MINIMUM DEVIATION, QUASI-LU FACTORIZATION OF NONSINGULAR MATRICES. *

M.I. BUENO† AND C.R. JOHNSON‡

Abstract. Not all matrices enjoy the existence of an LU factorization. For those that do not, a number of "repairs" are possible. For nonsingular matrices we offer here a permutation-free repair in which the matrix is factored $\tilde{L}\tilde{U}$, with $\tilde{L}$ and $\tilde{U}$ collectively as near as possible to lower and upper triangular (in a natural sense defined herein). Such factorization is not generally unique in any sense. In the process, we investigate further the structure of matrices without LU factorization and permutations that produce an LU factorization.

Key words. LU factorization, LPU factorization, quasi-LU factorization, sparsity pattern.

AMS subject classifications. 15A21, 15A23, 05A05, 05B25.

1. Introduction. Factorization of an m-by-m matrix A into a lower triangular matrix L and an upper triangular matrix U ("LU factorization") is important for a variety of computational, theoretical and applied reasons. Although the LU factorization is well known for its applications to the solution of linear systems of equations, there are many other applications of this factorization. Consider for instance the use of the LU factorization to compute the singular values of bidiagonal matrices [9] or to prove determinantal inequalities [1]. In any case, the LU factorization is an important mathematical concept by itself, and a generalization of this idea is studied here. Unfortunately, the LU factorization does not always exist. Characterizations of those A for which an LU factorization does exist are given in [8]. When an LU factorization does not exist, several repairs have been explored, both theoretically and computationally, such as allowing a permutation of the rows or columns of A [7], or allowing a (sub-)permutation between L and U [3, 6]. Each of these gives universal existence and is of some benefit.

Here we explore another form of repair that requires no explicit permutation. It is in the spirit of the characterization [8], but leads to a number of new ideas. If A is nonsingular and has no LU factorization, how may A be factored into $\tilde{L}$ and $\tilde{U}$ that are "nearly" lower and upper triangular ("quasi-LU factorization")? To be precise, we need an explicit measure of nearness to a matrix being triangular (upper or lower). We define the lower excess of a matrix M as

$$exc_L(M) = \sum_{s_j<j} (j - s_j) \tag{1.1}$$

where the (s_j, j) denote the positions of the "highest" nonzero entries in the columns of M, that is, for all $j = 1, ..., m$, $M(s_j, j) \neq 0$, and $M(i, j) = 0$ for $i < s_j$. Notice from (1.1) that only the pairs (s_j, j) above the main diagonal of M are considered in

*The work of the first author has been supported by Dirección General de Investigación (Ministerio de Ciencia y Tecnología) of Spain under grant MTM2006-06671/, as well as by the Postdoctoral Fellowship EX2004-0658 provided by Ministerio de Educación y Ciencia of Spain.

†Mathematics Department, The University of California, South Hall 6607, Santa Barbara, CA 93106 USA(mbueno@math.ucsb.edu).

‡Mathematics Department, The College of William and Mary, P.O. Box 8795, Williamsburg, VA 23187-8795 USA (crjohnso@math.wm.edu)

the summation. Equivalently, we define the upper excess of a matrix M as

$$exc_U(M) = \sum_{t_i<i}(i - t_i), \tag{1.2}$$

where the (i, t_i) denote the positions of the "leftmost" nonzero entries in the rows of M. Again, notice from (1.2) that only the pairs (i, t_i) below the main diagonal of M are considered in the summation. In each case, the excess is a measure of how far the columns (the rows) vary from the indicated form of triangularity. Let A be an m-by-m nonsingular matrix with no LU factorization, and consider any quasi-LU factorization $A = \tilde{L}\tilde{U}$ of A. We measure the deviation of this factorization from an LU factorization in the following way:

$$dev(\tilde{L}, \tilde{U}) = exc_L(\tilde{L}) + exc_U(\tilde{U}), \tag{1.3}$$

that is, we measure how far $\tilde{L}$ and $\tilde{U}$ are, respectively, from being a lower and an upper triangular matrix. In this paper, we give the minimum deviation from an LU factorization among all the quasi-LU factorizations of a matrix A. We express this minimum in terms of the nullity of the leading principal submatrices of A. Our "deviation" is closely related to a concept for a single matrix, rather than a pair, that arises in sparse matrix analysis [2, 4]. Given any variable-band matrix A, they define the lower (upper) semibandwidth of A as the smallest integer d_l (d_u) such that $a_{ij} = 0$ whenever $i - j > d_l$ ($j - i > d_u$). Then, the bandwidth of A is $d_l + d_u + 1$. Gaussian elimination without interchanges preserves the band structure and band matrix methods provide an easy way to exploit zeros in a matrix. When storing variable-band matrices in a computer, for each column every coefficient between the first entry in the column (row) and the diagonal is stored. The total number of coefficients stored is called the profile. If we denote the profile of an m-by-m matrix A by $p(A)$, note that $p(A) = exc_L(A) + exc_U(A) + m$. Variable-band matrices are also known as skyline, profile and envelope matrices.

The necessary and sufficient conditions for a matrix A to have an LU factorization given in [8] are

$$rank(A_{kk}) + k \geq rank(A_{km}) + rank(A_{mk}), \text{ for all } k = 1, ..., m-1, \tag{1.4}$$

in which A_{rs} (as throughout the paper) denotes the submatrix of A containing the first r rows and the first s columns of A. We define the k-th failure of A as n_k, given by

$$n_k = \max\{0, rank(A_{km}) + rank(A_{mk}) - rank(A_{kk}) - k\}. \tag{1.5}$$

Notice that, when A is a nonsingular matrix, $n_k = k - rank(A_{kk})$ is the nullity of A_{kk}. Moreover, A has an LU factorization if and only if $n_k = 0$ for $k = 1, ..., m$. Therefore, it is natural that when these conditions fail, the nullities n_k must play a role. We call the nullities of an m-by-m matrix A the set of numbers $\{n_k : k = 1, ..., m\}$. In Theorem 3.13 we show (which appears to be subtle) that the deviation of any quasi-LU factorization of a nonsingular matrix A from an LU factorization cannot be smaller than the sum of its nullities, that is, if $A = \tilde{L}\tilde{U}$, then

$$dev(\tilde{L}, \tilde{U}) \geq \sum_{k=1}^{m} n_k. \tag{1.6}$$

It is known (Theorem 2.1) that if A is a nonsingular matrix with no LU factorization then, A can be written as LPU where L is unit lower triangular, P is a permutation matrix, and U is upper triangular (LU factorization for arbitratry matrices). As we will show, if $A = LPU$, and we set $\tilde{L} = LP$, then $dev(\tilde{L}, U) = exc_L(\tilde{L}) = \sum_{k=1}^{m} n_k$. We could, alternatively, push P into U. The key question is: could we do better? In this paper, we show that it is not possible.

