Numer Algor (2018) 79:657–677
https://doi.org/10.1007/s11075-018-0473-7

CrossMark

ORIGINAL PAPER

Some refined bounds for the perturbation of the orthogonal projection and the generalized inverse

Wen Li[1] · Yanmei Chen[2,1] · Seakweng Vong[3] · Qilun Luo[1]

Received: 28 July 2017 / Accepted: 3 January 2018 / Published online: 18 January 2018

Abstract In this paper, we consider the perturbation of the orthogonal projection and the generalized inverse for an $n \times n$ matrix A and present some perturbation bounds for the orthogonal projections on the rang spaces of A and A^*, respectively. A combined bound for the orthogonal projection on the rang spaces of A and A^* is also given. The proposed bounds are sharper than the existing ones. From the combined bounds of the orthogonal projection on the rang spaces of A and A^*, we derived new perturbation bounds for the generalized inverse, which always improve the existing ones. The combined perturbation bound for the orthogonal projection and the generalized inverse is also given. Some numerical examples are given to show the advantage of the new bounds.

The work of authors were supported in part by National Natural Science Foundations of China (No. 11671158, 11771159, 11571124, 11601347), Major Project (No.2016KZDXM025), and Innovation Team Project (No. 2015KCXTD007) of Guangdong Provincial Universities and University of Macau (No. MYRG2017-00098-FST).

✉ Yanmei Chen
chch1980@163.com

Wen Li
liwen@scnu.edu.cn

Seakweng Vong
swvong@umac.mo

Qilun Luo
rockyalan@m.scnu.edu.cn

[1] School of Mathematical Sciences, South China Normal University, Guangzhou, 510631 China

[2] School of Mathematics and Systems Science, Guangdong Polytechnic Normal University, Guangzhou, 510665 China

[3] Department of Mathematics, University of Macau, Macau, China

Keywords Singular value decomposition · Orthogonal projection · Perturbation · Generalized inverse

Mathematics Subject Classification (2010) 65F10

1 Introduction

Let $C^{m\times n}$ and $C_r^{m\times n}$ be the set of all $m\times n$ complex matrices and the set of all $m\times n$ complex matrices with rank r, respectively. For a matrix $A\in C^{m\times n}$, by A^*, P_A, $rang(A)$, $\|A\|$, $\|A\|_F$, and $\|A\|_2$, we denote the conjugate transpose, the orthogonal projection, the rang space, the general unitarily invariant norm, the Frobenius norm (F-norm), and the spectral norm (2-norm) of A, respectively. We use I_m to denote the identity matrix of order m .

Let $A\in C_r^{m\times n}$ and $B\in C_s^{m\times n}$ have the following singular value decompositions (SVDs):

$$A=U\begin{pmatrix}\Sigma_1 & 0\\ 0 & 0\end{pmatrix}V^* \text{ and } B=\tilde{U}\begin{pmatrix}\tilde{\Sigma}_1 & 0\\ 0 & 0\end{pmatrix}\tilde{V}^*, \tag{1.1}$$

where $U=(\,U_1\;\;U_2\,)$, $\tilde{U}=(\,\tilde{U}_1\;\;\tilde{U}_2\,)\in C^{m\times m}$ and $V=(\,V_1\;\;V_2\,)$, $\tilde{V}=(\,\tilde{V}_1\;\;\tilde{V}_2\,)\in C^{n\times n}$ are unitary matrices, $U_1\in C^{m\times r}$, $\tilde{U}_1\in C^{m\times s}$, $V_1\in C^{n\times r}$, $\tilde{V}_1\in C^{n\times s}$, $\Sigma_1=diag(\sigma_1,\ldots,\sigma_r)$, $\tilde{\Sigma}_1=diag(\tilde{\sigma}_1,\ldots,\tilde{\sigma}_s)$, $\sigma_1\ge\cdots\ge\sigma_r>0$ and $\tilde{\sigma}_1\ge\cdots\ge\tilde{\sigma}_s>0$. Let $A^\dagger$ denote the Moore-Penrose inverse of A. It is easy to get

$$A=U_1\Sigma_1V_1^*,\, B=\tilde{U}_1\tilde{\Sigma}_1\tilde{V}_1^*, \tag{1.2}$$

$$A^\dagger=V_1\Sigma_1^{-1}U_1^*,\;\; B^\dagger=\tilde{V}_1\tilde{\Sigma}_1^{-1}\tilde{U}_1^*, \tag{1.3}$$

$$\|A^\dagger\|_2=\frac{1}{\sigma_r} \text{ and } \|B^\dagger\|_2=\frac{1}{\tilde{\sigma}_s}. \tag{1.4}$$

By (1.2) and (1.3), we have

$$E=B-A=\tilde{U}_1\tilde{\Sigma}_1\tilde{V}_1^*-U_1\Sigma_1V_1^*, \tag{1.5}$$

$$P_A=AA^\dagger=U_1U_1^*,\;\; P_{A^*}=A^*A^{*\dagger}=A^\dagger A=V_1V_1^* \tag{1.6}$$

and

$$P_B=\tilde{U}_1\tilde{U}_1^*,\;\; P_{B^*}=V_1\tilde{V}_1^*. \tag{1.7}$$

In this paper, we frequently use the inequality as follows: if C is an $n\times n$ matrix, σ_n and σ_1 are the smallest and the largest singular values of C, respectively, then

$$\sigma_1\|A\|_F\ge\|CA\|_F\ge\sigma_n\|A\|_F,\;\sigma_1\|A\|_F\ge\|AC\|_F\ge\sigma_n\|A\|_F.$$

The generalized inverse and the orthogonal projection play important roles in matrix computations. In particular, it can be applied to analyze sensitivity for solving least square problems. Recently, some researchers have studied their perturbation bounds, e.g., see [1–3, 6–8]. The classical Frobenius norm bounds for the orthogonal projection are listed below (e.g., see [8]): Let $A\in C_r^{m\times n}$ and $B=A+E\in C_s^{m\times n}$. Then

$$\|P_B-P_A\|_F\le\sqrt{\|A^\dagger\|_2^2+\|B^\dagger\|_2^2}\|E\|_F. \tag{1.8}$$

In particular, if $r = s$, then

$$\|P_B - P_A\| \le \sqrt{2}\min\{\|A^\dagger\|_2, \|B^\dagger\|_2\}\|E\|_F. \tag{1.9}$$

Let $A \in C_r^{m\times n}$, $B = A + E \in C_s^{m\times n}$. Recently, Chen, Chen and Li [2] presented the following Frobenius norm bound for the orthogonal projection:

$$\|P_B - P_A\|_F \le \sqrt{\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2}. \tag{1.10}$$

If $r = s$, then

$$\|P_B - P_A\|_F \le \sqrt{2}\min\{\|EA^\dagger\|_F, \|EB^\dagger\|_F\}. \tag{1.11}$$

The bounds (1.10) and (1.11) improve the corresponding ones in (1.8) and (1.9). On the other hand, they gave a combined bound as follows (see the bound (2.24) of Theorem 2.8 in [2]):

$$\|P_B - P_A\|_F^2 + \|P_{B^*} - P_{A^*}\|_F^2 \le \frac{4}{\sigma_r^2 + \tilde{\sigma}_s^2}\|E\|_F^2. \tag{1.12}$$

For the perturbation bound of the generalized inverse, Meng and Zheng [6] presented the following results: Let $A \in C_r^{m\times n}$ and $B = A + E \in C_s^{m\times n}$. Then

$$\|B^\dagger - A^\dagger\|_F \le \frac{1}{\min\{\sigma_r^2, \tilde{\sigma}_s^2\}}\|E\|_F. \tag{1.13}$$

If $r = s$, then

$$\|B^\dagger - A^\dagger\|_F \le \frac{1}{\sigma_r\tilde{\sigma}_r}\|E\|_F. \tag{1.14}$$

In order to illustrate our motivation of this paper, we give a simple example as follows. Let

$$A = \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} (1+\varepsilon)U & 0 \\ 0 & 0 \end{pmatrix},$$

where U is a $r \times r$ unitary, $\varepsilon > 0$. Then

$$EA^\dagger = \begin{pmatrix} \varepsilon I & 0 \\ 0 & 0 \end{pmatrix}, \quad EB^\dagger = \begin{pmatrix} \frac{\varepsilon}{1+\varepsilon} I & 0 \\ 0 & 0 \end{pmatrix}.$$

