Nečas Center for Mathematical Modeling

# *On the conditioning of factors in the SR decomposition*

*H. Fassbender, M. Rozložník*

Preprint no. 2016-04

http://ncmm.karlin.mff.cuni.cz/

# On the conditioning of factors in the SR decomposition

H. Faßbender [§]   M. Rozložník [†¶]

February 16, 2016

**Abstract**

Almost every nonsingular matrix $A \in \mathbb{R}^{2m,2m}$ can be decomposed into the product of a symplectic matrix $S$ and an upper $J$-triangular matrix $R$. This decomposition is not unique. In this paper we analyze the freedom of choice in the symplectic and the upper $J$-triangular factors and review several existing suggestions on how to choose the free parameters in the SR decomposition. In particular we consider two choices leading to the minimization of the condition number of the diagonal blocks in the upper $J$-triangular factor and to the minimization of the conditioning of the corresponding blocks in the symplectic factor. We develop bounds for the extremal singular values of the whole upper $J$-triangular factor and the whole symplectic factor in terms of the spectral properties of even-dimensioned principal submatrices of the skew-symmetric matrix associated with the SR decomposition. The theoretical results are illustrated on two small examples.

## 1 Introduction

For each natural number $m$ we define the skew-symmetric matrix

$$J_{2m} = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix} \in \mathbb{R}^{2m,2m},$$

where $I_m \in \mathbb{R}^{m,m}$ denotes the identity matrix of the order $m$. It is clear that $J_{2m}$ is nonsingular with $J_{2m}^{-1} = J_{2m}^T = -J_{2m}$. If $I_{2m} = [e_1, \ldots e_{2m}]$ is the identity matrix of order $2m$ we define a permutation matrix $P_{2m} \in \mathbb{R}^{2m,2m}$ as

$$P_{2m} = [e_1, e_3, \ldots, e_{2m-1}, e_2, e_4, \ldots, e_{2m}].$$

[§]Institut Computational Mathematics, AG Numerik, Technische Universität Braunschweig, Pockelsstr. 14, 38106 Braunschweig, Germany. E-mail: h.fassbender@tu-braunschweig.de

[†]Institute of Computer Science, Czech Academy of Sciences, Pod vodárenskou věží 2, CZ-182 07 Prague 8, Czech Republic. E-mail: miro@cs.cas.cz

[¶]This research is supported by the Grant Agency of the Czech Republic under the project 13-06684S.

It follows that $P_{2m}^{-1} = P_{2m}^T = [e_1, e_{m+1}, e_2, e_{m+2}, \ldots, e_m, e_{2m}]$. Using the permutation matrix $P_{2m}$, the matrix $J_{2m}$ can be permuted to the block diagonal matrix $\hat{J}_{2m} \in \mathbb{R}^{2m,2m}$ such that

$$\hat{J}_{2m} = P_{2m} J_{2m} P_{2m}^T = \mathrm{diag}(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}).$$

**Definition 1.** *A real square matrix $S_{2m} \in \mathbb{R}^{2m,2m}$ is a symplectic matrix if $S_{2m}^T J_{2m} S_{2m} = J_{2m}$. Similarly, a real rectangular matrix $S_{2m,2n} \in \mathbb{R}^{2m,2n}$ is called semi-symplectic if $S_{2m,2n}^T J_{2m} S_{2m,2n} = J_{2n}$.*

**Definition 2.** *A matrix $R_{2m} = \begin{pmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{pmatrix} \in \mathbb{R}^{2m,2m}$ is an upper $J$-triangular matrix if $R_{1,1}, R_{1,2}$, and $R_{2,2} \in \mathbb{R}^{m,m}$ are upper triangular matrices and $R_{2,1} \in \mathbb{R}^{m,m}$ is strictly upper triangular matrix.*

Note that if $R_{2m}$ is an upper $J$-triangular matrix, then the matrix $\hat{R}_{2m} = P_{2m} R_{2m} P_{2m}^T \in \mathbb{R}^{2m,2m}$ is an upper triangular matrix of the form

$$\hat{R}_{2m} = \begin{pmatrix} \hat{\mathrm{R}}_{1,1} & \ldots & \hat{\mathrm{R}}_{1,m} \\ 0 & \ddots & \vdots \\ 0 & 0 & \hat{\mathrm{R}}_{m,m} \end{pmatrix},$$

where $\hat{\mathrm{R}}_{i,j} \in \mathbb{R}^{2,2}$ for $i = 1, \ldots j$; $j = 1, \ldots, m$ and $\hat{\mathrm{R}}_{j,j}$ is upper triangular for $j = 1, \ldots, m$.

The question whether a square matrix of order $2m$ can be decomposed into the product of a symplectic matrix and an upper $J$-triangular matrix is answered in the following theorem.

**Theorem 3.** *(SR decomposition) [3, Theorem 3.8] Let $A_{2m} =\in \mathbb{R}^{2m,2m}$ be nonsingular and let $\hat{A}_{2m} = A_{2m} P_{2m}^T =\in \mathbb{R}^{2m,2m}$. There exists a decomposition $A_{2m} = S_{2m} R_{2m}$, where $S_{2m} \in \mathbb{R}^{2m,2m}$ is symplectic and $R_{2m}$ is upper $J$-triangular if and only if all leading principal minors of even dimension of the matrix $\hat{C}_{2m} = \hat{A}_{2m}^T J_{2m} \hat{A}_{2m}$ are nonzero. Then we have also the decomposition $\hat{A}_{2m} = \hat{S}_{2m} \hat{R}_{2m}$ where the factor $\hat{S}_{2m} = S_{2m} P_{2m}^{-1}$ satisfies $\hat{S}_{2m}^T J_{2m} \hat{S}_{2m} = \hat{J}_{2m}$ and the upper triangular factor $\hat{R}_{2m} = P_{2m} R_{2m} P_{2m}^T$ satisfies*

$$\begin{aligned} \hat{C}_{2m} &= \hat{A}_{2m}^T J_{2m} \hat{A}_{2m} = P_{2m} R_{2m}^T J_{2m} R_{2m} P_{2m}^T \\ &= P_{2m} R_{2m}^T P_{2m}^T \hat{J}_{2m} P_{2m} R P_{2m}^T = \hat{R}_{2m}^T \hat{J}_{2m} \hat{R}_{2m}. \end{aligned}$$

It follows from [3] that the SR decomposition is unique up to a special trivial (symplectic and upper $J$-triangular) factor that is characterized by the following corollary.

**Corollary 4.** *[3, Remark 3.9] Let $A_{2m} = S_{2m} R_{2m}$ and $A_{2m} = \tilde{S}_{2m} \tilde{R}_{2m}$ be two SR decompositions of a nonsingular matrix $A_{2m} \in \mathbb{R}^{2m,2m}$. Then there exists a matrix $D_{2m} = \begin{pmatrix} D_{1,1} & D_{1,2} \\ 0 & D_{1,1}^{-1} \end{pmatrix} \in \mathbb{R}^{2m,2m}$, where the matrices $D_{1,1} =$*

$\mathrm{diag}(d_{1,1},\ldots,d_{m,1}) \in \mathbb{R}^{m,m}$ *and* $D_{1,2} = \mathrm{diag}(d_{1,2},\ldots,d_{m,2}) \in \mathbb{R}^{m,m}$ *are such that* $\tilde{S}_{2m} = S_{2m}D_{2m}^{-1}$ *and* $\tilde{R}_{2m} = D_{2m}R_{2m}$.

Hence there is a freedom in the choice of the symplectic and the upper $J$-triangular factor. The SR decomposition is unique up to a special factor $D_{2m}$ that is at the same time symplectic and upper $J$-triangular leaving thus $2m$ parameters free.

In order to obtain a unique decomposition, several options have been proposed in the literature. Della-Dora [4] suggests to restrict the symplectic factor $S_{2m}$ to symplectic matrices with column sum equal to 1 in order to obtain a unique decomposition. This is also investigated in [7].

Mehrmann [7] suggests to restrict the upper $J$-triangular factor $R_{2m} = \left(\begin{smallmatrix} R_{1,1} & R_{1,2} \\ R_{2,1} & R_{2,2} \end{smallmatrix}\right)$ with additional constraints making the diagonal elements of $R_{1,2}$ zero and setting the diagonal of $R_{1,1}$ positive and in absolute value equal to the corresponding diagonal elements of $R_{2,2}$. As shown in [7] (under the assumption that all leading principal minors of even dimension of $\hat{C}_{2m}$ are nonzero as in Theorem 3), such an SR decomposition does always exist and is unique. This choice of the parameters will be denoted by MEH in the following discussions. Based on [7], Bunse-Gerstner [3] suggests to restrict the upper $J$-triangular factor to an upper $J$-triangular matrix where the upper right hand triangular part $R_{1,2}$ has a zero diagonal and the diagonal elements of $R_{1,1}$ are set to one in order to obtain a unique decomposition.

These restrictions have been proposed solely in order to obtain a unique SR decomposition for theoretical purposes. For the computation of the SR decomposition the degree of freedom should rather be used to make the computation as easy and stable as possible.

