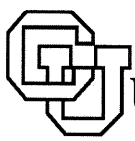
On Accurate Generalized Singular Value Computation in Floating-point Arithmetic

Zlatko Drmac
E. R. Jessup

CU-CS-811-96



University of Colorado at Boulder
DEPARTMENT OF COMPUTER SCIENCE

1

On accurate generalized singular value computation in floating–point arithmetic

Zlatko Drmač E. R. Jessup

CU-CS-811-96 October 1996

University of Colorado at Boulder

Technical Report CU-CS-811-96
Department of Computer Science
Campus Box 430
University of Colorado
Boulder, Colorado 80309

On accurate generalized singular value computation in floating–point arithmetic *

Zlatko Drmač† E. R. Jessup‡

October 9, 1996

Abstract

In this paper we present a new algorithm for floating–point computation of the generalized singular value decomposition of an arbitrary matrix pair $(A, B) \in \mathbf{R}^{m\times n} \times \mathbf{R}^{p\times n}$. In the case of full column rank A, the new algorithm computes all finite generalized singular values with high relative accuracy if $\min\{\kappa_2(AD),\ D \text{ diagonal}\}$ is moderate and if an accurate rank revealing LU factorization of B is possible. Numerical experiments show that, in that case, the new algorithm computes the generalized singular values of all pairs $\{(AD, D_1BD_2),\ D, D_1, D_2 \text{ diagonal matrices}\}$ with nearly the same relative accuracy.

1 Introduction

In [35], Van Loan introduces a new matrix decomposition of a general matrix pair $(A, B) \in \mathbf{C}^{m\times n} \times \mathbf{C}^{p\times n}$ $(m \geq n)$. He proves that there always exist unitary matrices U, V and a nonsingular matrix X such that U^*AX and V^*BX are diagonal matrices, and he defines the B–singular values of A as the elements of the set $\{\sigma \geq 0 \ : \ \det(A^*A - \sigma^2 B^*B) = 0\}$. Paige and Saunders [27] remove the minor constraint $m \geq n$ and reformulate the original decomposition to avoid the non–unitary matrix X. They show that there exist unitary matrices U, V, Q, diagonal matrices Σ_A, Σ_B, and a nonsingular triangular matrix R such that $U^*AQ = \Sigma_A[\mathbf{O}, R]$, $V^*BQ = \Sigma_B[\mathbf{O}, R]$. This form is equivalent to Van Loan's with $X = Q(I \oplus R^{-1})$. In either formulation, the new decomposition is called the *generalized singular value decomposition* (GSVD) of (A, B), and the B–singular values of A are the *generalized singular values* of (A, B). If B is square and nonsingular then the GSVD of (A, B) is equivalent to the ordinary singular value decomposition (SVD) of AB^{-1}.

The GSVD is a powerful tool in both theoretical analysis and numerical solution of problems like regularization and various types of constrained least squares [7], [18], [36], [37], [25]. It also arises in the symmetric definite generalized eigenvalue problem $Kx = \lambda Mx$, where the positive definite matrices K and M are factored as $K = A^*A$ and $M = B^*B$, respectively. The generalized singular values of (A, B) are then the square roots of the eigenvalues of $K - \lambda M$. An important advantage of using (A, B) instead of the pencil $K - \lambda M$ is that $\kappa_2(A) = \sqrt{\kappa_2(K)}$, $\kappa_2(B) = \sqrt{\kappa_2(M)}$. (Here $\kappa_2(A) = \|A\|_2\|A^\dagger\|_2$ is the spectral condition number, where $A^\dagger$ is the Moore–Penrose generalized inverse and $\|\cdot\|_2$ is operator norm induced by the Euclidean vector norm.)

The central issue in this paper is how to compute the generalized singular values of a real pair $(A, B) \in \mathbf{R}^{m\times n} \times \mathbf{R}^{p\times n}$ with high relative accuracy in floating–point arithmetic. This computation

*This research was supported by National Science Foundation grants ACS–9357812 and ASC–9625912, Department of Energy grant DE–FG03–94ER25215, and the Intel Corporation.

†Department of Computer Science, University of Colorado, Boulder CO 80309–0430. (zlatko@cs.colorado.edu)

‡Department of Computer Science, University of Colorado, Boulder CO 80309–0430. (jessup@cs.colorado.edu)

is not always possible. For instance, if ε is the roundoff unit and

$$A = \begin{bmatrix} 1 & 1+\zeta \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \tag{1.1}$$

then, for $|\zeta| < \varepsilon$, A is stored as exactly singular. In this case the smallest singular value of (A, B) is incorrectly computed as 0 due to rounding error in A_{12}. In fact, if each entry of A from (1.1) has an initial relative uncertainty of at most ε, then even the exact computation of the smallest generalized singular value of (A, B) provides no useful information.

There are, however, matrix pairs for which certain small matrix changes introduce only small generalized singular value perturbations. Those cases are identified by perturbation theory. (Cf. [2], [3], [8], [12], [14], [22], [23], [24], [31], [32].) In this paper we assume that the pair (A, B) determines all its generalized singular values well in the sense that their initial uncertainty due to the uncertainty in A and B is small. Furthermore, we assume that it is of interest to have approximations of the generalized singular values of (A, B) that are as accurate as possible. Such pairs do appear in the practice; see [5], [10] for more detailed discussion and examples. A desirable property of an algorithm is then to approximate the generalized singular values with a relative error not much larger than the initial uncertainty in (A, B).

There are several numerically attractive algorithms for the GSVD computation. The first one is based on a simple connection between the GSVD and the Cosine–Sine decomposition (CSD) of a partitioned orthonormal matrix, [29], [30], [38]. This algorithm first computes the QR factorization $\mathcal{G} \equiv \begin{bmatrix} A \\ B \end{bmatrix} = \mathcal{Q}R$ and then it computes the CSD of $\mathcal{Q} = \begin{bmatrix} \mathcal{Q}_1 \\ \mathcal{Q}_2 \end{bmatrix}$, where $\mathcal{Q}$ is partitioned in accordance with the partition of $\mathcal{G}$ so that $\mathcal{Q}_1 \in \mathbf{R}^{m\times n}$, $\mathcal{Q}_2 \in \mathbf{R}^{p\times n}$.

The second algorithm avoids the use of the $(m+p)\times n$ matrix $\mathcal{G}$ and transforms A and B separately. It has two phases: *(i)* using an algorithm of Bai and Zha [4], a general pair (A, B) is reduced to an equivalent pair of upper triangular matrices $(A_\triangleright, B_\triangleright)$ with nonsingular $B_\triangleright$; *(ii)* using an algorithm of Paige [26], implemented carefully by Bai and Demmel [3], the procedure completes by the GSVD computation of $(A_\triangleright, B_\triangleright)$. It is shown in [3], [4] that both phases of the algorithm are backward stable in the Frobenius matrix norm. That is, floating–point computation is equivalent to exact computation with $(A+\delta A, B+\delta B)$, where $\|\delta A\|_F/\|A\|_F$ and $\|\delta B\|_F/\|B\|_F$ are of order machine precision times a moderate function of matrix dimensions. This algorithm is superior to the CSD approach, and it is implemented as the LAPACK [1] procedure `SGGSVD()` for GSVD computation.

Both the CSD and the LAPACK algorithm are designed to use only orthogonal transformations. This restriction seems to be unnecessary because the generalized singular values are in fact invariant under the more general transformation $(A, B) \mapsto (A', B') = (U^\tau AX, V^\tau BX)$, where U, V are arbitrary orthogonal matrices and X is an arbitrary nonsingular matrix.

The first method for generalized singular value computation using nonorthogonal transformations is proposed in [9]. It is an implicit variant of the Falk–Langemeyer method [16] for the diagonalization of matrix pencils. An error analysis for the full column rank case is given in [8], and the method is further analyzed and modified in [14]. The second method that is not entirely based on orthogonal transformations is given in [14], [15]. In this method, a pair (A, B) of full column rank matrices is replaced by an equivalent pair (A', B'), and the SVD of the explicitly computed matrix $A'B'^{-1}$ is computed using the Jacobi SVD method [19], [12], [13].

Although based on nonorthogonal transformations, these two methods approximate the generalized singular values of a pair (A, B) of full column rank matrices with an error bound (cf. [8], [14], [15])

$$\max_{1\le i\le n} \frac{|\delta\sigma_i|}{\sigma_i} \le g(m,n,p)\cdot\varepsilon\cdot K_c(A,B), \quad K_c(A,B) = \kappa_2(AD_A^{-1}) + \kappa_2(BD_B^{-1}),$$

where $g(\cdot,\cdot,\cdot)$ is a modestly growing function of matrix dimensions, and D_A, D_B are diagonal matrices of Euclidean column norms of A and B, respectively. It is a remarkable fact that these

methods maintain the same accuracy in the family of all pairs (AD_1, BD_2), where D_1 and D_2 are arbitrary diagonal nonsingular matrices. This accuracy property is shared neither by the CSD nor the LAPACK procedure.

In this work we present a new algorithm that is capable of achieving the high relative accuracy on a set of matrix pairs that is much larger than the set of all (A, B) with moderate $K_c(A, B)$. For instance, we consider matrix pairs (A, B) where A is of full column rank with moderate $\kappa_2(A_c)$ and B is a matrix that can be accurately factored using the LU factorization with rank revealing (total) pivoting. We show how to reduce such a pair to an equivalent pair (A', B') of full column rank matrices with $K_c(A', B')$ not much larger than $\kappa_2(A_c)$. The generalized singular values of the new pair (A', B') are then computed by the method from [15].

The rest of the paper is organized as follows: In § 2 we show how condition numbers for the generalized singular value perturbations depend on different types of floating point matrix perturbations. In § 3 we briefly analyze LAPACK's [1] procedure `SGGSVD()` for the GSVD computation. In § 4 we present our new algorithm, and in § 5 we give detailed error and perturbation analyses that identify a set of input pairs for which computation with high relative accuracy is possible. We also show that our algorithm is capable of achieving that accuracy. Finally, in § 6 we present the results of rigorous numerical testing that demonstrate numerical robustness of our software. The numerical results correspond to the analysis from § 5, and they also indicate that, in the case of full column rank A and full (column or row) rank B, the algorithm computes the generalized singular values of all pairs $\{(AD, D_1BD_2),\ D, D_1, D_2$ diagonal matrices$\}$ with nearly the same relative accuracy. We recommend our algorithm as the method of choice for GSVD computation in floating–point arithmetic.

2 Floating point error analysis of a GSVD algorithm

A floating point algorithm for GSVD computation usually generates a sequence of matrix pairs $(A^{(k)}, B^{(k)})$, $k = 0, 1, \ldots$, that can be connected by commutative diagrams. Commutative diagrams provide systematic way to represent floating–point errors in a form that can easily be used in perturbation theory. In Figure 1, the computed pair $(A^{(k)}, B^{(k)})$ is the result of floating–point computation with the input $(A^{(k-1)}, B^{(k-1)})$. Equivalently, it is the result of exact computation starting with certain pair $(A^{(k-1)} + \delta A^{(k-1)}, B^{(k-1)} + \delta B^{(k-1)})$, where an estimate of $\delta A^{(k-1)}$, $\delta B^{(k-1)}$ is obtained by backward error analysis. In the forward error analysis, it is of interest to determine the distance between the floating-point result $(A^{(k+1)}, B^{(k+1)})$ and the exact result $(\bar{A}^{(k+1)}, \bar{B}^{(k+1)})$ obtained with the same input $(A^{(k)}, B^{(k)})$.

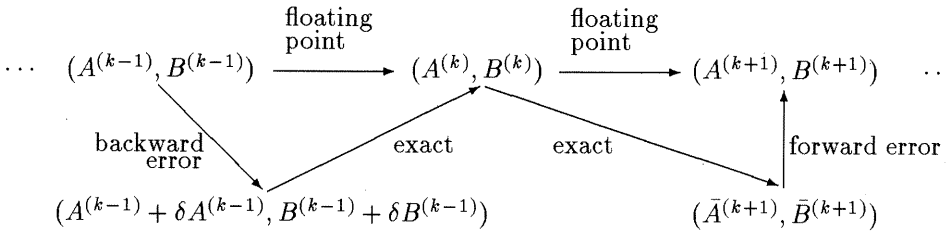


Figure 1: The backward and the forward error.

The relative accuracy of the computed generalized singular value approximation depends on the perturbations $\delta A^{(k)}$, $\delta B^{(k)}$ and on the corresponding condition numbers of $A^{(k)}$, $B^{(k)}$, $k = 0, 1, \ldots$. Here we note that floating–point errors and corresponding condition numbers strongly depend on the details of the algorithm. For instance, *(i)* if $A^{(k)}$ is obtained by changing only a few columns of $A^{(k-1)}$, then additional information on the zero pattern of $\delta A^{(k-1)}$ can be used to derive tight bounds for the condition number; *(ii)* if the new pair $(A^{(k)}, B^{(k)})$ is obtained by scaling the columns

of $(A^{(k-1)}, B^{(k-1)})$, then we have small elementwise relative perturbation; *(iii)* if $B^{(k)}$ is a triangular matrix obtained by the QR factorization of $B^{(k-1)}$, then the backward error analysis from [14] provides an estimate for each column of $\delta B^{(k-1)}$, $\|\delta B^{(k-1)} e_i\|_2 \ll \|B^{(k-1)} e_i\|_2$, $1 \le i \le n$, with e_i being the ith column of the identity matrix I.

The next task in the analysis is to derive a sharp estimate of the relevant condition number for a particular type of perturbation. The tool of trade in certain regular cases is the variational characterization of the generalized singular values: if B is full column rank matrix, then the nonincreasingly ordered generalized singular values $\sigma_1 \ge \cdots \ge \sigma_n$ of (A, B) satisfy

$$\sigma_i = \min_{\mathcal{X}_{n-i+1}} \max_{\substack{x \in \mathcal{X}_{n-i+1} \\ x \neq 0}} \frac{\|Ax\|_2}{\|Bx\|_2} = \max_{\mathcal{Y}_i} \min_{\substack{x \in \mathcal{Y}_i \\ x \neq 0}} \frac{\|Ax\|_2}{\|Bx\|_2}, \quad 1 \le i \le n, \tag{2.2}$$

where $\mathcal{X}_{n-i+1}$, $\mathcal{Y}_i$ are arbitrary $n-i+1$ and i dimensional subspaces of $\mathbf{R}^n$.

