Available online at www.sciencedirect.com

SciVerse ScienceDirect

Advances in Mathematics 244 (2013) 678–707

ADVANCES IN Mathematics

www.elsevier.com/locate/aim

# The entropic discriminant

Raman Sanyal[a], Bernd Sturmfels[b], Cynthia Vinzant[c,*]

[a] Institut für Mathematik, Freie Universität Berlin, Germany
[b] Department of Mathematics, University of California, Berkeley, USA
[c] Department of Mathematics, University of Michigan, Ann Arbor, USA

Received 16 October 2011; accepted 30 May 2013
Available online 20 June 2013

Communicated by Michael Hopkins

**Abstract**

The entropic discriminant is a non-negative polynomial associated to a matrix. It arises in contexts ranging from statistics and linear programming to singularity theory and algebraic geometry. It describes the complex branch locus of the polar map of a real hyperplane arrangement, and it vanishes when the equations defining the analytic center of a linear program have a complex double root. We study the geometry of the entropic discriminant, and we express its degree in terms of the characteristic polynomial of the underlying matroid. Singularities of reciprocal linear spaces play a key role. In the corank-one case, the entropic discriminant admits a sum of squares representation derived from the discriminant of a characteristic polynomial of a symmetric matrix.


*Keywords:* Matroid; Discriminant; Hyperplane arrangement; Ramification locus; Fully real system

## 1. Introduction

Entropy maximization for log-linear models in statistics leads to the optimization problem

$$\text{maximize } |x_1 x_2 \cdots x_n| \quad \text{subject to } A\mathbf{x} = \mathbf{b}. \tag{1}$$

Here $A$ is a fixed real $d \times n$-matrix of rank $d$ none of whose columns are zero. The right hand side vector $\mathbf{b} \in \mathbb{R}^d$ is a parameter that is allowed to vary. The problem (1) has a unique local solution in the interior of each bounded region of the hyperplane arrangement $\{x_i = 0\}_{i \in [n]}$ inside

* Corresponding author.
*E-mail addresses:* sanyal@math.fu-berlin.de (R. Sanyal), bernd@math.berkeley.edu (B. Sturmfels), vinzant@umich.edu (C. Vinzant).

0001-8708/$ - see front matter 
http://dx.doi.org/10.1016/j.aim.2013.05.019

the $(n-d)$-dimensional affine space $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}\}$. The bounded regions are $(n-d)$-dimensional convex polytopes. The number of bounded regions in this arrangement is constant for an open, dense set of vectors $\mathbf{b}$. This number, $\mu(A)$, is a quantity known in matroid theory as the *Möbius invariant*. The local optima of (1) are the *analytic centers* of these $\mu(A)$ polytopes. They are characterized by

$$A \cdot \mathbf{x} = \mathbf{b} \quad \text{and} \quad \left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_n}\right) \text{ lies in the row space of } A. \tag{2}$$

This translates into a system of polynomial equations in the variables $x_1, \ldots, x_n$. It is known [23,27] that all complex solutions of this system actually lie in $\mathbb{R}^n$. Thus $\mu(A)$ is the algebraic degree of (2).

The aim of this article is to address the following question: Under what condition on the right hand side $\mathbf{b}$ do two of the $\mu(A)$ solutions of polynomial equations represented by (2) come together? The set of all complex right hand side vectors $\mathbf{b} \in \mathbb{C}^d$ for which this happens is an algebraic variety $H_A$ in $\mathbb{C}^d$, called the *entropic discriminant*. Under mild hypotheses on the matrix $A$, the entropic discriminant $H_A$ is a hypersurface and we identify it with its defining polynomial, denoted by $H_A(\mathbf{b})$. This is a non-negative polynomial whose real zeros lie in certain linear subspaces of codimension 2.

**Example 1.** Let $d = 3$ and $n = 5$. The following $3 \times 5$-matrix has Möbius invariant $\mu(A) = 4$:

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

The entropic discriminant of $A$ is a homogeneous polynomial $H_A(b_1, b_2, b_3)$ of degree 8. It equals

$$\begin{aligned} &288b_2^2b_3^2(b_1^2b_2^2 + b_1^2b_3^2 + b_2^2s_1^2 + b_2^2s_2^2 + b_2^2s_3^2 + b_3^2s_1^2 + b_3^2s_2^2 + b_3^2s_3^2) + 1773b_2^4b_3^4 \\ &\quad + 720b_2^2b_3^2(s_1^2s_2^2 + b_1^2s_3^2) + 192(b_1^2b_2^4s_1^2 + b_2^4s_2^2s_3^2 + b_1^2b_3^4s_2^2 + b_3^4s_1^2s_3^2) \\ &\quad + 1216(b_1^2b_2^2b_3^2s_1^2 + b_1^2b_2^2b_3^2s_2^2 + b_2^2b_3^2s_1^2s_3^2 + b_2^2s_2^2s_3^2b_3^2) + 256b_1^2s_1^2s_2^2s_3^2 \\ &\quad + 320(b_1^2b_2^2s_1^2s_2^2 + b_1^2b_2^2s_1^2s_3^2 + b_1^2b_2^2s_2^2s_3^2 + b_1^2b_3^2s_1^2s_2^2 + b_1^2b_3^2s_1^2s_3^2 + b_1^2b_3^2s_2^2s_3^2 \\ &\quad + b_2^2s_1^2s_2^2s_3^2 + b_3^2s_1^2s_2^2s_3^2), \end{aligned}$$

where $s_1 = b_1 - b_2$, $s_2 = b_1 - b_3$, and $s_3 = b_1 - b_2 - b_3$. Thus $H_A(\mathbf{b})$ is a sum of squares of quartics.

It coincides with the discriminant of the following system of equations in three unknowns:

$$\begin{aligned} &1/z_1 + 1/(z_1 + z_2) + 1/(z_1 + z_3) = b_1, \\ &1/z_2 + 1/(z_1 + z_2) = b_2, \\ &1/z_3 + 1/(z_1 + z_3) = b_3. \end{aligned}$$

These equations are equivalent to (2) if we take $(z_1, z_2, z_3)$ to be coordinates for the row space of $A$. There are four solutions for any $\mathbf{b} = (b_1, b_2, b_3) \in \mathbb{C}^3$. They are distinct if and only if $H_A(\mathbf{b}) \neq 0$. The entropic discriminant $H_A(\mathbf{b})$ is a non-negative polynomial having precisely four real zeros:

$$V_{\mathbb{R}}(H_A) = \big\{(0:1:0),\ (0:0:1),\ (1:1:0),\ (1:0:1)\big\} \subset \mathbb{P}^2. \tag{3}$$

The complex variety $V_{\mathbb{C}}(H_A)$ is a curve of degree 8 in the projective plane with coordinates $(b_1 : b_2 : b_3)$. That curve is singular at its four real points. In addition, it has 16 isolated complex singularities. $\square$

We shall study the systems (2) for arbitrary $d$, $n$, and $A$. The following is our main result.

**Theorem 2.** *Let $A$ be a real $d \times n$-matrix of rank $d$ whose columns span $\geq d+1$ distinct lines. The entropic discriminant is a hypersurface, defined by a homogeneous polynomial $H_A(\mathbf{b})$ of degree*

$$\deg\, H_A(\mathbf{b}) = 2(-1)^d \cdot (d\chi(0) + \chi'(0)), \tag{4}$$

*where $\chi(t)$ is the characteristic polynomial of the rank $d$ matroid of $A$. For generic matrices $A$, this degree equals $2(n-d)\binom{n-1}{d-2}$. The polynomial $H_A(\mathbf{b})$ is non-negative for all arguments in $\mathbb{R}^d$.*

The generic degree $2(n-d)\binom{n-1}{d-2}$ is always an upper bound on the degree of the entropic discriminant, and the equality holds when the matroid of $A$ is uniform; cf. Proposition 33. For example, for generic matrices $A$ of size $3 \times 5$, the degree of $H_A(\mathbf{b})$ equals 16, and not 8 as in Example 1.

This article is organized as follows. In Section 2 we examine the polar map of a product of linear forms. The entropic discriminant is shown to coincide with the branch locus of that polar map. For example, consider the polar map of the binary form $f(z_1, z_2) = z_1(z_1 + 2z_2)(z_1 + 3z_2)(z_1 + az_2)$:

$$\nabla_f : \mathbb{P}^1 \to \mathbb{P}^1, \qquad (z_1 : z_2) \mapsto \left( \frac{\partial f}{\partial z_1}(z_1, z_2) : \frac{\partial f}{\partial z_2}(z_1, z_2) \right).$$

The branch locus of this map consists of the four zeros of the binary quartic $H_A(b_1, b_2)$ in Example 3. This connects our study of $H_A(\mathbf{b})$ to the topological theory of hyperplane arrangements [4,5], and to topics in classical algebraic geometry that are found in Chapter 1 of Dolgachev's book [6].

Section 3 is concerned with the important special case $n = d + 1$. Here the entropic discriminant has expected degree $d(d-1)$ and we can write it explicitly as a sum of squares. This expression is derived from known results on the discriminant of the characteristic polynomial of a symmetric matrix [2,14,15,18]. We then apply this to resolve two problems left open in the literature, namely the Sottile–Mukhin Conjecture [1] on the discriminant of the derivative of a univariate polynomial, and Conjecture 7.9 in [25] concerning real critical double eigenvalues of a net of symmetric matrices.

For any linear subspace $\mathcal{L}$ of $\mathbb{C}^n$, its *reciprocal* $\mathcal{L}^{-1}$ is defined as the Zariski closure of the set

$$\left\{ \left( \frac{1}{u_1}, \frac{1}{u_2}, \dots, \frac{1}{u_n} \right) \in \mathbb{C}^n : (u_1, u_2, \dots, u_n) \in \mathcal{L} \cap (\mathbb{C}^*)^n \right\}. \tag{5}$$

In Section 5 we study the geometry of the *reciprocal plane* $\mathcal{L}^{-1}$, further extending the line of work from Proudfoot–Speyer [20] to Huh–Katz [13]. We identify a minimal system of defining equations for $\mathcal{L}^{-1}$, we characterize the singular locus of $\mathcal{L}^{-1}$, and we determine all tangent cones. The relationship between that singular locus, the ramification locus of the map $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$, and the entropic discriminant $H_A(\mathbf{b})$ is studied in detail in Section 7. In Corollary 37 we show that the real variety defined by the polynomial $H_A(\mathbf{b})$ is a union of linear spaces of codimension 2 in $\mathbb{P}^{d-1}$. We saw this already for one instance in Example 1, where $d = 3$ and the real variety is finite.

Theorem 2 is proved in Section 6. However, one subtle but essential point needs to be taken care of before that proof. In order for (4) to be the correct degree, a more refined notion of entropic discriminant is required. Namely, we shall define $H_A(\mathbf{b})$ as the polynomial defining the cycle-theoretic branch locus of the restriction to $\mathcal{L}^{-1}$ of the linear map $A : \mathbb{C}^n \to \mathbb{C}^d$, where $\mathcal{L}$ is the row space of $A$. The following example justifies this "fine print" in Definition 28.

**Example 3.** Let $d = 2, n = 4$ and $A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & a \end{pmatrix}$ where $a$ is a real parameter. For general values of $a$, the entropic discriminant is irreducible and has degree 4, as predicted by Theorem 2:

$$\begin{aligned} H_A(b_1, b_2) = {} & (2268a^4 - 9720a^3 + 11664a^2)b_1^4 - (3000a^4 - 12528a^3 \\ & + 12960a^2 + 5184a)b_1^3 b_2 + (1744a^4 - 7980a^3 + 10584a^2 \\ & - 2160a + 5184)b_1^2 b_2^2 - (500a^4 - 2612a^3 + 4680a^2 \\ & - 3888a + 4320)b_1 b_2^3 + (63a^4 - 400a^3 + 999a^2 - 1350a + 1188)b_2^4. \end{aligned}$$

For special values of the parameter $a$, this expression factors over $\mathbb{Q}$. For $a = 6$, it is the square $972(36b_1^2 - 24b_1 b_2 + 5b_2^2)^2$. Thus, here the four points of $V_{\mathbb{C}}(H_A)$ in $\mathbb{P}^1$ are two double points. □

Our initial motivation for embarking on this project was a model in theoretical neuroscience proposed by Hillar and Wibisono [11]. These authors investigate the *retina equations* which characterize the maximum entropy distribution for a graphical model $G$ with $n$ edges having continuous random variables on $d$ nodes that represent the firing pattern of $d$ neurons. Their equations are

$$\sum_{j \in \mathcal{N}(i)} \frac{1}{z_i + z_j} = b_i \quad \text{for } i = 1, 2, \ldots, d, \tag{6}$$

where $\mathcal{N}(i)$ is the set of all nodes that are adjacent to the node $i$. The real numbers $b_1, b_2, \ldots, b_d$ are parameters that serve as the sufficient statistics of the desired maximum entropy distribution.

