Linear Algebra and its Applications 666 (2023) 62–95

Contents lists available at ScienceDirect

Linear Algebra and its Applications

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

New lower bounds on the minimum singular value of a matrix ☆

Avleen Kaur *, S.H. Lui

Department of Mathematics, University of Manitoba, Winnipeg, MB R3T 2N2, Canada

A R T I C L E   I N F O

Article history:
Received 1 April 2022
Accepted 17 February 2023
Available online 23 February 2023
Submitted by V. Mehrmann

MSC:
15A18
15A42
15A60
15B57

Keywords:
Block matrix
Eigenvalue
Lower bound
Friedrichs angle
Real symmetric matrices
Singular value

A B S T R A C T

The study of constraining the eigenvalues of the sum of two symmetric matrices, say $P + Q$, in terms of the eigenvalues of P and Q, has a long tradition. It is closely related to estimating a lower bound on the minimum singular value of a matrix, which has been discussed by a great number of authors. To our knowledge, no study has yielded a positive lower bound on the minimum eigenvalue, $\lambda_{\min}(P+Q)$, when $P + Q$ is symmetric positive definite with P and Q singular positive semi-definite. We derive two new lower bounds on $\lambda_{\min}(P+Q)$ in terms of the minimum positive eigenvalues of P and Q. The bounds take into account geometric information by utilizing the Friedrichs angles between certain subspaces. The basic result is when P and Q are two non-zero singular positive semi-definite matrices such that $P+Q$ is non-singular, then $\lambda_{\min}(P + Q) \geq (1 - \cos\theta_F) \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$, where $\lambda_{\min}$ represents the minimum positive eigenvalue of the matrix, and θ_F is the Friedrichs angle between the range spaces of P and Q. Such estimates lead to new lower bounds on the minimum singular value of full rank 1×2, 2×1, and 2×2 block matrices. Some examples provided in this paper further

☆ Funding: This work is supported by a University of Manitoba Graduate Fellowship (UMGF) from the University of Manitoba and a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC) under grant number RGPIN-2022-03665.

* Corresponding author.
E-mail addresses: kaura349@myumanitoba.ca (A. Kaur), shaun.lui@umanitoba.ca (S.H. Lui).

https://doi.org/10.1016/j.laa.2023.02.013

highlight the simplicity of applying the results in comparison to some existing lower bounds.

1. Introduction

The spectral problem of a symmetric matrix sum estimates the eigenvalues of a sum of two symmetric matrices $P+Q$, in terms of the eigenvalues of P and Q. Fundamental results, like Weyl's inequality in [25, p. 239], and several other works collected in [15], have addressed this problem. Another substantial contribution is Horn's conjecture proved in [32,33]. The present work is focused on the case when P and Q are symmetric positive semi-definite (PSD) matrices, which impacts numerous areas-such as computational economics, graph theory, perturbation theory, semi-definite programming, spectrum of self-adjoint operators, among others. As variance-covariance matrices are PSD, this problem appears in statistics, and more recently in statistical machine learning and spectral methods for data science, discussed in [4] and [13], respectively.

Singular values have been investigated for more than a century. For a square real matrix, its minimum singular value is less than or equal to its absolute minimum eigenvalue. Thus, formulation of a lower bound for the minimum singular value is an influential problem appearing in several studies including the condition number estimates of a matrix, resonant frequencies, population dynamics, principal component analysis, etc. Since the singular values of a matrix are the square-root of eigenvalues of its corresponding gram matrix, the singular values of a general block matrix are associated with the spectral problem of a sum of symmetric matrices. Although a myriad of research has been done on these topics, however, when the symmetric matrices are both singular PSD, we could not find a result providing a positive lower bound even if their sum is non-singular.

In practice, we often come across symmetric positive definite (SPD) matrices represented as a sum of two singular PSD matrices. To illustrate, let us estimate the minimum singular value of a full rank block column matrix, say $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$ so that A_1 and A_2 are rank deficient. This problem is equivalent to finding the minimum eigenvalue of $A^T A = A_1^T A_1 + A_2^T A_2$, an SPD matrix which is a sum of two singular PSD matrices. We derive a positive lower bound on the minimum singular value of A in terms of the minimum positive singular values of A_1 and A_2 in Corollary 3.7.

In this work, we desire a positive lower bound on the minimum eigenvalue of an SPD matrix $P + Q$, where $P, Q \in \mathbb{R}^{n\times n}$ are PSD matrices. Two positive lower bounds on the smallest eigenvalue of $P + Q$, framed in terms of the smallest positive eigenvalues of P and Q, are presented in Theorems 3.1 and 3.5. These estimates of the minimum eigenvalue employ the Friedrichs angle between certain subspaces, i.e., some principal angle between them, as given in Proposition 2.4. Moreover, the two new lower bounds lead to useful outcomes when applied to a 2×2 non-singular block matrix $\mathcal{X}$ as:

$$\mathcal{X}^T\mathcal{X} = \begin{bmatrix} A^T & C^T \\ B^T & D^T \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A^TA & A^TB \\ B^TA & B^TB \end{bmatrix} + \begin{bmatrix} C^TC & C^TD \\ D^TC & D^TD \end{bmatrix}.$$

Here $\mathcal{X}^T\mathcal{X}$ is a full rank matrix expressed as a sum of two PSD matrices. Therefore, the above expression admits a lower bound on the minimum singular value of $\mathcal{X}$, in terms of the minimum positive singular values of its blocks A, B, C, and D (Theorem 3.9). Finally, the above expression and $\mathcal{X}\mathcal{X}^T$ are used again to get two lower bounds on other singular values of $\mathcal{X}$ in Theorem 3.14.

We now summarize the notation used in this paper. Let $A \in \mathbb{R}^{n\times n}$ be a PSD matrix and let $\Lambda(A)$ denote the spectrum of A, that is, the set of eigenvalue of A. If $\mathfrak{r} = \text{rank}(A)$, then its eigenvalues $\lambda_1(A) \geq \lambda_2(A) \geq \ldots \geq \lambda_n(A) \geq 0$, are such that $\lambda_{\mathfrak{r}} > 0$ and $\lambda_i = 0$ for all $\mathfrak{r} + 1 \leq i \leq n$. Moreover, we define the minimum positive eigenvalue of A as

$$\lambda_{\min}(A) := \begin{cases} \lambda_{\mathfrak{r}}(A), & \text{if } A \neq O, \\ \infty, & \text{if } A = O. \end{cases} \tag{1.1}$$

For a matrix $A \in \mathbb{R}^{m\times n}$, $\text{rank}(A) = \text{rank}(A^TA) = \text{rank}(AA^T)$. The set of singular values of A is denoted by $\sigma(A)$. Let $\mathfrak{r} = \text{rank}(A)$, then its singular values $\sigma_1(A) \geq \sigma_2(A) \geq \ldots \geq \sigma_{\min(m,n)}(A)$ are such that $\sigma_{\mathfrak{r}}(A) > 0$ and $\sigma_i = 0$ for all $\mathfrak{r} + 1 \leq i \leq \min(m, n)$. Again, we define the minimum positive singular value of A as

$$\sigma_{\min}(A) := \begin{cases} \sigma_{\mathfrak{r}}(A), & \text{if } A \neq O, \\ \infty, & \text{if } A = O. \end{cases} \tag{1.2}$$

The above expressions for $\lambda_{\min}$ and $\sigma_{\min}$ are defined for the convenience of notation for results derived in section 3 and their value is set as infinity for zero matrices to ignore the zeros while calculating the minimum as it was required by the formulations derived for zero matrices. The 2-norm of a vector is denoted by $|\cdot|$, whereas $\|\cdot\|$ represents 2-norm of a matrix. The range space and null space of A are denoted by $\mathcal{R}(A)$ and $\mathcal{N}(A)$, respectively. For a scalar k, the constant vector of n components equal to k is denoted by $\mathbf{k}_n$. Also, for $A_i \in \mathbb{R}^{n_i\times n_i}$ with $n_i \in \mathbb{N}$, $\text{diag}(A_1, A_2, \ldots, A_k)$ denotes a square diagonal matrix of size $\sum_{i=1}^k n_i$ with diagonal blocks A_i, for $1 \leq i \leq k$.

This paper is organized as follows. In section 2, some important results that will be used to prove the main results are listed. In section 3, we prove some new positive lower bounds. Additionally, some examples and special cases for these results are discussed in section 4. Finally, some ideas for future work are mentioned in section 5.

2. Existing results

There is an abundance of results related to the study of the minimum singular value of a non-singular matrix in the literature. We attempt to summarize some of the existing focal results related to the ones developed in section 3.

2.1. Minimum eigenvalue of the sum of two PSD matrices

The problem of estimating a lower bound on the minimum eigenvalue of the sum of two symmetric matrices say $P, Q \in \mathbb{R}^{n\times n}$, has been investigated for many years. Several results describe an upper bound on the spectrum of $P+Q$ in terms of the spectrum of P and Q. One of the most fundamental results is the set of Weyl's inequalities, given in [25, p. 239], which is stated as follows:

$$\lambda_j(P) + \lambda_n(Q) \leq \lambda_j(P+Q) \leq \lambda_j(P) + \lambda_1(Q), \quad j = 1, 2 \ldots, n. \tag{2.1}$$

Other results include trace identity, given as $\sum \lambda(P) + \sum \lambda(Q) = \sum \lambda(P+Q)$, Ky-Fan inequalities, Lidskii and Wielandt inequalities, and Horn's conjecture as proved in [33,32]. In general, it is challenging to improve upon a lower bound on the minimum eigenvalue, $\lambda_n(P+Q) \geq \lambda_n(P) + \lambda_n(Q)$, that is set forth by (2.1).

A more systematic and theoretical analysis was conducted for a specific case when P and Q are PSD. Note that for any $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{\ell\times n}$, their corresponding gram matrices, defined as $A^T A, B^T B \in \mathbb{R}^{n\times n}$, respectively, are PSD; thus, the following lower bounds on eigenvalues of the sum of two PSD matrices were established in [9] and quoted in [8, p. 904]:

$$\lambda_j(A^T A + B^T B) \geq 2\sigma_j(AB^T), \qquad j = 1, 2, \ldots, \min(\ell, m, n). \tag{2.2}$$

R. Bhatia and F. Kittaneh also speculated a generalization of the arithmetic-geometric mean inequality in [10], which stated that for two PSD matrices $P, Q \in \mathbb{R}^{n\times n}$, $\lambda_j(P+Q) \geq 2\sqrt{\sigma_j(PQ)}$, for all $j = 1, 2, \ldots, n$, that were later proved by S. Drury in [17]. However, all these results give a trivial lower bound for the case in which $P+Q$ is SPD and both P and Q are singular PSD matrices.

2.2. Spectrum of saddle point matrices

One of the most commonly seen 2×2 block matrices of the form $\mathcal{X} = \begin{bmatrix} A & B_1 \\ B_2^T & -C \end{bmatrix}$, where $A \in \mathbb{R}^{m\times m}$ and one or both of $B_1, B_2 \in \mathbb{R}^{m\times n}$ are non-zero, is called a saddle point matrix. For a good survey of results on saddle point matrices see [5]. In particular, check Theorem 3.5 in [5, p. 21], which estimates the spectrum for the case when A is SPD, $B_1 = B_2$ is full rank, and $C = O$. A noteworthy improvement was presented in the form of Theorem 1 in [2, p. 341] with a positive or negative semi-definite matrix C. Note that these results provide singular values of a 2×2 block matrix from its corresponding gram matrix; however, estimating the parameters defined in these theorems could be difficult due to their complicated expressions. Another applicable result for the spectrum of a preconditioned saddle point matrix can be found in [43].

To formulate the spectrum of a more generalized saddle point matrix, several advancements have been considered, such as defining $B = B_1^T = -B_2$ in [6,7,42,1]. Another step forward was to have a symmetric indefinite leading block A. The first of such cases was proved in [22], by imposing the condition that A is SPD on $\mathcal{N}(B)$, which was eliminated in [3]. Recently, in [26], A has been considered to be a non-symmetric matrix with a positive definite symmetric part with $C = O$, which originates from discretized Navier-Stokes equations.

2.3. Lower bound on the minimum singular value

Several techniques are reported in the literature for formulating a lower bound on the minimum singular value of a particular type of matrices; however, we attempt to mention seminal contributions to this problem for a general non-singular matrix. An initial result for the special case of diagonally dominant matrices is derived in [46], and for a non-singular matrix, a consequential approach is Gerschgorin-type lower bounds formulated in [41,27]. The results evolved gradually into several stronger versions, as seen in [28,37,29,50]. Also, [24] devised a lower bound in terms of the determinant, the 2-norms of the rows, and columns of the matrix. Some later advancements of this result include [48,49,38] and the references therein.

It is well-known that for a 2×2 block matrix, the maximum singular value is bounded above by 2-norm of the matrix consisting of 2-norms of its blocks, see Theorem 1(f) in [44, p. 2630]. Thus, for $\mathcal{X} \in \mathbb{R}^{n\times n}$,

$$\mathcal{X} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \qquad \|\mathcal{X}\| \leq \left\| \begin{bmatrix} \|A\| & \|B\| \\ \|C\| & \|D\| \end{bmatrix} \right\|.$$

Since $\sigma_{\min}(\mathcal{X}) = \sigma_{\max}(\mathcal{X}^{-1})^{-1}$, so on applying this result to $\mathcal{X}^{-1}$ calculated in terms of its blocks, an estimate of $\sigma_{\min}(\mathcal{X})$ is obtained. One drawback of this method is that the expression for $\mathcal{X}^{-1}$ can be quite problematic.

2.4. Separation between subspaces

A cardinal component of this study is the concept of orthogonality. Let V be an inner-product space and $U \subseteq V$ be a subspace of V, then $U^{\perp} = \{v \in V \mid \langle v, u\rangle = 0, \forall\, u \in U\}$ is the orthogonal complement of U in V, where $\langle \cdot, \cdot \rangle$ is the standard inner product on $\mathbb{R}^n$. A pivotal result, as seen in [39, p. 409], states that

$$U^{\perp} + V^{\perp} = (U \cap V)^{\perp}, \tag{2.3}$$

whenever U, V be two subspaces of an inner-product space X. Our analysis is limited to finite dimensions by using orthogonal projections. Let $U \subseteq \mathbb{R}^n$ be a subspace. A matrix $P \in \mathbb{R}^{n\times n}$ is the orthogonal projection onto U if $\mathcal{R}(P) = U$, $P^2 = P$, and $P^T = P$. Moreover, if $x \in \mathbb{R}^n$, then $Px \in U$ and $(I - P)x \in U^{\perp}$. For a detailed discussion, we refer the reader to [20].

