Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

THE EIGENVALUE DECOMPOSITION OF NORMAL MATRICES BY THE SKEW-SYMMETRIC PART*

SIMON MATAIGNE† AND KYLE A. GALLIVAN‡

Abstract. We propose a new method for computing the eigenvalue decomposition of a dense real normal matrix A through the decomposition of its skew-symmetric part. The method relies on algorithms that are known to be efficiently implemented, such as the bidiagonal singular value decomposition and the symmetric eigenvalue decomposition. The advantages of this method stand for normal matrices with few real eigenvalues, such as random orthogonal matrices. We provide a stability and a complexity analysis of the method. The numerical performance is compared with existing algorithms. In most cases, the method has the same operation count as the Hessenberg factorization of a dense matrix. Finally, we provide experiments for the application of computing a Riemannian barycenter on the special orthogonal group.

Key words. Normal matrices, Eigenvalue decomposition, Orthogonal matrices, Schur decomposition, Matrix logarithm, Riemannian barycenter, Karcher mean.

AMS subject classifications. 65F15, 15B10, 15B57, 62R30.

1. Introduction.

Computing the eigenvalues and the eigenvectors of a matrix is one of the most fundamental problems in linear algebra. It is well known that certain classes of matrices allow fast algorithms with convergence and accuracy guarantees. For example, Hermitian matrices ($A = \overline{A^\top} =: A^*$) enjoy a wide scope of applications, and it is known that the QR algorithm [38, 39] converges globally [79, 80, 81]. Additionally, for Hermitian matrices, a divide and conquer method and a multisectioning method were designed [24, 31, 56]. Robust symmetric eigensolvers are available in `LAPACK` [6, 30]. All these methods consist of reducing the Hermitian matrix to a tridiagonal form under unitary ($UU^* = U^*U = I$) similarity transformations [58, 59, 63, 78] and then solving the tridiagonal eigenvalue problem (EVP). For general matrices, the QR algorithm remains a state of the art approach to obtain the Schur form (upper triangular matrix with eigenvalues on the diagonal).

Although not nearly as popular as the symmetric EVP, the real skew-symmetric EVP ($A = -A^\top$) is known to be even more efficiently solvable [60, 75] because, first, skew-symmetric matrices can also be reduced to a tridiagonal form [52, 75]. Second, the zero diagonal makes the EVP equivalent to a bidiagonal singular value problem with half the size of the original EVP. It can be solved both rapidly and accurately [27, 28, 36]. This procedure exploits that eigenvalues arise by pairs of conjugate purely imaginary numbers. Unfortunately, in the complex setting, skew-Hermitian matrices turn into Hermitian matrices when multiplied by the imaginary unit. This strategy cannot be extended; real-valued matrices are more structured.

The focus of this paper, normal matrices ($AA^* = A^*A$), includes the two previous classes. It was shown that the shifted QR algorithm converges for normal matrices [10]. However, normal matrices are

*Received by the editors on October 7, 2025. Accepted for publication on March 30, 2026. Handling Editor: Vanni Noferini. Corresponding Author: Simon Mataigne.

†ICTEAM Institute, UCLouvain, Louvain-la-Neuve, Belgium (simon.mataigne@uclouvain.be). Simon Mataigne is a Research Fellow of the Fonds de la Recherche Scientifique - FNRS. This work was supported by the Fonds de la Recherche Scientifique - FNRS under Grant no T.0001.23.

‡Department of Mathematics, Florida State University, Tallahasse, FL, USA (kgallivan@fsu.edu). This work was initiated while Kyle A. Gallivan was a Visiting Professor at the ICTEAM Institute, UCLouvain.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

diagonalizable under unitary similarity transformations, i.e., the Schur form is diagonal. Consequently, the costly QR iterations operate on a dense upper Hessenberg matrix while converging to a diagonal matrix without exploiting this property. The key observation for a better method is the following: the Hermitian part $H := \frac{A+A^*}{2}$ and the skew-Hermitian part $\Omega := \frac{A-A^*}{2}$ of a normal matrix A commute: $H\Omega = \Omega H$. It is known that two normal matrices H and Ω that commute are jointly diagonalizable, i.e., there exists a unitary matrix U such that $U^*HU = \Lambda_H$ and $U^*\Omega U = \Lambda_\Omega$ [51, Thm. 2.5.5], and thus $A = U(\Lambda_H + \Lambda_\Omega)U^*$. The latter decomposition can be obtained by the *simultaneous diagonalization* of two Hermitian matrices (here H and $i\Omega$) using the Jacobi method [19, 40]. These algorithms were notably developed for generalized EVPs [23, 77]. Another approach, called the *do-one-then-do-the-other* (DODO), considers the sequential diagonalization of H and Ω [66]. Moreover, a method leveraging random combinations of H and $i\Omega$ has been proposed recently [48]. Finally, the normal EVP can be transformed into a complex symmetric EVP, and more details can be found in [9, 37, 71]. However, all previously cited methods consider *complex* matrices. For real matrices, a Jacobi method [83] was designed. The aim of this paper is to show that the additional structure of real matrices enables a specific and computationally efficient approach.

Unitary matrices, called *orthogonal matrices* when real-valued, are an important subset of normal matrices. A rich literature has emerged from the representation of $n \times n$ unitary Hessenberg matrices in their *Schur parametric form*, i.e., as the product of n elementary Householder matrices (or equivalently as Givens rotations):

$$H = H_1 H_2 ... H_n \text{ where}$$
$$H_k = \operatorname{diag}\left[I_{k-1}, \begin{bmatrix} -\alpha_k & \beta_k \\ \beta_k & \alpha_k^* \end{bmatrix}, I_{n-k-3}\right] \text{ with } |\alpha_k|^2 + |\beta_k|^2 = 1 \text{ for } k = 1, 2, ..., n-1,$$

and $H_n = \operatorname{diag}[I_{n-1}, -\alpha_n]$ with $|\alpha_n| = 1$. This compact storage format allows cheap and stable QR iterations, called UHQR, requiring only $\mathcal{O}(n)$ flops per iteration when only the eigenvalues are desired [44, 65]. This method can be implemented as a bulge-chasing algorithm [8], and it was shown that appropriate shifts produce quadratic convergence [73, 74]. The success of the Schur parametric form also led to the development of alternative algorithms such as divide-and-conquer [43, 46] and SVD-based methods [4, 21]. The common step for all previously cited methods is the initial reduction to Hessenberg form. Our method avoids this computationally expensive step.

Contributions. In this paper, we show that real normal matrices, with orthogonal matrices as a special case, admit a fast factorization to the EVD by exploiting the structure of their invariant spaces. This structure only exists in the real setting. In most cases,[1] if a normal matrix $A \in \mathbb{R}^{n\times n}$ has r real eigenvalues (r does not need to be known in advance), its EVD can be obtained by solving the rank $n-r$ EVP of the skew-symmetric part of A, plus a $r \times r$ symmetric EVP. The method is designed for matrices where $r \ll n$, a frequent property of random orthogonal matrices (e.g., Haar-distributed). We provide a stability and a complexity analysis of the method and compare its numerical performance with existing algorithms. When all eigenvalues have distinct imaginary parts, i.e., well-separated in floating point arithmetic, the method has the same operation count as the Hessenberg factorization of a dense matrix. When the imaginary parts of the eigenvalues are clustered, we propose refinement steps based on theoretical bounds that allow to maintain a user-specified accuracy of the decomposition. We corroborate our results in a suite of experiments. The method facilitates efficiently computing the matrix logarithm of orthogonal matrices, the main computational

[1]The precise meaning of "most cases" is described in section 6.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

bottleneck for calculating Riemannian logarithms and barycenters on matrix manifolds. In particular, in these applications, the EVP is an intermediate computation used in iterative algorithms where speed is more important than accuracy to machine precision. In addition, the normal EVP also finds applications in generalized EVPs [20, 23] and computation of Pisarenko frequency estimates [5, 25].

Overview. Section 2 introduces the essential properties of the normal EVP and related concepts. Section 3 presents theoretical results that lead to the exact arithmetic formulation of our method for the normal EVP: Algorithm 3.1. In section 4, we analyze the impact of floating-point arithmetic through a sensitivity analysis of the invariant spaces of the skew-symmetric part with respect to perturbations of the original matrix; this yields Algorithm 4.1. Section 5 corroborates the perturbation analysis through a series of experiments. In section 6, we examine the computational complexity of the algorithm and validate the analysis with timing experiments. Finally, in section 7, we apply Algorithm 4.1 to enhance the performance of Riemannian gradient descent for computing the Riemannian barycenter on $\mathrm{SO}(n)$.

Reproducibility. An implementation of the method, as well as the codes to perform all experiments presented in this paper, is available at https://github.com/smataigne/NormalEVP.jl.

Notation. We write $A = [a_{ij}]_{i,j=1}^n$ to denote matrix entries. We use $\mathrm{sym}(A) \coloneqq \frac{A+A^\top}{2}$ and $\mathrm{skew}(A) \coloneqq \frac{A-A^\top}{2}$ for, respectively, the symmetric part and the skew-symmetric part of the matrix A. $A_{a:b}$ denotes the selection of *columns* indexed by a to b included, and $A_{a:b,c:d}$ refers to the submatrix of *rows* a to b and *columns* c to d included. The Frobenius norm and the spectral norm are denoted $\|\cdot\|_\mathrm{F}$ and $\|\cdot\|_2$, respectively. The identity matrix of size $n \times n$ is denoted I_n. Since orthogonal transformations are central to normal EVPs, we recall the notation for the orthogonal group $\mathrm{O}(n)$ defined as

$$\mathrm{O}(n) \coloneqq \{Q \in \mathbb{R}^{n\times n} \;| Q^\top Q = QQ^\top = I_n\}.$$

$\mathrm{O}(n)$ admits two connected components: matrices with determinant 1 and matrices with determinant -1, sometimes, respectively, referred to as the set of rotations and reflections. Rotations are best known as the special orthogonal group $\mathrm{SO}(n) \coloneqq \{Q \in \mathrm{O}(n) \mid \det(Q) = 1\}$. Finally, the set of $n \times p$ matrices, $n \geq p$, with orthonormal columns is called the Stiefel manifold $\mathrm{St}(n,p)$, namely

$$\mathrm{St}(n,p) \coloneqq \{V \in \mathbb{R}^{n\times p} \mid V^\top V = I_p\}.$$

In the special case $n = p$, we have $\mathrm{St}(n,n) = \mathrm{O}(n)$.

2. Preliminaries.

2.1. The normal eigenvalue problem.

In the context of *real* normal matrices, eigenvalues, and eigenvectors are, in general, complex-valued. Nonetheless, it is preferred for most eigenvalue algorithms to compute with real arithmetic. For a normal matrix A, obtaining the real Schur decomposition (RSD), i.e., $A = QSQ^\top$ with $Q \in \mathrm{O}(n)$, is equivalent to obtaining the EVD. Indeed, the real Schur form S is block-diagonal with 2×2 and/or 1×1 blocks [51, Thm 2.5.8]. Each 2×2 block has the form, $\lambda \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$, $\lambda > 0,\ \theta \in (0,\pi)$, and is associated with the conjugate pair of eigenvalues $\{\lambda e^{\pm i\theta}\}$. The transition from the RSD to the EVD uses for each 2×2 block the relation

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

$$\lambda \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ i & -i \end{bmatrix} \begin{bmatrix} \lambda e^{-i\theta} & 0 \\ 0 & \lambda e^{i\theta} \end{bmatrix} \begin{bmatrix} 1 & -i \\ 1 & i \end{bmatrix}. \tag{2.1}$$

In this paper, we consider a permuted version of the real Schur form S, introduced in Property 2.1. The latter permuted version allows writing the matrix transformations easily and succinctly. Therefore, we will refer to (2.2) as the "real Schur decomposition," although it differs from the usual block-diagonal definition by a permutation.

PROPERTY 2.1 ([51, Thm. 2.5.8]). *Every normal matrix $A \in \mathbb{R}^{n\times n}$ is similar to a real Schur form $S \in \mathbb{R}^{n\times n}$ under orthogonal transformation by a matrix $Q \in \mathrm{O}(n)$ such that*

$$A = QSQ^\top = Q \begin{bmatrix} \Lambda\cos(\Theta) & 0 & -\Lambda\sin(\Theta) \\ 0 & \breve{\Lambda} & 0 \\ \Lambda\sin(\Theta) & 0 & \Lambda\cos(\Theta) \end{bmatrix} Q^\top, \tag{2.2}$$

where $\breve{\Lambda} = \mathrm{diag}(\breve{\lambda}_1, ..., \breve{\lambda}_r)$ contains the r real eigenvalues of A. Let $p := \frac{n-r}{2}$,[2] then $\Lambda = \mathrm{diag}(\lambda_1, ..., \lambda_p)$ and $\Theta = \mathrm{diag}(\theta_1, ..., \theta_p)$ are real, can be assumed positive with $\theta_j \in (0, \pi)$ w.l.o.g. and are associated with the pairs of eigenvalues $\{\lambda_j e^{\pm i\theta_j}\}$ for $j = 1, ..., p$. The columns of the matrix Q are called Schur vectors. *Special case 1: if A is symmetric, $r = n$. Special case 2: if A is skew-symmetric, $\cos(\Theta) = 0$ and $\breve{\Lambda} = 0_r$ [51, Cor. 2.5.11].*

REMARK 2.2. The Schur vectors of A associated with real eigenvalues are also eigenvectors of A. Moreover, in view of (2.1), if $q_1, q_2 \in \mathrm{St}(n, 1)$ is a pair of orthogonal Schur vectors associated with the same (λ, θ)-block, then $\mathrm{span}([q_1\ q_2])$ is an invariant subspace of A and $\frac{\sqrt{2}}{2}(q_1 + iq_2)$ and $\frac{\sqrt{2}}{2}(q_1 - iq_2)$ are eigenvectors of A.

2.2. The orthogonal groups. The Lie groups $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ are particularly important subsets of normal matrices for applications [1, 22, 53]. Computing the RSD is a key primitive to obtain the matrix logarithm, which is required for computing Riemannian logarithms [57, 61, 67, 84, 85] and Riemannian barycenters [15, 16, 54, 55, 82, 85], notably on the Stiefel manifold. The principal matrix logarithm can be efficiently computed via the relation

$$\log(A) = Q\log(S)Q^\top = Q \begin{bmatrix} \log(\Lambda) & 0 & -\Theta \\ 0 & \log(\breve{\Lambda}) & 0 \\ \Theta & 0 & \log(\Lambda) \end{bmatrix} Q^\top. \tag{2.3}$$

For both $\mathrm{O}(n)$ and $\mathrm{SO}(n)$, the property that eigenvalues lie on the unit circle yields $\Lambda = I_p$ and $\breve{\Lambda} = \mathrm{diag}(\pm 1, ..., \pm 1)$. In particular, we have $\log(\Lambda) = 0$. Note that $\log(\breve{\Lambda})$ admits a real element if and only if negative eigenvalues arise by pairs since $\left[\begin{smallmatrix} 0 & -\pi \\ \pi & 0 \end{smallmatrix}\right] \in \log\left(\left[\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right]\right)$.

2.3. Invariance groups. The decomposition from (2.2) is often not unique. It is key to this paper to describe the admissible transformations of the Schur vectors Q such that (2.2) remains satisfied. To this end, we introduce the notion of the *invariance group* of a matrix in Definition 2.3.

DEFINITION 2.3. *Given a matrix $A \in \mathbb{R}^{n\times n}$, the* invariance group *of A, written $\mathrm{ig}(A)$, is defined as the set*

$$\mathrm{ig}(A) := \{Q \in \mathrm{O}(n) \mid QAQ^\top = A\}. \tag{2.4}$$

[2]Note that p is always an integer since r is odd if and only if n is odd.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

An equivalent definition of $\mathrm{ig}(A)$ is the set of orthogonal matrices that commute with A. It is easily verified that $\mathrm{ig}(A)$ is a group under matrix multiplication since for all $Q_1, Q_2 \in \mathrm{ig}(A)$, we have $Q_1Q_2 \in \mathrm{ig}(A)$. The set $\mathrm{ig}\left(\left[\begin{smallmatrix} 0 & -I_m \\ I_m & 0 \end{smallmatrix}\right]\right)$ is well known as the *ortho-symplectic group* $\mathrm{OSp}(2m)$, see, e.g., [32].

Let A be a normal matrix and $A = QSQ^\top$ be an RSD. For every $R \in \mathrm{ig}(S)$, we have

$$A = (QR)(R^\top SR)(QR)^\top = (QR)S(QR)^\top.$$

Therefore, QR is a also a matrix of Schur vectors. If $r < n$ (A is not symmetric) or if $\breve{\Lambda}$ has repeated diagonal entries, then $\mathrm{ig}(S)$ is uncountable. Indeed, for each 2×2 (λ, θ)-block of S and each real eigenvalue $\breve{\lambda}$ of multiplicity $m > 1$, we, respectively, have

$$\mathrm{ig}\left(\lambda \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}\right) = \mathrm{SO}(2) \quad \text{and,} \quad \mathrm{ig}(\breve{\lambda} I_m) = \mathrm{O}(m).$$

Finally, we highlight a property of invariance groups that is important in section 3.

Lemma 2.4. *For all pairs of matrices $A, B \in \mathbb{R}^{n\times n}$, we have*

$$\mathrm{ig}(A) \cap \mathrm{ig}(B) \subseteq \mathrm{ig}(A + B).$$

Proof. Let $Q \in \mathrm{ig}(A) \cap \mathrm{ig}(B)$. Therefore, we have $QAQ^\top = A$ and $QBQ^\top = B$, and by summation, $Q(A + B)Q^\top = A + B$. Hence, $Q \in \mathrm{ig}(A + B)$. ∎

2.4. The even-odd permutation.

To conclude the preliminaries, we introduce the notions of even-odd permutation and bidiagonalization of a tridiagonal skew-symmetric matrix. These notions are essential to the skew-symmetric EVP and are used in section 3.

Definition 2.5. *An even–odd permutation matrix $P_{\mathrm{eo}} \in \mathbb{R}^{n\times n}$ is a permutation such that for every vector $[a_1\ a_2\ \dots\ a_n] \in \mathbb{R}^{1\times n}$, we have*

$$[a_1\ a_2\ \dots\ a_n]P_{\mathrm{eo}} = [a_1\ a_3\ \dots\ a_{\lceil \frac{n}{2} \rceil}\ a_2\ a_4\ \dots\ a_{\lfloor \frac{n}{2} \rfloor}].$$

The even-odd permutation is well known in the field of discrete mathematics, and it is related to the notion of Red-Black ordering [68]. This permutation allows the reduction of the skew-symmetric EVP to a bidiagonal SVD problem by a similarity tranformation, described in Property 2.6.

Property 2.6 ([75]). *Let $T \in \mathbb{R}^{n\times n}$ be a skew-symmetric tridiagonal matrix and $P_{\mathrm{eo}} \in \mathbb{R}^{n\times n}$ be an even-odd permutation matrix, then*

$$P_{\mathrm{eo}}^\top T P_{\mathrm{eo}} = \begin{bmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{bmatrix}, \tag{2.5}$$

where $\widetilde{B} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor \times \lceil \frac{n}{2} \rceil}$ is bidiagonal.

3. The eigenvalue decomposition of normal matrices.

In this section, we describe our method in exact arithmetic for computing the RSD of a normal matrix A. We show how the skew-symmetric part of a normal matrix reveals the structure of its Schur vectors and how these computations can be organized to design a fast RSD/EVD algorithm. In section 4, we analyze the effect of floating point arithmetic. In particular, the method relies on the assumption that the three following routines are available:

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

- Routine 1: returns the SVD of a real bidiagonal matrix.
- Routine 2: returns the RSD/EVD of a real symmetric matrix.
- Routine 3: returns the RSD of a (small) real normal matrix.

In the context of exact arithmetic, these routines can be viewed as oracles, whereas in practice, they are efficiently implemented, for example, in `LAPACK`. Indeed, Routine 1 and Routine 2 are, respectively, available as `bdsqr`/`bdsdc` and `syev`/`syevd`/`syevr` routines. For the coherence of this discussion, the calls to Routine 3 should be exceptional and only applied on small-size problems. Routine 3 may be chosen to be `LAPACK`'s `gees` routine since the shifted QR algorithm is known to converge for normal matrices [10]. Routine 3 may also be any of the previously cited methods for normal or orthogonal EVPs, e.g., [4, 74, 83]. All of these `LAPACK`'s routines are stable, see [6].

3.1. The skew-symmetric eigenvalue problem. The method starts by obtaining the RSD of $\Omega \coloneqq \mathrm{skew}(A)$. It is known to be efficiently done in three steps [75].

First, we compute $\Omega = \widetilde{Q}T\widetilde{Q}^\top$ with $\widetilde{Q} \in \mathrm{O}(n)$ and T skew-symmetric and tridiagonal. This decomposition can be obtained in at least two ways, both adapted versions of symmetric tridiagonalization. Either by using Householder reflectors [78], see, e.g., SkewLinearAlgebra.jl, or by using the Lanczos algorithm [52]. Householder tridiagonalization is preferred for dense matrices for its stability and use of level 3 `BLAS` operations [17, 62, 42]. Lanczos algorithm finds applications for large sparse matrices.

Secondly, an even-odd permutation matrix P_{eo} yields

$$\Omega = \widetilde{Q}T\widetilde{Q}^\top = \widetilde{Q}P_{\mathrm{eo}} \begin{bmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{bmatrix} P_{\mathrm{eo}}^\top \widetilde{Q}^\top,$$

where $\widetilde{B} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor \times \lceil \frac{n}{2} \rceil}$ is bidiagonal. In the context of exact arithmetic, a transformation $G \in \mathrm{O}(n)$ can isolate the central square matrix 0_r in (3.1). G is obtained by a finite process described in https://github.com/smataigne/NormalEVP.jl. In practice, we compute the SVD of $\widetilde{B}$ directly. For theoretical purposes, we consider the transformation G in this section. This leads to

$$\Omega = \widetilde{Q}P_{\mathrm{eo}}G \begin{bmatrix} 0 & 0 & -B^\top \\ 0 & 0_r & 0 \\ B & 0 & 0 \end{bmatrix} G^\top P_{\mathrm{eo}}^\top \widetilde{Q}^\top, \tag{3.1}$$

where $B \in \mathbb{R}^{p \times p}$ is bidiagonal.

The third and final step is a call to Routine 1 to compute the SVD of the bidiagonal block $B = U\Sigma V^\top$. This yields

$$\Omega = \widetilde{Q}P_{\mathrm{eo}}G \begin{bmatrix} V & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & U \end{bmatrix} \begin{bmatrix} 0 & 0 & -\Sigma \\ 0 & 0_r & 0 \\ \Sigma & 0 & 0 \end{bmatrix} \begin{bmatrix} V^\top & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & U^\top \end{bmatrix} G^\top P_{\mathrm{eo}}^\top \widetilde{Q}^\top. \tag{3.2}$$

By defining the matrix $\widehat{Q} \coloneqq \widetilde{Q}P_{\mathrm{eo}}G \left[\begin{smallmatrix} V & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & U \end{smallmatrix}\right]$, the decomposition (3.2) becomes an RSD of Ω. This decomposition can be compared with the skew-symmetric part of the (unknown) RSD of $A = QSQ^\top$ from (2.2):

$$\Omega = Q\mathrm{skew}(S)Q^\top = Q \begin{bmatrix} 0 & 0 & -\Lambda\sin(\Theta) \\ 0 & 0_r & 0 \\ \Lambda\sin(\Theta) & 0 & 0 \end{bmatrix} Q^\top. \tag{3.3}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

By comparing (3.2) and (3.3), it must hold that $\Sigma = \Lambda \sin(\Theta)$ if we assume that both are sorted, e.g., in decreasing order. Unfortunately, it does not hold true for the Schur vectors: $Q \neq \widehat{Q}$ in general. However, by equating (3.2) and (3.3), it holds that

$$\widehat{Q}^\top Q \mathrm{skew}(S) Q^\top \widehat{Q} = \mathrm{skew}(S) \quad \Longrightarrow \quad \widehat{Q}^\top Q \in \mathrm{ig}(\mathrm{skew}(S)).$$

We recall that $\widehat{Q}$ is a matrix of Schur vectors of A if and only if $\widehat{Q}^\top Q \in \mathrm{ig}(S)$. We have thus not yet computed the Schur vectors of A since, in general, $\mathrm{ig}(\mathrm{skew}(S)) \not\subseteq \mathrm{ig}(S)$. Designing the next steps of the algorithm requires finding the relation between $\mathrm{ig}(\mathrm{skew}(S))$ and $\mathrm{ig}(S)$. We show in the next subsection that it depends on the number r of real eigenvalues and the diagonal entries of $\Lambda \sin(\Theta)$.

