Accuracy of the Hari–Zimmermann method for the generalized singular value decomposition[☆]

Sanja Singer[a,*], Saša Singer[b]

[a] *University of Zagreb, Faculty of Mechanical Engineering and Naval Architecture, I. Lučića 5, 10000 Zagreb, Croatia*

[b] *University of Zagreb, Faculty of Science, Department of Mathematics, Bijenička cesta 30, 10000 Zagreb, Croatia*

Abstract

We develop an error analysis for the generalized singular values of the pair (F, G) computed by the implicit Hari–Zimmermann algorithm. We obtain that, after one step of independent transformations, the generalized singular values are computed with high relative accuracy if the norms of columns of F and G have similar magnitudes. Similar holds for the generalized singular values obtained after completion of the algorithm.

Keywords: generalized singular values, Hari–Zimmermann algorithm, accuracy
2000 MSC: 65F15, 65Y20, 65G50

1. Introduction

The Generalized Singular Value Decomposition (GSVD) of the matrix pair (F, G) can be computed in various ways. For example, LAPACK `double precision` routine `DGGSVD` consists of the preprocessing step (subroutine `DGGSVP`, to make F and G triangular and G nonsingular), and the modified Kogbetliantz algorithm (subroutine `DTGSJA`).

The GSVD of a matrix pair (F, G) with G nonsingular can be viewed as the Generalized Eigenvalue Problem (GEP) for the pair $(A, B) := (F^*F, G^*G)$. Therefore all the methods for the GEP can implicitly be applied to the pair (F, G).

In [6], implicit application of the Hari–Zimmermann method (normalized variant of the Falk–Langemeyer method) for the computation of the GSVD is studied. Numerical tests show that the proposed method can be significantly faster and more accurate than the corresponding LAPACK routine `DTGSJA`. In this paper we derive a proof of accuracy of the implicit Hari–Zimmermann method.

[☆] This work has been fully supported by Croatian Science Foundation under the project IP-2014-09-3670.

[*] Corresponding author.
Email addresses: `ssinger@math.hr` (Sanja Singer), `singer@math.hr` (Saša Singer)

The paper is organized as follows. Section 2 contains brief description of the one-sided Hari–Zimmermann algorithm for the computation of the generalized singular values. In Section 3 we give perturbation bounds for a single Hari–Zimmermann transformation, while in Section 4 we present the perturbation bound for the sequence of independent transformations, as well as the main perturbation theorem.

2. The one-sided Hari–Zimmermann algorithm for the GSVD

For the sake of completeness, we will present here the Hari–Zimermann method constructed in [3] for the GEP, and modified in [6] for the GSVD.

To apply the method, matrix G in the pair (F, G) must be of full column rank, and for an efficient implementation, matrices F and G should be square. If pair (F, G) is not already in the required form, that can be achieved by the applying the LAPACK subroutine `DGGSVP`.

Since G is of full column rank, pair (F, G) can be scaled from the right-hand side to make the norms of the columns of the newly computed matrix $G^{(1)}$ equal to one. This task is easily performed by taking

$$D = \operatorname{diag}\left(\frac{1}{\|g_{11}\|_2}, \frac{1}{\|g_{22}\|_2}, \ldots, \frac{1}{\|g_{kk}\|_2}\right),$$

and the pair (F, G) is transformed into

$$F^{(1)} := FD, \quad G^{(1)} := GD.$$

Note that the new pair $(F^{(1)}, G^{(1)})$ has the same generalized singular values as the pair (F, G).

The one-sided Hari–Zimmermann method constructs a sequence of matrix pairs

$$F^{(\ell+1)} = F^{(\ell)} Z_\ell, \quad G^{(\ell+1)} = G^{(\ell)} Z_\ell, \quad \ell \in \mathbb{N}. \tag{2.1}$$

In step ℓ, matrix Z_ℓ in (2.1) is chosen to orthogonalize columns i and j in the matrix pair $(F^{(\ell)}, G^{(\ell)})$, and to keep the column norms of $G^{(\ell)}$ equal to one. If the matrix Z, that orthogonalizes the columns of the pair (F, G) is needed, the accumulation procedure for Z is very similar to (2.1), i.e.,

$$Z^{(1)} = I, \quad Z^{(\ell+1)} = Z^{(\ell)} Z_\ell.$$

Instead of one unknown angle in a plane rotation, matrix Z_ℓ has two unknown angles, to be able to orthogonalize two pairs of columns—one pair in each matrix F and G. The transformations are easier to be written in terms of elements of the pair $(A^{(\ell)}, B^{(\ell)}) := ([F^{(\ell)}]^T F^{(\ell)}, [G^{(\ell)}]^T G^{(\ell)})$. If the pivot columns of $F^{(\ell)}$ and $G^{(\ell)}$ are denoted by $f_i^{(\ell)}$, $f_j^{(\ell)}$,

$g_i^{(\ell)}$, and $g_j^{(\ell)}$, then the elements of $\widehat{A}^{(\ell)}$ and $\widehat{B}^{(\ell)}$ are the inner products of pivot columns of $F^{(\ell)}$ and $G^{(\ell)}$, respectively,

$$a_{ii}^{(\ell)} = \|f_i^{(\ell)}\|_2^2, \quad a_{jj}^{(\ell)} = \|f_j^{(\ell)}\|_2^2, \quad a_{ij}^{(\ell)} = [f_i^{(\ell)}]^T f_j^{(\ell)}, \quad b_{ij}^{(\ell)} = [g_i^{(\ell)}]^T g_j^{(\ell)}. \tag{2.2}$$

Then, the corresponding 2×2 pivot pair of $(A^{(\ell)}, B^{(\ell)})$ is denoted by $(\widehat{A}^{(\ell)}, \widehat{B}^{(\ell)})$, where

$$\widehat{A}^{(\ell)} = \begin{bmatrix} a_{ii}^{(\ell)} & a_{ij}^{(\ell)} \\ a_{ij}^{(\ell)} & a_{jj}^{(\ell)} \end{bmatrix}, \qquad \widehat{B}^{(\ell)} = \begin{bmatrix} b_{ii}^{(\ell)} & b_{ij}^{(\ell)} \\ b_{ij}^{(\ell)} & b_{jj}^{(\ell)} \end{bmatrix}. \tag{2.3}$$

