Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 422 (2007) 164–185
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Alexandrov's inequality and conjectures on some Toeplitz matrices ☆

Ivo Klemeš

Department of Mathematics and Statistics, McGill University, 805 Sherbrooke Street West, Montréal, Qué., Canada H3A 2K6

Received 1 April 2006; accepted 24 September 2006
Available online 16 November 2006
Submitted by S. Fallat

Abstract

We study determinant inequalities for certain Toeplitz-like matrices over **C**. For fixed n and $N \geqslant 1$, let Q be the $n \times (n + N - 1)$ zero–one Toeplitz matrix with $Q_{ij} = 1$ for $0 \leqslant j - i \leqslant N - 1$ and $Q_{ij} = 0$ otherwise. We prove that $\det(QQ^*)$ is the minimum of $\det(RR^*)$ over all complex matrices R with the same dimensions as Q satisfying $|R_{ij}| \geqslant 1$ whenever $Q_{ij} = 1$ and $R_{ij} = 0$ otherwise. Although R has a Toeplitz-like band structure, it is not required to be actually Toeplitz. Our proof involves Alexandrov's inequality for polarized determinants and its generalizations. This problem is motivated by Littlewood's conjecture on the minimum 1-norm of N-term exponential sums on the unit circle. We also discuss polarized Bazin–Reiss–Picquet identities, some connections with k-tree enumeration, and analogous conjectured inequalities for the elementary symmetric functions of QQ^*.

AMS classification: 15A15; 15A57; 42A05; 05C50

Keywords: Toeplitz matrix; Exponential sum; 1-norm; Inequality; Determinant; Mixed discriminant; Minor; Tree

1. Introduction

Let $f(z) = f_0 + f_1 z + f_2 z^2 + \cdots + f_m z^m$ be a complex polynomial of degree m and let n be a positive integer. Define the nth Toeplitz matrix of f to be the $n \times (n + m)$ matrix $T_n(f)$

☆ Research supported by the Natural Sciences and Engineering Research Council of Canada.
E-mail address: klemes@math.mcgill.ca

0024-3795/$ - see front matter
doi:10.1016/j.laa.2006.09.020

with $(T_n(f))_{ij} = f_{j-i}$ for $0 \leqslant j - i \leqslant m$ and $(T_n(f))_{ij} = 0$ otherwise. For each integer $N \geqslant 1$ let δ_N denote the polynomial

$$\delta_N(z) = 1 + z + z^2 + \cdots + z^{N-1}. \tag{1}$$

Thus, one example of a $T_n(f)$ is

$$T_5(\delta_4) = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

In general, a matrix M is said to be *Toeplitz* if $M_{ij} = M_{kl}$ whenever $(i - j) = (k - l)$. Clearly, $T_n(f)$ is Toeplitz for any f.

For a complex rectangular matrix R, denote by R^* the conjugate transpose. The matrix RR^* is then Hermitian. Given any complex $n \times n$ matrix A, denote by $S_k(A)$, $1 \leqslant k \leqslant n$, the degree k elementary symmetric polynomials of the eigenvalues of A. The S_k satisfy the identity

$$\det(I + tA) = 1 + \sum_{k=1}^{n} S_k(A)t^k. \tag{2}$$

In particular $S_1(A) = \operatorname{tr}(A)$ and $S_n(A) = \det(A)$. We will consider the following problem from [11, Conjecture 2.4] (where it was posed for the case $|f_{m_j}| = 1$).

Conjecture 1.1. *Let $N \geqslant 1$ and $0 \leqslant m_0 < \cdots < m_{N-1}$ be integers. Let $f(z) = \sum_{j=0}^{N-1} f_{m_j} z^{m_j}$ be a polynomial with complex coefficients f_{m_j} satisfying $|f_{m_j}| \geqslant 1$. If $n \geqslant 1$, $R = T_n(f)$, and $Q = T_n(\delta_N)$, then*

$$S_k(RR^*) \geqslant S_k(QQ^*) \tag{3}$$

for all $1 \leqslant k \leqslant n$.

It was noted in [11] that this would imply a sharp form of the Littlewood conjecture on the 1-norm of exponential sums. The Littlewood conjecture (in the non-sharp form) is the statement that for some $C > 0$, for all N, the above polynomials f satisfy $\|f\|_1 \geqslant C \log N$ on the unit circle, $z = \mathrm{e}^{\mathrm{i}\theta}$. This was proved by Konyagin [12] and independently by McGehee et al. [16]. The sharp form is the conjecture that $\|f\|_1 \geqslant \|\delta_N\|_1$, which is still unsolved. Conjecture 1.1 was proved in [11, Theorem 1.2] for the case $f(z) = \sum_{j=0}^{N-1} \pm z^j$, in other words when $m_j = j$ ("no gaps"), and the coefficients are ± 1. The proof gave the stronger result that the inequalities (3) hold for all k, for all R satisfying $R_{ij} = \pm Q_{ij}$ (arbitrary signs at each (i, j), thus not necessarily Toeplitz). Moreover, the same proof works for arbitrary odd integers in place of ± 1, or more generally any integers which are congruent to 1 modulo some fixed integer $m > 1$. The essence of the proof was to note that any square submatrix of Q has determinant 0, 1, or -1 (see [5, p. 853], or [11, Lemma 2.3]). A matrix with this property is said to be totally unimodular.

Definition. Let $N, n \geqslant 1$ be integers and let Q be the $n \times (n + N - 1)$ matrix $Q = T_n(\delta_N)$. Then $\mathscr{R}_n(N)$ denotes the set of all complex $n \times (n + N - 1)$ matrices R such that $|R_{ij}| \geqslant 1$ whenever $Q_{ij} = 1$, and $R_{ij} = 0$ whenever $Q_{ij} = 0$. (Note that R is not required to be Toeplitz.)

The following result concerning Conjecture 1.1 will be proved in this paper.

Theorem 1.2. *Let $N, n \geqslant 1$ and let $Q = T_n(\delta_N)$. Then*

$$\det(RR^*) \geqslant \det(QQ^*) \tag{4}$$

for all $R \in \mathscr{R}_n(N)$.

The special case $R = T_n(f)$ of Theorem 1.2 gives the case $k = n$ in (3) of Conjecture 1.1 for polynomials f having "no gaps", i.e. with $m_j = j$.

The proof of Theorem 1.2 is given in §3. We devote §2 to a review of polarized determinants (also called "mixed discriminants"; see [2]; [20, Section 2.5]), and to some generalizations of Alexandrov's inequality [1]; [20, Theorem 6.8.1]; [6, Theorem 4]; [10, Theorem 3], as far as will be needed for the proof. There is a vague similarity between (4) and the van der Waerden Conjecture, which states that $\mathrm{per}(U) \geqslant \mathrm{per}(\mathbf{1}/n)$ for any doubly stochastic $n \times n$ matrix U, where $\mathbf{1}$ is the $n \times n$ all ones matrix. By a curious coincidence, the solution of the latter also relied upon Alexandrov's inequality [20, p. 388, Note 1]. By an additional coincidence, it was solved at about the same time ($\approx$1980) as the Littlewood conjecture discussed above.

In §4.1, a gap version of Theorem 1.2 will also be discussed. In §4.2, we give an example $R \in \mathscr{R}_n(N)$ with $N = 4$ and $n = 7$ showing that the inequality $S_k(RR^*) \geqslant S_k(QQ^*)$ fails for some elementary symmetric function S_k, $k < n$. Since our example R is not Toeplitz, it is not a counter-example to Conjecture 1.1. In §4.4 we sketch methods and results of a more combinatorial nature. In particular, we show how "polarized Bazin–Reiss–Picquet identities" can be used in place of Alexandrov inequalities in some cases of the proof of Theorem 1.2.

2. Polarized determinants

For a fixed positive integer n, the polarized determinant D_n (also called the "mixed discriminant") is the function on ordered n-tuples $(A_1, \dots, A_n)$ of $n \times n$ complex matrices defined by $D_n(A_1, \dots, A_n) = \frac{1}{n!}$ times the coefficient of $\lambda_1 \cdots \lambda_n$ in $\det(\lambda_1 A_1 + \cdots + \lambda_n A_n)$, where the λ_i are scalars [20, Section 2.5]; [2]. If n is clear from the context, we may write D instead of D_n. D is symmetric and multilinear in the A_i, and D is real valued if all the A_i are Hermitian. Moreover, $D \geqslant 0$ if each A_i is Hermitian and nonnegative. (This can be seen by representing each A_i as $A_i = M_i M_i^*$ and applying Remark 2.2.) It is convenient to abbreviate $D(A_1, \dots, A_1, A_2, \dots, A_2, \dots, A_k, \dots, A_k)$ to $D(A_1^{\langle n_1 \rangle}, A_2^{\langle n_2 \rangle}, \dots, A_k^{\langle n_k \rangle})$ if each A_i occurs $n_i \geqslant 0$ times, with $n_1 + \cdots + n_k = n$. Then, a "multinomial theorem" holds for any number $k \geqslant 1$ of $n \times n$ matrices A_i;

$$\begin{aligned} &D_n((\lambda_1 A_1 + \cdots + \lambda_k A_k)^{\langle m \rangle}, B_1, \dots, B_l) \\ &\quad = \sum_{m_1 + \cdots + m_k = m} \frac{m!}{m_1! \cdots m_k!} D_n(A_1^{\langle m_1 \rangle}, \dots, A_k^{\langle m_k \rangle}, B_1, \dots, B_l) \lambda_1^{m_1} \cdots \lambda_k^{m_k}, \end{aligned} \tag{5}$$

where $m + l = n$ and $B_1, \dots, B_l$ are any $n \times n$ matrices. The case $l = 0$ in (5) gives

$$\begin{aligned} &\det(\lambda_1 A_1 + \cdots + \lambda_k A_k) = D_n((\lambda_1 A_1 + \cdots + \lambda_k A_k)^{\langle n \rangle}) \\ &\quad = \sum_{n_1 + \cdots + n_k = n} \frac{n!}{n_1! \cdots n_k!} D_n(A_1^{\langle n_1 \rangle}, \dots, A_k^{\langle n_k \rangle}) \lambda_1^{n_1} \cdots \lambda_k^{n_k}. \end{aligned} \tag{6}$$

Remark 2.1. If $\operatorname{rank}(A_1) < m_1 \leqslant n$ then

$$D_n(A_1^{\langle m_1\rangle}, A_2, \ldots) = 0 \tag{7}$$

for any $n \times n$ matrices $A_1, A_2, \ldots, A_{n-m_1+1}$.

The remark follows directly from the definition of D and the fact that any m_1 columns of A_1 must be linearly dependent.

