

Jordan Structures of Totally Nonnegative Matrices

Shaun M. Fallat and Michael I. Gekhtman

Abstract. An $n \times n$ matrix is said to be totally nonnegative if every minor of A is nonnegative. In this paper we completely characterize all possible Jordan canonical forms of irreducible totally nonnegative matrices. Our approach is mostly combinatorial and is based on the study of weighted planar diagrams associated with totally nonnegative matrices.

1 Introduction and Main Results

In this paper we give a complete solution to the inverse spectral problem for irreducible totally nonnegative matrices. Recall that an $n \times n$ matrix A is called *totally positive*, TP (*totally nonnegative*, TN) if every minor of A is positive (nonnegative). A is called *irreducible* if there is no permutation matrix P such that

$$PAP^T = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

where 0 is an $(n-r) \times r$ zero matrix ($1 \leq r \leq n-1$). The main results of the paper are summarized in the following theorem

Theorem 1 *Let A be an irreducible $n \times n$ TN matrix. Then*

- (1) *A has at least one non-zero eigenvalue. Furthermore, the non-zero eigenvalues of A are positive and distinct.*
- (2) *Let p be the number of non-zero eigenvalues of A . For each $i \geq 1$, let $m_i = m_i(A)$ denote the number of Jordan blocks of size i corresponding to the zero eigenvalue in the Jordan canonical form of A . Then the numbers m_i satisfy the following conditions:*
 - (a) $\sum im_i = n - p$.
 - (b) $m_i = 0$ for $i > p$.
 - (c) $\sum m_i \geq \frac{n-p}{p}$.
- (3) *Conversely, let $n \geq 2$, let $1 \leq p \leq n$, let $\lambda_1, \dots, \lambda_p$ be arbitrary distinct positive numbers and let m_1, m_2, \dots be nonnegative integers satisfying conditions in 2.*

Received by the editors January 22, 2003; revised May 6, 2003.

The first author's research was supported in part by an NSERC research grant.

AMS subject classification: 15A21, 15A48, 05C38.

Keywords: totally nonnegative matrices, planar diagrams, principal rank, Jordan canonical form.

©Canadian Mathematical Society 2005.

Then there exists an irreducible TN matrix A having non-zero eigenvalues and such that $m_i(A) = m_i$ for all $i \geq 1$.

The theorem above extends the classical result of Gantmacher and Krein [9, 10], who introduced the concept of total positivity and proved that the eigenvalues of every TP matrix are positive and distinct. The same is true for every *oscillatory matrix*, that is a totally nonnegative matrix such that some positive integer power of it is totally positive. Since the set of TN matrices coincides with the closure of the set of TP matrices, it follows that the eigenvalues of a TN matrix are real and nonnegative.

Gantmacher and Krein also proved that a TN matrix is oscillatory if and only if it is invertible and irreducible. Thus, their result on spectral properties of oscillatory matrices coincides with the first statement of Theorem 1 with an additional assumption that A is invertible. In its general form, part 1 of Theorem 1 was proved in our earlier paper [6] with C. R. Johnson, where we also formulated several conjectures on the possible structures of Jordan blocks corresponding to the zero eigenvalue in the Jordan canonical form of an irreducible TN matrix. All of these conjectures will be verified below and will also serve as steps in the proof of parts 2 and 3 of Theorem 1.

It should be noted that if one does not require irreducibility, the problem of describing possible Jordan canonical forms of TN matrices becomes trivial: a Jordan canonical form of a matrix is TN if and only if all its eigenvalues are real and nonnegative.

The paper is organized as follows: in the next section we will discuss a combinatorial approach to TN matrices based on their realization as path matrices of weighted planar diagrams. Section 3 reviews in more detail known results on nonzero eigenvalues of TN matrices. In addition, we give a description of nonzero spectrum of TN matrices associated with a given planar diagram, and we verify item (1) of Theorem 1. In section 4 we prove conjectures formulated in [6] concerning necessary conditions that must be satisfied by rank, principal rank and index of the zero eigenvalue of an arbitrary irreducible TN matrix. Results in this section lead to a proof of item (2) in Theorem 1. Finally, in Section 5, we complete the proof of the main theorem.

Readers interested in learning more about a vast array of applications pertaining to TN matrices in various areas of pure and applied mathematics are encouraged to consult the following books and review articles: [1, 3, 5, 7, 11, 17, 15].

2 Planar Networks, Bidiagonal Factorizations and Total Nonnegativity

It is an easy corollary of the classical Cauchy-Binet identity for determinants that the set of TN matrices forms a multiplicative semigroup. With this in mind, one way to better understand TN matrices is through a study of generators of this semigroup. In other words, one would like to find a decomposition of an arbitrary TN matrix into a product of certain “basic” or “elementary” TN matrices.

The first result of this kind was obtained in 1952 by A. Whitney [18], who showed that under a particular elimination scheme and assuming no accidental cancellation (although this turns out not to be a difficult issue, see [4]), one can reduce a TN matrix to an upper triangular TN matrix. Applying similar reasoning to the transpose

of the resulting upper triangular matrix produces a factorization of A in terms of totally nonnegative bidiagonal matrices of the form $E_k(\alpha) = I + \alpha E_{k,k-1}$, $E_j^T(\alpha)$ (with E_{ij} being the elementary standard basis matrix whose only nonzero entry is a one in position (i, j)) and a nonnegative diagonal factor. Such a factorization is called an *elementary bidiagonal factorization* of A .

In general, we define an *elementary bidiagonal matrix* as an upper or lower triangular matrix that has a form $D + \alpha E_{k,k\pm 1}$, where D is a diagonal matrix. Observe that an elementary bidiagonal matrix is TN if and only if it is entrywise nonnegative. There has been a significant amount of work done on factorizations of TN matrices into elementary bidiagonal matrices (see [2, 4, 12, 14]). In the nonsingular case, the most commonly used factorization is described by the next result.

Theorem 2 *Let A be an $n \times n$ nonsingular TN matrix. Then A can be written as*

$$(1) \quad A = E_2(l_k)(E_3(l_{k-1})E_2(l_{k-2})) \cdots \\ \cdots (E_n(l_{n-1}) \cdots E_3(l_2)E_2(l_1)) D (E_2^T(u_1)E_3^T(u_2) \cdots E_n^T(u_{n-1})) \cdots \\ \cdots (E_2^T(u_{k-2})E_3^T(u_{k-1})) E_2^T(u_k),$$

where $k = \binom{n}{2}$; $l_i, u_j \geq 0$ for all $i, j \in \{1, 2, \dots, k\}$; and D is a positive diagonal matrix.

