SIAM J. MATRIX ANAL. APPL.
Vol. 30, No. 4, pp. 1520–1545

SYMBOLIC AND EXACT STRUCTURE PREDICTION FOR SPARSE GAUSSIAN ELIMINATION WITH PARTIAL PIVOTING*

LAURA GRIGORI†, JOHN R. GILBERT‡, AND MICHEL COSNARD§

Abstract. In this paper we consider two structure prediction problems of interest in Gaussian elimination with partial pivoting of sparse matrices. First, we consider the problem of determining the nonzero structure of the factors L and U during the factorization. We present an exact prediction of the structure that identifies some numeric cancellations appearing during Gaussian elimination. The numeric cancellations are related to submatrices of the input matrix A that are structurally singular, that is, singular due to the arrangements of their nonzeros, and independent of their numerical values. Second, we consider the problem of estimating upper bounds for the structure of L and U prior to the numerical factorization. We present tight exact bounds for the nonzero structure of L and U of Gaussian elimination with partial pivoting $PA = LU$ under the assumption that the matrix A satisfies a combinatorial property, namely, the Hall property, and that the nonzero values in A are algebraically independent of each other. This complements existing work showing that a structure called the row merge graph represents a tight bound for the nonzero structure of L and U under a stronger combinatorial assumption, namely, the strong Hall property. We also show that the row merge graph represents a tight symbolic bound for matrices satisfying only the Hall property.

Key words. sparse LU factorization, partial pivoting, structure prediction, characterization of fill

AMS subject classifications. 65F50, 65F05, 68R10

DOI. 10.1137/050629343

1. Introduction.

In this paper we consider the problem of structure prediction when solving a linear system $Ax = b$ by Gaussian elimination with partial pivoting, where A is an $n \times n$ sparse, nonsingular, and nonsymmetric matrix and b is an n-vector. This elimination, also called LU factorization, involves explicit factorization of the matrix A into the product of L and U, where L is a unit lower triangular matrix and U is an upper triangular matrix.

One of the main characteristics of the sparse LU factorization is the notion of fill. "Fill" denotes a nonzero entry in the factors that was a zero in matrix A. When Gaussian elimination without pivoting is used, the nonzero structure of the factors can be computed without referring to the numerical values of the matrix and is determined before performing the numerical computation of the factors themselves. Knowledge of this structure is used to allocate memory, set up data structures, schedule parallel tasks, and save time [16] by avoiding operations on zeros. When pivoting is used for numerical stability [13], the structure of L and U depends not only on the structure of A but also on the row interchanges. As the row interchanges are determined while doing the numerical factorization, the computation of the structure of the factors has to be interleaved with the computation of the numerical values of the factors. Prior to the numerical factorization, only upper bounds of the structure of L and U can be determined.

*Received by the editors April 18, 2005; accepted for publication (in revised form) by E. Ng July 17, 2008; published electronically December 19, 2008.
http://www.siam.org/journals/simax/30-4/62934.html
†INRIA Saclay - Ile-de-France, Laboratoire de Recherche en Informatique, Paris-Sud 11 University, Bat 490, 91405 Orsay, France (Laura.Grigori@inria.fr).
‡Computer Science Department, University of California at Santa Barbara, Santa Barbara, CA 93106-5110 (gilbert@cs.ucsb.edu).
§INRIA Sophia Antipolis, 2004 route des Lucioles BP93, 06902 Sophia Antipolis, France (Michel.Cosnard@inria.fr).

We discuss in this paper two structure prediction problems. The first problem considers the computation of the nonzero structure of the factors during Gaussian elimination with row interchanges. The second problem is to obtain tight bounds of the structure of L and U prior to the numerical factorization. For both problems, we study relations between the combinatorial properties of the nonzero structure of the matrix A and the LU factorization.

Two kinds of structure prediction and two combinatorial properties of the input matrix are usually considered for these problems. The two structure predictions are called *symbolic* and *exact* [11]. *Symbolic* structure prediction assumes that the addition or subtraction of two nonzero results always yields a nonzero result. It ignores possible numeric cancellations occurring during the LU factorization. *Exact* structure prediction assumes that the nonzero values in A are algebraically independent from each other; in other words, it assumes that any computed zero is due to combinatorial properties of the nonzero structure. The two combinatorial properties of the input matrix are called the *strong Hall* property and the *Hall* property. The *strong Hall* property is an irreducibility condition. The *Hall* property is a weaker combinatorial assumption and is related to matrices with full-column rank. We will define these two properties in more detail later in the paper. A matrix that satisfies the Hall property can be decomposed using the Dulmage–Mendelsohn decomposition [2, 17, 18] into a block upper triangular form such that every block on the diagonal satisfies the strong Hall property. However, in practice this decomposition is not always used, and hence it is interesting to understand the structure prediction for matrices satisfying either the strong Hall property or the Hall property.

Much of the research has been aimed at predicting the structure and bounds of the factors L and U as tightly as possible [9, 10, 11, 12, 20]. The existing results for determining the nonzero structure of L and U during Gaussian elimination with partial pivoting $PA = LU$ are symbolic [20]. Under several additional conditions, this structure prediction is exact [11]. But in general it ignores possible numeric cancellations during the factorization for matrices satisfying the strong Hall property or only the Hall property. For the problem of predicting bounds for the structure of L and U prior to the numerical factorization, the existing results in the literature assume that A satisfies the strong Hall property. The results assume the LU factorization with partial pivoting is seen as

$$A = P_1 L_1 P_2 L_2 \dots P_{n-1} L_{n-1} U,$$

where P_i is an $n \times n$ elementary permutation matrix identifying the row interchanges at step i. L_i is an $n \times n$ elementary lower triangular matrix whose ith column contains the multipliers at step i. U is an $n \times n$ upper triangular matrix. $\widetilde{L}$ is the $n \times n$ matrix whose ith column is the ith column of L_i so that $\widetilde{L} - I = \sum_i (L_i - I)$. Note that this $\widetilde{L}$ is not the same as the factor L obtained from the factorization $PA = LU$. Both matrices are unit lower triangular, and they contain the same nonzero values but in different positions. The factor L has its rows in the order described by the entire row permutations. The factor $\widetilde{L}$ has the rows of its ith column in the order described by only the first i row interchanges. George and Ng [9] predict an upper bound of the nonzero structure of $\widetilde{L}$ and U, called the row merge graph, that contains the nonzeros in $\widetilde{L}$ and U for all possible row permutations which can later appear in the numerical factorization due to pivoting. Gilbert and Ng [11] showed that this is a tight exact bound for a square matrix with nonzero diagonal which satisfies the strong Hall property.

In this paper we provide answers to several open questions related to the two structure prediction problems considered here. For the first problem, we identify the exact structure prediction of L and U during LU factorization with partial pivoting. For the second problem, we describe the exact bounds of the factors obtained from the factorization $PA = LU$, when the matrix A satisfies only the Hall property. These exact bounds are not symbolic bounds. Then we show that the row merge graph represents symbolic bounds for the structure of $\widetilde{L}$ and U.

The exact structure prediction is based on the following approach: all of the elements of the factors L and U can be computed using the determinants of two submatrices of the input matrix A (see, for example, Gantmacher [8]). Consider, for example, the element in position (i, j) of U, where i and j are two indices with $i \leq j$. Let A_{i-1} be the submatrix of A formed by the first $(i-1)$ columns and the first $(i-1)$ rows of A. Let K be the $i \times i$ submatrix of A that includes the first i rows, the first $i-1$ columns, and column j of A. Then the value in position (i, j) of the factor U is given by the quotient of the determinant of K and the determinant of A_{i-1}. A similar relation exists for the elements of L. Our new results identify when the submatrix K is structurally singular, that is, singular due to the arrangements of its nonzeros, and independent of the numerical values. In exact arithmetic, the determinant of K is zero, and hence the element in position (i, j) corresponds to a numeric cancellation. This numeric cancellation is identified in our new results on exact structure prediction. However, in a backward stable factorization $A + E = \hat{L}\hat{U}$, the computed factors $\hat{L}$ and $\hat{U}$ are not necessarily close to the exact $A = LU$ factors, even though the norm of E is small. In particular, a zero in L or U may, in principle, be large in $\hat{L}$ or $\hat{U}$, so rounding it to zero may cause backward instability.

The rest of the paper is organized as follows. In section 2 we present background and several new results used throughout the paper. In section 3 we consider the problem of determining the nonzero structure of the factors L and U during Gaussian elimination with partial pivoting. We present new results that give an exact characterization of the fill occurring in the LU factorization. We show how the theoretical results can be used in an algorithm for computing fill-ins.

In sections 4 and 5 we consider the problem of predicting bounds for the structure of L and U prior to the numerical factorization. In section 4 we present an exact analysis for matrices that satisfy the Hall property. We present tight exact bounds for the nonzero structure of L and U of Gaussian elimination with partial pivoting $PA = LU$. In section 5 we present a symbolic analysis, and we show that the row merge graph is a lower symbolic bound for the factors $\widetilde{L}$ and U of the factorization $A = P_1 L_1 P_2 L_2 \ldots P_{n-1} L_{n-1} U$. In other words, for every edge of the row merge graph of a Hall matrix, there is a permutation such that this edge corresponds to a symbolic nonzero in $\widetilde{L}$ or U. By a simple counterexample, we will show that the row merge graph is not a tight bound for the factors L and U in the exact sense. These results are of practical interest since the row merge graph is used by several solvers implementing the sparse LU factorization with partial pivoting. In solvers like the sequential and shared memory versions of SuperLU [5, 6], the row merge graph is used to estimate the memory needs prior to the LU factorization. In solvers proposed in [9, 21], the numerical computation of the factors L and U is performed on the row merge graph, and some operations involve zero elements. Finally, section 6 presents concluding remarks.

2. Graphs of matrices and their properties.

In this section we provide the necessary notions to study the structure prediction of the sparse LU factorization with partial pivoting. We give definitions, previously published results, and two new results (Lemmas 2.6 and 2.7) that are needed by our subsequent proofs.

Let A be a sparse $n \times n$ matrix. A_{ij} denotes the element at row i and column j of A. We refer to the determinant of matrix A as $\det(A)$. We denote the submatrix of A formed by elements of row indices from i to j and column indices from d to e as $A(i\colon j, d\colon e)$. When the indices are not consecutive, we use the following notation: $A([i\colon j, k], d\colon e)$ denotes the submatrix of A formed by elements of row indices from i to j and k and column indices from d to e. We refer to the submatrix $A(1\colon i, 1\colon i)$ as the principal minor of order i of A.

Two graphs are used to predict the nonzero structure of the factors L and U from the structure of A. The first graph is the directed graph of A and is denoted by $G(A)$. This graph has n vertices and an edge $\langle i, j\rangle$ for each nonzero element A_{ij}. We say that the edge $\langle i, j\rangle$ is incident on the vertices i and j.

The second graph is the bipartite graph of A, denoted by $H(A)$. This graph is undirected and has n row vertices, n column vertices, and an edge $\langle i', j\rangle$ if and only if the element A_{ij} is nonzero. Note that whenever possible, we use prime to distinguish between row vertices and column vertices in a bipartite graph. Also we use $i, j, k, d,$ and e to denote a vertex of H for which it is known if it is a column or a row vertex. That is, i' stands for a row vertex and i for a column vertex. We use v and w to denote a generic vertex of H, that is, a vertex that can be a row vertex or a column vertex.

A *path* is a sequence of distinct vertices $\mathcal{Q} = (v_0, v_1, \ldots, v_{q-1}, v_q)$ such that for each two consecutive vertices v_i, v_{i+1} there is an edge from v_i to v_{i+1}. The length of this path is q. The vertices $v_1, \ldots, v_{q-1}$ are called intermediate vertices.

Let H be a bipartite graph with m row vertices and n column vertices. A matching M on H is a set of edges, no two of which are incident on the same vertex. A vertex is covered or matched by M if it is an end point of an edge of M. A matching is called column-complete if it has n edges, row-complete if it has m edges, and perfect if $m = n$ and it is both row- and column-complete. Given a graph H and a column vertex i, we denote by $H - i$ the subgraph of H induced by all of the row vertices and all of the column vertices except i.

The next lemma identifies a matching in the bipartite graph H of A such that if the edges of M become the diagonal elements, the values chosen make the permuted matrix strongly diagonally dominant. It will be used in section 4 to prove our results on exact structure prediction for Hall matrices.