Now suppose that R is an $n \times M$ complex matrix. We recall the Binet–Cauchy "expansion"

$$\det(RR^*) = \sum |\det(S)|^2,$$

where the sum is over all $n \times n$ submatrices S of R [14, p. 503]. Let $a_1, \ldots, a_k \geqslant 0$ be integers with $\sum a_i = M$, and let R be partitioned as $R = [R(1)|\cdots|R(k)]$ where $R(i)$ is $n \times a_i$. The $(a_1, \ldots, a_k)$ will be called the "block lengths" defining the partition. Given integers $n_1, \ldots, n_k \geqslant 0$ with $\sum n_i = n$ and $n_i \leqslant a_i$, define

$$\left(\begin{matrix} a_1 \\ n_1 \end{matrix}\,\middle|\,\begin{matrix} \cdots \\ \cdots \end{matrix}\,\middle|\,\begin{matrix} a_k \\ n_k \end{matrix}\right)_R := \sum_{(S_1,\ldots,S_k)} |\det[S_1|\cdots|S_k]|^2, \tag{8}$$

where the sum is over all $(S_1, \ldots, S_k)$ such that each block S_i consists of some n_i columns from $R(i)$. Furthermore, given integers $b_i \leqslant a_i$ with $\sum b_i = M - n$, it will be convenient to define

$$\left(\begin{matrix} a_1 \\ \widehat{b_1} \end{matrix}\,\middle|\,\begin{matrix} \cdots \\ \cdots \end{matrix}\,\middle|\,\begin{matrix} a_k \\ \widehat{b_k} \end{matrix}\right)_R := \left(\begin{matrix} a_1 \\ a_1 - b_1 \end{matrix}\,\middle|\,\begin{matrix} \cdots \\ \cdots \end{matrix}\,\middle|\,\begin{matrix} a_k \\ a_k - b_k \end{matrix}\right)_R,$$

viewing this as a sum like (8) but with each S_i defined by *deleting* some b_i columns from block $R(i)$. For example, if $R = T_7(\delta_4)$ then

$$\left(\begin{matrix} 4 \\ \widehat{2} \end{matrix}\,\middle|\,\begin{matrix} 2 \\ \widehat{0} \end{matrix}\,\middle|\,\begin{matrix} 4 \\ \widehat{1} \end{matrix}\right)_R = \left(\begin{matrix} 4 \\ 2 \end{matrix}\,\middle|\,\begin{matrix} 2 \\ 2 \end{matrix}\,\middle|\,\begin{matrix} 4 \\ 3 \end{matrix}\right)_R$$

is $\sum |\det(S)|^2$ summed over all 7×7 matrices S obtained by deleting 2 columns from the first block $R(1)$, 0 columns from the second block $R(2)$, and 1 column from the third block $R(3)$, in the following partition of R:

$$R = [R(1)|R(2)|R(3)] = \left[\begin{array}{cccc|cc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{array}\right]. \tag{9}$$

Remark 2.2. Let $A_i = R(i)R(i)^*$. If $\sum n_i = n$ and $0 \leqslant n_i \leqslant a_i$, then

$$\left(\begin{matrix} a_1 \\ n_1 \end{matrix}\,\middle|\,\begin{matrix} \cdots \\ \cdots \end{matrix}\,\middle|\,\begin{matrix} a_k \\ n_k \end{matrix}\right)_R = \frac{n!}{n_1!\cdots n_k!} D_n\big(A_1^{\langle n_1\rangle}, \ldots, A_k^{\langle n_k\rangle}\big). \tag{10}$$

Hence, if $\sum b_i = M - n$ and $0 \leqslant b_i \leqslant a_i$, then

$$\left(\begin{matrix} a_1 \\ \widehat{b_1} \end{matrix}\,\middle|\,\begin{matrix} \cdots \\ \cdots \end{matrix}\,\middle|\,\begin{matrix} a_k \\ \widehat{b_k} \end{matrix}\right)_R = \frac{n!}{(a_1 - b_1)!\cdots(a_k - b_k)!} D_n\big(A_1^{\langle a_1 - b_1\rangle}, \ldots, A_k^{\langle a_k - b_k\rangle}\big). \tag{11}$$

Moreover, if some $n_i > a_i$ then the right hand side of (10) is 0.

Proof. The proof is a well-known argument. Eq. (10) results from expanding (6) using the Binet–Cauchy expansion instead of the multinomial theorem, then comparing coefficients of $\lambda_1^{n_1}\cdots\lambda_k^{n_k}$ (provided that we first re-parametrize by setting $\lambda_i = t_i^2$). To do this, insert real parameters $(t_1,\ldots,t_k) =: t$ into the blocks $R(i)$, obtaining the matrix $R_t = [t_1R(1)|\cdots|t_kR(k)]$. Observe that $R_tR_t^* = t_1^2A_1 + \cdots + t_k^2A_k$. By the Binet–Cauchy expansion, $\det(R_tR_t^*) = \sum|\det(S)|^2$ where the sum is over all $n\times n$ submatrices S of R_t. Clearly, those S which contribute to the coefficient of $t_1^{2n_1}\cdots t_k^{2n_k}$ correspond precisely to the column choices described in (8), and this proves (10). Moreover, since each $R(i)$ is an $n\times a_i$ matrix, the term $t_1^{2n_1}\cdots t_k^{2n_k}$ can occur in the Binet–Cauchy expansion only if each $n_i \leqslant a_i$. This proves the last part of the remark. (Alternatively, this follows from Remark 2.1 and the fact that $\operatorname{rank}(A_i)\leqslant a_i$.) □

Let us record the case when all $t_i = 1$ in the above proof

$$\det(RR^*) = \sum|\det(S)|^2 = \sum_{n_1+\ldots+n_k=n}\begin{pmatrix} a_1 & \cdots & a_k \\ n_1 & \cdots & n_k\end{pmatrix}_R = \sum_{b_1+\cdots+b_k=M-n}\begin{pmatrix} a_1 & \cdots & a_k \\ \widehat{b_1} & \cdots & \widehat{b_k}\end{pmatrix}_R, \tag{12}$$

where the a_i are fixed with $\sum a_i = M$, and the summations are over the indices satisfying $\sum n_i = n$, $0\leqslant n_i\leqslant a_i$ and $\sum b_i = M-n$, $0\leqslant b_i\leqslant a_i$ respectively (which may be called the "admissible" indices for this partition). This identity is merely a grouping of terms in the sum $\sum|\det(S)|^2$. Now suppose that a given block of the partition, say $R(1)$, is partitioned further as $R(1) = [C|D]$ where C is $n\times c$, D is $n\times d$, and $c+d = a_1$. Then clearly

$$\begin{pmatrix} a_1 & a_2 & \cdots & a_k \\ \widehat{b_1} & \widehat{b_2} & \cdots & \widehat{b_k}\end{pmatrix}_R = \sum_{i+j=b_1}\begin{pmatrix} c & d & a_2 & \cdots & a_k \\ \widehat{i} & \widehat{j} & \widehat{b_2} & \cdots & \widehat{b_k}\end{pmatrix}_R, \tag{13}$$

where the sum is restricted to $0\leqslant i\leqslant c$ and $0\leqslant j\leqslant d$.

The next theorem is a special case of a result of Gårding [6, Theorem 4] on polarized hyperbolic polynomials, as pointed out by Khovanskiĭ in [10, Theorem 3]. The case $k = 2$ (for real symmetric matrices) is part of a theorem of Alexandrov [1]; [20, Theorem 6.8.1]. In these references the results involve certain strictly positive (that is, positive definite) matrices or forms. The nonnegative (= positive semidefinite) cases that we need, such as the next theorem, follow by standard continuity arguments (using Rouché's theorem). A more detailed study of the nonnegative case can be found in [2,18].

Theorem 2.3 [6,10]. *Let $n\geqslant k\geqslant 1$. If A is any Hermitian $n\times n$ matrix and $B, C_1,\ldots,C_{n-k}$ are nonnegative Hermitian $n\times n$ matrices, then all the zeros of the polynomial $p(\lambda) = D_n((A+\lambda B)^{\langle k\rangle}, C_1,\ldots,C_{n-k})$ are on the real line (or p is the zero polynomial).*

In the case $k = 2$, the polynomial p has real zeros if and only if

$$D_n(A,B,C_1,\ldots,C_{n-2})^2 \geqslant D_n(A,A,C_1,\ldots,C_{n-2})D_n(B,B,C_1,\ldots,C_{n-2}). \tag{14}$$

We will call this Alexandrov's inequality under the hypotheses of Theorem 2.3.

Corollary 2.4. *If $n\geqslant k\geqslant 1$ and $A, B, C_1,\ldots,C_{n-k}$ are all nonnegative Hermitian $n\times n$ matrices, then all the zeros of the polynomial $p(\lambda) = D_n((A+\lambda B)^{\langle k\rangle}, C_1,\ldots,C_{n-k})$ are in the real interval $(-\infty, 0]$, or p is the zero polynomial.*

To deduce the corollary from Theorem 2.3, note that in the corollary all the coefficients of p are nonnegative, so that a strictly positive real number cannot be a zero of p if p is not the zero polynomial.

The following result is well-known.

Theorem 2.5. *Let* $p(\lambda) = \sum_{i=0}^{d} \binom{d}{i} c_i \lambda^i$ *be a polynomial (not necessarily of degree* d*) such that all the zeros of* p *are in the real interval* $(-\infty, 0]$, *or* p *is the zero polynomial. Then the sequence* $\log |c_i|$ *is concave (using the convention* $\log(0) = -\infty$ *if necessary). In particular, if* $d \geqslant 1$ *then*

$$|c_i| \geqslant |c_0|^{\frac{d-i}{d}} |c_d|^{\frac{i}{d}}, \quad 0 \leqslant i \leqslant d.$$

Theorem 2.5 follows easily from Alexandrov's inequality (14) by noting that $p(\lambda)$ is a constant multiple of $\det(K + \lambda I)$ where $-K$ is a diagonal matrix containing the roots of p on its diagonal, and I is an identity matrix. Alternatively, a direct proof is given in [7, Theorem 51]. We remark that although d was allowed to be bigger than the degree of p in Theorem 2.5, the conclusion (that $\{\log |c_i|\}$ is concave) gets stronger as d approaches the degree. Similarly, c_0 is allowed to be zero, but the conclusion gets stronger if the polynomial is "shifted" by considering $p(\lambda)/\lambda$ instead of $p(\lambda)$. This will be seen in the next Alexandrov-type inequality, Corollary 2.6, which is stronger than what would be obtained from Alexandrov's inequality directly (see §4.3).

We continue with the notation used above, whereby R is any complex $n \times M$ matrix, we are given integers $a_1, \ldots, a_k \geqslant 0$ with $\sum a_i = M$, R is partitioned as $R = [R(1)| \cdots |R(k)]$ where $R(i)$ is $n \times a_i$, and $b_1, \ldots, b_k \geqslant 0$ are integers with $\sum b_i = M - n$ and $b_i \leqslant a_i$.