Salam considers in [8] symplectic Gram-Schmidt like algorithms. In that context the SR decomposition of a matrix $A = [a_1 \ a_2] \in \mathbb{R}^{2n\times 2}$ is considered. The SR decomposition $A = SR, S = [s_1 \ s_2] \in \mathbb{R}^{2n\times 2}$ and $R = \left(\begin{smallmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{smallmatrix}\right)$ was called the elementary SR decomposition (ESR). Clearly, $r_{11}r_{22} = a_1^T J_2 a_2$. Three different versions (ESR1, ESR2 and ESR3) have been discussed:

**ESR1** $r_{11} = \|a_1\|$, $r_{12}$ arbitrary,

**ESR2** $r_{11} = \|a_1\|$, $r_{12} = a_1^T a_2/r_{11}$,

**ESR3** $r_{11} = \sqrt{|a_1^T J_2 a_2|}$, $r_{12}$ arbitrary.

It is shown that ESR1 results in $\|s_1\| = 1, \|s_2\| \geq 1$, while ESR2 results in $\|s_1\| = 1$, and $\|s_2\|$ has minimal 2-norm. For ESR3 it is immediate that $r_{22} = \pm r_{11}$ depending on the sign of $a_1^T J_2 a_2$. The special case ESR3 with $r_{12} = 0$ was also suggested in [7]. It will play an important role in our findings, hence we introduce it here as

**ESR4** $r_{11} = \sqrt{|a_1^T J_2 a_2|}$, $r_{12} = 0$.

For this case, we have $\|s_1\| = \|a_1\|/\sqrt{|a_1^T J_2 a_2|}$ and $\|s_2\| = \|a_2\|/\sqrt{|a_1^T J_2 a_2|}$, that is, $\frac{\|a_1\|}{\|a_2\|} = \frac{\|s_1\|}{\|s_2\|}$.

In the context of the Hamiltonian Lanczos process it has been proposed to choose $\|s_1\| = 1, s_2^T s_1 = 0, s_1^T J s_2 = 1$ (that is, $s_1$ and $s_2$ are chosen orthogonal as well as $J$-orthogonal), see, e.g. [6, 1, 2]. In [1] it is shown that this minimizes the condition number of $S$ in case $\|s_1\| = 1$. This is the same as ESR2.

To the best of our knowledge there are no further theoretical results on how to make use of the freedom of choice in the symplectic and the upper $J$-triangular factors in the SR decomposition. The SR decomposition is usually not computed explicitly, algorithms based on the SR decomposition work in an implicit fashion as in the QR context. The choice of freedom in the symplectic and the upper $J$-triangular factors translates into the choice of freedom in the symplectic transformation and in the parameters of the upper Hessenberg-like form used. More knowledge about the effects of different choice for the freedom in the SR decomposition will help in choosing the free parameters in the implicit SR algorithms, see, e.g. [5] and the references therein. Moreover, algorithms based on the SR decomposition need some sort of re-orthogonalization with respect to the bilinear form imposed by $J$. The SR decomposition can be seen such an orthogonalization process. Hence, the choice of freedom in the SR decomposition means freedom in normalization.

The goal of this paper is to analyze how to make use of the freedom of choice in the symplectic and the upper $J$-triangular factors in the SR decomposition in order to yield stable algorithms. In particular, we will try to interpret some of the existing suggestions on how to choose the free parameters in terms of the conditioning of parts of the symplectic factor $S$ or the triangular factor $R$.

The organization of the paper is as follows. In Section 2 the minimization of the condition number of the diagonal blocks $\hat{\mathrm{R}}_{j,j}, j = 1, \ldots, m$ in the triangular factor $\hat{R}_{2m}$ is considered. It turns out that the particular choice of parameters denoted by MEH [7] or ESR4 [8] yields this minimum. Section 3 is devoted to the orthogonalization with respect to the bilinear form that is induced by the skew-symmetric matrix $J_{2m}$. In particular, we discuss how to choose $\hat{\mathrm{R}}_{j,j}, j = 1, \ldots, m$ in order to minimize the conditioning of the corresponding blocks $\hat{\mathrm{S}}_j = [\hat{s}_{2j-1}, \hat{s}_{2j}]$ in the symplectic factor $S_{2m}$. It turns out that ESR2 is locally worse than ESR3. A new choice (denoted here as ESR5) which is better than ESR2 and as good as ESR3 will be introduced. In Section 4 we develop bounds for the extremal singular values of the upper $J$-triangular factor $R_{2m}$ and the symplectic factor $S_{2m}$ in terms of the spectral properties of principal submatrices of the matrix $\hat{C}_{2m}$. In Section 5 we illustrate our results on two small model examples.

## 2 Minimization of the condition number of the diagonal blocks $\hat{\mathrm{R}}_{n,n}$ in the triangular factor

In this section, it will be shown that ESR4 as well as MEH minimizes the condition number of the diagonal blocks $\hat{\mathrm{R}}_{n,n}$ in the triangular factor $\hat{R}_{2m}$.

For $n = 1, \ldots, m$ we introduce the rectangular submatrices

$$A_{2m,2n} = [a_1, \ldots, a_{2n}] \in \mathbb{R}^{2m,2n}$$

and

$$\hat{A}_{2m,2n} = A_{2m,2n}P_{2n}^T = [\hat{a}_1, \ldots, \hat{a}_{2n}] \in \mathbb{R}^{2m,2n}$$

for the matrices $A_{2m}$ and $\hat{A}_{2m}$ as in Theorem 3. In addition, we define the rectangular matrices $S_{2m,2n} = [s_1, \ldots, s_{2n}] \in \mathbb{R}^{2m,2n}$ and $\hat{S}_{2m,2n} = S_{2m,2n}P_{2n}^{-1} = [\hat{s}_1, \ldots, \hat{s}_{2n}] \in \mathbb{R}^{2m,2n}$. From $\hat{S}_{2m}^T J_{2m}\hat{S}_{2m} = \hat{J}_{2m}$ and from the $2 \times 2$ block diagonal structure of the matrix $\hat{J}_{2m}$ it follows that

$$\hat{S}_{2m,2n}^T J_{2m}\hat{S}_{2m,2n} = \hat{J}_{2n}. \tag{1}$$

Consequently, the matrices $S_{2m,2n}$ are semi-symplectic with

$$S_{2m,2n}^T J_{2m} S_{2m,2n} = P_{2n}^T \hat{S}_{2m,2n}^T J_{2m}\hat{S}_{2m,2n}P_{2n} = P_{2n}^T\hat{J}_{2n}P_{2n} = J_{2n}.$$

Moreover, from the upper triangular structure of the factor $\hat{R}_{2m}$ in the decomposition $\hat{A}_{2m} = \hat{S}_{2m}\hat{R}_{2m}$ one can easily conclude that the same holds also for the $R$ factor in the SR decomposition of each rectangular submatrix $\hat{A}_{2m,2n} = \hat{S}_{2m,2n}\hat{R}_{2n}$. Here $\hat{R}_{2n} \in \mathbb{R}^{2n,2n}$ is the leading principal submatrix of order $2n$ of the upper triangular factor $\hat{R}_{2m}$ and thus it is also upper triangular. The factor $\hat{R}_{2n}$ can be computed from the Cholesky-like decomposition of the skew-symmetric matrix

$$\hat{C}_{2n} = \begin{pmatrix} \hat{\mathrm{C}}_{1,1} & \ldots & \hat{\mathrm{C}}_{1,n} \\ \vdots & & \vdots \\ \hat{\mathrm{C}}_{n,1} & \ldots & \hat{\mathrm{C}}_{n,n} \end{pmatrix} = \hat{A}_{2m,2n}^T J_{2m}\hat{A}_{2m,2n} = \hat{R}_{2n}^T\hat{J}_{2n}\hat{R}_{2n} \tag{2}$$

as (1) holds (please note, that a corresponding observation has been made for $\hat{R}_{2m}$ in Theorem 3). Here $\hat{C}_{2n}$ is partitioned into $2 \times 2$ blocks $\hat{\mathrm{C}}_{i,j}$.

We also have the decomposition $A_{2m,2n} = \hat{A}_{2m,2n}P_{2n}^{-T} = \hat{S}_{2m,2n}\hat{R}_{2n}P_{2n}^{-T} = S_{2m,2n}R_{2n}$, where $R_{2n}$ is defined as $R_{2n} = P_{2n}^{-1}\hat{R}_{2n}P_{2n}^{-T}$. Then $R_{2n}$ is a factor from the Cholesky-like decomposition $A_{2m,2n}^T J_{2m}A_{2m,2n} = R_{2n}^T J_{2n}R_{2n}$.

We consider the following partitioning of the skew-symmetric matrix $\hat{C}_{2n}$ and the upper triangular matrix $\hat{R}_{2n}$

$$\hat{C}_{2n} = \left(\begin{array}{ccc|c} & \hat{C}_{2(n-1)} & & \begin{matrix}\hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n}\end{matrix} \\ \hline \hat{\mathrm{C}}_{n,1} & \ldots & \hat{\mathrm{C}}_{n,n-1} & \hat{\mathrm{C}}_{n,n} \end{array}\right), \quad \hat{\mathrm{C}}_{j,n} = -\hat{\mathrm{C}}_{n,j}^T, \tag{3}$$

$$\hat{R}_{2n} = \left(\begin{array}{c|c} \hat{R}_{2(n-1)} & \begin{matrix}\hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n}\end{matrix} \\ \hline 0 & \hat{\mathrm{R}}_{n,n} \end{array}\right)$$

and the partitioning of the matrix $\hat{J}_{2n} = \mathrm{diag}(\hat{J}_{2(n-1)}, \hat{J}_2)$. Given the matrix $\hat{R}_{2(n-1)}$ at the step $n-1$, it follows then from (2) that the off-diagonal blocks