3.2. Invariant spaces of the skew-symmetric part: preparatory lemmas. The rest of section 3 is dedicated to the description of the relation between the invariance groups $\mathrm{ig}(\mathrm{skew}(S))$ and $\mathrm{ig}(S)$. By completing this analysis, we will be able to find a matrix $R \in \mathrm{O}(n)$ such that $(\widehat{Q}R)^\top Q \in \mathrm{ig}(S)$, and thus $A = (\widehat{Q}R)S(\widehat{Q}R)^\top$. Let us start by dividing the eigenvalues of A in three sets:

- Set 1: The eigenvalues with distinct nonzero imaginary parts.
- Set 2: The eigenvalues with repeated nonzero imaginary parts.
- Set 3: The real eigenvalues.

A preliminary result is showing that R features a block structure related to the aforementioned sets of eigenvalues of A. This is done in Lemmas 3.1 and 3.2. To this end, we partition Q and $\widehat{Q}$ column-wise to match the blocks of the decomposition (2.2):

$$Q \coloneqq [Q_{\mathrm{c},1}\ Q_{\mathrm{r}}\ Q_{\mathrm{c},2}] \quad \text{and,} \quad \widehat{Q} \coloneqq [\widehat{Q}_{\mathrm{c},1}\ \widehat{Q}_{\mathrm{r}}\ \widehat{Q}_{\mathrm{c},2}],$$

where the notation follows because, by definition, Q_{r} contains the Schur vectors associated with the *real* eigenvalues and $[Q_{\mathrm{c},1}\ Q_{\mathrm{c},2}]$ contains the Schur vectors associated with *complex* eigenvalues. Lemma 3.1 shows the equivalence between the real eigenspace of A and the nullspace of $\mathrm{skew}(A)$.

LEMMA 3.1. *Let $A \in \mathbb{R}^{n\times n}$ be a normal matrix with exactly $r > 0$ real eigenvalues and $\breve{\Lambda} \in \mathbb{R}^{r\times r}$ be a diagonal matrix with these eigenvalues as diagonal entries. If a matrix $V \in \mathrm{St}(n,r)$ satisfies $AV = V\breve{\Lambda}$, then $\mathrm{skew}(A)V = 0$. Conversely, if $\mathrm{skew}(A)V = 0$, then there exists a matrix $\breve{R} \in \mathrm{O}(r)$ such that $A(V\breve{R}) = (V\breve{R})\breve{\Lambda}$.*

Proof. ($\Longrightarrow$) Assume $AV = V\breve{\Lambda}$. Then, by definition, there is an RSD $A = QSQ^\top$ such that $V^\top Q = [0_{r\times p}\ I_{r\times r}\ 0_{r\times p}]$. It follows that $\mathrm{skew}(A)V = Q\mathrm{skew}(S)Q^\top V = Q0 = 0$.

($\Longleftarrow$) Assume $\mathrm{skew}(A)V = 0$ and let $A = QSQ^\top$ be an RSD. Then it follows that $\ker(\mathrm{skew}(A)) = \mathrm{span}(Q_{\mathrm{r}})$. Moreover, since $\mathrm{skew}(A)V = 0$, $\mathrm{span}(V) \subseteq \ker(\mathrm{skew}(A))$. Since both Q_{r} and V have column rank r, their columns span the same subspace of $\mathbb{R}^n$ and there is $\breve{R} \in \mathbb{R}^{r\times r}$ such that $Q_{\mathrm{r}} = V\breve{R}$. Finally, since $Q_{\mathrm{r}} \in \mathrm{St}(n,r)$, $I_r = Q_{\mathrm{r}}^\top Q_{\mathrm{r}} = \breve{R}^\top V^\top V \breve{R} = \breve{R}^\top \breve{R}$. Hence, $\breve{R} \in \mathrm{O}(r)$. ∎

Since $\widehat{Q}$ satisfies $\mathrm{skew}(A)\widehat{Q} = \widehat{Q}\mathrm{skew}(S)$ by (3.2), it holds in particular that $\widehat{Q}_{\mathrm{r}}$ spans the kernel of $\mathrm{skew}(A)$: $\mathrm{skew}(A)\widehat{Q}_{\mathrm{r}} = 0$. Therefore, Lemma 3.1 ensures the existence of $\breve{R} \in \mathrm{O}(r)$ such that $Q_{\mathrm{r}} = \widehat{Q}_{\mathrm{r}}\breve{R}$. As a corollary of Lemma 3.1, it also holds that

$$Q_{\mathrm{r}}^\top [\widehat{Q}_{\mathrm{c},1}\ \widehat{Q}_{\mathrm{c},2}] = 0 \quad \text{and,} \quad \widehat{Q}_{\mathrm{r}}^\top [Q_{\mathrm{c},1}\ Q_{\mathrm{c},2}] = 0.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Therefore, we can conclude that the matrix R such that $(\widehat{Q}R)^\top Q \in \mathrm{ig}(S)$ has the block structure

$$R = \begin{bmatrix} R_{\mathrm{a}} & 0 & R_{\mathrm{b}} \\ 0 & \breve{R} & 0 \\ R_{\mathrm{c}} & 0 & R_{\mathrm{b}} \end{bmatrix} \in \mathrm{O}(n), \tag{3.4}$$

where $\left[\begin{smallmatrix} R_{\mathrm{a}} & R_{\mathrm{b}} \\ R_{\mathrm{c}} & R_{\mathrm{d}} \end{smallmatrix}\right] \in \mathrm{O}(2p)$, $R_{\mathrm{a}} \in \mathbb{R}^{p\times p}$ and $\breve{R} \in \mathrm{O}(r)$.

We can show that R is even more structured by considering Set 1 and Set 2 separately. Lemma 3.2 shows that, for each set, R acts on independent sets of columns of $\widehat{Q}$.

Lemma 3.2. *For every normal matrix $A \in \mathbb{R}^{n\times n}$ that has $m \geq 1$ eigenvalues with imaginary part $\sigma > 0$, it holds for every $V \in \mathrm{St}(n, 2m)$ satisfying $\mathrm{skew}(A)V = V \left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]$ that there exists a matrix $R \in \mathrm{O}(2m)$ such that*

$$A(VR) = (VR)\begin{bmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{bmatrix},$$

where $D_m = \mathrm{diag}(d_1, ..., d_m)$ and $d_j \pm i\sigma$ is an eigenvalue of A for $j = 1, ..., m$.

Proof. Let $k := 2(p - m)$ and $A = \breve{Q}\breve{S}\breve{Q}^\top$ be an RSD with

$$\breve{S} = \begin{bmatrix} D_m & -\sigma I_m & 0 & 0 & 0 \\ \sigma I_m & D_m & 0 & 0 & 0 \\ 0 & 0 & \breve{\Lambda} & 0 & 0 \\ 0 & 0 & 0 & \Lambda_k \cos(\Theta_k) & -\Lambda_k \sin(\Theta_k) \\ 0 & 0 & 0 & \Lambda_k \sin(\Theta_k) & \Lambda_k \cos(\Theta_k) \end{bmatrix}, \tag{3.5}$$

where Λ_k, Θ_k are $k \times k$ diagonal matrices.[3] It follows that

$$\mathrm{skew}(A)V = V\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix} \iff \mathrm{skew}(\breve{S})\breve{Q}^\top V = \breve{Q}^\top V\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}. \tag{3.6}$$

Let us divide $\breve{Q}^\top V$ row-wise in three blocks such that

$$\breve{Q}^\top V := \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \text{ with } X \in \mathbb{R}^{2m\times 2m}, Y \in \mathbb{R}^{r\times 2m} \text{ and } Z \in \mathbb{R}^{2(p-m)\times 2m}.$$

Equation (3.6) implies that $0 = Y \left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]$ and therefore $Y = 0$. Moreover, letting $Z = \left[\begin{smallmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{smallmatrix}\right]$ with appropriate blocks, we have

$$\begin{bmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{bmatrix} Z = Z \begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}$$

$$\iff \begin{cases} \Lambda_k \sin(\Theta_k) Z_{21} = -\sigma Z_{12}, \\ \Lambda_k \sin(\Theta_k) Z_{12} = -\sigma Z_{21}, \end{cases} \text{ and, } \begin{cases} \Lambda_k \sin(\Theta_k) Z_{11} = \sigma Z_{22}, \\ \Lambda_k \sin(\Theta_k) Z_{22} = \sigma Z_{11}. \end{cases} \tag{3.7}$$

Since all diagonal entries of $\Lambda_k \sin(\Theta_k)$ are nonzero and distinct from σ, (3.7) is possible if and only if $Z = 0$. Finally, since $\breve{Q} \in \mathrm{O}(n)$ and $V \in \mathrm{St}(n, 2m)$, we have $\breve{Q}^\top V \in \mathrm{St}(n, 2m)$ and thus $X \in \mathrm{O}(2m)$. Letting $R = X^\top$, we have $\breve{Q}_{1:2m} = VR$ and this concludes the proof. ∎

[3] There is a permutation matrix P such that if $A = QSQ^\top$ is an RSD from (2.2), then $\breve{Q} = QP$ and $\breve{S} = P^\top SP$.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Lemmas 3.1 and 3.2 show that the matrix R such that $\widehat{Q}R$ is a matrix of Schur vectors acts independently on three sets of columns of $\widehat{Q}$. This fact is the basis for the main structure of the algorithm: the separate computation of the three sets of eigenvalues and eigenvectors, as detailed in the next subsections.

3.3. The eigenvalues with distinct nonzero imaginary parts. In this subsection, we show in Theorem 3.3 that the Schur vectors of A for eigenvalues with distinct nonzero imaginary parts can be obtained by computing the Schur vectors of the skew-symmetric part Ω. Intermediate lemmas are given in Appendix A.

THEOREM 3.3. *Let $A \in \mathbb{R}^{n\times n}$ be a normal matrix. If $\Lambda_k \sin(\Theta_k)$ is diagonal and contains k distinct nonzero imaginary parts of the non-repeated eigenvalues of* $\mathrm{skew}(A)$*, then, for all $V \in \mathrm{St}(n, 2k)$ such that*

$$\mathrm{skew}(A)V = V \begin{bmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{bmatrix},$$

we have

$$AV = V \begin{bmatrix} \Lambda_k \cos(\Theta_k) & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & \Lambda_k \cos(\Theta_k) \end{bmatrix} =: VS_{2k}.$$

If, additionally, $k = \frac{n}{2}$, then $A = VS_nV^\top$ is a real Schur decomposition (2.2).

Proof. Let $A = QSQ^\top$ be an RSD and take a permutation matrix P such that $(P^\top SP)_{1:2k,1:2k} = S_{2k}$. By Lemma 3.2, it follows that $QP = [VR \ \ (QP)_{(k+1):n}]$ for some $R \in \mathrm{O}(2k)$ and this yields

$$\begin{aligned}
\begin{bmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{bmatrix} &= V^\top \mathrm{skew}(A) V && \text{By assumption.} \\
&= V^\top Q \mathrm{skew}(S) Q^\top V \\
&= V^\top (QP)(P^\top \mathrm{skew}(S) P)(QP)^\top V \\
&= [R \ 0_{2k\times(n-2k)}](P^\top \mathrm{skew}(S)P)[R \ 0_{2k\times(n-2k)}]^\top \\
&= R \begin{bmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{bmatrix} R^\top.
\end{aligned}$$

Therefore, we have $R \in \mathrm{ig}\left(\left[\begin{smallmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{smallmatrix}\right]\right)$. Leveraging Lemma A.2, there is a diagonal matrix $\Phi \in \mathbb{R}^{k\times k}$ such that $R = \left[\begin{smallmatrix} \cos(\Phi) & -\sin(\Phi) \\ \sin(\Phi) & \cos(\Phi) \end{smallmatrix}\right]$. Moreover, it is easily verified that every such matrix R also belongs to $\mathrm{ig}\left(\left[\begin{smallmatrix} \Lambda_k \cos(\Theta_k) & 0 \\ 0 & \Lambda_k \cos(\Theta_k) \end{smallmatrix}\right]\right)$. This yields

$$R \in \begin{bmatrix} 0 & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & 0 \end{bmatrix} \subset \begin{bmatrix} \Lambda_k \cos(\Theta_k) & 0 \\ 0 & \Lambda_k \cos(\Theta_k) \end{bmatrix}.$$

Finally, it follows by Lemma 2.4 that

$$R \in \mathrm{ig}\left(\begin{bmatrix} \Lambda_k \cos(\Theta_k) & -\Lambda_k \sin(\Theta_k) \\ \Lambda_k \sin(\Theta_k) & \Lambda_k \cos(\Theta_k) \end{bmatrix}\right) =: \mathrm{ig}(S_{2k}).$$

Consequently, since $QP = [VR \ \ (QP)_{(k+1):n}]$, we have

$$AV = (QP)(P^\top SP)(QP)^\top V = \begin{bmatrix} VR & (QP)_{(k+1):n} \end{bmatrix} \begin{bmatrix} S_{2k} & 0 \\ 0 & * \end{bmatrix} \begin{bmatrix} R^\top \\ 0_{(n-2k)\times 2k} \end{bmatrix} = VRS_{2k}R^\top = VS_{2k}.$$

When $2k = n$, $V \in \mathrm{O}(n)$ and the RSD is readily obtained since $S_n = S$ and $A = VSV^\top$. ∎

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

The conclusion of Theorem 3.3 is that for all eigenvalues that have distinct, nonzero imaginary parts, the invariant subspaces of $\mathrm{skew}(A)$ are always invariant subspaces of A. However, this does not hold true for Set 2 and Set 3 and this implies additional computations, as shown next.

3.4. The eigenvalues with repeated nonzero imaginary parts. Assume there is a diagonal entry σ in $\Lambda\sin(\Theta)$ that has multiplicity $m > 1$. Then, the associated $2m \times 2m$ block in S can be isolated using a permutation P such that

$$(P^\top SP)_{1:2m,1:2m} = \begin{bmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{bmatrix}, \tag{3.8}$$

where D_m is a diagonal submatrix of $\Lambda\cos(\Theta)$. The proof of Theorem 3.3 relies on the following inclusion of invariant groups:

$$\mathrm{ig}\left(\begin{bmatrix} 0 & -\Lambda_k\sin(\Theta_k) \\ \Lambda_k\sin(\Theta_k) & 0 \end{bmatrix}\right) \subset \mathrm{ig}\left(\begin{bmatrix} \Lambda_k\cos(\Theta_k) & 0 \\ 0 & \Lambda_k\cos(\Theta_k) \end{bmatrix}\right).$$

Unfortunately, Lemma A.3 shows that the block $\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}$ admits a larger invariance group than when the eigenvalues have distinct imaginary parts, i.e., if Σ_m is an $m\times m$ diagonal matrix with distinct nonzero entries, then $\mathrm{ig}\left(\begin{bmatrix} 0 & -\Sigma_m \\ \Sigma_m & 0 \end{bmatrix}\right) \subset \mathrm{ig}\left(\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}\right)$. However, if $D_m \neq dI_m$, then $\mathrm{ig}\left(\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}\right) \not\subseteq \mathrm{ig}\left(\begin{bmatrix} D_m & 0 \\ 0 & D_m \end{bmatrix}\right)$. This relation is illustrated in Fig. 1.

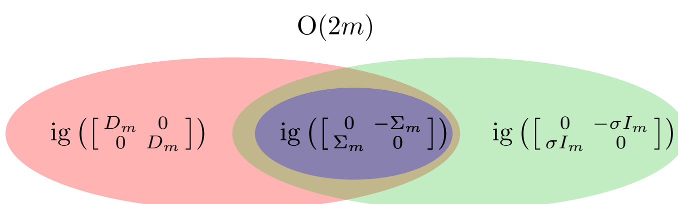

FIGURE 1. *Venn diagram of the sets* $\mathrm{ig}(\begin{bmatrix} D_m & 0 \\ 0 & D_m \end{bmatrix})$, $\mathrm{ig}(\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix})$, $\mathrm{ig}(\begin{bmatrix} 0 & -\Sigma_m \\ \Sigma_m & 0 \end{bmatrix})$, *where* D_m *and* Σ_m *are diagonal with nonzero entries and* $\sigma > 0$.

As shown in Theorem 3.4, it follows from Fig. 1 that if $V \in \mathrm{St}(n, 2m)$ spans the invariant subspace of $\mathrm{skew}(A)$ associated to σ, then $V^\top AV$ is not a real Schur form. Instead, $V^\top AV$ reveals a very structured *symmetric skew-Hamiltonian* matrix.

THEOREM 3.4. *For every normal matrix* $A \in \mathbb{R}^{n\times n}$ *that has* $m > 1$ *eigenvalues with imaginary part* $\sigma > 0$, *for all* $V \in \mathrm{St}(n, 2m)$ *such that*

$$\mathrm{skew}(A)V = V\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix},$$

we have

$$AV = V\begin{bmatrix} \widetilde{H} & -\widetilde{\Omega} - \sigma I_m \\ \widetilde{\Omega} + \sigma I_m & \widetilde{H} \end{bmatrix},$$

where $\widetilde{H} \in \mathbb{R}^{m\times m}$ *is symmetric and* $\widetilde{\Omega} \in \mathbb{R}^{m\times m}$ *is skew-symmetric.*

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Proof. Let $A = \breve{Q}\breve{S}\breve{Q}^\top$ be an RSD with the structure of (3.5). By Lemma 3.2, there is $R \in \mathrm{O}(2m)$ such that $\breve{Q}_{1:2m} = VR$. Since $\breve{Q}\mathrm{skew}(\breve{S})\breve{Q}^\top V = V \left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]$, it follows that

$$R \begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix} R^\top = \begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}. \tag{3.9}$$

Hence, $R \in \mathrm{ig}\left(\left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]\right)$. In Lemma A.3, we recall the following equivalence:

$$R \in \mathrm{ig}\left(\begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix}\right) \iff R \in \mathrm{O}(2m) \text{ and } R = \begin{bmatrix} E & -F \\ F & E \end{bmatrix}, \tag{3.10}$$

where $E, F \in \mathbb{R}^{m\times m}$. Recall that such matrices are called *ortho-symplectic*, see, e.g., [32, Eq. 5.1]. In view of (3.10), it holds that

$$\begin{aligned} AV &= \mathrm{sym}(A)V + \mathrm{skew}(A)V \\ &= \breve{Q}\mathrm{sym}(\breve{S})\breve{Q}^\top V + \mathrm{skew}(A)V \qquad (3.11) \\ &= VR \begin{bmatrix} D_m & 0 \\ 0 & D_m \end{bmatrix} R^\top + V \begin{bmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{bmatrix} \\ &= V \begin{bmatrix} ED_mE^\top + FD_mF^\top & (FD_mE^\top - ED_mF^\top)^\top - \sigma I_m \\ FD_mE^\top - ED_mF^\top + \sigma I_m & ED_mE^\top + FD_mF^\top \end{bmatrix}. \qquad (3.12) \end{aligned}$$

Defining $\widetilde{H} \coloneqq ED_mE^\top + FD_mF^\top$ and $\widetilde{\Omega} \coloneqq FD_mE^\top - ED_mF^\top$ concludes the proof. ∎

Let us consider again Q, the Schur vectors of A, $\widehat{Q}$, the Schur vectors of $\mathrm{skew}(A)$, and the permutation matrix P from (3.8). In view of Lemma 3.2, $(QP)_{1:2m} = (\widehat{Q}P)_{1:2m}R$ with $R \in \mathrm{O}(2m)$ and Theorem 3.4 holds with $V = (\widehat{Q}P)_{1:2m}$. Therefore, for each diagonal entry σ in $\Lambda\sin(\Theta)$ that has multiplicity $m > 1$, it follows from Theorem 3.4 that we must compute $V^\top AV$ and solve an additional $2m \times 2m$ EVP given by

$$V^\top AV = \begin{bmatrix} \widetilde{H} & -\widetilde{\Omega} - \sigma I_m \\ \widetilde{\Omega} + \sigma I_m & \widetilde{H} \end{bmatrix} = R \begin{bmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{bmatrix} R^\top, \tag{3.13}$$

where $R \in \mathrm{O}(2m)$ and $D_m \in \mathbb{R}^{m\times m}$ are to be computed. Solving (3.13) is enough to obtain the correct Schur vectors by $(QP)_{1:2m} \coloneqq VR$.

1. If $D_m = dI_m$, $d \in \mathbb{R}$, then it is the unique case where $\mathrm{ig}\left(\left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]\right) \subset \mathrm{ig}\left(\left[\begin{smallmatrix} dI_m & 0 \\ 0 & dI_m \end{smallmatrix}\right]\right)$. Consequently, $\widetilde{H} = dI_m$ and $\widetilde{\Omega} = 0$. No additional work is needed, one can assume $R = I_{2m}$ without loss of generality.
2. If D_m is not a multiple of the identity, we can recover an ortho-symplectic eigenvector matrix R from the *symmetric skew-Hamiltonian* EVP of $\left[\begin{smallmatrix} \widetilde{H} & -\widetilde{\Omega} \\ \widetilde{\Omega} & \widetilde{H} \end{smallmatrix}\right]$. Several methods have been designed for Hamiltonian and skew-Hamiltonian EVPs, see, e.g., [12, 76, 13, 70, 35], notably a QR algorithm relying exclusively on ortho-symplectic similarity transformations [70]. For completeness, we briefly discuss in Appendix B how the m first steps of Lanczos tridiagonalization, with restarting and reorthogonalization for stability, are enough to construct an ortho-symplectic matrix Y such that

$$\begin{bmatrix} \widetilde{H} & -\widetilde{\Omega} \\ \widetilde{\Omega} & \widetilde{H} \end{bmatrix} = Y \begin{bmatrix} T_m & 0 \\ 0 & T_m \end{bmatrix} Y^\top = R \begin{bmatrix} D_m & 0 \\ 0 & D_m \end{bmatrix} R^\top, \tag{3.14}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

where T_m is an $m \times m$ symmetric tridiagonal matrix. Consequently, once the EVP of $T_m = ZD_mZ^\top$ is solved with a call to Routine 2, we obtain $R \coloneqq Y \left[\begin{smallmatrix} Z & 0 \\ 0 & Z \end{smallmatrix}\right]$.
As studied in section 4, in practice, imaginary parts of the eigenvalues will not be exactly repeated but rather clustered and therefore, the symmetric skew-Hamiltonian structure cannot be guaranteed. This jeopardizes the attractive computational route based on (3.14). In order to maintain the accuracy of the global RSD, it is more robust to call Routine 3 to solve (3.13) directly. This choice has no influence on the computational complexity if m remains small, which is expected.

3.5. The real eigenvalues. We conclude this section by considering the computation of the r real eigenvalues of A and the associated eigenvectors. In Theorem 3.5, we show that an additional $r \times r$ symmetric EVP needs to be solved to obtain the eigenvectors of A.

THEOREM 3.5. *For every normal matrix $A \in \mathbb{R}^{n\times n}$ that has $r > 0$ real eigenvalues and for all $V \in \mathrm{St}(n,r)$ such that $\mathrm{skew}(A)V = 0$, we have*

$$AV = VH,$$

where the matrix $H \in \mathbb{R}^{r\times r}$ is symmetric.

