Mathematical
Inequalities
& Applications
Volume 20, Number 2 (2017), 427–439 doi:10.7153/mia-20-29

A NOTE ON THE SENSITIVITY ANALYSIS FOR THE SYMPLECTIC QR FACTORIZATION

HANYU LI AND PENG LV

(*Communicated by S. Puntanen*)

Abstract. In this note, the rigorous perturbation bounds for R factor of the implicit Bunch form of the symplectic QR factorization under normwise perturbation are derived by using the block matrix-vector equation approach, the technique of Lyapunov majorant function, and the Banach fixed point principle. These bounds are tighter than the one in [Li *et al.* Linear Multilinear Algebra, **63**, (2015), 78–96] and can be regarded as the rigorous versions of the optimal first-order perturbation bounds in [Li *et al.* J. Franklin Inst., **353**, 5 (2016), 1186–1205].

1. Introduction

Let $\mathbb{R}^{m\times n}$ be the set of $m\times n$ real matrices and $\mathbb{R}_r^{m\times n}$ be the subset of $\mathbb{R}^{m\times n}$ with rank r. For a matrix $Q\in\mathbb{R}_{2m}^{2m\times 2m}$, it is said to be symplectic if $Q^TJQ=J$, where Q^T denotes the transpose of Q and

$$J=\operatorname{diag}(J_0,J_0,\cdots,J_0)\in\mathbb{R}^{2m\times 2m}$$

with

$$J_0=\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}.$$

Let $A\in\mathbb{R}^{2m\times 2n}$ with $m\geqslant n$. If A^TJA is nonsingular, then A has the following factorization

$$A=Q\begin{bmatrix}R\\ 0\end{bmatrix}P^T=[Q_1,Q_2]\begin{bmatrix}R\\ 0\end{bmatrix}P^T=Q_1RP^T, \tag{1.1}$$

Mathematics subject classification (2010): 15A23, 15A45.
Keywords and phrases: Symplectic QR factorization, rigorous perturbation bound, Lyapunov majorant function, Banach fixed point principle.
This research is supported by the National Natural Science Foundation of China (Grant Nos. 11671060; 11201507) and the Fundamental Research Funds for the Central Universities (Grant No. 106112015CDJXY100003).

where $Q \in \mathbb{R}_{2m}^{2m\times 2m}$ is symplectic, $Q_1 \in \mathbb{R}_{2n}^{2m\times 2n}$ satisfies $Q_1^T J Q_1 = J_1$ with $J_1 = J(1:2n, 1:2n)$, $R = (r_{ij}) \in \mathbb{R}^{2n\times 2n}$ is upper triangular with 2×2 main diagonal blocks:

$$\begin{bmatrix} r_{2i-1,2i-1} & 0 \\ 0 & r_{2i,2i} \end{bmatrix}, \quad r_{2i-1,2i-1} = r_{2i,2i} > 0, \ \ i = 1,2,\cdots,n,$$

and P is a permutation matrix. The above factorization is unique and is called an implicit Bunch form of the symplectic QR factorization [12]. Besides, some authors also considered other forms of the symplectic QR factorization [4, 10, 13]. In this paper, we mainly discuss the symplectic QR factorization in (1.1).

The symplectic QR factorization, combined with the Pietzsch algorithm, can be used to accurately compute the eigenvalues of some classes of skew-symmetric or skew-Hermitian matrices [13] and is also a useful tool in the computation of some optimal control problems [2, 3, 8, 15]; see also [12] for a detailed introduction. Its algorithms, stability of algorithms, and perturbation analysis have been considered in [1, 2, 3, 4, 8, 10, 12, 13]. Some first-order and rigorous perturbation bounds were presented [1, 4, 10, 13]. Recently, Li *et al.* [12] derived the optimal first-order perturbation bounds of the symplectic QR factorization (1.1) under normwise perturbation and presented its normwise condition numbers.

In this paper, we will combine the block matrix-vector equation approach (see [11]) with the technique of Lyapunov majorant function (see [6, Definition 5.4]) and the Banach fixed point principle (see [6, Appendix D]) to study the rigorous perturbation bounds for R factor of the symplectic QR factorization (1.1) when the original matrix has the normwise perturbation. The new bounds will correspond to the optimal first-order perturbation bounds in [12] and are tighter than the one given in [10].

The rest of this paper is organized as follows. Section 2 presents some notation and preliminaries. The new rigorous perturbation bounds are given in Section 3. In Section 4, we give a numerical example to illustrate the results derived in Sections 3. The last section provides the concluding remarks of the whole paper.

2. Notation and preliminaries

Most of the notation and preliminaries given in this section can also be found in [11]. For the convenience of readers, we still exhibit them here.

Given a matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$, its spectral norm and Frobenius norm are denoted by $\|A\|_2$ and $\|A\|_F$, respectively. For these two matrix norms, the following inequalities hold (see [14, pp.80]):

$$\|XYZ\|_2 \leqslant \|X\|_2 \|Y\|_2 \|Z\|_2, \quad \|XYZ\|_F \leqslant \|X\|_2 \|Y\|_F \|Z\|_2, \tag{2.1}$$

whenever the matrix product XYZ is well-defined.