To use the relation (2.2) in perturbation estimates, first note that for full column rank A and B and sufficiently small $\|\delta A A^\dagger\|_2$ and $\|\delta B B^\dagger\|_2$ and all nonzero vectors x it holds that

$$\frac{1 - \|\delta A A^\dagger\|_2}{1 + \|\delta B B^\dagger\|_2} \frac{\|Ax\|_2}{\|Bx\|_2} \le \frac{\|(A+\delta A)x\|_2}{\|(B+\delta B)x\|_2} \le \frac{1 + \|\delta A A^\dagger\|_2}{1 - \|\delta B B^\dagger\|_2} \frac{\|Ax\|_2}{\|Bx\|_2}.$$

Now the variational characterization (2.2) implies that the ordered generalized singular values $\sigma_1 \ge \cdots \ge \sigma_n$ and $\tilde{\sigma}_1 \ge \cdots \ge \tilde{\sigma}_n$ of (A, B) and $(A+\delta A, B+\delta B)$, respectively, satisfy

$$\frac{1 - \|\delta A A^\dagger\|_2}{1 + \|\delta B B^\dagger\|_2} \le \frac{\tilde{\sigma}_i}{\sigma_i} \le \frac{1 + \|\delta A A^\dagger\|_2}{1 - \|\delta B B^\dagger\|_2}, \quad (\tilde{\sigma}_i = \sigma_i + \delta\sigma_i) \quad 1 \le i \le n. \tag{2.3}$$

An application of the relation (2.3) depends on how δA, δB are measured relative to A, B, respectively. For example, if $|\delta A| \le \eta |A|$, $|\delta B| \le \eta |B|$ are small elementwise relative perturbations of A and B then the relative perturbations $\delta\sigma_i$, $1 \le i \le n$, are bounded by

$$\frac{|\delta\sigma_i|}{\sigma_i} \le \eta \frac{\| \, |A| \cdot |A^\dagger| \, \|_2 + \| \, |B| \cdot |B^\dagger| \, \|_2}{1 - \eta \| \, |B| \cdot |B^\dagger| \, \|_2}, \quad 1 \le i \le n, \tag{2.4}$$

provided that η is sufficiently small. (The absolute value and inequalities are taken elementwise.) If we allow the more general relative perturbation $\|\delta A e_i\|_2 \le \eta \|A e_i\|_2$, $\|\delta B e_i\|_2 \le \eta \|B e_i\|_2$, $1 \le i \le n$, then the relation (2.3) implies

$$\frac{|\delta\sigma_i|}{\sigma_i} \le \frac{\frac{\|\delta A_c\|_2}{\|A_c\|_2}\kappa_2(A_c) + \frac{\|\delta B_c\|_2}{\|B_c\|_2}\kappa_2(B_c)}{1 - \frac{\|\delta B_c\|_2}{\|B_c\|_2}\kappa_2(B_c)} \le \sqrt{n}\eta \frac{\|A_c^\dagger\|_2 + \|B_c^\dagger\|_2}{1 - \sqrt{n}\eta \|B_c^\dagger\|_2}, \quad 1 \le i \le n, \tag{2.5}$$

where we use diagonal scalings $D_A = \mathrm{diag}\,(\|Ae_i\|_2)$, $D_B = \mathrm{diag}\,(\|Be_i\|_2)$ to define $A_c = A D_A^{-1}$, $\delta A_c = \delta A D_A^{-1}$, $B_c = B D_B^{-1}$, $\delta B_c = \delta B D_B^{-1}$. If the size of the perturbation is bounded only by $\|\delta A\|_2 \le \eta \|A\|_2$, $\|\delta B\|_2 \le \eta \|B\|_2$, then the relative error bound reads

$$\frac{|\delta\sigma_i|}{\sigma_i} \le \eta \frac{\kappa_2(A) + \kappa_2(B)}{1 - \eta\kappa_2(B)}, \quad 1 \le i \le n. \tag{2.6}$$

Relations (2.4), (2.5), (2.6) reveal three different condition numbers that relate the perturbation of the pair (A, B) of full column rank matrices to the generalized singular value perturbation. In Table (2.7), we summarize the above discussion and we also display some important relations between the three condition numbers derived from (2.3).

Perturbation $(\delta A, \delta B)$	Condition number $K.(A,B)$
$\|\delta A\| \le \eta\|A\|,\ \|\delta B\| \le \eta\|B\|$ $\|\|\delta A e_i\|_2 \le \eta\|\|Ae_i\|_2,\ \|\|\delta B e_i\|_2 \le \eta\|\|Be_i\|_2$ $\|\|\delta A\|_2 \le \eta\|\|A\|_2,\ \|\|\delta B\|_2 \le \eta\|\|B\|_2$	$K_{BS}(A,B) = \|\ \|A\|\cdot\|A^\dagger\|\ \|_2 + \|\ \|B\|\cdot\|B^\dagger\|\ \|_2$ $K_c(A,B) = \kappa_2(A_c) + \kappa_2(B_c)$ $K_2(A,B) = \kappa_2(A) + \kappa_2(B)$
Generalized singular value perturbation: $\max_{1\le i\le n} \frac{\lvert\delta\sigma_i\rvert}{\sigma_i} \le K.(A,B)\eta + O(\eta^2)$	
Always: $\|\ \|A\|\cdot\|A^\dagger\|\ \|_2 \le n \min_{D=\mathrm{diag}}\ \kappa_2(AD)$, $\kappa_2(A_c) \le \sqrt{n}\min_{D=\mathrm{diag}}\ \kappa_2(AD)$	
Possible: $\|\ \|A\|\cdot\|A^\dagger\|\ \|_2 \ll \kappa_2(A_c)$, $\kappa_2(A_c) \ll \kappa_2(A)$	

(2.7)

The estimate $\kappa_2(A_c) \le \sqrt{n}\min_{D=\mathrm{diag}}\ \kappa_2(AD)$ in (2.7), where the minimum is taken over the set of $n \times n$ diagonal nonsingular matrices, is proven by van der Sluis [34]. The proof of the remaining relations in Table (2.7) is straightforward. The subscript "BS" in K_{BS} indicates the dependence on the Bauer–Skeel condition numbers of A and B, cf. [6], [28]. Similarly, the subscript "c" is a reminder that K_c depends on the spectral condition number of the column scaled matrices A_c and B_c.

An important difference between $K_{BS}(A,B)$, $K_c(A,B)$ and $K_2(A,B)$ is that $K_{BS}(A,B) = K_{BS}(AD_1, BD_2)$ and $K_c(A,B) = K_c(AD_1, BD_2)$ for any diagonal nonsingular matrices D_1, D_2, while, on the other hand, $K_2(A,B)$ depends on such diagonal scalings. Furthermore, $K_{BS}(A,B)$ and $K_c(A,B)$ are never much larger than and can be much smaller than $K_2(A,B)$.

Hence, we expect better numerical properties (reliability, high accuracy in larger domain of input pairs) of an algorithm that produces floating–point errors yielding to an application of (2.4) and (2.5) rather than (2.6).

3 LAPACK's GSVD algorithm

In the LAPACK library [1], the procedure `SGGSVD` for the GSVD computation has two stages: *(i)* reduction of a general pair (A,B) to a regular pair (A',B') of upper triangular matrices; *(ii)* GSVD computation of the regular pair (A',B'). (The pair (A',B') is called *regular* if B' is a full column rank matrix.) Stage *(i)* of `SGGSVD` uses an algorithm of Bai and Zha [4], while stage *(ii)* is a careful implementation of an algorithm of Paige [26], [3]. Working with a regular pair has several advantages in Paige's algorithm because its implementation in the case of an irregular triangular pair is quite complicated [4], [3]. Bai and Zha's reduction algorithm is based on the QR factorization and the URV decomposition

$$A = U \begin{bmatrix} \mathbf{O} & R \\ \mathbf{O} & \mathbf{O} \end{bmatrix} V^T, \quad U, V \text{ orthogonal}, \quad R \text{ triangular nonsingular.}$$

Since Paige's algorithm is based on plane rotations, the whole process is therefore completed using solely orthogonal transformations:

ALGORITHM **3.1** *(Description of the LAPACK's procedure* `SGGSVD`*)*

Input $(A^{(0)}, B^{(0)}) \equiv (A,B) \in \mathbf{R}^{m\times n} \times \mathbf{R}^{p\times n}$, $\mathrm{rank}\, A = r_A$, $\mathrm{rank}\, B = r_B$.

Stage *(i)* Reduction. (Bai and Zha [4])

Step 1 Compute the URV decomposition of $B^{(0)}$ and replace $B^{(0)}$ by

$$B^{(0)} \overset{URV}{\longmapsto} \begin{bmatrix} \mathbf{O} & B_{12}^{(1)} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad B_{12}^{(1)} \in \mathbf{R}^{r_B \times r_B}.$$

Update and partition $A^{(0)}$ accordingly. Replace the pair $(A^{(0)}, B^{(0)})$ by the equivalent pair

$$(A^{(1)}, B^{(1)}) = ([A_{11}^{(1)}, A_{12}^{(1)}], \begin{bmatrix} \mathbf{O} & B_{12}^{(1)} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}), \quad A_{11}^{(1)} \in \mathbf{R}^{m \times (n-r_B)}.$$

Step 2 If $A_{11}^{(1)}$ is not empty ($r_B < n$), compute the URV decomposition of $A_{11}^{(1)}$ and replace $A_{11}^{(1)}$ by

$$A_{11}^{(1)} \overset{URV}{\longmapsto} \begin{bmatrix} \mathbf{O} & A_{12}^{(2)} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad A_{12}^{(2)} \in \mathbf{R}^{r_A^{(1)} \times r_A^{(1)}}, \quad r_A^{(1)} = \operatorname{rank} A_{11}^{(1)}.$$

Update and partition $A_{12}^{(1)}$ accordingly. Repartition $B^{(1)}$. The new pair reads

$$(A^{(2)}, B^{(2)}) = (\begin{bmatrix} \mathbf{O} & A_{12}^{(2)} & A_{13}^{(2)} \\ \mathbf{O} & \mathbf{O} & A_{23}^{(2)} \end{bmatrix}, \begin{bmatrix} \mathbf{O} & \mathbf{O} & B_{13}^{(2)} \\ \mathbf{O} & \mathbf{O} & \mathbf{O} \end{bmatrix}), \quad B_{13}^{(2)} = B_{12}^{(1)}.$$

Step 3 Compute the QR factorization of $A_{23}^{(2)}$ to get an upper triangular or upper trapezoidal $s \times r_B$ matrix $A_{23}^{(3)}$. If $s \geq r_B$, define $A_{23\,\triangleright}^{(3)}$ to be the leading $r_B \times r_B$ submatrix of $A_{23}^{(3)}$. If $s < r_B$, append $r_B - s$ zero rows to $A_{23}^{(3)}$ and denote the resulting matrix by $A_{23\,\triangleright}^{(3)}$.

Return The reduced regular pair reads $(A_{23\,\triangleright}^{(3)}, B_{13}^{(3)})$.

Stage *(ii)* GSVD of a regular pair of triangular matrices. (Paige [26], Bai and Demmel [3])

Step 1 Apply the algorithm from [3] to compute the GSVD of the regular pair $(A_{23\,\triangleright}^{(3)}, B_{13}^{(3)})$,

$$U_1^T A_{23\,\triangleright}^{(3)} Q_1 = \Xi_A R_1, \quad V_1^T B_{13}^{(3)} Q_1 = \Xi_B R_1.$$

Return The matrices U_1, V_1, Q_1, Ξ_A, Ξ_B.

Output Assemble all transformations to get the GSVD in the form introduced by Paige and Saunders:

$$U^T A Q = \Sigma_A [\mathbf{O}, R], \quad V^T B Q = \Sigma_B [\mathbf{O}, R], \quad \text{where} \tag{3.8}$$

$$\Sigma_A = \begin{bmatrix} I & \mathbf{O} \\ \mathbf{O} & \Xi_A \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad \Sigma_B = \begin{bmatrix} \mathbf{O} & \Xi_B \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad R = \begin{bmatrix} A_{12}^{(3)} & A_{13}^{(3)} Q_1 \\ \mathbf{O} & R_1 \end{bmatrix}.$$

For Van Loan's decomposition postmultiply (3.8) by $I \oplus R^{-1}$.

REMARK **3.1** The URV decomposition in Algorithm 3.1 is computed using the QR factorization with column pivoting and the RQ factorization.

3.1 An example of forward instability

Algorithm 3.1 is backward stable; see [3], [4]. However, backward stability in the matrix norm sense ($\|\delta A\|_F / \|A\|_F \ll 1$, $\|\delta B\|_F / \|B\|_F \ll 1$) does not guarantee high relative accuracy. We illustrate this fact via a simple example.

EXAMPLE **3.1** Let

$$A = \begin{bmatrix} 1 & -\alpha \\ 1 & \alpha \end{bmatrix}, \quad B = [\beta, \ \beta]. \tag{3.9}$$

The URV decomposition of B is achieved by a single rotation G with the angle $\pi/4$:

$$G \equiv \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}, \quad (A^{(1)}, B^{(1)}) = (AG, BG) = (\begin{bmatrix} \frac{1+\alpha}{\sqrt{2}} & \frac{1-\alpha}{\sqrt{2}} \\ \frac{1-\alpha}{\sqrt{2}} & \frac{1+\alpha}{\sqrt{2}} \end{bmatrix}, [0,\ \sqrt{2}\beta]).$$

Using $BA^{-1} = [\beta/\sqrt{2}, \beta/(\sqrt{2}/\alpha)]G$, we easily compute the generalized singular values of (A, B): $\sigma_1 = +\infty$, $\sigma_2 = \sqrt{2}\alpha/\beta(1+\alpha^2)^{-1/2}$. Note that σ_1 and σ_2 are perfectly well determined by α and β in the sense that a perturbation $\alpha \mapsto \alpha + \delta\alpha$, $\beta \mapsto \beta + \delta\beta$ with $|\delta\alpha/\alpha| \ll 1$, $|\delta\beta/\beta| \ll 1$ does not change σ_1, and the relative change in σ_2 is in the bounded by $|\delta\alpha/\alpha| + |\delta\beta/\beta|$ to first order. Hence, if the parameters α, β are known to a certain relative accuracy, it makes sense to aim to approximate σ_2 with correspondingly high relative accuracy.