An instinct for formulating a non-trivial lower bound on the minimum eigenvalue of a non-singular sum, $P + Q$, of two PSD matrices encouraged us to gauge the separation between the range spaces of matrices P and Q. In this case, a suitable measure is the principal angle between subspaces.

Definition 2.1 *(Principal angles [19]).* Let $U, V \subseteq \mathbb{R}^n$ be subspaces with $p = \dim(U) \geq \dim(V) = q \geq 1$. The principal angles $0 \leq \theta_1 \leq \theta_2 \leq \ldots \leq \theta_q \leq \frac{\pi}{2}$ between U and V are recursively defined for $k = 1, 2, \ldots, q$ by

$$\cos(\theta_k) = \max_{\substack{u \in U,\, v \in V \\ |u|=|v|=1}} |u^T v| =: u_k^T v_k,$$

subject to the constraints

$$u_i^T u = 0, \quad v_i^T v = 0, \quad i = 1, 2, \ldots, k-1.$$

The vectors $\{u_1, \ldots, u_q\}$, $\{v_1, \ldots, v_q\}$ are called principal vectors of the pair of spaces. The angle θ_1 is also called the minimal principal angle.

It is worth noting that $\theta_1 = 0$ if and only if $U \cap V \neq \{0\}$, and $\theta_1 = \frac{\pi}{2}$ if and only if $U \perp V$. The following result describes a beautiful connection between the orthogonal projections and the principal angles between their subspaces.

Theorem 2.2 *([47,19]). Let $U, V \subseteq \mathbb{R}^n$ be subspaces with $p = \dim(U) \geq \dim(V) = q \geq 1$, θ_i be the principal angles between U and V, and r be the number of angles θ_i such that $0 < \theta_i < \frac{\pi}{2}$. Let $\mathcal{P}, \mathcal{Q} \in \mathbb{R}^{n\times n}$ be orthogonal projections onto U, V, respectively, then the set of singular values of $\mathcal{P} \pm \mathcal{Q}$ are*

1. $\sigma(\mathcal{P} + \mathcal{Q}) = \{\mathbf{2}_k, 1 \pm \cos(\theta_{k+i})(i = 1, \ldots, r), \mathbf{1}_{n_1+n_2}, \mathbf{0}_{n_3}\}$,
2. $\sigma(\mathcal{P} - \mathcal{Q}) = \{\mathbf{1}_{n_1+n_2}, \mathbf{sin}(\boldsymbol{\theta_{k+i}})_2(i = 1, \ldots, r), \mathbf{0}_{k+n_3}\}$,

where $k = \dim(U \cap V)$, $n_1 = \dim(U \cap V^\perp) = p - k - r$, $n_2 = \dim(U^\perp \cap V) = q - k - r$, and $n_3 = \dim(U^\perp \cap V^\perp) = n - p - q + k$.

A notable application of principal angles is canonical correlations of matrix pairs given in [21], and in many other areas, namely eigenspaces, functional analysis, matrix perturbation theory, statistics, etc., are found in [34,45,16,14], respectively. The spectral problem of a sum of two PSD matrices is closely related to the first aforementioned application of canonical correlations, and its dependence upon the Friedrichs angle, discussed next, elucidates geometric aspects of spectral theory.

2.5. The Friedrichs angle

Finally, we attempt to shed light on another essential tool of this paper. Note that if $\dim(U\cap V) = k$, then $\theta_i = 0$, for all $1 \leq i \leq k$, and $\theta_{k+1} > 0$, whenever $k+1 \leq \min(p, q)$. This angle θ_{k+1}, has been studied widely in the context of angles between subspaces of Hilbert spaces and is referred to as the Friedrichs angle, see [16]. The Friedrichs angle between the subspaces M and N of a Hilbert space $(H, \langle\cdot,\cdot\rangle)$ is the angle $a(M, N)$ in $[0, \frac{\pi}{2}]$ whose cosine is given by

$$c(M, N) = \sup\{|\langle x, y\rangle| \,\big|\, x \in M \cap (M \cap N)^{\perp},\ |x| \leq 1, y \in N \cap (M \cap N)^{\perp}, |y| \leq 1\}.$$

A remarkable property of the Friedrichs angle is $c(M, N) < 1$ or $a(M, N) > 0$ if and only if $M + N$ is closed. Thus, the following definition is adapted for finite dimensions.

Definition 2.3 *(Friedrichs angle [16]).* The angle $\theta_F \in \left(0, \frac{\pi}{2}\right]$ between subspaces $U, V \subseteq \mathbb{R}^n$, whose cosine is defined by

$$\cos\theta_F := \sup\left\{|u^T v| \,\big|\, u \in U \cap (U \cap V)^{\perp}, |u| \leq 1,\ v \in V \cap (U \cap V)^{\perp}, |v| \leq 1\right\},$$

is called the Friedrichs angle.

We summarize some significant properties of the Friedrichs angle, which will be used later to prove some results.

Proposition 2.4. *Let $U, V \subseteq \mathbb{R}^n$ be subspaces, as defined in Theorem 2.2. Let $\mathcal{P}, \mathcal{Q}$ be orthogonal projections onto U and V, respectively, and let θ_F denote the Friedrichs angle between subspaces U and V, then the following results hold.*

1. $\theta_F = \theta_1\left(U \cap (U \cap V)^{\perp}, V \cap (U \cap V)^{\perp}\right)$.
2. $\theta_F = \theta_1(U, V)$ *if and only if* $U \cap V = \{0\}$.
3. $\theta_F = \theta_1\left(U, V \cap (U \cap V)^{\perp}\right) = \theta_1\left(U \cap (U \cap V)^{\perp}, V\right)$.
4. $\cos\theta_F = \|\mathcal{P}\mathcal{Q} - \mathcal{P}_{U\cap V}\|$.
5. $\theta_F = \theta_{k+1}(U, V)$, *if* θ_{k+1} *exists.*

See [16, p. 110] for the first four properties of Proposition 2.4. A proof of Property 4 is also provided in [12, p. 1430], which along with $\|\mathcal{P}\mathcal{Q} - \mathcal{P}_{U\cap V}\| = \cos\theta_{k+1}$, proved in [19, p. 245], implies Property 5. Several more interesting results on the Friedrichs angle are stated in [18, p. 242].

3. Main results

As discussed in section 1, a positive lower bound on the minimum eigenvalue of a non-singular sum of two PSD matrices, say $P, Q \in \mathbb{R}^{n\times n}$, is the key tool for the development of a positive lower bound on the minimum singular value of some full rank block matrices. Note that $\mathcal{N}(P) \cap \mathcal{N}(Q) = \{0\}$ when $P+Q$ is SPD, however, the range spaces of P and Q may intersect. Let $k = \dim(\mathcal{R}(P) \cap \mathcal{R}(Q))$, then the first k principal angles between $\mathcal{R}(P)$ and $\mathcal{R}(Q)$ vanish: $\theta_i = 0$ for $i = 1, 2, \ldots, k$. Therefore, if θ_{k+1} exists then it could contribute in estimating the minimum eigenvalue of $P + Q$ in terms of the minimum positive eigenvalues of P and Q. Even when θ_{k+1} does not exist, this idea serves as a motivation for the following theorem for a pair of two PSD matrices with a non-singular sum.

Theorem 3.1. *Let $P, Q \in \mathbb{R}^{n\times n}$ be PSD matrices of rank $p, q \le n$, respectively, so that $P + Q$ is non-singular. Then*

$$\lambda_{\min}(P+Q) \ge c(P,Q) \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\},$$

where $c(P,Q)$ is defined by

$$c(P,Q) = \begin{cases} 2, & r = 0,\ p+q = 2n, \\ 1, & r = 0,\ p+q < 2n, \\ 1-\cos(\theta_{k+1}), & r > 0, \end{cases} \tag{3.1}$$

where $k = p + q - n$, $0 \le \theta_1 \le \theta_2 \le \ldots \le \theta_{\min(p,q)} \le \frac{\pi}{2}$ represent the principal angles between $\mathcal{R}(P), \mathcal{R}(Q) \subseteq \mathbb{R}^n$, and r is the number of angles θ_i so that $0 < \theta_i < \frac{\pi}{2}$, for $1 \le i \le \min(p,q)$.

Proof. Since P, Q are PSD matrices, there exist matrices $\mathcal{A}_1 \in \mathbb{R}^{p\times n}$ and $\mathcal{A}_2 \in \mathbb{R}^{q\times n}$, so that $P = \mathcal{A}_1^T\mathcal{A}_1$, $Q = \mathcal{A}_2^T\mathcal{A}_2$. Moreover, $\mathcal{N}(P) = \mathcal{N}(\mathcal{A}_1)$ and $\mathcal{N}(Q) = \mathcal{N}(\mathcal{A}_2)$. Define $M_1 := \mathcal{R}(P) = \mathcal{R}(P^T) = \mathcal{N}(P)^\perp = \mathcal{N}(\mathcal{A}_1)^\perp$ and similarly define $M_2 := \mathcal{R}(Q) = \mathcal{N}(\mathcal{A}_2)^\perp$. Let $P_i \in \mathbb{R}^{n\times n}$ be the orthogonal projection on M_i, for $i = 1, 2$. Therefore, $\mathcal{R}(P_i) = M_i$ and $\mathcal{R}(I - P_i) = M_i^\perp = \mathcal{N}(\mathcal{A}_i)$ for $i = 1, 2$. The variational characterization of the smallest eigenvalue of a symmetric matrix implies

$$\lambda_{\min}(P+Q) = \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{x^T\mathcal{A}_1^T\mathcal{A}_1x + x^T\mathcal{A}_2^T\mathcal{A}_2x}{|x|^2}.$$

Since any $x \in \mathbb{R}^n \setminus \{0\}$ can be represented as $x = (I - P_i)x + P_ix$, for $i = 1, 2$, thus

$$\begin{aligned} x^T\mathcal{A}_i^T\mathcal{A}_ix &= [\mathcal{A}_i((I-P_i)x + P_ix)]^T [\mathcal{A}_i((I-P_i)x + P_ix)] \\ &= (\mathcal{A}_iP_ix)^T (\mathcal{A}_iP_ix) \qquad \text{(as } (I-P_i)x \in \mathcal{N}(\mathcal{A}_i)\text{)} \\ &= (P_ix)^T\mathcal{A}_i^T\mathcal{A}_i(P_ix), \end{aligned}$$

therefore,

$$\lambda_{\min}(P+Q) = \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{(P_1x)^T\mathcal{A}_1^T\mathcal{A}_1(P_1x) + (P_2x)^T\mathcal{A}_2^T\mathcal{A}_2(P_2x)}{|x|^2}. \tag{3.2}$$

Note that the minimum positive eigenvalue of $\mathcal{A}_i^T\mathcal{A}_i$ is identified by the variational characterization as follows,

$$\lambda_{\min}(\mathcal{A}_i^T\mathcal{A}_i) = \inf_{x\in\mathcal{N}(\mathcal{A}_i)^\perp} \frac{x^T\mathcal{A}_i^T\mathcal{A}_ix}{|x|^2}.$$

Since for any $x \in \mathbb{R}^n\setminus\{0\}$, $P_ix \in M_i = \mathcal{N}(\mathcal{A}_i)^\perp$, therefore the above expression gives

$$(P_ix)^T\mathcal{A}_i^T\mathcal{A}_i(P_ix) \geq \lambda_{\min}(\mathcal{A}_i^T\mathcal{A}_i)|P_ix|^2,$$

hence (3.2) provides the following estimate

$$\begin{aligned}
\lambda_{\min}(P+Q) &\geq \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\lambda_{\min}(\mathcal{A}_1^T\mathcal{A}_1)|P_1x|^2 + \lambda_{\min}(\mathcal{A}_2^T\mathcal{A}_2)|P_2x|^2}{|x|^2} \\
&= \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\lambda_{\min}(P)|P_1x|^2 + \lambda_{\min}(Q)|P_2x|^2}{|x|^2} \qquad (3.3)\\
&\geq \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\} \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{|P_1x|^2 + |P_2x|^2}{|x|^2} \\
&=: \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\} \inf_{x\in\mathbb{R}^n\setminus\{0\}} \Delta(x), \qquad (3.4)
\end{aligned}$$

where $\Delta(x) := \dfrac{|P_1x|^2 + |P_2x|^2}{|x|^2}$, for $x \in \mathbb{R}^n \setminus \{0\}$. Since the parallelogram identity for inner-product spaces states that

$$|P_1x|^2 + |P_2x|^2 = \frac{1}{2}\left[|(P_1+P_2)x|^2 + |(P_1-P_2)x|^2\right], \tag{3.5}$$

leading to a lower bound,

$$\inf_{x\in\mathbb{R}^n\setminus\{0\}} \Delta(x) \geq \frac{1}{2}\left[\sigma_{\min}^2(P_1+P_2) + \sigma_{\min}^2(P_1-P_2)\right], \tag{3.6}$$

hence the set of singular values of $P_1 \pm P_2$ need to be analyzed. To this end, note that $M_1^\perp \cap M_2^\perp = \mathcal{N}(P)\cap\mathcal{N}(Q) = \{0\}$, as

$$x \in \mathcal{N}(P)\cap\mathcal{N}(Q) \quad \Leftrightarrow \quad (P+Q)x = 0 \quad \Leftrightarrow \quad x = 0,$$

since $P+Q$ is non-singular. Also, (2.3) gives

$$M_1 + M_2 = (M_1^\perp \cap M_2^\perp)^\perp = \{0\}^\perp = \mathbb{R}^n,$$

consequently, Theorem 2.2 gives the set of singular values of $P_1 \pm P_2$ as

$$\begin{aligned} \sigma(P_1+P_2) &= \{\mathbf{2}_k, 1 \pm \cos(\theta_{k+i})(i=1,\ldots,r), \mathbf{1}_{n_1+n_2}\}, \\ \sigma(P_1-P_2) &= \{\mathbf{1}_{n_1+n_2}, \boldsymbol{\sin(\theta_{k+i})}_2(i=1,\ldots,r), \mathbf{0}_k\}, \end{aligned} \tag{3.7}$$

where,

$$\begin{aligned} k &= \dim(M_1 \cap M_2) = \dim(M_1) + \dim(M_2) - \dim(M_1+M_2) = p+q-n, \\ n_1 &= \dim(M_1 \cap M_2^{\perp}) = p-k-r, \\ n_2 &= \dim(M_1^{\perp} \cap M_2) = q-k-r, \\ n_3 &= \dim(M_1^{\perp} \cap M_2^{\perp}) = 0, \\ n &= n_1+n_2+k+2r = p+q-k. \end{aligned} \tag{3.8}$$

Let us estimate (3.6), thus (3.4) in terms of the following cases.

