hexagon:2611.00433v1

AN EXPLICIT DETERMINANTAL FORMULA FOR SINKHORN LIMITS OF MATRICES AND TENSORS

SHANKAR SIVARAJAN

Department of Computational Medicine and Bioinformatics
University of Michigan, Ann Arbor, MI, USA
sivarajn@umich.edu

Abstract. We give an explicit formula, as a single determinant, for the polynomial satisfied by any entry of the limit obtained by scaling a positive tensor of format $n_1 \times \cdots \times n_r$ to arbitrary prescribed one-dimensional margins (Sinkhorn–Knopp scaling, Kruithof's method, iterative proportional fitting). The polynomial is the multilinear resultant, on the Segre variety $\mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_r-1}$, of the linear forms expressing the margin conditions. Its degree is the multinomial coefficient $N = \binom{d}{n_1-1,\ldots,n_r-1}$ with $d = \sum_k (n_k - 1)$, and for generic tensors its roots are the corresponding entries of all complex solutions of the scaling equations. A Sylvester matrix turns it into an explicit $N \times N$ generalized eigenvalue problem, whose null vector recovers the whole limit. The leading coefficient is the principal A-determinant of the product of simplices, that is, the product of the hyperdeterminants of all faces of the tensor. The constant term is likewise a product of face hyperdeterminants, and the faces avoiding the chosen entry split off as a factor independent of that entry. For matrices we also give a short proof, in Plücker coordinates, of a diagonal form of the eliminant. For every positive tensor, the limit is the unique root for which the Sylvester matrix has a positive null vector of monomials. Generically, the other roots are the corner entries of the solutions of the same scaling equations with complex scaling factors.

1. Introduction

Let T be a positive tensor of format $n_1 \times \cdots \times n_r$, and for each mode k let $\mu^{(k)} \in \mathbb{R}^{n_k}_{>0}$ be target margins with a common total τ. Repeatedly rescaling the slices of T so that its mode-k margins equal $\mu^{(k)}$ converges to a tensor

$$P = T \circ \left(u^{(1)} \otimes \cdots \otimes u^{(r)}\right), \qquad \sum_{c:\, c_k = i} P_c = \mu^{(k)}_i \quad (1 \le k \le r,\ 1 \le i \le n_k),$$

which is the unique tensor of this form with these margins [9, 12, 7]. For $r = 2$ this is matrix scaling, known as the Sinkhorn–Knopp algorithm, as Kruithof's double-factor method, and, in statistics, as iterative proportional fitting; it also arises in entropically regularized optimal transport, where the marginals are arbitrary. Everything below allows arbitrary positive margins; the special targets $\mu^{(1)} = \mathbf{1}$, $\mu^{(2)} = \frac{m}{n}\mathbf{1}$ give what Rowland and Wu call the Sinkhorn limit.

The entries of P are algebraic over the entries of T and the margins. For 2×2 matrices they involve square roots [10, 5]. Rowland and Wu [11] conjectured, on the basis of extensive computation, that the top-left entry of the limit of an $m \times n$ matrix is a root of an explicit polynomial of degree $\binom{m+n-2}{m-1}$; Fang [5] proved the degree bound by viewing the scaling factors as a point of a Segre variety, and later proved the full conjecture [6].

We extend this to tensors of any order with arbitrary margins. Write $y = u^{(1)} \otimes \cdots \otimes u^{(r)}$. The margin conditions are *linear* in y, and y ranges over the cone over the Segre variety $\Sigma = \mathbb{P}^{n_1-1} \times \cdots \times \mathbb{P}^{n_r-1}$, of dimension $d = \sum_k (n_k - 1)$. After normalization the conditions are d

multilinear forms $\ell_1, \dots, \ell_d$; adding the form ℓ_x that fixes the corner entry to x gives $d+1$ forms on a d-dimensional variety. For generic T they have a common zero exactly when x is the corner entry of a solution. The number of complex solutions, $\deg \Sigma$, is known: it is the maximum likelihood degree of the scaled Segre variety [1]. Our main result (Theorem 3.2) is that the multilinear resultant $\mathrm{Res}_\Sigma(\ell_1, \dots, \ell_d, \ell_x)$ is, up to a factor independent of x, the product of $x - x_i$ over all complex solutions; by a theorem of Sturmfels and Zelevinsky [13] it is a single Sylvester determinant, of size $(d+1)N$ with x in exactly $N = \binom{d}{n_1-1,\dots,n_r-1} = \deg \Sigma$ columns, so that the corner entry is an eigenvalue of an explicit $N \times N$ pencil (Corollary 3.3). Section 4 identifies the leading coefficient with the principal A-determinant (a product of hyperdeterminants of faces), shows that the constant term is also such a product, and determines the content. Section 5 treats matrices, where the pencil has a diagonal form; this form was conjectured by Rowland and Wu [11] and proved by Fang [6], and we give a short proof in Plücker coordinates. Theorem 3.4 characterizes the limit among the roots for every positive T: it is the only root for which Φ has a left null vector of monomials at a positive point. Section 6 accounts for the other roots, which come from the same scaling equations with complex factors (equivalently, from the other critical points of the dual potential); the real ones are positive scalings of sign-twisted copies of T. Appendix A gives numerical examples, including cases in which several roots lie in the range of possible values of the corner entry and a degenerate tensor of reduced degree.

2. The Segre formulation

Fix the format $\mathbf{n} = (n_1, \dots, n_r)$ with all $n_j \geq 2$, write $k_j = n_j - 1$, $d = k_1 + \cdots + k_r$, and

$$N = \binom{d}{k_1, \dots, k_r} = \deg \Sigma.$$

Index cells by $c = (c_1, \dots, c_r)$, write $\mathbf{1} = (1, \dots, 1)$ for the corner, t_c for the entries of T, and $y_c = u^{(1)}_{c_1} \cdots u^{(r)}_{c_r}$ for the Segre coordinates. Let $\sigma(y) = \sum_c t_c y_c$ and define the multilinear forms

$$\ell_{k,i}(y) = \sum_c \Big(\delta_{c_k,i} - \frac{\mu^{(k)}_i}{\tau}\Big) t_c\, y_c \quad (1 \leq k \leq r,\ 2 \leq i \leq n_k), \qquad \ell_x(y) = t_{\mathbf{1}} y_{\mathbf{1}} - \frac{x}{\tau}\,\sigma(y).$$

There are d forms $\ell_{k,i}$; we list them as $\ell_1, \dots, \ell_d$.

Lemma 2.1. *A point $y \in \Sigma$ with $\sigma(y) \neq 0$ is a common zero of $\ell_1, \dots, \ell_d$ if and only if $P = \frac{\tau}{\sigma(y)}\, T \circ y$ has margins $\mu^{(1)}, \dots, \mu^{(r)}$. In that case $\ell_x(y) = 0$ exactly when $x = P_{\mathbf{1}}$.*

Proof. The mode-k margin of P at i is $\frac{\tau}{\sigma(y)} \sum_{c_k = i} t_c y_c$, and $\ell_{k,i}(y) = \frac{\sigma(y)}{\tau}\big(\text{margin} - \mu^{(k)}_i\big)$. The forms with $i = 1$ are omitted because the margins in each mode sum to τ automatically. Finally, $\ell_x(y) = \frac{\sigma(y)}{\tau}(P_{\mathbf{1}} - x)$. □

Thus the complex solutions of the scaling problem are the points of $Z = \Sigma \cap \{\ell_1 = \cdots = \ell_d = 0\}$ with $\sigma \neq 0$, and the scaling limit is the unique positive one.

