An SVD-Like Matrix Decomposition and Its Applications

Hongguo Xu*

Abstract

A matrix $S \in \mathbb{C}^{2m\times 2m}$ is symplectic if $SJS^* = J$, where $J = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$. Symplectic matrices play an important role in the analysis and numerical solution of matrix problems involving the indefinite inner product $x^*(iJ)y$. In this paper we provide several matrix factorizations related to symplectic matrices. We introduce a singular value-like decomposition $B = QDS^{-1}$ for any real matrix $B \in \mathbb{R}^{n\times 2m}$, where Q is real orthogonal, S is real symplectic, and D is permuted diagonal. We show the relation between this decomposition and the canonical form of real skew-symmetric matrices and a class of Hamiltonian matrices. We also show that if S is symplectic it has the structured singular value decomposition $S = UDV^*$, where U, V are unitary and symplectic, $D = \operatorname{diag}(\Omega, \Omega^{-1})$ and Ω is positive diagonal. We study the BJB^T factorization of real skew-symmetric matrices. The BJB^T factorization has the applications in solving the skew-symmetric systems of linear equations, and the eigenvalue problem for skew-symmetric/symmetric pencils. The BJB^T factorization is not unique, and in numerical application one requires the factor B with small norm and condition number to improve the numerical stability. By employing the singular value-like decomposition and the singular value decomposition of symplectic matrices we give the general formula for B with minimal norm and condition number.

Keywords. Skew-symmetric matrix, symplectic matrix, orthogonal(unitary) symplectic matrix, Hamiltonian matrix, eigenvalue problem, singular value decomposition (SVD), SVD-like decomposition, BJB^T factorization, Schur form, Jordan canonical form.
AMS subject classification. 15A18, 65F35.

1 Introduction

Let $J_m = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix} \in \mathbb{R}^{2m\times 2m}$. We will use J when the size is clear from the context. A matrix $S \in \mathbb{C}^{2m\times 2m}$ is *symplectic* if

$$SJS^* = J.$$

*Department of Mathematics, University of Kansas, Lawrence, KS 66045, USA. `xu@math.ukans.edu`.
This author is supported by NSF under Grant No.EPS-9874732 and matching support from the State of Kansas.

A matrix $S \in \mathbb{C}^{2m\times 2m}$ is *unitary symplectic* if S is both unitary and symplectic. In the real case S is *orthogonal symplectic.* A matrix $A \in \mathbb{C}^{2m\times 2m}$ is *Hamiltonian* if

$$(AJ)^* = AJ.$$

A matrix $A \in \mathbb{C}^{2m\times 2m}$ is *skew-Hamiltonian* if

$$(AJ)^* = -AJ.$$

A symplectic matrix is also called J-orthogonal. Symplectic similarity transformations preserve the structures of Hamiltonian, skew-Hamiltanian and symplectic matrices. Based on this fact symplectic matrices are used as the basic tool in the analysis and the numerical solution of Hamiltonian, skew-Hamiltonian and symplectic eigenvalue problems [21, 7, 9, 17, 14, 1, 10]. Recently in [4, 5] it is showed that every real matrix $B \in \mathbb{R}^{2m\times 2m}$ has the symplectic URV factorization $B = URV^T$, where U, V are orthogonal symplectic and R is block upper triangular. Based on this factorization and its generalization several numerically stable and structure preserving methods have be developed [4, 5, 3].

In this paper we study some other matrix decompositions related to symplectic matrices. Our purpose is to provide some new insights about the symplectic matrices. We show that for any real matrix $B \in \mathbb{R}^{n\times 2m}$ there exists a real orthogonal matrix Q and a real symplectic matrix S such that

$$B = QDS^{-1}, \tag{1}$$

where

$$D = \begin{array}{c} \\ p \\ q \\ p \\ n-2p-q \end{array} \begin{array}{c} \begin{array}{cccccc} p & q & m-p-q & p & q & m-p-q \end{array} \\ \left(\begin{array}{cccccc} \Sigma & 0 & 0 & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Sigma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \end{array},$$

and Σ is positive diagonal. We call this decomposition the *SVD-like decomposition*, because it has the form and properties very similar to the singular value decomposition (SVD). If $A = UDV^*$ is the SVD of the matrix A then

$$AA^* = U\Sigma^2U^*, \qquad A^*A = V\Sigma^2V^*,$$

which are the Schur forms of the positive semidefinite matrices AA^*, A^*A respectively. Similarly if the real matrix B has the SVD-like decomposition (1), by the symplectic property $JS^{-T} = SJ$, one has

$$BJB^T = Q \left[\begin{array}{cc|cc} 0 & 0 & \Sigma^2 & 0 \\ 0 & 0 & 0 & 0 \\ \hline -\Sigma^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] Q^T, \tag{2}$$

and

$$JB^TB = S \left[\begin{array}{ccc|ccc} 0 & 0 & 0 & \Sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \hline -\Sigma^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & -I_q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] S^{-1}. \tag{3}$$

Let $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_p)$. With appropriate permutations (2) gives the real Schur form of the real skew-symmetric matrix BJB^T,

$$\mathrm{diag}\left(\left[\begin{array}{cc} 0 & \sigma_1^2 \\ -\sigma_1^2 & 0 \end{array} \right], \ldots, \left[\begin{array}{cc} 0 & \sigma_p^2 \\ -\sigma_p^2 & 0 \end{array} \right]; 0, \ldots, 0 \right);$$

and (3) gives the real Jordan canonical form of the real *Hamiltonian* matrix JB^TB,

$$\mathrm{diag}\left(\left[\begin{array}{cc} 0 & \sigma_1^2 \\ -\sigma_1^2 & 0 \end{array} \right], \ldots, \left[\begin{array}{cc} 0 & \sigma_p^2 \\ -\sigma_p^2 & 0 \end{array} \right]; \underbrace{\left[\begin{array}{cc} 0 & 0 \\ -1 & 0 \end{array} \right], \ldots \left[\begin{array}{cc} 0 & 0 \\ -1 & 0 \end{array} \right]}_{q}; 0, \ldots, 0 \right).$$

Therefore the SVD-like decomposition (1) contains the whole information about the eigenstructures of the real skew-symmetric matrices and the real Hamiltonian matrices in the product forms BJB^T and JB^TB, respectively. The structured Jordan canonical form for general Hamiltonian matrices is given in [15].

We also show that every complex symplectic matrix S has the SVD

$$S = U \left[\begin{array}{cc} \Omega & 0 \\ 0 & \Omega^{-1} \end{array} \right] V^*, \tag{4}$$

where U, V are unitary symplectic, Ω is positive diagonal.

Finally we apply the SVD-like decomposition (1) and the structured SVD (4) to analyze the norm and condition number of the factor B in the BJB^T factorization. In [7, 2] it is shown that for the real skew-symmetric matrix $K \in \mathbb{R}^{n\times n}$ there exists a matrix $B \in \mathbb{R}^{n\times 2m}$ with full column rank such that

$$K = BJB^T.$$

The BJB^T factorization is similar to the full rank factorization $A = BB^T$ for a real symmetric positive semidefinite matrix A. Both factorizations are not unique. If $A = BB^T$ then for every orthogonal matrix Q, $A = (BQ)(QB)^T$ is another full rank factorization. Similarly if $K = BJB^T$ then for every symplectic matrix S, $K = (BS)J(BS)^T$ is another BJB^T factorization. However, these two factorizations have a major difference. For the full rank factorization all different factors have the same singular values. But for the BJB^T factorization, because of the non-orthogonality of symplectic matrices, different factors may have different singular values. One may seek for the factors with minimal norm and condition number. Such minimization problems have numerical significance. For instance, to solve the system of linear equations $Kx = b$ with K real skew-symmetric one may use the BJB^T factorization of K ([7]). Another example is that in [3] a numerical method is given for solving the eigenvalue problem of a skew-Hamiltonian/Hamiltonian pencil, which is equivalent to a skew-symmetric/symmetric pencil $\alpha K - \beta M$. The main task of the method is to compute certain condensed forms of B, JB^T and M simultaneously, where B is the factor of $K = BJB^T$. In both examples for numerical stability a factor B with small norm and condition number is required. In this paper we give the general form for the factors B with minimal norm and condition number, respectively.