$\hat{\mathrm{R}}_{i,n}$ can be computed via forward substitution

$$\begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix} = -\hat{J}_{2(n-1)} \hat{R}_{2(n-1)}^{-T} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \tag{4}$$

as $\hat{R}_{2(n-1)}^{-T}$ is a lower triangular matrix. It is also easy to see from (2) and (4) that the diagonal block $\hat{\mathrm{R}}_{n,n}$ satisfies

$$\begin{aligned} \hat{\mathrm{R}}_{n,n}^T \hat{J}_2 \hat{\mathrm{R}}_{n,n} &= \hat{\mathrm{C}}_{n,n} - \begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix}^T \hat{J}_{2(n-1)} \begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix} \\ &= \hat{\mathrm{C}}_{n,n} - \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix}^T \hat{R}_{2(n-1)}^{-1} \hat{J}_{2(n-1)} \hat{R}_{2(n-1)}^{-T} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \\ &= \hat{\mathrm{C}}_{n,n} - \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix}^T \hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \qquad (5) \\ &= \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \\ \hat{\mathrm{C}}_{n,n} \end{pmatrix}^T \begin{bmatrix} -\hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \\ I_2 \end{bmatrix} \end{aligned}$$

as $\hat{R}_{2(n-1)}^T \hat{J}_{2(n-1)} \hat{R}_{2(n-1)} = \hat{C}_{2(n-1)}$. Hence $\hat{\mathrm{R}}_{n,n}$ can be recovered from the Cholesky-like decomposition of the $2\times 2$ Schur-complement matrix[1] $\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}$, i.e.

$$\hat{\mathrm{R}}_{n,n}^T \hat{J}_2 \hat{\mathrm{R}}_{n,n} = \hat{C}_{2n} \backslash \hat{C}_{2(n-1)}. \tag{6}$$

The matrices $\hat{C}_{2n}$ and $\hat{C}_{2(n-1)}$ are skew-symmetric and therefore the Schur-complement matrix $\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}$ is also skew-symmetric with

$$\hat{C}_{2n} \backslash \hat{C}_{2(n-1)} = \begin{pmatrix} 0 & \pm\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\| \\ \mp\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\| & 0 \end{pmatrix} = \pm\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\| \hat{J}_2.$$

[1]For $X = \begin{pmatrix} \mathrm{X}_{1,1} & \mathrm{X}_{1,2} \\ \mathrm{X}_{2,1} & \mathrm{X}_{2,2} \end{pmatrix} \in \mathbb{R}^{n+m,n+m}$, a nonsingular $\mathrm{X}_{1,1} \in \mathbb{R}^{n,n}, \mathrm{X}_{1,2} \in \mathbb{R}^{n,m}, \mathrm{X}_{2,1} \in \mathbb{R}^{m,n}, \mathrm{X}_{2,2} \in \mathbb{R}^{m,m}$ we denote the Schur complement by $X \backslash \mathrm{X}_{1,1} = \mathrm{X}_{2,2} - \mathrm{X}_{2,1} \mathrm{X}_{1,1}^{-1} \mathrm{X}_{1,2}$.

Denoting the nonzero elements of $\hat{\mathrm{R}}_{n,n}$ by $r_{11}$, $r_{12}$ and $r_{22}$ we get

$$\begin{aligned}\hat{\mathrm{R}}_{n,n}^T \hat{J}_2 \hat{\mathrm{R}}_{n,n} &= \begin{pmatrix} r_{11} & 0 \\ r_{12} & r_{22} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix} \\ &= \begin{pmatrix} 0 & r_{11}r_{22} \\ -r_{11}r_{22} & 0 \end{pmatrix}.\end{aligned}$$

The freedom of $2m$ free parameters in the SR decomposition discussed in Corollary 4 is thus reflected in a free choice of the off-diagonal element $r_{12}$ and the choice of one of the diagonal elements $r_{11}$ or $r_{22}$ for each block $\hat{\mathrm{R}}_{n,n}$. This fixes the other diagonal element as $r_{11}r_{22} = \pm\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|$ has to be satisfied for each block $\hat{\mathrm{R}}_{n,n}$, where $n = 1, \ldots, m$.

The condition number of the diagonal matrix block $\hat{\mathrm{R}}_{n,n}$ can be expressed by the eigenvalues of the symmetric positive definite matrix

$$\hat{\mathrm{R}}_{n,n}^T \hat{\mathrm{R}}_{n,n} = \begin{pmatrix} r_{11} & 0 \\ r_{12} & r_{22} \end{pmatrix} \begin{pmatrix} r_{11} & r_{12} \\ 0 & r_{22} \end{pmatrix} = \begin{pmatrix} r_{11}^2 & r_{11}r_{12} \\ r_{11}r_{12} & r_{12}^2 + r_{22}^2 \end{pmatrix}.$$

Obviously, $\kappa(\hat{\mathrm{R}}_{n,n})$ can be expressed in terms of the parameters $r_{11}$, $r_{22}$ and $r_{12}$. The eigenvalues of $\hat{\mathrm{R}}_{n,n}^T \hat{\mathrm{R}}_{n,n}$ are

$$\frac{r_{11}^2 + r_{12}^2 + r_{22}^2}{2} \pm \sqrt{\frac{(r_{11}^2 + r_{12}^2 + r_{22}^2)^2}{4} - r_{11}^2 r_{22}^2}.$$

As $r_{11}r_{22} = \pm\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|$ is a fixed constant and $\|\hat{\mathrm{R}}_{n,n}\|_F^2 = r_{11}^2 + r_{12}^2 + r_{22}^2$, we have

$$\kappa^2(\hat{\mathrm{R}}_{n,n}) = \frac{\|\hat{\mathrm{R}}_{n,n}\|_F^2 + \sqrt{\|\hat{\mathrm{R}}_{n,n}\|_F^4 - 4\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|^2}}{\|\hat{\mathrm{R}}_{n,n}\|_F^2 - \sqrt{\|\hat{\mathrm{R}}_{n,n}\|_F^4 - 4\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|^2}}.$$

Indeed $\kappa(\hat{\mathrm{R}}_{n,n})$ is an increasing function of $\|\hat{\mathrm{R}}_{n,n}\|_F^2$, so if we want to keep the condition number of $\hat{\mathrm{R}}_{n,n}$ small, we must minimize its Frobenius norm $\|\hat{\mathrm{R}}_{n,n}\|_F^2$. This leads to $r_{12} = 0$ and to minimization of $r_{11}^2 + \frac{\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|^2}{r_{11}^2}$ as a function of $r_{11}$. This function is minimized for

$$r_{11} = \pm r_{22} = \pm\sqrt{\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\|}. \tag{7}$$

Hence, while $\hat{\mathrm{R}}_{n,n}^T \hat{J}_2 \hat{\mathrm{R}}_{n,n} = \pm\|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\| \hat{J}_2$ has to hold (6), in this optimal case the matrix $\hat{\mathrm{R}}_{n,n}^T \hat{\mathrm{R}}_{n,n}$ satisfies

$$\hat{\mathrm{R}}_{n,n}^T \hat{\mathrm{R}}_{n,n} = \|\hat{C}_{2n} \backslash \hat{C}_{2(n-1)}\| I_2. \tag{8}$$

Note that this choice exactly corresponds to the choice MEH [7] and to the version ERS4, that is, ESR3 with the choice $r_{12} = 0$ in [8]. It is also clear that the scaling that minimizes the condition number of $\hat{\mathrm{R}}_{n,n}$ is not necessarily the one that minimizes the condition number of the block $\hat{\mathrm{S}}_n = [\hat{s}_{2n-1}, \hat{s}_{2n}]$ in the semi-symplectic factor $\hat{S}_{2m,2n}$. This aspect will be further discussed in Section 3.

## 3 Minimization of the condition number of the blocks $\hat{\mathrm{S}}_n$ in the semi-symplectic factor

The SR decomposition of the rectangular matrix $\hat{A}_{2m,2n} = \hat{S}_{2m,2n}\hat{R}_{2n}$ can be seen as an orthogonalization process with respect a bilinear form induced by the skew-symmetric matrix $J_{2m}$. Indeed, if this process is applied to certain blocks of two column vectors of $\hat{A}_{2m,2n}$, then as its outputs we get the block columns of the semi-symplectic factor $\hat{S}_{2m,2n}$ and the upper triangular factor $\hat{R}_{2n}$ that contains the orthogonalization coefficients. If we consider the two-vector partitioning of matrices $\hat{A}_{2m,2n} = [\hat{A}_{2m,2(n-1)}, \hat{\mathrm{A}}_n]$ with $\hat{\mathrm{A}}_n = [\hat{a}_{2n-1}, \hat{a}_{2n}]$ and $\hat{S}_{2m,2n} = [\hat{S}_{2m,2(n-1)}, \hat{\mathrm{S}}_n]$ with $\hat{\mathrm{S}}_n = [\hat{s}_{2n-1}, \hat{s}_{2n}]$, then the last two columns of $\hat{A}_{2m,2n} = \hat{S}_{2m,2n}\hat{R}_{2n}$ represent the recurrence for computing the vectors $\hat{\mathrm{S}}_n$ and the diagonal block $\hat{\mathrm{R}}_{n,n}$ as

$$\hat{S}_{2m,2(n-1)} \begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix} + \hat{\mathrm{S}}_n\hat{\mathrm{R}}_{n,n} = \hat{\mathrm{A}}_n.$$