Lemma 2.2. *Fix margins with $\tau \neq 0$ and all $\mu^{(k)}_i \neq 0$ (for instance positive). For T outside a proper Zariski-closed subset, Z lies in the dense torus Σ° of Σ and consists of N simple points, at each of which $\sigma \neq 0$.*

Proof. Each $y \in \Sigma$ lies in the torus orbit O_F of a unique face $F = F_1 \times \cdots \times F_r$ ($\emptyset \neq F_j \subseteq [n_j]$), and $y_c \neq 0$ exactly for $c \in F$. For fixed $y \in O_F$ the $\ell_{k,i}(y)$ are linear forms in $(t_c)_{c \in F}$ with coefficient vectors $y_c\big(\chi_{k,i} - \frac{\mu^{(k)}_i}{\tau}\mathbf{1}\big)|_F$, where $\chi_{k,i}$ is the indicator of the slice $c_k = i$. The vectors

$\chi_{k,i}|_F$ ($i \in F_k$, all k) span a space of dimension $\dim F+1$ containing $\mathbf{1}_F$. If $F \neq [n_1]\times\cdots\times[n_r]$, choose k with $F_k \neq [n_k]$: either some $i \geq 2$ is not in F_k, and then the (k,i) vector is $-\frac{\mu_i^{(k)}}{\tau}\mathbf{1}_F \neq 0$, or $F_k = \{2,\dots,n_k\}$, and then the (k,i) vectors for $i \geq 2$ sum to $\frac{\mu_1^{(k)}}{\tau}\mathbf{1}_F \neq 0$. In both cases $\mathbf{1}_F$, and hence every $\chi_{k,i}|_F$, lies in the span of our vectors, so the forms impose $\dim F+1$ conditions on T. The incidence variety $\{(T,y) : y \in O_F,\ \ell_1(y) = \cdots = \ell_d(y) = 0\}$ therefore has dimension less than the number of cells, and for generic T we get $Z \cap O_F = \emptyset$. For F the whole box the same computation gives d independent conditions, so $I = \{(T,y) : y \in \Sigma^\circ,\ \ell_1 = \cdots = \ell_d = 0\}$ is a vector bundle over Σ°, smooth and irreducible, of dimension equal to the number of cells. Its projection to T is dominant, since for the rank-one tensor $T_c = \tau^{1-r}\prod_k \mu_{c_k}^{(k)}$ the point $y = \mathbf{1}$ lies in Z and is a simple zero (for positive μ, every positive T has a positive solution), so it is generically finite and, by generic smoothness, has reduced generic fibers. With no zeros on the boundary of Σ, Bernstein's theorem [2] counts the torus zeros, with multiplicity, as the normalized volume of $\Delta_{n_1-1}\times\cdots\times\Delta_{n_r-1}$, which is $\deg\Sigma = N$. Finally, a point of $Z \cap \Sigma^\circ$ with $\sigma = 0$ annihilates $f = \sum_c t_c \prod_j u_{c_j}^{(j)}$ and all of its logarithmic derivatives $u_i^{(k)}\partial f/\partial u_i^{(k)}$, so it is a singular point of $\{f = 0\}$ in the torus; that happens only for T in a proper closed subset (the closure of this set is the dual variety). □

3. The main theorem

For a multidegree $a = (a_1,\dots,a_r)$ let $S(a)$ be the space of polynomials of multidegree a in the variables $u^{(1)},\dots,u^{(r)}$, so $\dim S(a) = \prod_j \binom{a_j+k_j}{k_j}$. Put $K_0 = 0$, $K_j = k_1+\cdots+k_j$ ($j \geq 1$), and

$$m = (m_1,\dots,m_r), \qquad m_j = 1 + K_{j-1}.$$

Each ℓ_s has multidegree $\mathbf{1}$. The *Sylvester map* is

$$\Phi(x):\ S(m-\mathbf{1})^{\oplus(d+1)} \longrightarrow S(m), \qquad (g_1,\dots,g_d,g_x) \longmapsto \sum_{s=1}^{d} g_s\ell_s + g_x\ell_x.$$

Explicitly, in monomial bases (rows indexed by monomials u^a of multidegree m, columns by pairs (s,u^b) with u^b of multidegree $m-\mathbf{1}$), the entry in row u^a and column (s,u^b) is the coefficient of y_c in ℓ_s if $u^a = u^b\,y_c$ for a cell c, and 0 otherwise. Thus every entry is 0, $(\delta_{c_k,i} - \mu_i^{(k)}/\tau)\,t_c$, or $\delta_{c,\mathbf{1}}t_{\mathbf{1}} - (x/\tau)\,t_c$.

Lemma 3.1. *$\Phi(x)$ is square, of size $(d+1)N$, and x enters only the $N = \dim S(m-\mathbf{1})$ columns of the last summand, linearly.*

Proof. $\dim S(m-\mathbf{1}) = \prod_j \binom{K_j}{k_j} = N$. Since $\binom{1+K_j}{k_j} = \binom{K_j}{k_j}\frac{1+K_j}{1+K_{j-1}}$, the product telescopes: $\dim S(m) = N\prod_j \frac{1+K_j}{1+K_{j-1}} = (d+1)N$. □

Theorem 3.2. *For every T and μ, $\det\Phi(x) = \pm\operatorname{Res}_\Sigma(\ell_1,\dots,\ell_d,\ell_x)$, and when T and μ are positive this vanishes at $x = P_{\mathbf{1}}$. Suppose moreover that Z is contained in the dense torus of Σ (by Lemma 2.2 this holds for generic T). Then*

$$\det\Phi(x) = \pm\operatorname{Res}_\Sigma(\ell_1,\dots,\ell_d,\ell_x) = C\prod_{\xi\in Z}\ell_x(\xi)^{\operatorname{mult}_\xi} = C'\prod_{\xi\in Z,\ \sigma(\xi)\neq 0}\big(x - x_\xi\big)^{\operatorname{mult}_\xi},$$

where mult_ξ is the intersection multiplicity, $\sum_\xi \operatorname{mult}_\xi = N$, $x_\xi = \tau\,t_{\mathbf{1}}\xi_{\mathbf{1}}/\sigma(\xi)$ is the corner entry of the solution ξ, and $C \neq 0$ and C' do not depend on x. For generic T the polynomial $\det\Phi(x)$ has degree exactly N, with the corner entries of the N complex solutions as its roots.