In what follows, we will need the following extension of Theorem 2, due to Cryer [4]:

Theorem 3 *Any $n \times n$ TN matrix A can be written as*

$$(2) \quad A = \prod_{i=1}^M L^{(i)} \prod_{j=1}^N U^{(j)},$$

where the matrices $L^{(i)}$ and $U^{(j)}$ are, respectively, lower and upper elementary bidiagonal TN matrices.

Before we discuss connections between TN matrices and weighted digraphs we need to set forth some notation. For an $n \times n$ matrix $A = [a_{ij}]$, $\alpha, \beta \subseteq \{1, 2, \dots, n\}$, the submatrix of A lying in rows indexed by α and the columns indexed by β will be denoted by $A[\alpha|\beta]$. If, in addition, $\alpha = \beta$, then the principal submatrix $A[\alpha|\alpha]$ is abbreviated to $A[\alpha]$.

An excellent treatment of the combinatorial and algebraic aspects of bidiagonal factorizations of TN matrices along with generalizations for totally positive elements in reductive Lie groups is given in [2, 7, 8]. One of the main tools used in these papers is a graphical representation of the bidiagonal factorization in terms of planar diagrams (or networks) that can be described as follows.

An $n \times n$ diagonal matrix $\text{diag}(d_1, d_2, \dots, d_n)$ is represented by the diagram on the left in Figure 1 while an elementary lower (upper) bidiagonal matrix $E_k(l)$ ($E_j^T(u)$) is the diagram middle (right) of Figure 1.

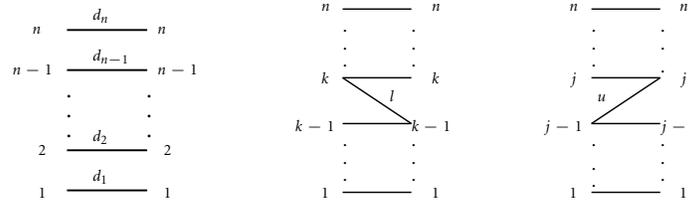


Figure 1: Diagrams

Each horizontal edge of the last two diagrams has a weight of 1. It is not difficult to verify that if A is a matrix represented by any one of the diagrams above, then $\det A[\{i_1, i_2, \dots, i_t\}|\{j_1, j_2, \dots, j_t\}]$ is nonzero if and only if in the corresponding diagram there is a family of t vertex-disjoint paths joining the vertices $\{i_1, i_2, \dots, i_t\}$ on the left-side of the diagram with the vertices $\{j_1, j_2, \dots, j_t\}$ on the right side. Moreover, in this case this family of paths is unique and

$$\det A[\{i_1, i_2, \dots, i_t\}|\{j_1, j_2, \dots, j_t\}]$$

is equal to the product of all the weights assigned to the edges that form this family.

Now, given a product $A = A_1 A_2 \cdots A_l$ in which each matrix A_i is either a diagonal matrix or an elementary (upper or lower) bidiagonal matrix, a corresponding diagram \mathcal{D} is obtained by concatenation left to right of the diagrams associated with the matrices A_1, A_2, \dots, A_l . Then the Cauchy-Binet formula for determinants applied to the matrix A above implies the next two results. Given a diagram \mathcal{D} and a path connecting a vertex on the left side of \mathcal{D} to a vertex in the right side of \mathcal{D} , the *weight of this path* is defined to be the product of all the weights of the edges along this path.

Proposition 4 Suppose $A = A_1 A_2 \cdots A_l$ in which each matrix A_i is either a diagonal matrix or an elementary (upper or lower) bidiagonal matrix. Then

- (1) the (i, j) -entry of A is equal to the sum of the weights of all paths joining the vertex i on the left side of the obtained diagram \mathcal{D} with the vertex j on the right side.
- (2) For index sets $\alpha = \{i_1, i_2, \dots, i_t\}$ and $\beta = \{j_1, j_2, \dots, j_t\}$, consider a collection $P(\alpha, \beta)$ of all families of vertex-disjoint paths joining the vertices $\{i_1, i_2, \dots, i_t\}$ on the left of the diagram \mathcal{D} with the vertices $\{j_1, j_2, \dots, j_t\}$ on the right. For $\pi \in P(\alpha, \beta)$, let $w(\pi)$ be the product of all the weights assigned to edges that form a family π . Then

$$\det A[\alpha|\beta] = \sum_{\pi \in P(\alpha, \beta)} w(\pi).$$

Theorems 2 and 3 imply that every TN matrix can be represented by a weighted planar diagram constructed from building blocks as in Figure 1 (in the case of degenerate matrices some of the horizontal edges may have to be erased). As the most important example, consider the bidiagonal factorization of an arbitrary TP matrix

A from Theorem 2. This factorization translates into the diagram in Figure 2 (here $k = \frac{n(n-1)}{2}$).

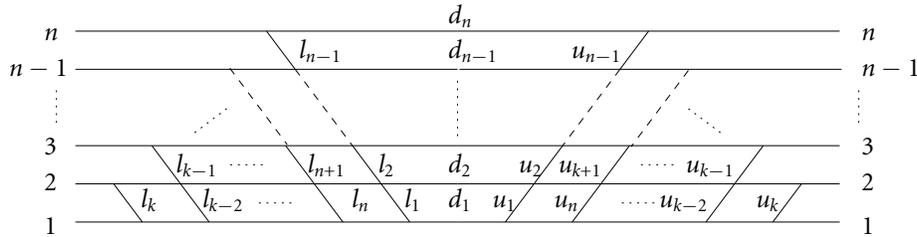


Figure 2: General $n \times n$ diagram.

More general weighted planar digraphs can also be associated to TN matrices (see, e.g., [3, 13]). A *planar diagram of order n* is a planar acyclic digraph \mathcal{D} with all edges oriented from left to right and $2n$ distinguished boundary vertices: n sources on the left and n sinks on the right with both sources and sinks labeled $1, \dots, n$ from bottom to top. To each edge of a planar diagram \mathcal{D} we assign a positive weight. We denote a collection of all assigned weights by W , and call the pair (\mathcal{D}, W) a *weighted planar diagram of order n* . Clearly, the diagrams in Figures 1 and 2 are weighted planar diagrams of order n .