The paper is organized as follows: In Section 2, we define the so-called key positions of a nonsingular matrix and relate them to the LPU factorization. We show that the rank of any leading submatrix of A equals the number of key positions in it which leads us to an expression for the nullities of A in terms of the number of key positions in each leading principal submatrix of A. We also include a new characterization of the existence of an LU factorization of a matrix that arises from the definition of the key positions. In Section 3 we prove the main result. We also show that the lower bound given in (1.6) is attainable and we give a set of minimum deviation quasi-LU factorizations of any nonsingular matrix. Moreover, we include an algorithm to compute one of those factorizations, in which $\tilde{U}$ is upper triangular. Section 4 contains auxiliary results related with the lower and upper excess of a matrix that are used in the proof of the main result (Theorem 3.13). Some seem interesting on their own.

2. The key positions of a nonsingular m-by-m matrix. In this section we define the key positions of a nonsingular matrix A. This definition arises in a very natural way from the so-called LPU factorization of A. The nullities (1.5) of A can be expressed in terms of these key positions (Lemma 2.10). In Section 3 this result will be useful to give a set of quasi-LU factorizations in which the minimum deviation (1.3) is attained.

As we mentioned in the introduction, not every m-by-m matrix A has an LU factorization. In [5, 3, 6] the following generalization of the LU factorization is considered.

THEOREM 2.1. *Given any m-by-m nonsingular matrix A, there exists a unit lower triangular matrix L, a nonsingular upper triangular matrix U and a permutation matrix P such that $A = LPU$. Moreover, P is the unique permutation matrix such that*

$$rank(A_{rs}) = rank(P_{rs}), \quad \text{for every } r, s \in \{1, ..., m\}. \tag{2.1}$$

Although Gohberg and Goldberg called it the LU factorization of arbitrary matrices in [6], we refer to the factorization presented in Theorem 2.1 as an LPU factorization.

In Theorem 2.1 it is not discussed if L and U are also unique. In fact, they are not as the example below shows.

Let $A = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 2 & 5 \\ 1 & 0 & -6 \end{bmatrix}$. The matrix A may be factorized, among infinitely many possibilities, in the following two ways:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 3 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix},$$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & -2 & -6 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

In general, if we consider any two LPU factorizations L_1PU_1 and L_2PU_2 of a nonsingular matrix A, where L_1 and L_2 are unit nonsingular lower triangular matrices, then

$$(L_2^{-1}L_1)P = P(U_2U_1^{-1}).$$

Let us denote by $L := L_2^{-1}L_1$ and $U := U_2U_1^{-1}$. Notice that if $P = I$, which implies that A has an LU factorization, then $L = U$. This means that L and U are diagonal matrices and we deduce that the LU factorization of A is unique up to diagonal scaling. However, if $P \neq I$, then $H := LP = PU$ is not a diagonal matrix anymore. Let $\{(i, l_i), i = 1, ..., m\}$ be the positions in which the nonzero entries of P are located. Reordering them, we may also denote these positions as $\{(t_j, j), j = 1, ..., m\}$. Then, H presents the following pattern:

$$H(i,j) = \begin{cases} 1, & j = l_i, \\ 0, & i < t_j \text{ or } j < l_i. \end{cases}$$

Therefore, $L_1 = L_2(HP^t)$, and $U_1 = (H^{-1}P)U_2$. Notice that HP^t and $H^{-1}P$ are diagonal matrices such that $H^{-1}P = (HP^t)^{-1}$ if and only if $P = I$. Then, if $P \neq I$, the factorization LPU of A is no longer unique.

Note that the following is an equivalence relation defined in the set of nonsingular matrices: Given two nonsingular matrices A and B, we say that A and B are T-congruent if there exists a nonsingular lower triangular matrix L and a nonsingular upper triangular matrix U such that $A = LBU$. According to Theorem 2.1, every nonsingular matrix is T-congruent to a unique permutation matrix. Therefore, P can be considered the canonical form of A under T-congruence.

Definition 2.2. *Let A be any nonsingular matrix and let P be the unique permutation matrix given in Theorem 2.1. We call P the key permutation matrix associated with A. The positions where the nonzero entries of P occur are called the key positions of A. We also say that the entries of A in the key positions are the key entries.*

From (2.1), we can say that the key permutation matrix P associated with a nonsingular matrix through an LPU factorization can be seen as an skeleton of A in the following sense: the rank of every leading submatrix of A equals the rank of the same leading submatrix in P.

Example 2.3. *Let* $A = \begin{bmatrix} 0 & 0 & 1 & 2 \\ 1 & 1 & -1 & 3 \\ 2 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 \end{bmatrix}$.
Then, the key permutation matrix associated with A is

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

The key positions of A are $\{(1,3),(2,1),(3,2),(4,4)\}$ while the corresponding key entries are, respectively, $\{1,1,0,0\}$. Notice that the key entries may be zero.

The next three results are simple consequences of the definition of the key positions. For that reason, we omit their proofs.

PROPOSITION 2.4. *If P is a permutation matrix, then P is its own key permutation matrix and its key positions are the locations in which the nonzero entries occur.*

PROPOSITION 2.5. *If A is a nonsingular matrix with LU factorization, its key permutation matrix is the identity matrix. Therefore, the key positions of A are $\{(i,i),\ i=1,...,m\}$.*

PROPOSITION 2.6. *Let A be any nonsingular matrix and let P be its key permutation matrix. Then A and P have the same key positions.*

Taking into account (2.1), the rank of any leading submatrix of A can be given in terms of the number of key entries it contains.

LEMMA 2.7. *Let A be any m-by-m nonsingular matrix. If P denotes the key permutation matrix associated with A, and p_{rs} denotes the number of key entries in A_{rs}, then*

$$rank(A_{rs}) = p_{rs}.$$

Proof. Assume that A is a permutation matrix. Consider the submatrix A_{rs} of A. Then, the number of key entries in A_{rs} equals the number of nonzero entries. Therefore, the result follows for permutation matrices.

Let us consider any nonsingular matrix A as well as the submatrix A_{rs}. If P denotes the key permutation matrix associated with A, since $rank(A_{rs}) = rank(P_{rs})$, the number of key entries in both submatrices is the same, and the statement follows. □

In the sequel, the following notations will be used: Given a matrix M, $M(i,:)$ and $M(:,j)$ denote the i-th row of M and the j-th column of M, respectively. Moreover, $M([1,...,i],:)$ and $M(:,[1,...,j])$ denote, respectively, the submatrix of M containing the first i rows of M and the first j columns of M. Finally, $M([1,2,..,k],[c_1,c_2,...,c_k])$ denotes the submatrix of M containing the first k rows of M and the columns of M given by $\{c_1,c_2,...,c_k\}$.

Theorem 2.1 and Lemma 2.7 imply the next two results in a straightforward way. They give alternative characterizations of the key positions of a nonsingular matrix A.

PROPOSITION 2.8. *[6] Let A be any m-by-m nonsingular matrix. Then $\{(i,l_i),\ i=1,...,m\}$ is the set of key positions of A if and only if for $i \geq 1$, l_i is the index such that*

$$rankA_{i-1,l_i-1} = rankA_{i,l_i-1} = rankA_{i-1,l_i} = rankA_{i,l_i} - 1,$$

in which we take $rankA_{i,0} = 0$, and $rankA_{0,j} = 0$ by convention.