Now let α be such that $\boldsymbol{fl}(1+\alpha) = \max\{1, \alpha\}$, e.g., $|\alpha| \notin [\boldsymbol{\varepsilon}, 1/\boldsymbol{\varepsilon}]$. If we set $\tau = \boldsymbol{fl}(1/\sqrt{2})$ and $\eta = \boldsymbol{fl}(\alpha/\sqrt{2})$ then

$$\tilde{A}^{(1)} \equiv \boldsymbol{fl}(A^{(1)}) = \begin{bmatrix} \tau & \tau \\ \tau & \tau \end{bmatrix} \quad \text{or} \quad \tilde{A}^{(1)} \equiv \boldsymbol{fl}(A^{(1)}) = \begin{bmatrix} \eta & -\eta \\ -\eta & \eta \end{bmatrix}.$$

Note that, in either case, the matrix $\tilde{A}^{(1)}$ is exactly singular, while the exact matrix $A^{(1)}$ is nonsingular. Hence, the information about the finite singular value is lost as soon as $\tilde{A}^{(1)}$ is computed and stored. It is too late even for exact computation because the reduced regular pair reads ([0], $[\sqrt{2}\beta]$) and the *computed* set of generalized singular values of (A, B) is $\{0, +\infty\}$.

Now consider the backward error. If $\boldsymbol{fl}(\sqrt{2}\beta) = \sqrt{2}\beta(1+\epsilon_1)$, $|\epsilon_1| \leq \boldsymbol{\varepsilon}$, we can write

$$[0, \boldsymbol{fl}(\sqrt{2}\beta)] = [(1+\epsilon_2)\beta, (1+\epsilon_3)\beta]\tilde{G}, \quad \tilde{G} = \begin{bmatrix} \tilde{c} & \tilde{s} \\ -\tilde{s} & \tilde{c} \end{bmatrix}, \quad \tilde{c}^2 + \tilde{s}^2 = 1,$$

where $|\epsilon_2|, |\epsilon_3| \approx O(\boldsymbol{\varepsilon})$ and $|G - \tilde{G}| \leq O(\boldsymbol{\varepsilon})|G|$. This calculation shows that the backward error δB is elementwise small: $|\delta B| \leq O(\boldsymbol{\varepsilon})|B|$. Consider, for example, the case $|\alpha| \leq \boldsymbol{\varepsilon}$. The backward error δA is obtained as the solution of the matrix equation

$$\begin{bmatrix} \tau & \tau \\ \tau & \tau \end{bmatrix} = (A + \delta A)\tilde{G}.$$

An easy calculation shows that δA_{11} and δA_{21} are of order $\boldsymbol{\varepsilon}$ and that

$$\delta A_{12} = \alpha + \tau(\tilde{c} - \tilde{s}), \quad \delta A_{22} = -\alpha + \tau(\tilde{c} - \tilde{s}).$$

Thus, the backward error δA is small in the matrix norm sense, $\|\delta A\|_F \leq O(\boldsymbol{\varepsilon})\|A\|_2$. However, the change in the second column of A simply deletes the parameter α and replaces it by roundoff noise or by an exact zero in the case $\epsilon_1 = \epsilon_2 = \epsilon_3$. Moreover, due to the singularity of $\tilde{A}^{(1)}$, the perturbation δA necessarily makes the matrix A exactly singular. (Note that the columns of A are mutually orthogonal.)

To illustrate the numerical instability described above, we run LAPACK's `SGGSVD()` procedure with different choices of α and $\beta = \alpha$. In the following table σ_2 denotes the value computed by stable formula $\sigma_2 = \sqrt{2/(1+\alpha^2)}$, $\tilde{\sigma}_2$ denotes the approximation of σ_2 computed by `SGGSVD()`, $\boldsymbol{\varepsilon}$ is the roundoff unit as computed by `SLAMCH('Epsilon')` (cf. [1]) and $\tilde{\boldsymbol{\varepsilon}} = \boldsymbol{\varepsilon}(1+\sqrt{\boldsymbol{\varepsilon}})$.

(A, B) from (3.9): SGGSVD versus exact formula for σ_2		
α ($\beta = \alpha$)	$\tilde{\sigma}_2$	$\sigma_2 \approx \sqrt{2/(1+\alpha^2)}$
$1/\varepsilon \approx$ 0.16777216E+08	0.42146848E-07	0.84293696E-07
$1/\tilde{\varepsilon} \approx$ 0.16773121E+08	0.84314280E-07	0.84314273E-07
$1/\sqrt{\varepsilon} \approx$ 0.40960000E+04	0.34526698E-03	0.34526698E-03
0.10000000E+01	0.10000000E+01	0.10000000E+01
$\sqrt{\varepsilon} \approx$ 0.24414062E-03	0.14140410E+01	0.14142135E+01
$\varepsilon \approx$ 0.59604645E-07	0.14142135E+01	0.14142135E+01
$\tilde{\varepsilon} \approx$ 0.59619197E-07	0.70693421E+00	0.14142135E+01
$\varepsilon/10 \approx$ 0.59604646E-08	0.00000000E+00	0.14142135E+01
$\varepsilon/100 \approx$ 0.59604643E-09	0.70710678E+02	0.14142135E+01
$\varepsilon =$ SLAMCH('Epsilon'), $\tilde{\varepsilon} \approx \varepsilon(1+\sqrt{\varepsilon})$		

The fact that stage *(ii)* of Algorithm 3.1 runs on a 1×1 regular pair for this example means that the large relative error in $\tilde{\sigma}_2$ is committed in stage *(i)*.

As a second test, we compute the matrix $GA = \text{diag}\,(\sqrt{2}, \sqrt{2}\alpha)$ and run Algorithm 3.1 with the pair (GA, B). This leads to an improvement of the accuracy but only in the cases of small α:

$\left(\begin{bmatrix}\sqrt{2} & 0\\ 0 & \sqrt{2}\alpha\end{bmatrix}, [\alpha, \alpha]\right)$: SGGSVD versus exact formula for σ_2		
α	$\tilde{\sigma}_2$	$\sigma_2 \approx \sqrt{2/(1+\alpha^2)}$
$1/\varepsilon \approx$ 0.16777216E+08	0.12644054E-06	0.84293696E-07
$1/\tilde{\varepsilon} \approx$ 0.16773121E+08	0.12647142E-06	0.84314273E-07
$1/\sqrt{\varepsilon} \approx$ 0.40960000E+04	0.34522486E-03	0.34526698E-03
0.10000000E+01	0.10000000E+01	0.10000000E+01
$\sqrt{\varepsilon} \approx$ 0.24414062E-03	0.14142135E+01	0.14142135E+01
$\varepsilon \approx$ 0.59604645E-07	0.14142135E+01	0.14142135E+01
$\tilde{\varepsilon} \approx$ 0.59619197E-07	0.14142135E+01	0.14142135E+01
$\varepsilon/10 \approx$ 0.59604646E-08	0.14142134E+01	0.14142135E+01
$\varepsilon/100 \approx$ 0.59604643E-09	0.14142135E+01	0.14142135E+01
$\varepsilon =$ SLAMCH('Epsilon'), $\tilde{\varepsilon} \approx \varepsilon(1+\sqrt{\varepsilon})$		

The relative accuracy is lost in the postmultiplication of A by the rotation G. In floating–point arithmetic, this operation transforms a matrix with mutually orthogonal columns (A) into a nearly or even exactly singular matrix.

Another source of errors in Algorithm 3.1 can be illustrated by running it on the pair (B, A). If $|\alpha| \leq \varepsilon$ or $|\alpha| \geq 1/\varepsilon$, the matrix A is declared rank deficient and the well–defined finite singular value is lost.

4 Reduction algorithm based on LU factorization

In this section, we present a new GSVD reduction algorithm. The main innovation of our approach is that we replace the URV decomposition by a combination of LU factorization and certain nonorthogonal triangular transformations. We also carefully scale the initial matrix pair to prevent uncontrolled condition growth during the reduction process.

4.1 QRT and LUT factorizations

The key feature of the new algorithm is a new simple and elegant factorization of a general matrix. It is based on pivoted QR or LU factorization and, in the case of a full column rank matrix, it reduces to QR or LU, respectively. In the general case, it provides a simple way to cancel out columns that are identified in pivoted QR or LU as linearly dependent on the remaining ones. More precisely, we have:

THEOREM **4.1** *Let $B \in \mathbf{R}^{p\times n}$ and $r_B = \operatorname{rank} B$. Then there exist an orthogonal $p \times p$ matrix Q, an $n \times n$ permutation matrix Π, an $r_B \times (n - r_B)$ matrix X, and an $r_B \times r_B$ upper triangular nonsingular matrix R such that*

$$B\Pi = Q\begin{bmatrix} R & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}\begin{bmatrix} I & X \\ \mathbf{O} & I \end{bmatrix}. \tag{4.10}$$

Furthermore, there exist permutation matrices P_1, P_2, a unit lower trapezoidal $p \times r_B$ matrix L, a unit upper triangular $r_B \times r_B$ matrix U, a diagonal nonsingular $r_B \times r_B$ matrix Δ, and an $r_B \times (n - r_B)$ matrix Y such that

$$P_1 B P_2 = L\Delta[U, \mathbf{O}]\begin{bmatrix} I & Y \\ \mathbf{O} & I \end{bmatrix}. \tag{4.11}$$

The factorizations (4.10) and (4.11) define the QRT and the LU factorizations of B, respectively.

Proof Let

$$B\Pi = Q\begin{bmatrix} R & \hat{R} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad Q^T Q = QQ^T = I,$$

be any rank revealing QR factorization of B, where Π is a permutation matrix and R is an $r_B \times r_B$ upper triangular nonsingular matrix. Define $X = R^{-1}\hat{R}$. Then

$$\begin{bmatrix} R & \hat{R} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}\begin{bmatrix} I & -X \\ \mathbf{O} & I \end{bmatrix} = \begin{bmatrix} R & \hat{R} - RX \\ \mathbf{O} & \mathbf{O} \end{bmatrix} = \begin{bmatrix} R & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

Similarly, let $P_1 B P_2 = L\Delta[U, \hat{U}]$ be any rank revealing LU factorization. Define $Y = U^{-1}\hat{U}$ and note that

$$\begin{bmatrix} U & \hat{U} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}\begin{bmatrix} I & -Y \\ \mathbf{O} & I \end{bmatrix} = \begin{bmatrix} U & \hat{U} - UY \\ \mathbf{O} & \mathbf{O} \end{bmatrix} = \begin{bmatrix} U & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}.$$

Q.E.D.

In Theorem 4.1 we use elementary triangular transformations that have the following useful properties.

PROPOSITION **4.1** *Let*

$$T \equiv T(X) = \begin{bmatrix} I & X \\ \mathbf{O} & I \end{bmatrix}. \tag{4.12}$$

Then $T(X)^{-1} = T(-X)$ and $\max\{\|T(X)\|_2, \|T(X)^{-1}\|_2\} \le 1 + \|X\|_2$. Furthermore, if R, U, X, Y are as in Theorem 4.1, and if D_R and D_U are any diagonal matrices that satisfy $|D_R| \ge |\mathbf{diag}(R)|$, $|D_U| \ge |\mathbf{diag}(U)|$, where $\mathbf{diag}(R) \stackrel{def}{=} \operatorname{diag}(R_{11}, \ldots, R_{r_B, r_B})$, then

$$\|X\|_2 \le \sqrt{r_B(n - r_B)}\,\|R^{-1}D_R\|_2, \quad \|Y\|_2 \le \sqrt{r_B(n - r_B)}\,\|U^{-1}D_U\|_2.$$

4.2 The algorithm

The structure of our algorithm is similar to the one of Algorithm 3.1. The main differences are: *(i)* we use an initial prescaling that is crucial in preserving the numerical stability in subsequent steps; *(ii)* we replace the URV factorization with the LUT factorization; *(iii)* we use the algorithm from [15] to compute the SVD of a single matrix instead of using simultaneous transformations of a pair of matrices. The new algorithm reads as follows.

ALGORITHM **4.1** *(LU-based GSVD computation)*

Input $(A, B) \in \mathbf{R}^{m\times n} \times \mathbf{R}^{p\times n}$, $\operatorname{rank} A = r_A$, $\operatorname{rank} B = r_B$.

Stage A Reduction.

Step 0 Scaling. Define $\Delta_A = \mathrm{diag}\,(\|Ae_i\|_2)$. If some column of A is zero, then replace the corresponding diagonal entry in the definition of Δ_A by any nonzero scalar. Compute $A^{(0)} = A\Delta_A^{-1}$, $B^{(0)} = B\Delta_A^{-1}$. The new pair $(A^{(0)}, B^{(0)})$ is equivalent to (A, B).

Step 1 Compute the LU factorization with total pivoting of of $B^{(0)}$:

$$\Pi_1 B^{(0)} \Pi_2 = \begin{bmatrix} L \\ \hat{L} \end{bmatrix} [U^{(1,1)}, U^{(1,2)}], \quad L, U^{(1,1)} \in \mathbf{R}^{r_B \times r_B},$$

where L is unit lower triangular and $U^{(1,1)}$ is upper triangular and nonsingular. Partition $A^{(0)}\Pi_2$ accordingly,

$$A^{(0)}\Pi_2 = [A_{11}^{(0)}, A_{12}^{(0)}].$$

Step 2 Compute $X = A_{11}^{(0)}(U^{(1,1)})^{-1}$ and

$$A^{(1)} \equiv [A_{11}^{(1)}, A_{12}^{(1)}] = [X, A_{12}^{(0)} - XU^{(1,2)}].$$

If the right generalized singular vectors are needed, compute $T(Y)$ with $Y = -(U^{(1,1)})^{-1}U^{(1,2)}$. Set $U^{(1,2)}$ to zero and $U^{(1,1)}$ to I_{r_B}.

Step 3 If $A_{12}^{(1)}$ is not empty ($r_B < n$), compute a rank revealing QR factorization of $A_{12}^{(1)}$,

$$A_{12}^{(1)}\Pi_3 = Q_{12}^{(1)} \begin{bmatrix} A_{12}^{(2)} & \hat{A}_{12}^{(2)} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \quad A_{12}^{(2)} \in \mathbf{R}^{r_A^{(1)} \times r_A^{(1)}}, \quad r_A^{(1)} = \mathrm{rank}\, A_{12}^{(1)},$$

where $A_{12}^{(2)}$ is upper triangular and nonsingular. Set $\hat{A}_{12}^{(2)}$ to zero. Update $A_{11}^{(1)}$ by $A_{11}^{(1)} \mapsto A_{11}^{(2)} = (Q_{12}^{(2)})^T A_{11}^{(1)}$. With an appropriate row partition of $A_{11}^{(2)}$ and with conformal column repartition the new pair reads

$$(A^{(2)}, B^{(2)}) = \left(\begin{bmatrix} A_{11,1}^{(2)} & A_{12}^{(2)} & \mathbf{O} \\ A_{11,2}^{(2)} & \mathbf{O} & \mathbf{O} \end{bmatrix}, \begin{bmatrix} L \\ \hat{L} \end{bmatrix} [I_{r_B}, \mathbf{O}, \mathbf{O}] \right).$$

Return At the end of Step 3 the reduced regular pair reads

$$\left(A_{11,2}^{(2)}, \begin{bmatrix} L \\ \hat{L} \end{bmatrix}\right). \tag{4.13}$$

Stage B GSVD of the regular pair (4.13). Use the algorithm from [15].