Proof. For every RSD $A = QSQ^\top$ where S follows the structure of (2.2), we know by Lemma 3.1 that $Q = [Q_{1:p} \mid V\breve{R} \mid Q_{(p+r+1):n}]$ for some $\breve{R} \in \mathrm{O}(r)$. Therefore, it holds that

$$\begin{aligned} AV &= QSQ^\top V \\ &= Q \begin{bmatrix} \Lambda\cos(\Theta) & 0 & -\Lambda\sin(\Theta) \\ 0 & \breve{\Lambda} & 0 \\ \Lambda\sin(\Theta) & 0 & \Lambda\cos(\Theta) \end{bmatrix} \begin{bmatrix} 0_{p\times r} \\ \breve{R}^\top \\ 0_{p\times r} \end{bmatrix} \\ &= Q \begin{bmatrix} 0_{p\times r} \\ \breve{\Lambda}\breve{R}^\top \\ 0_{p\times r} \end{bmatrix} \\ &= V\breve{R}\breve{\Lambda}\breve{R}^\top. \end{aligned}$$

Defining the symmetric matrix $H \coloneqq \breve{R}\breve{\Lambda}\breve{R}^\top$ completes the proof. ∎

3.6. The algorithm in exact arithmetic. We have gathered all pieces to design an algorithm in exact arithmetic. First, we obtain $\widehat{Q}$ from RSD of $\Omega = \mathrm{skew}(A)$ (3.2). By Theorem 4.4, the invariant spaces associated to Set 1 need no further action. For each invariant space of Ω associated to Set 2, i.e., a complex eigenvalue with a repeated imaginary part, an additional symmetric skew-Hamiltonian EVP needs to be solved. This gives the transformation $[Q_{\mathrm{c},1}\ Q_{\mathrm{c},2}] = [\widehat{Q}_{\mathrm{c},1}\ \widehat{Q}_{\mathrm{c},2}] \left[\begin{smallmatrix} R_\mathrm{a} & R_\mathrm{b} \\ R_\mathrm{c} & R_\mathrm{d} \end{smallmatrix}\right]$ from (3.4). Finally, by Theorem 4.7, the decomposition $H \coloneqq \widehat{Q}_\mathrm{r}^\top A \widehat{Q}_\mathrm{r} = \breve{R}\breve{\Lambda}\breve{R}^\top$ using a call to Routine 2 completes the determination of the RSD

$$A = \widehat{Q} \begin{bmatrix} R_\mathrm{a} & 0 & R_\mathrm{b} \\ 0 & \breve{R} & 0 \\ R_\mathrm{c} & 0 & R_\mathrm{d} \end{bmatrix} \begin{bmatrix} \Lambda\cos(\Theta) & 0 & -\Lambda\sin(\Theta) \\ 0 & \breve{\Lambda} & 0 \\ \Lambda\sin(\Theta) & 0 & \Lambda\cos(\Theta) \end{bmatrix} \begin{bmatrix} R_\mathrm{a} & 0 & R_\mathrm{b} \\ 0 & \breve{R} & 0 \\ R_\mathrm{c} & 0 & R_\mathrm{d} \end{bmatrix}^\top \widehat{Q}^\top.$$

The complete method in exact arithmetic is summarized in Algorithm 3.1. In a broad set of problems, steps 5 and 7 of Algorithm 3.1 are either not necessary or very simple. Indeed, for Haar-distributed random

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

orthogonal matrices [7, 64], it holds with probability 1 that $r \in \{0,1\}$ if $\det(A) = 1$ and $r \in \{1,2\}$ if $\det(A) = -1$ [34]. In particular, if $r = 1$, $H = \pm 1$, and $\breve{R} = 1$. Step 7 of Algorithm 3.1 is thus trivial. Moreover, all complex eigenvalues have distinct imaginary parts with probability 1 and thus step 5 is skipped. For general normal matrices, the distribution of the eigenvalues, and in particular the value of r, is specific to the context of the problem. Nonetheless, the case where all or a large majority of the eigenvalues belong to Set 1 is common.

Algorithm 3.1 Exact arithmetic algorithm for the real Schur decomposition of a normal matrix

Input: A normal matrix $A \in \mathbb{R}^{n\times n}$.
Output: Q and S where $Q \in \mathrm{O}(n)$ and $A = QSQ^\top$ is an RSD.
step 1: Compute the skew-symmetric part $\Omega \coloneqq \frac{1}{2}(A - A^\top)$.
step 2.1: Compute a tridiagonal reduction $\Omega = \widetilde{Q}T\widetilde{Q}^\top$.
step 2.2: Apply the permutation $\begin{bmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{bmatrix} \coloneqq P_{\mathrm{eo}}^\top T P_{\mathrm{eo}}$ where $\widetilde{B} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor \times \lceil \frac{n}{2} \rceil}$ is bidiagonal.
step 3: Compute the (permuted) SVD of $\widetilde{B} = U \begin{bmatrix} 0 & 0_{\lfloor \frac{r}{2} \rfloor \times \lceil \frac{r}{2} \rceil} \\ \Sigma & 0 \end{bmatrix} V^\top$ where $\mathrm{diag}(\Sigma) > 0$.
step 4: Define $\widehat{Q} \coloneqq \widetilde{Q}P_{\mathrm{eo}}\left[\begin{smallmatrix} V & 0 \\ 0 & U \end{smallmatrix}\right]$, $\Lambda\sin(\Theta) \coloneqq \Sigma$ and $p \coloneqq \frac{n-r}{2}$.
for each nonzero diagonal entry σ of Σ with multiplicity $m > 1$ **do**
  step 5.1: Define V as the $2m$ columns of $\widehat{Q}$ such that $\Omega V = V \left[\begin{smallmatrix} 0 & -\sigma I_m \\ \sigma I_m & 0 \end{smallmatrix}\right]$.
  step 5.2: Compute the matrix $V^\top AV \eqqcolon \begin{bmatrix} \widetilde{H} & -\widetilde{\Omega}-\sigma I_m \\ \widetilde{\Omega}+\sigma I_m & \widetilde{H} \end{bmatrix}$ where $\widetilde{H} = \widetilde{H}^\top$ and $\widetilde{\Omega} = -\widetilde{\Omega}^\top$.
  step 5.3: Compute the RSD $\begin{bmatrix} \widetilde{H} & -\widetilde{\Omega}-\sigma I_m \\ \widetilde{\Omega}+\sigma I_m & \widetilde{H} \end{bmatrix} = R \left[\begin{smallmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{smallmatrix}\right] R^\top$.
  step 5.4: Replace the columns of $\widehat{Q}$ selected at step 5.1 by VR.
end for
step 6: Define the diagonal matrix $\Lambda\cos(\Theta) \coloneqq \widehat{Q}_{1:p}^\top A\widehat{Q}_{1:p}$.
if $r > 0$ **then**
  step 7.1: Define $V \coloneqq \widehat{Q}_{(p+1):(p+r)}$ such that $\Omega V = 0$.
  step 7.2: Compute the symmetric matrix $H \coloneqq V^\top AV$.
  step 7.3: Compute the EVD of $H = \breve{R}\breve{\Lambda}\breve{R}^\top$.
  step 7.4: Replace $\widehat{Q}_{(p+1):(p+r)}$ by $V\breve{R}$.
end if
return $Q \coloneqq \widehat{Q}$ and $S \coloneqq \begin{bmatrix} \Lambda\cos(\Theta) & 0 & -\Lambda\sin(\Theta) \\ 0 & \breve{\Lambda} & 0 \\ \Lambda\sin(\Theta) & 0 & \Lambda\cos(\Theta) \end{bmatrix}$.

3.7. The eigenvalue decomposition by the symmetric part. Algorithm 3.1 shows that the RSD of real normal matrix can be obtained from its skew-symmetric part. It is natural to ask if this approach can be adapted to start from the *symmetric part*. The answer is positive, but several drawbacks arise compared to the skew-symmetric approach.

- Drawback 1: The symmetric part does not allow separating real from complex eigenvalues easily. Indeed, given an EVD of $\mathrm{sym}(A) = V\Lambda V^\top$, one cannot distinguish a real eigenvalue of A from the real part of the complex eigenvalue, except if its multiplicity is 1. Making this separation requires computing the product $V^\top AV$. In Algorithm 3.1, there is a direct correspondence between the real eigenspace of A and the kernel of $\mathrm{skew}(A)$.
- Drawback 2: The aforementioned separation between real and complex eigenvalues cannot be obtained if some real parts of complex eigenvalues are equal to real eigenvalues of A. Then, the

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

associated sub-block of $V^\top AV$ is a normal matrix with no additional structure.

- Drawback 3: The symmetric eigensolver for $\mathrm{sym}(A)$ does not take advantage of the even multiplicity of the real parts of complex eigenvalues, while the skew-symmetric part uses this advantage by computing a half-the-size bidiagonal SVD.
- Drawback 4: For every two-dimensional invariant subspace of A, the problem arises that

$$\mathrm{ig}\left(\begin{bmatrix}\lambda\cos(\theta) & -\lambda\sin(\theta)\\ \lambda\sin(\theta) & \lambda\cos(\theta)\end{bmatrix}\right) = \mathrm{SO}(2) \subset \mathrm{ig}\left(\begin{bmatrix}\lambda\cos(\theta) & 0\\ 0 & \lambda\cos(\theta)\end{bmatrix}\right) = \mathrm{O}(2).$$

  Since an element of $\mathrm{O}(2)$ belongs to $\mathrm{SO}(2)$ with probability $\frac{1}{2}$, there is a probability $\frac{1}{2}$ for each base to be incorrectly oriented.

The aforementioned drawbacks matter if a significant proportion of the eigenvalues are complex. If most eigenvalues are real, then, the approach by the symmetric part can be efficient. However, applications, e.g., with orthogonal matrices, are usually dominated by complex eigenvalues. This makes the approach by the skew-symmetric-part more relevant for applications to the best of the authors' knowledge.

3.8. Exact versus floating point arithmetic: an example. We have not yet adressed an essential algorithmic question: how accurate is Algorithm 3.1 in floating point arithmetic and how should it be implemented in order to always obtain a desired accuracy? Indeed, floating point arithmetic brings an additional challenge. In some cases, the invariant subspaces of Ω are less sensitive to perturbations of A than the invariant subspaces of A themselves, and, if not analyzed and addressed in the design of the algorithm, this can be a source of inaccuracies. We show next an illustrative example.

Consider a variable $\tau \geq 0$, a $p \times p$ random symmetric matrix E, $\|E\|_2 = 1$, and the normal matrix $A(\tau)$ defined by

$$A(\tau) = \begin{bmatrix}\sin(\tau E) & -\cos(\tau E)\\ \cos(\tau E) & \sin(\tau E)\end{bmatrix} \quad \text{such that,} \quad A(0) = \begin{bmatrix}0 & -I_p\\ I_p & 0\end{bmatrix}.$$

Notice that $A(0) = \Omega(0)$. For τ small enough, Taylor series yield $\cos(\tau E) \approx I_p - \frac{\tau^2}{2}E^2$ and $\sin(\tau E) \approx \tau E$ such that, using the spectral norm $\|\cdot\|_2$, we have

$$\|A(\tau) - A(0)\|_2 \approx \tau \quad \text{and,} \quad \|\Omega(\tau) - \Omega(0)\|_2 \approx \frac{\tau^2}{2}.$$

Around $\tau = 0$, it follows that $\Omega(\tau)$ is less sensitive to small variations of τ than $A(\tau)$. Let ε_{m} denote the machine precision. Then, applying Algorithm 3.1 on $A(\sqrt{\varepsilon_{\mathrm{m}}})$ or $A(0)$ will yield similar Schur vectors in output since, in floating point arithmetic, $\Omega(\sqrt{\varepsilon_{\mathrm{m}}}) \approx \Omega(0)$. Therefore, $\mathcal{O}(\sqrt{\varepsilon_{\mathrm{m}}})$-errors will be made in the computation of the Schur vectors of $A(\sqrt{\varepsilon_{\mathrm{m}}})$. This is verified in practice, here for $p = 100$ and $\varepsilon_{\mathrm{m}} = 2.22 \cdot 10^{-16}$, we have

$$\text{Naive implementation of Algorithm 3.1:} \quad \|A(\sqrt{\varepsilon_{\mathrm{m}}})\widehat{Q} - \widehat{Q}\widehat{S}\|_{\mathrm{F}} = 6.52 \cdot 10^{-9}\|A(\sqrt{\varepsilon_{\mathrm{m}}})\|_{\mathrm{F}}.$$

For methods relying on the diagonalization of commuting matrices, it is well known that sensitivity analyses depend on eigenvalue gaps. In [47], which studies the simultaneous diagonalization of symmetric matrices, a probabilistic accuracy bound is derived in terms of the smallest eigenvalue gap among the family of matrices. Similarly, the do-one-then-do-the-other algorithm of [66, Alg. 2] requires a refinement of the invariant subspaces that is governed by the eigenvalue gaps. As shown in section 4, the specific setting

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

of real normal matrices, together with Algorithm 4.1, allows for more precise statements, again involving eigenvalue gaps.

In the next section, we study how the loss of accuracy occurs and how Algorithm 3.1 should be implemented in order to obtain a desired accuracy μ.

4. Analysis of numerical stability and accuracy. In this section, we analyze the effect of floating point arithmetic on the accuracy of Algorithm 3.1. We rely on standard assumptions of floating point arithmetic. In particular, ε_{m} denotes the machine precision and, for every scalar $\alpha \in \mathbb{C}$, $\mathrm{fl}(\alpha) = \alpha(1+\Delta\alpha)$ with $|\Delta\alpha| \leq \varepsilon_{\mathrm{m}}$ models its float representation. The *absolute gap* between two numbers $\alpha, \widehat{\alpha} \in \mathbb{C}$ is defined as $|\alpha - \widehat{\alpha}|$ while the *relative gap* refers to $\frac{|\alpha-\widehat{\alpha}|}{|\alpha|+|\widehat{\alpha}|}$. The notation $\alpha \lesssim \beta$ means that there is a moderately growing function $\varphi(n)$, e.g., n or n^2, independent of the values of α and β, such that $\alpha \leq \varphi(n)\beta$.

We will analyze step by step the numerical properties of the algorithm. This will determine how to implement Algorithm 3.1 to obtain high accuracy on both the eigenvalues and the Schur vectors/invariant spaces. The floating point arithmetic version of the method is given in Algorithm 4.1. The algorithm requires to identify clusters of eigenvalues. A definition for a cluster of eigenvalues, justified by Theorems 4.4, 4.6 and 4.7, is given by Definition 4.1.

Definition 4.1. *Given $\delta > 0$, a matrix A, a set of $2m > 0$ eigenvalues $\{\pm i\sigma_j\}_{j=1}^m$ of $\Omega = \mathrm{skew}(A)$ with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_m \geq 0$ is a δ*-cluster *of Ω if and only if*

$$|\sigma_j - \sigma_{j+1}| \leq \delta\|A\|_{\mathrm{F}} \quad \textit{and} \quad |\sigma_k - \sigma| > \delta\|A\|_{\mathrm{F}} \quad \textit{for } j = 1,...,m-1, \textit{ and, } k = 1,...,m,$$

where $i\sigma$ is any eigenvalue of Ω not in $\{\pm i\sigma_j\}_{j=1}^m$.

Definition 4.1 is similar to the clustering criterion used in [66, Alg. 2]. Instead of the Frobenius norm $\|A\|_{\mathrm{F}}$, the results of [66] employ the *spread* of the eigenvalues, i.e., the maximum distance between the eigenvalues. This difference in definitions reflects the distinct hypotheses and analytical frameworks underlying the respective results.

4.1. Analysis of steps 1 to 4. Step 1 already entails some subtleties since it consists of n^2 subtractions. Some rare pathological cases exist where this operation might be an initial source of inaccuracy. Indeed, letting $\Omega = \mathrm{skew}(A)$ and $\widehat{\Omega} = \mathrm{skew}(\mathrm{fl}(A))$ such that $\widehat{\omega}_{ij} = \frac{\mathrm{fl}(a_{ij})-\mathrm{fl}(a_{ji})}{2}$, we obtain

$$\left|\frac{\omega_{ij} - \widehat{\omega}_{ij}}{\omega_{ij}}\right| = \left|\frac{a_{ij}\Delta a_{ij} - a_{ji}\Delta a_{ji}}{a_{ij} - a_{ji}}\right| \leq \varepsilon_{\mathrm{m}} \frac{|a_{ij}| + |a_{ji}|}{|a_{ij} - a_{ji}|}. \tag{4.1}$$

Based on the elementwise relation (4.1) and acknowledging that $\mathrm{Tr}(|A|^2) \leq \|A\|_{\mathrm{F}}^2$ by Cauchy–Schwarz inequality, we can also deduce the more practical matrix relation

$$\frac{\|\Omega - \widehat{\Omega}\|_{\mathrm{F}}}{\|\Omega\|_{\mathrm{F}}} \leq \varepsilon_{\mathrm{m}} \frac{2\|A\|_{\mathrm{F}}}{\|\Omega\|_{\mathrm{F}}}. \tag{4.2}$$

In view of (4.1), whenever symmetrically positioned entries of A are *relatively* close, the numerical approximation $\widehat{\omega}_{ij}$ may become inaccurate. However, relative inaccuracies only occur for ω_{ij} close to zero. Therefore, if $\|\Omega\|_{\mathrm{F}} \gg \varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, then $\widehat{\Omega}$ is accurate in the sense of (4.2), even if some entries are not. An example of a matrix set where such entrywise inaccuracies occur in general with negligible effect is

$$\{U \in \mathrm{O}(2m) \mid U = \exp\left[\begin{smallmatrix} 0 & -M^\top \\ M & 0 \end{smallmatrix}\right],\ M \in \mathbb{R}^{m\times m}\}. \tag{4.3}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Algorithm 4.1 Floating point arithmetic algorithm for the real Schur decomposition of a normal matrix (`nrmschur`)

Input: A normal matrix $A \in \mathbb{R}^{n\times n}$, $\delta, \delta_r = \sqrt{\varepsilon_{\mathrm{m}}}$ and $\mu = \varepsilon_{\mathrm{m}} t$, $t \geq 1$.
Output: $\widehat{Q}$ and $\widehat{S}$ where $d_{\mathrm{O}(n)}(\widehat{Q}) \lesssim \varepsilon_{\mathrm{m}}$ and $\|A\widehat{Q} - \widehat{Q}\widehat{S}\|_{\mathrm{F}} \lesssim \mu\|A\|_{\mathrm{F}}$.
step 1: Compute the skew-symmetric part $\widehat{\Omega} \approx \frac{1}{2}(A - A^\top)$.
step 2.1: Tridiagonalize $\widehat{\Omega} \approx \widetilde{Q}T\widetilde{Q}^\top$.
step 2.2: Apply the permutation $\left[\begin{smallmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{smallmatrix}\right] := P_{\mathrm{eo}}^\top T P_{\mathrm{eo}}$ where $\widetilde{B} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor \times \lceil \frac{n}{2} \rceil}$ is bidiagonal.
step 3: (Routine 1) Compute the (permuted) SVD of $\widetilde{B} \approx \widehat{U} \left[\begin{smallmatrix} 0 & 0_{\lfloor \frac{\widehat{r}}{2} \rfloor \times \lceil \frac{\widehat{r}}{2} \rceil} \\ \widehat{\Sigma} & 0 \end{smallmatrix}\right] \widehat{V}^\top$ where $\mathrm{diag}(\widehat{\Sigma}) > 0$ and $\widehat{r}$ is the size of the δ_r-cluster of eigenvalues of $\widehat{\Omega}$ around 0.
step 4: Compute $\widehat{Q} \approx \widetilde{Q}P_{\mathrm{eo}}\left[\begin{smallmatrix} V & 0 \\ 0 & U \end{smallmatrix}\right]$, set $\Lambda\sin(\Theta) := \widehat{\Sigma}$ and $\widehat{p} := \frac{n-\widehat{r}}{2}$.
for each δ-cluster of $2m$ (nonzero) eigenvalues of $\widehat{\Omega}$ **do**
  step 5.1: Set V as the $2m$ columns of $\widehat{Q}$ associated to the cluster and compute $M \approx V^\top AV$.
  step 5.2: (Routine 3) Compute the RSD $M \approx R\widetilde{S}_M R^\top$.
  step 5.3: Replace the $2m$ columns of $\widehat{Q}$ from step 5.1 by VR.
end for
step 6: Set $\Lambda\cos(\Theta)$ as the diagonal of $\widehat{Q}_{1:\widehat{p}}^\top \mathrm{sym}(A)\widehat{Q}_{1:\widehat{p}}$.
if there is a δ_r-cluster of $\widehat{r}$ eigenvalues of Ω around 0, **then**
  step 7.1: Compute the matrix $H \approx \widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}^\top A\widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}$.
  if $\|\mathrm{skew}(H)\|_{\mathrm{F}} \leq \varepsilon_{\mathrm{m}}\|H\|_{\mathrm{F}}$, **then**
    step 7.2a: (Routine 2) Compute the EVD of $\mathrm{sym}(H) \approx \breve{R}\breve{\Lambda}\breve{R}^\top$.
  else
    step 7.2b: (Routine 3) Compute the RSD of $H \approx \breve{R}\breve{S}_H\breve{R}^\top$.
  end if
  step 7.3: Set $\widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}} \approx \widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}\breve{R}$.
end if
step 8: (Optional, see subsection 4.5) Subspaces correction steps. For each $\frac{1}{t}$-cluster of eigenvalues of $\widehat{\Omega}$, repeat step 5 with $\delta = \frac{1}{t}$ and consider a Jacobi algorithm [83] for Routine 3.
return $\widehat{Q}$ and $\widehat{S} := \left[\begin{smallmatrix} \Lambda\cos(\Theta) & 0 & -\Lambda\sin(\Theta) \\ 0 & \breve{\Lambda} & 0 \\ \Lambda\sin(\Theta) & 0 & \Lambda\cos(\Theta) \end{smallmatrix}\right]$.

The matrix set (4.3) notably arises in the fundamental Riemannian computation of geodesics on the Grassmann manifold [11]. Entrywise inaccuracies can also be neglected if $\|\Omega\|_{\mathrm{F}} \leq \varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$. In this case, the matrix A can be numerically considered as symmetric. In conclusion, computing a skew-symmetric part is dubious when $\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}} \ll \|\Omega\|_{\mathrm{F}} \ll \|A\|_{\mathrm{F}}$. However, it is shown in section 6 that Algorithm 4.1 can handle small skew-symmetric parts.

Step 2 is the tridiagonal reduction using Householder reflectors. It is known to be backward stable [41, p. 416], i.e.,

$$\widehat{\Omega} + \Delta\widehat{\Omega} = \widetilde{Q}P_{\mathrm{eo}}\begin{bmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{bmatrix} P_{\mathrm{eo}}^\top\widetilde{Q}^\top \text{ with } \|\Delta\widehat{\Omega}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|\widehat{\Omega}\|_{\mathrm{F}}, \tag{4.4}$$

where $\Delta\widehat{\Omega}$ is exactly skew-symmetric, $\widetilde{Q} \in \mathrm{O}(n)$ is the exact product of the Householder reflectors numerically computed during the reduction and P_{eo} is an even–odd permutation (2.5). Moreover, Householder reflectors can be assembled very accurately as $\widetilde{Q}_{\mathrm{asm}}$ such that $\|\widetilde{Q}_{\mathrm{asm}} - \widetilde{Q}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}$ [49, Eq. 19.13]. Defining the distance

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

to orthogonality of a matrix V by

$$d_{\mathrm{O}(n)}(V) := \min_{Q\in \mathrm{O}(n)} \|Q - V\|_{\mathrm{F}}, \tag{4.5}$$

we can write $d_{\mathrm{O}(n)}(\widetilde{Q}_{\mathrm{asm}}) \lesssim \varepsilon_{\mathrm{m}}$.