Note that $\hat{\mathrm{S}}_n \in \mathbb{R}^{2m,2}$ is semi-symplectic with

$$\hat{\mathrm{S}}_n^T J_{2m}\hat{\mathrm{S}}_n = \hat{J}_2,$$

i.e. we have $\hat{s}_{2n-1}^T J_{2m}\hat{s}_{2n} = 1$. Introducing the matrix

$$\hat{\mathrm{U}}_n = [\hat{u}_{2n-1}, \hat{u}_{2n}] = \hat{\mathrm{S}}_n\hat{\mathrm{R}}_{n,n} \tag{9}$$

we can write

$$\hat{\mathrm{U}}_n = \hat{\mathrm{A}}_n - \hat{S}_{2m,2(n-1)} \begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix}, \tag{10}$$

where the off-diagonal blocks $\hat{\mathrm{R}}_{1,n}, \ldots, \hat{\mathrm{R}}_{n-1,n}$ with (4) and (2) satisfy

$$\begin{aligned} \begin{pmatrix} \hat{\mathrm{R}}_{1,n} \\ \vdots \\ \hat{\mathrm{R}}_{n-1,n} \end{pmatrix} &= \hat{J}_{2(n-1)}^{-1}\hat{R}_{2(n-1)}^{-T}\hat{A}_{2m,2(n-1)}^T J_{2m}\hat{\mathrm{A}}_n \\ &= \hat{J}_{2(n-1)}^{-1}\hat{S}_{2m,2(n-1)}^T J_{2m}\hat{\mathrm{A}}_n \end{aligned} \tag{11}$$

as $\hat{S}_{2m,2(n-1)} = \hat{A}_{2m,2(n-1)}\hat{R}_{2(n-1)}^{-1}$. Consequently, the columns of $\hat{\mathrm{U}}_n$ can be formally seen as the result of the orthogonalization of the columns of $\hat{\mathrm{A}}_n$ against the previously computed vectors of the semi-symplectic matrix $\hat{S}_{2m,2(n-1)}$ with respect to the bilinear form induced by the skew-symmetric matrix $J_{2m}$

$$\hat{\mathrm{U}}_n = \hat{\mathrm{A}}_n - \hat{S}_{2m,2(n-1)}\hat{J}_{2(n-1)}^{-1}\hat{S}_{2m,2(n-1)}^T J_{2m}\hat{\mathrm{A}}_n.$$

The last term can be rewritten using (2) and $\hat{S}_{2m,2(n-1)} = \hat{A}_{2m,2(n-1)}\hat{R}^{-1}_{2(n-1)}$

$$\hat{S}_{2m,2(n-1)}\hat{J}^{-1}_{2(n-1)}\hat{S}^T_{2m,2(n-1)}J_{2m}\hat{\mathrm{A}}_n = \hat{A}_{2m,2(n-1)}\hat{C}^{-1}_{2(n-1)}\begin{pmatrix}\hat{\mathrm{C}}_{1,n}\\ \vdots \\ \hat{\mathrm{C}}_{n-1,n}\end{pmatrix}.$$

Hence, from (10) and (11) we obtain for $\hat{\mathrm{U}}_n$

$$\hat{\mathrm{U}}_n = [\hat{A}_{2m,2(n-1)}, \hat{\mathrm{A}}_n]\begin{bmatrix} -\hat{C}^{-1}_{2(n-1)}\begin{pmatrix}\hat{\mathrm{C}}_{1,n}\\ \vdots \\ \hat{\mathrm{C}}_{n-1,n}\end{pmatrix} \\ I_2\end{bmatrix}. \tag{12}$$

Clearly, $\hat{\mathrm{U}}_n$ depends on $\hat{A}_{2m,2n}$ and $\hat{C}_{2n}$ (except of $\hat{\mathrm{C}}_{n,n}$).

Under the assumptions of Theorem 3 the diagonal elements of $\hat{R}_{2m}$ have to be nonzero. Hence the diagonal elements of each $\hat{\mathrm{R}}_{n,n}$ are nonzero. From

$$\hat{\mathrm{U}}_n^T J_{2m}\hat{\mathrm{U}}_n = \hat{\mathrm{R}}^T_{n,n}\hat{\mathrm{S}}^T_n J_{2m}\hat{\mathrm{S}}_n\hat{\mathrm{R}}_{n,n} = \hat{\mathrm{R}}^T_{n,n}\hat{J}_2\hat{\mathrm{R}}_{n,n},$$

we obtain the Cholesky-like decomposition

$$\begin{pmatrix} r_{11} & r_{12}\\ 0 & r_{22}\end{pmatrix}^T\begin{pmatrix} 0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix} r_{11} & r_{12}\\ 0 & r_{22}\end{pmatrix} = \begin{pmatrix} 0 & \hat{u}^T_{2n-1}J_{2m}\hat{u}_{2n}\\ \hat{u}^T_{2n}J_{2m}\hat{u}_{2n-1} & 0\end{pmatrix}$$

which implies $\hat{u}^T_{2n-1}J_{2m}\hat{u}_{2n} \neq 0$. Moreover, due to (6)

$$\hat{\mathrm{U}}_n^T J_{2m}\hat{\mathrm{U}}_n = \hat{C}_{2n}\backslash\hat{C}_{2(n-1)}.$$

The computation of $\hat{R}_{n,n}$ can be viewed as a kind of normalization in the orthogonalization process whereas

$$r_{11}r_{22} = \hat{u}^T_{2n-1}J_{2m}\hat{u}_{2n} = \pm\|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\|. \tag{13}$$

Note that $\hat{u}_{2n-1}$ and $\hat{u}_{2n}$ are given from previous projections in (10). They depend on $\hat{A}_{2m,2n}$, $\hat{C}_{2(n-1)}$ and $\hat{C}_{i,n}$ for $i = 1,\ldots,n-1$, see (12). Moreover, they determine the elements $r_{11}$ and $r_{12}$, whereas $r_{22}$ already depends on $r_{11}$ and the corresponding Schur complement $\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}$.

As already noted there is a freedom in the choice of $r_{12}$ and one of the elements $r_{11}$ and $r_{22}$. Choosing $r_{11}$ fixes $r_{22}$ due to the restriction (13). Considering the formula (9) with the fixed left-hand side $\hat{\mathrm{U}}_n$, we now ask for $\hat{\mathrm{R}}_{n,n}$ such that the conditioning of $\hat{\mathrm{S}}_n$ is minimal. Recall, in Section 2 we minimized the condition number of $\hat{\mathrm{R}}_{n,n}$, that is, we minimized the upper bound for its condition number in the bound $\kappa(\hat{\mathrm{S}}_n) \leq \kappa(\hat{\mathrm{U}}_n)\kappa(\hat{\mathrm{R}}_{n,n})$.

Now we want to choose $\hat{\mathrm{R}}_{n,n}$ such that $\kappa(\hat{\mathrm{S}}_n)$ is minimal. The eigenvalues of the matrix $\hat{\mathrm{S}}^T_n\hat{\mathrm{S}}_n$ satisfy the quadratic equation

$$(\|\hat{s}_{2n-1}\|^2 - \lambda)(\|\hat{s}_{2n}\|^2 - \lambda) - (\hat{s}_{2n-1}, \hat{s}_{2n})^2 = 0, \tag{14}$$

and therefore, for the squares of the singular values of $\hat{S}_n$ we have

$$\|\hat{S}_n\|^2, \quad \sigma^2_{min}(\hat{S}_n) = \frac{\|\hat{S}_n\|_F^2}{2} \pm \sqrt{\frac{\|\hat{S}_n\|_F^4}{4} - \det(\hat{S}_n^T\hat{S}_n)}.$$

It follows from (9) that $\hat{S}_n^T\hat{S}_n = \hat{R}_{n,n}^{-T}(\hat{U}_n^T\hat{U}_n)\hat{R}_{n,n}^{-1}$ and hence using the identity $\det(\hat{U}_n^T\hat{U}_n) = \|\hat{U}_n\|^2\sigma^2_{min}(\hat{U}_n)$ and (13) we see that

$$\begin{aligned}\det(\hat{S}_n^T\hat{S}_n) &= \det(\hat{R}_{n,n}^{-T})\det(\hat{U}_n^T\hat{U}_n)\det(\hat{R}_{n,n}^{-1}) \\ &= \frac{\det(\hat{U}_n^T\hat{U}_n)}{(r_{11}r_{22})^2} \\ &= \left(\frac{\|\hat{U}_n\|\sigma_{min}(\hat{U}_n)}{(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})}\right)^2\end{aligned}$$

does not depend on the parameters $r_{11}$ and $r_{12}$. Here $(\cdot,\cdot)$ denotes the standard inner product, $(x,y) = x^Ty$. Consequently, the condition number

$$\kappa^2(\hat{S}_n) = \frac{\|\hat{S}_n\|_F^2 + \sqrt{\|\hat{S}_n\|_F^2 - 4\det(\hat{S}_n^T\hat{S}_n)}}{\|\hat{S}_n\|_F^2 - \sqrt{\|\hat{S}_n\|_F^2 - 4\det(\hat{S}_n^T\hat{S}_n)}}$$

is minimized if we minimize $\|\hat{S}_n\|_F^2 = \|\hat{s}_{2n-1}\|^2 + \|\hat{s}_{2n}\|^2$ as a function of $r_{11}$ and $r_{12}$. Noting that $\hat{S}_n^T\hat{S}_n = (\hat{U}_n\hat{R}_{n,n}^{-1})^T(\hat{U}_n\hat{R}_{n,n}^{-1})$ also yields

$$\|\hat{s}_{2n-1}\|^2 = \|\hat{u}_{2n-1}\|^2/r_{11}^2, \tag{15}$$

$$\|\hat{s}_{2n}\|^2 = \|r_{12}\hat{u}_{2n-1} - r_{11}\hat{u}_{2n}\|^2/(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2, \tag{16}$$

and

$$(\hat{s}_{2n-1}, \hat{s}_{2n}) = -\frac{(\hat{u}_{2n-1}, r_{12}\hat{u}_{2n-1} - r_{11}\hat{u}_{2n})}{r_{11}(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})}. \tag{17}$$

