Mathematical
Inequalities
& Applications
Volume 22, Number 1 (2019), 133–149 doi:10.7153/mia-2019-22-10

IMPROVED RIGOROUS MULTIPLICATIVE PERTURBATION BOUNDS FOR THE GENERALIZED CHOLESKY FACTORIZATION AND THE CHOLESKY–LIKE FACTORIZATION

AAMIR FAROOQ, MAHVISH SAMAR, HANYU LI* AND CHUNLAI MU

(*Communicated by S. Puntanen*)

Abstract. Some improved rigorous multiplicative perturbation bounds for the generalized Cholesky factorization and the Cholesky-like factorization which are two generalizations of the classic Cholesky factorization are obtained by bringing together the modified matrix-vector equation approach with the method of Lyapunov majorant function and the Banach fixed point theorem. These bounds are continually tighter than the corresponding ones given in the literature.

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 $\mathbb{R}^{m\times n}$ consisting of matrices with rank r. Let I_r be the identity matrix of order r. For a matrix $A\in\mathbb{R}^{m\times n}$, we denote by A^T and $A[<i>]$ the transpose and the i-th leading principal submatrix of A, respectively.

Consider the following symmetric quasi-definite matrix $K\in\mathbb{R}^{(m+n)\times(m+n)}$

$$K=\begin{bmatrix} A & B^T \\ B & -C \end{bmatrix}, \tag{1.1}$$

here $A\in\mathbb{R}_m^{m\times m}$ is symmetric positive definite, $B\in\mathbb{R}_n^{n\times m}$, and $C\in\mathbb{R}^{n\times n}$ is symmetric positive semi-definite. The matrix K has the following factorization

$$K=LJ_{m+n}L^T, \tag{1.2}$$

where

$$L=\begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix},\ J_{m+n}=\begin{bmatrix} I_m & 0 \\ 0 & -I_n \end{bmatrix},$$

Mathematics subject classification (2010): 15A23, 15A45.
Keywords and phrases: Generalized Cholesky factorization, Cholesky-like factorization, multiplicative perturbation, Lyapunov majorant fuction, Banach fixed point theorem, rigorous perturbation bound.
This research is supported by the National Natural Science Foundation of China (Grant Nos. 11671060; 11771062).
* Corresponding author.

with $L_{11} \in \mathbb{R}_m^{m\times m}$ and $L_{22} \in \mathbb{R}_n^{n\times n}$ being lower triangular, and $L_{21} \in \mathbb{R}_n^{n\times m}$. It is easy to check that

$$A = L_{11}L_{11}^T, \;\; B = L_{21}L_{11}^T, \;\; C + L_{21}L_{21}^T = L_{22}L_{22}^T.$$

The factorization (1.2) is called the generalized Cholesky factorization and L is referred to as the generalized Cholesky factor [20].

Now, we consider the skew-symmetric matrix $B \in \mathbb{R}^{2n\times 2n}$. If all even leading principal submatrices of B are nonsingular, i.e., $B[\langle 2i\rangle]$ $(i = 1,\cdots,n)$ are nonsingular, then B has the following factorization

$$B = R^T \widehat{J}_{2n} R, \tag{1.3}$$

where $R = (r_{ij}) \in \mathbb{R}^{2n\times 2n}$ is upper triangular with $r_{2j-1,2j} = 0, r_{2j-1,2j-1} > 0$, $r_{2j,2j} = \pm r_{2j-1,2j-1}$ for $j = 1,2,\cdots,n$ and

$$\widehat{J}_{2n} = \mathrm{diag}(J_0,\cdots,J_0), \; J_0 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

That is, R has 2×2 blocks of the form $\begin{bmatrix} r & 0 \\ 0 & \pm r \end{bmatrix}$ running down the main diagonal. The factorization (1.3) is called the Cholesky-like factorization and R is referred to as the Cholesky-like factor [1].

In recent years, algorithms, error analysis of algorithms, and perturbation analysis for these two factorizations had been studied by some authors [1, 5, 9, 10, 11, 17, 18, 20]. Among these, Li and Yang [10] considered the multiplicative rigorous perturbation bounds using the matrix equation and the refined matrix equation approaches. The multiplicative perturbation has some advantages compared with the additive perturbation; see e.g. [3, 10, 19] for detailed explanation. In this paper, we continue the research on multiplicative perturbation bounds for the above two factorizations using the (block) matrix-vector equation, the method of Lyapunov majorant fuction (e.g., [8, Chapter 5]), and the Banach fixed point theorem (e.g., [8, Appendix 5]). The obtained bounds will enhance the corresponding results given in [10].

2. Notation and preliminaries

The majority of the notation and preliminaries endorsed in this section are from [2, 11, 14]. We still present them here to make easier for readers.

In a given matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$, its spectral norm and Frobenius norm are betoken by $\|A\|_2$ and $\|A\|_F$, respectively. For these two matrix norms, the following inequalities clasp (see [16]):

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

whenever the matrix product XYZ is well-defined.

For the given matrix $A=[a_1,a_2,\ldots,a_n]=(a_{ij})\in\mathbb{R}^{n\times n}$, we denote the vector of the first i elements of a_j by $a_j^{(i)}$. With these, we adopt the operators as in [2],

$$\mathrm{uvec}(A):=\begin{bmatrix} a_1^{(1)} \\ a_2^{(2)} \\ \vdots \\ a_n^{(n)} \end{bmatrix}\in\mathbb{R}^{\nu_1},\mathrm{vec}(A):=\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}\in\mathbb{R}^{n^2},$$

$$\mathrm{up}(A):=\begin{bmatrix} \frac{1}{2}a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & \frac{1}{2}a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{2}a_{nn} \end{bmatrix},\quad \mathrm{ut}(A):=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix},$$

where $\nu_1=n(n+1)/2$. Considering the structures of these operators, we have

$$\Pi_{mn}\mathrm{vec}(A)=\mathrm{vec}(A^T),\quad \mathrm{uvec}(A)=M_{\mathrm{uvec}}\mathrm{vec}(A),$$
$$\mathrm{vec}(\mathrm{ut}(A))=M_{\mathrm{ut}}\mathrm{vec}(A),\quad \mathrm{vec}(\mathrm{up}(A))=M_{\mathrm{up}}\mathrm{vec}(A), \tag{2.2}$$

where Π_{mn} is the vec-permutation matrix depending only on the dimension of matrix and

$$M_{\mathrm{uvec}}=\mathrm{diag}(J_1,J_2,\cdots,J_n)\in\mathbb{R}^{\nu_1\times n^2},\ J_i=\left[I_i,0_{i\times(n-i)}\right]\in\mathbb{R}^{i\times n},$$
$$M_{\mathrm{ut}}=\mathrm{diag}\left(\hat{J}_1,\hat{J}_2,\cdots,\hat{J}_n\right)\in\mathbb{R}^{n^2\times n^2},\ \hat{J}_i=\mathrm{diag}\left(I_i,0_{(n-i)\times(n-i)}\right)\in\mathbb{R}^{n\times n},$$
$$M_{\mathrm{up}}=\mathrm{diag}\left(\tilde{J}_1,\tilde{J}_2,\cdots,\tilde{J}_n\right)\in\mathbb{R}^{n^2\times n^2},\ \tilde{J}_i=\mathrm{diag}\left(I_{i-1},1/2,0_{(n-i)\times(n-i)}\right)\in\mathbb{R}^{n\times n}.$$