LEMMA 2.1 (Gilbert and Ng [11]). *Suppose the bipartite graph H has a perfect matching M. Let A be a matrix with $H(A) = H$ such that $A_{ij} > n$ for $\langle i', j\rangle \in M$ and $0 < A_{ij} < 1$ for $\langle i', j\rangle \notin M$. If A is factored by Gaussian elimination with partial pivoting, then the edges of M will be the pivots.*

If M is a matching on H, an alternating path with respect to M is a path on which every second edge is an element of M. A c-alternating path is a path that follows matching edges from rows to columns. An r-alternating path is a path that follows matching edges from columns to rows. Suppose the last vertex of one c-alternating path is the first vertex of another c-alternating path. The path obtained by their concatenation is also a c-alternating path. The same result holds for r-alternating paths. Suppose $\mathcal{Q}$ is an alternating path from an unmatched vertex v to a different vertex w. If the last vertex w on $\mathcal{Q}$ is unmatched or the last edge on $\mathcal{Q}$ belongs to M, then a new matching M_1 can be obtained from M by alternating along path $\mathcal{Q}$. The set of edges of M_1 is given by $M \oplus \mathcal{Q} = (M \cup \mathcal{Q}) - (M \cap \mathcal{Q})$. If w is matched by M, then v is matched and w is unmatched by M_1 and $|M_1| = |M|$. If w is unmatched by M, then both v and w are matched by M_1, $|M_1| = |M| + 1$, and $\mathcal{Q}$ is called an augmenting path with respect to M.

2.1. Hall and strong Hall graphs. We briefly review the Hall and the strong Hall properties and related results. For a detailed description of Hall and strong Hall matrices and their properties, the reader is directed to [2, 3, 11].

A bipartite graph with m rows and n columns has the *Hall property* if every set of k column vertices is adjacent to at least k row vertices, for all $1 \leq k \leq n$. The next theorem and corollary relate the Hall property to column-complete matchings and matrices with full-column rank. In Corollary 2.3 [11] it is shown that if H is Hall and given a matrix A with $H = H(A)$, then the set of ways to fill in its values to make it singular has measure zero. Hence almost all matrices A with $H = H(A)$ have full-column rank.

THEOREM 2.2 (Hall's theorem). *A bipartite graph has a column-complete matching if and only if it has the Hall property.*

COROLLARY 2.3 (Gilbert and Ng [11]). *If a matrix A has full-column rank, then $H(A)$ is Hall. Conversely, if H is Hall, then almost all matrices A with $H = H(A)$ have full-column rank.*

Known results in structure prediction were obtained under an additional assumption, called the strong Hall property. A bipartite graph with m rows and $n \leq m$ columns satisfies the *strong Hall property* if

(i) $m = n > 1$ and every set of k column vertices is adjacent to *more than* k row vertices, for all $1 \leq k < n$, or

(ii) $m > n$ and every set of k column vertices is adjacent to *more than* k row vertices, for all $1 \leq k \leq n$.

LEMMA 2.4 (Gilbert and Ng [11]). *If H is strong Hall and has more nonzero rows than columns and M is any column-complete matching on H, then from every row or column vertex v of H there is a c-alternating path to some unmatched row vertex i' (which depends on v and M).*

The next theorem relates alternating paths and matchings in strong Hall graphs. This theorem was used in several structure prediction results, in the context of sparse LU factorization by Gilbert and Ng in [11], as well as in the sparsity analysis of QR factorization by Coleman, Edenbrandt, and Gilbert in [4] and Hare et al. in [15]. In this paper we will use it in Lemma 2.6 to derive a new result on alternating paths and matchings in strong Hall graphs.

THEOREM 2.5 (alternating paths, Gilbert [12]). *Let H be a strong Hall graph with at least two rows, let i be a column vertex of H, and let v be any row or column vertex of H such that a path exists from i to v. Then H has a column-complete matching for which there exists a c-alternating path from i to v (or, equivalently, an r-alternating path from v to i).*

The next lemma is new. Given a path in a bipartite graph H between a column vertex and a row vertex or between two row vertices, the lemma shows that there is an alternating path with respect to a column-complete matching of H which excludes a row vertex at the extremity of the path. We will use it in sections 3 and 4 to estimate the nonzero structure of the factors L and U.

LEMMA 2.6. *Let H be a strong Hall graph with more nonzero rows than columns, let v be a row or column vertex of H, and let i' be any row vertex of H such that a path exists from v to i'. Then H has a column-complete matching which excludes vertex i' and for which there exists a c-alternating path from v to i'.*

Proof. We distinguish two different cases.

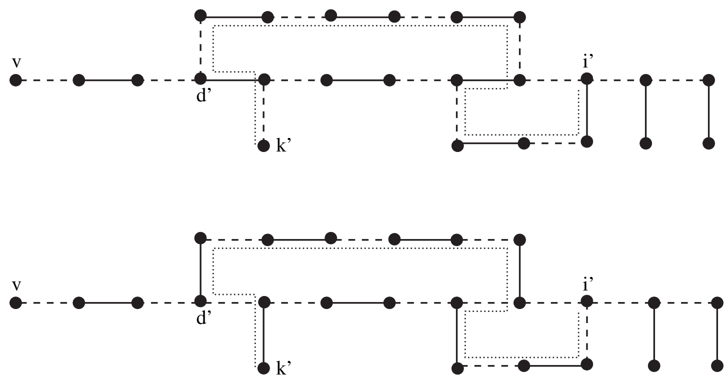

FIG. 2.1. *Case* 1 *of Lemma* 2.6. *In the upper graph, the solid edges are the matching M; path $\mathcal{P}$ is the horizontal path from v to i'; path $\mathcal{Q}$ is the light dotted line from i' to k'. In the lower graph, the solid edges are the matching M_1. The path obtained by concatenating $\mathcal{P}[v\colon d']$ and $\mathcal{Q}[d'\colon i']$ is c-alternating with respect to M_1.*

Case 1 (v is a column vertex). By hypothesis, there is a path from v to i'. As H is strong Hall, the alternating path Theorem 2.5 applies and says that H has a column-complete matching M for which there exists a c-alternating path $\mathcal{P}$ from v to i'. If i' is not covered by M, then M is the column-complete matching searched. Otherwise, Lemma 2.4 implies that there is an unmatched row vertex k' and a c-alternating path $\mathcal{Q}$ from i' to k'. Now obtain matching M_1 from M by alternating along path $\mathcal{Q}$, where i' is unmatched in M_1.

If $\mathcal{P}$ and $\mathcal{Q}$ have no vertices in common (except row vertex i'), then $\mathcal{P}$ is still c-alternating from v to i' with respect to M_1. If the only vertex in common for $\mathcal{P}$ and $\mathcal{Q}$ (except row vertex i') is column vertex v, then let e' be the row vertex matched by M to v that belongs to the path $\mathcal{Q}$. The path formed by $\langle v, e'\rangle$ followed by $\mathcal{Q}[e'\colon i']$ is c-alternating with respect to M_1.

If $\mathcal{P}$ and $\mathcal{Q}$ have intermediate vertices in common, let d' be the first (row) vertex of $\mathcal{P}$ (starting from v) which belongs to $\mathcal{Q}$. The path obtained by the concatenation of $\mathcal{P}[v\colon d']$ and $\mathcal{Q}[d'\colon i']$ is c-alternating with respect to M_1, and this ends the proof for this case. This case is illustrated in Figure 2.1.

Case 2 (v is a row vertex). We denote the row vertex v as v'. By hypothesis, there is a path from v' to i'. Suppose $v' \neq i'$; otherwise there is nothing to prove. Let d be the first column vertex on this path, that is, the next vertex after v'. H is a strong Hall graph that has a path from column vertex d to row vertex i'. The first case of this theorem, that we have just proved, says that there is a column-complete matching M that excludes vertex i' and for which there exists a c-alternating path $\mathcal{P}$ from d to i'. We distinguish four cases.

Case 2.1 (v' is not matched by M). Let e' be the row vertex matched by M to the column vertex d. We obtain a new matching M_1 by unmatching row vertex e' and matching row vertex v' to row vertex d. The path formed by $\langle v', d\rangle$ followed by $\mathcal{P}$ is c-alternating from v' to i' with respect to M_1. Note that M_1 excludes row vertex i', and this is the path searched.

Case 2.2 (v' is matched by M to the column vertex d). The path obtained by $\langle v', d\rangle$ followed by $\mathcal{P}$ is c-alternating from v' to i' with respect to the matching M, and the matching M excludes row vertex i'.

Case 2.3 (v' is matched by M and belongs to the path $\mathcal{P}$). Then $\mathcal{P}[v'\colon i']$ is a c-alternating path with respect to the matching M.

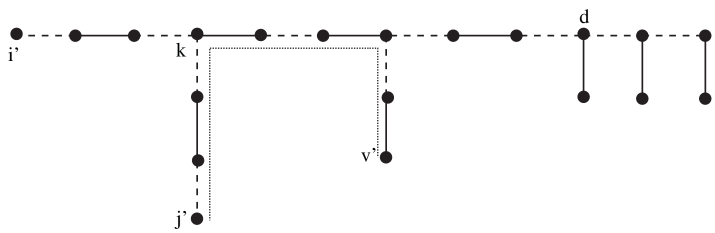

FIG. 2.2. *Case* 2.4 *of Lemma* 2.6. *The solid edges are the matching M; path $\mathcal{P}$ is the horizontal path from i' to d; path $\mathcal{R}$ is the light dotted line from v' to j'. Here, $\mathcal{P}$ and $\mathcal{R}$ have at least one common vertex. Vertex k is the last vertex on $\mathcal{P}$ (starting from d) that belongs to $\mathcal{R}$. The path obtained by concatenating $\mathcal{R}[v'\!:k]$ and $\mathcal{P}[k\!:i']$ is c-alternating with respect to M, and M excludes row vertex i'.*

Case 2.4 (v' is matched by M to a different column vertex than d and does not belong to the path $\mathcal{P}$). Lemma 2.4 applies and says that there is an unmatched row vertex j' and a c-alternating path $\mathcal{R}$ from v' to j'.

If $\mathcal{P}$ and $\mathcal{R}$ have no vertices in common, then obtain matching M_1 from M by alternating along path $\mathcal{R}$. As v' is matched in M, then v' is unmatched in M_1 and j' is matched in M_1. From here we proceed as in Case 2.1, and we obtain a matching that excludes vertex i' and with respect to which there is a c-alternating path from v' to i'.

If $\mathcal{P}$ and $\mathcal{R}$ have at least one vertex in common, then let k be the last vertex of $\mathcal{P}$ (starting from d) which belongs to $\mathcal{R}$. Note that k has to be a column vertex. The path obtained by concatenating $\mathcal{R}[v'\!:k]$ and $\mathcal{P}[k\!:i']$ is c-alternating with respect to M, M excludes the row vertex i', and this ends the proof. This case is illustrated in Figure 2.2. ∎

2.2. Hall sets and their properties. For a bipartite graph H with m row vertices and $n \leq m$ column vertices, a set of k column vertices, $1 \leq k \leq n$, forms a *Hall set* if these columns are adjacent to exactly k rows [15].

Under the assumption that A satisfies the Hall property, the union of two Hall sets is a Hall set, so there exists a unique Hall set of maximum cardinality in any given set of columns. The set of maximum cardinality might be empty. Let C_j be the Hall set of maximum cardinality in the first j columns; we define $C_0 = \emptyset$. Let R_j be the set of all row indices covered by the columns of C_j; thus C_j and R_j have the same cardinality. Note that if we assume all diagonal entries of A are nonzero, then $R_j = \{i' : 1 \leq i \leq j \textit{ and } i \in C_j\}$.

The Hall sets of maximum cardinality are useful to partition a Hall graph into two subgraphs: one that satisfies the Hall property and another one that satisfies the strong Hall property. Let H be a bipartite graph with m row vertices and $n < m$ column vertices that satisfies the Hall property. Let C be the Hall set of maximum cardinality in H, and let R be the set of row vertices covered by column vertices of C. The first subgraph $\widetilde{H}$ is induced by all of the row vertices in R and all of the column vertices in C. This subgraph satisfies the Hall property. The second subgraph $\widehat{H}$ is induced by all of the row vertices except those in R and all of the column vertices except those in C. This subgraph is strong Hall because its Hall set of maximum cardinality is empty.

In a similar way, we can partition the edges of a column-complete matching M of H into edges belonging to the graph $\widetilde{H}$ and edges belonging to the graph $\widehat{H}$. This is expressed in a more general way in the following lemma.