Proof. Res_Σ is the resultant of type $(k_1,\dots,k_r;1,\dots,1)$: an irreducible polynomial in the coefficients of $d+1$ multilinear forms that vanishes exactly when they have a common zero on Σ [8, Ch. 13, §2A]. Since $\min(k_j,1)=1$ for all j, by Sturmfels and Zelevinsky [13] (see [8, Ch. 13, Prop. 2.2 and (2.3)]) the resultant equals the determinant of the Sylvester map $S(m^\pi-\mathbf{1})\otimes\mathbb{C}^{d+1}\to S(m^\pi)$ for each of the $r!$ vectors $m^\pi_j=1+\sum_{\pi(i)<\pi(j)}k_i$, one for each permutation π of the factors. For the identity permutation this is our m, and $\det\Phi=\pm\,\mathrm{Res}_\Sigma$ identically in the coefficients of the forms; being a polynomial identity, it holds for our specialized coefficients. At $x=P_{\mathbf{1}}$ the forms have the common zero y^* given by the limit, so the resultant vanishes.

When Z lies in the torus (it is then finite, being closed in Σ), the Poisson formula for sparse resultants [4, Thm. 1.1] gives $\mathrm{Res}_\Sigma(\ell_1,\dots,\ell_d,f)=C\prod_{\xi\in Z}f(\xi)^{\mathrm{mult}_\xi}$ for every f of multidegree $\mathbf{1}$. Here C is a product of powers of directional resultants of $\ell_1,\dots,\ell_d$; these are nonzero because Z meets no boundary orbit, and C does not depend on f. By Bernstein's theorem [2] the total multiplicity is the normalized volume of $\Delta_{n_1-1}\times\cdots\times\Delta_{n_r-1}$, that is, $\deg\Sigma=N$. For $f=\ell_x$ each factor is $\ell_x(\xi)=t_{\mathbf{1}}\xi_{\mathbf{1}}-\frac{x}{\tau}\sigma(\xi)$, which is $-\frac{\sigma(\xi)}{\tau}(x-x_\xi)$ if $\sigma(\xi)\neq0$ and independent of x otherwise. By Lemma 2.1 the x_ξ are the corner entries of the complex solutions. The last sentence follows from Lemma 2.2. □

Corollary 3.3 (Eigenvalue form)**.** *Write $\Phi(x)=[\,M_0\mid K-xJ\,]$, where M_0 collects the first dN columns and $K-xJ$ the last N. Whenever $\det\Phi\not\equiv0$ (in particular for generic T), M_0 has full column rank; let the columns of Z_0, with entries in $\mathbb{Q}(T,\mu)$, span its left kernel. Then*

$$\det\Phi(x)=c\,\det\big(Z_0^\top K-x\,Z_0^\top J\big),$$

so $P_{\mathbf{1}}$ is a generalized eigenvalue of an explicit $N\times N$ pencil. More generally, let $\xi\in Z$ with $\sigma(\xi)\neq0$, and suppose x_ξ is a simple root of $\det\Phi$. Then the left null vector of $\Phi(x_\xi)$ is, up to scale, the evaluation $h\mapsto h(\xi)$ on $S(m)$. So it determines every ratio $u_i^{(k)}/u_{i'}^{(k)}$ of ξ, and hence the whole solution $P_\xi=\frac{\tau}{\sigma(\xi)}T\circ\xi$, with entries in $\mathbb{Q}(T,\mu)(x_\xi)$. For $\xi=y^*$ this recovers the limit from $P_{\mathbf{1}}$.

Proof. Multiply Φ on the left by an invertible matrix whose last N rows are $Z_0^\top$; the result is block triangular. The functional $h\mapsto h(\xi)$ is nonzero, and it kills every $\sum g_s\ell_s+g_x\ell_{x_\xi}$ because $\ell_1,\dots,\ell_d$ and ℓ_{x_ξ} all vanish at ξ. When the root is simple the left kernel is one-dimensional. Since $m_k\geq1$ for every k, the monomials of $S(m)$ include pairs differing by one variable of each mode, and the corresponding ratios of coordinates of the null vector are the ratios $u_i^{(k)}/u_{i'}^{(k)}$. □

Theorem 3.4 (The root of the limit)**.** *Let T and the margins be positive, let P^* be the limit, and let $y^*\in\Sigma$ be its Segre point. For $y\in\Sigma$ write $\mathrm{ev}_y=(y^a)_a$ for the vector of monomials of multidegree m evaluated at y, a row vector indexed like the rows of Φ.*

1. *P^* is the only solution with all entries positive, and y^* is a simple point of Z.*
2. *For $x\in\mathbb{C}$, the left kernel of $\Phi(x)$ contains ev_y for some $y\in\Sigma$ with all coordinates positive if and only if $x=P^*_{\mathbf{1}}$, and then $y=y^*$. In particular $\det\Phi(P^*_{\mathbf{1}})=0$.*
3. *If $P^*_{\mathbf{1}}$ is a simple root of $\det\Phi$, its left null vector is, up to scale, ev_{y^*}, which has all entries positive. If moreover Z lies in the torus and all roots are simple, then $P^*_{\mathbf{1}}$ is the only root of $\det\Phi$ whose left null vector can be scaled to have all entries positive.*

Proof. (1) Existence and uniqueness of the positive scaling are classical [12, 7]. To show that y^* is a simple zero, scale y^* so that $T\circ y^*=P^*$, and use the coordinates $a=\log u$, writing tangent vectors as $\alpha=(\alpha^{(1)},\dots,\alpha^{(r)})$. The directions with $\alpha^{(k)}=\lambda_k\mathbf{1}$ for every k are tangent to the fibers of the cone over Σ, so $T_{y^*}\Sigma$ is the quotient by them. The derivative of the map $a\mapsto$ (all margins of $T\circ e^a$) at y^* is the symmetric matrix A of the form $q(\alpha)=\sum_c P^*_c\big(\sum_k\alpha^{(k)}_{c_k}\big)^2$.

Since $P^* > 0$, A is positive semidefinite, and $A\alpha = 0$ exactly when $\sum_k \alpha^{(k)}_{c_k} = 0$ for all c, that is, when $\alpha^{(k)} = \lambda_k \mathbf{1}$ with $\sum_k \lambda_k = 0$.

Now let α satisfy $d\ell_{k,i}(\alpha) = 0$ for $i \geq 2$. Summing over i gives $\sum_i \ell_{k,i} = 0$ identically when $\ell_{k,1}$ is defined by the same formula, so also $d\ell_{k,1}(\alpha) = 0$. As $\ell_{k,i} = M_{k,i} - \frac{\mu^{(k)}_i}{\tau}\sigma$ with $M_{k,i} = \sum_{c_k=i} t_c y_c$, this reads $(A\alpha)^{(k)}_i = \frac{\mu^{(k)}_i}{\tau}\, d\sigma(\alpha)$ for all k, i. Let e be the vector with $e^{(1)} = \mathbf{1}$ and $e^{(k)} = 0$ for $k \geq 2$, so that $Ae = (\mu^{(1)}, \dots, \mu^{(r)})$. Then $A\big(\alpha - \frac{d\sigma(\alpha)}{\tau} e\big) = 0$, so every $\alpha^{(k)}$ is a multiple of $\mathbf{1}$, and α is zero in $T_{y^*}\Sigma$. Thus $\ell_1, \dots, \ell_d$ have independent differentials at y^*, and y^* is a simple zero.

(2) The entry of $\mathrm{ev}_y\, \Phi(x)$ in the column (s, u^b) is $y^b \ell_s(y)$, with $\ell_{d+1} = \ell_x$. If y is positive then $y^b > 0$, so ev_y lies in the left kernel exactly when $\ell_1(y) = \cdots = \ell_d(y) = \ell_x(y) = 0$. Since $\sigma(y) > 0$, Lemma 2.1 shows that this means $\frac{\tau}{\sigma(y)} T \circ y$ is a positive solution with corner entry x, hence equals P^* by (1).