Taking the partial derivatives

$$\begin{aligned}\frac{\partial\|\hat{S}_n\|_F^2}{\partial r_{11}} &= -2\frac{\|\hat{u}_{2n-1}\|^2}{r_{11}^3} - 2r_{12}\frac{(\hat{u}_{2n-1}, \hat{u}_{2n})}{(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2} + 2r_{11}\frac{\|\hat{u}_{2n}\|^2}{(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2}, \\ \frac{\partial\|\hat{S}_n\|_F^2}{\partial r_{12}} &= 2r_{12}\frac{\|\hat{u}_{2n-1}\|^2}{(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2} - 2r_{11}\frac{(\hat{u}_{2n-1}, \hat{u}_{2n})}{(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2},\end{aligned}$$

we obtain two nonlinear equations for the stationary point in the form

$$r_{12}\|\hat{u}_{2n-1}\|^2 - r_{11}(\hat{u}_{2n-1}, \hat{u}_{2n}) = 0, \tag{18}$$

$$r_{11}\|\hat{u}_{2n}\|^2 - r_{12}(\hat{u}_{2n-1}, \hat{u}_{2n}) = \frac{\|\hat{u}_{2n-1}\|^2(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})^2}{r_{11}^3}.$$

This leads to the formulas for the minimizer

$$r_{11} = \frac{\|\hat{u}_{2n-1}\||(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}}{(\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n))^{1/2}}, \tag{19}$$

$$r_{12} = \frac{(\hat{u}_{2n-1}, \hat{u}_{2n})|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}}{\|\hat{u}_{2n-1}\|(\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n))^{1/2}}. \tag{20}$$

This choice implies that (17) gives

$$(\hat{s}_{2n-1}, \hat{s}_{2n}) = 0.$$

Hence, the necessary condition (18) implies the orthogonality of $\hat{s}_{2n}$ with respect to $\hat{s}_{2n-1}$. Further, from (15), (16), (19) and (20) it follows, that

$$\|\hat{s}_{2n-1}\|^2 = \frac{\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n)}{|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|}, \qquad \|\hat{s}_{2n}\|^2 = \frac{\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n)}{|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|}.$$

and therefore,

$$\hat{\mathrm{S}}_n^T\hat{\mathrm{S}}_n = \frac{\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n)}{|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|} I_2. \tag{21}$$

Thus the block $\hat{\mathrm{S}}_n$ is well-conditioned with $\kappa(\hat{\mathrm{S}}_n) = 1$. Recall that here we considered the general case $r_{11} \neq r_{22}$ and no restriction on $r_{12}$ together with the restriction (13). This lead to the choice (19) and (20), which will be denoted by ESR5 from now on;

**ESR5** $r_{11} \neq r_{22}$, $r_{11}$ as in (19) and $r_{12}$ as in (20).

Note that both choices ESR2 and ESR5 lead to $\hat{\mathrm{S}}_n$ with orthogonal columns, but in ESR5 the norms of $\hat{s}_{2n-1}$ and $\hat{s}_{2n}$ are equilibrated, while in ESR2 we have $\|\hat{s}_{2n-1}\| = 1$ and $\|\hat{s}_{2n}\| \geq 1$ (see [8]) that leads often to much larger condition number of $\hat{\mathrm{S}}_n$ than in ESR5. This will be illustrated also in Section 5.

# 4 Condition number of factors in the SR decomposition

It follows from Section 2 that for ESR4, that is, under the optimal choice (7) of the parameters $r_{11}$ and $r_{12}$

$$|r_{11}| = |r_{22}|, \qquad r_{11}r_{22} = (\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n}), \qquad r_{12} = 0,$$

the norms of $\hat{\mathrm{R}}_{n,n}$ and $\hat{\mathrm{R}}_{n,n}^{-1}$ are determined by the norm of the Schur complement matrix $\|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\| = |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|$ whereas

$$\|\hat{\mathrm{R}}_{n,n}\| = |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}, \quad \|\hat{\mathrm{R}}_{n,n}^{-1}\| = |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{-1/2}. \tag{22}$$

This gives $\kappa(\hat{\mathrm{R}}_{n,n}) = 1$. Moreover, using $\hat{\mathrm{S}}_n = \hat{\mathrm{U}}_n \hat{\mathrm{R}}_{n,n}^{-1}$ we get

$$(23) \quad \|\hat{\mathrm{S}}_n\| = |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{-1/2}\|\hat{\mathrm{U}}_n\|, \quad \|\hat{\mathrm{S}}_n^\dagger\| = |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}\|\hat{\mathrm{U}}_n^\dagger\|$$

leading to $\kappa(\hat{\mathrm{S}}_n) = \kappa(\hat{\mathrm{U}}_n)$. Here $X^\dagger$ denotes the Moore-Penrose pseudoinverse of a (possibly rectangular) matrix $X$.

On the other hand, it follows from Section 3 that for ESR5, that is, under the optimal choice (7) of the parameters $r_{11}$ and $r_{12}$

$$r_{11} = \frac{\|\hat{u}_{2n-1}\||(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}}{(\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n))^{1/2}}, \quad r_{12} = \frac{(\hat{u}_{2n-1}, \hat{u}_{2n})|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}}{\|\hat{u}_{2n-1}\|(\|\hat{\mathrm{U}}_n\|\sigma_{min}(\hat{\mathrm{U}}_n))^{1/2}},$$

with $r_{22}$ given from the the condition $r_{11}r_{22} = (\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})$, the norms of $\hat{\mathrm{S}}_n$ and $\hat{\mathrm{S}}_n^\dagger$ satisfy

$$\begin{aligned} \|\hat{\mathrm{S}}_n\| &= \frac{\|\hat{\mathrm{U}}_n\|}{|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}\kappa^{1/2}(\hat{\mathrm{U}}_n)}, \\ (24) \qquad \|\hat{\mathrm{S}}_n^\dagger\| &= \frac{|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}\|\hat{\mathrm{U}}_n^\dagger\|}{\kappa^{1/2}(\hat{\mathrm{U}}_n)}. \end{aligned}$$

This gives $\kappa(\hat{\mathrm{S}}_n) = 1$. Moreover, from $\hat{\mathrm{U}}_n^T\hat{\mathrm{U}}_n = \hat{\mathrm{R}}_{n,n}^T\hat{\mathrm{S}}_n^T\hat{\mathrm{S}}_n\hat{\mathrm{R}}_{n,n}$ and (21) we get

$$\begin{aligned} \|\hat{\mathrm{R}}_{n,n}\| &= |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{1/2}\kappa^{1/2}(\hat{\mathrm{U}}_n), \\ (25) \qquad \|\hat{\mathrm{R}}_{n,n}^{-1}\| &= |(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})|^{-1/2}\kappa^{1/2}(\hat{\mathrm{U}}_n) \end{aligned}$$

leading to $\kappa(\hat{\mathrm{R}}_{n,n}) = \kappa(\hat{\mathrm{U}}_n)$.

We have $\max_{n=1,\ldots,m}\|\hat{\mathrm{S}}_n\| \leq \|\hat{S}_{2m}\|$. A similar statement holds for the minimal singular values $\min_{n=1,\ldots,m}\sigma_{min}(\hat{\mathrm{S}}_n) \geq \sigma_{min}(\hat{S}_{2m})$. It is also clear that for each $n = 1, \ldots, m$ we have

$$\|\hat{S}_{2m,2n}\| \leq \|\hat{S}_{2m}\|, \quad \sigma_{min}(\hat{S}_{2m,2n}) \geq \sigma_{min}(\hat{S}_{2m}).$$

Thus $\kappa(\hat{S}_{2m,2n}) \leq \kappa(S_{2m})$. For the symplectic factor $\hat{S}_{2m}$ we have that $\hat{S}_{2m}^T J_{2m}\hat{S}_{2m} = \hat{J}_{2m}$ and thus

$$\hat{S}_{2m}^{-1} = \hat{J}_{2m}^{-1}\hat{S}_{2m}^T\hat{J}_{2m}.$$

Therefore, the condition number of $\hat{S}_{2m}$ is given by the square of its norm as $\kappa(\hat{S}_{2m}) = \|\hat{S}_{2m}\|^2$. The norm of $\hat{S}_{2m}$ can be bounded as

$$\|\hat{S}_{2m}\| \leq \sum_{n=1}^{m}\|\hat{\mathrm{S}}_n\|.$$

The matrix $\hat{R}_{2m}$ is block upper triangular and therefore the norms of all diagonal blocks $\hat{\mathrm{R}}_{n,n}$ represent the lower bounds for its norm;

$$\max_{n=1,\ldots,m}\|\hat{\mathrm{R}}_{n,n}\| \leq \|\hat{R}_{2m}\|.$$

A similar statement can be given for the inverses $\max_{n=1,\ldots,m} \|\hat{\mathrm{R}}_{n,n}^{-1}\| \leq \|\hat{R}_{2m}^{-1}\|$. In the following theorems we give upper bounds for the norm of $\hat{R}_{2m}^{-1}$ and $\hat{R}_{2m}$ in terms of the norms of the diagonal blocks $\hat{\mathrm{R}}_{n,n}$ and $\hat{\mathrm{R}}_{n,n}^{-1}$, $n = 1, \ldots, m$.

