}, \quad \Sigma_{bj} = B_{jj}, \quad \Sigma_a = \widetilde{A}_{jj}, \quad \Sigma_b = \widetilde{B}_{jj}, \]

and set

\[
\begin{cases}
\Lambda \equiv \text{diag} (\Lambda_1, \Lambda_2) \overset{\text{def}}{=} \text{diag} (\alpha_1, \ldots, \alpha_n), \\
\Omega \equiv \text{diag} (\Omega_1, \Omega_2) \overset{\text{def}}{=} \text{diag} (\beta_1, \ldots, \beta_n),
\end{cases}
\]

\[ \Lambda_1, \Omega_1 \in \mathbb{C}^{t \times t}. \]

Similarly, we define $\tilde{\Lambda}, \tilde{\Omega}, \tilde{\Lambda}_{jj}, \tilde{\Omega}_{jj}$, $j = 1, 2$. It is easy to see that

\[ \Lambda_1 = \Sigma_{a1}, \quad \tilde{\Lambda}_1 = \tilde{\Sigma}_{a1}, \quad \Omega_2 = \Sigma_{b2}, \quad \tilde{\Omega}_2 = \tilde{\Sigma}_{b2}. \]

We first consider the square case, i.e., $m = p = n$. 

Now, besides (7.3), we also have \( \Lambda_2 = \Sigma_{a_2}, \tilde{\Lambda}_2 = \tilde{\Sigma}_{a_2}, \Omega_1 = \Sigma_{b_1}, \tilde{\Omega}_1 = \tilde{\Sigma}_{b_1} \). Identity (4.14) from §4 and

\[
(7.4) \quad \Lambda V^H \tilde{V} \tilde{\Omega} - \Omega U^H \tilde{U} \tilde{\Lambda} = \begin{pmatrix}
\Lambda_1 V_1^H \tilde{V}_1^H \tilde{\Omega}_1 - \Omega_1 U_1 \tilde{U}_1 \tilde{\Lambda}_1 & \Lambda_1 V_1^H \tilde{V}_2^H \tilde{\Omega}_2 - \Omega_1 U_1 \tilde{U}_2 \tilde{\Lambda}_2 \\
\Lambda_2 V_2^H \tilde{V}_1^H \tilde{\Omega}_1 - \Omega_2 U_2 \tilde{U}_1 \tilde{\Lambda}_1 & \Lambda_2 V_2^H \tilde{V}_2^H \tilde{\Omega}_2 - \Omega_2 U_2 \tilde{U}_2 \tilde{\Lambda}_2
\end{pmatrix}
\]

yields

\[
(7.5a) \quad \Lambda_1 V_1^H \tilde{V}_2^H \tilde{\Omega}_2 - \Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2
\]

\[
= (\Omega_1, -\Lambda_1) \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U}_2 \\ \tilde{V}_2 \end{pmatrix} \begin{pmatrix} \tilde{\Lambda}_2 \\ \tilde{\Omega}_2 \end{pmatrix} \overset{\text{def}}{=} E.
\]

Similarly, by (4.15) we have (cf. (7.4))

\[
(7.5b) \quad \Lambda_1 U_1^H \tilde{U}_2^H \tilde{\Omega}_2 - \Omega_1 V_1^H \tilde{V}_2 \tilde{\Lambda}_2
\]

\[
= (\Lambda_1, \Omega_1) \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (P_Z - P_{\tilde{Z}}) \begin{pmatrix} \tilde{U}_2 \\ \tilde{V}_2 \end{pmatrix} \begin{pmatrix} \tilde{\Lambda}_2 \\ -\tilde{\Omega}_2 \end{pmatrix} \overset{\text{def}}{=} F.
\]

Similarly to (4.13) one may also get

\[
B^H V U^H \tilde{A} - A^H U V^H \tilde{B} = -(B^H, -A^H) \begin{pmatrix} V U^H \\ U V^H \end{pmatrix} (Z - \tilde{Z})
\]

and thus

\[
(7.6a) \quad \Lambda V^H \tilde{V} \tilde{\Omega} - \Omega U^H \tilde{U} \tilde{\Lambda} = (\Omega, -\Lambda) \begin{pmatrix} U^H \\ V^H \end{pmatrix} (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \tilde{Q}^H.
\]

where \( \tilde{Q} = \tilde{Q}^{-1}(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \in \mathcal{U}_n \) (refer to the proof of Theorem 3.3 in §4). Therefore

\[
(7.6a) \quad \Lambda_1 V_1^H \tilde{V}_2 \tilde{\Omega}_2 - \Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2
\]

\[
= (\Omega_1, -\Lambda_1) \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (Z - \tilde{Z})(\tilde{Z}^H \tilde{Z})^{-\frac{1}{2}} \tilde{Q}^H \begin{pmatrix} 0 \quad n \cdot \Omega_{1-n} \end{pmatrix} \overset{\text{def}}{=} \tilde{E}.
\]

By treating \( \{A, B\} \) and \( \{\tilde{A}, \tilde{B}\} \) equally, (7.6a) produces

\[
\tilde{A} \tilde{V}^H V \Omega - \tilde{\Omega} \tilde{U}^H U \Lambda = (\tilde{\Omega}, -\tilde{\Lambda}) \begin{pmatrix} \tilde{U}^H \\ \tilde{V}^H \end{pmatrix} (\tilde{Z} - Z)(Z^H Z)^{-\frac{1}{2}} \tilde{Q}^H
\]

and thus

\[
(7.6b) \quad \Lambda U^H \tilde{U} \tilde{\Omega} - \Omega V^H \tilde{V} \tilde{\Lambda} = (\tilde{\Omega}^T, 0) \tilde{Q}(Z^H Z)^{-\frac{1}{2}} (Z - \tilde{Z})^H \begin{pmatrix} \tilde{U} \\ \tilde{V} \end{pmatrix} \begin{pmatrix} \tilde{\Omega} \\ -\tilde{\Lambda} \end{pmatrix},
\]

where \( \tilde{Q} = Q^{-1}(Z^H Z)^{-\frac{1}{2}} \in \mathcal{U}_n \). So

\[
(7.7b) \quad \Lambda_1 U_1^H \tilde{U}_2 \tilde{\Omega}_2 - \Omega_1 V_1^H \tilde{V}_2 \tilde{\Lambda}_2
\]

\[
= (I^{(\ell)}, 0) \tilde{Q}(Z^H Z)^{-\frac{1}{2}} (Z - \tilde{Z})^H \begin{pmatrix} \tilde{U}_2^H \\ \tilde{V}_2^H \end{pmatrix} \begin{pmatrix} \tilde{\Omega}_2 \\ -\tilde{\Lambda}_2 \end{pmatrix} \overset{\text{def}}{=} \tilde{F}.
holds.

We are now ready for the proofs. First, we prove Theorem 6.1.