(3) A simple root has a one-dimensional left kernel, which by (2) is spanned by ev_{y^*}. For another simple root x_ξ, the left null vector is ev_ξ by Corollary 3.3. If it can be scaled to be positive, then all ratios of monomials at ξ are positive. These include all ratios $u^{(k)}_i / u^{(k)}_{i'}$, so ξ is positive and $\xi = y^*$. □

Thus the limit is distinguished from the other roots by positivity of the whole solution, read off from the null vector of Φ. When Z lies in the torus, (1) shows that the factor $x - P^*_{\mathbf{1}}$ occurs in $\det \Phi$ with multiplicity one unless another complex solution has the same corner entry.

Since Theorem 3.2 applies to any entry after permuting the indices, every entry of the limit is a root of a determinant of this kind; when $P^*_{\mathbf{1}}$ is a simple root of $\det \Phi$, a single corner determinant and its left null vector already recover all entries of P^* via Corollary 3.3.

Remark 3.5. Theorem 3.2 needs no genericity beyond the assumption on Z; there are $r!$ admissible multidegrees, one for each ordering of the factors, and all give the same determinant up to sign.

4. Leading and constant coefficients

Write $R(x) = \det \Phi(x)$. We treat the entries t_c as indeterminates and the margins as coordinates on the linear space of margin vectors with a common total τ, so the coefficients of R lie in $\mathcal{R} = \mathbb{Q}[T, \mu][\tau^{-1}]$; "up to a unit" means up to a factor $q\tau^e$ with $q \in \mathbb{Q}^\times$, $e \in \mathbb{Z}$. Let $A = \{e_c\} \subset \mathbb{Z}^{n_1} \oplus \cdots \oplus \mathbb{Z}^{n_r}$ be the vertex set of $\Delta_{n_1-1} \times \cdots \times \Delta_{n_r-1}$. For a face $F = F_1 \times \cdots \times F_r$ ($\emptyset \neq F_j \subseteq [n_j]$) let T_F be the corresponding subtensor and $\mathrm{Disc}_F(T)$ its A-discriminant $\Delta_{A\cap F}$: t_c for a cell, the $k \times k$ minor for a square face of a matrix, Cayley's hyperdeterminant for a $2 \times 2 \times 2$ face, and in general the hyperdeterminant of T_F when it is nontrivial (and 1 otherwise) [8]. The principal A-determinant is $E_A(T) = \prod_F \mathrm{Disc}_F(T)$, with all multiplicities equal to one because the toric variety $X_A = \Sigma$ is smooth [8, Ch. 10, Thm. 1.2].

Theorem 4.1. *Up to a unit, the leading coefficient of $R(x)$ is $E_A(T)$; in particular it does not depend on the margins except through a power of τ.*

Proof. The coefficient of x^N is $(-1/\tau)^N \mathrm{Res}_\Sigma(\ell_1, \dots, \ell_d, \sigma)$. A resultant is unchanged by adding to one form a multiple of another of the same multidegree, so we may replace $\ell_{k,i}$ by $M_{k,i}(y) = \sum_{c_k=i} t_c y_c$, which removes μ. Write $f(u) = \sum_c t_c \prod_j u^{(j)}_{c_j}$, the polynomial with support A and coefficients T. Then $\sigma = f$ and $M_{k,i} = u^{(k)}_i\, \partial f / \partial u^{(k)}_i$, and together these span the space of all logarithmic derivatives of f and f itself (the omitted $i = 1$ derivatives are determined

by the Euler relations $\sum_i u_i^{(k)} \partial f/\partial u_i^{(k)} = f$). The A-resultant is the resultant on the toric variety $X_A = \Sigma$, so $\mathrm{Res}_\Sigma(M, \sigma)$ is, up to a nonzero constant, the principal A-determinant $E_A(f) = R_A\big(u\,\partial f/\partial u, f\big)$ as defined in [8, Ch. 10, (1.1)]. By [8, Ch. 10, Thm. 1.2], since $X_A = \Sigma$ is smooth, it is the product of the discriminants of all faces, each to the first power. □

Corollary 4.2. *The degree of R is less than N exactly on the hypersurface $E_A(T) = 0$, that is, when some face of T has vanishing hyperdeterminant (or minor, or entry).*

The faces through the corner and the others play different roles.

Proposition 4.3. *If F does not contain the corner, then $\mathrm{Disc}_F(T)$ divides every coefficient of $R(x)$.*

Proof. For generic T with $\mathrm{Disc}_F(T) = 0$, the restriction of f to the face F has a singular point in its torus, since by definition $\{\mathrm{Disc}_F = 0\}$ is the closure of the set of such coefficient vectors [8, Ch. 9, (1.1) and Def. 1.2]. This gives a point $\xi \in \Sigma$ in the torus orbit of F at which f and all its logarithmic derivatives vanish, that is, all margins of $T \circ \xi$ are zero. So $\sigma(\xi) = 0$ and every $\ell_{k,i}(\xi) = 0$. Since $\mathbf{1} \notin F$, also $\xi_{\mathbf{1}} = 0$, hence $\ell_x(\xi) = 0$ for every x, and $R \equiv 0$. As this holds on a dense subset of the hypersurface $\mathrm{Disc}_F = 0$, the irreducible polynomial Disc_F divides R as a polynomial in T and x, and being free of x it divides each coefficient. □

Proposition 4.4. *For $1 \le k \le r$ let $T^{(k)}$ be the slab $\{t_c : c_k \ge 2\}$, of format $n_1 \times \cdots \times (n_k - 1) \times \cdots \times n_r$, let A_k be its vertex configuration, and let $N_k = \binom{d-1}{k_1,\ldots,k_k-1,\ldots,k_r}$ be the degree of $\Sigma_k = \Sigma \cap \{u_1^{(k)} = 0\}$. Up to a unit,*

$$R(0) = t_{\mathbf{1}}^N \prod_{k=1}^{r} \big(\mu_1^{(k)}\big)^{N_k} E_{A_k}\big(T^{(k)}\big).$$

Proof. At $x = 0$, $\ell_x = t_{\mathbf{1}}\, u_1^{(1)} \cdots u_1^{(r)}$. The resultant is multiplicative in its last argument, and $\mathrm{Res}_\Sigma(\ell_1, \ldots, \ell_d, u_1^{(k)}) = \pm\,\mathrm{Res}_{\Sigma_k}(\ell_1|_{\Sigma_k}, \ldots, \ell_d|_{\Sigma_k})$, so $R(0) = \pm t_{\mathbf{1}}^N \prod_k R_k$ with R_k the latter resultant. On Σ_k only cells of $T^{(k)}$ survive. Write σ_k and $M_{j,i}^{(k)} = \sum_{c \in T^{(k)},\, c_j = i} t_c y_c$ for the restricted total and margin forms. For mode k the restricted forms $\ell_{k,i}|_{\Sigma_k} = M_{k,i}^{(k)} - \frac{\mu_i^{(k)}}{\tau}\sigma_k$ $(2 \le i \le n_k)$ sum to $\frac{\mu_1^{(k)}}{\tau}\sigma_k$, and the matrix taking $(M_{k,i}^{(k)})_{i \ge 2}$ to them has determinant $\mu_1^{(k)}/\tau$. So, after a triangular change of basis of determinant $\mu_1^{(k)}/\tau$, the d restricted forms become σ_k, $M_{k,i}^{(k)}$ $(i \ge 3)$ and $M_{j,i}^{(k)}$ $(j \ne k,\ i \ge 2)$; the other modes are handled by subtracting multiples of σ_k. Since Res_{Σ_k} has degree N_k in each argument, $R_k = \pm(\mu_1^{(k)}/\tau)^{N_k}\,\mathrm{Res}_{\Sigma_k}(M^{(k)}, \sigma_k)$, and the last resultant is, up to a nonzero constant, $E_{A_k}(T^{(k)})$, exactly as in the proof of Theorem 4.1. □