Now, if (\mathcal{D}, W) is an arbitrary weighted planar diagram of order n , then the first statement in Proposition 4 can be used as a *definition* of an $n \times n$ matrix $A(\mathcal{D}, W)$ associated with (\mathcal{D}, W) . It was first observed by Karlin and MacGregor [16], that the second statement remains valid in this more general situation and, in particular, $A(\mathcal{D}, W)$ is TN.

We denote by $\mathcal{M}(\mathcal{D})$ the set of all TN matrices A such that $A = A(\mathcal{D}, W)$ for some choice W of positive weights. In what follows, a *path from i to j in \mathcal{D}* will mean a path from the i th source to j th sink in a planar diagram \mathcal{D} . We will also use terms *the highest* and *the lowest* path in \mathcal{D} for the unique paths from n to n and from 1 to 1 such that the entire diagram is enclosed between these two paths.

We conclude this section with a well-known criterion for irreducibility of a TN matrix. A diagram-based proof is provided to illustrate the technique to be used in the following sections.

Proposition 5 *An $n \times n$ TN matrix $A = [a_{ij}]$ is irreducible if and only if $a_{ij} > 0$ for all i, j such that $|i - j| = 1$.*

Proof Only necessity needs to be established, as the sufficiency of the condition is obvious. Suppose that for some i and j , $a_{ij} = 0$. If A is irreducible, it contains no zero rows or columns. Thus, there exists i' such that $a_{i'j} > 0$. We claim that if $i' < i$ then $a_{kl} = 0$ for all $k \geq i, l \leq j$, and if $i' > i$ then $a_{kl} = 0$ for all $k \leq i, l \geq j$. Since both cases can be treated in a similar way, we consider only the former.

Let \mathcal{D} be a planar diagram that corresponds to A . In \mathcal{D} there is a path P from i' to j , but no paths from i to j . If for $a_{kl} > 0$ for some $k \geq i, l \leq j$, then there is also a path Q from k to l in \mathcal{D} , and P and Q intersect at some vertex v_1 . Since A has no nonzero rows, there is a path R that starts at i , and since $i' < i < k$, R intersects either P or Q at a vertex v_0 that lies to the left of v_1 . Then, following R from i to v_0 , then either P or Q from v_0 to v_1 and then P from v_1 to j , we obtain a path from i to j , which is a contradiction.

Now, if we assume that $a_{i+1,i} = 0$ for some i , then $a_{kl} = 0$ either for all $k > i$ and $l < i$ or for all $k < i$ and $l > i$, and in both cases, A is reducible. The case $a_{i+1,i} = 0$ can be treated similarly. ■

Note that if A is TN and $a_{i+1,i}, a_{i,i+1} > 0$ for each i , then certainly $a_{ii} > 0$ for each i .

3 Positive Eigenvalues

From this point on, we restrict our attention to irreducible TN matrices, which we denote by ITN.

As mentioned in the introduction, Gantmacher and Krein [10] proved that any oscillatory (equivalently, invertible ITN) matrix has distinct positive eigenvalues. Thus we turn our attention to singular ITN matrices. The following result obtained in [6] can be viewed as an extension of Gantmacher and Krein's original result on the positive eigenvalues of oscillatory matrices.

Theorem 6 *Let A be an $n \times n$ irreducible TN matrix. Then the positive eigenvalues of A are distinct.*

To make the presentation more self-contained, we provide an outline of the proof of Theorem 6. Since TN matrices have only nonnegative eigenvalues and ITN matrices have positive trace, it is clear that every ITN matrix must have at least one nonzero (positive) eigenvalue. Starting with an $n \times n$ ITN matrix A , we construct an ITN matrix A' that satisfies either:

- (i) A' is $n \times n$, similar to A , and has at least one more zero entry; or
- (ii) A' is $(n - 1) \times (n - 1)$ and has the same nonzero eigenvalues A .

The matrix A' is obtained from A via similarity transformation by elementary bidiagonal matrix of the form $E_k(\alpha)$ (or $E_k^T(\alpha)$) that appears in an appropriately chosen bidiagonal factorization of A . If the result \tilde{A} of the similarity transformation does not contain a zero column (row), we set $A' = \tilde{A}$ and show that A' satisfies (i) above. Otherwise, A' is obtained from \tilde{A} by deleting the zero column (row) and the corresponding row (column). In this case, A' satisfies (ii). Observe that operations described above clearly preserve nonzero eigenvalues of a matrix. Verifying that A' is irreducible requires a bit more work.

Repeatedly applying this reduction algorithm, one finally produces an irreducible tridiagonal (i.e., with only nonzero entries in positions (i, j) such that $|i - j| \leq 1$) TN matrix T that, by construction, has the same nonzero eigenvalues as A . It is

well-known that an irreducible tridiagonal TN matrix must have distinct positive eigenvalues (see, e.g., [10]). Hence A must have distinct positive eigenvalues.

Thus the first statement in Theorem 1 has been verified. To prove the remainder of Theorem 1, we will need a refinement of Theorem 6.

First, recall that the *rank* of a given $m \times n$ matrix A , denoted by $\text{rank}(A)$, is the size of the largest invertible square submatrix of A . Naturally, the *principal rank* of an $n \times n$ matrix A , denoted by $\text{p-rank}(A)$, is defined as the size of the largest invertible principal submatrix of A . Clearly

$$1 \leq \text{p-rank}(A) \leq \text{rank}(A) \leq n.$$

It is not difficult to show that the principal rank of a TN matrix is equal to the number of its nonzero (positive) eigenvalues. Indeed, if a TN matrix A has p nonzero eigenvalues $\lambda_1, \dots, \lambda_p$, then the characteristic polynomial of A , $\det(\lambda I - A)$, has the form $\det(\lambda I - A) = \lambda^{n-p} \sum_{j=0}^p (-1)^j c_{n-j} \lambda^{p-j}$ and $c_{n-p} = \lambda_1 \cdots \lambda_p > 0$. On the other hand, the coefficient of λ^{n-k} in the characteristic polynomial of A is equal to $(-1)^k$ times the sum of all $k \times k$ principal minors of A . Since all the minors are nonnegative, it follows that all principal minors of A of size greater than p are zero and there is at least one nonzero principal minor of A of size p . In other words, $\text{p-rank}(A) = p$.

Thus, for an $n \times n$ TN matrix A , $(n - \text{p-rank}(A))$ is equal to the sum of the sizes of the Jordan blocks corresponding to the eigenvalue zero. Since $(n - \text{rank}(A))$ is equal to the number of Jordan blocks corresponding to zero, we conclude that, for ITN matrices, if k is the smallest positive integer such that $\text{rank}(A^k) = \text{p-rank}(A)$, then k is equal to the size of the largest Jordan block corresponding to the eigenvalue zero.