Step 3 is the bidiagonal SVD. Algorithms such as bdsqr provide a mixed forward–backward stable SVD of $\widetilde{B} \in \mathbb{R}^{\lfloor \frac{n}{2} \rfloor \times \lceil \frac{n}{2} \rceil}$:

$$\widetilde{B} + \Delta\widetilde{B} = \widehat{U}\widehat{\Sigma}\widehat{V}^\top \quad \text{and,} \quad \|\Delta\widetilde{B}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|\widetilde{B}\|_{\mathrm{F}},$$

such that $\max\{d_{\mathrm{O}(\lfloor \frac{n}{2} \rfloor)}(\widehat{U}), d_{\mathrm{O}(\lceil \frac{n}{2} \rceil)}(\widehat{V})\} \lesssim \varepsilon_{\mathrm{m}}$ [6, 28].

REMARK 4.2. The singular values of $\widetilde{B}$ are computed with high *relative* accuracy by algorithms such as bdsqr [28, Thm. 6 and 7]. This holds because relatively small perturbations of $\widetilde{B}$ yield relatively perturbed singular values [30, Chap. 4.3]. Moreover, it was conjectured [28] and proved [26, Thm. 6.1] that left and right singular vectors of $\widetilde{B}$ are also very accurately computed. However, the Bauer–Fike theorem [51, Thm. 6.3.2] and its corollary [51, Cor. 6.3.4] indicate that, due to the perturbation $\Delta\widehat{\Omega}$ in (4.4), relative accuracy of the eigenvalues of Ω cannot be achieved. Consequently, the relative accuracy of the singular values of $\widetilde{B}$ is not relevant.

Step 4 is a product of matrices. Define $\widehat{Q}_{\mathrm{exact}}$ as the *exact* matrix product of $\widetilde{Q}_{\mathrm{asm}}P_{\mathrm{eo}}$ by $\left[\begin{smallmatrix} \widehat{V} & 0 \\ 0 & \widehat{U} \end{smallmatrix}\right]$. Since the permutation is performed exactly, w.l.o.g., we assume $P_{\mathrm{eo}} = I_n$. We obtain that

$$\begin{aligned} d_{\mathrm{O}(n)}(\widehat{Q}_{\mathrm{exact}}) &\le \|\widetilde{Q}_{\mathrm{asm}} \left[\begin{smallmatrix} \widehat{V} & 0 \\ 0 & \widehat{U} \end{smallmatrix}\right] - \widetilde{Q} \left[\begin{smallmatrix} V & 0 \\ 0 & U \end{smallmatrix}\right] \|_{\mathrm{F}} \\ &= \|\widetilde{Q}_{\mathrm{asm}} \left[\begin{smallmatrix} \widehat{V}-V+V & 0 \\ 0 & \widehat{U}-U+U \end{smallmatrix}\right] - \widetilde{Q} \left[\begin{smallmatrix} V & 0 \\ 0 & U \end{smallmatrix}\right] \|_{\mathrm{F}} \\ &\le \|\widetilde{Q}_{\mathrm{asm}} - \widetilde{Q}\|_{\mathrm{F}} + \|\widetilde{Q}_{\mathrm{asm}} \left[\begin{smallmatrix} \widehat{V}-V & 0 \\ 0 & \widehat{U}-U \end{smallmatrix}\right] \|_{\mathrm{F}} \\ &\lesssim \varepsilon_{\mathrm{m}}. \end{aligned}$$

Finally, it holds by [49, Eq. 3.13] that $\|\widehat{Q} - \widehat{Q}_{\mathrm{exact}}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|\widetilde{Q}_{\mathrm{asm}}\|_{\mathrm{F}}\| \left[\begin{smallmatrix} V & 0 \\ 0 & U \end{smallmatrix}\right] \|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}$. Therefore, by the triangular inequality, we have $d_{\mathrm{O}(n)}(\widehat{Q}) \lesssim \varepsilon_{\mathrm{m}}$. By summing all errors from steps 1 to 4, we obtain the skew-symmetric error matrix $\Delta_2\widehat{\Omega}$:

$$\begin{aligned} \Delta_2\widehat{\Omega} := \Delta\widehat{\Omega} &+ \widetilde{Q} \left[\begin{smallmatrix} 0 & -\Delta\widetilde{B}^\top \\ \Delta\widetilde{B} & 0 \end{smallmatrix}\right] \widetilde{Q}^\top \\ &+ (\widetilde{Q}_{\mathrm{asm}} - \widetilde{Q}) \left[\begin{smallmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{smallmatrix}\right] \widetilde{Q}^\top - \widetilde{Q} \left[\begin{smallmatrix} 0 & -\widetilde{B}^\top \\ \widetilde{B} & 0 \end{smallmatrix}\right]^\top (\widetilde{Q}_{\mathrm{asm}} - \widetilde{Q})^\top \\ &+ (\widehat{Q} - \widehat{Q}_{\mathrm{exact}}) \left[\begin{smallmatrix} 0 & -\widehat{\Sigma} \\ \widehat{\Sigma} & 0 \end{smallmatrix}\right] \widehat{Q}^\top - \widehat{Q} \left[\begin{smallmatrix} 0 & -\widehat{\Sigma} \\ \widehat{\Sigma} & 0 \end{smallmatrix}\right]^\top (\widehat{Q} - \widehat{Q}_{\mathrm{exact}})^\top + \mathcal{O}(\varepsilon_{\mathrm{m}}^2). \end{aligned}$$

$\mathcal{O}(\varepsilon_{\mathrm{m}}^2)$ gathers terms where, e.g., $\widetilde{Q}_{\mathrm{asm}} - \widetilde{Q}$ appears twice. If we consider $(\widehat{Q}, \widehat{\Sigma})$ as the output of the skew-symmetric RSD, this yields the mixed forward–backward stable decomposition of the skew-symmetric part $\widehat{\Omega}$:

$$\widehat{\Omega} + \Delta_2\widehat{\Omega} = \widehat{Q} \begin{bmatrix} 0 & -\widehat{\Sigma} \\ \widehat{\Sigma} & 0 \end{bmatrix} \widehat{Q}^\top \quad \text{with} \quad \|\Delta_2\widehat{\Omega}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|\widehat{\Omega}\|_{\mathrm{F}} \quad \text{and} \quad d_{\mathrm{O}(n)}(\widehat{Q}) \lesssim \varepsilon_{\mathrm{m}}. \tag{4.6}$$

Mixed forward–backward stability refers to the fact that $\widehat{Q}$ is near orthogonal. In this context, the decomposition would be said *backward* stable if $\widehat{Q}$ was exactly orthogonal. As shown in section 5, $d_{\mathrm{O}(n)}(\widehat{Q})$ is very close to machine precision in practice.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

4.2. Sensitivity of the invariant subspaces: Set 1. Let A be a normal matrix with an exact RSD $A = QSQ^\top$. We first assume that the matrix A has no real eigenvalues ($r = 0, n = 2p$) and that all eigenvalues have distinct imaginary parts. Given a backward stable RSD of $\widehat{\Omega}$, i.e., $\widehat{\Omega}\widehat{Q} - \widehat{Q}\text{skew}(\widehat{S}) = \widehat{E}$ such that $\|\widehat{E}\|_\mathrm{F} \lesssim \varepsilon_\mathrm{m}\|\widehat{\Omega}\|_\mathrm{F}$, we quantify $\|A\widehat{Q} - \widehat{Q}S\|_\mathrm{F}$ in Theorem 4.4. We first assume that $\widehat{Q}$ is *exactly* orthogonal such that Theorem 4.4 focuses on the important conclusions. We address deviation from orthogonality in Corollary 4.5. This quantification is key for two points: first, characterizing how accurately eigenvalues are computed at step 6 by the Rayleigh quotients and second, designing correction steps to achieve a specified accuracy on the Schur vectors, if needed.

We start by showing in Lemma 4.3 that it is equivalent to consider the numerical output $\widehat{\Omega}\widehat{Q} - \widehat{Q}\text{skew}(\widehat{S}) = \widehat{E}$ with $\|\widehat{E}\|_\mathrm{F} \leq \widehat{\tau}\varepsilon_\mathrm{m}\|\widehat{\Omega}\|_\mathrm{F}$ from (4.6) and the more practical relation $\Omega\widehat{Q} - \widehat{Q}\text{skew}(S) = E$ with $\|E\|_\mathrm{F} \leq \tau\|A\|_\mathrm{F}$ and $\tau > 0$. In the next lemma, we assume that S and $\widehat{S}$ are organized such that $\text{skew}(S)$ and $\text{skew}(\widehat{S})$ have their eigenvalues sorted, e.g., in decreasing order.

Lemma 4.3. *Let $A = QSQ^\top$ be a normal matrix, $\Omega := \text{skew}(A)$ and $\widehat{\Omega} := \text{skew}(\text{fl}(A))$. If the orthogonal matrix $\widehat{Q} \in \mathrm{O}(2p)$ is such that $\widehat{\Omega}\widehat{Q} - \widehat{Q}\text{skew}(\widehat{S}) = \widehat{E}$ is a backward stable RSD of $\widehat{\Omega}$, namely, if $\|\widehat{E}\|_\mathrm{F} \leq \widehat{\tau}\varepsilon_\mathrm{m}\|\widehat{\Omega}\|_\mathrm{F}$ for $\widehat{\tau} > 0$, then it follows that*

$$\|\Omega\widehat{Q} - \widehat{Q}\text{skew}(S)\|_\mathrm{F} \leq (2\widehat{\tau} + 4)\varepsilon_\mathrm{m}\|A\|_\mathrm{F} + \mathrm{o}(\varepsilon_\mathrm{m}).$$

Proof. Let us relate $\widehat{Q}$ to Ω instead of $\widehat{\Omega}$.

$$\begin{aligned} &\widehat{\Omega}\widehat{Q} - \widehat{Q}\text{skew}(\widehat{S}) = \widehat{E} \\ \iff\ &\Omega\widehat{Q} - \widehat{Q}\text{skew}(S) = \widehat{E} - (\widehat{\Omega} - \Omega)\widehat{Q} - \widehat{Q}\text{skew}(S - \widehat{S}). \end{aligned}$$

Now let $\text{skew}(\widehat{S}_2)$ be the *exact* real Schur form of $\widehat{\Omega}$. Then, we have

$$\begin{aligned} \|\Omega\widehat{Q} - \widehat{Q}\text{skew}(S)\|_\mathrm{F} &\leq \|\widehat{E}\|_\mathrm{F} + \|(\widehat{\Omega} - \Omega)\widehat{Q}\|_\mathrm{F} + \|\widehat{Q}\text{skew}(S - \widehat{S})\|_\mathrm{F} \\ &= \|\widehat{E}\|_\mathrm{F} + \|\widehat{\Omega} - \Omega\|_\mathrm{F} + \|\text{skew}(S - \widehat{S})\|_\mathrm{F} \\ &\leq \|\widehat{E}\|_\mathrm{F} + \|\widehat{\Omega} - \Omega\|_\mathrm{F} + \|\text{skew}(S - \widehat{S}_2)\|_\mathrm{F} + \|\text{skew}(\widehat{S}_2 - \widehat{S})\|_\mathrm{F}. \end{aligned}$$

By [50, Thm. 1], we have $\|\text{skew}(S - \widehat{S}_2)\|_\mathrm{F} \leq \|\widehat{\Omega} - \Omega\|_\mathrm{F}$. Moreover, since $\widehat{\Omega} = \widehat{Q}\text{skew}(\widehat{S})\widehat{Q}^\top + \widehat{E}\widehat{Q}^\top$, again by [50, Thm. 1], we have $\|\text{skew}(\widehat{S}_2 - \widehat{S})\|_\mathrm{F} \leq \|\widehat{E}\widehat{Q}^\top\|_\mathrm{F} = \|\widehat{E}\|_\mathrm{F}$. By (4.2), we have $\|\widehat{\Omega} - \Omega\|_\mathrm{F} \leq 2\varepsilon_\mathrm{m}\|A\|_\mathrm{F}$ and $\|\widehat{E}\|_\mathrm{F} \leq \widehat{\tau}\varepsilon_\mathrm{m}\|\widehat{\Omega}\|_\mathrm{F} \leq \widehat{\tau}\varepsilon_\mathrm{m}(\|\Omega\|_\mathrm{F} + \varepsilon_\mathrm{m}2\|A\|_\mathrm{F})$. This yields

$$\begin{aligned} \|\Omega\widehat{Q} - \widehat{Q}\text{skew}(S)\|_\mathrm{F} &\leq 2\|\widehat{E}\|_\mathrm{F} + 2\|\widehat{\Omega} - \Omega\|_\mathrm{F} \\ &\leq 2\widehat{\tau}\varepsilon_\mathrm{m}(\|\Omega\|_\mathrm{F} + \varepsilon_\mathrm{m}2\|A\|_\mathrm{F}) + 2(2\varepsilon_\mathrm{m}\|A\|_\mathrm{F}) \\ &\leq (2\widehat{\tau} + 4)\varepsilon_\mathrm{m}\|A\|_\mathrm{F} + 4\widehat{\tau}\varepsilon_\mathrm{m}^2\|A\|_\mathrm{F}. \end{aligned}$$ ∎

We can now prove the stability result of Theorem 4.4.

Theorem 4.4. *Let $A = QSQ^\top \in \mathbb{R}^{2p\times 2p}$ be a normal matrix with eigenvalues $\lambda_j e^{\pm i\theta_j}$, $\lambda_j > 0$, $\theta_j \in (0, \pi)$ for $j = 1, ..., p$ and $\Omega := \text{skew}(A)$. Assume moreover that the imaginary parts of the eigenvalues are all distinct. Then, if the matrix $\widehat{Q} \in \mathrm{O}(2p)$ satisfies $\Omega\widehat{Q} - \widehat{Q}\text{skew}(S) = E$ with $\|E\|_\mathrm{F} \leq \tau\varepsilon_\mathrm{m}\|A\|_\mathrm{F}$, we have*

$$\|A\widehat{Q} - \widehat{Q}S\|_\mathrm{F} \leq \tau\varepsilon_\mathrm{m}\|A\|_\mathrm{F}\left(1 + \max_{\substack{i,j=1,..,p \\ i\neq j}}\left|\frac{\lambda_i\cos(\theta_i) - \lambda_j\cos(\theta_j)}{\lambda_i\sin(\theta_i) - \lambda_j\sin(\theta_j)}\right|\right). \tag{4.7}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Moreover, if A is orthogonal, then

$$\max_{\substack{i,j=1,..,p\\ i\neq j}} \left| \frac{\lambda_i \cos(\theta_i) - \lambda_j \cos(\theta_j)}{\lambda_i \sin(\theta_i) - \lambda_j \sin(\theta_j)} \right| = \max_{\substack{i,j=1,..,p\\ i\neq j}} \left| \tan\left(\frac{\theta_i+\theta_j}{2}\right)\right|. \tag{4.8}$$

Proof. Let us introduce that $\Omega = Q\mathrm{skew}(S)Q^\top$. This leads to

$$\Omega\widehat{Q} - \widehat{Q}\mathrm{skew}(S) = E \iff \mathrm{skew}(S)X - X\mathrm{skew}(S) = \widetilde{E}, \tag{4.9}$$

where we defined $X \coloneqq Q^\top\widehat{Q}$ and $\widetilde{E} \coloneqq Q^\top E$. Recall that $S = \left[\begin{smallmatrix} \Lambda\cos(\Theta) & -\Lambda\sin(\Theta) \\ \Lambda\sin(\Theta) & \Lambda\cos(\Theta)\end{smallmatrix}\right]$ such that (4.9) is the concatenation of $p^2 = \frac{n^2}{4}$ small 2×2 Sylvester equations. Indeed, for every pair $i,j = 1,...,p$, using the short-hand notation $s_i \coloneqq \sin(\theta_i)$ and $c_i \coloneqq \cos(\theta_i)$, we have

$$\begin{bmatrix} 0 & -\lambda_i s_i \\ \lambda_i s_i & 0 \end{bmatrix} X_{ij} - X_{ij}\begin{bmatrix} 0 & -\lambda_j s_j \\ \lambda_j s_j & 0 \end{bmatrix} = \widetilde{E}_{ij}, \tag{4.10}$$

where X_{ij} and $\widetilde{E}_{ij}$ are nonoverlapping 2×2 blocks of X and $\widetilde{E}$, respectively. If $i \neq j$, by [14, Thm. VII.2.8], we have $\|X_{ij}\|_{\mathrm{F}} \le \frac{1}{|\lambda_i s_i - \lambda_j s_j|}\|\widetilde{E}_{ij}\|_{\mathrm{F}}$. Therefore, acknowledging that $\|A\widehat{Q} - \widehat{Q}S\|_{\mathrm{F}} = \|SX - XS\|_{\mathrm{F}} \le \|\mathrm{sym}(S)X - X\mathrm{sym}(S)\|_{\mathrm{F}} + \|\widetilde{E}\|_{\mathrm{F}}$, we obtain

$$\begin{aligned}
\|A\widehat{Q} - \widehat{Q}S\|_{\mathrm{F}}^2 &= \sum_{i,j=1}^{p} \left\| \lambda_i \begin{bmatrix} c_i & -s_i \\ s_i & c_i \end{bmatrix} X_{ij} - X_{ij}\lambda_j \begin{bmatrix} c_j & -s_j \\ s_j & c_j\end{bmatrix}\right\|_{\mathrm{F}}^2 \\
&\le \sum_{i,j=1}^{p} \Big(|\lambda_i c_i - \lambda_j c_j|\, \|X_{ij}\|_{\mathrm{F}} + \|\widetilde{E}_{ij}\|_{\mathrm{F}}\Big)^2 \\
&\le \left(1 + \max_{\substack{i,j=1,..,p\\ i\neq j}} \left|\frac{\lambda_i c_i - \lambda_j c_j}{\lambda_i s_i - \lambda_j s_j}\right|\right)^2 \sum_{i,j=1}^{p} \|\widetilde{E}_{ij}\|_{\mathrm{F}}^2.
\end{aligned}$$

The claim (4.7) follows directly by observing that $\sqrt{\sum_{i,j=1}^p \|\widetilde{E}_{ij}\|_{\mathrm{F}}^2} = \|\widetilde{E}\|_{\mathrm{F}} = \|E\|_{\mathrm{F}} \le \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$. If A is orthogonal, then $\lambda_j = 1$ for $j = 1,...,p$. (4.8) follows from simple trigonometric identities. ∎

Generalizing Theorem 4.4 when the matrix $\widehat{Q}$ is not exactly orthogonal is simple and does not change its conclusion. This is shown in Corollary 4.5.

Corollary 4.5. *Assume in Theorem 4.4 that $\widehat{Q} \notin \mathrm{O}(2p)$ but $d_{\mathrm{O}(n)}(\widehat{Q}) \le \kappa\varepsilon_{\mathrm{m}}$ with $\kappa > 0$. Then, the matrix $\widehat{Q}_* \in \arg\min_{Q\in\mathrm{O}(n)} \|Q - \widehat{Q}\|_{\mathrm{F}}$ satisfies the hypotheses of Theorem 4.4 and $\|A\widehat{Q} - \widehat{Q}S\|_{\mathrm{F}} \le 2\kappa\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}} + \|A\widehat{Q}_* - \widehat{Q}_*S\|_{\mathrm{F}}$.*

Proof. First, we have

$$\begin{aligned}
\|A\widehat{Q} - \widehat{Q}S\|_{\mathrm{F}} &\le \|A(\widehat{Q} - \widehat{Q}_*) - (\widehat{Q} - \widehat{Q}_*)S\|_{\mathrm{F}} + \|A\widehat{Q}_* - \widehat{Q}_*S\|_{\mathrm{F}} \\
&\le 2\|\widehat{Q} - \widehat{Q}_*\|_{\mathrm{F}}\|A\|_{\mathrm{F}} + \|A\widehat{Q}_* - \widehat{Q}_*S\|_{\mathrm{F}} \\
&\le 2\kappa\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}} + \|A\widehat{Q}_* - \widehat{Q}_*S\|_{\mathrm{F}}.
\end{aligned}$$

Moreover, $\widehat{Q}_*$ satisfies the hypotheses of Theorem 4.4:

$$\begin{aligned}
\|\Omega\widehat{Q}_* - \widehat{Q}_*\mathrm{skew}(S)\|_{\mathrm{F}} &\le \|\widehat{\Omega}(\widehat{Q}_* - \widehat{Q}) - (\widehat{Q}_* - \widehat{Q})\mathrm{skew}(S)\|_{\mathrm{F}} + \|\widehat{\Omega}\widehat{Q} - \widehat{Q}\mathrm{skew}(S)\|_{\mathrm{F}} \\
&\le \|\widehat{Q}_* - \widehat{Q}\|_{\mathrm{F}}(\|\Omega\|_{\mathrm{F}} + \|\mathrm{skew}(S)\|_{\mathrm{F}}) + \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}} \\
&\le (2\kappa + \tau)\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}.
\end{aligned}$$

∎

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

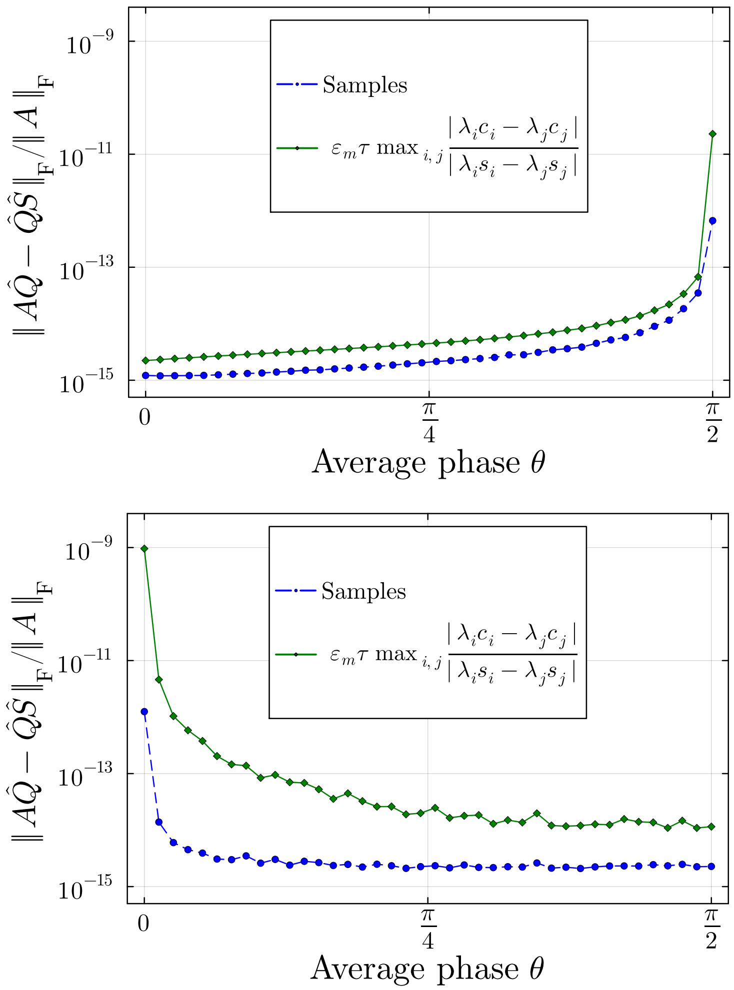

FIGURE 2. *Accuracy in double precision of Algorithm 4.1 (without any correction step 8) for 100×100 orthogonal (left plot) and normal (right plot) matrices when the average phase of the eigenvalues varies from $\theta = 0$ to $\theta = \frac{\pi}{2}$. The Schur vectors are sampled with Haar distribution on $\mathrm{O}(100)$. The eigenvalues of the matrices are sampled with phase uniformly distributed in $[\theta - \Delta\theta, \theta + \Delta\theta]$, with $\Delta\theta = 0.005$. For normal matrices, the moduli of the eigenvalues are uniformly sampled in $[0, 2]$. Results are averaged over 50 runs. The accuracy follows the behavior predicted by Theorem 4.4.*

Theorem 4.4 is very informative on the accuracy that Algorithm 4.1 will achieve when it is allowed to take its fastest route by skipping steps 5 and 7. We recall that this setting occurs with very high probability for Haar-distributed orthogonal matrices, and more generally, for random orthogonal matrices. In particular, if the bound in (4.7) is not satisfactory to the user, i.e., if $\max_{\substack{i,j=1,\dots,p \\ i\neq j}} \left|\frac{\lambda_i \cos(\theta_i) - \lambda_j \cos(\theta_j)}{\lambda_i \sin(\theta_i) - \lambda_j \sin(\theta_j)}\right| \geq t$, the sets of eigenvalues responsible for the loss of accuracy can be identified and the associated invariant subspaces can be corrected. We detail this approach in subsection 4.5. In Fig. 2, we present the results of a numerical experiment corroborating the bounds of Theorem 4.4 and assessing their relevance.