For the matrix $A = (A_{ij}) \in \mathbb{R}^{2n\times 2n}$, where $A_{ij} \in \mathbb{R}^{2\times 2}$, $i, j = 1,2,\cdots,n$, we define the following operators:

$$\text{uvecb}(A) = \begin{bmatrix} \text{vec}(A_{11}) \\ \vdots \\ \text{vec}(A_{1n}) \\ \hline \text{vec}(A_{22}) \\ \vdots \\ \text{vec}(A_{2n}) \\ \hline \vdots \\ \hline \text{vec}\left(A_{(n-1)(n-1)}\right) \\ \text{vec}\left(A_{(n-1)n}\right) \\ \hline \text{vec}(A_{nn}) \end{bmatrix} \in \mathbb{R}^{\nu_1}, \quad \text{vecb}(A) = \begin{bmatrix} \text{vec}(A_{11}) \\ \vdots \\ \text{vec}(A_{1n}) \\ \hline \vdots \\ \hline \text{vec}(A_{n1}) \\ \vdots \\ \text{vec}(A_{nn}) \end{bmatrix} \in \mathbb{R}^{4n^2},$$

$$\text{upb}(A) = \begin{bmatrix} \frac{1}{2}A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & \frac{1}{2}A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{2}A_{nn} \end{bmatrix}, \quad \text{utb}(A) = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ 0 & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix},$$

where $\nu_1 = 2n(n+1)$ and the operator 'vec' stacks the columns of a matrix one underneath the other.

Let $\mathbb{D}_{2n} \in \mathbb{R}^{2n\times 2n}$ denote the set of diagonal positive definite matrices with 2×2 main diagonal blocks $s_i I_2$, where $s_i > 0$, $i = 1,2,\cdots,n$. Hereafter, I_r is the identity matrix of order r. Then, for any $D \in \mathbb{D}_{2n}$,

$$\text{upb}(AD) = \text{upb}(A)D, \quad D\text{upb}(A) = D\text{upb}(A). \tag{2.2}$$

Making use of the structures of the operators defined above, we have

$$\text{uvecb}(A) = M_{\text{uvecb}}\text{vecb}(A), \quad \text{vecb}(\text{utb}(A)) = M_{\text{utb}}\text{vecb}(A),$$
$$\text{vecb}(\text{upb}(A)) = M_{\text{upb}}\text{vecb}(A), \tag{2.3}$$

where

$$\begin{aligned} &M_{\text{uvecb}} = \text{diag}(S_1, S_2, \cdots, S_n) \in \mathbb{R}^{\nu_1 \times 4n^2}, \\ &S_i = \left[0_{4(n-i+1)\times 4(i-1)}, I_{4(n-i+1)}\right] \in \mathbb{R}^{4(n-i+1)\times 4n}, \\ &M_{\text{utb}} = \text{diag}\left(\hat{S}_1, \hat{S}_2, \cdots, \hat{S}_n\right) \in \mathbb{R}^{4n^2\times 4n^2}, \\ &\hat{S}_i = \text{diag}\left(0_{4(i-1)\times 4(i-1)}, I_{4(n-i+1)}\right) \in \mathbb{R}^{4n\times 4n}, \\ &M_{\text{upb}} = \text{diag}\left(\tilde{S}_1, \tilde{S}_2, \cdots, \tilde{S}_n\right) \in \mathbb{R}^{4n^2\times 4n^2}, \\ &\tilde{S}_i = \text{diag}\left(0_{4(i-1)\times 4(i-1)}, 1/2 I_4, I_{4(n-i)}\right) \in \mathbb{R}^{4n\times 4n}. \end{aligned}$$

Moreover,

$$M_{\text{uvecb}}M_{\text{uvecb}}^T = I_{v_1}, \quad M_{\text{uvecb}}^T M_{\text{uvecb}} = M_{\text{utb}}. \tag{2.4}$$

Thus, letting $\text{uvecb}^\dagger : \mathbb{R}^{v_1} \to \mathbb{R}^{2n\times 2n}$ be the right inverse of the operator 'uvecb' such that $\text{uvecb}\cdot\text{uvecb}^\dagger = 1_{v_1\times v_1}$ and $\text{uvecb}^\dagger\cdot\text{uvecb} = \text{utb}$. Then the matrix of the operator 'uvecb' is M_{uvecb}^T. That is, $\text{uvecb}^\dagger(A) = M_{\text{uvecb}}^T\text{vecb}(A)$.

Let $A = (A_{ij}) \in \mathbb{R}^{2m\times 2n}$ with $A_{ij} \in \mathbb{R}^{2\times 2}$, $i = 1,2,\cdots,m$, $j = 1,2,\cdots,n$. Like the result for the regular operator 'vec', the following result holds for 'vecb':

$$\hat{\Pi}_{m,n}\text{vecb}(A) = \text{vecb}(A^T), \tag{2.5}$$

where $\hat{\Pi}_{m,n} = (\Pi_{m,n} \otimes \Pi_{2,2}) \in \mathbb{R}^{4mn\times 4mn}$ with $\Pi_{m,n} = \sum_{i=1}^{m}\sum_{j=1}^{n}(E_{ij}\otimes E_{ij}^T)$. In these expressions, $\otimes$ denotes the Kronecker product [5, Chapter 4] and the matrix $E_{ij} \in \mathbb{R}^{m\times n}$ has entry 1 in the (i,j)-th position and zeros elsewhere. Given another matrix B, the block Kronecker product between B and A is defined by

$$B \boxtimes A = \begin{bmatrix} B\otimes A_{11} & B\otimes A_{12} & \cdots & B\otimes A_{1n} \\ B\otimes A_{21} & B\otimes A_{22} & \cdots & B\otimes A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ B\otimes A_{m1} & B\otimes A_{m2} & \cdots & B\otimes A_{mn} \end{bmatrix}.$$

For the block Kronecker product, the following results hold [7]

$$\text{vecb}\,(ACB) = (B^T \boxtimes A)\text{vecb}\,(C), \tag{2.6}$$

$$\|B\boxtimes A\|_2 = \|B\|_2\|A\|_2, \tag{2.7}$$

$$(B\boxtimes A)(C\boxtimes G) = (BC\boxtimes AG), \tag{2.8}$$

$$(B\boxtimes A)^{-1} = B^{-1}\boxtimes A^{-1}, \text{ if } B \text{ and } A \text{ are nonsingular.} \tag{2.9}$$

Here, the matrices C and G are of suitable orders and are partitioned appropriately.