Corollary 2.6. *Let* $k \geqslant 2$ *and suppose that* $1 \leqslant (b_1 + b_2) \leqslant \min(a_1, a_2)$. *Then*

$$\left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{b_1} & \widehat{b_2} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array} \right)_R \geqslant \frac{(b_1+b_2)!}{b_1! b_2!} \left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{(b_1+b_2)} & \widehat{0} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array} \right)_R^{b_1/(b_1+b_2)}$$
$$\times \left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{0} & \widehat{(b_1+b_2)} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array} \right)_R^{b_2/(b_1+b_2)}$$

Proof. Consider the polynomial

$$f(\lambda) = D_n\big((A_1 + \lambda A_2)^{\langle (a_1+a_2)-(b_1+b_2) \rangle}, A_3^{\langle a_3-b_3 \rangle}, \ldots, A_k^{\langle a_k-b_k \rangle}\big),$$

where $A_i = R(i)R(i)^*$. By Corollary 2.4, the zeros of f are all in $(-\infty, 0]$, or f is identically zero. We now look at the coefficients of f. Let $m = (a_1 + a_2) - (b_1 + b_2)$, and let $c_i = a_i - b_i$. Then

$$f(\lambda) = \sum_{r+s=m} \frac{m!}{r!s!} D_n\left(A_1^{\langle r \rangle}, (\lambda A_2)^{\langle s \rangle}, A_3^{\langle c_3 \rangle}, \ldots, A_k^{\langle c_k \rangle}\right).$$

But the terms with $r > a_1$ or $s > a_2$ are zero (by Remark 2.1), hence we can restrict the summation to the interval $m - a_1 \leqslant s \leqslant a_2$. Also, using (11) we have

$$D_n(A_1^{\langle r \rangle}, A_2^{\langle s \rangle}, A_3^{\langle c_3 \rangle}, \ldots, A_k^{\langle c_k \rangle}) = \frac{r!s!c_3! \cdots c_k!}{n!} \left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{a_1 - r} & \widehat{a_2 - s} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array} \right)_R$$

Thus, defining $c := \frac{m!c_3! \ldots c_k!}{n!}$, we have

$$f(\lambda) = c \sum_{r+s=m} \left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{a_1 - r} & \widehat{a_2 - s} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array} \right)_R \lambda^s$$

where the summation is over the interval $m - a_1 \leqslant s \leqslant a_2$, that is $a_2 - (b_1 + b_2) \leqslant s \leqslant a_2$. Making the change of indices $t := a_1 - r$ and $u := a_2 - s$ we see that $t + u = (b_1 + b_2)$, the range is $0 \leqslant t \leqslant (b_1 + b_2)$, and $s = t + a_2 - (b_1 + b_2)$. Hence

$$f(\lambda) = c\lambda^{a_2-(b_1+b_2)} \sum_{t+u=(b_1+b_2)} \left(\begin{array}{c|c|c|c|c} a_1 & a_2 & a_3 & \cdots & a_k \\ \widehat{t} & \widehat{u} & \widehat{b_3} & \cdots & \widehat{b_k} \end{array}\right)_R \lambda^t$$

Letting $p(\lambda) = f(\lambda)/(c\lambda^{a_2-(b_1+b_2)})$ we can apply Theorem 2.5 to the (normalized) coefficients of p. In the case of the coefficient of λ^t with $t = b_1$, this is the desired inequality since that coefficient must be divided by the binomial coefficient $\frac{(b_1+b_2)!}{b_1!b_2!}$, whereas the first and last binomial coefficients are both equal to 1. □

Corollary 2.6 can be applied to any two blocks of a partition instead of $R(1)$, $R(2)$, since the order in which blocks are labelled is immaterial here. We want to apply the result to the first and last blocks $R(1)$, $R(3)$ of certain 3-block partitions $[R(1)|R(2)|R(3)]$ of R defined as follows. If $R \in \mathscr{R}_n(N)$, let $x = \min(n-1, N)$ and $y = (n + N - 1) - 2x$. Then define the three block lengths of the partition by $(a_1, a_2, a_3) := (x, y, x)$. An example with $n - 1 > N$ was exhibited in (9). The following is an example with $n - 1 < N$. Let $R = Q = T_4(\delta_7)$. Then $n = 4$, $N = 7$, $x = 3$, $y = 4$, so that $(a_1, a_2, a_3) = (3, 4, 3)$ and the partition looks like

$$R = [R(1)|R(2)|R(3)] = \left[\begin{array}{ccc|cccc|ccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array}\right]. \tag{15}$$

Exercise. If $x = \min(n-1, N)$ and $y = (n + N - 1) - 2x$ then either $x = N$ or $x + y = N$.

Application of Corollary 2.6 to the first and last blocks of such partitions gives the following.

Lemma 2.7. *Let $n, N \geqslant 1$, $R \in \mathscr{R}_n(N)$, $x = \min(n-1, N)$, and $y = (n + N - 1) - 2x$. If $a, b, c \geqslant 0$ are integers with $a + b + c = N - 1$, $1 \leqslant a + c \leqslant x$, and $b \leqslant y$, then*

$$\left(\begin{array}{c|c|c} x & y & x \\ \widehat{a} & \widehat{b} & \widehat{c} \end{array}\right)_R \geqslant \frac{(a+c)!}{a!c!} \left(\begin{array}{c|c|c} x & y & x \\ \widehat{a+c} & \widehat{b} & \widehat{0} \end{array}\right)_R^{a/(a+c)} \left(\begin{array}{c|c|c} x & y & x \\ \widehat{0} & \widehat{b} & \widehat{a+c} \end{array}\right)_R^{c/(a+c)}.$$

The lemma will be useful because the inequality becomes equality when $R = Q$ (see Corollary 3.8). The lemma can also be iterated, by the next remark.

Remark 2.8. Let $n, N \geqslant 1$, $R \in \mathscr{R}_n(N)$, $x = \min(n-1, N)$, and $y = (n + N - 1) - 2x$. If $a, b \geqslant 0$ are integers with $a + b = N - 1$, $a \leqslant x$, and $b \leqslant y$, then

$$\left(\begin{array}{c|c|c} x & y & x \\ \widehat{a} & \widehat{b} & \widehat{0} \end{array}\right)_R \geqslant \left(\begin{array}{c|c} x & y \\ \widehat{a} & \widehat{b} \end{array}\right)_S \quad \text{and} \quad \left(\begin{array}{c|c|c} x & y & x \\ \widehat{0} & \widehat{b} & \widehat{a} \end{array}\right)_R \geqslant \left(\begin{array}{c|c} y & x \\ \widehat{b} & \widehat{a} \end{array}\right)_T,$$

where S denotes the submatrix of R defined by the intersection of the first $n - x$ rows with the first $x + y$ columns, and T denotes the submatrix of R defined by the intersection of the last $n - x$ rows with the last $x + y$ columns. Moreover, S and T are in $\mathscr{R}_{n-x}(N)$.

Proof. First consider T and the second inequality. Recall that the notation $\widehat{0}$ in the first block of x columns of R means that none of these columns is deleted when forming the square $(n \times n)$ submatrices of R. The hypothesis $R \in \mathscr{R}_n(N)$ insures that R is "upper diagonal" on this first block, that is $R_{ij} = 0$ for $i > j$. Also $|R_{ii}| \geqslant 1$, hence the second inequality follows. The proof of the first inequality is similar, by symmetry. Finally, it is clear from the definitions that S and T are in $\mathscr{R}_{n-x}(N)$. $\square$

3. Proof of Theorem 1.2

Theorem 1.2 is an immediate consequence of the following more specific result to be proved in this section.

Theorem 3.1. *Let* $n, N \geqslant 1$, $Q = T_n(\delta_N)$, $R \in \mathscr{R}_n(N)$, $x = \min(n-1, N)$, *and* $y = (n + N - 1) - 2x$. *Then*

$$\begin{pmatrix} x & y & x \\ \widehat{a} & \widehat{b} & \widehat{c} \end{pmatrix}_R \geqslant \begin{pmatrix} x & y & x \\ \widehat{a} & \widehat{b} & \widehat{c} \end{pmatrix}_Q \tag{16}$$

for all admissible a, b, c (*i.e. for all integers* $a, b, c \geqslant 0$ *with* $a + b + c = N - 1$, $a, c \leqslant x$, *and* $b \leqslant y$).

To deduce Theorem 1.2 from Theorem 3.1, sum both sides of (16) over all admissible (a, b, c), thus obtaining $\det(RR^*)$ on the left and $\det(QQ^*)$ on the right (by the special case of identity (12) for a partition with $k = 3$ blocks). Consequently, the remainder of this section concerns the proof of Theorem 3.1.

Lemma 3.2 [5]. *For every n and N, the matrix* $Q = T_n(\delta_N)$ *is totally unimodular, that is: Every square submatrix of Q has determinant* 0, 1, *or* −1.

Proof. See [5, p. 853] or [11, Lemma 2.3]. $\square$

Lemma 3.3. *Let the columns* Q_j *of* $Q = T_n(\delta_N)$ *be enumerated from left to right in the usual way, so that* $Q = [Q_1, \ldots, Q_{n+N-1}]$. *Let* $S = [Q_{j_1}, \ldots, Q_{j_n}]$ *be an* $n \times n$ *submatrix of Q. If the elements of the set* $\{1, \ldots, n + N - 1\} \backslash \{j_1, \ldots, j_n\}$ *are not distinct mod N, then* $\det(S) = 0$.

Proof. Fix the integers $n, N \geqslant 1$. Let $\mathbf{e}$ denote the column n-vector whose entries are all equal to 1. Consider the residue classes $\mathscr{C}_k$ mod N on the interval $I = \{1, 2, \ldots, n + N - 1\}$, defined by $\mathscr{C}_k = (k + N\mathbf{Z}) \cap I$, $k = 1, \ldots, N$. The key observation is that for each k we have

$$\sum_{j \in \mathscr{C}_k} Q_j = \mathbf{e}, \tag{17}$$

which is easy to verify by inspection of Q. Now note that the set $T := \{1, \ldots, n + N - 1\} \backslash \{j_1, \ldots, j_n\}$ contains exactly $N - 1$ elements, say $T = \{i_1, \ldots, i_{N-1}\}$. Suppose that two of these elements are in the same residue class $\mathscr{C}_k$. Then T intersects at most $N - 2$ of the residue classes $\mathscr{C}_1, \ldots, \mathscr{C}_N$. Therefore at least 2 of the residue classes, say $\mathscr{C}_a, \mathscr{C}_b$, are completely contained in the complement of T, i.e. in $I \backslash T = \{j_1, \ldots, j_n\}$. But the vectors Q_j, $j \in \mathscr{C}_a \cup \mathscr{C}_b$ are linearly dependent since $\sum_{j \in \mathscr{C}_a} Q_j = \mathbf{e} = \sum_{j \in \mathscr{C}_b} Q_j$ by (17). Hence $\det(S) = 0$. $\square$

The lemma has a converse, but we postpone it since it is not needed in the proof of Theorem 3.1 (see Corollary 3.8).

