Japan J. Indust. Appl. Math. (2014) 31:513–539
DOI 10.1007/s13160-014-0145-5

ORIGINAL PAPER **Area 2**

Verified bounds for all the singular values of matrix

Shinya Miyajima

Received: 14 May 2013 / Revised: 27 April 2014 / Published online: 17 May 2014

Abstract Algorithms for computing verified bounds of all singular values of matrix are proposed. It is shown that the proposed bounds are equal or tighter than previous bounds. The computational costs of the proposed algorithms are equal or smaller than those of the previous algorithms. Numerical results show the properties of the proposed algorithms. As an application of the proposed algorithms, algorithms for computing verified bounds of all generalized singular values are also sketched.

Keywords Verified bounds · Singular values of matrix · Generalized singular values

Mathematics Subject Classfication (2010) 15A18 · 65F15 · 65G20

1 Introduction

A matrix factorization having great importance in numerical linear algebra is the singular value decomposition (SVD), which is based on the following theorem:

Theorem 1 (E.g. Golub and Van Loan [4]) *Let $A \in \mathbb{R}^{m\times n}$ be given. There exist orthogonal $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ such thatx*

$$U^T AV = \Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_q), \quad q = \min(m, n),$$
$$\sigma_1 \geq \cdots \geq \sigma_{r^*} > \sigma_{r^*+1} = \cdots = \sigma_q = 0, \quad r^* = \mathrm{rank}(A).$$

This research was partially supported by Grant-in-Aid for Scientific Research (C) (23560066, 2011–2015) from the Ministry of Education, Science, Sports and Culture of Japan.

S. Miyajima (✉)
Faculty of Engineering, Gifu University, 1-1 Yanagido, Gifu-shi, Gifu 501-1193, Japan
e-mail: miyajima@gifu-u.ac.jp

The nonnegative real numbers σ_i, $i = 1, \ldots, q$ are called the singular values of A, which play important roles in application areas. It is well known that σ_i^2 are the eigenvalues of the symmetric pencils $A^T A - \lambda I_n$ and $AA^T - \lambda I_m$, where I_n denotes the $n \times n$ identity matrix. See [4] regarding to the definition of a pencil and its eigenvalues. The important issue is the computations for *few* of the largest or smallest singular values. On the other hand, the computation of *all* the singular values is also important. In fact, the computation of all the singular values appear in several applications, and the algorithms (e.g. [3,8,11]) for numerically computing all the singular values have been proposed.

In this paper, we consider computing verified bounds of all the singular values. Oishi [9] firstly proposed such an algorithm utilizing numerical full SVD. Recently Rump [14] proposed two algorithms. The first and second algorithms utilize the full SVD and numerical eigen-decomposition, respectively.

The purpose of this paper is to propose four algorithms for computing verified bounds of all the singular values. It is shown that the proposed first and second algorithms give equal or tighter bounds than those by the algorithm in [9] and the first algorithm in [14], respectively. It is proved that the proposed fourth algorithm yields equal or tighter bounds than those by the second algorithm in [14] when $m \geq n$. The computational costs of the proposed first, second and fourth algorithms are equal or smaller than those of the algorithm in [9], and the first and second algorithms in [14], respectively. The computational cost of the third algorithm is smaller than those of the proposed first and second algorithms. In the proposed first and third algorithms, numerical economy SVD is utilized instead of the full SVD. This enables us to reduce computational costs significantly when $m \gg n$ or $m \ll n$ (see Sects. 3, 4 for details).

As an application of the proposed algorithms, we describe algorithms for computing verified bounds for generalized singular values, which are defined by the following theorem:

Theorem 2 (Van Loan [15]) *Let $A \in \mathbb{R}^{m \times n}$ with $m \geq n$ and $B \in \mathbb{R}^{p \times n}$ be given. There exist orthogonal $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{p \times p}$ and a nonsingular $X \in \mathbb{R}^{n \times n}$ such that*

$$
\begin{aligned}
&U^T AX = \Sigma_A = \mathrm{diag}(c_1, \ldots, c_n),\ \ V^T BX = \Sigma_B = \mathrm{diag}(s_1, \ldots, s_q),\ \ q = \min(p, n),\\
&0 \leq c_1 \leq \cdots \leq c_n \leq 1,\ \ 1 \geq s_1 \geq \cdots \geq s_{r^*} > s_{r^*+1} = \cdots = s_q = 0,\ \ r^* = \mathrm{rank}(B),\\
&c_i^2 + s_i^2 = 1,\ \ i = 1, \ldots, q.
\end{aligned}
$$

This matrix factorization and the quotients $\mu_j = c_j/s_j$, $j = 1, \ldots, r^*$ are called the generalized singular value decomposition (GSVD) and the generalized singular values of A and B, respectively. Note that μ_j^2 are the eigenvalues of the symmetric pencil $A^T A - \lambda B^T B$ (see [4]). Although more general definition of the GSVD can be found in [12], in this paper, we define the GSVD by Theorem 2 for simplicity. The GSVD is a tool used in many applications, such as damped least squares, least squares with equality constraints, certain generalized eigenvalue problems and weighted least squares [15]. An algorithm for computing verified bounds of c_j, s_j and j-th columns of U, V and X for *specified* $j \in \{1, \ldots, r^*\}$ has been proposed in [1]. As far as the author knows, on the other hand, an algorithm giving verified bounds of μ_j for *all* $j = 1, \ldots, r^*$ has not been written down in literatures. We thus extend the proposed algorithms and sketch eight algorithms for this purpose. The former and latter four algorithms are applicable if $B^T B$ and $A^T A$ are nonsingular, respectively. We do not assume but prove these nonsingularities during the executions of these algorithms.

This paper is organized as follows: In Sect. 2, notations and theories utilized in this paper are introduced. In Sect. 3, theories for computing verified bounds of all the singular values are constructed. In Sect. 4, numerical results are reported. In Sect. 5, the theories in Sect. 3 are extended for computing verified bounds of all the generalized singular values. Let B^+ denote the pseudo-inverse of B. In Sect. 6, an algorithm for computing an upper bound of $\|B^+\|_2$, which is required in some of the algorithms based on the theories in Sect. 5, is introduced. In Sect. 7, numerical results for the generalized singular values are reported. Finally Sect. 8 summarizes the results in this paper and highlights possible extensions and future work.

2 Preliminaries

In this section, we introduce some notations and theories utilized hereafter. Let I_n and O_n be the $n \times n$ identity and zero matrices, respectively. For $M = (M_{ij}) \in \mathbb{R}^{m\times n}$, let $M^T := (M_{ji})$, $|M| := (|M_{ij}|)$, M^+ denote the pseudo-inverse of M and $M^{+T} := (M^T)^+$. If $m = n$, especially, let $M^{-T} := (M^T)^{-1}$ and $\mathrm{diag}(M) := (M_{11}, \ldots, M_{nn})^T$. For $v \in \mathbb{R}^n$, v_i denotes the i-th element of v. For $c, r \in \mathbb{R}$ where $r \geq 0$, $< c, r >$ denotes the interval whose center and radius are c and r, respectively. For $A \in \mathbb{R}^{m\times n}$ and $q = \min(m, n)$, let $\sigma_i(A)$, $i = 1, \ldots, q$ be the singular values of A such that $\sigma_1(A) \geq \cdots \geq \sigma_q(A)$. For $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, $B \in \mathbb{R}^{p\times n}$ and $r^* = \mathrm{rank}(B)$, let $\sigma_i(A, B)$, $i = 1, \ldots, r^*$ be the generalized singular values of A and B such that $\sigma_1(A, B) \geq \cdots \geq \sigma_{r^*}(A, B)$. For symmetric $A \in \mathbb{R}^{n\times n}$, let $\lambda_i(A)$, $i = 1, \ldots, n$ be the eigenvalues of A such that $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$. For symmetric $A \in \mathbb{R}^{n\times n}$, symmetric non-negative definite $B \in \mathbb{R}^{n\times n}$ and $r^* = \mathrm{rank}(B)$, let $\lambda_i(A, B)$, $i = 1, \ldots, r^*$ be the generalized eigenvalues of A and B such that $\lambda_1(A, B) \geq \cdots \geq \lambda_{r^*}(A, B)$. Denote a relative rounding error unit and an underflow constant by $\mathbf{u}$ and $\underline{\mathbf{u}}$, respectively. For IEEE754 double precision, we have $\mathbf{u} = 2^{-53}$ and $\underline{\mathbf{u}} = 2^{-1074}$. Let $e^{(n,i)}$, $i = 1, \ldots, n$ be the i-th column of I_n, $\gamma_n := n\mathbf{u}/(1 - n\mathbf{u})$ and $s^{(n)} := (1, \ldots, 1)^T \in \mathbb{R}^n$. We cite Lemmas 1 and 2, and Theorem 3.

Lemma 1 (Rump [14]) *Let $X \in \mathbb{R}^{m\times n}$ be given and $E := I_n - X^T X$. If $\|E\|_2 < 1$, then $m \geq n$, X has full rank, and*

$$\sqrt{1 - \|E\|_2} \leq \sigma_i(X) \leq \sqrt{1 + \|E\|_2} \quad \textit{and} \quad \frac{1}{\sqrt{1 + \|E\|_2}} \leq \sigma_i(X^+) \leq \frac{1}{\sqrt{1 - \|E\|_2}}$$

for $i = 1, \ldots, n$. In particular,

$$\sqrt{1 - \|E\|_2} \leq \|X\|_2 \leq \sqrt{1 + \|E\|_2} \quad \textit{and} \quad \frac{1}{\sqrt{1 + \|E\|_2}} \leq \|X^+\|_2 \leq \frac{1}{\sqrt{1 - \|E\|_2}}.$$

Lemma 2 (E.g. Horn and Johnson [6]) *Let A, $E \in \mathbb{R}^{m\times n}$ be given. Then $|\sigma_i(A+E)-\sigma_i(A)| \le \|E\|_2$ and $\sigma_i(AE^T) \le \sigma_i(A)\|E\|_2$ hold for $i=1,\dots,\min(m,n)$.*

Theorem 3 (E.g. Parlett [10]) *Let $A \in \mathbb{R}^{n\times n}$ be symmetric, $y \in \mathbb{R}^n$ be a unit vector, $\theta := y^T Ay$, α be the eigenvalue of A closest to θ, and $\mathrm{gap}(A,\theta) := \min|\lambda_i(A)-\theta|$ over all $\lambda_i(A) \ne \alpha$. Then it follows that*

$$|\theta-\alpha| \le \frac{\|Ay-\theta y\|_2^2}{\mathrm{gap}(A,\theta)}.$$

3 Verification theories

In this section, we refer previous theories for computing verified bounds of all the singular values and construct theories for this purpose. We cite Theorems 4, 5 and 6 regarding to the previous theories.

Theorem 4 (Oishi [9]) *Let $A \in \mathbb{R}^{m\times n}$, $U \in \mathbb{R}^{m\times m}$, $\Sigma \in \mathbb{R}^{m\times n}$ and $V \in \mathbb{R}^{n\times n}$ be given, Σ be diagonal with $\Sigma_{11} \ge \cdots \ge \Sigma_{qq} \ge 0$ where $q := \min(m,n)$, $E := U\Sigma V^T - A$, $F := V^TV - I_n$ and $G := U^TU - I_m$. If $\|F\|_2 < 1$ and $\|G\|_2 < 1$, $\underline{\delta}_i \le \sigma_i(A) \le \overline{\delta}_i$ holds for $i=1,\dots,q$, where*

$$\underline{\delta}_i := \Sigma_{ii} - \delta_i, \quad \overline{\delta}_i := \Sigma_{ii} + \delta_i, \quad \delta_i := \Sigma_{ii}\max(\|F\|_2, \|G\|_2) + \|E\|_2.$$

Remark 1 Comparing to the 1-norm and ∞-norm, it is not computationally efficient to obtain verified upper bounds of the 2-norm. It is well known (e.g. [4]) for $M \in \mathbb{R}^{m\times n}$ that $\|M\|_2 \le \sqrt{\|M\|_1\|M\|_\infty}$. If M is symmetric in particular, we have $\|M\|_2 \le \|M\|_\infty$. Thus in Sects. 4 and 7, the compared algorithms avoid the direct computation of the 2-norm by utilizing these properties.

Let $q := \min(m,n)$ and $Q := \max(m,n)$. The computational cost of the algorithm based on Theorem 4 is $4m^3 + 4m^2n + 8mn^2 + 13n^3 + 6Qq^2 + \mathcal{O}(m^2+n^2)$ divided into

$4m^2n + 8mn^2 + 9n^3$ full SVD to obtain U, Σ and V (c.f. [4, Sect. 5.4.5]),
$6Qq^2$ inclusion of $(U\Sigma)V^T$ when $m \ge n$, or $U(\Sigma V^T)$ when $m < n$,
$4m^3$ inclusion of U^TU,
$4n^3$ inclusion of V^TV.