4.3. Sensitivity of the invariant subspaces: Set 2. From now on, we use the notation c, s for $\cos(\theta)$ and $\sin(\theta)$. Assume still that $A = QSQ^\top \in \mathbb{R}^{2p \times 2p}$ $(r = 0)$ is a normal matrix with eigenvalues $\lambda_j e^{\pm i\theta_j}$, $\lambda_j > 0$, $\theta_j \in (0, \pi)$ for $j = 1, \dots, p$. We next consider the case of eigenvalues that have repeated imaginary

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

parts (Set 2). Assume a set of indices $\mathcal{J} \subset \{1, 2, .., p\}$, $|\mathcal{J}| =: m$, such that for all $j \in \mathcal{J}$, we have $\lambda_j s_j = \sigma > 0$. In other words, we can numerically compute $\widehat{V}$ such that $\widehat{\Omega}\widehat{V} - \widehat{V}\widehat{\sigma}\left[\begin{smallmatrix} 0 & -I_m \\ I_m & 0 \end{smallmatrix}\right] = \widehat{E}$ with $\|\widehat{E}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|\widehat{\Omega}\|_{\mathrm{F}}$. In view of Corollary 4.5, w.l.o.g., we can consider $\widehat{V} \in \mathrm{St}(n, 2m)$. We set $\mathcal{I} := \{1, 2, .., p\} \setminus \mathcal{J}$. The proof of Theorem 4.4 fails when [14, Thm. VII.2.8] is used after (4.10) because, for $j_1, j_2 \in \mathcal{J}$, $\lambda_{j_1} s_{j_1} = \lambda_{j_2} s_{j_2}$.

As expected, the accuracy of the invariant subspace computation depends on the separation between $\pm i\sigma$ and the other eigenvalues values of Ω. In subsection 4.5, we use Theorem 4.6 to design an algorithm that satisfies a desired accuracy.

THEOREM 4.6. *Let $A = QSQ^\top \in \mathbb{R}^{2p\times 2p}$ be a normal matrix with eigenvalues $\lambda_j e^{\pm i\theta_j}$, $\lambda_j > 0$, $\theta_j \in (0, \pi)$ for $j = 1, ..., p$ and $\Omega := \mathrm{skew}(A)$. Assume there is $\mathcal{J} \subset \{1, ..., p\}$, $|\mathcal{J}| =: m$, such that for all $j \in \mathcal{J}$, we have $\lambda_j s_j = \sigma$. Set $\mathcal{I} := \{1, ..., p\} \setminus \mathcal{J}$. If the matrix $\widehat{V} \in \mathrm{St}(2p, 2m)$ satisfies $\Omega\widehat{V} - \widehat{V}\sigma\left[\begin{smallmatrix} 0 & -I_m \\ I_m & 0 \end{smallmatrix}\right] = E$ with $\|E\|_{\mathrm{F}} \le \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, then we have*

$$\begin{aligned}&\min_{R\in \mathrm{O}(2m)} \|A(\widehat{V}R) - (\widehat{V}R)\left[\begin{smallmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{smallmatrix}\right]\|_{\mathrm{F}} \\ &\le \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}\left(1 + \frac{\max\limits_{i\in\mathcal{I}, j\in\mathcal{J}} |\lambda_i c_i - \lambda_j c_j|}{\min\limits_{i\in\mathcal{I}} |\lambda_i s_i - \sigma|} + \varepsilon_{\mathrm{m}}\tau\|A\|_{\mathrm{F}}\frac{\max\limits_{j_1,j_2\in\mathcal{J}} |\lambda_{j_1} c_{j_1} - \lambda_{j_2} c_{j_2}|}{\min\limits_{i\in\mathcal{I}} |\lambda_i s_i - \sigma|^2}\right).\end{aligned}$$

Proof. Define $X := Q^\top\widehat{V}$, then it follows that

$$\mathrm{skew}(S)X - X\sigma\begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} = Q^\top E =: \widetilde{E}.$$

There exist two permutations $P_{\mathrm{l}} \in \mathrm{O}(n), P_{\mathrm{r}} \in \mathrm{O}(2m)$ such that $P_{\mathrm{l}}^\top XP_{\mathrm{r}} = \left[\begin{smallmatrix} X_{\mathcal{I},\mathcal{J}} \\ X_{\mathcal{J},\mathcal{J}} \end{smallmatrix}\right]$ where the block $X_{\mathcal{I},\mathcal{J}}$ comprises all 2×2 blocks X_{ij} such that for $i \in \mathcal{I}, j \in \mathcal{J}$. We have

$$\begin{bmatrix} 0 & -\lambda_i s_i \\ \lambda_i s_i & 0 \end{bmatrix} X_{ij} - X_{ij}\begin{bmatrix} 0 & -\sigma \\ \sigma & 0 \end{bmatrix} = (P_{\mathrm{l}}^\top\widetilde{E}P_{\mathrm{r}})_{ij}. \tag{4.11}$$

If it hold that $\widetilde{E} = 0$, then we would have $X_{\mathcal{I},\mathcal{J}} = 0$ and $X_{\mathcal{J},\mathcal{J}} \in \mathrm{O}(2m)$. However, since $\widetilde{E} \ne 0$ and $X \in \mathrm{St}(n, 2m)$, we have

$$X_{\mathcal{I},\mathcal{J}}^\top X_{\mathcal{I},\mathcal{J}} + X_{\mathcal{J},\mathcal{J}}^\top X_{\mathcal{J},\mathcal{J}} = I_{2m} \Longrightarrow \|X_{\mathcal{J},\mathcal{J}}^\top X_{\mathcal{J},\mathcal{J}} - I_{2m}\|_{\mathrm{F}} \le \|X_{\mathcal{I},\mathcal{J}}\|_{\mathrm{F}}^2$$

$$\text{and by (4.11)}, \quad \|X_{\mathcal{I},\mathcal{J}}\|_{\mathrm{F}}^2 \le \sum_{i\in\mathcal{I}, j\in\mathcal{J}} \frac{1}{|\lambda_i s_i - \sigma|^2}\|(P_{\mathrm{l}}^\top\widetilde{E}P_{\mathrm{r}})_{ij}\|_{\mathrm{F}}^2 \le \frac{\|E\|_{\mathrm{F}}^2}{\min_{i\in\mathcal{I}} |\lambda_i s_i - \sigma|^2}.$$

The last inequality is obtained using [14, Thm. VII.2.8]. The deviation of $X_{\mathcal{J},\mathcal{J}}$ from orthogonality can thus be quantified. In addition, by taking the SVD $X_{\mathcal{J},\mathcal{J}} = U_X\Sigma_X V_X^\top$, we can write $\|X_{\mathcal{J},\mathcal{J}}^\top X_{\mathcal{J},\mathcal{J}} - I_{2m}\|_{\mathrm{F}} = \|\Sigma_X^2 - I_{2m}\|_{\mathrm{F}} \ge \|\Sigma_X - I_{2m}\|_{\mathrm{F}}$, where the last inequality stands because $|\alpha^2 - 1| \ge |\alpha - 1|$ for all $\alpha \ge 0$. Define $\left[\begin{smallmatrix} \widetilde{S}_{2(p-m)} & 0 \\ 0 & \widetilde{S}_{2m} \end{smallmatrix}\right] := P_{\mathrm{l}}^\top SP_{\mathrm{l}}$, then for all $R \in \mathrm{O}(2m)$, we have

$$\begin{aligned}\|A\widehat{V}R - \widehat{V}R\widetilde{S}_{2m}\|_{\mathrm{F}} &\le \|\mathrm{sym}(S)XR - XR\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} + \|E\|_{\mathrm{F}} \\ &\le \|\mathrm{sym}(\widetilde{S}_{2(p-m)})X_{\mathcal{I},\mathcal{J}}R - X_{\mathcal{I},\mathcal{J}}R\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} \\ &\quad + \|\mathrm{sym}(\widetilde{S}_{2m})X_{\mathcal{J},\mathcal{J}}R - X_{\mathcal{J},\mathcal{J}}R\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} + \|E\|_{\mathrm{F}}.\end{aligned} \tag{4.12}$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

The first term of (4.12) can be bounded by

$$\begin{aligned}\|\mathrm{sym}(\widetilde{S}_{2(p-m)})X_{\mathcal{I},\mathcal{J}}R - X_{\mathcal{I},\mathcal{J}}R\,\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} &\leq \max_{i\in\mathcal{I},j\in\mathcal{J}} |\lambda_i c_i - \lambda_j c_j| \|X_{\mathcal{I},\mathcal{J}}R\|_{\mathrm{F}} \\ &\leq \frac{\max_{i\in\mathcal{I},j\in\mathcal{J}} |\lambda_i c_i - \lambda_j c_j|}{\min_{i\in\mathcal{I}} |\lambda_i s_i - \sigma|} \|E\|_{\mathrm{F}}. \end{aligned} \tag{4.13}$$

Moreover, the second term can also be bounded by

$$\begin{aligned}&\|\mathrm{sym}(\widetilde{S}_{2m})X_{\mathcal{J},\mathcal{J}}R - X_{\mathcal{J},\mathcal{J}}R\,\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} \\ &= \|\mathrm{sym}(\widetilde{S}_{2m})(X_{\mathcal{J},\mathcal{J}}R - I_{2m}) - (X_{\mathcal{J},\mathcal{J}}R - I_{2m})\mathrm{sym}(\widetilde{S}_{2m})\|_{\mathrm{F}} \\ &\leq \max_{j_1,j_2\in\mathcal{J}} |\lambda_{j_1} c_{j_1} - \lambda_{j_2} c_{j_2}| \|X_{\mathcal{J},\mathcal{J}}R - I_{2m}\|_{\mathrm{F}}. \end{aligned} \tag{4.14}$$

In particular, for $R = V_X U_X^\top$ which minimizes $\|X_{\mathcal{J},\mathcal{J}}R - I_{2m}\|_{\mathrm{F}}$, we have

$$\|X_{\mathcal{J},\mathcal{J}}R - I_{2m}\|_{\mathrm{F}} = \|\Sigma_X - I_{2m}\|_{\mathrm{F}} \leq \frac{\|E\|_{\mathrm{F}}^2}{\min_{i\in\mathcal{I}} |\sigma - \lambda_i s_i|^2}. \tag{4.15}$$

The theorem follows directly by inserting (4.13), (4.14) and (4.15) in (4.12) and acknowledging that $\|E\|_{\mathrm{F}} \leq \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$. ∎

4.4. Sensitivity of the invariant subspaces: Set 3. The case of real eigenvalues (Set 3) can be handled in a similarly to the case of repeated imaginary parts. Indeed, the proof of Theorem 4.7 is very close to that of Theorem 4.6, if $\left[\begin{smallmatrix} D_m & -\sigma I_m \\ \sigma I_m & D_m \end{smallmatrix}\right]$ is replaced by $\breve{\Lambda}$. For this reason, the proof is given in Appendix C.

THEOREM 4.7. *Let $A = QSQ^\top \in \mathbb{R}^{n\times n}$ be a normal matrix with real eigenvalues $\breve{\lambda}_k$ for $k \in \mathcal{K} := \{1, ..., r\}$ and $\Omega := \mathrm{skew}(A)$. Moreover, assume $\lambda_j e^{\pm i\theta_j}$ is an eigenvalue of A for $j \in \mathcal{I}$. If the matrix $\widehat{V} \in \mathrm{St}(n, r)$ satisfies $\Omega\widehat{V} = \widehat{E}$ with $\|E\|_{\mathrm{F}} \leq \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, then we have*

$$\begin{aligned}&\min_{\breve{R}\in\mathrm{O}(2m)} \|A\widehat{V}\breve{R} - \widehat{V}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} \\ &\leq \tau\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}} \left(1 + \frac{\max\limits_{i\in\mathcal{I},k\in\mathcal{K}} |\lambda_i c_i - \breve{\lambda}_k|}{\min\limits_{i\in\mathcal{I}} |\lambda_i s_i|} + \varepsilon_{\mathrm{m}}\tau\|A\|_{\mathrm{F}} \frac{\max\limits_{k_1,k_2\in\mathcal{K}} |\breve{\lambda}_{k_1} - \breve{\lambda}_{k_2}|}{\min\limits_{i\in\mathcal{I}} |\lambda_i s_i|^2}\right).\end{aligned}$$

Proof. See Appendix C. ∎

4.5. Analysis of steps 5 to 8. The previous subsections determined the sensitivity of the computation of the Schur vectors of A by the Schur vectors of Ω. We can now provide conditions on steps 5 to 8 of Algorithm 4.1 to achieve a desired accuracy.

Step 5 requires identifying every approximately invariant subspace of Ω spanned by $\widehat{V}$, i.e., $\widehat{V}$ such that $\|\Omega\widehat{V} - \widehat{V}\sigma J_{2m}\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, $m > 1$. In view of Corollary 4.5, we assume $\widehat{V} \in \mathrm{St}(n, 2m)$. To recover the Schur vectors of A, we need to compute $M := \widehat{V}^\top A\widehat{V}$. In exact arithmetic, $2m$ is the multiplicity of the singular value σ of Ω. In floating point arithmetic, $2\widehat{m}$ must be the size of the cluster of singular values of $\widehat{\Omega}$. If we take $\widehat{V}$ as the $2\widehat{m}$ computed Schur vectors associated to a δ-cluster of $\widehat{\Omega}$, in view of Theorem 4.6, it holds that $\|A\widehat{V} - \widehat{V}M\|_{\mathrm{F}} \lesssim \frac{\varepsilon_{\mathrm{m}}}{\delta}\|A\|_{\mathrm{F}}$. If δ is chosen such that the δ-cluster is well-separated from the others and if the RSD $M \approx R\widetilde{S}_{2m}R^\top$ is such that $\|M - R\widetilde{S}_{2m}R^\top\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, then after step 5, we have $\|A\widehat{V}R - \widehat{V}R\widetilde{S}_{2m}\|_{\mathrm{F}} \lesssim \frac{\varepsilon_{\mathrm{m}}}{\delta}\|A\|_{\mathrm{F}}$ and $d_{\mathrm{St}(n,2m)}(\widehat{V}R) \lesssim \varepsilon_{\mathrm{m}}$. Notice here that the effective accuracy that is

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

obtained is in fact function of the largest δ such that the δ-cluster of $\widehat{\Omega}$ remains the same, rather than the user specified value δ. As mentioned in section 3, several stable algorithms exist for Routine 3 at step 5.2.

Step 6 obtains the real parts of the eigenvalues by computing Rayleigh quotients. It is known that the Rayleigh quotient offers the square of the accuracy of the eigenvector, see, e.g., [69, Lecture 27]. Therefore, by Theorem 4.4, step 6 provides a worst-case $\mathcal{O}(\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}})$-absolute accuracy on the real parts of the eigenvalues if $\max_{i,j}\left|\frac{\lambda_i c_i-\lambda_j c_j}{\lambda_i s_i-\lambda_j s_j}\right| \lesssim \frac{1}{\sqrt{\varepsilon_{\mathrm{m}}}}$. It suggests considering δ-clusters of Ω with $\delta \leq \sqrt{\varepsilon_{\mathrm{m}}}$ to guarantee high absolute accuracy on the eigenvalues.

Step 7 computes the real eigenvalues and the eigenvectors of A. This starts by estimating r by $\widehat{r}$. By Theorem 4.7, considering a δ_r-cluster of the eigenvalues of $\widehat{\Omega}$ around zero will provide a worst-case $\mathcal{O}(\frac{\varepsilon_{\mathrm{m}}}{\delta_r}\|A\|_{\mathrm{F}})$-absolute accuracy on the real eigenspace of A. It is essential not to underestimate r, i.e., to have $\widehat{r} \geq r$. Therefore, δ_r should always be increased until the δ_r-cluster around zero is well separated from the other eigenvalues, if needed at the cost of the symmetry of $H \coloneqq \widehat{Q}^\top_{\widehat{p}+1:\widehat{p}+\widehat{r}} A \widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}$. If $\|\mathrm{skew}(H)\|_{\mathrm{F}} \leq \varepsilon_{\mathrm{m}}\|H\|_{\mathrm{F}}$, then H can be considered exactly symmetric without loss of accuracy since the skew-symmetric part can be considered as a negligible perturbation. If H is symmetric, the decomposition at step 7.2a using, e.g., syevr, yields

$$H + \Delta H = \breve{R}\breve{\Lambda}\breve{R}^\top \quad \text{with} \quad \|\Delta H\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|H\|_{\mathrm{F}},$$

with $d_{\mathrm{O}(\widehat{r})}(\breve{R}) \lesssim \varepsilon_{\mathrm{m}}$ [6]. This yields $\|A\widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}\breve{R} - \widehat{Q}_{\widehat{p}+1:\widehat{p}+\widehat{r}}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} \lesssim \frac{\varepsilon_{\mathrm{m}}}{\delta_r}\|A\|_{\mathrm{F}}$ with $d_{\mathrm{St}(n,\widehat{r})}(\widehat{Q}_{p+1:p+\widehat{r}}\breve{R}) \lesssim \varepsilon_{\mathrm{m}}$. Only in certain circumstances will symmetric eigensolvers achieve relative accuracy [29, 30, 33, 72]. If H is not numerically symmetric, it means that $\widehat{r}$ overestimated r, Routine 3 should then be preferred to Routine 2 to maintain accuracy.

Step 8 can finally be considered to improve the accuracy of the RSD to a desired tolerance $\mu = \varepsilon_{\mathrm{m}} t$, $t \geq 1$ if this accuracy has not yet been attained. By combining Theorems 4.4, 4.6 and 4.7, we deduce that $\|A\widehat{Q} - \widehat{Q}\widehat{S}\|_{\mathrm{F}} \lesssim \left(1 + \max_{i\neq j,k}\{\frac{|\lambda_i c_i-\lambda_j c_j|}{|\lambda_i s_i-\lambda_j s_j|}, \frac{|\breve{\lambda}_k-\lambda_j c_j|}{|\lambda_j s_j|}\}\right)\varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$. If $\max_{i\neq j,k}\{\frac{|\lambda_i c_i-\lambda_j c_j|}{|\lambda_i s_i-\lambda_j s_j|}, \frac{|\breve{\lambda}_k-\lambda_j c_j|}{|\lambda_j s_j|}\} \geq t$, the groups of eigenvalues responsible for making the constant larger than t can be identified. Then, to restore the desired accuracy, one should repeat step 5 with $\delta = \frac{1}{t}$ on the invariant subspaces associated with $\frac{1}{t}$-clusters of $\widehat{\Omega}$. Because the matrix $\widehat{Q}$ is already close to the Schur vectors of A, the matrix M at step 5 is a small perturbation of the Schur form. For such matrices, a Jacobi algorithm such as [83] is efficient since it converges quadratically. One sweep is enough in practice to obtain the desired accuracy. However, for specific matrices and t too small, the $\frac{1}{t}$-clusters may include many eigenvalues. The user should be aware that step 8 may then require a significant additional computational work. In double precision, the authors recommend using $t \geq 100$ based on empirical experience. If the output of Routine 3 satisfies $\|M - R\widetilde{S}_{2m}R^\top\|_{\mathrm{F}} \lesssim \varepsilon_{\mathrm{m}}\|A\|_{\mathrm{F}}$, then finally, we obtain the decomposition $\|A\widehat{Q} - \widehat{Q}\widehat{S}\|_{\mathrm{F}} \lesssim \mu\|A\|_{\mathrm{F}}$ with $d_{\mathrm{O}(n)}(\widehat{Q}) \lesssim \varepsilon_{\mathrm{m}}$.

5. Numerical experiments on the accuracy. In this section, we present the results of several numerical experiments on the accuracy of Algorithm 4.1 for a diverse set of $n \times n$ problems. The results are given in Table 1 and Fig. 3. For five different distributions of the eigenvalues, we report the average accuracy of Algorithm 4.1 over 100 runs. Codes reproducing the numerical experiments can be found at https://github.com/smataigne/NormalEVP.jl.

- Experiment E1 verifies that in the setting of Theorem 4.4, if the bound of (4.7) remains small, then Algorithm 4.1 achieves high accuracy on the RSD.
- Experiment E2 tests a random normal scenario representing an "average normal case." In this case, a moderate loss of accuracy of the Schur decomposition occurs as n grows because the separation

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

TABLE 1

Numerical experiment on the accuracy in double precision of `nrmschur` *(Algorithm 4.1) and* `schur` *(LAPACK's* `gees`*) to compute $A \approx \widehat{Q}\widehat{S}\widehat{Q}^\top$. Step 8 is considered only for the last case where exactly one Jacobi sweep is performed. We chose $\delta, \delta_r, \frac{1}{t}$ equal to $\sqrt{\varepsilon_\mathrm{m}}$ to define clusters. S is the ground truth real Schur form. Each reported accuracy is averaged over 100 runs. The normal matrices A are randomly sampled with eigenvalues of different distributions specified in the table. The Schur vectors are always sampled with a Haar distribution on $\mathrm{O}(n)$. For all $a, b \in \mathbb{R}$, ae-b denotes $a \cdot 10^{-b}$.*

n	$\frac{\Vert A\widehat{Q}-\widehat{Q}\widehat{S}\Vert_\mathrm{F}}{\Vert A\Vert_\mathrm{F}}$		$\frac{\Vert \widehat{Q}^\top\widehat{Q}-I\Vert_\mathrm{F}}{\sqrt{n}}$		$\frac{\Vert \mathrm{diag}(S-\widehat{S})\Vert_\mathrm{F}}{1+\Vert \mathrm{diag}(S)\Vert_\mathrm{F}}$	
	`nrmschur`	`schur`	`nrmschur`	`schur`	`nrmschur`	`schur`
E1: $\lambda = 1,\ \theta \sim \mathcal{U}(0, \frac{\pi}{4})$: Best accuracy scenario on $\mathrm{SO}(n)$.						
10	8.0e-16	1.1e-15	6.7e-16	1.2e-15	3.8e-16	4.6e-16
32	1.3e-15	1.9e-15	1.2e-15	2.0e-15	6.6e-16	8.1e-16
100	1.5e-15	3.5e-15	1.6e-15	3.7e-15	6.6e-16	1.6e-15
316	1.5e-15	7.2e-15	2.0e-15	7.8e-15	5.8e-16	4.3e-15
1000	1.7e-15	2.3e-14	2.8e-15	2.3e-14	6.4e-16	1.9e-14
E2: $\lambda \sim \mathcal{U}(0, 2),\ \theta \sim \mathcal{U}(0, \pi)$: Random normal scenario.						
10	2.7e-15	1.4e-15	6.2e-16	1.3e-15	3.4e-16	5.1e-16
32	1.6e-14	2.6e-15	1.1e-15	2.2e-15	6.2e-16	1.0e-15
100	1.2e-13	4.2e-15	1.5e-15	3.8e-15	7.2e-16	1.9e-15
316	4.0e-13	7.4e-15	1.9e-15	7.0e-15	6.0e-16	3.9e-15
1000	1.6e-12	1.3e-14	2.6e-15	1.2e-14	6.2e-16	8.4e-15
E3: $\lambda, \check{\lambda} \sim \mathcal{U}(0, 2),\ \theta \sim \mathcal{U}(0, \pi)$: Random normal scenario, 20% real eigenvalues.						
10	2.3e-15	1.4e-15	5.7e-16	1.3e-15	2.9e-16	4.9e-16
32	1.3e-14	2.6e-15	1.2e-15	2.2e-15	5.8e-16	1.0e-15
100	6.7e-14	4.1e-15	3.8e-15	3.7e-15	6.5e-16	1.9e-15
316	2.7e-13	6.9e-15	1.2e-14	6.5e-15	5.8e-16	3.7e-15
1000	8.7e-13	1.1e-14	2.9e-14	1.1e-14	5.7e-16	7.3e-15
E4: $\lambda \sim \mathcal{U}(0, 2),\ \theta \sim \mathcal{U}(0, \pi)$: Random normal scen., 20% repeated imaginary parts.						
10	3.5e-15	1.3e-15	6.6e-16	1.3e-15	3.4e-16	4.6e-16
32	2.4e-14	2.0e-15	1.1e-15	1.9e-15	5.8e-16	7.5e-16
100	1.2e-13	3.6e-15	1.5e-15	3.5e-15	6.9e-16	1.4e-15
316	4.1e-13	7.4e-15	2.1e-15	7.9e-15	8.4e-16	4.2e-15
1000	1.5e-12	2.3e-14	3.1e-15	2.4e-14	1.2e-15	1.9e-14
E5: $\lambda \sim \mathcal{U}(0, 2),\ \theta \sim \pi\sqrt{\varepsilon_\mathrm{m}}\mathcal{N}(1, 1)$: Worst-case normal scenario.						
10	5.1e-10 $\xrightarrow{\text{step 8}}$ 1.2e-15	1.1e-15	9.9e-16	1.1e-15	5.3e-16	5.3e-16
32	2.1e-9 → 2.9e-15	1.8e-15	2.2e-15	1.9e-15	1.2e-15	6.1e-16
100	5.6e-9 → 5.8e-15	3.7e-15	4.1e-15	3.6e-15	2.9e-15	1.7e-15
316	9.7e-9 → 1.4e-14	6.1e-15	6.8e-15	6.5e-15	3.5e-15	2.5e-15
1000	1.1e-8 → 8.4e-14	9.7e-15	1.1e-14	1.0e-14	5.8e-15	3.8e-15

between the imaginary parts of the eigenvalues decreases with n. However, the absolute accuracy on the eigenvalues remains close to machine precision thanks to the Rayleigh quotients.

