UDC 519.62

Chi-Ye Wu (Jinan Univ. and School Math. Sci., Univ. Electron. Sci. and Technology, China),
Ting-Zhu Huang (School Math. Sci., Univ. Electron. Sci. and Technology, China)

PERTURBATION AND ERROR ANALYSES OF PARTITIONED LU FACTORIZATION FOR BLOCK TRIDIAGONAL LINEAR SYSTEMS*

АНАЛІЗ ЗБУРЕНЬ ТА ПОХИБОК РОЗБИТОЇ НА ЧАСТИНИ LU ФАКТОРИЗАЦІЇ ДЛЯ БЛОЧНО-ТРИДІАГОНАЛЬНИХ ЛІНІЙНИХ СИСТЕМ

The perturbation and backward error analyses of the partitioned LU factorization for block tridiagonal matrices are presented. Moreover, we consider the perturbation bounds for the partitioned LU factorization for block tridiagonal linear systems. Finally, numerical examples are given to verify our results.

Наведено аналіз збурень та зворотний аналіз похибок для розбитої на частини LU факторизації блочно-тридіагональних матриць. Крім того, вивчаються границі збурень для розбитої на частини LU факторизації блочно-тридіагональних лінійних систем. Також наведено числові експерименти, які підтверджують справедливість даних результатів.

1. Introduction. We consider the linear system $Ax = b$ when A is a nonsingular block tridiagonal matrix as follows:

$$A = \begin{pmatrix} A_1 & C_1 & & & \\ B_2 & A_2 & C_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & C_{s-1} \\ & & & B_s & A_s \end{pmatrix}, \tag{1.1}$$

where $A_i \in \mathbb{R}^{k_i \times k_i}$, $B_i \in \mathbb{R}^{k_i \times k_{i-1}}$, and $C_i \in \mathbb{R}^{k_i \times k_{i+1}}$ for all $1 \le i \le s$.

For a nonsingular block tridiagonal matrix as above, our interest is to solve the linear system $Ax = b$ efficiently and accurately. Applying partitioned LU factorization for a general matrix, the representation of partitioned LU factorization for nonsingular block tridiagonal matrices is as follows:

$$A = \begin{pmatrix} L_{11} & & & \\ B_2 U_{11}^{-1} & I_2 & & \\ & & \ddots & \\ & & & I_s \end{pmatrix} \begin{pmatrix} I_1 & \\ & S_1 \end{pmatrix} \begin{pmatrix} U_{11} & L_{11}^{-1} C_1 & & \\ & I_2 & & \\ & & \ddots & \\ & & & I_s \end{pmatrix} = L_1 D_1 U_1,$$

where

$$S_1 = \begin{pmatrix} A_2 - B_2 U_{11}^{-1} L_{11}^{-1} C_1 & C_2 & & & \\ B_3 & A_3 & C_3 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & C_{s-1} \\ & & & B_s & A_s \end{pmatrix}.$$

* This research was supported by NSFC (No. 11301223), GDNSF (No. S2012040007333), financial engineering and algorithm research center of Jinan University, high level university construction and overall planning project of Jinan University (Shenzhen Tourism College) (No. 88015201) and teaching reform research project of Jinan University (No. 55610196).

If $A_2 - B_2U_{11}^{-1}L_{11}^{-1}$ can be factorized as follows:

$$A_2 - B_2U_{11}^{-1}L_{11}^{-1} = L_{22}U_{22},$$

then D_1 satisfies

$$D_1 = L_2D_2U_2 = \begin{pmatrix} I_1 & & & & \\ & L_{22} & & & \\ & B_3U_{22}^{-1} & I_3 & & \\ & & & \ddots & \\ & & & & I_s \end{pmatrix} \times$$

$$\times \begin{pmatrix} I_1 & & \\ & I_2 & \\ & & S_2 \end{pmatrix} \begin{pmatrix} I_1 & & & & \\ & U_{22} & L_{11}^{-1}C_2 & & \\ & & I_3 & & \\ & & & \ddots & \\ & & & & I_s \end{pmatrix},$$

where the form of S_2 is similar to that of S_1 so that we here ignore it. For a given i, if the first block of S_i can be factorized, then the partitioned LU factorization can run to the $(i+1)$st step. Otherwise, the factorization must break down in the ith step. Suppose that the factorization can run to completion. Then we have

$$A = L_1 \ldots L_{s-1}L_sU_sU_{s-1}\ldots U_1,$$

where

$$D_{s-1} = L_sU_s.$$

Note that there are different form and content between the partitioned LU factorization and the general block LU factorization, because every step in process the former is more one LU factorization than the latter, and the factors both L_i and U_i of the former are triangular forms which are not satisfied for the latter.