The matrix Z_ℓ is the identity matrix, except in the plane (i, j), where its restriction $\widehat{Z}_\ell$ is equal to

$$\widehat{Z}_\ell = \frac{1}{\sqrt{1 - (b_{ij}^{(\ell)})^2}} \begin{bmatrix} \cos\varphi_\ell & \sin\varphi_\ell \\ -\sin\psi_\ell & \cos\psi_\ell \end{bmatrix}, \tag{2.4}$$

with

$$\begin{aligned} \cos\varphi_\ell &= \cos\vartheta_\ell + \xi_\ell(\sin\vartheta_\ell - \eta_\ell\cos\vartheta_\ell), & \cos\psi_\ell &= \cos\vartheta_\ell - \xi_\ell(\sin\vartheta_\ell + \eta_\ell\cos\vartheta_\ell), \\ \sin\varphi_\ell &= \sin\vartheta_\ell - \xi_\ell(\cos\vartheta_\ell + \eta_\ell\sin\vartheta_\ell), & \sin\psi_\ell &= \sin\vartheta_\ell + \xi_\ell(\cos\vartheta_\ell - \eta_\ell\sin\vartheta_\ell), \\ \xi_\ell &= \frac{b_{ij}^{(\ell)}}{\sqrt{1 + b_{ij}^{(\ell)}} + \sqrt{1 - b_{ij}^{(\ell)}}}, & \eta_\ell &= \frac{b_{ij}^{(\ell)}}{\left(1 + \sqrt{1 + b_{ij}^{(\ell)}}\right)\left(1 + \sqrt{1 - b_{ij}^{(\ell)}}\right)}, \\ \tan(2\vartheta_\ell) &= \frac{2a_{ij}^{(\ell)} - (a_{ii}^{(\ell)} + a_{jj}^{(\ell)})b_{ij}^{(\ell)}}{(a_{jj}^{(\ell)} - a_{ii}^{(\ell)})\sqrt{1 - (b_{ij}^{(\ell)})^2}}, & -\frac{\pi}{4} &< \vartheta_\ell \le \frac{\pi}{4}. \end{aligned} \tag{2.5}$$

There are some special cases. If $a_{ij}^{(\ell)} = b_{ij}^{(\ell)} = 0$, obviously both pivot submatrices are already diagonal, and we can choose $\vartheta_\ell = 0$ to make the transformation identity, i.e., not to apply it. If $a_{ii}^{(\ell)} = a_{jj}^{(\ell)}$, and $2a_{ij}^{(\ell)} = (a_{ii}^{(\ell)} + a_{jj}^{(\ell)})b_{ij}^{(\ell)}$, then the matrices $\widehat{A}^{(\ell)}$ and $\widehat{B}^{(\ell)}$ are proportional. Hence, we set $\vartheta_\ell = \frac{\pi}{4}$, unless $a_{ij}^{(\ell)} = b_{ij}^{(\ell)} = 0$.

Hari in his Ph.D. thesis proved [3, Proposition 2.4] that

$$\min\{\cos\varphi_\ell, \cos\psi_\ell\} > 0, \tag{2.6}$$

and

$$-1 < \tan\varphi_\ell \tan\psi_\ell \le 1, \tag{2.7}$$

with $\tan\varphi_\ell \tan\psi_\ell = 1$, if and only if $\vartheta_\ell = \frac{\pi}{4}$. Therefore $\widehat{Z}_\ell$ is nonsingular. More precisely, it holds (see [6])

$$\det(\widehat{Z}_\ell) = \frac{1}{\sqrt{1 - (b_{ij}^{(\ell)})^2}} > 0. \tag{2.8}$$

It is easy to show that the condition number of $\widehat{Z}_\ell$, the restriction of the transformation matrix, is

$$\|\widehat{Z}_\ell\|_2 = \frac{|\cos(\psi_\ell - \varphi_\ell)|}{1 - |\sin(\psi_\ell - \varphi_\ell)|} = \frac{|(1 - \xi_\ell\eta_\ell)^2 - \xi_\ell^2|}{1 - |2\xi_\ell(1 - \xi_\ell\eta_\ell)|}.$$

The last equality shows that it does not depend explicitly on ϑ_ℓ, since the previous formula can be written in terms of $b_{ij}^{(\ell)}$. As we expected, the condition number depends only on the distance of the submatrix $\widehat{B}^{(\ell)}$ from the singularity, i.e., how far is $|b_{ij}^{(\ell)}|$ from 1 (the effect of the size of $|b_{ij}^{(\ell)}|$ on the condition number of $\widehat{Z}_\ell$ is shown in Figure 2.1). The one-sided Hari–Zimmerman algorithm is given in Algorithm 2.1.

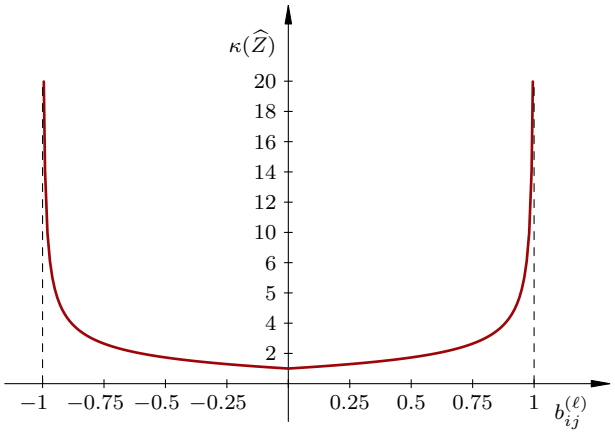

Figure 2.1: Effect of size of the element $b_{ij}^{(\ell)}$ on the condition number of $\widehat{Z}$.

3. Backward error analysis of the implicit Hari–Zimmermann algorithm

In this section we will prove that the implicit Hari–Zimmermann method is backward stable, and, if the matrices permit, computes the generalized singular values with small relative errors.

Let $\mathbb{F} \subset \mathbb{R}$ denote the set of exactly representable floating–point numbers, without the gradual underflow range. In other words, if $\gamma \in \mathbb{F}$, then $\gamma = 0$, or $|\gamma|$ lies within the underflow and overflow limits. Let $fl(\ )$ denote the computed value of the argument, in the finite precision floating–point arithmetic with the unit roundoff error ε. We assume that fl satisfies the following relations (see [5])

$$fl(x \circ y) = (1 + \varepsilon_\circ)(x \circ y), \quad |\varepsilon_\circ| \leq \varepsilon, \tag{3.1}$$

$$fl(\sqrt{x}) = (1 + \varepsilon_{\sqrt{}})\sqrt{x}, \quad |\varepsilon_{\sqrt{}}| \leq \varepsilon, \tag{3.2}$$

$$fl(x \cdot y + z)) = (1 + \varepsilon_{\text{fma}})(x \cdot y + z)), \quad |\varepsilon_{\text{fma}}| \leq \varepsilon, \tag{3.3}$$

Algorithm 2.1: Implicit cyclic Hari–Zimmermann algorithm.

Description: Algorithm `HZ_Orthog` computes the generalized singular values of a matrix pair (F, G) (full GSVD algorithm is given in [6]). Constant *max_sw* denotes the maximal number of allowed sweeps. No transformations in a sweep means that all the transformations have been computed as the identity matrices.

```
HZ_Orthog(inout :: F, G, Z,  in :: k, acc, max_sw);
begin
    it = 0;
    repeat   // sweep loop
        it = it + 1;
        for all pairs (i, j), 1 ≤ i < j ≤ k do
            compute Â and B̂ from (2.2);
            compute the elements of Ẑ by using (2.5);
            // transform F and G
            [f_i, f_j] = [f_i, f_j] · Ẑ;
            [g_i, g_j] = [g_i, g_j] · Ẑ;
        end for
    until (no transf. in this sweep) or (it ≥ max_sw);
    // compute the generalized singular values of (F, G)
    for i = 1 to k do
        σ_i = ||f_i||_2;
    end for
end
```

where $\circ$ is any of the four elementary arithmetic operations, whenever the input arguments of $f\ell$ belong to $\mathbb{F}$, and the correctly rounded result is within the range of $\mathbb{F}$. The operation in (3.3) is known as the "fused multiply and add" (FMA), performed with only one rounding, which is implemented in many modern processors. If FMA is not supported, then the error in (3.3) is bounded by 2ε, and all the results below are still valid, but with slightly larger numerical constants in all bounds. From now on, $\varepsilon_1, \varepsilon_2, \ldots$, with numerical subscripts, denote the relative errors of individual operations, i.e., the quantities absolutely bounded by the unit roundoff error.