PROPOSITION 2.9. *Let A be any m-by-m nonsingular matrix . Then, $\{(i,l_i),\ i=1,...,m\}$ are the key positions of A if and only if l_k is the minimum index such that*

$$rank(A([1,2,..,k],[l_1,l_2,...,l_k])) = k, \quad \text{for all } k=1,...,m. \tag{2.2}$$

As we mentioned in the introduction, one of our goals is to bound the deviation of any quasi-LU factorization of a nonsingular matrix A in terms of the nulllities of A (1.5). It turns out that the nullity n_k of A can be expressed in terms of the number of key entries in the submatrix A_{kk} as follows:

LEMMA 2.10. *Let A be any m-by-m nonsingular matrix. Then, the nullity n_k of A can also be computed as*

$$n_k = k - p_{kk},$$

where p_{kk} denotes the number of key positions in A_{kk}.

Proof. The result is a consequence of the definition of n_k and Lemma 2.7. □

The next example shows that, although the key positions charaterize the nullities of A, the converse does not happen. This example gives two matrices with different key positions and the same nullities.

EXAMPLE 2.11. *The two matrices given below have the same nullities: $n_1 = 1$, $n_2 = 1$, $n_3 = 1$, $n_4 = 0$. However the key positions corresponding to each of them are different.*

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 2 & -1 & 3 \\ 1 & 2 & -1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 2 & 0 & 1 & 4 \end{bmatrix}.$$

In the sequel we will often use the following permutation matrix.

$$T = \begin{bmatrix} 0 & 0 & ... & 0 & 1 \\ 0 & 0 & ... & 1 & 0 \\ ... & ... & ... & ... & ... \\ 0 & 1 & ... & 0 & 0 \\ 1 & 0 & ... & 0 & 0 \end{bmatrix}. \tag{2.3}$$

This is the so-called backward identity matrix and we denote it by T. The following proposition gives the key positions of the matrices A^t and $TA^{-1}T$, in terms of the key positions of A. This result will be used to prove the main result (Theorem 3.13).

PROPOSITION 2.12. *Let A be any nonsingular matrix and let P be its key permutation matrix. Then,*

1. *the key permutation matrix associated with A^t is P^t.*
2. *the key permutation matrix associated with the matrix $TA^{-1}T$ is given by TP^tT.*

Proof.

According to Theorem 2.1, there exist a nonsingular unit lower triangular matrix L and a nonsingular upper triangular matrix U such that $A = LPU$. Compute the transpose of this matrix and we get

$$A^t = U^t P^t L^t.$$

Then, taking into account Theorem 2.1 again, the first claim follows.

On the other hand,

$$TA^{-1}T = TU^{-1}P^tL^{-1}T = (TU^{-1}T)(TP^tT)(TL^{-1}T),$$

and the second claim follows. □

The key positions of a nonsingular matrix can also be viewed as pivot positions in the following sense.

PROPOSITION 2.13. *Let A be any m-by-m nonsingular matrix, and let P be its key permutation matrix. Then, both matrices $\tilde{A} = AP^t$ and $\hat{A} = P^tA$ have an LU factorization.*

Proof. Notice that if $\{(i, l_i) : i = 1, ..., m\}$ denote the key positions of A, for $1 \leq k \leq m$,

$$\tilde{A}_{kk} = A([1, .., k], :)P^t(:, [1, .., k]) = A([1, .., k], [l_1, .., l_k]).$$

According to Proposition 2.9, $rank A([1, .., k], [l_1, .., l_k]) = k$, therefore the submatrix $\tilde{A}_{kk}$ has full rank for all k, and the first claim follows.

According to Proposition 2.12, P^t is the key permutation matrix of A^t. Using the result we have just proven, A^tP has an LU factorization, and therefore $(A^tP)^t = P^tA$ has also an LU factorization. □

Using the definition of the key positions, we give a new characterization of matrices with an LU factorization. The next proposition will also be important in order to prove the main result of this paper.

PROPOSITION 2.14. *Let A be any m-by-m nonsingular matrix. Then, A has an LU factorization if and only if*

$$rank((\hat{P}A)_{rs}) \geq rank(\hat{P}_{rs}), \quad \textit{for all}\, r, s \in \{1, ..., m\}$$

and for every permutation matrix $\hat{P}$.

Proof.

Assume that A has an LU factorization and $\hat{P}$ is any permutation matrix. Let $\{(t_i, i) : i = 1, ..., m\}$ be the key positions of $\hat{P}$. According to Proposition 2.5, the key permutation matrix of A is the identity matrix, and therefore, the key entries of A appear in positions $\{(t_i, i) : i = 1, ..., m\}$ of $\hat{P}A$.

Consider the submatrix $\hat{P}_{rs}$ of $\hat{P}$. According to Lemma 2.7, the rank of $\hat{P}_{rs}$ equals the number of key entries in this submatrix. If $rank(\hat{P}_{rs}) = 0$, then the result follows in a straightforward way. Assume that $rank(\hat{P}_{rs}) \neq 0$. Then, the rows in $(\hat{P}A)_{rs}$ containing key entries of A are rows of the $s \times s$ leading principal submatrix of A. Since this submatrix is nonsingular, the rows are linearly independent vectors and the result follows.

Assume now that

$$rank((\hat{P}A)_{rs}) \geq rank(\hat{P}_{rs}), \quad \text{for every}\, r, s \in \{1, ..., m\},$$

for any permutation matrix $\hat{P}$. In particular, if $\hat{P} = I$ and $r = s$, we get

$$rank(A_{rr}) \geq rank(I_{rr}),$$

which implies the statement. □

From the previous proposition we get another characterization of the key permutation matrix of a nonsingular matrix A.

COROLLARY 2.15. *Let A be any m-by-m nonsingular matrix. The matrix P is the key permutation matrix associated with A if and only if*

$$rank((\hat{P}A)_{rs}) \geq rank((\hat{P}P)_{rs}), \quad \textit{for all}\, r, s \in \{1, ..., m\}$$

and for every permutation matrix $\hat{P}$.

Proof. According to Proposition 2.13, the matrix $B = P^t A$ has an LU factorization. Therefore, from Proposition 2.14,

$$rank((\tilde{P}B)_{rs}) \geq rank((\tilde{P})_{rs}), \quad \text{for all } r, s \in \{1, ..., m\}$$

and for every permutation matrix $\tilde{P}$. Then,

$$rank((\tilde{P}P^t A)_{rs}) \geq rank((\tilde{P})_{rs}).$$

If we denote by $\hat{P} := \tilde{P}P^t$, our claim follows. □

3. Minimum deviation quasi-LU factorization. This section is devoted to our main result. We prove that the deviation from an LU factorization (see 1.3) of any quasi-LU factorization of a nonsingular matrix A is always at least equal to the sum of the nullities of A, i.e., the sum of the $n'_k s$, defined in (1.5). We also prove that this lower bound can be attained. Any nonsingular matrix A with key permutation matrix P can be factorized as $A = \tilde{L}U$ where $\tilde{L} = LP$ is an almost lower triangular matrix (see Definition 3.1) and U is an upper triangular matrix. Any of the quasi-LU factorizations of A obtained in this way satisfies that its deviation from an LU factorization equals the sum of the nullities of A. We call each of these factorizations a left key quasi-LU factorization of A. In this section, we present the MATLAB code of an algorithm that computes one of these factorizations. Similarly the set of right key quasi-LU factorizations of A can be defined, and alike results to those obtained for the left key factorizations can be proven. We also show that although both kind of factorizations achieve the minimum deviation, one of them produces, in general, fewer extra diagonals than the other one.