It follows from the assumption \((\alpha_i, \beta_i) \in \Gamma_1, i = 1, 2, \ldots, \ell\) and \((\tilde{\alpha}_j, \tilde{\beta}_j) \in \Gamma_2, j = \ell + 1, \ldots, n\) that

\[
\begin{align*}
\min_{1 \leq i \leq \ell} |\alpha_i| &\leq \cos \theta_1, \\
\min_{\ell + 1 \leq j \leq n} |\tilde{\beta}_j| &\geq \sin(\theta_1 + \psi), \\
\max_{1 \leq i \leq \ell} |\beta_i| &\leq \sin \theta_1, \\
\max_{\ell + 1 \leq j \leq n} |\tilde{\alpha}_j| &\leq \cos(\theta_1 + \psi).
\end{align*}
\]

(7.8)

We claim that

\[
\begin{align*}
\max \left\{ \|\Lambda_1 V_1^H \tilde{V}_2 \tilde{\Omega}_2 - \Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2\|, \|\Lambda_1 U_1^H \tilde{U}_2 \tilde{\Omega}_2 - \Omega_1 V_1^H \tilde{V}_2 \tilde{\Lambda}_2\| \right\} \\
\geq \eta \max \left\{ \|V_1^H \tilde{V}_2\|, \|U_1^H \tilde{U}_2\| \right\} = \eta \max_{i = 1, 2} \|\sin \Theta_i\|.
\end{align*}
\]

(7.9)

If (7.9) holds, then (6.3) is trivial. We have to prove (7.9). In fact, if \(\|V_1^H \tilde{V}_2\| \geq \|U_1^H \tilde{U}_2\|\), then with the help of (7.8), we have

\[
\begin{align*}
\|\Lambda_1 V_1^H \tilde{V}_2 \tilde{\Omega}_2 - \Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2\| \\
\geq \|\Lambda_1 V_1^H \tilde{V}_2 \tilde{\Omega}_2\| - \|\Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2\| \\
\geq \|\Lambda_1^{-1}\| \|V_1^H \tilde{V}_2\| \|\tilde{\Omega}_2^{-1}\| \|\tilde{\Lambda}_2\| - \|\Omega_1\| \|U_1^H \tilde{U}_2\| \|\tilde{\Lambda}_2\| \\
\geq \|V_1^H \tilde{V}_2\| \sin \psi \\
= \eta \|\sin \Theta_2\|.
\end{align*}
\]

(7.10a)

If, on the other hand, \(\|V_1^H \tilde{V}_2\| \leq \|U_1^H \tilde{U}_2\|\), then again with the help of (7.8), we obtain

\[
\|\Lambda_1 U_1^H \tilde{U}_2 \tilde{\Omega}_2 - \Omega_1 V_1^H \tilde{V}_2 \tilde{\Lambda}_2\| \geq \eta \|\sin \Theta_1\|.
\]

(7.10b)

Combining inequalities (7.10a) and (7.10b) leads to (7.9). As in the proof of Lemma 4.4, to prove (6.2), we see that

\[
\begin{align*}
\left( \Lambda_1 - \Lambda_1 \right) \tilde{I} X \left( \tilde{\Omega}_2 \tilde{\Omega}_2 \right) - \left( \Omega_1 \Omega_1 \right) \tilde{I} X \left( \tilde{\Lambda}_2 - \tilde{\Lambda}_2 \right) = \tilde{I} \left( E \quad F \right) \tilde{I},
\end{align*}
\]

(7.11)

where unitary matrices \(\tilde{I}\) and \(\tilde{I}\) are defined by (4.10) with \(I\) having appropriate dimensions, and

\[
X = \frac{1}{\sqrt{2}} \begin{pmatrix} U_1^H \tilde{U}_2 & U_1^H \tilde{U}_2 \\ V_1^H \tilde{V}_2 & -V_1^H \tilde{V}_2 \end{pmatrix} = \begin{pmatrix} U_1^H \tilde{U}_2 \\ V_1^H \tilde{V}_2 \end{pmatrix} \tilde{I}.
\]

It is evident that (7.8) still holds with \(\Lambda_1, \Omega_1, \tilde{\Omega}_2, \) and \(\tilde{\Lambda}_2\) replaced by

\[
\begin{align*}
\left( \Lambda_1 - \Lambda_1 \right), \quad \left( \Omega_1 \Omega_1 \right), \quad \left( \tilde{\Omega}_2 \tilde{\Omega}_2 \right), \quad \left( \tilde{\Lambda}_2 - \tilde{\Lambda}_2 \right),
\end{align*}
\]
respectively. Therefore, as in (7.10a), one obtains from (7.11) that

\[(7.12) \quad \left\| \begin{pmatrix} U_1^H \tilde{U}_2 \\ V_1^H \tilde{V}_2 \end{pmatrix} \right\| \leq \frac{1}{\eta} \left\| \begin{pmatrix} E \\ F \end{pmatrix} \right\|.\]

All the SVs of $U_1^H \tilde{U}_2$ and those of $V_1^H \tilde{V}_2$ are the same as those of $\sin \Theta_1$ and those of $\sin \Theta_2$, respectively, so the SVs of

\[
\begin{pmatrix} U_1^H \tilde{U}_2 \\ V_1^H \tilde{V}_2 \end{pmatrix}
\]

are the same as those of

\[
\begin{pmatrix} \sin \Theta_1 \\ \sin \Theta_2 \end{pmatrix};
\]

thus

\[(7.13) \quad \left\| \begin{pmatrix} U_1^H \tilde{U}_2 \\ V_1^H \tilde{V}_2 \end{pmatrix} \right\| \equiv \left\| \begin{pmatrix} \sin \Theta_1 \\ \sin \Theta_2 \end{pmatrix} \right\|.
\]

It follows from Lemma 4.5 and

\[
\begin{pmatrix} E \\ F \end{pmatrix} = \begin{pmatrix} \Omega_1 & -\Lambda_1 \\ \Lambda_1 & \Omega_1 \end{pmatrix} \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (P_Z - \tilde{P}_Z) \begin{pmatrix} \tilde{U}_2 \\ \tilde{V}_2 \end{pmatrix} \begin{pmatrix} \tilde{\Lambda}_2 \\ \tilde{\Omega}_2 \end{pmatrix},
\]

and

\[
\begin{pmatrix} \Omega_1 & -\Lambda_1 \\ \Lambda_1 & \Omega_1 \end{pmatrix} \in \mathcal{U}_{2}, (\begin{pmatrix} \tilde{\Lambda}_2 \\ \tilde{\Omega}_2 \end{pmatrix} - \Lambda_2) \in \mathcal{U}_{(n-\ell)}
\]

that

\[(7.14) \quad \left\| \begin{pmatrix} E \\ F \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (P_Z - \tilde{P}_Z) \begin{pmatrix} \tilde{U}_2 \\ \tilde{V}_2 \end{pmatrix} \right\|.
\]

Equation (6.2) is a consequence of (7.12), (7.13), and (7.14).

Now, we prove Theorem 6.2. Denote by $V_1^H \tilde{V}_2 = (v_{ij})$, $U_1^H \tilde{U}_2 = (u_{ij})$, then,

\[
\begin{align*}
\| \Lambda_1 V_1^H \tilde{V}_2 \tilde{\Omega}_2 - \Omega_1 U_1^H \tilde{U}_2 \tilde{\Lambda}_2 \|_F^2 &+ \| \Lambda_1 U_1^H \tilde{U}_2 \tilde{\Omega}_2 - \Omega_1 V_1^H \tilde{V}_2 \tilde{\Lambda}_2 \|_F^2 \\
&= \sum_{i,j} \left[ |\alpha_i \tilde{\beta}_j v_{ij} - \beta_i \tilde{\alpha}_j u_{ij}|^2 + |\alpha_i \tilde{\beta}_j u_{ij} - \beta_i \tilde{\alpha}_j v_{ij}|^2 \right] \\
&= \sum_{i,j} \left[ (\alpha_i^2 \tilde{\beta}_j^2 + \beta_i^2 \tilde{\alpha}_j^2) |v_{ij}|^2 + |u_{ij}|^2 - 4 \Re \alpha_i \beta_i \tilde{\alpha}_j \tilde{\beta}_j u_{ij} v_{ij} \right] \\
&\geq \sum_{i,j} (\alpha_i \tilde{\beta}_j - \beta_i \tilde{\alpha}_j)^2 (|v_{ij}|^2 + |u_{ij}|^2) \\
&\geq \eta^2 \left( \| V_1^H \tilde{V}_2 \|_F^2 + \| U_1^H \tilde{U}_2 \|_F^2 \right).
\end{align*}
\]