The SVD-like decomposition (1) also provides a different numerical way to solve the eigenvalue problem for the matrices BJB^T and JB^TB. For a matrix in product/quotient form it is a common idea to perform numerical computations on matrix factors rather than on the explicitly generated matrix. In this way the numerical accuracy usually can be improved. One classical example is the QZ algorithm for computing the eigenvalues of the matrix $B^{-1}A$. There are many numerical methods following this idea, e.g., [19, 6, 13, 16, 11]. The same idea can be used for solving the eigenvalue problems for the matrices BJB^T and JB^TB by computing the SVD-like form of B. Following this idea a numerical method is proposed in [20]. Here we only present the SVD-like decomposition for theoretical purpose.

Some basic properties about symplectic matrices and skew-symmetric matrices will be provided in Section 2. The SVD-like decomposition (1) will be proved in Section 3. The structured SVD (4) will be given in Section 4. For the BJB^T factorization the general forms of the factors with minimal norm and minimal condition number, respectively, will be given in Section 5. Some numerical methods for the BJB^T factorization will be discussed in Section 6. Finally our conclusion will be given in Section 7.

In this paper $\|B\|$ denotes the spectral norm of B. $\kappa(B) = \|B\| \left\|B^\dagger\right\|$ denotes the condition number of B, where $B^\dagger$ is the pseudo-inverse of B.

2 Preliminaries

We list the following properties for real symplectic matrices, which can be found in [8, 17].

Proposition 1

1. *The matrix* $J = \begin{bmatrix} 0 & I_m \\ -I_m & 0 \end{bmatrix}$ *is symplectic, and* $J^{-1} = J^T = -J$.
2. *If* $X \in \mathbb{C}^{m\times m}$ *is nonsingular then* $\begin{bmatrix} X & 0 \\ 0 & X^{-*} \end{bmatrix}$ *is symplectic. (Here* $X^{-*} = (X^{-1})^*$*.)*
3. *If* $Y \in \mathbb{C}^{m\times m}$ *is Hermitian then the matrix* $\begin{bmatrix} I_m & Y \\ 0 & I_m \end{bmatrix}$ *is symplectic.*
4. *If* $X, Y \in \mathbb{C}^{m\times m}$ *satisfy* $XY^* = YX^*$ *and* $\det X \neq 0$*, then the matrix* $\begin{bmatrix} X & Y \\ 0 & X^{-*} \end{bmatrix}$ *is symplectic.*

5. *If* $S \in \mathbb{C}^{2m\times 2m}$ *is symplectic then* S^*, $S^{-*}(= JSJ^*)$, $S^{-1}(= J^*S^*J)$ *are symplectic.*

6. *If* $S_1, S_2 \in \mathbb{C}^{2m\times 2m}$ *are symplectic then* S_1S_2 *is symplectic.*

The following properties can also be found in [8, 17].

Proposition 2

1. *A* $2m \times 2m$ *real orthogonal symplectic matrix has the block form* $\left[\begin{smallmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{smallmatrix}\right]$ *with* $U_1, U_2 \in \mathbb{R}^{m\times m}$.

2. *Let* $C \in \mathbb{R}^{2m\times r}$. *There exists a real orthogonal symplectic matrix* U *such that when* $r > m$,
$$C = UR = U \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},$$
with $R_{11}, R_{21} \in \mathbb{R}^{m\times m}$ *and* $R_{12}, R_{22} \in \mathbb{R}^{m\times(r-m)}$; *and when* $r \le m$,
$$C = UR = U \begin{bmatrix} R_{11} \\ 0 \\ R_{21} \\ 0 \end{bmatrix},$$
with $R_{11}, R_{21} \in \mathbb{R}^{r\times r}$. *In both cases* R_{11} *is upper triangular and* R_{21} *is strictly upper triangular.*

The following lemmas give some factorizations which will be used later.

Lemma 1 *Suppose that* $C \in \mathbb{R}^{n\times 2m}$ *and* $CJC^T = 0$. *Then* $\operatorname{rank} C = r \le \min\{n, m\}$, *and there exists a real symplectic matrix* Z *and a real orthogonal matrix* Q *such that*
$$Q^T CZ = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Proof. See Appendix. □

Lemma 2 *Suppose that* $X \in \mathbb{R}^{2n\times 2m}$ *and* $XJ_mX^T = J_n$. *Then there exists a real symplectic matrix* Z *such that*
$$XZ = \begin{array}{c} \\ n \\ n \end{array}\!\!\begin{array}{c} \begin{array}{cccc} n & m-n & n & m-n \end{array} \\ \begin{pmatrix} I & 0 & 0 & 0 \\ 0 & 0 & I & 0 \end{pmatrix} \end{array}.$$

Proof. See Appendix. □

It is well known that every real skew-symmetric matrix K is orthogonally similar to a matrix
$$\operatorname{diag}\left(\begin{bmatrix} 0 & \sigma_1^2 \\ -\sigma_1^2 & 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 & \sigma_m^2 \\ -\sigma_m^2 & 0 \end{bmatrix}; 0, \ldots, 0\right),$$

where $\sigma_1, \ldots, \sigma_m > 0$.

With an appropriate permutation K has a real Schur-like decomposition

$$K = Q \begin{bmatrix} 0 & \Sigma^2 & 0 \\ -\Sigma^2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} Q^T, \tag{5}$$

where Q is real orthogonal and $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_m) > 0$.

With this Schur-like form we have the following properties.

Proposition 3 *Let K be skew-symmetric and let (5) be the Schur-like form of K. Then*

1. $\|K\| = \|\Sigma\|^2$;
2. $\|K^\dagger\| = \|\Sigma^{-1}\|^2$;
3. $\kappa(K) = \kappa^2(\Sigma)$.
4. $\operatorname{rank} K = 2m$

Proof. Immediate. $\square$

3 SVD-like decomposition

The SVD-like decomposition (1) is presented in the following theorem.

Theorem 3 *If $B \in \mathbb{R}^{n\times 2m}$, then there exists a real symplectic matrix S and a real orthogonal matrix Q such that*

$$Q^T B S = D = \begin{array}{c} \\ p \\ q \\ p \\ n-2p-q \end{array} \begin{array}{c} \begin{array}{cccccc} p & q & m-p-q & p & q & m-p-q \end{array} \\ \begin{pmatrix} \Sigma & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Sigma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \end{array},$$

where $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_p) > 0$. Moreover, $\operatorname{rank} B = 2p + q$.