Definition 3.4. Let $N \geqslant 1$ and $0 \leqslant b_i \leqslant a_i, i = 1, \ldots, k$ be integers. Let $I_1 < I_2 < \cdots < I_k$ be adjacent disjoint intervals of integers with cardinalities $|I_i| = a_i$, i.e. sets of the form $I_i = [c_{i-1}, c_i) \cap \mathbf{Z}$ for some integers c_i with $c_i - c_{i-1} = a_i$. Define

$$\left\{\begin{matrix} a_1 \\ b_1 \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} a_k \\ b_k \end{matrix}\right\}_N$$

to be the number of subsets $S \subset \cup_{i=1}^k I_i$ such that the elements of S are distinct mod N and $|S \cap I_i| = b_i, i = 1, \ldots, k$.

It is clear that this number is well defined (i.e. the intervals may begin at any point c_0 without affecting the result). The number is also clearly invariant under reflection, that is, we get the same result if we reverse the order of both the a_i and the b_i.

Corollary 3.5. *Let $Q = T_n(\delta_N)$, and consider the partition of Q defined by arbitrary $(a_1, \ldots, a_k)$ with $\sum a_i = n + N - 1$. Let $0 \leqslant b_i \leqslant a_i$ with $\sum b_i = N - 1$. Then*

$$\left(\begin{matrix} a_1 \\ \widehat{b_1} \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} a_k \\ \widehat{b_k} \end{matrix}\right)_Q \leqslant \left\{\begin{matrix} a_1 \\ b_1 \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} a_k \\ b_k \end{matrix}\right\}_N. \tag{18}$$

Proof. This follows immediately from Lemmas 3.3 and 3.2. □

Lemma 3.6. *Let $N \geqslant 1, x, y, a, b, c \geqslant 0$ be integers such that $a, c \leqslant x$ and $b \leqslant y$. Suppose that $x = N$ or $x + y = N$. If $1 \leqslant a + c \leqslant x$ then*

$$\begin{aligned}\left\{\begin{matrix} x \\ a \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ c \end{matrix}\right\}_N &= \frac{(a+c)!}{a!c!}\left\{\begin{matrix} x \\ a+c \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ 0 \end{matrix}\right\}_N = \frac{(a+c)!}{a!c!}\left\{\begin{matrix} x \\ 0 \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ a+c \end{matrix}\right\}_N \\ &= \frac{(a+c)!}{a!c!}\left(\left\{\begin{matrix} x \\ a+c \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ 0 \end{matrix}\right\}_N\right)^{a/(a+c)}\left(\left\{\begin{matrix} x \\ 0 \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ a+c \end{matrix}\right\}_N\right)^{c/(a+c)}.\end{aligned} \tag{19}$$

If $a + c > x$ then $\left\{\begin{smallmatrix} x \\ a \end{smallmatrix} \middle| \begin{smallmatrix} y \\ b \end{smallmatrix} \middle| \begin{smallmatrix} x \\ c \end{smallmatrix}\right\}_N = 0.$

Proof. Suppose that $1 \leqslant a + c \leqslant x$. The second equality in (19) holds by reflection symmetry, and this implies the third equality since $a/(a+c) + c/(a+c) = 1$. Thus it suffices to consider the first equality. Let $I_1 < I_2 < I_3$ be three adjacent and disjoint intervals in $\mathbf{Z}$ with cardinalities x, y, x respectively, as required in the definition.

First, suppose that $x = N$. Then we need to prove that

$$\left\{\begin{matrix} N \\ a \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} N \\ c \end{matrix}\right\}_N = \frac{(a+c)!}{a!c!}\left\{\begin{matrix} N \\ a+c \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} N \\ 0 \end{matrix}\right\}_N. \tag{$*$}$$

Since $|I_1| = N = |I_3|$, each of the intervals I_1, I_3 is a complete set of residues mod N. Let $\psi : I_1 \to I_3$ be the bijection determined by $\psi(t) \equiv t \bmod N$. Let $S \subset I_1 \cup I_2$ be a set whose elements are distinct mod N with $|S \cap I_1| = a + c$ and $|S \cap I_2| = b$. Then S corresponds to

$\frac{(a+c)!}{a!c!}$ different sets $T \subset I_1 \cup I_2 \cup I_3$ such that the elements of T are distinct mod N and $|T \cap I_1| = a$, $|T \cap I_2| = b$, $|T \cap I_3| = c$. This can be seen via the following natural correspondence: Choose a subset $K \subset S \cap I_1$ with $|K| = c$ and use ψ to map K into the interval I_3. That is, let $L = (S \cap I_1)\backslash K$ and define $T := L \cup (S \cap I_2) \cup \psi(K)$. The number of different sets T we can obtain this way from a given S is clearly $\frac{(a+c)!}{a!c!}$, the number of ways of choosing the c-subset K from the set $S \cap I_1$. Moreover, given a set T with the above properties, it is easy to see that it is obtained from exactly one such S. (Just map the set $T \cap I_3$ back into I_1, thus defining S by $S := (T \cap I_1) \cup \psi^{-1}(T \cap I_3) \cup (T \cap I_2)$. The hypothesis that the elements of T are distinct mod N implies that S has the same cardinality as T and that the elements of S are distinct mod N.) This proves (∗).

Now suppose that $x + y = N$. Then we have $I_1 + N = I_3$ and $|I_1| = x = |I_3|$, with $x \leqslant N$. Hence, we can define a map $\psi : I_1 \to I_3$ by $\psi(t) = t + N$. Clearly, ψ is a bijection and $\psi(t) \equiv t \bmod N$. The proof is now similar to the first case, using this ψ to define the correspondence.

Finally, suppose that $a + c > x$. If $x = N$ then obviously there do not exist $a + c$ integers which are all distinct mod N, so by definition the bracket $\left\{\begin{matrix} x \\ a \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ c \end{matrix}\right\}_N$ is zero. Similarly, if $x + y = N$ then the first and last intervals I_1 and I_3 are congruent mod N as remarked above, and so together contain only x distinct residues mod N. Thus the latter bracket is again zero. □

We note the analogue of (13): If say $a_1 = c + d$ and $c, d \geqslant 0$ then clearly

$$\left\{\begin{matrix} a_1 \\ b_1 \end{matrix} \middle| \begin{matrix} a_2 \\ b_2 \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} a_k \\ b_k \end{matrix}\right\}_N = \sum_{i+j=b_1} \left\{\begin{matrix} c \\ i \end{matrix} \middle| \begin{matrix} d \\ j \end{matrix} \middle| \begin{matrix} a_2 \\ b_2 \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} a_k \\ b_k \end{matrix}\right\}_N, \tag{20}$$

where the sum is restricted to $0 \leqslant i \leqslant c$ and $0 \leqslant j \leqslant d$. Similarly, we may "split" any of the other a_i.

Theorem 3.7. *Let* $n, N \geqslant 1$, $R \in \mathscr{R}_n(N)$, $x = \min(n-1, N)$, *and* $y = (n + N - 1) - 2x$. *If* $a, b, c \geqslant 0$ *are integers with* $a + b + c = N - 1$, $a, c \leqslant x$, *and* $b \leqslant y$, *then*

$$\left(\begin{matrix} x \\ \widehat{a} \end{matrix} \middle| \begin{matrix} y \\ \widehat{b} \end{matrix} \middle| \begin{matrix} x \\ \widehat{c} \end{matrix}\right)_R \geqslant \left\{\begin{matrix} x \\ a \end{matrix} \middle| \begin{matrix} y \\ b \end{matrix} \middle| \begin{matrix} x \\ c \end{matrix}\right\}_N. \tag{21}$$

Proof. Fix $N \geqslant 1$. The plan is to use induction on n in the range $n \geqslant N$, and to give a direct proof for each n in the initial range $1 \leqslant n \leqslant N$. Given $n \geqslant N$, let the nth induction hypothesis, $H(n)$, state that: If $1 \leqslant k \leqslant n$, $R \in \mathscr{R}_k(N)$, $x = \min(k-1, N)$, $y = (k + N - 1) - 2x$, and $a, b, c \geqslant 0$ are integers with $a + b + c = N - 1$, and $a, c \leqslant x$, and $b \leqslant y$, then (21) holds.

Step 1. We prove $H(N)$. Let $1 \leqslant n \leqslant N$, $R \in \mathscr{R}_n(N)$, $x = \min(n-1, N)$, $y = (n + N - 1) - 2x$, and let $a, b, c \geqslant 0$ be integers with $a + b + c = N - 1$, and $a, c \leqslant x$, and $b \leqslant y$. Thus, $x = n - 1$ and $y = N - (n - 1)$.

Case (1.1) Suppose $a = 0$ or $c = 0$. By symmetry it suffices to consider one of these cases, say $c = 0$. Remark 2.8 gives

$$\left(\begin{matrix} x \\ \widehat{a} \end{matrix} \middle| \begin{matrix} y \\ \widehat{b} \end{matrix} \middle| \begin{matrix} x \\ \widehat{0} \end{matrix}\right)_R \geqslant \left(\begin{matrix} x \\ \widehat{a} \end{matrix} \middle| \begin{matrix} y \\ \widehat{b} \end{matrix}\right)_S, \tag{22}$$

where S is the $1 \times N$ matrix $S = [R_{11}, \ldots, R_{1N}]$. Since each of these $|R_{1j}|^2 \geqslant 1$, the right-hand side of (22) is at least $\sum 1$ summed over all choices of a-subsets and b-subsets from the two corresponding blocks of S, and this by definition is

$$\begin{Bmatrix} x \\ a \end{Bmatrix}\!\!\begin{vmatrix} y \\ b \end{vmatrix}_N = \begin{Bmatrix} x & | & y & | & x \\ a & | & b & | & 0 \end{Bmatrix}_N,$$

which proves (21) in Case (1.1). (In fact, we could have been more specific. Since $a + b = N - 1$ and $x + y = N$, then there are only two possibilities: Either $a = x$, $b = y - 1$, or $a = x - 1$, $b = y$. In the first case (22) $\geqslant y$ and in the second case (22) $\geqslant x$.)

Case (1.2) Now we may suppose $a \geqslant 1$ and $c \geqslant 1$. If $a + c > x$ there is nothing to prove since the right-hand side of (21) is zero, by Lemma 3.6. So assume that $1 \leqslant a + c \leqslant x$ and apply Lemma 2.7 to the left-hand side of (21). Then applying Case (1.1) on the resulting two right-hand factors, followed by identity (19), gives (21).

Step 2. Let $n \geqslant N + 1$ and assume that $H(n - 1)$ holds. Let $R \in \mathscr{R}_n(N)$, $x = \min(n - 1, N)$, $y = (n + N - 1) - 2x$, and let $a, b, c \geqslant 0$ be integers with $a + b + c = N - 1$, and $a, c \leqslant x$, and $b \leqslant y$. Then $x = N$ and $y = (n - 1) - N \geqslant 0$. Hence $a + c < x$.