Equation (6.7) follows in a straightforward way from (7.7) and (7.15). Identities (7.5a) and (7.5b) and inequalities (7.14) and (7.15) then yield (6.6).
We now prove Theorem 6.3 for the square case. Again, (7.11) is the key. For the fact that \( \alpha_i, \beta_i, \bar{\alpha}_j, \bar{\beta}_j \geq 0 \), we easily see that

\[
\rho \left( -\alpha_i, \beta_i, (\bar{\alpha}_j, \bar{\beta}_j) \right) = \rho \left( \alpha_i, \beta_i, (\bar{\alpha}_j, \bar{\beta}_j) \right) \geq \rho \left( (\alpha_i, \beta_i), (\bar{\alpha}_j, \bar{\beta}_j) \right) \geq \eta
\]

for \( 1 \leq i \leq \ell, \ell + 1 \leq j \leq n \). So from Li [18, §4, Lemma 4.1], which was proved by the author using results due to [1] and [3], it follows that there exists a constant \( c \geq 1 \), which can be chosen such that (6.10) holds, so that (note \( \hat{E} = E \) and \( \hat{F} = F \))

\[
\left\| \begin{pmatrix} U^H \hat{U}_2 \\ V_1^H \hat{V}_2 \end{pmatrix} \right\| \leq c \cdot \frac{1}{\eta} \left\| \begin{pmatrix} E & F \\ F & E \end{pmatrix} \right\| = c \cdot \frac{1}{\eta} \left\| \begin{pmatrix} \hat{E} & \hat{F} \\ \hat{F} & \hat{E} \end{pmatrix} \right\|.
\]

This together with (7.13) and (7.14) yields (6.8), and with (7.13),

\[
\left\| \begin{pmatrix} \hat{E} \\ \hat{F} \end{pmatrix} \right\| \leq \left\| \begin{pmatrix} \hat{E} \\ 0 \end{pmatrix} \right\| + \left\| \begin{pmatrix} 0 \\ \hat{F} \end{pmatrix} \right\| = \| \hat{E} \| + \| \hat{F} \|
\]

and (7.7) produce (6.9).

We treat the case when \( A \) or \( B \) is not square in the same way as we did in §4. For an example, suppose that \( m > n \) and \( p > n \). Let \( \hat{A}, \hat{B}, \tilde{A}, \tilde{B} \) be as defined by (4.21). It is easy to verify that

\[
\hat{U}^H \tilde{A} \hat{Q} \overset{\text{def}}{=} \begin{pmatrix} I \\ U^H \end{pmatrix} \tilde{A} \begin{pmatrix} I \\ Q \end{pmatrix} = \begin{pmatrix} I \\ \Lambda_1 \end{pmatrix}_{\Sigma_{\alpha_2}, 0} \overset{\text{def}}{=} \begin{pmatrix} \hat{\Lambda}_1 \\ \hat{\Lambda}_2 \end{pmatrix},
\]

\[
\hat{V}^H \tilde{B} \hat{Q} \overset{\text{def}}{=} \begin{pmatrix} V^H \\ I \end{pmatrix} \tilde{B} \hat{Q} = \begin{pmatrix} 0, \Sigma_{\beta_1} \\ \Omega_2 \end{pmatrix}_I \overset{\text{def}}{=} \begin{pmatrix} \hat{\Omega}_1 \\ \hat{\Omega}_2 \end{pmatrix},
\]

where

\[
\hat{\Lambda}_1 = \begin{pmatrix} I \\ \Lambda_1 \end{pmatrix}, \quad \hat{\Lambda}_2 = (\Sigma_{\alpha_2}, 0), \quad \hat{\Omega}_1 = (0, \Sigma_{\beta_1}), \quad \hat{\Omega}_2 = (\Omega_2, I),
\]

and

\[
\tilde{U}^H \tilde{A} \tilde{Q} \overset{\text{def}}{=} \begin{pmatrix} I \\ \tilde{U}^H \end{pmatrix} \tilde{A} \begin{pmatrix} I \\ \tilde{Q} \end{pmatrix} = \begin{pmatrix} I \\ \tilde{\Lambda}_1 \end{pmatrix}_{\tilde{\Sigma}_{\alpha_2}, 0} \overset{\text{def}}{=} \begin{pmatrix} \tilde{\Lambda}_1 \\ \tilde{\Lambda}_2 \end{pmatrix},
\]

\[
\tilde{V}^H \tilde{B} \tilde{Q} \overset{\text{def}}{=} \begin{pmatrix} \tilde{V}^H \\ I \end{pmatrix} \tilde{B} \tilde{Q} = \begin{pmatrix} 0, \tilde{\Sigma}_{\beta_1} \\ \tilde{\Omega}_2 \end{pmatrix}_I \overset{\text{def}}{=} \begin{pmatrix} \tilde{\Omega}_1 \\ \tilde{\Omega}_2 \end{pmatrix},
\]

where

\[
\tilde{\Lambda}_1 = \begin{pmatrix} I \\ \tilde{\Lambda}_1 \end{pmatrix}, \quad \tilde{\Lambda}_2 = (\tilde{\Sigma}_{\alpha_2}, 0), \quad \tilde{\Omega}_1 = (0, \tilde{\Sigma}_{\beta_1}), \quad \tilde{\Omega}_2 = (\tilde{\Omega}_2, I).
\]

Moreover,

\[
\sigma \{ \tilde{\Lambda}_1, \tilde{\Omega}_1 \} = \{(1, 0), \ldots, (1, 0)\} \bigcup_{p-n} \sigma \{ \tilde{A}_{11}, \tilde{B}_{11} \},
\]

\[
\sigma \{ \tilde{\Lambda}_2, \tilde{\Omega}_2 \} = \sigma \{ \tilde{A}_{22}, \tilde{B}_{22} \} \bigcup_{m-n} \{(0, 1), \ldots, (0, 1)\}.
\]

Therefore, if the conditions of Theorem 6.1 or those of Theorem 6.2 hold for \( \{ A, B \} \) and \( \{ \hat{A}, \hat{B} \} \), they also hold for \( (m+p-n, m+p-n, m+p-n) \)-GMP \( \{ \hat{A}, \hat{B} \} \) and \( \{ \tilde{A}, \tilde{B} \} \) with \( \ell = p - n + \ell \). The reader must finish the details.
8. More perturbation theorems for GSSSs with comparisons. In this section, we present not only more bounds for $\sin \Theta_1$ and $\sin \Theta_2$, but also bounds for $\sin \Theta_3$ and $\sin \Theta_4$.

A very interesting feature of Theorems 6.1–6.3 is that none of the bounds for $\sin \Theta_1$ and $\sin \Theta_2$ bear any relation to $P$ and $Q$, or to $\Theta_3$ and $\Theta_4$. This is not the case for the theorems we will develop below. Another different feature is that here we start by defining four residual matrices due to Sun [30] instead of (7.5)–(7.7) in §7. But here, our derivation of perturbation equations from the four residuals seems more elegant.