Proof. Since BJB^T is skew-symmetric, by (5) it has a Schur-like form

$$BJB^T = U \begin{bmatrix} 0 & \Sigma^2 & 0 \\ -\Sigma^2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} U^T,$$

where U is real orthogonal and $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_p) > 0$. Let $\Gamma = \operatorname{diag}(\Sigma, \Sigma, I_{n-2p})$ and $X := \Gamma^{-1} U^T B$. Then

$$XJX^T = \begin{bmatrix} 0 & I_p & 0 \\ -I_p & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{6}$$

Partition $X = [X_1^T, X_2^T, X_3^T]^T$ conformably. Then (6) gives

$$[X_1^T, X_2^T]^T J [X_1^T, X_2^T] = J_p.$$

By Lemma 2 there exists a real symplectic matrix $S_1 \in \mathbb{R}^{2m\times 2m}$ such that

$$[X_1^T, X_2^T]^T S_1 = \begin{bmatrix} I_p & 0 & 0 & 0 \\ 0 & 0 & I_p & 0 \end{bmatrix}.$$

Since $(XS_1)J(XS_1)^T = XJX^T$, by (6) and above block form

$$XS_1 = \begin{bmatrix} I_p & 0 & 0 & 0 \\ 0 & 0 & I_p & 0 \\ 0 & X_{32} & 0 & X_{34} \end{bmatrix},$$

and $[X_{32}, X_{34}]J_{m-p}[X_{32}, X_{34}]^T = 0$. By Lemma 1 there exists a symplectic matrix Z and an orthogonal matrix V such that $V^T[X_{32}, X_{34}]Z = \left[\begin{smallmatrix} I_q & 0 \\ 0 & 0 \end{smallmatrix}\right]$, and $q \le \min\{n-2p, m-p\}$. Let $Z = [Z_{ij}]_{2\times 2}$ with $Z_{ij} \in \mathbb{R}^{(m-p)\times(m-p)}$ for $i, j = 1, 2$. Define the symplectic matrix

$$S_2 = \begin{bmatrix} I_p & 0 & 0 & 0 \\ 0 & Z_{11} & 0 & Z_{12} \\ 0 & 0 & I_p & 0 \\ 0 & Z_{21} & 0 & Z_{22} \end{bmatrix}.$$

Then

$$XS_1S_2 = \begin{bmatrix} I_{2p} & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} I_p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_p & 0 & 0 \\ 0 & I_q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that Γ and $\operatorname{diag}(I_{2p}, V)$ commute. Let $S := S_1S_2$ and $Q := U \operatorname{diag}(I_{2p}, V)P$, where $P = \operatorname{diag}\left(I_p, \left[\begin{smallmatrix} 0 & I_p \\ I_q & 0 \end{smallmatrix}\right], I_{n-2p-q}\right)$. Since $B = U\Gamma X$, one has

$$Q^TBS = Q^TU\Gamma XS = P^T\Gamma \begin{bmatrix} I_p & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I_p & 0 & 0 \\ 0 & I_q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \Sigma & 0 & 0 & 0 & 0 & 0 \\ 0 & I_q & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Sigma & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

□

4 SVD of symplectic matrices

Every matrix has an SVD [12, p. 70]. But for a symplectic matrix the SVD has a special structure, which will be given below. For completeness we consider both the complex and real symplectic matrices.

Theorem 4 *Every symplectic matrix $S \in \mathbb{C}^{2m\times 2m}$ has an SVD of the form*

$$S = U \begin{bmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{bmatrix} V^*,$$

where U, V are unitary symplectic and $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_m)$ with $\omega_1 \geq \ldots \geq \omega_m \geq 1$.

Proof. Let

$$S = WDZ^* \tag{7}$$

be an SVD. Since S is symplectic we have $S = JS^{-*}J^*$. Then

$$S = J(WDZ^*)^{-*}J^* = (JW)D^{-1}(JZ)^* \tag{8}$$

is another SVD of S. If ω is a singular value of S obviously ω^{-1} is also a singular value of S. So D can be expressed as $D = \mathrm{diag}(\Omega, \Omega^{-1})$, where Ω is positive diagonal with diagonal elements arranged in decreasing order and bounded below by 1. Moreover, the two SVDs (7) and (8) imply that

$$D(Z^*JZ)D = W^*JW, \qquad D(W^*JW)D = Z^*JZ, \tag{9}$$

from which

$$D^2(Z^*JZ)D^2 = D(D(Z^*JZ)D)D = D(W^*JW)D = Z^*JZ.$$

Without loss of the generality let $\Omega = \mathrm{diag}(\omega_1 I, \ldots, \omega_p I, I)$ and $\omega_1 > \ldots > \omega_p > 1$. From the equation $D^2(Z^*JZ)D^2 = Z^*JZ$ and the fact that Z^*JZ is skew-Hermitian we have

$$Z^*JZ = \left[\begin{array}{cccc|cccc} 0 & & & & X_1 & & & \\ & \ddots & & & & \ddots & & \\ & & 0 & & & & X_p & \\ & & & Y_{11} & & & & Y_{12} \\ \hline -X_1^* & & & & 0 & & & \\ & \ddots & & & & \ddots & & \\ & & -X_p^* & & & & 0 & \\ & & & -Y_{12}^* & & & & Y_{22} \end{array}\right]. \tag{10}$$

Since Z and J are unitary so is the matrix Z^*JZ. This implies that $\begin{bmatrix} 0 & X_i \\ -X_i^* & 0 \end{bmatrix}$ $(i = 1, \ldots, p)$ are unitary. Therefore $X_i^*X_i = I$ for $i = 1, \ldots, p$. Moreover, $Y := \begin{bmatrix} Y_{11} & Y_{12} \\ -Y_{12}^* & Y_{22} \end{bmatrix}$ is also skew-Hermitian and unitary. Note that Z^*JZ is similar to J. So Z^*JZ has m identical eigenvalues i and $-i$ respectively. Note also that from the block form (10) each matrix $\begin{bmatrix} 0 & X_i \\ -X_i^* & 0 \end{bmatrix}$ has the same number of eigenvalues i and $-i$. Then Y must have the same number of eigenvalues i and $-i$, too. With this, as well as from its skew-Hermitian and unitary properties, the matrix Y is unitarily similar to $\mathrm{diag}(iI, -iI)$. On the other hand the matrix J is also unitarily similar to $\mathrm{diag}(iI, -iI)$. So there exists a unitary matrix P such that

$$Y = PJP^* = P \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} P^*.$$

Partition $P = [P_{ij}]_{2\times 2}$ conformably. Define the unitary matrix $\tilde{Z} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{21} & \tilde{Z}_{22} \end{bmatrix}$, where

$$\begin{aligned} \tilde{Z}_{11} &= \mathrm{diag}(X_1, \ldots, X_p, P_{11}), \qquad \tilde{Z}_{12} = \mathrm{diag}(0, \ldots, 0, P_{12}), \\ \tilde{Z}_{22} &= \mathrm{diag}(I, \ldots, I, P_{22}), \qquad \tilde{Z}_{21} = \mathrm{diag}(0, \ldots, 0, P_{21}). \end{aligned}$$

Then $(Z\tilde{Z})^* J (Z\tilde{Z}) = J$, indicating that $V := Z\tilde{Z}$ is unitary symplectic.

Using the fact that D and $\tilde{Z}$ commute and D is symplectic, from the first equation in (9), we have

$$(W\tilde{Z})^* J (W\tilde{Z}) = \tilde{Z}^* W^* J W \tilde{Z} = \tilde{Z}^* D (Z^* J Z) D \tilde{Z} = D V^* J V D = DJD = J,$$

i.e., the matrix $U := W\tilde{Z}$ is also unitary symplectic. Finally we have

$$S = WDZ^* = WD\tilde{Z}\tilde{Z}^* Z^* = (W\tilde{Z}) D (\tilde{Z}^* Z^*) = UDV^*.$$

□

For real symplectic matrices we have the following real version.

Theorem 5 *Every real symplectic matrix $S \in \mathbb{R}^{2m\times 2m}$ has a real SVD,*

$$S = U \begin{bmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{bmatrix} V^T,$$

where U, V are real orthogonal symplectic, $\Omega = \mathrm{diag}(\omega_1, \ldots, \omega_m)$ with $\omega_1 \geq \ldots \geq \omega_m \geq 1$.

Proof. The proof is analogous to that of Theorem 4. □

The results show that for the SVD of a symplectic matrix all matrices U, V and D can be chosen symplectic. Moreover, the singular values are in (ω, ω^{-1}) pairs.