Case 1: Suppose $r=0$. Note that (3.8) implies $n_1 = p-k$, $n_2 = q-k$ for this case. Firstly, consider $k=0$ and $n_1=n_2=0$, then (3.8) implies $p=q=0$, thus $P=Q=O$. Since $P+Q$ is required to be non-singular, thus this case is rejected.

For $k=0$ and $n_1+n_2>0$, (3.7) and (3.8) yield $\sigma(P_1 \pm P_2) = \{\mathbf{1}_{n_1+n_2}\} = \{\mathbf{1}_n\}$, hence (3.6) gives

$$\inf_{x \in \mathbb{R}^n \setminus \{0\}} \Delta(x) \geq \frac{1}{2}\left[1^2+1^2\right] = 1. \tag{3.9}$$

For $k>0$ and $n_1=n_2=0$, (3.8) implies that $k=p=q=n$ or both P and Q are non-singular. Therefore, $M_1=M_2=\mathbb{R}^n$, hence (3.4) gives

$$\inf_{x \in \mathbb{R}^n \setminus \{0\}} \Delta(x) = \inf_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|x|^2+|x|^2}{|x|^2} = 2. \tag{3.10}$$

For $k, n_1 > 0$, and $n_2 = 0$, (3.8) results in $p=n$ and $q=k$, that is, $M_2 \subseteq M_1 = \mathbb{R}^n$ or Q is non-singular. Hence, (3.4) becomes

$$\inf_{x \in \mathbb{R}^n \setminus \{0\}} \Delta(x) = \inf_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|x|^2+|P_2x|^2}{|x|^2} \geq 1.$$

Similarly, for $k, n_2 > 0$ and $n_1 = 0$, the same lower bound as above is derived which also coincides with (3.9). Finally, consider $k, n_1, n_2 > 0$, since $\dim(M_1 \cap M_2) = k > 0$, therefore $M_1 \cap M_2$ is a non-trivial subspace. Define $M_3 := M_2 \cap (M_1 \cap M_2)^{\perp}$, then

$$\begin{aligned} \dim(M_3) &= n - \dim(M_3^{\perp}) \\ &= n - \dim(M_2^{\perp} + M_1 \cap M_2) && \text{(by (2.3))} \\ &= n - (n - \dim(M_2)) - \dim(M_1 \cap M_2) \end{aligned}$$

$$= n - (n-q) - k = q - k,$$

or $\dim(M_3) = q - k = n_2 > 0$, thus it is a non-trivial subspace. Let P_3 and P_U be the orthogonal projections onto the subspace M_3 and some subspace U of $\mathbb{R}^n$, respectively. The following was proved in [12, p. 1429],

$$\begin{aligned}P_3 &= P_{M_2 \cap (M_1 \cap M_2)^\perp} = P_{M_2} P_{(M_1 \cap M_2)^\perp} \\ &= P_{M_2}\left(I - P_{M_1 \cap M_2}\right) \\ &= P_{M_2} - P_{M_2} P_{M_1 \cap M_2} \\ &= P_2 - P_{M_1 \cap M_2},\end{aligned}$$

or $P_2 = P_{M_1 \cap M_2} + P_3$, which implies for any $x \in \mathbb{R}^n \setminus \{0\}$,

$$P_2 x = P_{M_1 \cap M_2} x + P_3 x. \tag{3.11}$$

By definition, $M_1 \cap M_2$ and M_3 are mutually orthogonal subspaces. Therefore, for any $x \in \mathbb{R}^n \setminus \{0\}$,

$$|P_2 x|^2 = |P_{M_1 \cap M_2} x|^2 + |P_3 x|^2, \tag{3.12}$$

and (3.4) becomes

$$\begin{aligned}\inf_{x \in \mathbb{R}^n \setminus \{0\}} \Delta(x) &= \inf_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|P_1 x|^2 + |P_{M_1 \cap M_2} x|^2 + |P_3 x|^2}{|x|^2} \\ &\geq \inf_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|P_1 x|^2 + |P_3 x|^2}{|x|^2} \\ &= \frac{1}{2} \inf_{x \in \mathbb{R}^n \setminus \{0\}} \frac{|(P_1 + P_3)x|^2 + |(P_1 - P_3)x|^2}{|x|^2} \qquad \text{(by (3.5))} \\ &\geq \frac{1}{2}\left[\sigma^2_{\min}(P_1 + P_3) + \sigma^2_{\min}(P_1 - P_3)\right]. \end{aligned} \tag{3.13}$$

By Theorem 2.2,

$$\begin{aligned}\sigma(P_1 + P_3) &= \{\mathbf{2}_{\tilde{k}}, 1 \pm \cos(\alpha_{\tilde{k}+i})(i = 1, \ldots, r), \mathbf{1}_{\tilde{n}_1 + \tilde{n}_2}, 0_{\tilde{n}_3}\}, \\ \sigma(P_1 - P_3) &= \{\mathbf{1}_{\tilde{n}_1 + \tilde{n}_2}, \mathbf{\sin(\boldsymbol{\alpha}_{\tilde{k}+i})}_2 (i = 1, \ldots, r), \mathbf{0}_{\tilde{k} + \tilde{n}_3}\},\end{aligned} \tag{3.14}$$

where $0 \leq \alpha_1 \leq \alpha_2 \leq \ldots \leq \alpha_{\min(p, q-k)} \leq \frac{\pi}{2}$ represent the principal angles between the subspaces M_1 and M_3, and r is the number of principal angles that satisfy $0 < \alpha_i < \frac{\pi}{2}$. Observe that r is the same number for M_1 and M_2 by Definition 2.1, or see [19, p. 231] for more details. Thus, $r = 0$ and (3.8) gives the following parameters

$$\begin{aligned} \tilde{k} &= \dim(M_1 \cap M_3) = \dim(M_1 \cap (M_2 \cap (M_1 \cap M_2)^{\perp})) = 0, \\ \tilde{n}_1 &= p - \tilde{k} - r = p, \\ \tilde{n}_2 &= (q-k) - \tilde{k} - r = q - k, \\ \tilde{n_3} &= n - p - (q-k) + \tilde{k} = n - (p+q-k) = 0. \end{aligned} \tag{3.15}$$

Hence, $\sigma_{\min}(P_1 \pm P_3) = 1$, therefore (3.13) implies that $\inf_{x\in\mathbb{R}^n\setminus\{0\}} \Delta(x) \geq 1$, which coincides with (3.9).

In conclusion, for $r = 0$ and $p + q < 2n$, $\lambda_{\min}(P + Q) \geq \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$ holds, or $c(P, Q) = 1$. Whereas, for $r = 0$ and $p + q = 2n$, $\lambda_{\min}(P + Q) \geq 2\min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$ holds, or $c(P, Q) = 2$.

Case 2: Suppose $r > 0$. For $k = 0$ and any $n_1, n_2 \geq 0$, (3.7) implies that $\sigma_{\min}(P_1+P_2) = 1 - \cos\theta_1$ and $\sigma_{\min}(P_1 - P_2) = \sin\theta_1$. By (3.6),

$$\begin{aligned} \inf_{x\in\mathbb{R}^n\setminus\{0\}} \Delta(x) &\geq \frac{1}{2}\left[(1-\cos\theta_1)^2 + \sin^2\theta_1\right] \\ &= 1 - \cos\theta_1. \end{aligned}$$

Therefore, (3.4) gives $\lambda_{\min}(P + Q) \geq (1 - \cos\theta_1)\min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$.

For $k > 0$ and $n_1, n_2 \geq 0$, thus $M_1 \cap M_2$ is non-trivial. Consider the subspaces M_1 and M_3, as defined earlier for $k, n_1, n_2 > 0$ in Case 1. Note that M_3 is always non-trivial, as $n_1, n_2 \geq 0$ (3.8) implies $p \geq k + r$ and $q \geq k + r$, thus $\dim M_3 = q - k \geq r > 0$. Also, the set of parameters for M_1 and M_3 are given by (3.15) and $r > 0$ as follows,

$$\begin{aligned} \tilde{k} &= \dim(M_1 \cap M_3) = 0, \\ \tilde{n}_1 &= p - \tilde{k} - r = p - r, \\ \tilde{n}_2 &= (q-k) - \tilde{k} - r = q - k - r, \\ \tilde{n}_3 &= n - p - (q-k) + \tilde{k} = 0. \end{aligned} \tag{3.16}$$

Let θ_F be the Friedrichs angle between M_1 and M_2, then by Property 3 of Proposition 2.4, it is equal to the minimal angle between M_1 and M_3 in (3.14), that is,

$$\alpha_1 = \theta_F = \theta_{k+1}, \tag{3.17}$$

where the last equality follows by Property 5 of Proposition 2.4. Therefore, (3.14) implies that $\sigma_{\min}(P_1 + P_3) = 1 - \cos\alpha_1 = 1 - \cos\theta_{k+1}$ and $\sigma_{\min}(P_1 - P_3) = \sin\alpha_1 = \sin\theta_{k+1}$, so (3.13) yields

$$\begin{aligned} \inf_{x\in\mathbb{R}^n\setminus\{0\}} \Delta(x) &\geq \frac{1}{2}\left[(1-\cos\theta_{k+1})^2 + \sin^2\theta_{k+1}\right] \\ &= 1 - \cos\theta_{k+1}. \end{aligned} \tag{3.18}$$

Thus, (3.4) implies $\lambda_{\min}(P+Q) \geq (1-\cos\theta_{k+1}) \min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$, or $c(P,Q) = 1-\cos\theta_{k+1}$, which is consistent for $r > 0$ and $k = 0$. □

Remark 3.2. Recall that θ_{k+1} is the Friedrichs angle between $\mathcal{R}(P)$ and $\mathcal{R}(Q)$ as stated in Property 5 of Proposition 2.4. It can be easily calculated by using a result by A. Björck and G. Golub given in [11]. Let $Q_1 \in \mathbb{R}^{n\times p}$ and $Q_2 \in \mathbb{R}^{n\times q}$ represent orthogonal bases for $\mathcal{R}(P)$ and $\mathcal{R}(Q)$, respectively. Define $\mathcal{M} = Q_1^T Q_2$, then $\cos\theta_i = \sigma_i(\mathcal{M})$, where $i = 1, 2, \ldots, \min(p, q)$.

Remark 3.3. When $r = 0$ and $p+q = 2n$, which occurs when both P and Q are SPD, the term $c(P,Q) = 2$ is convenient for future use. It can be easily strengthened by using (2.1). The significance of Theorem 3.1 is that it gives a positive lower bound on $\lambda_{\min}(P+Q)$ for the case when both P and Q are rank deficient, for which the estimate of the theorem reads $\lambda_{\min}(P+Q) \geq (1-\cos\theta_F)\min\{\lambda_{\min}(P), \lambda_{\min}(Q)\}$. Also, (3.18) is an optimal lower bound for $k = 0$, however, there is a scope of improvement when $k, r > 0$. This bound results from (3.12), which is carefully selected through analysis described in section 2 of [30], which also summarizes the cases for $r = 0$ in section 1.

Note that (3.1) utilizes the angle between the range spaces of matrices $P, Q \in \mathbb{R}^{n\times n}$, thus it can be extended to rectangular matrices with the same number of rows. The results stated below are relevant for describing the lower bounds on the minimum singular value of certain block matrices with full rank.

Proposition 3.4. *For distinct non-zero matrices $A \in \mathbb{R}^{n\times p}$, $B \in \mathbb{R}^{n\times q}$*

1. $c(A,B) = c(AA^T, BB^T)$.
2. $c(A,B) = c(B,A)$.
3. $c(A, O_{n\times k}) = 1$.
4. $c(A,B) = 1-\cos\theta_F$, *when both A, B are rank-deficient, where θ_F is the Friedrichs angle between $\mathcal{R}(A)$ and $\mathcal{R}(B)$.*

The following result is derived to complete the analysis of PSD matrices P, Q. It reduces to (2.1) when at least one of P and Q is SPD. However, the result may be weaker or stronger than the estimate given by Theorem 3.1.

Theorem 3.5. *Let $P, Q \in \mathbb{R}^{n\times n}$ be PSD matrices of rank $p, q \leq n$, respectively, so that $P+Q$ is non-singular, then*

$$\lambda_{\min}(P+Q) \geq \psi(P,Q),$$

where, for $r = 0$

$$\psi(P,Q) := \begin{cases} a^2, & p = n, q < n, \\ b^2, & p < n, q = n, \\ a^2 + b^2, & p = q = n, \\ \min\left\{a^2, b^2\right\}, & \textit{otherwise}, \end{cases}$$

and for $r > 0$,

$$\psi(P,Q) := \frac{1}{2}\left[a^2 + b^2 - \frac{1}{2}(a+b)\sqrt{(a+b)^2 - 4ab\sin^2\theta_{k+1}}\right.$$
$$\left. - \frac{1}{2}|a-b|\sqrt{(a-b)^2 + 4ab\sin^2\theta_{k+1}}\right],$$

where $a = \sqrt{\lambda_{\min}(P)}$, $b = \sqrt{\lambda_{\min}(Q)}$, $k = p + q - n$, $0 \le \theta_1 \le \theta_2 \le \ldots \le \theta_{\min(p,q)} \le \frac{\pi}{2}$ *represent the principal angles between* $\mathcal{R}(P), \mathcal{R}(Q) \subseteq \mathbb{R}^n$, *and* r *is the number of principal angles* θ_i *so that* $0 < \theta_i < \frac{\pi}{2}$, *for* $1 \le i \le \min(p,q)$.