Most of the proofs of the results in this section involve auxiliary lemmas given in Section 4. For the sake of clarity, we include these auxiliary results at the end of the paper.

Definition 3.1. *An m-by-m matrix* $\tilde{L}$ *is said to be almost lower triangular with upper semibandwidth* d_u *if* $\tilde{L}(i,j) = 0$ *whenever* $j > i + d_u$ *and* $\tilde{L}(i, i + d_u) \neq 0$ *for some* i. *An almost upper triangular matrix* $\tilde{U}$ *with lower semibandwidth* d_l *is defined similarly.*

Definition 3.2. *Let* A *be any m-by-m nonsingular matrix, and let* P *be its key permutation matrix. Consider any nonsingular lower triangular matrix* L *and any nonsingular upper triangular matrix* U *such that* $A = LPU$. *Let* $A = \tilde{L}U$, *where* $\tilde{L} = LP$. *We say that* $A = \tilde{L}U$ *is a left key quasi-LU factorization of* A.

In general, if $\tilde{L}$ *denotes any almost lower triangular matrix, we say that* $A = \tilde{L}U$ *is a left quasi-LU factorization of* A.

In Proposition 3.4, we express the deviation of any left key quasi-LU factorization of a nonsingular matrix A in terms of the lower excess of its key permutation matrix. The next lemma shows that the lower (1.1) and the upper (1.2) excess of a permutation matrix are equal and, we call this common value the excess of a permutation matrix.

Lemma 3.3. *Let* P *be a permutation matrix. If* $\{(i, l_i) : i = 1, ..., m\}$ *denote the key positions of* P, *then,*

$$exc_L(P) = exc_U(P) = \sum_{i=1}^{m} \frac{|l_i - i|}{2}.$$

Proof. Notice that, by definition

$$exc_U(P) = \sum_{l_i < i} (i - l_i), \quad exc_L(P) = \sum_{i < l_i} (l_i - i). \tag{3.1}$$

Since $(l_1, ..., l_m)$ is a permutation of $(1, 2, ..., m)$, we get

$$\sum_{i=1}^{m} l_i = \sum_{i=1}^{m} i.$$

Therefore,

$$0 = \sum_{i=1}^{m}(l_i - i) = \sum_{l_i > i}(l_i - i) + \sum_{l_i < i}(l_i - i) = exc_L(P) - exc_U(P).$$

Moreover,

$$\sum_{i=1}^{m} \frac{|l_i - i|}{2} = \sum_{l_i > i} \frac{(l_i - i)}{2} + \sum_{l_i < i} \frac{(i - l_i)}{2} = \frac{exc_L(P)}{2} + \frac{exc_U(P)}{2} = exc_L(P),$$

and the result follows. □

Since $exc_L(P) = exc_U(P)$ for every permutation matrix P, in the sequel we will denote the common value by $exc(P)$.

PROPOSITION 3.4. *Let A be any nonsingular matrix and let $A = \tilde{L}U$ be a left key quasi-LU factorization of A. If P denotes the key permutation matrix associated with A, then*

$$dev(\tilde{L}, U) = exc_L(\tilde{L}) = exc(P).$$

Proof. The result follows from (1.3) and Lemma 4.2. □

The next corollary gives the value of the deviation from an LU factorization of any left key quasi-LU factorization of a matrix A in terms of the key positions of A. Furthermore, it also gives the upper semibandwidth of $\tilde{L}$. This corollary will also be the key to prove that the deviation of any left key quasi-LU factorization can be computed as the sum of the nullities of A, that is, the sum of the n_k's.

COROLLARY 3.5. *Let A be any m-by-m nonsingular matrix, and let $A = \tilde{L}U$ be a left key quasi-LU factorization of A. If P denotes the key permutation matrix of A and $\{(i, l_i) : i = 1, ..., m\}$ denote the key positions of A, then the deviation of $\tilde{L}U$ from an LU factorization is given by*

$$dev(\tilde{L}, U) = \sum_{i < l_i}(l_i - i). \tag{3.2}$$

Moreover, $\tilde{L}$ is an almost lower triangular matrix with upper semibandwidth d_u, where $d_u = \max_i\{l_i - i\}$.

Proof. Taking into account Proposition 3.4, and (3.1), the result in (3.2) follows in a straightforward way.

Considering Definition 3.1, the result regarding the upper semibandwidth can be obtained. □

THEOREM 3.6. *Let A be any m-by-m nonsingular matrix, and let $\{n_1, n_2, ..., n_m\}$ be the nullities of A. If $A = \tilde{L}U$ denotes a left key quasi-LU factorization of A, then*

$$dev(\tilde{L}, U) = \sum_{k=1}^{m} n_k,$$

i.e., the deviation from an LU factorization produced by any left key quasi-LU factorization of A equals the sum of its nullities.

Proof. Let us denote by p_{kk} the number of key positions in the submatrix A_{kk}. Then, by Lemma 2.10,

$$\sum_{k=1}^{m} n_k = \sum_{k=1}^{m} (k - p_{kk}).$$

Let us denote by t_k the number of leading principal submatrices in which the key position (k, l_k) appears. Then,

$$t_k = m + 1 - \max\{k, l_k\} = \begin{cases} m+1-k, & \text{if} \quad l_k \leq k, \\ m+1-l_k, & \text{if} \quad l_k > k. \end{cases}.$$

Hence,

$$\sum_{k=1}^{m} (m+1-k-t_k) = \sum_{l_k > k} (m+1-k-t_k) = \sum_{l_k > k} (l_k - k).$$

Therefore, taking into account Corollary 3.5, we want to prove that

$$\sum_{k=1}^{m} (m+1-k-t_k) = \sum_{k=1}^{m} (k - p_{kk}).$$

But, since $\sum_{k=1}^{m} (m+1-k) = \sum_{k=1}^{m} k$, it suffices to prove that $\sum_{k=1}^{m} t_k = \sum_{k=1}^{m} p_{kk}$.

Let us denote by b_k the number of key positions in A_{kk} that are not in $A_{k-1,k-1}$. Then,

$$p_{kk} = p_{11} + \sum_{j=2}^{k} b_j, \quad \text{for } k \geq 2,$$

and, we get

$$\sum_{k=1}^{m} p_{kk} = mp_{11} + \sum_{j=2}^{m} [(m-j+1)b_j] = \sum_{k=1}^{m} t_k.$$

□

Observe the following two almost lower triangular matrices.

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 0 \\ -2 & 1 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Notice that although they have a different “profile”, they have the same lower excess.