5 Error and perturbation analysis of the algorithm

In this section we a give detailed analysis of the numerical properties of Algorithm 4.1. Since we are interested in cases where all finite generalized singular values can be computed with relative accuracy, we restrict our analysis in this section to the following full rank case: $r_A = n$, $r_B = \min\{p, n\}$. For simplicity, we consider only the case $r_B = p$. Since the scaling in Step 0 is elementwise backward and forward stable for any diagonal scaling (it introduces relative uncertainty at most of the order of the round–off in each nonzero entry), we assume that $(A^{(0)}, B^{(0)}) = (A, B)$. For ease of notation, we also assume that the matrix B is previously pre– and postmultiplied by suitable permutation matrices so that no row or column interchanges are necessary for a rank revealing LU factorization.

We use the following partition of $B = LU$:

$$B = [B^{(1,1)}, B^{(1,2)}] = L[U^{(1,1)}, U^{(1,2)}], \quad B^{(1,1)}, L, U^{(1,1)} \in \mathbf{R}^{p\times p}.$$

In the same way we write

$$B + \delta B = [B^{(1,1)} + \delta B^{(1,1)}, B^{(1,2)} + \delta B^{(1,2)}] = \tilde{L}\tilde{U} = \tilde{L}[\tilde{U}^{(1,1)}, \tilde{U}^{(1,2)}].$$

A structure of the analysis is given in the form of a commutative diagram on Figure 2 where vertical arrows ($\uparrow$, $\downarrow$) denote exact computations, horizontal arrows ($\leftarrow$, $\rightarrow$) denote backward perturbations, $\Longrightarrow$ denotes a forward perturbation, and diagonal arrows ($\swarrow$, $\searrow$, $\nearrow$) denote floating–point computations.

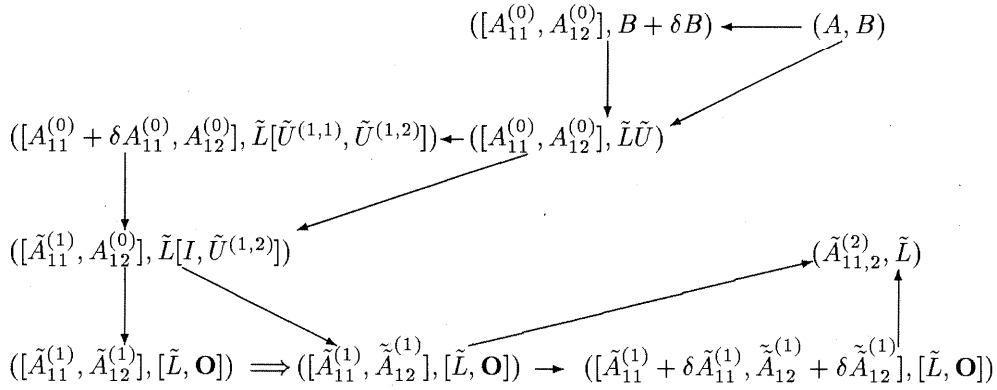


Figure 2: Commutative diagram of the floating–point algorithm.

In Figure 2, $\tilde{L}$, $\tilde{U}$ are computed triangular factors of B, and $\tilde{L}\tilde{U} = B + \delta B$, for some backward error δB. Hence, the pair $(A^{(0)}, \tilde{L}\tilde{U})$ is obtained from $(A^{(0)}, B + \delta B)$ by an exact LU factorization of $B + \delta B$. We estimate the relative difference between the generalized singular values of (A, B) and $(A, \tilde{L}\tilde{U})$ in § 5.1. We analyze Step 2 using both backward and forward error analysis. First, the approximation $\tilde{A}_{11}^{(1)}$ of $A_{11}^{(1)}$ is represented as the exact product $\tilde{A}_{11}^{(1)} = (A_{11}^{(0)} + \delta A_{11}^{(0)})(\tilde{U}^{(1,1)})^{-1}$. Next, starting with $[\tilde{A}_{11}^{(1)}, A_{12}^{(0)}]$ we compute the matrix $\tilde{\tilde{A}}_{12}^{(1)}$ and estimate the difference $\delta\tilde{\tilde{A}}_{12}^{(1)}$ (forward error) between $\tilde{\tilde{A}}_{12}^{(1)}$ and the exact matrix $\tilde{\tilde{A}}_{12}^{(1)} = A_{12}^{(0)} - \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)}$. The generalized singular value perturbations caused by $\delta A_{11}^{(0)}$ and $\delta\tilde{A}_{12}^{(1)}$ are analyzed separately in § 5.2. The error analysis of Step 3 in § 5.3 is essentially the backward error analysis of the QR factorization. We conclude the analysis in § 5.4 where we give estimates of condition numbers of matrices defined in Algorithm 4.1. An important conclusion in § 5.4 is that the relevant condition numbers of all matrices generated by Algorithm 4.1 remain controlled by the condition numbers of initial matrices A and B.

5.1 Error analysis of Step 1

The accuracy of Step 1 of Algorithm 4.1 depends on the accuracy of the computed triangular factors of B. Therefore, we start the analysis with perturbation estimates for the floating–point LU factorization. Our estimates are based on the backward error estimate from [18] and on the perturbation analysis of the LU factorization from [33].

The LU factorization in floating–point arithmetic has a very strong form of backward stability that allows an elementwise bound on the backward error. More precisely:

THEOREM **5.1** *[18, Theorem 3.3.1] Let $B = \boldsymbol{fl}(B)$ be an $p \times n$ matrix, and let its LU factorization be computed by the outer product version of the Gaussian elimination [18, Algorithm 3.2.3]. If no*

zero pivots are encountered during the elimination process, the computed factors $\tilde{L}$ and $\tilde{U}$ satisfy

$$\tilde{L}\tilde{U} = B + \delta B, \quad |\delta B| \leq \varepsilon_{LU}(|B| + |\tilde{L}| \cdot |\tilde{U}|) + O(\varepsilon^2), \tag{5.14}$$

where $0 \leq \varepsilon_{LU} = \varepsilon_{LU}(p,n) \leq 3(\min\{p,n\} - 1)\varepsilon$.

In the next theorem we analyze how δB from (5.14) changes the exact factors of $B = LU$, i.e., we bound the errors $\delta L = \tilde{L} - L$ and $\delta U = \tilde{U} - U$. We use the following elementwise operators:

$$(\mathbf{tril}(A))_{ij} = \begin{cases} A_{ij}, & i > j, \\ 0, & i \leq j, \end{cases} \qquad (\overline{\mathbf{tril}}(A))_{ij} = \begin{cases} A_{ij}, & i \geq j, \\ 0, & i < j, \end{cases}$$

and $\mathbf{triu}(A) = (\mathbf{tril}(A^T))^T$, $\overline{\mathbf{triu}}(A) = (\overline{\mathbf{tril}}(A^T))^T$.

THEOREM **5.2** *Let the matrix B in Theorem 5.1 have full row rank, and let $B = LU$ be its LU factorization, where L is a unit lower triangular matrix and U is a full row rank upper trapezoidal matrix. If the spectral radius of*

$$E_B = |\tilde{L}^{-1}\delta B^{(1,1)}(\tilde{U}^{(1,1)})^{-1}| \tag{5.15}$$

is less than one, there exist a strictly lower triangular E_L and an upper triangular E_U such that

$$\tilde{L} = (I + E_L)L, \quad \tilde{U} = U(I + E_U), \tag{5.16}$$

and

$$|E_L| \leq \varepsilon_{LU}|\tilde{L}|\mathbf{tril}(|\tilde{L}^{-1}|(|LU^{(1,1)}| + |\tilde{L}| \cdot |\tilde{U}^{(1,1)}|)|(\tilde{U}^{(1,1)})^{-1}|)|L^{-1}| + O(\varepsilon^2), \tag{5.17}$$

$$E_U = \begin{bmatrix} E_U^{(1,1)} & E_U^{(1,2)} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}, \tag{5.18}$$

$$\begin{aligned} |E_U^{(1,1)}| &\leq \varepsilon_{LU}|(U^{(1,1)})^{-1}|\overline{\mathbf{triu}}(|\tilde{L}^{-1}(|LU^{(1,1)}| \\ &+ |\tilde{L}| \cdot |\tilde{U}^{(1,1)}|)|(\tilde{U}^{(1,1)})^{-1}|)|\tilde{U}^{(1,1)}| + O(\varepsilon^2), \end{aligned} \tag{5.19}$$

$$\begin{aligned} |E_U^{(1,2)}| &\leq \varepsilon_{LU}|(U^{(1,1)})^{-1}|\{|\tilde{L}^{-1}|(|LU^{(1,2)}| + |\tilde{L}| \cdot |\tilde{U}^{(1,2)}|) \\ &+ \mathbf{tril}\,(|\tilde{L}^{-1}|(|LU^{(1,1)}| + |\tilde{L}| \cdot |\tilde{U}^{(1,1)}|)|(U^{(1,1)})^{-1}|)|U^{(1,2)}|\} + O(\varepsilon^2). \end{aligned} \tag{5.20}$$

Hence, in Theorem 5.1, the backward perturbed matrix can be represented as $B + \delta B = (I + E_L)B(I + E_U)$.

Proof First, note that

$$B^{(1,1)} + \delta B^{(1,1)} = \tilde{L}\tilde{U}^{(1,1)}, \quad \delta U^{(1,2)} = \tilde{L}^{-1}(\delta B^{(1,2)} - \delta L\, U^{(1,2)}), \tag{5.21}$$

and that

$$|\delta B^{(i,i)}| \leq \varepsilon_{LU}(|B^{(1,i)}| + |\tilde{L}| \cdot |\tilde{U}^{(1,i)}|), \quad i = 1,2. \tag{5.22}$$

Now an application of [33, Theorem 5.1] to the first equation in (5.21) yields an estimate of $E_L = \delta L\, L^{-1}$. Namely,

$$|E_L| \leq |\tilde{L}|\mathbf{tril}((I - E_B)^{-1}E_B)|L^{-1}|. \tag{5.23}$$

Writing the Neumann expansion of $(I - E_B)^{-1}$ and using (5.15), (5.22) yields (5.17). Similarly, [33, Theorem 5.1] implies that

$$|(U^{(1,1)})^{-1}\delta U^{(1,1)}| \leq |(U^{(1,1)})^{-1}|\overline{\mathbf{triu}}(E_B(I - E_B)^{-1})|\tilde{U}^{(1,1)}|, \tag{5.24}$$

and (5.19) follows. To estimate E_U first note that

$$[\tilde{U}^{(1,1)}, \tilde{U}^{(1,2)}] = [U^{(1,1)}, U^{(1,2)}] \begin{bmatrix} I + (U^{(1,1)})^{-1}\delta U^{(1,1)} & (U^{(1,1)})^{-1}\delta U^{(1,2)} \\ \mathbf{O} & I \end{bmatrix}.$$

Hence, it remains to estimate $|(U^{(1,1)})^{-1}\delta U^{(1,2)}|$. From the second equation in (5.21) it follows that

$$\begin{aligned} |(U^{(1,1)})^{-1}\delta U^{(1,2)}| &\leq |(U^{(1,1)})^{-1}\tilde{L}^{-1}\delta B^{(1,2)}| + |(U^{(1,1)})^{-1}\tilde{L}^{-1}\delta L\, U^{(1,2)}| \\ &\leq \varepsilon_{LU}|(U^{(1,1)})^{-1}| \cdot |\tilde{L}^{-1}|(|LU^{(1,2)}| + |\tilde{L}| \cdot |\tilde{U}^{(1,2)}|) \\ &+ |(U^{(1,1)})^{-1}| \cdot |\mathbf{tril}(\tilde{L}^{-1}\delta B^{(1,1)}(U^{(1,1)})^{-1})| \cdot |U^{(1,2)}|. \end{aligned}$$

Now, using (5.22) implies (5.20). Q.E.D.

In Algorithm 4.1, we replace the original pair (A, B) by $(A, B+\delta B)$, that is, by $(A, \tilde{L}\tilde{U})$, where δB, $\tilde{L}$, and $\tilde{U}$ are as in Theorem 5.2. In the following proposition, we estimate the generalized singular value perturbations caused by δB.

PROPOSITION **5.1** *Let the assumptions of Theorem 5.2 hold, and let $\sigma_1 \geq \cdots \geq \sigma_p$ and $\sigma_1' \geq \cdots \geq \sigma_p'$ be the finite generalized singular values of (A, B) and $(A, B+\delta B)$, respectively. If A is a full column rank matrix with the QR factorization $A = QR$, and if $R^{(1,1)}$ is the $p \times p$ upper left main submatrix of R, then*

$$\max_{1\leq i\leq p} \frac{|\sigma_i' - \sigma_i|}{\sqrt{\sigma_i'\sigma_i}} \leq \frac{1}{2} \frac{\|E_L\|_2 + \|R^{(1,1)}[E_U^{(1,1)}, E_U^{(1,2)}]R^{-1}\|_2}{1 - (1/32)\|E_L\|_2\|R^{(1,1)}[E_U^{(1,1)}, E_U^{(1,2)}]R^{-1}\|_2}, \tag{5.25}$$

provided that the right–hand side in (5.25) is strictly positive.

Proof Let $A = QR$ be the QR factorization of A with a nonsingular upper triangular matrix R. Without loss of generality we may consider matrix pairs (B, R) and $((I + E_L)B(I + E_U), R)$ or, equivalently, the matrices BR^{-1} and $(I + E_L)BR^{-1}(I + RE_UR^{-1})$. By a result of Ren–Cang Li [23, Theorem 5.2], the distance between the exact and the perturbed generalized singular values is bounded by

$$\max_{1\leq i\leq p} \frac{|\sigma_i' - \sigma_i|}{\sqrt{\sigma_i'\sigma_i}} \leq \frac{1}{2} \frac{\|E_L\|_2 + \|RE_UR^{-1}\|_2}{1 - (1/32)\|E_L\|_2\|RE_UR^{-1}\|_2}. \tag{5.26}$$

Now the special structure of E_U implies the desired bound. Q.E.D.