Consider the following illustrative example.

Example 7 Consider the $n \times n$ lower Hessenberg (0,1)-matrix

$$H = \begin{bmatrix} 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 1 & \cdots & 1 & 1 \\ 1 & 1 & \cdots & 1 & 1 \end{bmatrix}.$$

Then H is an irreducible TN matrix with $\text{rank}(H) = n - 1$ and $\text{p-rank}(H) = \lceil \frac{n}{2} \rceil$ (cf. [6]).

If an $n \times n$ TN matrix A is realized as $A(\mathcal{D}, W)$ for some weighted planar diagram (\mathcal{D}, W) , then notions of rank, principal rank, and irreducibility can be conveniently interpreted in terms of \mathcal{D} . Namely, A is irreducible if and only if, for every $i = 1, \dots, n$, there exist paths in \mathcal{D} from i to $i \pm 1$ (ignoring $i - 1$, when $i = 1$ and $i + 1$, when $i = n$). Similarly, since rank and principal rank are defined in terms of nonsingular submatrices, it follows that the rank (resp. principal rank) of A can be

interpreted as the largest number of vertex disjoint paths in \mathcal{D} beginning on the left and terminating on the right (resp. the largest number of vertex disjoint paths which begin and terminate in the same index set). Since this interpretation obviously does not depend on particular values of weights, it makes perfect sense to say that a diagram \mathcal{D} is irreducible of rank r and principal rank p whenever some (and therefore all) $A \in \mathcal{M}(\mathcal{D})$ has these properties. The following observation will be particularly useful for us: if $A \in \mathcal{M}(\mathcal{D})$, then, for any positive integer k , the rank of A^k is equal to the rank of the diagram \mathcal{D}^k obtained from \mathcal{D} by gluing left to right k copies of the diagram \mathcal{D} .

We now come to the main result of this section

Theorem 8 *Let \mathcal{D} be an irreducible diagram of order n with rank n . Then, for every $0 < \lambda_1 < \lambda_2 < \dots < \lambda_n$ there is a TN matrix $A \in \mathcal{M}(\mathcal{D})$ with eigenvalues $\lambda_1, \dots, \lambda_n$.*

Proof First observe, that if $A = A(\mathcal{D}, W)$ is in $\mathcal{M}(\mathcal{D})$, then so is DA , where $D = \text{diag}(d_1, \dots, d_n)$ is any positive diagonal matrix. Indeed, to obtain DA , one needs to multiply by d_i every weight in W assigned to an edge starting at the i th source. With this in mind, we will reduce the statement to a solvable case of the multiplicative inverse eigenvalue problem.

Since \mathcal{D} is of rank n we can select a collection of n vertex disjoint paths, P_1, \dots, P_n in \mathcal{D} such that P_i joins the i th source to the i th sink in \mathcal{D} . To every edge that is contained in one of the paths P_i we assign a weight 1 and to every other edge in \mathcal{D} we assign a weight $\epsilon > 0$. Call this weight assignment $W(\epsilon)$. Since paths P_i do not intersect, for every $i \neq j$, a path from i th to j th sink in \mathcal{D} (if it exists) must contain an edge that does not belong to any of the paths P_i . This implies that $A(\mathcal{D}, W(\epsilon)) = I + O(\epsilon)$, where I is the identity matrix, and $O(\epsilon)$ denotes a matrix whose entries are polynomials in ϵ with zero constant terms. Multiplying $A(\mathcal{D}, W(\epsilon))$ from the left by the inverse of its diagonal part, we obtain a matrix $A(\epsilon) = (a_{ij}(\epsilon))_{i,j=1}^n \in \mathcal{M}(\mathcal{D})$ such that $a_{ii}(\epsilon) = 1$ and $a_{ij}(\epsilon) = O(\epsilon)$ for $i \neq j$.

Now, fix n distinct positive numbers $\lambda_1 < \lambda_2 < \dots < \lambda_n$. By choosing ϵ small enough, we can ensure that

$$\frac{\min(\lambda_2 - \lambda_1, \dots, \lambda_n - \lambda_{n-1})}{\lambda_n} > 2 \sum_{i \neq j} a_{ij}(\epsilon).$$

Then Theorem 4.4.11 of [19] guarantees a solution to the multiplicative inverse eigenvalue problem for $A(\epsilon)$ and $\lambda_1, \dots, \lambda_n$, i.e., an existence of a positive diagonal matrix D such that $DA(\epsilon)$ has eigenvalues $\lambda_1, \dots, \lambda_n$. (It should be noted that the theorem we refer to is formulated for positive definite rather than TN matrices. However its proof requires only that every matrix $B = (b_{ij})$ in the set $\{DA(\epsilon) : D \text{ is positive diagonal}\}$ (i) has distinct positive eigenvalues $\mu_1 < \dots < \mu_n$ and (ii) satisfies $\mu_1 \leq b_{ii} \leq \mu_n$. (i) is obvious since B is nonsingular ITN matrix and (ii) is just a consequence of the interlacing inequalities for TN matrices (Theorem 14 in [10] applied to 1×1 principal submatrices of a TN matrix B .) Since $DA(\epsilon) \in \mathcal{M}(\mathcal{D})$, the proof is complete. ■

Corollary 9 *Let \mathcal{D} be an irreducible diagram of order n and of principal rank p . Then, for every $0 < \lambda_1 < \lambda_2 < \dots < \lambda_p$ there is a TN matrix $A \in \mathcal{M}(\mathcal{D})$ with positive eigenvalues $\lambda_1, \dots, \lambda_p$.*

Proof We will fix p and use induction on $n \geq p$. The base case $n = p$ coincides with the statement of Theorem 8. Now let $n > p$ and let P_1 be the lowest path from 1st source to the 1st sink, and let P_n be the highest path from n source to the n th sink. Since $p < n$, the rank of \mathcal{D} is strictly less than n . This means that there exists a “vertical” line L that cuts through \mathcal{D} in such a way that (i) L intersects both P_1 and P_n ; (ii) L intersects every path in \mathcal{D} from i to j at most once; and (iii) the total number, m , of intersections of L with \mathcal{D} is strictly less than n .