Theorem 5 (Rump [14]) *Let A, U, V, F, G and q be as in Theorem 4. Define D, $E \in \mathbb{R}^{m\times n}$ so that $U^TAV = D + E$ and D is diagonal. If $\|F\|_2 < 1$ and $\|G\|_2 < 1$, then there is a numbering $\nu : \{1,\dots,q\} \to \{1,\dots,q\}$ with $\underline{\varepsilon}_i \le \sigma_{\nu(i)}(A) \le \overline{\varepsilon}_i$ for $i=1,\dots,q$, where*

$$\underline{\varepsilon}_i := \frac{|D_{ii}| - \|E\|_2}{\sqrt{(1+\|F\|_2)(1+\|G\|_2)}}, \quad \overline{\varepsilon}_i := \frac{|D_{ii}| + \|E\|_2}{\sqrt{(1-\|F\|_2)(1-\|G\|_2)}}.$$

Remark 2 Note that the definitions of D and E imply that $E_{ii} = 0, i = 1, \ldots, q$. The matrices E and $\hat{E}$ in Theorems 6 or 11, Remark 3, or Corollaries 2, 4, 6, or 8 have the similar property.

The computational cost of the algorithm based on Theorem 5 is $4m^3 + 4m^2n + 8mn^2 + 13n^3 + 4Q^2q + 6Qq^2 + \mathcal{O}(m^2 + n^2)$ divided into

$4m^2n + 8mn^2 + 9n^3$ the full SVD,
$4Q^2q + 6Qq^2$ inclusion of $(U^TA)V$ when $m \geq n$, or $U^T(AV)$ when $m < n$,
$4m^3$ the inclusion of U^TU,
$4n^3$ the inclusion of V^TV.

Theorem 6 (Rump [14]) *Let A, V, F and q be as in Theorem 4. Define D, $E \in \mathbb{R}^{n\times n}$ so that $(AV)^TAV = D + E$ and D is diagonal. If $\|F\|_2 < 1$, there is a numbering $\nu : \{1, \ldots, q\} \to \{1, \ldots, q\}$ with $\underline{\zeta}_i \leq \sigma_{\nu(i)}(A) \leq \overline{\zeta}_i$ for $i = 1, \ldots, q$, where*

$$\underline{\zeta}_i := \sqrt{\frac{D_{ii} - \|E\|_2}{1 + \|F\|_2}}, \quad \overline{\zeta}_i := \sqrt{\frac{D_{ii} + \|E\|_2}{1 - \|F\|_2}}.$$

The computational cost of the algorithm based on Theorem 6 is $14mn^2 + 13n^3 + \mathcal{O}(m^2 + n^2)$ divided into

$2mn^2$ approximate computation of A^TA,
$9n^3$ eigen-decomposition to obtain V (c.f. [4, Algorithm 8.3.3]),
$4n^3$ the inclusion of V^TV,
$4mn^2$ inclusion of AV,
$8mn^2$ inclusion of $(AV)^TAV$.

We formulate and prove Theorems 7, 9, 10 and 11 to develop the proposed algorithms, and establish Theorems 8 and 12 to clarify the relations between the proposed bounds and the previous bounds.

Theorem 7 *Let A, U, Σ, V and q be as in Theorem 4, $\hat{U}$ and $\hat{V}$ be the submatrices which consist of the first q columns of U and V, respectively, $\hat{\Sigma} \in \mathbb{R}^{q\times q}$ be diagonal with $\hat{\Sigma}_{ii} = \Sigma_{ii}, i = 1, \ldots, q$, $\hat{E} := \hat{U}\hat{\Sigma}\hat{V}^T - A$, $\hat{F} := \hat{V}^T\hat{V} - I_q$ and $\hat{G} := \hat{U}^T\hat{U} - I_q$. If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, then $\underline{\delta}_i^M \leq \sigma_i(A) \leq \overline{\delta}_i^M$ holds for $i = 1, \ldots, q$, where*

$$\underline{\delta}_i^M := \hat{\Sigma}_{ii}\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)} - \|\hat{E}\|_2,$$
$$\overline{\delta}_i^M := \hat{\Sigma}_{ii}\sqrt{(1 + \|\hat{F}\|_2)(1 + \|\hat{G}\|_2)} + \|\hat{E}\|_2.$$

Proof The inequalities $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, and Lemma 1 show that $\hat{U}$ and $\hat{V}$ have full rank. From Lemmas 1 and 2, we obtain

$$\begin{aligned}
\sigma_i(A) &= \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T) + \sigma_i(A) - \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T), \\
\sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T) &\le \|\hat{U}\|_2 \sigma_i(\hat{\Sigma})\|\hat{V}\|_2 \le \hat{\Sigma}_{ii}\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)}, \\
\hat{\Sigma}_{ii} &= \sigma_i(\hat{\Sigma}) = \sigma_i(\hat{U}^+\hat{U}\hat{\Sigma}\hat{V}^T\hat{V}^{+T}) \le \|\hat{U}^+\|_2\sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T)\|\hat{V}^+\|_2 \\
&\le \frac{\sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T)}{\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)}}, \\
|\sigma_i(A) - \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^T)| &\le \|\hat{E}\|_2,
\end{aligned}$$

which prove the result. □

The computational cost of the algorithm based on Theorem 7 is $24Qq^2 + 12q^3 + \mathcal{O}(m^2+n^2)$ divided into

$14Qq^2 + 8q^3$ economy SVD to obtain $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}$ (c.f. [4, Sect. 5.4.5]),
$6Qq^2$ inclusion of $(\hat{U}\hat{\Sigma})\hat{V}^T$ when $m \ge n$, or $\hat{U}(\hat{\Sigma}\hat{V}^T)$ when $m < n$,
$4Qq^2$ inclusion of $\hat{U}^T\hat{U}$ when $m \ge n$, or $\hat{V}^T\hat{V}$ when $m < n$,
$4q^3$ inclusion of $\hat{V}^T\hat{V}$ when $m \ge n$, or $\hat{U}^T\hat{U}$ when $m < n$.

Theorem 8 *Let q, $\underline{\delta}_i$ and $\overline{\delta}_i$ be as in Theorem 4, and $\underline{\delta}_i^M$ and $\overline{\delta}_i^M$ be as in Theorem 7. Then $\underline{\delta}_i \le \underline{\delta}_i^M$ and $\overline{\delta}_i^M \le \overline{\delta}_i$ follow for $i = 1, \dots, q$.*

Proof Let Σ, E, F and G be as in Theorem 4, and $\hat{\Sigma}$, $\hat{E}$, $\hat{F}$ and $\hat{G}$ be as in Theorem 7. Since $\hat{F}$ and $\hat{G}$ are the submatrices of F and G, respectively, $\|\hat{F}\|_2 \le \|F\|_2$ and $\|\hat{G}\|_2 \le \|G\|_2$ hold. These inequalities and $\hat{\Sigma}_{ii} = \Sigma_{ii}$ yield

$$\begin{aligned}
\hat{\Sigma}_{ii}\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)} &\le \Sigma_{ii}\sqrt{(1+\|F\|_2)(1+\|G\|_2)} \\
&\le \Sigma_{ii}(1+\max(\|F\|_2, \|G\|_2)), \\
\hat{\Sigma}_{ii}\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)} &\ge \Sigma_{ii}\sqrt{(1-\|F\|_2)(1-\|G\|_2)} \\
&\ge \Sigma_{ii}(1-\max(\|F\|_2, \|G\|_2)).
\end{aligned}$$

These inequalities and $\hat{E} = E$ show the result. □

Remark 3 Let A, U, V, q, F and G be as in Theorem 4, D, E, $\underline{\varepsilon}_i$ and $\overline{\varepsilon}_i$ be as in Theorem 5, and $\hat{U}$, $\hat{V}$, $\hat{F}$ and $\hat{G}$ be as in Theorem 7. Define $\hat{D}, \hat{E} \in \mathbb{R}^{q\times q}$ so that $\hat{U}^T A\hat{V} = \hat{D} + \hat{E}$ and $\hat{D}$ is diagonal. From Lemmas 1 and 2, if $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, we have

$$\begin{aligned}
\sigma_i(\hat{D}) - \|\hat{E}\|_2 \le \sigma_i(\hat{D}+\hat{E}) = \sigma_i(\hat{U}^T A\hat{V}) &\le \|\hat{U}\|_2\sigma_i(A)\|\hat{V}\|_2 \\
&\le \sigma_i(A)\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)},
\end{aligned}$$

so that there is a numbering $\nu : \{1, \dots, q\} \to \{1, \dots, q\}$ with

$$\underline{\varepsilon}_i^{\tilde{M}} \le \sigma_{\nu(i)}(A), \quad \underline{\varepsilon}_i^{\tilde{M}} := \frac{|\hat{D}_{ii}| - \|\hat{E}\|_2}{\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)}}.$$

The lower bound $\underline{\varepsilon}_i^{\tilde{M}}$ satisfies $\underline{\varepsilon}_i^{\tilde{M}} \geq \underline{\varepsilon}_i$, since $\|\hat{E}\|_2 \leq \|E\|_2$, $\|\hat{F}\|_2 \leq \|F\|_2$, $\|\hat{G}\|_2 \leq \|G\|_2$, and $\hat{D}_{ii} = D_{ii}$ hold. On the other hand, the inequality

$$\sigma_{\nu(i)}(A) \leq \overline{\varepsilon}_i^{\tilde{M}}, \quad \overline{\varepsilon}_i^{\tilde{M}} := \frac{|\hat{D}_{ii}| + \|\hat{E}\|_2}{\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)}}$$

does not always follow. For example, let $m = 2n$,

$$A = \begin{pmatrix} I_n \\ O_n \end{pmatrix}, \quad U = \begin{pmatrix} O_n & I_n \\ I_n & O_n \end{pmatrix}, \quad V = I_n.$$

We then obtain

$$\hat{U} = \begin{pmatrix} O_n \\ I_n \end{pmatrix}, \quad \hat{V} = I_n, \quad \hat{F} = \hat{G} = O_n, \quad \hat{D} = \hat{E} = O_n,$$

so that $\overline{\varepsilon}_i^{\tilde{M}} = 0$, although $\sigma_i(A) = 1$, $i = 1, \ldots, n$. Note that Theorem 5 is still valid even in this case, since

$$F = O_n, \quad G = O_m, \quad D = \begin{pmatrix} O_n \\ O_n \end{pmatrix}, \quad E = \begin{pmatrix} O_n \\ I_n \end{pmatrix},$$

so that $\underline{\varepsilon}_i = -1$ and $\overline{\varepsilon}_i = 1$ follow for all i.

Theorem 9 *Adding with the definitions and the assumptions in Theorem 5, for $i \in \{1, \ldots, q\}$, if $< |D_{ii}|, \|E\|_2 >$ is isolated from $\bigcup_{j=1, j\neq i}^{q} < |D_{jj}|, \|E\|_2 >$ and $|D_{ii}| > \|E\|_2$, then define*

$$\rho_i := \begin{cases} \min\limits_{1 \leq j \leq q, j \neq i} ||D_{jj}| - |D_{ii}|| - \|E\|_2 & (\text{if } m = n) \\ \min\left(\min\limits_{1 \leq j \leq q, j \neq i} ||D_{jj}| - |D_{ii}|| - \|E\|_2, |D_{ii}| \right) & (\text{otherwise}) \end{cases},$$

$$\mathsf{e}^{(i)} := \begin{pmatrix} E^T e^{(m,i)} \\ E e^{(n,i)} \end{pmatrix},$$

$$\xi_i := \frac{\|\mathsf{e}^{(i)}\|_2^2}{2\rho_i}, \quad \omega_i := \min(\xi_i, \|E\|_2).$$

Otherwise let $\omega_i := \|E\|_2$. Then $\underline{\varepsilon}_i^M \leq \sigma_{\nu(i)}(A) \leq \overline{\varepsilon}_i^M$ holds for $i = 1, \ldots, q$, where

$$\underline{\varepsilon}_i^M := \frac{|D_{ii}| - \omega_i}{\sqrt{(1 + \|F\|_2)(1 + \|G\|_2)}}, \quad \overline{\varepsilon}_i^M := \frac{|D_{ii}| + \omega_i}{\sqrt{(1 - \|F\|_2)(1 - \|G\|_2)}}.$$

Remark 4 Since $< |D_{ii}|, \|E\|_2 >$ is isolated from the other intervals and $|D_{ii}| > \|E\|_2$, $\min_{1 \leq j \leq q, j \neq i} ||D_{jj}| - |D_{ii}|| - \|E\|_2 > 2\|E\|_2 - \|E\|_2 = \|E\|_2 \geq 0$ and

$|D_{ii}| > \|E\|_2 \geq 0$ hold. Hence ρ_i is always positive. The real numbers ρ_i in Theorem 11 have the similar property.