REMARK **5.1** In relation (5.26), we can easily estimate $\|RE_UR^{-1}\|_2$ by

$$\|RE_UR^{-1}\|_2 \leq \kappa_2(R)\|E_U\|_2 = \kappa_2(A)\|E_U\|_2.$$

It is important to note that the matrices A and R both have columns of unit Euclidean norm because of the scaling in Step 0 of Algorithm 4.1.

5.2 Error analysis of Step 2

In Step 2 of Algorithm 4.1, the transformation of the matrix $\tilde{U}$ is error–free and effortless. It remains to analyze the computation of the matrix $\tilde{A}^{(1)}$. We first consider the computation of $\tilde{A}_{11}^{(1)}$.

PROPOSITION **5.2** *Let $\tilde{A}_{11}^{(1)}$ be a computed approximation of the solution of the matrix equation $X\tilde{U}^{(1,1)} = A_{11}^{(0)}$. Then there exists an $\delta A_{11}^{(0)}$ such that $\tilde{A}_{11}^{(1)} = (A_{11}^{(0)} + \delta A_{11}^{(0)})(\tilde{U}^{(1,1)})^{-1}$, and*

$$|\delta A_{11}^{(0)}| \leq r_B\varepsilon|A_{11}^{(0)}| \cdot |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,1)}| + O(\varepsilon^2). \tag{5.27}$$

Hence, for $1 \leq i \leq r_B$,

$$\|\delta A_{11}^{(0)} e_i\|_2 \leq r_B\varepsilon\||(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,1)}|e_i\|_1 + O(\varepsilon^2). \tag{5.28}$$

Proof Using an analysis of Wilkinson [39] we know that there exist matrices $\delta\tilde{U}_k^{(1,1)}$, $1 \le k \le m$, such that

$$e_k^T \tilde{A}_{11}^{(1)}(\tilde{U}^{(1,1)} + \delta\tilde{U}_k^{(1,1)}) = e_k^T A_{11}^{(0)}, \quad |\delta\tilde{U}_k^{(1,1)}|_{ij} \le (j-i+1)\varepsilon|\tilde{U}^{(1,1)}|_{ij}, \quad 1 \le i \le j \le r_B.$$

Hence the residual matrix $\delta A_{11}^{(0)}$ defined by

$$\tilde{A}_{11}^{(1)}\tilde{U}^{(1,1)} - A_{11}^{(0)} = \delta A_{11}^{(0)} \tag{5.29}$$

satisfies (5.27). Now we rewrite (5.29) as $\tilde{A}_{11}^{(1)} = (A_{11}^{(0)} + \delta A_{11}^{(0)})(\tilde{U}^{(1,1)})^{-1}$. Finally, relation (5.28) follows from (5.27) because $A_{11}^{(0)}$ has unit columns. Q.E.D.

REMARK **5.2** Note that the bounds (5.27), (5.28) are invariant under row scaling of $\tilde{U}^{(1,1)}$. Furthermore, if $|\tilde{U}^{(1,1)}|_{ii} \ge |\tilde{U}^{(1,1)}|_{ij}$, $j > i$, then for all $j \ge i$ it holds that $(|(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,1)}|)_{ij} \le 2^{j-i}$; see [20, Lemma 3.1]. Note also that the factor r_B can be removed by using double precision accumulation.

In the next theorem we estimate the uncertainty in the generalized singular values of the new pair $([\tilde{A}_{11}^{(1)}, A_{12}^{(0)}], \tilde{L}[I, \tilde{U}^{(1,2)}])$.

THEOREM **5.3** *Let $\sigma_1' \ge \cdots \ge \sigma_n'$ and $\sigma_1'' \ge \cdots \ge \sigma_n''$ be the generalized singular values of the pairs $([A_{11}^{(0)}, A_{12}^{(0)}], \tilde{L}[\tilde{U}^{(1,1)}, \tilde{U}^{(1,2)}])$ and $([\tilde{A}_{11}^{(1)}, A_{12}^{(0)}], \tilde{L}[I, \tilde{U}^{(1,2)}])$, respectively. Furthermore, let $[A_{12}^{(0)}, A_{11}^{(0)}] = [W^{[1]}, W^{[2]}]T$ be the QR factorization of $[A_{12}^{(0)}, A_{11}^{(0)}]$ with upper triangular T, and let $T^{[2,2]} = (W^{[2]})^T A_{11}^{(0)}$. Then for all i, either $\sigma_i'' = \sigma_i' = \infty$ or*

$$1 - \|\delta A_{11}^{(0)}(T^{[2,2]})^{-1}\|_2 \le \frac{\sigma_i''}{\sigma_i'} \le 1 + \|\delta A_{11}^{(0)}(T^{[2,2]})^{-1}\|_2 \tag{5.30}$$

Furthermore, if $A_{11}^{(0)} = VT_{11}^{(0)}$ is the QR factorization of $A_{11}^{(0)}$ and if ψ is the maximal acute principal angle between $\mathcal{R}(A_{11}^{(0)})$ and $\mathcal{R}(A_{12}^{(0)})^{\perp} \bigcap \mathcal{R}(A^{(0)})$ then

$$\|\delta A_{11}^{(0)}(T^{[2,2]})^{-1}\|_2 \le r_B\varepsilon \frac{\| \, |A_{11}^{(0)}| \cdot |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,1)}| \cdot |(T_{11}^{(0)})^{-1}| \, \|_2}{\cos\psi}, \tag{5.31}$$

and similarly

$$\|\delta A_{11}^{(0)}(T^{[2,2]})^{-1}\|_2 \le r_B\varepsilon \frac{\| \, |A_{11}^{(0)}| \, \|_2 \|(A_{11}^{(0)})^{\dagger}\|_2}{\cos\psi} \| \, |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,1)}| \, \|_2. \tag{5.32}$$

Proof We can equivalently consider the pair

$$(\tilde{L}[\tilde{U}^{(1,2)}, I], ([W^{[1]}, W^{[2]}] + [\mathbf{O}, \delta A_{11}^{(0)}]T^{-1})T) \tag{5.33}$$

and compare its singular values to those of $(\tilde{L}[\tilde{U}^{(1,1)}, I], T)$. The variational characterization of the generalized singular values immediately implies (5.30). Now note that $[\mathbf{O}, \delta A_{11}^{(0)}]T^{-1} = [\mathbf{O}, \delta A_{11}^{(0)}(T^{[2,2]})^{-1}]$ and that $T^{[2,2]} = ((W^{[2]})^T V)T_{11}^{(0)}$. Finally, note that $\sigma_{\min}((W^{[2]})^T V) = \cos\psi > 0$ so that relation (5.31) and (5.32) follow from relation (5.27) using matrix norm inequalities. Q.E.D.

Next we analyze the computation of $\tilde{A}_{12}^{(1)}$:

$$([\tilde{A}_{11}^{(1)}, A_{12}^{(0)}], \tilde{L}[I, \tilde{U}^{(1,2)}]) \longmapsto ([\tilde{A}_{11}^{(1)}, \boldsymbol{fl}(A_{12}^{(0)} - \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)})], \tilde{L}[I, \mathbf{O}]).$$

Since this operation involves only matrix sum and multiply operations, the error analysis is simple, at least if we use the standard matrix multiply algorithm.

PROPOSITION **5.3** *Let* $\tilde{A}_{12}^{(1)} = A_{12}^{(0)} - \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)}$, $\tilde{\tilde{A}}_{12}^{(1)} \equiv \boldsymbol{fl}(\tilde{A}_{12}^{(1)}) = \tilde{A}_{12}^{(1)} + \delta\tilde{A}_{12}^{(1)}$. *Then*

$$\begin{aligned} |\delta\tilde{A}_{12}^{(1)}| &\leq \varepsilon|\tilde{A}_{12}^{(1)}| + \varepsilon_P|\tilde{A}_{11}^{(1)}| \cdot |\tilde{U}^{(1,2)}| + O(\varepsilon^2) \\ &\leq \varepsilon|\tilde{A}_{12}^{(1)}| + \varepsilon_P|\tilde{A}_{11}^{(0)}| \cdot |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,2)}| + O(\varepsilon^2), \end{aligned}$$

where $\varepsilon_P \approx O(r_B\varepsilon)$.

Proof Straightforward. Indeed,

$$\begin{aligned} \boldsymbol{fl}(\tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)}) &= \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)} + E_1, \quad |E_1| \leq \varepsilon_P|\tilde{A}_{11}^{(1)}| \cdot |\tilde{U}^{(1,2)}|, \\ \boldsymbol{fl}(A_{12}^{(0)} - \boldsymbol{fl}(\tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)})) &= A_{12}^{(0)} - \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)} - E_1 + E_2, \end{aligned}$$

where $|E_2| \leq \varepsilon|A_{12}^{(0)} - \tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)} - E_1| \leq \varepsilon|\tilde{A}_{12}^{(1)}| + O(\varepsilon^2)$. Q.E.D.

Now we can estimate generalized singular value perturbations due to the error $\delta\tilde{A}_{12}^{(1)}$.

THEOREM **5.4** *Let* $\sigma_1'' \geq \cdots \geq \sigma_n''$ *and* $\sigma_1''' \geq \cdots \geq \sigma_n'''$ *be the generalized singular values of* $([\tilde{A}_{11}^{(1)}, A_{12}^{(0)}], \tilde{L}[I, \tilde{U}^{(1,2)}])$ *and* $([\tilde{A}_{11}^{(1)}, \tilde{\tilde{A}}_{12}^{(1)}], [\tilde{L}, \mathbf{O}])$, *respectively. Furthermore, let* $[\tilde{A}_{11}^{(1)}, \tilde{A}_{12}^{(1)}] = [\tilde{W}^{(1)}, \tilde{W}^{(2)}]\tilde{T}$ *be the QR factorization of* $[\tilde{A}_{11}^{(1)}, \tilde{A}_{12}^{(1)}]$, *where* $\tilde{T}$ *is upper triangular and nonsingular. Let* $\tilde{T}^{(2,2)} = (\tilde{W}^{(2)})^T\tilde{A}_{12}^{(1)}$. *Then, for all* i, *either* $\sigma_i''' = \sigma_i'' = \infty$ *or*

$$1 - \|\delta\tilde{A}_{12}^{(1)}(\tilde{T}^{(2,2)})^{-1}\|_2 \leq \frac{\sigma_i'''}{\sigma_i''} \leq 1 + \|\delta\tilde{A}_{12}^{(1)}(\tilde{T}^{(2,2)})^{-1}\|_2. \tag{5.34}$$

Furthermore, let $\tilde{A}_{12}^{(1)} = \tilde{V}T_{12}^{(1)}$ *be the QR factorization of* $\tilde{A}_{12}^{(1)}$, *let* ψ_i *be the angle between* $\mathcal{R}(A_{12}^{(0)}e_i)$ *and* $\mathcal{R}(\tilde{A}_{11}^{(1)}\tilde{U}^{(1,2)}e_i)$, *and let* ϕ *be the maximal acute principal angle between* $\mathcal{R}(\tilde{A}_{12}^{(1)})$ *and* $\mathcal{R}(\tilde{A}_{11}^{(1)})^{\perp} \bigcap \mathcal{R}(\tilde{A}^{(1)})$. *Then in (5.34)*

$$\begin{aligned} \|\delta\tilde{A}_{12}^{(1)}(\tilde{T}^{(2,2)})^{-1}\|_2 &\leq \varepsilon\frac{\| \, |\tilde{A}_{12}^{(1)}| \cdot |(T_{12}^{(1)})^{-1}| \, \|_2}{\cos\phi} \\ &+ \varepsilon_P\frac{\| \, |A_{11}^{(0)}| \cdot |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,2)}| \, \|_2}{\min_i \sin\psi_i}\frac{\|(\tilde{A}_{12}^{(1)})_c^{\dagger}\|_2}{\cos\phi} + O(\varepsilon^2). \end{aligned} \tag{5.35}$$

Proof Similarly to the proof of Theorem 5.3, we create a new pair

$$([\tilde{L}, \mathbf{O}], ([\tilde{W}^{(1)}, \tilde{W}^{(2)}] + [\mathbf{O}, \delta\tilde{A}_{12}^{(1)}]\tilde{T}^{-1})\tilde{T})$$

and immediately obtain (5.34). Let us now prove (5.35). First, note that $\delta\tilde{A}_{12}^{(1)} = E_2 - E_1$, where E_1 and E_2 are defined in Proposition 5.3. We easily find that $(\tilde{T}^{(2,2)})^{-1} = (T_{12}^{(1)})^{-1}((\tilde{W}^{(2)})^T\tilde{V})^{-1}$ and thus

$$\|E_2(\tilde{T}^{(2,2)})^{-1}\|_2 \leq \frac{\|E_2(T_{12}^{(1)})^{-1}\|_2}{\sigma_{\min}((\tilde{W}^{(2)})^T\tilde{V})} \leq \varepsilon\frac{\| \, |\tilde{A}_{12}^{(1)}| \cdot |(T_{12}^{(1)})^{-1}| \, \|_2}{\cos\phi} + O(\varepsilon^2).$$

It remains to estimate $E_1(\tilde{T}^{(2,2)})^{-1}$. It is not hard to show, using Pythagoras' theorem that, for all i, $\|\tilde{A}_{12}^{(1)}e_i\|_2 \geq \sin\psi_i$. Hence

$$\|E_1(T_{12}^{(1)})^{-1}\|_2 \leq \| \, |(E_1)_c| \, \|_2\|(\tilde{A}_{12}^{(1)})_c^{\dagger}\|_2 \leq \varepsilon_P\frac{\| \, |\tilde{A}_{11}^{(1)}| \cdot |\tilde{U}^{(1,2)}| \, \|_2}{\min_i \sin\psi_i}\|(\tilde{A}_{12}^{(1)})_c^{\dagger}\|_2,$$

where we have also used Proposition 5.3, and column scaling of E_1 is performed using column norms of $\tilde{A}_{12}^{(1)}$. Now an application of Proposition 5.2 yields the desired estimate. Q.E.D.