Thus both extreme coefficients of R are products of face discriminants. In particular $R(0) \ne 0$ for generic T and any margins with $\mu_1^{(k)} \ne 0$.

Theorem 4.5. *Fix margins with $\tau \ne 0$ and all $\mu_i^{(k)} \ne 0$. As a polynomial in x, $R(x)$ has content, up to a unit,*

$$c(T) = t_{\mathbf{1}} \prod_{F \not\ni \mathbf{1}} \mathrm{Disc}_F(T),$$

the product over all faces not containing the corner (including the vertices t_c, $c \ne \mathbf{1}$). Consequently $R = c(T)\,\tilde{E}_T(x)$, where $\tilde{E}_T$ is primitive of degree N; up to a unit, its leading coefficient

and constant term are

$$\prod_{F\ni\mathbf{1},\ F\neq\{\mathbf{1}\}} \mathrm{Disc}_F(T) \qquad \text{and} \qquad t_{\mathbf{1}}^{N-1}\prod_k \big(\mu_1^{(k)}\big)^{N_k}\prod_{F\not\ni\mathbf{1}} \mathrm{Disc}_F(T)^{\,e_F-1},$$

where $e_F = \#\{k : 1 \notin F_k\}$. *Moreover* $\tilde{E}_T$ *is a power of an irreducible polynomial in* $\mathbb{C}[T, x]$.

Proof. The content divides the leading coefficient $E_A(T)$ (Theorem 4.1). This is squarefree and free of μ, and it is a product of irreducible face discriminants that are pairwise non-associate, since each Disc_F involves exactly the variables t_c, $c \in F$ (see below). So $c = \prod_{F\in S} \mathrm{Disc}_F$ for some set S of faces. Proposition 4.3 gives $F \in S$ for every $F \not\ni \mathbf{1}$. At $t_{\mathbf{1}} = 0$ we have $\ell_x = -\frac{x}{\tau}\sigma$, so $R = (-x/\tau)^N E_A(T)|_{t_{\mathbf{1}}=0} = 0$, since $t_{\mathbf{1}}$ is a factor of E_A; hence $\{\mathbf{1}\} \in S$. Now let $F \ni \mathbf{1}$ be a face of positive dimension with nontrivial discriminant. Disc_F is irreducible [8, Ch. 9, Def. 1.2] and invariant up to sign under permutations of the indices within each mode of F. These permutations act transitively on the cells of F, so the nonconstant polynomial Disc_F involves every t_c with $c \in F$, in particular $t_{\mathbf{1}}$. By Proposition 4.4, $R(0)$ is a nonzero polynomial whose only dependence on $t_{\mathbf{1}}$ is the factor $t_{\mathbf{1}}^N$, since the slabs do not contain the corner. An irreducible polynomial involving $t_{\mathbf{1}}$ and not associate to $t_{\mathbf{1}}$ cannot divide it. Hence $F \notin S$.

The formula for $\tilde{E}_T(0) = R(0)/c(T)$ follows from Proposition 4.4. Every face $F \not\ni \mathbf{1}$ lies in exactly the $e_F \geq 1$ slabs $T^{(k)}$ with $1 \notin F_k$, and $E_{A_k}(T^{(k)})$ is the product of the discriminants of the faces of $T^{(k)}$.

For the last claim, let I be the incidence variety of Lemma 2.2. It is irreducible, and its projection to T has degree N with reduced generic fibers. By Theorem 3.2, $\tilde{E}_T$ is, up to its leading coefficient, the norm of $X - x$ from the function field of I to $\mathbb{C}(T)$, which is a power of the minimal polynomial of x. □

Remark 4.6 (Degenerate tensors)**.** For special positive T the full determinant R can vanish identically. For instance, by Proposition 4.3 this happens for a 3×3 matrix with a vanishing 2×2 minor away from the $(1, 1)$ entry. The right object is therefore $\tilde{E}_T$. Since the limit depends continuously on T, the identity $\tilde{E}_T(P_{\mathbf{1}}) = 0$, valid for generic T, holds for *every* positive T. The coefficients of $\tilde{E}_T$ have no common factor, so $\tilde{E}_T \not\equiv 0$ outside a subset of codimension at least two.

5. Matrices

For $r = 2$, Theorem 3.2 gives an $N \times N$ pencil with $N = \binom{m+n-2}{m-1}$ whose determinant is the eliminant of the top-left entry. The faces with nontrivial discriminant are then exactly the square submatrices, so the coefficients in Section 4 are products of minors: the leading coefficient of $\tilde{E}_T$ is the product of the minors of size at least two that contain t_{11}, and the minors avoiding both the first row and the first column enter its constant term. Rowland and Wu [11] found, and Fang [6] proved, a diagonal form of this eliminant, in which the matrix enters only through minors on a diagonal and the margins only through an integer-coefficient matrix. Two regular pencils with the same simple eigenvalues are equivalent, so for generic T the two forms differ by a change of basis; we give a short proof of this formula.

Plücker coordinates and derivations. Let $e_1, \ldots, e_M$ be the standard basis of $\mathbb{C}^M$, totally ordered. For a k-subset $J = \{j_1 < \cdots < j_k\}$ write $e_J = e_{j_1} \wedge \cdots \wedge e_{j_k}$. The Plücker vector of the row space V of a full-rank $k \times M$ matrix W is $\omega = \sum_J \omega_J e_J$, with ω_J the maximal minor of W on the columns J. A matrix L acts on $\wedge^k\mathbb{C}^M$ as the derivation $D_L(v_1 \wedge \cdots \wedge v_k) = \sum_i v_1 \wedge \cdots \wedge Lv_i \wedge \cdots \wedge v_k$.