The pointwise one-sided transformations on the matrices $F^{(\ell)}$ and $G^{(\ell)}$ in Algorithm 2.1 are equivalent to the respective congruence (2.1) on the matrix pair

$$(A^{(\ell)}, B^{(\ell)}) := ((F^{(\ell)})^T F^{(\ell)}, (G^{(\ell)})^T G^{(\ell)}).$$

These transformations do not alter the generalized eigenvalues of A and B, i.e., the generalized singular values of F and G, regardless of how we compute the parameters of the

transformation matrix $\widehat{Z}_\ell$ in (2.4), as long as the transformation has the same structure as in (2.4) and is nonsingular. Therefore, the stability of the algorithm depends solely on how much the *computed* matrices after the transformation differ from *exact* results of *some* transformation (not necessarily the same) with the structure (2.4). In the backward stability terms, the computed results are interpreted as the *exact* results of a "structure preserving" transformation from (2.4), applied on somewhat perturbed initial matrices $F + \delta F$ and $G + \delta G$, with perturbed generalized singular values. Structure preservation implies that the "backwardly" perturbed generalized singular values are exactly equal to the computed ones.

3.1. Backward stability of a single Hari–Zimmermann transformation

The Hari–Zimmermann transformation $\widehat{Z}_\ell$ in (2.4) is determined by the following three parameters: $\sin\varphi_\ell$, $\sin\psi_\ell$, and $b_{ij}^{(\ell)}$. From (2.6), both cosines are positive, i.e., they are uniquely determined by the sines

$$\cos\varphi_\ell = \sqrt{1-\sin^2\varphi_\ell}, \quad \cos\psi_\ell = \sqrt{1-\sin^2\psi_\ell}, \tag{3.4}$$

thus avoiding the first formula in (2.5). In practice, it pays off to compute the cosines by using (3.4), to preserve the essential structure of the transformation in presence of rounding errors, because (3.4) yields a slightly better error bound than (2.5). In exact arithmetic, $B^{(\ell)}$ is always positive definite, so $|b_{ij}^{(\ell)}| < 1$, and from (2.7) it follows that the transformation is nonsingular.

Quite generally, regardless of the pivot strategy, each step of the one-sided algorithm can be viewed as a multiplication of both working matrices from the right-hand side, by a certain Hari–Zimmermann transformation W. If (i, j) is the pivot pair at that step, then W is equal to the identity matrix, except for the pivot submatrix $\widehat{W}$ of order 2 in the (i, j)-plane, where

$$\widehat{W} = \begin{bmatrix} \widehat{w}_{11} & \widehat{w}_{12} \\ \widehat{w}_{21} & \widehat{w}_{22} \end{bmatrix} = \frac{1}{\sqrt{1-b^2}} \begin{bmatrix} \cos\tilde{\varphi} & \sin\tilde{\varphi} \\ -\sin\tilde{\psi} & \cos\tilde{\psi} \end{bmatrix}. \tag{3.5}$$

To simplify the notation, we will omit the step index ℓ, while discussing the accuracy of a single step. The step begins with working matrices F and G. First, it computes the transformation matrix, and then transforms the pivot columns i and j, to obtain the transformed matrices FW and GW. In floating–point arithmetic, each computation involves rounding errors, so let $W' = fl(W)$ be the actually computed transformation matrix, and let $F' = fl(FW')$ and $G' = fl(GW')$ be the computed matrices after the transformation.

Similarly to $\widehat{Z}_\ell$ above, the pivot submatrix $\widehat{W}$ is determined by the following three parameters: $\sin\tilde{\varphi}$, $\sin\tilde{\psi}$, and b. These parameters are actually computed in the earlier part of the step, from the annihilation relations (2.5), but, for the transformation itself, they can be regarded simply as given values, that are exactly representable floating–point numbers, i.e., they belong to $\mathbb{F}$. From (3.5), these three values must satisfy

$$fl(1-b^2) > 0, \quad |\sin\tilde{\varphi}| \le 1, \quad |\sin\tilde{\psi}| \le 1. \tag{3.6}$$

Moreover, to preserve the full column rank of F and G, and for the backward error bound, it is necessary that the exact $\widehat{W}$ is nonsingular, i.e.,

$$\cos\tilde{\varphi}\cdot\cos\tilde{\psi}+\sin\tilde{\varphi}\cdot\sin\tilde{\psi}=\cos(\tilde{\varphi}-\tilde{\psi})\neq 0. \tag{3.7}$$

To compute the elements of $\widehat{W}$, first we compute the cosines, as in (3.4), and then we divide all four elements by $\sqrt{1-b^2}$. In this process, an expression of the form $\sqrt{1-\gamma^2}$ is evaluated for three different arguments γ (all three parameters of $\widehat{W}$). The following proposition gives the relative error bound for the computed results.

Proposition 3.1. *Let $\gamma\in\mathbb{F}$, such that $|\gamma|\leq 1$. If the expression $\sqrt{1-\gamma^2}$ is evaluated in the floating–point arithmetic, the computed value is*

$$f\ell\left(\sqrt{1-\gamma^2}\right)=(1+\varepsilon_\gamma)\sqrt{1-\gamma^2},\quad |\varepsilon_\gamma|\leq 1.5\varepsilon.$$

Proof. By using (3.1)–(3.3) and the standard linearization, i.e., by neglecting the terms of order $O(\varepsilon^2)$, we obtain

$$\begin{aligned} f\ell\left(\sqrt{1-\gamma^2}\right) &= (1+\varepsilon_{\sqrt{\ }})\sqrt{(1+\varepsilon_{\mathrm{fma}})(1-\gamma^2)}\approx(1+\varepsilon_{\sqrt{\ }})\left(1+\frac{\varepsilon_{\mathrm{fma}}}{2}\right)\sqrt{1-\gamma^2}\\ &\approx\left(1+\varepsilon_{\sqrt{\ }}+\frac{\varepsilon_{\mathrm{fma}}}{2}\right)\sqrt{1-\gamma^2}=(1+\varepsilon_\gamma)\sqrt{1-\gamma^2}, \end{aligned}$$

where $|\varepsilon_\gamma|\leq 1.5\varepsilon$. □

The same type of linearized error analysis will be used henceforward. Proposition 3.1 implies that the cosines are computed with the relative error bounded by 1.5ε. Since $f\ell(\sqrt{1-b^2})$ is also computed with the relative error ε_b bounded by 1.5ε, the final division of all four elements by this value increases their respective relative error bounds by 2.5ε. A detailed proof of this fact is as follows. Let $\tilde{\gamma}$ denote the computed value of either the sine or the cosine in (3.5). Then $\tilde{\gamma}=(1+\varepsilon_\gamma)\gamma$, where the relative errors for sines are $\varepsilon_\gamma=0$, and for cosines $|\varepsilon_\gamma|\leq 1.5\varepsilon$. For the division by $f\ell(\sqrt{1-b^2})$, from (3.1) it follows that

$$f\ell\left(\frac{\tilde{\gamma}}{f\ell\,(\sqrt{1-b^2})}\right)=(1+\varepsilon_/)\frac{(1+\varepsilon_\gamma)\gamma}{(1+\varepsilon_b)(\sqrt{1-b^2})}\approx(1+\varepsilon_/+\varepsilon_\gamma-\varepsilon_b)\frac{\gamma}{\sqrt{1-b^2}},$$

where $|\varepsilon_/|\leq\varepsilon$, so $|\varepsilon_/+\varepsilon_\gamma-\varepsilon_b|\leq|\varepsilon_\gamma|+2.5\varepsilon$.