REMARK **5.3** Note that, if the columns of A are fairly linearly independent ($\kappa_2(A_c)$ moderate) and if $\| \, |(\tilde{U}^{(1,1)})^{-1}| \cdot |\tilde{U}^{(1,2)}| \, \|_2$ is moderate, $\min_i \sin\psi_i$ has reasonably large lower bound.

REMARK **5.4** We cannot guarantee that even a "reasonably implemented" fast Level 3 BLAS (cf. [1, § 4.13]) can speed up Step 2 without sacrificing relative accuracy. The reason is that the elementwise error estimates in Proposition 5.2 and Proposition 5.3 do not hold if "fast algorithms" are used (cf. [21]).

5.3 Error analysis of Step 3

In the last step of the reduction phase we compute the QR factorization of $\tilde{\tilde{A}}_{12}^{(1)}$. Floating–point error analysis of the QR factorization implies that the computed upper triangular factor $\tilde{A}_{12}^{(2)}$ satisfies the relation

$$\tilde{Q}_{12}^{(2)} \tilde{A}_{12}^{(2)} = \tilde{\tilde{A}}_{12}^{(1)} + \delta\tilde{\tilde{A}}_{12}^{(1)}, \quad \|\delta\tilde{\tilde{A}}_{12}^{(1)} e_i\|_2 \leq \varepsilon_{QR} \|\tilde{\tilde{A}}_{12}^{(1)} e_i\|_2, \quad 1 \leq i \leq n - r_B,$$

where $\tilde{Q}_{12}^{(2)}$ is a certain orthogonal matrix, and $\delta\tilde{\tilde{A}}_{12}^{(1)}$ is a backward error that depends on the details of the floating–point implementation of the QR factorization. The quantity ε_{QR} is bounded by ε times a modestly growing function of matrix dimensions; see e.g. [17], [14]. The same sequence of (nearly) orthogonal transformations is applied to $\tilde{A}_{11}^{(1)}$ so that the resulting matrix $\tilde{A}_{11}^{(2)}$ satisfies

$$\tilde{Q}_{12}^{(2)} \tilde{A}_{11}^{(2)} = \tilde{A}_{11}^{(1)} + \delta\tilde{A}_{11}^{(1)}, \quad \|\delta\tilde{A}_{11}^{(1)} e_i\|_2 \leq \varepsilon_{QR} \|\tilde{A}_{11}^{(1)} e_i\|_2, \quad 1 \leq i \leq r_B.$$

Hence, with an appropriate row partition of $\tilde{A}_{11}^{(2)}$,

$$\begin{bmatrix} \tilde{A}_{11,1}^{(2)} & \tilde{A}_{12}^{(2)} \\ \tilde{A}_{11,2}^{(2)} & \mathbf{O} \end{bmatrix} = (\tilde{Q}_{12}^{(2)})^T [\tilde{A}_{11}^{(1)} + \delta\tilde{A}_{11}^{(1)}, \tilde{\tilde{A}}_{12}^{(1)} + \delta\tilde{\tilde{A}}_{12}^{(1)}].$$

Thus, the output of the reduction phase is the pair $(\tilde{A}_{11,2}^{(2)}, \tilde{L})$. We may proceed with this pair as with the exact one but must keep in mind that its singular values differ from the singular values of $([A_{11}^{(1)}, A_{12}^{(1)}], [\tilde{L}\tilde{U}^{(1,1)}, \mathbf{O}])$. The relative differences between the maching generalized singular values of the two pairs are estimated in the following theorem.

THEOREM **5.5** *Let $\sigma'''_{\ell+1} \geq \cdots \geq \sigma'''_n$ and $\sigma''''_{\ell+1} \geq \cdots \geq \sigma''''_n$ be the finite generalized singular values of $([\tilde{A}_{11}^{(1)}, \tilde{\tilde{A}}_{12}^{(1)}], [\tilde{L}, \mathbf{O}])$ and $(\tilde{A}_{11,2}^{(2)}, \tilde{L})$, respectively. (The remaining $\ell = n - r_B$ values are infinite.) If $\eta \equiv \|(([\tilde{A}_{11}^{(1)}, \tilde{\tilde{A}}_{12}^{(1)}])_c)^\dagger\|_2 \, \|[(\delta\tilde{A}_{11}^{(1)})_c, (\delta\tilde{\tilde{A}}_{12}^{(1)})_c]\|_2 < 1$, then, for $\ell + 1 \leq i \leq n$,*

$$1 - \eta \leq \frac{\sigma''''_i}{\sigma'''_i} \leq 1 + \eta. \tag{5.36}$$

Proof Similar to the proof of Theorem 5.4. The only difference is that we cannot use any special zero pattern of $[\delta\tilde{A}_{11}^{(1)}, \delta\tilde{\tilde{A}}_{12}^{(1)}]$. Q.E.D.

REMARK **5.5** Finally, we may combine the estimate (5.36) with the estimates (5.26), (5.30), (5.34) to get an estimate of σ''''_i / σ_i.

5.4 Analysis of condition numbers

In this section we analyze the condition numbers that determine the relative accuracy of the generalized singular value approximations computed by Algorithm 4.1 in floating–point arithmetic.

For the sake of simplicity, we neglect the fact that condition numbers have their own condition numbers, and thus we consider the condition numbers of matrices obtained in exact computation. We use the notation from Algorithm 4.1.
We first analyze the condition numbers related to the matrix A.

PROPOSITION **5.4** *Let $A^{(1)}$, $U^{(1,1)}$, $U^{(1,2)}$ be as in Algorithm 4.1, and let $\Delta_U = \mathbf{diag}(U^{(1,1)})$, $(U^{(1,1)})_d = \Delta_U^{-1}U^{(1,1)}$, $(U^{(1,2)})_d = \Delta_U^{-1}U^{(1,2)}$. Then*

$$\kappa_2((A^{(1)})_c) \leq \sqrt{n}\kappa_2(A^{(0)})\kappa_2((U^{(1,1)})_d)(1 + \|(U^{(1,2)})_d\|_2)^2. \tag{5.37}$$

Proof Note that

$$\kappa_2(A^{(1)}\Delta_U) \leq \kappa_2(A^{(0)})\kappa_2((U^{(1,1)})_d)(1 + \|(U^{(1,2)})_d\|_2)(1 + \frac{\|(U^{(1,2)})_d\|_2}{\|(U^{(1,1)})_d\|_2}). \tag{5.38}$$

Now a result of van der Sluis [34] implies that

$$\kappa_2((A^{(1)})_c) \leq \sqrt{n} \min_{\Delta=\mathbf{diag}(\Delta)} \kappa_2(A^{(1)}\Delta) \leq \sqrt{n}\kappa_2(A^{(1)}\Delta_U)$$

and (5.37) follows. Q.E.D.

REMARK **5.6** It trivially holds that for $A^{(1)} = [A_{11}^{(1)}, A_{12}^{(1)}]$ the condition numbers satisfy

$$\max\{\kappa_2((A_{11}^{(1)})_c), \kappa_2((A_{12}^{(1)})_c)\} \leq \kappa_2((A^{(1)})_c).$$

The generalized singular values of the pair $(A_{11,2}^{(2)}, \begin{bmatrix} L \\ \hat{L} \end{bmatrix})$ are computed with high relative accuracy by the algorithm from [15] if the condition numbers $\kappa_2((A_{11,2}^{(2)})_c)$ and $\kappa_2((\begin{bmatrix} L \\ \hat{L} \end{bmatrix})_c)$ are moderate. An estimate of $\kappa_2((A_{11,2}^{(2)})_c)$ is given in the following proposition.

PROPOSITION **5.5** *Let*

$$\mathcal{T} = \{\begin{bmatrix} \Delta & \mathbf{O} \\ \Omega & \Xi \end{bmatrix}; \Delta \textit{ diagonal nonsingular, } \Xi \textit{ square nonsingular, } \Omega \textit{ arbitrary}\}$$

where the partition of each $T \in \mathcal{T}$ is conformal to the partition of $A^{(2)}$. Furthermore, let $\mathcal{T}_A = \{T \in \mathcal{T};\ \Delta = \mathrm{diag}\,(\|A_{11,2}^{(2)}e_i\|_2)^{-1}\}$. Then

$$\kappa_2((A_{11,2}^{(2)})_c) \leq \min\{\sqrt{r_B} \min_{T\in\mathcal{T}} \kappa_2(A^{(1)}T),\ \min_{T\in\mathcal{T}_A} \kappa_2(A^{(1)}T)\}. \tag{5.39}$$

Proof Note that for any $T \in \mathcal{T}$

$$\begin{bmatrix} A_{11,1}^{(2)} & A_{12}^{(2)} \\ A_{11,2}^{(2)} & \mathbf{O} \end{bmatrix} T = \begin{bmatrix} \star & \star \\ A_{11,2}^{(2)}\Delta & \mathbf{O} \end{bmatrix}$$

Hence

$$\kappa_2((A_{11,2}^{(2)})_c) \leq \min_{T\in\mathcal{T}_A} \kappa_2(A^{(2)}T), \qquad \min_{D=\mathbf{diag}(D)} \kappa_2(A_{11,2}^{(2)}D) \leq \min_{T\in\mathcal{T}} \kappa_2(A^{(2)}T)$$

and (5.39) follows from the orthogonal invariance and near optimality [34] of $\kappa_2((\cdot)_c)$. Q.E.D.

Thus, if the original matrix A has moderate $\kappa_2(A_c)$, and if $\kappa_2(U_d^{(1,1)})$ is moderate, then in the computed reduced regular pair $(\tilde{A}_{11,2}^{(2)}, \tilde{L})$ the condition number $\kappa_2((\tilde{A}_{11,2}^{(2)})_c)$ is moderate.

Finally, since the LU factorization of B is computed with a rank revealing pivoting, the numbers $\kappa_2(U_d^{(1,1)})$ and $\kappa_2((\begin{bmatrix} L \\ \hat{L} \end{bmatrix})_c)$ are almost never large. (In the description of Algorithm 4.1 we do not specify the choice of pivoting. In practice, we use the best one that is available. The relative accuracy of Algorithm 4.1 improves as our ability to factor B accurately as LU improves.)

6 On software implementation of the algorithm

The major part of our software implementation of Algorithm 4.1 is based on LAPACK and BLAS 3 libraries. We also use the Jacobi SVD algorithm [19], [12], [13] because it is the fastest known method capable of achieving the high relative accuracy. We also use a procedure that computes the LU factorization of B with total pivoting. The output of our algorithm is the GSVD of (A, B) in Van Loan's form. That is, we compute orthogonal matrices V and W and a nonsingular matrix X such that V^TAX and W^TBX are in diagonal form. An interesting feature of our algorithm is an option to return V and W in factored form, using products of the Householder reflections. In that case, information necessary to retrieve the reflections that define V and W is overwritten on the arrays A and B, respectively. Hence, we can compute and use V and W without additional square arrays. This saves $m^2 + p^2$ memory locations which is attractive if $m \gg n$ or $p \gg n$.

6.1 Test results

We use several different types of test pairs. The first type is taken from [15] and it contains full column rank matrices with controlled spectral condition number $\kappa_2((\cdot)_c)$ of the column equilibrated matrix. For the reader's convenience we give a detailed description of the test pair generation.

EXAMPLE **6.1** We generate random full column rank matrices A_c and B_c with given $\kappa_2(A_c)$ and $\kappa_2(B_c)$ and apply scalings $A = A_c\Delta_A$, $B = B_c\Delta_B$, where Δ_A, Δ_B are random diagonal, nonsingular matrices with given spectral condition numbers.
Each 4–tuple $(\kappa_2(A_c), \kappa_2(\Delta_A), \kappa_2(B_c), \kappa_2(\Delta_B))$ is chosen from a 4–dimensional mesh of condition numbers

$$\mathcal{C} = \left\{\kappa_{ijkl} = (10^i, 10^j, 10^k, 10^l) \ : \ (i,j,k,l) \in \mathcal{I} \times \mathcal{J} \times \mathcal{K} \times \mathcal{L} \subset \mathbf{N}^4\right\},$$

where $\mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}$ are determined at the very beginning of the test and kept fixed. For each fixed κ_{ijkl}, we generate A_c, Δ_A, B_c, Δ_B using different distributions of their singular vlues. We use all admissible values of the parameter **MODE** in LAPACK's **DLATM1()** procedure [11]. Hence, for each 4–tuple $(A_c, \Delta_A, B_c, \Delta_B)$ we can choose the singular value distribution modes from the set

$$\mathcal{M} = \{\mu_{i'j'k'l'} = (\mu_{i'}, \mu_{j'}, \mu_{k'}, \mu_{l'})\} \subseteq \mathcal{P}_1 \times \mathcal{P}_2 \times \mathcal{P}_3 \times \mathcal{P}_4 \subseteq \{\pm 1, \ldots, \pm 6\}^4,$$

For each fixed $(\kappa_{ijkl}, \mu_{i'j'k'l'})$ we generate random pairs using different random number generators as specified by the parameter **IDIST** in **DLATM1()** procedure. Thus, our set of random number distributions is $\mathcal{R} \subseteq \{\mathcal{U}(-1,1), \mathcal{U}(0,1), \mathcal{N}(0,1)\}$, where $\mathcal{U}(-1,1)$, $\mathcal{U}(0,1)$ are uniform distributions on $(-1,1)$ and $(0,1)$, respectively, and $\mathcal{N}(0,1)$ is the normal distribution. For each fixed distribution $\chi \in \mathcal{R}$ we generate a set $\mathcal{E}^{\chi}_{\kappa_{ijkl},\mu_{i'j'k'l'}}$ of different pairs, where the cardinality of $\mathcal{E}^{\chi}_{\kappa_{ijkl},\mu_{i'j'k'l'}}$ is fixed at the very beginning of the test. This process makes a total of

$$\tau \equiv |\mathcal{I}|\,|\mathcal{J}|\,|\mathcal{K}|\,|\mathcal{L}|\,|\mathcal{M}|$$

different classes and $\tau \prod_{\chi\in\mathcal{R}} |\mathcal{E}^{\chi}_{\kappa_{ijkl},\mu_{i'j'k'l'}}|$ different matrix pairs.

Each test pair is generated in double precision and its generalized singular values are computed using a double precision procedure. The generalized singular values computed by the double precision procedure are then taken as reference for the single precision procedure run on the original pair rounded to single precision.

For a test pair (A, B) with well–conditioned A_c and B_c, the value of

$$\zeta_1(A,B) = \varepsilon \max\{\kappa_2(A_c), \kappa_2(B_c)\}$$

gives a good estimate of relative errors in the generalized singular values computed by the algorithm from [15].