Here, $0_{s\times t}$ is the $s\times t$ zero matrix. Moreover,

$$M_{\mathrm{uvec}}M_{\mathrm{uvec}}^T=I_{\nu_1},\quad M_{\mathrm{uvec}}^TM_{\mathrm{uvec}}=M_{\mathrm{ut}}. \tag{2.3}$$

Let $\mathrm{uvec}^\dagger:\mathbb{R}^{\nu_1}\to\mathbb{R}^{m\times n}$ be the right inverse of the operator 'uvec' such that $\mathrm{uvec}\cdot\mathrm{uvec}^\dagger=1_{\nu_1\times\nu_1}$ and $\mathrm{uvec}^\dagger\cdot\mathrm{uvec}=\mathrm{ut}$. Then the matrix of the operator 'uvec' is M_{uvec}^T. That is, $\mathrm{uvec}^\dagger(A)=M_{\mathrm{uvec}}^T\mathrm{vec}(A)$.

Let $\mathbb{D}_{m+n}\in\mathbb{R}^{(m+n)\times(m+n)}$ denote the set of $(m+n)\times(m+n)$ positive definite diagonal matrices. Then, for any $D_{m+n}=\mathrm{diag}(\delta_1,\delta_2,\ldots,\delta_{m+n})\in\mathbb{D}_{m+n}$ and $A\in\mathbb{R}^{(m+n)\times(m+n)}$,

$$\mathrm{up}(AD_{m+n})=\mathrm{up}(A)D_{m+n},\quad D_{m+n}\mathrm{up}(A)=D_{m+n}\mathrm{up}(A). \tag{2.4}$$

Furthermore, from [4] we have

$$\left\|\mathrm{up}(A)+D_{m+n}^{-1}\mathrm{up}\left(A^T\right)D_{m+n}\right\|_F\leqslant\sqrt{1+\varsigma_{D_{m+n}}^2}\|A\|_F, \tag{2.5}$$

where $\varsigma_{D_{m+n}}=\max\limits_{1\leqslant i<j\leqslant m+n}\{\delta_j/\delta_i\}$.

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 adopt the following operators:

$$\operatorname{uvecb}(A)=\left[\begin{array}{c}\operatorname{vec}(A_{11})\\ \vdots\\ \operatorname{vec}(A_{1n})\\ \hline \operatorname{vec}(A_{22})\\ \vdots\\ \operatorname{vec}(A_{2n})\\ \hline \vdots\\ \hline \operatorname{vec}\left(A_{(n-1)(n-1)}\right)\\ \operatorname{vec}\left(A_{(n-1)n}\right)\\ \hline \operatorname{vec}(A_{nn})\end{array}\right]\in\mathbb{R}^{v_2},\quad \operatorname{vecb}(A)=\left[\begin{array}{c}\operatorname{vec}(A_{11})\\ \vdots\\ \operatorname{vec}(A_{1n})\\ \hline \vdots\\ \hline \operatorname{vec}(A_{n1})\\ \vdots\\ \operatorname{vec}(A_{nn})\end{array}\right]\in\mathbb{R}^{4n^2},$$

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

where $v_2=2n(n+1)$. Making use of the structures of these operators, we have

$$\hat{\Pi}_{nn}\operatorname{vecb}(A)=\operatorname{vecb}(A^T),\quad \operatorname{uvecb}(A)=M_{\text{uvecb}}\operatorname{vecb}(A),$$
$$\operatorname{vecb}(\operatorname{utb}(A))=M_{\text{utb}}\operatorname{vecb}(A),\quad \operatorname{vecb}(\operatorname{upb}(A))=M_{\text{upb}}\operatorname{vecb}(A),\tag{2.6}$$

where $\hat{\Pi}_{nn}=(\Pi_{nn}\otimes\Pi_{22})\in\mathbb{R}^{4nn\times 4nn}$ with $\otimes$ denoting the Kronecker product whose definition is given later in this section and

$$\begin{aligned}
&M_{\text{uvecb}}=\operatorname{diag}(S_1,S_2,\cdots,S_n)\in\mathbb{R}^{v_2}\times 4n^2,\\
&S_i=\left[0_{4(n-i+1)\times 4(i-1)},I_{4(n-i+1)}\right]\in\mathbb{R}^{4(n-i+1)\times 4n},\\
&M_{\text{utb}}=\operatorname{diag}\left(\hat{S}_1,\hat{S}_2,\cdots,\hat{S}_n\right)\in\mathbb{R}^{4n^2\times 4n^2},\\
&\hat{S}_i=\operatorname{diag}\left(0_{4(i-1)\times 4(i-1)},I_{4(n-i+1)}\right)\in\mathbb{R}^{4n\times 4n},\\
&M_{\text{upb}}=\operatorname{diag}\left(\tilde{S}_1,\tilde{S}_2,\cdots,\tilde{S}_n\right)\in\mathbb{R}^{4n^2\times 4n^2},\\
&\tilde{S}_i=\operatorname{diag}\left(0_{4(i-1)\times 4(i-1)},1/2I_4,I_{4(n-i)}\right)\in\mathbb{R}^{4n\times 4n}.
\end{aligned}$$

Moreover,

$$M_{\text{uvecb}}M_{\text{uvecb}}^T=I_{v_1},\quad M_{\text{uvecb}}^T M_{\text{uvecb}}=M_{\text{utb}}.\tag{2.7}$$

Thus, letting $\operatorname{uvecb}^{\dagger}:\mathbb{R}^{v_2}\to\mathbb{R}^{2n\times 2n}$ be the right inverse of the operator 'uvecb' such that $\operatorname{uvecb}\cdot\operatorname{uvecb}^{\dagger}=1_{v_2\times v_2}$ and $\operatorname{uvecb}^{\dagger}\cdot\operatorname{uvecb}=\operatorname{utb}$. Then the matrix of the operator 'uvecb' is M_{uvecb}^T. That is, $\operatorname{uvecb}^{\dagger}(A)=M_{\text{uvecb}}^T\operatorname{vecb}(A)$.

Let $\mathbb{D}_{2n} \in \mathbb{R}^{2n\times 2n}$ denote the set of $2n \times 2n$ diagonal positive definite matrices with 2×2 main diagonal blocks $s_i I_2$, where $s_i > 0$, $i = 1,2,\cdots,n$. Then, for any $D_{2n} \in \mathbb{D}_{2n}$ and $A = (A_{ij}) \in \mathbb{R}^{2n\times 2n}$,

$$\mathrm{upb}(AD_{2n}) = \mathrm{upb}(A)D_{2n}, \quad \mathrm{upb}(D_{2n}A) = D_{2n}\mathrm{upb}(A). \tag{2.8}$$

Furthermore, from [10, Lemma 1.1],

$$\left\|\mathrm{upb}\,(A) + D_{2n}^{-1}\mathrm{upb}\left(A^T\right)D_{2n}\right\|_F \leqslant \sqrt{1+\varsigma_{D_{2n}}^2}\,\|A\|_F, \tag{2.9}$$

where $\varsigma_{D_{2n}} = \max\limits_{1\leqslant i<j\leqslant n}\{s_j/s_i\}$.