In addition, the following inequality from [10, Lemma 2.1] is also necessary.

LEMMA 2.1. *For any matrix $A = (A_{ij}) \in \mathbb{R}^{2n\times 2n}$ with $A_{ij} \in \mathbb{R}^{2\times 2}$ and $D \in \mathbb{D}_{2n}$, we have*

$$\left\|\text{upb}\,(A) - D^{-1}\text{upb}\left(A^T\right)D\right\|_F \leqslant \sqrt{1+\varsigma_D^2}\|A\|_F, \tag{2.10}$$

where $\varsigma_D = \max_{1\leqslant i<j\leqslant n}\{s_j/s_i\}$.

3. Perturbation bounds for the symplectic QR factorization

Assume that the matrices A, Q_1 and R in (1.1) are perturbed as

$$A \to A + \Delta A, \ \ Q_1 \to Q_1 + \Delta Q_1, \ \ R \to R + \Delta R,$$

where $\Delta A \in \mathbb{R}^{2m\times 2n}$, $\Delta Q_1 \in \mathbb{R}^{2m\times 2n}$ is such that $(Q_1+\Delta Q_1)^T J (Q_1+\Delta Q_1) = J_1$, and $\Delta R \in \mathbb{R}^{2n\times 2n}$ is upper triangular such that $R+\Delta R$ has the same structure as that of R. Thus, the perturbed symplectic QR factorization of A in (1.1) is

$$A + \Delta A = (Q_1 + \Delta Q_1)(R + \Delta R) P^T. \tag{3.1}$$

Then

$$(A+\Delta A)^T J (A+\Delta A) = P(R+\Delta R)^T J_1 (R+\Delta R) P^T. \tag{3.2}$$

In the following, we regard ΔR as the unknown matrix of the matrix equation (3.2), and obtain the condition under which Eq. (3.2) has the unique solution.

As done in [11] and [12], from (3.2), we have

$$\begin{aligned} J_1 \Delta R R^{-1} = &\ \mathrm{upb}\left(Q_1^T J \Delta A P R^{-1} - \left(Q_1^T J \Delta A P R^{-1}\right)^T\right) + \mathrm{upb}\left(R^{-T} P^T \Delta A^T J \Delta A P R^{-1}\right) \\ &- \mathrm{upb}\left(R^{-T} \Delta R^T J_1 \Delta R R^{-1}\right). \end{aligned} \tag{3.3}$$

Applying the operator 'vecb' to (3.3) and using (2.3), (2.6) and (2.5) gives

$$\begin{aligned} \left(R^{-T} \boxtimes J_1\right) \mathrm{vecb}(\Delta R) = &\ M_{\mathrm{upb}}\left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right] \mathrm{vecb}(\Delta A P) \\ &+ M_{\mathrm{upb}}\left(R^{-T} \boxtimes R^{-T}\right) \mathrm{vecb}\left(P^T \Delta A^T J \Delta A P - \Delta R^T J_1 \Delta R\right). \end{aligned}$$

As done in [11] and noting $J_1^{-1} = J_1^T$ and (2.9), we can obtain

$$\begin{aligned} \mathrm{vecb}(\Delta R) = &\ \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right] \mathrm{vecb}(\Delta A P) \\ &+ \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left(R^{-T} \boxtimes R^{-T}\right) \mathrm{vecb}\left(P^T \Delta A^T J \Delta A P - \Delta R^T J_1 \Delta R\right) \end{aligned} \tag{3.4}$$

and show that Eq. (3.4) is equivalent to

$$\begin{aligned} &\mathrm{uvecb}(\Delta R) \\ &= M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right] \mathrm{vecb}(\Delta A P) \\ &\quad + M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left(R^{-T} \boxtimes R^{-T}\right) \mathrm{vecb}\left(P^T \Delta A^T J \Delta A P - \Delta R^T J_1 \Delta R\right). \end{aligned} \tag{3.5}$$

As a matter of convenience, let

$$\begin{aligned} G_R &= M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right], \\ H_R &= M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left(R^{-T} \boxtimes R^{-T}\right). \end{aligned}$$

Thus, applying the operator '$\mathrm{uvecb}^\dagger$' to (3.5) leads to

$$\Delta R = \mathrm{uvecb}^\dagger\left[G_R \mathrm{vecb}(\Delta A P) + H_R \mathrm{vecb}\left(P^T \Delta A^T J \Delta A P - \Delta R^T J_1 \Delta R\right)\right].$$

The above equation can be written as an operator equation for ΔR:

$$\begin{aligned}\Delta R &= \Phi(\Delta R, \Delta A) \\ &= \mathrm{uvecb}^{\dagger}\left[G_R \mathrm{vecb}(\Delta A P) + H_R \mathrm{vecb}\left(P^T \Delta A^T J \Delta A P - \Delta R^T J_1 \Delta R\right)\right].\end{aligned} \tag{3.6}$$

As done in [11] and [9], in the following, we will apply the technique of Lyapunov majorant function and the Banach fixed point principle to investigate the rigorous perturbation bounds for ΔR based on the operator equation (3.6). For completeness of the method and convenience of readers, we include the detailed process here though some steps are the same as the ones in [11].