We conclude that the computed pivot submatrix $\widetilde{W}=f\ell(\widehat{W})$ satisfies

$$\widetilde{W}=\begin{bmatrix}\widetilde{w}_{11} & \widetilde{w}_{12}\\ \widetilde{w}_{21} & \widetilde{w}_{22}\end{bmatrix}=\begin{bmatrix}(1+\varepsilon_{bc1})\widehat{w}_{11} & (1+\varepsilon_{bs1})\widehat{w}_{12}\\ (1+\varepsilon_{bs2})\widehat{w}_{21} & (1+\varepsilon_{bc2})\widehat{w}_{22}\end{bmatrix}, \tag{3.8}$$

with componentwise relative errors bounded by

$$|\varepsilon_{bc1}|,\ |\varepsilon_{bc2}|\leq 4\varepsilon,\quad |\varepsilon_{bs1}|,\ |\varepsilon_{bs2}|\leq 2.5\varepsilon. \tag{3.9}$$

Now we analyze the errors introduced by applying the computed transformation W' to the original matrices F and G in floating–point arithmetic. The computed matrices $F' = f\ell(FW')$ and $G' = f\ell(GW')$ can be written as perturbed results of the exact transformation by the original matrix W

$$F' = FW + \delta F', \quad G' = GW + \delta G', \tag{3.10}$$

where $\delta F'$ and $\delta G'$ denote the forward perturbations caused by rounding errors throughout the computation, starting from W, F and G. Since both W and W' differ from the identity matrix only in the (i, j)-plane, the transformation in (3.10) changes only the pivot columns i, j in F and G. Consequently, only these two columns are possibly nonzero in both perturbation matrices $\delta F'$ and $\delta G'$.

The analysis for F and G is exactly the same, so we will do it only for F, and G will be mentioned only when necessary. The transformation of pivot columns is determined by the pivot submatrices of W and W'

$$[f'_i, f'_j] = f\ell([f_i, f_j] \cdot \widetilde{W}) = [f_i, f_j] \cdot \widehat{W} + [\delta f'_i, \delta f'_j]. \tag{3.11}$$

The computation of each element is organized in such a way to minimize the error bound. First, the element of $\widetilde{W}$ with the smaller error bound (originating from the sine) is involved in an extra multiplication, and the final value is computed by a single FMA operation, as follows

$$\begin{aligned} f'_{pi} &= f\ell(\widetilde{w}_{11} f_{pi} + \widetilde{w}_{21} f_{pj}) = f\ell\left(\widetilde{w}_{11} f_{pi} + f\ell(\widetilde{w}_{21} f_{pj})\right), \\ f'_{pj} &= f\ell(\widetilde{w}_{12} f_{pi} + \widetilde{w}_{22} f_{pj}) = f\ell\left(f\ell(\widetilde{w}_{12} f_{pi}) + \widetilde{w}_{22} f_{pj}\right). \end{aligned}$$

In terms of the original values in F and $\widehat{W}$, and rounding errors caused by these operations, from (3.1), (3.3) and (3.8), we have

$$\begin{aligned} f'_{pi} &= (1 + \varepsilon_{\text{fma1}})[(1 + \varepsilon_{bc1})\widehat{w}_{11} f_{pi} + (1 + \varepsilon_1)(1 + \varepsilon_{bs2})\widehat{w}_{21} f_{pj}], \\ f'_{pj} &= (1 + \varepsilon_{\text{fma2}})[(1 + \varepsilon_2)(1 + \varepsilon_{bs1})\widehat{w}_{12} f_{pi} + (1 + \varepsilon_{bc2})\widehat{w}_{22} f_{pj}]. \end{aligned} \tag{3.12}$$

From (3.11) and (3.12), it follows that the corresponding perturbations $\delta f'_{pi}$ and $\delta f'_{pj}$ are

$$\begin{aligned} \delta f'_{pi} &\approx (\varepsilon_{\text{fma1}} + \varepsilon_{bc1})\widehat{w}_{11} f_{pi} + (\varepsilon_{\text{fma1}} + \varepsilon_1 + \varepsilon_{bs2})\widehat{w}_{21} f_{pj}, \\ \delta f'_{pj} &\approx (\varepsilon_{\text{fma2}} + \varepsilon_2 + \varepsilon_{bs1})\widehat{w}_{12} f_{pi} + (\varepsilon_{\text{fma2}} + \varepsilon_{bc2})\widehat{w}_{22} f_{pj}. \end{aligned}$$

By using (3.9) and $|\varepsilon_{\text{fma1}}|, |\varepsilon_{\text{fma2}}| \leq \varepsilon$, we get the following pointwise absolute error bounds

$$\begin{aligned} |\delta f'_{pi}| &\leq \varepsilon(5\,|\widehat{w}_{11}| \cdot |f_{pi}| + 4.5\,|\widehat{w}_{21}| \cdot |f_{pj}|), \\ |\delta f'_{pj}| &\leq \varepsilon(4.5\,|\widehat{w}_{12}| \cdot |f_{pi}| + 5\,|\widehat{w}_{22}| \cdot |f_{pj}|). \end{aligned}$$

Finally, it is easy to see that the forward perturbations of pivot columns satisfy the following normwise bounds

$$\begin{aligned}\|\delta f_i'\|_2 &\le \varepsilon(5\,|\widehat{w}_{11}|\cdot\|f_i\|_2 + 4.5\,|\widehat{w}_{21}|\cdot\|f_j\|_2),\\ \|\delta f_j'\|_2 &\le \varepsilon(4.5\,|\widehat{w}_{12}|\cdot\|f_i\|_2 + 5\,|\widehat{w}_{22}|\cdot\|f_j\|_2).\end{aligned}$$

By using the second equality in (3.5), and the fact that both cosines are nonnegative, we get a more convenient form of these bounds

$$\begin{aligned}\|\delta f_i'\|_2 &\le \frac{\varepsilon}{\sqrt{1-b^2}}(5\cos\tilde{\varphi}\cdot\|f_i\|_2 + 4.5\,|\sin\tilde{\psi}|\cdot\|f_j\|_2),\\ \|\delta f_j'\|_2 &\le \frac{\varepsilon}{\sqrt{1-b^2}}(4.5\,|\sin\tilde{\varphi}|\cdot\|f_i\|_2 + 5\cos\tilde{\psi}\cdot\|f_j\|_2).\end{aligned} \tag{3.13}$$

The same relations hold for the perturbations $\delta g_i'$, $\delta g_j'$ of pivot columns in G', with the addition that the initial pivot columns g_i, g_j have unit norms.