Let \mathcal{D}_l and \mathcal{D}_r be, respectively, the left and the right parts into which L cuts \mathcal{D} . Clearly, all sources (resp. sinks) of \mathcal{D} belong to \mathcal{D}_l (resp. \mathcal{D}_r). Construct a new diagram, $\tilde{\mathcal{D}}$ by attaching \mathcal{D}_l to \mathcal{D}_r from the right in such a way that i th source is glued to the i th sink. Then it follows that the order m of $\tilde{\mathcal{D}}$ is strictly less than n , $\tilde{\mathcal{D}}$ is irreducible since \mathcal{D} is, and the principal rank of $\tilde{\mathcal{D}}$ is p .

By the induction assumption, for every choice of $0 < \lambda_1 < \lambda_2 < \dots < \lambda_p$, there is a weight assignment \tilde{W} for $\tilde{\mathcal{D}}$ such that an $m \times m$ matrix $A(\tilde{\mathcal{D}}, \tilde{W})$ has positive eigenvalues $\lambda_1, \dots, \lambda_p$. By our construction, $\tilde{A} = A(\tilde{\mathcal{D}}, \tilde{W})$ can be factored into $\tilde{A} = A(\mathcal{D}_r, \tilde{W}_r)A(\tilde{\mathcal{D}}_l, \tilde{W}_l)$, where weight assignments \tilde{W}_r, \tilde{W}_l are defined in a natural way. But then the matrix $A = A(\mathcal{D}_l, \tilde{W}_l)A(\tilde{\mathcal{D}}_r, \tilde{W}_r)$ has the same positive eigenvalues as $A(\tilde{\mathcal{D}}, \tilde{W})$. On the other hand, $A \in \mathcal{M}(\mathcal{D})$ with a corresponding weight assignment W obtained from \tilde{W} as follows: \mathcal{D} is contained in \mathcal{D}_l or \mathcal{D}_r , then the weight of this edge in W coincides with that in \tilde{W} . If L intersects an edge $v_l \rightarrow v_r$ in \mathcal{D} at an interior point v , then the weight assigned to $v_l \rightarrow v_r$ in W is $w_l w_r$, where w_l (resp. w_r) is the weight of the edge $v_l \rightarrow v$ (resp. $v \rightarrow v_r$) in \tilde{W} . ■

4 Index of the Eigenvalue Zero

Recall that the *index* of an eigenvalue λ corresponding to a matrix A is the smallest positive integer k such that $\text{rank}((A - \lambda I)^k) = \text{rank}((A - \lambda I)^{k+1})$. Equivalently, the index of an eigenvalue is equal to the size of the largest Jordan block corresponding to that eigenvalue. It was observed in [6] that in many examples of ITN matrices, the size of the largest Jordan block corresponding to zero does not exceed the principal rank. Based on the evidence, this relation was conjectured in [6] to be true in general. Moreover, the conjecture was verified for $n \leq 7$ and also in the cases when $\text{p-rank}(A) \geq \lceil \frac{n}{2} \rceil$; when $\text{p-rank}(A) \geq \text{rank}(A) - 1$; and when $\text{p-rank}(A) = 1$ or 2 .

Since for ITN matrices the number of positive eigenvalues is equal to the principal rank, the claim that the index of zero does not exceed the principal rank is equivalent to

$$\text{rank}(A^{\text{p-rank}(A)}) = \text{p-rank}(A).$$

for any ITN matrix A .

The central result of this section is a proof of the conjecture formulated in [6].

Theorem 10 *Let A be an $n \times n$ irreducible TN matrix with a principal rank equal to k*

with $1 \leq k < n$. Then $\text{rank}(A^k) = \text{p-rank}(A) = k$. In particular, the size of the largest Jordan block corresponding to zero is at most k .

Proof We will use induction on k and n . The case of $k = 1$ and arbitrary n was treated in [6]. Suppose that A is an $n \times n$ irreducible TN matrix with principal rank equal to k . Let \mathcal{D} be an irreducible diagram that represents A and P be the highest path from n to n in \mathcal{D} (see Figure 3). Choose the largest possible vertex m

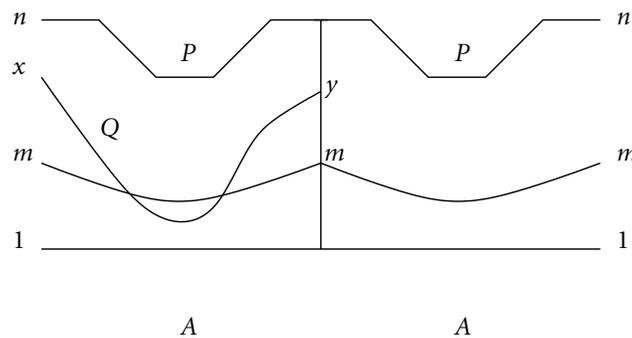


Figure 3: Planar Diagram \mathcal{D} .

($m < n$) such that there exists a path from m to m that does not intersect P . Observe that such a vertex exists because the principal rank of A is greater than 1. Define a subdiagram \mathcal{D}' of \mathcal{D} in the following way. First delete all edges that belong to P or have a common vertex with P . Then \mathcal{D}' is a union of all paths from i to j ($1 \leq i, j \leq m$) in the resulting diagram. Note that \mathcal{D}' is irreducible since \mathcal{D} is. Furthermore, by the maximality of m , the principal rank of \mathcal{D}' is equal to $k - 1$. (A matrix A' represented by \mathcal{D}' should not be confused with an $m \times m$ principal submatrix of A).

Suppose $m < n - 1$, and let x be any vertex such that $m < x < n$. Then one of the following must hold: Any path Q from x must either: (i) intersect P , or (ii) terminate at a sink y , where $y \leq m$.

Suppose neither of these cases hold, *i.e.*, assume Q begins at x and terminates at y with $y > m$, and does not intersect P . Then if $x < y$, it follows that there exists a path T from x to x that does not intersect P . Indeed, since A is irreducible there exists a path R from x to x , and by the maximality of m , this path R must intersect P . Thus R must intersect Q . To construct T (a path from x that does not intersect P), follow R and Q until they part, then follow Q until they meet up again, and then follow R until x (note that R and Q may meet at x). On the other hand if $y < x$, then applying similar reasoning it follows that there exists a path from y to y that does not intersect P . Since both $x, y > m$, both cases contradict the maximality of m .