It is easy to get

$$\|P_B - P_A\|_F = 0,$$

and the right hand sides of the bounds (1.10) and (1.11) are

$$\sqrt{\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2} = \sqrt{\left(1 + \frac{1}{(1+\varepsilon)^2}\right)\varepsilon^2 r}$$

and $\sqrt{2}\min\{\|EA^\dagger\|_F, \|EB^\dagger\|_F\} = \frac{\varepsilon}{1+\varepsilon}\sqrt{2r}$, respectively. This example shows that the variation of the orthogonal projection is independent of the perturbation ε. However, the bounds in (1.10) and (1.11) depend on the perturbation ε. By this motivation, we explore some new variations for the perturbation of the orthogonal projection; the new bounds may rub down the weakness of the bounds (1.10) and (1.11). From the proposed bounds, we also derive the new perturbation bounds of the generalized inverse, which can always improve the bound in (1.13). The idea is based on the elaborate analysis on the matrix decompositions. The contributions of this article are given below:

- Improves the existing Frobenius norm bounds for perturbation of orthogonal projections
- Improves the perturbation bound for the generalized inverse by the perturbation result of orthogonal projections

The rest of the paper is organized as follows. In Section 2, we present the perturbation bounds for the orthogonal projection on $rang(A)$ and $rang(A^*)$, respectively. Also, we give a combined bound for the orthogonal projection on $rang(A)$ and $rang(A^*)$. Comparing with some existing ones, our bounds are sharper. In Section 3, we give the perturbation bounds for the generalized inverse, which always improve the existing one. In Section 4, we consider the combined perturbation bound for the orthogonal projection and the generalized inverse. In Section 5, we plot figures of some examples to show the advantage of the new bounds. The final section gives some concluding remarks.

2 The perturbation bound for orthogonal projections

In this section, we consider improving some existing perturbation bounds for the orthogonal projection. The following two lemmas were given in [2], which will be used in the sequel.

Lemma 2.1 *Let $U = (U_1\ U_2)$ and $V = (V_1\ V_2)$ be unitary matrices of order n, where $U_1, V_1 \in \mathcal{C}^{n\times r}$. Then for any unitarily invariant norm $||\cdot||$, we have*

$$\|U_1^* V_2\| = \|U_2^* V_1\|. \tag{2.1}$$

Lemma 2.2 *Let $A \in C_r^{m\times n},\ B \in C_s^{m\times n}$ have the SVDs (1.1). Then*

$$\begin{aligned} \|P_B - P_A\|_F^2 &= \|\tilde{U}_1^* U_2\|_F^2 + \|U_1^* \tilde{U}_2\|_F^2 \\ \|P_{B^*} - P_{A^*}\|_F^2 &= \|\tilde{V}_1^* V_2\|_F^2 + \|V_1^* \tilde{V}_2\|_F^2. \end{aligned} \tag{2.2}$$

If $r = s$, then

$$\|P_B - P_A\|_F^2 = 2\|\tilde{U}_1^* U_2\|_F^2 \ \text{ and } \|P_{B^*} - P_{A^*}\|_F^2 = 2\|\tilde{V}_1^* V_2\|_F^2. \tag{2.3}$$

In the following, we discuss the perturbation of the orthogonal projection.

Theorem 2.3 *Let $A \in C_r^{m\times n}$, and $B = A + E \in C_s^{m\times n}$ be a perturbed matrix of A. Then*

$$\|P_{B^*} - P_{A^*}\|_F^2 \le \frac{1}{\tilde{\sigma}_s^2}(\|E\|_F^2 - \sigma_r^2\|EA^\dagger\|_F^2) + \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|EB^\dagger\|_F^2).$$

If $r = s$, then

$$\|P_{B^*} - P_{A^*}\|_F^2 \le 2\min\left\{\frac{1}{\tilde{\sigma}_r^2}(\|E\|_F^2 - \sigma_r^2\|EA^\dagger\|_F^2), \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|EB^\dagger\|_F^2\right\}.$$

Proof By (1.5) and (1.3), it is easy to get

$$\tilde{U}^* E A^\dagger U = \begin{pmatrix} \tilde{\Sigma}_1 \tilde{V}_1^* V_1 \Sigma_1^{-1} - \tilde{U}_1^* U_1 & 0 \\ -\tilde{U}_2^* U_1 & 0 \end{pmatrix}.$$

Hence,

$$\begin{aligned} \|EA^\dagger\|_F^2 &= \|\tilde{\Sigma}_1 \tilde{V}_1^* V_1 \Sigma_1^{-1} - \tilde{U}_1^* U_1\|_F^2 + \|\tilde{U}_2^* U_1\|_F^2 \\ &= \|(\tilde{\Sigma}_1 \tilde{V}_1^* V_1 - \tilde{U}_1^* U_1 \Sigma_1)\Sigma_1^{-1}\|_F^2 + \|\tilde{U}_2^* U_1\|_F^2 \\ &\le \frac{1}{\sigma_r^2}\|\tilde{\Sigma}_1 \tilde{V}_1^* V_1 - \tilde{U}_1^* U_1 \Sigma_1\|_F^2 + \|\tilde{U}_2^* U_1\|_F^2, \end{aligned} \tag{2.4}$$

from which one can deduce that

$$\|\tilde{\Sigma}_1 \tilde{V}_1^* V_1 - \tilde{U}_1^* U_1 \Sigma_1\|_F^2 \ge \sigma_r^2(\|EA^\dagger\|_F^2 - \|\tilde{U}_2^* U_1\|_F^2). \tag{2.5}$$

By (1.5), we have

$$\begin{aligned} \tilde{U}^* E V &= \tilde{U}^*(\tilde{U}_1 \tilde{\Sigma}_1 \tilde{V}_1^* - U_1 \Sigma_1 V_1^*)V \\ &= \begin{pmatrix} \tilde{\Sigma}_1 \tilde{V}_1^* V_1 - \tilde{U}_1^* U_1 \Sigma_1 & \tilde{\Sigma}_1 \tilde{V}_1^* V_2 \\ -\tilde{U}_2^* U_1 \Sigma_1 & 0 \end{pmatrix}, \end{aligned} \tag{2.6}$$

which together with (2.5) gives

$$\begin{aligned} \|E\|_F^2 &= \|\tilde{\Sigma}_1 \tilde{V}_1^* V_1 - \tilde{U}_1^* U_1 \Sigma_1\|_F^2 + \|\tilde{U}_2^* U_1 \Sigma_1\|_F^2 + \|\tilde{\Sigma}_1 \tilde{V}_1^* V_2\|_F^2 \\ &\ge \sigma_r^2(\|EA^\dagger\|_F^2 - \|\tilde{U}_2^* U_1\|_F^2) + \sigma_r^2\|\tilde{U}_2^* U_1\|_F^2 + \tilde{\sigma}_s^2\|\tilde{V}_1^* V_2\|_F^2 \\ &= \sigma_r^2\|EA^\dagger\|_F^2 + \tilde{\sigma}_s^2\|\tilde{V}_1^* V_2\|_F^2. \end{aligned}$$

Hence,

$$\|\tilde{V}_1^* V_2\|_F^2 \le \frac{1}{\tilde{\sigma}_s^2}(\|E\|_F^2 - \sigma_r^2\|EA^\dagger\|_F^2). \tag{2.7}$$

Interchanging the matrices A and B in (2.7) gives

$$\|V_1^* \tilde{V}_2\|_F^2 \le \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|EB^\dagger\|_F^2). \tag{2.8}$$

By Lemma 2.2, we have

$$\|P_{B^*} - P_{A^*}\|_F^2 = \|\tilde{V}_2^* V_1\|_F^2 + \|\tilde{V}_1^* V_2\|_F^2,$$

which together (2.7) and (2.8) gives

$$\|P_{B^*} - P_{A^*}\|_F^2 \le \frac{1}{\tilde{\sigma}_s^2}(\|E\|_F^2 - \sigma_r^2\|EA^\dagger\|_F^2) + \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|EB^\dagger\|_F^2).$$

If $r = s$, by Lemma 2.2, (2.7), and (2.8), we get

$$\|P_{B^*} - P_{A^*}\|_F^2 \le 2\min\left\{\frac{1}{\tilde{\sigma}_r^2}(\|E\|_F^2 - \sigma_r^2\|EA^\dagger\|_F^2), \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|EB^\dagger\|_F^2)\right\}.$$

This proves the theorem. □

We replace A and B by A^* and B^* in Theorem 2.3, respectively; it is easy to get the following corollary.