Assume now that we have (5.2) and (5.3) for $\{A, B\}$ and $\{\bar{A}, \bar{B}\}$. Define
\[
R_A \buildrel \text{def} \over = \bar{A}Q_1 - U_1 A_{11} = (\bar{A} - A)Q_1,
\]
\[
R_B \buildrel \text{def} \over = \bar{B}Q_1 - V_1 B_{11} = (\bar{B} - B)Q_1,
\]
\[
R_{A^H} \buildrel \text{def} \over = \bar{A}^H U_1 - P_1 A_{11}^H = (\bar{A} - A)^H U_1,
\]
\[
R_{B^H} \buildrel \text{def} \over = \bar{B}^H V_1 - P_1 B_{11}^H = (\bar{B} - B)^H V_1.
\]

Suppose more strongly that (5.2) and (5.3) are as described in the beginning of §7 up to (7.3). Assume for the moment that $m = p = n$. Then
\[
A_{jj} = \Lambda_j, \quad B_{jj} = \Omega_j, \quad \bar{A}_{jj} = \bar{\Lambda}_j, \quad \bar{B}_{jj} = \bar{\Omega}_j, \quad j = 1, 2.
\]

From (8.1) it follows that
\[
R_A \buildrel \text{def} \over = \bar{U}_2^H R_A = \bar{\Lambda}_2 \bar{P}_2^H Q_1 - \bar{U}_2^H U_1 \Lambda_1,
\]
\[
R_B \buildrel \text{def} \over = \bar{V}_2^H R_B = \bar{\Omega}_2 \bar{P}_2^H Q_1 - \bar{V}_2^H V_1 \Omega_1,
\]
\[
R_{A^H} \buildrel \text{def} \over = \bar{Q}_2^H R_{A^H} = \bar{\Lambda}_2 \bar{U}_2^H U_1 - \bar{Q}_2^H P_1 \Lambda_1,
\]
\[
R_{B^H} \buildrel \text{def} \over = \bar{Q}_2^H R_{B^H} = \bar{\Omega}_2 \bar{V}_2^H V_1 - \bar{Q}_2^H P_1 \Omega_1.
\]

Therefore, we get
\[
C_1 \stackrel{\text{def}}{=} -\bar{\Omega}_2 R_A + \bar{\Lambda}_2 R_B = \bar{\Lambda}_2 \bar{U}_2^H U_1 \Lambda_1 - \bar{\bar{\Lambda}}_2 \bar{V}_2^H V_1 \Omega_1,
\]
\[
C_2 \stackrel{\text{def}}{=} R_{A^H} \Omega_1 - R_{B^H} \Lambda_1 = \bar{\bar{\Lambda}}_2 \bar{U}_2^H U_1 \Omega_1 - \bar{\bar{\bar{\Lambda}}} \bar{V}_2^H V_1 \Lambda_1,
\]
\[
C_3 \stackrel{\text{def}}{=} \bar{\bar{\Lambda}}_2 R_A + R_{A^H} \Lambda_1 = \bar{\bar{\Lambda}}^2 \bar{P}_2^H Q_1 - \bar{\bar{\bar{Q}}} \bar{P}_2^H P_1 \Lambda_1^2,
\]
\[
C_4 \stackrel{\text{def}}{=} \bar{\bar{\Omega}}_2 R_B + R_{B^H} \Omega_1 = \bar{\bar{\Omega}}^2 \bar{P}_2^H Q_1 - \bar{\bar{\bar{Q}}} \bar{P}_2^H P_1 \Omega_1^2.
\]
\[
C_5 \stackrel{\text{def}}{=} C_3 \Omega_1^2 - C_4 \Lambda_1^2 = \bar{\bar{\bar{\Lambda}}} \bar{P}_2^H Q_1 \Omega_1^2 - \bar{\bar{\bar{\bar{Q}}} \bar{P}_2^H Q_1 \Omega_1^2},
\]
\[
C_6 \stackrel{\text{def}}{=} \bar{\bar{\Omega}}_2 C_3 + \bar{\bar{\bar{\Lambda}}} \bar{C}_4 = \bar{\bar{\Omega}}^2 \bar{P}_2^H P_1 \Lambda_1^2 - \bar{\bar{\bar{\bar{Q}}} \bar{P}_2^H P_1 \Omega_1^2}.
\]

**Theorem 8.1.** The conditions and notations are as described in Theorem 6.1. Then
\[
\max_{i=1,2} \|\sin \Theta_i\| \leq \frac{1}{\eta} \max \left\{ \left\| \left( \begin{array}{c} \bar{U}_2^H \\ \bar{V}_2^H \end{array} \right) (Z - \bar{Z}) Q_{10} \right\| \left\| Z^+ \right\|_2, \right. \\
\left. \left\| \left( \begin{array}{c} U_1^H \\ V_1^H \end{array} \right) (Z - \bar{Z}) Q_{20} \right\| \left\| \bar{Z}^+ \right\|_2 \right\},
\]
\[
\max_{i=1,2} \|\sin \Theta_i\| \leq \frac{1}{\eta} \left\{ \left\| \left( \begin{array}{c} \bar{U}_2^H \\ \bar{V}_2^H \end{array} \right) (Z - \bar{Z}) Q_{10} \right\| \left\| Z^+ \right\|_2, \right. \\
\left. \left\| \left( \begin{array}{c} U_1^H \\ V_1^H \end{array} \right) (Z - \bar{Z}) Q_{20} \right\| \left\| \bar{Z}^+ \right\|_2 \right\},
\]
\[
\max_{i=1,2} \|\sin \Theta_i\| \leq \frac{1}{\eta} \left\{ \left\| \left( \begin{array}{c} \bar{U}_2^H \\ \bar{V}_2^H \end{array} \right) (Z - \bar{Z}) Q_{10} \right\| \left\| Z^+ \right\|_2, \right. \\
\left. \left\| \left( \begin{array}{c} U_1^H \\ V_1^H \end{array} \right) (Z - \bar{Z}) Q_{20} \right\| \left\| \bar{Z}^+ \right\|_2 \right\},
\]
and

\begin{equation}
\| \sin \Theta_3 \| \leq \frac{1}{\eta} \| \tilde{Z}^+ \|_2 \| Z \|_2 \left\{ \left\| \begin{pmatrix} \tilde{U}_2^H (\tilde{A} - A) Q_{10} \\
\tilde{V}_2^H (\tilde{B} - B) Q_{10} \end{pmatrix} \right\| \| Z^+ \|_2 \\
+ \frac{1}{2} \left\| \begin{pmatrix} U_1^H \\
V_1^H \end{pmatrix} (Z - \tilde{Z}) \tilde{Q}_{20} \right\| \| \tilde{Z}^+ \|_2 \right\},
\end{equation}

\begin{equation}
\| \sin \Theta_4 \| \leq \frac{1}{\eta} \| \tilde{Z} \|_2 \| Z^+ \|_2 \left\{ \frac{1}{2} \left\| \begin{pmatrix} \tilde{U}_2^H \\
\tilde{V}_2^H \end{pmatrix} (Z - \tilde{Z}) Q_{10} \right\| \| Z^+ \|_2 \\
+ \left\| \begin{pmatrix} U_1^H (\tilde{A} - A) \tilde{Q}_{20} \\
V_1^H (\tilde{B} - B) \tilde{Q}_{20} \end{pmatrix} \right\| \| \tilde{Z}^+ \|_2 \right\},
\end{equation}

where \( Q_{j0}, \ldots \) are defined by (5.7),

\begin{equation}
\eta \triangleq \frac{\min_{(\alpha, \beta) \in \Gamma_1} \frac{|\alpha^2 \beta^2 - \beta^2 \tilde{\alpha}^2|}{(\alpha^2 + \beta^2)(\tilde{\alpha}^2 + \tilde{\beta}^2)}}{\min_{(\alpha, \beta) \in \Gamma_2} \frac{|\alpha^2 \beta^2 - \beta^2 \tilde{\alpha}^2|}{(\alpha^2 + \beta^2)(\tilde{\alpha}^2 + \tilde{\beta}^2)}}.
\end{equation}