Observe that an immediate consequence of the above claim is that any two paths that begin and end in $\{m + 1, \dots, n\}$ must intersect. Furthermore, if there is a path

in \mathcal{D} that starts (resp. terminates) in $\{m + 1, \dots, n - 1\}$ and does not intersect P , then there is no path that terminates (resp. starts) in $\{m + 1, \dots, n - 1\}$ and does not intersect P , for if both such paths exist, they have to intersect (since both intersect a path from m to m) and thus, can be used to produce a path that starts and terminates in $\{m + 1, \dots, n - 1\}$ and does not intersect P . Therefore, we can assume that \mathcal{D} does not contain a path that terminates in $\{m + 1, \dots, n - 1\}$ and does not intersect P (otherwise one can deal with the transpose of A instead of A).

Next, consider a matrix A^k . It can be represented by the diagram \mathcal{D}^k , obtained by a concatenation of k copies of \mathcal{D} . Let r be the rank of A^k . Then there exists a family \mathcal{F} of r vertex disjoint paths in \mathcal{D}^k . Since P was chosen to be the highest path from n to n in \mathcal{D} , at most one of these paths can intersect a path P^k defined in \mathcal{D}^k in an obvious way. Without a loss of generality, we may assume that $P^k \in \mathcal{F}$. Let Q be any other path in \mathcal{F} and let Q_i be the part of Q that lies in the i th copy of \mathcal{D} in \mathcal{D}^k (counting from left to right). Since none of the Q_i 's intersect P , by the previous discussion, the end points of Q_i must belong to the index set $\{1, 2, \dots, m\}$. This means that the path obtained by gluing Q_2, \dots, Q_k belongs to $(\mathcal{D}')^{k-1}$. In other words, \mathcal{F} gives rise to a family of $r - 1$ vertex disjoint paths in $(\mathcal{D}')^{k-1}$ and, therefore,

$$r = \text{rank}(A^k) \leq 1 + \text{rank}((A')^{k-1}).$$

By the induction assumption, $\text{rank}((A')^{k-1}) = \text{p-rank}(A') = k - 1$ and we obtain

$$k = \text{p-rank}(A) = \text{p-rank}(A^k) \leq \text{rank}(A^k) \leq k,$$

which completes the proof. ■

The next result is an immediate consequence of Theorem 10.

Theorem 11 *Let A be an $n \times n$ irreducible TN matrix. Then*

$$\text{p-rank}(A) \geq \left\lceil \frac{n}{n - \text{rank}(A) + 1} \right\rceil,$$

Proof For convenience of notation, let $p = \text{p-rank}(A)$, $r = \text{rank}(A)$, and J_0 denote the number of Jordan blocks of A corresponding to 0. Then the inequality above is equivalent to

$$p \geq \left\lceil \frac{n}{J_0 + 1} \right\rceil,$$

since $n - \text{rank}(A)$ represents the nullity of A . By Theorem 10 the size of the largest Jordan block corresponding to 0 is at most p , hence since p is equal to the number of distinct positive eigenvalues of A , it is clear that the minimal possible number of Jordan blocks corresponding to 0 is at least $\lceil \frac{n-p}{p} \rceil$. Consequently, $J_0 \geq \lceil \frac{n-p}{p} \rceil$, which implies $p \geq \lceil \frac{n}{J_0+1} \rceil$. This completes the proof. ■

It is worth noting here that results of this section constitute a proof of the statement (2) of Theorem 1. This claim will be made more precise in the next section.

5 Jordan Structures

For an arbitrary $n \times n$ ITN matrix A of rank r and principal rank p , let m_i ($i = 1, 2, \dots$) denote the number of $i \times i$ Jordan blocks that correspond to the zero eigenvalue in the Jordan canonical form of A . From Theorems 10 and 11 we know that

$$(3) \quad p \geq \frac{n}{n-r+1}$$

and $m_i = 0$ for $i > p$. Furthermore, by Theorem 6, A has p distinct positive eigenvalues $0 < \lambda_1 < \dots < \lambda_p$. It follows that non-negative integers m_1, \dots, m_p are subject to conditions

$$(4) \quad m_1 + \dots + m_p = n - r, \quad m_1 + 2m_2 + \dots + pm_p = n - p.$$

Combining the first equality in (4) with an inequality (3) re-written as $n - r \geq \frac{n-p}{p}$, we obtain the condition (c) of the second part of Theorem 1.

Denote

$$(5) \quad J(A) = (n, r, p; m_1, \dots, m_p).$$

We call $J(A)$ the *combinatorial Jordan data* of A .

A tuple $J = (n, r, p; m_1, \dots, m_p)$ is called *admissible* if it satisfies conditions (3) and (4). Furthermore, J is called *realizable* if there exists an ITN matrix A such that $J = J(A)$. As discussed above, an admissible tuple $J = (n, r, p; m_1, \dots, m_p)$ satisfies conditions (a)–(c) in item (2) of Theorem 1. Hence the only item left to prove is the third part of the main theorem. Given the terminology we have just adopted, item (3) of Theorem 1 can be re-stated as follows.

Theorem 12 For an arbitrary admissible $J = (n, r, p; m_1, \dots, m_p)$ and arbitrary distinct positive numbers $\lambda_1 < \dots < \lambda_p$, there exists an ITN matrix A with the combinatorial Jordan data J and positive eigenvalues $\lambda_1, \dots, \lambda_p$.

The remainder of this section is devoted to proving Theorem 12, and thus completing a proof of our main result.

Lemma 13 If

$$J = (n, r, p; m_1, \dots, m_{p-l}, \underbrace{0, \dots, 0}_l)$$

is admissible, then $J' = (n - l, r - l, p - l; m_1, \dots, m_{p-l})$ is admissible as well.

Proof Since J is admissible,

$$(6) \quad \begin{aligned} m_1 + \dots + m_{p-l} &= n - r = (n - l) - (r - l) \\ m_1 + 2m_2 \dots + (p - l)m_{p-l} &= n - p = (n - l) - (r - l) \end{aligned}$$

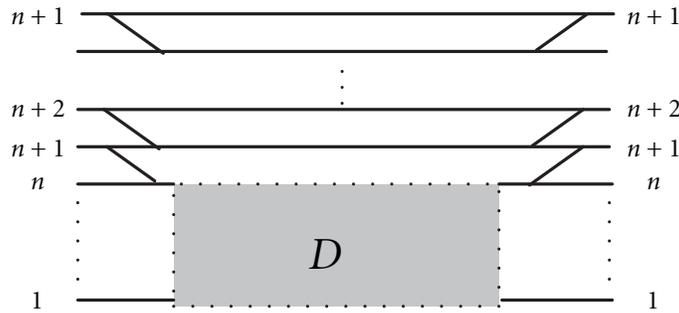
and thus, J' satisfies (4). Moreover, it follows from (6) that $(p-l)(n-r) - (n-p) \geq 0$. Therefore, $(p - l) - \frac{n-l}{n-r+1} = \frac{(p-l)(n-r) - (n-p)}{n-r+1} \geq 0$, and J' is admissible. ■

Lemma 14 If $J = (n, r, p; m_1, \dots, m_p)$ is realizable, then

$$J' = (n + l, r + l, p + l; m_1, \dots, m_p, \underbrace{0, \dots, 0}_l)$$

is realizable as well.