- Experiments E3 and E4 are similar to E2, but enforce nonempty sets Set 2 and Set 3, respectively. The accuracy of the RSD remains equivalent to E2, as well as the absolute accuracy on the eigenvalues.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

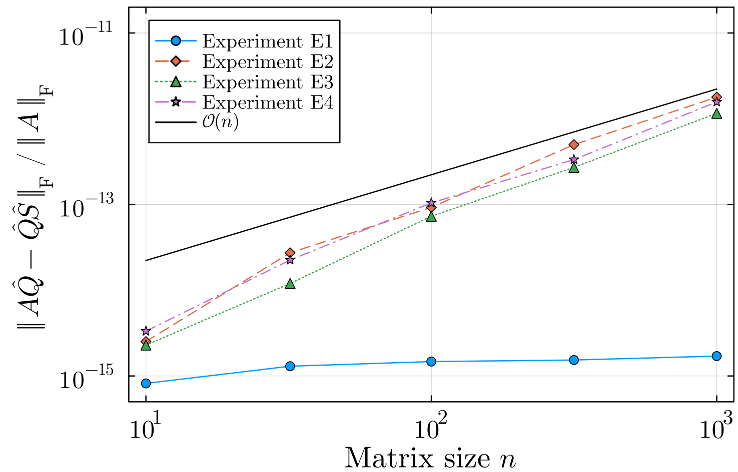

FIGURE 3. *Evolution of the accuracy in double precision of* `nrmschur` *(Algorithm 4.1) without correction step 8. The matrix size n varies from 10 to 1000. The plots shows results for experiments E1 to E4 from section 5. The results are averaged over 100 runs.*

- Finally, experiment E5 tests the worst case for Algorithm 4.1 in terms of accuracy. The skew-symmetric part Ω is small such that the eigenvalues of Ω are relatively close and induce sensitivity of the invariant subspaces. Moreover, Ω is large enough such that eigenvalues of Ω are not considered as $\sqrt{\varepsilon_{\mathrm{m}}}$-clusters and the accuracy is not restored by step 5. However, the high accuracy on the eigenvalues is maintained by the Rayleigh quotients. We show that one sweep of the Jacobi method [83] allows to restore high accuracy of the RSD.

In Table 1, the deviation from orthogonality is measured by the simple criterion $\|\widehat{Q}^\top \widehat{Q} - I\|_{\mathrm{F}}$. We recall in Lemma A.4 that this measure is equivalent to $d_{\mathrm{O}(n)}(\widehat{Q})$ when it is small. The results of Fig. 3 show that for the experiments E2 to E4, the accuracy of the RSD grows proportionally to the matrix size n. Moreover, for the experiment E1, the accuracy is almost insensitive to n.

6. Complexity analysis and running time performance. We analyze the best and worst-case computational complexities of Algorithm 4.1. We exclude the cases where the matrix A is either purely skew-symmetric or symmetric since Algorithm 4.1 reduces to obtaining, respectively, the skew-symmetric RSD with Routine 1 or the symmetric EVD with Routine 2. These ideal cases are not relevant to our discussion. By analogy with algorithms such as quicksort, we will show that worst-case complexity analysis is not the most relevant here, as worst cases are very unlikely to occur.

Since the cost of the method is $\mathcal{O}(n^3)$, the $\mathcal{O}(n^2)$ costs are neglected. Following the notation from [69, Lec. 8], we say that a method requires $\sim g(n,p,r)$ flops if

$$\lim_{n,p,r\to\infty} \frac{\text{number of flops}}{g(n,p,r)} = 1. \tag{6.1}$$

The numerical experiments presented in this section are performed on a processor *Intel(R) Core(TM) i7-8750H CPU @ 2.20GHz* in single-threaded framework. Codes reproducing the numerical experiments can be found at https://github.com/smataigne/NormalEVP.jl.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

6.1. The intuitive (but not) best-case complexity. Intuitively, the best case for Algorithm 4.1 is when all eigenvalues are complex and none of the imaginary parts are clustered (relatively to ε_{m}). The assumptions of Theorem 3.3 are then matched with $k = \lfloor \frac{n}{2} \rfloor = p$ ($r = 1$ for n odd). We recall that for Haar-distributed special orthogonal matrices [7, 64], this case occurs with probability 1 [34]. Step 1 is negligible since it requires $\mathcal{O}(n^2)$ flops.

The cost of step 2 is $\sim \frac{4}{3}n^3$ for the skew-symmetric tridiagonal reduction [69, Lec. 26], plus an additional cost of $\sim \frac{4}{3}n^3$ to assemble the Householder reflectors [41, Sec. 5.1.6]. Indeed, given $M \in \mathbb{R}^{k\times k}$ and $v \in \mathbb{R}^k$, performing $(I - vv^\top)M$ requires $\sim 4k^2$ flops. Therefore, assembling all reflectors of size $k = 1, ..., n-1$ has a cost $\sim \sum_{k=1}^{n-1} 4k^2 \sim \frac{4}{3}n^3$ flops.

At step 3, the call to Routine 1 can be neglected since an $\mathcal{O}(p^2)$ implementation exists [45].

Step 4 requires $\sim 4np^2$ flops, step 5 is skipped, and step 6 costs $\sim 2n^2p$ flops. Step 7 is skipped (or costs $\mathcal{O}(n^2)$ if $r = 1$). The total cost is thus

$$\sim \frac{8}{3}n^3 + 4np^2 + 2n^2p \text{ flops.} \tag{6.2}$$

Given $p = \lfloor \frac{n}{2} \rfloor$, the total cost is $\sim \frac{14}{3}n^3$, among which $\sim \frac{6}{3}n^3$ are level 3 `BLAS` operations [18]. Moreover, the $\sim \frac{8}{3}n^3$ flops from the tridiagonalization also admit blocked implementations, maximizing the use of level 3 `BLAS` (see, e.g., SkewLinearAlgebra.jl).

In the best case, the total cost of Algorithm 4.1 is thus equivalent to a Hessenberg factorization, which is known to cost $\frac{10}{3}n^3$ flops to obtain the H factor [69, Lec. 26] and $\sim \frac{4}{3}n^3$ to assemble the Householder reflectors, hence $\sim \frac{14}{3}n^3$ flops in total. This result is Algorithm 4.1's best advantage over methods that include an initial reduction of the matrix A to Hessenberg form [4, 43, 73, 74], while the latter methods usually offer better accuracy on the eigenvectors in general.

Figure 4 illustrates the performance of `nrmschur` (Algorithm 4.1) compared to `Julia` calls to `LAPACK`'s Hessenberg factorization (`gehrd`+`orghr`) and `schur` routine (`gees`). We also include the performance of our `Julia` implementation of the UHQR algorithm [44]. UHQR's computational complexity for the eigenvalues is $\sim \frac{10}{3}n^3$ for the Hessenberg reduction plus $\mathcal{O}(n^2)$ for QR iterations, thus faster than `LAPACK's` eigenvalue routine (hseqr). However, if the Schur/eigen-vectors are desired, the complexity of UHQR is $\sim \frac{14}{3}n^3$ for the Hessenberg reduction plus $\mathcal{O}(n^3)$ because Givens rotations of the QR are accumulated.

With due caution in interpreting complexity from timing experiments, the numerical results in Fig. 4 corroborate that the cost of Algorithm 4.1 is comparable to that of a Hessenberg factorization. Moreover, it outperforms `gees` by a factor 10 for $n \approx 100$ and a factor 2 for $n = 10000$. We can observe that UHQR (only for eigenvalues) lies above Algorithm 4.1; however, a `FORTRAN` implementation of UHQR would be faster. Nonetheless, this highlights that because Algorithm 4.1 relies exclusively on high-performance software, namely `LAPACK`, it can be efficiently implemented in any computer language where an interface to this software is available.

6.2. The true best-case complexity. If we consider $r > 1$ real eigenvalues, then there is an additional cost of $\sim 2n^2r + 2nr^2$ flops at step 7.1. The call to Routine 2 at step 7.2a requires $\sim (\frac{4}{3} + \frac{4}{3})r^3$ flops for tridiagonalization plus $\mathcal{O}(r^2)$ flops to solve the tridiagonal EVP [30]. Finally, step 7.3 demands $\sim 2nr^2$ flops. However, notice that $p := \frac{n-r}{2}$ decreases when r rises. Hence, all costs from (6.2) depending on p decrease accordingly. For r sufficently small, we show that the decrease of p has a larger effect on the operation count

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

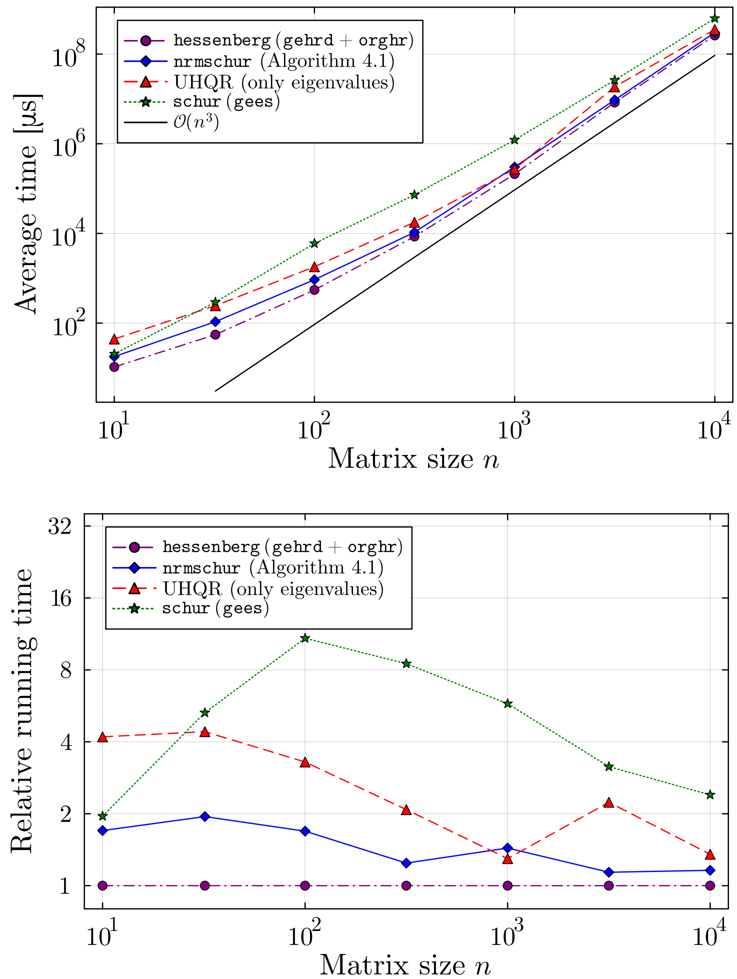

FIGURE 4. *Top: Benchmark of* `nrmschur` *(Algorithm 4.1) on Haar-distibuted matrices of* $\mathrm{SO}(n)$ *w.r.t. the* `Julia` *routines* `schur` *(LAPACK's* `gees`*) and* `hessenberg` *(LAPACK's* `gehrd` *and* `orghr`*). The performance of the UHQR algorithm for eigenvalues [44] as well as a representative* $\mathcal{O}(n^3)$ *is also drawn. The plot shows experiments for* $n \in \{10, 32, 100, 316, 1000, 3162, 10000\}$*. Bottom: the same results are displayed but divided by the computational time of* `hessenberg`*.*

than the increase of r. For $r > 0$, a pessimistic (but realistic, see Remark 6.1) total cost of Algorithm 4.1 is

$$\sim \frac{8}{3}(n^3 + r^3) + n(n-r)^2 + n^2(n-r) + 2n^2 r + 4nr^2 \text{ flops.} \tag{6.3}$$

By setting $r = \alpha n$ with $\alpha \in [0, 1]$, we obtain the complexity

$$\sim \left(\frac{8}{3}\alpha^3 + 5\alpha^2 - \alpha + \frac{14}{3}\right) n^3 \text{ flops.} \tag{6.4}$$

The polynomial in α achieves a local minimum at $\alpha = \frac{\sqrt{5^2+8}-5}{8} \approx 0.093$ for an optimal cost of $4.61n^3$ flops, instead of $\sim \frac{14}{3}n^3 \approx 4.66n^3$ when $\alpha = 0$. Although the variation of the constant is too small to yield significant gain in performance, the take-home conclusion is that "a few" real eigenvalues (up to 20%) do not impact the performance negatively. The numerical experiment of Fig. 5 confirms the relevance of this operation count. The average runtime fits the expected behavior when α varies.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

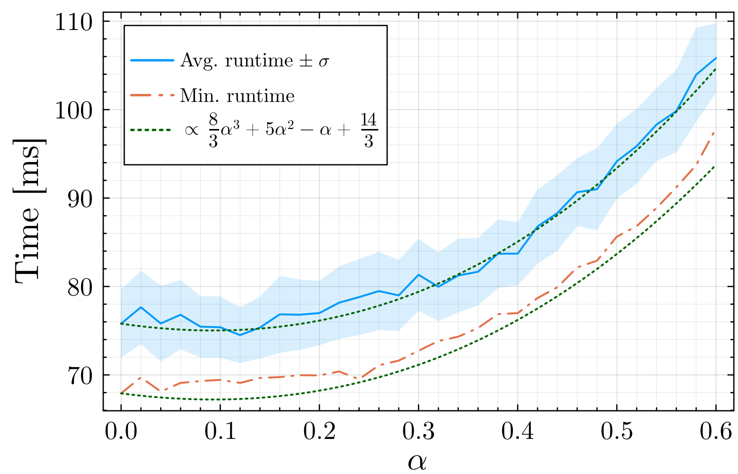

FIGURE 5. *Numerical experiments over* 600×600 *random normal matrices with proportion* α *of real eigenvalues. For each value* α ($\Delta\alpha = 0.02$)*, 200 matrices are sampled. The Schur vectors are Haar-distributed on* $\mathrm{O}(n)$*, the real eigenvalues are normally distributed with unit variance, and pairs of conjugate eigenvalues are sampled with uniformly distributed phases in* $[0, \pi]$ *and radii in* $[0, 2]$*. The theoretical complexity from* (6.4) *is represented by the dotted green line. The polynomial* $\frac{8}{3}\alpha^3 + 5\alpha^2 - \alpha + \frac{14}{3}$ *is scaled by a constant factor to match the intercept* ($\alpha = 0$) *of the numerical experiments.*

REMARK 6.1. When the number r of real eigenvalues grows, exact arithmetic could use only $n - r$ Householder reflectors to tridiagonalize $\mathrm{skew}(A)$ instead of $n - 1$. Indeed, since $\mathrm{skew}(A)$ has rank $n - r$ and since each column of a tridiagonal matrix with a nonzero subdiagonal element is independent from all previous columns, the (exact) tridiagonal form cannot have more than $n - r$ nonzero subdiagonal elements. In short, tridiagonalization is a rank-revealing method. The most expensive reflectors being the $n - k$ ones of largest size, the worst-case complexity in exact arithmetic of the tridiagonalization is $\sim \frac{8}{3}(n^3 - r^3)$ instead of $\sim \frac{8}{3}n^3$. However, due to numerical roundoff, logical zeros are perturbed. Given random matrices $U, V \in \mathbb{R}^{n\times l}$, $l \ll n$, it can be verified that `LAPACK`'s methods `sytrd` on $UU^\top$ and `gehrd` on $UV^\top$ produce $n - 1$ nontrivial reflectors. This validates in practice the cost $\sim \frac{8}{3}n^3$ considered in (6.3).

6.3. The worst-case complexity. The worst case for Algorithm 4.1 is when all eigenvalues are complex (all but one if n is odd) and their imaginary parts form two conjugate clusters. This case maximizes the additional work of Routine 3 at step 5. For simplicity, assume n even such that $r = 0$ and at step 5, $\Sigma = \sigma I_p$ such that $V = \widehat{Q}$. Essentially, all steps but step 5.2 are a waste a computational resources.

Assuming we ignore the (potentially inaccurate) symmetric skew-Hamiltonian structure, step 5.1 requires $\sim 4n^3$ flops. Step 5.2 requires $\sim \frac{14}{3}n^3$ and $\mathcal{O}(n^3)$ flops for, respectively, the Hessenberg reduction and Routine 3. Step 5.3 demands $\sim 2n^3$ flops. Algorithm 4.1 returns after step 5 since $\Lambda\cos(\Theta)$ was computed at step 5.2. Taking into account the $\sim \frac{8}{3}n^3 + \mathcal{O}(n^2) + n^3$ flops of steps 2 to 4, the total cost is $\sim \frac{43}{3}n^3 + \mathcal{O}(n^3)$, thus a large increase of the computational cost.

Figure 6 reports an experiment that is similar to the best case (Fig. 4) but for the worst case. Without surprise, the average runtime of Algorithm 4.1 exceeds that of `gees` since `gees` becomes a subroutine of Algorithm 4.1. The practical advantage of Algorithm 4.1 stems from the unlikelihood of having $\Sigma = \sigma I_p$. Most likely, the multiplicity of the entries of Σ, if any, remains small compared to n. Then, the cost of step 5 remains negligible, and the performance does not significantly differ from the $\sim \frac{14}{3}n^3$ flops of the best case.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

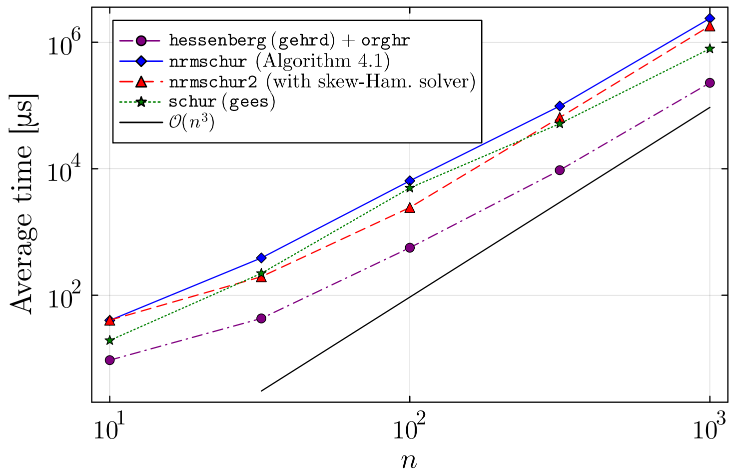

FIGURE 6. *Worst-case benchmark, as in Fig. 4, with $n \times n$ random normal matrices such that $\Sigma = \sigma I_p$, $\sigma \in \mathbb{R}$, and $D_p \sim \mathcal{N}(0,1)$. The Schur vectors are sampled with Haar-distribution on* $\mathrm{O}(n)$ *[64]. The plot shows $n \in \{10, 32, 100, 316, 1000\}$.*

7. Application: Riemannian barycenter on manifolds. The Riemannian barycenter is a classical problem from statistics on manifolds, in particular on $\mathrm{SO}(n)$ and $\mathrm{St}(n,p)$, see, e.g., [15, 54, 55, 82, 85]. For both manifolds, intensive computations of matrix logarithms of orthogonal matrices are required, as detailed next for $\mathrm{SO}(n)$. Given a collection of N matrices $\{X_i\}_{i=1}^N$ where $X_i \in \mathrm{SO}(n)$ for $i = 1, \dots, N$, one is interested in finding a Riemannian center of mass $X_{\mathrm{c}} \in \mathrm{SO}(n)$ of the collection, i.e., a matrix satisfying

$$X_{\mathrm{c}} \in \underset{X \in \mathrm{SO}(n)}{\operatorname{argmin}} \frac{1}{2} \sum_{i=1}^{N} d(X_i, X)^2, \tag{7.1}$$

where $d(X_{\mathrm{a}}, X_{\mathrm{b}}) \coloneqq \|\log(X_{\mathrm{a}}^\top X_{\mathrm{b}})\|_{\mathrm{F}}$ is the Riemannian distance function on $\mathrm{SO}(n)$. Well-posedness of the problem, existence and uniqueness of X_{c} are discussed in [2]. The definition (7.1) extends the usual notion of center of mass in the Euclidean space. To solve (7.1), we consider the Riemannian gradient descent [3, 54], i.e., for $k = 1, 2, \dots$ we compute

$$X_{\mathrm{c}}^{k+1} = X_{\mathrm{c}}^{k} \exp\left(-\frac{1}{N} \sum_{i=1}^{N} \log(X_i^\top X_{\mathrm{c}}^k) \right). \tag{7.2}$$

To test the performance of Algorithm 4.1, we perform 100 iterations of Riemannian gradient descent using (7.2). We perform a fixed number of iterations instead of taking a convergence criterion to evaluate more clearly the impact of the RSD routine. The Riemannian gradient step is computed using the relation $\log(A) = Q \log(S) Q^\top$ (see (2.3)). The RSD is computed using either `schur` (`LAPACK`'s `gees`) or `nrmschur` (Algorithm 4.1). We vary the size n of $\mathrm{SO}(n)$, and the number N of samples in the collection. The results are displayed in Table 2. The larger n and N, the larger the improvement. Indeed, the factor of improvement in running time varies from 1.72 in the smallest configuration to 4.38 for $(N = 64, n = 200)$.

8. Conclusion. In this paper, we proposed a new method for computing the RSD of a dense real normal matrix, or equivalently, its eigenvalue decomposition. We showed that the skew-symmetric part of the matrix is a valuable tool for obtaining this decomposition, as its two-dimensional invariant subspaces coincide with those of the original matrix.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

TABLE 2
Time in seconds to perform 100 *steps of Riemannian gradient descent (RGD) for the Riemannian barycenter problem. The sizes n of* SO(n) *vary from* 25 *to* 200 *and the number of samples from* 16 *to* 64. *Orthogonal EVDs are, respectively, computed with* `schur` *and* `nrmschur`.