Proof It is sufficient to prove $|D_{ii}| - \xi_i \leq \sigma_{\nu(i)}(D+E) \leq |D_{ii}| + \xi_i$ in the case when $< |D_{ii}|, \|E\|_2 >$ is isolated from the other intervals and $|D_{ii}| > \|E\|_2$. Let

$$\mathsf{D} := \begin{pmatrix} O_n & D^T + E^T \\ D + E & O_m \end{pmatrix}, \quad \mathsf{y}_+^{(i)} := \frac{1}{\sqrt{2}} \begin{pmatrix} e^{(n,i)} \\ e^{(m,i)} \end{pmatrix} \quad \text{and} \quad \mathsf{y}_-^{(i)} := \frac{1}{\sqrt{2}} \begin{pmatrix} e^{(n,i)} \\ -e^{(m,i)} \end{pmatrix}.$$

It is well known (e.g. [4]) that $\sigma_i(D+E)$, $i = 1, \ldots, q$ are the eigenvalues of D, since the characteristic polynomial of D coincides with that of

$$\mathrm{diag}(\sigma_1(D+E), \ldots, \sigma_q(D+E), -\sigma_1(D+E), \ldots, -\sigma_q(D+E), \underbrace{0, \ldots, 0}_{|m-n|}).$$

Moreover $\mathsf{y}_+^{(i)}$ and $\mathsf{y}_-^{(i)}$ are unit vectors. We thus prove the above inequality by applying Theorem 3 setting A and y in this theorem as D and $\mathsf{y}_+^{(i)}$, or D and $\mathsf{y}_-^{(i)}$, respectively.

If $D_{ii} \geq 0$, from Remark 2, we have

$$\begin{aligned}
\mathsf{y}_+^{(i)^T} \mathsf{D} \mathsf{y}_+^{(i)} &= \frac{1}{2}(e^{(m,i)^T} D e^{(n,i)} + e^{(m,i)^T} E e^{(n,i)} + e^{(n,i)^T} D^T e^{(m,i)} \\
&\quad + e^{(n,i)^T} E^T e^{(m,i)}) \\
&= \frac{1}{2}(D_{ii} + D_{ii}) = D_{ii} = |D_{ii}|, \\
\mathsf{D}\mathsf{y}_+^{(i)} - |D_{ii}| \mathsf{y}_+^{(i)} &= \frac{1}{\sqrt{2}} \begin{pmatrix} (D^T + E^T) e^{(m,i)} - |D_{ii}| e^{(n,i)} \\ (D+E) e^{(n,i)} - |D_{ii}| e^{(m,i)} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} D_{ii} e^{(n,i)} + E^T e^{(m,i)} - |D_{ii}| e^{(n,i)} \\ D_{ii} e^{(m,i)} + E e^{(n,i)} - |D_{ii}| e^{(m,i)} \end{pmatrix} \\
&= \frac{1}{\sqrt{2}} \begin{pmatrix} |D_{ii}| e^{(n,i)} + E^T e^{(m,i)} - |D_{ii}| e^{(n,i)} \\ |D_{ii}| e^{(m,i)} + E e^{(n,i)} - |D_{ii}| e^{(m,i)} \end{pmatrix} = \frac{1}{\sqrt{2}} \mathsf{e}^{(i)}.
\end{aligned}$$

From Lemma 2, $\sigma_{\nu(i)}(D+E) \in < |D_{ii}|, \|E\|_2 >$ holds for $i = 1, \ldots, q$. Since $< |D_{ii}|, \|E\|_2 >$ is isolated from $\bigcup_{j=1, j \neq i}^{q} < |D_{jj}|, \|E\|_2 >$ and $|D_{ii}| > \|E\|_2$, additionally, $0 \notin < |D_{ii}|, \|E\|_2 >$ follows and $< |D_{ii}|, \|E\|_2 >$ is also isolated from $\bigcup_{j=1}^{q} < -|D_{jj}|, \|E\|_2 >$. Therefore the eigenvalue of D closest to $|D_{ii}|$ is $\sigma_{\nu(i)}(D+E)$. When $m = n$, we hence obtain

$$\begin{aligned}
&\mathrm{gap}(\mathsf{D}, |D_{ii}|) \\
&\quad = \min\left(\min_{1 \leq j \leq q, j \neq \nu(i)} |\sigma_j(D+E) - |D_{ii}||, \min_{1 \leq j \leq q} |-\sigma_j(D+E) - |D_{ii}|| \right) \\
&\quad = \min_{1 \leq j \leq q, j \neq \nu(i)} |\sigma_j(D+E) - |D_{ii}|| \geq \min_{1 \leq j \leq q, j \neq i} ||D_{jj}| - |D_{ii}|| - \|E\|_2 = \rho_i.
\end{aligned}$$

When $m \neq n$, we similarly obtain

$$\begin{aligned}
&\mathrm{gap}(\mathsf{D}, |D_{ii}|)\\
&\quad = \min\left(\min_{1\le j\le q, j\ne\nu(i)} |\sigma_j(D+E) - |D_{ii}||, \min_{1\le j\le q} |-\sigma_j(D+E) - |D_{ii}||, |0-|D_{ii}||\right)\\
&\quad = \min\left(\min_{1\le j\le q, j\ne\nu(i)} |\sigma_j(D+E) - |D_{ii}||, |D_{ii}|\right)\\
&\quad \ge \min\left(\min_{1\le j\le q, j\ne i} ||D_{jj}| - |D_{ii}|| - \|E\|_2, |D_{ii}|\right) = \rho_i.
\end{aligned}$$

These discussion and Theorem 3 give $||D_{ii}| - \sigma_{\nu(i)}(D+E)| \le \xi_i$, which proves the above inequality.

If $D_{ii} < 0$, it holds that

$$\begin{aligned}
\mathsf{y}_-^{(i)^T} \mathsf{D} \mathsf{y}_-^{(i)} &= -D_{ii} = |D_{ii}|,\\
\mathsf{D}\mathsf{y}_-^{(i)} - |D_{ii}|\mathsf{y}_-^{(i)} &= \frac{1}{\sqrt{2}}\begin{pmatrix} -D_{ii}e^{(n,i)} - E^T e^{(m,i)} - |D_{ii}|e^{(n,i)} \\ D_{ii}e^{(m,i)} + Ee^{(n,i)} + |D_{ii}|e^{(m,i)} \end{pmatrix}\\
&= \frac{1}{\sqrt{2}}\begin{pmatrix} |D_{ii}|e^{(n,i)} - E^T e^{(m,i)} - |D_{ii}|e^{(n,i)} \\ -|D_{ii}|e^{(m,i)} + Ee^{(n,i)} + |D_{ii}|e^{(m,i)} \end{pmatrix}\\
&= \frac{1}{\sqrt{2}}\begin{pmatrix} -E^T e^{(m,i)} \\ Ee^{(n,i)} \end{pmatrix}.
\end{aligned}$$

This and the discussion similar to that in the case when $D_{ii} \ge 0$ yield $||D_{ii}| - \sigma_{\nu(i)}(D+E)| \le \xi_i$, which completes the proof. □

The computational cost of the algorithm based on Theorem 9 is similar to that of the algorithm based on Theorem 5. It is obvious that $\underline{\varepsilon}_i \le \underline{\varepsilon}_i^M$ and $\overline{\varepsilon}_i^M \le \overline{\varepsilon}_i$ follow for $i = 1, \dots, q$.

Remark 5 Let A, q, F and G be as in Theorem 4, D and E be as in Theorem 5, $\hat{F}$ and $\hat{G}$ be as in Theorem 7, and $\hat{D}$ and $\hat{E}$ be as in Remark 3. Theorem 9 is based on

$$\frac{\sigma_i(D+E)}{\sqrt{(1+\|F\|_2)(1+\|G\|_2)}} \le \sigma_i(A) \le \frac{\sigma_i(D+E)}{\sqrt{(1-\|F\|_2)(1-\|G\|_2)}}, \quad i = 1, \dots, q.$$

If it follows that

$$\frac{\sigma_i(\hat{D}+\hat{E})}{\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)}} \le \sigma_i(A) \le \frac{\sigma_i(\hat{D}+\hat{E})}{\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)}}, \quad i = 1, \dots, q,$$

therefore, improvement of Theorem 9 utilizing the economy SVD will be possible. From the first half of Remark 3, $\sigma_i(\hat{D}+\hat{E})/\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)} \le \sigma_i(A)$ holds. As the example in the second half of Remark 3 shows, however, $\sigma_i(A) \le \sigma_i(\hat{D}+\hat{E})$

$/\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)}$ does not always follow. Hence the improvement seems to be difficult. We leave the improvement as an open challenge.

Theorem 10 *Let A and q be as in Theorem 4, and $\hat{U}$, $\hat{\Sigma}$, $\hat{V}$, $\hat{F}$ and $\hat{G}$ be as in Theorem 7. Define*

$$\underline{\Sigma} := \begin{cases} \dfrac{\sqrt{1-\|\hat{G}\|_2}}{\sqrt{1+\|\hat{F}\|_2}}\hat{\Sigma} & (when\ m \geq n) \\ \dfrac{\sqrt{1-\|\hat{F}\|_2}}{\sqrt{1+\|\hat{G}\|_2}}\hat{\Sigma} & (when\ m < n) \end{cases}, \quad \overline{\Sigma} := \begin{cases} \dfrac{\sqrt{1+\|\hat{G}\|_2}}{\sqrt{1-\|\hat{F}\|_2}}\hat{\Sigma} & (when\ m \geq n) \\ \dfrac{\sqrt{1+\|\hat{F}\|_2}}{\sqrt{1-\|\hat{G}\|_2}}\hat{\Sigma} & (when\ m < n) \end{cases},$$

$$\rho := \begin{cases} \dfrac{\|A\hat{V}-\hat{U}\hat{\Sigma}\|_2}{\sqrt{1-\|\hat{F}\|_2}} & (when\ m \geq n) \\ \dfrac{\|\hat{U}^TA-\hat{\Sigma}\hat{V}^T\|_2}{\sqrt{1-\|\hat{G}\|_2}} & (when\ m < n) \end{cases}.$$

If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, $\underline{\Sigma}_{ii} - \rho \leq \sigma_i(A) \leq \overline{\Sigma}_{ii} + \rho$ follows for $i = 1, \ldots, q$.

Proof Similarly to the proof of Theorem 7, $\hat{V}$ and $\hat{U}$ have full rank. When $m \geq n$, it holds from Lemmas 1 and 2 that

$$\begin{aligned} \sigma_i(A) &= \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1}) + \sigma_i(A) - \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1}), \\ \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1}) &\leq \|\hat{U}\|_2\sigma_i(\hat{\Sigma})\|\hat{V}^{-1}\|_2 \leq \overline{\Sigma}_{ii}, \\ \hat{\Sigma}_{ii} &= \sigma_i(\hat{\Sigma}) = \sigma_i(\hat{U}^+\hat{U}\hat{\Sigma}\hat{V}^{-1}\hat{V}) \leq \|\hat{U}^+\|_2\sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1})\|\hat{V}\|_2 \\ &\leq \frac{\sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1})\sqrt{1+\|\hat{F}\|_2}}{\sqrt{1-\|\hat{G}\|_2}}, \\ |\sigma_i(A) - \sigma_i(\hat{U}\hat{\Sigma}\hat{V}^{-1})| &\leq \|A - \hat{U}\hat{\Sigma}\hat{V}^{-1}\|_2 = \|(A\hat{V} - \hat{U}\hat{\Sigma})\hat{V}^{-1}\|_2 \\ &\leq \|A\hat{V} - \hat{U}\hat{\Sigma}\|_2\|\hat{V}^{-1}\|_2 \leq \rho, \end{aligned}$$

showing $\underline{\Sigma}_{ii} - \rho \leq \sigma_i(A) \leq \overline{\Sigma}_{ii} + \rho$. When $m < n$, we obtain

$$\begin{aligned} \sigma_i(A) &= \sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T) + \sigma_i(A) - \sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T), \\ \sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T) &\leq \|\hat{U}^{-1}\|_2\sigma_i(\hat{\Sigma})\|\hat{V}\|_2 \leq \overline{\Sigma}_{ii}, \\ \hat{\Sigma}_{ii} &= \sigma_i(\hat{U}^T\hat{U}^{-T}\hat{\Sigma}\hat{V}^T\hat{V}^{+T}) \leq \|\hat{U}\|_2\sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T)\|\hat{V}^+\|_2 \\ &\leq \frac{\sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T)\sqrt{1+\|\hat{G}\|_2}}{\sqrt{1-\|\hat{F}\|_2}}, \end{aligned}$$

$$|\sigma_i(A) - \sigma_i(\hat{U}^{-T}\hat{\Sigma}\hat{V}^T)| \le \|A - \hat{U}^{-T}\hat{\Sigma}\hat{V}^T\|_2 = \|\hat{U}^{-T}(\hat{U}^T A - \hat{\Sigma}\hat{V}^T)\|_2$$
$$\le \|\hat{U}^{-T}\|_2 \|\hat{U}^T A - \hat{\Sigma}\hat{V}^T\|_2 \le \rho,$$

giving $\underline{\Sigma}_{ii} - \rho \le \sigma_i(A) \le \overline{\Sigma}_{ii} + \rho$. □

The computational cost of the algorithm based on Theorem 10 is $22Qq^2 + 12q^3 + \mathcal{O}(m^2 + n^2)$ divided into

$14Qq^2 + 8q^3$ the economy SVD,
$4Qq^2$ inclusion of $A\hat{V}$ when $m \ge n$, or $\hat{U}^T A$ when $m < n$,
$4Qq^2$ the inclusion of $\hat{U}^T\hat{U}$ or $\hat{V}^T\hat{V}$,
$4q^3$ the inclusion of $\hat{V}^T\hat{V}$ or $\hat{U}^T\hat{U}$.