Proof Let A be an ITN matrix such that $J(A) = J$ and D be a planar diagram that represents A . Then the matrix represented by the diagram below is ITN and has a



combinatorial Jordan data J' . ■

Theorem 15 Every admissible $J = (n, r, p; m_1, \dots, m_p)$ with $m_p \neq 0$ is realizable. Moreover, there exists an ITN matrix $A = [a_{ij}]$, such that $J(A) = J$ and $a_{ij} \in \{0, 1\}$ ($i, j = 1, \dots, n$).

Proof We prove the statement by constructing a diagram $D(J)$ that represents an ITN matrix A such that $J(A) = J$ and $a_{ij} \in \{0, 1\}$ ($i, j = 1, \dots, n$). This diagram is built from smaller diagrams of the form in Figure 4.

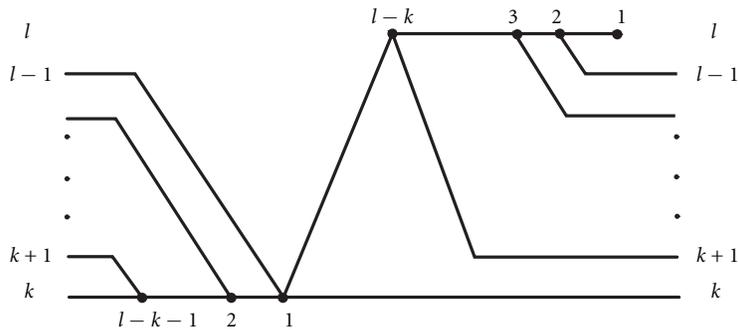


Figure 4: Diagram $D(k, l)$.

Here we assume that $l - k \geq 2$. Given two diagrams $D(k, l)$ and $D(l, m)$, one obtains a new diagram, denoted by $D(k, l, m)$, by gluing the bottom level of $D(l, m)$ to the top level of $D(k, l)$ in such a way that for every $i = 1, \dots, \min(m - l - 1, l - k)$, the block numbered i on the bottom level of $D(l, m)$ is glued to the the block numbered i on the top level of $D(k, l)$. In the same fashion, we construct a diagram $D(k_1, k_2, \dots, k_{s-1}, k_s)$ for arbitrary $k_1 < k_1 + 2 \leq k_2 < \dots < k_{s-1} + 2 \leq k_s$ by gluing $D(k_{s-1}, k_s)$ to $D(k_1, k_2, k_3, \dots, k_{s-1})$ according to the same rule. For example, the diagram $D(1, 3, 5, 8, 11)$ is shown in Figure 5.

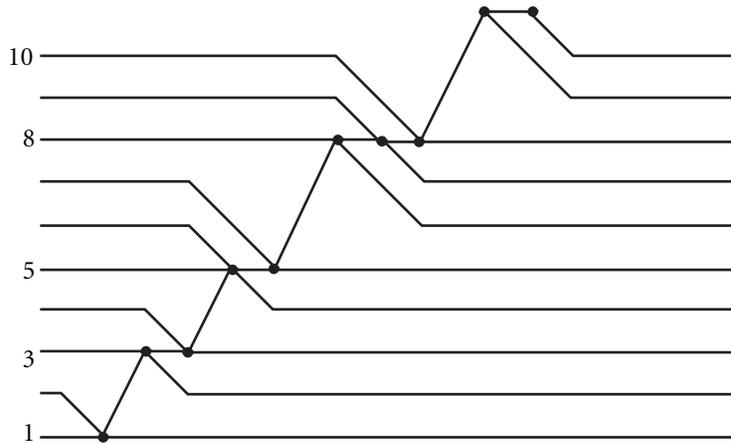


Figure 5: Diagram $D(1, 3, 5, 8, 11)$.

Note that $D(k_1, \dots, k_s)$, in fact, represents a $(k_s - 1) \times (k_s - 1)$ matrix. Let us now put $k_1 = 1, k_s = m + 1$ and consider an $m \times m$ TN matrix $A(k_1, \dots, k_s) = (a_{ij})_{i,j=1}^m$ obtained from $D(k_1, \dots, k_s)$ with weights along all the edges assumed to be equal to 1. A brief inspection of the diagram $D(k, l)$ shows that for every pair i, j of indices, there is at most one path in the diagram $D(k_1, \dots, k_s)$ starting at the level i on the left and ending at the level j on the right. Therefore, every entry of $A(k_1, \dots, k_s)$ is either 0 or 1. Furthermore, it follows from the construction of $D(k_1, \dots, k_s)$, that $a_{1m} = 1$ and $a_{i+1,i} = 1$ for $i = 1, \dots, m - 1$. This implies that $A(k_1, \dots, k_s)$ is irreducible.

Now, let us consider an admissible tuple $J = (n, r, p; m_1, \dots, m_p)$ with $m_p \neq 0$ and define $r_0 = n + 1, r_{p+1} = 0$ and

$$(7) \quad r_i = n - i + 1 - \sum_{\alpha=1}^{i-1} \alpha m_\alpha - i \sum_{\alpha=i}^p m_\alpha \quad (i = 1, \dots, p).$$

Note that $r_1 = r, r_p = 1$ and that solving (7) for m_i results in

$$(8) \quad m_i = r_{i+1} - 2r_i + r_{i-1} \quad (i = 1, \dots, p).$$

This implies, in particular, that

$$(9) \quad r_{i-1} - r_i \geq r_i - r_{i+1} \quad (i = 1, \dots, p).$$

Consider a diagram $D(J) = D(1, r_{p-1}, \dots, r_1, n+1)$ and a matrix $A(J) = A(1, r_{p-1}, \dots, r_1, n+1)$ that corresponds to $D(J)$. For example, the diagram in Figure 5 corresponds to $J = (10, 8, 4; 0, 1, 0, 1)$. We shall prove that

- (i) $\text{rank}(A(J)) = r$;
- (ii) $\text{rank}(A(J)^2) = 1 + \text{rank}(A(1, r_{p-1}, \dots, r_1))$.