Let $A = (A_{ij}) \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times q}$. The Kronecker product between A and B is defined by [6, Chapter 4],

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

For the Kronecker product, the following results hold [6, Chapter 4]

$$\mathrm{vec}\,(ACB) = (B^T \otimes A)\mathrm{vec}\,(C)\,, \tag{2.10}$$

$$\|B\otimes A\|_2 = \|B\|_2\|A\|_2, \tag{2.11}$$

$$(B\otimes A)(C\otimes G) = (BC)\otimes(AG), \tag{2.12}$$

$$(B\otimes A)^{-1} = B^{-1}\otimes A^{-1}, \text{ if } B \text{ and } A \text{ are nonsingular.} \tag{2.13}$$

In above expression, the matrices C and G are of suitable orders.

Let $A = (A_{ij}) \in \mathbb{R}^{2m\times 2n}$ with $A_{ij} \in \mathbb{R}^{2\times 2}$, $i = 1,2,\cdots,m$, $j = 1,2,\cdots,n$. 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]

$$\mathrm{vecb}\,(ACB) = (B^T \boxtimes A)\mathrm{vecb}\,(C)\,, \tag{2.14}$$

$$\|B\boxtimes A\|_2 = \|B\|_2\|A\|_2, \tag{2.15}$$

$$(B\boxtimes A)(C\boxtimes G) = (BC)\boxtimes(AG), \tag{2.16}$$

$$(B\boxtimes A)^{-1} = B^{-1}\boxtimes A^{-1}, \text{ if } B \text{ and } A \text{ are nonsingular.} \tag{2.17}$$

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

3. Perturbation bounds for the generalized Cholesky factorization

Assume that the matrices K and L in (1.2) are perturbed as

$$K \to QKQ^T, \quad L \to L + \Delta L,$$

where $Q = I_{m+n} + E \in \mathbb{R}^{(m+n)\times(m+n)}$ and $\Delta L \in \mathbb{R}^{(m+n)\times(m+n)}$ is lower triangular such that $L + \Delta L$ has the same structure as that of L. So,

$$\begin{aligned} &\widetilde{K} = QKQ^T \\ &(I_{m+n} + E)K(I_{m+n} + E)^T = (L + \Delta L)J_{m+n}(L + \Delta L)^T. \end{aligned} \tag{3.1}$$

Expanding (3.1) and using (1.2), we have

$$\begin{aligned} LJ_{m+n}L^T E^T + ELJ_{m+n}L^T + ELJ_{m+n}L^T E^T = LJ_{m+n}\Delta L^T + \Delta LJ_{m+n}L^T \\ + \Delta LJ_{m+n}\Delta L^T. \end{aligned} \tag{3.2}$$

Premultipling the above equation by L^{-1} and postmultipling it by L^{-T} leads to

$$\begin{aligned} J_{m+n}\Delta L^T L^{-T} + L^{-1}\Delta LJ_{m+n} &= J_{m+n}\Delta L^T L^{-T} + (J_{m+n}\Delta L^T L^{-T})^T \\ &= J_{m+n}L^T E^T L^{-T} + L^{-1}ELJ_{m+n} + L^{-1}ELJ_{m+n}L^T E^T L^{-T} \\ &\quad - L^{-1}\Delta LJ_{m+n}\Delta L^T L^{-T}. \end{aligned} \tag{3.3}$$

As done in [12, 13], from (3.3), we have

$$\begin{aligned} J_{m+n}\Delta L^T L^{-T} &= \operatorname{up}\left(J_{m+n}L^T E^T L^{-T} + L^{-1}ELJ_{m+n}\right) \\ &\quad + \operatorname{up}\left(L^{-1}\left(ELJ_{m+n}L^T E^T L^{-T} - L^{-1}\Delta LJ_{m+n}\Delta L^T\right)L^{-T}\right). \end{aligned} \tag{3.4}$$

Applying the operator ‘vec’ to (3.4), and using (2.10) and (2.2) gives

$$\begin{aligned} &\left(L^{-1} \otimes J_{m+n}\right)\operatorname{vec}\left(\Delta L^T\right) \\ &= M_{\text{up}}\left((LJ_{m+n})^T \otimes L^{-1} + \left(L^{-1} \otimes (J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)}\right)\operatorname{vec}(E) \\ &\quad + M_{\text{up}}\left(L^{-1} \otimes L^{-1}\right)\operatorname{vec}\left(ELJ_{m+n}L^T E^T - \Delta LJ_{m+n}\Delta L^T\right). \end{aligned}$$

As done in [12], we can obtain

$$\begin{aligned} &\operatorname{vec}\left(\Delta L^T\right) \\ &= \left(L \otimes J_{m+n}^{-1}\right)M_{\text{up}}\left((LJ_{m+n})^T \otimes L^{-1} + \left(L^{-1} \otimes (J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)}\right)\operatorname{vec}(E) \\ &\quad + \left(L \otimes J_{m+n}^{-1}\right)M_{\text{up}}\left(L^{-1} \otimes L^{-1}\right)\operatorname{vec}\left(ELJ_{m+n}L^T E^T - \Delta LJ_{m+n}\Delta L^T\right). \end{aligned} \tag{3.5}$$

and show that (3.5) is equivalent to

$$\begin{aligned} &\operatorname{uvec}\left(\Delta L^T\right) \\ &= M_{\text{uvec}}\left(L \otimes J_{m+n}^{-1}\right)M_{\text{up}}\left(\left((LJ_{m+n})^T \otimes L^{-1}\right) + \left(L^{-1} \otimes (J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)}\right)\operatorname{vec}(E) \\ &\quad + M_{\text{uvec}}\left(L \otimes J_{m+n}^{-1}\right)M_{\text{up}}\left(L^{-1} \otimes L^{-1}\right)\operatorname{vec}\left(ELJ_{m+n}L^T E^T - \Delta LJ_{m+n}\Delta L^T\right). \end{aligned} \tag{3.6}$$

As an issue of comfort, let

$$G_L = M_{\text{uvec}}\left(L\otimes J_{m+n}^{-1}\right) M_{\text{up}}\left(\left((LJ_{m+n})^T\otimes L^{-1}\right)+\left(L^{-1}\otimes (J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)}\right),$$
$$H_L = M_{\text{uvec}}\left(L\otimes J_{m+n}^{-1}\right) M_{\text{up}}\left(L^{-1}\otimes L^{-1}\right).$$

Thus, applying the operator 'uvec†' to (3.6) leads

$$\Delta L^T = \text{uvec}^\dagger\left(G_L\text{vec}(E)+H_L\text{vec}\left(ELJ_{m+n}L^TE^T-\Delta LJ_{m+n}\Delta L^T\right)\right).$$

The above equation can be written as an operator equation for ΔL^T :

$$\begin{aligned}\Delta L^T &= \Phi\left(\Delta L^T, E\right)\\ &= \text{uvec}^\dagger\left(G_L\text{vec}(E)+H_L\text{vec}\left(ELJ_{m+n}L^TE^T-\Delta LJ_{m+n}\Delta L^T\right)\right).\end{aligned}\tag{3.7}$$

We will execute the technique of Lyapunov majorant function and the Banach fixed point theorem to probe the rigorous perturbation bounds for ΔL^T based on the operator equation (3.7) as done in [12]. To make it easy and clear for readers and for plenum of the method, we comprehend the detail process here through some steps which are same as in [12].