Theorem 11 *Let A and q be as in Theorem 4. Let also V and F be as in Theorem 4 if $m \ge n$. Otherwise let $V \in \mathbb{R}^{m\times m}$ be given and $F := V^T V - I_m$. Define $\hat{D}$, $\hat{E} \in \mathbb{R}^{q\times q}$ so that*

$$\hat{D} + \hat{E} = \begin{cases} (AV)^T AV & (when\ m \ge n) \\ (A^T V)^T A^T V & (when\ m < n) \end{cases}$$

and $\hat{D}$ is diagonal. Let $s^{(q)} := (1, \dots, 1)^T \in \mathbb{R}^q$ and $f := |\hat{E}|s^{(q)}$, and assume $\|F\|_2 < 1$. For $i \in \{1, \dots, q\}$, if $< \hat{D}_{ii}, f_i >$ is isolated from $\bigcup_{j=1, j\ne i}^q < \hat{D}_{jj}, f_j >$, then define $\rho_i := \min_{1\le j\le q, j\ne i}(|\hat{D}_{ii} - \hat{D}_{jj}| - f_j)$, $g_i := \|\hat{E}e^{(q,i)}\|_2^2/\rho_i$, $h_i := \min(f_i, g_i)$,

$$\underline{\zeta}_i^M := \begin{cases} \sqrt{\dfrac{\hat{D}_{ii} - h_i}{1 + \|F\|_2}} & (if\ \hat{D}_{ii} \ge h_i) \\ 0 & (otherwise) \end{cases}, \quad \overline{\zeta}_i^M := \sqrt{\frac{\hat{D}_{ii} + h_i}{1 - \|F\|_2}}.$$

Otherwise let $\bigcup_{j=1}^k < \hat{D}_{i_j i_j}, f_{i_j} >$, where $i \in \{i_1, \dots, i_k\} \subseteq \{1, \dots, q\}$, be the connected intervals which is isolated from the other intervals, and define

$$\underline{\xi}_i := \max\left(\min_{j\in\{i_1,\dots,i_k\}}(\hat{D}_{jj} - f_j), \hat{D}_{ii} - \|\hat{E}\|_\infty\right),$$
$$\overline{\xi}_i := \min\left(\max_{j\in\{i_1,\dots,i_k\}}(\hat{D}_{jj} + f_j), \hat{D}_{ii} + \|\hat{E}\|_\infty\right),$$
$$\underline{\zeta}_i^M := \begin{cases} \sqrt{\dfrac{\underline{\xi}_i}{1 + \|F\|_2}} & (if\ \underline{\xi}_i \ge 0) \\ 0 & (otherwise) \end{cases}, \quad \overline{\zeta}_i^M := \sqrt{\frac{\overline{\xi}_i}{1 - \|F\|_2}}.$$

Then there is a numbering $\nu : \{1, \dots, q\} \to \{1, \dots, q\}$ with $\underline{\zeta}_i^M \le \sigma_{\nu(i)}(A) \le \overline{\zeta}_i^M$ for $i = 1, \dots, q$.

Proof When $m \geq n$, Lemmas 1 and 2 yield

$$\begin{aligned}\sigma_i(A)^2 &= \sigma_i(A^TA) = \sigma_i(V^{-T}V^TA^TAVV^{-1})\\ &\leq \|V^{-1}\|_2^2\sigma_i(V^TA^TAV) \leq \frac{\lambda_i((AV)^TAV)}{1-\|F\|_2} = \frac{\lambda_i(\hat{D}+\hat{E})}{1-\|F\|_2},\\ \lambda_i(\hat{D}+\hat{E}) &= \sigma_i(V^TA^TAV) \leq \|V\|_2^2\sigma_i(A^TA) \leq (1+\|F\|_2)\sigma_i(A)^2.\end{aligned}$$

When $m < n$, we similarly obtain

$$\frac{\lambda_i(\hat{D}+\hat{E})}{1+\|F\|_2} \leq \sigma_i(A)^2 \leq \frac{\lambda_i(\hat{D}+\hat{E})}{1-\|F\|_2}. \tag{3.1}$$

Since $\lambda_i(\hat{D}+\hat{E}) = \sigma_i(\hat{D}+\hat{E})$ and $\lambda_i(\hat{D}) = \sigma_i(\hat{D})$, Lemma 2 yields $|\lambda_i(\hat{D}+\hat{E}) - \lambda_i(\hat{D})| \leq \|\hat{E}\|_2 \leq \|\hat{E}\|_\infty$. Hence the numbering ν such that $\lambda_{\nu(i)}(\hat{D}) = \hat{D}_{ii}$ satisfies

$$|\lambda_{\nu(i)}(\hat{D}+\hat{E}) - \hat{D}_{ii}| \leq \|\hat{E}\|_\infty, \quad i = 1,\ldots,q. \tag{3.2}$$

Gershgorin circle theorem (e.g. [4]) gives that $\lambda_i(\hat{D}+\hat{E}), i = 1,\ldots,q$ are included in the set $\bigcup_{j=1}^q < \hat{D}_{jj}, f_j >$. If k of the intervals form a connected domain which is isolated from the other intervals, moreover, then there are precisely k eigenvalues.

Hence if $< \hat{D}_{ii}, f_i >$ is isolated from $\bigcup_{j=1, j\neq i}^q < \hat{D}_{jj}, f_j >$, then this interval contains precisely one eigenvalue of $\hat{D}+\hat{E}$ and the eigenvalue uniquely contained in $< \hat{D}_{ii}, f_i >$ must be $\lambda_{\nu(i)}(\hat{D}+\hat{E})$ [1], so that $|\hat{D}_{ii} - \lambda_{\nu(i)}(\hat{D}+\hat{E})| \leq f_i$ holds. The isolation of $< \hat{D}_{ii}, f_i >$ additionally enables us to apply Theorem 3 by setting A and y in this theorem as $\hat{D}+\hat{E}$ and $e^{(q,i)}$, respectively. We have $e^{(q,i)^T}(\hat{D}+\hat{E})e^{(q,i)} = \hat{D}_{ii}$ from Remark 2. The eigenvalue of $\hat{D}+\hat{E}$ closest to $\hat{D}_{ii}$ is $\lambda_{\nu(i)}(\hat{D}+\hat{E})$. It holds that $\mathrm{gap}(\hat{D}+\hat{E}, \hat{D}_{ii}) = \min_{1\leq j\leq q, j\neq\nu(i)} |\lambda_j(\hat{D}+\hat{E}) - \hat{D}_{ii}| \geq \rho_i$, and $(\hat{D}+\hat{E})e^{(q,i)} - \hat{D}_{ii}e^{(q,i)} = \hat{E}e^{(q,i)}$. These relations and Theorem 3 yield $|\hat{D}_{ii} - \lambda_{\nu(i)}(\hat{D}+\hat{E})| \leq g_i$.

If $< \hat{D}_{ii}, f_i >$ is not isolated from the other intervals, from the definition of $\bigcup_{j=1}^k < \hat{D}_{i_ji_j}, f_{i_j} >$ and the Gershgorin theorem, this set contains k eigenvalues and the contained eigenvalues must be $\lambda_j(\hat{D}+\hat{E})$, $j = \nu(i_1),\ldots,\nu(i_k)$, so that $\lambda_{\nu(i)}(\hat{D}+\hat{E})$ is included in this union. This and (3.2) give $\lambda_{\nu(i)}(\hat{D}+\hat{E}) \in \bigcup_{j=1}^k < \hat{D}_{i_ji_j}, f_{i_j} > \cap < \hat{D}_{ii}, \|\hat{E}\|_\infty >$. The result follows from these eigenvalue properties and (3.1). □

The computational cost of the algorithm based on Theorem 11 is $14Qq^2 + 13q^3 + \mathcal{O}(m^2+n^2)$ divided into

$2Qq^2$ approximate computation of A^TA or AA^T,
$9q^3$ eigen-decomposition to obtain V,

[1] Otherwise the union of $\nu(i)-1$ intervals $\bigcup_{\hat{D}_{jj}>\hat{D}_{ii}} < \hat{D}_{jj}, f_j >$ contains $\nu(i)$ eigenvalues $\lambda_j(\hat{D}+\hat{E})$, $j = 1,\ldots,\nu(i)$, or the union of $q-\nu(i)$ intervals $\bigcup_{\hat{D}_{jj}<\hat{D}_{ii}} < \hat{D}_{jj}, f_j >$ contains $q-\nu(i)+1$ eigenvalues $\lambda_j(\hat{D}+\hat{E})$, $j = \nu(i),\ldots,q$. They contradict the Gershgorin theorem.

$4q^3$ the inclusion of $V^T V$,
$4Qq^2$ inclusion of AV when $m \geq n$, or $A^T V$ when $m < n$,
$8Qq^2$ inclusion of $(AV)^T AV$ when $m \geq n$, or $(A^T V)^T A^T V$ when $m < n$.

Theorem 12 *Let F and q be as in Theorem 4, D and E be as in Theorem 6, and $\underline{\zeta}_i^M$ and $\overline{\zeta}_i^M$ be as in Theorem 11. If $m \geq n$ and $D_{ii} \geq \|E\|_\infty$ for $i \in \{1, \ldots, q\}$, it holds that*

$$\sqrt{\frac{D_{ii} - \|E\|_\infty}{1 + \|F\|_2}} \leq \underline{\zeta}_i^M, \quad \overline{\zeta}_i^M \leq \sqrt{\frac{D_{ii} + \|E\|_\infty}{1 - \|F\|_2}}.$$

Remark 6 As shown in Theorem 6, $\underline{\zeta}_i$ and $\overline{\zeta}_i$ contain not $\|E\|_\infty$ but $\|E\|_2$. In the practical application of the algorithm based on Theorem 6, on the other hand, $\|E\|_2$ in $\underline{\zeta}_i$ and $\overline{\zeta}_i$ are replaced by $\|E\|_\infty$ in order to avoid the direct computation of 2-norm. Hence Theorem 12 clarifies the relation between the bounds in Theorems 6 and 11 in the practical applications when $m \geq n$.

Remark 7 If $D_{ii} < \|E\|_\infty$, $\sqrt{(D_{ii} - \|E\|_\infty)/(1 + \|F\|_2)}$ is complex, so that this quantity becomes meaningless.

Proof Let $\hat{D}$, $\hat{E}$ and f be as in Theorem 11. If $m \geq n$, $\hat{D}$ and $\hat{E}$ coincides with D and E, respectively. This and $f_i \leq \|E\|_\infty$ prove the theorem. □

4 Numerical results

In this section, we report numerical results to show the properties of the proposed algorithms and performances of our implementation. We used a computer with Intel Xeon 2.66GHz Dual CPU, 4.00GB RAM and MATLAB 7.5 with Intel Math Kernel Library and IEEE 754 double precision. In order to assess the quality of the verified bounds, we define the radius as follows:

$$\text{the radius} := \frac{\text{the upper bound} - \text{the lower bound}}{2}.$$

The compared algorithms are as follows:

M1: The algorithm based on Theorem 7,
M2: The algorithm based on Theorem 9,
M3: The algorithm based on Theorem 10,
M4: The algorithm based on Theorem 11,
O: The algorithm based on Theorem 4,
R1: The algorithm based on Theorem 5,
R2: The algorithm based on Theorem 6,
V: VERSOFT [13] function `versingval`.