Lemma 5.1. *In the basis (e_J), $(D_L)_{J,J} = \sum_{a\in J} L_{aa}$; if $J' = J\setminus\{b\}\cup\{a\}$ with $b\in J$, $a\notin J$, then $(D_L)_{J,J'} = \operatorname{sgn}(J,J')L_{ba}$ with $\operatorname{sgn}(J,J') = (-1)^{\#\{z\in J\cap J' : \min(a,b)<z<\max(a,b)\}}$; all other entries vanish.*

Proof. The term of $D_L e_{J'}$ contributing to e_J replaces e_a by e_b; sorting moves e_b past exactly the elements of $J\cap J'$ strictly between a and b. □

Lemma 5.2. *If $g\in V$ then $D_{gq^\top}\omega = (q\cdot g)\,\omega$; if $q\perp V$ then $D_{gq^\top}\omega = 0$.*

Proof. With $\omega = v_1\wedge\cdots\wedge v_k$, $D_{gq^\top}\omega = \sum_i (q\cdot v_i)\, v_1\wedge\cdots\wedge g\wedge\cdots\wedge v_k$. If $q\perp V$ every term vanishes; if $g\in V$ take $v_1 = g$. □

Let $T\in\mathbb{R}^{m\times n}_{>0}$ with row targets $\rho = \mu^{(1)}$ and column targets $\kappa = \mu^{(2)}$. Index $\mathbb{C}^{m+n}$ by $1<\cdots<m<1'<\cdots<n'$, let $\omega_J(B)$ be the maximal minor of $[I_m \mid B]$ on the columns J, and put

$$\alpha = (\mathbf{1}_m,\kappa),\qquad \beta = (-\rho,\mathbf{1}_n),\qquad L = \alpha\beta^\top + \rho_1\alpha e_1^\top - \kappa_1 e_{1'}\beta^\top.$$

The *link* $\mathcal{L}$ is the set of m-subsets J with $1'\in J$, $1\notin J$; it corresponds to pairs $I = \{2,\dots,m\}\setminus J$ and $K = \{k\in\{2,\dots,n\} : k'\in J\}$ with $|I| = |K|$, so $|\mathcal{L}| = N$. For $J\in\mathcal{L}$ let $J^* = J\setminus\{1'\}\cup\{1\}$, $Q = (D_L-\rho_1)|_{\mathcal{L}\times\mathcal{L}}$, and $s_J = (D_L)_{J,J^*}$. By Lemma 5.1, $Q_{J,J} = \sum_{k\in K}\kappa_k - \sum_{i\in[m]\cap J}\rho_i - \rho_1$, $Q_{J,J'} = \operatorname{sgn}(J,J')\alpha_b\beta_a$ for exchanges avoiding $1,1'$, and $s_J = (-1)^{m-1-|I|}\rho_1\kappa_1$.

Theorem 5.3 (Fang [6])**.** *With $\hat\Delta_J = \omega_J(T)$ and $\hat\Gamma_J = s_J t_{11}\,\omega_{J^*}(T)$, the top-left entry $x = p_{11}$ of the limit is a root of*

$$F(x) = \det\big(\operatorname{diag}(\hat\Gamma) + x\,\operatorname{diag}(\hat\Delta)\,Q\big).$$

Proof. Let V be the row space of $[I\mid P]$ with Plücker vector ω. The margins say $\alpha\in V$ and $\beta\perp V$, so Lemma 5.2 gives $D_L\omega = (\beta\cdot\alpha+\rho_1)\omega = \rho_1\omega$. For $J\in\mathcal{L}$ the coordinates $J'\notin\mathcal{L}$ that can occur in $(D_L\omega)_J$ come from $b = 1'$ or $a = 1$, where $L_{1',a} = 0$ $(a\neq 1)$ and $L_{b,1} = 0$ $(b\neq 1')$; hence $\sum_{J'\in\mathcal{L}} Q_{J,J'}\omega_{J'}(P) + s_J\omega_{J^*}(P) = 0$. Since $[I\mid P] = \operatorname{diag}(\lambda)[I\mid T]\operatorname{diag}(\lambda^{-1},\nu)$, each $\omega_J(P)$ is $\omega_J(T)$ times a torus character, and $\omega_{J^*}(P)\,\omega_J(T) = \frac{t_{11}}{x}\,\omega_{J^*}(T)\,\omega_J(P)$. Multiplying the row by $x\,\omega_J(T)$ shows that $w = (\omega_J(P))_{J\in\mathcal{L}}\neq 0$ lies in the kernel. □

Proposition 5.4. *Let $E = \{2,\dots,m\}\cup\{2',\dots,n'\}$ and $\Theta = \operatorname{diag}((-1)^{\#(J\cap[m])})$. Under $J\mapsto J\setminus\{1'\}$, $\Theta Q\Theta = \varepsilon(\alpha_E)\iota(\beta_E) - \rho_1$ on $\wedge^{m-1}\mathbb{C}^E$, where $\varepsilon(v) = v\wedge\cdot$ and $\iota(\beta)$ is contraction with β, and $(\varepsilon\iota)^2 = (\beta_E\cdot\alpha_E)\varepsilon\iota$ with $\beta_E\cdot\alpha_E = \rho_1-\kappa_1$. Hence Q has eigenvalues $-\kappa_1$ and $-\rho_1$ with multiplicities $\binom{m+n-3}{m-2}$ and $\binom{m+n-3}{m-1}$ (understood as a single eigenvalue of multiplicity N when $\rho_1=\kappa_1$).*

Proof. Away from $1,1'$, $L = \alpha\beta^\top$, and the sign difference caused by $1'$ is corrected by Θ; $D_{\alpha\beta^\top} = \varepsilon(\alpha)\iota(\beta)$. From $\iota\varepsilon+\varepsilon\iota = \beta\cdot\alpha$ and $\varepsilon^2 = 0$ we get $(\varepsilon\iota)^2 = (\beta\cdot\alpha)\varepsilon\iota$, and the trace of $\varepsilon\iota$ on $\wedge^{m-1}\mathbb{C}^{m+n-2}$ is $(\beta\cdot\alpha)\binom{m+n-3}{m-2}$. □

In particular Q is invertible, so the leading coefficient of F is $\det Q\prod_{J\in\mathcal{L}}\omega_J(T)$, with $\det Q = \pm\kappa_1^{\binom{m+n-3}{m-2}}\rho_1^{\binom{m+n-3}{m-1}}$. Up to sign, $\omega_J(T)$ is the minor of T on the rows $[m]\setminus J$ and the columns $J\cap\{1',\dots,n'\}$. As J runs over $\mathcal{L}$, these are exactly the N square submatrices containing the entry t_{11}, including t_{11} itself. Since F and $\tilde E_T$ have the same degree N and the same roots for generic T, comparing leading coefficients shows that $F(x)$ equals $t_{11}(\det Q)\,\tilde E_T(x)$ up to a unit, leaving the product of the minors of size at least two through the corner as the leading coefficient of $\tilde E_T$, in agreement with Theorem 4.5.

6. The other roots

Kruithof's method and Sinkhorn's algorithm look for positive factors $u^{(1)}, \dots, u^{(r)}$ such that $P(u) = T \circ (u^{(1)} \otimes \cdots \otimes u^{(r)})$ has the prescribed margins. The other roots of $\det \Phi$ arise in the same way once the factors are allowed to be arbitrary nonzero complex numbers. Write $G = \{(\lambda_1 \mathbf{1}, \dots, \lambda_r \mathbf{1}) : \prod_k \lambda_k = 1\}$ for the torus of rescalings that do not change $P(u)$, and consider the dual potential

$$f(u) = \sum_{k=1}^{r} \sum_{i=1}^{n_k} \mu_i^{(k)} \log u_i^{(k)} - \sum_c P(u)_c$$

on $(\mathbb{C}^\times)^{n_1+\cdots+n_r}$. The logarithms are multivalued, but df is not: $u_i^{(k)}\, \partial f / \partial u_i^{(k)} = \mu_i^{(k)} - \sum_{c_k = i} P(u)_c$.