Assume that $Z\in\mathbb{R}^{(m+n)\times(m+n)}$ is upper triangular with the same structure as that of ΔL^T, $\|Z\|_F\leqslant\rho$ for some $\rho\geqslant 0$, and $\|E\|_F=\delta$. Then it follows from the definition of the operator 'uvec†' and (2.1) that

$$\|\Phi(Z,E)\|_F\leqslant\|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2+\|H_L\|_2\rho^2.\tag{3.8}$$

From (3.8), we have the Lyapunov majorant function of the operator equation (3.7)

$$h(\rho,\delta)=\|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2+\|H_L\|_2\rho^2$$

and the Lyapunov majorant equation

$$h(\rho,\delta)=\rho,\ \text{ i.e., }\ \|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2+\|H_L\|_2\rho^2=\rho.\tag{3.9}$$

Assume that $\delta\in\Omega=\left\{\delta\geqslant 0: 1-4\|H_L\|_2\left(\|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2\right)\geqslant 0\right\}$. Then, the Lyapunov majorant equation (3.9) has two nonnegative roots: $\rho_1(\delta)\leqslant\rho_2(\delta)$ with

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

Let the set $B(\delta)$ be

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

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

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

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

$$h'_\rho(f(\delta),\delta) = 1-\sqrt{1-4\|H_L\|_2\left(\|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2\right)} < 1$$

when $\delta\in\Omega_1=\left\{\delta\geqslant 0: 1-4\|H_L\|_2\left(\|G_L\|_2\delta+\|L\|_2^2\|H_L\|_2\delta^2\right)>0\right\}$. Then the operator $\Phi(\cdot,E)$ is contractive on the set $B(\delta)$ for $\delta\in\Omega_1$. Thus, from the Banach fixed point theorem, we have that the operator equation (3.7), i.e., the matrix equation (3.1), has a unique solution in the set $B(\delta)$. As a result, $\left\|\Delta L^T\right\|_F\leqslant f(\delta)$ for $\delta\in\Omega_1$. In summary, we have the following main theorem.

THEOREM 3.1. *Suppose that $K\in\mathbb{R}^{(m+n)\times(m+n)}$ is defined by* (1.1) *and has the factorization* (1.2). *Let $Q=I_{m+n}+E\in\mathbb{R}^{(m+n)\times(m+n)}$. If*

$$\|H_L\|_2\left(\|G_L\|_2\|E\|_F+\|L\|_2^2\|H_L\|_2\|E\|_F^2\right)<\frac{1}{4}, \tag{3.10}$$

then $\widetilde{K}=QKQ^T$ has the unique generalized Cholesky factorization

$$\widetilde{K}=QKQ^T=(L+\Delta L)J_{m+n}(L+\Delta L)^T,$$

and

$$\|\Delta L\|_F \leqslant \frac{2\left(\|G_L\|_2\|E\|_F+\|L\|_2^2\|H_L\|_2\|E\|_F^2\right)}{1+\sqrt{1-4\|H_L\|_2\left(\|G_L\|_2\|E\|_F+\|L\|_2^2\|H_L\|_2\|E\|_F^2\right)}} \tag{3.11}$$

$$\leqslant 2\left(\|G_L\|_2\|E\|_F+\|L\|_2^2\|H_L\|_2\|E\|_F^2\right) \tag{3.12}$$

$$<\left(\|L\|_2+2\|G_L\|_2\right)\|E\|_F. \tag{3.13}$$

Proof. It is anything but difficult to see that the condition (3.10) is the same as the one in Ω_1. Thus, from the discussions before "Theorem 3.1." , it suffices to obtain the bound (3.13). This can be done by noting (3.12) and the fact

$$2\|L\|_2\|H_L\|_2\|E\|_F\leqslant\sqrt{1+\frac{\|G_L\|_2^2}{\|L\|_2^2}}-\frac{\|G_L\|_2}{\|L\|_2}<1. \tag{3.14}$$

which can be derived from (3.10). □

REMARK 3.1. In [10], by composite of the classic and refined matrix equation approaches, the following rigorous multiplicative perturbation bound was obtained,

$$\frac{\|\Delta L\|_F}{\|L\|_2}\leqslant(\sqrt{3}+\sqrt{6})\inf_{D_{m+n}\in\mathbb{D}_{m+n}}\sqrt{1+\zeta_{D_{m+n}}^2}\kappa(D_{m+n}L^{-1})\|Q-I_{m+n}\|_F. \tag{3.15}$$

under the condition

$$\kappa(L)\|E\|_F < (\sqrt{6}-2)/2. \tag{3.16}$$

In the accompanying, we will demonstrate that the bound (3.13) is sharper than (3.15). Actually, like the evidence of Corollary 3.4 in [15], for any $D_{m+n} \in \mathbb{D}_{m+n}$ and $X \in \mathbb{R}^{(m+n)\times(m+n)}$, using (2.13), (2.12), (2.11) and (2.1), we have

$$\|G_L\|_2 = \left\|M_{\text{uvec}}\,(L\otimes J_{m+n})\,(D_{m+n}^{-1}\otimes I_{m+n})\,(D_{m+n}\otimes I_{m+n})M_1\right\|_2 \text{ with}$$

$$\begin{aligned} M_1 &= M_{\text{up}}\left((LJ_{m+n})^T\otimes L^{-1} + \left(L^{-1}\otimes(J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)}\right) \\ &= \left\|M_{\text{uvec}}\left((LD_{m+n}^{-1})\otimes J_{m+n}\right)M_{\text{up}}\left(\left((D_{m+n}(LJ_{m+n})^T)\otimes L^{-1}\right)+M_2\right)\right\|_2 \text{ with} \end{aligned}$$

$$\begin{aligned} M_2 &= \left((D_{m+n}L^{-1})\otimes(J_{m+n}L^T)\right)\Pi_{(m+n)(m+n)} \\ &\leqslant \left\|LD_{m+n}^{-1}\right\|_2\left\|M_{\text{up}}\left(\left((D_{m+n}(LJ_{m+n})^T)\otimes L^{-1}\right)+M_2\right)\right\|_2 \\ &= \left\|LD_{m+n}^{-1}\right\|_2 \max_{\|\text{vec}(X)\|_2=1}\left\|M_{\text{up}}\left(\left((D_{m+n}(LJ_{m+n})^T)\otimes L^{-1}\right)+M_2\right)\text{vec}\,(X)\right\|_2. \end{aligned} \tag{3.17}$$