In the literature there are lots of papers dealing with the perturbation bounds for usual, or pointwise, LU, Cholesky or QR factorizations. References relevant to this problem include Barrlund [1], Stewart [2–4], Chang and Paige [5], and Dopico and Molera [6], etc. First-order perturbation bounds are frequently used, such as Chang, Paige and Stewart [7], and Stewart [2, 3]. Dopico and Molera [6] presented expressions for the terms of any order in the series expressions of the perturbed LU and Cholesky factors. When the above factorization for the original matrix A in (1.1) runs to completion, the question is whether the perturbed matrix $A+E$ exists the partitioned LU factorization. If E satisfies

$$|E| \le \epsilon|A|,$$

where ϵ is small sufficiently and $|A|$ stands for a matrix of absolute values of entries of A, then the partitioned LU factorization for the perturbed matrix $A+E$ exists. The relations between $S_{ij}^{(k)}(A+E)$, $D_{ij}^{(k)}(A+E)$ and $S_{ij}^{(k)}(A)$, $D_{ij}^{(k)}(A)$, respectively, are considered, where $S_{ij}^{(k)}(A)$ and

$D_{ij}^{(k)}(A)$ stand for block (i,j) of S_k and D_k, respectively. Moreover, perturbation bounds for the factors are also proposed.

The error analysis is one of the most powerful tools for studying the accuracy and stability of numerical algorithms. References relevant to this problem include Higham [8–10], Amodio and Mazzia [11], Demmel, Higham, and Shreiber [12], Zhao, Wang, and Ren [13], Mattheij [14], Forsgren, Gill, and Shinnerl [15], and Bueno and Dopico [16], etc. In this paper, applying the special property that the factors L_i and U_i are triangular forms, then some assumptions on the BLAS3 that can not be applied in the error analysis of general block LU factorization can be used in that of the partitioned LU factorization. Hence, error analysis of the partitioned LU factorization for block tridiagonal linear systems can be considered. Comparing the results of Higham [8], Demmel and Higham [17] with those of this paper, the distinction between the former and the latter is conspicuous. Based on the assumptions, the latter conditions are weaker than those of the former. Finally, two numerical examples are considered to illustrate our theory results, where the mentioned matrices generated from the discretization of partial differential equation $-\Delta u = f$ and random block tridiagonal matrices by MATLAB 6.5, respectively.

2. Perturbation theory. In this section our interest is to present perturbation analysis of the factors of the partitioned LU factorization.

2.1. Some properties. We first consider the relation between the first block of $S_k(A+E)$ and that of $S_k(A)$.

Theorem 2.1. *Let the partitioned LU factorization for the block tridiagonal matrix A in* (1.1) *run to completion. Assume that ϵ is small enough that $|E| \le \epsilon|A|$. Then*

$$S_{11}^{(k)}(A+E) = S_{11}^{(k)}(A) + T_k + O(\epsilon^2),$$

where T_k, $1 \le k \le s$, satisfy

$$T_1 = E_{11}, \qquad T_k = \begin{pmatrix} B_k \left(S_{11}^{(k-1)}\right)^{-1} & I_k \end{pmatrix} \begin{pmatrix} -T_{k-1} & -E_{k-1,k} \\ -E_{k,k-1} & E_{k,k} \end{pmatrix} \begin{pmatrix} \left(S_{11}^{(k-1)}\right)^{-1} C_k \\ I_k \end{pmatrix}.$$

Proof. To save clutter we will omit "$+O(\epsilon^2)$". The proof is essentially inductive. For $k = 1$, we have

$$S_{11}^{(1)}(A+E) = (A_2 + E_{22}) - (B_2 + E_{21}) U_{11}^{-1}(A+E) L_{11}^{-1}(A+E)(C_1 + E_{12}).$$

Since $A^{-1} = U^{-1}L^{-1}$ and $|E| \le \epsilon|A|$. Then

$$S_{11}^{(1)}(A+E) = (A_2 + E_{22}) - (B_2 + E_{21})(A_1^{-1} + A_1^{-1}E_{11}A_1^{-1})(C_1 + E_{12}) =$$
$$= A_2 - B_2A_1^{-1}C_1 + E_{22} - E_{21}A_1^{-1}C_1 - B_2A_1^{-1}E_{11}A_1^{-1}C_1 - B_2A_1^{-1}E_{12} =$$
$$= S_{11}^{(1)}(A) + \begin{pmatrix} B_2A_1^{-1} & I_2 \end{pmatrix} \begin{pmatrix} -E_{11} & -E_{12} \\ -E_{21} & E_{22} \end{pmatrix} \begin{pmatrix} A_1^{-1}C_1 \\ I_2 \end{pmatrix}.$$

For $k = i-1$, by the assumption, it follows that

$$S_{11}^{(i-1)}(A+E) = S_{11}^{(i-1)}(A) + T_i,$$

where, from the structure of T_i, we obtain $T_i = O(\epsilon)$. For $k = i$, we get

$$S_{11}^{(i)}(A+E) = (A_{i+1} + E_{i+1,i+1}) - (B_{i+1} + E_{i+1,i}) \left(S_{11}^{(i-1)} + T_i\right)^{-1} (C_i + E_{i,i+1}) =$$

$$= A_{i+1} - B_{i+1} U_{ii}^{-1} C_i + E_{i+1,i+1} - E_{i+1,i} \left(S_{11}^{(i-1)}\right)^{-1} C_i - B_{i+1} \left(S_{11}^{(i-1)}\right)^{-1} T_i \left(S_{11}^{(i-1)}\right)^{-1} C_i =$$

$$= S_{11}^{(i)}(A) + \begin{pmatrix} B_{i+1} \left(S_{11}^{(i-1)}\right)^{-1} & I_{i+1} \end{pmatrix} \begin{pmatrix} -T_i & -E_{i,i+1} \\ -E_{i+1,i} & E_{i+1,i+1} \end{pmatrix} \begin{pmatrix} \left(S_{11}^{(i-1)}\right)^{-1} C_i \\ I_{i+1} \end{pmatrix}.$$

Theorem 2.1 is proved.