Following our theory, we compute another *a priori* relative error estimate in the following way. For computed triangular factors $\tilde{L}$, $\tilde{U}$ of $\Pi_1 B \Pi_2$, we compute the residual $\delta B = \tilde{L}\tilde{U} - \Pi_1 B \Pi_2$ and the values

$$\begin{aligned} \tilde{\varepsilon}_L(B) &= \| \, |\tilde{L}| \, \mathbf{tril}(\delta B) \, |\tilde{L}^{-1}| \, \|_1, \\ \tilde{\varepsilon}_U(B) &= \| \, |\tilde{U}^{-1}| \, \overline{\mathbf{triu}}(\delta B) \, |\tilde{U}| \, \|_1, \end{aligned}$$

(cf. (5.23) and (5.24) in the proof of Theorem 5.2) and

$$\zeta_2(A,B) = \max\{ \, \varepsilon \kappa_2(A_S), \tilde{\varepsilon}_L(B), \tilde{\varepsilon}_U(B) \, \}.$$

(Here $\| \cdot \|_1$ is the operator norm induced by the ℓ_1 vector norm. We use $\| \cdot \|_1$ instead of $\| \cdot \|_2$ because $\| \cdot \|_1$ is easier to compute.) If $\zeta_1(A,B)$ and $\zeta_2(A,B)$ are realistic and sharp enough to be used in the practice, then the values of

$$\theta_1(A,B) = \frac{\max_{\sigma \in \sigma(A,B)} \frac{|\delta\sigma|}{\sigma}}{\zeta_1(A,B)}, \quad \theta_2(A,B) = \frac{\max_{\sigma \in \sigma(A,B)} \frac{|\delta\sigma|}{\sigma}}{\zeta_2(A,B)}$$

should be bounded by a moderate function of m, n, p and should not be much less than one. (A value of $\theta_i(A,B)$ that is below one means that $\zeta_i(A,B)$ overestimates the actual relative error.)

We also use the following measure for the accuracy of our algorithm:

$$\varepsilon(i,k) = \max_{\kappa_2(A_c)=10^i, \kappa_2(B_c)=10^k} \max_{\sigma \in \sigma(A,B)} \frac{|\delta\sigma|}{\sigma}, \quad (i,k) \in \mathcal{I} \times \mathcal{K},$$

that is, we compute the maximal relative error over all generalized singular values of all matrix pairs with fixed $\kappa_2(A_c) = 10^i$, $\kappa_2(B_c) = 10^k$. Note that $-\log_{10} \varepsilon(i,k)$ gives an approximate minimal number of correct digits in the computed approximations of the generalized singular values of the test pairs with fixed "coordinates" $(i,k) \in \mathcal{I} \times \mathcal{K}$. According to the theory from [15], we can expect $-\log_{10} \varepsilon(i,k)$ to be roughly $7 - \max\{i,k\}$.

To inspect the values of some relevant condition numbers, we also compute

$$\theta_3(B) = \frac{\kappa_2(B_c)}{\max\{\tilde{\varepsilon}_L(B), \tilde{\varepsilon}_U(B)\}/\varepsilon}, \quad \theta_4(B) = \| \, |U^{-1}| \cdot |U| \, \|_1.$$

An example of the above described values is given in Figures 3, 4, 5. The input data are

$$\begin{aligned} \mathcal{I} &= \{2, \ldots, 7\}, \ \mathcal{K} = \mathcal{I}, \\ \mathcal{J} &= \{4, 8, 10, 12, 14, 16\}, \ \mathcal{L} = \mathcal{J}, \\ \mathcal{M} &= \{(5, 4, -5, 3), (3, -4, 5, -3), (4, 5, 3, -4)\}, \ \mathcal{R} = \{\mathcal{U}(-1, 1)\}. \end{aligned}$$

For each node of $\mathcal{C} \times \mathcal{M} \times \mathcal{R}$ we performed one test on a randomly generated pair. As a reference, we use the double precision algorithm from [15]. The values of θ_1 and θ_2 in Figure 3 are bounded by 100 (roughly), which means that the accumulated round–off enters the error linearly in matrix dimensions. Both ζ_1 and ζ_2 provide good relative error estimates, although ζ_1 is slightly more pessimistic. The number of correct digits shown in Figure 4 corresponds to the predicted theoretical behavior. The values of θ_3 in Figure 5 show that, in this example, the condition numbers $\kappa_2(B_c)$ and $\max\{\tilde{\varepsilon}_L(B), \tilde{\varepsilon}_U(B)\}/\varepsilon$ differ by a factor on the order of the dimensions of the problem. This means that in this example the method from [15] is as accurate as Algorithm 4.1. Finally, the values of θ_4 are, as expected, bounded by a factor of dimensionality.

Our next example illustrates a substantial difference between the direct application of the algorithm from [15] and our new approach that starts with the LU factorization of B.

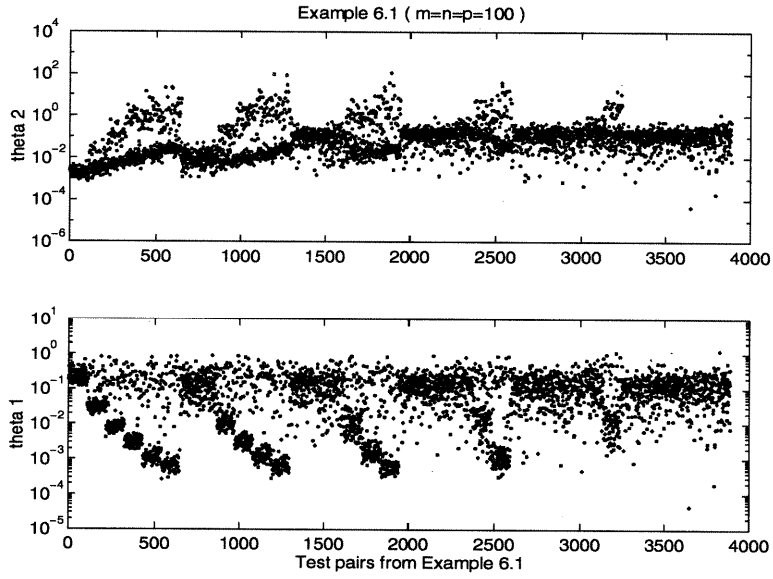


Figure 3: The values of $\theta_1(\cdot,\cdot)$ and $\theta_2(\cdot,\cdot)$.

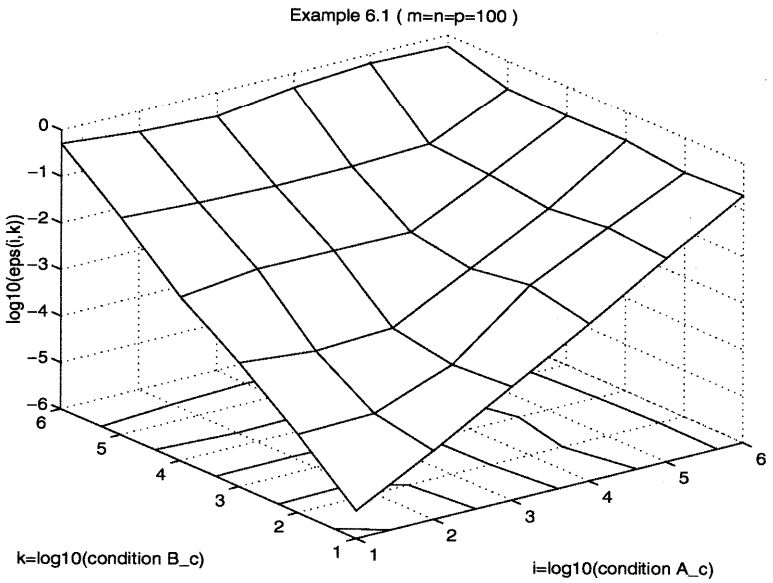


Figure 4: The values of $\log_{10}\varepsilon(i,k)$, $(i,k) \in \mathcal{I} \times \mathcal{K}$. Observe that $-\log_{10}\varepsilon(i,k)$ behaves like $7 - \max\{i,k\}$ (roughly).

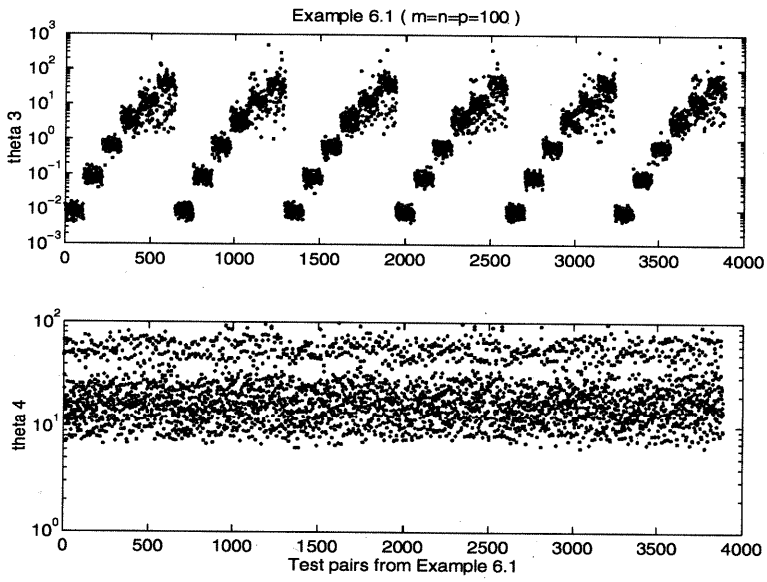


Figure 5: The values of $\theta_3(\cdot,\cdot)$ and $\theta_4(\cdot,\cdot)$.

EXAMPLE **6.2** In this example, the test matrix generator follows the scheme described in Example 6.1 and additionally scales the rows of each generated B by a diagonal ill–conditioned matrix D. That is, we first generate a random matrix B_S in the same way as the matrix B_c in Example 6.1. Then we generate diagonal matrices Δ_B, $\Delta_B^{(1)}$ with the same spectral condition number and compute $B = \Delta_B^{(1)} B_S \Delta_B$. In this way we obtain a matrix B with high $\kappa_2(B_c)$. Tests with LAPACK's `SGGSVD()` and the algorithm from [15] show that neither of those procedures is capable of achieving high relative accuracy with such a matrix. Since we use $\kappa_2(\Delta_B^{(1)})$ up to 10^{16}, we cannot use the double precision algorithm from [15] as a reference. We use a double precision implementation of Algorithm 4.1 instead.

The input data in this example are the same as in Example 6.1. The test results are given in Figures 6, 7, 8. Note that the values of θ_2 show the same behavior as in Example 6.1, while the values of θ_1 are much smaller. This indicates a rather pessimistic estimate if we use ζ_1, and together with large values of θ_3 illustrates the superiority of Algorithm 4.1 to the direct application of the algorithm from [15]. The number of correct digits, shown in Figure 7, confirms that the accuracy is as good as in Example 6.1.

EXAMPLE **6.3** In this example we first generate an $m \times n$ matrix A and an $n \times p$ matrix C in the same way as A and B, respectively, in Example 6.1. Then we define $B = C^T$. In this way we control the size of $\kappa_2(B_r)$, where B_r is obtained from B by scaling its rows to have unit Euclidean norm. Note that in this example $\kappa_2(B_r)$ is used instead of $\kappa_2(B_c)$ in the definitions of $\zeta_1(A,B)$ and $\theta_3(B)$. Also note that the algorithm from [15] is not applicable because B does not have full column rank.

The test results are given in Figures 9, 10, 11. We see that the comments about the relative accuracy from Example 6.1 apply here as well.

EXAMPLE **6.4** We generate A and B in the same way as in Example 6.3, but with larger dimensions, $m = 300$, $n = 150$, $p = 50$, and with $\mathcal{I} = \mathcal{K} = \{2,\ldots,6\}$ and $\mathcal{J} = \mathcal{L} = \{8,12,14,16\}$. For

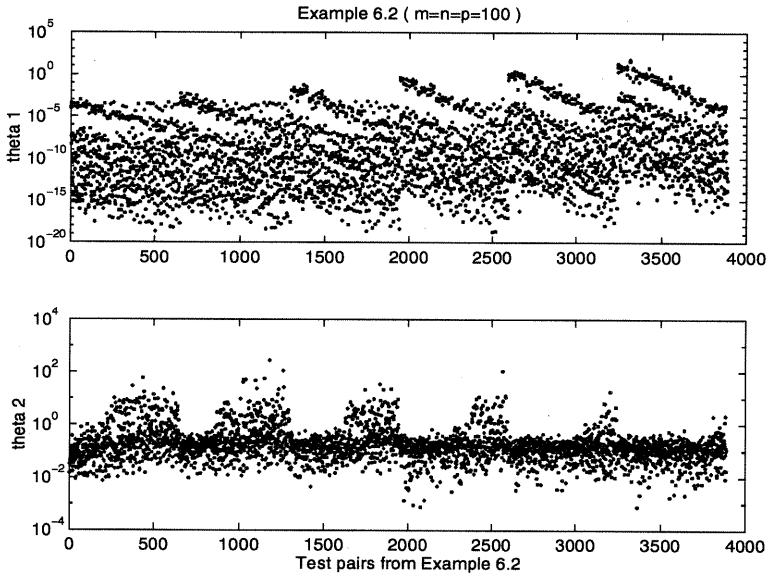


Figure 6: The values of $\theta_1(\cdot,\cdot)$ and $\theta_2(\cdot,\cdot)$.

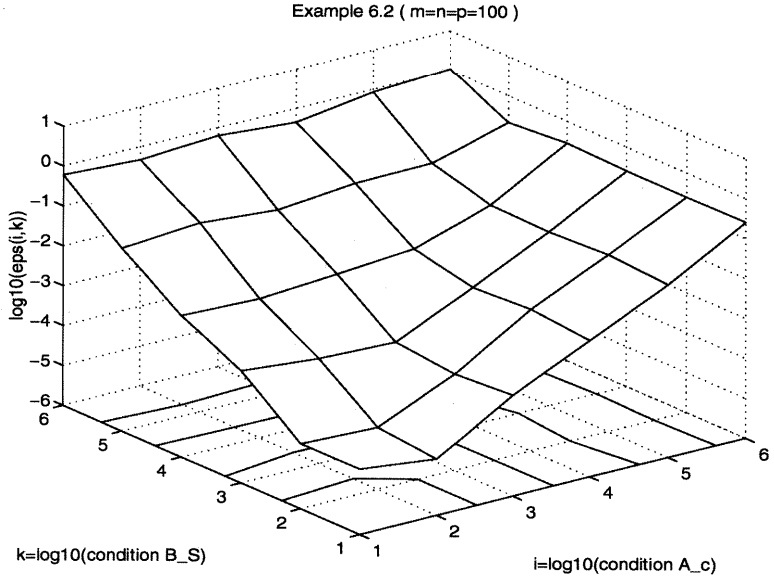


Figure 7: The values of $\log_{10}\varepsilon(i,k)$, $(i,k)\in\mathcal{I}\times\mathcal{K}$.