N \ n	25		50		100		200	
	`schur`	`nrmschur`	`schur`	`nrmschur`	`schur`	`nrmschur`	`schur`	`nrmschur`
16	0.307	0.178	1.31	0.663	10.1	2.56	42.9	11.6
32	0.570	0.327	2.90	1.35	21.0	4.97	77.4	17.8
64	1.02	0.527	5.87	2.23	43.2	9.56	185	42.2

Our approach was the following. First, we developed the matrix theory results that motivate the algorithm. This naturally led to an exact arithmetic version of the method, which we detail in Algorithm 3.1. Next, we examined the method's robustness under floating point arithmetic, resulting in the floating point variant Algorithm 4.1.

By analyzing the sensitivity of the invariant subspaces of the skew-symmetric part, we theoretically quantified in section 4 and numerically corroborated in section 5 how an unfavorable distribution of eigenvalues can reduce the numerical accuracy of the RSD if no additional computations are performed. While this loss of accuracy may often be negligible in applications such as Riemannian computations, we demonstrated how it can be mitigated by applying correction steps to selected inaccurate invariant subspaces, using the bounds provided in Theorems 4.4, 4.6 and 4.7. The numerical accuracy of Algorithm 4.1 was verified in section 5.

We then analyzed the computational complexity of the method. In particular, we showed that in scenarios such as Haar-distributed orthogonal matrices, the cost of Algorithm 4.1 is comparable to that of performing a Hessenberg reduction followed by assembling the Householder reflectors, i.e., $\sim \frac{14}{3}n^3$.

Finally, we highlighted the relevance of this method in the context of statistics on manifolds and computations on the Stiefel manifold and the orthogonal group, where the Riemannian logarithm plays a central role and requires an efficient eigenvalue decomposition.

Appendix A. Special invariance groups. In this section, we provide important lemmas about invariance groups, and we characterize two of them in Lemmas A.2 and A.3.

LEMMA A.1. *For every $A \in \mathbb{R}^{n\times n}$, all $Q \in \mathrm{ig}(A)$ and all $V \in \mathrm{O}(n)$, we have $VQV^\top \in \mathrm{ig}(VAV^\top)$. In particular, V can be a permutation matrix.*

Proof. Since $QAQ^\top = A$ and $V^\top V = I_n$, we have $(VQV^\top)(VAV^\top)(VQV^\top) = (VAV^\top)$. ∎

LEMMA A.2. *Let $\Sigma \in \mathbb{R}^{p\times p}$ be a diagonal matrix with distinct nonzero diagonal values. An orthogonal matrix $Q \in \mathrm{O}(2p)$ is such that*

$$Q\begin{bmatrix}0 & -\Sigma\\ \Sigma & 0\end{bmatrix}Q^\top = \begin{bmatrix}0 & -\Sigma\\ \Sigma & 0\end{bmatrix}, \tag{A.1}$$

if and only if $Q = \begin{bmatrix}\cos(\Phi) & -\sin(\Phi)\\ \sin(\Phi) & \cos(\Phi)\end{bmatrix}$ for some diagonal matrix $\Phi \in \mathbb{R}^{p\times p}$.

Proof. We proceed by induction. Assume first $p = 1$, i.e., $\Sigma = \sigma \in \mathbb{R}$. Then, for every $Q \in \mathrm{O}(2)$, there is $\phi \in \mathbb{R}$ such that either $Q = \begin{bmatrix}\cos(\phi) & -\sin(\phi)\\ \sin(\phi) & \cos(\phi)\end{bmatrix}$ (Q is a rotation) or $Q = \begin{bmatrix}\cos(\phi) & \sin(\phi)\\ \sin(\phi) & -\cos(\phi)\end{bmatrix}$ (Q is a reflection). It

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

is easily verified that only rotations belong to ig $\left(\left[\begin{smallmatrix} 0 & -\sigma \\ \sigma & 0 \end{smallmatrix}\right]\right)$. Thus, Lemma A.2 holds for $p = 1$.

Assume Lemma A.2 holds for $p \geq 1$ and let $\Sigma \in \mathbb{R}^{p\times p}$, $S_1 = \left[\begin{smallmatrix} 0 & -\Sigma \\ \Sigma & 0 \end{smallmatrix}\right]$ and $S_2 = \left[\begin{smallmatrix} 0 & -\sigma \\ \sigma & 0 \end{smallmatrix}\right]$ with $\sigma \in \mathbb{R}$ be nonzero and distinct from all entries of Σ. Let $Q = \left[\begin{smallmatrix} Q_1 & Q_2 \\ Q_3 & Q_4 \end{smallmatrix}\right] \in \mathrm{O}(2p+2)$ with $Q_1 \in \mathbb{R}^{2p\times 2p}$ and $Q_4 \in \mathbb{R}^{2\times 2}$. Then, if one imposes

$$Q \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} Q^\top = \begin{bmatrix} S_1 & 0 \\ 0 & S_2 \end{bmatrix} \text{ and } Q \in \mathrm{O}(2p+2),$$

it must hold in particular that

$$\begin{cases} Q_3 S_1 Q_3^\top + Q_4 S_2 Q_4^\top = S_2, \\ Q_3 S_1 Q_1^\top + Q_4 S_2 Q_2^\top = 0, \end{cases} \quad \text{and} \quad \begin{cases} Q_1^\top Q_1 + Q_3^\top Q_3 = I_{2p}, \\ Q_1^\top Q_2 + Q_3^\top Q_4 = 0. \end{cases}$$

By multiplying the second equation by Q_1 on the right, we obtain $Q_3 S_1 Q_1^\top Q_1 + Q_4 S_2 Q_2^\top Q_1 = 0$. Leveraging both orthogonality constraints, we obtain $Q_3 S_1 (I_{2p} - Q_3^\top Q_3) - Q_4 S_2 Q_4^\top Q_3 = 0$. Introducing the first equation, we obtain $Q_3 S_1 = S_2 Q_3$. Let $Q_3 := [Q_{31}\ Q_{32}]$ such that

$$\begin{aligned} Q_3 S_1 = S_2 Q_3 &\iff Q_{32}\Sigma = S_2 Q_{31} \text{ and } -Q_{31}\Sigma = S_2 Q_{32}. \\ &\iff Q_{31}\Sigma^2 = \sigma^2 Q_{31} \text{ and } Q_{32}\Sigma^2 = \sigma^2 Q_{32}. \end{aligned}$$

Since σ is distinct from all entries of Σ, $Q_3 = 0$. Hence, $Q_4 \in \mathrm{O}(2)$ and $Q_2 = 0$. Finally, $Q \in$ ig $\left(\left[\begin{smallmatrix} S_1 & 0 \\ 0 & S_2 \end{smallmatrix}\right]\right)$ if and only if it is block-diagonal, $Q_1 \in \mathrm{ig}(S_1)$ and $Q_4 \in \mathrm{ig}(S_2) = \mathrm{SO}(2)$.

To conclude, we can define $\widetilde{\Sigma} := \mathrm{diag}(\Sigma, \sigma)$ and take a permutation matrix P such that $P \left[\begin{smallmatrix} S_1 & 0 \\ 0 & S_2 \end{smallmatrix}\right] P^\top = \left[\begin{smallmatrix} 0 & -\widetilde{\Sigma} \\ \widetilde{\Sigma} & 0 \end{smallmatrix}\right]$. Applying Lemma A.1 is enough to conclude that Lemma A.2 holds for $p+1$. Therefore, it holds for all $p \geq 1$. ∎

Lemma A.3. *An orthogonal matrix* $Q = \begin{bmatrix} A & C \\ B & D \end{bmatrix} \in \mathrm{O}(2m)$, $A \in \mathbb{R}^{m\times m}$, *is such that*

$$Q \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} Q^\top = \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix}, \tag{A.2}$$

if and only if $A = D$ *and* $B = -C$. *The matrix* Q *is called* ortho-symplectic *[32].*

Proof. It is enough to notice that (A.2) is equivalent to

$$Q \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} = \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} Q \iff \begin{bmatrix} C & -A \\ D & -B \end{bmatrix} = \begin{bmatrix} -B & -D \\ A & C \end{bmatrix}. \tag{A.3}$$

Therefore $Q \in \mathrm{O}(2m)$, $A = D$ and $B = -C$ are necessary and sufficient conditions. ∎

Finally, we briefly recall the relation between the distance to the orthogonal group $d_{\mathrm{O}(n)}(\widehat{Q})$ from (4.5) and $\|Q^\top Q - I\|_{\mathrm{F}}$. These measures are equivalent for $d_{\mathrm{O}(n)}(\widehat{Q})$ small.

Lemma A.4. *For all matrices* $\widehat{Q} \in \mathbb{R}^{n\times n}$, *it holds that*

$$d_{\mathrm{O}(n)}(\widehat{Q}) \leq \|\widehat{Q}^\top \widehat{Q} - I\|_{\mathrm{F}} \leq 2 d_{\mathrm{O}(n)}(\widehat{Q}) + d_{\mathrm{O}(n)}(\widehat{Q})^2.$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Proof. Let Σ be the diagonal matrix of the singular values of $\widehat{Q}$. Then, it is well known that $d_{\mathrm{O}(n)}(\widehat{Q}) = \|\Sigma - I\|_{\mathrm{F}}$. Moreover, $\|\widehat{Q}^\top\widehat{Q} - I\|_{\mathrm{F}} = \|\Sigma^2 - I\|_{\mathrm{F}}$. It follows that $\|\Sigma^2 - I\|_{\mathrm{F}} = \|(\Sigma - I)(\Sigma + I)\|_{\mathrm{F}} \geq \|\Sigma - I\|_{\mathrm{F}}$. Moreover,

$$\|\Sigma^2 - I\|_{\mathrm{F}} = \|(\Sigma - I + I)^\top(\Sigma - I + I) - I\|_{\mathrm{F}} \leq 2\|\Sigma - I\|_{\mathrm{F}} + \|\Sigma - I\|_{\mathrm{F}}^2.$$

This concludes the proof. ∎

Appendix B. Symmetric skew-Hamiltonian matrices. In this section, we highlight two important properties of the symmetric skew-Hamiltonian matrix $\left[\begin{smallmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{smallmatrix}\right]$ from Theorem 3.4. This matrix appears when a normal matrix has complex eigenvalues with repeated imaginary parts.

Lemma B.1. *A matrix $A \in \mathbb{R}^{2m\times 2m}$ is symmetric skew-Hamitonian, i.e., $A = \left[\begin{smallmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{smallmatrix}\right]$ with $\widetilde{H}$ symmetric and $\widetilde{\Omega}$ skew-symmetric, if and only if there is an eigenvalue decomposition $A = Q\left[\begin{smallmatrix}\Lambda & 0\\ 0 & \Lambda\end{smallmatrix}\right]Q^\top$ where Q is ortho-symplectic and Λ is diagonal.*

Proof. ($\Longrightarrow$) Assume $A = \left[\begin{smallmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{smallmatrix}\right]$. Since A is symmetric, all eigenvalues are real. Moreover, $\left[\begin{smallmatrix}v_1\\ v_2\end{smallmatrix}\right]$ is a normed eigenvector of A if and only if $\left[\begin{smallmatrix}-v_2\\ v_1\end{smallmatrix}\right]$ is an eigenvector of A. Indeed,

$$\begin{bmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{bmatrix}\begin{bmatrix}v_1\\ v_2\end{bmatrix} = \lambda\begin{bmatrix}v_1\\ v_2\end{bmatrix} \iff \begin{cases}(\widetilde{H} - \lambda I)v_1 = \widetilde{\Omega}v_2,\\ (\widetilde{H} - \lambda I)v_2 = -\widetilde{\Omega}v_1.\end{cases} \iff \begin{bmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{bmatrix}\begin{bmatrix}-v_2\\ v_1\end{bmatrix} = \lambda\begin{bmatrix}-v_2\\ v_1\end{bmatrix}. \tag{B.1}$$

Therefore, the multiplicity of every eigenvalue is even and greater or equal to two. Since (B.1) holds for every eigenvector, the eigenvector's matrix always admits a form $Q = \left[\begin{smallmatrix}E & -F\\ F & E\end{smallmatrix}\right]$ with $A = Q\left[\begin{smallmatrix}\Lambda & 0\\ 0 & \Lambda\end{smallmatrix}\right]Q^\top$. By Lemma A.3, Q is ortho-symplectic.

($\Longleftarrow$) Since Q is ortho-symplectic, $Q = \left[\begin{smallmatrix}E & -F\\ F & E\end{smallmatrix}\right]$ by Lemma A.3. Multiplying $Q\left[\begin{smallmatrix}\Lambda & 0\\ 0 & \Lambda\end{smallmatrix}\right]Q^\top$ such as in (3.12) concludes the proof. ∎

Theorem B.2. *For every symmetric skew-Hamiltonian matrix $A \in \mathbb{R}^{2m\times 2m}$, i.e., $A = \left[\begin{smallmatrix}\widetilde{H} & -\widetilde{\Omega}\\ \widetilde{\Omega} & \widetilde{H}\end{smallmatrix}\right]$ with $\widetilde{H}$ symmetric and $\widetilde{\Omega}$ skew-symmetric, if* $\operatorname{rank}(A) = 2m$ *and every eigenvalue of A has multiplicity* 2*, then the m first steps of Lanczos tridiagonalization are enough in exact arithmetic to build an ortho-symplectic matrix Y where $A = Y\left[\begin{smallmatrix}T_m & 0\\ 0 & T_m\end{smallmatrix}\right]Y^\top$ and T_m is an $m \times m$ tridiagonal symmetric matrix.*

Proof. We denote $K = [b\ Ab\ \ldots A^{2m-1}b]$ a Krylov basis of A generated by b. Choose any $v \in \mathbb{R}^n$ such that $b \coloneqq Av \neq 0$ and $K_{1:m}$ has column rank m.[4] By Lemma B.1, we know that there is an ortho-symplectic matrix Q such that $A = Q\left[\begin{smallmatrix}\Lambda & 0\\ 0 & \Lambda\end{smallmatrix}\right]Q^\top$ and $\Lambda \in \mathbb{R}^{m\times m}$ is diagonal. Let k_j denote the jth column of K, it follows that

$$k_j = Q\begin{bmatrix}\Lambda^j & 0\\ 0 & \Lambda^j\end{bmatrix}Q^\top v.$$

[4]Notice that $\operatorname{rank}(K) \leq m$ since every eigenvalue has multiplicity 2 by Lemma B.1.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Moreover, for any integer pair i, j, with $1 \leq i, j \leq n$, we have

$$
\begin{aligned}
k_i^\top \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} k_j &= v^\top Q \begin{bmatrix} \Lambda^i & 0 \\ 0 & \Lambda^i \end{bmatrix} Q^\top \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} Q \begin{bmatrix} \Lambda^j & 0 \\ 0 & \Lambda^j \end{bmatrix} Q^\top v \\
&= v^\top Q \begin{bmatrix} \Lambda^i & 0 \\ 0 & \Lambda^i \end{bmatrix} \begin{bmatrix} 0 & -I_m \\ I_m & 0 \end{bmatrix} \begin{bmatrix} \Lambda^j & 0 \\ 0 & \Lambda^j \end{bmatrix} Q^\top v \\
&= v^\top Q \begin{bmatrix} 0 & -\Lambda^{i+j} \\ \Lambda^{i+j} & 0 \end{bmatrix} Q^\top v \\
&= 0. \qquad \text{(B.2)}
\end{aligned}
$$

The last equality stands because $v^\top \Omega v = 0$ for every skew-symmetric matrices Ω. Letting $J_{2m} \coloneqq \left[\begin{smallmatrix} 0 & -I_m \\ I_m & 0 \end{smallmatrix}\right]$, it follows from (B.2) that $K_{1:m}^\top J_{2m} K_{1:m} = 0$. Therefore, the matrix $[K_m \;\; J_{2m}K_{1:m}]$ is full rank. Consequently, the Gram-Schmidt QR factorization of $K_{1:m}$, namely, $K_{1:m} = Y_{1:m}R$, generates $Y \coloneqq [Y_{1:m} \;\; J_{2m}Y_{1:m}]$ where Y is ortho-symplectic. Moreover, we also have

$$
k_i^\top A J_{2m} k_j = k_{i+1}^\top J_{2m} k_j = 0, \qquad \text{(B.3)}
$$

and thus $K_{1:m}^\top A J_{2m} K_{1:m} = 0$. Finally, it is the praised property of Lanczos algorithm that $Y_{1:m}^\top A Y_{1:m} = T_m$ where T_m is tridiagonal and symmetric [41, Thm. 9.1.1]. Combining (B.2) and (B.3), we obtain

$$
\begin{aligned}
Y^\top A Y &= \begin{bmatrix} Y_{1:m}^\top \\ Y_{1:m}^\top J_{2m}^\top \end{bmatrix} A [Y_{1:m} \;\; J_{2m}Y_{1:m}] \\
&= \begin{bmatrix} Y_{1:m}^\top A Y_{1:m} & Y_{1:m}^\top A J_{2m} Y_{1:m} \\ Y_{1:m}^\top J_{2m}^\top A Y_{1:m} & Y_{1:m}^\top J_{2m}^\top A J_{2m} Y_{1:m} \end{bmatrix} \\
&= \begin{bmatrix} T_m & 0 \\ 0 & T_m \end{bmatrix}.
\end{aligned}
$$

We used that $J_{2m}^\top A J_{2m} = A$. This concludes the proof. ∎

Algorithm B.1 is a simple modification of the usual modified Gram-Schmidt to build a matrix that is not only orthogonal but also symplectic. In floating point arithmetic, the matrix Y is approximately orthogonal but *exactly* symplectic by construction. Instead of orthogonalizing any matrix, Algorithm B.1 orthogonalizes the Krylov basis generated by A symmetric skew-Hamiltonian, we directly obtain Lanczos's procedure with full re-orthogonalization.

In Fig. 7, we evaluate how much Theorem B.2 can improve the performance compared to `LAPACK`'s symmetric eigenvalue routine for small EVP's. The method using only m steps of Lanczos tridiagonalization is called `wxeigen` and is available at wxeigen.jl in the repository. We observe that `wxeigen` with full re-orthogonalization achieves good running time performance and accuracy.

Appendix C. Proof of Theorem 4.7.

Proof. Let $A = QSQ^\top \in \mathbb{R}^{n\times n}$ be a normal matrix with real eigenvalue $\breve{\lambda}_k$ for $k \in \mathcal{K} \coloneqq \{1, ..., r\}$. Moreover, assume $\lambda_j e^{\pm i\theta_j}$ is an eigenvalue of A for $j \in \mathcal{I}$. Finally, assume we have computed $\widehat{V} \in \mathrm{St}(n, r)$ such that $\Omega\widehat{V} = E$ with $\|E\|_\mathrm{F} \leq \tau\varepsilon_\mathrm{m}\|A\|_\mathrm{F}$. We have

$$
\Omega\widehat{V} = E \iff \mathrm{skew}(S)Q^\top\widehat{V} = Q^\top E \eqqcolon \widetilde{E}. \qquad \text{(C.1)}
$$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

Algorithm B.1 Re-orthogonalized Lanczos for symmetric skew-Hamiltonian matrix.

INPUT: $v \in \mathbb{R}^{2m}$, $A \in \mathbb{R}^{2m\times 2m}$.
$\beta_0 = 0$, $y_1 = Av/\|Av\|_2$ and $y_{m+1} = J_{2m}y_1$.
for $i = 2, ..., m$ **do**
  $y_i = Ay_{i-1}$ and $\alpha_{i-1} = y_{i-1}^\top y_i$.
  $y_i = y_i - \alpha_{i-1}y_{i-1} - \beta_{i-2}y_{i-2}$. Lanczos step.
  for $j = 1, .., i-1$ **do**
    $y_i = y_i - (y_i^\top y_j)y_j - (y_i^\top y_{m+j})y_{m+j}$. Full re-orthogonalization.
  end for
  $\beta_i = \|y_i\|_2$.
  $y_i = y_i/\beta_i$ and $y_{m+i} = J_{2m}y_i$. Build symplectic basis.
end for
$\alpha_m = y_m A y_m$.
return $Y \in \mathrm{O}(2m) \cap \mathrm{OSp}(2m)$, $\{\alpha_i\}_{i=1}^m$ and $\{\beta_i\}_{i=1}^{m-1}$.