If, instead of Description I, we use its equivalence, Description II, to describe the separation between \( \sigma \{ A_{11}, B_{11} \} \) and \( \sigma \{ A_{22}, B_{22} \} \), i.e., there exist \( \tilde{\alpha} \geq 0, \delta > 0, \alpha + \delta \leq 1 \) such that

\begin{equation}
\max_{(\tilde{\alpha}, \tilde{\beta}) \in \sigma \{ A_{22}, B_{22} \}} \rho((\tilde{\alpha}, \tilde{\beta}), (0, 1)) \leq \alpha, \quad \min_{(\alpha, \beta) \in \sigma \{ A_{11}, B_{11} \}} \rho((\alpha, \beta), (0, 1)) \geq \alpha + \delta,
\end{equation}

then

\begin{equation}
\| \sin \Theta_3 \| \leq \frac{1}{\sqrt{1 - \alpha^2}} \| \tilde{Z}^+ \|_2 \| Z \|_2 \left\{ \left\| \tilde{U}_2^H (\tilde{B} - B) Q_{10} \right\| \| Z^+ \|_2 + \sqrt{1 - (\alpha + \delta)^2} \| \sin \Theta_2 \| \right\},
\end{equation}

\begin{equation}
\| \sin \Theta_4 \| \leq \frac{1}{\alpha + \delta} \| \tilde{Z} \|_2 \| Z^+ \|_2 \left\{ \left\| \tilde{U}_2^H (\tilde{A} - A) Q_{20} \right\| \| \tilde{Z}^+ \|_2 + \alpha \| \sin \Theta_1 \| \right\},
\end{equation}

\begin{equation}
\| \sin \Theta_4 \| \leq \frac{1}{\alpha + \delta} \| \tilde{Z} \|_2 \| Z^+ \|_2 \left\{ \left\| \tilde{U}_2^H (\tilde{A} - A) Q_{20} \right\| \| \tilde{Z}^+ \|_2 + \alpha \| \sin \Theta_1 \| \right\}.
\end{equation}

Proof. Without loss of generality, assume that \( \{ A, B \} \) and \( \{ \tilde{A}, \tilde{B} \} \) have decompositions as described in the beginning of §7 up to (7.3).

For the case of \( m = p = n \), (8.6) follows from (5.6), (7.9), and (8.4), since by (8.1) and (8.3), we have

\begin{align*}
C_1 &= (-\tilde{\Omega}_2, \tilde{\Lambda}_2) \begin{pmatrix} R_A \\ R_B \end{pmatrix} = \| C_1 \| \leq \| (-\tilde{\Omega}_2, \tilde{\Lambda}_2) \|_2 \left\| \begin{pmatrix} R_A \\ R_B \end{pmatrix} \right\| = \left\| \begin{pmatrix} R_A \\ R_B \end{pmatrix} \right\|, \\
C_2^H &= (\Omega_1, -\Lambda_1) \begin{pmatrix} R_A^H \\ R_B^H \end{pmatrix} = \| C_2^H \| \leq \| (\Omega_1, -\Lambda_1) \|_2 \left\| \begin{pmatrix} R_A^H \\ R_B^H \end{pmatrix} \right\| = \left\| \begin{pmatrix} R_A^H \\ R_B^H \end{pmatrix} \right\|,
\end{align*}
and
\[
\begin{align*}
\left\| \begin{pmatrix} R_A \\ V_{B}^{H} \end{pmatrix} \right\| & = \left\| \begin{pmatrix} U_{2}^{H} \\ \bar{V}_{2}^{H} \end{pmatrix} (Z - \bar{Z})Q_{1} \right\| \leq \left\| \begin{pmatrix} U_{2}^{H} \\ \bar{V}_{2}^{H} \end{pmatrix} (Z - \bar{Z})Q_{10} \right\| Z^{+} \left\|_{2}, \\
\left\| \begin{pmatrix} R_{A}^{H} \\ R_{B}^{H} \end{pmatrix} \right\| & = \left\| \begin{pmatrix} U_{1}^{H} \\ V_{1}^{H} \end{pmatrix} (Z - \bar{Z})\bar{Q}_{2} \right\| \leq \left\| \begin{pmatrix} U_{1}^{H} \\ V_{1}^{H} \end{pmatrix} (Z - \bar{Z})\bar{Q}_{20} \right\| \bar{Z}^{+} \left\|_{2}.
\end{align*}
\]

Here we have used the following facts in the above derivation.

\[
\begin{align*}
\left\| \begin{pmatrix} -\bar{\Omega}_{2}, \bar{\Lambda}_{2} \end{pmatrix} \right\|_{2} = & \max_{j} \sqrt{\alpha_{j}^{2} + \beta_{j}^{2}} = 1, \quad \left\| (\Omega_{i}, -\Lambda_{i}) \right\|_{2} = \max_{i} \sqrt{\alpha_{i}^{2} + \beta_{i}^{2}} = 1, \\
Q_{1} = & Q_{10}(Q_{1}^{H}Q_{1})^{-\frac{1}{2}}(Q_{1}^{H}Q_{1})^{\frac{1}{2}}, \quad Q_{2} = \bar{Q}_{20}Q_{2}^{H}\bar{Q}_{2}, \\
\left\| (Q_{1}^{H}Q_{1})^{\frac{1}{2}} \right\|_{2} = & \left\| Q_{1} \right\|_{2} \leq \left\| (Q_{1}, Q_{2}) \right\|_{2} \leq \left\| Z^{+} \right\|_{2} \quad \text{(by (6.6))}, \\
\left\| \bar{Q}_{2}^{H}\bar{Q}_{2} \right\|_{2} = & \left\| \bar{Z}^{+} \right\|_{2}.
\end{align*}
\]