Definition 3.7. *Let A be any m-by-m nonsingular matrix. We call lower profile of A the set $\{(r_j, j) : j = 1, ..., m\}$ of positions of A such that*

1. *$A(r_j, j) \neq 0$ for $j = 1, ..., m$.*
2. *$A(i, j) = 0$ if $i < r_j$.*

Equivalently, the upper profile can be defined.

Note that the term "profile" in this paper has a slightly different meaning than the same term when used in the context of sparse matrices. [2, 4].

LEMMA 3.8. *Let $A = LP$, where L is a lower triangular matrix and P is a permutation matrix. Then, A and P have the same lower profile. Similarly, if $A = PU$, where U is an upper triangular matrix, then A and P have the same upper profile.*

Then, a natural question arises from Theorem 3.6. If $A = \tilde{L}U$ denotes any left quasi-LU factorization of A, is it possible to find different lower profiles for $\tilde{L}$ such that the deviation of the corresponding quasi-LU factorizations is exactly the sum of the nullities of A? The answer is negative.

THEOREM 3.9. *Let A be a nonsingular matrix and let P be its key permutation matrix. If $\{n_k\}$ denote the nullities of A, and $A = \tilde{L}U$ denotes any left quasi-LU factorization of A, then there exists a unique lower profile for $\tilde{L}$ such that $dev(\tilde{L}, U) = \sum_{k=1}^{m} n_k$. Moreover, it is the lower profile corresponding to the left key quasi-LU factorizations of A.*

Proof. Let L and U be, respectively, a lower and an upper triangular matrix such that $A = LPU$. If $A = \tilde{L}U_1$ denotes any left quasi-LU factorization of A,

$$\tilde{L}U_1 = LPU,$$

which implies

$$\tilde{L} = LP(UU^{-1}).$$

Taking into account Theorem 2.1, we deduce that the key permutation matrix of $\tilde{L}$ is P. According to Definition 4.3 and Lemma 4.4, $\tilde{L} \geq_L P$, and the statement follows. □

Next we give the MATLAB code of an algorithm that computes the matrices $\tilde{L}$ and U corresponding to a left key quasi-LU factorization of A. Notice that no permutation is involved in the process.

ALGORITHM 1. *Given an m-by-m nonsingular matrix A, this algorithm computes a left key quasi-LU factorization of A:*

```
function [Lt,U]=LtildeU(A)

[m,m]=size(A); Lt=zeros(m); U=zeros(m);

for j=1:m
 for i=1:m
  if A(i,j)~=0
        U(j,:)=A(i,:)
        Lt(:,j)=A(:,j)/A(i,j)
        A=A-Lt(:,j)*U(j,:)
    end
  end
end
```

The set of right key quasi-LU factorizations of A can also be defined. If A denotes a nonsingular matrix and $A = LPU$ then, $A = L\tilde{U}$, where $\tilde{U} = PU$, is said to be a right key quasi-LU factorization of A.

THEOREM 3.10. *Let A be any m-by-m nonsingular matrix, and let $\{n_1, n_2, ..., n_m\}$ be the nullities of A. If $A = L\tilde{U}$ denotes a right key quasi-LU factorization of A, then*

$$dev(L, \tilde{U}) = \sum_{k=1}^{m} n_k,$$

i.e., the deviation from an LU factorization produced by any right key quasi-LU factorization of A equals the sum of the nullities of A.

Proof. Taking into account Proposition 2.12, $A^t = (U^t P^t) L^t$ is a left key quasi-LU factorization of A^t. Notice that, since $rank(A_{kr}) = rank(A^t_{kr})$, the matrices A and A^t have the same set of nullities. Then, taking into account Theorem 3.6, Proposition 3.4, and Lemma 3.3

$$\sum_{k=1}^{m} n_k = dev(\tilde{U}^t, L^t) = exc(P^t) = exc(P).$$

Finally, by definition of deviation and taking into account Lemma 4.2

$$dev(L, \tilde{U}) = exc_U(\tilde{U}) = exc(P),$$

and the result follows. □

Notice that although both the left and the right key quasi-LU factorizations of a nonsingular matrix achieve minimum deviation, the almost triangular factor of one of them may present larger semibandwidth than the other when the key permutation matrix P is not symmetric.

EXAMPLE 3.11. *Consider the matrix A given in Example 2.3. According to Corollary 3.5, if $A = \tilde{L}U$ and $A = L\tilde{U}$ are, respectively, a left and a right key quasi-LU factorization of A, then*

$$dev(\tilde{L}, U) = dev(L, \tilde{U}) = exc(P) = 2.$$

However, the upper semibandwidth of $\tilde{L}$ is two while the lower semibandwidth of $\tilde{U}$ is one.

Next we show that if $A = \tilde{L}\tilde{U}$ denotes any quasi-LU factorization of A, then the deviation of this factorization from an LU factorization cannot be smaller than the sum of the nullities of A.

3.1. The main result. First, we introduce an auxiliary lemma that shows how to express any quasi-LU factorization of a nonsingular matrix in terms of a left key quasi-LU factorization of the same matrix.

LEMMA 3.12. *If A is a nonsingular matrix, $A = \tilde{L}_1\tilde{U}_1$ denotes any quasi-LU factorization of A, and $A = \tilde{L}U$ denotes a left key quasi-LU factorization of A, then there exists a nonsingular matrix H such that*

$$\tilde{L}_1 = \tilde{L}H, \quad \textit{and} \quad \tilde{U}_1 = H^{-1}U.$$

Proof. Let $H = \tilde{L}^{-1}\tilde{L}_1$. Then,

$$A = \tilde{L}_1\tilde{U}_1 = \tilde{L}H\tilde{U}_1.$$

Taking into account that $\tilde{L}$ is nonsingular and $A = \tilde{L}U$,

$$H\tilde{U}_1 = U,$$

and the result follows. □

THEOREM 3.13. *Let A be any m-by-m nonsingular matrix and let $\{n_k\}_{k=1}^m$ denote the nullities of A. If P denotes the key permutation matrix of A, and $A = \tilde{L}_1\tilde{U}_1$ denotes any quasi-LU factorization of A, then*

$$dev(\tilde{L}_1, \tilde{U}_1) \geq \sum_{k=1}^{m} n_k.$$

Proof. If $A = \tilde{L}U$ denotes any left key quasi-LU factorization of A, from Lemma 3.12, there exists a nonsingular matrix H such that

$$\tilde{L}_1 = \tilde{L}H, \quad \tilde{U}_1 = H^{-1}U.$$

Therefore, by (1.3)

$$dev(\tilde{L}_1, \tilde{U}_1) = exc_L(\tilde{L}H) + exc_U(H^{-1}U). \tag{3.3}$$

Since $\tilde{L} = LP$ for some lower triangular matrix L, and taking into account Lemma 4.2, we get

$$dev(\tilde{L}_1, \tilde{U}_1) = exc_L(PH) + exc_U(H^{-1}). \tag{3.4}$$

Let P_H be the key permutation matrix associated with H^{-1}. Then, by Corollary 4.6,

$$exc_U(H^{-1}) \geq exc(P_H). \tag{3.5}$$

By Proposition 2.12, the matrix TP_H^tT is the key permutation matrix associated with THT, where T is the matrix given in (2.3). Therefore, according to Proposition 2.13, the matrix $(TP_HT)(THT) = TP_HHT$ has an LU factorization. Let $\hat{P} := PP_H^tT$. Then, by Corollary 4.17,

$$exc_L(PH) = exc_L(\hat{P}TP_HHTT) \geq exc(\hat{P}T). \tag{3.6}$$

which implies that

$$exc_L(PH) \geq exc(PP_H^t). \tag{3.7}$$

From (3.4), (3.5) and (3.7),

$$dev(\tilde{L}_1, \tilde{U}_1) \geq exc(PP_H^t) + exc(P_H). \tag{3.8}$$

Since $exc(P_H) = exc(P_H^t)$, from Lemma 4.1, Proposition 3.4 and Theorem 3.6, the result follows.