**Theorem 5.** *Let $\hat{A}_{2m} \in \mathbb{R}^{2m,2m}$ and $\hat{A}_{2m,2n} \in \mathbb{R}^{2m,2n}$ be the matrices defined as in Theorem 3 such that no leading principal minor of even dimension of the matrix $\hat{C}_{2m} = \hat{A}_{2m}^T J_{2m} \hat{A}_{2m}$ vanishes (i.e., $\hat{C}_{2n} = \hat{A}_{2m,2n}^T J_{2m} \hat{A}_{2m,2n}$ is nonsingular for all $n = 1, \ldots, m$). In the SR decomposition of $\hat{A}_{2m}$ the triangular factor $\hat{R}_{2m}$ is chosen such that (13) holds for each diagonal $2 \times 2$ block $\hat{\mathrm{R}}_{n,n}$, $n = 1, \ldots, m$. Then the norm of the inverse of the triangular factor $R_{2m}$ is bounded by*

$$\|\hat{R}_{2m}^{-1}\| \leq \|\hat{\mathrm{R}}_{1,1}^{-1}\| + \sum_{n=2}^{m} \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n\| \|\hat{\mathrm{R}}_{n,n}^{-1}\|$$

*Proof.* Observe that

$$\hat{R}_{2n} = \left( \begin{array}{c|c} \hat{R}_{2(n-1)} & \hat{R}_{2(n-1)} \hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{C}_{1,n} \\ \vdots \\ \hat{C}_{n-1,n} \end{pmatrix} \\ \hline 0 & \hat{\mathrm{R}}_{n,n} \end{array} \right)$$

as the off-diagonal blocks $\hat{\mathrm{R}}_{i,n}$ satisfy (4) and $\hat{C}_{2(n-1)} = \hat{R}_{2(n-1)}^T \hat{J}_{2(n-1)} \hat{R}_{2(n-1)}$. Hence, the inverse of the triangular matrix $\hat{R}_{2n}$ can be expressed as

$$\hat{R}_{2n}^{-1} = \left( \begin{array}{c|c} \hat{R}_{2(n-1)}^{-1} & -\hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \hat{\mathrm{R}}_{n,n}^{-1} \\ \hline 0 & \hat{\mathrm{R}}_{n,n}^{-1} \end{array} \right). \tag{26}$$

Considering the identity (12) the previous expression can be rewritten as follows

$$\hat{R}_{2n}^{-1} = \left( \begin{array}{c|c} \begin{matrix} \hat{R}_{2(n-1)}^{-1} \\ 0 \end{matrix} & \hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n \hat{\mathrm{R}}_{n,n}^{-1} \end{array} \right).$$

Using the triangular inequality the norm of $\hat{R}_{2n}^{-1}$ can be bounded by

$$\begin{aligned} \|\hat{R}_{2n}^{-1}\| &\leq \|\hat{R}_{2(n-1)}^{-1}\| + \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n \hat{\mathrm{R}}_{n,n}^{-1}\| \\ &\leq \|\hat{R}_{2(n-1)}^{-1}\| + \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n\| \|\hat{\mathrm{R}}_{n,n}^{-1}\|. \end{aligned}$$

It follows immediately

$$\begin{aligned} \|\hat{R}_{2m}^{-1}\| &\leq \|\hat{R}_{2(m-2)}^{-1}\| + \sum_{n=m-1}^{m} \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n\| \|\hat{\mathrm{R}}_{n,n}^{-1}\| \\ &\leq \cdots \\ &\leq \|\hat{\mathrm{R}}_{1,1}^{-1}\| + \sum_{n=2}^{m} \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n\| \|\hat{\mathrm{R}}_{n,n}^{-1}\|. \end{aligned}$$

Therefore, we obtain the statement of our theorem. □

Since the matrix $\hat{R}_{2m}$ represents the upper triangular factor in the Cholesky-like decomposition $\hat{C}_{2m} = \hat{R}_{2m}^T \hat{J}_{2m} \hat{R}_{2m}$ we can bound its norm via $\|\hat{R}_{2m}\| \leq \|\hat{C}_{2m}\| \|\hat{R}_{2m}^{-1}\|$ and the bound for $\|\hat{R}_{2m}^{-1}\|$ formulated in Theorem 5. The norm of $\hat{R}_{2m}$ can be also bounded directly in terms of the norms of diagonal blocks $\hat{\mathrm{R}}_{n,n}$, $n = 1, \ldots, m$.

**Theorem 6.** *Under the same assumptions as in Theorem 5 the square of the norm of the triangular factor $\hat{R}_{2m}$ can be bounded as*

$$\|\hat{R}_{2m}\| \leq \sum_{n=1}^{m} \left( \|\hat{\mathrm{R}}_{n,n}\| + \|\hat{\mathrm{R}}_{n,n}^{-1}\| \|\hat{A}_{2m,2n}^{\dagger} \hat{\mathrm{U}}_n\| \|\hat{C}_{2m}\| \right).$$

*Proof.* Consider the following partitioning of $\hat{R}_{2m}$

$$\hat{R}_{2m} = \left( \begin{array}{c|ccc} & \hat{\mathrm{R}}_{1,n} & \cdots & \hat{\mathrm{R}}_{1,m} \\ \hat{R}_{2(n-1)} & \vdots & & \vdots \\ & \hat{\mathrm{R}}_{n-1,n} & \cdots & \hat{\mathrm{R}}_{n-1,m} \\ \hline 0 & & \breve{R}_{2(m-n+1)} & \end{array} \right)$$

with the trailing principal submatrix $\breve{R}_{2(m-n+1)} \in \mathbb{R}^{2(m-n+1),2(m-n+1)}$ for $n = 1, \ldots m$. Then due to $\breve{R}_{2m} = \hat{R}_{2m}$ $(n = 1)$ and (11) we have

$$\begin{aligned} \breve{R}_{2(m-n+1)} &= \left( \begin{array}{c|c} \hat{\mathrm{R}}_{n,n} & \hat{\mathrm{R}}_{n+1,m} \cdots \hat{\mathrm{R}}_{n,m} \\ \hline 0 & \breve{R}_{2(m-n)} \end{array} \right) \\ &= \left( \begin{array}{c|c} \hat{\mathrm{R}}_{n,n} & -\hat{J}_2^{-1} \hat{\mathrm{S}}_n^T J_{2m} \left( \hat{\mathrm{A}}_{n+1} \cdots \hat{\mathrm{A}}_m \right) \\ \hline 0 & \breve{R}_{2(m-n)} \end{array} \right). \end{aligned}$$

Using the triangular inequality the norm of $\breve{R}_{2(m-n+1)}$ can be then bounded by

$$\begin{aligned} \|\breve{R}_{2(m-n+1)}\| &\leq \|\breve{R}_{2(m-n)}\| + \|\hat{\mathrm{R}}_{n,n}\| + \|\hat{J}_2^{-1} \hat{\mathrm{S}}_n^T J_{2m} (\hat{\mathrm{A}}_{n+1} \cdots \hat{\mathrm{A}}_m)\| \\ &\leq \|\breve{R}_{2(m-n)}\| + \|\hat{\mathrm{R}}_{n,n}\| + \|\hat{\mathrm{R}}_{n,n}^{-1}\| \|\hat{\mathrm{U}}_n^T J_{2m} (\hat{\mathrm{A}}_{n+1} \cdots \hat{\mathrm{A}}_m)\| \end{aligned}$$

as $\hat{\mathrm{S}}_n = \hat{\mathrm{U}}_n \hat{\mathrm{R}}_{n,n}^{-1}$. Similarly, it follows immediately from the previous recurrence that

$$\begin{aligned} \|\hat{R}_{2m}\| &\leq \|\breve{R}_{2(m-2)}\| + \sum_{n=1}^{2} \|\hat{\mathrm{R}}_{n,n}\| + \|\hat{\mathrm{R}}_{n,n}^{-1}\| \|\hat{\mathrm{U}}_n^T J_{2m} (\hat{\mathrm{A}}_{n+1} \cdots \hat{\mathrm{A}}_m)\| \\ &\leq \cdots \\ &\leq \sum_{n=1}^{m} \left( \|\hat{\mathrm{R}}_{n,n}\| + \|\hat{\mathrm{R}}_{n,n}^{-1}\| \|\hat{\mathrm{U}}_n^T J_{2m} (\hat{\mathrm{A}}_{n+1} \cdots \hat{\mathrm{A}}_m)\| \right). \end{aligned}$$

The statement of our theorem now follows from the identity

$$\hat{\mathrm{U}}_n^T J_{2m}(\hat{\mathrm{A}}_{n+1}\cdots\hat{\mathrm{A}}_m) = (\hat{A}_{2m,2n}^{\dagger}\hat{\mathrm{U}}_n)^T \begin{pmatrix} \hat{\mathrm{C}}_{1,n+1} & \cdots & \hat{\mathrm{C}}_{1,m} \\ \vdots & & \vdots \\ \hat{\mathrm{C}}_{n,n+1} & \cdots & \hat{\mathrm{C}}_{n,m} \end{pmatrix}$$

considering that the norm of its rightmost matrix can be bounded by the norm of the matrix $\hat{C}_{2m}$. $\square$

The norm of the matrix $\hat{A}_{2m,2n}^{\dagger}\hat{\mathrm{U}}_n$ that appears in the statements of Theorem 5 and Theorem 6 can be bounded as

$$\begin{aligned} \|\hat{A}_{2m,2n}^{\dagger}\hat{\mathrm{U}}_n\| &\leq 1 + \|\hat{C}_{2(n-1)}^{-1}\begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix}\| \\ &\leq 1 + \|\hat{C}_{2(n-1)}^{-1}\|\|\begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix}\| \\ &\leq 1 + \|\hat{C}_{2(n-1)}^{-1}\|\|\hat{C}_{2n}\|. \end{aligned}$$

Similarly, it follows from (5) that the norm of the Schur complement matrix $|(\hat{u}_{2n-1}, J_{2m}\hat{u}_{2n})| = \|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\|$ can be bounded from above by

$$\begin{aligned} \|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\| &\leq \|\begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \\ \hat{\mathrm{C}}_{n,n} \end{pmatrix}\|(1 + \|\hat{C}_{2(n-1)}^{-1}\begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix}\|) \\ &\leq \|\hat{C}_{2n}\|(1 + \|\hat{C}_{2(n-1)}^{-1}\|\|\hat{C}_{2n}\|). \end{aligned}$$

The norm of its inverse $(\hat{C}_{2n}\backslash\hat{C}_{2(n-1)})^{-1}$ can be bounded by the norm of $\hat{C}_{2n}^{-1}$, as the inverse of the Schur complement $(\hat{C}_{2n}\backslash\hat{C}_{2(n-1)})^{-1}$ is the lower right block in the inverse of the matrix $\hat{C}_{2n}^{-1}$. This leads to the lower bound

$$1/\|\hat{C}_{2n}^{-1}\| \leq \|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\|.$$

If the principal submatrix $\hat{C}_{2(n-1)}$ is ill-conditioned, then one may observe a significant growth of entries in the Schur complement matrix which leads to large norm of the diagonal block $\hat{\mathrm{R}}_{n,n}$. If in such case the principal submatrix $\hat{C}_{2n}$ is well-conditioned, the lower bound for $\|\hat{C}_{2n}\backslash\hat{C}_{2(n-1)}\|$ is a large underestimate of the actual size of the norm of the Schur complement matrix. Conversely, if $\hat{C}_{2(n-1)}$ is well-conditioned and $\hat{C}_{2n}$ is ill-conditioned, the size of the norm of the Schur complement matrix can be rather small. However, this leads to an increase of the norm of $\hat{\mathrm{R}}_{n,n}^{-1}$. Similar considerations can be done also for the norms of the blocks $\hat{\mathrm{S}}_n$.