To fit the system (6) into our framework, we introduce new unknowns $x_{ij} = 1/(z_i + z_j)$ for all edges $\{i, j\} \in E(G)$. This translates (6) into the linear system $A \cdot \mathbf{x} = \mathbf{b}$, where $A$ is the node-edge incidence matrix of $G$ and $\mathbf{x} = (x_{ij} : \{i, j\} \in E(G))$ is a column vector of unknowns. Of course, these unknowns obey the additional constraints that $\mathbf{x}$ must lie in the reciprocal plane $\mathcal{L}^{-1}$, where $\mathcal{L}$ is the row space of $A$. Thus the retina equations of Hillar and Wibisono fit our format (2):

$$A \cdot \mathbf{x} = \mathbf{b} \quad \text{and} \quad \mathbf{x} \in \mathcal{L}^{-1}. \tag{7}$$

The entropic discriminant $H_A(\mathbf{b})$ characterizes measurements $\mathbf{b}$ for which the retina equations (6) or (7) have multiple roots. Of particular interest is the case $n = \binom{d}{2}$, when $G = K_d$ is the complete graph, and the sum in (6) is over $j \in \{1, \ldots, n\} \setminus \{i\}$. The characteristic polynomial $\chi_d(t)$ of the corresponding matroid was computed by Zaslavsky [30], in his work of colorings of signed graphs:

$$\chi_d(t) = \sum_{k=0}^{d} \left( \left\{ \begin{matrix} d \\ k \end{matrix} \right\} + d \left\{ \begin{matrix} d-1 \\ k \end{matrix} \right\} \right) (t-1)_k^{(2)}. \tag{8}$$

Here $\left\{ \begin{smallmatrix} d \\ k \end{smallmatrix} \right\}$ is the Stirling number of the second kind and $(x)_{k+1}^{(2)} = x(x-2)\cdots(x-2k)$ is the generalized falling factorial. One can also compute $\chi_d(t)$ with the exponential generating function

$$\sum_{d\geq 0} \chi_d(t) \cdot \frac{x^d}{d!} = (1+x) \cdot \big(2 \cdot \exp(x) - 1\big)^{(t-1)/2}, \tag{9}$$

found in [24, Exercise 5.25]. Using these formulas, one obtains the first few values of the degree of $H_A(\mathbf{b})$ and of the number of solutions of the retina equations on the complete graph $G = K_d$:

$$\begin{array}{rcccccc} d = 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ \deg(H_A(\mathbf{b})) = 22 & 270 & 3148 & 38990 & 524858 & 7705572 & 123087958. \\ \mu(A) = 7 & 51 & 431 & 4208 & 46824 & 586141 & 8161237 \end{array} \tag{10}$$

The requisite combinatorics is developed in Section 4. It covers material from matroid theory, focusing on geometric interpretations of the characteristic polynomial and the Möbius invariant. For instance, the third row in (10) is computed from the series in (9) for $t = 0$, using formula (28).

## 2. The polar map of a product of linear forms

The $d \times n$-matrix $A = (a_{ij})$ determines a product of linear forms in $d$ unknowns $\mathbf{z} = (z_1, \ldots, z_d)$:

$$f(\mathbf{z}) = \prod_{j=1}^{n} \left( \sum_{i=1}^{d} a_{ij} z_i \right). \tag{11}$$

The hypersurface $V_{\mathbb{C}}(f)$ is an arrangement of $n$ hyperplanes in the complex projective space $\mathbb{P}^{d-1}$. The *polar map* of this hypersurface is the rational map

$$\nabla_f : \mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{d-1}, \qquad \mathbf{z} \mapsto \left( \frac{\partial f}{\partial z_1}(\mathbf{z}) : \frac{\partial f}{\partial z_2}(\mathbf{z}) : \cdots : \frac{\partial f}{\partial z_d}(\mathbf{z}) \right).$$

The base locus of $\nabla_f$ is the singular locus of $V_{\mathbb{C}}(f)$, and this is the union of all codimension-2 strata in the hyperplane arrangement. If the columns of $A$ are linearly independent then $\nabla_f$ is the Cremona transformation of classical algebraic geometry, and, in general, the polar map $\nabla_f$ is also known as the *polar Cremona transformation* [5]. The Jacobian of $\nabla_f$ is the Hessian of the polynomial $f$, that is, the symmetric matrix of second derivatives. We consider its determinant

$$\mathrm{Hess}(f) = \det \left( \frac{\partial^2 f}{\partial z_i \partial z_j} \right)_{1\leq i,j\leq d}.$$

This is a homogeneous polynomial of degree $d(n-2)$. Its zero set in $\mathbb{P}^{d-1}$, denoted by $V_{\mathbb{C}}(\mathrm{Hess}(f))$, is also referred to as the Hessian of $f$. We are interested in the image of that hypersurface under $\nabla_f$.

**Proposition 4.** *The entropic discriminant equals the image of the Hessian under the polar map:*

$$V_{\mathbb{C}}(H_A) = \textit{closure of } \nabla_f\big(V_{\mathbb{C}}(\mathit{Hess}(f)) \setminus V_{\mathbb{C}}(f)\big). \tag{12}$$

**Proof.** Let $\mathcal{L}^{-1}$ denote the reciprocal of the subspace $\mathcal{L}$ spanned by the rows of $A$, regarded as a subvariety of $\mathbb{P}^{n-1}$. The variety $\mathcal{L}^{-1}$ is the closure of the image of the map $\mathbb{P}^{d-1} \dashrightarrow \mathbb{P}^{n-1}$ that takes a general point $\mathbf{z} = (z_1 : \cdots : z_d)$ in $\mathbb{P}^{d-1}$ to $(\mathbf{z}A)^{-1} = \big((\sum_{i=1}^d a_{i1}z_i)^{-1} : \cdots : (\sum_{i=1}^d a_{in}z_i)^{-1}\big)$ in $\mathbb{P}^{n-1}$. The polar map is the composition of this map with the linear projection $\mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{d-1}$, $\mathbf{x} \mapsto A\mathbf{x}$. In symbols, we have $\nabla_f(\mathbf{z}) = A\big((\mathbf{z}A)^{-1}\big)$. This observation shows that the fiber of $\nabla_f$ over a general real point $\mathbf{b} \in \mathrm{Im}(\nabla_f)$ consists of $\mu(A)$ real points in $\mathbb{P}^{d-1}$, namely, the points represented by the analytic centers in the arrangement defined by the coordinate hyperplanes in the affine space $\{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{b}\}$. This result was also obtained in dual form by Dimca and Papadima in [4, Corollary 4 (1)].

For special complex points $\mathbf{b} \in \mathbb{P}^{d-1}$, two of its $\mu(A)$ preimages under $\nabla_f$ may coincide. At such a preimage $\mathbf{z}$ of multiplicity $\geq 2$, the Jacobian of $\nabla_f$ drops rank, and the Hessian of $f$ vanishes at $\mathbf{z}$. Conversely, points $\mathbf{z}$ outside the hyperplane arrangement $V_{\mathbb{C}}(f)$ at which the polynomial $\mathrm{Hess}(f)$ vanishes must be double roots of the system of equations $\nabla_f(\mathbf{z}) = \mathbf{b}$. Since the parametrization $\mathbf{z} \mapsto \mathbf{x} = (\mathbf{z}A)^{-1}$ maps $\mathbb{P}^{d-1}$ birationally onto the reciprocal plane $\mathcal{L}^{-1}$, such double roots appear if and only if the intersection $\mathcal{L}^{-1} \cap \{\mathbf{x} \in \mathbb{P}^{n-1} : A\mathbf{x} = \mathbf{b},\ x_1x_2\cdots x_n \neq 0\}$ has a point of multiplicity $\geq 2$. This condition on $\mathbf{b}$ is the geometric definition of the entropic discriminant $H_A$. □

We have not yet addressed the question whether the entropic discriminant actually has codimension 1, and this may in fact not be the case. For instance, if $A$ is the identity matrix and $f = z_1z_2\cdots z_d$ then $\nabla_f$ is the classical Cremona transformation on $\mathbb{P}^{d-1}$ and $\mathrm{Hess}(f) = (-1)^{d-1}(d-1)f^{d-2}$. Here, the Hessian coincides with the hyperplane arrangement, and the entropic discriminant is not a hypersurface. We shall see that this is essentially the only exceptional case.

The matrix $A = (a_{ij})$ is called *basic* if its column rays lie on $d$ distinct lines in $\mathbb{R}^d$. Since $A$ has rank $d$ and no zero columns, this means that the distinct column directions form a basis of $\mathbb{R}^d$.

**Corollary 5.** *If $A$ is not basic then the entropic discriminant is a hypersurface in $\mathbb{P}^{d-1}$.*

**Proof.** A classical formula [16, p. 660, Example 10] for the Hessian determinant of $f$ states that

$$\mathrm{Hess}(f) = (-1)^{d-1}(n-1)f^{d-2} \cdot \sum_{I \in \binom{[n]}{d}} \det(A_I)^2 \prod_{k \in [n]\setminus I} (a_{1k}z_1 + a_{2k}z_2 + \cdots + a_{dk}z_d)^2 \tag{13}$$

where $A_I$ denotes the $d \times d$-submatrix of $A$ with column indices $I$. If $A$ is not basic, then at least two summands are not scalar multiples of each other. This implies that the Hessian hypersurface is not contained in the hyperplane arrangement $V_{\mathbb{C}}(f)$. The polar map $\nabla_f$ is a finite-to-one morphism on the open set $\mathbb{P}^{d-1} \setminus V_{\mathbb{C}}(f)$, and hence it maps the Hessian to a hypersurface in $\mathbb{P}^{d-1}$, namely $H_A$. □

**Corollary 6.** *For any non-basic $A$, the polynomial $H_A(\mathbf{b})$ is homogeneous and nonnegative on $\mathbb{R}^d$.*

**Proof.** It suffices to prove this for the square-free polynomial $\tilde{H}_A(\mathbf{b})$ that vanishes on $V_{\mathbb{C}}(H_A)$. Indeed, if $\tilde{H}_A(\mathbf{b})$ is homogeneous and nonnegative then so is any real product of its factors.

Homogeneity is straightforward since the geometric definition ensures that $\mathbf{b} \in V_{\mathbb{C}}(H_A)$ implies $\lambda\mathbf{b} \in V_{\mathbb{C}}(H_A)$. To show non-negativity, let $K$ denote the subfield of $\mathbb{R}$ which is generated by the entries of $A$. We regard the entries of $\mathbf{b} = (b_1, \ldots, b_d)$ as indeterminates over $K$. Let $L$ be the algebraic closure of the rational function field $K(b_1, \ldots, b_d)$. Then the equation $\nabla_f(\mathbf{z}) = \mathbf{b}$ has $\mu(A)$ distinct solutions with coordinates in $L$. We substitute these solutions into the sum in (13) and we take their product in the field $L$. The result is a sum of squares in $L$ that is a symmetric polynomial in the roots. It is invariant under the action of the Galois group of $L$ over $K(b_1, \ldots, b_d)$ and thus lies in $K(b_1, \ldots, b_d)$. The sum of squares representation over $L$ ensures that this rational function is non-negative under all specializations of $\mathbf{b}$ to $\mathbb{R}$ at which it does not have a pole. The numerator of this rational function is a product of the factors of $\tilde{H}_A(\mathbf{b})$. We conclude that $\tilde{H}_A(\mathbf{b})$ does not change signs on $\mathbb{R}^d$. Hence, either $\tilde{H}_A(\mathbf{b})$ or $-\tilde{H}_A(\mathbf{b})$ is non-negative on $\mathbb{R}^d$. $\square$

The above argument shows that $H_A(\mathbf{b})$ is non-negative but it does not furnish a representation of $H_A(\mathbf{b})$ as a sum of squares of polynomials. We also note that the computation of $H_A(\mathbf{b})$ from $\mathrm{Hess}(f)$ is a task of elimination theory that is quite non-trivial even for moderate values of $d$ and $n$.

One case where the elimination problem can be solved more easily is $d = 2$. Here $f(z_1, z_2)$ is a binary form of degree $n$ enjoying the property that all its zeros on the line $\mathbb{P}^1$ are defined over $\mathbb{R}$. The polar map $\nabla_f$ takes the complex projective line $\mathbb{P}^1$ to itself. This map has degree $n-1$, i.e. the fiber over a general point $\mathbf{b} \in \mathbb{P}^1$ consists of $n-1$ points. We are interested in those points $\mathbf{b}$ on the line $\mathbb{P}^1$ for which two or more of the points in its fiber collide. The Hessian of $f$ equals

$$\det\begin{pmatrix} \dfrac{\partial^2 f}{\partial z_1^2} & \dfrac{\partial^2 f}{\partial z_1 \partial z_2} \\ \dfrac{\partial^2 f}{\partial z_1 \partial z_2} & \dfrac{\partial^2 f}{\partial z_2^2} \end{pmatrix} = (1-n) \cdot \sum_{1 \le i < j \le n} (a_{1i}a_{2j} - a_{1j}a_{2i})^2 \cdot \prod_{k \in [n]\setminus\{i,j\}} (a_{1k}z_1 + a_{2k}z_2)^2.$$

This is a binary form of degree $2n-4$, so it defines a configuration of $2n-4$ points in $\mathbb{P}^1$. All points have non-real coordinates. The entropic discriminant of $f$ is the image of these $2n-4$ points under the polar map $\nabla_f$. Proposition 4 gives the following rule for computing the entropic discriminant:

$$H_A(b_1, b_2) = \mathrm{Resultant}_{\mathbf{z}}\left(\mathrm{Hess}(f(\mathbf{z})), b_2 \frac{\partial f}{\partial z_1}(\mathbf{z}) - b_1 \frac{\partial f}{\partial z_2}(\mathbf{z})\right). \tag{14}$$

This formula can be rewritten as the discriminant of a binary form:

$$H_A(b_1, b_2) = \mathrm{Discriminant}_{\mathbf{z}}\left(b_2 \frac{\partial f}{\partial z_1}(\mathbf{z}) - b_1 \frac{\partial f}{\partial z_2}(\mathbf{z})\right). \tag{15}$$

The binary form $H_A(b_1, b_2)$ has degree $2n-4$ provided no two columns of $A$ are parallel. Being nonnegative, the entropic discriminant is a sum of squares of binary forms of degree $n-2$ over $\mathbb{R}$.