Whereas, combining (2.10), (2.2), (2.4), (2.5) and (2.1) gives

$$\begin{aligned} &\max_{\|\text{vec}(X)\|_2=1}\left\|M_{\text{up}}\left(\left((D_{m+n}(LJ_{m+n})^T)\otimes L^{-1}\right)+M_2\right)\text{vec}\,(X)\right\|_2 \\ &= \max_{\|\text{vec}(X)\|_2=1}\left\|M_{\text{up}}\text{vec}\left(J_{m+n}L^TX^TL^{-T}D_{m+n} + L^{-1}XLJ_{m+n}D_{m+n}\right)\right\|_2 \\ &= \max_{\|\text{vec}(X)\|_2=1}\left\|\text{vec}\left(\text{up}\left(J_{m+n}L^TX^TL^{-T}D_{m+n} + D_{m+n}^{-1}(D_{m+n}L^{-1}XLJ_{m+n})D_{m+n}\right)\right)\right\|_2 \\ &= \max_{\|X\|_F=1}\left\|\left(\text{up}(J_{m+n}L^TX^TL^{-T}D_{m+n}) + D_{m+n}^{-1}\text{up}(D_{m+n}L^{-1}XLJ_{m+n})D_{m+n}\right)\right\|_F \\ &\leqslant \max_{\|X\|_F=1}\sqrt{1+\zeta_{D_{m+n}}^2}\left\|(D_{m+n}L^{-1}XLJ_{m+n})^T\right\|_F \\ &\leqslant \sqrt{1+\zeta_{D_{m+n}}^2}\left\|D_{m+n}L^{-1}\right\|_2\|L\|_2. \end{aligned} \tag{3.18}$$

Thus, plugging (3.18) into (3.17) yields

$$\|G_L\|_2 \leqslant \left(\inf_{D_{m+n}\in\mathbb{D}_{m+n}}\sqrt{1+\zeta_{D_{m+n}}^2}\,\kappa\left(D_{m+n}L^{-1}\right)\right)\|L\|_2, \tag{3.19}$$

which together with the fact $\kappa\left(D_{m+n}L^{-1}\right) \geqslant 1$ shows that the bound (3.13) is indeed tighter than (3.15).

REMARK 3.2. We can obtain the first-order multiplicative perturbation bound from (3.11) as follows:

$$\|\Delta L\|_F \lesssim \|G_L\|_2\,\|Q-I_{m+n}\|_F = \|G_L\|_2\,\|E\|_F. \tag{3.20}$$

Considering (3.19), it is easy to see that the above first-order multiplicative perturbation bound is tighter than the one given in [10, (2.18)].

4. Perturbation bounds for the Cholesky-like factorization

The method of getting the rigorous multiplicative perturbation bounds for the Cholesky-like factorization is very like the one for the generalized Cholesky factorization. The main difference is that we will use the block matrix-vector equation approach. Specifically, let the matrices B and R in (1.3) be perturbed as

$$B \to S^T BS, \;\; R \to R + \Delta R,$$

where $S = I_{2n} + F$ and ΔR is upper triangular such that $R + \Delta R$ has the same structure as that of R. Then the perturbed Cholesky-like factorization of B is

$$\begin{aligned}&\widetilde{B} = S^T KS\\ &(I_{2n}+F)^T K (I_{2n}+F) = (R+\Delta R)^T J_{2n} (R+\Delta R).\end{aligned} \tag{4.1}$$

Expanding (4.1) and using (1.3), we have

$$R^T J_{2n} RF + F^T R^T J_{2n} R + F^T R^T J_{2n} RF = R^T J_{2n} \Delta R + \Delta R^T J_{2n} R + \Delta R^T J_{2n} \Delta R.$$

Premultipling the above equation by R^{-T} and postmultipling it by R^{-1} leads to

$$\begin{aligned}J_{2n}\Delta R R^{-1} + R^{-T}\Delta R^T J_{2n} &= J_{2n}\Delta R R^{-1} + (J_{2n}\Delta R R^{-1})^T\\ &= J_{2n} RFR^{-1} + R^{-T} F^T R^T J_{2n} + R^{-T} F^T R^T J_{2n} RFR^{-1}\\ &\quad - R^{-T}\Delta R^T J_{2n}\Delta R R^{-1}.\end{aligned} \tag{4.2}$$

As done in [11, 14], from (4.2), we have

$$\begin{aligned}J_{2n}\Delta R R^{-1} &= \mathrm{upb}\left(J_{2n} RFR^{-1} + R^{-T} F^T R^T J_{2n}\right)\\ &\quad + \mathrm{upb}\left(R^{-T}\left(F^T R^T J_{2n} RFR^{-1} - R^{-T}\Delta R^T J_{2n}\Delta R\right) R^{-1}\right).\end{aligned} \tag{4.3}$$

Applying the operator 'vecb' to (4.3) and using (2.14) and (2.6) gives

$$\begin{aligned}\left(R^{-T}\boxtimes J_{2n}\right)\mathrm{vecb}\,(\Delta R) &= M_{\mathrm{upb}}\left(\left(R^{-T}\boxtimes (J_{2n}R)\right) + \left((J_{2n}R)\boxtimes R^{-T}\right)\hat{\Pi}_{2n}\right)\mathrm{vecb}\,(F)\\ &\quad + M_{\mathrm{upb}}\left(R^{-T}\boxtimes R^{-T}\right)\mathrm{vecb}\left(F^T R^T J_{2n} RF - \Delta R^T J_{2n}\Delta R\right).\end{aligned}$$

As done in [11], we can obtain

$$\begin{aligned}\mathrm{vecb}\,(\Delta R) &= \left(R^T\boxtimes J_{2n}^{-1}\right) M_{\mathrm{upb}}\left(\left(R^{-T}\boxtimes (J_{2n}R)\right) + \left((J_{2n}R)\boxtimes R^{-T}\right)\hat{\Pi}_{2n}\right)\mathrm{vecb}\,(F)\\ &\quad + \left(R^T\boxtimes J_{2n}^{-1}\right) M_{\mathrm{upb}}\left(R^{-T}\boxtimes R^{-T}\right)\mathrm{vecb}\left(F^T R^T J_{2n} RF - \Delta R^T J_{2n}\Delta R\right).\end{aligned} \tag{4.4}$$

and show that (4.4) is equivalent to

$$\begin{aligned}&\mathrm{uvecb}\,(\Delta R)\\ &= M_{\mathrm{uvecb}}\left(R^T\boxtimes J_{2n}^{-1}\right) M_{\mathrm{upb}}\left(\left(R^{-T}\boxtimes (J_{2n}R)\right) + \left((J_{2n}R)\boxtimes R^{-T}\right)\hat{\Pi}_{2n}\right)\mathrm{vecb}\,(F)\\ &\quad + M_{\mathrm{uvecb}}\left(R^T\boxtimes J_{2n}^{-1}\right) M_{\mathrm{upb}}\left(R^{-T}\boxtimes R^{-T}\right)\mathrm{vecb}\left(F^T R^T J_{2n} RF - \Delta R^T J_{2n}\Delta R\right).\end{aligned} \tag{4.5}$$

As an issue of comfort, let

$$G_R = M_{\text{uvecb}}\left(R^T \boxtimes J_{2n}^{-1}\right) M_{\text{upb}}\left(\left(R^{-T} \boxtimes (J_{2n}R)\right) + \left((J_{2n}R) \boxtimes R^{-T}\right)\hat{\Pi}_{2n}\right),$$
$$H_R = M_{\text{uvecb}}\left(R^T \boxtimes J_{2n}^{-1}\right) M_{\text{upb}}\left(R^{-T} \boxtimes R^{-T}\right)$$

Thus, applying the operator 'uvecb†' to (4.5) leads

$$\Delta R = \text{uvecb}^\dagger\left(G_R \text{vecb}(F) + H_R \text{vecb}\left(F^T R^T J_{2n} R F - \Delta R^T J_{2n} \Delta R\right)\right).$$

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

$$\begin{aligned} \Delta R &= \Phi(\Delta R, F) \\ &= \text{uvecb}^\dagger\left(G_R \text{vecb}(F) + H_R \text{vecb}\left(F^T R^T J_{2n} R F - \Delta R^T J_{2n} \Delta R\right)\right). \end{aligned} \tag{4.6}$$

In the following, we will apply the same technique in Section 3 to study the rigorous perturbation bounds for ΔR based on the operator equation (4.6). For completeness of the method and convenience of readers, we include the detailed process here.