Inequalities (8.7) and (8.8) follow from the proof of (7.9) in §7 (see also Li [18]), (8.5), (8.13), (8.14), Lemmas 1.1 and 1.2, and (5.6), since

\[
\begin{align*}
\left\| \begin{pmatrix} \bar{P}_{2}^{H}Q_{1} \end{pmatrix} \right\| & \geq \left\| \begin{pmatrix} \bar{P}_{2}^{H}\bar{P}_{2} \end{pmatrix}^{-\frac{1}{2}} \right\| \left\| \begin{pmatrix} \bar{P}_{20}^{H}Q_{10} \end{pmatrix} \right\| \left\| (Q_{1}^{H}Q_{1})^{-\frac{1}{2}} \right\|_{2}^{-1} \left\| \sin \Theta_{3} \right\| \left\| Q^{-1} \right\|_{2}^{-1}, \\
\Rightarrow \left\| \sin \Theta_{3} \right\| & \leq \left\| Z \right\|_{2} \left\| \bar{Z}^{+} \right\|_{2} \left\| \begin{pmatrix} \bar{P}_{2}^{H}Q_{1} \end{pmatrix} \right\|, \\
\left\| \sin \Theta_{4} \right\| & \leq \left\| Z^{+} \right\|_{2} \left\| \bar{Z} \right\|_{2} \left\| \begin{pmatrix} \bar{Q}_{2}^{H}P_{1} \end{pmatrix} \right\|,
\end{align*}
\]

and

\[
\begin{align*}
C_{5} = & (\bar{\Lambda}_{2}R_{A} + R_{A}^{H}\Lambda_{1})\Omega_{1}^{2} - (\bar{\Omega}_{2}R_{B} + R_{B}^{H}\Omega_{1})\Lambda_{1}^{2} \\\n= & (\bar{\Lambda}_{2} - \bar{\Omega}_{2}) \begin{pmatrix} R_{A} \\ R_{B} \end{pmatrix} \begin{pmatrix} \Omega_{1}^{2} \\ \Lambda_{1}^{2} \end{pmatrix} + (R_{A}^{H}, R_{B}^{H}) \begin{pmatrix} -\Lambda_{1} \end{pmatrix} \Omega_{1} \Lambda_{1}, \\
C_{6} = & -\bar{\Lambda}_{2}(\bar{\Lambda}_{2}R_{A} + R_{A}^{H}\Lambda_{1}) + \bar{\Omega}_{2}(\bar{\Omega}_{2}R_{B} + R_{B}^{H}\Omega_{1}) \\\n= & \bar{\Lambda}_{2}\bar{\Omega}_{2}(-\bar{\Lambda}_{2}, \bar{\Lambda}_{2}) \begin{pmatrix} R_{A} \\ R_{B} \end{pmatrix} + (\bar{\Omega}_{2}, \bar{\Lambda}_{2}) \begin{pmatrix} R_{A}^{H} \\ R_{B}^{H} \end{pmatrix} \begin{pmatrix} -\Lambda_{1} \end{pmatrix} \Omega_{1},
\end{align*}
\]

which produce

\[
\begin{align*}
\left\| C_{5} \right\| \leq & \left\| \begin{pmatrix} R_{A} \\ R_{B} \end{pmatrix} \right\| \left\| \begin{pmatrix} \Omega_{1}^{2} \\ \Lambda_{1}^{2} \end{pmatrix} \right\|_{2} + \left\| (R_{A}^{H}, R_{B}^{H}) \right\| \left\| \Omega_{1} \Lambda_{1} \right\|_{2} \\
\leq & \left\| \begin{pmatrix} R_{A} \\ R_{B} \end{pmatrix} \right\| + \frac{1}{2} \left\| \begin{pmatrix} R_{A}^{H} \\ R_{B}^{H} \end{pmatrix} \right\|, \\
\left\| C_{6} \right\| \leq & \frac{1}{2} \left\| \begin{pmatrix} R_{A} \\ R_{B} \end{pmatrix} \right\| + \left\| \begin{pmatrix} R_{A}^{H} \\ R_{B}^{H} \end{pmatrix} \right\|;
\end{align*}
\]
and
\[
\left\| \begin{pmatrix} R_A & R_B \\ \overline{R}_A & \overline{R}_B \end{pmatrix} \right\| = \left\| \begin{pmatrix} \overline{U}_2^H (\overline{A} - A) Q_{10} \\ \overline{V}_2^H (\overline{B} - B) Q_{10} \end{pmatrix} \right\| Z^+ \|_2, \\
\left\| \begin{pmatrix} \overline{R}_A^H \\ \overline{R}_B^H \end{pmatrix} \right\| = \left\| \begin{pmatrix} U_1^H (\overline{A} - A) \overline{Q}_{20} \\ V_1^H (\overline{B} - B) \overline{Q}_{20} \end{pmatrix} \right\| \| \overline{Z}^+ \|_2.
\]

In the derivation of (8.18), we have noted
\[
\| \Omega_1 \Lambda_1 \|_2 = \max_i \alpha_i \beta_i \leq \frac{1}{2}, \quad \left\| \begin{pmatrix} \Omega_1^2 \\ \Lambda_1^2 \end{pmatrix} \right\|_2 = \max_i \sqrt{\alpha_i^4 + \beta_i^4} \leq 1,
\]
and some other inequalities of the same kinds. In proving the last relation in (8.15), we need
\[
\left\| (Q_1^H Q_1)^{-\frac{1}{2}} \right\|^{-1} = \sigma_{\min} (Q_1) \geq \sigma_{\min} (Q),
\]
where \( \sigma_{\min} (\cdot) \) denotes the smallest singular value of a matrix. Inequalities (8.11) and (8.12) follow from the second and the third equations in (8.3), (8.16), (8.17), and (8.10), since \( \alpha + \delta \geq \eta, \sqrt{1 - \alpha^2} \geq \eta \) by (5.11), and
\[
\| \Lambda^{-1} \|_2 \geq \alpha + \delta, \quad \| \Omega_1 \|_2 \leq \sqrt{1 - (\alpha + \delta)^2}, \quad \| \hat{\Lambda} \|_2 \leq \alpha, \quad \| \hat{\Omega}_1^{-1} \|_2 \geq \sqrt{1 - \alpha^2},
\]
also
\[
\sqrt{1 - \alpha^2} \| \hat{P}_Q Q_1 \| \leq \| \hat{\Omega}^{-1}_1 \|_2 \| \hat{P}_Q Q_1 \| \leq \| \hat{\Omega} \|_2 \| \hat{P}_Q Q_1 \|
\]
\[
\leq \| R_B \| + \| V_2^H V_1 \| \sqrt{1 - (\alpha + \delta)^2},
\]

(8.19) \[
\left\| \begin{pmatrix} \sin \Theta_1 \\ \sin \Theta_2 \end{pmatrix} \right\| \leq c \cdot \frac{1}{\eta} \left\| \begin{pmatrix} \overline{U}_2^H \\ \overline{V}_2^H \end{pmatrix} (Z - \overline{Z}) Q_{10} \right\| Z^+ \|_2 \\
+ \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (Z - \overline{Z}) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]

and
(8.20) \[
\| \sin \Theta_3 \| \leq c \cdot \frac{1}{\eta} \| \overline{Z}^+ \|_2 \| Z \|_2 \left\| \begin{pmatrix} \overline{U}_2^H (\overline{A} - A) Q_{10} \\ \overline{V}_2^H (\overline{B} - B) Q_{10} \end{pmatrix} \right\| Z^+ \|_2 \\
+ \frac{1}{2} \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (Z - \overline{Z}) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]

(8.21) \[
\| \sin \Theta_4 \| \leq c \cdot \frac{1}{\eta} \| \overline{Z} \|_2 \| Z^+ \|_2 \left\| \begin{pmatrix} \overline{U}_2^H \\ \overline{V}_2^H \end{pmatrix} (Z - \overline{Z}) Q_{10} \right\| Z^+ \|_2 \\
+ \frac{1}{2} \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (\overline{A} - A) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]

The proof of the theorem for the case of \( m \neq n \) or \( p \neq n \) can be given in an analogous way to those in §7. We omit the details here. \( \Box \)

The following theorem, in a somewhat similar spirit to Theorems 6.2 and 6.3, can also be proved analogously.