**Example 7.** Let $n = 3$ and consider a general binary cubic with real zeros:

$$f = (a_{11}z_1 + a_{21}z_2)(a_{12}z_1 + a_{22}z_2)(a_{13}z_1 + a_{23}z_2).$$

The sum of squares representation in (13) tells us that the Hessian of $f$ equals

$$(a_{11}a_{22} - a_{21}a_{12})^2(a_{13}z_1 + a_{23}z_2)^2 + (a_{11}a_{23} - a_{21}a_{13})^2(a_{12}z_1 + a_{22}z_2)^2 \\ + (a_{12}a_{23} - a_{22}a_{13})^2(a_{11}z_1 + a_{21}z_2)^2.$$

For any invertible matrix $U$, the entropic discriminant satisfies $H_{UA}(U\mathbf{b}) = H_A(\mathbf{b})$. This implies that $H_A(\mathbf{b})$ can be written in the $2 \times 2$ minors $p_{ij}$ of the matrix $2 \times 4$-matrix $(A, \mathbf{b})$. We have

$$H_A(\mathbf{b}) = (p_{12} \cdot p_{34})^2 + (p_{13} \cdot p_{24})^2 + (p_{23} \cdot p_{14})^2. \tag{16}$$

For $n = 4$, an expression for $H_A(\mathbf{b})$ in terms of the $2 \times 2$ minors of the $2 \times 5$-matrix $(A, \mathbf{b})$ is

$$(p_{12}^2 p_{34} p_{35} p_{45})^2 + (p_{13}^2 p_{24} p_{25} p_{45})^2 + (p_{14}^2 p_{23} p_{25} p_{35})^2 + (p_{14} p_{23}^2 p_{15} p_{45})^2 \\ + (p_{13} p_{24}^2 p_{15} p_{35})^2 + (p_{12} p_{34}^2 p_{15} p_{25})^2 + \frac{7}{2}(p_{23} p_{24} p_{34} p_{15}^2)^2 \\ + \frac{7}{2}(p_{13} p_{14} p_{34} p_{25}^2)^2 + \frac{7}{2}(p_{12} p_{14} p_{24} p_{35}^2)^2 + \frac{7}{2}(p_{12} p_{13} p_{23} p_{45}^2)^2. \tag{17}$$

At present we do not know how to extend the formulas (16) and (17) to $n \geq 5$. $\square$

It is natural to ask how the formulas (14) and (15) would generalize to $d \geq 3$, and the answer is given by the projective duality between the entropic discriminant and the *Steinerian hypersurface* [6, Section 1.1.6]. If $f$ is any homogeneous polynomial of degree $n$ in $\mathbf{z} = (z_1, \ldots, z_d)$ then its Steinerian is

$$\mathrm{St}_f(c_1, c_2, \ldots, c_d) = \mathrm{Discriminant}_{\mathbf{z}}\left(c_1 \frac{\partial f}{\partial z_1}(\mathbf{z}) + c_2 \frac{\partial f}{\partial z_2}(\mathbf{z}) + \cdots + c_d \frac{\partial f}{\partial z_d}(\mathbf{z})\right). \tag{18}$$

In this formula, we are taking the discriminant of a form of degree $n - 1$, namely, the polar of $f$ with respect to a generic point $\mathbf{c}$. Corollary 1.2.2 in [6] tells us that the hypersurface defined by $\mathrm{St}_f(\mathbf{c})$ is dual to the image of the hypersurface defined by $\mathrm{Hess}(f(\mathbf{z}))$ under the polar map $\nabla_f$.

In our situation, the given form $f$ is a product of linear forms as in (11), and some care needs to be taken in removing contributions from singularities. Indeed, the Steinerian $\mathrm{St}_f$ of a hyperplane arrangement is supported on that same hyperplane arrangement plus an extra component. It is this extra component we are interested in. We call this hypersurface the *residual Steinerian* of $f$.

**Corollary 8.** *The entropic discriminant of a $d \times n$-matrix $A$ is the hypersurface in $\mathbb{P}^{d-1}$ projectively dual to the residual Steinerian of the arrangement of $n$ hyperplanes given by the columns of $A$.*

Let us briefly revisit the case $d = 2$ from this point of view. We saw that the entropic discriminant consists of $2n - 4$ points on a projective line with coordinates $(b_1 : b_2)$. The Steinerian consists of $2n - 4$ points on the dual projective line with coordinates $(c_1 : c_2)$. In our formulas (14) and (15) we tacitly identified these two lines and their point configurations via $(c_1 : c_2) = (-b_2 : b_1)$.

For $d \geq 3$, the formula (18) is less useful for the purpose of computing $H_A(\mathbf{b})$ because dualizing the residual Steinerian in a computer algebra system is hard. Instead, we find it preferable to use

$$\langle H_A(\mathbf{b})\rangle = \big(\langle \operatorname{Hess}(f(\mathbf{z}))\rangle + \langle 2\times 2-\text{ minors of the } 2\times d-\text{ matrix } (\mathbf{b}, \nabla_f)\rangle\big) : \langle \nabla_f \rangle^{\infty}. \tag{19}$$

This ideal-theoretic reformulation of (12) is the direct generalization of (14) to $d \geq 3$.

Nevertheless, the (residual) Steinerian of a hyperplane arrangement remains a beautiful topic in geometry, and its interplay with the combinatorics of the entropic discriminant certainly deserves further study. We close this section with an illustration of this for lines in the plane $\mathbb{P}^2$.

**Example 9.** This example was worked out with help from Igor Dolgachev. Let $d = 3$ and suppose the matroid of $A$ is uniform. Thus $V_{\mathbb{C}}(f)$ is an arrangement of $n$ lines in general position in $\mathbb{P}^2$. By Theorem 2, the entropic discriminant $H_A$ is a curve of degree $2(n-1)(n-3)$. Its singular locus consists of the $n$ columns of $A$. By dualizing, we obtain the Steinerian $\mathrm{St}_f$, a curve of degree $3(n-2)^2$. Each of the $n$ lines occurs with multiplicity $n-2$ in the Steinerian. Removing these lines, we find that the residual Steinerian $H_A^{\vee}$ is a curve of degree $3(n-2)^2 - n(n-2) = 2(n-2)(n-3)$. □

## 3. The codimension-1 case

The discriminant of the characteristic polynomial of a symmetric matrix is non-negative because real symmetric matrices have only real eigenvalues. The study of this discriminant is a classical subject in mathematics, going back to an 1846 paper by Borchardt [2]. Explicit representations of this discriminant as a sum of squares were also presented in work of Newell [18], Ilyushechkin [14], and Lax [15]. See [25, Section 7.5] for an exposition, and work of Domokos [7] for the state of the art.

In this section we establish a relationship between this subject and the entropic discriminant. We focus on the case $n = d+1$, and we express $H_A(\mathbf{b})$ as a specialization of the discriminant of the characteristic polynomial of a symmetric matrix. We shall use this to derive the following result.

**Theorem 10.** *Let $A$ be a non-basic matrix with $d$ rows and $n = d+1$ columns. Then the entropic discriminant $H_A(\mathbf{b})$ is a sum of squares of polynomials. Moreover, if the entries of $A$ are rational numbers then $H_A(\mathbf{b})$ is a sum of squares in $\mathbb{Q}[b_1, \ldots, b_d]$.*

**Example 11.** If $d = 3, n = 4$ and $A = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$ then $H_A(b_1, b_2, b_3)$ is the sum of 10 squares

$$\frac{7}{4}b_1^4(b_2-b_3)^2 + \frac{56}{27}(b_1-b_2)^2 b_1^2 b_2^2 + \frac{1}{108}(5b_1b_2 - 9b_1b_3 - 14b_2^2 + 18b_2b_3)^2 b_1^2$$
$$+ \frac{1}{27}(5b_1b_2 - 3b_1b_3 - 8b_2^2 + 6b_2b_3)^2 b_1^2 + \frac{1}{9}(b_1b_2 + b_1b_3 - 2b_2b_3)^2(b_1 - 2b_2)^2$$
$$+ \frac{7}{108}(5b_1b_2 + 3b_1b_3 - 2b_2^2 - 6b_2b_3)^2 b_1^2 + \frac{1}{216}(13b_1b_2 - 21b_1b_3$$
$$- 7b_2^2 - 12b_2b_3 + 27b_3^2)^2 b_1^2 + \frac{1}{36}(5b_1^2b_2 - 7b_1^2b_3 - 7b_1b_2^2 + 4b_1b_2b_3$$

$$+ 9b_1b_3^2 + 14b_2^2b_3 - 18b_2b_3^2)^2 + \frac{1}{216}(5b_1b_2 - 21b_1b_3 + b_2^2 - 12b_2b_3 + 27b_3^2)^2b_1^2$$
$$+ \frac{1}{36}(5b_1^2b_2 - b_1^2b_3 - 4b_1b_2^2 - 8b_1b_2b_3 + 8b_2^2b_3)^2.$$

This expression is derived from the sum of 10 squares found at the top of p. 97 in [25]. $\square$

**Proof of Theorem 10.** Let $A$ be a non-basic $d \times (d+1)$-matrix and let $v \in \mathbb{R}^{d+1}$ span the kernel of $A$. If $v$ has a zero coordinate, say $v_{d+1} = 0$, then we can reduce our analysis to a smaller case, namely, a $(d-1) \times d$-matrix obtained by taking the columns of $A$ modulo the last column. Hence we may assume that all coordinates of $v$ are non-zero.

Next, we claim that it suffices to prove our assertions for the special case where

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & -1 \\ 0 & 1 & 0 & \cdots & 0 & -1 \\ 0 & 0 & 1 & \cdots & 0 & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix} \quad \text{and} \quad v = (1, 1, 1, \ldots, 1)^T. \tag{20}$$

That this suffices is ensured by the following transformation rule for the entropic discriminant:

$$H_{UAD}(\mathbf{b}) = H_A(U^{-1}\mathbf{b}). \tag{21}$$

This identity holds for any invertible $d \times d$-matrix $U$ and any invertible diagonal $n \times n$-matrix $D$, and its validity is easily seen from the geometric definition of $H_A$. We here use this for $n = d+1$.

We now fix $A$ and $v$ as in (20). Then $\mathcal{L} = \text{rowspace}(A)$ is the hyperplane $x_1+x_2+\cdots+x_n = 0$. Its reciprocal $\mathcal{L}^{-1}$ is the hypersurface of degree $d$ in $\mathbb{P}^d$ that is defined by the polynomial

$$\sum_{i=1}^{n} \prod_{j \neq i} x_j = \det \begin{pmatrix} x_1 + x_n & x_n & \cdots & x_n \\ x_n & x_2 + x_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & x_n \\ x_n & \cdots & x_n & x_{n-1} + x_n \end{pmatrix}. \tag{22}$$

This symmetric determinantal representation of the $(n-1)$st elementary symmetric polynomial is taken from [21]. The linear system $A\mathbf{x} = \mathbf{b}$ is equivalent to

$$x_i = b_i + x_n \quad \text{for } i = 1, 2, \ldots, n-1. \tag{23}$$

Thus the points satisfying (2) can be computed by substituting (23) into (22) and equating the resulting univariate polynomial to zero. Setting $t = x_n$, the solutions to (2) correspond to zeros of

$$p_{\mathbf{b}}(t) = \det\big(tE + \text{diag}(b_1, b_2, \ldots, b_d)\big) \quad \text{where } E = \begin{pmatrix} 2 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 1 & \cdots & 1 \\ 1 & 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 2 \end{pmatrix}. \tag{24}$$

In particular, $H_A(\mathbf{b})$ equals the discriminant of the univariate polynomial $p_{\mathbf{b}}(t)$. The following proposition applied to $E$ and $X = -\text{diag}(b_1, \ldots, b_d)$ completes the proof of the theorem. $\square$

**Proposition 12.** *Let $E \in \mathbb{R}^{m\times m}$ be a symmetric positive definite matrix and $X$ a symmetric matrix of indeterminates. Then the discriminant of the* generalized *characteristic polynomial* $\det(tE - X)$ *with respect to $t$ is a sum of squares in $\mathbb{Q}(E_{ij} : 1 \le i, j \le m)[X_{ij} : 1 \le i, j \le m]$.*

**Proof.** Since $E$ has a Cholesky factorization $E = MM^T$, it follows that

$$\det(tE - X) = \det(E) \cdot \det(tI - M^{-1}XM^{-T}). \tag{25}$$

We get a sum of squares formula from the known representations of the discriminant of the characteristic polynomial of a real symmetric matrix. However, our emphasis lies on the *rationality* of the desired formula. Following [25, Section 7.5], let $\hat{X} = M^{-1}XM^{-T}$ and consider the linear map

$$\wedge_2 \mathbb{R}^m \to \mathrm{Sym}_2\mathbb{R}^m, \qquad Z \mapsto [\hat{X}, Z] = \hat{X}Z - Z\hat{X}$$

that takes a skew-symmetric matrix to the commutator with $\hat{X}$.

Let $\{W_{ij} = e_i \wedge e_j : 1 \le i < j \le m\}$ be the standard basis for the space of skew-symmetric matrices and likewise $\{S_{ij} = e_i \cdot e_j : 1 \le i \le j \le m\}$ the standard basis for the space of symmetric matrices. Let $\Phi$ be the $\binom{m+1}{2} \times \binom{m}{2}$-matrix representing the linear map in the chosen bases. By choosing suitable bases, it can be seen that the eigenvalues of $\Phi^T\Phi$ are the squared pairwise differences of the eigenvalues of $\hat{X}$. Hence the determinant of $\Phi^T\Phi$ is the discriminant of $\det(tI - \hat{X})$. The sum of squares representation can be obtained by applying the Binet–Cauchy theorem.