Case (2.1) Suppose $a = 0$ or $c = 0$. By symmetry it suffices to consider one of these cases, say $c = 0$. Remark 2.8 again gives

$$\begin{pmatrix} x & | & y & | & x \\ \widehat{a} & | & \widehat{b} & | & \widehat{0} \end{pmatrix}_R \geqslant \begin{pmatrix} x & | & y \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S,$$

where $S \in \mathscr{R}_{n-x}(N)$ is the submatrix of R defined by the intersection of the first $n - x$ rows with the first $x + y$ columns. But $x = N$, so if we let $k := n - x = n - N$, then $y = (n + N - 1) - 2x = (n + N - 1) - 2N = k - 1$, $S \in \mathscr{R}_k(N)$, $1 \leqslant k \leqslant n - 1$, and

$$\begin{pmatrix} x & | & y \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S = \begin{pmatrix} N & | & k-1 \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S.$$

Define $x' = \min(k - 1, N)$ and $y' = (k + N - 1) - 2x'$. There are two subcases to consider; either (**I**) $k \leqslant N$ or (**II**) $k \geqslant N + 1$.

(**I**) If $k \leqslant N$ then $x' = k - 1$ and $y' = N - x'$, so we may "split" the first block of N columns of S as $N = x' + y'$. Therefore, by (13), the induction hypothesis $H(n - 1)$, and by (20), we have:

$$\begin{aligned} \begin{pmatrix} N & | & k-1 \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S &= \begin{pmatrix} x'+y' & | & x' \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S = \sum_{i+j=a} \begin{pmatrix} x' & | & y' & | & x' \\ \widehat{i} & | & \widehat{j} & | & \widehat{b} \end{pmatrix}_S \\ &\geqslant \sum_{i+j=a} \begin{Bmatrix} x' & | & y' & | & x' \\ i & | & j & | & b \end{Bmatrix}_N = \begin{Bmatrix} x'+y' & | & x' \\ a & | & b \end{Bmatrix}_N \\ &= \begin{Bmatrix} x & | & y \\ a & | & b \end{Bmatrix}_N = \begin{Bmatrix} x & | & y & | & x \\ a & | & b & | & 0 \end{Bmatrix}_N, \end{aligned}$$

which proves (21) in the case (**I**) of Case (2.1).

(**II**) If $k \geqslant N + 1$ then $x' = N$ and $y' = (k - 1) - x'$, so we may split the last block of $(k - 1)$ columns of S as $(k - 1) = y' + x'$. Therefore, by (13), the induction hypothesis $H(n - 1)$, and by (20), we have

$$\begin{aligned} \begin{pmatrix} N & | & k-1 \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S &= \begin{pmatrix} x' & | & y'+x' \\ \widehat{a} & | & \widehat{b} \end{pmatrix}_S = \sum_{i+j=b} \begin{pmatrix} x' & | & y' & | & x' \\ \widehat{a} & | & \widehat{i} & | & \widehat{j} \end{pmatrix}_S \\ &\geqslant \sum_{i+j=b} \begin{Bmatrix} x' & | & y' & | & x' \\ a & | & i & | & j \end{Bmatrix}_N = \begin{Bmatrix} x' & | & y'+x' \\ a & | & b \end{Bmatrix}_N = \begin{Bmatrix} x & | & y \\ a & | & b \end{Bmatrix}_N, \end{aligned}$$

which proves (21) in the case (**II**) of Case (2.1).

Case (2.2) Now we may suppose $a \geqslant 1$ and $c \geqslant 1$. Since $1 \leqslant a + c \leqslant x$, we may apply Lemma 2.7 to the left-hand side of (21). Then applying Case (2.1) on the resulting two right-hand factors, followed by identity (19), gives (21). This proves $H(n)$ and hence the theorem. □

Proof of Theorem 3.1. Applying first Theorem 3.7, then Corollary 3.5, gives

$$\begin{pmatrix} x \\ \widehat{a} \end{pmatrix}\!\!\left|\begin{matrix} y \\ \widehat{b} \end{matrix}\right|\!\!\begin{pmatrix} x \\ \widehat{c} \end{pmatrix}_R \geqslant \left\{\begin{matrix} x \\ a \end{matrix}\,\middle|\,\begin{matrix} y \\ b \end{matrix}\,\middle|\,\begin{matrix} x \\ c \end{matrix}\right\}_N \geqslant \begin{pmatrix} x \\ \widehat{a} \end{pmatrix}\!\!\left|\begin{matrix} y \\ \widehat{b} \end{matrix}\right|\!\!\begin{pmatrix} x \\ \widehat{c} \end{pmatrix}_Q. \qquad \square$$

Corollary 3.8. *The converse of Lemma* 3.3 *holds. More specifically, if* $Q = T_n(\delta_N)$, $1 \leqslant j_1 < \cdots < j_n \leqslant n + N - 1$, *and the elements of the set* $\{1, \ldots, n + N - 1\}\backslash\{j_1, \ldots, j_n\}$ *are distinct* mod N, *then*

$$\det[Q_{j_1}, \ldots, Q_{j_n}] = \pm 1.$$

Consequently, equality holds in (18) *of Corollary* 3.5. *In particular*,

$$\det QQ^* = \begin{pmatrix} n + N - 1 \\ \widehat{N - 1} \end{pmatrix}_Q = \begin{Bmatrix} n + N - 1 \\ N - 1 \end{Bmatrix}_N.$$

Proof. Corollary 3.5 and the case $R = Q$ of Theorem 3.7 show that

$$\begin{pmatrix} x \\ \widehat{a} \end{pmatrix}\!\!\left|\begin{matrix} y \\ \widehat{b} \end{matrix}\right|\!\!\begin{pmatrix} x \\ \widehat{c} \end{pmatrix}_Q = \left\{\begin{matrix} x \\ a \end{matrix}\,\middle|\,\begin{matrix} y \\ b \end{matrix}\,\middle|\,\begin{matrix} x \\ c \end{matrix}\right\}_N \tag{23}$$

for $x = \min(n - 1, N)$, $y = (n + N - 1) - 2x$, and all admissible a, b, c. In the notation of Lemma 3.3, suppose that $S = [Q_{j_1}, \ldots, Q_{j_n}]$, the elements of the set $\{1, \ldots, n + N - 1\}\backslash\{j_1, \ldots, j_n\}$ are distinct mod N, but $\det S \neq \pm 1$. Then $\det S = 0$ since Q is totally unimodular (see Lemma 3.2). It follows (by Lemma 3.3 and total unimodularity of Q) that there is strict inequality (i.e. $<$) in (23) when a, b, c are the specific integers corresponding to this S. (Any $n \times n$ submatrix S is obtained by deleting some number a, b, c of columns from the blocks $Q(1), Q(2), Q(3)$ respectively of the partition $Q = [Q(1)|Q(2)|Q(3)]$ defined by the block lengths (x, y, x).) This contradiction concludes the proof. □

A direct proof of the converse of Lemma 3.3 can also be given (not relying on Theorem 3.7). That is, one can show by direct calculation that $\det[Q_{j_1}, \ldots, Q_{j_n}] = \pm 1$ whenever the complementary column numbers are distinct mod N. We leave this as an exercise.

4. Further remarks

4.1. A "gaps" version of Theorem 1.2

Let $N \geqslant 1$ and fix a set E of N integers of the form $E = \{m_1 < m_2 < \cdots < m_N\}$ where $m_1 = 1$. Put $K = m_N$. Let $\mathscr{G}_n(E)$ be the set of all complex $n \times (n + K - 1)$ matrices G such that $|G_{ij}| \geqslant 1$ whenever $j - i + 1 \in E$, and $G_{ij} = 0$ otherwise. If $Q = T_n(\delta_N)$ as previously, we can ask whether

$$\det(GG^*) \geqslant \det(QQ^*) \tag{24}$$

for all $G \in \mathscr{G}_n(E)$. Theorem 1.2 is the special case $m_j = j$ (the case of "no gaps"). The methods of this paper easily give a positive answer to this question whenever the number of rows n is such that the number of columns of G is divisible by K, and at the same time the number of columns of Q is divisible by N. (This condition on n reduces to the condition that $n-1$ be divisible by both K and N; clearly there exist infinitely many such n.) For such n, the main steps in the proof of (24) are as follows. Let $n \geqslant 1$ satisfy $n-1=(p-1)K=(q-1)N$ for some integers $p, q \geqslant 1$, so that G has $n+K-1=pK$ columns and Q has $n+N-1=qN$ columns. Consider the partition of G into p blocks of K columns each. Then by (12),

$$\det(GG^*) = \sum_{b_1+\cdots+b_p=K-1} \left(\begin{matrix} K \\ \widehat{b_1} \end{matrix} \middle| \begin{matrix} K \\ \widehat{b_2} \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} K \\ \widehat{b_p} \end{matrix}\right)_G. \tag{25}$$

For each term of this sum, repeated use of Corollary 2.6 (on the first and last blocks) shows that,

$$\left(\begin{matrix} K \\ \widehat{b_1} \end{matrix} \middle| \begin{matrix} K \\ \widehat{b_2} \end{matrix} \middle| \begin{matrix} \cdots \\ \cdots \end{matrix} \middle| \begin{matrix} K \\ \widehat{b_p} \end{matrix}\right)_G \geqslant \frac{(K-1)!}{b_1!b_2!\cdots b_p!}\rho(G),$$

where $\rho(G)=\min_i \sum_j |G_{ij}|^2 \geqslant N$. Thus, by the multinomial theorem we get

$$\det(GG^*) \geqslant p^{K-1}N.$$

When $G=Q$ (so that $K=N$, $p=q$), there are equalities in all steps of this reasoning. Indeed, the number of ways of choosing $N-1$ integers from the interval $[1, qN]$ which are distinct as residues mod N is clearly $q^{N-1}N$. Thus by Corollary 3.8, $\det(QQ^*)=q^{N-1}N$. The proof of (24) is now concluded by checking the elementary inequality $p^{K-1}N \geqslant q^{N-1}N$ given that $K(p-1)=N(q-1)$ and $K \geqslant N$. We leave this as an exercise.

This proof can be refined to give the stronger result that for each $1 \leqslant i \leqslant N$ we have

$$\begin{aligned} \left(\begin{matrix} N \\ i \end{matrix} \middle| \begin{matrix} n-1 \\ n-i \end{matrix}\right)_Q &= (q-1)^{i-1}\binom{N-1}{i-1}N \\ &\leqslant (p-1)^{i-1}\binom{K-1}{i-1}N \leqslant \left(\begin{matrix} K \\ i \end{matrix} \middle| \begin{matrix} n-1 \\ n-i \end{matrix}\right)_G, \end{aligned}$$

where notation (8) is being used in the first and last brackets, and the middle two brackets are ordinary binomial coefficients.