In the algorithms M1, M2, M3, O, and R1, the economy or full SVD are executed by the MATLAB function `svd`. In M4 and R2, the eigen-decomposition is executed by

Table 1 Computing times (s) in Sect. 4.1

m	n	M1	M2	M3	M4	O	R1	R2	V
1000	300	0.7606	2.1017	0.7123	0.4478	1.8745	2.0879	0.4523	21610
3000	300	1.5187	28.293	1.3742	0.9300	26.359	28.260	0.9454	MO
10000	300	4.3709	MO	3.9379	2.5639	MO	MO	2.5925	MO
1000	1000	17.181	17.929	16.765	7.7360	17.125	17.830	7.7315	23026
3000	3000	467.78	485.30	458.94	195.87	467.98	484.69	195.92	MO
300	1000	0.7927	2.1006	0.7027	0.4396	1.9132	2.0909	4.8780	21539
300	3000	1.6637	28.379	1.4008	0.9426	26.710	28.300	112.31	MO
300	10000	5.9769	MO	3.9487	2.5950	MO	MO	MO	MO

the function `eig`. For $A \in \mathbb{R}^{m\times n}$ having full rank, we define the condition number $\kappa(A) := \|A\|_2\|A^+\|_2$. When A is rank deficient, $\kappa(A) := \infty$. In tables below, the notation MO and fail1 mean that the algorithms failed because of memory over, and R2 failed since $D_{ii} - \|E\|_2$ in Theorem 6 could not be verified to be positive for some i, respectively.

4.1 Example 1

In this example, we observe computing times of the algorithms for various m and n. Consider the case when $A \in \mathbb{R}^{m\times n}$ is generated by `A = randn(m,n);`. The function `randn` generates a random matrix whose elements are uniformly distributed in $[-1, 1]$. Table 1 displays the computing times of the algorithms for various m and n.

The computing times of M1 and M4 were approximately equal or smaller than those of O and R2, respectively. The computing times of M2 were approximately equal to those of R1. These results coincide with the discussion in Sect. 3. When $m < n$, especially, M4 was much faster than R2. The reason is that M4 does not involve matrix multiplications of $n \times n$ matrices, although R2 involves them. From the similar reason, M1 succeeded even when $(m, n) = (10000, 300), (300, 10000)$, as opposed that O failed because of memory over.

4.2 Example 2

In this example, we observe the magnitudes of the radii for various $\kappa(A)$. Consider the case when $A \in \mathbb{R}^{1000\times 10}$ is generated by `A = gallery('randsvd',[1000, 10],cnd);`. We used the Higham's test matrix `randsvd` [5]. Then $\kappa(A) \approx$ `cnd` holds approximately. Table 2 displays the maximum and minimum radii obtained by the algorithms for various `cnd`.

The radii by M1, M2 and M4 were smaller than those by O, R1 and R2, respectively. This result coincides with the discussion in Sect. 3. The algorithm M4 gave large maximum radii when `cnd` was large. The reason is as follows: The smallest singular

Table 2 Maximum (upper part) and minimum (lower part) radii in Sect. 4.2

cnd	M1	M2	M3	M4	O	R1	R2	V
1e+0	3.1e–14	5.4e–13	3.0e–14	2.9e–14	3.1e–13	5.4e–13	2.9e–14	5.3e–14
1e+4	3.8e–14	1.6e–13	3.8e–14	2.2e–14	3.1e–13	2.7e–13	1.2e–11	2.2e–14
1e+8	3.6e–14	1.6e–13	3.7e–14	2.2e–10	3.1e–13	3.4e–13	fail1	2.0e–14
1e+12	2.9e–14	1.6e–13	2.9e–14	1.3e–8	2.9e–13	3.1e–13	fail1	`NaN`
1e+16	5.7e–14	2.1e–13	5.3e–14	9.4e–9	3.2e–13	3.7e–13	fail1	`NaN`
1e+0	3.1e–14	5.3e–13	3.0e–14	1.2e–14	3.1e–13	5.3e–13	1.2e–14	5.3e–14
1e+4	1.3e–14	3.2e–16	1.4e–14	5.5e–17	1.3e–14	1.2e–13	9.9e–15	1.7e–16
1e+8	1.4e–14	2.2e–16	1.4e–14	5.1e–17	1.4e–14	1.8e–13	fail1	1.4e–16
1e+12	4.9e–15	2.1e–16	5.2e–15	4.3e–17	4.9e–15	1.5e–13	fail1	`NaN`
1e+16	2.4e–14	2.8e–16	2.0e–14	1.2e–16	2.4e–14	2.1e–13	fail1	`NaN`

Table 3 Obtained radii in Sect. 4.3

M1			M2		
1.1e–14	6.0e–15	6.0e–15	1.4e–14	8.1e–15	8.2e–15
M3			M4		
9.5e–15	4.5e–15	4.5e–15	6.3e–15	4.3e–8	4.3e–8
R1			R2		
2.1e–14	8.1e–15	8.2e–15	6.7e–15	fail1	fail1
O			V		
1.7e–14	6.0e–15	6.0e–15	2.7e–15	1.8e–15	1.8e–15

value of A becomes close to zero in this case. Thus $(\hat{D}_{ii} + h_i)/(1 - \|F\|_2)$ for some i in Theorem 11 becomes also close to zero, so that this quantity is enlarged by taking square root. For instance, let $\hat{D}_{ii} = 0$, $\|F\|_2 = 0$ and h_i =1e–16. Then $\overline{\zeta}_i^M$ =1e–8, although $(\hat{D}_{ii} + h_i)/(1 - \|F\|_2)$ =1e–16.

4.3 Example 3

In this example, we observe the radii when A is rank deficient. Consider the case when $A \in \mathbb{R}^{10\times 3}$ is generated by `A = repmat(randn(10,1),1,3);`. Then $\sigma_2(A) = \sigma_3(A) = 0$ holds strictly. Table 3 displays all the radii obtained by the algorithms, where the first, second and third radii correspond to $\sigma_1(A)$, $\sigma_2(A)$ and $\sigma_3(A)$, respectively.

Table 3 shows that all the algorithms except R2 could compute verified bounds for all singular values, even if there exist zero and multiple singular values. We can confirm the similar tendencies to those in Sect. 4.2 regarding to the magnitudes of the radii.

4.4 Example 4

In this example, we observe the magnitudes of the radii and computing times for matrices in the university of Florida sparse matrix collection [2]. Table 4 shows names, m, n, $\mathrm{rank}(A)$ and $\kappa(A)$ of the matrices being used. Tables 5 and 6 display the similar quantities to Tables 1 and 2, respectively. Tables 5 and 6 showed the similar tendencies to those in Sects. 4.1, 4.2 and 4.3.

Table 4 Properties of matrices

Name	m	n	$\mathrm{rank}(A)$	$\kappa(A)$
`Franz1`	2240	768	768	2.7e+15
`lp_adlittle`	56	138	56	4.6e+2
`lp_brandy`	220	303	193	∞
`Maragal_2`	555	350	220	∞
`west0497`	497	497	497	4.6e+11

Table 5 Computing times (s) in Sect. 4.4

name	M1	M2	M3	M4	O	R1	R2	V
`Franz1`	8.5057	21.757	8.0845	4.8320	19.246	21.628	4.7589	MO
`lp_adlittle`	0.0115	0.0171	0.0100	0.0065	0.0159	0.0161	0.0225	19.289
`lp_brandy`	0.1601	0.2036	0.1503	0.1047	0.1820	0.1963	0.1887	737.23
`Maragal_2`	0.5467	0.7875	0.5068	0.3613	0.6852	0.7607	0.3588	6249.7
`west0497`	1.1153	1.2117	1.0344	0.7465	1.0947	1.1913	0.7535	8016.2

Table 6 Maximum (upper part) and minimum (lower part) radii in Sect. 4.4

Name	M1	M2	M3	M4	O	R1	R2	V
`Franz1`	3.6e–12	4.8e–12	2.7e–12	3.7e–7	6.0e–12	5.0e–12	fail1	MO
`lp_adlittle`	4.0e–12	4.3e–12	3.3e–12	1.3e–12	6.6e–12	4.9e–12	1.4e–10	`NaN`
`lp_brandy`	5.9e–11	5.7e–11	5.5e–11	9.6e–6	7.9e–11	6.4e–11	fail1	`NaN`
`Maragal_2`	2.5e–12	2.2e–12	2.2e–12	3.1e–7	3.6e–12	3.1e–12	fail1	`NaN`
`west0497`	1.2e–7	1.2e–7	1.1e–7	8.0e–3	1.3e–7	1.2e–7	fail1	`NaN`
`Franz1`	1.7e–12	1.9e–12	7.1e–13	6.5e–13	1.7e–12	1.9e–12	fail1	MO
`lp_adlittle`	1.2e–12	9.5e–15	5.0e–13	3.1e–15	1.3e–12	5.6e–13	2.3e–12	`NaN`
`lp_brandy`	1.1e–11	9.1e–15	5.7e–12	3.9e–15	1.1e–11	7.4e–12	fail1	`NaN`
`Maragal_2`	7.3e–13	7.5e–15	4.3e–13	2.7e–15	7.3e–13	8.7e–13	fail1	`NaN`
`west0497`	1.2e–8	1.2e–17	3.2e–9	6.3e–13	1.2e–8	4.1e–9	fail1	`NaN`

5 Extension to generalized singular values

In this section, we extend the presented theorems in Sect. 3 for computing verified bounds of all the generalized singular values. Corollaries 1, 2, 3 and 4 are extensions of Theorems 7, 9, 10 and 11, respectively, which are applicable when $B^T B$ is nonsingular.

Corollary 1 *Let $A \in \mathbb{R}^{m\times n}$ with $m \geq n$, $B \in \mathbb{R}^{p\times n}$ with $p \geq n$, $\hat{U} \in \mathbb{R}^{m\times n}$, $\hat{\Sigma} \in \mathbb{R}^{n\times n}$ and $\hat{V} \in \mathbb{R}^{n\times n}$ be given, $\hat{\Sigma}$ be diagonal with $\hat{\Sigma}_{11} \geq \cdots \geq \hat{\Sigma}_{nn}$, $\hat{E} := \hat{U}\hat{\Sigma}\hat{V}^T B^T B - A$, $\hat{F} := \hat{V}^T B^T B\hat{V} - I_n$, $\hat{G} := \hat{U}^T\hat{U} - I_n$ and $\beta \geq \|B^+\|_2$. If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, then $\mathrm{rank}(B) = n$ and $\underline{\delta}_i^A \leq \sigma_i(A, B) \leq \overline{\delta}_i^A$ hold for $i = 1, \ldots, n$, where*

$$\underline{\delta}_i^A := \hat{\Sigma}_{ii}\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)} - \beta\|\hat{E}\|_2,$$
$$\overline{\delta}_i^A := \hat{\Sigma}_{ii}\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)} + \beta\|\hat{E}\|_2.$$

Proof The inequality $\|\hat{F}\|_2 < 1$ and Lemma 1 give $\mathrm{rank}(B\hat{V}) = n$, which implies $\mathrm{rank}(B) = n$. Since $p \geq n$ and $\mathrm{rank}(B) = n$, $B^T B$ is positive definite, so that there exists the Cholesky factorization of $B^T B$ such that $B^T B = LL^T$, where $L \in \mathbb{R}^{n\times n}$ is nonsingular lower triangular. It hence follows for $i = 1, \ldots, n$ that

$$\begin{aligned}\sigma_i(A, B)^2 &= \lambda_i(A^T A, B^T B) = \lambda_i(L^{-1}A^T AL^{-T}) = \lambda_i((AL^{-T})^T AL^{-T})\\ &= \sigma_i(AL^{-T})^2,\end{aligned}$$

which shows $\sigma_i(A, B) = \sigma_i(AL^{-T})$. From this, $\|(L^T\hat{V})^T L^T\hat{V} - I_n\|_2 = \|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, we can apply Theorem 7 for bounding $\sigma_i(A, B)$ by setting A, $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}$ in this theorem as AL^{-T}, $\hat{U}$, $\hat{\Sigma}$ and $L^T\hat{V}$, respectively. We then obtain $\underline{\delta}_i^{A^*} \leq \sigma_i(A, B) \leq \underline{\delta}_i^{A^*}$, where

$$\underline{\delta}_i^{A^*} := \hat{\Sigma}_{ii}\sqrt{(1-\|\hat{F}\|_2)(1-\|\hat{G}\|_2)} - \|\hat{U}\hat{\Sigma}(L^T\hat{V})^T - AL^{-T}\|_2,$$
$$\overline{\delta}_i^{A^*} := \hat{\Sigma}_{ii}\sqrt{(1+\|\hat{F}\|_2)(1+\|\hat{G}\|_2)} + \|\hat{U}\hat{\Sigma}(L^T\hat{V})^T - AL^{-T}\|_2.$$