From the result as above, we have the following theorem.

Theorem 2.2. *Let the partitioned LU factorization for the block tridiagonal matrix A in* (1.1) *run to completion. Assume that ϵ is small enough that $|E| \le \epsilon |A|$. Then*

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij},$$

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \beta_i(\alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}),$$

where

$$\beta_i = \begin{cases} 1, & k \le i \le s-1, \\ 0, & 1 \le i < k, \end{cases} \qquad \alpha_{ij} = \begin{cases} 1, & i = j = 1, \\ 0, & \textit{others}. \end{cases}$$

Proof. By the partitioned LU factorization, we have

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + E_{ij}, \qquad i, j \ne 1. \tag{2.1}$$

Combining (2.1) with Theorem 2.1 gives

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}, \tag{2.2}$$

where

$$\alpha_{ij} = \begin{cases} 1, & i = j = 1, \\ 0, & \text{others}. \end{cases}$$

By the form of D_k, it follows that

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A), \quad 1 \le i < k. \tag{2.3}$$

From (2.2), we have

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}, \quad k \le i \le s-1. \tag{2.4}$$

For (2.3) and (2.4), it follows that

$$D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + \beta_i \left(\alpha_{ij}(T_k + O(\epsilon^2)) + (1-\alpha_{ij})E_{ij}\right),$$

where

$$\beta_i = \begin{cases} 1, & k \le i \le s-1, \\ 0, & 1 \le i < k. \end{cases}$$

Theorem 2.2 is proved.

***Corollary* 2.1.** *Let the partitioned LU factorization for the block tridiagonal matrix A in* (1.1) *run to completion. Assume that ϵ is small enough that $|E| \leq \epsilon|A|$. Then*

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + O(\epsilon), \qquad D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + O(\epsilon).$$

Proof. From the proof of Theorem 2.1 and the form of T_k, it follows that $T_k = O(\epsilon)$, then $T_k + E_{ij} + O(\epsilon^2) = O(\epsilon)$. Therefore

$$S_{ij}^{(k)}(A+E) = S_{ij}^{(k)}(A) + O(\epsilon), \qquad D_{ij}^{(k)}(A+E) = D_{ij}^{(k)}(A) + O(\epsilon).$$

From $|E| \leq \epsilon|A|$, if ϵ is sufficiently small, then the spectral radius $\rho(L^{-1}EU^{-1}) < 1$ holds. Therefore it has a unique block LU factorization (see Theorem 12.1 in [8] for details). In this case the question is whether the matrices $S_{11}^{(k)}$, $1 \leq k \leq s-1$, exist the LU factorization, that is, whether the perturbed matrix $A+E$ exists the partitioned LU factorization. By Theorem 2.1, it follows that

$$S_{11}^{(k)}(A+E) = S_{11}^{(k)}(A) + T_k + O(\epsilon^2).$$

For the assumption of Theorem 2.1, we have

$$S_{11}^{(k)}(A+E) = L_{k+1,k+1}U_{k+1,k+1} + T_k + O(\epsilon^2) =$$
$$= L_{k+1,k+1}\left(I_{k+1} + L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}\right)U_{k+1,k+1}.$$

Since

$$T_k + O(\epsilon^2) = O(\epsilon), \qquad \rho\left(L^{-1}EU^{-1}\right) < 1.$$

Then

$$\rho\left(L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}\right) < 1,$$

that is, $I_{k+1} + L_{k+1,k+1}^{-1}(T_k + O(\epsilon^2))U_{k+1,k+1}^{-1}$ exists the LU factorization. Thus $S_{11}^{(k)}(A+E)$ $(1 \leq k \leq s-1)$ have the LU factorization. Hence the perturbed matrix $A+E$ exists the partitioned LU factorization. Based on the above mentioned, we have the following theorem.

Theorem 2.3. *Let the partitioned LU factorization for the block tridiagonal matrix A in* (1.1) *run to completion. Assume that ϵ is small enough that $|E| \leq \epsilon|A|$. Then the perturbed matrix $A+E$ has the partitioned LU factorization.*

2.2. Perturbation bounds for the factors. In this subsection we present the bounds for the factors. First of all, we consider the bound for $S_{ij}^{(k)}$. Obviously, we are easy to get the following componentwise perturbation bound for $S_{11}^{(1)}$ by applying Theorem 2.1:

$$\left|S_{11}^{(1)}(A+E) - S_{11}^{(1)}(A)\right| \leq \epsilon|A_1|.$$

Unless otherwise stated, let in this section an unsubscripted norm $\|.\|$ be an arbitrary subordinate and monotone matrix norm. For $S_{11}^{(k)}$, we have the following theorem.