□

4. Auxiliary results: the lower and upper excess of matrices. This section is dedicated to the study of properties involving the lower and upper excess of nonsingular matrices. It includes several technical lemmas that are necessary to prove the main result.

LEMMA 4.1. *If P and Q denote two m-by-m permutation matrices, then*

$$exc(PQ) + exc(Q) \geq exc(P).$$

Proof. Let $\{(p_i, i) : i = 1, ..., m\}$ and $\{(i, q_i) : i = 1, ..., m\}$ be the key positions of P and Q, respectively. Then, the key positions of PQ are $\{(p_i, q_i) : i = 1, ..., m\}$ and, according to Lemma 3.3

$$exc(PQ) + exc(Q) = \sum_{i=1}^{m} \frac{|p_i - q_i|}{2} + \sum_{i=1}^{m} \frac{|q_i - i|}{2}.$$

Applying the triangular inequality,

$$\sum_{i=1}^{m} \frac{|p_i - q_i|}{2} + \sum_{i=1}^{m} \frac{|q_i - i|}{2} \geq \sum_{i=1}^{m} \frac{|p_i - i|}{2} = exc(P).$$

□

LEMMA 4.2. *If B is any nonsingular matrix, L is a nonsingular lower triangular matrix and U is a nonsingular upper triangular matrix, then*

$$exc_L(LB) = exc_L(B), \quad exc_U(BU) = exc_U(B).$$

Proof. In order to prove the first statement, it is sufficient to prove that B and LB have the same lower profile. Let us denote $A := LB$, then

$$A(1, :) = L(1, :)B = L(1, 1)B(1, :).$$

Since $L(1, 1) \neq 0$,

$$A(1, i) \neq 0 \quad \text{if and only if} \quad B(1, i) \neq 0, \quad \text{for} \quad i = 1, ..., m.$$

If $B(1, i) \neq 0$ for all i, the lemma has been proven. Otherwise, suppose that $B(1, i) = 0$ for some i. Assume that

$$B(r, i) = 0 = A(r, i), \quad \text{for} \quad r = 1, ..., k.$$

Then,

$$A(k+1, i) = L(k+1, :)B(:, i) = L(k+1, 1 : k+1)B(1 : k+1, i) = L(k+1, k+1)B(k+1, i).$$

Therefore,

$$A(k + 1, i) \neq 0 \quad \text{if and only if} \quad B(k + 1, i) \neq 0.$$

The second statement may be proven in a similar way.

□

In order to compare the lower excess or the upper excess of two matrices we define the following two partial orders on the set of m-by-m nonsingular matrices.

DEFINITION 4.3. *Let A and B be two m-by-m matrices, and let $\{(a_i, i) : i = 1, ..., m\}$ and $\{(b_i, i) : i = 1, ..., m\}$(resp. $\{(i, a_i) : i = 1, ..., m\}$ and $\{(i, b_i) : i = 1, ..., m\}$) be the lower profile (resp. the upper profile) of A and B, respectively. Then,*

$$A \geq_L B \text{ (resp. } A \geq_U B) \quad \text{if and only if} \quad a_i \leq b_i, \quad \text{for} \quad i = 1, ..., m,$$

LEMMA 4.4. *If A is any m-by-m nonsingular matrix and P denotes its key permutation matrix, then*

$$A \geq_L P, \quad A \geq_U P.$$

Proof. If the key entries in A are nonzero numbers, then the result is trivial.

If any key entry (i, l_i) in A is zero, according to Proposition 2.8, there must be a nonzero entry in $A(i, [1, ..., l_i - 1])$ and there must be a nonzero entry in $A([1, ..., i - 1], l_i)$. Therefore, the result follows in a straightforward way. □

LEMMA 4.5. *Given two matrices A and B such that $A \geq_L B$ (resp. $A \geq_U B$) , then*

$$exc_L(A) \geq exc_L(B) \text{ (resp. } exc_U(A) \geq exc_U(B)).$$

Proof. This may be deduced directly from Definition 4.3.

□

COROLLARY 4.6. *Let A be any m-by-m nonsingular matrix and let P be the corresponding key permutation matrix. Then*

$$exc_L(A) \geq exc(P), \quad \text{and} \quad exc_U(A) \geq exc(P).$$

Proof. It is enough to consider Lemmas 4.4 and 4.5. □

LEMMA 4.7. *Let A and B be any two matrices and let $\hat{P}$ be any permutation matrix. If $A \geq_L B$ (resp. $A \geq_U B$) , then*

$$A\hat{P} \geq_L B\hat{P} \text{ (resp. } \hat{P}A \geq_U \hat{P}B).$$

Proof. We prove the first result. The second one can be proven in a similar way.

Let $\{(a_i, i) : i = 1, ..., m\}$ and $\{(b_i, i) : i = 1, ..., m\}$ be the lower profile of A and B, respectively. Assume that $\{(a_i, j_i) : i = 1, ..., m\}$ and $\{(b_i, j_i) : i = 1, ..., m\}$ are the lower profile of $A\hat{P}$ and $B\hat{P}$, respectively. Then, since $a_i \leq b_i$ for every $i = 1, ..., m$, the result follows in a straightforward way. □

In the rest of this section, we first introduce a new partial order between matrices aimed to simplify the statement and the proof of Lemma 4.15. This lemma allows us to prove Corollary 4.17. This result has an important role in the proof of the main result.

DEFINITION 4.8. *Let M be any m-by-m matrix. We define the T-lower excess of M (see 2.3 for a definition of the matrix T) as follows*

$$exc_{LT}(M) = exc_L(MT).$$

EXAMPLE 4.9. *Consider the matrix* $M = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 1 \\ 3 & 1 & 2 \end{bmatrix}$. *Then,* $MT = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix}$. *Note that* $exc_L(M) = 1$ *while* $exc_{LT}(M) = 2$. *In plain words, the lower excess measures the excess with respect to the main diagonal of* M *while the T-lower excess measures the excess with respect to the main skew diagonal.*

DEFINITION 4.10. *Let* B *and* C *be two* m*-by-*m *matrices. We say that* $B \geq_{LT} C$ *if and only* $BT \geq_L CT$.