4.2. A counter-example

Theorem 1.2 may suggest that Conjecture 1.1 holds for all $R \in \mathscr{R}_n(N)$, not just for the Toeplitz case. The following example shows that this is false. We take $n=7$, $k=6$, $N=4$. Let x be real and define

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

Recalling that Q is the matrix R with $x=1$, it can be verified (using a computer if necessary) that $S_6(RR^*)-S_6(QQ^*)=178x^2-358x+180=2(x-1)(89x-90)$. Hence any $1<x<90/89$ gives

$$S_6(RR^*) < S_6(QQ^*),$$

thus contradicting (3).

In spite of this counter-example, some positive results for S_k have been obtained by the author: For $R \in \mathscr{R}_n(N)$, we have proved the result $S_k(RR^*) \geqslant S_k(QQ^*)$ for various cases "lower" than this counter-example such as (a) all $1 \leqslant k \leqslant n$ provided $n \leqslant N+1$, (b) $N = 3$, for all n, all $1 \leqslant k \leqslant n$, and (c) $N = 4$, for all $1 \leqslant k \leqslant n \leqslant 6$. (All cases with $N = 1$ and $N = 2$ also hold, trivially.) The counter-example was found thanks to the fact that there seemed to be a difficulty in the proof of the "next" case; $n = 7$, $k = 6$, $N = 4$. The proofs of (a)–(c) involved the same technique of grouping certain collections of minors (i.e. subdeterminants of R), this time in the Binet–Cauchy expansion for S_k, and applying Alexandrov inequalities. Some ideas involving "product versions" of these inequalities (e.g. (∗∗) and (++) in §4.4) were also used. The proofs suggest that the only role of the Toeplitz hypothesis, in this method, will be that certain "small" minors occur with a "large" multiplicity, allowing for more possible groupings. In particular, in the present case of $n = 7$, $k = 6$, $N = 4$, our work suggests that something like

$$S_6(RR^*) + (|R_{24}|^2 - |R_{13}|^2) + (|R_{67}|^2 - |R_{78}|^2) \geqslant S_6(QQ^*)$$

holds even if R is not Toeplitz. If so, then we could say that certain 1×1 minors were missing from the hoped-for inequality $S_6(RR^*) \geqslant S_6(QQ^*)$, and that they would not be missing if R is Toeplitz.

4.3. On Corollary 2.6 and Alexandrov's inequality

We will use an example to compare Corollary 2.6 to a direct application of Alexandrov's inequality. In the notation of §2, let $R = [R(1)|R(2)]$ be a complex 4×6 matrix partitioned with the block lengths $(a_1, a_2) = (3, 3)$, and let $A_i = R(i)R(i)^*$. Then Corollary 2.6 gives

$$\left(\begin{matrix}3\\ \hat{1}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{1}\end{matrix}\right)_R \geqslant 2\left(\begin{matrix}3\\ \hat{2}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{0}\end{matrix}\right)_R^{1/2}\left(\begin{matrix}3\\ \hat{0}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{2}\end{matrix}\right)_R^{1/2}.$$

By (11) of Remark 2.2 we have

$$\left(\begin{matrix}3\\ \hat{1}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{1}\end{matrix}\right)_R = \frac{4!}{2!2!}D_4(A_1, A_1, A_2, A_2),$$
$$\left(\begin{matrix}3\\ \hat{2}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{0}\end{matrix}\right)_R = \frac{4!}{1!3!}D_4(A_1, A_2, A_2, A_2),$$
$$\left(\begin{matrix}3\\ \hat{0}\end{matrix}\,\middle|\,\begin{matrix}3\\ \hat{2}\end{matrix}\right)_R = \frac{4!}{3!1!}D_4(A_1, A_1, A_1, A_2),$$

so the above inequality states that

$$D_4(A_1, A_1, A_2, A_2) \geqslant \frac{4}{3}D_4(A_1, A_1, A_1, A_2)^{1/2}D_4(A_1, A_2, A_2, A_2)^{1/2}.$$

On the other hand, the direct application of Alexandrov's inequality (14) gives

$$D_4(A_1, A_1, A_2, A_2) \geqslant D_4(A_1, A_1, A_1, A_2)^{1/2}D_4(A_1, A_2, A_2, A_2)^{1/2},$$

which is weaker. The reason is of course that Alexandrov's inequality is sharp mainly in the positive definite case, and here the A_i are of rank $\leqslant 3 < 4$, hence semidefinite. Refs. [2,18] give a more detailed look at semidefinite cases of Alexandrov's inequality.

4.4. Other methods: polarized Bazin–Reiss–Picquet identities

In this section we discuss some ideas from the initial phases of the work on Theorem 1.2. For the purposes of Theorem 1.2, these ideas were finally replaced by the use of Alexandrov inequalities for the polarized determinant (i.e. Corollary 2.6). They are still of possible interest because they give a detailed picture of the extremality of Q and they yield other results that Alexandrov's inequality does not give directly. The ideas can be described loosely as opportunistic use of various classical determinant identities. For a nice overview of such identities we refer the reader to a paper of Leclerc [13]. In particular, we use the Bazin–Reiss–Picquet identity [13, Eq. (15)]; [17, p. 193, Section 202, Eq. (1)], reproduced below for the reader's convenience:

Bazin–Reiss–Picquet Identity. *Let $n = a + b + r$ where $a, b \geqslant 1, r \geqslant 0$ are integers. Let A, B be $n \times n$ matrices whose last r columns are the same, that is $A_j = B_j$, $a + b < j \leqslant n$. Let C be an $\binom{a+b}{a} \times \binom{a+b}{a}$ matrix with rows and columns indexed by the a-subsets of the interval $[1, a + b]$ (listed in the same order for both rows and columns) and with entries $C_{I,J}$ defined as follows. For each ordered pair (I, J) of a-subsets, $C_{I,J} = \det \gamma(I, J)$ where $\gamma(I, J)$ is the $n \times n$ matrix obtained by replacing in A the columns A_j, $j \in [1, a + b]\backslash I$ by the columns B_j, $j \in [1, a + b]\backslash J$ of B (ordered such that j increases from left to right). Then*

$$\det C = (\det A)^{\alpha}(\det B)^{\beta},$$

where

$$\alpha = \binom{a+b-1}{a-1}, \quad \beta = \binom{a+b-1}{b-1}.$$

The case $a = 1$ or $b = 1$ is called Bazin's identity. The case $r > 0$ is often referred to as an "extensional" of the case $r = 0$, and can in fact be deduced from the latter by Muir's "law of extensionals" [17, p. 179, Section 187]. There are also identities for the determinant of certain square submatrices of the above C having a product of determinants on the right hand side, each of which is a certain mixture of the columns of the original A and B. These have been called "polarized" Bazin–Reiss–Picquet identities [13, Proposition 5.7]. (For an example, see Theorem B in §4.4.3.) A complete description of the possible "polarizations" is apparently still an open problem, as we have not been able to find the identity conjectured in the last paragraph of §4.4.3. However, in principle, any true identity must follow from the Plücker relations [8, p. 312, Eq. (4)], since these are known to generate all relations (between the maximal minors of a rectangular matrix).

4.4.1. Bazin–Reiss–Picquet identities

As a simple example, consider Theorem 1.2 for $N = 3, n = 3$, so that

$$Q = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}, \quad R = \begin{bmatrix} a & w & p & 0 & 0 \\ 0 & b & x & q & 0 \\ 0 & 0 & c & y & r \end{bmatrix},$$

where $|a|, \ldots, |r| \geqslant 1$. Let the 3×3 minors of R be denoted by $(ijk)_R := \det[R_i R_j R_k]$. To show directly that

$$\begin{aligned} &|(134)_R|^2 + |(135)_R|^2 + |(234)_R|^2 + |(235)_R|^2 \\ &\quad \geqslant |(134)_Q|^2 + |(135)_Q|^2 + |(234)_Q|^2 + |(235)_Q|^2 = 2, \end{aligned} \tag{$*$}$$

(which is just Theorem 3.1 for $a = 1, b = 0, c = 1$), arrange these determinants into 2×2 matrices M_R and M_Q as follows:

$$M_R := \begin{bmatrix} (134)_R & (135)_R \\ (234)_R & (235)_R \end{bmatrix}, \quad M_Q := \begin{bmatrix} (134)_Q & (135)_Q \\ (234)_Q & (235)_Q \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

By a case of Bazin's identity we have $\det M_R = (123)_R(345)_R$. (This is also a Plücker relation; see [8, p. 312, Eq. (4)]; [13, Section 2.1]; [19].) For our special matrix R this gives $|\det M_R| = |(123)_R(345)_R| = |abcpqr| \geqslant 1$, and on the other hand Hadamard's inequality (or Cauchy–Schwarz) gives

$$|\det M_R| \leqslant \left(|(134)_R|^2 + |(135)_R|^2\right)^{1/2} \left(|(234)_R|^2 + |(235)_R|^2\right)^{1/2}.$$

Thus,

$$\begin{aligned} &(|(134)_R|^2 + |(135)_R|^2)^{1/2}(|(234)_R|^2 + |(235)_R|^2)^{1/2} \\ &\quad \geqslant 1 = (|(134)_Q|^2 + |(135)_Q|^2)^{1/2}(|(234)_Q|^2 + |(235)_Q|^2)^{1/2}, \end{aligned} \qquad (**)$$

from which $(*)$ follows by the arithmetic–geometric mean inequality applied to the product on the left-hand side. The result $(**)$ is formally stronger than Theorem 3.1. We will refer to $(**)$ and other similar results as "product versions" of Theorem 3.1. The same reasoning shows that

$$(\epsilon|(134)_R|^2 + |(135)_R|^2)^{1/2} \left(|(234)_R|^2 + \frac{1}{\epsilon}|(235)_R|^2\right)^{1/2} \geqslant 1$$

for any $\epsilon > 0$. Let us note that the product version, say $(**)$, can actually be deduced from the proof of Theorem 3.1 as follows. For any 3×5 matrix R, Corollary 2.6 gives

$$\left(\begin{matrix}2\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{1}\end{matrix}\right)_R \geqslant 2 \left(\begin{matrix}2\\ \widehat{2}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{0}\end{matrix}\right)_R^{1/2} \left(\begin{matrix}2\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{2}\end{matrix}\right)_R^{1/2}. \qquad (***)$$

Applying this to the matrix $[c_1R_1, c_2R_2, R_3, R_4, R_5]$ where c_1, c_2 are adjustable parameters, one finds that the choices

$$c_1 = (|(234)_R|^2 + |(235)_R|^2)^{1/2}, \quad c_2 = (|(134)_R|^2 + |(135)_R|^2)^{1/2}$$

lead to the result $(**)$.