LEMMA 2.7. *Let A be an $m \times n$ Hall matrix, $m \geq n$. Let C be a Hall set of cardinality p in A, where $p \leq n$, and let R be the set of all row indices covered by the columns of C. Suppose M is a column-complete matching in the bipartite graph $H(A)$. Then each column vertex j of C is matched by M to a row vertex i' of R.*

Proof. The proof is immediate. ∎

3. Nonzero structure of *L* and *U* during Gaussian elimination with partial pivoting. Let A be an $n \times n$ nonsingular matrix. In this section we consider the problem of determining the nonzero structure of the factors L and U during Gaussian elimination with partial pivoting. In the first part of this section we consider the LU factorization without pivoting. We first present a brief overview of several well-known results described in the literature. Then we describe why these results ignore numeric cancellations related to submatrices of A that are structurally singular. In section 3.1 we present new results that identify some numeric cancellation occurring during Gaussian elimination and caused by submatrices of A that are structurally singular. In section 3.2 we describe how the new results can be used in the Gaussian elimination with partial pivoting. We also present an algorithm that uses the new results to compute the nonzero structure of the factors L and U.

The main result in the structure prediction of Gaussian elimination without pivoting is the fill path Lemma 3.1. This lemma relates paths in the directed graph $G(A)$ and the nonzero elements that appear in the factors L and U, represented in the so-called filled graph $G^+(A)$.

LEMMA 3.1 (fill path (Rose and Tarjan [20])). *Let G be a directed or undirected graph whose vertices are the integers 1 through n, and let G^+ be its filled graph. Then $\langle i, j\rangle$ is an edge of G^+ if and only if there is a path in G from i to j whose intermediate vertices are all smaller than* $\min(i, j)$.

The filled graph $G^+(A)$ represents a symbolic bound for the factors L and U; that is, it ignores possible numeric cancellation during the factorization. The next lemma represents an example of conditions under which this structure prediction is exact, by taking into account the values of the nonzeros in the matrix. In this lemma, a square Hall submatrix of A denotes a square submatrix of A which satisfies the Hall property and which is formed by a subset of rows and columns of A that can be different and noncontiguous.

LEMMA 3.2 (Gilbert and Ng [11]). *Suppose A is square and nonsingular and has a triangular factorization $A = LU$ without pivoting. Suppose also that all of the diagonal elements of A, except possibly the last one, are nonzero and that every square Hall submatrix of A is nonsingular. Then $G(L+U) = G^+(A)$; that is, every nonzero predicted by the filled graph of A is actually nonzero in the factorization.*

We are interested in fill when the diagonal may contain zeros (perhaps due to pivoting), but Lemma 3.2 does not hold in this case. An example showing this was given by Brayton, Gustavson, and Willoughby [1]. We give a slightly different example in Figure 3.1, where we display a matrix A, its bipartite graph $H(A)$, and its directed graph $G(A)$. Note that $H(A)$ satisfies the strong Hall property. Since there is a path from 5 to 4 through lower numbered vertices in $G(A)$, the edge $\langle 5, 4\rangle$ belongs to the filled graph $G^+(A)$, but $L_{54} = 0$ regardless of the nonzero values of A. That is because after the first step of elimination, the elements in column positions 2 and 4 of the rows 2 and 5 are linearly dependent. At the second step of elimination the element L_{54} is zeroed.

A simpler way of understanding this numeric cancellation is to consider the two submatrices $A([1\colon 3, 5], 1\colon 4)$ and $A(1\colon 3, 1\colon 3)$ and their determinants that determine

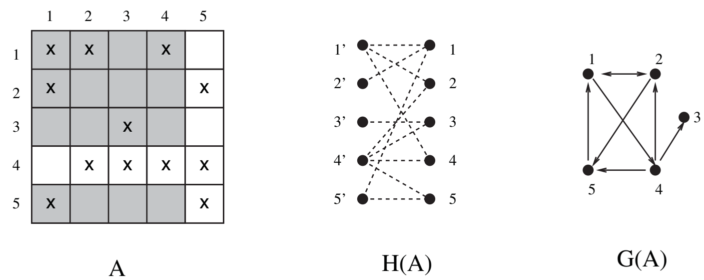

FIG. 3.1. *Example showing that the fill path Lemma* 3.2 *does not predict exactly the nonzero structure of L and U when factorizing without pivoting the strong Hall matrix A. Details are given in the text following Lemma* 3.1.

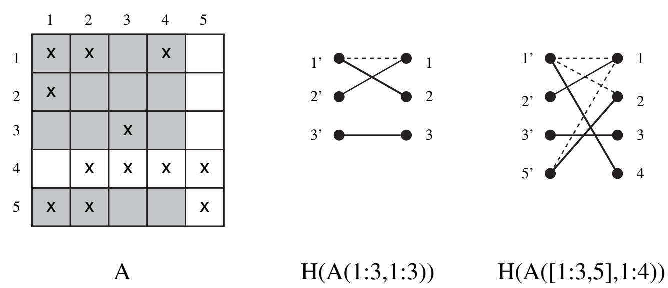

FIG. 3.2. *Example for Theorem* 3.4 $(1 \rightarrow 2)$. *Consider the strong Hall matrix A and the matrix* $B = A([1\colon 3,5], 1\colon 4)$ *displayed in patterned gray. The Hall set of maximum cardinality of* $H(A([1\colon 3,5],1\colon 3))$ *is* $C_{35} = \{3\}$ *and* $R_{35} = \{3'\}$. *Suppose element* L_{54} *is nonzero. The perfect matching* M_k *of matrix* $A_k = A(1\colon 3, 1\colon 3)$ *is formed by the edges* $\langle 1', 2\rangle$, $\langle 2', 1\rangle$, *and* $\langle 3', 3\rangle$. *The perfect matching* M_B *of B is formed by the edges* $\langle 1', 4\rangle$, $\langle 2', 1\rangle$, $\langle 3', 3\rangle$, *and* $\langle 5', 2\rangle$. *Form a path by starting at* $5'$ *and by following one edge in* M_B *and one edge in* M_k. *This yields the path* $(5', 2, 1', 4)$.

the value of L_{54}. The submatrix $A([1\colon 3,5], 1\colon 4)$ (displayed in light gray in Figure 3.1) has three columns $(2, 3,$ and $4)$ with nonzero elements in only two rows (1 and 3). This submatrix does not satisfy the Hall property, and its determinant is zero. This is the approach we use to identify some numeric cancellations in the LU factorization.

The following lemma describes the above observation. Assuming that the LU factorization exists, this lemma relates the value of an element of the factors L and U to the singularity of a submatrix of A.

LEMMA 3.3 (Gilbert and Ng [11]). *Suppose A is square and nonsingular and has a triangular factorization* $A = LU$ *without pivoting. Let i be a row index and j a column index of A, and let B be the submatrix of A consisting of rows* 1 *through* $\min(i,j)-1$ *and i, and columns* 1 *through* $\min(i,j)-1$ *and j. Then* $(L+U)_{ij}$ *is zero if and only if B is singular.*

3.1. New results. Theorem 3.4 is the first new result of this section and provides necessary and sufficient conditions, in terms of paths in the bipartite graph $H(A)$ for a fill element to occur in exact arithmetic during Gaussian elimination. It is illustrated in Figures 3.2 and 3.3. Consider the nonzero structure of L. Suppose that the factorization exists until the step $j-1$ of factorization; that is, the principal minor of order $j-1$ is nonzero. The theorem uses the fact that L_{ij} is nonzero if and only if the determinant of the submatrix $A([1\colon j-1, i], 1\colon j)$ is nonzero.

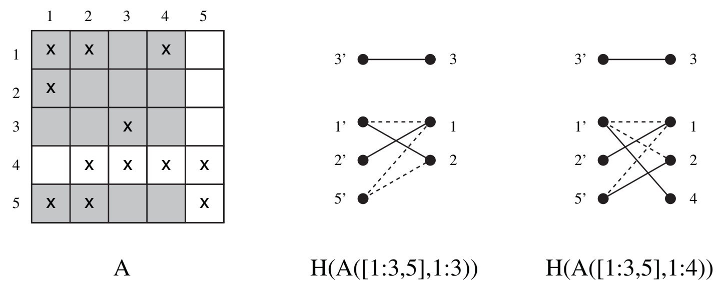

FIG. 3.3. *Example for Theorem* 3.4 ($3 \to 1$). *Consider the strong Hall matrix A as in Figure* 3.2, *the matrix $B = A([1{:}3,5],1{:}4)$, $C_{35} = \{3\}$, and $R_{35} = \{3'\}$. Consider the path $\mathcal{Q} = (5',1,1',4)$ that has no vertex in $C_{35} \cup R_{35}$. The graph $H(A([1{:}3,5],1{:}3))$ is partitioned into two subgraphs. The subgraph induced by column vertex 3 and row vertex $3'$ satisfies the Hall property, and it has a perfect matching $\widetilde{M} = \{\langle 3',3\rangle\}$. The subgraph induced by column vertices $\{1,2\}$ and row vertices $\{1',2',5'\}$ satisfies the strong Hall property and has a column-complete matching $\widehat{M} = \{\langle 1',2\rangle,\langle 2',1\rangle\}$ for which there is a c-alternating path $\mathcal{R} = (1',2,5')$. The matching M is formed by the edges of $\widetilde{M}$ and $\widehat{M}$ and is presented by solid edges in the graph $H(A([1{:}3,5],1{:}3))$. The matching obtained by alternating along path $\mathcal{R}$ is a perfect matching of $H(B)$ and is presented at the right of the figure.*

THEOREM 3.4. *Let A be an $n \times n$ nonsingular matrix that has a triangular factorization $A = LU$. Suppose that every square Hall submatrix of A is nonsingular. Let i be a row of A, j be a column of A, and $k = \min(i,j) - 1$. Let M_k be a perfect matching of $A(1{:}k,1{:}k)$. Let C_{ki} be the Hall set of maximum cardinality in $H(A([1{:}k,i],1{:}k))$ and R_{ki} be the set of all row indices covered by columns of C_{ki}. Then the following three statements are equivalent:*

1. *$(L+U)_{ij}$ is nonzero.*
2. *There is an r-alternating path in the bipartite graph $H(A)$ from row vertex i' to column vertex j with respect to the matching M_k.*
3. *There is a path in the bipartite graph $H(A)$ from i' to j whose intermediate vertices are smaller than or equal to k and that has no vertex in $C_{ki} \cup R_{ki}$.*

Proof. Let A_k be the leading $(k \times k)$ principal submatrix of A and $\det(A_k)$ be its determinant. As we suppose the factorization exists, $\det(A_k)$ is nonzero. This implies that A_k satisfies the Hall property and has a perfect matching M_k. The matching M_k also represents a column-complete matching in the graph $H(A([1{:}k,i],1{:}k))$. Lemma 2.7 applies with respect to the graph $H(A([1{:}k,i],1{:}k))$ and the Hall set C_{ki} and says that each row vertex of R_{ki} has to be matched by M_k to one of the column vertices in C_{ki}. Since i' is not a row vertex matched by M_k, then $i' \notin R_{ki}$.

Let B be the submatrix of A consisting of columns 1 through k and j and rows 1 through k and row i. Suppose edge $\langle i',j\rangle$ does not belong to $H(A)$; otherwise the proof is trivial. We will prove now that the three statements are equivalent.

$1 \to 2$. As $(L+U)_{ij}$ is nonzero by hypothesis, then Lemma 3.3 applies and shows that B is a nonsingular matrix. Hence its bipartite graph $H(B)$ satisfies the Hall property, and there is a perfect matching M_B in $H(B)$.

Consider now the row vertex i' in the bipartite graph $H(A)$. Recall we assume that edge $\langle i',j\rangle$ does not belong to $H(A)$. Row vertex i' is matched by M_B to column vertex j_0. Since $i' \notin R_{ki}$, we can deduce that $j_0 \notin C_{ki}$. Column vertex j_0 is matched by M_k to some row vertex i'_0, where $i'_0 \neq i'$ since i' is not matched by M_k. Also we have that $i'_0 \notin R_{ki}$. Row vertex i'_0 is matched by M_B to some column vertex j_1, where $j_1 \neq j_0$ since j_0 is matched in M_B to i'. If $j_1 = j$, then we stop. Otherwise, we continue our reasoning. For each row vertex we consider its matched column vertex by M_B;

then for each column vertex we consider its matched row vertex by M_k. Continuing inductively, we arrive at vertex j. The vertices followed during our reasoning are vertices $i', j_0, i'_0, j_1, i'_1, \dots, i'_t, j$. Edge $\langle i', j_0 \rangle$ and edges $\langle i'_q, j_{q+1} \rangle$ are edges of $H(B)$ which belong to the perfect matching M_B. Edge $\langle i'_t, j \rangle$ and edges $\langle j_q, i'_q \rangle$ are edges of $H(A_k)$ which belong to the perfect matching M_k. This yields a path in $H(A)$ from row vertex i' to column vertex j that is r-alternating with respect to the matching M_k.