Corollary 2.4 *Let $A \in C_r^{m\times n}$ and $B = A + E \in C_s^{m\times n}$. Then*

$$\|P_B - P_A\|_F^2 \le \frac{1}{\tilde{\sigma}_s^2}(\|E\|_F^2 - \sigma_r^2\|A^\dagger E\|_F^2) + \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|B^\dagger E\|_F^2). \tag{2.9}$$

If $r = s$, then

$$\|P_B - P_A\|_F^2 \le 2\min\left\{\frac{1}{\tilde{\sigma}_r^2}(\|E\|_F^2 - \sigma_r^2\|A^\dagger E\|_F^2), \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_r^2\|B^\dagger E\|_F^2)\right\}. \tag{2.10}$$

Remark 2.1 It is easy to see that the bounds in Corollary 2.4 are always sharper than the corresponding one in (1.8) and (1.9) of [8]. In order to show that the bound (2.10) is the sharpest, taking matrices A and B as given in Section 1, then the bound (2.10) gives

$$\begin{aligned}\frac{1}{\tilde{\sigma}_r^2}(\|E\|_F^2 - \sigma_r^2\|A^\dagger E\|_F^2) &= \left(\frac{1}{1+\varepsilon}\right)^2(\varepsilon^2 r - \varepsilon^2 r) = 0,\\ \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_r^2\|B^\dagger E\|_F^2) &= \varepsilon^2 r - (1+\varepsilon)^2\frac{\varepsilon^2}{(1+\varepsilon)^2}r = 0.\end{aligned}$$

Hence, the equality in the bound (2.10) can be achieved.

In order to compare (2.9) with (1.10), consider

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1+\varepsilon & 0 \\ 0 & \varepsilon \\ 0 & 0 \end{pmatrix}, \tag{2.11}$$

where $0 < \varepsilon \le 1$. Then

$$EA^\dagger = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad A^\dagger E = \begin{pmatrix} \varepsilon & 0 \\ 0 & 0 \end{pmatrix}, \quad B^\dagger E = \begin{pmatrix} \frac{\varepsilon}{1+\varepsilon} & 0 \\ 0 & 1 \end{pmatrix}, \quad EB^\dagger = \begin{pmatrix} \frac{\varepsilon}{1+\varepsilon} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

The bound (1.10) is

$$\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2 = 1 + \varepsilon^2 + \left(\frac{\varepsilon}{1+\varepsilon}\right)^2.$$

The bound (2.9) is

$$\frac{1}{\tilde{\sigma}_s^2}(\|E\|_F^2 - \sigma_r^2\|A^\dagger E\|_F^2) + \frac{1}{\sigma_r^2}(\|E\|_F^2 - \tilde{\sigma}_s^2\|B^\dagger E\|_F^2) = 1 + \varepsilon^2 - \frac{\varepsilon^4}{(1+\varepsilon)^2},$$

which shows that the bound (2.9) is sharper than the one in (1.10). For more general case, we will give some comparison in Section 5.

For the perturbation with $rank(A) = rank(B)$, the combined bound of the orthogonal projection on $rang(A)$ and $rang(A^*)$ was given in [2]. However, for general perturbation, no combined bound has been given so far. Next, we present the combined perturbation bounds.

Theorem 2.5 *Let $A \in C_r^{m\times n}$ and $B = A + E \in C_s^{m\times n}$. Then*

$$\begin{aligned}&\min\{\tilde{\sigma}_s^2, \sigma_r^2\}(\|P_{B^*} - P_{A^*}\|_F^2 + \|P_B - P_A\|_F^2)\\ &\le 2\|E\|_F^2 - \sigma_r^2\tilde{\sigma}_s^2(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2).\end{aligned} \tag{2.12}$$

If $r = s$, then we have

$$\begin{aligned}&(\tilde{\sigma}_r^2 + \sigma_r^2)(\|P_{B^*} - P_{A^*}\|_F^2 + \|P_B - P_A\|_F^2)\\ &\le 4\|E\|_F^2 - 2\sigma_r^2\tilde{\sigma}_s^2(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2).\end{aligned} \tag{2.13}$$

Proof By (1.5) and (1.3), we have

$$\begin{aligned}B^\dagger E A^\dagger &= \tilde{V}_1\tilde{V}_1^* V_1\Sigma_1^{-1}U_1^* - \tilde{V}_1\tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_1 U_1^*,\\ \tilde{V}^* B^\dagger E A^\dagger U &= \begin{pmatrix}\tilde{V}_1^* V_1\Sigma_1^{-1} - \tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_1 & 0\\ 0 & 0\end{pmatrix}.\end{aligned}$$

Hence, we have

$$\begin{aligned}\|B^\dagger E A^\dagger\|_F^2 &= \|\tilde{V}_1^* V_1\Sigma_1^{-1} - \tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_1\|_F^2\\ &= \|\tilde{\Sigma}_1^{-1}(\tilde{\Sigma}_1\tilde{V}_1^* V_1 - \tilde{U}_1^* U_1\Sigma_1)\Sigma_1^{-1}\|_F^2\\ &\le \sigma_r^{-2}\tilde{\sigma}_s^{-2}\|\tilde{\Sigma}_1\tilde{V}_1^* V_1 - \tilde{U}_1^* U_1\Sigma_1\|_F^2.\end{aligned} \tag{2.14}$$

By (2.6) and (2.14), we have

$$\begin{aligned}\|E\|_F^2 &= \|\tilde{\Sigma}_1\tilde{V}_1^* V_1 - \tilde{U}_1^* U_1\Sigma_1\|_F^2 + \|\tilde{U}_2^* U_1\Sigma_1\|_F^2 + \|\tilde{\Sigma}_1\tilde{V}_1^* V_2\|_F^2\\ &\ge \sigma_r^2\tilde{\sigma}_s^2\|B^\dagger E A^\dagger\|_F^2 + \sigma_r^2\|\tilde{U}_2^* U_1\|_F^2 + \tilde{\sigma}_s^2\|\tilde{V}_1^* V_2\|_F^2.\end{aligned}$$

So

$$\sigma_r^2\|\tilde{U}_2^* U_1\|_F^2 + \tilde{\sigma}_s^2\|\tilde{V}_1^* V_2\|_F^2 \le \|E\|_F^2 - \sigma_r^2\tilde{\sigma}_s^2\|B^\dagger E A^\dagger\|_F^2.$$

Similarly, we have

$$\tilde{\sigma}_s^2\|U_2^*\tilde{U}_1\|_F^2 + \sigma_r^2\|V_1^*\tilde{V}_2\|_F^2 \le \|E\|_F^2 - \sigma_r^2\tilde{\sigma}_s^2\|A^\dagger E B^\dagger\|_F^2.$$

By Lemma 2.2, we have

$$\begin{aligned}&\min\{\tilde{\sigma}_s^2, \sigma_r^2\}(\|P_{B^*} - P_{A^*}\|_F^2 + \|P_B - P_A\|_F^2)\\ &\le 2\|E\|_F^2 - \sigma_r^2\tilde{\sigma}_s^2(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2),\end{aligned}$$

which implies the first bound of the theorem.

If $r = s$, then by Lemma 2.2, we have

$$\begin{aligned}&(\tilde{\sigma}_r^2 + \sigma_r^2)(\|P_{B^*} - P_{A^*}\|_F^2 + \|P_B - P_A\|_F^2)\\ &\le 4\|E\|_F^2 - 2\sigma_r^2\tilde{\sigma}_s^2(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2),\end{aligned}$$

which implies the second bound of the theorem. This proves the theorem. □

Remark 2.2 It is easy to see that the bound (2.13) in Theorem 2.5 always improves the one (1.12) given in [2]. On the other hand, the equality in (2.13) can be achieved. In fact, let A and B be unitary matrices. Then the equality of the bound (2.13) holds. Furthermore, we get the new combined bound (2.12) which does not impose restriction on rank.

3 The perturbation bound for generalized inverses

In this section, we use the bound (2.12) given in Section 2 to get a new perturbation bound for generalized inverses.