Proof. Define $M_1 := \mathcal{R}(P)$, $M_2 := \mathcal{R}(Q)$ and $P_i \in \mathbb{R}^{n\times n}$ to be the orthogonal projection onto M_i, for $i = 1, 2$. On following the proof of Theorem 3.1, (3.3) gives

$$\begin{aligned} \lambda_{\min}(P+Q) &\ge \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{\lambda_{\min}(P)|P_1x|^2 + \lambda_{\min}(Q)|P_2x|^2}{|x|^2} \\ &= \inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{a^2|P_1x|^2 + b^2|P_2x|^2}{|x|^2} \qquad (3.19) \\ &= \frac{1}{2}\inf_{x\in\mathbb{R}^n\setminus\{0\}} \frac{|(aP_1+bP_2)x|^2 + |(aP_1-bP_2)x|^2}{|x|^2} \qquad \text{(by (3.5))} \\ &\ge \frac{1}{2}\left[\sigma^2_{\min}(aP_1+bP_2) + \sigma^2_{\min}(aP_1-bP_2)\right]. \qquad (3.20) \end{aligned}$$

The results given in [19, p. 247] and [19, pp. 234–235] give the following expression,

$$aP_1 + bP_2 = Z\text{diag}\left((a+b)I_k, aS + bE(\theta_{k+i})(i = 1,\ldots,r), aI_{n_1}, bI_{n_2}, O_{n_3}\right)Z^T,$$

where Z is an orthogonal matrix, $S = \left[\begin{smallmatrix}1 & 0\\ 0 & 0\end{smallmatrix}\right]$, $E(\theta) = \left[\begin{smallmatrix}\cos^2\theta & \cos\theta\sin\theta\\ \cos\theta\sin\theta & \sin^2\theta\end{smallmatrix}\right]$, and the expressions for k, n_1, n_2, n_3, n are given by (3.8). Thus, the set of singular values of $aP_1 \pm bP_2$ are

$$
\begin{aligned}
\sigma(aP_1+bP_2) &= \left\{\frac{1}{2}\left[(a+b)\pm\sqrt{(a+b)^2-4ab\sin^2\theta_{k+i}}\right](i=1,\dots,r),\right.\\
&\qquad\left.\boldsymbol{(a+b)}_k,\boldsymbol{a}_{n_1},\boldsymbol{b}_{n_2},\boldsymbol{0}_{n_3}\right\},\\
\sigma(aP_1-bP_2) &= \left\{\frac{1}{2}\left[\sqrt{(a-b)^2+4ab\sin^2\theta_{k+i}}\pm|a-b|\right](i=1,\dots,r),\right.\\
&\qquad\left.\boldsymbol{|a-b|}_k,\boldsymbol{a}_{n_1},\boldsymbol{b}_{n_2},\boldsymbol{0}_{n_3}\right\},
\end{aligned}
\tag{3.21}
$$

resulting in the following cases.

Case 1: Suppose $r=0$, then (3.8) yields $n_1=p-k$ and $n_2=q-k$ for this case. Firstly, consider $k=n_1=0$ and $n_2>0$, then (3.8) implies $p=0$ and $q=n$. Thus, $P=O$ and $P+Q=Q$ is non-singular. Therefore, (3.19) leads to $\lambda_{\min}(P+Q)=\lambda_{\min}(Q)=b^2$. Similarly, $Q=O$ for $k=n_2=0$ and $n_1>0$, thus $\lambda_{\min}(P+Q)=\lambda_{\min}(P)=a^2$.

For $k=0$ and $n_1,n_2>0$, (3.21) gives $\sigma_{\min}(aP_1\pm bP_2)=\min\{a,b\}$. Thus, (3.20) gives $\lambda_{\min}(P+Q)\geq\min\{a^2,b^2\}$.

For $k>0$ and $n_1=n_2=0$, (3.8) implies $k=p=q=n$, that is, both P and Q are non-singular. By (3.21), $\sigma_{\min}(aP_1+bP_2)=a+b$, $\sigma_{\min}(aP_1-bP_2)=|a-b|$, therefore (3.20) gives

$$
\lambda_{\min}(P+Q)\geq\frac{1}{2}\left[(a+b)^2+(a-b)^2\right]=a^2+b^2.
$$

For $k,n_2>0$ and $n_1=0$, (3.8) gives $p=k$ and $q=n$, which imply $M_1\subseteq M_2=\mathbb{R}^n$, or Q is non-singular. Thus, (3.19) becomes

$$
\lambda_{\min}(P+Q)=\inf_{x\in\mathbb{R}^n\setminus\{0\}}\frac{a^2|P_1x|^2+b^2|x|^2}{|x|^2}\geq b^2=\lambda_{\min}(Q),
$$

similarly $k,n_1>0$ and $n_2=0$ gives $\lambda_{\min}(P+Q)\geq a^2=\lambda_{\min}(P)$. Finally, consider $k,n_1,n_2>0$, since $\dim(M_1\cap M_2)=k>0$, then $M_1\cap M_2$ is a non-trivial subspace. Consider $M_3=M_2\cap(M_1\cap M_2)^\perp\neq\{0\}$ as defined in the proof of Theorem 3.1, then (3.12) in (3.19) gives

$$
\begin{aligned}
\lambda_{\min}(P+Q) &= \inf_{x\in\mathbb{R}^n\setminus\{0\}}\frac{a^2|P_1x|^2+b^2|P_{M_1\cap M_2}x|^2+b^2|P_3x|^2}{|x|^2}\\
&\geq \inf_{x\in\mathbb{R}^n\setminus\{0\}}\frac{a^2|P_1x|^2+b^2|P_3x|^2}{|x|^2}\\
&= \frac{1}{2}\inf_{x\in\mathbb{R}^n\setminus\{0\}}\frac{|(aP_1+bP_3)x|^2+|(aP_1-bP_3)x|^2}{|x|^2} \qquad \text{(by (3.5))}\\
&\geq \frac{1}{2}\left[\sigma^2_{\min}(aP_1+bP_3)+\sigma^2_{\min}(aP_1-bP_3)\right].
\end{aligned}
\tag{3.22}
$$

The set of singular values of $aP_1 \pm bP_3$ are given by (3.21) as follows

$$\begin{aligned}\sigma(aP_1 + bP_3) &= \left\{ \frac{1}{2}\left[(a+b) \pm \sqrt{(a+b)^2 - 4ab\sin^2\alpha_{\tilde{k}+i}}\right] (i = 1, \dots, r), \right.\\ &\qquad\qquad \left. \boldsymbol{(a+b)}_{\tilde{k}}, \boldsymbol{a}_{\tilde{n}_1}, \boldsymbol{b}_{\tilde{n}_2}, \boldsymbol{0}_{\tilde{n}_3} \right\},\\ \sigma(aP_1 - bP_3) &= \left\{ \frac{1}{2}\left[\sqrt{(a-b)^2 + 4ab\sin^2\alpha_{\tilde{k}+i}} \pm |a-b|\right] (i = 1, \dots, r), \right.\\ &\qquad\qquad \left. \boldsymbol{|a-b|}_{\tilde{k}}, \boldsymbol{a}_{\tilde{n}_1}, \boldsymbol{b}_{\tilde{n}_2}, \boldsymbol{0}_{\tilde{n}_3} \right\},\end{aligned} \tag{3.23}$$

where the parameters are the same as (3.15) with $r = 0$. Thus, $\sigma_{\min}(aP_1 \pm bP_3) = \min\{a, b\}$. By (3.22), $\lambda_{\min}(P + Q) \geq \min\{a^2, b^2\}$.

Case 2: Suppose $r > 0$. Then for $k = 0$ and $n_1, n_2 \geq 0$, by (3.21)

$$\begin{aligned}\sigma_{\min}(aP_1 + bP_2) &= \frac{1}{2}\left[(a+b) - \sqrt{(a+b)^2 - 4ab\sin^2\theta_1}\right],\\ \sigma_{\min}(aP_1 - bP_2) &= \frac{1}{2}\left[\sqrt{(a-b)^2 + 4ab\sin^2\theta_1} - |a-b|\right].\end{aligned}$$

Thus, (3.20) gives

$$\begin{aligned}\lambda_{\min}(P+Q) \geq \frac{1}{2}\Big[a^2 + b^2 &- \frac{1}{2}(a+b)\sqrt{(a+b)^2 - 4ab\sin^2\theta_1}\\ &- \frac{1}{2}|a-b|\sqrt{(a-b)^2 + 4ab\sin^2\theta_1}\Big].\end{aligned}$$

For $k > 0$ and $n_1, n_2 \geq 0$, (3.16), (3.17), and (3.23) yield

$$\begin{aligned}\sigma_{\min}(aP_1 + bP_3) &= \frac{1}{2}\left[(a+b) + \sqrt{(a+b)^2 - 4ab\sin^2\theta_{k+1}}\right],\\ \sigma_{\min}(aP_1 - bP_3) &= \frac{1}{2}\left[\sqrt{(a-b)^2 + 4ab\sin^2\theta_{k+1}} - |a-b|\right],\end{aligned}$$

hence (3.22) implies

$$\begin{aligned}\lambda_{\min}(P+Q) \geq \frac{1}{2}\Big[a^2 + b^2 &- \frac{1}{2}(a+b)\sqrt{(a+b)^2 - 4ab\sin^2\theta_{k+1}}\\ &- \frac{1}{2}|a-b|\sqrt{(a-b)^2 + 4ab\sin^2\theta_{k+1}}\Big]. \qquad \square\end{aligned}$$

In the proofs of Theorems 3.1 and 3.5, a technique similar to the case of $k = 0$ and $n_1, n_2 > 0$ can be applied to the cases $k = n_1 = 0$ and $n_2 > 0$, and $k = n_2 = 0$ and

$n_1 > 0$, to get another positive lower bound; however, they turn out to be weaker than the stated results. On combining Theorems 3.1 and 3.5, another positive lower bound on $\lambda_{\min}(P+Q)$ is given as follows.

Corollary 3.6. *Let $P, Q \in \mathbb{R}^{n\times n}$ be PSD matrices of rank $p, q \le n$, respectively, so that $P+Q$ is non-singular. Then*

$$\lambda_{\min}(P+Q) \ge \max\left[c(P,Q)\min\left\{\lambda_{\min}(P), \lambda_{\min}(Q)\right\}, \psi(P,Q)\right].$$

For convenience of notation, define the function Ψ for matrices $A \in \mathbb{R}^{n\times p}$, $B \in \mathbb{R}^{n\times q}$:

$$\Psi(A,B) = \sqrt{\psi(AA^T, BB^T)}, \tag{3.24}$$

where ψ is defined by Theorem 3.5. Note that $\psi(AA^T, BB^T)$ is a function defined in terms of $a = \sqrt{\lambda_{\min}(AA^T)} = \sigma_{\min}(A)$, $b = \sqrt{\lambda_{\min}(BB^T)} = \sigma_{\min}(B)$, and principal angles between $\mathcal{R}(AA^T) = \mathcal{R}(A)$ and $\mathcal{R}(BB^T) = \mathcal{R}(B)$. Thus, $\Psi(A,B)$ is a function defined in terms of positive singular values of A and B, and principal angles between $\mathcal{R}(A)$ and $\mathcal{R}(B)$.

A positive lower bound, defined by Corollary 3.6, could be useful in several circumstances, such as for a full rank block 2×1 matrix, with rank deficient sub-blocks. Hence, the following applications are presented.

Corollary 3.7. *For $m \ge n$, let $A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \in \mathbb{R}^{m\times n}$ be full rank, then*

$$\sigma_{\min}(A) \ge \max\left[\sqrt{c(A_1^T, A_2^T)}\min\left\{\sigma_{\min}(A_1), \sigma_{\min}(A_2)\right\}, \Psi(A_1^T, A_2^T)\right].$$

Proof. Since A is a full rank matrix, $\sigma^2_{\min}(A) = \lambda_n(A^TA) = \lambda_{\min}(A_1^TA_1 + A_2^TA_2)$, thus Corollary 3.6 implies

$$\sigma^2_{\min}(A) \ge \max\left[c(A_1^TA_1, A_2^TA_2)\min\left\{\sigma^2_{\min}(A_1), \sigma^2_{\min}(A_2)\right\}, \psi\left(A_1^TA_1, A_2^TA_2\right)\right],$$

which gives the desired result after applying Property 1 of Proposition 3.4 and (3.24). □

Corollary 3.8. *For $m \le n$, let $A = [A_1 \quad A_2] \in \mathbb{R}^{m\times n}$ be full rank, then*

$$\sigma_{\min}(A) \ge \max\left[\sqrt{c(A_1, A_2)}\min\left\{\sigma_{\min}(A_1), \sigma_{\min}(A_2)\right\}, \Psi(A_1, A_2)\right].$$

After securing the above lower bounds, our subsequent aim is to extend them to a non-singular 2×2 block matrix. While it could be tedious to estimate the singular values of 2×2 block matrices, it is easier to find the singular values of its blocks which are of smaller size. Thus, another significant application of Theorems 3.1 and 3.5 is the following result, which give four estimates on the minimum singular value of a non-singular matrix.