This and

$$\begin{aligned}\|\hat{U}\hat{\Sigma}(L^T\hat{V})^T - AL^{-T}\|_2 &= \|(\hat{U}\hat{\Sigma}\hat{V}^T LL^T - A)L^{-T}\|_2 = \|\hat{E}L^{-T}\|_2\\ &\leq \|\hat{E}\|_2\|L^{-T}\|_2 = \|\hat{E}\|_2\sqrt{\|L^{-T}L^{-1}\|_2}\\ &= \|\hat{E}\|_2\sqrt{\|(LL^T)^{-1}\|_2}\\ &= \|\hat{E}\|_2\sqrt{\|(B^T B)^{-1}\|_2} = \|\hat{E}\|_2\|B^+\|_2 \leq \beta\|\hat{E}\|_2 \end{aligned} \tag{5.1}$$

prove the corollary. □

The computational cost of the algorithm based on Corollary 1 is $73mn^2/3 + 62n^3/3 + 4n^2p + \mathcal{O}(m^2 + n^2 + p^2)$ divided into

$4n^2p$ inclusion of B^TB,
$n^3/3$ Cholesky factorization of the center of the inclusion,
$mn^2/3$ approximate computation of AL^{-T} via forward substitution,
$14mn^2 + 8n^3$ economy SVD to obtain $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}_S$ such that $AL^{-T} \approx \hat{U}\hat{\Sigma}\hat{V}_S^T$,
$n^3/3$ approximate computation of $L^{-T}\hat{V}_S$ to obtain $\hat{V}$ via backward substitution,
$6n^3$ inclusion of $\hat{V}^TB^TB$,
$6mn^2$ inclusion of $\hat{U}\hat{\Sigma}\hat{V}^TB^TB$,
$6n^3$ inclusion of $\hat{V}^TB^TB\hat{V}$,
$4mn^2$ inclusion of $\hat{U}^T\hat{U}$,

except the computation of β.

Corollary 2 *Let A and B be as in Corollary 1, $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ be given, D and E be defined similarly to those in Theorem 5, $F := V^TB^TBV - I_n$ and $G := U^TU - I_m$. If $\|F\|_2 < 1$ and $\|G\|_2 < 1$, then $\operatorname{rank}(B) = n$ and there is a numbering $\nu : \{1, \ldots, q\} \to \{1, \ldots, q\}$ with $\underline{\varepsilon}_i^A \le \sigma_{\nu(i)}(A, B) \le \overline{\varepsilon}_i^A$ for $i = 1, \ldots, n$, where $\underline{\varepsilon}_i^A$ and $\overline{\varepsilon}_i^A$ are defined similarly to $\underline{\varepsilon}_i^M$ and $\overline{\varepsilon}_i^M$ in Theorem 9, respectively.*

Proof Let L be as in the proof of Corollary 1. The relation $\sigma_i(A, B) = \sigma_i(AL^{-T})$ and application of Theorem 9 by setting A, U and V in this theorem as AL^{-T}, U and L^TV, respectively, prove the result. □

The computational cost of the algorithm based on Corollary 2 is $4m^3 + 8m^2n + 43mn^2/3 + 65n^3/3 + 4n^2p + \mathcal{O}(m^2 + n^2 + p^2)$ divided into

$4n^2p$ the inclusion of B^TB,
$n^3/3$ the Cholesky factorization,
$mn^2/3$ the approximate computation of AL^{-T},
$4m^2n + 8mn^2 + 9n^3$ full SVD to obtain U, Σ and V_S such that $AL^{-T} \approx U\Sigma V_S^T$,
$n^3/3$ the approximate computation of $L^{-T}V_S$ to obtain V,
$12n^3$ inclusion of V^TB^TBV,
$4m^3$ inclusion of U^TU,
$4m^2n + 6mn^2$ inclusion of $(U^TA)V$.

Corollary 3 *Let A, B, $\hat{U}$, $\hat{\Sigma}$, $\hat{V}$, $\hat{F}$ and $\hat{G}$ be as in Corollary 1. Define*

$$\underline{\Sigma} := \frac{\sqrt{1 - \|\hat{G}\|_2}}{\sqrt{1 + \|\hat{F}\|_2}}\hat{\Sigma}, \quad \overline{\Sigma} := \frac{\sqrt{1 + \|\hat{G}\|_2}}{\sqrt{1 - \|\hat{F}\|_2}}\hat{\Sigma}, \quad \rho := \frac{\|A\hat{V} - \hat{U}\hat{\Sigma}\|_2}{\sqrt{1 - \|\hat{F}\|_2}}.$$

If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, then $\operatorname{rank}(B) = n$ and $\underline{\Sigma}_{ii} - \rho \le \sigma_i(A, B) \le \overline{\Sigma}_{ii} + \rho$ follow for $i = 1, \ldots, n$.

Proof Let L be as in the proof of Corollary 1. The coincidence $\sigma_i(A, B) = \sigma_i(AL^{-T})$ and application of Theorem 10 by setting A, $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}$ in this theorem as AL^{-T}, $\hat{U}$, $\hat{\Sigma}$ and $L^T\hat{V}$, respectively, prove the result. □

The computational cost of the algorithm based on Corollary 3 is $67mn^2/3 + 62n^3/3 + 4n^2p + \mathcal{O}(m^2 + n^2 + p^2)$ divided into

$4n^2p$ the inclusion of B^TB,
$n^3/3$ the Cholesky factorization,
$mn^2/3$ the approximate computation of AL^{-T},
$14mn^2 + 8n^3$ the economy SVD to obtain $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}_S$,
$n^3/3$ the approximate computation of $L^{-T}\hat{V}_S$,
$4mn^2$ the inclusion of $A\hat{V}$,
$12n^3$ the inclusion of $\hat{V}^TB^TB\hat{V}$,
$4mn^2$ the inclusion of $\hat{U}^T\hat{U}$.

Corollary 4 *Let A, B, $\hat{V}$ and $\hat{F}$ be as in Corollary 1. Define $\hat{D}, \hat{E} \in \mathbb{R}^{n\times n}$ so that $\hat{D} + \hat{E} = (A\hat{V})^TA\hat{V}$ and $\hat{D}$ is diagonal. Let $\underline{\zeta}_i^A$ and $\overline{\zeta}_i^A$ be defined similarly to $\underline{\zeta}_i^M$ and $\overline{\zeta}_i^M$ in Theorem 11, respectively. If $\|\hat{F}\|_2 < 1$, then* $\mathrm{rank}(B) = n$ *and there is a numbering $\nu : \{1, \ldots, n\} \to \{1, \ldots, n\}$ with $\underline{\zeta}_i^A \le \sigma_{\nu(i)}(A, B) \le \overline{\zeta}_i^A$ for $i = 1, \ldots, n$.*

Proof Let L be as in the proof of Corollary 1. The relation $\sigma_i(A, B) = \sigma_i(AL^{-T})$ and application of Theorem 11 by setting A and V in this theorem as AL^{-T} and $L^T\hat{V}$, respectively, show the inequality. □

The computational cost of the algorithm based on Corollary 4 is $14mn^2 + 26n^3 + 4n^2p + \mathcal{O}(m^2 + n^2 + p^2)$ divided into

$2mn^2$ approximate computation of A^TA,
$4n^2p$ the inclusion of B^TB,
$14n^3$ generalized eigen-decomposition via Cholesky-QR algorithm (e.g. [4, Algorithm 8.7.1]) to obtain $\hat{V}$,
$4mn^2$ inclusion of $A\hat{V}$,
$8mn^2$ inclusion of $(A\hat{V})^TA\hat{V}$,
$12n^3$ the inclusion of $\hat{V}^TB^TB\hat{V}$.

Corollaries 5, 6, 7 and 8 are extensions of Theorems 7, 9, 10 and 11, respectively, which are applicable when A^TA is nonsingular. The computational costs of the algorithms based on Corollaries 5, 6, 7 and 8 are analogous to those of the algorithms based on Corollaries 1, 2, 3 and 4, respectively.

Corollary 5 *Let $A \in \mathbb{R}^{m\times n}$ with $m \ge n$, $B \in \mathbb{R}^{p\times n}$, $\hat{U} \in \mathbb{R}^{p\times q}$, $\hat{\Sigma} \in \mathbb{R}^{q\times q}$ and $\hat{V} \in \mathbb{R}^{n\times q}$ be given where $q :=$ $\min(p, n)$, $\hat{\Sigma}$ be diagonal with $\hat{\Sigma}_{11} \ge \cdots \ge \hat{\Sigma}_{qq}$, $\hat{E} := \hat{U}\hat{\Sigma}\hat{V}^TA^TA - B$, $\hat{F} := \hat{V}^TA^TA\hat{V} - I_q$, $\hat{G} := \hat{U}^T\hat{U} - I_q$, $r^* :=$ $\mathrm{rank}(B)$ and $\beta \ge \|A^+\|_2$. Assume $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, and let r_δ denote an index satisfying*

$$\hat{\Sigma}_{r_\delta r_\delta} > \frac{\beta\|\hat{E}\|_2}{\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)}} \ge \hat{\Sigma}_{r_{\delta+1} r_{\delta+1}}.$$

Then $\mathrm{rank}(A) = n$ *and* $1/\overline{\delta}_i^B \le \sigma_{r^*-i+1}(A, B) \le 1/\underline{\delta}_i^B$ *hold for* $i = 1, \dots, r_\delta$, *where*

$$\underline{\delta}_i^B := \hat{\Sigma}_{ii}\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)} - \beta\|\hat{E}\|_2,$$
$$\overline{\delta}_i^B := \hat{\Sigma}_{ii}\sqrt{(1 + \|\hat{F}\|_2)(1 + \|\hat{G}\|_2)} + \beta\|\hat{E}\|_2.$$

If $r_\delta = q$ *in particular,* $1/\overline{\delta}_i^B \le \sigma_{q-i+1}(A, B) \le 1/\underline{\delta}_i^B$ *holds for* $i = 1, \dots, q$.

Proof Similarly to the proof of Corollary 1, there exists the Cholesky factorization of $A^T A$ such that $A^T A = LL^T$, where $L \in \mathbb{R}^{n\times n}$ is nonsingular lower triangular. It thus follows for $i = 1, \dots, r^*$ that

$$\sigma_{r^*-i+1}(A, B)^2 = \lambda_{r^*-i+1}(A^T A, B^T B) = \frac{1}{\lambda_i(B^T B, A^T A)} = \frac{1}{\lambda_i(L^{-1}B^T BL^{-T})}$$
$$= \frac{1}{\lambda_i((BL^{-T})^T BL^{-T})} = \frac{1}{\sigma_i(BL^{-T})^2},$$

which yields $\sigma_{r^*-i+1}(A, B) = 1/\sigma_i(BL^{-T})$. By applying Theorem 7 setting A, $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}$ in this theorem as BL^{-T}, $\hat{U}$, $\hat{\Sigma}$ and $L^T\hat{V}$, respectively, we obtain $\underline{\delta}_i^{B^*} \le \sigma_i(BL^{-T}) \le \overline{\delta}_i^{B^*}$, where

$$\underline{\delta}_i^{B^*} := \hat{\Sigma}_{ii}\sqrt{(1 - \|\hat{F}\|_2)(1 - \|\hat{G}\|_2)} - \|\hat{U}\hat{\Sigma}(L^T\hat{V})^T - BL^{-T}\|_2,$$
$$\overline{\delta}_i^{B^*} := \hat{\Sigma}_{ii}\sqrt{(1 + \|\hat{F}\|_2)(1 + \|\hat{G}\|_2)} + \|\hat{U}\hat{\Sigma}(L^T\hat{V})^T - BL^{-T}\|_2.$$

This and the similar derivation to (5.1) give $\underline{\delta}_i^B \le \sigma_i(BL^{-T}) \le \overline{\delta}_i^B$. This inequality and $\sigma_{r^*-i+1}(A, B) = 1/\sigma_i(BL^{-T})$ prove the corollary. □

Remark 8 Note that the algorithm based on Corollary 5 does not give verified bounds for all the generalized singular values if $r_\delta < r^*$. The algorithms based on Corollaries 6, 7 and 8 have the analogous property.