Theorem 3.1 *Let $A \in C_r^{m\times n}$ and $B = A + E \in C_s^{m\times n}$. Then*

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\min\{\sigma_r^4, \tilde{\sigma}_s^4\}}\|E\|_F^2 - \frac{1}{2}\left(\frac{\max\{\sigma_r^2, \tilde{\sigma}_s^2\}}{\min\{\sigma_r^2, \tilde{\sigma}_s^2\}} - 1\right)(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2). \quad (3.1)$$

Proof By the first equation of (2.14), we have

$$\begin{aligned}\|B^\dagger E A^\dagger\|_F^2 &= \|\tilde{V}_1^* V_1 \Sigma_1^{-1} - \tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_1\|_F^2 \\ &= \|\tilde{V}_1^*(B^\dagger - A^\dagger)U_1\|_F^2.\end{aligned}$$

Clearly, we have

$$\begin{aligned}\|\tilde{V}_1^*(B^\dagger - A^\dagger)U_2\|_F^2 &= \|\tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_2\|_F^2 \\ \|\tilde{V}_2^*(B^\dagger - A^\dagger)U_1\|_F^2 &= \|\tilde{V}_2^* V_1 \Sigma_1^{-1}\|_F^2 \\ \|\tilde{V}_2^*(B^\dagger - A^\dagger)U_2\|_F^2 &= 0.\end{aligned}$$

Hence, it follows that

$$\begin{aligned}\|B^\dagger - A^\dagger\|_F^2 &= \|B^\dagger E A^\dagger\|_F^2 + \|\tilde{\Sigma}_1^{-1}\tilde{U}_1^* U_2\|_F^2 + \|\tilde{V}_2^* V_1 \Sigma_1^{-1}\|_F^2 \\ &\le \|B^\dagger E A^\dagger\|_F^2 + \frac{1}{\tilde{\sigma}_s^2}\|\tilde{U}_1^* U_2\|_F^2 + \frac{1}{\sigma_r^2}\|\tilde{V}_2^* V_1\|_F^2.\end{aligned}$$

Similarly, we have

$$\|B^\dagger - A^\dagger\|_F^2 \le \|A^\dagger E B^\dagger\|_F^2 + \frac{1}{\sigma_r^2}\|U_1^*\tilde{U}_2\|_F^2 + \frac{1}{\tilde{\sigma}_s^2}\|V_2^*\tilde{V}_1\|_F^2.$$

Hence, we have

$$\begin{aligned}2\sigma_r^2\tilde{\sigma}_s^2\|B^\dagger - A^\dagger\|_F^2 \le\ & \sigma_r^2\tilde{\sigma}_s^2(\|B^\dagger E A^\dagger\|_F^2 + \|A^\dagger E B^\dagger\|_F^2) \\ & + \max\{\sigma_r^2, \tilde{\sigma}_s^2\}[\|U_1^*\tilde{U}_2\|_F^2 + \|\tilde{U}_1^* U_2\|_F^2 \\ & + \|V_2^*\tilde{V}_1\|_F^2 + \|\tilde{V}_2^* V_1\|_F^2].\end{aligned}$$

which together with (2.12) gives

$$\begin{aligned}2\sigma_r^2\tilde{\sigma}_s^2\|B^\dagger - A^\dagger\|_F^2 \le\ & \sigma_r^2\tilde{\sigma}_s^2(\|B^\dagger E A^\dagger\|_F^2 + \|A^\dagger E B^\dagger\|_F^2) \\ & + \frac{\max\{\sigma_r^2, \tilde{\sigma}_s^2\}}{\min\{\sigma_r^2, \tilde{\sigma}_s^2\}}[2\|E\|_F^2 - \sigma_r^2\tilde{\sigma}_s^2(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2)].\end{aligned}$$

Consequently,

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{\|E\|_F^2}{\min\{\sigma_r^4, \tilde{\sigma}_s^4\}} - \frac{1}{2}\left(\frac{\max\{\sigma_r^2, \tilde{\sigma}_s^2\}}{\min\{\sigma_r^2, \tilde{\sigma}_s^2\}} - 1\right)(\|A^\dagger E B^\dagger\|_F^2 + \|B^\dagger E A^\dagger\|_F^2),$$

which proves the theorem. □

Next we give the perturbation bounds of the generalized inverse for perturbation of a full rank matrix A.

Theorem 3.2 *Let $A \in C_m^{m\times n}(m \le n)$, and $B = A + E \in C_r^{m\times n}$. Then*

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\sigma_m^2 + \tilde{\sigma}_r^2}\left[\|A^\dagger E\|_F^2 + \|B^\dagger E\|_F^2 + \frac{\tilde{\sigma}_r^2}{\sigma_m^2}(m - r)\right]. \tag{3.2}$$

Furthermore, if $B \in C_m^{m\times n}(m \le n)$, then

$$\|B^\dagger - A^\dagger\|_F \le \min\{\|B^\dagger\|_2\|A^\dagger E\|_F, \|A^\dagger\|_2\|B^\dagger E\|_F\}. \tag{3.3}$$

If $A \in C_n^{m\times n}$, then

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\sigma_n^2 + \tilde{\sigma}_r^2}\left[\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2 + \frac{\tilde{\sigma}_r^2}{\sigma_n^2}(n - r)\right]. \tag{3.4}$$

In addition, if $rank(B) = n$, then

$$\|B^\dagger - A^\dagger\|_F \le \min\{\|B^\dagger\|_2\|EA^\dagger\|_F, \|A^\dagger\|_2\|EB^\dagger\|_F\}. \tag{3.5}$$

Proof Let $rank(B) = r \le m$. By (1.2), we have

$$A = U\Sigma_1 V_1^*, \ B = \tilde{U}_1\tilde{\Sigma}_1\tilde{V}_1^*.$$

By (1.3), we have

$$\tilde{V}^*(B^\dagger - A^\dagger)U\Sigma_1 = \begin{pmatrix} \tilde{\Sigma}_1^{-1}\tilde{U}_1^*U\Sigma_1 - \tilde{V}_1^*V_1 \\ -\tilde{V}_2^*V_1 \end{pmatrix}.$$

Then, we have

$$\|(B^\dagger - A^\dagger)U\Sigma_1\|_F^2 = \|\tilde{\Sigma}_1^{-1}\tilde{U}_1^*U\Sigma_1 - \tilde{V}_1^*V_1\|_F^2 + \|\tilde{V}_2^*V_1\|_F^2.$$

Hence,

$$\|\tilde{\Sigma}_1^{-1}\tilde{U}_1^*U\Sigma_1 - \tilde{V}_1^*V_1\|_F^2 + \|\tilde{V}_2^*V_1\|_F^2 \ge \sigma_m^2\|B^\dagger - A^\dagger\|_F^2. \tag{3.6}$$

A simple computation gives

$$\tilde{V}^*B^\dagger EV = \begin{pmatrix} \tilde{V}_1^*V_1 - \tilde{\Sigma}_1^{-1}\tilde{U}_1^*U\Sigma_1 & \tilde{V}_1^*V_2 \\ 0 & 0 \end{pmatrix},$$

which together with (3.6) gives

$$\begin{aligned}\|B^\dagger E\|_F^2 &= \|\tilde{\Sigma}_1^{-1}\tilde{U}_1^*U\Sigma_1 - \tilde{V}_1^*V_1\|_F^2 + \|\tilde{V}_1^*V_2\|_F^2 \\ &\ge \sigma_m^2\|(B^\dagger - A^\dagger)\|_F^2 + \|\tilde{V}_1^*V_2\|_F^2 - \|\tilde{V}_2^*V_1\|_F^2. \end{aligned} \tag{3.7}$$

Similarly, we have

$$V^*(B^\dagger - A^\dagger)\tilde{U}_1\tilde{\Sigma}_1 = \begin{pmatrix} V_1^*\tilde{V}_1 - \Sigma_1^{-1}U^*\tilde{U}_1\tilde{\Sigma}_1 \\ V_2^*\tilde{V}_1 \end{pmatrix},$$

and then

$$\|V_1^*\tilde{V}_1 - \Sigma_1^{-1}U^*\tilde{U}_1\tilde{\Sigma}_1\|_F^2 + \|V_2^*\tilde{V}_1\|_F^2 \ge \tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_1\|_F^2. \tag{3.8}$$

By similar arguments as above, we may obtain

$$\|A^\dagger E\|_F^2 = \|\Sigma_1^{-1}U^*\tilde{U}_1\tilde{\Sigma}_1 - V_1^*\tilde{V}_1\|_F^2 + \|V_1^*\tilde{V}_2\|_F^2,$$

which together with (3.8) gives

$$\|A^\dagger E\|_F^2 \ge \tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_1\|_F^2 + \|V_1^*\tilde{V}_2\|_F^2 - \|V_2^*\tilde{V}_1\|_F^2$$

Since

$$\begin{aligned}
\tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_1\|_F^2 &= \tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_1\|_F^2 + \tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_2\|_F^2 \\
&\quad -\tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_2\|_F^2 \\
&= \tilde{\sigma}_r^2\|B^\dagger - A^\dagger\|_F^2 - \tilde{\sigma}_r^2\|(B^\dagger - A^\dagger)\tilde{U}_2\|_F^2 \\
&= \tilde{\sigma}_r^2\|B^\dagger - A^\dagger\|_F^2 - \tilde{\sigma}_r^2\|\Sigma_1^{-1}U^*\tilde{U}_2\|_F^2 \\
&\ge \tilde{\sigma}_r^2\|B^\dagger - A^\dagger\|_F^2 - \frac{\tilde{\sigma}_r^2}{\sigma_m^2}(m-r),
\end{aligned}$$

we have

$$\|A^\dagger E\|_F^2 \ge \tilde{\sigma}_r^2\|B^\dagger - A^\dagger\|_F^2 - \frac{\tilde{\sigma}_r^2}{\sigma_m^2}(m-r) + \|V_1^*\tilde{V}_2\|_F^2 - \|V_2^*\tilde{V}_1\|_F^2. \tag{3.9}$$

Hence, by (3.7) and (3.9), we have

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\sigma_m^2 + \tilde{\sigma}_r^2}[\|A^\dagger E\|_F^2 + \|B^\dagger E\|_F^2 + \frac{\tilde{\sigma}_r^2}{\sigma_m^2}(m-r)].$$

which proves the bound (3.2).