Theorem 3.9. *For a non-singular matrix*

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \in \mathbb{R}^{n\times n},$$

where $A_{11} \in \mathbb{R}^{p\times k}$*,* $A_{22} \in \mathbb{R}^{q\times \ell}$*, for* $1 \le p, q, k, \ell \le n$*, the following hold*

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\left\{\sigma_{\min}\left([A_{11}, A_{12}]\right), \sigma_{\min}\left([A_{21}, A_{22}]\right)\right\}, \tag{3.25a}$$

$$\sigma_{\min}(A) \ge \Psi\left([A_{11}, A_{12}]^T, [A_{21}, A_{22}]^T\right), \tag{3.25b}$$

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\{r_1, r_2\}, \tag{3.25c}$$

where

$$r_1 := \max\left[c_1 \min\{\sigma_{\min}(A_{11}), \sigma_{\min}(A_{12})\}, \Psi(A_{11}, A_{12})\right],$$
$$r_2 := \max\left[c_2 \min\{\sigma_{\min}(A_{21}), \sigma_{\min}(A_{22})\}, \Psi(A_{21}, A_{22})\right],$$

where $c_1 = \sqrt{c(A_{11}, A_{12})}$*,* $c_2 = \sqrt{c(A_{21}, A_{22})}$*, and* $\theta \in \left(0, \frac{\pi}{2}\right]$ *is the minimum principal angle between* $\mathcal{R}\left([A_{11}, A_{12}]^T\right), \mathcal{R}\left([A_{21}, A_{22}]^T\right) \subseteq \mathbb{R}^n$*. Moreover,*

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\{c_1, c_2\}\cdot \min_{1\le i,j\le 2}\{\sigma_{\min}(A_{ij})\}. \tag{3.26}$$

Proof. Since $A_{11} \in \mathbb{R}^{p\times k}$, $A_{22} \in \mathbb{R}^{q\times \ell}$, then $p + q = n = k + \ell$. Let $R_1 = [A_{11} \quad A_{12}] \in \mathbb{R}^{p\times n}$, and $R_2 = [A_{21} \quad A_{22}] \in \mathbb{R}^{q\times n}$, then by a direct calculation

$$A^T A = \begin{bmatrix} R_1^T & R_2^T \end{bmatrix} \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} = R_1^T R_1 + R_2^T R_2. \tag{3.27}$$

Note that $\operatorname{rank}(R_1) \le p$ and $\operatorname{rank}(R_2) \le q$, thus $\operatorname{rank}(R_1) + \operatorname{rank}(R_2) \le p + q = n$, also

$$\begin{aligned} \operatorname{rank}(R_1) + \operatorname{rank}(R_2) &= \operatorname{rank}(R_1^T R_1) + \operatorname{rank}(R_2^T R_2) \\ &\ge \operatorname{rank}(R_1^T R_1 + R_2^T R_2) \\ &= \operatorname{rank}(A^T A) = \operatorname{rank}(A) = n. \end{aligned}$$

Therefore, $\operatorname{rank}(R_1)+\operatorname{rank}(R_2) = n$, which implies that $\operatorname{rank}(R_1) = p$ and $\operatorname{rank}(R_2) = q$, that is, R_1 and R_2 are full rank matrices. And, $A^T A$ is an SPD matrix expressed a sum of two singular PSD matrices, thus by Theorem 3.1

$$\begin{aligned} \sigma^2_{\min}(A) &= \lambda_{\min}(A^T A) \\ &\ge c(R_1^T R_1, R_2^T R_2) \min\left\{\lambda_{\min}(R_1^T R_1), \lambda_{\min}(R_2^T R_2)\right\} \\ &= c(R_1^T, R_2^T) \min\left\{\sigma^2_{\min}(R_1), \sigma^2_{\min}(R_2)\right\}, \end{aligned} \tag{3.28}$$

where the last equality results from Property 1 of Proposition 3.4. Let θ be the minimum principal angle between $\mathcal{R}(R_1^T)$ and $\mathcal{R}(R_2^T)$. Since A is non-singular, (3.27) gives $\mathcal{N}(R_1)\cap\mathcal{N}(R_2)=\{0\}$, thus (2.3) implies $\mathcal{R}(R_1^T)+\mathcal{R}(R_2^T)=\mathcal{N}(R_1)^\perp+\mathcal{N}(R_2)^\perp=(\mathcal{N}(R_1)\cap\mathcal{N}(R_2))^\perp=\{0\}^\perp=\mathbb{R}^n$. Therefore,

$$\begin{aligned}\dim(\mathcal{R}(R_1^T)\cap\mathcal{R}(R_2^T)) &= \dim(\mathcal{R}(R_1^T))+\dim(\mathcal{R}(R_2^T))-\dim(\mathcal{R}(R_1^T)+\mathcal{R}(R_2^T))\\ &= \operatorname{rank}(R_1)+\operatorname{rank}(R_2)-\dim(\mathbb{R}^n)\\ &= p+q-n=0,\end{aligned}$$

or, $\mathcal{R}(R_1^T)\cap\mathcal{R}(R_2^T)=\{0\}$. Hence, $\mathbb{R}^n=\mathcal{R}(R_1^T)\oplus\mathcal{R}(R_2^T)$, that is, $\mathcal{R}(R_1^T)$ and $\mathcal{R}(R_2^T)$ are complementary subspaces so that $0<\theta\le\frac{\pi}{2}$, which implies if $r=0$ then $\theta=\frac{\pi}{2}$. Therefore, (3.1) can simply be expressed as $c(R_1^T,R_2^T)=1-\cos\theta$, thus (3.28) yields

$$\sigma_{\min}^2(A)\ge(1-\cos\theta)\min\left\{\sigma_{\min}^2(R_1),\sigma_{\min}^2(R_2)\right\},$$

which leads to (3.25a). Also, applying Theorem 3.5 to (3.27) implies

$$\sigma_{\min}^2(A)=\lambda_{\min}\left(R_1^TR_1+R_2^TR_2\right)\ge\psi\left(R_1^TR_1,R_2^TR_2\right),$$

which by (3.24) gives (3.25b). Since R_1 has full rank, Corollary 3.8 for estimating $\sigma_{\min}(R_1)$ leads to

$$\sigma_{\min}(R_1)\ge\max\left[c_1\min\left\{\sigma_{\min}(A_{11}),\sigma_{\min}(A_{12})\right\},\Psi\left(A_{11},A_{12}\right)\right]=:r_1,$$

where $c_1=\sqrt{c(A_{11},A_{12})}$. Similarly, a lower bound on $\sigma_{\min}(R_2)$ is

$$\sigma_{\min}(R_2)\ge\max\left[c_2\min\left\{\sigma_{\min}(A_{21}),\sigma_{\min}(A_{22})\right\},\Psi\left(A_{21},A_{22}\right)\right]=:r_2,$$

where $c_2=\sqrt{c(A_{21},A_{22})}$, and hence (3.25a) is expressed as

$$\sigma_{\min}(A)\ge\sqrt{1-\cos\theta}\cdot\min\left\{r_1,r_2\right\},$$

which on further simplification gives,

$$\begin{aligned}\sigma_{\min}(A)&\ge\sqrt{1-\cos\theta}\cdot\min\left[c_1\min\left\{\sigma_{\min}(A_{11}),\sigma_{\min}(A_{12})\right\},\right.\\ &\qquad\qquad\left.c_2\min\left\{\sigma_{\min}(A_{21}),\sigma_{\min}(A_{22})\right\}\right]\\ &\ge\sqrt{1-\cos\theta}\cdot\min\left\{c_1,c_2\right\}\cdot\min_{1\le i,j\le 2}\left\{\sigma_{\min}(A_{ij})\right\}.\quad\square\end{aligned}$$

As $\sigma_{\min}(A)=\sigma_{\min}(A^T)$, the above estimates yield the following result framed in terms of block columns of A.

Corollary 3.10. *For a non-singular matrix*

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \in \mathbb{R}^{n\times n},$$

where $A_{11} \in \mathbb{R}^{p\times k}$, $A_{22} \in \mathbb{R}^{q\times \ell}$, *for* $1 \le p, q, k, \ell \le n$,

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\left\{\sigma_{\min}\left(\left[\begin{smallmatrix} A_{11} \\ A_{21}\end{smallmatrix}\right]\right), \sigma_{\min}\left(\left[\begin{smallmatrix} A_{12} \\ A_{22}\end{smallmatrix}\right]\right)\right\}, \tag{3.29a}$$

$$\sigma_{\min}(A) \ge \Psi\left(\left[\begin{smallmatrix} A_{11} \\ A_{21}\end{smallmatrix}\right], \left[\begin{smallmatrix} A_{12} \\ A_{22}\end{smallmatrix}\right]\right), \tag{3.29b}$$

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\{s_1, s_2\}, \tag{3.29c}$$

where

$$s_1 := \max\left[c_1 \min\{\sigma_{\min}(A_{11}), \sigma_{\min}(A_{21})\}, \Psi\left(A_{11}^T, A_{21}^T\right)\right],$$
$$s_2 := \max\left[c_2 \min\{\sigma_{\min}(A_{12}), \sigma_{\min}(A_{22})\}, \Psi\left(A_{12}^T, A_{22}^T\right)\right],$$

where $c_1 = \sqrt{c(A_{11}^T, A_{21}^T)}$, $c_2 = \sqrt{c(A_{12}^T, A_{22}^T)}$, *and* $\theta \in \left(0, \frac{\pi}{2}\right]$ *is the minimum principal angle between* $\mathcal{R}\left(\left[\begin{smallmatrix} A_{11} \\ A_{21}\end{smallmatrix}\right]\right), \mathcal{R}\left(\left[\begin{smallmatrix} A_{12} \\ A_{22}\end{smallmatrix}\right]\right) \subseteq \mathbb{R}^n$. *Moreover,*

$$\sigma_{\min}(A) \ge \sqrt{1-\cos\theta}\cdot \min\{c_1, c_2\}\cdot \min_{1\le i,j\le 2}\{\sigma_{\min}(A_{ij})\}. \tag{3.30}$$

Remark 3.11. The estimate given by (3.25a) is stronger than (3.25c), which is greater than (3.26), however, the sharpness of (3.25b) varies for different matrices (see Example 5). The inequality (3.26) gives a lower bound on the minimum singular value of a non-singular 2×2 block matrix in terms of the minimum positive singular value of its blocks. The estimates from Theorem 3.9 and Corollary 3.10 may differ, so in practice, one may use the maximum of all of the bounds obtained from both of them. A MATLAB® implementation for any 2×2 block matrix is given in [30].

Now, we simplify Theorem 3.9 for the special case of a saddle point matrix as follows.

Corollary 3.12. *For a non-singular saddle point matrix*

$$\mathcal{X} = \begin{bmatrix} A & B \\ B^T & O \end{bmatrix} \in \mathbb{R}^{(m+n)\times(m+n)},$$

where $A \in \mathbb{R}^{m\times m}$ *is non-singular and* $B \in \mathbb{R}^{m\times n}$ *is full rank,*

$$\sigma_{\min}(\mathcal{X}) \ge \sqrt{1-\cos\theta}\cdot \min\{\sigma_{\min}(A), \sigma_{\min}(B)\},$$

where θ *is the minimum principal angle between* $\mathcal{R}\left([A, B]^T\right)$ *and* $\mathcal{R}\left([B^T, O]^T\right)$.

Proof. Since $\mathcal{X}$ is non-singular, according to (3.25a),

$$\sigma_{\min}(\mathcal{X}) \geq \sqrt{1-\cos\theta} \cdot \min\left\{\sigma_{\min}([A,B]), \sigma_{\min}([B^T,O])\right\}.$$

Since $\sigma^2_{\min}([A,B]) = \lambda_{\min}([A,B][A,B]^T) = \lambda_{\min}(AA^T + BB^T) \geq \sigma^2_{\min}(A)$, and similarly $\sigma^2_{\min}([B^T,O]) = \sigma^2_{\min}(B)$, hence the desired result. □

In [31], we apply Corollary 3.12 to the global space-time spectral operator for the Stokes problem in an unsteady state, which is a saddle point matrix of the form $\mathcal{X}$ with a non-symmetric leading block A. Also, Corollary 3.10 can be applied to $\mathcal{X}$ to get an analogous result to Corollary 3.12.

After discussing lower bounds on the minimum singular value of a non-singular 2×2 block matrix, we divert our attention to constructing a lower bound on some other singular values such as the following result.

Theorem 3.13 *([23]). Let $A_{ij} \in \mathbb{R}^{m\times n}$ for $i,j = 1,2$, and*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},$$

then

$$\sigma_j(A) \geq \sqrt{2\sigma_j\left(A_{11}A_{12}^T + A_{21}A_{22}^T\right)}, \qquad j = 1,2,\ldots,\min(m,n).$$

We provide simpler proof for two similar lower bounds with more general sizes for their sub-blocks.

Theorem 3.14. *For a block matrix*

$$A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right],$$

where $A_{11} \in \mathbb{R}^{m\times \ell}$ and $A_{22} \in \mathbb{R}^{n\times p}$, then

$$\sigma_j(A) \geq \sqrt{2\sigma_j\left(A_{11}A_{21}^T + A_{12}A_{22}^T\right)}, \qquad j = 1,2,\ldots,\min(m,n,\ell+p),$$

also,

$$\sigma_j(A) \geq \sqrt{2\sigma_j\left(A_{11}^TA_{12} + A_{21}^TA_{22}\right)}, \qquad j = 1,2,\ldots,\min(m+n,\ell,p).$$

Proof. Let $R_1 = [A_{11} \quad A_{12}]$ and $R_2 = [A_{21} \quad A_{22}]$, so that $A = \left[\begin{smallmatrix} R_1 \\ R_2 \end{smallmatrix}\right]$. For $j = 1,2,\ldots,\min(m+n,\ell+p)$, $\sigma_j^2(A) = \lambda_j(A^TA) = \lambda_j(R_1^TR_1 + R_2^TR_2)$. Therefore, by (2.2), for $j = 1,2,\ldots,\min(m,n,\ell+p)$

$$\sigma_j^2(A) \geq 2\sigma_j(R_1 R_2^T) = 2\sigma_j\left(A_{11}A_{21}^T + A_{12}A_{22}^T\right).$$

Also, let $C_1 = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$ and $C_2 = \begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}$, so that $A = [C_1 \quad C_2]$, then $\sigma_j^2(A) = \lambda_j(AA^T) = \lambda_j(C_1C_1^T + C_2C_2^T)$, for $j = 1, 2, \ldots, \min(m+n, \ell+p)$. Therefore, by (2.2), for $j = 1, 2, \ldots, \min(m+n, \ell, p)$

$$\sigma_j^2(A) \geq 2\sigma_j(C_1^T C_2) = 2\sigma_j\left(A_{11}^T A_{12} + A_{21}^T A_{22}\right). \quad \square$$

Corollary 3.15. *For a saddle point matrix*

$$\mathcal{X} = \begin{bmatrix} A & B \\ B^T & O \end{bmatrix},$$

where $A \in \mathbb{R}^{m\times m}$ is non-singular and $B \in \mathbb{R}^{m\times n}$ is full rank,

$$\sigma_j(\mathcal{X}) \geq \sqrt{2\sigma_j(AB)}, \qquad j = 1, 2, \ldots, \min(m, n).$$

4. Examples

In this section, we present some toy examples to illustrate the main results in section 3. We explore the behavior of the new results, compare their performance with some existing ones in the literature, and apply these results to matrices occurring in well-known applications.