Theorem 2.4. *Let the partitioned LU factorization for the block tridiagonal matrix A in* (1.1) *run to completion, and let*

$$\chi = \max_k\left\{\left\|\left(S_{11}^{(k-1)}\right)^{-1}C_k\right\|\|A_{k,k-1}\| + \left\|B_k\left(S_{11}^{(k-1)}\right)^{-1}\right\|\|A_{k-1,k}\| + \|A_{kk}\|\right\},$$

$$\omega = \max_k \left\{ \left\| B_k \left(S_{11}^{(k-1)}\right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)}\right)^{-1} C_k \right\| \right\}$$

with

$$\left\| B_k \left(S_{11}^{(k-1)}\right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)}\right)^{-1} C_k \right\| \neq 1.$$

Assume that ϵ is small enough that $|E| \leq \epsilon |A|$. Then

$$\left\| S_{11}^{(k)}(A+E) - S_{11}^{(k)}(A) \right\| \leq \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon + O(\epsilon^2).$$

Proof. We first consider the bound for T_k. From Theorem 2.1 it follows that

$$T_k = \begin{pmatrix} B_k \left(S_{11}^{(k-1)}\right)^{-1} & I_k \end{pmatrix} \begin{pmatrix} -T_{k-1} & -E_{k-1,k} \\ -E_{k,k-1} & E_{k,k} \end{pmatrix} \begin{pmatrix} \left(S_{11}^{(k-1)}\right)^{-1} C_k \\ I_k \end{pmatrix} =$$

$$= -B_k \left(S_{11}^{(k-1)}\right)^{-1} T_{k-1} \left(S_{11}^{(k-1)}\right)^{-1} C_k - E_{k,k-1} \left(S_{11}^{(k-1)}\right)^{-1} C_k -$$

$$- B_k \left(S_{11}^{(k-1)}\right)^{-1} E_{k-1,k} + E_{k,k}.$$

Taking the monotone norm on both sides yields

$$\|T_k\| \leq \left\| B_k \left(S_{11}^{(k-1)}\right)^{-1} \right\| \left\| \left(S_{11}^{(k-1)}\right)^{-1} C_k \right\| \|T_{k-1}\| + \epsilon \left\| \left(S_{11}^{(k-1)}\right)^{-1} C_k \right\| \|A_{k,k-1}\| +$$

$$+ \epsilon \left\| B_k \left(S_{11}^{(k-1)}\right)^{-1} \right\| \|A_{k-1,k}\| + \epsilon \|A_{kk}\|.$$

Rearranging the above inequality, we have

$$\|T_k\| + \frac{\chi\epsilon}{\omega-1} \leq \omega \left(\|T_{k-1}\| + \frac{\chi\epsilon}{\omega-1} \right) \leq$$

$$\leq \omega^{k-1} \left(\|T_1\| + \frac{\chi\epsilon}{\omega-1} \right) \leq$$

$$\leq \omega^{k-1} \left(\|A_1\| + \frac{\chi}{\omega-1} \right) \epsilon.$$

Then

$$\|T_k\| \leq \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon.$$

Theorem 2.4 is proved.

From Theorems 2.2 and 2.4, it is easy to propose the perturbation bounds for $S_{ij}^k(A)$ and $D_{ij}^k(A)$, i.e.,

$$\left\| S_{ij}^{(k)}(A+E) - S_{ij}^{(k)}(A) \right\| \leq \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon + 2\|A_{ij}\|\epsilon + O(\epsilon^2),$$

$$\left\| D_{ij}^{(k)}(A+E) - D_{ij}^{(k)}(A) \right\| \leq \omega^{k-1} \left(\|A_1\| + \frac{\chi(\omega^{k-1}-1)}{\omega^{k-1}(\omega-1)} \right) \epsilon + 2\|A_{ij}\|\epsilon + O(\epsilon^2).$$

3. Error analysis. Throughout, we use the conventional error model of floating-point arithmetic in which the evaluation of an expression in floating-point arithmetic is denoted by $fl(\cdot)$, with

$$fl(a\ o\ b) = (a\ o\ b)(1+\delta), \qquad |\delta| \le u, \quad o = +, -, *, /$$

(see, for example, [8]). Here u is the unit roundoff of the machine being employed.

Unless otherwise mentioned, in this section an unsubscripted norm denotes

$$\|A\| := \max_{i,j}|a_{ij}|.$$

Note that for this norm, the best inequality is

$$\|AB\| \le n\|A\|\|B\|,$$

where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$. It is well known that this norm is not consistent, but for sparse matrices it is simple and proper choice.

Based on fast matrix multiplication techniques, the use of BLAS3 affects the stability only insofar as it increases the constant terms in the normwise backward error bounds (see [17] for details). We have the following assumptions about the underlying level-3 BLAS.

(a) The computed approximation $\hat{C}$ to $C = AB$, where $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times p}$, satisfies

$$\hat{C} = AB + \Delta C, \qquad \|\Delta C\| \le c_1(m,n,p)u\|A\|\|B\| + O(u^2),$$

where $c_1(m,n,p)$ denotes a constant depending on m, n, and p.

(b) If $T \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{m\times p}$, then computed solution $\hat{X}$ to the triangular systems $TX = B$ satisfies

$$T\hat{X} = B + \Delta B, \qquad \|\Delta B\| \le c_2(m,p)u\|T\|\|\hat{X}\| + O(u^2),$$

where $c_2(m,p)$ denotes a constant depending on m and p.

Assumption (b) on the BLAS3 can not be applied in the error analysis of general block LU factorization because the factor U is not triangular form. In view of this consideration, the partitioned LU for block tridiagonal matrices is presented because the factors L and U are triangular form, i.e., errors incurred at the process of the partitioned LU factorization and substitution can be presented by applying assumption (b). We first recall error analyses of the partitioned LU factorization for a general partitioned matrix $A \in \mathbb{R}^{n\times n}$ and of the corresponding computed solution to $Ax = b$.