To get a rational representation we apply the above reasoning to the slightly altered map

$$Z \mapsto [Z, X]_E := E^{-1}XZ - ZXE^{-1}.$$

It is clear that a representation in the standard basis is over $\mathbb{Q}(E_{ij})$ and hence yields an appropriate sum of squares. To see that this actually yields the discriminant for the generalized characteristic polynomial, choose bases $W'_{ij} = M^{-T}W_{ij}M^{-1}$ and $S'_{ij} = M^{-T}S_{ij}M^{-1}$ and verify

$$[X, W'_{ij}]_E = M^{-T}[\hat{X}, W_{ij}]M^{-1}.$$

Hence, a representation in the new bases is given by $\Phi$ above. $\square$

Evaluating the discriminant of $p_{\mathbf{b}}(t)$ in (24) leads to the following data concerning the monomial expansion of the entropic discriminant $H_A(\mathbf{b})$ of the particular matrix $A$ in (20):

| $d$ | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|
| degree of $H_A(\mathbf{b})$ | 2 | 6 | 12 | 20 | 30 |
| number of monomials | 3 | 19 | 201 | 3081 | 62683 |
| leading (lex) monomial | $b_1^2$ | $b_1^4b_2^2$ | $b_1^6b_2^4b_3^2$ | $b_1^8b_2^6b_3^4b_4^2$ | $b_1^{10}b_2^8b_3^6b_4^4b_5^2$. |

(26)

**Remark 13.** The entropic discriminant $H_A(\mathbf{b})$ is a symmetric polynomial in $b_1, \ldots, b_d$ since the set of rows of the matrix $A$ is invariant under permutations. It thus admits a unique representation as a polynomial in the elementary symmetric polynomials

$$e_k = \sum_{1\le i_1<\cdots<i_k\le n} b_{i_1}b_{i_2}\cdots b_{i_k} \quad \text{for } k = 1, 2, \ldots, d.$$

For example, if $d = 4$ then the 201 terms in $b_1, b_2, b_3, b_4$ translate into only 16 terms in $e_1, e_2, e_3, e_4$:

$$\begin{aligned}H_A(\mathbf{b}) = {} & 432e_1^4e_4^2 - 432e_1^3e_2e_3e_4 + 128e_1^3e_3^3 + 108e_1^2e_2^3e_4 - 36e_1^2e_2^2e_3^2 - 2160e_1^2e_2e_4^2 \\ & + 1800e_1e_2^2e_3e_4 + 120e_1^2e_3^2e_4 - 540e_1e_2e_3^3 - 405e_2^4e_4 + 135e_2^3e_3^2 \\ & + 2400e_1e_3e_4^2 + 1800e_2^2e_4^2 - 2700e_2e_3^2e_4 + 675e_3^4 - 2000e_4^3.\end{aligned}$$

As an application of our theory, we are now able to answer two questions from the literature. The first deals with the discriminant of the derivative of a univariate polynomial. According to Alexandersson and Shapiro [1, Theorem 1.4], Frank Sottile and Eugene Mukhin formulated this conjecture at the AIM meeting "Algebraic systems with only real solutions" in October 2010.

**Corollary 14.** *The discriminant of the derivative of a univariate polynomial $f(t)$ of degree $n$ is a sum of squares of polynomials in the differences of the roots of $f(t)$.*

**Proof.** Let $\mathcal{D}_n = \mathrm{discr}_t\big(f'(t)\big)$. We shall write $\mathcal{D}_n$ as a specialization of the entropic discriminant and use the sum of squares decomposition given in Theorem 10. Consider the univariate polynomial $f(t) = \prod_{i=1}^n (t - a_i)$. Notice that $x_i = t - a_i$ provides a parametrization for the one-dimensional affine space $\{A\mathbf{x} = \mathbf{b}\}$, where we take $A$ as in (20) and $b_i = a_n - a_i$ for $i = 1, \ldots, n-1$. We plug this parametrization into the polynomial (22) that defines $\mathcal{L}^{-1}$. This yields the derivative $f'(t) = \sum_{j=1}^n \prod_{i \neq j}(t - a_i)$. Thus $f'(t)$ equals the polynomial $p_{\mathbf{b}}(t)$ of (24) whose discriminant (with respect to $t$) equals $H_A(\mathbf{b})$. We conclude that $\mathcal{D}_n$ equals the entropic discriminant $H_A((a_n - a_i)_{i\in[n-1]})$. Using Theorem 10, we conclude that $\mathcal{D}_n$ is a sum of squares in $\mathbb{Q}[\mathbf{b}] = \mathbb{Q}[(a_n - a_i) : i \in [n-1]]$. $\square$

Our techniques can also be applied to answer a question that was left open in [25, Section 7.5]. Namely, we conclude this section by proving Conjecture 7.9 of [25].

**Corollary 15.** *There exist three real symmetric $d \times d$-matrices $C_0, C_1, C_2$ such that all $\binom{d+1}{3}$ pairs of complex numbers $(x, y)$ for which $C_0 + xC_1 + yC_2$ has a critical double eigenvalue are real.*

**Proof.** Consider the symmetric matrix $\hat{X} = M^{-1}XM^{-T}$ with $X = \mathrm{diag}(b_1, \ldots, b_d)$ in the proof of Proposition 12. Its entries are linear forms in $\mathbf{b} = (b_1, b_2, \ldots, b_d)$. We replace the unknowns $b_i$ by generic real affine-linear forms in two variables $x$ and $y$, say $b_i = w_i + u_i x + v_i y$ for $i = 1, \ldots, d$. The symmetric matrix resulting from this substitution is a net of real symmetric $d \times d$-matrices:

$$\hat{X} = C_0 + x\,C_1 + y\,C_2.$$

The real values of $(x, y)$ for which this matrix has a critical double eigenvalue corresponds to the intersections of this affine plane with the real variety of $H_A(\mathbf{b})$, with $A$ given in (20).

We claim that the real radical of the entropic discriminant is the codimension-2 ideal

$$\sqrt[\mathbb{R}]{\langle H_A(\mathbf{b})\rangle} = \bigcap_{1\le i<j\le d} \langle b_i, b_j\rangle \cap \bigcap_{1\le i<j<k\le d} \langle b_i - b_j, b_j - b_k\rangle. \tag{27}$$

This identity follows from the geometric description of $H_A(\mathbf{b})$ in terms of colliding analytic centers. Indeed, the hyperplane arrangement defined by $\{x_i = 0\}$ in $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$ consists of

$n = d + 1$ points on a line. They form $d$ bounded segments. The analytic centers of two segments collide if and only if three of the $d + 1$ points coincide. There are such $\binom{d+1}{3}$ triples, each imposing a condition of codimension 2. They are expressed by the $\binom{d+1}{3}$ prime ideals in the intersection (27).

Since the real variety of $H_A(\mathbf{b})$ is a union of $\binom{d+1}{3}$ real linear spaces, each of its intersection points with the plane $\mathbf{b} = \mathbf{w} + \mathbf{u}x + \mathbf{v}y$ is also defined over the reals. Therefore all $\binom{d+1}{3}$ symmetric matrices with a critical double eigenvalue in the net $C_0 + xC_1 + yC_2$ have real entries. □

**Remark 16.** In general, the issue of determining $\sqrt[\mathbb{R}]{\langle H_A(\mathbf{b})\rangle}$ is very subtle. The validity of the identity (27) above rests formally on the prime decomposition of the real radical ideal $\sqrt[\mathbb{R}]{\langle H_A(\mathbf{b})\rangle}$ described in Corollary 37. See also Example 32 for the particular matrix $A$ in (20).

## 4. Matroids and graphs

In this section we discuss the notions from matroid theory which are needed for the statement and proof of Theorem 2. We also discuss various matroids arising from graphs, including those representing the Hillar–Wibisono model (6). Matroid theory is a classical subject in combinatorics with many (axiomatic) paths leading to it. For us, matroids come in the form of matrices and hence we take the concrete approach via realizable matroids. For more on this subject see [19,24].

Our given matrix $A = (A_1, A_2, \dots, A_n) \in \mathbb{R}^{d\times n}$ is identified with an ordered collection of $n$ vectors that span $d$-space. The corresponding matroid $M = M(A)$ records all linear dependencies among these vectors. A subset $I \subseteq [n] = \{1, 2, \dots, n\}$ is called *(in)dependent* whenever $A_I = (A_i : i \in I)$ is linearly (in)dependent. The *rank* $\mathrm{rk}(I)$ of $I$ is the rank of $A_I$. The rank of the matroid $M$ is the rank of $A$. A *circuit* is an inclusion-minimal dependent subset, and $I$ is independent if it does not contain a circuit. A subset $F \subseteq [n]$ is a *flat* if $F$ equals $\{i \in [n] : A_i \in \mathrm{span}(A_F)\}$, that is, if $F$ precisely indexes a collection of vectors contained in some linear subspace. Equivalently, $F$ is a flat if and only if it meets every circuit in at least 2 elements or not at all. Flats will play an important role in Sections 5 and 6. The *lattice of flats* $L(M)$ is the collection of all flats, ordered by inclusion, with minimal element $\hat{0} = \{i : A_i = 0\}$ and maximal element $\hat{1} = [n]$. The lattice of flats represents combinatorial information about the containment relations of the various subspaces spanned by subsets of columns of $A$. It is one of the central objects in the enumerative theory of matroids.

A different but equivalent perspective on $M(A)$ and $L(M)$ is by means of the hyperplane arrangements alluded to in Section 2. The $n$ columns of $A$ are normal to $n$ linear (not necessarily distinct) hyperplanes $h_1, h_2, \dots, h_n \subseteq \mathbb{R}^d$. In this context, a subset $I \subseteq [n]$ is independent if and only if the intersection of $\{h_i : i \in I\}$ has codimension $|I|$. The collection of linear subspaces obtained by intersections of these hyperplanes is isomorphic to the lattice of flats $L(M)$ when partially ordered by *reverse* inclusion. For a generic vector $\mathbf{b}$, the matroid associated to $(A, \mathbf{b})$ is called the *free extension* of $M(A)$. The hyperplane arrangement corresponding to the free extension is obtained by adding a hyperplane such that intersections with flats are transverse.

At the beginning of Section 1, we considered a different arrangement of affine hyperplanes associated to $A$. To relate this to $M(A)$, observe that the $n$ coordinate hyperplanes $\{x_i = 0\}$ in $\mathbb{R}^n$ induce a hyperplane arrangement in $\ker(A) \cong \mathbb{R}^{n-d}$. This arrangement corresponds to the *dual* matroid to $A$, namely $M(B)$, where $B$ is an $(n - d) \times n$ matrix whose rows form a basis for ker$(A)$. The hyperplane arrangement $\{g_1, g_2, \ldots, g_n\}$ in $\mathbb{R}^{n-d}$ associated to the columns of $B$ is linearly isomorphic to the arrangement of the $n$ coordinate hyperplanes in ker $A$. Dually, the hyperplane arrangement $\{h_1, h_2, \ldots, h_n\}$ given by the columns of $A$ yields a linearly isomorphic representation of the arrangement of coordinate hyperplanes inside ker$(B)$.

The matroid dual to the free extension by $\mathbf{b}$ is called the *free co-extension*, which corresponds to the linear arrangement of the $n+1$ coordinate hyperplanes in ker$((A, \mathbf{b}))$. Here we distinguish the last hyperplane $g_\infty$ as the hyperplane "at infinity". Restricting the arrangement to $g_\infty$ recovers the original arrangement in ker$(A)$. The arrangement that will be central to our cause is the arrangement of $n$ affine hyperplanes given by the intersection of coordinate hyperplanes in $\{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$, for generic $\mathbf{b}$. This is the restriction of the $g_i$ to some parallel displacement $g_\infty + t$ (for some generic $t \notin g_\infty$). Alternatively, this is the affine arrangement $\{\hat{g}_i = g_i + t_i \subset \mathbb{R}^{n-d} : i = 1, \ldots, n\}$ where the displacements $t_i \in \mathbb{R}^{n-d}$ are generic. Thus, the arrangement of coordinate hyperplanes in $\{A\mathbf{x} = \mathbf{b}\}$ can be obtained by a generic, parallel perturbation of the hyperplanes $g_1, g_2, \ldots, g_n$.

Associated to $L = L(M)$ is its *Möbius function* $\mu_L : L \times L \to \mathbb{Z}$, which is defined by $\mu_L(F, F) = 1$,

$$\mu_L(F, H) = -\sum_{F \subseteq G \subset H} \mu_L(F, G)$$

if $F \subseteq H$, and $\mu_L(F, H) = 0$ otherwise. The *characteristic polynomial* of $M$ is defined by

$$\chi_M(t) = \sum_{F \in L(M)} \mu_{L(M)}(\hat{0}, F)\, t^{rk(M)-rk(F)}.$$

The *(unsigned) Möbius invariant* of $M$, or of the matrix $A$, is the positive integer

$$\mu(A) = \mu(M(A)) = |\mu_L(\hat{0}, \hat{1})| = (-1)^d \chi_M(0). \tag{28}$$

Here the last equality comes from Rota's Sign Theorem.