Assume that $Z \in \mathbb{R}^{2n\times 2n}$ is upper triangular with the same structure as that of ΔR, $\|Z\|_F \leqslant \rho$ for some $\rho \geqslant 0$, and $\|\Delta A\|_F = \delta$. Then it follows from the definition of the operator 'uvecb†' and (2.1) that

$$\|\Phi(Z, \Delta A)\|_F \leqslant \|G_R\|_2 \delta + \|H_R\|_2 \delta^2 + \|H_R\|_2 \rho^2. \tag{3.7}$$

From (3.7), we have the Lyapunov majorant function of the operator equation (3.6)

$$h(\rho, \delta) = \|G_R\|_2 \delta + \|H_R\|_2 \delta^2 + \|H_R\|_2 \rho^2$$

and the Lyapunov majorant equation

$$h(\rho, \delta) = \rho, \text{ i.e., } \|G_R\|_2 \delta + \|H_R\|_2 \delta^2 + \|H_R\|_2 \rho^2 = \rho. \tag{3.8}$$

Assume that $\delta \in \Omega = \left\{\delta \geqslant 0 : 1 - 4\|H_R\|_2\left(\|G_R\|_2 \delta + \|H_R\|_2 \delta^2\right) \geqslant 0\right\}$. Then, the Lyapunov majorant equation (3.8) has two nonnegative roots: $\rho_1(\delta) \leqslant \rho_2(\delta)$ with

$$\rho_1(\delta) = f(\delta) = \frac{2\left(\|G_R\|_2 \delta + \|H_R\|_2 \delta^2\right)}{1 + \sqrt{1 - 4\|H_R\|_2 (\|G_R\|_2 \delta + \|H_R\|_2 \delta^2)}}.$$

Let the set $\mathscr{B}(\delta)$ be

$$\mathscr{B}(\delta) = \{Z \in \mathbb{R}^{2n\times 2n} : \text{Having the same strure as that of } \Delta R \text{ and } \|Z\|_F \leqslant f(\rho)\},$$

which is closed and convex. We can check that the operator $\Phi(\cdot, \Delta A)$ maps the set $\mathscr{B}(\delta)$ into itself and for $Z, \tilde{Z} \in \mathscr{B}(\delta)$,

$$\left\|\Phi(Z, \Delta A) - \Phi\left(\tilde{Z}, \Delta A\right)\right\|_F \leqslant h'_\rho(f(\delta), \delta)\left\|Z - \tilde{Z}\right\|_F.$$

Since the derivative of the function $h(\rho, \delta)$ relative to ρ at $f(\delta)$ satisfies

$$h'_\rho(f(\delta), \delta) = 1 - \sqrt{1 - 4\|H_R\|_2 (\|G_R\|_2 \delta + \|H_R\|_2 \delta^2)} < 1$$

when $\delta \in \Omega_1 = \left\{\delta \geqslant 0 : 1 - 4\|H_R\|_2\left(\|G_R\|_2 \delta + \|H_R\|_2 \delta^2\right) > 0\right\}$. Then the operator $\Phi(\cdot, \Delta A)$ is contractive on the set $\mathscr{B}(\delta)$ for $\delta \in \Omega_1$. Thus, from the Banach fixed point principle, we have that the operator equation (3.6), i.e., the matrix equation (3.2), has a unique solution in the set $\mathscr{B}(\delta)$. As a result, $\|\Delta R\|_F \leqslant f(\delta)$ for $\delta \in \Omega_1$. In this case, the unknown matrix ΔQ_1 in (3.1) is also determined uniquely. This fact can be justified by noting that ΔR is determined uniquely and $R+\Delta R$ is nonsingular, where the latter can be derived from

$$\left\|\Delta R R^{-1}\right\|_2 \leqslant 2\left\|H_R\right\|_2\|\Delta R\|_F \leqslant 2\left\|H_R\right\|_2 f(\delta)$$
$$=\frac{4\left\|H_R\right\|_2\left(\left\|G_R\right\|_2 \delta+\left\|H_R\right\|_2 \delta^2\right)}{1+\sqrt{1-4\left\|H_R\right\|_2\left(\left\|G_R\right\|_2 \delta+\left\|H_R\right\|_2 \delta^2\right)}}<1.$$

To obtain the above inequality, we have used the following inequality

$$\left\|H_R\right\|_2 \geqslant \frac{1}{2}\left\|R^{-1}\right\|_2,$$

which can be verified by noting the structures of the matrices M_{uvecb}, M_{upb}, and H_R.

In summary, we have the following main theorem.

THEOREM 3.1. *Let the unique symplectic QR factorization of $A \in \mathbb{R}^{2m \times 2n}$ be as in (1.1), $\Delta A \in \mathbb{R}^{2m \times 2n}$, and*

$$G_R=M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left[\left(R^{-T} \boxtimes Q_1^T J\right)-\left(Q_1^T J \boxtimes R^{-T}\right) \hat{\Pi}_{m, n}\right],$$
$$H_R=M_{\mathrm{uvecb}}\left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}}\left(R^{-T} \boxtimes R^{-T}\right).$$

If

$$\left\|H_R\right\|_2\left(\left\|G_R\right\|_2\|\Delta A\|_F+\left\|H_R\right\|_2\|\Delta A\|_F^2\right)<\frac{1}{4}, \tag{3.9}$$

then $A+\Delta A$ has the unique symplectic QR factorization (3.1). Moreover,

$$\|\Delta R\|_F \leqslant \frac{2\left(\left\|G_R\right\|_2\|\Delta A\|_F+\left\|H_R\right\|_2\|\Delta A\|_F^2\right)}{1+\sqrt{1-4\left\|H_R\right\|_2\left(\left\|G_R\right\|_2\|\Delta A\|_F+\left\|H_R\right\|_2\|\Delta A\|_F^2\right)}} \tag{3.10}$$
$$\leqslant 2\left(\left\|G_R\right\|_2\|\Delta A\|_F+\left\|H_R\right\|_2\|\Delta A\|_F^2\right) \tag{3.11}$$
$$<\left(1+2\left\|G_R\right\|_2\right)\|\Delta A\|_F. \tag{3.12}$$