Proposition 6.1.

(1) *For every T, if $u^{(k)} \in (\mathbb{C}^\times)^{n_k}$ are such that $P(u)$ has the prescribed margins, then $P(u)_{\mathbf{1}}$ is a root of $\det \Phi$, and the left kernel of $\Phi(P(u)_{\mathbf{1}})$ contains the evaluation at $u^{(1)} \otimes \cdots \otimes u^{(r)}$.*

(2) *For generic T there are exactly N such complex solutions modulo G, all simple, and their corner entries are the roots of $\det \Phi$, counted with multiplicity. When $\tilde{E}_T$ is irreducible these roots are the Galois conjugates of $P^*_{\mathbf{1}}$ over $\mathbb{Q}(T, \mu)$.*

(3) *The complex solutions are exactly the critical points of f, modulo G.*

(4) *If T and μ are positive and a solution P is real, there are signs $s^{(k)} \in \{\pm 1\}^{n_k}$ with $\operatorname{sign}(P_c) = \prod_k s_{c_k}^{(k)}$ for all c. Then P is a* positive scaling, *with the prescribed margins, of the sign-twisted tensor $T^s = T \circ (s^{(1)} \otimes \cdots \otimes s^{(r)})$, and a critical point of*

$$f_s(a) = \sum_{k,i} \mu_i^{(k)} a_i^{(k)} - \sum_c T^s_c\, e^{a^{(1)}_{c_1} + \cdots + a^{(r)}_{c_r}}, \qquad a \in \mathbb{R}^{n_1+\cdots+n_r}.$$

For trivial s, f_s is strictly concave modulo G and its unique critical point is the limit. The twisted tensors T^s have the same complex solutions as T, and hence the same roots.

Proof. (1) The point $\xi = u^{(1)} \otimes \cdots \otimes u^{(r)}$ lies in the torus of Σ, and $\ell_1, \dots, \ell_d$ and ℓ_x with $x = P(u)_{\mathbf{1}}$ vanish at ξ by Lemma 2.1. So $\mathrm{ev}_\xi \neq 0$ lies in the left kernel of $\Phi(x)$, as in the proof of Theorem 3.4(2). (2) This is Theorem 3.2 together with Lemma 2.2; the roots of an irreducible polynomial are conjugate. (3) This follows from the formula for df. (4) If P is real then so is the rank-one tensor $P/T = u^{(1)} \otimes \cdots \otimes u^{(r)}$, so after a rescaling in G every $u^{(k)}$ is real. Write $u^{(k)} = s^{(k)} e^{a^{(k)}}$ coordinatewise. For trivial s, f_s is linear minus a positive combination of exponentials of the linear forms $a^{(1)}_{c_1} + \cdots + a^{(r)}_{c_r}$, which vanish simultaneously only in the directions of G. Finally, $T^s \circ (u^{(1)} \otimes \cdots \otimes u^{(r)}) = T \circ \big((s^{(1)} u^{(1)}) \otimes \cdots \otimes (s^{(r)} u^{(r)})\big)$. □

Algebraically, the limit is on the same footing as the other complex solutions; only positivity selects it (Theorem 3.4). The number N of complex solutions is the maximum likelihood degree of the scaled Segre variety [1].

7. Remarks

Remark 7.1 (Higher marginals)**.** Iterative proportional fitting with prescribed marginals of higher order (a hierarchical log-linear model with configuration A') has the same structure: the margin conditions are linear in y, and y ranges over the toric variety $X_{A'}$. The corner entry of the limit is then a root of the A'-resultant of the margin forms and ℓ_x, which is the

Chow form of $X_{A'}$ evaluated at the margin space and has degree at most $\deg X_{A'}$ in x. In the Segre case, Sturmfels and Zelevinsky [13] make this a single determinant. In general one needs Macaulay-type quotient formulas for sparse resultants [3].

Remark 7.2 (Open questions)**.** For matrices, the diagonal form of Theorem 5.3 separates the dependence on T (Plücker coordinates) from the dependence on the margins (the operator Q). It would be interesting to find a basis with the same property for tensors. The irreducibility of $\tilde{E}_T$ in every format (equivalently, by Theorem 4.5, whether the corner entries x_ξ of the N generic complex solutions are distinct), the Galois group of $\tilde{E}_T$ over $\mathbb{Q}(T,\mu)$, and the behavior for tensors with prescribed zero patterns are also open.

Appendix A. Examples

We illustrate Theorem 3.4 and Proposition 6.1. The corner entry of every positive tensor with margins $\mu^{(k)}$ lies in $I = (0, \min_k \mu_1^{(k)})$ (since the entries are positive and the slice sums through the corner equal $\mu_1^{(k)}$). In Examples A.1–A.3, $\tilde{E}_T$ has several roots in I. For each real root x we list the sign vectors $s^{(1)}, \dots, s^{(r)}$ of the real solution with corner entry x (Proposition 6.1(4)), normalized by $s_1^{(k)} = +$ for $k \geq 2$. By Theorem 3.4, the corner entry of the limit is the unique root whose sign vectors are all positive. Polynomials printed in full are normalized as in Theorem 4.5. Numerical values are rounded to the digits shown.

Example A.1. Let

$$T = \begin{pmatrix} 3 & 9 & 1 \\ 3 & 2 & 9 \\ 5 & 3 & 4 \end{pmatrix}, \qquad \rho = \kappa = (1,1,1),$$

as in [11, Example 1]. Then $N = 6$, $I = (0,1)$, and

$$\tilde{E}_T(x) = 30354912\,x^6 - 17851428\,x^5 - 68393079\,x^4 - 10189476\,x^3 + 17082657\,x^2 + 1162026\,x - 997272,$$

which is the polynomial given by [11, Theorem 3] for this T. The leading coefficient is $(-21)\cdot 24\cdot(-36)\cdot 7\cdot 239$, the product of the four 2×2 minors containing t_{11} and $\det T$. The constant term is $t_{11}^5\rho_1^3\kappa_1^3 \cdot t_{22}t_{23}t_{32}t_{33}\cdot(t_{22}t_{33}-t_{23}t_{32}) = 3^5\cdot 216\cdot(-19)$. The polynomial is irreducible over $\mathbb{Q}$, and its roots are

x	$s^{(1)}$	$s^{(2)}$
-0.7859741748	$(-,+,+)$	$(+,-,+)$
-0.7765360630	$(-,-,+)$	$(+,-,-)$
-0.3038869398	$(-,-,+)$	$(+,-,-)$
0.2766771162	$(+,+,+)$	$(+,+,+)$
0.3503294291	$(+,-,+)$	$(+,+,-)$
1.827480889	$(+,-,-)$	$(+,+,-)$

Two roots lie in I, and only one has all sign vectors positive. By Theorem 3.4, $P_{11}^* = 0.2766771162$, as computed in [11], and the left null vector of $\Phi(P_{11}^*)$ gives the whole limit (Corollary 3.3):

$$P^* = \begin{pmatrix} 0.27668 & 0.64805 & 0.07528 \\ 0.25194 & 0.13114 & 0.61692 \\ 0.47138 & 0.22082 & 0.30780 \end{pmatrix}.$$