Evaluations of the characteristic polynomial have nice combinatorial interpretations in terms of hyperplane arrangements [10,24]. The Möbius invariant $\mu(A)$ equals the number of bounded regions of the restriction of the $n$ coordinate hyperplanes to $\{A\mathbf{x} = \mathbf{b}\}$, for generic $\mathbf{b}$. This fact played an important role in [3, Section 3]. The proof is a straightforward deletion–contraction argument, using that $\mu(A)$ and the number of bounded regions in $\{A\mathbf{x} = \mathbf{b}\}$ adhere to the same recurrence relations. This number is related to the *beta invariant* of the free extension $(A, \mathbf{b})$,

$$\beta(A, \mathbf{b}) := (-1)^{\mathrm{rk}(A,\mathbf{b})} \sum_{I \subseteq [n+1]} (-1)^{|I|} \mathrm{rk}(A, \mathbf{b})_I = (-1)^{\mathrm{rk}(A)} \sum_{F \in L} \mu_L(\hat{0}, F) = \mu(A)$$

where the middle equality is taken from [28, Proposition 7.3.1]. The geometric content of this statement was proved by Greene and Zaslavsky [10, Eq. (3.1)] and, in a more algebro-geometric context, in [4]. The beta invariant is unchanged under duality of matroids and thus $\beta(A, \mathbf{b}) = \mu(A)$ is the number of bounded regions for the coordinate arrangement in $\{A\mathbf{x} = \mathbf{b}\}$ when $\mathbf{b}$ is generic.

The proof of the following observation illustrates the typical line of arguments in matroid theory.

**Proposition 17.** *The Möbius invariant $\mu(A)$ equals 1 if and only if the matrix $A$ is basic (defined in Corollary 5) if and only if its geometric lattice $L(M)$ is the Boolean lattice of all subsets of $[d]$.*

**Proof.** An equivalent statement appears in [4, Corollary 4 (2) (b1)]. For completeness, we here include a combinatorial proof. By the definition of the lattice of flats, we can assume that $A$ has no zero columns and that no two columns are proportional. Hence, the matrix $A$ is basic if and only if $M(A)$ is isomorphic to the uniform matroid $U_{d,d}$ whose lattice of flats is the Boolean lattice of all subsets of $[d]$. The Möbius invariant of the matroid $U_{d,d}$ is $\mu(A) = 1$; see Example 18. Conversely, if $A$ is non-basic, there is a column $e \in [n]$ that is not an *isthmus*, that is, not contained in every basis. For such an element $e$, the Möbius invariant satisfies the deletion–contraction identity

$$\mu(A) = \mu(A \setminus e) + \mu(A/e).$$

By Rota's Sign Theorem, the Möbius invariant is always a positive integer and hence $\mu(A) \geq 2$. $\square$

We have now defined the combinatorial ingredients for the degree (4) of the entropic discriminant. With this in place, we derive the value of that degree for generic matrices $A$ stated in Theorem 2.

**Example 18** (*Uniform Matroids*)**.** A generic $d \times n$-matrix $A$ with $d \leq n$ gives rise to the *uniform matroid* $M = U_{d,n}$ in which every subset of cardinality $\leq d$ is independent. The corresponding lattice of flats is a truncated Boolean lattice in which a subset $F \subseteq [n]$ is a flat if and only if $|F| < d$ or $F = [n]$. The Möbius function on the Boolean lattice is $\mu(F, G) = (-1)^{|G \setminus F|}$ for $F \subseteq G$. Hence

$$\chi_{U_{d,n}}(t) = t^d - nt^{d-1} + \cdots + (-1)^{d-1} \binom{n}{d-1} t + (-1)^d \binom{n-1}{d-1}.$$

Note that $t = 1$ is always a zero of the characteristic polynomial. The number of solutions of Eqs. (2) for generic $A$ equals $\mu(A) = \binom{n-1}{d-1}$. The degree (4) of the entropic discriminant equals

$$2(-1)^d \cdot (d\chi_{U_{d,n}}(0) + \chi'_{U_{d,n}}(0)) = 2\left[d\binom{n-1}{d-1} - \binom{n}{d-1}\right] = 2(n-d)\binom{n-1}{d-2}.$$

As we will see in Proposition 33, this quantity is an upper bound for fixed $n$ and $d$. $\square$

*Graphical matroids* are an important class of examples. Let $G$ be a graph on $d$ nodes with $n$ edges and $c$ connected components. For an arbitrary but fixed orientation of the edges, let $A_G$ be the $d \times n$ *incidence matrix* of node-edge pairs, with entries $+1, -1, 0$ if the node is incoming, out-going, or non-incident for the edge. Reorienting an edge of $G$ results in scaling the corresponding column of $A_G$ by $-1$ and hence leaves the matroid $M_G = M(A_G)$ invariant. Note that $A_G$ has rank $d-c$ and a matrix representation of full rank can be obtained by selecting a node in every connected component of $G$ and deleting the corresponding rows. The matroid concepts above have natural interpretations in graph-theoretic terms: circuits correspond to cycles and independent sets to forests. The characteristic polynomial $\chi_G(t) = \chi_{M_G}(t)$ in this context is also called the *tension polynomial* and $t^c \chi_G(t)$ counts the number of proper $t$-colorings of $G$ where $t \in \mathbb{Z}_+$. Returning to the setting of Section 2, the hyperplane arrangement given by the columns of $A_G$ is the graphic arrangement associated with $G$, which has the defining polynomial

$$f_G(\mathbf{z}) = \prod_{(i,j) \in E(G)} (z_i - z_j).$$

The entropic discriminant $H_G(\mathbf{b})$ is the equation of the branch locus of the gradient map $\nabla_{f_G}$. As $A_G$ does not have full rank, we assume $z_i = 0$ for the rows $i$ that were deleted when passing from $A_G$ to a rank $d-c$ matrix with $d-c$ rows. The gradient map $\nabla_{f_G}$ is discussed in [12, Remark 8].

**Example 19** (*Cycles*)**.** Let $G = C_{d+1}$ be the cycle with $n = d+1$ edges. Every collection of $d$ or fewer edges is independent and $M_G$ has a unique circuit. The truncated matrix $A_{C_{d+1}}$ has corank 1 and $M_G$ is the uniform matroid $U_{d,d+1}$. The reciprocal plane $\mathcal{L}_{A_G}^{-1}$ is a hypersurface of degree $d$, and the entropic discriminant $H_{C_{d+1}}(\mathbf{b})$ is the polynomial of degree $d(d-1)$ seen in Section 3. $\square$

**Example 20** (*Complete Graphs*)**.** As the name says, the complete graph $G = K_{d+1}$ has all possible edges on $d+1$ nodes. The characteristic polynomial is the chromatic polynomial divided by $t$:

$$\chi_{K_{d+1}}(t) = (t-1)(t-2)\cdots(t-d).$$

The reciprocal plane $\mathcal{L}_{K_{d+1}}^{-1}$ is a projective variety of degree $(-1)^d \chi_{K_{d+1}}(0) = d!$. We find that

$$\deg\ H_{K_{d+1}}(\mathbf{b}) = 2 \cdot \left( d - 1 - \frac{1}{2} - \frac{1}{3} - \frac{1}{4} \cdots - \frac{1}{d} \right) \cdot d\,!$$

is the value of the matroid invariant (4) for the incidence matrix $A_{K_n}$ of the complete graph $K_n$.

For example, for $d = 3$ we get the complete graph on 4 nodes, with node-edge incidence matrix

$$A_{K_4} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{pmatrix}.$$

The reciprocal plane $\mathcal{L}_{K_4}^{-1}$ is a surface of degree 6 in $\mathbb{P}^5$. Its homogeneous prime ideal is generated by four quadrics, one for each of the 3-cycles in $K_4$. The entropic discriminant $H_{A_{K_4}}$ defines a curve in the projective plane $\mathbb{P}^2$. That curve has degree 14 and it has precisely six real points. $\square$

The matroids associated with the retina equations (6) are different from the matroids $M_G$ above. Their matroids correspond to *all-negative graphs* in Zaslavsky's theory of signed graphs [31]. Here, an all-negative graph $-G$ is an ordinary graph with all edges marked by $-1$. The incidence matrix $A_{-G}$ of $-G$ has entries in $\{0, 1\}$ where a 1 signifies an incident node-edge pair. The corresponding matroid $M(-G) = M(A_{-G})$ is the *unoriented cycle matroid*. The matroid-theoretic notions for $M(-G)$ translate to (signed) graph concepts but the transitions are more involved. For all-negative graphs, the circuits correspond to *even primitive walks*, that is, even cycles or pairs of odd cycles connected by a simple path (of length possibly 0); cf. [31, Corollary 7D.3(e)]. For the state of the art on algebraic properties of the circuits of $A_{-G}$ see the recent work of Tatakis and Thoma [26]. Evaluations of the characteristic polynomial have interpretations in terms of *signed colorings* [30].

For example, the all-negative complete graph $-K_4$ on four nodes has the incidence matrix

$$A_{-K_4} = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}. \tag{29}$$

Note that this matrix has rank 4. Its matroid has the characteristic polynomial

$$\chi_{-K_4}(t) = t^4 - 6\,t^3 + 15\,t^2 - 17\,t + 7.$$

The characteristic polynomials for the all-negative complete graphs on any number of nodes were computed by Zaslavsky [30, Eq. (5.8)]. We presented his formula in the introduction in (8). An equivalent formula in terms of generating functions due to Stanley [24, Example 5.25] was shown in (9).

For the matrix (29), the reciprocal variety $\mathcal{L}_{-K_4}^{-1}$ is defined by the three cubic equations

$$\begin{aligned}
&x_{12}x_{13}x_{24} - x_{12}x_{13}x_{34} - x_{12}x_{24}x_{34} + x_{13}x_{24}x_{34} = 0,\\
&x_{13}x_{14}x_{23} - x_{13}x_{14}x_{24} - x_{13}x_{23}x_{24} + x_{14}x_{23}x_{24} = 0,\\
&x_{12}x_{14}x_{23} - x_{12}x_{14}x_{34} - x_{12}x_{23}x_{34} + x_{14}x_{23}x_{34} = 0.
\end{aligned}$$

The task in (6) is to solve these cubic equations together with linear equations $A_{-K_4} \cdot \mathbf{x} = \mathbf{b}$ for the six unknowns $x_{12}, \dots, x_{34}$. The number of solutions to this system is $\mu(M_{-K_4}) = 7$, and all seven solutions are real when the $b_i$ are real. One of the solutions has only positive coordinates if and only if the column vector $(b_1, b_2, b_3, b_4)$ of parameters lies in the convex polyhedral cone spanned by the columns of $A_{-K_4}$. The entropic discriminant $H_{-K_4}(b_1, b_2, b_3, b_4)$ characterizes parameter values for which the number of solutions is less than 7. It is a surface in $\mathbb{P}^3$ of degree $2(4 \cdot 7 - 17) = 22$. The Möbius invariant $\mu(M_{-K_d})$ and the degree of $H_{-K_d}$ for larger values of $d$ are displayed in (10).

We close this section with the remark that the study of characteristic polynomials of matroids is an active area of current research in combinatorics. The coefficients of $\chi(t)$ have interpretations as face numbers of broken circuit complexes and form a log-concave sequence. This log-concavity was a longstanding conjecture recently resolved by Huh [12] for graphs and in its full generality by Huh–Katz [13]. Their methods of proof are based on the geometry of reciprocal planes, our topic in the next section. Specifically, a key player in [13] is the tropicalization of the graph of $\mathcal{L} \dashrightarrow \mathcal{L}^{-1}$.

## 5. Geometry of reciprocal planes

Entropic discriminants arise as branch loci from projecting reciprocal planes. This was already hinted at in the proof of Proposition 4. We shall make this precise in Section 6, where it will be our main ingredient in the proof of Theorem 2. In this section we build up to this proof by deriving some results on reciprocal planes. We believe that these results are of interest in their own right.

We fix a $d \times n$-matrix $A$ of rank $d$ with no zero columns. Its rows span a $(d-1)$-dimensional subspace $\mathcal{L}$ in the projective space $\mathbb{P}^{n-1}$. Let $T$ denote the dense torus in $\mathbb{P}^{n-1}$, i.e. the complement of the $n$ coordinate hyperplanes $\{x_i = 0\}$. The *reciprocal plane* $\mathcal{L}^{-1}$ is the Zariski closure of the coordinate-wise inverse of $\mathcal{L} \cap T$, as in (5). It is an irreducible projective variety of dimension $d-1$. The inversion map from $\mathcal{L}$ to $\mathcal{L}^{-1}$ is birational and it is an isomorphism on $\mathcal{L} \cap T$. The coordinate ring of the reciprocal plane $\mathbb{C}[\mathbf{x}]/I(\mathcal{L}^{-1})$ is isomorphic to the Orlik–Terao algebra, studied in [22].

Proudfoot and Speyer [20] showed that $\mathcal{L}^{-1}$ is stratified by the flats of the matroid $M(A)$. Recall that $J \subseteq [n]$ is a flat of $M(A)$ if and only if $\mathrm{rk}(A_J) < \mathrm{rk}(A_{J'})$ for all $J' \supsetneq J$. Here $A_J$ denotes the column-induced submatrix of $A$. For a flat $J \subseteq [n]$, the corresponding stratum

$\mathcal{L}^{-1} \cap \mathbb{P}^J = \{p \in \mathcal{L}^{-1} : \operatorname{supp}(p) \subseteq J\}$ is isomorphic to $\mathcal{L}_J^{-1}$, the reciprocal plane associated to the restriction $A_J$. We shall investigate these boundary strata and the singular locus $\operatorname{Sing}(\mathcal{L}^{-1})$ of $\mathcal{L}^{-1}$.