Furthermore, if $rank(B) = m$, then by Lemma 2.1, we have $\|\tilde{V}_1^* V_2\|_F^2 = \|\tilde{V}_2^* V_1\|_F^2$. It follows from (3.7) that

$$\|B^\dagger - A^\dagger\|_F \le \|A^\dagger\|_2\|B^\dagger E\|_F. \tag{3.10}$$

Interchanging the role of A and B in the above inequality gives

$$\|B^\dagger - A^\dagger\|_F \le \|B^\dagger\|_2\|A^\dagger E\|_F,$$

which together with (3.12) yields (3.3).

The proofs of (3.4) and (3.5) are analogical. This proves the theorem. □

The following corollary follows immediately from the proof of Theorem 3.2.

Corollary 3.3 *Let $A \in C_m^{m\times n}$, and $B = A + E$ be any perturbed matrix of A. Then*

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{\|A^\dagger E\|_F^2 + \|B^\dagger E\|_F^2}{\sigma_m^2}. \tag{3.11}$$

If $A \in C_n^{m\times n}$, then

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{\|EA^\dagger\|_F^2 + \|EB^\dagger\|_F^2}{\sigma_n^2}. \tag{3.12}$$

By Theorem 3.2 it is easy to get a relative perturbation bound for the generalized inverse.

Theorem 3.4 *Let $A,\ B \in C^{m\times n}$, and $B = A + E$. If $rank(A) = rank(B) = \min\{m, n\}$, then*

$$\frac{\|B^\dagger - A^\dagger\|_F}{\|A^\dagger\|_F} \le \|B^\dagger\|_2 \|E\|_2. \tag{3.13}$$

Remark 3.1 The bound (3.13) can reduce to Theorem 1.1 of [8] when both A and B are nonsingular matrices. Clearly, the new bound in (3.1) is sharper than the corresponding one in (1.13). However, for the case that $r = s < \min\{m, n\}$, we only obtain the following bound by the same technique as in Theorem 3.1:

$$\|B^\dagger - A^\dagger\|_F \le \|A^\dagger\|_2 \|B^\dagger\|_2 \|E\|_F, \tag{3.14}$$

which is the same as the existing bound (1.14). But for $r = s = \min\{m, n\}$, from (3.2) and (3.4), we can derive the bound

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\sigma_m^2 + \tilde{\sigma}_m^2}[\|A^\dagger E\|_F^2 + \|B^\dagger E\|]_F^2$$

or

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\sigma_n^2 + \tilde{\sigma}_n^2}[\|EA^\dagger\|_F^2 + \|EB^\dagger\|]_F^2.$$

These bounds are smaller than (3.14), but larger than (3.3) and (3.5), respectively.

The following examples show the difference of the new bound in Theorem 3.1 and the existing bounds. Let

$$A = \begin{pmatrix} U & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ B = \begin{pmatrix} (1+\varepsilon)U & 0 & 0 \\ 0 & \varepsilon V & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $0 < \varepsilon < 1$, U and V are $r \times r$ unitary matrices. Then

$$\|B^\dagger - A^\dagger\|_F^2 = \left(\frac{\varepsilon^2}{(1+\varepsilon)^2} + \frac{1}{\varepsilon^2}\right) r$$

By the bound (3.1), we have

$$\|B^\dagger - A^\dagger\|_F^2 \le \left(\frac{2}{\varepsilon^2} + \frac{\varepsilon^2}{(1+\varepsilon)^2} - 1\right) r.$$

The bound (1.13) is

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{2r}{\varepsilon^2}.$$

Now, we give another example. Let A and B be given by (2.11). Then

$$B^\dagger E A^\dagger = A^\dagger E B^\dagger = \begin{pmatrix} \frac{\varepsilon}{1+\varepsilon} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

By the bound (3.1), we get

$$\|B^\dagger - A^\dagger\|_2^2 \le \frac{2}{\varepsilon^2} - \frac{1-\varepsilon}{1+\varepsilon},$$

while by the bound (1.13), we have

$$\|B^\dagger - A^\dagger\|_2^2 \le \frac{2}{\varepsilon^2}.$$

Remark 3.2 For A and its perturbed matrix B being full rank matrices, it is noted that the bounds (3.3) and (3.5) are always sharper than the one in (3.14). The following example shows the difference of (3.3) and (3.14). Let

$$A = \begin{pmatrix} U & 0 & 0 \\ 0 & V & 0 \end{pmatrix}, \quad B = \begin{pmatrix} (1+\varepsilon)U & 0 & 0 \\ 0 & \varepsilon V & 0 \end{pmatrix},$$

where $0 < \varepsilon < 1$, U and V are $r \times r$ unitary matrices. Then by (3.14), we have

$$\|B^\dagger - A^\dagger\|_F^2 \le \left(1 + \frac{(\varepsilon-1)^2}{\varepsilon^2}\right) r.$$

By (3.3), we have

$$\|B^\dagger - A^\dagger\|_F^2 \left(\frac{\varepsilon^2}{(1+\varepsilon)^2} + \frac{(\varepsilon-1)^2}{\varepsilon^2}\right) r,$$

which shows that the bound (3.3) is sharper.

For the perturbation of a full rank matrix, it is difficult to compare the perturbation bound for the generalized inverse in theory. However, the following example shows that the new bounds in Theorem 3.2 is the best. Let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where $0 < \varepsilon \le 1$. Then $\|B^\dagger - A^\dagger\|_F^2 = 1 + \frac{(1-\varepsilon)^2}{\varepsilon^2}$.

By (1.13), we have

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\varepsilon^4}[1 + (1-\varepsilon)^2].$$

The bound (3.1) gives

$$\|B^\dagger - A^\dagger\|_F^2 \le \frac{1}{\varepsilon^4} + \frac{(\varepsilon-1)^2}{\varepsilon^2}$$

and (3.2) is

$$\begin{aligned} \|B^\dagger - A^\dagger\|_F^2 &\le 1 + \frac{(1-\varepsilon)^2}{1+\varepsilon^2}(1 + \frac{1}{\varepsilon^2}) \\ &= 1 + \frac{(1-\varepsilon)^2}{\varepsilon^2}. \end{aligned}$$

Finally, it follows from (3.11) that

$$\|B^{\dagger} - A^{\dagger}\|_F^2 \leq 1 + (1-\varepsilon)^2 + \frac{(1-\varepsilon)^2}{\varepsilon^2}.$$

For small ε, it is easy to see that

$$(3.2) < (3.11) < (3.1) < (1.13),$$

where we have used the equation numbers to represent the corresponding bounds. This example also shows that the equality of the bound (3.2) can be achieved. Further numerical comparison of the proposed bounds and existing bounds will be given in Section 5.

Remark 3.3 The following example shows that (3.3) and (3.5) cannot hold in the case $r = s < \min\{m, n\}$, let

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

then $E = B - A$, $A^{\dagger} = A$, $B^{\dagger} = B$, and

$$A^{\dagger}E = EA^{\dagger} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, B^{\dagger}E = EB^{\dagger} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence, $\|B^{\dagger} - A^{\dagger}\|_F = \sqrt{2}$, but

$$\min\{\|B^{\dagger}\|_2\|A^{\dagger}E\|_F, \|A^{\dagger}\|_2\|B^{\dagger}E\|_F\} = \min\{\|B^{\dagger}\|_2\|EA^{\dagger}\|_F, \|A^{\dagger}\|_2\|EB^{\dagger}\|_F\} = 1,$$

which contradicts to those in (3.3) and (3.5).

4 Combined perturbation bounds for the orthogonal projection and the generalized inverse

The idea of the combined bound was first used in [4, 5]. Following this idea, we consider combined perturbation bounds for the orthogonal projection and the generalized inverse. The main result of this section is given below.