Corollary 6 *Let* A, B, q *and* r^* *be as in Corollary 5,* $U \in \mathbb{R}^{p\times p}$ *and* $V \in \mathbb{R}^{n\times n}$ *be given,* $F := V^T A^T AV - I_n$ *and* $G := U^T U - I_p$. *Define* $D, E \in \mathbb{R}^{p\times n}$ *so that* $D + E = U^T BV$ *and* D *is diagonal. Let* $\nu : \{1, \dots, q\} \to \{1, \dots, q\}$ *and* r_ε *denote a numbering and an index satisfying* $|D_{\nu(1)\nu(1)}| \ge \cdots \ge |D_{\nu(q)\nu(q)}|$ *and* $|D_{\nu(r_\varepsilon)\nu(r_\varepsilon)}| > \|E\|_2 \ge |D_{\nu(r_\varepsilon+1)\nu(r_\varepsilon+1)}|$, *respectively. If* $\|F\|_2 < 1$ *and* $\|G\|_2 < 1$, *then* $\mathrm{rank}(A) = n$ *and* $1/\overline{\varepsilon}_{\nu(i)}^B \le \sigma_{r^*-i+1}(A, B) \le 1/\underline{\varepsilon}_{\nu(i)}^B$ *hold for* $i = 1, \dots, r_\varepsilon$, *where* $\underline{\varepsilon}_i^B$ *and* $\overline{\varepsilon}_i^B$ *are defined similarly to* $\underline{\varepsilon}_i^M$ *and* $\overline{\varepsilon}_i^M$ *in Theorem 9, respectively. If* $r_\varepsilon = q$ *in particular,* $1/\overline{\varepsilon}_{\nu(i)}^B \le \sigma_{q-i+1}(A, B) \le 1/\underline{\varepsilon}_{\nu(i)}^B$ *holds for* $i = 1, \dots, q$.

Proof Let L be as in the proof of Corollary 5. The coincidence $\sigma_{r^*-i+1}(A, B) = 1/\sigma_i(BL^{-T})$ and the application of Theorem 9 by setting A, U and V in this theorem as BL^{-T}, U and $L^T V$, respectively, prove the result. □

Corollary 7 *Let A, B, $\hat{U}$, $\hat{\Sigma}$, $\hat{V}$, q, $\hat{F}$, $\hat{G}$, r^* and β be as in Corollary 5. Define*

$$\underline{\Sigma} := \begin{cases} \dfrac{\sqrt{1-\|\hat{G}\|_2}}{\sqrt{1+\|\hat{F}\|_2}}\hat{\Sigma} & (when\ p \geq n) \\ \dfrac{\sqrt{1-\|\hat{F}\|_2}}{\sqrt{1+\|\hat{G}\|_2}}\hat{\Sigma} & (when\ p < n) \end{cases}, \quad \overline{\Sigma} := \begin{cases} \dfrac{\sqrt{1+\|\hat{G}\|_2}}{\sqrt{1-\|\hat{F}\|_2}}\hat{\Sigma} & (when\ p \geq n) \\ \dfrac{\sqrt{1+\|\hat{F}\|_2}}{\sqrt{1-\|\hat{G}\|_2}}\hat{\Sigma} & (when\ p < n) \end{cases},$$

$$\rho := \begin{cases} \dfrac{\|B\hat{V} - \hat{U}\hat{\Sigma}\|_2}{\sqrt{1-\|\hat{F}\|_2}} & (when\ p \geq n) \\ \dfrac{\beta\|\hat{U}^T B - \hat{\Sigma}\hat{V}^T A^T A\|_2}{\sqrt{1-\|\hat{G}\|_2}} & (when\ p < n) \end{cases}.$$

Let r_ρ denote an index satisfying $\underline{\Sigma}_{r_\rho r_\rho} > \rho \geq \underline{\Sigma}_{r_\rho+1 r_\rho+1}$. If $\|\hat{F}\|_2 < 1$ and $\|\hat{G}\|_2 < 1$, then $\operatorname{rank}(A) = n$ *and $1/(\overline{\Sigma}_{ii} + \rho) \leq \sigma_{r^*-i+1}(A, B) \leq 1/(\underline{\Sigma}_{ii} - \rho)$ follow for $i = 1, \ldots, r_\rho$. If $r_\rho = q$ in particular, $1/(\overline{\Sigma}_{ii} + \rho) \leq \sigma_{q-i+1}(A, B) \leq 1/(\underline{\Sigma}_{ii} - \rho)$ holds for $i = 1, \ldots, q$.*

Proof Let L be as in the proof of Corollary 5. The relation $\sigma_{r^*-i+1}(A, B) = 1/\sigma_i(BL^{-T})$, the application of Theorem 10 by setting A, $\hat{U}$, $\hat{\Sigma}$ and $\hat{V}$ in this theorem as BL^{-T}, $\hat{U}$, $\hat{\Sigma}$ and $L^T\hat{V}$, respectively, and the similar derivation to (5.1) give the result. □

Corollary 8 *Let A, B, $\hat{V}$, $\hat{F}$, q and r^* be as in Corollary 5. Define $\hat{D}, \hat{E} \in \mathbb{R}^{n\times n}$ so that $\hat{D} + \hat{E} = (B\hat{V})^T(B\hat{V})$ and $\hat{D}$ is diagonal. Let $\nu : \{1, \ldots, n\} \to \{1, \ldots, n\}$ and r_ζ denote a numbering and an index satisfying $\hat{D}_{\nu(1)\nu(1)} \geq \cdots \geq \hat{D}_{\nu(n)\nu(n)}$ and $\hat{D}_{\nu(r_\zeta)\nu(r_\zeta)} > \|\hat{E}\|_\infty \geq \hat{D}_{\nu(r_\zeta+1)\nu(r_\zeta+1)}$, respectively. Let $\underline{\zeta}_i^B$ and $\overline{\zeta}_i^B$ be defined similarly to $\underline{\zeta}_i^M$ and $\overline{\zeta}_i^M$ in Theorem 11, respectively. If $\|\hat{F}\|_2 < 1$, then* $\operatorname{rank}(A) = n$ *and $1/\overline{\zeta}_{\nu(i)}^B \leq \sigma_{r^*-i+1}(A, B) \leq 1/\underline{\zeta}_{\nu(i)}^B$ follow for $i = 1, \ldots, r_\zeta$. If $r_\zeta = q$ in particular, $1/\overline{\zeta}_{\nu(i)}^B \leq \sigma_{q-i+1}(A, B) \leq 1/\underline{\zeta}_{\nu(i)}^B$ follows for $i = 1, \ldots, q$.*

Proof Let L be as in the proof of Corollary 5. The relation $\sigma_{r^*-i+1}(A, B) = 1/\sigma_i(BL^{-T})$ and application of Theorem 11 by setting A and V in this theorem as BL^{-T} and $L^T\hat{V}$, respectively, yield the result. □

6 An algorithm to obtain β

For computing verified bounds for the generalized singular values based on Corollaries 1, 5 and 7, we need to obtain β in these corollaries. In this section, we introduce a fast algorithm to obtain β in Corollary 1. We can obtain β in Corollaries 5 and 7 completely analogously by this algorithm. We present Theorem 13 for this purpose.

Theorem 13 *Let $B \in \mathbb{R}^{p \times n}$ with $p \geq n$,* $\operatorname{rank}(B) = n$*, $q \in \{1, \infty\}$ and $B_c \in \mathbb{R}^{n \times n}$ be symmetric. Assume the floating point Cholesky factorization $B_c \approx \tilde{L}\tilde{L}^T$ runs to completion, where $\tilde{L} \in \mathbb{R}^{n \times n}$ is lower triangular. Let $X_L \in \mathbb{R}^{n \times n}$ be an approximate inverse of $\tilde{L}$ whose columns $X_L e^{(n,i)}$ are computed by substitution, in any order, of n linear systems $L(X_L e^{(n,i)}) = e^{(n,i)}$, $i = 1, \ldots, n$, $s^{(n)} := (1, \ldots, 1)^T \in \mathbb{R}^n$,*

$$\rho_q := \gamma_n \| |X_L||\tilde{L}|s^{(n)}\|_q + \frac{n\underline{\mathbf{u}}}{1 - n\mathbf{u}} \|ns^{(n)} + \operatorname{diag}(|\tilde{L}|)\|_q, \quad \tau_q := \frac{\|X_L\|_q}{1 - \rho_q},$$

$$\tau_C := \gamma_{n+1} \| |\tilde{L}||\tilde{L}^T|s^{(n)}\|_\infty + \frac{n\underline{\mathbf{u}}}{1 - (n-1)\mathbf{u}} \|(n-1)s^{(n)} + \operatorname{diag}(|\tilde{L}|)\|_\infty.$$

If $\tau_1 \tau_\infty (\tau_C + \|B^T B - B_c\|_\infty) < 1$, it holds that

$$\|B^+\|_2 \leq \sqrt{\frac{\tau_1 \tau_\infty}{1 - \tau_1 \tau_\infty (\tau_C + \|B^T B - B_c\|_\infty)}}.$$

Proof From $\tau_1 \tau_\infty \tau_C < 1$ and [7, Theorem 10], we have $\|B_c^{-1}\|_2 \leq \tau_1 \tau_\infty / (1 - \tau_1 \tau_\infty \tau_C)$. This, $p \geq n$, $\operatorname{rank}(B) = n$ and Lemma 2 yield

$$\begin{aligned}
\sigma_n(B)^2 &= \sigma_n(B^T B) = \sigma_n(B_c) + \sigma_n(B^T B) - \sigma_n(B_c) \\
&\geq \sigma_n(B_c) - |\sigma_n(B^T B) - \sigma_n(B_c)| \geq \sigma_n(B_c) - \|B^T B - B_c\|_2 \\
&\geq \sigma_n(B_c) - \|B^T B - B_c\|_\infty = \frac{1}{\|B_c^{-1}\|_2} - \|B^T B - B_c\|_\infty \\
&\geq \frac{1 - \tau_1 \tau_\infty \tau_C}{\tau_1 \tau_\infty} - \|B^T B - B_c\|_\infty.
\end{aligned}$$

This and $\|B^+\|_2 = 1/\sigma_n(B)$ prove the theorem. □

In the practical application of Theorem 13, B_c is the center of the inclusion of $B^T B$. Thus B_c, $\tilde{L}$ and the inclusion of $B^T B$ can be obtained in "the other parts" of the algorithm based on Corollary 1. By utilizing these matrices, computational cost for obtaining β is $n^3/3 + \mathcal{O}(n^2)$, since the computation of X_L requires $n^3/3$ operations.

7 Numerical results for generalized singular values

In this section, we report numerical results for the generalized singular values. We used the same computer as that in Sect. 4. Throughout this section, let r_δ, r_ε, r_ρ and r_ζ be as in Corollaries 5, 6, 7 and 8, respectively. The compared algorithms are as follows:

M1A: The algorithm based on Corollary 1,
M2A: The algorithm based on Corollary 2,
M3A: The algorithm based on Corollary 3,
M4A: The algorithm based on Corollary 4,
M1B: The algorithm based on Corollary 5,

M2B: The algorithm based on Corollary 6,
M3B: The algorithm based on Corollary 7,
M4B: The algorithm based on Corollary 8.

The functions `svd` and `eig` are adopted analogously to Sect. 4. The notation MO is defined similarly to Sect. 4. The Cholesky factorization is executed via the function `chol`. In tables below, the notation fail2 means that the algorithms failed because $\|\hat{F}\|_2 < 1$ or $\|F\|_2 < 1$ could not be verified, where $\hat{F}$ and F are defined as in Corollaries 1, 2, 5 or 6.

7.1 Example 1

In this example, we observe computing times of the algorithms for various m, n and p. Consider the case when $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times n}$ are generated by `A = randn(m,n);` and `B = randn(p,n);`, respectively. Table 7 displays the computing times of the algorithms for various m, n and p. When $(m, n, p) = (3000, 1000, 300), (3000, 3000, 300), (10000, 3000, 300)$, we executed M1B, M2B, M3B and M4B only, since the other algorithms are not applicable. In this example, $r_\delta = r_\varepsilon = r_\rho = r_\zeta = \min(p, n)$ followed in all the cases when the algorithms succeeded.

The computing times of M1A and M3A were comparable. The same can be said for M1B and M3B. The computing times of M2A and M2B were larger than those of M1A and M3A, and M1B and M3B, respectively. This result coincides with the discussion in Sect. 5. When $n \le p$, M4A and M4B were faster than M1A, M2A and M3A, and M1B, M2B and M3B, respectively. When $n > p$, on the other hand, M4B was slower than M1B, M2B and M3B.