On the other hand, in backward terms, the computed matrices in (3.10) can be viewed as exact results of the transformation W, but applied to perturbed original matrices F and G

$$F' = (F+\delta F)W, \quad G' = (G+\delta G)W, \tag{3.14}$$

where δF and δG now denote the backward perturbations, provided that such matrices exist. Again, only the pivot columns are possibly nonzero in δF and δG. For the pivot columns of F, (3.14) reduces to the exact transformation by the pivot submatrix $\widehat{W}$

$$[f_i', f_j'] = ([f_i, f_j] + [\delta f_i, \delta f_j])\widehat{W}. \tag{3.15}$$

From (3.11) and (3.15) we see that the forward and backward perturbations are related by $[\delta f_i', \delta f_j'] = [\delta f_i, \delta f_j]\widehat{W}$. By assumptions (3.6) and (3.7), the exact matrix $\widehat{W}$ is nonsingular. Hence, (3.15) can be written as

$$[\delta f_i, \delta f_j] = [\delta f_i', \delta f_j']\widehat{W}^{-1}, \tag{3.16}$$

which proves the existence of backward perturbations in (3.14).

From (3.5) it follows that

$$\widehat{W}^{-1} = \frac{\sqrt{1-b^2}}{\cos\tilde{\varphi}\cos\tilde{\psi} + \sin\tilde{\varphi}\sin\tilde{\psi}}\begin{bmatrix}\cos\tilde{\psi} & -\sin\tilde{\varphi}\\ \sin\tilde{\psi} & \cos\tilde{\varphi}\end{bmatrix} = \frac{\sqrt{1-b^2}}{\cos(\tilde{\varphi}-\tilde{\psi})}\begin{bmatrix}\cos\tilde{\psi} & -\sin\tilde{\varphi}\\ \sin\tilde{\psi} & \cos\tilde{\varphi}\end{bmatrix},$$

and (3.16) now becomes

$$[\delta f_i, \delta f_j] = \frac{\sqrt{1-b^2}}{\cos(\tilde{\varphi}-\tilde{\psi})}[\cos\tilde{\psi}\cdot\delta f_i' + \sin\tilde{\psi}\cdot\delta f_j',\ -\sin\tilde{\varphi}\cdot\delta f_i' + \cos\tilde{\varphi}\cdot\delta f_j'].$$

For brevity, let $c_{ij} = 1/|\cos(\tilde{\varphi} - \tilde{\psi})|$. The Euclidean norm of backward perturbations is then bounded in terms of forward perturbations

$$\|\delta f_i\|_2 \leq c_{ij}\sqrt{1-b^2}\,(\cos\tilde{\psi}\cdot\|\delta f_i'\|_2 + |\sin\tilde{\psi}|\cdot\|\delta f_j'\|_2),$$
$$\|\delta f_j\|_2 \leq c_{ij}\sqrt{1-b^2}\,(|\sin\tilde{\varphi}|\cdot\|\delta f_i'\|_2 + \cos\tilde{\varphi}\cdot\|\delta f_j'\|_2).$$

First, we substitute the bounds from (3.13) into these inequalities. Since both cosines are nonnegative, we can take $|\tilde{\varphi}|, |\tilde{\psi}| \leq \pi/2$. By using the standard trigonometric identities (like $|\sin\gamma| = \sin|\gamma|$, for $|\gamma| \leq \pi$), we obtain the normwise bounds for the backward perturbations in terms of the original pivot columns

$$\begin{aligned}
\|\delta f_i\|_2 &\leq \varepsilon c_{ij}[(5\cos\tilde{\varphi}\cdot\cos\tilde{\psi} + 4.5\,|\sin\tilde{\varphi}|\cdot|\sin\tilde{\psi}|)\cdot\|f_i\|_2 + 9.5\cos\tilde{\psi}\cdot|\sin\tilde{\psi}|\cdot\|f_j\|_2] \\
&\leq \varepsilon c_{ij}[5\cos(|\tilde{\varphi}| - |\tilde{\psi}|)\cdot\|f_i\|_2 + 4.25\sin(2|\tilde{\psi}|)\cdot\|f_j\|_2] \\
&\leq \varepsilon c_{ij}(5\|f_i\|_2 + 4.25\|f_j\|_2), \\
\|\delta f_j\|_2 &\leq \varepsilon c_{ij}[9.5\cos\tilde{\varphi}\cdot|\sin\tilde{\varphi}|\cdot\|f_i\|_2 + (5\cos\tilde{\varphi}\cdot\cos\tilde{\psi} + 4.5\,|\sin\tilde{\varphi}|\cdot|\sin\tilde{\psi}|)\cdot\|f_j\|_2] \\
&\leq \varepsilon c_{ij}[4.25\sin(2|\tilde{\varphi}|)\cdot\|f_i\|_2 + 5\cos(|\tilde{\varphi}| - |\tilde{\psi}|)\cdot\|f_j\|_2] \\
&\leq \varepsilon c_{ij}(4.25\|f_i\|_2 + 5\|f_j\|_2).
\end{aligned} \tag{3.17}$$

The same bounds hold for $\|\delta g_i\|_2$ and $\|\delta g_j\|_2$, with $\|g_i\|_2 = \|g_j\|_2 = 1$, i.e.,

$$\|\delta g_i\|_2 \leq 9.25\varepsilon c_{ij} = 9.25\varepsilon c_{ij}\|g_i\|_2, \qquad \|\delta g_j\|_2 \leq 9.25\varepsilon c_{ij} = 9.25\varepsilon c_{ij}\|g_j\|_2. \tag{3.18}$$

Theorem 3.2. *Let F and G be of full column rank k, and let all columns of G be of unit Euclidean norm. Let W be a nonsingular Hari–Zimmermann transformation as in (3.5). Then each step of the one-sided Hari–Zimmermann algorithm is backward stable.*

Proof. We have shown already that there exist perturbation matrices δF and δG such that (3.14) holds, and the only possible nonzero columns in these matrices are bounded in norm by (3.17) and (3.18), with respect to the original pivot columns in F and G. □

3.2. Backward stability with annihilation parameters

Note that Theorem 3.2 holds regardless of how the initial parameters $\sin\tilde{\varphi}$, $\sin\tilde{\psi}$, and b of $\widehat{W}$ are given, as long as they are exactly representable and satisfy the assumptions (3.6) and (3.7). The actual values of these parameters in step ℓ of the Hari–Zimmermann algorithm are computed from (2.5), i.e.,

$$b = f\ell(b_{ij}^{(\ell)}), \quad \sin\tilde{\varphi} = f\ell(\sin\varphi_\ell), \quad \sin\tilde{\psi} = f\ell(\sin\psi_\ell), \tag{3.19}$$

with the aim to annihilate $a_{ij}^{(\ell+1)}$ and $b_{ij}^{(\ell+1)}$ after the transformation. In one-sided algorithms, these aims are equivalent to the orthogonalization of pivot columns in the newly computed matrices $F^{(\ell+1)}$ and $G^{(\ell+1)}$. One of the main advantages of the one-sided algorithm is that it does *not* involve the explicit annihilation, meaning that none of the

computed elements is actually set to zero. Instead, the new pivot columns are computed, and there is no big harm if they are not *perfectly* orthogonal. Therefore, the parameters of the "annihilator" transformation $\widehat{W} = \widehat{Z}_\ell$ do not have to be computed with the ultimate precision—the pivot columns can be "fully" orthogonalized in the later steps. In that respect, the one-sided algorithm is self-correcting, in contrast to the two-sided algorithm.