We can identify each circuit $C$ of the matroid $M(A)$ with a vector $v \in \mathbb{R}^n$ in the kernel of $A$ with support $\operatorname{supp}(v) = C$. Let $\mathcal{C}(A) \subseteq \mathbb{R}^n$ denote the set of representative vectors for all circuits of $M(A)$. To each $v \in \mathcal{C}(A)$ we associate a polynomial

$$h_v(\mathbf{x}) = \sum_{i \in \operatorname{supp}(v)} v_i \prod_{j \neq i} x_j = \mathbf{x}^{\operatorname{supp}(v)} \sum_{i \in \operatorname{supp}(v)} \frac{v_i}{x_i}. \tag{30}$$

These *circuit polynomials* cut out the variety $\mathcal{L}^{-1}$. In fact, Proudfoot and Speyer proved the much stronger result that $\{h_v : v \in \mathcal{C}(A)\}$ is a *universal Gröbner basis* for the prime ideal of $\mathcal{L}^{-1}$.

As the set of all circuits is typically rather large, one might be interested in a smaller set of polynomials to cut out $\mathcal{L}^{-1}$. The following characterizes subsets of the set of circuit polynomials that cut out $\mathcal{L}^{-1}$ set-theoretically. As we saw above, the boundary of $\mathcal{L}^{-1} \setminus T$ in $\mathcal{L}^{-1}$ is described by flats of $M(A)$. Recall that $J \subseteq [n]$ is a flat if and only if $|J^c \cap \operatorname{supp}(v)| \neq 1$ for every circuit $v \in \mathcal{C}(A)$. We say that a non-flat $J \subset [n]$ is *exposed* by a circuit $v \in \mathcal{C}(A)$ if $|J^c \cap \operatorname{supp}(v)| = 1$.

**Proposition 21.** *Let $\mathcal{B} \subseteq \mathcal{C}(A)$ be a subset of the set of circuits. The corresponding set of circuit polynomials $\{h_v : v \in \mathcal{B}\}$ cuts out $\mathcal{L}^{-1}$ set-theoretically if and only if $\mathcal{B}$ exposes every non-flat.*

**Proof.** Suppose that $J$ is a non-flat that is not exposed by any $v \in \mathcal{B}$. Then, for each $v \in \mathcal{B}$, either $|J^c \cap \operatorname{supp}(v)| \geq 2$, in which case $h_v$ is identically zero on $\mathbb{P}^J$, or $J^c \cap \operatorname{supp}(v) = \emptyset$, in which case $v$ is a circuit of $A_J$ and $h_v$ vanishes on $\mathcal{L}_J^{-1}$. This shows that the subvariety of $\mathbb{P}^{n-1}$ cut out by $\{h_v : v \in \mathcal{B}\}$ contains $\mathcal{L}_J^{-1}$ and hence is strictly larger than $\mathcal{L}^{-1}$.

Conversely, assume that $\mathcal{B}$ exposes every non-flat. Let $p$ be any zero of $\{h_v : v \in \mathcal{B}\}$, and let $J = \operatorname{supp}(p)$. Suppose that $J$ is a non-flat of $M(A)$. Then there exists $v \in \mathcal{B}$ that exposes $J$. This means that exactly one of the terms of $h_v$ is non-zero at $p$, and hence $h_v(p) \neq 0$. We conclude that $J$ is a flat of $M(A)$. Since $\mathcal{L}_J^{-1}$ is a boundary stratum of $\mathcal{L}^{-1}$, it is sufficient to prove that $p_J^{-1} \in \operatorname{rowspan} A_J$. For this, we shall prove that the kernel of $A_J$ is spanned by $\{v_J : v \in \mathcal{B} \text{ and } \operatorname{supp}(v) \subseteq J\}$. Let $J_0 \subseteq J$ be a basis of $M(A_J)$. If $J_0$ is not a flat, then there is a circuit $v_1 \in \mathcal{B}$ supported on $J$ such that $\operatorname{supp}(v_1) \setminus J_0 = \{j_1\}$. Set $J_1 = J_0 \cup \{j_1\}$ and repeat the procedure. This process terminates after $k = |J| - \operatorname{rk}(A_J) = \dim \ker A_J$ many steps. The matrix of the resulting circuits $v_1, \ldots, v_k$ is lower-triangular and hence gives a basis for $\ker(A_J)$. $\square$

This previous result highlights the connection of our study to tropical geometry.

**Remark 22.** Combining Proposition 21 with the results of [29], we infer that a collection of circuits cuts out the reciprocal plane $\mathcal{L}^{-1}$ set-theoretically if and only if it constitutes a *tropical basis* for the tropicalization of the linear space $\mathcal{L}$. Yu and Yuster [29, Section 2.2] showed that different inclusion-minimal tropical bases for $\mathcal{L}$ need not have the same cardinality. Specifically, the uniform matroid $U_{2,5}$ has inclusion-minimal tropical bases of sizes 5 and 6. From this we can infer that Proposition 21 holds only set-theoretically and not in the ideal-theoretic or scheme-theoretic sense.

**Example 23.** If the matroid $M(A)$ is uniform, then the prime ideal of the reciprocal plane $\mathcal{L}^{-1}$ is minimally generated by $\binom{n-1}{d}$ polynomials of degree $d$. This can be seen as follows. The initial ideal of $\mathcal{L}^{-1}$ with respect to the reverse lexicographic term order is generated by the square-free monomials representing broken circuits. These are $x_{i_1}x_{i_2}\cdots x_{i_d}$ where $1 \le i_1 < \cdots < i_d \le n-1$, so their number is $\binom{n-1}{d}$. By [29, Lemma 5], the basic circuits obtained by adding the last element $n$ form an inclusion-minimal tropical basis for $\mathcal{L}^{-1}$. Hence, by Remark 22, the corresponding $h_v$ minimally cut out $\mathcal{L}^{-1}$. It follows that they form a minimal generating set for the ideal of $\mathcal{L}^{-1}$. □

We now come to the main result in this section, namely, the characterization of the tangent cone of the reciprocal plane $\mathcal{L}^{-1}$ at any point. For the sake of convenience, we here identify the $(d-1)$-dimensional projective variety $\mathcal{L}^{-1}$ with the corresponding $d$-dimensional affine variety in $\mathbb{C}^n$.

The *tangent cone* $\mathrm{TC}_p X$ of a variety $X \subset \mathbb{C}^n$ at a point $p$ is a scheme that describes the local behavior of $X$ around $p$. For a polynomial $f \in \mathbb{C}[\mathbf{x}]$, the *initial form* $\mathrm{in}_{-\mathbf{1}}(f)$ is the non-zero homogeneous component of $f$ of minimal degree. The tangent cone $\mathrm{TC}_p X$ is defined by the ideal

$$I(\mathrm{TC}_p X) = \langle \mathrm{in}_{-\mathbf{1}}(f(\mathbf{x}+p)) : f \in I(X)\rangle. \tag{31}$$

The following result shows that the tangent cone of $\mathcal{L}^{-1}$ at any point is reduced and irreducible. Here we use $\mathcal{L}_{A/J}$ to denote the $(d - \mathrm{rank}(J))$-dimensional linear space $\mathcal{L}/\mathcal{L}_J$ in $\mathbb{C}^n/\mathcal{L}_J \simeq \mathbb{C}^{n-|J|}$.

**Theorem 24.** *Let $A \in \mathbb{R}^{d\times n}$ be a matrix of full row rank $d$ and let $\mathcal{L}^{-1}$ be its reciprocal plane in $\mathbb{C}^n$. For any point $p \in \mathcal{L}^{-1}$ with support $J$, the tangent cone is isomorphic to the direct product*

$$\mathrm{TC}_p\mathcal{L}^{-1} \cong \mathcal{L}_J \times \mathcal{L}^{-1}_{A/J}, \tag{32}$$

*where "$\cong$" denotes the equality of affine schemes after a linear transformation in $\mathbb{C}^n$.*

**Proof.** We inspect the initial forms of the circuit polynomials that define $\mathcal{L}^{-1}$. Let $v \in \mathcal{C}(A)$ be a circuit with support $C = \mathrm{supp}(v)$ and circuit polynomial $h_v(\mathbf{x})$ as in (30). First suppose that $C \not\subseteq J$. We write $v = v' + v''$ where $\mathrm{supp}(v') = C \cap J$ and $\mathrm{supp}(v'') = C \setminus J$. Then $v''$ is a circuit of the matroid $M(A/J)$ obtained from $M(A)$ by contraction at $J$. The following identity holds:

$$h_v(\mathbf{x}+p) = \mathbf{x}^{C\setminus J}\cdot h_{v'}(\mathbf{x}+p) + (\mathbf{x}+p)^{C\cap J}\cdot h_{v''}(\mathbf{x}).$$

Every term of $\mathbf{x}^{C\setminus J}h_{v'}(\mathbf{x}+p)$ has degree at least $|C\setminus J|$ while

$$\mathrm{in}_{-\mathbf{1}}\big((\mathbf{x}+p)^{C\cap J}h_{v''}(\mathbf{x})\big) = p^{C\cap J}\cdot h_{v''}(\mathbf{x})$$

has degree $|C\setminus J|-1$. This means that $h_{v''}(\mathbf{x})$ is the initial form of $h_v(\mathbf{x}+p)$. As every circuit $w$ of the contraction $M(A/J)$ is the restriction $v''$ of some circuit $v$ of $M(A)$, we conclude that the tangent cone ideal at $p$ contains the prime ideal $\langle h_w(\mathbf{x}) : w \in \mathcal{C}(A/J)\rangle$ that defines $\mathcal{L}^{-1}_{A/J}$.

Next suppose that $C \subseteq J$. Then $p$ is a regular point on the hypersurface $\{h_v = 0\}$, and the initial form $\mathrm{in}_{-\mathbf{1}}(h_v(\mathbf{x}+p))$ is the differential $D_p h_v$. The differential of $h_v$ at the point $p$ is

$$D_p h_v(\mathbf{x}) = \sum_{i=1}^{n}\frac{\partial h_v}{\partial x_i}(p)\,x_i = p^C\sum_{i\in C}\frac{x_i}{p_i}\left(\sum_{j\in C\setminus i}\frac{v_j}{p_j}\right) = -\sum_{i\in C}\frac{v_i}{p_i^2}x_i.$$

The second equality holds because $p^{-1}$ lies in rowspan$(A_J) \cap (\mathbb{C}^*)^J$, and the third equality follows from the fact that $-v_i/p_i = \sum_{j \neq i} v_j/p_j$, since $v$ is a circuit for $A_J$. Thus, $D_p h_v$ vanishes on the rowspan of $A_J \mathrm{diag}(p_J)^2$, denoted by $\mathcal{L}_J(p)$, and all circuits vanishing on this row span arise this way.

We have shown that the prime ideal of the irreducible variety $\mathcal{L}_J(p) \times \mathcal{L}^{-1}_{A/J}$ is contained in the ideal of the tangent cone of $\mathcal{L}^{-1}$ at $p$. Since both ideals have the same height, and the former is prime, it follows that they are equal. This proves the equality of schemes that was claimed. □

A closer inspection of the proof reveals that the initial forms of $h_v(\mathbf{x} + p)$ for $v \in \mathcal{C}(A)$ furnish a universal Gröbner basis for the tangent cone of $\mathcal{L}^{-1}$ at $p$. In particular, we obtain a simple description of the tangent space of $\mathcal{L}^{-1}$ at a point $p$ by taking those initial forms that are linear.

**Corollary 25.** *For a point $p \in \mathcal{L}^{-1}$ with support $J$, the tangent space is orthogonal to the space spanned by the circuits of the $d \times |J|$-matrix $A_J \mathrm{diag}(p_J)^2$ and the circuits of $A/J$ of size 2.*

**Proof.** The tangent space is cut out by the linear forms in the ideal of the tangent cone. From the initial forms in the proof of Theorem 24, we see that $\mathrm{in}_{-\mathbf{1}}(h_v)$ is linear whenever $|\mathrm{supp}(v) \cap J^c| \leq 2$. If $C \subseteq J$, then $\mathrm{in}_{-\mathbf{1}}(h_v)$ corresponds to a circuit of $A_J \mathrm{diag}(p_J)^2$. Otherwise, the two elements of $C \setminus J$ are parallel in the contraction $A/J$ and the corresponding circuit polynomial is linear. □

This is closely related to [22, Theorem 2.3], which investigates the quadratic component of the ideal $I(\mathcal{L}^{-1})$. Our discussion shows that the dimension of the tangent space is constant on each stratum of $\mathcal{L}^{-1}$. We obtain the following characterization of the singular locus of the reciprocal plane $\mathcal{L}^{-1}$.

**Corollary 26.** *The singular locus of the reciprocal plane $\mathcal{L}^{-1}$ is pure of codimension 2. It is the union of all boundary strata $\mathcal{L}^{-1}_J$ such that the contraction $M(A/J)$ is a non-basic matroid.*

**Proof.** A point $p \in \mathcal{L}^{-1}$ is smooth if and only if the codimension of the tangent space equals $\mathrm{codim}(\mathcal{L}^{-1}) = n - d$. The description of the tangent space in terms of the matroids $M(A_J)$ and $M(A/J)$ in Corollary 25 shows that its codimension is $|J| - \mathrm{rk} A_J + \mathrm{Par}(A/J)$ where $\mathrm{Par}(A/J)$ is the dimension of the space of 2-circuits of $A/J$. Suppose $M(A/J)$ has $r$ distinct 1-flats (or lines), and let $\lambda_1, \ldots, \lambda_r$ be the sizes of these parallelism classes. The circuits of each parallelism class span a linear space of dimension $\lambda_i - 1$. As these circuits are disjoint, we have $\mathrm{Par}(A/J) = \sum_{i=1}^{r} (\lambda_i - 1) = |J^c| - r$. The number of parallelism classes of $M(A/J)$ is at least $\mathrm{rk}(A/J)$. Thus the codimension of the tangent space is $\leq n - d$, and equality holds if and only if $M(A/J)$ is basic (cf. Proposition 17).