Theorem 4.1 *Let $A \in C_r^{m\times n}$ and $B = A + E \in C_r^{m\times n}$. Then*

$$2\tilde{\sigma}_r^2\sigma_r^2\|P_{B^*}(B^{\dagger} - A^{\dagger})P_A\|_F^2 + \tilde{\sigma}_r^2\|P_{B^*} - P_{A^*}\|_F^2 + \sigma_r^2\|P_B - P_A\|_F^2 \leq 2\|E\|_F^2$$

or

$$2\tilde{\sigma}_r^2\sigma_r^2\|P_{A^*}(B^{\dagger} - A^{\dagger})P_B\|_F^2 + \sigma_r^2\|P_{B^*} - P_{A^*}\|_F^2 + \tilde{\sigma}_r^2\|P_B - P_A\|_F^2 \leq 2\|E\|_F^2.$$

Proof By (1.2) and (1.3), we have

$$\begin{aligned} B(B^{\dagger} - A^{\dagger})A &= \tilde{U}_1\tilde{U}_1^*U_1\Sigma_1V_1^* - \tilde{U}_1\tilde{\Sigma}_1\tilde{V}_1^*V_1V_1^* \\ &= \tilde{U}_1(\tilde{U}_1^*U_1\Sigma_1 - \tilde{\Sigma}_1\tilde{V}_1^*V_1)V_1^*. \end{aligned}$$

Then

$$\tilde{U}^* B(B^\dagger - A^\dagger)AV = \begin{pmatrix} \tilde{U}_1^* U_1 \Sigma_1 - \tilde{\Sigma}_1 \tilde{V}_1^* V_1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{4.1}$$

Since

$$\begin{aligned} \tilde{U}_1^* U_1 \Sigma_1 - \tilde{\Sigma}_1 \tilde{V}_1^* V_1 &= \tilde{U}_1^* (U_1 \Sigma_1 V_1^* - \tilde{U}_1 \tilde{\Sigma}_1 \tilde{V}_1^*) V_1 \\ &= -\tilde{U}_1^* E V_1, \end{aligned}$$

by (4.1), we have

$$\|B(B^\dagger - A^\dagger)A\|_F^2 = \|\tilde{U}_1^* E V_1\|_F^2.$$

By (2.6), we have

$$\tilde{U}_2^* E V_1 = -\tilde{U}_2^* U_1 \Sigma_1, \;\; \tilde{U}_1^* E V_2 = \tilde{\Sigma}_1 \tilde{V}_1^* V_2 \text{ and} \tilde{U}_2^* E V_2 = 0.$$

Hence, we get

$$\|B(B^\dagger - A^\dagger)A\|_F^2 + \|\tilde{U}_2^* U_1 \Sigma_1\|_F^2 + \|\tilde{\Sigma}_1 \tilde{V}_1^* V_2\|_F^2 = \|\tilde{U}^* E V\|_F^2. \tag{4.2}$$

A simple computation gives that

$$\begin{aligned} \|B(B^\dagger - A^\dagger)A\|_F &= \|\tilde{U}_1 \tilde{\Sigma}_1 \tilde{V}_1^* (B^\dagger - A^\dagger) U_1 \Sigma_1 V_1^*\|_F \\ &= \|\tilde{\Sigma}_1 \tilde{V}_1^* (B^\dagger - A^\dagger) U_1 \Sigma_1\|_F \\ &\geq \tilde{\sigma}_r \sigma_r \|\tilde{V}_1^* (B^\dagger - A^\dagger) U_1\|_F \end{aligned}$$

and

$$\begin{aligned} \|\tilde{V}_1^* (B^\dagger - A^\dagger) U_1\|_F^2 &= tr(U_1^* (B^\dagger - A^\dagger)^* \tilde{V}_1 \tilde{V}_1^* (B^\dagger - A^\dagger) U_1) \\ &= tr((B^\dagger - A^\dagger)^* \tilde{V}_1 \tilde{V}_1^* (B^\dagger - A^\dagger) U_1 U_1^*) \\ &= tr((B^\dagger - A^\dagger)^* P_{B^*} (B^\dagger - A^\dagger) P_A) \\ &= tr(P_A (B^\dagger - A^\dagger)^* P_{B^*} P_{B^*} (B^\dagger - A^\dagger) P_A) \\ &= \|P_{B^*} (B^\dagger - A^\dagger) P_A\|_F^2, \end{aligned}$$

which together with (4.2) gives

$$\|P_{B^*} (B^\dagger - A^\dagger) P_A\|_F^2 + \sigma_r^2 \|\tilde{U}_2^* U_1\|_F^2 + \tilde{\sigma}_r^2 \|\tilde{V}_1^* V_2\|_F^2 \leq \|E\|_F^2. \tag{4.3}$$

It follows from Lemma 2.2 that the first bound in Theorem 4.1 holds. The second bound is given by interchanging the role of A and B in the just given bound. This proves the theorem. □

5 Numerical results

In order to show the differences between the new bounds and the existing bounds, we give some numerical examples and plot their figures in the following two subsections.

5.1 Comparisons of the new bounds with the existing bounds for orthogonal projections

In this subsection, we will compare the new bound with the existing ones for orthogonal projections by some examples. The first example is for perturbation with different ranks.

Example 1 Let

$$A = U \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} V^* \in C^{m\times n},\ E = U \begin{pmatrix} I_{r+1} & 0 \\ 0 & 0 \end{pmatrix} V^* \in C^{m\times n},\ B = A + \epsilon E,$$

where U and V are unitary matrices.

Now taking $\epsilon = 0.1 : 0.1 : 1$, and $m = 3, n = 2, r = 1$, then we plot the figure as follows (Fig. 1):

In this example, we can see that the bound (2.9) is sharper than the corresponding ones in (1.8) and (1.10). The next example is for the case of $rank(A) = rank(B)$.

Example 2 Let $A = U_1 \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V_1^* \in C^{5\times 4}$, $B_0 = U_2 \begin{pmatrix} \Sigma_2 \\ 0 \end{pmatrix} V_2^* \in C^{5\times 4}$, and $B = \epsilon B_0$, where Σ_1, Σ_2 are 4×4 positive diagonal matrices and U_1, U_2, V_1, V_2 are unitary matrices.

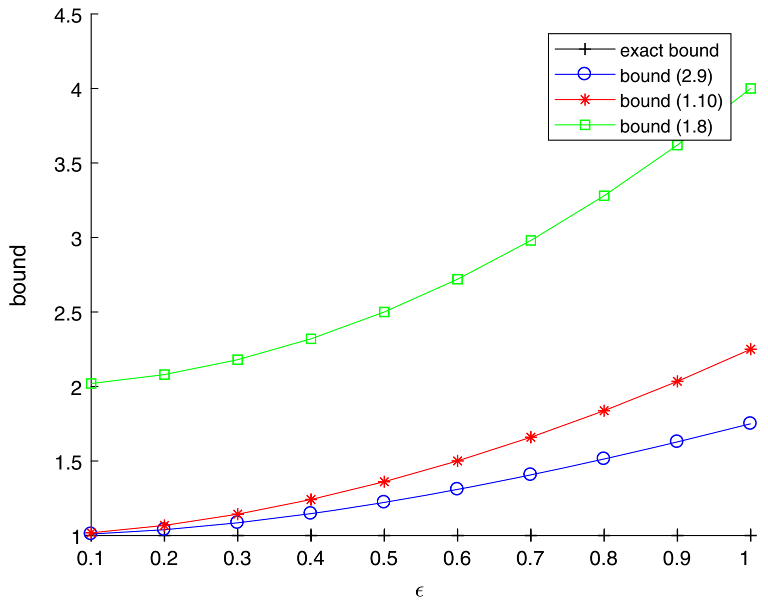

Fig. 1 Comparison of (2.9) with existing bounds

Let the diagonal entries of Σ_1, Σ_2 be taken by the uniform distribution in the interval $(1, 2)$, and let $\epsilon = 0.1 : 0.1 : 1$. Then we plot the figure as follows:

By Fig. 2, it is seen that the bound (2.10) is the best. This confirms the theoretical analysis in Remark 2.1.

Next, we take 100 matrix pairs A and B in $\mathbb{R}^{5\times 4}$ for comparing these bounds for the case of $rank(A) = rank(B)$.

Example 3 Let $A = U_1 \begin{pmatrix} \Sigma_1 \\ 0 \end{pmatrix} V_1^* \in C^{5\times 4}$, $B_0 = U_2 \begin{pmatrix} \Sigma_2 \\ 0 \end{pmatrix} V_2^* \in C^{5\times 4}$, and $B = \epsilon B_0$, where U_1, U_2, V_1, V_2 are taken unitary matrices randomly and Σ_1, Σ_2 are 4×4 diagonal matrices whose diagonal entries are generated independently by uniform distributions in $(1, 2)$.