$2 \rightarrow 3$. Consider the r-alternating path $(i', j_0, i'_0, j_1, i'_1, \dots, i'_t, j)$ from i' to j with respect to the matching M_k. All of the intermediate vertices on this path are smaller than or equal to k. Because $i' \notin R_{ki}$, we can deduce that $j_0 \notin C_{ki}$. Continuing inductively, we can deduce that this path does not include any vertex in $C_{ki} \cup R_{ki}$.

$3 \rightarrow 1$. Let d' be the last row vertex on $\mathcal{Q}$, that is, the vertex just before j on $\mathcal{Q}$. We partition the graph $H(A([1\colon k, i], 1\colon k))$ into two subgraphs. The first subgraph, induced by the row vertices in R_{ki} and the column vertices in C_{ki}, satisfies the Hall property and has a perfect matching $\widetilde{M}$. The second subgraph, induced by the row vertices $1, \dots k'$ and i', except row vertices in R_{ki}, and the column vertices 1 through k, except column vertices in C_{ki}, is strong Hall. Lemma 2.6 says that there is a column-complete matching $\widehat{M}$ which excludes row vertex i' and for which there exists a c-alternating path $\mathcal{R}$ from d' to i'.

Let the matching M be formed by the edges of $\widehat{M}$ and the edges of $\widetilde{M}$. This matching represents a column-complete matching in $H(A([1\colon k, i], 1\colon k))$. We now show that the graph $H(B)$ satisfies the Hall property. Recall that column vertex j and row vertex i' are not matched by the matching M. Consider path $\mathcal{R}$ from d' to i' that is c-alternating with respect to matching M. Obtain a new matching $M \oplus \mathcal{R}$ from M by alternating along path $\mathcal{R}$. As i' is not matched in M and d' is matched in M, then i' is matched in $M \oplus \mathcal{R}$ and d' is not matched in $M \oplus \mathcal{R}$. Add to matching $M \oplus \mathcal{R}$ the edge $\langle d', j \rangle$.

Thus we obtain a perfect matching in $H(B)$; that is, $H(B)$ satisfies the Hall property. By hypothesis, every square Hall submatrix of A is nonsingular, and thus B is nonsingular and its determinant is nonzero. Therefore $(L+U)_{ij}$ is nonzero. ∎

The next theorem uses Hall sets of maximum cardinality associated with subsets of columns of A to restrict paths corresponding to nonzero elements of L and U. In this paper we use this theorem in section 4 to determine upper bounds for the factorization $PA = LU$, where the matrix A satisfies only the Hall property. Note that for a matrix satisfying the strong Hall property, the Hall set of maximum cardinality of a subset of columns is always empty. Thus Theorem 3.5 is relevant to matrices satisfying only the Hall property. This theorem can also be useful in the algorithm described in section 3.2. The Hall sets involved can be computed prior to the factorization using an algorithm as, for example, the one proposed in [15].

THEOREM 3.5. *Let A be an $n \times n$ nonsingular matrix that is factored by Gaussian elimination as $A = LU$. Suppose that $(L+U)_{ij}$ is nonzero. Let $k = min(i, j) - 1$, and let C_k be the Hall set of maximum cardinality in the first k columns and R_k be the set of all row indices covered by columns of C_k. Then there is a path in the bipartite graph $H(A)$ from row vertex i' to column vertex j whose intermediate vertices are smaller than or equal to k and that has no vertex in $C_k \cup R_k$.*

Proof. Let C_{ki} be the Hall set of maximum cardinality in $H(A([1\colon k, i], 1\colon k))$ and R_{ki} be the set of all row indices covered by columns of C_{ki}. It can be easily shown that $C_k \subseteq C_{ki}$ and $R_k \subseteq R_{ki}$. The third statement of Theorem 3.4 implies that this theorem holds. ∎

Note that Theorem 3.5 provides only a necessary condition for fill to occur during the elimination. Figure 3.4 (as well as Theorem 3.4) shows that the condition is

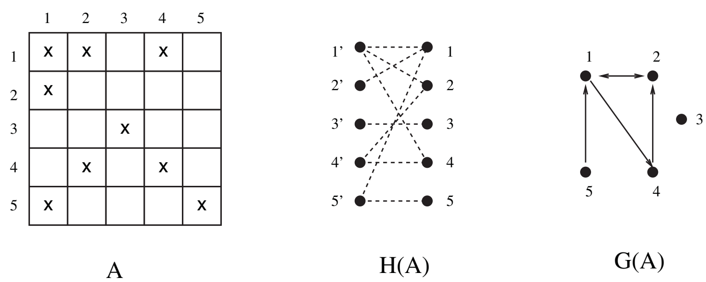

FIG. 3.4. *Example showing that the converse of Theorem* 3.5 *is not true.*

not sufficient. Consider the Hall matrix A in Figure 3.4. The Hall set of maximum cardinality is $C_3 = \{3\}$, and it covers the row index $R_3 = \{3'\}$. The Hall set of maximum cardinality in $H(A([1\colon 3, 5], 1\colon 3))$ is $C_{35} = \{2, 3\}$, and it covers the row indices $R_{35} = \{1', 3'\}$. There is a path $(5', 1, 1', 4)$ in $H(A)$ that has no vertex in C_3. However, the element $L_{54} = 0$ because of numeric cancellation.

3.2. Computing the nonzero structure of the factors *L* and *U* during Gaussian elimination with partial pivoting. In this section we present an algorithm that uses the results of the previous section to compute the nonzero structure of the factors L and U during the LU factorization with partial pivoting. The algorithm computes one column of L and one row of U at a time.

First, we present Theorem 3.6 that describes explicitly how Theorem 3.4 can be used during the LU factorization with partial pivoting of a matrix A. This theorem supposes that the first $j-1$ steps of the LU factorization exist, and it gives the necessary results to compute the structure of column j of L and of row j of U at the jth step of factorization.

THEOREM 3.6. *Let A be an $n \times n$ nonsingular matrix that is to be decomposed using LU factorization with partial pivoting. Suppose that the first $j-1$ steps of LU factorization with partial pivoting of A exist and have been executed. Let $P_{J-1} = P_{j-1}P_{j-2}\dots P_1$ be the permutations performed during the first $j-1$ steps of elimination, and let M_{j-1} be a perfect matching of $(P_{J-1}A)(1\colon j-1, 1\colon j-1)$. Suppose that every square Hall submatrix of A is nonsingular. At the jth step of decomposition, the element L_{ij} is nonzero if and only if there is a c-alternating path in the bipartite graph $H(P_{J-1}A)$ from column vertex j to row vertex i' with respect to the matching M_{j-1}. The element U_{ji} is nonzero if and only if there is an r-alternating path in the bipartite graph $H(A)$ from row vertex j' to column vertex i with respect to the matching M_{j-1}.*

Proof. The proof is similar to the proof of Theorem 3.4. ∎

Algorithm 1 uses Theorem 3.6 and sketches the factorization $PA = LU$, where $P = P_{n-1}\dots P_1$ and each P_j reflects the permutation of two rows at step j of factorization. At each step j, the structure of column j of L is determined, and then its numerical values are computed. The element of maximum magnitude in column j of L is chosen as the pivot. Let L_{kj} be this element. The algorithm interchanges rows k and j of L and rows k and j of A. Then the structure of row j of U is determined, followed by the computation of its numerical values.

The structure of column j of L is computed by finding all of the c-alternating paths with respect to M_{j-1} from column vertex j to some row vertex i'. This can be achieved in a similar way to the augmenting path technique, used in finding maximum matchings in bipartite graphs and described, for example, in [7]. This technique ensures that each edge of the bipartite graph of A is traversed at most once. The structure of row i of U is computed in a similar way. Since the jth diagonal element corresponds to a nonzero, Theorem 3.4 ensures that there is a c-alternating path $\mathcal{Q}$ from column vertex j to row vertex j' with respect to the matching M_{j-1}. During the computation of the structure of column j of L, we store for each row vertex i' the column vertex just before i' on a c-alternating path with respect to M_{j-1} from j to i'. This allows us to retrace $\mathcal{Q}$. The algorithm computes a new matching M_j by alternating along path $\mathcal{Q}$.

The overall complexity of computing the structure of L and the structure of U in Algorithm 1 is hence bounded by $O(n \cdot nnz(A))$, where n is the order and $nnz(A)$ is the number of nonzeros of matrix A.

Algorithm 1. LU factorization with partial pivoting, aware of some cancellations

---

$M_0 = \emptyset$
for $j := 1$ to n **do**
  if $j < n$ **then**
    1. Compute structure of $L(j\colon n, j)$. This is formed by all row vertices $i' \geq j$ such that there is a c-alternating path in $H(A)$ with respect to M_{j-1} from column vertex j to row vertex i'.
    2. Compute numerical values of $L(j\colon n, j)$.
    3. Find k such that $|L_{kj}| = \max |L(j\colon n, j)|$. Let $v = L_{kj}$.
    4. Interchange $L(j, :)$ with $L(k, :)$ and $A(j, :)$ with $A(k, :)$. Let $\mathcal{Q}[j\colon j']$ be the c-alternating path in $H(A)$ with respect to M_{j-1} that corresponds to L_{jj}.
    5. Scale: $L(:, j) = L(:, j)/v$.
  end if
  6. Compute structure of $U(j, j+1\colon n)$. This is formed by all column vertices $i \geq j$ such that there is an r-alternating path in $H(A)$ from row vertex j' to column vertex i with respect to the matching M_{j-1}.
  7. Compute numerical values of $U(j, j+1\colon n)$. Let $U_{jj} = v$.
  if $j = 1$ **then**
    $M_1 = \mathcal{Q}$
  else
    $M_j = M_{j-1} \oplus \mathcal{Q}$
  end if
end for

Several aspects need to be investigated and remain as open questions. The first important aspect is related to the practical interest of using this algorithm, which depends on the utility of identifying numeric cancellations and on the number of numeric cancellations that appear in real world applications. The second aspect is related to the complexity of Algorithm 1, which is equivalent to the complexity of one of the first algorithms for computing the structure of the factors L and U, denoted as the FILL2 algorithm in [20]. The algorithms proposed more recently for computing fill-ins [10] are faster in practice than FILL2. Since we expect Algorithm 1 to have a similar run time to FILL2, further investigation is required to make it competitive with respect to the new algorithms.

4. Tight exact bounds for the structure prediction of $PA = LU$, when A satisfies only the Hall property.

Let A be an $n \times n$ matrix that satisfies the Hall property. Suppose A is factored by Gaussian elimination with row interchanges as $PA = LU$. In this section we discuss the problem of predicting bounds for the factors L and U prior to the numerical factorization. We consider exact results; that is, the upper bounds do not include elements that correspond to numeric cancellations due to submatrices of A structurally singular.

The next three theorems give tight exact bounds for the nonzero structure of the factors L and U. Theorem 4.1 gives upper bounds for the structure of L and U in terms of paths in the bipartite graph $H(A)$. Theorems 4.2 and 4.3 show that this bound is the tightest possible for Gaussian elimination with row interchanges of a matrix that satisfies the Hall property. That is, for every predicted element of the upper bound, there is a permutation and a choice of the values of matrix A such that this element corresponds to a nonzero in the factors L or U.

THEOREM 4.1. *Let A be an $n \times n$ nonsingular matrix that is factored by Gaussian elimination with row interchanges as $PA = LU$. Let i be an index, j be a column index, and $q = \min(i, j) - 1$. Let C_q be the Hall set of maximum cardinality in the first q columns and R_q be the set of all row indices covered by columns of C_q. If L_{ij} is nonzero, then there is a path in the bipartite graph $H(A)$ from row vertex k' to column vertex j whose intermediate column vertices are all in $\{1, \ldots, q\}$ and that has no vertex in $C_q \cup R_q$, where k is the row of A that corresponds to row i of PA. If U_{ij} is nonzero, then there is a path in the bipartite graph $H(A)$ from column vertex i to column vertex j whose intermediate column vertices are all in $\{1, \ldots, q\}$ and that has no vertex in $C_q \cup R_q$.*

Proof of Case 1 ($i \geq j$ (*structure of* L)). Due to Theorem 3.5, there is a path $\mathcal{Q}$ in $H(A)$ from row vertex k' to column vertex j whose intermediate column vertices are all in $\{1, \ldots, j-1\}$ and that has no vertex in $C_{j-1} \cup R_{j-1}$. This is the path searched in the theorem. □

Proof of Case 2 ($i < j$ (*structure of* U)). According to Theorem 3.5, there is a path $\mathcal{Q}$ in $H(A)$ from row vertex k' to column vertex j whose intermediate column vertices are all in $\{1, \ldots, i-1\}$ and that has no vertex in $C_{i-1} \cup R_{i-1}$.