**Theorem 8.2.** The conditions and notations are as described in Theorem 6.2. Then there exists a constant \( c \geq 1 \) such that for any unitarily invariant norm \( \| \cdot \| \) we have

(8.19) \[
\left\| \begin{pmatrix} \sin \Theta_1 \\ \sin \Theta_2 \end{pmatrix} \right\| \leq c \cdot \frac{1}{\eta} \left\| \begin{pmatrix} \overline{U}_2^H \\ \overline{V}_2^H \end{pmatrix} (Z - \overline{Z}) Q_{10} \right\| Z^+ \|_2 \\
+ \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (Z - \overline{Z}) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]

and
(8.20) \[
\| \sin \Theta_3 \| \leq c \cdot \frac{1}{\eta} \| \overline{Z}^+ \|_2 \| Z \|_2 \left\| \begin{pmatrix} \overline{U}_2^H (\overline{A} - A) Q_{10} \\ \overline{V}_2^H (\overline{B} - B) Q_{10} \end{pmatrix} \right\| Z^+ \|_2 \\
+ \frac{1}{2} \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (Z - \overline{Z}) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]

(8.21) \[
\| \sin \Theta_4 \| \leq c \cdot \frac{1}{\eta} \| \overline{Z} \|_2 \| Z^+ \|_2 \left\| \begin{pmatrix} \overline{U}_2^H \\ \overline{V}_2^H \end{pmatrix} (Z - \overline{Z}) Q_{10} \right\| Z^+ \|_2 \\
+ \frac{1}{2} \left\| \begin{pmatrix} U_1^H \\ V_1^H \end{pmatrix} (\overline{A} - A) Q_{20} \right\| \| \overline{Z}^+ \|_2,
\]
where

\[
\hat{\eta} \overset{\text{def}}{=} \min_{\ell+1 \leq i \leq n} \frac{|\alpha_i^2 \beta_i^2 - \beta_i^2 \alpha_i^2|}{(\alpha_i + \beta_i^2)(\alpha_i^2 + \beta_i^2)},
\]

\[
\hat{\eta}_1 \overset{\text{def}}{=} \min_{\ell+1 \leq i \leq n} \frac{|1^2 \beta_i^2 - 0^2 \alpha_i^2|}{(1^2 + 0^2)(\alpha_i^2 + \beta_i^2)},
\]

\[
\hat{\eta}_2 \overset{\text{def}}{=} \min_{1 \leq i \leq t} \frac{|\alpha_i^2 0^2 - \beta_i^2 1^2|}{(\alpha_i^2 + \beta_i^2)(0^2 + 1^2)},
\]

and

\[
0 < \tilde{\eta} \overset{\text{def}}{=} \begin{cases} 
\hat{\eta} & \text{if } m \leq n \text{ and } p \leq n, \\
\min\{\hat{\eta}, \hat{\eta}_1\} & \text{if } m \leq n < p, \\
\min\{\hat{\eta}, \hat{\eta}_2\} & \text{if } m > n \geq p, \\
\min\{\hat{\eta}, \hat{\eta}_1, \hat{\eta}_2\} & \text{if } m > n \text{ and } p > n.
\end{cases}
\]

Generally, \( c \) can be chosen so that (6.10) holds, but for \( \| \cdot \| = \| \cdot \|_F \), (8.19)–(8.21) can be improved in the following way: replacing \( c \) in (8.20) and (8.21) by 0 and 1 respectively.

\[
\| \sin \Theta_1 \|_2^2 + \| \sin \Theta_2 \|_2^2 \leq \frac{1}{\tilde{\eta}} \left[ \left\| \begin{pmatrix} U^H_2 & V^H_2 \end{pmatrix} (Z - \tilde{Z}) Q_{10} \right\|_F^2 \left\| Z^+ + \tilde{Z}^+ \right\|_2^2 \\
+ \left\| \begin{pmatrix} U^H_1 & V^H_1 \end{pmatrix} (Z - \tilde{Z}) Q_{20} \right\|_F^2 \left\| Z^+ + \tilde{Z}^+ \right\|_2^2 \right]^{\frac{1}{2}}.
\]

We have illustrated in detail in §6 how much weaker the gap \( \tilde{\eta} \) (defined by (6.5)) is than \( \eta \) (defined by (5.9)). Similar arguments hold for the weakness of \( \tilde{\eta} \) defined by (8.23) with respect to \( \tilde{\eta} \) defined by (8.9). There is a straightforward interpretation of the appearance of \( \tilde{\eta} \) (cf. \( \tilde{\eta} \)). We now have an idea of how the generalized singular value problem closely relates to the generalized eigenvalue problem for definite pencils. Specifically, \((\alpha, \beta) \in \sigma(A, B) \Rightarrow (\alpha_i^2, \beta_i^2) \in \lambda(A^H A - \lambda B^H B)\), the subspace spanned by the column vectors of \( Q_{10} \), is an eigenspace of the definite pencil \( A^H A - \lambda B^H B \) belonging to the generalized eigenvalues \((\alpha_i^2, \beta_i^2), i = 1, 2, \ldots, \ell\), and so on. \( \tilde{\eta} \) (cf. \( \tilde{\eta} \)) should then be regarded as a gap in a sense of the chordal metric between two certain subsets of the spectra of the corresponding definite pencils, since for \( \alpha, \beta, \alpha, \beta \geq 0, \)

\[
\frac{1}{\sqrt{2}} \rho((\alpha, \beta), (\alpha^2, \beta^2), (\alpha^2, \beta^2)) \leq \frac{|\alpha^2 \beta^2 - \beta^2 \alpha^2|}{(\alpha^2 + \beta^2)(\alpha^2 + \beta^2)} \leq \rho((\alpha^2, \beta^2), (\alpha^2, \beta^2)).
\]

In this case, bounds (8.7), (8.8), (8.20), and (8.21) reflect on the hidden relation between the problems, though it might be impossible to get these bounds via those for the eigenspaces for definite pencils [17], [18], [32].

As to the sharpness of our bounds in §§6 and 8, to our knowledge, it may be impossible to know exactly which one is better than others. But generally we propose that, apart from weak assumptions on gaps between certain sets of GSVs, bounds
(6.2), (6.3), (6.6), (6.7), (6.8), (6.9), (6.8), and (8.6) are equally effective. They may be only a constant factor apart. For \( \sin \Theta_3 \) and \( \sin \Theta_4 \), bounds (8.7) and (8.8) may be much less sharp than (8.11) and (8.12). To see this, we note

\[
\eta \geq \tilde{\eta} \geq \eta^2,
\]

and \( \tilde{\eta} \) approaching \( \eta^2 \) indeed occurs in some cases. Thus, if \( \eta \) is also very small, then the right-hand sides of (8.7) and (8.8) must be much larger than those of (8.11) and (8.12). Why should we bother to establish (8.7) and (8.8)? We argue that (8.7) and (8.8) are of theoretical interest. Also, the perturbation equations (8.5) used to prove (8.7) and (8.8) can be used to prove (8.20) and (8.21), which are valid under weaker assumptions that certain subsets of GSVs of corresponding matrix pairs are not required to be well separated, as shown by Fig. 2 and (6.1).