LEMMA 4.11. *Let* B *and* C *be two* m*-by-*m *matrices such that* $B \geq_{LT} C$. *Then,* $exc_{LT}(B) \geq exc_{LT}(C)$.

Proof. This result is obtained in a straightforward way taking into account Lemma 4.5. □

For some purposes, it will also be useful to introduce a partial order in the set of m-by-1 vectors.

DEFINITION 4.12. *Let* u *and* v *be two* m*-by-*1 *vectors. Let* $u(r)$ *and* $v(s)$ *be, respectively, the leading entries of* u *and* v, *that is,* $u(r) \neq 0$, $v(s) \neq 0$, $u(i) = 0$, *for* $i < r$, *and* $v(i) = 0$ *for* $i < s$. *Then,*

$$u \geq v \quad \text{if and only if} \quad r \leq s.$$

If u *is not greater than or equal to* v, *then we say that* $u < v$.

Moreover, we define the T*-lower excess of the* k*-th column* $M(:, k)$ *of a nonsingular matrix* M *as* $exc_{LT}(M(:, k)) := \max\{0, r - m - 1 + k\}$, *where* $M(r, k)$ *is the leading entry in that column.*

LEMMA 4.13. *Let* A *and* B *be two* m*-by-*m *nonsingular matrices. If* $A(:, k) \geq B(:, k)$, *then* $exc_{LT}(A(:, k) \geq exc_{LT}(B(:, k))$.

DEFINITION 4.14. *Let* A *be any* m*-by-*m *nonsingular matrix with* LU *factorization and let* $\hat{P}$ *be any permutation matrix. Assume that the key positions of* $\hat{P}$ *occur in positions* $\{(t_i, i) : i = 1, ..., m\}$. *Then,*

- *the* i*-th column of* $\hat{P}A$ *is of type* 1 *if* $(\hat{P}A)(t_i, i) \neq 0$.
- *the* i*-th column of* $\hat{P}A$ *is of type* 2 *if* $(\hat{P}A)(t_i, i) = 0$ *and* $(\hat{P}A)([1, ..., t_i - 1], i) \neq 0$.
- *the* i*-th column of* $\hat{P}A$ *is of type* 3 *if* $(\hat{P}A)(t_i, i) = 0$ *and* $(\hat{P}A)([1, ..., t_i - 1], i) = 0$.

It is worth to make some comments about the previous definition. Notice that if A is a nonsingular matrix with LU factorization, the key positions of A are $\{(i, i) : i = 1, ..., m\}$. If the key positions of $\hat{P}$ occur in positions $\{(t_i, i) : i = 1, ..., m\}$, then the entries of $\hat{P}A$ in the key positions of $\hat{P}$ are the key entries of A, i.e., $(\hat{P}A)(t_i, i) = A(i, i)$. Next we show an example.

Let

$$\hat{P} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}. \tag{4.1}$$

Then, the matrix $\hat{P}A$ is given by

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

We have circled the key entries of A in $\hat{P}A$, which occur in the key positions of $\hat{P}$. Notice that columns $1, 4, 5$ of $\hat{P}A$ are columns of type 1, the third column is of type 2 and the second column is of type 3. In general, if the k-th column of $\hat{P}A$ is of type 1 or type 2, then

$$(\hat{P}A)(:,k) \geq \hat{P}(:,k),$$

which, according to Lemma 4.13, implies that

$$exc_{LT}(\hat{P}A)(:,k) \geq exc_{LT}(\hat{P}(:,k)).$$

However, if the k-th column of $\hat{P}A$ is of type 3, then

$$(\hat{P}A)(:,k) < \hat{P}(:,k).$$

Next we give a result whose proof is very subtle. An sketch of the proof would be the following: We show that, given an m-by-m nonsingular matrix A with an LU factorization and given any permutation matrix $\hat{P}$, if $\hat{P}A$ has columns of type 3, then it is possible to permute the columns of A in such a way that the new matrix A_m has also an LU factorization, all the columns of $\hat{P}A_m$ are of type 1 or 2, and $exc_{LT}(\hat{P}A) \geq exc_{LT}(\hat{P}A_m)$.

In the proof of this lemma, we mention the cycle decomposition of a permutation matrix. By this we mean the following: consider, for example, the permutation matrix

$$\hat{P} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

The key positions $\{(1,3),(2,5),(3,1),(4,2),(5,4)\}$ of $\hat{P}$ constitute a permutation of the set $\{1,2,3,4,5\}$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 2 & 4 \end{pmatrix}.$$

The cycle decomposition of this permutation is given by $(1,3)(2,5,4)$, which leads to the cycle decomposition of the permutation matrix $\hat{P}$ as

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

LEMMA 4.15. *Let A be any m-by-m nonsingular matrix with an LU factorization and let $\hat{P}$ be any permutation matrix. Then, there exists a sequence of matrices $A_1, ..., A_m$ satisfying the following conditions:*

1. $exc_{LT}(\hat{P}A_k) \geq exc_{LT}(\hat{P}A_{k+1})$ *for all* $k \in \{1, ..., m\}$.
2. A_k *has an LU factorization for all* $k \in \{1, ..., m\}$.
3. The first k *columns of* $\hat{P}A_k$ *are of type* 1 *or* 2.

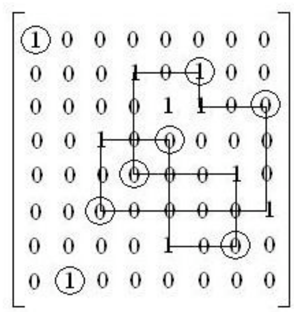
Proof. If all the columns of $\hat{P}A$ are of type 1 or 2, then $A_1 = ... = A_m = A$ and the result follows. So let us assume that there is at least one column in $\hat{P}A$ of type 3. We will prove the result by induction on k.

Let $\{(t_i, i)\}$ be the key positions of $\hat{P}$. Notice that the entries of $\hat{P}A$ in positions $\{(t_i, i)\}$ correspond to the key entries of A, i.e., the entries of A in positions $\{(i, i)\}$. Let $A_1 = A$. Since A has an LU factorization, the submatrix A_{11} is nonsingular and therefore, the first column of $\hat{P}A$ is of type 1, and A_1 satisfies conditions 2 and 3.