Suppose 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 $\|F\|_F = \delta$. Then it follows from the definition of the operator 'uvecb†' and (2.1) that

$$\|\Phi(Z,F)\|_F \leqslant \|G_R\|_2 \delta + \|R\|_2^2 \|H_R\|_2 \delta^2 + \|H_R\|_2 \rho^2. \tag{4.7}$$

From (4.7), we have the Lyapunov majorant function of the operator equation (4.6)

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

and the Lyapunov majorant equation

$$h(\rho,\delta) = \rho, \ \text{ i.e., } \ \|G_R\|_2 \delta + \|R\|_2^2 \|H_R\|_2 \delta^2 + \|H_R\|_2 \rho^2 = \rho. \tag{4.8}$$

Assume that $\delta \in \Omega = \left\{\delta \geqslant 0 : 1 - 4\|H_R\|_2\left(\|G_R\|_2 \delta + \|R\|_2^2 \|H_R\|_2 \delta^2\right) \geqslant 0\right\}$. Then, the Lyapunov majorant equation (3.7) has two nonnegative roots: $\rho_1(\delta) \leqslant \rho_2(\delta)$ with

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

Let the set $B(\delta)$ be

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

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

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

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

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

when $\delta \in \Omega_1 = \left\{\delta \geqslant 0 : 1-4\|H_R\|_2\left(\|G_R\|_2\delta + \|R\|_2^2\|H_R\|_2\delta^2\right) > 0\right\}$. Then the operator $\Phi(\cdot,F)$ is contractive on the set $B(\delta)$ for $\delta\in\Omega_1$. Thus, from the Banach fixed point theorem, we have that the operator equation (4.6), i.e., the matrix equation (4.1), has a unique solution in the set $B(\delta)$. As a result, $\|\Delta R\|_F \leqslant f(\delta)$ for $\delta\in\Omega_1$. In summary, we have the following main theorem.

THEOREM 4.1. *Suppose that the skew-symmetric matrix $B\in\mathbb{R}^{(2n)\times(2n)}$ is defined by* (1.1) *and has the factorization* (1.2)*. Let $S = I_{2n}+F\in\mathbb{R}^{(2n)\times(2n)}$. If*

$$\|H_R\|_2\left(\|G_R\|_2\|F\|_F + \|R\|_2^2\|H_R\|_2\|F\|_F^2\right) < \frac{1}{4}, \tag{4.9}$$

then $\widetilde{B} = S^TBS$ has the unique Cholesky-like factorization

$$\widetilde{B} = S^TBS = (R+\Delta R)^T J_{2n}(R+\Delta R),$$

and

$$\|\Delta R\|_F \leqslant \frac{2\left(\|G_R\|_2\|F\|_F + \|R\|_2^2\|H_R\|_2\|F\|_F^2\right)}{1+\sqrt{1-4\|H_R\|_2\left(\|G_R\|_2\|F\|_F+\|R\|_2^2\|H_R\|_2\|F\|_F^2\right)}} \tag{4.10}$$

$$\leqslant 2\left(\|G_R\|_2\|F\|_F + \|R\|_2^2\|H_R\|_2\|F\|_F^2\right) \tag{4.11}$$

$$< (\|R\|_2 + 2\|G_R\|_2)\|F\|_F. \tag{4.12}$$

Proof. It is anything but difficult to see that the condition (4.9) is the same as the one in Ω_1. Thus, from the discussions before "Theorem 4.1.", it suffices to obtain the bound (4.12). This can be done by noting (4.9) and the fact,

$$2\|R\|_2\|H_R\|_2\|F\|_F \leqslant \sqrt{1+\frac{\|G_R\|_2^2}{\|F\|_2^2}} - \frac{\|G_R\|_2}{\|R\|_2} < 1. \quad \square \tag{4.13}$$

REMARK 4.1. In [10], by composite of the classic and refined matrix equation approaches, the following rigorous multiplicative perturbation bound was obtained,

$$\frac{\|\Delta R\|_F}{\|R\|_2} \leqslant (\sqrt{3}+\sqrt{6})\inf_{D_{2n}\in\mathbb{D}_{2n}}\sqrt{1+\zeta_{D_{2n}}^2}\,\kappa(D_{2n}^{-1}R)\|S-I_{2n}\|_F. \tag{4.14}$$

under the condition

$$\kappa(R)\|F\|_F < (\sqrt{6}-2)/2. \tag{4.15}$$

In the accompanying, we will demonstrate that the bound (4.12) is sharper than (4.14). Actually, like the evidence of Corollary 3.4 in [15], for any $D_{2n} \in \mathbb{D}_{2n}$ and $X \in \mathbb{R}^{(2n)\times(2n)}$, using (2.17), (2.16), (2.15) and (2.1), we have

$$\|G_R\|_2 = \left\| M_{\text{uvecb}} \left(R^T \boxtimes J_{2n}^{-1}\right) \left(D_{2n}^{-1} \boxtimes I_{2n}\right) \left(D_{2n} \boxtimes I_{2n}\right) M_3 \right\|_2$$

with

$$\begin{aligned} M_3 &= M_{\text{upb}} \left(\left(R^{-T} \boxtimes (J_{2n}R)\right) + \left((J_{2n}R) \boxtimes R^{-T}\right) \hat{\Pi}_{2n} \right) \\ &= \left\| M_{\text{uvecb}} \left((R^T D_{2n}^{-1}) \boxtimes J_{2n}^{-1} \right) M_{\text{upb}} \left(\left((D_{2n}R^{-T}) \boxtimes (J_{2n}R)\right) + \left((D_{2n}J_{2n}R) \boxtimes R^{-T}\right) \hat{\Pi}_{2n} \right) \right\|_2 \\ &\leqslant \left\| R^T D_{2n}^{-1} \right\|_2 \left\| M_{\text{upb}} \left(\left((D_{2n}R^{-T}) \boxtimes (J_{2n}R)\right) + \left((D_{2n}(J_{2n}R)) \boxtimes R^{-T}\right) \hat{\Pi}_{2n} \right) \right\|_2 \\ &= \left\| R^T D_{2n}^{-1} \right\|_2 \max_{\|\text{vecb}(X)\|_2=1} \left\| M_{\text{upb}} \left(\left((D_{2n}R^{-T}) \boxtimes (J_{2n}R)\right) + M_4 \right) \text{vecb}(X) \right\|_2 \end{aligned} \tag{4.16}$$