5 BJB^T factorizations

It is shown in [7, 2] that every real skew-symmetric matrix $K \in \mathbb{R}^{n\times n}$ has a factorization BJB^T with B. The Cholesky-like factorization methods are also provided in [7, 2]. To simplify our analysis we will only consider the factorizations with $B \in \mathbb{R}^{n\times 2m}$ of full column rank, where $m = (\mathrm{rank}\, K)/2$. (The rank of K is an even number by Proposition 3.) In principle one can consider $K = BJ_sB^T$ with any $s \geq m$. But the problem can be analyzed in a similar way.

A BJB^T factorization can also be determined by the real Schur-like form (5) of K. In fact, let Q be real orthogonal and Σ be positive diagonal, both defined as in (5). Define

$$B_o = Q \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \\ 0 & 0 \end{bmatrix}. \tag{11}$$

Then $K = B_oJB_o^T$. As mentioned in Section 1 a skew-symmetric matrix K may have many BJB^T factorizations. In this section we will consider the factors with minimal norm and minimal condition number.

We first give the general form of B.

Proposition 4 *Let K be skew-symmetric and have the Schur-like form (5). If B satisfies $K = BJ_mB^T$ then there exists a real symplectic matrix S such that*

$$B = Q\begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \\ 0 & 0 \end{bmatrix} S = B_oS. \tag{12}$$

Proof. Let B satisfy $K = BJ_mB^T$. By the Schur-like form (5), for $\Gamma = \mathrm{diag}(\Sigma, \Sigma, I)$ one has

$$\Gamma^{-1}Q^TKQ\Gamma^{-1} = \Gamma^{-1}Q^TBJ_m(\Gamma^{-1}Q^TB)^T = \begin{bmatrix} J_m & 0 \\ 0 & 0 \end{bmatrix},$$

Partition $\Gamma^{-1}Q^TB = \left[\begin{smallmatrix} S \\ T \end{smallmatrix}\right]$ with $S \in \mathbb{R}^{2m\times 2m}$. Then $SJ_mS^T = J_m$ and $T = 0$. Therefore

$$B = Q\Gamma\begin{bmatrix} S \\ 0 \end{bmatrix} = Q\begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \\ 0 & 0 \end{bmatrix} S = B_oS.$$

□

Note that (12) is an SVD-like decomposition.

To get the general form of optimal factors we need the following results.

Lemma 6 *Suppose that $B \in \mathbb{R}^{n\times 2m}$ has the form (12), where Q is real orthogonal, S is real symplectic and $\Sigma > 0$ arranged as $\Sigma = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_r I_{n_r})$ with $\sigma_1 > \ldots > \sigma_r > 0$. Suppose that the singular values of the symplectic matrix S are arranged as $\omega_m^{-1} \leq \ldots \leq \omega_1^{-1} \leq 1 \leq \omega_1 \leq \ldots \leq \omega_m$. Then the following results hold.*

(i) $\|B\| \geq \sigma_1\sqrt{\frac{\omega_{n_1}^2+\omega_{n_1}^{-2}}{2}} \geq \sigma_1$.

Moreover, $\|B\| = \sigma_1$ if and only if S has the SVD

$$S = \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & U_1 & 0 & U_2 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & -U_2 & 0 & U_1 \end{bmatrix}\begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & \hat{\Omega} & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & 0 & 0 & \hat{\Omega}^{-1} \end{bmatrix} V^T,$$

where $U := \left[\begin{smallmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{smallmatrix}\right]$ and V are orthogonal symplectic, $\hat{\Omega} = \mathrm{diag}(\omega_{n_1+1}, \ldots, \omega_m)$; and U, $\hat{\Omega}$ satisfy

$$\left\|\begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma} \end{bmatrix} U \begin{bmatrix} \hat{\Omega} & 0 \\ 0 & \hat{\Omega}^{-1} \end{bmatrix}\right\| \leq \sigma_1$$

with $\hat{\Sigma} = \mathrm{diag}(\sigma_2 I_{n_2}, \ldots, \sigma_r I_{n_r})$.

(ii) The smallest singular value of B satisfies $\sigma_{2m}(B) \le \sigma_r \sqrt{\frac{2}{\omega_{n_r}^2+\omega_{n_r}^{-2}}} \le \sigma_r$. Moreover, $\sigma_{2m}(B) = \sigma_r$ if and only if S has the SVD

$$S = \begin{bmatrix} U_1 & 0 & U_2 & 0 \\ 0 & I_{n_r} & 0 & 0 \\ -U_2 & 0 & U_1 & 0 \\ 0 & 0 & 0 & I_{n_r} \end{bmatrix} \begin{bmatrix} \hat{\Omega} & 0 & 0 & 0 \\ 0 & I_{n_r} & 0 & 0 \\ 0 & 0 & \hat{\Omega}^{-1} & 0 \\ 0 & 0 & 0 & I_{n_r} \end{bmatrix} V^T,$$

where $U := \left[\begin{smallmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{smallmatrix}\right]$ and V are orthogonal symplectic, $\hat{\Omega} = \mathrm{diag}(\omega_{n_r+1}, \ldots, \omega_m)$; and U, $\hat{\Omega}$ satisfy

$$\left\| \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma} \end{bmatrix} U \begin{bmatrix} \hat{\Omega} & 0 \\ 0 & \hat{\Omega}^{-1} \end{bmatrix} \right\| \ge \sigma_r$$

with $\hat{\Sigma} = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_{r-1} I_{n_{r-1}})$.

Proof. See Appendix. □

We now give the general formula for the factors of the BJB^T factorization with minimal norm.

Theorem 7 *Let $K \in \mathbb{R}^{n\times n}$ be skew-symmetric and* rank $K = 2m$. *Then $\|B\| \ge \sqrt{\|K\|}$ for every factor B satisfying $K = BJ_mB^T$. Moreover, suppose that K has the Schur-like form (5), where Q is real orthogonal, $\Sigma = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_r I_{n_r})$ with $\sigma_1 > \ldots > \sigma_r > 0$. A factor B satisfies $\|B\| = \sqrt{\|K\|}$ if and only if it has the form*

$$B = Q \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \\ 0 & 0 \end{bmatrix} \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & U_1 & 0 & U_2 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & -U_2 & 0 & U_1 \end{bmatrix} \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & \hat{\Omega} & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & 0 & 0 & \hat{\Omega}^{-1} \end{bmatrix} V^T,$$

where $U := \left[\begin{smallmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{smallmatrix}\right]$ and V are real orthogonal symplectic, $\hat{\Omega}$ is positive diagonal; and U, $\hat{\Omega}$ satisfy

$$\left\| \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma} \end{bmatrix} U \begin{bmatrix} \hat{\Omega} & 0 \\ 0 & \hat{\Omega}^{-1} \end{bmatrix} \right\| \le \sqrt{\|K\|}$$

with $\hat{\Sigma} = \mathrm{diag}(\sigma_2 I_{n_2}, \ldots, \sigma_r I_{n_r})$.

Proof. The first part is trivial. The second part follows directly from Proposition 4, (i) of Lemma 6, and Proposition 3. □

The general formula for B with minimal condition number is given in the next theorem.