To prove (i), let us observe that, by construction of $D(J)$, every path that starts above or at the level r_i terminates strictly above the level r_{i+1} . Moreover, due to inequalities (9), for every $i = 1, \dots, p-1$, indices $r_{i-1} - 1, \dots, r_{i+1} - r_i + r_{i-1}$ on the left are joined by a collection of vertex-disjoint paths with indices $r_i, \dots, r_{i+1} + 1$ on the left. Combining these two observations, we see that there is a collection of vertex-disjoint paths whose right end points are indices 1 through $r_1 = r$. On the other hand, any two paths that end at or above r_1 intersect. Thus, $\text{rank}(A(J)) = r$.

To establish (ii), recall that the rank of $A(J)^2$ is equal to the maximal number of vertex-disjoint paths in the diagram $D(J)^2$ obtained by concatenation of two copies of $D(J)$. Let P be the highest possible path from n to n in $D(J)$. Then the concatenation of two copies of P , which we denote by P^2 , is the highest possible path from n to n in $D(J)^2$. By construction every path in $D(J)$ that does not intersect P , terminates below the level r_1 and if we want this path to continue through $D(J)^2$ without intersecting P^2 , then its part that belongs to the second copy of $D(J)^2$ must belong to the sub-diagram $D(1, r_{p-1}, \dots, r_1)$. Thus the left-hand side in (ii) cannot exceed the right-hand side. To show that they are equal, simply repeat the argument used in proving (i).

It follows from (i) and (ii) that $\text{rank}(A(J)^2) = 1 + \text{rank}(A(1, r_{p-1}, \dots, r_1)) = 1 + r_2$. Arguing in exactly the same way implies that

$$\text{rank}(A(J)^i) = i - 1 + r_k \quad (k = 1, \dots, p)$$

and so, by (7),

$$n - \text{rank}(A(J)^i) = n - i + 1 - r_i = \sum_{\alpha=1}^{i-1} \alpha m_\alpha + i \sum_{\alpha=i}^p m_\alpha \quad (k = 1, \dots, p)$$

The left-hand side of the last equality describes the number of Jordan blocks that correspond to the zero eigenvalue in the Jordan form of $A(J)^i$, while the right-hand side corresponds with the number of zero Jordan blocks of any matrix with a combinatorial Jordan data J . This completes the proof. ■

Theorem 16 Every admissible J is realizable.

Proof Let l be the smallest number such that $m_{p-l} \neq 0$. By Lemma 13, if

$$J = (n, r, p; m_1, \dots, m_{p-l}, \underbrace{0, \dots, 0}_l)$$

is admissible then $J' = (n-l, r-l, p-l; m_1, \dots, m_{p-l})$ is admissible and therefore, by the previous theorem, J' is realizable. Then Lemma 14 implies that J is realizable as well. ■

Now to complete a proof of Theorem 12, consider a diagram \mathcal{D} that realizes admissible combinatorial Jordan data J . Then, by Corollary 9, for arbitrary $0 < \lambda_1 < \dots < \lambda_p$, there exists a matrix $A \in \mathcal{M}(\mathcal{D})$ with positive eigenvalues $\lambda_1, \dots, \lambda_p$. Since every matrix in $\mathcal{M}(\mathcal{D})$ has the same combinatorial Jordan data, the proof is complete.

Acknowledgments This article continues our previous joint work with Charles Johnson [6]. We would like to thank Charles for introducing us to this problem and for his constant interest and encouragement. We also thank the referee for helpful suggestions on how to improve the presentation.

References

- [1] T. Ando, *Totally positive matrices*. Linear Algebra Appl. **90**(1987), 165–219.
- [2] A. Berenstein, S. Fomin and A. Zelevinsky, *Parameterizations of canonical bases and totally positive matrices*. Adv. Math. **122**(1996), 49–149.
- [3] F. Brenti, *Combinatorics and total positivity*. J. Combin. Theory Ser. A **71**(1996), 175–218.
- [4] C. W. Cryer, *Some properties of totally positive matrices*. Linear Algebra and Appl. **15**(1976), 1–25.
- [5] S. M. Fallat, *Bidiagonal factorizations of totally nonnegative matrices*. Amer. Math. Monthly **109**(2001), 697–712.
- [6] S. M. Fallat, M. I. Gekhtman, and C. R. Johnson, *Spectral structures of irreducible totally nonnegative matrices*. SIAM J. Matrix Anal. Appl. **22**(2000), 627–645.
- [7] S. Fomin and A. Zelevinsky, *Total positivity: tests and parameterizations*. Math. Intelligencer **22**(2000), 23–33.
- [8] ———, *Double Bruhat cells and total positivity*. J. Amer. Math. Soc. **12**(1999), 335–380.
- [9] F. R. Gantmacher and M. G. Krein, *Sur les matrices complement non-negatives et oscillatoires*. Comp. Math. **4**(1937), 445–476.
- [10] F. R. Gantmacher and M. G. Krein, *Oscillation matrices and kernels and small vibrations of mechanical systems*, AMS, Providence, RI, 2002.
- [11] M. Gasca and C. A. Micchelli, eds. *Total positivity and its applications*, Kluwer Academic, Dordrecht, 1996.
- [12] M. Gasca and J.M. Peña, *On factorizations of totally positive matrices*. In: Total positivity and its applications, Kluwer Academic, Dordrecht, 1996. pp. 109–130.
- [13] I. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*. Adv. in Math. **58**(1985), 300–321.
- [14] C. Loewner, *On totally positive matrices*. Math. Z. **63**(1955), 338–340.
- [15] S. Karlin, *Total positivity, I*, Stanford University Press, Stanford, 1968.
- [16] S. Karlin and J. McGregor, *Coincidence probabilities*. Pacific J. Math. **9**(1959), 1141–1164.
- [17] G. Lusztig, *Total positivity in reductive groups*. In: Lie theory and geometry, Birkhäuser, Boston, MA, 1994, pp. 531–568.

- [18] A. Whitney, *A Reduction theorem for totally positive matrices*. J. Analyse Math. **2**(1952), 88–92.
- [19] Shu-fang Xu, *An introduction to inverse algebraic eigenvalue problems*, Peking University Press, Beijing, 1998.

Department of Mathematics and Statistics
University of Regina
Regina, Saskatchewan
S4S 0A2
e-mail: sfallat@math.uregina.ca

Department of Mathematics
University of Notre Dame
Notre Dame, IN. 46556-5683
U.S.A.
e-mail: Michael.Gekhtman.1@nd.edu