Example A.2. Let T be the $2 \times 2 \times 3$ tensor with slices

$$(t_{1jk}) = \begin{pmatrix} 4 & 3 & 8 \\ 1 & 6 & 4 \end{pmatrix}, \qquad (t_{2jk}) = \begin{pmatrix} 7 & 9 & 6 \\ 9 & 5 & 4 \end{pmatrix},$$

and let $\mu^{(1)} = (9, 15)$, $\mu^{(2)} = (7, 17)$, $\mu^{(3)} = (7, 7, 10)$. Then $N = 12$ and $I = (0, 7)$. The four nonreal roots of $\tilde{E}_T$ are

$$-10.41457216 \pm 12.96393243\, i, \qquad 2.276903026 \pm 9.430689749\, i;$$

by Proposition 6.1 they are the corner entries of solutions with nonreal scaling factors. The real roots are

x	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$
-14.79904486	$(-,-)$	$(+,-)$	$(+,+,-)$
0.2885708706	$(+,-)$	$(+,-)$	$(+,-,+)$
0.5554305597	$(+,-)$	$(+,-)$	$(+,-,+)$
0.5715181134	$(+,+)$	$(+,+)$	$(+,+,+)$
1.058504652	$(+,-)$	$(+,-)$	$(+,-,+)$
1.751282005	$(+,-)$	$(+,-)$	$(+,-,+)$
10.20633540	$(+,-)$	$(+,-)$	$(+,-,+)$
13.98321693	$(+,-)$	$(+,-)$	$(+,-,-)$

Five roots lie in I, and only one has all sign vectors positive. Hence $P^*_{111} = 0.5715181134$.

Example A.3. Let

$$T = \begin{pmatrix} 8 & 2 & 9 & 7 \\ 9 & 1 & 4 & 2 \\ 3 & 7 & 4 & 5 \\ 2 & 6 & 9 & 9 \end{pmatrix}, \qquad \rho = (2, 12, 4, 6), \qquad \kappa = (1, 15, 2, 6).$$

Then $N = \binom{6}{3} = 20$ and $I = (0, 1)$. The polynomial $\tilde{E}_T$ is irreducible over $\mathbb{Q}$; we omit its coefficients. Two of its roots are not real, namely $-0.7139207340 \pm 0.2868301405\, i$. Of the eighteen real roots, seven are negative, four are greater than 1, and seven lie in I, of which only one has all sign vectors positive:

x	$s^{(1)}$	$s^{(2)}$
0.01731304145	$(+,+,-,-)$	$(+,-,+,+)$
0.04413940135	$(+,+,-,+)$	$(+,-,+,+)$
0.05702300275	$(+,+,+,+)$	$(+,+,+,+)$
0.09535084645	$(+,+,-,-)$	$(+,-,+,-)$
0.2014801225	$(+,+,-,+)$	$(+,-,+,-)$
0.2743580597	$(+,+,+,-)$	$(+,+,+,-)$
0.7484409356	$(+,+,-,+)$	$(+,-,+,-)$

Hence $P^*_{11} = 0.05702300275$.

Example A.4. Let T be the $2 \times 2 \times 2$ tensor with $t_{ijk} = 4(i-1) + 2(j-1) + k$, that is, with slices

$$(t_{1jk}) = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \qquad (t_{2jk}) = \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix},$$

and let $\mu^{(1)} = (1, 3)$, $\mu^{(2)} = (2, 2)$, $\mu^{(3)} = (3, 1)$. The 2×2 minors of T are nonzero, but its hyperdeterminant vanishes, as it does for every $2 \times 2 \times 2$ tensor whose entries are affine

functions of the indices. Hence the x^N coefficient in Theorem 4.5 vanishes, so $\deg \tilde{E}_T$ drops from 6 to 5:

$$\tilde{E}_T(x) = -8192\,x^5 - 142272\,x^4 - 780288\,x^3 - 1686368\,x^2 - 1085952\,x + 387072,$$

which is irreducible over $\mathbb{Q}$. Its constant term is $t_{111}^5(1\cdot 2\cdot 3)^2 \cdot t_{122}t_{212}t_{221}\cdot t_{222}^2 = 1\cdot 36\cdot 168\cdot 64$, as in Theorem 4.5. Its roots are real and simple:

x	$s^{(1)}$	$s^{(2)}$	$s^{(3)}$
-9.395305470	$(-,+)$	$(+,-)$	$(+,-)$
-3.082353684	$(-,+)$	$(+,-)$	$(+,+)$
-2.773617481	$(-,+)$	$(+,+)$	$(+,-)$
-2.364676371	$(-,-)$	$(+,-)$	$(+,-)$
0.2487655058	$(+,+)$	$(+,+)$	$(+,+)$

Hence $P_{111}^* = 0.2487655058$.

References

[1] C. Améndola, N. Bliss, I. Burke, C. Gibbons, M. Helmer, S. Hoşten, E. Nash, J. Rodriguez, D. Smolkin, *The maximum likelihood degree of toric varieties*, J. Symbolic Comput. 92 (2019) 222–242, doi:10.1016/j.jsc.2018.04.016.

[2] D. N. Bernshtein, *The number of roots of a system of equations*, Funct. Anal. Appl. 9 (1975) 183–185, doi:10.1007/BF01075595.

[3] C. D'Andrea, *Macaulay style formulas for sparse resultants*, Trans. Amer. Math. Soc. 354 (2002) 2595–2629, doi:10.1090/S0002-9947-02-02910-0.

[4] C. D'Andrea, M. Sombra, *A Poisson formula for the sparse resultant*, Proc. Lond. Math. Soc. (3) 110 (2015) 932–964, doi:10.1112/plms/pdu069.

[5] M. C. Fang, *Closed form of a generalized Sinkhorn limit*, arXiv:2506.06338 (2025).

[6] M. C. Fang, *The eliminant of the generalized Sinkhorn limit: a resolution of the Rowland–Wu conjectures*, preprint, Zenodo, doi:10.5281/zenodo.22864938 (2026).

[7] J. Franklin, J. Lorenz, *On the scaling of multidimensional matrices*, Linear Algebra Appl. 114/115 (1989) 717–735, doi:10.1016/0024-3795(89)90490-4.

[8] I. M. Gelfand, M. M. Kapranov, A. V. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*, Birkhäuser, Boston, 1994.

[9] J. Kruithof, *Telefoonverkeersrekening*, De Ingenieur 52 (1937) E15–E25.

[10] M. B. Nathanson, *Alternate minimization and doubly stochastic matrices*, Integers 20A (2020), #A10, doi:10.5281/zenodo.10657507.

[11] E. Rowland, J. Wu, *The entries of the Sinkhorn limit of an $m \times n$ matrix*, arXiv:2409.02789 (2024).

[12] R. Sinkhorn, *Diagonal equivalence to matrices with prescribed row and column sums*, Amer. Math. Monthly 74 (1967) 402–405.

[13] B. Sturmfels, A. Zelevinsky, *Multigraded resultants of Sylvester type*, J. Algebra 163 (1994) 115–127, doi:10.1006/jabr.1994.1007.