Proof. It is easy to see that the condition (3.9) is the same as the one in Ω_1. Thus, from the discussions before Theorem 3.1, it suffices to obtain the bound (3.12). This can be done by noting (3.11) and the fact

$$2\left\|H_R\right\|_2\|\Delta A\|_F \leqslant \sqrt{1+\left\|G_R\right\|_2^2}-\left\|G_R\right\|_2<1.$$

which can be derived from (3.9). □

REMARK 3.1. From (3.10), by omitting the higher-order terms, we can get the first-order perturbation bound of R factor in (1.1):

$$\|\Delta R\|_F \leqslant\left\|G_R\right\|_2\|\Delta A\|_F+O\left(\|\Delta A\|_F^2\right), \tag{3.13}$$

under the condition which ensures that the unique symplectic QR factorization of $A+\Delta A$ exists. This condition like (3.2) in [12] or (3.1) in [10] will be weaker than (3.9). In [12, Theorem 3.2], the authors presented the following optimal first-order bound for R factor in (1.1):

$$\|\Delta R\|_F \leqslant \left\| \left(R^T \otimes J_1^T\right) \mathscr{D} K_{QR} \right\|_2 \|\Delta A\|_F + O\left(\|\Delta A\|_F^2\right), \tag{3.14}$$

where

$$\mathscr{D} = \mathrm{diag}\left(\mathscr{D}_1, \mathscr{D}_1, \mathscr{D}_2, \mathscr{D}_2, \cdots, \mathscr{D}_n, \mathscr{D}_n\right) \in \mathbb{R}^{4n^2 \times 4n^2}$$

with

$$\mathscr{D}_k = \mathrm{diag}\left(\underbrace{1, \cdots, 1}_{2k-2}, \frac{1}{2}, \frac{1}{2}, \underbrace{0, \cdots, 0}_{2n-2k} \right) \in \mathbb{R}^{2n \times 2n}, \ k = 1, 2, \cdots, n,$$

and

$$K_{QR} = \left(R^{-T} P^T \otimes Q_1^T J\right) - \left(Q_1^T J \otimes R^{-T} P^T\right) \Pi_{2m,2n}.$$

Now we show that the bound (3.14) is the same as (3.13). In fact, according to the above definition and the definition of the operator 'upb', we can check that for any matrix $X \in \mathbb{R}^{2n \times 2n}$,

$$\mathscr{D}\mathrm{vec}(X) = \mathrm{vec}(\mathrm{upb}(X)).$$

Thus, for any matrix $Y \in \mathbb{R}^{2m \times 2n}$, using the fact that $\|\mathrm{vec}(Y)\|_2 = \|\mathrm{vecb}(Y)\|_2$, (2.6), and (2.3), we have

$$\begin{aligned}
&\left\| \left(R^T \otimes J_1^T\right) \mathscr{D} K_{QR} \mathrm{vec}(Y) \right\|_2 \\
&= \left\| \left(R^T \otimes J_1^T\right) \mathscr{D} \mathrm{vec}\left(Q_1^T J Y P R^{-1} - R^{-T} P^T Y^T J^T Q_1\right) \right\|_2 \\
&= \left\| \left(R^T \otimes J_1^T\right) \mathrm{vec}\left[\mathrm{upb}\left(Q_1^T J Y P R^{-1} - R^{-T} P^T Y^T J^T Q_1\right)\right] \right\|_2 \\
&= \left\| \mathrm{vec}\left[J_1^T \mathrm{upb}\left(Q_1^T J Y P R^{-1} - R^{-T} P^T Y^T J^T Q_1\right) R\right] \right\|_2 \\
&= \left\| \mathrm{vecb}\left[J_1^T \mathrm{upb}\left(Q_1^T J Y P R^{-1} - R^{-T} P^T Y^T J^T Q_1\right) R\right] \right\|_2 \\
&= \left\| \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}} \left[\mathrm{vecb}\left(Q_1^T J Y P R^{-1} - R^{-T} P^T Y^T J^T Q_1\right)\right] \right\|_2 \\
&= \left\| \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}} \left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right) \hat{\Pi}_{m,n}\right] \mathrm{vecb}(YP) \right\|_2.
\end{aligned} \tag{3.15}$$

From the definitions of the matrices M_{utb} and M_{upb}, we can verify that

$$M_{\mathrm{utb}} \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}} = \left(R^T \boxtimes J_1^T\right) M_{\mathrm{upb}},$$

which together with (3.15) and (2.4) gives

$$\left\| \left(R^T \otimes J_1^T\right) \mathscr{D} K_{QR} \mathrm{vec}(Y) \right\|_2 = \left\| M_{\mathrm{uvecb}}^T G_R \mathrm{vecb}(YP) \right\|_2 = \|G_R \mathrm{vecb}(YP)\|_2.$$

Thus, from the definition of spectral norm and the fact that $\|\mathrm{vec}(Y)\|_2 = \|Y\|_F = \|\mathrm{vecb}(YP)\|_2$, we get

$$\begin{aligned}\left\|\left(R^T \otimes J_1^T\right)\mathscr{D}K_{QR}\right\|_2 &= \max_{\|\mathrm{vec}(Y)\|_2=1}\left\|\left(R^T \otimes J_1^T\right)\mathscr{D}K_{QR}\mathrm{vec}(Y)\right\|_2\\ &= \max_{\|\mathrm{vec}(Y)\|_2=1}\|G_R\mathrm{vecb}(YP)\|_2\\ &= \max_{\|\mathrm{vecb}(YP)\|_2=1}\|G_R\mathrm{vecb}(YP)\|_2 = \|G_R\|_2.\end{aligned}$$

So the bounds (3.14) and (3.13) are the same. Therefore, the rigorous bounds in Theorem 3.1 can be regarded as the rigorous versions of the optimal first-order perturbation bound given in [12].