The condition number of $\hat{\mathrm{U}}_n$ plays an important role when comparing the singular values of the blocks $\hat{\mathrm{R}}_{n,n}$ and $\hat{\mathrm{S}}_n$ in ESR4 and ESR5, respectively. It follows from (22) and (24) that the norms of $\hat{\mathrm{R}}_{n,n}$ and $\hat{\mathrm{R}}_{n,n}^{-1}$ are by a factor of $\kappa^{1/2}(\hat{\mathrm{U}}_n)$ larger in ESR5 than in ESR4, while the identities (23) and (25) show that that the norms of $\hat{\mathrm{S}}_n$ and $\hat{\mathrm{S}}_n^{\dagger}$ in ESR5 are by a factor of $\kappa^{1/2}(\hat{\mathrm{U}}_n)$ smaller than those in ESR4. Using (9) the extremal singular values of $\hat{\mathrm{U}}_n$ can be bounded by

$$\|\hat{\mathrm{U}}_n\| \leq \|\hat{A}_{2m,2n}\| \left( 1 + \|\hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix} \| \right),$$

and

$$\begin{aligned} \sigma_{min}(\hat{\mathrm{U}}_n) &\geq \sigma_{min}(\hat{A}_{2m,2n}) \sqrt{1 + \sigma_{min}^2(\hat{C}_{2(n-1)}^{-1} \begin{pmatrix} \hat{\mathrm{C}}_{1,n} \\ \vdots \\ \hat{\mathrm{C}}_{n-1,n} \end{pmatrix})} \\ &\geq \sigma_{min}(\hat{A}_{2m,2n}). \end{aligned}$$

# 5 Numerical examples

Recall from Theorem 3 that the upper triangular factor $\hat{R}_{2m}$ of the SR decomposition $\hat{A}_{2m} = \hat{S}_{2m}\hat{R}_{2m}$ can be obtained from the Cholesky-like decomposition

$$\hat{C}_{2m} = \hat{A}_{2m}^T J_{2m} \hat{A}_{2m} = \hat{R}_{2m}^T \hat{J}_{2m} \hat{R}_{2m}.$$

It was shown in [9] that given the fixed skew-symmetric matrix $\hat{C}_{2n}$, the factorization $\hat{C}_{2n} = \hat{A}_{2m,2n}^T J_{2m} \hat{A}_{2m,2n}$, where $\hat{A}_{2m,2n} \in \mathbb{R}^{2m,2n}$ is an arbitrary rectangular matrix is not unique. Several classes of matrices have been discussed in [9] including the analysis of the conditioning of such factors that must always satisfy $\kappa^2(\hat{A}_{2m,2n}) \geq \kappa(\hat{C}_{2n})$. The author also gives the characterization of matrices with minimal condition number that lead to the equality $\kappa^2(\hat{A}_{2m,2n}) = \kappa(\hat{C}_{2n})$. We will consider a well-conditioned $\hat{C}_2$ and, among the potentially many matrices $\hat{A}_{4,2}$ that lead to this same $\hat{C}_2$, two examples:

- an ill-conditioned $\hat{A}_{4,2}$ with $\kappa^2(\hat{A}_{4,2}) \gg \kappa(\hat{C}_2)$, where ESR4 and ESR5 behave very differently: for ESR4 a well-conditioned $\hat{R}_2$ and ill-conditioned $\hat{S}_2$ is obtained, while a well-conditioned $\hat{S}_2$ and ill-conditioned $\hat{R}_2$ is obtained for ESR5.
- a well-conditioned $\hat{A}_{4,2}$ with $\kappa(\hat{A}_{4,2}) \approx 1$, where we will show that both $\hat{S}_2$ and $\hat{R}_2$ are not well-conditioned but quite similar for ESR4 and ESR5.

In both examples we also look at the conditioning of factors in the ESR2 case.

In particular, we will consider the $2 \times 2$ skew-symmetric matrix $\hat{C}_2$

$$\hat{C}_2 = \begin{pmatrix} 0 & \varepsilon \\ -\varepsilon & 0 \end{pmatrix},$$

where $\varepsilon$ is a small positive parameter. It appears that both singular values of $\hat{C}_2$ are equal to $\varepsilon$ and thus $\hat{C}_2$ itself is well-conditioned with $\kappa(\hat{C}_2) = 1$.

First we consider the $4 \times 2$ matrix

$$\hat{A}_{4,2} = \begin{pmatrix} \varepsilon & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{pmatrix}$$

such that $\hat{A}_{4,2}^T J_4 \hat{A}_{4,2} = \hat{C}_2$. It is easy to show that $\hat{A}_{4,2}$ is ill-conditioned with $\kappa(\hat{A}_{4,2}) \approx \frac{2}{\varepsilon}$, whereas $\|\hat{A}_{4,2}\| = \sqrt{2}$ and $\sigma_{min}(\hat{A}_{4,2}) = \frac{\varepsilon}{\sqrt{2}}$.

**Case ESR4:** First we will consider the ESR4 setting. Due to (2) the factor $\hat{R}_2$ has to satisfy $\hat{R}_2^T \hat{J}_2 \hat{R}_2 = \hat{C}_2$. As $r_{12} = 0$ this yields the decomposition

$$\begin{pmatrix} 0 & \varepsilon \\ -\varepsilon & 0 \end{pmatrix} = \hat{R}_2^T \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \hat{R}_2 \quad \Rightarrow \quad \hat{R}_2 = \begin{pmatrix} \sqrt{\varepsilon} & 0 \\ 0 & \sqrt{\varepsilon} \end{pmatrix}.$$

The singular values of the diagonal matrix $\hat{R}_2$ are both equal to $\sqrt{\varepsilon}$, but the factor $\hat{R}_2$ remains perfectly conditioned with $\kappa(\hat{R}_2) = 1$. The factor $\hat{S}_2$ has the following form

$$\hat{S}_2 = \begin{pmatrix} \sqrt{\varepsilon} & 0 \\ 0 & 0 \\ \frac{1}{\sqrt{\varepsilon}} & \frac{1}{\sqrt{\varepsilon}} \\ 0 & 0 \end{pmatrix}.$$

The norm of $\hat{S}_2$ is equal to $\|\hat{S}_2\| = \frac{\sqrt{2}}{\sqrt{\varepsilon}}$ and the minimal singular value of $\hat{S}_2$ is equal to its reciprocal $\sigma(\hat{S}_2) = \frac{\sqrt{\varepsilon}}{\sqrt{2}}$. Thus $\hat{S}_2$ is ill-conditioned with the condition number given as $\kappa(\hat{S}_2) = \kappa(\hat{A}_{4,2}) = \frac{2}{\varepsilon}$.

**Case ESR5:** The situation changes dramatically when we consider the case ESR5. The elements of the factor $\hat{R}_2$ are given by

$$\hat{R}_2 = \begin{pmatrix} \sqrt{1+\varepsilon^2} & \frac{\sqrt{\varepsilon}}{\sqrt{1+\varepsilon^2}} \\ 0 & \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} \end{pmatrix}.$$

The singular values of $\hat{R}_2$ are equal to $\|\hat{R}_2\| = \sqrt{2}$ and $\sigma_{min}(\hat{R}_2) = \frac{\varepsilon}{\sqrt{2}}$. The factor $\hat{R}_2$ is ill-conditioned with the condition number given as $\kappa(\hat{R}_2) = \kappa(\hat{A}_{4,2}) =$

$\frac{2}{\varepsilon}$. The factor $\hat{S}_2$ is given by

$$\hat{S}_2 = \begin{pmatrix} \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} & -\frac{1}{\varepsilon}\frac{1}{\sqrt{1+\varepsilon^2}} \\ 0 & 0 \\ \frac{1}{\varepsilon}\frac{1}{\sqrt{1+\varepsilon^2}} & \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} \\ 0 & 0 \end{pmatrix}.$$

The columns of $\hat{S}_2$ are orthogonal with respect to the standard inner product and have the same norm. Thus $\hat{S}_2$ is perfectly conditioned with $\kappa(\hat{S}_2) = 1$.

**Case ESR2:** Note that due to the property $\|\hat{C}_2\| = \|\hat{A}_2\|\sigma_{min}(\hat{A}_2) = \varepsilon$ the scheme ESR2 coincides with the scheme ESR5 on this example.

In our second example the $4 \times 2$ matrix $\hat{A}_{4,2}$ is chosen as

$$\hat{A}_{4,2} = \begin{pmatrix} \varepsilon & \varepsilon \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$$

such that $\hat{A}_{4,2}^T J_4 \hat{A}_{4,2} = \hat{C}_2$. However, $\hat{A}_{4,2}$ is now well-conditioned with $\kappa(\hat{A}_{4,2}) = \|\hat{A}_{4,2}\| = \sqrt{1+2\varepsilon^2}$ and $\sigma_{min}(\hat{A}_{4,2}) = 1$.