Finally, to see that the singular locus is pure of codimension 2, we note that if $M$ is any non-basic matroid of rank $r \geq 3$, then there is an element $e$ such that $M/e$ is non-basic. To show this, we can assume that $M$ is non-basic on $r + 1$ elements, each representing a different line. If $M = M_1 \oplus M_2$ is not connected and $M_1$ is non-basic then any $e \in M_2$ will work. Otherwise, $M$ is a uniform matroid and the contraction is uniform of rank $r - 1 \geq 2$ on $r$ elements. By Example 18, the uniform matroid $U_{n,d}$ is non-basic if and only if $n > d > 1$. Therefore, if $J$ is a flat of $M(A)$ such that $M(A/J)$ is non-basic of rank $\geq 3$, then there is a flat $J' \supset J$ such that $M(A/J')$ is non-basic. □

## 6. Ramification locus

The entropic discriminant describes the locus of points $\mathbf{b} \in \mathbb{P}^{d-1}$ such that the zero-dimensional scheme defined by the constraints $\mathbf{x} \in \mathcal{L}^{-1}$ and $A\mathbf{x} = \mathbf{b}$ is not reduced. Equivalently, the entropic discriminant is the defining polynomial of the branch locus of the map $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$. We begin with the observation that this map has no base points and is hence a projective morphism.

**Lemma 27.** *The variety $\mathcal{L}^{-1}$ is disjoint from the center of the projection $A : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{d-1}$.*

**Proof.** Our claim states that $\mathcal{L}^{-1} \cap \ker(A) = \{0\}$ holds in $\mathbb{C}^n$. Let $p$ be a vector in $\mathcal{L}^{-1} \cap \ker(A)$ and $J = \mathrm{supp}(p)$. Then $p_J^{-1} = z\,A_J$ for some $z \in \mathbb{C}^d$, and $p \in \ker(A)$ implies $0 = z\,(A\,p) = (z\,A)\,p = \sum_{j\in J} p_j{}^{-1} p_j = |J|$. It follows that $J = \emptyset$ and $p = 0$. $\square$

We now focus on the *ramification locus* of the dominant projective morphism $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$. By definition, this is the Zariski closure of the set of regular points $p \in \mathcal{L}^{-1}$ for which

$$\mathcal{L} + \mathrm{rowspan}\,\mathrm{Jac}(\mathcal{L}^{-1})(p) \neq \mathbb{C}^n. \tag{33}$$

Here $\mathrm{Jac}(\mathcal{L}^{-1})$ is the Jacobian matrix of $\mathcal{L}^{-1}$, whose row vectors are $\nabla h_v(\mathbf{x})$ for $v \in \mathcal{C}(A)$, as in (30). This condition states that the intersection of $\mathcal{L}^{-1}$ and $\{\mathbf{x} : A\mathbf{x} = Ap\}$ is not transverse at $p$.

The *ramification scheme* $\mathcal{R}_A = \mathrm{Proj}(\mathbb{C}[\mathbf{x}]/J_A)$ is defined by the following ideal in $\mathbb{C}[x_1, \ldots, x_n]$:

$$J_A = \left( I(\mathcal{L}^{-1}) + \left\langle n \times n \text{ minors of } \begin{pmatrix} A \\ \mathrm{Jac}(\mathcal{L}^{-1}) \end{pmatrix} \right\rangle \right)$$
$$: \left\langle (n-d) \times (n-d) \text{ minors of Jac } (\mathcal{L}^{-1}) \right\rangle^{\infty}. \tag{34}$$

By the Zariski–Nagata Purity Theorem [17], the ramification locus is pure of codimension 1 in $\mathcal{L}^{-1}$. Hence the ramification scheme $\mathcal{R}_A$ is either empty or has codimension 1 in $\mathcal{L}^{-1}$. The former happens when $A$ is basic, and the latter happens when $A$ is non-basic. We prove in Section 7 that $\mathcal{R}_A$ contains the singular locus of $\mathcal{L}^{-1}$ and hence that the saturation step in (34) is redundant.

**Definition 28.** Let $A \in \mathbb{R}^{d\times n}$ be a non-basic matrix of rank $d$. The *ramification cycle* is the algebraic cycle of dimension $d-2$ in $\mathbb{P}^{n-1}$ defined by the ramification scheme $\mathcal{R}_A$. By Corollary 5, the push-forward of the ramification cycle under the morphism $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$ is a cycle of codimension 1. We define the *entropic discriminant* of $A$ to be the homogeneous polynomial $H_A(\mathbf{b})$ that represents this cycle in $\mathbb{P}^{d-1}$. It is unique up to multiplication by a non-zero constant.

The following example shows that the ramification cycle may not be reduced.

**Example 29.** Let $A$ be the matrix in Example 3. For $a \neq 0, 2, 3$, the prime ideal of $\mathcal{L}^{-1}$ equals

$$I(\mathcal{L}^{-1}) = \langle 2x_1x_2 - 3x_1x_3 + x_2x_3,\ 2x_1x_2 - ax_1x_4 + (a-2)x_2x_4,$$
$$3x_1x_3 - ax_1x_4 + (a-3)x_3x_4\rangle.$$

The ramification ideal $J_A$ is the sum of $I(\mathcal{L}^{-1})$ and the ideal of $4 \times 4$ minors of the matrix

$$\begin{pmatrix} A \\ \mathrm{Jac}(\mathcal{L}^{-1}) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 2 & 3 & a \\ 2x_2 - 3x_3 & 2x_1 + x_3 & -3x_1 + x_2 & 0 \\ 2x_2 - ax_4 & 2x_1 - (a-2)x_4 & 0 & -ax_1 + (a-2)x_2 \\ 3x_3 - ax_4 & 0 & 3x_1 + (a-3)x_4 & -ax_1 + (a-3)x_3 \\ 0 & x_3 - (a-2)x_4 & x_2 + (a-3)x_4 & -(a-2)x_2 + (a-3)x_3 \end{pmatrix}.$$

The ramification cycle is a zero-dimensional cycle of degree 4 in $\mathbb{P}^3$. For the special value $a = 6$, it is twice the reduced cycle of degree 2 defined by $\langle 2x_2 - 3x_3 + 6x_4, 2x_1 - x_3 + 4x_4, x_3^2 - 4x_3x_4 + 8x_4^2 \rangle$. The push-forward of this cycle under $\mathbb{P}^3 \dashrightarrow \mathbb{P}^1$ is defined by the binary quartic in Example 3. $\square$

Since the projection $A : \mathbb{P}^{n-1} \dashrightarrow \mathbb{P}^{d-1}$ has no base points on the subscheme $\mathcal{R}_A$ (by Lemma 27), the push-forward by $A$ preserves the degree of the ramification cycle. Thus, in order to establish the degree formula in Theorem 2, it suffices to show that the degree of $\mathcal{R}_A$ equals $2(-1)^d (d\chi(0) + \chi'(0))$. In order to compute its degree, we use a slightly different description of $\mathcal{R}_A$. Let $T$ denote the dense torus $\{x_1 x_2 \cdots x_n \neq 0\}$ in the projective space $\mathbb{P}^{n-1}$. Inside $T$, the variety $\mathcal{L}^{-1}$ is a complete intersection. Namely, it is defined by $B \cdot \mathbf{x}^{-1} = 0$, where $B = (B_1, \ldots, B_n)$ is a *Gale transform* for $A$, that is, an $(n-d) \times n$-matrix whose rows span the kernel of $A$. Consider the polynomial

$$g_A(\mathbf{x}) = \det \begin{pmatrix} & A & \\ B_1 x_1^{-2} & \cdots & B_n x_n^{-2} \end{pmatrix} \cdot \prod_{i=1}^{n} x_i^2 = \det \begin{pmatrix} A_1 x_1^2 & \cdots & A_n x_n^2 \\ & B & \end{pmatrix}.$$

The $n \times n$-matrix above now plays the same role as the Jacobian matrix did in (34). Thus the hypersurface defined by $g_A(\mathbf{x}) = 0$ inside $\mathcal{L}^{-1} \cap T$ is the restricted ramification locus $\mathcal{R}_A \cap T$.

If $g_A$ is zero at a point $p \in T$ then the intersection $\ker(A\mathrm{diag}(p)^2) \cap \ker(B)$ contains a non-zero vector. The kernel of $B$ is spanned by the rows of $A$, so the $d \times d$-matrix $A\mathrm{diag}(p)^2 A^T$ also drops rank. Hence $g_A(\mathbf{x})$ divides $\det(A\mathrm{diag}(\mathbf{x})^2 A^T)$. Both polynomials have the same degree $2d$, and hence they are equal (up to a scalar, which we ignore). Using the Cauchy–Binet Formula, this gives

$$g_A(\mathbf{x}) = \det(A\mathrm{diag}(\mathbf{x})^2 A^T) = \sum_{I \in \binom{[n]}{d}} \det(A_I)^2 \prod_{i \in I} x_i^2. \tag{35}$$

We next define similar polynomials that cut out $\mathcal{R}_A$ on the non-singular boundary strata of $\mathcal{L}^{-1}$. Let $J \subset [n]$ be any proper flat of rank $r$ in $M(A)$ and set $\mathbf{x}_J = (x_j : j \in J)$. Let $\hat{A}_J$ now denote any $r \times |J|$ submatrix of $A_J = (A_j : j \in J)$ whose rows are linearly independent. We define

$$g_{A_J}(\mathbf{x}_J) = \det(\hat{A}_J \mathrm{diag}(\mathbf{x}_J)^2 \hat{A}_J^T) = \sum_{I \in \binom{J}{r}} \det(\hat{A}_I)^2 \prod_{i \in I} x_i^2. \tag{36}$$

Here $\hat{A}_I$ denotes the square submatrix of $\hat{A}_J$ induced on the $r$ columns indexed by $I \subset J$.

**Lemma 30.** *Let $p$ be a smooth point on the reciprocal plane $\mathcal{L}^{-1}$ with* $\mathrm{supp}(p) = J$. *Then the ramification locus $\mathcal{R}_A$ contains the point $p$ if and only if $g_{A_J}(p_J) = 0$.*

**Proof.** Since $p$ is smooth, the condition (33) reduces to $\ker(A_J) \cap \ker(\mathrm{Jac}(\mathcal{L}_J^{-1})) \neq \{0\}$. From the argument prior to (35) we see that, for $p_J \in (\mathbb{C}^*)^J$, this is equivalent to $g_{A_J}(p_J) = 0$. $\square$

**Remark 31.** This characterization shows that the ramification locus $\mathcal{R}_A$ equals the closure of its intersection with the torus, $\mathcal{R}_A \cap T$. To see this, suppose that $\mathcal{R}_A$ has some component $Z$ contained in the boundary of the torus $\{x_1 \cdots x_n = 0\}$. Then $Z$ is contained in $\mathcal{L}_J^{-1}$ for some proper flat $J$, where $\dim(\mathcal{L}_J^{-1}) = \operatorname{rank}(A_J) - 1$. Since $\mathcal{R}_A$ is pure of codimension one in $\mathcal{L}^{-1}$, we see that $\dim(Z) = d-2$. It follows that $\operatorname{rank}(A_J) = d-1$ and $Z = \mathcal{L}_J^{-1}$. However, $M(A/J)$ has rank 1 and is therefore basic. Lemma 30 then tells us that $\mathcal{L}_J^{-1}$ is not contained in $\mathcal{R}_A$. This shows that to define the ideal $J_A$ in (34), we could instead saturate with respect to the ideal $\langle x_1 x_2 \cdots x_n \rangle$.

We shall now use the polynomial $g_A(\mathbf{x})$ to compute the degree of the ramification cycle.

**Proof of Theorem 2.** Let $A$ be a non-basic real $d \times n$-matrix of rank $d$ and $\chi(t)$ the characteristic polynomial of the matroid $M(A)$. We shall prove that the degree of the algebraic cycle underlying the $(d-2)$-dimensional subscheme $\mathcal{R}_A$ of $\mathbb{P}^{n-1}$ equals the matroid invariant (4). Lemma 27 then implies that $H_A(\mathbf{b})$ has the same degree, and this will complete the proof of Theorem 2.

From above, we know that the scheme $\mathcal{R}_A$ is contained in the hypersurface $\{g_A = 0\}$ of $\mathbb{P}^{n-1}$. Let $\widehat{\mathcal{R}}_A$ denote the scheme-theoretic intersection of the reciprocal plane with this hypersurface:

$$\widehat{\mathcal{R}}_A = \operatorname{Proj}\big(\mathbb{C}[\mathbf{x}]/(I(\mathcal{L}^{-1}) + \langle g_A \rangle)\big). \tag{37}$$

The $(d-2)$-dimensional scheme $\widehat{\mathcal{R}}_A$ is the intersection of the $(d-1)$-dimensional irreducible variety $\mathcal{L}^{-1}$ and the hypersurface $g_A$. By Bézout's Theorem [9, Theorem 1.4.4], its degree equals

$$\deg\big(\widehat{\mathcal{R}}_A\big) = \deg(g_A) \cdot \deg(\mathcal{L}^{-1}). \tag{38}$$

We claim that $\widehat{\mathcal{R}}_A$ decomposes into $\#\mathrm{Hyp}(A)+1$ components of dimension $d-2$, one of which is $\mathcal{R}_A$. Here $\mathrm{Hyp}(A)$ denotes the set of hyperplane flats, that is, flats $J$ such that $\operatorname{rk}(A_J) = d-1$. We see that $\mathcal{R}_A$ and $\widehat{\mathcal{R}}_A$ agree in the torus $T$, so their difference must lie in the coordinate hyperplanes. Recall from Section 5 that the reciprocal plane intersects the dense torus $T^J$ of $\mathbb{P}^J$ if and only if $J$ is a flat, and if so, the closure of that intersection is the reciprocal plane $\mathcal{L}_J^{-1}$. Such a stratum has dimension $d-2$ in $\mathbb{P}^{n-1}$ if and only if $J$ is a hyperplane flat. Since $J \in \mathrm{Hyp}(A)$ does not contain a basis of $M(A)$, each summand in the formula (35) for $g_A$ vanishes on $T^J$. To be precise, $g_A$ vanishes to order exactly 2 on the torus $T^J$, since $J$ is only one element away from containing a basis.