Example 1. Here, we consider four pairs of PSD matrices P, Q, so that $P+Q$ is SPD. The exact value of $\lambda_{\min}(P+Q)$ is compared with the lower bounds given by Theorems 3.1 and 3.5. The existing results in the literature give a trivial lower bound. See Remark 3.2 for the definition of matrix $\mathcal{M}$ used in these examples.

1. Let $P = \operatorname{diag}(5, 0, 0)$, $Q = \operatorname{diag}(0, 4, 9)$, so that $\operatorname{rank} P = 1$ and $\operatorname{rank} Q = 2$, and $P + Q = \operatorname{diag}(5, 4, 9)$ is SPD. Note that $\mathcal{R}(P) = \mathcal{R}(Q)^\perp$, thus $k = r = 0$ and $p + q = 3 < 6$, so that $c(P, Q) = 1$. Therefore, both Theorems 3.1 and 3.5 give the lower bound 4, which is optimal as $\lambda_{\min}(P + Q) = 4$.
2. Let $P = \operatorname{diag}(1, 1, 0)$, $Q = \operatorname{diag}(0, 1, 3)$, so that $\operatorname{rank} P = \operatorname{rank} Q = 2$, and $P + Q = \operatorname{diag}(1, 2, 3)$ is SPD. Clearly, $\lambda_{\min}(P + Q) = 1$, let us calculate the lower bounds. Note that principal angles between $\mathcal{R}(P)$ and $\mathcal{R}(Q)$ are $\theta_1 = 0$ and $\theta_2 = \frac{\pi}{2}$, so that $k = 1$, $r = 0$ and $p + q = 4 < 6$, thus $c(P, Q) = 1$. Thus, both Theorems 3.1 and 3.5 give the optimal lower bound 1.
3. Consider the following PSD matrices P and Q so that $\operatorname{rank} P = \operatorname{rank} Q = 1$ and $P + Q$ is SPD,

$$P = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}, \; Q = \operatorname{diag}(6, 0), \; \text{thus } P + Q = \begin{bmatrix} 8 & 2 \\ 2 & 2 \end{bmatrix}.$$

Since the eigen-decomposition of $P = \mathcal{E}\Lambda\mathcal{E}^T$ and $Q = IQI^T$, where $\mathcal{E} = \frac{\sqrt{2}}{2}\begin{bmatrix}1 & -1\\ 1 & 1\end{bmatrix}$ and $\Lambda = \operatorname{diag}(4, 0)$, we get $\mathcal{M} = \frac{\sqrt{2}}{2}\begin{bmatrix}1 & 1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix} = \frac{\sqrt{2}}{2}$. Note that $k = 0$ and $r > 0$, thus $\cos\theta_1 = \sigma_{\max}(\mathcal{M}) = \frac{\sqrt{2}}{2}$ and Theorem 3.1 implies that

$$\lambda_{\min}(P + Q) \geq \left(1 - \frac{\sqrt{2}}{2}\right) \cdot 4 \approx 1.1716,$$

which is stronger than the lower bound obtained by applying Theorem 3.5,

$$\lambda_{\min}(P + Q) \geq 1.1270.$$

Thus, Corollary 3.6 gives the former result, that is, $\lambda_{\min}(P + Q) \geq 1.1716$, whereas the exact value of $\lambda_{\min}(P + Q)$ is approximately 1.3944.

4. Consider the following PSD matrices P and Q so that $\operatorname{rank} P = \operatorname{rank} Q = 2$ and $P + Q$ is SPD,

$$P = \begin{bmatrix}10 & & \\ & 5 & \\ & & 0\end{bmatrix}, Q = \begin{bmatrix}12 & & \\ & 3 & 9\\ & 9 & 27\end{bmatrix}, \text{ thus } P + Q = \begin{bmatrix}22 & & \\ & 8 & 9\\ & 9 & 27\end{bmatrix}.$$

Note that $k = r = 1 > 0$, so the orthonormal bases for P and Q define $\mathcal{M}$ as follows,

$$\mathcal{M} = \begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}0 & 1\\ \frac{1}{\sqrt{10}} & 0\\ \frac{3}{\sqrt{10}} & 0\end{bmatrix} = \begin{bmatrix}0 & 1\\ \frac{1}{\sqrt{10}} & 0\end{bmatrix}.$$

Thus, $c(P, Q) = 1 - \cos\theta_2 = 1 - \sigma_2(\mathcal{M}) = 1 - \frac{1}{\sqrt{10}}$, consequently Theorem 3.1 implies that

$$\lambda_{\min}(P + Q) \geq \left(1 - \frac{1}{\sqrt{10}}\right) \cdot 5 \approx 3.4189,$$

which is weaker than the lower bound obtained by applying Theorem 3.5,

$$\lambda_{\min}(P + Q) \geq 3.7770.$$

Thus, Corollary 3.6 also gives the above lower bound for $\lambda_{\min}(P + Q)$, the exact value of which is approximately 4.4137.

Example 2 *(Block diagonal matrix).* For a non-singular block diagonal matrix $D = \left[\begin{smallmatrix}A & O\\ O & B\end{smallmatrix}\right]$, use (3.25a) to get

$$\sigma_{\min}(D) \geq \sqrt{1-\cos\theta} \cdot \min\left\{\sigma_{\min}([A, O]), \sigma_{\min}([O, B])\right\},$$

where θ is the minimum angle between $\mathcal{R}\left([A, O]^T\right)$ and $\mathcal{R}\left([O, B]^T\right)$. Note that $\sigma_{\min}([A, O]) = \sigma_{\min}(A)$ and $\sigma_{\min}([O, B]) = \sigma_{\min}(B)$, and it is straightforward to see that $\theta = \frac{\pi}{2}$. Therefore, the result becomes

$$\sigma_{\min}(D) \geq \min\left\{\sigma_{\min}(A), \sigma_{\min}(B)\right\}.$$

In fact, the above inequality is an equality, thus the lower bound is sharp. Similar results can be obtained for a non-singular block anti-diagonal square matrix.

Example 3 *(Block triangular matrix).* For a non-singular block upper triangular matrix $U = \left[\begin{smallmatrix} A_{11} & A_{12} \\ O & A_{22} \end{smallmatrix}\right]$, the most simplified result of Theorem 3.9 is given by (3.26):

$$\sigma_{\min}(U) \geq \sqrt{1-\cos\theta} \cdot \min_{i=1,2}\{c_i\} \cdot \min\left\{\sigma_{\min}\left(A_{11}\right), \sigma_{\min}\left(A_{12}\right), \sigma_{\min}\left(O\right), \sigma_{\min}\left(A_{22}\right)\right\},$$

where θ is the minimum angle between $\mathcal{R}\left([A_{11}, A_{12}]^T\right)$ and $\mathcal{R}\left([O, A_{22}]^T\right)$, $c_1 = \sqrt{c(A_{11}, A_{12})}$, and $c_2 = \sqrt{c(O, A_{22})} = 1 \geq c_1$ by Property 3 of Proposition 3.4. Also, $\sigma_{\min}(O) = \infty$ by (1.2). Hence,

$$\sigma_{\min}(U) \geq \sqrt{1-\cos\theta} \cdot c_1 \cdot \min\left\{\sigma_{\min}(A_{11}), \sigma_{\min}(A_{12}), \sigma_{\min}(A_{22})\right\}. \tag{4.1}$$

Similarly, an estimate for a non-singular block lower triangular matrix can be derived.

When every block is a square matrix, then [8, p. 352] gives the following expression

$$U^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ O & A_{22}^{-1} \end{bmatrix},$$

whose maximum singular value is $\left(\sigma_{\min}(U)\right)^{-1}$. This task could be challenging to perform due to the presence of the term $A_{11}^{-1}A_{12}A_{22}^{-1}$. Note that the blocks need not be square for (4.1). For instance, consider the matrix U given as follows:

$$U = \begin{bmatrix} 10 & 0 & 0 \\ 4 & 2 & 0 \\ 1 & 1 & 6 \end{bmatrix},$$

then [50] gives $\sigma_{\min}(U) \geq 1.7087$. Whereas, on placing the partitions on U to make it a 2×2 block matrix so that its $(1, 1)$ block is either of size 2×1 or 2×2, (3.25b) gives a stronger result $1.7473 \leq 1.8285 \approx \sigma_{\min}(U)$. A MATLAB® implementation is given in [30].

Example 4. Here, we explain the use of the new lower bounds with the help of two 2×2 block matrices denoted by A, where its (i, j)-th block is denoted by A_{ij} and its i-th block row is denoted by R_i, where $1 \leq i, j \leq 2$. Also, $\mathcal{M}$ represents the matrix defined in

Remark 3.2. Most of the existing results do not provide a bound in terms of the blocks of the matrices considered in this example.

1. First, we consider a non-singular matrix with singular blocks, given as follows:

$$A = \left[\begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right].$$

Note that $\sigma_{\min}(A) = 1 = \sigma_{\min}(A_{ij}) = \sigma_{\min}(R_i)$. It is straightforward to see that a basis for $\mathcal{R}(R_1^T)$ is $\{[0\ 1\ 0\ 0]^T,\ [0\ 0\ 0\ 1]^T\}$ and for $\mathcal{R}(R_2^T)$ is $\{[1\ 0\ 0\ 0]^T,\ [0\ 0\ 1\ 0]^T\}$. Since $\theta = \frac{\pi}{2}$ or $r = 0$, (3.25a) implies

$$\sigma_{\min}(A) \geq (1 - 0) \cdot \min\{1, 1\} = 1,$$

and (3.25b) yields

$$\sigma_{\min}(A) \geq \sqrt{\min\{1, 1\}} = 1.$$

Moreover, the same result is obtained on applying (3.25c), as $r = 0$ for both the pairs A_{1i} and A_{2i} for $i = 1, 2$, thus

$$\sigma_{\min}(A) \geq (1 - 0) \min\{1, 1\} = 1.$$

In order to use (3.26), note that a basis for $\mathcal{R}(A_{11}), \mathcal{R}(A_{21})$ is $\{[0\ 1]^T\}$, and for $\mathcal{R}(A_{12}), \mathcal{R}(A_{22})$ is $\{[1\, 0]^T\}$, therefore $c_1 = c_2 = 1$. Thus, (3.26) yields

$$\sigma_{\min}(A) \geq (1 - 0) \min\{1, 1\} \min\{1, 1, 1, 1\} = 1.$$

To sum up, Theorem 3.9 gives the optimal lower bound as $\sigma_{\min}(A) = 1$.

2. Let us analyze the new lower bounds on a non-singular non-symmetric saddle point matrix with an indefinite leading block, given as follows:

$$A = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & -1 & 1 \\ \hline 1 & 0 & 0 \end{array}\right]$$

For using (3.25a), we use the orthonormal basis for $\mathcal{R}(R_1^T)$ and $\mathcal{R}(R_2^T)$ to define the following matrix,

$$\mathcal{M} = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & \frac{2}{\sqrt{3}} \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ 1 \end{bmatrix},$$

which gives $\cos\theta = \sigma_{\max}(\mathcal{M}) = \sqrt{\frac{2}{3}} \approx 0.8165$. Therefore, (3.25a) yields

$$\sigma_{\min}(A) \geq \sqrt{1 - 0.8165} \cdot \min\{1, 1\} \approx 0.4284.$$

Moreover, the same lower bound is obtained on applying other results. Since $k = \text{rank}(R_1^T) + \text{rank}(R_2^T) - 3 = 0$, and $r = 1$ with $\cos\theta$ derived as above, (3.25b) implies

$$\sigma_{\min}(A) \geq \sqrt{\frac{1}{2}\left[2 - \sqrt{4(1 - \sin^2\theta)}\right]} = \sqrt{1 - 0.8165} \approx 0.4284.$$

Also, (3.25c) results in

$$\sigma_{\min}(A) \geq \sqrt{1 - 0.8165} \cdot \min\{1, 1\} \approx 0.4284.$$

For using (3.26), observe that $c_1 = c_2 = 1$ and $\sigma_{\min}(A_{11}) = 1$, $\sigma_{\min}(A_{12}) = \sqrt{2}$, $\sigma_{\min}(A_{21}) = 1$ and $\sigma_{\min}(A_{22}) = \infty$. Therefore, the inequality gives

$$\sigma_{\min}(A) \geq \sqrt{1 - 0.8165} \cdot 1 \cdot \min\{\min\{1, \sqrt{2}\}, \min\{1, \infty\}\} \approx 0.4284.$$

Thus, Theorem 3.9 gives the best lower bound of value 0.4284 for the $\sigma_{\min}(A)$, the exact value of which is 0.4450.