Taking $\epsilon = 0.1 : 0.1 : 1$. For each given ϵ, we generate 100 matrix pairs A and B_0 by the method in Example 3. For those chosen 100 matrix pairs, we can get the best bounds among (1.9), (1.11), and (2.10) and denote the number of the best bound among three bounds by BN(*.*). Then we draw the figure below:

From Fig. 3, it is easy to see that the bound (2.10) is the best in the most cases (above 80 percent). Next, two examples are for different ranks.

Example 4 Let

$$A = U_1 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V_1^* \in C^{5\times 4}, B_0 = U_2 \begin{pmatrix} \Omega_2 \\ 0 \end{pmatrix} V_2^* \in C^{5\times 4},$$

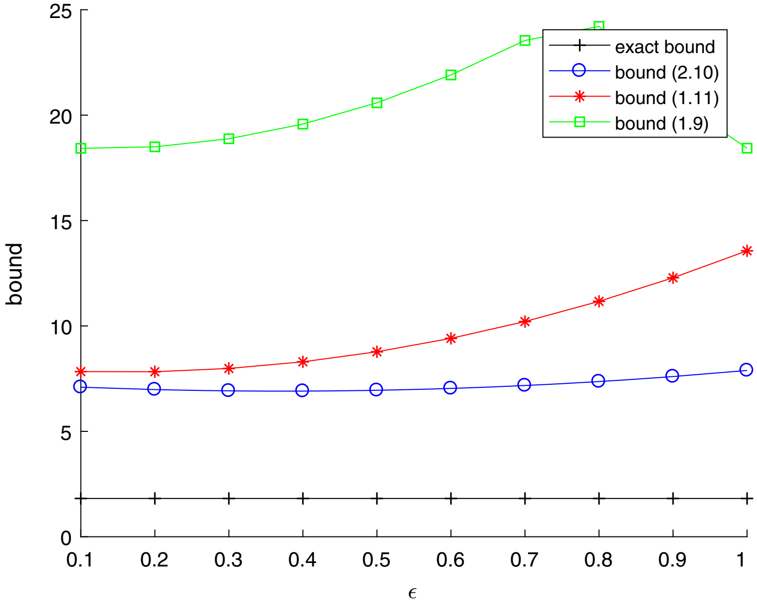

Fig. 2 Comparison of (2.10) with existing bounds

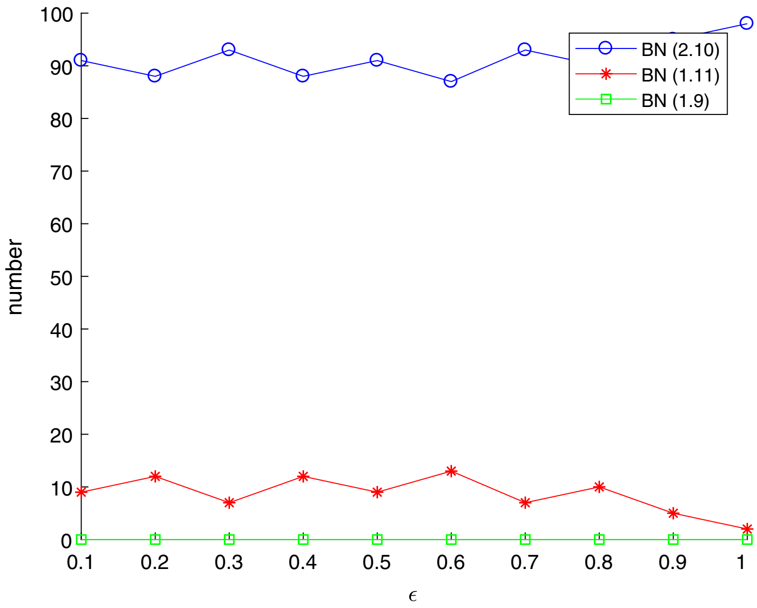

Fig. 3 Comparison of (2.10) with existing bounds by 100 matrix pairs

and $B = \epsilon B_0$, where U_1, U_2, V_1, V_2 are unitary matrices and Σ_1, Ω_2 are 2×2 and 4×4 positive diagonal matrices, respectively, whose diagonal entries are all generated independently by uniform distributions in the interval $(1, 2)$.

Taking $\epsilon = 0.1 : 0.1 : 1$, then we plot Fig. 4 for comparing bounds (1.8), (1.10), and (2.9) with the exact bound.

In Fig. 4, one can see the bound (2.9) is the closest to the exact bound.

Next we take 100 matrix pairs A and B in $\mathbb{R}^{5\times 4}$ for comparing bounds (1.8), (1.10), and (2.9) for the case of $rank(A) \neq rank(B)$.

Example 5 Let

$$A = U_1 \begin{pmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{pmatrix} V_1^* \in C^{5\times 4}, \; B_0 = U_2 \begin{pmatrix} \Omega_1 \\ 0 \end{pmatrix} V_2^* \in C^{5\times 4},$$

and $B = \epsilon B_0$, where U_1, U_2, V_1, V_2 are taken unitary matrices randomly and Σ_1, Ω_1 are 2×2 and 4×4 diagonal matrices whose diagonal entries are generated independently by uniform distributions in the interval $(1, 2)$, respectively.

Let $\epsilon = 0.1 : 0.1 : 1$. For each given ϵ, we generate 100 matrix pairs A and B_0 given by the method in Example 5. For those chosen 100 matrix pairs, we can get the best bounds among (1.8), (1.10), and (2.9) and denote the number of the best bound among three bounds by BN(*.*). Then we draw the figure below:

Figure 5 shows that the bound (2.9) is the best in most cases (above 70 percent).

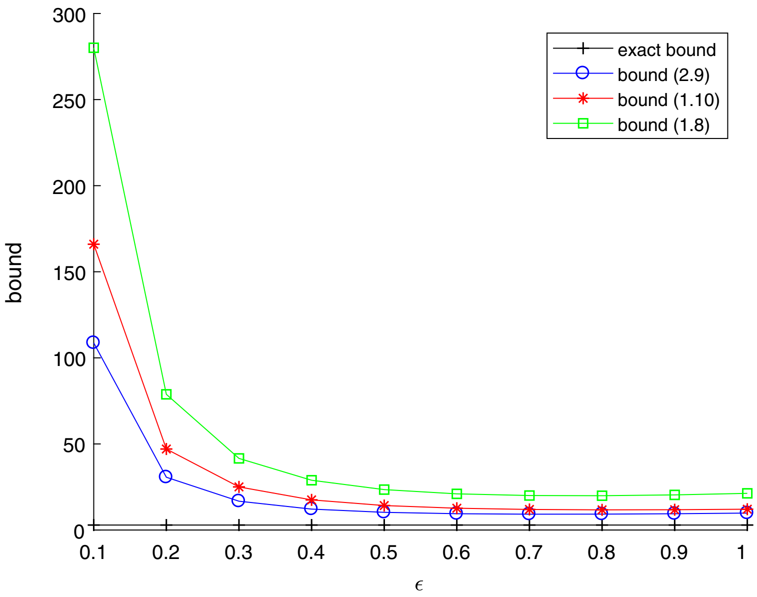

Fig. 4 Comparison of (2.9) with existing bounds for different ranks

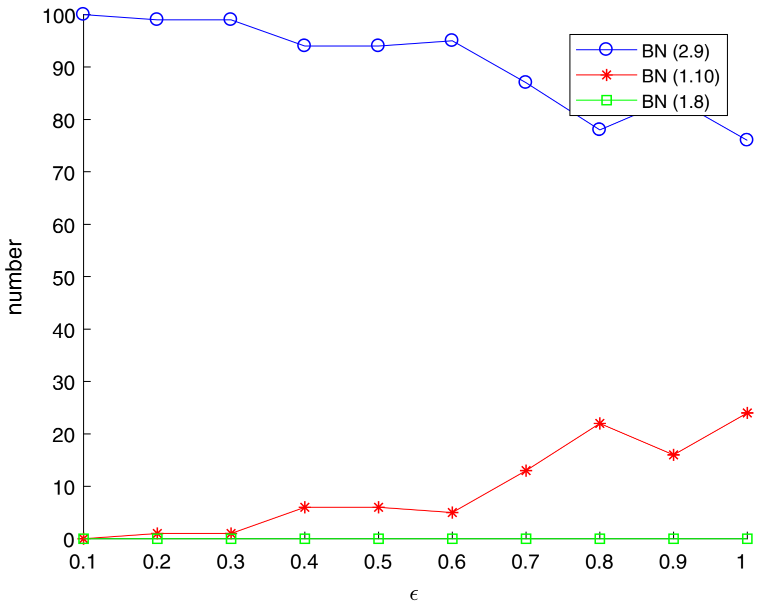

Fig. 5 Comparison of (2.9) with existing bounds by 100 random matrix pairs

5.2 Comparison of the new bounds with the existing bounds for generalized inverses

In this subsection, we will compare the proposed bound with the existing ones for generalized inverses by some numerical examples and see their difference from the exact bound.