By hypothesis, the factorization exists; thus the ith diagonal element of PA is nonzero. Theorem 3.5 applies with respect to this element and says that there is a path $\mathcal{R}$ in $H(A)$ from column vertex i to row vertex k' whose intermediate column vertices are all in $\{1, \ldots, i-1\}$ and that has no vertex in $C_{i-1} \cup R_{i-1}$.

Using the path $\mathcal{R}$ and the path $\mathcal{Q}$, we can form a path in $H(A)$ from column vertex i to column vertex j whose intermediate column vertices are all in $\{1, \ldots, i-1\}$ and that has no vertex in $C_{i-1} \cup R_{i-1}$. This is the path searched in the theorem. □

The next two theorems show that the upper bound defined in Theorem 4.1 for the structure of L and U is tight. First, Theorem 4.2 shows that the bound for the structure of L is tight, and it is illustrated in Figure 4.1. Second, Theorem 4.3 shows that the bound for U is tight, and it is illustrated in Figure 4.2.

The bound for L depends on the row permutations of A. It considers every row i of the original matrix A. The bound identifies all column indices j that correspond to elements of row i that can become potentially nonzeros during the factorization through permutations. The bound for U is independent of row permutations of A. It identifies potential nonzeros U_{ij} using paths that relate column vertex i to column vertex j in the bipartite graph of A. None of the results assumes that the input matrix A has a zero-free diagonal.

THEOREM 4.2. *Let H be the structure of a square Hall matrix. Let j be a column vertex, C_{j-1} be the Hall set of maximum cardinality in the first $j-1$ columns, R_{j-1} be the set of row indices covered by columns in C_{j-1}, and i' be any row vertex not in R_{j-1}. Suppose that H contains a path from i' to j whose intermediate column*

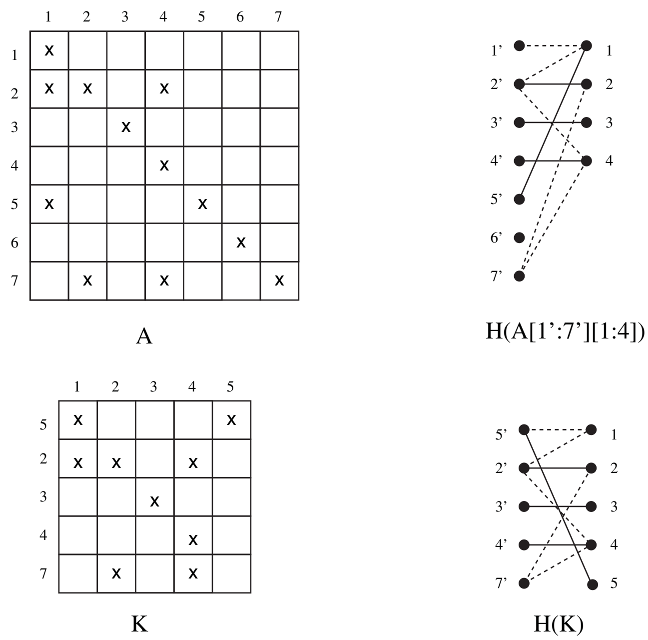

FIG. 4.1. *Example for Theorem* 4.2 *showing the construction that makes element* L_{75} *nonzero for the Hall matrix* A *at the top left. The last row vertex on the path* $(7', 4, 2', 1, 5', 5)$ *between* $i' = 7'$ *and* $j = 5$ *satisfying the conditions in Theorem* 4.2 *is* $e' = 5'$*. The graph* H_{j-1}*, presented at the top right, is the subgraph of* H *induced by column vertices* 1 *through* $j-1 = 4$ *and all of the row vertices. The solid edges represent a column-complete matching* M_{j-1} *that excludes row vertex* $7'$ *and with respect to which there is a c-alternating path* $\mathcal{R} = (5', 1, 2', 2, 7')$ *from* $5'$ *to* $7'$*. At the bottom right,* K *is the submatrix of* PA *with columns* 1 *through* $j = 5$ *and the rows in corresponding positions after four steps of pivoting. The fifth row of* K *is* $k' = i' = 7'$*. In* $H(K)$ *there is a maximum matching* $M_{j-1} \oplus \mathcal{R}$ *represented by solid edges at the bottom right. Thus the element* L_{75} *is nonzero.*

vertices are all in $\{1, \ldots, j-1\}$ *and which has no vertex in* $C_{j-1} \cup R_{j-1}$*. There exists a nonsingular matrix* A *with* $H(A) = H$ *and a permutation matrix* P *such that if* A *is factored by Gaussian elimination with row interchanges as* $PA = LU$*, then row* i *of* A *is permuted in some row position* k *of* PA*,* $k \geq j$ *and* $L_{kj} \neq 0$*.*

Proof. By hypothesis, there is a path in H from i' to j whose intermediate column vertices are all at most j. Consider H_{j-1} the subgraph of H induced by all row vertices and all column vertices from 1 to $j-1$. The graph H satisfies the Hall property, and hence H_{j-1} also satisfies the Hall property. We obtain a column-complete matching M_{j-1} in this graph which will induce the pivoting order for the first $j-1$ steps of elimination. We partition the graph H_{j-1} into two subgraphs. The first subgraph $\widetilde{H}_{j-1}$ satisfies the Hall property and is induced by all of the row vertices in R_{j-1} and all of the column vertices in C_{j-1}. Let $\widetilde{M}_{j-1}$ be a perfect matching in this subgraph. The second subgraph $\widehat{H}_{j-1}$ satisfies the strong Hall property and is induced by all of the row vertices except row vertices in R_{j-1} and all of the column vertices 1 through $j-1$ except column vertices in C_{j-1}. Let $\widehat{M}_{j-1}$ be a column-complete matching in this subgraph.

We distinguish two cases to determine $\widehat{M}_{j-1}$, depending on if $\langle i', j \rangle$ is an edge of $H(A)$ or not. First, assume that $\langle i', j \rangle$ is an edge of $H(A)$. Lemma 2.4 says that for any column-complete matching M of $\widehat{H}_{j-1}$ there is a c-alternating path $\mathcal{R}$ from i' to some unmatched row vertex. We denote by $\widehat{M}_{j-1}$ the matching obtained from

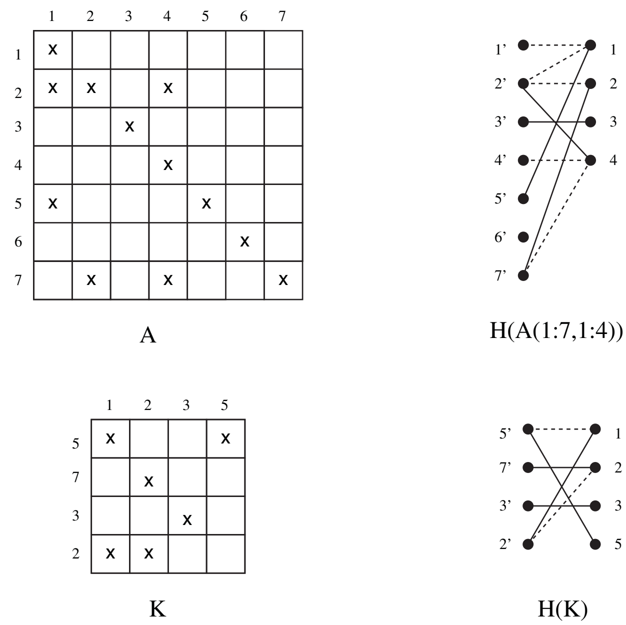

FIG. 4.2. *Example for Theorem 4.3 showing the construction that makes element U_{45} nonzero for the Hall matrix A at the top left. The Hall set of maximum cardinality in the first three columns is $C_3 = \{3\}$, $R_3 = \{3'\}$. Consider the path $\mathcal{Q} = (5, 5', 1, 2', 4)$ satisfying the conditions in Theorem 4.3 ($k' = 5'$ and $e' = 2'$). The graph $\widehat{H}_4$, presented at the top right, is the subgraph of H induced by column vertices 1 through $i = 4$ and all of the row vertices. The solid edges represent a column complete-matching M_4 that is formed by the edge $\langle 2', 4\rangle$, the matching $\widehat{M}_3$ (formed by the edges $\langle 5', 1\rangle, \langle 7', 2\rangle, \langle 2', 4\rangle$), and the matching $\widetilde{M}_3$ (formed by the edge $\langle 3', 3\rangle$). This matching determines the pivoting order for the first four steps of elimination. With respect to the matching $\widehat{M}_3$ there is a c-alternating path $\mathcal{R} = (5', 1, 2')$. Consider the matrix $K = A([5', 7', 3', 2'][1:3, 5])$ presented at the bottom left and its graph presented at the bottom right. The perfect matching M is presented by solid edges in the graph $H(K)$. Since K satisfies the Hall property and the minor of order-3 of PA is nonzero, then the element U_{45} is nonzero.*

M by alternating along path $\mathcal{R}$. With this choice, row vertex i' is not covered by the matching $\widehat{M}_{j-1}$. Second, assume that $\langle i', j\rangle$ is not an edge of $H(A)$. Let e' be the last row vertex on the path between i' and j, that is, the vertex just before j. Therefore Lemma 2.6 applies and says that there is a column-complete matching $\widehat{M}_{j-1}$ which excludes vertex i' and for which there exists a c-alternating path $\mathcal{R}$ from e' to i'.

Let the column-complete matching M_{j-1} be formed by the edges of $\widehat{M}_{j-1}$ and the edges of $\widetilde{M}_{j-1}$. We choose the values of A such that every square submatrix of A that is Hall, including A itself, is nonsingular. We can say that this is possible by using an argument as the one described in [11] (the determinant of a Hall submatrix is a polynomial in its nonzero values, not identically zero, since the Hall property implies a perfect matching). We choose the values of the nonzeros of A corresponding to edges of M_{j-1} to be larger than n and the values of the other nonzeros of A to be between 0 and 1. With this choice, Lemma 2.1 says that the first $j-1$ steps of elimination of A pivot on nonzeros corresponding to edges of M_{j-1}. Let P be the permutation matrix that describes these row interchanges.

Note that with our choice of M_{j-1}, row vertex i' is not covered by the matching M_{j-1}. Thus, after the first $j-1$ steps of elimination, row i of A was moved to a row in position k of PA, where $k \geq j$. We prove that this choice makes L_{kj} nonzero. If $\langle i', j\rangle$

is an edge of $H(A)$, then L_{kj} is nonzero. Otherwise, let K be the $j \times j$ submatrix of A that includes the first j columns and the rows 1 to $j-1$ in corresponding positions of PA and row i of A (that is, row k of PA). Thus the columns of K are those numbered 1 through j in $H(A)$. The first $j-1$ columns are matched by M_{j-1}, while the last column j is not matched by M_{j-1}. The first $j-1$ rows of K are those matched to columns 1 through $j-1$ of $H(A)$ by M_{j-1}. The last row of K is row number i of A.

To show that L_{kj} is nonzero, we still need to show that K satisfies the Hall property. Recall that column vertex j and row vertex i' are not matched by the matching M_{j-1} in $H(K)$. Consider path $\mathcal{R}$ from e' to i' that is c-alternating with respect to matching M_{j-1}. Obtain a new matching $M_{j-1} \oplus \mathcal{R}$ from M_{j-1} by alternating along path $\mathcal{R}$. As i' is not matched in M_{j-1} and e' is matched in M_{j-1}, then i' is matched in $M_{j-1} \oplus \mathcal{R}$ and e' is not matched in $M_{j-1} \oplus \mathcal{R}$. Add to matching $M_{j-1} \oplus \mathcal{R}$ the edge $\langle e', j\rangle$, and thus we get a perfect matching in $H(K)$; that is, $H(K)$ satisfies the Hall property. By our choice of values, every submatrix that satisfies the Hall property is nonsingular. Therefore element L_{kj} is nonzero. ∎

THEOREM 4.3. *Let H be the structure of a square Hall matrix. Let i and j be two column vertices, $i < j$, let C_{i-1} be the Hall set of maximum cardinality in the first $i-1$ columns, and let R_{i-1} be the row vertices covered by columns in C_{i-1}. Suppose that H contains a path from j to i whose intermediate column vertices are all in $\{1, \dots, i-1\}$ and that has no vertex in $C_{i-1} \cup R_{i-1}$. There exists a nonsingular matrix A with $H(A) = H$ and a permutation matrix P such that if A is factored by Gaussian elimination with row interchanges as $PA = LU$, then U_{ij} is nonzero.*

Proof. By hypothesis, there is a path $\mathcal{Q}$ in $H(A)$ from column vertex j to column vertex i whose intermediate column vertices are all at most $i-1$. Let k' be the first row vertex on $\mathcal{Q}$, that is, the vertex just after j on $\mathcal{Q}$. Let e' be the last row vertex on $\mathcal{Q}$, that is, the vertex just before i on $\mathcal{Q}$. Note that e' can be equal to k'.