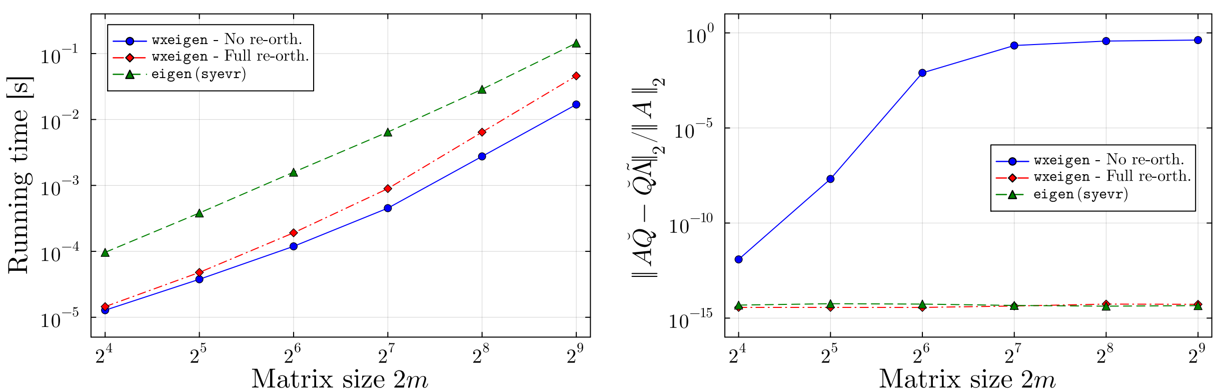

FIGURE 7. *Numerical performance evaluation of* `wxeigen`*, without and with full re-orthogonalization of the Lanczos vectors, compared to* `eigen` *(LAPACK's* `syevr`*). The plot on the left shows the running time performance. The plot on the right shows the accuracy performance in spectral norm. Taking advantage of the symmetric skew-Hamiltonian structure allows to design faster methods than the generic symmetric eigensolver.*

Defining $X := Q^\top \widehat{V}$, there are two permutation $P_\mathrm{l} \in \mathrm{O}(n), P_\mathrm{r} \in \mathrm{O}(r)$ such that $P_\mathrm{l}^\top X P_\mathrm{r} = \begin{bmatrix} X_{\mathcal{I},\mathcal{K}} \\ X_{\mathcal{K},\mathcal{K}} \end{bmatrix}$ where the block $X_{\mathcal{I},\mathcal{K}}$ comprises all 2×1 blocks X_{ik} such that for $i \in \mathcal{I}, k \in \mathcal{K}$, by (C.1), we have

$$\begin{bmatrix} 0 & -\lambda_i s_i \\ \lambda_i s_i & 0 \end{bmatrix} X_{ik} = (P_\mathrm{l}^\top \widetilde{E} P_\mathrm{r})_{ik}.$$

Since $X \in \mathrm{St}(n, r)$, we have

$$X_{\mathcal{I},\mathcal{K}}^\top X_{\mathcal{I},\mathcal{K}} + X_{\mathcal{K},\mathcal{K}}^\top X_{\mathcal{K},\mathcal{K}} = I_r \Longrightarrow \|X_{\mathcal{K},\mathcal{K}}^\top X_{\mathcal{K},\mathcal{K}} - I_r\|_\mathrm{F} \leq \|X_{\mathcal{I},\mathcal{K}}\|_\mathrm{F}^2$$
$$\text{and,} \quad \|X_{\mathcal{I},\mathcal{K}}\|_\mathrm{F}^2 = \sum_{i\in\mathcal{I},k\in\mathcal{K}} \frac{1}{|\lambda_i s_i|^2} \|(P_\mathrm{l}^\top \widetilde{E} P_\mathrm{r})_{ik}\|_\mathrm{F}^2 \leq \frac{\|E\|_\mathrm{F}^2}{\min_{i\in\mathcal{I}} |\lambda_i s_i|^2},$$

where [14, Thm. VII.2.8] is used. The deviation of $X_{\mathcal{K},\mathcal{K}}$ from orthogonality is quantified by that of $X_{\mathcal{I},\mathcal{K}}$ from zero. In addition, by taking the SVD $X_{\mathcal{K},\mathcal{K}} = U_X \Sigma_X V_X^\top$, we can write $\|X_{\mathcal{K},\mathcal{K}}^\top X_{\mathcal{K},\mathcal{K}} - I_r\|_\mathrm{F} = \|\Sigma_X^2 - I_r\|_\mathrm{F} \geq$

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

$\|\Sigma_X - I_r\|_{\mathrm{F}}$ by Lemma A.4. Define $\begin{bmatrix} \widetilde{S}_{n-r} & 0 \\ 0 & \breve{\Lambda} \end{bmatrix} := P_1^\top S P_1$. For all $\breve{R} \in \mathrm{O}(r)$, we have

$$\begin{aligned} \|A\widehat{V}\breve{R} - \widehat{V}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} &\leq \|\mathrm{sym}(S)XR - XR\breve{\Lambda}\|_{\mathrm{F}} + \|E\|_{\mathrm{F}} \\ &\leq \|\mathrm{sym}(\widetilde{S}_{n-r})X_{\mathcal{I},\mathcal{K}}\breve{R} - X_{\mathcal{I},\mathcal{K}}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} \\ &\quad + \|\breve{\Lambda}X_{\mathcal{K},\mathcal{K}}\breve{R} - X_{\mathcal{K},\mathcal{K}}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} + \|E\|_{\mathrm{F}}. \end{aligned} \tag{C.2}$$

The first term of (4.12) can be bounded by

$$\begin{aligned} \|\mathrm{sym}(\widetilde{S}_{n-r})X_{\mathcal{I},\mathcal{K}}\breve{R} - X_{\mathcal{I},\mathcal{K}}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} &\leq \max_{i\in\mathcal{I},k\in\mathcal{K}} |\lambda_i c_i - \breve{\lambda}_k| \|X_{\mathcal{I},\mathcal{K}}\breve{R}\|_{\mathrm{F}} \\ &\leq \frac{\max_{i\in\mathcal{I},k\in\mathcal{K}} |\lambda_i c_i - \breve{\lambda}_k|}{\min_{i\in\mathcal{I}} |\lambda_i s_i|} \|E\|_{\mathrm{F}}. \end{aligned} \tag{C.3}$$

Moreover, the second term can also be bounded by

$$\begin{aligned} \|\breve{\Lambda}X_{\mathcal{K},\mathcal{K}}\breve{R} - X_{\mathcal{K},\mathcal{K}}\breve{R}\breve{\Lambda}\|_{\mathrm{F}} &= \|\breve{\Lambda}(X_{\mathcal{K},\mathcal{K}}\breve{R} - I_r) - (X_{\mathcal{K},\mathcal{K}}\breve{R} - I_r)\breve{\Lambda}\|_{\mathrm{F}} \\ &\leq \max_{k_1,k_2\in\mathcal{K}} |\breve{\lambda}_{k_1} - \breve{\lambda}_{k_2}| \|X_{\mathcal{K},\mathcal{K}}\breve{R} - I_r\|_{\mathrm{F}}. \end{aligned}$$

In particular, for $\breve{R} = V_X U_X^\top$, which minimizes $\|X_{\mathcal{K},\mathcal{K}}\breve{R} - I_r\|_{\mathrm{F}}$, we have

$$\|X_{\mathcal{K},\mathcal{K}}\breve{R} - I_r\|_{\mathrm{F}} = \|\Sigma_X - I_r\|_{\mathrm{F}} \leq \frac{\|E\|_{\mathrm{F}}^2}{\min_{i\in\mathcal{I}} |\lambda_i s_i|^2}. \tag{C.4}$$

Inserting (C.3) and (C.4) in (C.2) results in Theorem 4.7. ∎

Acknowledgment. The authors would like to thank Paul Van Dooren and Pierre-Antoine Absil for their invaluable advice during the preparation of this manuscript.

REFERENCES

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, Princeton, NJ, 2008. http://press.princeton.edu/titles/8586.html.

[2] B. Afsari. Riemannian Lp center of mass: Existence, uniqueness, and convexity. *Proc. Amer. Math. Soc.*, 139(2):655–673, 2011. http://www.jstor.org/stable/41059320.

[3] B. Afsari, R. Tron, and R. Vidal. On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM J. Control Optim.*, 51(3):2230–2260, 2013. https://doi.org/10.1137/12086282X.

[4] G.S. Ammar, W.B. Gragg, and L. Reichel. On the eigenproblem for orthogonal matrices. In *1986 25th IEEE Conference on Decision and Control*, 1963–1966, 1986. https://doi.org/10.1109/CDC.1986.267356.

[5] G.S. Ammar, W.B. Gragg, and L. Reichel. Determination of Pisarenko frequency estimates as eigenvalues of an orthogonal matrix. In: F.T. Luk (editor), *Advanced Algorithms and Architectures for Signal Processing II*, Vol. 0826. International Society for Optics and Photonics, SPIE, 143–145, 1988. https://doi.org/10.1117/12.942026.

[6] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*, third edition. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.

[7] T.W. Anderson, I. Olkin, and L.G. Underhill. Generation of random orthogonal matrices. *SIAM J. Sci. Stat. Comput.*, 8(4):625–629, 1987. https://doi.org/10.1137/0908055.

[8] J.L. Aurentz, T. Mach, R. Vandebril, and D.S. Watkins. Fast and stable unitary QR algorithm. *Electron. Trans. Numer. Anal.*, 44:327–341, 2015. http://etna.math.kent.edu/vol.44.2015/pp327-341.dir/pp327-341.pdf.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

[9] J.L. Aurentz, T. Mach, R. Vandebril, and D.S. Watkins. Computing the eigenvalues of symmetric tridiagonal matrices via a Cayley transformation. *Electron. Trans. Numer. Anal.*, 46:447–459, 2017. http://etna.math.kent.edu/vol.46.2017/pp447-459.dir/pp447-459.pdf.

[10] S. Batterson. Convergence of the Francis shifted QR algorithm on normal matrices. *Linear Algebra Appl.*, 207:181–195, 1994. https://doi.org/10.1016/0024-3795(94)90010-8.

[11] T. Bendokat, R. Zimmermann, and P.-A. Absil. A Grassmann manifold handbook: Basic geometry and computational aspects. *Adv. Computat. Math.*, 50:Article number 6, 2024. https://doi.org/10.1007/s10444-023-10090-8.

[12] P. Benner and H. Fassbender. An implicitly restarted symplectic Lanczos method for the Hamiltonian eigenvalue problem. *Linear Algebra Appl.*, 263:75–111, 1997. https://doi.org/10.1016/S0024-3795(96)00524-1.

[13] P. Benner, D. Kressner, and V. Mehrmann. Skew-Hamiltonian and Hamiltonian eigenvalue problems: Theory, algorithms and applications. In: Z. Drmač, M. Marušić, and Z. Tutek (editors), *Proceedings of the Conference on Applied Mathematics and Scientific Computing*, Dordrecht, 2005. Springer Netherlands, 3–39. https://link.springer.com/chapter/10.1007/1-4020-3197-1_1.

[14] R. Bhatia. *Matrix Analysis*, Vol. 169. Springer New York, NY, 1997. https://doi.org/10.1007/978-1-4612-0653-8.

[15] D.A. Bini and B. Iannazzo. Computing the Karcher mean of symmetric positive definite matrices. *Linear Algebra Appl.*, 438(4):1700–1710, 2013. 16th ILAS Conference Proceedings, Pisa 2010. https://doi.org/10.1016/j.laa.2011.08.052.

[16] D.A. Bini, B. Iannazzo, B. Jeuris, and R. Vandebril. Geometric means of structured matrices. *Bit Numer. Math.*, 54(1):55–83, 2014. https://doi.org/10.1007/s10543-013-0450-4.

[17] C. Bischof and C. Van Loan. The WY representation for products of Householder matrices. *SIAM J. Sci. Stat. Comput.*, 8(1):s2–s13, 1987. https://doi.org/10.1137/0908009.

[18] L.S. Blackford, A. Petitet, R. Pozo, K. Remington, R.C. Whaley, J. Demmel, J. Dongarra, I. Duff, S. Hammarling, G. Henry, et al. An updated set of basic linear algebra subprograms (BLAS). *ACM Trans. Math. Softw.*, 28(2):135–151, 2002.

[19] A. Bunse-Gerstner, R. Byers, and V. Mehrmann. Numerical methods for simultaneous diagonalization. *SIAM J. Matrix Anal. Appl.*, 14, 1998. https://doi.org/10.1137/0614062.

[20] A. Bunse-Gerstner and L. Elsner. Schur parameter pencils for the solution of the unitary eigenproblem. *Linear Algebra Appl.*, 154–156:741–778, 1991. https://doi.org/10.1016/0024-3795(91)90402-I.

[21] D. Calvetti, L. Reichel, and H. Xu. A CS decomposition for orthogonal matrices with application to eigenvalue computation. *Linear Algebra Appl.*, 476:197–232, 2015. https://doi.org/10.1016/j.laa.2015.03.007.

[22] R. Chakraborty and B.C. Vemuri. Statistics on the Stiefel manifold: Theory and applications. *Ann. Stat.*, 2019. http://api.semanticscholar.org/CorpusID:13671005.

[23] J.-P. Charlier and P. van Dooren. A Jacobi-like algorithm for computing the generalized Schur form of a regular pencil. In: H.A. van der Vorst and P. van Dooren (editors), *Parallel Algorithms for Numerical Linear Algebra*, Vol. 1. *Advances in Parallel Computing*. North-Holland, 17–36, 1990. https://doi.org/10.1016/B978-0-444-88621-7.50007-4.

[24] J. Cuppen. A divide and conquer method for the symmetric eigenproblem. *Numer. Math.*, 36:177–195, 1980. https://doi.org/10.1007/BF01396757.

[25] G. Cybenko. Computing Pisarenko frequency estimates. In: *Proceedings of the Princeton Conference on Information Science*, Princeton, 1985.

[26] P. Deift, J. Demmel, L.-C. Li, and C. Tomei. The bidiagonal singular value decomposition and hamiltonian mechanics. *SIAM J. Numer. Anal.*, 28(5):1463–1516, 1991. https://doi.org/10.1137/0728076.

[27] J. Demmel and W Kahan. Computing small singular values of bidiagonal matrices with guaranteed high relative accuracy: LAPACK working note number 3, 1988. http://www.osti.gov/biblio/5039344.

[28] J. Demmel and W. Kahan. Accurate singular values of bidiagonal matrices. *SIAM J. Sci. Stat. Comput.*, 11(5):873–912, 1990. https://doi.org/10.1137/0911052.

[29] J. Demmel and K. Veselić. Jacobi's method is more accurate than QR. *SIAM J. Matrix Anal. Appl.*, 13(4):1204–1245, 1992. https://doi.org/10.1137/0613074.

[30] I.S. Dhillon. *A new $O(N^2)$ algorithm for the symmetric tridiagonal eigenvalue/eigenvector problem*. PhD thesis, 1997.

[31] J.J. Dongarra and D.C. Sorensen. A fully parallel algorithm for the symmetric eigenvalue problem. *SIAM J. Sci. Stat. Comput.*, 8(2):s139–s154, 1987. https://doi.org/10.1137/0908018.

[32] F.M. Dopico and C.R. Johnson. Parametrization of the matrix symplectic group and applications. *SIAM J. Matrix Anal. Appl.*, 31(2):650–673, 2009. https://doi.org/10.1137/060678221.

[33] S.C. Eisenstat and I.C.F. Ipsen. Three absolute perturbation bounds for matrix eigenvalues imply relative bounds. *SIAM J. Matrix Anal. Appl.*, 20(1):149–158, 1998. https://doi.org/10.1137/S0895479897323282.

[34] M. Fasi and L. Robol. Sampling the eigenvalues of random orthogonal and unitary matrices. *Linear Algebra Appl.*, 620:297–321, 2021. https://doi.org/10.1016/j.laa.2021.02.031.

[35] H. Fassbender, D.S. Mackey, N. Mackey, and H. Xu. Hamiltonian square roots of skew-Hamiltonian matrices. *Linear Algebra Appl.*, 287(1):125–159, 1999. https://doi.org/10.1016/S0024-3795(98)10137-4.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

[36] K. Fernando and B. Parlett. Accurate singular values and differential qd algorithms. *Numer. Math.*, 67:191–229, 03 1994. https://doi.org/10.1007/s002110050024.

[37] M. Ferranti and R. Vandebril. Computing eigenvalues of normal matrices via complex symmetric matrices. *J. Comput. Appl. Math.*, 259(PART A), 2013-03-01. https://doi.org/10.1016/j.cam.2013.08.036.

[38] J.G.F. Francis. The QR transformation A unitary analogue to the LR transformation - Part 1. *Comput. J.*, 4:265–271, 1961. https://doi.org/10.1093/comjnl/4.3.265.

[39] J.G.F. Francis. The QR transformation–Part 2. *Comput. J.*, 4(4):332–345, 1962. https://doi.org/10.1093/comjnl/4.4.332.

[40] H.H. Goldstine and L.P. Horwitz. A procedure for the diagonalization of normal matrices. *J. ACM*, 6(2):176–195, 1959. https://doi.org/10.1145/320964.320975.

[41] G.H. Golub and C.F. Van Loan. *Matrix Computations*, third edition. The Johns Hopkins University Press, Baltimore MD, 1996.

[42] K. Goto and R. Van De Geijn. High-performance implementation of the level-3 BLAS. *ACM Trans. Math. Softw.* 35(1):Article No.: 4, 2008. https://doi.org/10.1145/1377603.1377607.

[43] W. Gragg and L. Reichel. A divide and conquer method for unitary and orthogonal eigenproblems. *Numer. Math.*, 57:695–718, 12 1990. https://doi.org/10.1007/BF01386438.

[44] W.B. Gragg. The QR algorithm for unitary Hessenberg matrices. *J. Comput. Appl. Math.*, 16(1):1–8, 1986. https://doi.org/10.1016/0377-0427(86)90169-X.

[45] B. Großer and B. Lang. An O(n^2) algorithm for the bidiagonal SVD. *Linear Algebra Appl.*, 358(1):45–70, 2003. https://doi.org/10.1016/S0024-3795(01)00398-6.

[46] M. Gu, R. Guzzo, X.-b. Chi, and X.-Q. Cao. A stable divide and conquer algorithm for the unitary eigenproblem. *SIAM J. Matrix Anal. Appl.*, 25(2):385–404, 2003. https://doi.org/10.1137/S0895479899359539.

[47] H. He and D. Kressner. Randomized joint diagonalization of symmetric matrices. *SIAM J. Matrix Anal. Appl.*, 45(1):661–684, 2024. https://doi.org/10.1137/22M1541265.

[48] H. He and D. Kressner. A simple, randomized algorithm for diagonalizing normal matrices. *Calcolo*, 62(3):30, 2025. https://doi.org/10.1007/s10092-025-00654-z.

[49] N.J. Higham. *Accuracy and Stability of Numerical Algorithms*, second edition. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002. https://doi.org/10.1137/1.9780898718027.

[50] A.J. Hoffman and H.W. Wielandt. The variation of the spectrum of a normal matrix. *Duke Math. J.*, 20(1):37–39, 1953. https://doi.org/10.1215/S0012-7094-53-02004-3.

[51] R.A. Horn and C.R. Johnson. *Matrix Analysis*, second edition. Cambridge University Press, Cambridge, New York, 2013.

[52] J. Huang and Z. Jia. A skew-symmetric Lanczos bidiagonalization method for computing several extremal eigenpairs of a large skew-symmetric matrix. *SIAM J. Matrix Anal. Appl.*, 45(2):1114–1147, 2024. https://doi.org/10.1137/23M1553029.

[53] A.T. James. Normal multivariate analysis and the orthogonal group. *Ann. Math. Stat.*, 25(1):40–75, 1954. https://doi.org/10.1214/aoms/1177728846.

[54] K. Krakowski, K. Hüper, and J. Manton. On the computation of the Karcher mean on spheres and special orthogonal groups. 2007. http://www.researchgate.net/publication/228753783_On_the_computation_of_the_Karcher_mean_on_spheres_and_special_orthogonal_groups.

[55] Y. Lim and M. Pálfia. Matrix power means and the Karcher mean. *J. Funct. Anal.*, 262(4):1498–1514, 2012. https://doi.org/10.1016/j.jfa.2011.11.012.

[56] S.-S. Lo, B. Philippe, and A. Sameh. A multiprocessor algorithm for the symmetric tridiagonal eigenvalue problem. *SIAM J. Sci. Stat. Comput.*, 8(2):s155–s165, 1987. https://doi.org/10.1137/0908019.

[57] S. Mataigne, R. Zimmermann, and N. Miolane. An efficient algorithm for the Riemannian Logarithm on the Stiefel Manifold for a family of Riemannian Metrics. *SIAM J. Matrix Anal. Appl.*, 46(2):879–905, 2025. https://doi.org/10.1137/24M1647801.

[58] C.C. Paige. Computational Variants of the Lanczos Method for the Eigenproblem. *IMA J. Appl. Math.*, 10(3):373–381, 1972. https://doi.org/10.1093/imamat/10.3.373.

[59] B.N. Parlett and D.S. Scott. The Lanczos algorithm with selective orthogonalization. *Math. Comput.*, 33(145):217–238, 1979. http://www.jstor.org/stable/2006037.

[60] C. Penke, A. Marek, C. Vorwerk, C. Draxl, and P. Benner. High performance solution of skew-symmetric eigenvalue problems with applications in solving the Bethe-Salpeter eigenvalue problem. *Parallel Comput.*, 96:102639, 2020. https://doi.org/10.1016/j.parco.2020.102639.

[61] Q. Rentmeesters. *Algorithms for data fitting on some common homogeneous spaces*. PhD thesis, Catholic University of Louvain (UCLouvain), 2013. http://dial.uclouvain.be/pr/boreal/fr/object/boreal:132587.

[62] R. Schreiber and C. Van Loan. A storage-efficient WY representation for products of Householder transformations. *SIAM J. Sci. Stat. Comput.*, 10(1):53–57, 1989. https://doi.org/10.1137/0910005.

Electronic Journal of Linear Algebra, ISSN 1081-3810
A publication of the International Linear Algebra Society
Volume 42, pp. 349-386, April 2026.

[63] H. D. Simon. Analysis of the symmetric Lanczos algorithm with reorthogonalization methods. *Linear Algebra Appl.*, 61:101–131, 1984. https://doi.org/10.1016/0024-3795(84)90025-9.

[64] G.W. Stewart. The efficient generation of random orthogonal matrices with an application to condition estimators. *SIAM J. Numer. Anal.*, 17(3):403–409, 1980. https://doi.org/10.1137/0717034.

[65] M. Stewart. An error analysis of a unitary Hessenberg QR algorithm. *SIAM J. Matrix Anal. Appl.*, 28(1):40–67, 2006. https://doi.org/10.1137/04061948X.

[66] B.D. Sutton. Simultaneous diagonalization of nearly commuting Hermitian matrices: do-one-then-do-the-other. *IMA J. Numer. Anal.*, 44(2):1061–1089, 2023. https://doi.org/10.1093/imanum/drad033.

[67] T. Szwagier and X. Pennec. Rethinking the Riemannian logarithm on flag manifolds as an orthogonal alignment problem. In: F. Nielsen and F. Barbaresco (editors), *Geometric Science of Information*. Springer Nature Switzerland, Cham, 375–383, 2023.

[68] S. Tao and Y. Xu. Parallel implementation of red–black ordering preconditioners for accelerating gmres solutions in neutron transport code pandas-moc. *Ann. Nucl. Energy*, 193:110008, 2023. https://doi.org/10.1016/j.anucene.2023.110008.

[69] L.N. Trefethen and D. Bau. *Numerical Linear Algebra*. SIAM, Philadelphia, PA, 1997.

[70] C. Van Loan. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.*, 61:233–251, 1984. https://doi.org/10.1016/0024-3795(84)90034-X.

[71] R. Vandebril. A unitary similarity transform of a normal matrix to complex symmetric form. *Appl. Math. Lett.*, 24(2):160–164, 2011. https://doi.org/10.1016/j.aml.2010.08.037.

[72] K. Veselić and I. Slapničar. Floating-point perturbations of Hermitian matrices. *Linear Algebra Appl.*, 195:81–116, 1993. https://doi.org/10.1016/0024-3795(93)90258-P.

[73] T.-L. Wang and W.B. Gragg. Convergence of the shifted QR algorithm for unitary Hessenberg matrices. *Math. Comput.*, 71(240):1473–1496, 2002. `http://www.jstor.org/stable/4099960`.

[74] T.-L. Wang and W.B. Gragg. Convergence of the unitary QR algorithm with a unimodular Wilkinson shift. *Math. Comput.*, 72(241):375–385, 2003. `http://www.jstor.org/stable/4099996`.

[75] R.C. Ward and L.J. Gray. Eigensystem computation for skew-symmetric and a class of symmetric matrices. *ACM Trans. Math. Softw.*, 4(3):278–285, 1978. https://doi.org/10.1145/355791.355798.

[76] D.S. Watkins. On Hamiltonian and symplectic Lanczos processes. *Linear Algebra Appl.*, 385:23–45, 2004. Special Issue in honor of Peter Lancaster. https://doi.org/10.1016/j.laa.2002.11.001.

[77] J.F. Watters. Simultaneous quasi-diagonalization of normal matrices. *Linear Algebra Appl.*, 9:103–117, 1974. https://doi.org/10.1016/0024-3795(74)90030-5.

[78] J.H. Wilkinson. Householder's method for symmetric matrices. *Numer. Math.*, 4(1):354–361, 1962. https://doi.org/10.1007/BF01386332.

[79] J.H. Wilkinson. Convergence of the LR, QR, and related algorithms. *Comput. J.*, 8(1):77–84, 1965. https://doi.org/10.1093/comjnl/8.1.77.

[80] J.H. Wilkinson. The QR algorithm for real symmetric matrices with multiple eigenvalues. *Comput. J.*, 8(1):85–87, 1965. https://doi.org/10.1093/comjnl/8.1.85.

[81] J.H. Wilkinson. Global convergence of tridiagonal QR algorithm with origin shifts. *Linear Algebra Appl.*, 1(3):409–420, 1968. https://doi.org/10.1016/0024-3795(68)90017-7.

[82] S. Ying, H. Qin, Y. Peng, and Z. Wen. Compute Karcher means on SO(n) by the geometric conjugate gradient method. *Neurocomputing*, 215:169–174, 2016. SI: Stereo Data. https://doi.org/10.1016/j.neucom.2015.05.143.

[83] B.B Zhou and R.P Brent. An efficient method for computing eigenvalues of a real normal matrix. *J. Parallel Distrib. Comput.*, 63(6):638–648, 2003. https://doi.org/10.1016/S0743-7315(03)00007-8.

[84] R. Zimmermann. A matrix-algebraic algorithm for the Riemannian logarithm on the Stiefel manifold under the canonical metric. *SIAM J. Matrix Anal. Appl.*, 38(2):322–342, 2017. https://doi.org/10.1137/16M1074485.

[85] R. Zimmermann and K. Hüper. Computing the Riemannian logarithm on the Stiefel manifold: Metrics, methods, and performance. *SIAM J. Matrix Anal. Appl.*, 43(2):953–980, 2022. https://doi.org/10.1137/21M1425426.