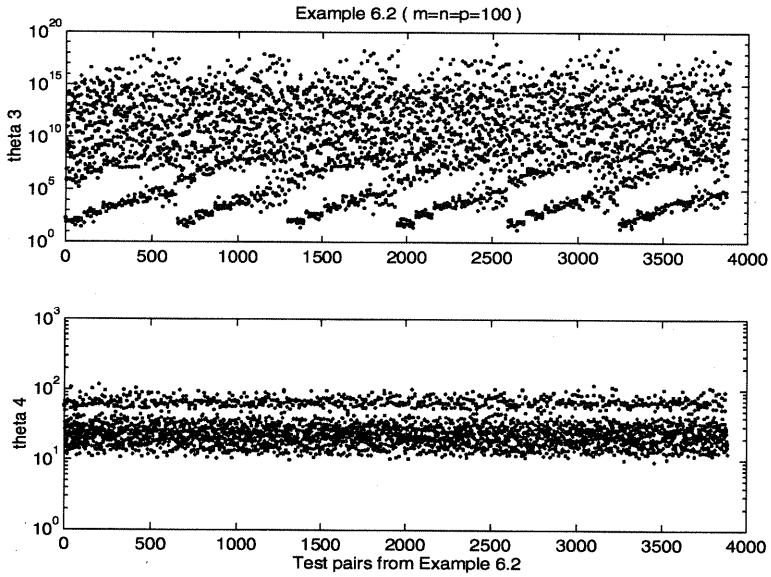


Figure 8: The values of $\theta_3(\cdot,\cdot)$ and $\theta_4(\cdot,\cdot)$.

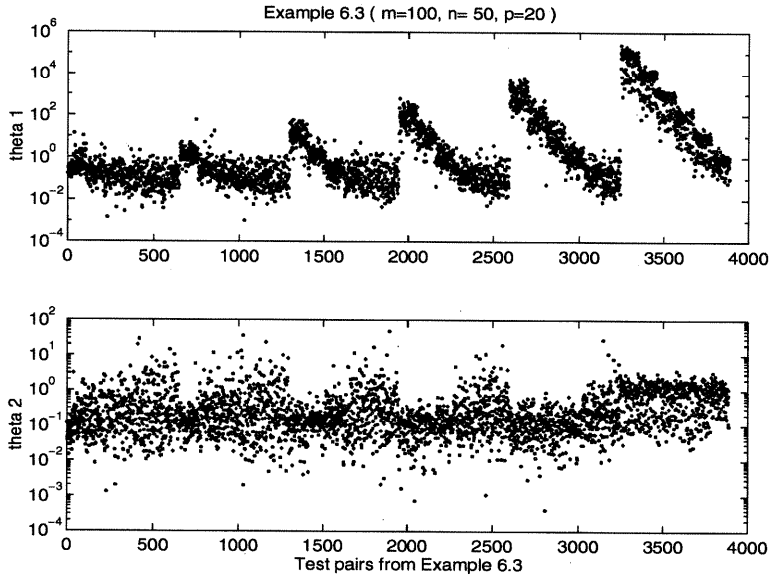


Figure 9: The values of $\theta_1(\cdot,\cdot)$ and $\theta_2(\cdot,\cdot)$.

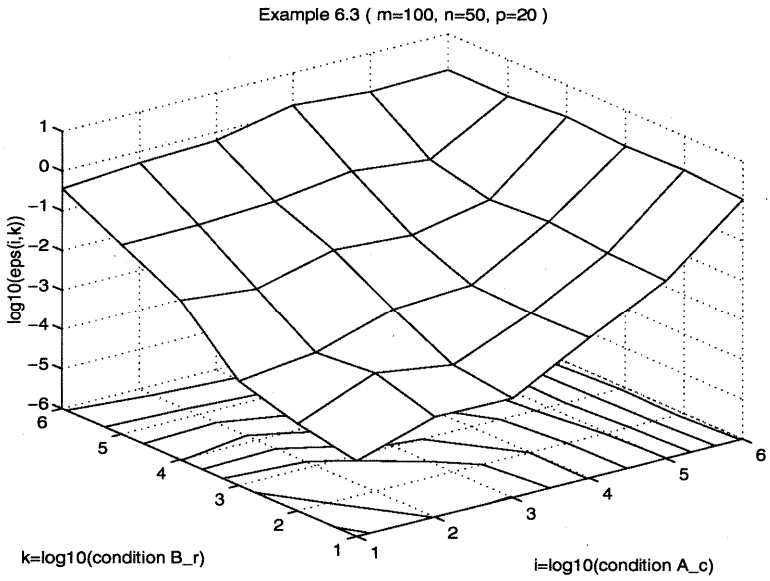


Figure 10: The values of $\log_{10} \varepsilon(i,k)$, $(i,k) \in \mathcal{I} \times \mathcal{K}$.

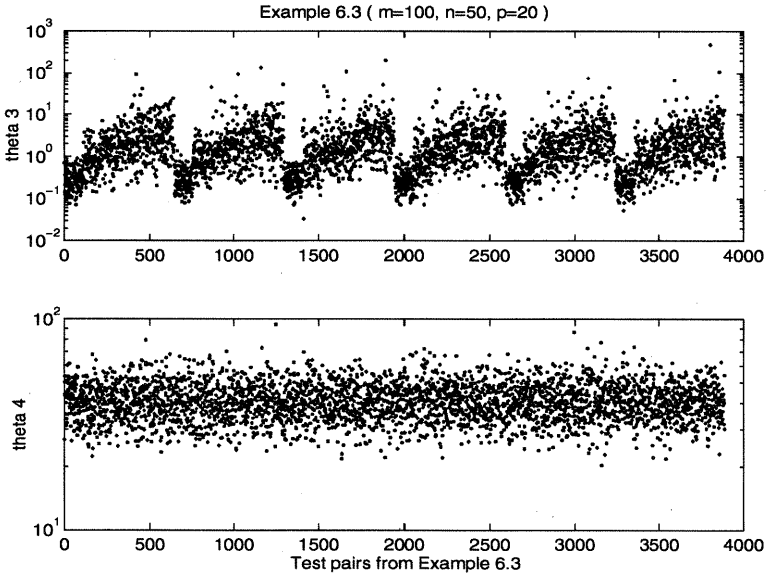


Figure 11: The values of $\theta_3(\cdot,\cdot)$ and $\theta_4(\cdot,\cdot)$.

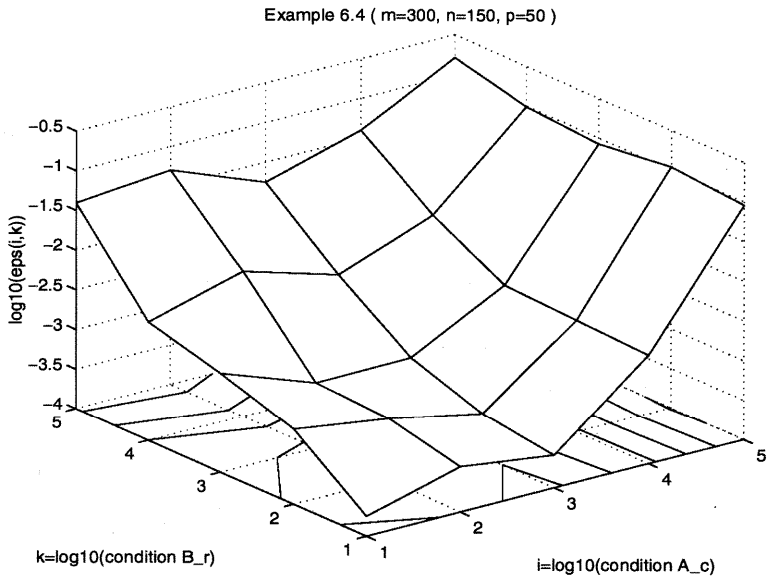


Figure 12: The values of $\log_{10} \varepsilon(i,k)$, $(i,k) \in \mathcal{I} \times \mathcal{K}$.

simplicity, we display only the values of $\varepsilon(i,k)$ (Figure 12). We can see the minimal number of correct digits shows the same behavior as in all previous examples.

References

[1] E. Anderson, Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. D. Croz, A. Greenbaum, S. Hammarling, A. McKenny, S. Ostrouchov, and D. Sorensen, *LAPACK users' guide*, SIAM, 1992.

[2] Z. Bai, *The CSD, GSVD, their applications and computations*, Technical Report 958, IMA Preprint Series, April 1992.

[3] Z. Bai and J. Demmel, *Computing the generalized singular value decomposition*, LAPACK Working Note 46, University of Tennessee, Computer Science Department, May 1992.

[4] Z. Bai and H. Zha, *A new preprocessing algorithm for the computation of the generalized singular value decomposition*, SIAM J. Sci. Stat. Comp., 14 (1993), pp. 1007–1012.

[5] J. Barlow and J. Demmel, *Computing accurate eigensystems of scaled diagonally dominant matrices*, SIAM J. Num. Anal., 27 (1990), pp. 762–791.

[6] F. Bauer, *Genauigkeitsfragen bei der Lösung linearer Gleichungssysteme*, ZAMM, 46 (1966), pp. 409–421.

[7] Å. Björk, *Numerical Methods for Least Squares Problems*, SIAM, 1996.

[8] A. Deichmöller, *Über die Berechnung verallgemeinerter singulärer Werte mittels Jacobi–änlicher Verfahren*, PhD thesis, Lehrgebiet Mathematische Physik, Fernuniversität Hagen, 1991.

[9] A. Deichmöller and K. Veselić, *Two algorithms for computing the symmetric positive definite generalized eigenvalue problem and the generalized singular values of full column rank matrices.* Preprint, LG Mathematische Physik, Fernuniversität Hagen, D-58084 Hagen, 1991.

[10] J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselić, and Z. Drmač, *Computing the singular value decomposition with high relative accuracy.* Document in preparation, 1996.

[11] J. Demmel and A. McKenney, *A test matrix generation suite*, LAPACK Working Note 9, Courant Institute, New York, March 1989.

[12] J. Demmel and K. Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 1204–1245.

[13] Z. Drmač, *Implementation of Jacobi rotations for accurate singular value computation in floating point arithmetic.* SIAM J. Sci. Comp., to appear.

[14] ——, *Computing the Singular and the Generalized Singular Values*, PhD thesis, Lehrgebiet Mathematische Physik, Fernuniversität Hagen, 1994.

[15] ——, *The tangent and the cosine–sine algorithm for computing the generalized eigenvalue and singular value decompositions.* Submitted to SIAM J. Numer. Anal., July 1995.

[16] S. Falk and P. Langemeyer, *Das Jacobische Rotationsverfahren für reellsymmetrische Matrizenpaare I, II*, ZAMM, 0 (1960), pp. 30–43.

[17] W. M. Gentleman, *Error analysis of QR decompositions by Givens transformations*, Linear Algebra Appl., 10 (1975), pp. 189–197.

[18] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 1989.

[19] M. R. HESTENES, *Inversion of matrices by biorthogonalization and related results*, J. SIAM, 6 (1958), pp. 51–90.

[20] N. J. HIGHAM, *The accuracy of solution to triangular systems*, SIAM J. Num. Anal., 26 (1989), pp. 1252–1265.

[21] ——, *Accuracy and Stability of Numerical Algorithms*, SIAM, 1996.

[22] R.-C. LI, *Bounds on perturbations of generalized singular values and of associated subspaces*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 195–234.

[23] ——, *Relative perturbation theory: (I) Eigenvalue and singular value variations*, technical report, Mathematical Science Section, Oak Ridge National Laboratory, Oak Ridge, TN 37831–6367, January 1996.

[24] C. C. PAIGE, *A note on a result of Sun Ji–guang: Sensitivity of the CS and GSV decompositions*, SIAM J. Num. Anal., 21 (1984), pp. 186–191.

[25] ——, *The general linear model and the generalized singular value decomposition*, Linear Algebra Appl., 70 (1985), pp. 269–284.

[26] ——, *Computing the generalized singular value decomposition*, SIAM J. Sci. Stat. Comp., 7 (1986), pp. 1126–1146.

[27] C. C. PAIGE AND M. A. SAUNDERS, *Towards a generalized singular value decomposition*, SIAM J. Num. Anal., 18 (1981), pp. 398–405.

[28] R. D. SKEEL, *Scaling for numerical stability in Gaussian elimination*, Journal of the Association for Computing Machinery, 26 (1979), pp. 494–526.

[29] G. W. STEWART, *Computing the CS decomposition of a partitioned orthonormal matrix*, Numer. Math., 40 (1982), pp. 297–306.

[30] ——, *A method for computing the generalized singular value decomposition*, in Matrix Pencils Proceedings of a Conference, Springer Verlag, 1982.

[31] J.-G. SUN, *Perturbation analysis for the generalized eigenvalue and the generalized singular value problem*, in Matrix Pencils Proceedings of a Conference, Springer Verlag, 1982.

[32] ——, *Perturbation analysis for the generalized singular value problem*, SIAM J. Num. Anal., 20 (1983), pp. 611–625.

[33] ——, *Componentwise perturbation bounds for some matrix decompositions*, BIT, 32 (1992), pp. 702–714.

[34] A. VAN DER SLUIS, *Condition numbers and equilibration of matrices*, Numer. Math., 14 (1969), pp. 14–23.

[35] C. F. VAN LOAN, *Generalized Singular Values with Algorithms and Applications*, PhD thesis, University of Michigan, 1973.

[36] ——, *Generalizing the singular value decomposition*, SIAM J. Num. Anal., 13 (1976), pp. 76–83.

[37] ——, *A generalized SVD analysis of some weighting methods for equality constrained least squares*, in Matrix Pencils Proceedings of a Conference, Springer Verlag, 1982.

[38] ——, *Computing the CS and the generalized singular value decomposition*, Numer. Math., 46 (1985), pp. 479–491.

[39] J. H. WILKINSON, *Rounding Errors in Algebraic Processes*, Prentice–Hall, Inc., 1963.