Let $\widehat{H}_{i-1}$ be the strong Hall subgraph of H induced by all of the row vertices except row vertices in R_{i-1} and all of the column vertices 1 through $i-1$ except column vertices in C_{i-1}. Lemma 2.6 says that there is a column-complete matching $\widehat{M}_{i-1}$ which excludes e' and for which there exists a c-alternating path $\mathcal{R}$ from k' to e'. (If $k' = e'$, then $\mathcal{R}$ is empty.) Let $\widetilde{H}_{i-1}$ be the subgraph of $H(A)$ induced by all of the row vertices in R_{i-1} and all of the column vertices in C_{i-1}. The graph $\widetilde{H}_{i-1}$ satisfies the Hall property, and Lemma 2.6 says that there is a perfect matching $\widetilde{M}_{i-1}$ in $\widetilde{H}_{i-1}$.

Consider H_i the subgraph of H induced by all of the row vertices and all of the column vertices 1 through i. The matching M_i formed by the edge $\langle e', i\rangle$, all of the edges of $\widehat{M}_{i-1}$, and all of the edges of $\widetilde{M}_{i-1}$ is a column-complete matching in H_i . We choose the values of A such that every square submatrix of A that is Hall, including A itself, is nonsingular. We set the values of the nonzeros of A corresponding to edges of M_i to be larger than n and the values of the other nonzeros of A to be between 0 and 1. With this choice the first i steps of elimination of A pivot on nonzeros corresponding to edges of M_i (Lemma 2.1). Let P be the permutation matrix that describes these row interchanges.

We prove that this pivoting choice makes U_{ij} nonzero. Let K be the submatrix $PA(1{:}i, [1{:}i-1, j])$. To show that U_{ij} is nonzero, we need to show that the graph $H(K)$ satisfies the Hall property. For this, consider again the matching $\widehat{M}_{i-1}$ and the c-alternating path $\mathcal{R}$ from k' to e'. Consider the path formed by the edge $\langle j, k'\rangle$ followed by $\mathcal{R}$, and consider the matching M obtained by alternating along this path. Since k' is matched by $\widehat{M}_{i-1}$ and j is unmatched by $\widehat{M}_{i-1}$, then both k' and j are matched by M, and its cardinality is $|\widehat{M}_{i-1}| + 1$. We add to matching M the edges of

$\widetilde{M}_{i-1}$. Thus M is a perfect matching in $H(K)$; that is, this matrix satisfies the Hall property, and its determinant is nonzero. This shows that U_{ij} is nonzero. ◻

We make one final note on the similarities between the exact structure prediction presented in this section and the sparsity analysis of the QR factorization for square matrices satisfying the Hall property. The structure prediction for the QR factorization of matrices satisfying only the Hall property was studied by Hare et al. in [15] and Pothen in [19]. It can be easily shown that the structure of Q represents a tight exact bound for the structure of L of the factorization $PA = LU$ and that the structure of R is a tight exact bound for the structure of U obtained from Gaussian elimination with row interchanges.

5. The row merge graph and structure prediction for $A = P_1L_1 \ldots P_{n-1}L_{n-1}U$. Let A be an $n \times n$ matrix with nonzero diagonal that satisfies the Hall property. Suppose A is factored by Gaussian elimination with row interchanges as $A = P_1L_1P_2L_2 \ldots P_{n-1}L_{n-1}U$ and $\widetilde{L}$ is the union of the L_i. An upper bound for the nonzero structure of $\widetilde{L}$ and U was proposed by George and Ng [9]. This upper bound, called the row merge graph, contains the nonzeros in the factors for all possible row permutations that can later appear in the numerical factorization due to pivoting. In this section we discuss the row merge graph as an upper bound for the nonzero structure of the factors $\widetilde{L}$ and U when the matrix A satisfies only the Hall property. Thus we extend the work of Gilbert and Ng who showed in [11] that the row merge graph is a tight upper bound for Gaussian elimination with row permutations of strong Hall matrices.

First, we consider an exact analysis; that is, we assume only that the nonzero values in A are algebraically independent of each other. By a simple counterexample we show that for matrices satisfying only the Hall property, the row merge graph is not a tight bound for the factors $\widetilde{L}$ and U in the exact sense. This means that the row merge graph predicts as nonzero elements of $\widetilde{L}$ and U that during the actual factorization are zeroed. Second, we relax the condition on the numerical values of the nonzeros of A by considering a symbolic analysis. This is a weaker analysis than the exact analysis performed in section 4, since we ignore the possibility of numeric cancellation during the factorization. With this assumption, we show that the row merge graph is a tight bound for the factors $\widetilde{L}$ and U. In other words, for every edge of the row merge graph of a Hall matrix, there is a permutation such that this edge corresponds to a symbolic nonzero in the factors $\widetilde{L}$ or U.

5.1. Existing results. The row merge graph was proposed by George and Ng [9] as an upper bound for the nonzero structure of $\widetilde{L}$ and U and is obtained as follows: at each step of elimination an upper bound of the structure of $\widetilde{L}$ and U is computed. Consider step i and all of the rows that are candidates to pivoting at this step. An upper bound of their structure is given by the union of their structures. Thus the structure of each row candidate to pivoting is replaced by this union. The bipartite graph that contains all of the edges of the upper bound of $\widetilde{L}$ and U is called the *row merge graph*, denoted by $H^{\times}(A)$. The matrix containing a nonzero element for each edge of $H^{\times}(A)$ is referred to as the row merge matrix of A, denoted as $A^{\times}$. Several results in the literature use a directed version of the row merge graph, denoted as $G^{\times}(A)$ or $G^{\times}(H)$. This graph has n vertices and an edge for each nonzero of $A^{\times}$. The next theorem proves the claim that the row merge graph is an upper bound for the structure of $\widetilde{L}$ and U.

THEOREM 5.1 (George and Ng [9]). *Let A be a nonsingular square matrix with nonzero diagonal. Suppose Gaussian elimination with row interchanges is performed*

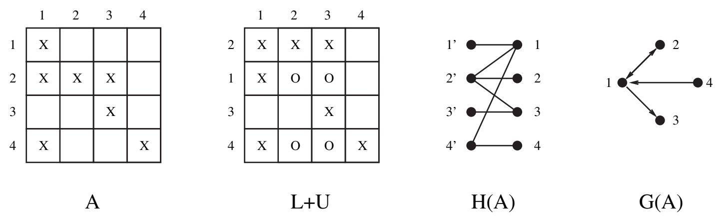

FIG. 5.1. *Example matrix A showing that the row merge graph is not an exact tight bound. The nonzero elements of A are denoted by* x*, and the fill elements of the factors $\widetilde{L}$ and U (corresponding to edges of $G^+(PA)$) are denoted by* o.

as $A = P_1L_1 \dots P_{n-1}L_{n-1}U$, and let $\widetilde{L}$ be the union of the L_i. Then

$$G(L+U) \subseteq G^\times(A).$$

When the matrix satisfies the strong Hall property, Gilbert and Ng [11] showed that this graph represents a tight exact bound for the structure of $\widetilde{L}$ and U. That is, having a strong Hall graph H, for every edge $\langle i', j\rangle$ in its row merge graph $H^\times$, there exists a nonsingular matrix A (depending on i' and j) with $H(A) = H$ such that the element in position (i, j) of $L+U$ is nonzero. Nothing is known for the case when the matrix satisfies only the Hall property, and this question is the subject of this section.

5.2. The row merge graph and counterexample for tight exact bounds. In Figure 5.1 we give a counterexample showing that the row merge graph is not tight in the exact sense. The edge $\langle 4', 3\rangle$ is an edge of the row merge graph $H^\times(A)$. We present a permutation that makes the entry in position $(4, 3)$ nonzero in the factor $\widetilde{L}$. At the first step of elimination we pivot on the element at position $(2, 1)$, while at the next steps of elimination we pivot on the diagonal. Let P be the matrix describing these permutations. The directed graph $G(PA)$ has a path $(4, 1, 3)$; therefore the element in position $(4, 3)$ fills in. Then the $\langle 4, 3\rangle$ entry in $G^+(PA)$ is nonzero, but $\widetilde{L}_{43} = 0$, regardless of the nonzero values of A. Note that there is no choice of pivot at the first step of elimination that fills the element at position $(4, 3)$. We conclude that there is no permutation that makes the element $\widetilde{L}_{43}$ nonzero.

5.3. The row merge graph as a tight symbolic bound. We now discuss a symbolic analysis; that is, we ignore the possibility of numeric cancellation during the factorization. With this assumption, we show that the row merge graph is a tight bound for the factors $\widetilde{L}$ and U.

An example of the construction of the row merge matrix is presented in Figure 5.2. At the first step of elimination, rows 1, 4, and 5 are candidates to pivoting. The union of their structure is formed, and it replaces the structure of each one of these rows. This is repeated at each step on the trailing matrix.

Row merge fill elements refer to elements that are zero in the original matrix A but are nonzero in the row merge matrix $A^\times$. Similarly, *row merge fill edges* refer to those edges that don't belong to $H(A)$ but belong to the row merge graph $H^\times(A)$. The row merge fill edges in the row merge graph $H^\times(A)$ are related to paths in the bipartite graph $H(A)$ by Definition 5.2 and Theorem 5.3.

DEFINITION 5.2 (Gilbert and Ng [11]). *A path $\mathcal{Q} = (i', j_1, i'_1, j_2, i'_2, \dots, j_t, i'_t, j)$ in $H(A)$ is a row merge fill path for LU elimination with partial pivoting if either*

	1	2	3	4	5	6	7	8	9	10
1	x	ō	x	ō	ō					
2		x	ō	ō	ō	ō		ō		ō
3			x	ō	ō	ō		ō	ō	ō
4	x	x	ō	x	ō	ō	ō	ō	ō	ō
5	x	ō	ō	x	x	ō	ō	ō	ō	ō
6						x	ō	ō	ō	ō
7				x	x	ō	x	ō	ō	ō
8								x	ō	ō
9			x	ō	ō	x	ō	ō	x	ō
10		x	ō	x	ō	x	ō	x	ō	x

FIG. 5.2. *Example to illustrate the construction of a row merge matrix* $A^{\times}$*, Theorems* 5.4 *and* 5.5. *The nonzero elements of* A *are denoted by* x, *and the row merge fill elements are denoted by* $\bar{\text{o}}$.

$t = 0$ *or the following conditions are satisfied:*

1. $j_k < j$ *and* $j_k \leq i'$ *for all* $1 \leq k \leq t$.
2. *Let* j_p *be the largest* j_k. *Then there is some* q *with* $p \leq q \leq t$, $j_p \leq i'_q \leq n$, *and the three paths* $\mathcal{Q}[i'\colon j_p]$, $\mathcal{Q}[j_p\colon i'_q]$, *and* $\mathcal{Q}[i'_q\colon j]$ *are also row merge fill paths in* $H(A)$.

The next theorem due to Gilbert and Ng gives a necessary and sufficient condition for fill to occur in the row merge graph $H^{\times}(A)$.

THEOREM 5.3 (Gilbert and Ng [11]). *For two vertices* i', j *of the bipartite graph* $H(A)$, *the edge* $\langle i', j\rangle$ *is an edge of* $H^{\times}(A)$ *if and only if there is a row merge fill path joining* i' *and* j *in* $H(A)$.

We present two algorithms that use Definition 5.2 to decompose a row merge fill path in paths and edges of the bipartite graph $H(A)$. Algorithm 2 decomposes the row merge fill path $\mathcal{Q}[i'\colon j]$ in subpaths by recursively applying Definition 5.2. The recursivity is stopped when a path is reduced to an edge. Its aim is to record for each intermediate column vertex j_p its corresponding row vertex of the middle path $\mathcal{Q}[j_p\colon i'_q]$ (that is, $MC[j_p] = i'_q$). Note that the vertices belonging to $\mathcal{Q}[i'\colon j]$ are distinct, and hence each intermediate vertex belongs to one and only one middle path.