Theorem 8 *Let $K \in \mathbb{R}^{n\times n}$ be skew-symmetric and* rank $K = 2m$. *Then $\kappa(B) \ge \sqrt{\kappa(K)}$ for every B satisfying $K = BJ_mB^T$. Moreover, suppose that K has the Schur-like form (5), where Q is real orthogonal, $\Sigma = \mathrm{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_r I_{n_r})$ with $\sigma_1 > \ldots > \sigma_r > 0$. The following are equivalent.*

(a) $\kappa(B)=\sqrt{\kappa(K)}$.
(b) $\|B\|=\sqrt{\|K\|}$ *and* $\left\|B^{\dagger}\right\|=\sqrt{\|K^{\dagger}\|}$.
(c) B *has the form*

$$Q\begin{bmatrix}\Sigma & 0\\ 0 & \Sigma\end{bmatrix}\left[\begin{array}{ccc|ccc} I_{n_1} & 0 & 0 & 0 & 0 & 0\\ 0 & U_1 & 0 & 0 & U_2 & 0\\ 0 & 0 & I_{n_r} & 0 & 0 & 0\\ \hline 0 & 0 & 0 & I_{n_1} & 0 & 0\\ 0 & -U_2 & 0 & 0 & U_1 & 0\\ 0 & 0 & 0 & 0 & 0 & I_{n_r}\end{array}\right]\left[\begin{array}{ccc|ccc} I_{n_1} & 0 & 0 & 0 & 0 & 0\\ 0 & \tilde{\Omega} & 0 & 0 & 0 & 0\\ 0 & 0 & I_{n_r} & 0 & 0 & 0\\ \hline 0 & 0 & 0 & I_{n_1} & 0 & 0\\ 0 & 0 & 0 & 0 & \tilde{\Omega}^{-1} & 0\\ 0 & 0 & 0 & 0 & 0 & I_{n_r}\end{array}\right]V^T,$$

where $U:=\left[\begin{smallmatrix}U_1 & U_2\\ -U_2 & U_1\end{smallmatrix}\right]$ *and* V *are real orthogonal symplectic,* $\tilde{\Omega}$ *is positive diagonal; and* U*,* $\tilde{\Omega}$ *satisfy*

$$\sigma_r\leq\left\|\begin{bmatrix}\tilde{\Sigma} & 0\\ 0 & \tilde{\Sigma}\end{bmatrix}U\begin{bmatrix}\tilde{\Omega} & 0\\ 0 & \tilde{\Omega}^{-1}\end{bmatrix}\right\|\leq\sigma_1$$

with $\tilde{\Sigma}=\operatorname{diag}(\sigma_2 I_{n_2},\ldots,\sigma_{r-1}I_{n_{r-1}})$.

Proof. Let $\sigma_{2m}(B)$ and $\sigma_{2m}(K)$ be the smallest singular value of B and K respectively. By Lemma 6,

$$\|B\|\geq\sigma_1=\sqrt{\|K\|},\qquad \sigma_{2m}(B)\leq\sqrt{\sigma_{2m}(K)}.$$

By the definition of pseudo-inverse, $\left\|B^{\dagger}\right\|=1/\sigma_{2m}(B)$ and $\left\|K^{\dagger}\right\|=1/\sigma_{2m}(K)$, we have

$$\kappa(B)=\|B\|\left\|B^{\dagger}\right\|=\|B\|/\sigma_{2m}(B)\geq\sqrt{\|K\|}/\sqrt{\sigma_{2m}(K)}=\sqrt{\kappa(K)}.$$

(a) $\Leftrightarrow$ (b) can be proved by using above inequalities.

(b) $\Leftrightarrow$ (c) is obtained by combining (i) and (ii) in Lemma 6. □

Obviously, the factor B_o in (12) always has minimal norm and minimal condition number.

Example 1 Consider the skew-symmetric matrix

$$K=\begin{bmatrix}0 & 0 & 4 & 0\\ 0 & 0 & 0 & 1\\ -4 & 0 & 0 & 0\\ 0 & -1 & 0 & 0\end{bmatrix},$$

which is nonsingular and is in Schur-like form. One has $\|K\|=4$ and $\kappa(K)=4$. The general form for B to satisfy $K=BJB^T$ is

$$B=\begin{bmatrix}2 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}S,$$

where S is symplectic.

The general form for B to satisfy $\|B\| = \sqrt{\|K\|} = 2$ is

$$B = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & c & 0 & s \\ 0 & 0 & 1 & 0 \\ 0 & -s & 0 & c \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \omega & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \omega^{-1} \end{bmatrix} V^T,$$

where $\frac{1}{2} \leq \omega \leq 2$, V is orthogonal symplectic, and the matrix $\left[\begin{smallmatrix} c & s \\ -s & c \end{smallmatrix}\right]$ is a Givens rotation, which is orthogonal symplectic.

The general form for B to satisfy $\kappa(B) = \sqrt{\kappa(K)} = 2$ is

$$B = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} V^T,$$

where V is orthogonal symplectic.

Lemma 6 can also be used to solve the following optimization problems.

Corollary 9 *Suppose that $B \in \mathbb{R}^{n \times 2m}$ and* $\operatorname{rank} B = 2p$. *Then B has the SVD-like decomposition as in (12). Moreover, let $\mathcal{S}$ be the set of $2m \times 2m$ real symplectic matrices. Then*

$$\begin{aligned} &\min_{Z \in \mathcal{S}} \|BZ\| = \|\Sigma\|, \\ &\max_{Z \in \mathcal{S}} \sigma_{2m}(BZ) = \sigma_{2m}(\Sigma), \\ &\min_{Z \in \mathcal{S}} \kappa(BZ) = \kappa(\Sigma), \end{aligned}$$

where Σ is positive diagonal defined in (12).

Proof. Since B has full column rank, by Theorem 3 it must have an SVD-like decomposition (12), i.e., $B = Q \left[\begin{smallmatrix} \Sigma & 0 \\ 0 & \Sigma \end{smallmatrix}\right] S^{-1}$, where Q is real orthogonal, S is real symplectic and Σ is positive diagonal.

For any symplectic matrix Z, $BZ = Q \left[\begin{smallmatrix} \Sigma & 0 \\ 0 & \Sigma \end{smallmatrix}\right] (S^{-1}Z)$. Since for any real symplectic matrix Z, $S^{-1}Z$ is symplectic, by (i) of Lemma 6, $\|BZ\| \geq \|\Sigma\|$. When $Z = S$ one has $\|BZ\| = \|\Sigma\|$. Therefore $\min_{Z \in \mathcal{S}} \|BZ\| = \|\Sigma\|$.

The second identity can be proved in a similar way. The third identity follows from the first two identities. □

Corollary 10 *Suppose that $A \in \mathbb{R}^{2m \times 2m}$ is symmetric positive definite, and $\rho(JA)$, $\rho((JA)^{-1})$ are the spectral radius of the matrices JA and $(JA)^{-1}$, respectively. Let $\mathcal{S}$ be the set of $2m \times 2m$ real symplectic matrices. Then*

$$\begin{aligned} &\min_{Z \in \mathcal{S}} \left\|Z^T A Z\right\| = \rho(JA), \\ &\max_{Z \in \mathcal{S}} \sigma_{2m}(Z^T A Z) = 1/\rho((JA)^{-1}), \\ &\min_{Z \in \mathcal{S}} \kappa(Z^T A Z) = \rho(JA)\rho((JA)^{-1}). \end{aligned}$$

Proof. Let $A = LL^T$ be the Cholesky factorization [12, p. 143]. Since A is positive definite, L^T is nonsingular. The results can be obtained by applying Corollary 9 to matrix L^T. The only thing that we still need to show is that $\rho(JA) = \|\Sigma\|^2$, $\rho((JA)^{-1}) = \|\Sigma^{-1}\|^2$. This follows from the fact that with $L^T = Q \left[\begin{smallmatrix} \Sigma & 0 \\ 0 & \Sigma \end{smallmatrix}\right] S^{-1}$, $JA = JLL^T = S \left[\begin{smallmatrix} 0 & \Sigma^2 \\ -\Sigma^2 & 0 \end{smallmatrix}\right] S^{-1}$. □

6 Methods for BJB^T factorizations

In this section we will discuss the norm and condition number of the factor B computed by several BJB^T factorization methods.