**Case ESR4:** The factor $\hat{R}_2$ for ESR4 is the same as in the first example

$$\hat{R}_2 = \begin{pmatrix} \sqrt{\varepsilon} & 0 \\ 0 & \sqrt{\varepsilon} \end{pmatrix}.$$

It is perfectly conditioned with $\kappa(\hat{R}_2) = 1$. The factor $\hat{S}_2$ is given by

$$\hat{S}_2 = \begin{pmatrix} \sqrt{\varepsilon} & \sqrt{\varepsilon} \\ \frac{1}{\sqrt{\varepsilon}} & 0 \\ 0 & \frac{1}{\sqrt{\varepsilon}} \\ 0 & 0 \end{pmatrix}$$

with the norm $\|\hat{S}_2\| = \frac{\sqrt{1+2\varepsilon}}{\sqrt{\varepsilon}}$. The minimal singular value of $\hat{S}_2$ is equal to $\sigma(\hat{S}_2) = \frac{1}{\sqrt{\varepsilon}}$. Nevertheless, the factor $\hat{S}_2$ remains well-conditioned for small values of $\varepsilon$ as its condition number is given by $\kappa(\hat{S}_2) = 1 + 2\varepsilon^2$.

**Case ESR5:** The factor $\hat{R}_2$ is equal to

$$\hat{R}_2 = \begin{pmatrix} \frac{\sqrt{1+\varepsilon^2}\sqrt{\varepsilon}}{\sqrt[4]{1+2\varepsilon^2}} & \frac{\varepsilon^2\sqrt{\varepsilon}}{\sqrt{1+\varepsilon^2}\sqrt[4]{1+2\varepsilon^2}} \\ 0 & \frac{\sqrt{\varepsilon}\sqrt[4]{1+2\varepsilon^2}}{\sqrt{1+\varepsilon^2}} \end{pmatrix}.$$

It is easy to show that the singular values of $\hat{R}_2$ are both approximately equal to $\|\hat{R}_2\| \approx \sigma_{min}(\hat{R}_2) \approx \sqrt{\varepsilon}$. The factor $\hat{R}_2$ remains also well-conditioned for small values of $\varepsilon$. The factor $\hat{S}_2$ is given by

$$\hat{S}_2 = \begin{pmatrix} \frac{\sqrt{\varepsilon}\sqrt[4]{1+2\varepsilon^2}}{\sqrt{1+\varepsilon^2}} & \frac{\sqrt{\varepsilon}}{\sqrt{1+\varepsilon^2}\sqrt[4]{1+2\varepsilon^2}} \\ \frac{\sqrt[4]{1+2\varepsilon^2}}{\sqrt{\varepsilon}\sqrt{1+\varepsilon^2}} & -\frac{\varepsilon\sqrt{\varepsilon}}{\sqrt{1+\varepsilon^2}\sqrt[4]{1+2\varepsilon^2}} \\ 0 & \frac{\sqrt{1+\varepsilon^2}}{\sqrt{\varepsilon}\sqrt[4]{1+2\varepsilon^2}} \\ 0 & 0 \end{pmatrix}.$$

As before, the columns of $\hat{S}_2$ are orthogonal with respect to the standard inner product and have the same norm. Thus $\hat{S}_2$ is perfectly conditioned, $\kappa(\hat{S}_2) = 1$.

**Case ESR2:** The situation is quite different from ESR5 if we consider the case ESR2. The elements of the factor $\hat{R}_2$ are given by

$$\hat{R}_2 = \begin{pmatrix} \sqrt{1+\varepsilon^2} & \frac{\varepsilon^2}{\sqrt{1+\varepsilon^2}} \\ 0 & \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} \end{pmatrix}.$$

The singular values of $\hat{R}_2$ are approximately equal to $\|\hat{R}_2\| = \sqrt{1+\varepsilon^2}$ and $\sigma_{min}(\hat{R}_2) = \varepsilon$. The factor $\hat{R}_2$ is ill-conditioned. Its the condition number is approximately equal to $\kappa(\hat{R}_2) = \frac{\sqrt{1+\varepsilon^2}}{\varepsilon}$. The factor $\hat{S}_2$ is given by

$$\hat{S}_2 = \begin{pmatrix} \frac{\varepsilon}{\sqrt{1+\varepsilon^2}} & \frac{1}{\sqrt{1+\varepsilon^2}} \\ \frac{1}{\sqrt{1+\varepsilon^2}} & -\frac{\varepsilon}{\sqrt{1+\varepsilon^2}} \\ 0 & \frac{\sqrt{1+\varepsilon^2}}{\sqrt{\varepsilon}} \\ 0 & 0 \end{pmatrix}.$$

The factor $\hat{S}_2$ has now orthogonal columns with respect to the standard inner product, but their norms are significantly different making the factor ill-conditioned. Its minimum singular value is equal to $\sigma_{min}(\hat{S}_2) = 1$ and its norm and condition number are equal to $\kappa(\hat{S}_2) = \|\hat{S}_2\| = \frac{\sqrt{1+\varepsilon+\varepsilon^2}}{\sqrt{\varepsilon}}$.

# 6 Conclusions

In this paper we have considered the decomposition $\hat{A}_{2m} = \hat{S}_{2m}\hat{R}_{2m}$ into a matrix $\hat{S}_{2m}$ with $\hat{S}_{2m}^T J_{2m} \hat{S}_{2m} = \hat{J}_{2m}$ and an upper triangular matrix $\hat{R}_{2m}$. Assuming that all principal minors of even dimension of the matrix $\hat{C}_{2m} = \hat{A}_{2m}^T J_{2m} \hat{A}_{2m}$ are nonzero, we have analyzed the conditioning of the triangular factor $\hat{R}_{2m}$ from Cholesky-like factorization of $\hat{C}_{2m} = \hat{R}_{2m}^T J_{2m} \hat{R}_{2m}$. The norms of $\hat{R}_{2m}$ and its inverse $\hat{R}_{2m}^{-1}$ can be bounded in terms of the norms of inverses of principal submatrices of even dimension and the block columns of $\hat{C}_{2m}$, and in terms of spectral properties of diagonal blocks $\hat{\mathrm{R}}_{n,n}$, $n = 1, \ldots, m$, in the resulting triangular factor $\hat{R}_{2m}$. Since there is a freedom in the choice of these blocks we have analyzed two choices: the choice ESR4 that minimizes the condition number of the diagonal block $\hat{\mathrm{R}}_{n,n}$ and the choice ESR5 that minimizes the condition number of the block $\hat{\mathrm{S}}_n$ in the symplectic factor $\hat{S}_{2m}$ at each step $n = 1, \dots m$. Since the decomposition $\hat{A}_{2m} = \hat{S}_{2m}\hat{R}_{2m}$ can be seen as an orthogonalization process with respect to a bilinear form induced by the skew-symmetric matrix $J_{2m}$, the computation of blocks $\hat{\mathrm{S}}_n$ and $\hat{\mathrm{R}}_{n,n}$ in the elementary SR decomposition (ESR) can be seen as a kind of normalization. The spectral properties of $\hat{\mathrm{S}}_n$ and $\hat{\mathrm{R}}_{n,n}$ depend on the conditioning of the matrix $\hat{\mathrm{U}}_n = \hat{\mathrm{S}}_n\hat{\mathrm{R}}_{n,n}$. Our analysis indicates that if the condition numbers of $\hat{\mathrm{U}}_n$ are large, there can be a significant difference in the conditioning of the resulting triangular and symplectic factors in ESR4 and ESR5. This difference can be small in the case of well-conditioned matrices $\hat{\mathrm{U}}_n$ as also indicated in one of our numerical examples.

# References

[1] P. Benner, *Structured Krylov subspace methods for eigenproblems with spectral symmetries*, in Workshop on Theoretical and Computational Aspects of Matrix Algorithms, Schloß Dagstuhl, October 1217, 2003, October 2003, Available from: http://archiv.tu-chemnitz.de/pub/2010/0085.

[2] P. Benner, H. Fassbender, and M. Stoll, *A Hamiltonian Krylov-Schur-type method based on the symplectic Lanczos process*, Linear Algebra and its Applications, 435 (2011), pp. 578–600.

[3] A. Bunse-Gerstner, *Matrix factorizations for symplectic QR-like methods*, Linear Algebra and its Applications, 83 (1986), pp. 49–77.

[4] J. Della Dora, *Numerical linear algorithms and group theory*, Linear Algebra and its Applications, 10 (1975), pp. 267–283.

[5] H. Fassbender, *Structured eigenvalue problems - structure-preserving algorithms, structured error analysis*, in Handbook of Linear Algebra, L. Hogben, ed., Discrete Mathematics and Its Applications, CRC Press, Taylor and Francis Group, second edition ed., 2014, ch. 63.

[6] W.R. Ferng, W.-W. Lin, and C.-S. Wang, *The shift-inverted J-Lanczos algorithm for the numerical solutions of large sparse algebraic Riccati equations*, Comput. Math. Appl., 33 (1997), pp. 23–40.

[7] V. Mehrmann, *Der SR-Algorithmus zur Berechnung der Eigenwerte einer Matrix*, Diplomarbeit, Universität Bielefeld, Bielefeld, FRG, 1979.

[8] A. Salam, *On theoretical and numerical aspects of symplectic gram-schmidt-like algorithms*, Numerical Algorithms, 39 (2005), pp. 437–462.

[9] H. Xu, *An SVD-like matrix decomposition and its applications*, Linear Algebra Appl., 368 (2003), pp. 1–24.