From (2.2), $b = f\ell(b_{ij}^{(\ell)})$ is computed as the inner product of the pivot columns $g_i^{(\ell)}$, $g_j^{(\ell)}$ of $G^{(\ell)}$, which have unit norms. The floating–point error bounds for this computation can be found in [4, Section 3.1.]. If b violates the first assumption in (3.6), i.e., $|b| \gtrapprox 1-\varepsilon$, then the pivot submatrix $\widehat{B}^{(\ell)} = [g_i^{(\ell)}, g_j^{(\ell)}]^T[g_i^{(\ell)}, g_j^{(\ell)}]$ in (2.3) is not positive definite, so the pivot columns are linearly dependent to the working precision. This can happen only when $G^{(\ell)}$ is severely ill-conditioned. Since the pivot submatrix is given by

$$\widehat{B}^{(\ell)} = \begin{bmatrix} 1 & b_{ij}^{(\ell)} \\ b_{ij}^{(\ell)} & 1 \end{bmatrix},$$

the Cauchy interlace theorem implies that

$$\sigma_{\min}^2(G^{(\ell)}) = \lambda_{\min}(B^{(\ell)}) \leq \lambda_{\min}(\widehat{B}^{(\ell)}) = 1 - |b_{ij}^{(\ell)}|,$$
$$\sigma_{\max}^2(G^{(\ell)}) = \lambda_{\max}(B^{(\ell)}) \geq \lambda_{\max}(\widehat{B}^{(\ell)}) = 1 + |b_{ij}^{(\ell)}|.$$

In exact arithmetic, if $|b_{ij}^{(\ell)}| \geq 1 - \varepsilon$, then the condition of $G^{(\ell)}$ is at least

$$\kappa(G^{(\ell)}) = \frac{\sigma_{\max}(G^{(\ell)})}{\sigma_{\min}(G^{(\ell)})} \geq \sqrt{\frac{2-\varepsilon}{\varepsilon}}.$$

For a moderately conditioned starting matrix G, it is reasonable to assume that *all* computed values of $b = f\ell(b_{ij}^{(\ell)})$ throughout the algorithm are uniformly bounded away from 1, so that $f\ell(1 - b^2) > 0$ in (3.6). Then there exists a (moderate) real constant c_b such that

$$\frac{1}{\sqrt{1-b^2}} \leq c_b. \tag{3.20}$$

Note that global convergence of the algorithm (see [3]) guarantees that $B^{(\ell)} \to I$, i.e., $b_{ij}^{(\ell)} \to 0$, so the highest value of $1/\sqrt{1-b^2}$ occurs in the first few sweeps of the algorithm.

The remaining two parameters of $\widehat{W}$ in (3.19), namely, $\sin\tilde{\varphi} = f\ell(\sin\varphi_\ell)$ and $\sin\tilde{\psi} = f\ell(\sin\psi_\ell)$, are computed from (2.5). A tedious analysis of rounding errors (skipped here for brevity) shows that both computed values have small absolute errors—within a small multiple of ε, compared to the exact values. The same also holds for $\cos\tilde{\varphi}$ and $\cos\tilde{\psi}$, if they are computed from (2.5), instead of (3.4).

When any one of the exact sines is very close to ± 1, in an extremely rare case of unfavorable rounding, we can get $|\sin\tilde{\varphi}| > 1$ or $|\sin\tilde{\psi}| > 1$, thus violating the assumptions (3.6). If this happens, we set $\sin\tilde{\varphi} = \pm 1$ or $\sin\tilde{\psi} = \pm 1$, and the signs are chosen to match the signs of the initially computed values. By doing so, we actually reduce the magnitude of the error (both absolute and relative) in the computed parameters of $\widehat{W}$. It should be said that such a case never occurred during the course of numerical testing.

Finally, we discuss the assumption (3.7) and possible sizes of the factors c_{ij} in (3.17) and (3.18). From (2.8), it follows that the exact annihilation parameters must satisfy

$$\cos(\varphi_\ell - \psi_\ell) = \cos\varphi_\ell \cos\psi_\ell + \sin\varphi_\ell \sin\psi_\ell = \sqrt{1 - (b_{ij}^{(\ell)})^2}. \tag{3.21}$$

Whenever the computed parameter $b = f\ell(b_{ij}^{(\ell)})$ satisfies $f\ell(1 - b^2) > 0$, the earlier analysis ensures that $\sqrt{1 - b^2}$ is computed with (at most) a small absolute error, even if it is close to zero. The same holds for the computed sines and cosines in (3.5). Therefore, (3.21) is also valid for the computed values, with (at most) a small absolute error

$$\cos(\tilde{\varphi} - \tilde{\psi}) = \cos\tilde{\varphi}\cos\tilde{\psi} + \sin\tilde{\varphi}\sin\tilde{\psi} \approx \sqrt{1 - b^2}.$$

For annihilating transformations $\widehat{W} = \widehat{Z}_\ell$, the scale factor c_{ij} in (3.17) and (3.18) satisfies

$$c_{ij} = \frac{1}{|\cos(\tilde{\varphi} - \tilde{\psi})|} \approx \frac{1}{\sqrt{1 - b^2}},$$

and from (3.20), it follows that all c_{ij} are (approximately) bounded from above by c_b. To get an exact bound, the value of c_b in (3.20) may have to be increased slightly. From now on, we assume that c_b has been increased already, if necessary, so that $c_{ij} \le c_b$ holds exactly, for all steps ℓ.

4. Accuracy of the implicit Hari–Zimmermann algorithm

4.1. Backward bound for a sequence of independent transformations

Backward bounds (3.17) and (3.18) are valid for a single step of the one-sided (implicit) Hari–Zimmermann algorithm with the pivot pair (i, j), irrespective of the chosen pivoting strategy in the whole algorithm. The important thing is that both perturbation matrices have at most two nonzero columns—these are exactly the pivot columns.

In the so-called parallel pivot strategies (e.g., the modified modulus strategy), we can independently orthogonalize $k/2$ pairs of pivot columns in both working matrices. Such a block of $k/2$ transformations, which transforms *all* columns in both working matrices, will be called a *stage* of the algorithm. For the discussion below, it does not matter if the stage is implemented in parallel, or performed sequentially.

To simplify the notation, let F and G denote the starting matrices at the beginning of a stage, i.e., we can take $F = F^{(\ell)}$ and $G = G^{(\ell)}$ at the end of any step ℓ of the

algorithm. The algorithm then performs a stage of $k/2$ independent transformations, and let F' and G' now denote the computed matrices at the end of the stage, so we can take $F' = F^{(\ell+k/2)}$ and $G' = G^{(\ell+k/2)}$. Then, by a repeated application of Theorem 3.2, there exist perturbation matrices δF and δG, such that

$$F' = (F + \delta F)W_s, \quad G' = (G + \delta G)W_s, \tag{4.1}$$

where W_s is the product of all transformations applied in the stage. To get nice bounds for these perturbations, we need the following lemma.