Lemma 3.1 [17]. *Under assumptions* (a), (b) *and*

$$\hat{L}_{11}\hat{U}_{11} = A_1 + \Delta A_1, \qquad \|\Delta A_1\| \le c_3(k_1)u\|\hat{L}_{11}\|\|\hat{U}_{11}\| + O(u^2),$$

the LU factors of $A \in \mathbb{R}^{n\times n}$ computed using the partitioned outer product form of LU factorization with block size k_1 satisfy $\hat{L}\hat{U} = A + \Delta A$, where

$$\|\Delta A\| \le u\left(\delta(n,k_1)\|A\| + \theta(n,k_1)\|\hat{L}\|\|\hat{U}\|\right) + O(u^2),$$

and where

$$\delta(n,k_1) = 1 + \delta(n-k_1,k_1), \quad \delta(k_1,k_1) = 0,$$

$$\theta(n,k_1) = \max\{c_3(k_1),\ c_2(k_1,n-k_1),\ 1 + c_1(n-k_1,k_1,n-k_1)+$$
$$+\delta(n-k_1,k_1) + \theta(n-k_1,k_1)\}, \qquad \theta(k_1,k_1) = 0.$$

The following lemma is Problem 12.6 in [8].

Lemma 3.2 [8]. *Under the conditions of Lemma* 3.1 *the computed solution to* $Ax = b$ *satisfies*

$$(A + \delta A)\hat{x} = b, \qquad \|\delta A\| \leq c_n u(\|A\| + \|\hat{L}\|\|\hat{U}\|) + O(u^2),$$

where c_n *is a constant depending on* n *and the block size.*

The corresponding error analyses of block tridiagonal matrix A in (1.1) and its linear systems are presented as follows.

Theorem 3.1. *Let the partitioned LU factorization for the block tridiagonal matrix* A *in* (1.1) *run to completion. Then, under assumptions* (a) *and* (b),

$$A + \Delta A = \hat{L}\hat{U}, \qquad \|\Delta A\| \leq (\xi_{ij}\|A\| + \zeta_{ij}\|L\|\|U\|)u + O(u^2),$$

where

$$\xi_{ij} = \begin{cases} 0, & i = j = 1, \\ 1, & i = j \neq 1, \\ c_2(k_i, k_i)k_i\kappa(L_{ii}), & i = j - 1, \\ c_2(k_i, k_i)k_i\kappa(U_{ii}), & i = j + 1, \end{cases} \qquad \zeta_{ij} = \begin{cases} c_2(k_1, k_1), & i = j = 1, \\ c_i, & i = j \neq 1, \\ 0, & \textit{others}, \end{cases}$$

$$c_i = \max\left\{1 + c_1(k_{i-1}, k_{i-1}, k_i), c_2(k_i, k_i)\right\}.$$

Proof. To save clutter we will omit "$+O(u^2)$". By the assumption (b), we have

$$\hat{\hat{U}}_{ii}^{-1}\hat{U}_{ii} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i, k_i)u\|\hat{\hat{U}}_{ii}^{-1}\|\|\hat{U}_{ii}\|, \tag{3.1}$$

where $\hat{\hat{U}}_{ii}^{-1}$ is the computed quantities for inverting $\hat{U}_{ii}$. Thus,

$$\hat{\hat{U}}_{ii}^{-1} = (I_i + \Delta I_i)\hat{U}_{ii}^{-1}.$$

Then

$$\|\hat{\hat{U}}_{ii}^{-1}\| \leq \|\hat{U}_{ii}^{-1}\| + O(u).$$

By representation (3.1), it follows that

$$\|\Delta I_i\| \leq c_2(k_i, k_i)u\|\hat{U}_{ii}^{-1}\|\|\hat{U}_{ii}\|,$$

Similarly, we get

$$\hat{U}_{ii}^{-1}\hat{U}_{ii} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i, k_i)u\|U_{ii}^{-1}\|\|U_{ii}\| = c_2(k_i, k_i)\kappa(U_{ii})u. \tag{3.2}$$

A similar assertion holds for the following case:

$$\hat{L}_{ii}\hat{\hat{L}}_{ii}^{-1} = I_i + \Delta I_i, \qquad \|\Delta I_i\| \leq c_2(k_i, k_i)u\|L_{ii}\|\|L_{ii}^{-1}\| = c_2(k_i, k_i)\kappa(L_{ii})u. \tag{3.3}$$

By the process of partitioned factorization gives

$$B_2\hat{\hat{U}}_{11}^{-1}\hat{U}_{11} = B_2 + \Delta B_2,$$
$$\hat{L}_{11}\hat{\hat{L}}_{11}^{-1}C_1 = C_1 + \Delta C_1. \tag{3.4}$$

From representations (2.2) – (2.4), we obtain

$$\|\Delta B_2\| \le c_2(k_1,k_1)k_1\kappa(U_{11})\|B_{22}\|u,$$
$$\|\Delta C_1\| \le c_2(k_1,k_1)k_1\kappa(L_{11})\|C_1\|u.$$