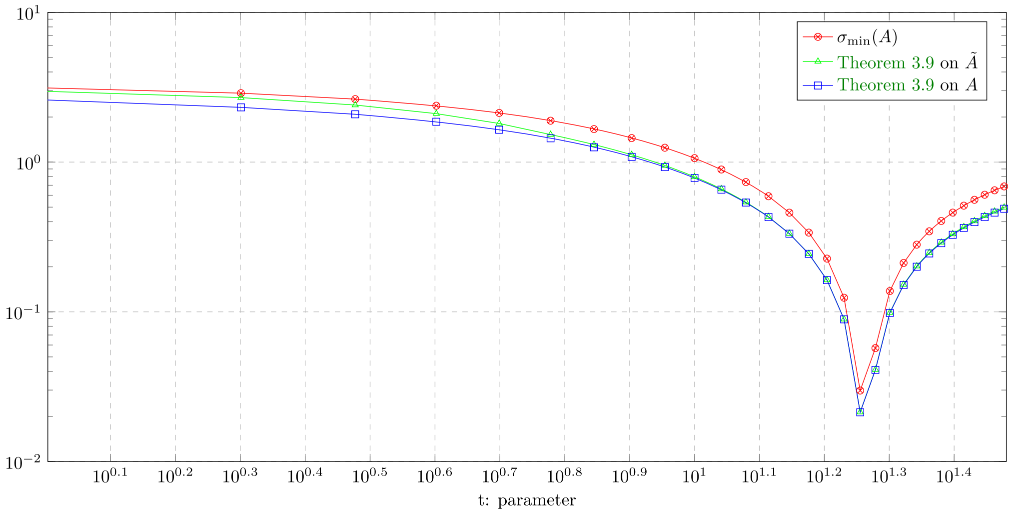
Example 5. Let us consider two different partitions on the same matrix as follows,

$$A = \left[\begin{array}{c|cc} t & 10 & 0 \\ \hline 3 & 2 & -2 \\ 2 & 0 & 6 \end{array}\right], \qquad \tilde{A} = \left[\begin{array}{c|cc} t & 10 & 0 \\ 3 & 2 & -2 \\ \hline 2 & 0 & 6 \end{array}\right], \qquad \text{where } t = 1, 2, \ldots, 30.$$

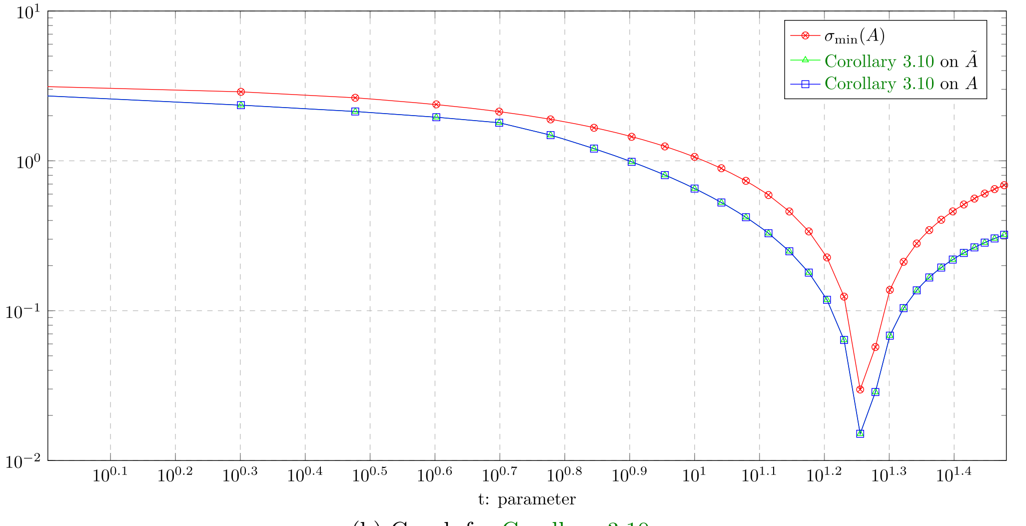
Fig. 1 displays the result of best lower bounds from Theorem 3.9 and Corollary 3.10 for both partitions, along with the exact value of $\sigma_{\min}(A)$.

Fig. 1a shows that Theorem 3.9 provides a decent estimate of $\sigma_{\min}(A)$. The best lower bound on $\sigma_{\min}(A)$ is given by (3.25b) for $t = 1$, and by (3.25a) for $t = 2, 3, \ldots 30$. For $\tilde{A}$, the largest lower bound is given by (3.25b) for $t = 1, 2, \ldots 5$, and by (3.25a) for $t = 6, 7, \ldots, 30$. On increasing the value of t, the lower bound obtained from (3.25a) improves up to $t = 18$, for which the absolute error in approximation for A is 0.008452 and for $\tilde{A}$ is 0.008389. The results from $\tilde{A}$ appear to be overall sharper than A. Thus, the sharpness of results may vary for distinct partitions of the same matrix.

Similarly, the trends for Corollary 3.10 are depicted by Fig. 1b. It is observed that Corollary 3.10 gives identical results for both matrix A and $\tilde{A}$. The minimum absolute error in approximation is 0.01469, which occurs when $t = 18$. The best lower bound is given by the first inequality of Corollary 3.10 for all t except for $t = 1$, for which it was obtained from the second inequality of Corollary 3.10. A Matlab® implementation is given in [30].



(a) Graph for Theorem 3.9.



(b) Graph for Corollary 3.10.

Fig. 1. Estimates of $\sigma_{\min}(A)$ for Example 5.

Example 6 *(M- and H-matrices).* The following matrices are considered in [40],

$$A = \begin{bmatrix} 8 & -2 & -1 \\ -5 & 7 & -3 \\ -3 & -4 & 5 \end{bmatrix}, \quad B = \begin{bmatrix} 7 & -3 & -2 \\ -2 & 5 & -1 \\ -3 & -4 & 9 \end{bmatrix}, \quad C = \begin{bmatrix} -5 & 2 & -4 \\ 3 & -6 & -2 \\ -1 & -4 & -8 \end{bmatrix},$$

where A and B are M-matrices and C is an H-matrix, for which upper bounds on the minimum singular values were devised. We calculate the best lower bounds secured from Theorem 3.9 and Corollary 3.10 in Table 1 through a MATLAB® implementation given

Table 1
Lower bounds for M- and H-matrices.

Matrix	$\sigma_{\min}$	Best lower bound	Size of leading block
A	0.7744	(3.29a):0.7354	1×2 or 2×2
B	1.8830	(3.29a):1.5855	1×2 or 2×2
C	0.9015	(3.29a):0.8770	1×1 or 2×1

Table 2
Comparison of new lower bounds with the existing results.

Matrix	$\sigma_{\min}$	Existing	New Result	Size of leading block
D	9.8608	9.6389	(3.25b):9.6932	1×2 or 2×1 or 2×2
E	9.0409	8.0731	(3.25b):8.1814	2×1 or 2×2
F	7.6233	5.6070	(3.25a):6.0553	1×1 or 1×2
G	6.7547	5.2107	(3.25a):5.2728	1×1 or 1×2
H	1.9619	1.4142	(3.29b):1.8651	1×1 or 2×1
I	1.0677	0.7898	(3.25a):0.8996	1×1 or 1×2
J	3.0786	2.2303	(3.25b):2.8220	1×1 or 1×2
K	2.5146	2.2170	(3.29b):2.3847	1×1 or 2×1

in [30]. In Table 1, the size leading block refers to the size of $(1,1)$ block of the matrix specifying the partition being placed, and more than one partition means that the same lower bound is obtained in all cases. It is evident that our results provide a good estimate for M- and H-matrices.

Example 7. In this example, we compare our results to some well-known existing results that give a lower bound on the minimum singular value of a matrix. The following matrices are strictly diagonally dominant (SDD) matrices, for which several lower bounds were analyzed in [40],

$$D=\begin{bmatrix}10&1&1\\1&20&1\\1&1&30\end{bmatrix},\ E=\begin{bmatrix}10&1&1\\1&20&1\\10&1&30\end{bmatrix},\ F=\begin{bmatrix}10&1&1\\1&20&1\\20&1&30\end{bmatrix},\ G=\begin{bmatrix}10&1&1\\10&20&1\\20&1&30\end{bmatrix}.$$

Also, some lower bounds for following matrices were compared in [38],

$$H=\begin{bmatrix}3&2&0\\1&9&5\\0&5&7\end{bmatrix},\quad I=\begin{bmatrix}2&-1&0\\2&1&0\\-4&-4&5\end{bmatrix},\quad J=\begin{bmatrix}5&0&0\\-4&9&4\\-1&7&9\end{bmatrix},\quad K=\begin{bmatrix}4&0&0\\-1&5&0\\0&5&4\end{bmatrix}.$$

The third column of Table 2 states the best among all lower bounds evaluated for the above matrices in [40,38]. The findings mentioned in Table 2 indicate that the results obtained from Theorem 3.9 and Corollary 3.10 provide a sharper lower bound on the minimum singular value of all SDD matrices considered in [40], albeit they may not be optimal for all SDD matrices. A MATLAB® implementation is given in [30]. In the above examples, we have listed the partitions that lead to the best estimates, which may not be feasible if the matrix is large. Based on our numerical experiments, choosing the leading block of the matrix to be a square matrix of a suitable size often results in a partition that gives fine results.

Example 8. One of the prominent problems consisting of a positive definite matrix expressed as a sum of two singular positive semi-definite matrices is solving the following linear system

$$(A + VV^T)x = b, \tag{4.2}$$

where $A \in \mathbb{R}^{n\times n}$ is PSD, $V \in \mathbb{R}^{n\times m}$, $b \in \mathbb{R}^n$, with $m < n$ so that the matrix $A + VV^T$ is non-singular, thus $\mathcal{N}(A) \cap \mathcal{N}(V^T) = \{0\}$. Such linear systems arise in problems such as KKT systems in constrained optimization, least squares problems, partial differential equations, etc. A lower bound on $\lambda_{\min}(A + VV^T)$ is required to estimate the condition number of $A + VV^T$, which describes the difficulty of solving this problem.

Let us analyze a special case of this problem, $\mathcal{A} + vv^T$, where we define $\mathcal{A} = \operatorname{diag}(0, 1, a_3, \dots, a_n)$ with $a_i \geq 1$ for all $i \geq 3$, and $v = [v_1 \ v_2 \ \dots \ v_n]^T \in \mathbb{R}^{n\times 1}$ with $|v| = 1$ and $v_1 \neq 0$. It is easily observed that $\lambda_{\min}(\mathcal{A}) = \lambda_{\min}(vv^T) = 1$. Note that an orthonormal basis for $\mathcal{R}(\mathcal{A})$ and $\mathcal{R}(vv^T)$ is given as $\{\mathbf{e}_2, \mathbf{e}_3, \dots, \mathbf{e}_n\}$ and $\{v\}$, respectively, where $\mathbf{e}_k$ represents the elementary basis vector of length n with the k-th component equals 1 and rest equal to 0. Therefore, $\mathcal{M} = [v_2 \, v_3 \, \dots \, v_n]$ implying $\cos\theta_1 = \sqrt{1 - v_1^2}$. Hence, Theorems 3.1 and 3.5 give the following result,

$$\lambda_{\min}(\mathcal{A} + vv^T) \geq 1 - \sqrt{1 - v_1^2} > 0.$$

Similarly, Theorems 3.1 and 3.5 can be applied for estimating the condition number of linear systems of the form (4.2).

Example 9. Consider the following Poisson problem, a second-order ordinary differential equation with some given boundary conditions, as stated in [36, p. 15]:

$$u''(x) = f(x) \text{ on } 0 < x < 1, \tag{4.3a}$$

$$u(0) = \alpha \text{ and } u(1) = \beta. \tag{4.3b}$$

Let N be a given discretization parameter and U_j represent the approximation to $u(x_j)$, where $x_j = jh$ and $h = N^{-1}$, for all $0 \leq j \leq N$. Then $U_0 = \alpha$ and $U_N = \beta$, and the rest of the approximations are obtained by solving the following second-order finite difference approximation to (4.3a),

$$\frac{U_{j-1} - 2U_j + U_{j+1}}{h^2} = f(x_j), \text{ for } 1 \leq j \leq N - 1, \tag{4.4}$$

which, according to [36, p. 34], can be written as follows for an odd N:

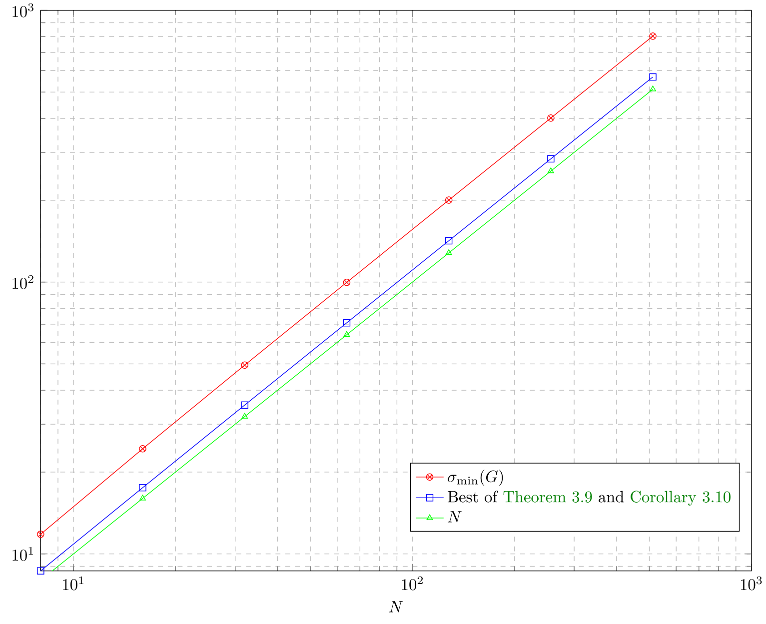


Fig. 2. Estimates of $\sigma_{\min}(G)$ for Example 9.

$$\frac{1}{h^2}\left[\begin{array}{ccccc|ccccc} -2 & & & & & 1 & & & & \\ & -2 & & & & 1 & 1 & & & \\ & & -2 & & & & 1 & 1 & & \\ & & & \ddots & & & & \ddots & \ddots & \\ & & & & -2 & & & & 1 & 1 \\ \hline 1 & 1 & & & & -2 & & & & \\ & 1 & 1 & & & & -2 & & & \\ & & \ddots & \ddots & & & & -2 & & \\ & & & 1 & 1 & & & & \ddots & \\ & & & & 1 & & & & & -2 \end{array}\right]\left[\begin{array}{c} U_1 \\ U_3 \\ U_5 \\ \vdots \\ U_{N-2} \\ \hline U_2 \\ U_4 \\ U_6 \\ \vdots \\ U_{N-1} \end{array}\right] = \left[\begin{array}{c} f(x_1) - \frac{\alpha}{h^2} \\ f(x_3) \\ f(x_5) \\ \vdots \\ f(x_{N-2}) \\ \hline f(x_2) \\ f(x_4) \\ f(x_6) \\ \vdots \\ f(x_{N-1}) - \frac{\beta}{h^2} \end{array}\right]. \tag{4.5}$$

Define the coefficient matrix for the above problem as G of size $(N-1) \times (N-1)$. Fig. 2 displays the results of the exact $\sigma_{\min}(G)$ and the best lower bound obtained from Theorem 3.9 and Corollary 3.10 for $N = 2^k$, where $3 \leq k \leq 9$. For $N = 16$ and 512, the best lower bounds were obtained from (3.25a) and (3.29b), whereas (3.25c) produced the best estimate for $N = 32$, and the rest of them were obtained from (3.25b). It is concluded that the new lower bounds give an optimal estimate for $\sigma_{\min}(G)$, that is, both the exact and approximate values are of $\mathcal{O}(N)$. See [30] for a MATLAB® implementation.