Lemma 4.1. *Let $\alpha, \beta, \nu_i, \nu_j \in \mathbb{R}$, such that $\alpha\beta \geq 0$. Then*

$$(\alpha\nu_i + \beta\nu_j)^2 + (\beta\nu_i + \alpha\nu_j)^2 \leq (\alpha + \beta)^2(\nu_i^2 + \nu_j^2).$$

Proof. By straightforward manipulation we obtain

$$\begin{aligned}(\alpha\nu_i + \beta\nu_j)^2 + (\beta\nu_i + \alpha\nu_j)^2 &= \alpha^2\nu_i^2 + 2\alpha\beta\nu_i\nu_j + \beta^2\nu_j^2 + \beta^2\nu_i^2 + 2\alpha\beta\nu_i\nu_j + \alpha^2\nu_j^2 \\ &= (\alpha + \beta)^2(\nu_i^2 + \nu_j^2) - 2\alpha\beta(\nu_i - \nu_j)^2 \\ &\leq (\alpha + \beta)^2(\nu_i^2 + \nu_j^2),\end{aligned}$$

where the inequality follows from $\alpha\beta \geq 0$. □

For any pivot pair (i, j) in a stage, consider the sum of squares of the two relations in (3.17). Let $\alpha = 5\varepsilon c_{ij}, \beta = 4.25\varepsilon c_{ij}$, $\nu_i = \|f_i\|_2$, and $\nu_j = \|f_j\|_2$. Then, Lemma 4.1 gives

$$\|\delta f_i\|_2^2 + \|\delta f_j\|_2^2 \leq (9.25\varepsilon c_{ij})^2(\|f_i\|_2^2 + \|f_j\|_2^2). \tag{4.2}$$

The same relation for G follows from (3.18), with $\|g_i\|_2^2 + \|g_j\|_2^2 = 2$.

Theorem 4.2. *After a stage of $k/2$ independent Hari–Zimmermann transformations, the backward perturbations δF and δG in (4.1) are bounded by*

$$\|\delta F\|_F \leq 9.25\varepsilon c_s\|F\|_F, \quad \|\delta G\|_F \leq 9.25\varepsilon c_s\|G\|_F = 9.25\varepsilon c_s\sqrt{k},$$

where $\|\ \|_F$ denotes the Frobenius norm of a matrix, and $c_s := \max c_{ij}$ over all pairs of pivot indices (i, j) at this stage of the algorithm.

Proof. Since the $k/2$ independent pivot pairs (i, j) in a stage cover all columns of both matrices, by summing (4.2) over these pivot pairs, and using $c_{ij} \leq c_s$, we get

$$\sum_{p=1}^{k} \|\delta f_p\|_2^2 \leq (9.25\varepsilon c_s)^2 \sum_{p=1}^{k} \|f_p\|_2^2.$$

For any matrix S with columns s_p, $p = 1, \ldots, k$, the Frobenius norm of S can be written in terms of the Euclidean norm of its columns as

$$\|S\|_F = \sqrt{\sum_{p=1}^{k} \|s_p\|_2^2}.$$

Hence, we obtain the following bound

$$\|\delta F\|_F \leq 9.25\varepsilon c_s \|F\|_F.$$

The bound for δG follows from $\|G\|_F = \sqrt{k}$. □

In terms of the spectral norm, by using the norm equivalence inequalities $\|S\|_2 \leq \|S\|_F \leq \sqrt{k}\|S\|_2$, the results of Theorem 4.2 can be expressed in a slightly weaker form

$$\|\delta F\|_2 \leq 9.25\varepsilon c_s \sqrt{k}\,\|F\|_2, \quad \|\delta G\|_2 \leq 9.25\varepsilon c_s \sqrt{k}.$$

4.2. Accuracy of the one-sided Hari–Zimmermann algorithm

A typical proof of the accuracy of a Jacobi-type algorithm (for example, of the Jacobi SVD algorithm, see [1]) uses the fact that the absolute value of the tangent of the rotation angle is less or equal to 1. This is required to establish a connection between the norms of the pivot columns before and after the transformation, so that the perturbation of each pivot column can be expressed in terms of the same column (instead of both columns). With such bounds it is easy to use some of the standard theorems about the relative accuracy of singular values.

Unfortunately, in the Hari–Zimmermann algorithm we have two "rotation" angles $\tilde{\varphi}$ and $\tilde{\psi}$, and from (2.7), there is no guarantee that the tangents of both angles are bounded. To get the required form of bounds, we simply use the actual ratio of norms of pivot columns.

For simplicity, we assume that the indices in each pivot pair (i, j) are ordered so that $\|f_i\|_2 \geq \|f_j\|_2$. Since F is of full column rank, then $\|f_j\|_2 = r_{ij}\|f_i\|_2$, with $0 < r_{ij} \leq 1$, and (3.17) can be written as

$$\|\delta f_i\|_2 \leq \varepsilon c_{ij}(5 + 4.25 r_{ij})\|f_i\|_2, \qquad \|\delta f_j\|_2 \leq \varepsilon c_{ij}\left(\frac{4.25}{r_{ij}} + 5\right)\|f_j\|_2. \tag{4.3}$$

Generally speaking, these bounds are quite pessimistic. A closer look at (3.17) reveals that the final bounds in (3.17), and, consequently, the bounds in (4.3) are obtained by taking

$$\sin(2|\tilde{\psi}|) \leq 1, \quad \sin(2|\tilde{\varphi}|) \leq 1.$$

In later stages of the algorithm, both angles tend to zero, which damps the effect of the other pivot column in (3.17), and the terms containing r_{ij} in (4.3).

The relative accuracy of the implicit Hari–Zimmermann method is based on the following perturbation result by Drmač [2, Corollary 2.8].

Theorem 4.3 (Drmač). *Let F and G be of full column rank k, and let the columns of perturbation matrices δF and δG satisfy the following bounds*

$$\|\delta f_p\|_2 \leq \epsilon \|f_p\|_2, \qquad \|\delta g_p\|_2 \leq \epsilon \|g_p\|_2, \qquad p = 1, \ldots, k, \tag{4.4}$$

for some constant ϵ, such that $0 \leq \epsilon < 1$. Then, the relative errors in the perturbed generalized singular values $\tilde{\sigma}_p$ of the pair $(F + \delta F, G + \delta G)$ are bounded by

$$\frac{|\tilde{\sigma}_p - \sigma_p|}{\sigma_p} \leq \left(1 + \frac{\sigma_{\min}(G_S)}{\sigma_{\min}(F_S)}\right) \frac{\epsilon \sqrt{q}}{\sigma_{\min}(G_S) - \epsilon \sqrt{q}}, \qquad p = 1, \ldots, k, \tag{4.5}$$

where $F_S = F \operatorname{diag}(\|f_p\|_2^{-1})$, $G_S = G \operatorname{diag}(\|g_p\|_2^{-1})$, and q is the maximal number of nonzero elements in any row of δF and δG.

Note that the column norms of scaled matrices F_S and G_S are all equal to 1. Since the original G in the Hari–Zimmermann algorithm is already scaled in such a way, we have $G_S = G$.

The result of Theorem 4.3 can be applied for one transformation with the pivot pair (i, j), and for a whole stage of $k/2$ independent transformations. To apply it for the whole process, we need to know the total number of transformations, or the total number of stages needed for "full" orthogonalization of columns in F and G to the working precision.