By assumption (b), we can get the bound for ΔA_i as follows:

$$\|\Delta A_1\| \le c_2(k_1,k_1)u\|L_{11}\|\|U_{11}\|.$$

For ΔB_3, ΔA_2 and ΔC_2, because of some errors incurred at the process of multiplication and subtraction of the matrices and the LU factorization for $A_2 - B\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}^{-1}C_1$, they are different from ΔB_2, ΔA_1 and ΔC_1, respectively. Let

$$L_{21}U_{12} = B_2U_{11}^{-1}L_{11}^{-1}C_1 = H.$$

Then the computed approximate $\hat{H}$ satisfies

$$B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1 + \Delta H = \hat{H}, \qquad \|\Delta H\| \le c_1(k_1,k_1,k_2)u\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\|.$$

Let

$$A_2 - H = G.$$

Then the computed approximate $\hat{G}$ to G satisfies

$$A_2 - \hat{H} + \Delta G' = \hat{G}, \qquad \|\Delta G'\| \le u(\|A_2\| + \|\hat{H}\|).$$

That is,

$$A_2 - B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1 + \Delta G = \hat{G},$$
$$\|\Delta G\| \le u(\|A_2\| + (1 + c_1(k_1,k_1,k_2))\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\|). \tag{3.5}$$

Applying the LU factorization for $\hat{G}$ gives

$$\hat{G} + \Delta G'' = \hat{L}_{22}\hat{U}_{22}, \qquad \|\Delta G''\| \le c_2(k_2,k_2)u\|L_{22}\|\|U_{22}\|.$$

Combining (3.1) with (3.2), it follows that

$$A_2 + \Delta A_2 = \hat{L}_{22}\hat{U}_{22} + B_2\hat{\hat{U}}_{11}^{-1}\hat{\hat{L}}_{11}^{-1}C_1,$$
$$\|\Delta A_2\| \le u(\|A_2\| + (1 + c_1(k_1,k_1,k_2))\|B_2U_{11}^{-1}\|\|L_{11}^{-1}C_1\| + c_2(k_2,k_2)\|L_{22}\|\|U_{22}\|).$$

From the factors $\hat{L}_{11}$, $\hat{U}_{11}$, $\hat{L}_{22}$ and $\hat{U}_{22}$ obtained in the process of the factorization, it is conspicuous that $\hat{L}_{11}$ and $\hat{U}_{11}$ are different from $\hat{L}_{22}$ and $\hat{U}_{22}$, respectively. For the latter contain the errors incurred at multiplication and subtraction beside the course of factorization. For ΔA_i, $3 \le i \le s$, the similar results hold

$$\|\Delta A_i\| \le u(\|A_i\| + (1 + c_1(k_{i-1}, k_{i-1}, k_i))\|B_i U_{i-1,i-1}^{-1}\| \times$$
$$\times \|L_{i-1,i-1}^{-1} C_{i-1}\| + c_2(k_i, k_i)\|L_{ii}\|\|U_{ii}\|) \le$$
$$\le u(\|A_i\| + c(\|L_{i,i-1}\|\|U_{i-1,i-1}\| + \|L_{ii}\|\|U_{ii}\|)),$$

where $c_i = \max\{1 + c_1(k_{i-1}, k_{i-1}, k_i), c_2(k_i, k_i)\}$. For ΔB_{i+1} and ΔC_i for all $2 \le i \le s-1$, we have

$$\|\Delta B_{i+1}\| \le c_2(k_i, k_i) k_i \kappa(U_{ii})\|B_{i+1,i+1}\|u,$$
$$\|\Delta C_i\| \le c_2(k_i, k_i) k_i \kappa(L_{ii})\|C_i\|u.$$

Therefore

$$\|\Delta A\| \le (\xi_{ij}\|A\| + \zeta_{ij}\|L\|\|U\|)u,$$

where

$$\xi_{ij} = \begin{cases} 0, & i = j = 1, \\ 1, & i = j \ne 1, \\ c_2(k_i, k_i) k_i \kappa(L_{ii}), & i = j - 1, \\ c_2(k_i, k_i) k_i \kappa(U_{ii}), & i = j + 1, \end{cases} \qquad \zeta_{ij} = \begin{cases} c_2(k_1, k_1), & i = j = 1, \\ c, & i = j \ne 1, \\ 0, & \text{others.} \end{cases}$$

Theorem 3.1 is proved.

Remark **3.1.** Comparing Lemma 3.1 with Theorem 3.1, we know the following comments.

1) The assumption of Lemma 3.1

$$\hat{L}_{11}\hat{U}_{11} = A_1 + \Delta A_1, \qquad \|\Delta A_1\| \le c_3(k_1)u\|\hat{L}_{11}\|\|\hat{U}_{11}\| + O(u^2),$$

is omitted in Theorem 3.1. Based on this point, the assumptions of the latter are weaker than those of the former.

2) It is conspicuous that the proof of the latter is different from that of the former.

3) In the result of the former, the computed approximate $\hat{L}$ and $\hat{U}$ are applied, however, the exact quantities L and U are used in that of the latter.

From Theorem 3.1, we have the following assertion on block tridiagonal linear systems. Note that ΔL and ΔU are produced in solving $Ly = b$ and $Ux = y$, respectively.