I. Schur method. The Schur form of K provides a way to compute B_o as in (11). Obviously B_o has both the minimal norm and condition number. Several numerical methods for computing the Schur form of K are available, e.g., [18, 12]. However the cost for computing the Schur-like form by the QR-like algorithm is about $O(10n^3)$ flops. (Here we borrowed the cost for the symmetric QR algorithm [12, p. 421].) So the Schur method is much more expensive than the following Cholesky-like methods, which require only about $O(n^3/3)$ flops [7, 2].

II. Cholesky-like methods. Such methods were developed in [7, 2]. Let us briefly explain the procedure provided in [2]. Suppose $K \neq 0$. One can determine a permutation P_1 such that $P_1^T K P_1 = \left[\begin{smallmatrix} K_{11} & K_{12} \\ -K_{12}^T & K_{22} \end{smallmatrix}\right]$, where $K_{11} \in \mathbb{R}^{2\times 2}$ is nonsingular. It can be written

$$P_1^T K P_1 = \begin{bmatrix} R_{11}^T & 0 \\ -K_{12}^T R_{11}^{-T} J_1^T & I \end{bmatrix} \begin{bmatrix} J_1 & 0 \\ 0 & K_1 \end{bmatrix} \begin{bmatrix} R_{11} & -J_1 R_{11}^{-1} K_{12} \\ 0 & I \end{bmatrix},$$

where matrix $K_1 = K_{22} + K_{12}^T K_{11}^{-1} K_{12}$ is still skew-symmetric. Repeat the reduction on K_1 and continue. One finally has

$$K = \tilde{P}\tilde{B}\,\mathrm{diag}(J_1, \ldots, J_1)\tilde{B}^T \tilde{P}^T =: (\tilde{P}\tilde{B})\tilde{J}(\tilde{P}\tilde{B})^T,$$

where $\tilde{P}$ is a permutation. (The procedure in [7] computes a factorization

$$K = (\tilde{P}\tilde{B})\,\mathrm{diag}(K_1, \ldots, K_m)(\tilde{P}\tilde{B})^T,$$

where $K_1, \ldots, K_m \in \mathbb{R}^{2\times 2}$ are skew-symmetric and nonsingular.) Matrix $\tilde{J}$ can be written $\tilde{J} = PJP^T$ for some permutation P. Hence the method computes the BJB^T factorization $K = (\tilde{P}\tilde{B}P)J(\tilde{P}\tilde{B}P)^T$. For numerical stability partial or complete pivoting strategy is implemented in the reduction procedure.

It is not easy to give a prior estimate about the norm or the condition number of the computed factor B. So it is not clear whether the methods can compute a factor with relatively small norm or condition number. We did some numerical experiments to test the method provided in [2] with complete pivoting. We chose several groups of random matrices K with different size n. For each size we tested 20 skew-symmetric matrices.

The maximum and minimum values of the quantities $\gamma_N := \frac{\|B\|}{\sqrt{\|K\|}}$ and $\gamma_C := \frac{\kappa(B)}{\sqrt{\kappa(K)}}$, and the norm and condition number of the corresponding matrices B and K are reported in Table 1. We observed that γ_N is between 1 and 7 and γ_C is between 2 and 13. We also tested some matrices K with large norm or condition number. Four groups of such matrices were tested. The matrices were generated by performing different scaling strategies to the randomly formed skew-symmetric matrices. In each case we tested 20 matrices with size 50×50. The numerical results are reported in Table 2. We observed that γ_N, γ_C are between 1 and 3. These two numerical examples show that the Cholesky-like factorization method with complete pivoting computes the factors with small norm and condition number.

n		γ_N	γ_C	$\|B\|$	$\|K\|$	$\kappa(B)$	$\kappa(K)$
10	max	1.95	3.30	4.34	4.94	7.79	5.57
	min	1.33	2.23	3.16	5.63	11.3	26.0
50	max	3.43	4.26	12.5	12.5	79.8	351.5
	min	2.76	5.05	9.91	12.9	39.9	64.4
100	max	4.70	5.63	20.2	18.7	165.6	865.6
	min	3.90	6.20	17.2	19.4	69.5	125.6
200	max	6.24	12.2	32.9	27.7	167.9	190.4
	min	5.55	11.6	28.8	27.0	185.5	253.8

Table 1: Numerical results for Cholesky-like factorization I.

Group		γ_N	γ_C	$\|B\|$	$\|K\|$	$\kappa(B)$	$\kappa(K)$
1	max	1.67	2.48	23.8	201.9	85.5	1.2×10^3
	min	1.03	1.28	192.6	3.5×10^4	1.1×10^3	7.1×10^5
2	max	1.58	2.20	38.8	601.7	131.9	3.6×10^3
	min	1.01	1.62	2.1×10^5	4.2×10^{10}	8.4×10^5	2.7×10^{11}
3	max	1.21	1.25	2.93	5.86	1.4×10^5	1.2×10^{10}
	min	1.0	1.22	2.74	7.52	1.5×10^4	1.5×10^8
4	max	2.13	2.64	14.3	44.8	2.6×10^5	9.5×10^9
	min	1.23	1.26	18.2	220.8	3.9×10^4	9.7×10^8

Table 2: Numerical results for Cholesky-like factorization II.

For the second example we also tested the method with partial pivoting. But γ_N and γ_C are much bigger than that with complete pivoting. When $\|K\|$ or $\kappa(K)$ is large γ_N can be 70 times bigger and γ_C can be 700 times bigger.

7 Conclusion

Several matrix decompositions related to symplectic matrices have been presented. The first one is an SVD-like decomposition. The canonical forms of the skew-symmetric matrix BJB^T and the Hamiltonian matrix JB^TB can be derived from such a decomposition for a matrix B. The second one is the structured SVD of a symplectic matrix. Applying these decompositions to the BJB^T factorization of a skew-symmetric matrix K, the general forms for the factors B with minimal norm and condition number, respectively, have been formulated. Several BJB^T factorization methods have been discussed. Numerical results show that the Cholesky-like method with complete pivoting is usually a simple way to compute the factors with small norm and condition number. Some other optimization problems related to symplectic matrices were also solved.

Appendix

Proof of Lemma 1. Let $C = Q\left[\begin{smallmatrix} R \\ 0 \end{smallmatrix}\right]$ be the QR decomposition of C, where $R \in \mathbb{R}^{r\times 2m}$ is of full row rank. Applying the symplectic QR decomposition (Proposition 2) to R^T there is an orthogonal symplectic matrix Z_1 such that if $r > m$,

$$RZ_1 = \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix},$$

and if $r \leq m$,

$$RZ_1 = \begin{bmatrix} R_{11} & 0 & R_{12} & 0 \end{bmatrix}, \tag{13}$$

where $R_{11} \in \mathbb{R}^{s\times s}$ is lower triangular and $R_{12} \in \mathbb{R}^{s\times s}$ is strictly lower triangular, with $s = \min\{r, m\}$; and if $r > m$, $R_{21}, R_{22} \in \mathbb{R}^{(r-m)\times m}$. The condition $CJC^T = 0$ implies that $(RZ_1)J(RZ_1)^T = 0$. If $r > m$ then one has

$$R_{11}R_{12}^T = R_{12}R_{11}^T, \quad R_{11}R_{22}^T = R_{12}R_{21}^T.$$

By comparing the components of the matrices in the first equation and using the fact that $[R_{11}, R_{12}]$ is of full row rank, one can get that R_{11} is nonsingular and $R_{12} = 0$. Applying this to the second equation one has $R_{22} = 0$. Now matrix $R = \left[\begin{smallmatrix} R_{11} & 0 \\ R_{21} & 0 \end{smallmatrix}\right]$ and $R_{11} \in \mathbb{R}^{m\times m}$ is nonsingular. But this implies that $r = \operatorname{rank} C = \operatorname{rank} R = m$, which is a contradiction. Consequently $r \leq m$ (and it is obvious that $r \leq n$). Now RZ_1 has the form (13). Similarly one can show that $R_{11} \in \mathbb{R}^{r\times r}$ is nonsingular and $R_{12} = 0$. Defining the symplectic matrix $Z_2 = \operatorname{diag}(R_{11}^{-1}, I_{m-r}; R_{11}^T, I_{m-r})$ one has $RZ_1Z_2 = \begin{bmatrix} I_r & 0 & 0 & 0 \end{bmatrix}$. Hence for $Z = Z_1Z_2$, we have $Q^TCZ = \left[\begin{smallmatrix} R \\ 0 \end{smallmatrix}\right] Z = \left[\begin{smallmatrix} I_r & 0 \\ 0 & 0 \end{smallmatrix}\right]$. □