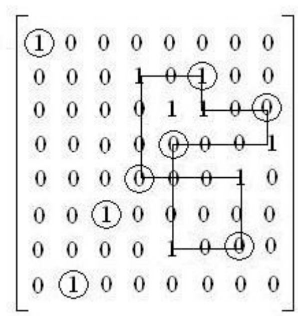
Assume that there exist matrices $A_1, ..., A_{k-1}$ satisfying the conditions in the statement of the lemma. If the k-th column of $\hat{P}A_{k-1}$ is of type 1 or 2, then $A_k = A_{k-1}$, and we are done. So let us suppose that the k-th column of $\hat{P}A_{k-1}$ is of type 3. This, in particular, implies that $A_{k-1}(k, k) = 0$. Since A_{k-1} has an LU factorization, the k-by-k leading principal submatrix of A_{k-1} is nonsingular, and there must be a nonzero term f in its determinant containing as factors a nonzero entry in $A_{k-1}(k, 1 : k-1)$ and a nonzero entry in $A_{k-1}(1 : k-1, k)$, or equivalently, in $(\hat{P}A_{k-1})(t_k, 1 : k-1)$ and $(\hat{P}A_{k-1})([t_1, ..., t_{k-1}], k)$. Consider the positions in A_{k-1} in which the factors of f occur and construct an m-by-m permutation matrix which has ones in those positions and has also ones in the positions (i, i) for $i = k+1, ..., m$. This permutation matrix can be expressed as a product of permutation cycles. It is obvious that one of these cycles (that we denote by C_{k-1}) contains both the nonzero entries in $A_{k-1}(k, 1 : k-1)$ and $A_{k-1}(1 : k-1, k)$ mentioned before. Consider the closed polygonal line that joins each nonzero entry in C_{k-1} which is not on the main diagonal with the corresponding main diagonal entries in the row and column where that nonzero entry is. Let us think in the corresponding polygonal line in $\hat{P}C_{k-1}$ (See example at the end of this proof). Now let us draw the same polygonal line on $\hat{P}A_{k-1}$. Any two consecutive edges of this line are perpendicular and connect vertices of different nature, that is, a key entry of A_{k-1} with a nonzero non-key entry (ones in C_{k-1}). Let us define as up-edges those whose bottom vertex is a key entry and, similarly, let us call down-edge to those whose bottom vertex is a non-key entry. The edge of the polygonal line on the k-th column of $\hat{P}A_{k-1}$ is a down-edge because the kth-column of $\hat{P}A_{k-1}$ is of type 3. Let us consider the up-edges in the polygonal line. It is obvious that the sum of the lengths of the up-edges is greater than or equal to the length of the down-edge on the k-th column. Then, if the bottom vertex of the down-edge on the k-th column is in position (r_k, k), consider the up-edges whose bottom vertex is below or on the r_k-th row and the top vertex is above the r_k-th row. Among all these columns pick the column with highest leading entry (s_i, i) and permute the i-th column with the k-th column. If $s_i \leq t_k$, the matrix obtained after applying the permutation is $\hat{P}A_k$. Notice that the new k-th column is of type 1 if $s_i = t_k$, or if $s_i < t_k$ and the entry in position (t_k, i) is nonzero. On the other hand, it will be of type 2 if $s_i < t_k$ and the entry in position (t_k, i) is zero. Moreover, the T-lower excess has not increased since the leading entry in the initial i-th column was higher than the leading entry in the initial k-th column. Finally, notice that vectors $(\hat{P}A_{k-1})([1, ..., t_i], i)$ and $(\hat{P}A_{k-1})([1, ..., t_i], k)$ are both nonzero and linearly independent. Then, taking into account Proposition 2.14, the matrix A_k obtained has an LU factorization.

If $s_i > t_k$, the resulting matrix after the permutation of columns i and k is again a nonsingular matrix with LU factorization and such that the first $k-1$ columns are of type 1 or 2 while the k-th column is of type 3 (See example at the end of the proof). Repeat the same process until the k-th column becomes of type 1 or 2. This process ends since the polygonal line is closed. The construction assures that the T-lower excess of the resulting matrix $\hat{P}A_k$ does not increase with respect to that of $\hat{P}A_{k-1}$. Taking into account Proposition 2.14, the final matrix A_k has an LU factorization and our claim is proven. □

EXAMPLE 4.16. *We consider a certain* 8*-by-*8 *nonsingular matrix* A *with LU factorization and a certain permutation matrix* $\hat{P}$*. Below we show the matrix* $\hat{P}A_7$ *on the left. The key positions of* $\hat{P}$ *have been circled. Notice that columns one through seven are of type* 1 *or* 2 *while the eighth column is of type* 3*. We have also drawn the closed polygonal line mentioned in the proof of Lemma 4.15. According to the proof, in order to get* $\hat{P}A_8$*, we pick the third column to be permuted with the eighth one. The matrix on the right is the one obtained after the permutation. The eighth column is still of type* 3 *although the leading entry in this column is higher than before. Therefore, we need to repeat the process. Notice that the permutation of columns* 4 *and* 8 *in the second matrix gives us* $\hat{P}A_8$*.*



Matrix $\hat{P}A_7$ *and the polygonal line*



Matrix obtained after permuting columns 3 *and* 8*.*

COROLLARY 4.17. *Let* A *be any* m*-by-*m *nonsingular matrix with LU factorization. If* $\hat{P}$ *is any permutation matrix, then*

$$exc_L(\hat{P}AT) \geq exc_L(\hat{P}T).$$

Proof. Taking into account Lemma 4.15, there exists a matrix A_m such that

$$exc_{LT}(\hat{P}A) \geq exc_{LT}(\hat{P}A_m).$$

Moreover, since $\hat{P}A_m$ only has columns of type 1 and 2,

$$\hat{P}A_m \geq_{LT} \hat{P}.$$

Taking into account Lemma 4.11,

$$exc_{LT}(\hat{P}A_m) \geq exc_{LT}(\hat{P})$$

and the result follows. □

The role of the matrix T in the statement of Corollary 4.17 may look superfluous. However, as we show with the next example, it is not true, in general, that

$$exc_L(\hat{P}A) \geq exc_L(\hat{P}).$$

Just consider the 2-by-2 matrices $\hat{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. It is easy to check that $exc_L(\hat{P}A) = 0$ while $exc_L(\hat{P}) = 1$.

Acknowledgments. The first author thanks Prof. Roberto Costas for his useful comments to improve the presentation of this paper.

The authors also thank an anonymous referee for all his suggestions to improve this paper.

REFERENCES

[1] W. Barrett and C. R. Johnson *Spanning-tree extensions of the Hadamard-Fischer inequalities*, Linear Algebra and its appl. 66(1985), 177–193.

[2] I. S. Duff, A. M. Erisman, and J.K. Reid *Direct methods for sparse matrices*, Oxford University Press, New York, 1989.

[3] L. Elsner, *On some algebraic problems in connection with general eigenvalue algorithms*, Linear Algebra and its applications, 26(1979), 123–138.

[4] C. A. Felippa, Solution of linear equations with skyline-stored symmetric matrix. Computers & Structures, 5(1975), 13–29.

[5] Gelfand, I. M.; Naimark, M. A. Unitarnye predstavleniya klassičeskih grupp. (Russian) [Unitary representations of the classical groups.] Trudy Mat. Inst. Steklov., vol. 36, Izdat. Nauk SSSR, Moscow-Leningrad, 1950. 288 pp.

[6] I. Gohberg, and S. Goldberg, *Finite dimensional Wiener-Hopf equations and factorizations of matrices*, Linear Algebra and its Applications, 48(1982), 219–236.

[7] R. A. Horn, and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1999.

[8] C. R. Johnson, and P. Okunev, *Necessary and sufficient conditions for LU factorability*, William and Mary NSF-REU report, 1997.

[9] B.N. Parlett, *The new qd algorithms*, Acta Numerica, 459–491(1995).