Theorem 3.2. *Let A be as in* (1.1), *and suppose that the partitioned LU factorization computes an approximate solution $\hat{x}$ to $Ax = b$, where $\hat{x}$ is the exact solution of the system $(A + \delta A)\hat{x} = b$. Then*

$$\|\delta A\| \le (\xi_{ij}\|A\| + \delta_{ij}\|L\|\|U\|)u + O(u^2),$$
$$\frac{\|\hat{x} - x\|}{\|\hat{x}\|} \le n\left(\left(\xi_{ij}\kappa(A) + \frac{\gamma_{ns}}{u}\kappa(U)\right) + \left(\zeta_{ij} + \frac{n\gamma_{ns}}{u}\right)\|L\|\|U\|\|A^{-1}\|\right)u + O(u^2),$$

where $\delta_{ij} = \zeta_{ij} + \gamma_{2ns}/u$.

Proof. By the assumption, it follows that

$$(\hat{L}+\Delta L)(\hat{U}+\Delta U)\hat{x}=b.$$

Then

$$\delta A=\Delta A+\Delta L\hat{U}+\hat{L}\Delta U+\Delta L\Delta U. \tag{3.6}$$

In the following proof we need the bounds for ΔL and ΔU. Applying the factorization and the result of [18], it follows that

$$(\hat{U}_1+\Delta U_1)x=y^{(1)},\qquad |\Delta U_1|\leq\frac{nu}{1-nu}|\hat{U}_1|.$$

For a given i, we have

$$(\hat{U}_i+\Delta U_i)y^{(i-1)}=y^{(i)},\qquad |\Delta U_i|\leq\frac{nu}{1-nu}|\hat{U}_i|.$$

Thus,

$$(\hat{U}_s+\Delta U_s)\ldots(\hat{U}_1+\Delta U_1)x=y,$$
$$|\Delta U|\leq\frac{nsu}{1-nu}|\hat{U}_s|\ldots|\hat{U}_1|\leq\gamma_{ns}|\hat{U}|,$$

where $\gamma_{ns}=nsu/(1-nsu)$. By the definition of norm, we get

$$\|\Delta U\|\leq\gamma_{ns}\|\hat{U}\|. \tag{3.7}$$

On the other hand, we obtain

$$(\hat{L}_1+\Delta L_1)\ldots(\hat{L}_s+\Delta L_s)y=b,\qquad \|\Delta L\|\leq\gamma_{ns}\|\hat{L}\|. \tag{3.8}$$

Combining (3.6), (3.7) with (3.8), by Theorem 3.1, it follows that

$$\|\delta A\|\leq(\xi_{ij}\|A\|+\zeta_{ij}\|L\|\|U\|)u+(2\gamma_{ns}+\gamma_{ns}^2)n\|\hat{L}\|\|\hat{U}\|\leq$$
$$\leq(\xi_{ij}\|A\|+\delta_{ij}\|L\|\|U\|)u,$$

where $\delta_{ij}=\zeta_{ij}+\gamma_{2ns}n/u$ and $2\gamma_{ns}+\gamma_{ns}^2\leq\gamma_{2ns}$ [8, 9]. The following proof refers to the relative error. By Higham [10], we have

$$\|\hat{x}-x\|\leq\|A^{-1}(\Delta A+\Delta L\hat{U})+\hat{U}^{-1}\Delta U\|\|\hat{x}\|. \tag{3.9}$$

Applying Theorem 3.1, from (3.7) and (3.8), it follows that

$$\frac{\|\hat{x}-x\|}{\|\hat{x}\|}\leq n\left(\left(\xi_{ij}\kappa(A)+\frac{\gamma_{ns}}{u}\kappa(U)\right)+\left(\zeta_{ij}+\frac{n\gamma_{ns}}{u}\right)\|L\|\|U\|\|A^{-1}\|\right)u.$$

Theorem 3.2 is proved.

Actually, for $k_i=1,\ 1\leq i\leq s$, there exists a relationship between $\kappa(U)$ and $\kappa(A)$ and $\|L\|\leq 1$ holds when the partial pivoting strategy is applied during the factorization, then the relative error mentioned above can be $O(\kappa(A)u)$. On the other hand, the triangular form of the factors L_i and U_i of the partitioned LU factorization in this paper advantages the relative error $\|\hat{x}-x\|/\|\hat{x}\|$.

***Remark* 3.2.** Comparing Lemma 3.2 with Theorem 3.2, we have the following comments besides the first comment in Remark 3.1.

1. The coefficient c_n of Lemma 3.2 is a faint constant, however, those of Theorem 3.2 are given exactly.

2. In the latter the relative error of solution is also considered, however, the form is not referred to.

4. Numerical experiments. In this section, applying MATLAB 6.5, we illustrate theory results on backward error generated from the partitioned LU factorization for block tridiagonal matrices and on the relative error of solution to linear systems.

***Example* 4.1.** Let block tridiagonal matrices be generated from the discretization of partial differential equation $-\Delta u = f$, where $A_i = \mathrm{tridiag}(-1, 4, -1)_{\mathrm{k_i \times k_i}}$. Some results corresponding to the example are listed in Table 4.1.