Another extremal property of Q which can be deduced in the above example concerns the number of bases in the set of columns of Q. We fix any field $\mathbf{F}$ and suppose that the variables $a, \ldots, r$ in the above matrix R are any nonzero elements of $\mathbf{F}$. Let

$$\left(\begin{matrix}2\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{1}\end{matrix}\right)_R^{\#}$$

denote the number of 3×3 minors of R which are nonzero in $\mathbf{F}$ and are of the type indicated (i.e. those obtained by removing 1, 0, 1 columns respectively from the three blocks of the partition of R defined by the block lengths (2, 1, 2); see §2 for related notation.)

Theorem A

$$\left(\begin{matrix}2\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{1}\end{matrix}\right)_R^{\#} \geqslant 2 = \left(\begin{matrix}2\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}1\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}2\\ \widehat{1}\end{matrix}\right)_Q^{\#}.$$

Proof. The above Bazin identity showed that $\det M_R \neq 0$ in the field $\mathbf{F}$. Hence each of the two columns of M_R has at least one nonzero entry. □

A proof of this using Alexandrov's inequality may also be possible, if for instance the following question has a positive answer:

Question: Consider the matroid defined by the above R viewed in $\mathbf{F}$. Can this matroid always be realized (i.e. parametrized) by some matrix R' over $\mathbf{C}$, with the same dimensions as R, such that every 3×3 minor of R' has modulus 1 or 0?

If R' exists, one could simply apply $(***)$ to R' to deduce Theorem A, thus in effect giving a proof via Alexandrov's inequality. The idea of using Alexandrov's inequality in counting problems already occurs in [21].

All of the above generalizes immediately to the case $n = N \geqslant 2, b = 0, a + c = N - 1$ of Theorem 3.1, i.e. to corresponding results concerning

$$\left(\begin{matrix} N-1 \\ \widehat{a} \end{matrix} \,\middle|\, \begin{matrix} 1 \\ \widehat{0} \end{matrix} \,\middle|\, \begin{matrix} N-1 \\ \widehat{c} \end{matrix}\right)_R, \quad R \in \mathscr{R}_N(N),$$

including the analogue of Theorem A. We shall confine ourselves to stating the initial step of the generalized arguments: For given a, c, we arrange the corresponding minors into a (square) matrix $M = M_R$ with rows indexed by the c-subsets I of the first block of $N - 1$ columns of R and columns indexed by the a-subsets J of the last block of $N - 1$ columns of R. For each such pair of subsets (I, J) the entry in M is $\det[I; R_N; J]$ where $[I; R_N; J]$ consists of the columns of R indicated (R_N being the middle column of R). Then the Bazin–Reiss–Picquet identity gives

$$\det M = \pm \det[R_1, \ldots, R_N]^{\alpha} \det[R_N, \ldots, R_{2N-1}]^{\gamma},$$

where

$$\alpha = \binom{N-2}{c-1}, \quad \gamma = \binom{N-2}{a-1}.$$

4.4.2. More Bazin identities

We can extend the ideas of §4.4.1 to some cases other than $n = N$. In this section we look at the case $Q = T_4(\delta_3)$, $R \in \mathscr{R}_4(3)$. More precisely, we discuss an alternative proof of

$$\left(\begin{matrix} 3 \\ \widehat{1} \end{matrix} \,\middle|\, \begin{matrix} 3 \\ \widehat{1} \end{matrix}\right)_R \geqslant \left(\begin{matrix} 3 \\ \widehat{1} \end{matrix} \,\middle|\, \begin{matrix} 3 \\ \widehat{1} \end{matrix}\right)_Q. \tag{+}$$

Thus R is 4×6 and it is partitioned as $R := [R_1, R_2, R_3 | R_4, R_5, R_6]$. We arrange the corresponding nine 4×4 minors $(ij|kl)$ to form the matrix

$$M_R := \begin{bmatrix} (12|45) & (12|46) & (12|56) \\ (13|45) & (13|46) & (13|56) \\ (23|45) & (23|46) & (23|56) \end{bmatrix} \Rightarrow M_Q = \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix}.$$

For later reference, we remark that M_Q can be interpreted as a signed version of an incidence matrix. More precisely, the transpose of M_Q (as well as M_Q itself in this example) is the matrix of the boundary operator ∂_1 from the linear span of the 1-dimensional simplexes (i.e. 2-subsets) to the linear span of the 0-dimensional simplexes (i.e. 1-subsets) on 3 points, with respect to their standard bases (with some fixed orientations and orderings). In fact, the labels (12|, (13|, (23| can be interpreted directly as the standard basis of the 2-subsets of $\{1, 2, 3\}$. Then the column labels |45), |46), |56) may be thought of in terms of their complements in |567), that is |45) =

$\widehat{6}$, $|46) = \widehat{5}$, $|56) = \widehat{4}$, and these may in turn be relabelled using their equivalents mod 3, that is $\widehat{6} = \widehat{3}, \widehat{5} = \widehat{2}, \widehat{4} = \widehat{1}$, and thus identified with the standard basis of the 1-subsets, {3}, {2}, {1} of {1, 2, 3} (evidently listed in reverse order here). It can be checked that orientations can also be assigned such that the matrix of ∂_1 in these two bases is then exactly the transpose of M_Q. Extremal properties of such operators ∂_k and associated "Laplacian" operators $(\partial_k\partial_k^*)$ have received much attention [4]. One such property is that the nonzero eigenvalues of $M_Q M_Q^*$ are all equal. In this example the eigenvalues are 3, 3, 0.

We now return to the proof of (+). Although M_R is square, its determinant gives no useful information (it is identically zero). Instead, we observe that every 2×2 minor of M_R can be computed by Bazin's identity. For example

$$\det\begin{bmatrix}(12|45) & (12|46)\\ (13|45) & (13|46)\end{bmatrix} = (123|4)\cdot(1|456).$$

Computing all nine of these gives us the 2-compound of M_R,

$$\mathscr{C}^{(2)}(M_R) = \begin{bmatrix}(123|4)(1|456) & (123|5)(1|456) & (123|6)(1|456)\\ (123|4)(2|456) & (123|5)(2|456) & (123|6)(2|456)\\ (123|4)(3|456) & (123|5)(3|456) & (123|6)(3|456)\end{bmatrix}.$$

Note that the latter is a product vw where v is a column vector and w is a row vector. It follows that M_R has rank at most 2 and thus $M_R M_R^*$ has at most two nonzero eigenvalues $\lambda_1, \lambda_2 \geqslant 0$. We now have

$$\begin{aligned}\left(\begin{matrix}3\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}3\\ \widehat{1}\end{matrix}\right)_R &= \|M_R\|_2^2 = \operatorname{tr} M_R M_R^* = \lambda_1 + \lambda_2\\ &\geqslant 2\sqrt{\lambda_1\lambda_2} = 2\sqrt{\operatorname{tr}\mathscr{C}^{(2)}(M_R M_R^*)}\\ &= 2\sqrt{\operatorname{tr}(\mathscr{C}^{(2)}(M_R)\mathscr{C}^{(2)}(M_R)^*)} = 2\sqrt{\|\mathscr{C}^{(2)}(M_R)\|_2^2}\\ &= 2\|v\|_2\|w\|_2 = 2\left(\begin{matrix}3\\ \widehat{2}\end{matrix}\,\middle|\,\begin{matrix}3\\ \widehat{0}\end{matrix}\right)_R^{1/2}\left(\begin{matrix}3\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}3\\ \widehat{2}\end{matrix}\right)_R^{1/2}.\end{aligned}$$

The argument to this point is valid for any complex 4×6 matrix R, so that we have in effect re-proved Lemma 2.7 for this example (without using Alexandrov's inequality). The proof of (+) is now completed by checking that the last term is $\geqslant 2\sqrt{3 \cdot 3} = 6$ and that the case $R = Q$ has equality in all of these steps. As in §4.4.1, with more work one can prove a stronger "product version" of (+), namely

$$\|M_R^{(1)}\|_2^2\|M_R^{(2)}\|_2^2\|M_R^{(3)}\|_2^2 \geqslant 8 = 2\cdot 2\cdot 2 = \|M_Q^{(1)}\|_2^2\|M_Q^{(2)}\|_2^2\|M_Q^{(3)}\|_2^2, \tag{++}$$

where $M_R^{(i)}$ denotes the ith row of M_R. (It also holds for columns.) Also, for matrices over fields one gets a corresponding Theorem A, namely

$$\left(\begin{matrix}3\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}3\\ \widehat{1}\end{matrix}\right)_R^{\#} \geqslant 6 = \left(\begin{matrix}3\\ \widehat{1}\end{matrix}\,\middle|\,\begin{matrix}3\\ \widehat{1}\end{matrix}\right)_Q^{\#},$$

by observing from $\mathscr{C}^{(2)}(M_R)$ above that no column of M_R can contain two zeros.

4.4.3. *Polarized Bazin–Reiss–Picquet identities*

We conclude with a slightly more complicated example exhibiting the combinatorial aspects of this approach in a more full-fledged form. We take $Q = T_5(\delta_4)$, $R \in \mathscr{R}_5(4)$ and discuss the alternative proof of

$$\left(\begin{matrix}4\\ \widehat{2}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{1}\end{matrix}\right)_R \geqslant \left(\begin{matrix}4\\ \widehat{2}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{1}\end{matrix}\right)_Q. \qquad (\dagger)$$

R is 5×8 and is partitioned as $R := [R_1, R_2, R_3, R_4 | R_5, R_6, R_7, R_8]$. Arrange the corresponding 5×5 minors $(ij|klm)$ to form the 6×4 matrix

$$M_R := \begin{bmatrix} (12|567) & (12|568) & (12|578) & (12|678) \\ (13|567) & (13|568) & (13|578) & (13|678) \\ (14|567) & (14|568) & (14|578) & (14|678) \\ (23|567) & (23|568) & (23|578) & (23|678) \\ (24|567) & (24|568) & (24|578) & (24|678) \\ (34|567) & (34|568) & (34|578) & (34|678) \end{bmatrix} \Rightarrow M_Q = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & -1 & -1 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}.$$

As before in §4.4.2, M_Q and its transpose can be viewed as standard matrices of appropriate boundary operators, and they have rank 3. (Specifically, $(M_Q)^t$ is the matrix of ∂_1 from 2-subsets to 1-subsets of a 4-point set, whereas M_Q is the matrix of ∂_2 from 3-subsets to 2-subsets of a 4-point set.)