ALGORITHM 2. Decomposition in subpaths

Input $\mathcal{Q} = (i', j_1, i'_1, \ldots, j_t, i'_t, j)$
Output MC **array updated**
if $t \neq 0$ **then**
1. decompose $\mathcal{Q}[i'\colon j]$ in $\mathcal{Q}[i'\colon j_p]$, $\mathcal{Q}[j_p\colon i'_q]$, and $\mathcal{Q}[i'_q\colon j]$ such that j_p is the largest j_k, where $1 \leq k \leq t$ and $j_p \leq i'_q \leq n$ and the three paths are also row merge fill paths (Definition 5.2).
2. $MC[j_p] = i'_q$.
3. decompose each of the three paths (which is not an edge) in sub-fill paths.

end if

Algorithm 3 decomposes the row merge fill path $\mathcal{Q}[i'\colon j]$ in an alternating sequence of edges and middle paths that we refer to as $ASEM$. It is easy to check that this algorithm returns the sequence $ASEM = \{\langle i', k_1\rangle, \mathcal{Q}[k_1\colon MC[k_1]], \langle MC[k_1], k_2\rangle, \mathcal{Q}[k_2\colon MC[k_2]], \ldots, \mathcal{Q}[k_u\colon MC[k_u]], \langle MC[k_u], j\rangle\}$, where $u \leq t$ and $k_1 = j_1$.

Consider an edge of the row merge graph $\langle i', j\rangle$ and its associated row merge fill path $\mathcal{Q} = (i', j_1, i'_1, \ldots, j_t, i'_t, j)$. We define a pivoting strategy relative to this path. At each elimination step k, if column vertex k is an intermediate vertex of

ALGORITHM 3. Decomposition in alternating sequence of edges and middle paths

Input $\mathcal{Q} = (i', j_1, i'_1, \ldots, j_t, i'_t, j)$
Output alternating sequence $ASEM$
if $t \neq 0$ **then**
1. decompose $\mathcal{Q}[i' \colon j]$ in $\mathcal{Q}[i' \colon j_p]$, $\mathcal{Q}[j_p \colon i'_q]$ and $\mathcal{Q}[i'_q \colon j]$ such that j_p is the largest j_k, $j_p \leq i'_q \leq n$, and the three paths are also row merge fill paths (Definition 5.2).
2. decompose $\mathcal{Q}[i' \colon j_p]$ in an alternating sequence and assign it to $ASEM$.
3. add the middle sub-fill path $\mathcal{Q}[j_p \colon i'_q]$ at the end of the sequence $ASEM$.
4. decompose $\mathcal{Q}[i'_q \colon j]$ in an alternating sequence and add it at the end of the sequence $ASEM$.
5. return the sequence $ASEM$.

else
6. return the edge $\langle i', j \rangle$.

end if

the path $\mathcal{Q}[i' \colon j]$, then we pivot on the element in position $(MC[k], k)$, and P_k is the elementary permutation matrix that describes this pivoting. If column vertex k is not an intermediate vertex of $\mathcal{Q}[i' \colon j]$, then we pivot on the diagonal element; that is, the elementary permutation matrix P_k is the identity. We call this strategy of pivoting the middle correspondent pivoting strategy with respect to the path $\mathcal{Q}[i' \colon j]$. In the next theorem we prove that such a strategy is valid; that is, the LU factorization exists in a symbolic sense.

LEMMA 5.4. *Let A be a square matrix with nonzero diagonal that satisfies the Hall property. Let $\langle i', j \rangle$ be an edge of the row merge graph $H^\times(A)$ and $\mathcal{Q}[i' \colon j]$ be its corresponding fill path in $H(A)$. Let $P = P_{n-1} \ldots P_2 P_1$ be the permutation matrix describing the middle correspondent pivoting strategy relative to $\mathcal{Q}[i' \colon j]$. Gaussian elimination $A = P_1 L_1 \ldots P_{n-1} L_{n-1} U$ exists in the symbolic sense.*

Proof. If the fill path $\mathcal{Q}[i' \colon j]$ corresponds to an edge of $H(A)$, then we choose P to be the identity matrix. As we assume the matrix A has a nonzero diagonal, the Gaussian elimination exists in the symbolic sense. In the rest of the proof, we assume that $\langle i', j \rangle$ is not an edge of $H(A)$.

As the case $j = 1$ is trivial, we will assume that $j > 1$. We will prove this by induction. At the first step of elimination, if row vertex $1'$ and column vertex 1 do not belong to $\mathcal{Q}[i' \colon j]$, then we pivot on the element in position $(1, 1)$. If the column vertex 1 belongs to $\mathcal{Q}[i' \colon j]$, then consider the fill path $\mathcal{Q}[1 \colon k']$, where $k' = MC[1]$ and $k' \geq 1$. We can see that $\mathcal{Q}[1 \colon k']$ is an edge of $H(A)$, and thus we can pivot on the element A_{k1}. Note that according to Definition 5.2, we cannot have that row vertex $1'$ belongs to $\mathcal{Q}[i' \colon j]$ and column vertex 1 does not belong to $\mathcal{Q}[i' \colon j]$.

Consider the kth step of elimination, where $k < n$. Suppose that at each elimination step prior to k, the middle correspondent pivoting strategy was valid; that is, the diagonal elements of the permuted matrix are nonzero. We show that at this step k we can apply the same pivoting strategy. Let P_{K-1} be the permutation matrix that describes the first $k-1$ row interchanges, that is, $P_{K-1} = P_{k-1} \ldots P_1$. Let A_k be the $k \times k$ principal submatrix of $P_{K-1}A$ that includes the first k columns and the rows in corresponding positions of $P_{K-1}A$. The columns of A_k are those numbered 1 through k in $H(A)$; the rows of A_k are those given by the permutation matrix P_{K-1}. We add to the matrix A_k all of the diagonal elements, except the last one, nonzero by our hypothesis. In the directed graph $G(A_k)$ we will number the vertices from 1 to k.

First, we will prove that the kth diagonal element of the permuted matrix $P_{K-1}A$ is nonzero. This corresponds to the last diagonal element of A_k. If k' is not an intermediate row vertex of the path $\mathcal{Q}[i'\colon j]$, then during the first $k-1$ steps of elimination row k was not permuted, and the last diagonal element of A_k is nonzero by our hypothesis. If k' is an intermediate row vertex on the path $\mathcal{Q}[i'\colon j]$, then row k was permuted during the first $k-1$ steps of elimination. We denoted by d the row that at the kth step of elimination is in position k of matrix A_k. We now trace the pivoting process to discover where row d comes from. Let k_1 be the middle correspondent vertex of k' ($MC[k_1] = k', k_1 \leq k'$). If $k_1 = k$, then $d' = k'$. Otherwise, according to our pivoting choice, the element in position (k', k_1) was used as the pivot at step k_1, and thus row k was interchanged with the row in position k_1. At this point, either row vertex k'_1 does not belong to $\mathcal{Q}[i'\colon j]$, and then $d = k_1$, or else it belongs, and then row k_1 was used as the pivot in some column k_2, where it was interchanged with some row $k_2 < k_1$. Extending the induction, we arrive at a row vertex $k'_q = d'$, which is not an intermediate row vertex of $\mathcal{Q}[i'\colon j]$. The vertices in $H(A)$ followed while tracing the pivoting process form the path $(k, k', k_1, k'_1, k_2, k'_2, \ldots, k_q, d')$. On this path, the edge $\langle k', k_1\rangle$ and the edges $\langle k'_p, k'_{p+1}\rangle$, with $1 \leq p < q$, correspond to diagonal elements of A_k. Hence this path can be transformed into the path $(k, k_1, \ldots, k_q, k)$ in $G(A_k)$. As $k > k_1 > \cdots > k_q$, according to Theorem 3.1 this path is a fill path in the directed graph $G(A_k)$, and the kth diagonal element of $P_{K-1}A$ corresponds to a symbolic nonzero.

Second, we show that at elimination step k we can apply the middle correspondent pivoting strategy. We distinguish two cases.

Case 1 (column vertex k is not an intermediate column vertex of $\mathcal{Q}[i'\colon j]$). We have just proved that the kth diagonal element of $P_{K-1}A$ is an edge of the filled graph $G^+(A_k)$. We use as the pivot the diagonal element.

Case 2 (column vertex k is an intermediate column vertex of $\mathcal{Q}[i'\colon j]$). Let e' be the middle path correspondent vertex of k, that is, $MC[k] = e'$ and $k \leq e'$. Let $\mathcal{Q}[e'\colon k]$ be the fill path between e' and k which is a subpath of our initial path $\mathcal{Q}[i'\colon j]$.

If $e' = k'$ (that is, $MC[k] = k'$), then row k' was not involved in any row permutation. We use as the pivot the diagonal element. If $e' > k$, then let K be the $(k+1) \times (k+1)$ submatrix of $P_{K-1}A$ that includes the first k columns and the rows in corresponding positions of $P_{K-1}A$ and column e and row e' of $P_{K-1}A$. We add to matrix K the first k diagonal elements, which correspond to symbolic nonzeros by our hypothesis. The vertices of the directed graph $G(K)$ are the vertices 1 through k and vertex e.

In the following, we want to show that $\langle e, k\rangle$ is an edge of the directed graph $G^+(K)$. If path $\mathcal{Q}[e'\colon k]$ is simply an edge, then $\langle e, k\rangle$ is an edge of $G^+(K)$. Otherwise, we decompose path $\mathcal{Q}[e'\colon k]$ into an alternating sequence of edges and middle paths using Algorithm 3. The following sequence is obtained: $\{\langle e', e_1\rangle, \mathcal{Q}[e_1\colon MC[e_1]], \langle MC[e_1], e_2\rangle, \mathcal{Q}[e_2\colon MC[e_2]], \ldots, \mathcal{Q}[e_q\colon MC[e_q]], \langle MC[e_q], k\rangle\}$. We can rewrite the sequence as a directed path from vertex e to vertex k of $G(K)$: $(e, e_1, e_2, \ldots, e_q, k)$. The intermediate vertices on this path are less than both e and k, because of the row merge fill paths Definition 5.2. Therefore $\langle e, k\rangle$ is an edge of $G^+(K)$, and thus it corresponds to a symbolic nonzero. This shows that we can choose as the pivot the element in position (e, k) at this step of elimination, and this ends our proof. ∎

The next theorem shows that the row merge graph represents a tight bound for the nonzero structure of $\widetilde{L}$ and U, in the symbolic sense. It is illustrated in Figures 5.3, 5.4, 5.5, and 5.6.

THEOREM 5.5. *Let A be a square matrix with nonzero diagonal that satisfies the Hall property. Let $\langle i', j\rangle$ be an edge of the row merge graph $H^\times(A)$. There is a*

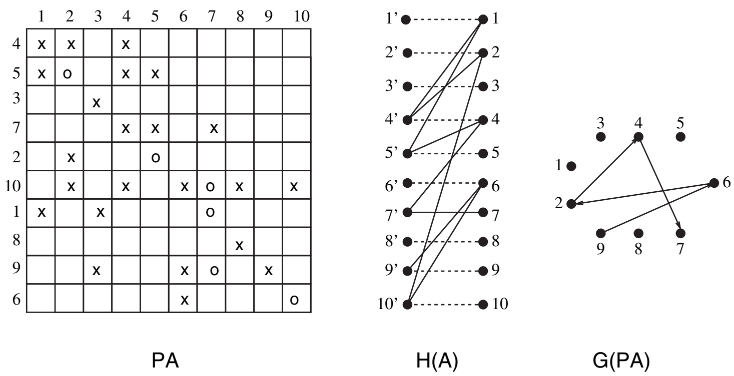

FIG. 5.3. *Example illustrating Theorem* 5.5 *and showing that* L_{97} *is nonzero for the row merge matrix presented in Figure* 5.2. *Consider the row merge fill path* $\mathcal{Q}[9'\colon 7] = (9', 6, 10', 2, 4', 1, 5', 4, 7', 7)$. *This path is displayed by solid edges in the bipartite graph* $H(A)$. *Figure* 5.4 *presents the decomposition of path* $\mathcal{Q}[9'\colon 7]$ *by Algorithms* 2 *and* 3. *First, Algorithm* 2 *decomposes* $\mathcal{Q}[9'\colon 7]$ *and obtains the following middle paths:* $\mathcal{Q}[6\colon 10'], \mathcal{Q}[2\colon 5'], \mathcal{Q}[4'\colon 1], \mathcal{Q}[4\colon 7']$. *This decomposition gives us the pivoting strategy, illustrated in the permuted matrix at the top left of Figure* 5.3. *Second, the fill path* $\mathcal{Q}[9'\colon 7]$ *is decomposed in an alternating sequence of edges and middle paths using Algorithm* 3. *This allows us to obtain the path* $(9, 6, 2, 4, 7)$ *which is a fill path in the directed graph of the permuted matrix* PA.