Proof of Lemma 2. By $XJ_mX^T = J_n$ one has $m \geq n$. Partition $X = \left[\begin{smallmatrix} X_1 \\ X_2 \end{smallmatrix}\right]$ with $X_1, X_2 \in \mathbb{R}^{n\times 2m}$. Then $XJX^T = J$ implies that $X_kJX_k^T = 0$ for $k = 1, 2$ and $X_1JX_2^T = I_n$. The last equation implies that $\operatorname{rank} X_1 = \operatorname{rank} X_2 = n$. Since $X_1JX_1^T = 0$ and $\operatorname{rank} X_1 = n$, by Lemma 1 there exists a symplectic matrix Z_1 such that $X_1 Z_1 = [I_n,\, 0;\, 0,\, 0]$. (The orthogonal matrix Q is replaced by I here because X_1 is already of full row rank.) Partition $X_2 Z_1 = [X_{21},\, X_{22},\, X_{23},\, X_{24}]$ conformably. From $(X_1 Z_1) J (X_2 Z_1)^T = I_n$ one has $X_{23} = I_n$. Define the symplectic matrix

$$Z_2 = \operatorname{diag}\left(\begin{bmatrix} I & 0 \\ X_{24}^T & I \end{bmatrix}, \begin{bmatrix} I & -X_{24} \\ 0 & I \end{bmatrix} \right).$$

Then

$$X Z_1 Z_2 = \begin{bmatrix} I_n & 0 & 0 & 0 \\ \tilde{X}_{21} & X_{22} & I_n & 0 \end{bmatrix},$$

where by $XJX^T = J$ matrix $\tilde{X}_{21}$ is symmetric. Finally using the symmetry of $\tilde{X}_{21}$ one can construct a symplectic matrix

$$Z_3 = \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ -\tilde{X}_{21} & -X_{22} & I & 0 \\ -X_{22}^T & 0 & 0 & I \end{bmatrix}.$$

Then for $Z := Z_1 Z_2 Z_3$ the matrix XZ is in the asserted form. $\square$

Proof of Lemma 6. (i) By Theorem 5 the symplectic matrix S has the symplectic SVD $S = W \left[\begin{smallmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{smallmatrix}\right] Z^T$, where W and Z are orthogonal symplectic, and $\Omega = \operatorname{diag}(\omega_1, \ldots, \omega_m)$, with $1 \le \omega_1 \le \ldots \le \omega_m$. By the SVD of S and the decomposition (12), the matrices B and $\left[\begin{smallmatrix} \Sigma & 0 \\ 0 & \Sigma \end{smallmatrix}\right] W \left[\begin{smallmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{smallmatrix}\right]$ have the same singular values. So the singular values of B are just the square root of the eigenvalues of the matrix

$$A = \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} W \begin{bmatrix} \Omega^2 & 0 \\ 0 & \Omega^{-2} \end{bmatrix} W^T \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} =: \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} H \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix}. \tag{14}$$

Clearly

$$\|A\| I - \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} H \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix} \ge 0$$

or

$$\|A\| \begin{bmatrix} \Sigma^{-2} & 0 \\ 0 & \Sigma^{-2} \end{bmatrix} \ge H.$$

Performing a congruence transformation on both matrices with J, and recognizing that matrix $H = W \left[\begin{smallmatrix} \Omega^2 & 0 \\ 0 & \Omega^{-2} \end{smallmatrix}\right] W^T$ is symplectic and symmetric, one has

$$\|A\| \begin{bmatrix} \Sigma^{-2} & 0 \\ 0 & \Sigma^{-2} \end{bmatrix} \ge H^{-1}.$$

Combining the above two inequalities one obtains

$$2\,\|A\| \begin{bmatrix} \Sigma^{-2} & 0 \\ 0 & \Sigma^{-2} \end{bmatrix} \ge H + H^{-1} = W \begin{bmatrix} \Omega^2 + \Omega^{-2} & 0 \\ 0 & \Omega^2 + \Omega^{-2} \end{bmatrix} W^T.$$

Let $E_x = \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & 0 & I_{n_1} & 0 \end{bmatrix}^T$. Since $\Sigma = \operatorname{diag}(\sigma_1 I_{n_1}, \ldots, \sigma_r I_{n_r})$, for any unit norm vector $x \in \operatorname{range} E_x$,

$$\frac{2\,\|A\|}{\sigma_1^2} = 2\,\|A\|\, x^T \begin{bmatrix} \Sigma^{-2} & 0 \\ 0 & \Sigma^{-2} \end{bmatrix} x \geq x^T W \begin{bmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{bmatrix} W^T x,$$

or equivalently for any unit norm vector $y = W^T x \in \operatorname{range} W^T E_x$,

$$\frac{2\,\|A\|}{\sigma_1^2} \geq y^T \begin{bmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{bmatrix} y. \tag{15}$$

Note that the matrix $\left[\begin{smallmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{smallmatrix}\right]$ has the eigenvalues $\omega_1^2+\omega_1^{-2}, \omega_1^2+\omega_1^{-2}, \ldots, \omega_m^2+\omega_m^{-2}, \omega_m^2+\omega_m^{-2}$ in non-decreasing order. This is because the function $f(\omega) = \omega^2+\omega^{-2}$ is increasing for $\omega \geq 1$ and the diagonal elements of Ω are in non-decreasing order $1 \leq \omega_1 \leq \ldots \leq \omega_m$. By the minmax theorem, e.g., [12, p.394], it follows that

$$\frac{2\,\|A\|}{\sigma_1^2} \geq \min_{\mathcal{S},\,\operatorname{rank}\mathcal{S}=2n_1} \max_{y\in\mathcal{S}} y^T \begin{bmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{bmatrix} y = \omega_{n_1}^2 + \omega_{n_1}^{-2}.$$

So we have

$$\|B\| = \sqrt{\|A\|} \geq \sigma_1 \sqrt{\frac{\omega_{n_1}^2+\omega_{n_1}^{-2}}{2}} \geq \sigma_1.$$

When $\|B\| = \sigma_1$, it is clear that $\omega_{n_1} = 1$. So at least n_1 diagonal elements of Ω are 1. Without loss of the generality we assume that $\Omega = \operatorname{diag}(I_t, \omega_{t+1}, \ldots, \omega_m)$, where $t \geq n_1$ and $1 < \omega_{t+1} \leq \ldots \leq \omega_m$. Since $\|B\| = \sqrt{\|A\|} = \sigma_1$, by (15) for every unit norm vector $y \in W^T E_x$, one has

$$y^T \begin{bmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{bmatrix} y = 2.$$