REMARK 3.2. For another form of the symplectic QR factorization, the following rigorous perturbation bound were derived in [10, Theorem 5.1],

$$\|\Delta R\|_F \leqslant \left(\sqrt{6}+\sqrt{3}\right)\left(\inf_{D\in\mathbb{D}_{2n}}\sqrt{1+\varsigma_D^2}\kappa\left(D^{-1}R\right)\right)\|Q_1\|_2\|\Delta A\|_F, \tag{3.16}$$

under the condition $\|Q\|_2\left\|R^{-1}\right\|_2\|\Delta A\|_F < \sqrt{3/2}-1$. Here, for a nonsingular matrix X, $\kappa(X)$ denotes its condition number and is defined by $\kappa(X) = \|X\|_2\left\|X^{-1}\right\|_2$. As pointed out in [12], the bound still holds for the factorization (1.1). In the following, we will show that the bound (3.12) is tighter than (3.16).

In fact, similar to the proof of Corollary 3.4 in [12], for any $D \in \mathbb{D}_{2n}$ and $X \in \mathbb{R}^{2m\times 2n}$, using (2.9), (2.8) , (2.7) and (2.1), we have

$$\begin{aligned}&\|G_R\|_2\\ &= \left\|M_{\mathrm{uvecb}}\left(R^T\boxtimes J_1^T\right)\left(D^{-1}\boxtimes I_{2n}\right)(D\boxtimes I_{2n})M_{\mathrm{upb}}\left[\left(R^{-T}\boxtimes Q_1^TJ\right)-\left(Q_1^TJ\boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right]\right\|_2\\ &= \left\|M_{\mathrm{uvecb}}\left(R^TD^{-1}\boxtimes J_1^T\right)M_{\mathrm{upb}}\left[\left(DR^{-T}\boxtimes Q_1^TJ\right)-\left(DQ_1^TJ\boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right]\right\|_2\\ &\leqslant \left\|R^TD^{-1}\right\|_2\left\|M_{\mathrm{upb}}\left[\left(DR^{-T}\boxtimes Q_1^TJ\right)-\left(DQ_1^TJ\boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right]\right\|_2\\ &= \left\|R^TD^{-1}\right\|_2\max_{\|\mathrm{vecb}(X)\|_2=1}\left\|M_{\mathrm{upb}}\left[\left(DR^{-T}\boxtimes Q_1^TJ\right)-\left(DQ_1^TJ\boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right]\mathrm{vecb}(X)\right\|_2.\end{aligned} \tag{3.17}$$

Whereas, combining (2.5), (2.6), (2.3), (2.2), (2.10) and (2.1) gives

$$\begin{aligned}&\max_{\|\mathrm{vecb}(X)\|_2=1}\left\|M_{\mathrm{upb}}\left[\left(DR^{-T}\boxtimes Q_1^TJ\right)-\left(DQ_1^TJ\boxtimes R^{-T}\right)\hat{\Pi}_{m,n}\right]\mathrm{vecb}(X)\right\|_2\\ &= \max_{\|\mathrm{vecb}(X)\|_2=1}\left\|M_{\mathrm{upb}}\mathrm{vecb}\left(Q_1^TJXR^{-1}D-R^{-T}X^TJ^TQ_1D\right)\right\|_2\\ &= \max_{\|\mathrm{vecb}(X)\|_2=1}\left\|\mathrm{vecb}\left(\mathrm{upb}\left(Q_1^TJXR^{-1}D-D^{-1}\left(Q_1^TJXR^{-1}D\right)^TD\right)\right)\right\|_2\\ &= \max_{\|X\|_F=1}\left\|\mathrm{upb}\left(Q_1^TJXR^{-1}D\right)-D^{-1}\mathrm{upb}\left(\left(\left(Q_1^TJXR^{-1}D\right)^T\right)D\right)\right\|_F\\ &\leqslant \max_{\|X\|_F=1}\sqrt{1+\varsigma_D^2}\left\|Q_1^TJXR^{-1}D\right\|_F \leqslant \sqrt{1+\varsigma_D^2}\left\|R^{-1}D\right\|_2\|Q_1\|_2.\end{aligned} \tag{3.18}$$

Thus, substituting (3.18) into (3.17) yields

$$\|G_R\|_2 \leqslant \left(\inf_{D \in \mathbb{D}_{2n}} \sqrt{1+\varsigma_D^2}\kappa\left(D^{-1}R\right) \right) \|Q_1\|_2.$$

Meanwhile, from the fact $Q_1^T J Q_1 = J_1$, it is easy to obtain that

$$\left(\inf_{D \in \mathbb{D}_{2n}} \sqrt{1+\varsigma_D^2}\kappa\left(D^{-1}R\right) \right) \|Q_1\|_2 \geqslant 1.$$

Thus, we have the claimed result that the bound (3.12) is indeed tighter than (3.16). Naturally, the bound (3.10) is also always tighter than (3.16). However, in general, the bound (3.10) is not attainable. This is because, to obtain the bound (3.10), we have used the triangle inequality and the submultiplicative inequality several times. Only when all the inequalities in those inequalities are equalities can the equality in the bound (3.10) be achieved. In general, this case is infrequent. Whereas, numerical results given in Section 4 show that the bound (3.12) and hence (3.10) can be much tighter than (3.16).

REMARK 3.3. As done in [12], from

$$\Delta Q_1 = \Delta A P R^{-1} - Q_1 \Delta R R^{-1} - \Delta Q_1 \Delta R R^{-1}, \tag{3.19}$$

which is derived from (3.1) and (1.1), by omitting the higher-order terms and using (2.6), (3.4), (2.8), we obtain

$$\operatorname{vecb}(\Delta Q_1) \approx G_{Q_1} \operatorname{vecb}(\Delta A P), \tag{3.20}$$

where

$$G_{Q_1} = \left(R^{-T} \boxtimes I_{2m}\right) - \left(I_{2n} \boxtimes Q_1 J_1^T\right) M_{\mathrm{upb}} \left[\left(R^{-T} \boxtimes Q_1^T J\right) - \left(Q_1^T J \boxtimes R^{-T}\right) \hat{\Pi}_{m,n}\right].$$

Then

$$\|\Delta Q_1\|_F \leqslant \|G_{Q_1}\|_2 \|\Delta A\|_F + O\left(\|\Delta A\|_F^2\right). \tag{3.21}$$

Similar to the discussions in Remark 3.1, we can verify that

$$\left\| \left(R^{-T} P^T \otimes I_{2m}\right) + \left(I_{2n} \otimes Q_1 J_1\right) \mathscr{D} K_{QR} \right\|_2 = \|G_{Q_1}\|_2.$$

So the bound (3.21) is same as the optimal one in [12, Theorem 3.6].