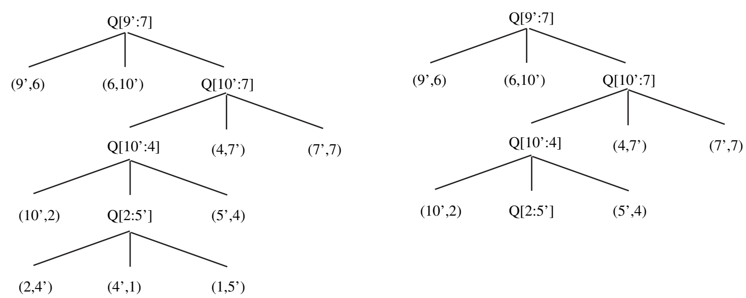

FIG. 5.4. *Example of the application of Algorithm* 2 *(left) and Algorithm* 3 *(right) on the path* $\mathcal{Q}[9'\colon 7] = (9', 6, 10', 2, 4', 1, 5', 4, 7', 7)$ *from Figure* 5.3.

permutation $P = P_1 P_2 \dots P_{n-1}$ *such that if* A *is factored by Gaussian elimination as* $A = P_1 L_1 \dots P_{n-1} L_{n-1} U$, *with* $\widetilde{L}$ *being the union of* L_i, *then* $(L+U)_{ij} \neq 0$, *in the symbolic sense.*

Proof. According to Theorem 5.3 there is a row merge fill path in $H(A)$ from row vertex i' to column vertex j. Let $\mathcal{Q}[i'\colon j]$ be formed by the vertices $(i', j_1, i'_1, \dots, j_t, i'_t, j)$. Assume that $t \neq 0$. At each step of elimination we pivot following the middle correspondent pivoting strategy with respect to the path $\mathcal{Q}[i'\colon j]$, as described in Lemma 5.4.

Assume now that we are at the jth step of elimination. Let P_{J-1} be the permutation matrix that describes the first $j-1$ row interchanges. Let K be the principal submatrix of $P_{J-1}A$ that includes the first j columns and column i and the rows in corresponding positions of PA (that is, if $i' \leq j$, then K is a $j \times j$ matrix; otherwise K is a $(j+1) \times (j+1)$ matrix). In matrix K we add diagonal elements, with $1 \leq i \leq j$, which are nonzero by our hypothesis. When $i > j$, we also add diagonal element (i', i) (row i was not permuted). The vertices of the directed graph $G(K)$ are numbered 1 through j and i.

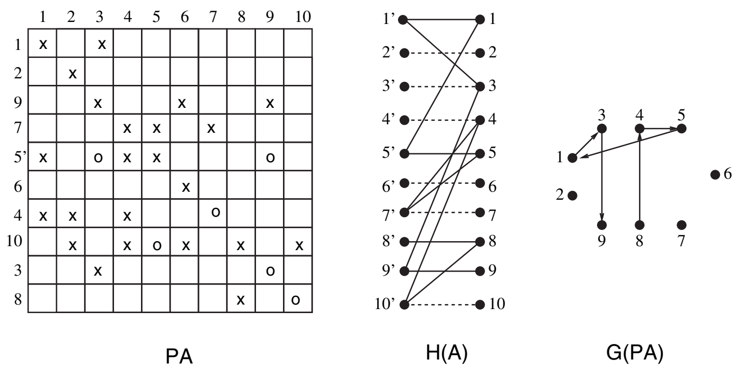

FIG. 5.5. *Example illustrating Theorem* 5.5 *and showing that* U_{89} *is nonzero for the row merge matrix presented in Figure* 5.2*. The row merge fill path* $Q[8'\colon 9] = (8', 8, 10', 4, 7', 5, 5', 1, 1', 3, 9', 9)$ *is displayed by solid edges in the bipartite graph* $H(A)$*. Figure* 5.6 *presents the decomposition of path* $Q[8'\colon 9]$ *by Algorithms* 2 *and* 3*. First, the path* $\mathcal{Q}[8'\colon 9]$ *is decomposed using Algorithm* 2*, and the following middle paths are obtained:* $\mathcal{Q}[8\colon 10']$, $\mathcal{Q}[4\colon 7']$, $\mathcal{Q}[5\colon 5']$, $\mathcal{Q}[9'\colon 3]$*, and* $\mathcal{Q}[1\colon 1']$*. This decomposition gives us the pivoting strategy, illustrated in the permuted matrix at the top left of Figure* 5.5*. Algorithm* 3 *decomposes the fill path* $\mathcal{Q}[8'\colon 9]$ *in an alternating sequence of edges and middle paths. This allows us to obtain the path* $(8, 4, 5, 1, 3, 9)$ *which is a fill path in the directed graph of the permuted matrix* PA*.*

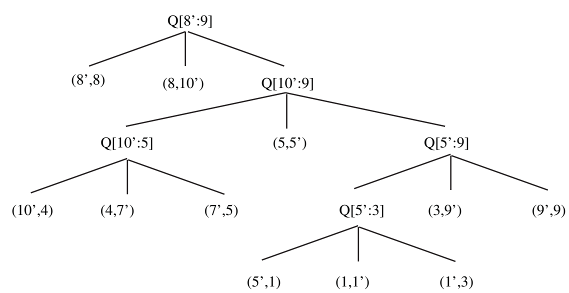

FIG. 5.6. *Example of the application of Algorithms* 2 *and* 3 *on the path* $\mathcal{Q}[8'\colon 9] = (8', 8, 10', 4, 7', 5, 5', 1, 1', 3, 9', 9)$ *from Figure* 5.5*. Both algorithms return the same result.*

Case 1 ($i > j$ (structure of $\widetilde{L}$)). The proof is similar to the proof of Lemma 5.4, in which a middle path becomes an edge of the filled graph of A, and we omit the details here.

Case 2 ($i < j$ (structure of U)). Consider row merge fill path $\mathcal{Q}[i'\colon j] = (i', j_1, i'_1, \ldots, j_t, i'_t, j)$. We distinguish two cases. If column vertex i is not a vertex of path $\mathcal{Q}[i'\colon j]$, then row i of A is not permuted during our pivoting strategy, and the proof is similar to the $\widetilde{L}$ case. If column vertex i is a vertex of path $\mathcal{Q}[i'\colon j]$, Definition 5.2 decomposes this path in the following three paths: $\mathcal{Q}[i'\colon i]$, $\mathcal{Q}[i\colon k']$, and $\mathcal{Q}[k'\colon j]$ such that $i < k' \le n$, and the three paths are also row merge fill paths in $H^{\times}(A)$. Our pivoting strategy interchanges rows i and k of A at the ith step of elimination. We use Algorithm 3 to decompose the path $\mathcal{Q}[k'\colon j]$ in an alternating sequence of edges and middle paths. This sequence is transformed into a path from i to j in the graph $G(K)$ which has all of the intermediate column vertices smaller than i. This corresponds to an edge in the filled graph $G^{+}(K)$. Thus the element U_{ij} corresponds to a symbolic nonzero, and this ends our proof. □

We make one note about the structure prediction of $A = P_1L_1 \ldots P_{n-1}L_{n-1} U$. The tight bound of U obtained for the structure prediction of $PA = LU$ (Theorem 4.3) also represents a tight bound for U obtained in $A = P_1L_1 \ldots P_{n-1}L_{n-1} U$. But there does not seem to be a simple way to express tight exact bounds for $\widetilde{L}$, where $\widetilde{L}$ is the union of the L_i obtained from $A = P_1L_1 \ldots P_{n-1}L_{n-1}U$.

6. Concluding remarks. In this paper we have discussed two aspects of interest in the structure prediction problem of sparse LU factorization with partial pivoting of a matrix A. The first aspect considers the computation of the nonzero structure of the factors during Gaussian elimination with row interchanges. We have presented new results that provide an exact structure prediction for matrices that satisfy the strong Hall property or only the Hall property. We then have used the theoretical results to derive an algorithm for computing fill-ins. The second aspect is to estimate tight bounds of the structure of L and U prior to the numerical factorization. We have introduced tight exact bounds for the nonzero structure of L and U of Gaussian elimination with partial pivoting $PA = LU$, under the assumption that the matrix A satisfies the Hall property. We have also shown that the row merge graph represents a tight symbolic bound for the structure of the factors $\widetilde{L}$ and U obtained from the factorization $A = P_1L_1 \ldots P_{n-1}L_{n-1}U$.

The practical usage of the exact structure prediction presented in this paper remains an open problem. Several aspects are of interest. One important question is to understand if rounding to zero elements that correspond to numeric cancellation in exact arithmetic leads to instability in the Gaussian elimination. A different aspect is to analyze on real world matrices how many numeric cancellations, that Theorem 3.4 identifies, occur during Gaussian elimination. Another aspect is to compare experimentally the bounds presented in this paper with the bounds provided by the row merge graph, knowing that the latter can be efficiently computed [14].

Acknowledgments. The authors thank the anonymous reviewers for their helpful comments and suggestions to improve the presentation of the paper.

REFERENCES

[1] R. K. Brayton, F. G. Gustavson, and R. A. Willoughby, *Some results on sparse matrices*, Math. Comp., 24 (1970), pp. 937–954.
[2] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, Cambridge, 1991.
[3] R. A. Brualdi and B. L. Shader, *Strong Hall matrices*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 359–365.
[4] T. F. Coleman, A. Edenbrandt, and J. R. Gilbert, *Predicting fill for sparse orthogonal factorization*, J. ACM, 33 (1986), pp. 517–532.
[5] J. W. Demmel, J. R. Gilbert, and X. S. Li, *An asynchronous parallel supernodal algorithm for sparse Gaussian elimination*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 915–952.
[6] J. W. Demmel, S. C. Eisenstat, J. R. Gilbert, X. S. Li, and J. W. H. Liu, *A supernodal approach to sparse partial pivoting*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 720–755.
[7] I. S. Duff, *On algorithms for obtaining a maximum transversal*, ACM Trans. Math. Software, 7 (1981), pp. 315–330.
[8] F. R. Gantmacher, *The Theory of Matrices*, Chelsea, New York, 1959.
[9] A. George and E. Ng, *Symbolic factorization for sparse Gaussian elimination with partial pivoting*, SIAM J. Sci. Stat. Comput., 8 (1987), pp. 877–898.
[10] J. R. Gilbert and J. W. H. Liu, *Elimination structures for unsymmetric sparse LU factors*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 334–352.
[11] J. R. Gilbert and E. G. Ng, *Predicting structure in nonsymmetric sparse matrix factorizations*, in Graph Theory and Sparse Matrix Computation, A. George, J. R. Gilbert, and J. W. H. Liu, eds., Springer-Verlag, New York, 1994, pp. 107–139.

[12] J. R. Gilbert, *An Efficient Parallel Sparse Partial Pivoting Algorithm*, Technical report 88/45052-1, Christian Michelsen Institute, Bergen, Norway, 1988.
[13] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, 2nd ed., Baltimore, MD, 1989.
[14] L. Grigori, M. Cosnard, and E. G. Ng, *On the row merge tree for sparse LU factorization with partial pivoting*, BIT, 47 (2007), pp. 45–76.
[15] D. R. Hare, C. R. Johnson, D. D. Olesky, and P. van den Driessche, *Sparsity analysis of the QR factorization*, SIAM J. Matrix Anal. Appl., 14 (1991), pp. 655–669.
[16] X. Li and J. Demmel, *A scalable distributed-memory sparse direct solver for unsymmetric linear systems*, ACM Trans. Math. Software, 29 (2003), pp. 110–140.
[17] L. Lovasz and M. D. Plummer, *Matching Theory*, North–Holland, Amsterdam, 1986.
[18] A. Pothen and C.-J. Fan, *Computing the block triangular form of a sparse matrix*, ACM Trans. Math. Software, 16 (1990), pp. 303–324.
[19] A. Pothen, *Predicting the structure of sparse orthogonal factors*, Linear Algebra Appl., 194 (1993), pp. 183–204.
[20] D. J. Rose and R. E. Tarjan, *Algorithmic aspects of vertex elimination on directed graphs*, SIAM J. Appl. Math., 34 (1978), pp. 176–197.
[21] K. Shen, T. Yang, and X. Jiao, *S+: Efficient 2d sparse LU factorization on parallel machines*, SIAM J. Matrix Anal. Appl., 22 (2000), pp. 282–305.