One can compute every 3×3 minor of M_R using Plücker relations. For example, we find that

$$\det \begin{bmatrix} (12|567) & (12|568) & (12|678) \\ (14|567) & (14|568) & (14|678) \\ (34|567) & (34|568) & (34|678) \end{bmatrix} = (1|5678)(4|5678)(1234|6).$$

Identities similar to this one have been called "polarized" Bazin–Reiss–Picquet identities [13, Propositions 3.4 and 5.7]. However, some of the 3×3 minors of M_R are identically zero – for example, any minor taken from the three rows "labelled" by (12|, (13|, (23|. The possibilities can be summarized as follows:

Theorem B. *Every* 3×3 *minor of* M_R *is either* 0 *or a product of the form* $\pm(i|5678)(j|5678)(1234|k)$ *for some* $i, j \in \{1, 2, 3, 4\}$, $k \in \{5, 6, 7, 8\}$. *The latter occurs iff the* 3 *row labels of the minor form a tree. Moreover, the Prüfer code (see [15]) of the tree consists of* i, j *in some order, and* k *is the unique index common to all three column labels of the minor.*

This leads to the result that the 3-compound $\mathscr{C}^{(3)}(M_R)$ factors as vw where v, w are certain column and row vectors with

$$\|v\|_2^2 = \left(\sum_i |(i|5678)|^2\right)^2 = \left(\begin{matrix}4\\ \widehat{3}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{0}\end{matrix}\right)_R^2, \quad \|w\|_2^2 = \left(\sum_k |(1234|k)|^2\right) = \left(\begin{matrix}4\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{3}\end{matrix}\right)_R.$$

In particular, one can imitate the reasoning in §4.4.2 to re-prove the inequality (seen in Lemma 2.7)

$$\left(\begin{matrix}4\\ \widehat{2}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{1}\end{matrix}\right)_R \geqslant 3 \left(\begin{matrix}4\\ \widehat{3}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{0}\end{matrix}\right)_R^{2/3} \left(\begin{matrix}4\\ \widehat{0}\end{matrix}\,\middle|\,\begin{matrix}4\\ \widehat{3}\end{matrix}\right)_R^{1/3},$$

and use this to verify (†). Product versions and field versions analogous to §4.4.2 can also be obtained.

Thus, the case $R = Q$ has connections to the classical counting of the number of labelled trees on $\{1, 2, 3, 4\}$ (i.e. the Cayley formula $N^{N-2} = 4^2$, which follows from the Prüfer code). This number showed up as $\|v\|_2^2 = 4^2$, and was, equivalently, the number of bases (of the row space of M_Q) in the set of rows of M_Q, viewed over the field **C** or **Q**. It would be interesting to work out the above computations for a general case of the form $Q = T_{N+1}(\delta_N)$, $R = [R_1, \ldots, R_N | R_{N+1}, \ldots, R_{2N}]$ where R is $(N + 1) \times 2N$, with the rows of M_R indexed by k-subsets of $\{1, \ldots, N\} =: [1, N]$ and columns of M_R indexed by $(N + 1 - k)$-subsets of $[N + 1, 2N]$ for a fixed k. Then $(M_Q)^t$ should be the standard matrix of the boundary operator ∂_{k-1} from k-subsets to the $(k - 1)$-subsets. We could now define a "$(k - 1)$-dimensional tree on $[1, N]$" to be a set of k-subsets of $[1, N]$ such that the corresponding rows of M_Q are a basis of its row space over **C**. It is known that the rank of M_Q over **C** is $r := \binom{N-1}{k-1}$, and thus a $(k - 1)$-dimensional tree on $[1, N]$ has r elements [9].

Theorem C. *It is not always the case that the nonzero $r \times r$ minors of M_Q have the values ± 1.*

It was ± 1 in all of the above examples (in short, because they involved only 1-dimensional trees). This theorem implies that if we take these same $r \times r$ minors $\det(S)$ inside M_R then we cannot always have a "polarized Bazin–Reiss–Picquet identity" of the form

$$\det(S) = \pm 1 \cdot \text{product of certain minors of } R \text{ of size } (N + 1) \times (N + 1).$$

Theorem C was essentially noted by Bolker (see [3, Eq. (21), p. 136]) and by Kalai in [9], both of whom gave the counter-example $N = 6, k = 3$, using the 2-dimensional tree $T = \mathbf{P}_2 = \{123, 134, 145, 156, 126, 235, 346, 245, 356, 246\}$, known as "the 6 point triangulation of the projective plane $\mathbf{P}_2$". Kalai's proofs show that $\det(S) = \pm 2$ for some 10×10 minor $\det(S)$ on the 10 rows of M_Q given by the 10 labels $(123|, (134|, \ldots$ corresponding to $T = \mathbf{P}_2$. The number "2" arises as the cardinality of the homology group $H_{k-2}(\mathbf{P}_2)$. In the same paper Kalai proves a generalization of the Cayley formula: If each labelled $(k - 1)$-dimensional tree T (on $[1, N]$) is counted with a *multiplicity* equal to the square of the number of elements in the homology group $H_{k-2}(T)$ over the integers, then the "total number" is $N^{\binom{N-2}{k-1}}$. For this to be consistent with a factorization $\mathscr{C}^{(r)}(M_R) = vw$ of the kind seen in Theorem B, the entries of v and w may need to correspond to trees with "multiplicities" in general.

Problem D. Are there polarized Bazin–Reiss–Picquet identities of the form

$$\det(S) = h \cdot \text{product of certain maximal minors of } R$$

whenever $\det(S)$ is an $r \times r$ minor inside M_R, where $h = h(S)$ is an appropriate integer "multiplicity"?

We suspect that the answer is yes whenever the sets A and B of the row and column labels of S are both trees in the above sense, and that in this case

$$h = \det(S_Q) = \pm |H_{k-2}(A)| \cdot |H_{N-k-1}(B)|,$$

where S_Q denotes the matrix S when $R = Q$. (If A or B is not a tree, then the identity is conjectured to be simply $\det(S) = 0$.)

A particular case of Problem D occurs in the above case $N = 6, k = 3$, as follows: Let $R = [1, 2, 3, 4, 5, 6|7, 8, 9, 10, 11, 12]$ be any complex 7×12 matrix (where $1, 2, 3, \ldots$ denote the columns of R), and let M_R be the matrix of minors of R of the form $(\alpha|\beta)$ where α is a 3-element subset of the columns $[1, 2, 3, 4, 5, 6]$ and β is a 4-element subset of the columns $[7, 8, 9, 10, 11, 12]$. Let $A = \mathbf{P}_2 = \{123, 134, 145, 156, 126, 235, 346, 245, 356, 246\}$, and let $B =$ all 4-subsets of $\{7, 8, 9, 10, 11, 12\}$ containing the vertex 7. Let S be the 10×10 submatrix of M_R indexed by the $(\alpha, \beta) \in A \times B$. One can check that $|H_{k-2}(A)| = |H_1(\mathbf{P}_2)| = 2$ and $|H_{N-k-1}(B)| = |H_2(B)| = 1$. The proposed Bazin–Reiss–Picquet identity is then

$$\det(S) = \pm 2x_1x_2x_3x_4x_5x_6y_7^4,$$

where $x_i = (i|7, 8, 9, 10, 11, 12)$, $1 \leqslant i \leqslant 6$; $y_j = (1, 2, 3, 4, 5, 6|j)$, $j = 7$, are the minors of R with the column numbers as indicated. We have verified this equation numerically for "generic" R. As far as the author is aware, identities of this kind, with a non-unit integer factor such as 2, are not considered in either [13] or [17].

Acknowledgments

Some of this research was done while the author was on sabbatical leave at the mathematics departments of University of Victoria, Canada, University of Canterbury, New Zealand, University of Western Australia, and University of New South Wales. I wish to thank these departments for their hospitality.

References

[1] A.D. Alexandrov, Die gemischten Diskriminanten und die gemischten Volumina,Mat. Sbornik N.S., 3 (45) (1938) 227–252 (German summary).
[2] R.B. Bapat, Mixed discriminants of positive semidefinite matrices, Linear Algebra Appl. 126 (1989) 107–124.
[3] E.D. Bolker, Simplicial geometry and transportation polytopes, Trans. Amer. Math. Soc. 217 (1976) 121–142.
[4] A.M. Duval, V. Reiner, Shifted simplicial complexes are Laplacian integral, Trans. Amer. Math. Soc. 354 (11) (2002) 4313–4344 (electronic).
[5] D.R. Fulkerson, O.A. Gross, Incidence matrices and interval graphs, Pacific J. Math. 15 (1965) 835–856.
[6] L. Gårding, An inequality for hyperbolic polynomials, J. Math. Mech. 8 (6) (1959) 957–965.
[7] G.H. Hardy, J.E. Littlewood, G. Pólya, Inequalities, second ed., Cambridge University Press, 2001.
[8] W.V.D. Hodge, D. Pedoe, Methods of Algebraic Geometry, vol. I, Cambridge University Press, 1953.
[9] G. Kalai, Enumeration of Q-acyclic simplicial complexes, Israel J. Math. 45 (4) (1983) 337–351.
[10] A.G. Khovanskiĭ, Analogues of the Alexandrov–Fenchel inequalities for hyperbolic forms, (Russian)Dokl. Akad. Nauk SSSR 276 (6) (1984) 1332–1334, English translation: Soviet Math. Dokl. 29 (3) (1984) 710–713.
[11] I. Klemeš, Finite Toeplitz matrices and sharp Littlewood conjectures, Algebra i Analiz 13 (1) (2001) 39–59. (English) translation in St. Petersburg Math. J. 13 (1) (2002) 27–40.
[12] S.V. Konyagin, On the Littlewood problem, (Russian)Izv. Akad. Nauk SSSR Ser. Mat. 45 (2) (1981) 243–265 463. English translation: Math. USSR-Izv. 45 (2) (1982) 205–225.
[13] B. Leclerc, On identities satisfied by minors of a matrix, Adv. Math. 100 (1) (1993) 101–132.
[14] A.W. Marshall, I. Olkin, Inequalities: Theory of Majorization and its Applications, Academic Press, New York–London, 1979.
[15] J.W. Moon, Counting labelled trees, in: 12th Biennial Seminar of the Canad. Math. Congr. (Vancouver, 1969). Canad. Math. Monogr., No. 1, Canad. Math. Congr., Montreal, Que., 1970.
[16] O.C. McGehee, L. Pigno, B. Smith, Hardy's inequality and the L^1 norm of exponential sums, Ann. of Math. (2) 113 (3) (1981) 613–618.
[17] T. Muir, A Treatise on the Theory of Determinants, Dover Publications, Inc., New York, 1960 (Revised and enlarged by W.H. Metzler).
[18] A.A. Panov, On mixed discriminants connected with positive semidefinite quadratic forms, (Russian)Dokl. Akad. Nauk SSSR 282 (2) (1985) 273–276, English translation: Soviet Math. Dokl. 31 (3) (1985) 465–467.

[19] G.-C. Rota, B. Sturmfels, Introduction to invariant theory in superalgebras, in: D. Stanton (Ed.), Invariant theory and tableaux (Minneapolis, MN, 1988), IMA Vol. Math. Appl., vol. 19, Springer, New York, 1990.
[20] R. Schneider, Convex bodies: the Brunn–Minkowski theory, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1993.
[21] R.P. Stanley, Two combinatorial applications of the Aleksandrov–Fenchel inequalities, J. Combin. Theory Ser. A 31 (1) (1981) 56–65.