Furthermore, using (3.19) and the results on R factor given in Theorem 3.1, we can obtain the rigorous perturbation bounds for Q_1 factor:

$$\begin{aligned}
\|\Delta Q_1\|_F \leqslant & \left(2+\sqrt{2}\right) \left\| \left(I_{2n} \boxtimes Q_1 J_1^T\right) M_{\mathrm{upb}} \left(R^{-T} \boxtimes R^{-T}\right) \right\|_2 \left(\|\Delta A\|_F^2 + \|\Delta R\|_F^2\right) \\
& + \left(2+\sqrt{2}\right) \|G_{Q_1}\|_2 \|\Delta A\|_F \\
\leqslant & \left(2+\sqrt{2}\right) \left[\|G_{Q_1}\|_2 + \left\| \left(R^{-T} \boxtimes Q_1\right) \right\|_2 \left(1 + \|G_R\|_2\right)\right] \|\Delta A\|_F,
\end{aligned}$$

which correspond to the optimal first-order perturbation bound (3.21). However, these two bounds are larger than the one given in [10, Theorem 5.1]. So we omit their detailed derivation.

4. Numerical experiments

In this section, a numerical example from [12] is used to compare the bounds derived in this paper with the corresponding ones given in [10] and [12].

EXAMPLE 4.1. This example is the same as Example 5.1 of [12]. Specifically, let $\varepsilon > 0$ be small enough, and

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & \varepsilon^2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

From the explanations in [12, Example 5.1], we have that A has the unique symplectic QR factorization (1.1), and $Q = Q_1 = \mathrm{diag}(\varepsilon^{-1}, \varepsilon, 1, 1), R = \mathrm{diag}(\varepsilon, \varepsilon, 1, 1)$ and $P = [e_1, e_3, e_2, e_4]$. Using the expressions of G_R and G_{Q_1}, we can verify that

$$G_R = \begin{bmatrix} G_{11} & 0 & 0 & 0 \\ 0 & G_{22} & G_{23} & 0 \\ 0 & 0 & 0 & G_{34} \end{bmatrix} \in \mathbb{R}^{12\times 16}, \quad G_{Q_1} = \begin{bmatrix} H_{11} & 0 & 0 & 0 \\ 0 & 0 & H_{23} & 0 \\ 0 & 0 & H_{33} & 0 \\ 0 & 0 & 0 & H_{44} \end{bmatrix} \in \mathbb{R}^{16\times 16},$$

where

$$G_{11} = \begin{bmatrix} \frac{\varepsilon}{2} & 0 & 0 & \frac{1}{2\varepsilon} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{\varepsilon}{2} & 0 & 0 & \frac{1}{2\varepsilon} \end{bmatrix}, G_{22} = \begin{bmatrix} \varepsilon & 0 & 0 & 0 \\ 0 & \frac{1}{\varepsilon} & 0 & 0 \\ 0 & 0 & \varepsilon & 0 \\ 0 & 0 & 0 & \frac{1}{\varepsilon} \end{bmatrix}, G_{23} = \begin{bmatrix} 0 & 0 & 0 & \frac{1}{\varepsilon} \\ 0 & -\frac{1}{\varepsilon} & 0 & 0 \\ 0 & 0 & -\frac{1}{\varepsilon} & 0 \\ \frac{1}{\varepsilon} & 0 & 0 & 0 \end{bmatrix}, G_{34} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix},$$

$$H_{11} = \begin{bmatrix} \frac{1}{2\varepsilon} & 0 & 0 & -\frac{1}{2\varepsilon^3} \\ 0 & \frac{1}{\varepsilon} & 0 & 0 \\ 0 & 0 & \frac{1}{\varepsilon} & 0 \\ -\frac{\varepsilon}{2} & 0 & 0 & \frac{1}{2\varepsilon} \end{bmatrix}, H_{23} = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{\varepsilon^2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\varepsilon^2} & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, H_{33} = \frac{1}{\varepsilon} I_4, H_{44} = \begin{bmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{bmatrix}.$$

Thus, upon computation, we have

$$b_{(3.13)} = \|G_R\|_2 = \frac{\sqrt{2}}{\varepsilon}, \quad b_{(3.14)} = \left\| \left(R^T \otimes J_1^T\right) \mathscr{D} K_{QR} \right\|_2 = \frac{\sqrt{2}}{\varepsilon},$$

$$b_{(3.21)} = \|G_{Q_1}\|_2 = \frac{1+\varepsilon^4}{2\varepsilon^3} = \left\| \left(R^{-T}P^T \otimes I_{2m}\right) + \left(I_{2n} \otimes Q_1 J_1\right) \mathscr{D} K_{QR} \right\|_2,$$

and

$$b_{(3.12)} = 1 + 2\|G_R\|_2 = 1 + \frac{2\sqrt{2}}{\varepsilon},$$

$$b_{(3.16)} = \left(\sqrt{6}+\sqrt{3}\right) \left(\inf_{D\in\mathbb{D}_{2n}} \sqrt{1+\varsigma_D^2 \kappa\left(D^{-1}R\right)} \right) \|Q_1\|_2 = (\sqrt{6}+\sqrt{3}) \frac{\sqrt{1+\varepsilon^2}}{\varepsilon^2}.$$