Example 6 Let $D \in C^{4\times 4}$ be a diagonal matrix whose diagonal entries are generated independently by a uniform distribution on the interval (0, 1), and let D_2 be also a diagonal matrix whose diagonal entries are taken from the first 2 diagonal ones of D. Denote

$$A = U \begin{pmatrix} D & 0 \end{pmatrix} V^* \in C_4^{4\times 6}, B_0 = U \begin{pmatrix} D_2 & 0 \\ 0 & 0 \end{pmatrix} V^* \in C_2^{4\times 6},$$

and $B = \epsilon B_0$, where U and V are unitary matrices.

Let $\epsilon = 0.5 : 0.05 : 1$. Then we plot the following figures for comparing the bounds (1.13), (3.1), (3.2), and (3.17) with the exact bound and comparing the bound (3.1) with (1.13) for a perturbation of a full rank matrix A, respectively.

In Fig. 6, it is seen that the bound (3.2) is the closest to the exact bound, and the bound (3.1) is very close to the bound (1.13). So we plot Fig. 7, in which one can see that the bound (3.1) is also better than (1.13). For A and its perturbed matrix B being full rank rectangular matrices, we compare the bounds (3.3) and (1.14) with the exact bound in the following example.

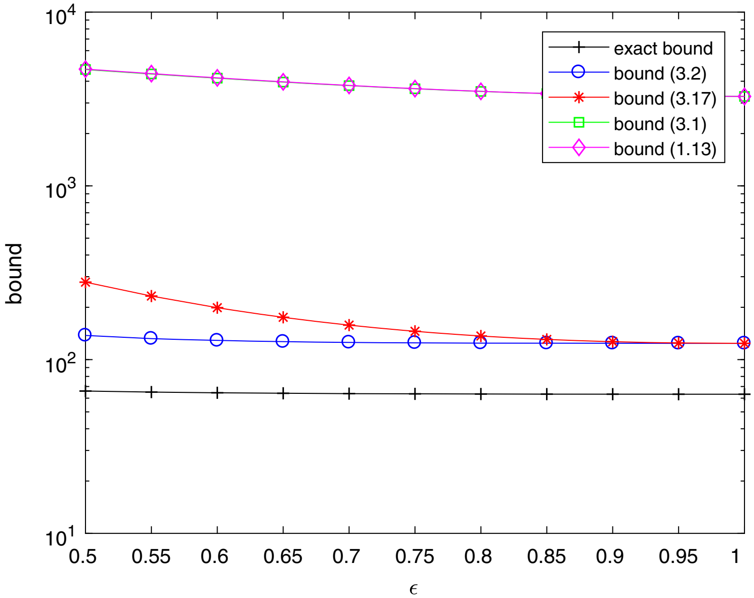

Fig. 6 Comparison of the proposed bounds with existing bounds

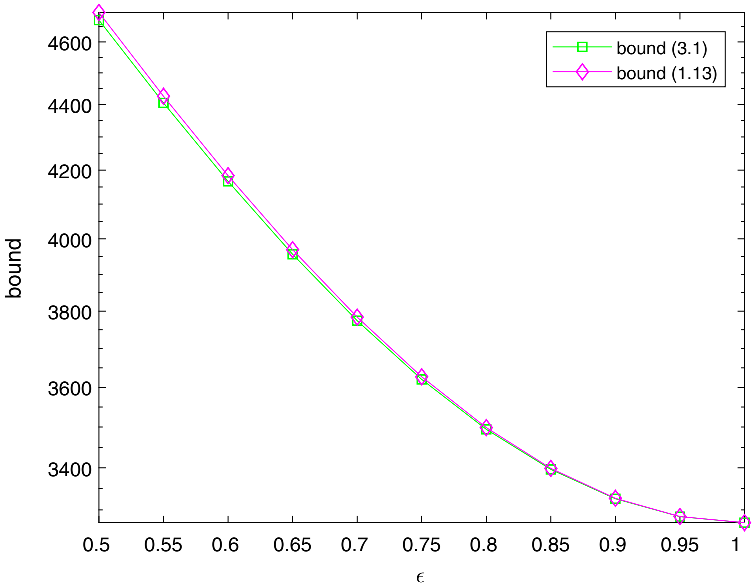

Fig. 7 Comparison of the bounds (3.1) with (1.13)

Example 7 Let

$$A = U \begin{pmatrix} D_1 & 0 \end{pmatrix} V^* \in C_4^{4\times 6}, B_0 = U \begin{pmatrix} D_2 & 0 \end{pmatrix} V^* \in C_4^{4\times 6},$$

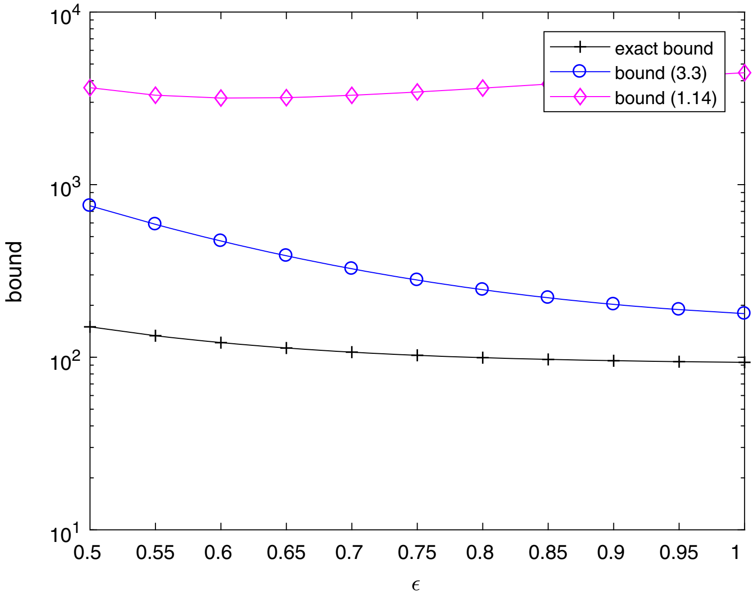

Fig. 8 Comparison bounds (1.14) and (3.3) with the exact bound

and $B = \epsilon B_0$, where U, V are unitary matrices and D_1, D_2 are 4×4 positive diagonal matrices whose diagonal entries are generated independently from uniform distribution on the interval (0, 1).

Taking $\epsilon = 0.5 : 0.05 : 1$, then we draw the figure below:

From Fig. 8, it is known that the bound (3.3) is better than (1.14) and is closed to the exact bound.

6 Concluding remarks

In the paper, we considered the perturbation bounds for the orthogonal projection and the generalized inverse, respectively. We have obtained the following results:

– The refined perturbation bounds for the orthogonal projection on $rang(A)$ and $rang(A^*)$, respectively
– Some new perturbation bounds for the generalized inverse
– A combined perturbation bound for the orthogonal projections on $rang(A)$ and $rang(A^*)$
– A combined perturbation bounds for the orthogonal projection and the generalized inverse

The new bounds improve the existing ones. Some numerical examples are given to justify the theoretical results.

In the future, we may employ the technique given in this paper to analyze the sensitivity of the generalized least square problems arisen from practical applications.

Acknowledgements The authors would like to thank the referees for their help comments.

References

1. Bhatia, R.: Matrix Analysis. Springer, New York (1997)
2. Chen, Y.M., Chen, X.S.: On perturbation bounds for orthogonal projections. Numer. Algorithm. **73**, 433–444 (2016)
3. Li, B.X., Li, W., Cui, L.B.: New bounds for perurbation of the orthogonal projection. Calcolo **50**, 69–78 (2013)
4. Li, W., Sun, W.W.: Combined perturbation bounds: I. Eigensystems and singular value decompositions. SIAM J. Matrix Anal. Appl. **29**(2), 543–655 (2007)
5. Li, W., Sun, W.W.: Combined perturbation bounds: II. Polar decompositions. Sci. China (Ser. A): Math. **50**(9), 1339–1346 (2007)
6. Meng, L.S., Zheng, B.: The optimal perturbation bounds of the Moore–Penrose inverse under the Frobenius norm. Linear Algebra Appl. **432**, 956–963 (2010)
7. Stewart, G.W.: On the perturbation of the pseudo-inverse, projections and linear squares problems. SIAM Rev. **19**, 634–662 (1977)
8. Sun, J.: Matrix Perturbation Analysis. 2nd edn. Science Press, Beijing (2001). (in Chinese)