As our derivation for bounds seems a little more elegant than that of Sun [30], it would be natural for us to expect that our bounds should be better. To enable us to compare ours with Sun's, we first simplify bounds due to Sun [30]. In our process of simplification, no essential amplification is involved, i.e., there is no essential loss of sharpness in the simplified bounds in comparison with the original ones of Sun [30]. We note that Sun [30] presented only bounds for \( \sin \Theta_i \), \( i = 1, 2, 3, 4 \), under the assumptions of Theorem 6.1. So now we assume that the conditions and notations are as described in Theorem 6.1, and Description II (refer to (8.10)), as a description equivalent to Description I, is also referred to when suitable; \( \eta \) is defined by (5.11). In our notation, Sun's bounds are read as

\[
\| \sin \Theta_1 \| \leq \frac{\tau_1}{\delta} \left( (r_a + \omega_1 r_b) \| Z^+ \|_2 + \omega_1 \omega_2 (\omega_3 r_a' + r_b') \| \tilde{Z}^+ \|_2 \right),
\]

(8.26)

\[
\| \sin \Theta_2 \| \leq \frac{\tau_2}{\delta} \left( \omega_2 \omega_4 (r_a + \omega_1 r_b) \| Z^+ \|_2 + (\omega_3 r_a' + r_b') \| \tilde{Z}^+ \|_2 \right),
\]

(8.27)

\[
\| \sin \Theta_3 \| \leq \frac{\tau_3}{\delta} \left( \omega_2 \omega_4 (r_a + \omega_1 r_b) \| Z^+ \|_2 + \omega_2 \omega_4 (\omega_3 r_a' + r_b') \| \tilde{Z}^+ \|_2 \right),
\]

(8.28)

\[
\| \sin \Theta_4 \| \leq \frac{\tau_4}{\delta} \left( \omega_4 (r_a + \omega_1 r_b) \| Z^+ \|_2 + (r_a' + \omega_1 \omega_2 \omega_4 r_b') \| \tilde{Z}^+ \|_2 \right),
\]

where

\[
\tau_1 = \frac{(\alpha + \delta)(1 - \alpha^2)}{2 \alpha + \delta}, \quad \tau_2 = \frac{(\alpha + \delta)^2 \sqrt{1 - \alpha^2}}{2 \alpha + \delta},
\]

\[
\omega_1 = \frac{\alpha}{\sqrt{1 - \alpha^2}}, \quad \omega_2 = \sqrt{\frac{1 - (\alpha + \delta)^2}{1 - \alpha^2}}, \quad \omega_2 = \frac{\sqrt{1 - (\alpha + \delta)^2}}{\alpha + \delta}, \quad \omega_4 = \frac{\alpha}{\alpha + \delta},
\]

\[
r_a = \left\| (\tilde{A} - A) Q_{10} \right\|, \quad r_b = \left\| (\tilde{B} - B) Q_{10} \right\|, \quad r_a' = \left\| U_{1}^{H} (\tilde{A} - A) \right\|, \quad r_b' = \left\| V_{1}^{H} (\tilde{B} - B) \right\|,
\]

Denote \( d = (\alpha + \delta) \sqrt{1 - \alpha^2} + \alpha \sqrt{1 - (\alpha + \delta)^2} \). For \( \sin \Theta_1 \) and \( \sin \Theta_2 \), it follows from

\[
\eta \frac{\tau_1}{\delta} = \frac{(\alpha + \delta) \sqrt{1 - \alpha^2}}{d \sqrt{1 - \alpha^2}}, \quad \eta \frac{\tau_1}{\delta} \omega_1 = \frac{(\alpha + \delta) \sqrt{1 - \alpha^2}}{d \alpha},
\]

\[
\frac{\tau_1}{\delta} \omega_1 \omega_2 = \frac{\alpha \sqrt{1 - (\alpha + \delta)^2}}{d (\alpha + \delta)}, \quad \frac{\tau_1}{\delta} \omega_1 \omega_2 \omega_3 = \frac{\alpha \sqrt{1 - (\alpha + \delta)^2}}{d \sqrt{1 - (\alpha + \delta)^2}},
\]
and
\[ \eta \frac{\tau_2}{\delta} = \frac{(\alpha + \delta) \sqrt{1 - \alpha^2}}{d} (\alpha + \delta), \quad \eta \frac{\tau_2}{\delta} \omega_2 = \frac{(\alpha + \delta) \sqrt{1 - \alpha^2}}{d} \sqrt{1 - (\alpha + \delta)^2}, \]
\[ \eta \frac{\tau_2}{\delta} \omega_3 \omega_4 = \frac{\alpha \sqrt{1 - (\alpha + \delta)^2}}{d} \sqrt{1 - \alpha^2}, \quad \eta \frac{\tau_2}{\delta} \omega_3 \omega_4 \omega_1 = \frac{\alpha \sqrt{1 - (\alpha + \delta)^2}}{d} \alpha \]
that
\[ \max_{i=1,2} \| \sin \Theta_i \| \leq \frac{1}{\eta} \max \left\{ \sqrt{r_a^2 + r_b^2} \| Z^\|_2, \sqrt{r_a^2 + r_b^2} \| \tilde{Z}^\|_2 \right\}. \]

The simplifications of (8.27) and (8.28) are in the same spirit. Here we give the final results and leave the details to the reader:

\[ \| \sin \Theta_3 \| \leq \frac{1}{\eta} \cdot 2 \| Z \|_2 \| \tilde{Z} \|_2 \max \left\{ \sqrt{r_a^2 + r_b^2} \| Z^+ \|_2, \sqrt{r_a^2 + r_b^2} \| \tilde{Z}^+ \|_2 \right\}, \]
\[ \| \sin \Theta_4 \| \leq \frac{1}{\eta} \cdot 2 \| \tilde{Z} \|_2 \| Z \|_2 \max \left\{ \sqrt{r_a^2 + r_b^2} \| Z^+ \|_2, \sqrt{r_a^2 + r_b^2} \| \tilde{Z}^+ \|_2 \right\}. \]

Now turning to our bounds (8.6), (8.11), and (8.12), we note that
\[ \left\| \begin{pmatrix} U_2^H & V_2^H \\ 
\end{pmatrix} (Z - \tilde{Z}) Q_{10} \right\| \leq \left\| \begin{pmatrix} A - \bar{A} \\ B - \bar{B} \end{pmatrix} Q_{10} \right\| \]
\[ \leq \sqrt{2} \sqrt{r_a^2 + r_b^2} \leq 2 \left\| \begin{pmatrix} A - \bar{A} \\ B - \bar{B} \end{pmatrix} Q_{10} \right\|, \]
\[ \left\| \begin{pmatrix} U_2^H & V_2^H \\ 
\end{pmatrix} (Z - \tilde{Z}) \tilde{Q}_{20} \right\| \leq \left\| \begin{pmatrix} U_1^H & V_1^H \\ 
\end{pmatrix} \begin{pmatrix} A - \bar{A} \\ B - \bar{B} \end{pmatrix} \right\| \]
\[ \leq \sqrt{2} \sqrt{r_a^2 + r_b^2} \leq 2 \left\| \begin{pmatrix} U_1^H & V_1^H \\ 
\end{pmatrix} \begin{pmatrix} A - \bar{A} \\ B - \bar{B} \end{pmatrix} \right\|. \]

Therefore (8.6) and (8.29), thus (8.25) and (8.26) as well, would be equally effective. By the way, we should say that the simplified bounds (8.29)–(8.31) of Sun’s bounds seem more favorable than their original (8.25)–(8.28).

**Acknowledgment.** The author acknowledges helpful advice from the referees. He also acknowledges that an anonymous referee pointed out that bounds for sin \( \Theta_3 \) are of great practical interest and suggested the inclusion of such bounds in the revised version of this paper. He feels indebted to Professor Axel Ruhe for reading the manuscript so carefully and for very detailed and constructive suggestions for revising this paper.

**REFERENCES**