5. Summary and future work

According to Sophie Germain, "Algebra is but written geometry and geometry is but figured algebra." In this paper, we attempt to solve a long-standing edge case of the spectral problem of a symmetric matrix sum by inspecting it from a geometric perspective. Precisely, our aim was to formulate a positive lower bound for the minimum eigenvalue of an SPD matrix expressed as a sum of two singular PSD matrices, say $P+Q$, by exploring the separation between the range spaces of the matrices P and Q as their corresponding null spaces are disjoint. The existing results for such a case, including Weyl's inequalities, yield a trivial lower bound on $\lambda_{\min}(P+Q)$. Thus, we resolve this case by formulating two results generating a positive lower bound on $\lambda_{\min}(P+Q)$, given in the form of Theorems 3.1 and 3.5 based upon the Friedrichs angle between $\mathcal{R}(P)$ and $\mathcal{R}(Q)$. Since Example 1 illustrates that either one of two lower bounds could be stronger than the other under varied circumstances, the best lower bound is provided by Corollary 3.6. A significant application of such an edge case is discussed in Example 8.

Direct use of Corollary 3.6 results in lower bounds on the minimum singular value of full rank block column and block row matrices, given by Corollaries 3.7 and 3.8, respectively. Moreover, Theorem 3.1 and Theorem 3.5 are applied recursively for developing eight lower bounds on the minimum singular value of a non-singular block 2×2 matrix, which are given by Theorem 3.9 and Corollary 3.10. As any full rank matrix can be partitioned in terms of block row or column matrix, and any non-singular matrix can be partitioned in terms of a block 2×2 matrix, the applications of the results derived for block matrices are numerous.

Example 5 illustrates that each partition yields different results for a matrix, thus several estimates can be obtained from the eight lower bounds formulated in Theorem 3.9 and Corollary 3.10, which sometimes are stronger than the ones existing in literature, as mentioned in Examples 3 and 6. These eight results satisfy the relations (3.25a) $\geq$ (3.25c) $\geq$ (3.26), and (3.29a) $\geq$ (3.29c) $\geq$ (3.30). No other relationships among the eight lower bounds can be established. Thus, we suggest taking the maximum of all lower bounds from Theorem 3.9 and Corollary 3.10, which for a partition described in Example 9 leads to an optimal estimate. Lastly, we give some estimates on the first few singular values of a block 2×2 matrix. We extend Theorem 3.13 to more general block sizes in Theorem 3.14 with the help of a much simpler proof.

The lower bounds described by Theorems 3.1 and 3.5 are sharp for the case of $\mathcal{R}(P) \cap \mathcal{R}(Q) = \{0\}$ so that the parameter $k = 0$. However, there is a scope for improvement in these results for the case when $r, k > 0$. It may be possible to incorporate $P_{M_1 \cap M_2}$ in (3.22) to improve the lower bound. Since changing the partition of a matrix changes the lower bounds given by Theorem 3.9 and Corollary 3.10, thus one can try to determine the best partition for a certain class of matrices yielding optimal estimates on the minimum singular value.

Some techniques exist for calculating the principal angles between two subspaces, see [35]. A more efficient algorithm can be designed for calculating the Friedrichs angle between two subspaces. Since it is defined for subspaces of Hilbert space in Definition 2.3, it may allow us to extend the main results to a more general setting.

CRediT authorship contribution statement

Avleen Kaur: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing – original draft. **S.H. Lui:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Software, Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have cited the link to our GitHub repository containing the codes used in the article.

Acknowledgements

We thank Richard Mikaël Slevinsky and Sooyeong Kim for a careful reading of the manuscript. We would like to thank the reviewer for taking the time and effort necessary to review the manuscript. We sincerely appreciate all valuable comments and suggestions, which helped us to improve the quality of the manuscript.

References

[1] O. Axelsson, Unified analysis of preconditioning methods for saddle point matrices, Numer. Linear Algebra Appl. 22 (2015) 233–253, https://doi.org/10.1002/nla.1947.
[2] O. Axelsson, M. Neytcheva, Eigenvalue estimates for preconditioned saddle point matrices, Numer. Linear Algebra Appl. 13 (2006) 339–360, https://doi.org/10.1002/nla.469.
[3] Z.Z. Bai, Eigenvalue estimates for saddle point matrices of Hermitian and indefinite leading blocks, J. Comput. Appl. Math. 237 (2013) 295–306, https://doi.org/10.1016/j.cam.2012.05.007.
[4] D. Barber, Bayesian Reasoning and Machine Learning, Bayesian Reasoning and Machine Learning, Cambridge University Press, 2012.
[5] M. Benzi, G. Golub, J. Liesen, Numerical solution of saddle point problems, Acta Numer. 14 (2005) 1–137, https://doi.org/10.1017/S0962492904000212.
[6] M. Benzi, V. Simoncini, On the eigenvalues of the class of saddle point matrices, Numer. Math. 103 (2006) 173–196, https://doi.org/10.1007/s00211-006-0679-9.
[7] L. Bergamaschi, On the eigenvalue distribution of constrained preconditioned symmetric saddle point matrices, Numer. Linear Algebra Appl. 19 (2012) 754–772, https://doi.org/10.1002/nla.806.
[8] D.S. Bernstein, Scalar, Vector, and Matrix Mathematics: Theory, Facts, and Formulas, revised and expanded edition, Princeton University Press, Princeton, NJ, 2018.
[9] R. Bhatia, F. Kittaneh, On singular values of product of operators, SIAM J. Matrix Anal. Appl. 11 (1990) 272–277, https://doi.org/10.1137/0611018.
[10] R. Bhatia, F. Kittaneh, The matrix arithmetic-geometric mean inequality revisited, Linear Algebra Appl. 428 (2008) 2177–2191, https://doi.org/10.1016/j.laa.2007.11.030.

[11] A. Björck, G. Golub, Numerical methods for computing angles between linear subspaces, Math. Comput. 27 (1973) 579–594, https://doi.org/10.2307/2005662.
[12] A. Böttcher, I. Spitkovsky, A gentle guide to the basics of two projections theory, Linear Algebra Appl. 432 (2010) 1412–1459, https://doi.org/10.1016/j.laa.2009.11.002.
[13] Y. Chen, Y. Chi, J. Fan, C. Ma, Spectral methods for data science: a statistical perspective, Found. Trends Mach. Learn. 14 (2021) 566–806, https://doi.org/10.1561/2200000079.
[14] J. Dauxois, G. Nkiet, Canonical analysis of two Euclidean subspaces and its applications, in: Sixth Special Issue on Linear Algebra and Statistics, Linear Algebra Appl. 264 (1997) 355–388, https://doi.org/10.1016/S0024-3795(96)00244-3.
[15] J.M. Day, W. So, R.C. Thompson, The spectrum of a Hermitian matrix sum, Linear Algebra Appl. 280 (1998) 289–332, https://doi.org/10.1016/s0024-3795(98)10019-8.
[16] F. Deutsch, The angle between subspaces of a Hilbert space, in: S.P. Singh (Ed.), Approximation Theory, Wavelets and Applications, Springer Netherlands, Dordrecht, 1995, pp. 107–130.
[17] S. Drury, On a question of Bhatia and Kittaneh, Linear Algebra Appl. 437 (2012) 1955–1960, https://doi.org/10.1016/j.laa.2012.04.040.
[18] A. Galántai, Projectors and Projection Methods, Advances in Mathematics, Springer, Boston, MA, 2004.
[19] A. Galántai, Subspaces, angles and pairs of orthogonal projections, Linear Multilinear Algebra 56 (2008) 227–260, https://doi.org/10.1080/03081080600743338.
[20] G.H. Golub, C.F. Van Loan, Matrix Computations, 4th ed., The Johns Hopkins University Press, Baltimore, 2013, https://jhupbooks.press.jhu.edu/title/matrix-computations.
[21] G.H. Golub, H. Zha, The canonical correlations of matrix pairs and their numerical computation, in: A. Bojanczyk, G. Cybenko (Eds.), Linear Algebra for Signal Processing, Springer New York, New York, NY, 1995, pp. 27–49.
[22] N. Gould, V. Simoncini, Spectral analysis of saddle point matrices with indefinite leading blocks, SIAM J. Matrix Anal. Appl. 31 (2010) 1152–1171, https://doi.org/10.1137/080733413.
[23] O. Hirzallah, Inequalities for sums and products of operators, Linear Algebra Appl. 407 (1990) 32–42, https://doi.org/10.1016/j.laa.2005.04.017.
[24] Y. Hong, C.-T. Pan, A lower bound for the smallest singular value, Linear Algebra Appl. 172 (1992) 27–32, https://doi.org/10.1016/0024-3795(92)90016-4.
[25] R.A. Horn, C.R. Johnson, Matrix Analysis, 2nd ed., Cambridge University Press, Cambridge, 2013.
[26] N. Huang, C.-F. Ma, On the eigenvalues of the saddle point matrices discretized from Navier-Stokes equations, Numer. Algorithms 79 (2018) 41–64, https://doi.org/10.1007/s11075-017-0427-5.
[27] C. Johnson, A Gersgorin-type lower bound for the smallest singular value, Linear Algebra Appl. 112 (1989) 1–7, https://doi.org/10.1016/0024-3795(89)90583-1.
[28] C. Johnson, T. Szulc, Further lower bounds for the smallest singular value, Linear Algebra Appl. 272 (1998) 169–179, https://doi.org/10.1016/S0024-3795(97)00330-3.
[29] C. Johnson, T. Szulc, D. Wojtera-Tyrakowska, Optimal Gersgorin-style estimation of extremal singular values, Linear Algebra Appl. 402 (2005) 46–60, https://doi.org/10.1016/j.laa.2005.01.024.
[30] A. Kaur, S.H. Lui, https://github.com/avleenk2312/lb_minsingular, 2022. (Accessed 21 November 2022).
[31] A. Kaur, S.H. Lui, Space-time spectral method for the Stokes problem, Appl. Numer. Math. (2023), https://doi.org/10.1016/j.apnum.2023.02.013.
[32] A.A. Klyachko, Stable bundles, representation theory and Hermitian operators, Sel. Math. New Ser. 4 (1998) 419–445, https://doi.org/10.1007/s000290050037.
[33] A. Knutson, T. Tao, Honeycombs and sums of Hermitian matrices, Not. Am. Math. Soc. 48 (2001) 175–186, http://www.ams.org/notices/200102/fea-knutson.pdf.
[34] A. Knyazev, Toward the optimal preconditioned eigensolver: locally optimal block preconditioned conjugate gradient method, SIAM J. Sci. Comput. (2001), https://doi.org/10.1137/s1064827500366124.
[35] A. Knyazev, M. Argentati, Principal angles between subspaces in an A-based scalar product: algorithms and perturbation estimates, SIAM J. Sci. Comput. 23 (2002) 2009–2041, https://doi.org/10.1137/S1064827500377332.
[36] R.J. LeVeque, Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems, SIAM, 2007.
[37] L. Li, Estimation for matrix singular value, Comput. Math. Appl. 37 (1999) 9–15, https://doi.org/10.1016/S0898-1221(99)00108-X.
[38] M. Lin, M. Xie, On some lower bounds for smallest singular value of matrices, Appl. Math. Lett. 121 (2021) 107411, https://doi.org/10.1016/j.aml.2021.107411.

[39] C.D. Meyer, Matrix Analysis and Applied Linear Algebra, 1st ed., SIAM, Philadelphia, 2000.
[40] N. Morača, Bounds for norms of the matrix inverse and the smallest singular value, Linear Algebra Appl. (2008), https://doi.org/10.1016/j.laa.2007.12.026.
[41] L. Qi, Some simple estimates for singular value of a matrix, Linear Algebra Appl. 56 (1984) 105–119, https://doi.org/10.1016/0024-3795(84)90117-4.
[42] S.-Q. Shen, L. Jian, W.-D. Bao, T.-Z. Huang, On the eigenvalue distribution of preconditioned nonsymmetric saddle point matrices, Numer. Linear Algebra Appl. 21 (2014) 557–568, https://doi.org/10.1137/100783509.
[43] D. Silvester, A. Wathen, Fast iterative solution of stabilised Stokes systems, part II: using general block preconditioners, SIAM J. Numer. Anal. 31 (1994) 1352–1367, https://doi.org/10.1137/0731070.
[44] A. Smoktunowicz, Block matrices and symmetric perturbations, Linear Algebra Appl. 429 (2008) 2628–2635, https://doi.org/10.1016/j.laa.2008.03.013.
[45] G.W. Stewart, J.-g. Sun, Matrix Perturbation Theory, Computer Science and Scientific Computing, Elsevier Science, 1990.
[46] J. Varah, A lower bound on the smallest singular value of a matrix, Linear Algebra Appl. 11 (1975) 3–5, https://doi.org/10.1016/0024-3795(75)90112-3.
[47] P.Å. Wedin, On angles between subspaces of a finite dimensional inner product space, in: B. Kågström, A. Ruhe (Eds.), Matrix Pencils, Springer Berlin Heidelberg, Berlin, Heidelberg, 1983, pp. 263–285.
[48] Y.-S. Yu, D.-H. Gu, A note on a lower bound for the smallest singular value, Linear Algebra Appl. 253 (1997) 25–38, https://doi.org/10.1016/0024-3795(95)00784-9.
[49] L. Zou, A lower bound for the smallest singular value, J. Math. Inequal. 6 (2012) 625–629, https://doi.org/10.7153/jmi-06-60.
[50] L. Zou, Y. Jiang, Estimation of the eigenvalues and the smallest singular value of matrices, Linear Algebra Appl. 433 (2010) 1203–1211, https://doi.org/10.1016/j.laa.2010.05.002
.