with

$$M_4 = \left((D_{2n}(J_{2n}R)) \boxtimes R^{-T}\right) \hat{\Pi}_{2n}.$$

Whereas, combining (2.14), (2.6), (2.8), (2.9) and (2.1) gives

$$\begin{aligned} &\max_{\|\text{vecb}(X)\|_2=1} \left\| M_{\text{upb}} \left(\left((D_{2n}R^{-T}) \boxtimes (J_{2n}R)\right) + \left((D_{2n}(J_{2n}R)) \boxtimes R^{-T}\right) \hat{\Pi}_{2n} \right) \text{vecb}(X) \right\|_2 \\ &= \max_{\|\text{vecb}(X)\|_2=1} \left\| M_{\text{upb}} \text{vecb} \left(J_{2n}RXR^{-1}D_{2n} + R^{-T}X^T R^T J_{2n}^T D_{2n} \right) \right\|_2 \\ &= \max_{\|\text{vecb}(X)\|_2=1} \left\| \text{vecb} \left(\text{upb} \left(J_{2n}RXR^{-1}D_{2n} + D_{2n}^{-1} (J_{2n}RXR^{-1}D_{2n})^T D_{2n} \right) \right) \right\|_2 \\ &= \max_{\|X\|_F=1} \left\| \text{upb}(J_{2n}RXR^{-1}D_{2n}) + D_{2n}^{-1} \text{upb}(J_{2n}RXR^{-1}D_{2n})^T D_{2n} \right\|_F \\ &\leqslant \max_{\|X\|_F=1} \sqrt{1+\varsigma_{D_{2n}}^2} \left\| J_{2n}RXR^{-1}D_{2n} \right\|_F \\ &\leqslant \sqrt{1+\varsigma_{D_{2n}}^2} \left\| R^{-1}D_{2n} \right\|_2 \|R\|_2 . \end{aligned} \tag{4.17}$$

Thus, plugging (4.17) into (4.16) yields

$$\|G_R\|_2 \leqslant \left(\inf_{D\in\mathbb{D}_{2n}} \sqrt{1+\varsigma_{D_{2n}}^2} \kappa\left(D_{2n}^{-1}R\right) \right) \|R\|_2, \tag{4.18}$$

which together with the fact $\kappa\left(D_{2n}^{-1}R\right) \geqslant 1$ shows that the bound (4.12) is indeed tighter than (4.14).

REMARK 4.2. The first-order multiplicative perturbation bound from (4.10) can be chalk out as under:

$$\|\Delta R\|_F \lesssim \|G_R\|_2 \|S - I_{2n}\|_F = \|G_R\|_2 \|F\|_F. \tag{4.19}$$

Considering (4.18), it is easy to see that the above first-order multiplicative perturbation bound is tighter than the one given in [10, (3.16)].

5. Numerical examples

Two examples are given in this segment to represent the outcomes inferred in the over two areas. Hereafter, the MATLAB notation will be utilized. For the first example, like the done in [2], we select the accompanying scaling matrix $D_{m+n} = D_c \equiv \operatorname{diag}(\|\mathrm{L}(:,\mathrm{i})\|_2)$, which makes the condition number $\kappa(D_{m+n}L^{-1})$ be nearly minimum. Our numerical experiments additionally propose that another choice for scaling matrix may give a decent estimate, that is, choosing $D_e = \operatorname{diag}(\delta_1, \delta_2, \ldots, \delta_{m+n})$ to approximately equilibrate the rows of $D_r L^{-1}$. To do this, take $\delta_1 = 1/\left\|(D_r L^{-1})(1,:)\right\|_2$, then for $j = 2,3,\ldots,m+n$ take $\delta_j = 1/\left\|(D_r L^{-1})(j,:)\right\|_2$ if

$$\left\|(D_r L^{-1})(j,:)\right\|_2 \geqslant \left\|(D_r L^{-1})(j-1,:)\right\|_2;$$

otherwise $\delta_j = \delta_{j-1}$. Here $D_r = \operatorname{diag}(\|\mathrm{L}(:,\mathrm{i})\|_1)$ and $\|X\|_1$ denotes the 1-norm of X. More on techniques and clarifications of picking the scaling matrices can be found in [2]. For the second example, we choose the scaling matrices D_r and D_e as done in [11]. That is, for D_r, we set its i-th 2×2 main diagonal block to be

$$\frac{\|R(2i-1,:)\|_2 + \|R(2i,:)\|_2}{2} I_2, \; i = 1,\ldots,n,$$

and for D_e, we characterize it as follows:

$$\delta_1 = \frac{2}{\|(D_c R(:,1)\|_2 + \|(D_c R)(:,2)\|_2};$$

for $j = 2,3,\ldots,n$:

$$\delta_j = \frac{2}{\|(D_c R)(:,2j-1)\|_2 + \|(D_c R)(:,2j)\|_2};$$

if

$$\|(D_c R)(:,2j-1)\|_2 + \|(D_c R)(:,2j)\|_2 \geqslant \|(D_c R)(:,2j-3)\|_2 + \|(D_c R)(:,2j-2)\|_2;$$

otherwise, $\delta_j = \delta_{j-1}$. Here $D_c = \operatorname{diag}((\|(D_c R)(:,2j-1)\|_2 + \|(D_c R)(:,2j)\|_2))/2I_2$, $j = 1,\ldots,n,$.

EXAMPLE 5.1. This example has been taken and unraveled by algorithm given in [20]. That is, we take $A_m = [a_{ij}] = H_m + I_m \in \mathbb{R}^{m\times m}$, where $H_m = \frac{1}{i+j-1}$ is an $m \times m$ Hilbert matrix. Also take $B = [b_{ij}] = [\max(i,j)] \in \mathbb{R}^{n\times m}$ and $C_n = [c_{ij}] = U_n \sigma_n U_n \in \mathbb{R}^{n\times n}$, where $U_n = I_n - \frac{2}{w^T w} w w^T$ with $w = (1:n)^T$, and $\sigma_n = \operatorname{diag}(1,2,\ldots,n-1,0)$. Upon computation in MATLAB Ra2016 on PC, with machine precision 2.2×10^{-16}, we have the numerical results for different values of m and n shown in *Table 1*.

In *Table 1*, we denote

$$\gamma = (\|L\|_2 + 2\|G_L\|_2), \; \gamma(D_x) = (\sqrt{3}+\sqrt{6})\sqrt{1+\zeta_{D_{m+n}}^2}\,\kappa(D_{m+n}L^{-1})\|L\|_2, \; x = c \text{ or } e,$$

and t_γ and $t_{\gamma(D_x)}$ the time costing for figuring γ and $\gamma(D_x)$, respectively.

From *Table 1*, it is anything but difficult to see that the bound (3.13), the column denoted by γ, is constantly more sharper than (3.15), the column denoted by $\gamma(D_x)$, regardless of which scaling matrix is chosen. In the interim, we can likewise observe that it is without a doubt more costly to gauge the bound (3.15); think about the column denoted by t_γ and $t_{\gamma(D_x)}$.