For one transformation W, we can take $F = F^{(\ell)}$ and $G = G^{(\ell)}$ as the starting matrices. Since $r_{ij} \leq 1$, from (4.3) and (3.18), it follows that the required bounds (4.4) hold with

$$\epsilon := \varepsilon c_{ij} \left(\frac{4.25}{r_{ij}} + 5\right),$$

as all other non-pivot columns in δF and δG are equal to zero. From (3.14), it follows that the generalized singular values of the computed pair $(F' = F^{(\ell+1)}, G' = G^{(\ell+1)})$ after the transformation W, are equal to those of the perturbed initial pair $(F + \delta F, G + \delta G)$. If $\epsilon < 1$, then (4.5) with $q = 2$ gives the relative error bound for the perturbation of the generalized singular values induced by that transformation (in both the forward and the backward sense).

For a stage of $k/2$ independent transformations, we proceed in exactly the same manner. Let $F = F^{(\ell)}$ and $G = G^{(\ell)}$ be the starting matrices at the beginning of the stage. From (4.1), it follows that the generalized singular values of the computed pair $(F' = F^{(\ell+k/2)}, G' = G^{(\ell+k/2)})$ after the stage, are equal to those of the perturbed initial pair $(F + \delta F, G + \delta G)$.

Similarly as in Theorem 4.2, let $r_s := \min r_{ij}$ over all pairs of pivot indices (i, j) at this stage of the algorithm. From (4.3) and (3.18), it follows that the bounds (4.4) hold with

$$\epsilon := \varepsilon c_s \left(\frac{4.25}{r_s} + 5\right).$$

In general, all the columns in δF and δG are nonzero. If $\epsilon < 1$, then (4.5) with $q = k$ gives the relative error bound for the perturbation of the generalized singular values caused by the product W_s of all transformations in the stage.

If in each transformation (stage), r_{ij} (r_s) is reasonably bounded from below, and c_{ij} (c_s) is reasonably bounded from above, the one-sided Hari–Zimmermann algorithm computes the generalized singular values with high relative accuracy, provided that the initial scaled matrices F_S and G_S are well-conditioned.

In particular, if the columns of G are nearly orthonormal, i.e., if the matrix $G^T G$ is near to the identity, then the algorithm inherits a good behavior from the ordinary one-sided Jacobi SVD algorithm—all values c_{ij} are close to 1, which, from the start, damps the effect of possibly small ratios r_{ij}.

From Theorem 4.2 it is easy to bound the perturbations for all stages of the one-sided Hari–Zimmermann algorithm. Note that Theorem 4.2 is valid for each stage of the orthogonalization process. In stage i, $i \geq 0$, formula (4.5) has the following form

$$\frac{|\tilde{\sigma}_p^{(i)} - \tilde{\sigma}_p^{(i-1)}|}{\tilde{\sigma}_p^{(i-1)}} \leq \left(1 + \frac{\sigma_{\min}(G_S^{(i-1)})}{\sigma_{\min}(F_S^{(i-1)})}\right) \frac{\epsilon \sqrt{q^{(i-1)}}}{\sigma_{\min}(G_S^{(i-1)}) - \epsilon \sqrt{q^{(i-1)}}} := C_i, \tag{4.6}$$

where $p = 1, \ldots, k$, $F_S^{(i)} = F^{(i)} \operatorname{diag}(\|f_p^{(i)}\|_2^{-1})$, $G_S^{(i)} = G^{(i)} \operatorname{diag}(\|g_p^{(i)}\|_2^{-1})$, and $q^{(i)}$ is the maximal number of nonzero elements in any row of $\delta F^{(i)}$ and $\delta G^{(i)}$. In addition, we define $\sigma_p = \tilde{\sigma}_p^{(0)}$.

Theorem 4.4. *Let F and G be of full column rank k, and let the columns of perturbation matrices $\delta F^{(i)}$ and $\delta G^{(i)}$ in each stage of the algorithm satisfy the following bounds*

$$\|\delta f_p^{(i)}\|_2 \leq \epsilon \|f_p^{(i)}\|_2, \qquad \|\delta g_p^{(i)}\|_2 \leq \epsilon \|g_p^{(i)}\|_2,$$

for $p = 1, \ldots, k$, and some constant ϵ, such that $0 \leq \epsilon < 1$. Then, the relative errors in the perturbed generalized singular values $\tilde{\sigma}_p := \sigma_p^{(N)}$ of the pair $(F + \delta F, G + \delta G)$ after the N stages of the algorithm are bounded

$$\frac{|\tilde{\sigma}_p - \sigma_p|}{\sigma_p} \leq C_1 + C_2(1 + C_1) + \cdots + C_N(1 + C_1) \cdots (1 + C_N), \tag{4.7}$$

for all generalized singular values σ_p, $p = 1, \ldots, k$.

Proof. From (4.6) it follows

$$(1 - C_i)\tilde{\sigma}_p^{(i-1)} \leq \tilde{\sigma}_p^{(i)} \leq (1 + C_i)\tilde{\sigma}_p^{(i-1)} \qquad i = 1, \ldots, N.$$

By repetition of the same argument, we obtain

$$\tilde{\sigma}_p^{(i)} \leq (1 + C_i)\tilde{\sigma}_p^{(i-1)} \leq (1 + C_i)(1 + C_{i-1})\tilde{\sigma}_p^{(i-2)} \leq \cdots \leq (1 + C_i) \cdots (1 + C_1)\sigma_p. \tag{4.8}$$

Directly form Theorem 4.3 it follows

$$\frac{|\tilde{\sigma}_p - \sigma_p|}{\sigma_p} \leq \frac{|\tilde{\sigma}_p^{(N)} - \tilde{\sigma}_p^{(N-1)}|}{\sigma_p} + \cdots + \frac{|\tilde{\sigma}_p^{(1)} - \tilde{\sigma}_p^{(0)}|}{\sigma_p} \leq C_N \frac{\tilde{\sigma}_p^{(N)}}{\sigma_p} + \cdots + C_2 \frac{\tilde{\sigma}_p^{(2)}}{\sigma_p} + C_1.$$

By substitution of $\tilde{\sigma}_p^{(i)}$ from (4.8) for $i = 1, \ldots, N$ into the previous equation, we immediately obtain (4.7). □

5. Conclusion

In this paper we proved that the implicit Hari–Zimmermann method for computation of the generalized singular values of matrix pairs, is backward stable, and, if the matrices permit, computes the generalized singular values with small relative errors.

Acknowledgment

The authors are indebted to Vedran Novaković, for his valuable help with preparing the final draft of the manuscript.

References

[1] J. W. Demmel, K. Veselić, Jacobi's method is more accurate than QR, SIAM J. Matrix Anal. Appl. 13 (4) (1992) 1204–1245.

[2] Z. Drmač, Computing the singular and the generalized singular values, Ph.D. thesis, FernUniversität–Gesamthochschule, Hagen (1994).

[3] V. Hari, On cyclic Jacobi methods for the positive definite generalized eigenvalue problem, Ph.D. thesis, FernUniversität–Gesamthochschule, Hagen (1984).

[4] N. J. Higham, Accuracy and Stability of Numerical Algorithms, 2nd ed., SIAM, Philadelphia, 2002.

[5] IEEE 754-2008, Standard for Floating-Point Arithmetic, New York, NY, USA (Aug. 2008).

[6] V. Novaković, S. Singer, S. Singer, Blocking and parallelization of the Hari–Zimmermann variant of the Falk–Langemeyer algorithm for the generalized SVD, Parallel Comput. 49 (2015) 136–152.