7.2 Example 2

In this example, we observe the magnitudes of the radii for various $\kappa(A)$. Consider the case when $A \in \mathbb{R}^{1000\times 10}$ and $B \in \mathbb{R}^{1000\times 10}$ are generated by `A =`

Table 7 Computing times (s) in Sect. 7.1

m	n	p	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
2000	300	1000	1.3104	10.809	1.2162	0.8197	0.9630	2.2682	0.8989	0.6319
1000	300	2000	0.9738	2.3588	0.9105	0.6273	1.2783	10.449	1.1967	0.8037
2000	300	2000	1.3593	10.879	1.2967	0.8613	1.3864	10.542	1.2858	0.8787
10000	300	2000	4.6824	MO	4.4253	2.7647	1.8935	11.015	1.7763	1.3769
2000	300	10000	1.9058	11.406	1.8139	1.3782	4.7023	MO	4.4114	2.8252
10000	300	10000	5.2191	MO	4.8989	3.2776	5.2155	MO	4.9094	3.3011
1000	1000	1000	18.514	18.613	17.557	9.9175	18.328	18.468	17.348	9.9185
2000	2000	2000	142.89	144.76	137.26	75.783	142.79	144.80	136.74	75.299
2000	1000	300	–	–	–	–	3.3226	5.2596	3.2534	8.2147
2000	2000	300	–	–	–	–	12.077	28.514	11.697	58.356
10000	2000	300	–	–	–	–	32.255	48.781	32.061	81.392

Table 8 Maximum (upper part) and minimum (lower part) radii in Sect. 7.2

`cndA`	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
1e+0	1.8e–14	8.8e–15	4.6e–15	4.2e–15	1.5e–14	8.0e–15	3.9e–15	5.6e–15
1e+2	6.8e–15	7.6e–15	3.6e–15	3.3e–15	1.6e–9	1.9e–12	1.9e–12	3.8e–12
1e+4	6.5e–15	7.7e–15	3.9e–15	3.4e–15	1.3e–3	2.0e–8	2.0e–8	4.0e–8
1e+6	5.1e–15	7.3e–15	3.2e–15	3.0e–15	8.5e–7	9.5e–5	9.5e–5	1.8e–4
1e+8	4.6e–15	7.2e–15	3.3e–15	5.1e–12	fail2	fail2	fail2	fail2
1e+0	1.7e–14	7.5e–15	3.9e–15	2.9e–15	1.1e–14	6.7e–15	3.2e–15	4.7e–15
1e+2	3.5e–15	9.0e–17	2.7e–16	3.1e–17	1.8e–13	1.9e–14	1.9e–14	3.8e–14
1e+4	3.1e–15	3.1e–18	4.5e–16	2.5e–18	1.4e–11	1.9e–12	1.9e–12	3.8e–12
1e+6	2.1e–15	2.0e–18	1.3e–16	1.8e–18	1.2e–9	9.4e–11	9.4e–11	1.8e–10
1e+8	1.6e–15	1.4e–18	1.7e–16	1.6e–18	fail2	fail2	fail2	fail2

Table 9 Obtained r_δ, r_ε, r_ρ and r_ζ in Sect. 7.2

`cndA`	r_δ	r_ε	r_ρ	r_ζ
1e+0	10	10	10	10
1e+2	10	10	10	10
1e+4	10	10	10	10
1e+6	3	10	10	10
1e+8	fail2	fail2	fail2	fail2

`gallery('randsvd', [1000,10],cndA);` and `B = randn(1000,10);`, respectively. Tables 8 and 9 display the maximum and minimum radii, and r_δ, r_ε, r_ρ and r_ζ, respectively, for various `cndA`.

The algorithms M1A, M2A, M3A and M4A gave small radii. The radii by M1B, M2B, M3B and M4B were large when `cndA` was large. The reason is that $\|\hat{F}\|_2$ and $\|\hat{E}\|_2$ in Corollary 5, $\|F\|_2$ and $|E_{ij}|$ in Corollary 6, ρ in Corollary 7 and $|\hat{E}_{ij}|$ in Corollary 8 were large when `cndA` was large. It is guessed that the enlargement of these values are caused by the loss of accuracy of $\hat{V}$ and V. The radii by M1B were larger than those by M2B, M3B and M4B, and r_δ became less than 10 when `cndA` = 1e+6. The reason is that M1B requires the upper bound for $\|A^+\|_2$, although M2B, M3B and M4B do not require it.

7.3 Example 3

In this example, we observe the magnitudes of the radii for various $\kappa(B)$. Consider the case when $A \in \mathbb{R}^{1000\times 10}$ and $B \in \mathbb{R}^{1000\times 10}$ are generated by `A = randn(1000,10);` and `B = gallery('randsvd',[1000,10],cndB);`, respectively. Tables 10 and 11 display the similar quantities to Tables 8 and 9, respectively, for various `cndB`.

Table 10 Maximum (upper part) and minimum (lower part) radii in Sect. 7.3

cndB	M1A	M2A	M3A	M4A	M1B	M2B	M3B	M4B
1e+0	1.2e–11	7.8e–12	3.8e–12	3.4e–12	1.7e–11	7.8e–12	3.8e–12	5.6e–12
1e+2	3.3e–6	3.1e–7	3.1e–7	3.1e–7	4.2e–8	7.5e–10	2.1e–9	5.8e–10
1e+4	1.1e+0	1.2e–1	1.2e–1	1.2e–1	2.9e–4	7.2e–8	5.4e–5	5.0e–8
1e+6	9.8e+5	1.5e+5	1.5e+5	1.5e+5	2.4e+0	7.4e–6	6.0e–1	1.1e–1
1e+8	fail2	fail2	fail2	fail2	5.1e+4	2.4e–3	1.3e+3	9.3e+5
1e+0	1.2e–11	6.9e–12	3.3e–12	2.4e–12	1.3e–11	6.7e–12	3.2e–12	4.8e–12
1e+2	3.0e–6	3.2e–9	3.2e–9	3.1e–9	7.9e–12	7.6e–12	3.8e–12	5.8e–12
1e+4	9.0e–1	1.2e–5	1.2e–5	1.2e–5	5.6e–12	6.8e–12	3.5e–12	4.7e–12
1e+6	8.3e+5	1.5e–1	1.5e–1	1.5e–1	5.8e–12	7.4e–12	4.0e–12	5.5e–12
1e+8	fail2	fail2	fail2	fail2	9.4e–12	7.5e–12	3.6e–12	5.6e–12

Table 11 Obtained r_δ, r_ε, r_ρ and r_ζ in Sect. 7.3

cndB	r_δ	r_ε	r_ρ	r_ζ
1e+0	10	10	10	10
1e+2	10	10	10	10
1e+4	10	10	10	10
1e+6	10	10	10	10
1e+8	10	10	10	9

The radii by M1A, M2A, M3A and M4A in this section showed tendencies analogous to those of the radii by M1B, M2B, M3B and M4B in Sect. 7.2, respectively. The radii by M1B, M2B, M3B and M4B in this section were larger than those by M1A, M2A, M3A and M4A in Sect. 7.2 when `cndB` was large. The reason is as follows: Let L be as in Corollary 5. The smallest singular value of BL^{-T} becomes close to zero as `cndB` increases. Hence $\underline{\delta}^B_{r_\delta}$, $\overline{\delta}^B_{r_\delta}$, $\underline{\varepsilon}^B_{\nu(r_\varepsilon)}$, $\overline{\varepsilon}^B_{\nu(r_\varepsilon)}$, $\underline{\Sigma}_{r_\rho r_\rho} - \rho$, $\overline{\Sigma}_{r_\rho r_\rho} + \rho$, $\underline{\zeta}^B_{\nu(r_\zeta)}$ and $\overline{\zeta}^B_{\nu(r_\zeta)}$ in Corollaries 5, 6, 7 and 8 become also close to zero, so that the radii are enlarged by taking reciprocals. For instance, let $\underline{\delta}^B_{r_\delta}$ =1.0e–8 $-$ 1.0e–15 and $\overline{\delta}^B_{r_\delta}$ =1.0e–8 $+$ 1.0e–15. Then $(1/\underline{\delta}^B_{r_\delta} - 1/\overline{\delta}^B_{r_\delta})/2 \approx$1.0e+1, although $(\overline{\delta}^B_{r_\delta} - \underline{\delta}^B_{r_\delta})/2$ =1.0e–15. When `cndB` $=$ 1e+8, r_ζ became less than 10, although $r_\delta = r_\varepsilon = r_\rho$ =10. The reason is that M4B is based on the bound for the eigenvalues of $L^{-1}B^TBL^{-T}$, i.e. *square* of the singular values of BL^{-T}, as opposed that M1B, M2B and M3B are based on the bound for the singular values of BL^{-T}.

7.4 Example 4

In this example, we observe the magnitudes of the radii when A is rank deficient. Consider the case when $A \in \mathbb{R}^{10\times 3}$ and $B \in \mathbb{R}^{10\times 3}$ are generated by `A = repmat(randn (10,1),1,3);` and `B = randn(10,3);`, respectively. We

Table 12 Obtained radii in Sect. 7.4

M1A			M2A		
8.6e–15	4.7e–15	4.7e–15	6.0e–15	1.1e–15	1.1e–15
M3A			M4A		
5.3e–15	1.4e–15	1.4e–15	3.8e–15	1.2e–8	1.2e–8

executed M1A, M2A, M3A and M4A only. Table 12 displays analogous quantities to Table 3 obtained by these algorithms, showing the similar tendency to that in Sect. 4.3.

7.5 Example 5

In this example, we observe the magnitudes of the radii when B is rank deficient. Consider the case when $A \in \mathbb{R}^{10\times 3}$ and $B \in \mathbb{R}^{10\times 3}$ are generated by `A = randn(10,3);` and `B = repmat(randn(10,1),1,3);`, respectively. In this case, A and B have only one generalized singular value. We executed M1B, M2B, M3B and M4B only. The radii obtained by M1B, M2B, M3B and M4B are 4.5e–15, 1.9e–15, 2.1e–15 and 2.8e–15, respectively. We moreover obtained $r_\delta = r_\varepsilon = r_\rho = r_\zeta = 1$. Thus the verified bound for the one generalized singular value could be computed by the executed algorithms.

8 Conclusion

In this paper, we proposed algorithms for computing verified bounds of all the singular values, and reported some numerical results. We moreover showed that the proposed algorithms give equal or tighter bounds than those by the previous algorithms, and introduced the application of the proposed algorithms.

By modifying the proposed algorithms slightly, verified bounds of all the singular values, where A are complex and/or interval, can be computed. By utilizing [7, Theorem 7], verified bounds of all columns of X in Theorem 2 can be obtained, since these columns are the eigenvectors corresponding to the generalized eigenvalues of A^TA and B^TB. Our future work will be the open challenge discussed in Remark 5, to develop an algorithm for computing verified bounds of the generalized singular values in the case when $m < n$ or the both of A^TA and B^TB are singular, and to propose a new algorithm for computing verified bounds for *few* of the largest or smallest singular values.

Acknowledgments The author wishes to thank Prof. Siegfried M. Rump in Hamburg University of Technology for his comments, and sincerely acknowledges the referee for the suggestion which significantly improved the first version of this paper.

References

1. Alefeld, G., Hoffmann, R., Mayer, G.: Verification algorithms for generalized singular values. Math. Nachr. **208**, 5–29 (1999)

2. Davis, T.A.: The university of Florida sparse matrix collection. http://www.cise.ufl.edu/research/sparse/matrices/
3. Fernando, K.V., Parlett, B.N.: Accurate singular values and differential qd algorithms. Numer. Math. **67**, 191–229 (1994)
4. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press, Baltimore (1996)
5. Higham, N.: Accuracy and Stability of Numerical Algorithms, 2nd edn. SIAM Publications, Philadelphia (2002)
6. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1994)
7. Miyajima, S., Ogita, T., Rump, S.M., Oishi, S.: Fast verification for all eigenpairs in symmetric positive definite generalized eigenvalue problems. Reliab. Comput. **14**, 24–45 (2010)
8. Nakatsukasa, Y., Aishima, K., Yamazaki, I.: dqds with aggressive early deflation. SIAM. J. Matrix Anal. Appl. **33**(1), 22–51 (2012)
9. Oishi, S.: Fast enclosure of matrix eigenvalues and singular values via rounding mode controlled computation. Linear Algebra Appl. **324**, 133–146 (2001)
10. Parlett, B.N.: The Symmetric Eigenvalue Problem. SIAM Publications, Philadelphia (1997)
11. Parlett, B.N., Marques, O.A.: An implementation of the dqds algorithm (positive case). Linear Algebra Appl. **309**, 217–259 (2000)
12. Paige, C.C., Sauders, M.A.: Towards a generalized singular value decomposition. SIAM J. Numer. Anal. **18**(3), 398–405 (1981)
13. Rohn, J.: VERSOFT: Verification software in MATLAB / INTLAB. http://uivtx.cs.cas.cz/rohn/matlab/
14. Rump, S.M.: Verified bounds for singular values, in particular for the spectral norm of a matrix and its inverse. BIT Numer. Math. **51**(2), 367–384 (2011)
15. Van Loan, C.F.: Generalizing the singular value decomposition. SIAM J. Numer. Anal. **13**(1), 76–83 (1976)