Table 1: *Comparison of multiplicative rigorous perturbation bounds* (3.13) *and* (3.15).

m,n	γ	t_γ	$\gamma(D_c)$	$t_{\gamma(D_c)}$	$\gamma(D_e)$	$t_{\gamma(D_e)}$
2,2	15.3506	0.345051	116.8539	0.204960	258.4705	0.313993
4,3	107.4681	0.352381	1.1056e+03	0.209144	4.4447e+03	0.309875
5,3	195.5623	0.405021	2.0708e+03	0.207219	1.0906e+04	0.326690
10,9	4.6491e+03	0.385043	6.1565e+04	0.211413	8.9710e+05	0.323771
11,10	6.8824e+03	0.423454	9.3648e+04	0.209297	1.5460e+06	0.336801
15,13	2.1999e+04	0.711031	3.2805e+05	0.196419	7.8187e+06	0.316718

EXAMPLE 5.2. Let $R \in \mathbb{R}^{(2n)\times(2n)}$ be a Kahan-like matrix:

$$R = \operatorname{diag}(I_2, S, S^2, \ldots, S^{n-1}) \begin{bmatrix} 1 & -c & \cdots & -c \\ & 1 & \cdots & -c \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix}$$

where $S = \operatorname{diag}(s,s)$ with $s = \sin(\theta)$, and $c = \cos(\theta)$. The numerical results for $n = 3,4,5,6,8,10$ with $\theta = \frac{\pi}{8}$ are shown in *Table 2*.

In *Table 2*, we denote

$$\gamma = (\|R\|_2 + 2\|G_R\|_2), \;\; \gamma(D_x) = (\sqrt{3}+\sqrt{6})\sqrt{1+\zeta_{D_{2n}}^2}\,\kappa(D_{2n}^{-1}R)\|R\|_2, \;\; x = r \text{ or } e.$$

Table 2: *Comparison of multiplicative rigorous perturbation bounds* (4.12) *and* (4.14) *for the* $2n\times 2n$ *Kahan-like matrices.*

n	γ	t_γ	$\gamma(D_r)$	$t_{\gamma(D_r)}$	$\gamma(D_e)$	$t_{\gamma(D_e)}$
3	18.6161	0.148354	407.7249	0.120412	501.8577	0.125007
4	30.2781	0.155032	2.1505e+03	0.128087	2.6386e+03	0.129356
5	49.1837	0.211037	1.0582e+04	0.170354	1.2752e+04	0.197575
6	74.0051	0.228509	4.9358e+04	0.176780	5.8332e+04	0.218767
8	159.4854	0.329462	9.6622e+05	0.178909	1.1050e+06	0.289606
10	326.7863	0.351635	1.7318e+07	0.205004	1.9335e+07	0.291463

From *Table 2*, it is easy to see that the bound (4.12), the column marked by γ, is always tighter than (4.14), the columns marked by $\gamma(D_x)$, no matter which scaling matrix is selected. Meanwhile, we can also see that it is indeed more expensive to estimate the bound (4.14); compare the columns marked by t_γ and $t_{\gamma(D_x)}$. Facilitate increasingly these bounds are constantly more keen at that point the those bounds displayed in [11, Table 1] for additive perturbation bounds in light of the fact that here matrix is seriously ill condition so multiplicative perturbation bounds are useful to acquire the more keen bound.

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

REFERENCES

[1] P. BENNER, R. BYERS, H. FASSBENDER, V. MEHRMANN AND D. WATKINS, *Cholesky-like factorizations of skew-symmetric matrices*, Electron. Trans. Numer. Anal. **11**, (2000), 85–93.

[2] X. W. CHANG, *Perturbation analysis of Some Matrix factorization*, PhD thesis, School of Computer Science, McGill University, Montreal, 1997.

[3] X. W. CHANG AND R. C. LI, *Multiplicative perturbation analysis for QR factorizations*, Numer. Algebra Control Optim. **1**, (2011), 301–316.

[4] X. W. CHANG, C. C. PAIGE AND G. W. STEWART, *Perturbation analysis for the QR factorization*, SIAM J. Matrix Anal. Appl. **18**, (1997), 775–791.

[5] H. B. GUO AND W. G. WANG, *Perturbation bounds for the Cholesky-like factorization of skew-symmetric matrix*, Nanjing Univ. J. Math. Biquarterly **26**, (2009), 43–50.

[6] R. A. HORN AND C. R. JOHNSON, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[7] R. H. KONING, H. NEUDECKER AND T. WANSBEEK, *Block Kronecker products and the vecb operator*, Linear Algebra Appl. **149**, (1991), 165–184.

[8] M. KONSTANTINOV, D. GU, V. MEHRMANN AND P. PETKOV, *Perturbation Theory for Matrix Equations*, Elsevier, Amsterdam, 2003.

[9] H. Y. LI, H. YANG, AND H. SHAO, *A note on the perturbation analysis for the generalized Cholesky factorization*, Appl. Math. Comput. **215**, (2010), 4022–4027.

[10] H. Y. LI AND Y. YANG, *Rigorous multiplicative perturbation bounds for the generalized Cholesky factorization and the Cholesky-like factorization*, J. Math. Inequal. **8**, (2014), 925–937.

[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] H. Y. LI AND Y. M. WEI, *Improved rigorous perturbation bounds for the LU and QR factorizations*, Numer. Linear Algebra Appl. **22**, (2015), 1115–1130.

[13] H. Y. LI AND P. LV, *New perturbation bounds and condition numbers for the hyperbolic QR factorization*, Linear Multilinear Algebra **68**, 8 (2017), 1540–1553.

[14] H. Y. LI AND P. LV, *A note on the sesitivity analysis for the symplectic QR factorization*, Math. Inequal. Appl. **20**, 2 (2017), 427–439.

[15] W. LI, Z. J. XIE AND S. W. VONG, *Sensitivity analysis for the symplectic QR factorization*, J. Franklin Inst. **353**, 5 (2016), 1186–1205.

[16] G. W. STEWART AND J. G. SUN, *Matrix Perturbation Theory*, Academic Press, Boston, 1990, **149**, (1991), 165–184.

[17] W. G. WANG AND J. X. ZHAO, *Perturbation analysis for the generalized Cholesky factorization*, Appl. Math. Comput. **147**, (2004), 601–606.

[18] W. G. WANG, *Perturbation bounds for the generalized Cholesky factorization*, Nanjing Univ. J. Math Biquarterly **20**, (2003), 113–119 (in Chinese).

[19] Y. YANG AND H. Y. LI, *Multiplicative perturbation bounds for the SR factorization*, J. Math. Res. Appl. **34**, 4 (2014), 423–434.

[20] J. X. ZHAO, *The generalized Cholesky factorization method for saddle point problems*, Appl. Math. Comput. **92**, (1998), 49–58.

(Received April 26, 2018)

Aamir Farooq
College of Mathematics and Statistics
Chongqing University
Chongqing, 401331, P. R. China
e-mail: aamirf88@yahoo.com

Mahvish Samar
College of Mathematics and Statistics
Chongqing University
Chongqing, 401331, P. R. China
e-mail: mahvishsamar@hotmail.com

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

Chunlai Mu
College of Mathematics and Statistics
Chongqing University
Chongqing, 401331, P. R. China
e-mail: clmu2005@163.com