Here, the results from [12, Example 5.1] are used to obtain $b_{(3.16)}$. It is easy to see that, for this example, the first-order perturbation bounds (3.13) and (3.21) are the same as the corresponding optimal ones given in [12], and the rigorous perturbation bound (3.12) is tighter than the one from [10], i.e., the bound (3.16). These results confirm the analysis given in Remarks 3.1, 3.2, and 3.3. To illustrate the differences between the bounds (3.12) and (3.16) clearly, a figure on the values of ε and the corresponding bounds (3.12) and (3.16) is given below. More specifically, we set $\varepsilon = 10^{-2}$, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}, and 10^{-8} and then compute the corresponding bounds (3.12) and (3.16). This figure shows that the bound (3.12) can be much tighter than (3.16).

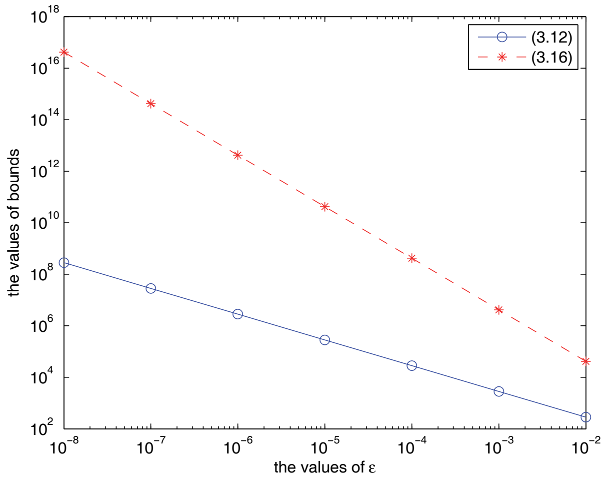

Figure 1: *Comparison of the bounds* (3.12) *and* (3.16)

5. Concluding remarks

In this note, we consider the rigorous perturbation bounds for R factor of the implicit Bunch form of the symplectic QR factorization when the original matrix has the normwise perturbation. Using the approach in this paper, we can also obtain the rigorous perturbation bounds for R factor with componentwise perturbation, i.e., the perturbation $\Delta A \in \mathbb{R}^{2m\times 2n}$ of A satisfies:

$$|\Delta A| \leqslant \varepsilon C\,|A|,$$

where $C = (c_{ij}) \in \mathbb{R}^{2m\times 2m}$ with $0 \leqslant c_{ij} \leqslant 1$, $\varepsilon \geqslant 0$ is a small constant, and for any matrix $X = (x_{ij})$, $|X|$ is define by $(|x_{ij}|)$. However, we cannot show that the obtained bounds are always tighter than the corresponding one in [10] in theory though the former behave better in numerical experiments. So these results are not presented in this note.

Acknowledgements. The authors would like to thank the editors, prof. S. Puntanen and Dr. Marjan Praljak, and two anonymous reviewers for their valuable comments and helpful suggestions, which improved the paper.

REFERENCES

[1] R. BHATIA, *Matrix factorizations and their perturbations*, Linear Algebra Appl. **197**, (1994), 245–276.
[2] A. BUNSE-GERSTNER, *Matrix factorizations for symplectic QR-like methods*, Linear Algebra Appl. **83**, (1986), 49–77.
[3] A. BUNSE-GERSTNER AND V. MEHRMANN, *A symplectic QR-like algorithm for the solution of the real algebraic Riccati equation*, IEEE Trans. Automat. Control **31**, 12 (1986), 1104–1113.
[4] X. W. CHANG, *On the sensitivity of the SR decomposition*, Linear Algebra Appl. **282**, (1998), 297–310.
[5] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
[6] M. KONSTANTINOV, D. GU, V. MEHRMANN AND P. PETKOV, *Perturbation Theory for Matrix Equations*, Elsevier, Amsterdam, 2003.
[7] R. H. KONING, H. NEUDECKER AND T. WANSBEEK, *Block Kronecker products and the vecb operator*, Linear Algebra Appl. **149**, (1991), 165–184.
[8] D. KRESSNER, *Block algorithms for orthogonal symplectic factorizations*, BIT **43**, (2003), 775–790.
[9] H. Y. LI AND Y. M. WEI, *Improved rigorous perturbation bounds for the LU and QR factorizations*, Linear Algebra Appl. **22**, (2015), 1115–1130.
[10] H. Y. LI, H. YANG AND H. SHAO, *Perturbation analysis for the symplectic QR factorization*, Linear Multilinear Algebra **63**, (2015), 78–96.
[11] H. Y. LI, Y. M. WEI AND Y. F. YANG, *New rigorous perturbation bounds for the Cholesky-like factorization of skew-symmetric matrix*, Linear Algebra Appl. **491**, (2016), 83–100.
[12] W. LI, Z. J. XIE AND S. W. VONG, *Sensitivity analysis for the symplectic QR factorization*, J. Franklin Inst. **353**, 5 (2016), 1186–1205.
[13] S. SINGER AND S. SINGER, *Rounding error and perturbation bounds for the symplectic QR factorization*, Linear Algebra Appl. **358**, (2003), 255–279.
[14] G. M. STEWART AND J. G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990.
[15] D. S. WATKINS AND L. ELSNER, *Self-similar flows*, Linear Algebra Appl. **110**, (1988), 213–242.

(Received April 23, 2016)

Hanyu Li
College of Mathematics and Statistics
Chongqing University
Chongqing 401331, P. R. China
e-mail: lihy.hy@gmail.com, hyli@cqu.edu.cn

Peng Lv
College of Mathematics and Statistics
Chongqing University
Chongqing 401331, P. R. China
e-mail: lvpeng@cqu.edu.cn