Table 4.1

Size	$\|A - \hat{L} * \hat{U}\|$	$\|x - \hat{x}\|/\|\hat{x}\|$
900×900	$1.7764e - 015$	$2.2204e - 015$
1600×1600	$2.6645e - 015$	$1.0880e - 014$
3600×3600	$3.5527e - 015$	$1.4655e - 014$

***Example* 4.2.** Let A be random block tridiagonal matrices, where A_i, B_i and C_i are random matrices with approximately $0.8 \times k_i \times k_i$, $0.2 \times k_i \times k_{i-1}$ and $0.2 \times k_{i-1} \times k_i$ uniformly distributed nonzero entries, respectively. The results listed in Table 2.

Table 4.2

Size	$\|A - \hat{L} * \hat{U}\|$	$\|x - \hat{x}\|/\|\hat{x}\|$
900×900	$5.6843e - 014$	$3.4195e - 013$
1600×1600	$1.2967e - 013$	$1.2765e - 012$
3600×3600	$8.1712e - 014$	$3.3598e - 012$

From the results as above, it shows that the errors $\|A - \hat{L} * \hat{U}\|$ and $\|x - \hat{x}\|/\|\hat{x}\|$ are very small. However, we can not say that the partitioned LU factorization must be stable, because the backward error contains $\|L\|$. For example,

$$A = \begin{pmatrix} \epsilon & 0 & 1 & 0 & & \\ 0 & \epsilon & 0 & 1 & & \\ 1 & 0 & \epsilon & 0 & 1 & 0 \\ 0 & 1 & 0 & \epsilon & 0 & 1 \\ & & 1 & 0 & 1 & 0 \\ & & 0 & 1 & 0 & 1 \end{pmatrix},$$

where ϵ is sufficient small. Applying the partitioned LU factorization in this paper gives

$$L = L_1L_2L_3 = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ \frac{1}{\epsilon} & & 1 & & & \\ & \frac{1}{\epsilon} & & 1 & & \\ & & \frac{\epsilon}{\epsilon^2-1} & & 1 & \\ & & & \frac{\epsilon}{\epsilon^2-1} & & 1 \end{pmatrix}.$$

Hence, $\|L\|$ is boundless when ϵ is sufficient small.

References

1. *Barrland A.* Perturbation bounds for the LDL^H and the LU factorizations // BIT. – 1991. – **31**. – P. 358 – 363.
2. *Stewart G. W.* Perturbation bounds for the QR factorizaton of a matrix // SIAM J. Numer. Anal. – 1977. – **14**. – P. 509 – 518.
3. *Stewart G. W.* On the perturbation of LU, Cholesky, and QR factorizations // SIAM J. Numer. Anal. – 1993. – **14**. – P. 1141 – 1145.
4. *Stewart G. W.* On the perturbation of LU and Cholesky factors // SIAM J. Numer. Anal. – 1997. – **17**. – P. 1 – 6.
5. *Xiao-Wen Chang, Christopher C. Paige.* On the sensitivity of the LU factorization // BIT. – 1998. – **38**. – P. 486 – 501.
6. *Dopico F. M., Molera J. M.* Perturbation theory of factorizations of LUF type through series expansions // SIAM J. Matrix Anal. and Appl. – 2005. – **27**. – P. 561 – 581.
7. *Xiao-Wen Chang, Christopher C. Paige, Stewart G. W.* Perturbation analyses for the QR factorization // SIAM J. Matrix Anal. and Appl. – 1997. – **18**. – P. 775 – 791.
8. *Higham N. J.* Accuracy and stability of numerical algorithm. – Philadelphia: SIAM, 1996.
9. *Higham N. J.* Accuracy and stability of algorithms in numerical linear algebra // Manchester Center Comput. Math. Numer. Anal. Rept. – 1998. – № 333.
10. *Higham N. J.* The accuracy of solutions to triangular systems // SIAM J. Numer. Anal. – 1989. – **26**. – P. 1252 – 1256.
11. *Amodio P., Mazzia F.* A new approach to the backward error analysis in the LU factorization algorithm // BIT. – 1999. – **39**. – P. 385 – 402.
12. *Demmel J. W., Higham N. J., Shreiber R. S.* Stability of block LU factorization // Numer. Linear Algebra and Appl. – 1995. – **2**. – P. 173 – 190.
13. *Jinxi Zhao, Weiguo Wang, Weiqing Ren.* Stability of the matrix factorization for solving block tridiagonal symmetric indefinite linear systems // BIT Numer. Math. – 2004. – **44**. – P. 181 – 188.
14. *Mattheij R. M. M.* The stability of LU-decompostions of block tridiagonal matrices // Bull. Austral. Math. Soc. – 1984. – **29**. – P. 177 – 205.
15. *Forsgren A., Gill P. E., Shinnerl J. R.* Stability of symmetric ill-conditioned systems arising in interior methods for constrained optimization // SIAM J. Matrix Anal. and Appl. – 1996. – **17**. – P. 187 – 211.
16. *Bueno M. I., Dopico F. M.* Stability and sensitivity of tridiagonal LU factorization without pivoting // BIT Numer. Math. – 2004. – **44**. – P. 651 – 673.
17. *Demmel J. W., Higham N. J.* Stability of block algorithms with fast level-3 BLAS // ACM Math. Software. – 1992. – **18**. – P. 274 – 291.
18. *Stoer J., Bulirsch R.* Introduction to numerical analysis: – 2nd ed. – New York: Springer-Verlag, 1993.

Received 25.04.13