The matrix $\left[\begin{smallmatrix} \Omega^2+\Omega^{-2} & 0 \\ 0 & \Omega^2+\Omega^{-2} \end{smallmatrix}\right]$ has $2t$ copies of the smallest eigenvalue 2 with the corresponding eigenspace = range $\begin{bmatrix} I_t & 0 & 0 & 0 \\ 0 & 0 & I_t & 0 \end{bmatrix}^T$. So every y must be in this eigenspace, or

$$\operatorname{range} W^T E_x \subseteq \operatorname{range} \begin{bmatrix} I_t & 0 \\ 0 & 0 \\ 0 & I_t \\ 0 & 0 \end{bmatrix}.$$

By this and the block structure of the orthogonal symplectic matrix W (Proposition 2),

$$W^T E_x = \begin{bmatrix} E_1 & E_2 \\ 0 & 0 \\ -E_2 & E_1 \\ 0 & 0 \end{bmatrix},$$

with $E_1, E_2 \in \mathbb{R}^{t\times n_1}$. Let $E = \left[\begin{smallmatrix} E_1 & E_2 \\ -E_2 & E_1 \end{smallmatrix}\right]$. Clearly

$$E^T J_t E = (W^T E_x)^T J_m (W^T E_x) = J_{n_1}, \quad E^T E = (W^T E_x)^T W^T E_x = I_{n_1}.$$

By these facts and the symplectic QR decomposition (Proposition 2) there exists an orthogonal symplectic matrix F such that

$$F^T E = \begin{bmatrix} I_{n_1} & 0 \\ 0 & 0 \\ 0 & I_{n_1} \\ 0 & 0 \end{bmatrix}. \tag{16}$$

(This can be proved in a similar way as in Lemma 2 by using the extra condition that $E^T E = I_{n_1}$.) Since F is orthogonal symplectic it has the form $\left[\begin{smallmatrix} F_1 & F_2 \\ -F_2 & F_1 \end{smallmatrix}\right]$ with $F_1, F_2 \in \mathbb{R}^{t\times t}$. Now define the orthogonal symplectic matrix

$$\hat{F} = \begin{bmatrix} F_1 & 0 & F_2 & 0 \\ 0 & I & 0 & 0 \\ -F_2 & 0 & F_1 & 0 \\ 0 & 0 & 0 & I \end{bmatrix}.$$

Then (16) implies that $\hat{F}^T W^T E_x = E_x$, or

$$W\hat{F} = \begin{bmatrix} I_{n_1} & 0 & 0 & 0 \\ 0 & U_1 & 0 & U_2 \\ 0 & 0 & I_{n_1} & 0 \\ 0 & -U_2 & 0 & U_1 \end{bmatrix},$$

where $\left[\begin{smallmatrix} U_1 & U_2 \\ -U_2 & U_1 \end{smallmatrix}\right] =: U$ is orthogonal symplectic. The matrix $\hat{F}$ commutes with $\mathrm{diag}(\Omega, \Omega^{-1})$. Therefore, by setting $V = Z\hat{F}$ one has $S = (W\hat{F}) \left[\begin{smallmatrix} \Omega & 0 \\ 0 & \Omega^{-1} \end{smallmatrix}\right] V^T$ as asserted. With this form

$$\left\| \begin{bmatrix} \hat{\Sigma} & 0 \\ 0 & \hat{\Sigma} \end{bmatrix} U \begin{bmatrix} \hat{\Omega} & 0 \\ 0 & \hat{\Omega} \end{bmatrix} \right\| \le \sigma_1$$

is obviously the necessary and sufficient condition for $\|B\| = \sigma_1$.

(ii) For A defined in (14), $\|A^{-1}\| = 1/(\sigma_{2m}(B))^2$ since

$$A^{-1} = \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix}^{-1} H^{-1} \begin{bmatrix} \Sigma & 0 \\ 0 & \Sigma \end{bmatrix}^{-1}$$

and

$$H^{-1} = W \begin{bmatrix} \Omega^{-2} & 0 \\ 0 & \Omega^2 \end{bmatrix} W^T.$$

Replacing H and A in (i) by H^{-1} and A^{-1} we can show that

$$(\sigma_{2m}(B))^{-1} = \sqrt{\|A^{-1}\|} \geq \sigma_r^{-1}\sqrt{\frac{\omega_{n_r}^2 + \omega_{n_r}^{-2}}{2}} \geq \sigma_r^{-1}.$$

Hence

$$\sigma_{2m}(B) \leq \sigma_r\sqrt{\frac{2}{\omega_{n_r}^2 + \omega_{n_r}^{-2}}} \leq \sigma_r.$$

When $\sigma_{2m}(B) = \sigma_r$ the block structure of S can be obtained in the same way. $\square$

Acknowledgments. The author gratefully acknowledge Volker Mehrmann and the anonymous reviewers for their valuable comments and suggestions.

References

[1] P. Benner. *Contributions to the Numerical Solution of Algebraic Riccati Equations and Related Eigenvalue Problems.* Logos–Verlag, Berlin, Germany, 1997. *Also:* Dissertation, Fakultät für Mathematik, TU Chemnitz–Zwickau, 1997.

[2] P. Benner, R. Byers, H. Faßbender, V. Mehrmann, and D. Watkins. Cholesky-like factorizations of skew-symmetric matrices. *Electr. Trans. Num. Anal.*, 11:85–93, 2000.

[3] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 24(1):165–190, 2002.

[4] P. Benner, V. Mehrmann, and H. Xu. A new method for computing the stable invariant subspace of a real Hamiltonian matrix. *J. Comput. Appl. Math.*, 86:17–43, 1997.

[5] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[6] A. Bojanczyk, G.H. Golub, and P. Van Dooren. The periodic Schur decomposition; algorithms and applications. In *Proc. SPIE Conference, vol. 1770*, pages 31–42, 1992.

[7] J.R. Bunch. A note on the stable decomposition of skew-symmetric matrices. *Math. J. Comput.*, 38:475–479, 1982.

[8] A. Bunse-Gerstner. Matrix factorization for symplectic QR-like methods. *Linear Algebra Appl.*, 83:49–77, 1986.

[9] R. Byers. A Hamiltonian QR-algorithm. *SIAM J. Sci. Statist. Comput.*, 7:212–229, 1986.

[10] H. Faßbender. *Symplectic Methods for Symplectic Eigenproblems.* Kluwer Academic, New York 2000.

[11] G. Golub, K. Sølna, and P. Van Dooren. Computing the SVD of a general matrix product/quotient. *SIAM J. Matrix Anal. Appl.*, 22:1–19, 2000.

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

[13] J.J. Hench and A.J. Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39:1197–1210, 1994.

[14] P. Lancaster and L. Rodman. *The Algebraic Riccati Equation.* Oxford University Press, Oxford, 1995.

[15] W.-W. Lin, V. Mehrmann, and H. Xu. Canonical forms for Hamiltonian and symplectic matrices and pencils. *Linear Algebra Appl.*, 301–303:469–533, 1999.

[16] R. Mathias. Accurate eigensystem computations by Jacobi methods. *SIAM J. Matrix Anal. Appl.*, 16:977–1003, 1995.

[17] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[18] M.H.C. Paardekooper. An eigenvalue algorithm for skew-symmetric matrices. *Numer. Math.*, 17:189–202, 1971.

[19] K. Veselič and V. Hari. A note on a one-sided Jacobi algorithm. *Numer. Math.*, 56:627–633, 1989.

[20] H. Xu. A numerical method for computing an SVD-like decomposition. Submitted to *SIAM J. Matrix Anal. Appl.*.

[21] V.A. Yakubovich and V.M. Starzhinskii. *Linear Differential Equations with Periodic Coefficients, Volume 1, 2.* John Wiley & Sons, New York, Toronto, 1975.