Furthermore, the strata $\mathcal{L}_J^{-1}$ are not contained in $\mathcal{R}_A$ for $J \in \mathrm{Hyp}(A)$. This follows from Lemma 30. Indeed, by Corollary 26, the points in $\mathcal{L}^{-1} \cap T^J$ are non-singular in $\mathcal{L}^{-1}$, and hence the polynomial $g_{A_J}(\mathbf{x}_J)$ is not identically zero on $\mathcal{L}_J^{-1}$. We conclude that the irreducible varieties $\mathcal{L}_J^{-1}$, for $J \in \mathrm{Hyp}(A)$, are components of dimension $d-2$ and multiplicity 2 in the scheme $\widehat{\mathcal{R}}_A$.

We have derived the following equidimensional decomposition of the cycle defined in (37):

$$\widehat{\mathcal{R}}_A = \mathcal{R}_A \cup \left( \bigcup_{J \in \mathrm{Hyp}(A)} 2 \cdot \mathcal{L}_J^{-1} \right). \tag{39}$$

Since the degree is additive on equidimensional cycles, we can use (38) to conclude that

$$\begin{aligned} \deg(\mathcal{R}_A) &= \deg(g_A) \cdot \deg(\mathcal{L}^{-1}) - 2 \sum_{J \in \mathrm{Hyp}(A)} \deg(\mathcal{L}_J^{-1}) \\ &= 2d \cdot \mu(A) - 2 \sum_{J \in \mathrm{Hyp}(A)} \mu(A_J). \end{aligned} \tag{40}$$

The coefficient of $t^i$ in the characteristic polynomial $\chi(t)$ equals $(-1)^{d-i}$ times the sum of the Möbius invariants $\mu(A_J)$ where $J$ runs over all flats of rank $d-i$. For $i=0$ this gives $\mu(A)=(-1)^d\chi(0)$, and for $i=1$ we get $\sum_{J\in\mathrm{Hyp}(A)}\mu(A_J)=(-1)^{d-1}\chi'(0)$. Hence the right hand side of (40) equals the desired matroid invariant (4). This completes the proof of Theorem 2. $\square$

The decomposition (39) can be used to compute the ideal of the ramification scheme. Namely, since all hyperplane strata $\mathcal{L}_J^{-1}$ lie in complement of the torus $T$, we have the algebraic identity

$$J_A=\big(I(\mathcal{L}^{-1})+\langle\, g_A\,\rangle\big):\langle x_1x_2\cdots x_n\rangle^{\infty}. \tag{41}$$

We illustrate the identity (41) and our proof of Theorem 2 for the codimension 1 case.

**Example 32.** Let $A$ be the matrix in Eq. (20) of Section 3. The reciprocal plane $\mathcal{L}^{-1}$ is the hypersurface defined by the elementary symmetric polynomial $e_{n-1}(x_1,x_2,\ldots,x_n)$. Eq. (35) defining the ramification locus in the torus is $g_A=e_{n-1}(x_1^2,x_2^2,\ldots,x_n^2)$. The scheme $\widehat{\mathcal{R}}_A$ in (37) is the complete intersection of these two hypersurfaces. Its ideal has the primary decomposition

$$\begin{aligned}&\big\langle e_{n-1}(x_1,x_2,\ldots,x_n),\,e_{n-1}(x_1^2,x_2^2,\ldots,x_n^2)\big\rangle\\ &\quad=\langle e_{n-1}(x_1,x_2,\ldots,x_n),\,e_{n-2}(x_1,x_2,\ldots,x_n)\rangle\cap\bigcap_{1\le i<j\le n}\langle\, x_i^2,x_i+x_j\,\rangle.\end{aligned} \tag{42}$$

This is the decomposition discussed after (38), with the first intersectand being the ideal $J_A$ that defines $\mathcal{R}_A$. This ideal is contained in the Jacobian ideal of the reciprocal plane $\mathcal{L}^{-1}$ because

$$e_{n-2}=\frac{1}{2}\sum_{i=1}^{n}\frac{\partial e_{n-1}}{\partial x_i}.$$

This identity proves the ideal-theoretical inclusion $J_A\subset I(\mathrm{Sing}(\mathcal{L}^{-1}))$. We conclude that $\mathcal{R}_A$ contains $\mathrm{Sing}(\mathcal{L}^{-1})$ when $n=d+1$. As we shall see in Theorem 35, the inclusion $\mathrm{Sing}(\mathcal{L}^{-1})\subset\mathcal{R}_A$ is always true, even if $n>d+1$. This inclusion implies, as argued in Corollary 37, that the real variety of $H_A(\mathbf{b})$ is indeed the union of codimension 2 planes given in (27). $\square$

We close this section with a combinatorial proof of the assertion, stated informally immediately after Theorem 2, that generic matrices maximize the degree of the entropic discriminant.

**Proposition 33.** *As $A$ ranges over all non-basic $d\times n$-matrices of rank $d$, the degree of the entropic discriminant $H_A(\mathbf{b})$ attains its maximal value $2(n-d)\binom{n-1}{d-2}$ when the matroid $M(A)$ is uniform.*

**Proof.** It follows from Theorem 2 and Example 18 that $2(n-d)\binom{n-1}{d-2}$ is the degree of the entropic discriminant when $M=M(A)$ is uniform. We must show that this number is a strict upper bound otherwise. The claim is an entirely matroid-theoretic statement, and so let us define $\delta(M)=2(-1)^{\mathrm{rk}(M)}(\mathrm{rk}(M)\chi_M(0)+\chi_M'(0))$ for all matroids $M$. The characteristic polynomial satisfies a deletion–contraction recurrence, namely, $\chi_M(t)=\chi_{M\setminus e}(t)-\chi_{M/e}(t)$ for $e\in M$ not an isthmus. It follows that the entropic degree satisfies a deletion–contraction recurrence plus a correction term:

$$\delta(M)=\delta(M\setminus e)+\delta(M/e)+\mu(M/e).$$

All three terms on the right hand side are non-negative. The desired inequality follows by induction on the rank $d$ and corank $n-d$. In rank 1 all simple matroids are uniform. Corank 1 is dealt with in Section 3. The same argument shows that $\mu(M) \leq \binom{n-1}{d-1}$ with equality if and only if $M = U_{d,n}$. $\square$

## 7. Real issues

Our point of departure for this paper was the observation that, for real $\mathbf{b}$, Eqs. (2) have only real solutions, namely, the $\mu(A)$ analytic centers of the bounded regions in the arrangement of $n$ coordinate hyperplanes in $\{A\mathbf{x} = \mathbf{b}\} \simeq \mathbb{R}^{n-d}$. It is thus natural to ask what it means for two such analytic centers to collide, and how this relates to the real points in the ramification locus and in the entropic discriminant. We shall prove that the real loci of these two complex varieties are both pure of codimension one. Our first step in this direction is the following lemma.

**Lemma 34.** *All real points in the ramification scheme are singular in the reciprocal plane:*

$$(\mathcal{R}_A)_{\mathbb{R}} \subseteq \mathrm{Sing}(\mathcal{L}^{-1})_{\mathbb{R}}. \tag{43}$$

**Proof.** The sum of squares formula in (35) reveals that $g_A(\mathbf{x}) = 0$ has no real solutions in the torus $T$. In symbols, $(\mathcal{R}_A \cap T)_{\mathbb{R}} = \emptyset$. Likewise, for any flat $J$ with $\mathcal{L}_J^{-1}$ nonsingular in $\mathcal{L}^{-1}$, the polynomial $g_{A_J}$ is a similar sum of squares, and hence $(\mathcal{R}_A \cap T^J)_{\mathbb{R}} = \emptyset$. Lemma 30 ensures that no regular point of $\mathcal{L}^{-1}$ with real coordinates lies in the ramification locus of the morphism $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$. $\square$

The following is our main result in this section. We find that the reverse inclusion holds in (43).

**Theorem 35.** *The ramification scheme $\mathcal{R}_A$ contains the singular locus of $\mathcal{L}^{-1}$, and we have*

$$(\mathcal{R}_A)_{\mathbb{R}} = \mathrm{Sing}(\mathcal{L}^{-1})_{\mathbb{R}}. \tag{44}$$

This theorem implies that the saturation in the formula (34) for the ramification ideal $J_A$ was unnecessary. Before presenting the proof, we shall derive two corollaries and discuss one example.

**Corollary 36.** *The Zariski closure of $(\mathcal{R}_A)_{\mathbb{R}}$ is pure of codimension 2 in $\mathcal{L}^{-1}$.*

**Proof.** Theorem 35 implies that the Zariski closure of the set $(\mathcal{R}_A)_{\mathbb{R}}$ of real ramification points equals the singular locus $\mathrm{Sing}(\mathcal{L}^{-1})$ of the reciprocal plane $\mathcal{L}^{-1}$. Corollary 26 represents $\mathrm{Sing}(\mathcal{L}^{-1})$ as a union of reciprocal linear spaces all of which are defined over $\mathbb{R}$ and have codimension 2 in $\mathcal{L}^{-1}$. $\square$

We now obtain the following characterization of the real locus of the entropic discriminant.

**Corollary 37.** *The Zariski closure of the set of real points in the hypersurface defined by the entropic discriminant $H_A(\mathbf{b})$ is pure of codimension 2 in $\mathbb{P}^{d-1}$. Its irreducible components are the linear spaces* $\mathrm{span}(A_j : j \in J)$, *where $J$ runs over all corank 2 flats of $M(A)$ for which the contraction $M(A/J)$ is a non-basic matroid.*

**Proof.** The real variety of $H_A$ is the image of the real points in $\mathcal{R}_A$ under the $\mu(A)$-to-one morphism $A : \mathcal{L}^{-1} \to \mathbb{P}^{d-1}$. Hence the Zariski closure of the real variety of $H_A$ is pure of codimension 2 in $\mathbb{P}^{d-1}$ as well. The description of its irreducible components now follows from that given in Corollary 26. $\square$

We now revisit our very first example to illustrate the previous corollary.

**Example 38.** For $d = 3$, the codimension-2 strata of $\mathcal{L}^{-1}$ are the $n$ coordinate points $e_i$ in $\mathbb{P}^{n-1}$. Their images under the map $A$ are the columns $A_1, \ldots, A_n$. For generic $A$, the points $e_1, \ldots, e_n$ comprise $\mathrm{Sing}(\mathcal{L}^{-1})$. Lemma 34 implies that $V_{\mathbb{R}}(H_A)$ is contained in $\{A_1, \ldots, A_n\}$, and Theorem 35 reveals that equality holds. For special $3 \times n$-matrices $A$, the matroid $M(A/i)$ may be basic for some $i$. If this happens then $e_i$ is a non-singular point in $\mathcal{L}^{-1}$ and its image $A_i$ does not belong to $V_{\mathbb{R}}(H_A)$. Looking back at Example 1, we notice that the matroid $M(A/i)$ is basic for $i = 1$ and it is non-basic for $i = 2, 3, 4, 5$. This explains our finding in (3) that the real variety $V_{\mathbb{R}}(H_A)$ consists of precisely the four points $A_2, A_3, A_4$ and $A_5$ in the projective plane $\mathbb{P}^2$. $\square$

We are now ready to present the proof of our main result in this section.

**Proof of Theorem 35.** We first note that the identity (44) follows immediately from Lemma 34 and the inclusion $\mathcal{R}_A \supseteq \mathrm{Sing}(\mathcal{L}^{-1})$ in the first assertion. Hence it suffices to prove that inclusion.

By Corollary 26, the singular locus of $\mathcal{L}^{-1}$ is a reducible variety whose irreducible components are the boundary strata $\mathcal{L}_J^{-1}$ where $M(A/J)$ is a non-basic matroid of rank 2. We consider one such component $\mathcal{L}_J^{-1}$, regarded as a subvariety of $\mathbb{C}^J \times \{0\}$ inside of $\mathbb{C}^n = \mathbb{C}^J \times \mathbb{C}^{J^c}$. A generic point of $\mathcal{L}_J^{-1}$ has the form $(p, 0)$ where $p \in (\mathbb{C}^*)^J$. Our goal is to show that this point lies in the ramification locus $\mathcal{R}_A$ by producing a sequence of points in $\mathcal{R}_A$ that converges to $(p, 0)$.

We may assume that $J = \{1, \ldots, k\}$ is a flat of rank $d - 2$ and our matrix $A$ has the block form

$$A = \left(\begin{array}{c|c} \hat{A} & * \\ \hline 0 & B \end{array}\right)$$

where $\hat{A} \in \mathbb{R}^{(d-2)\times k}$ and $B \in \mathbb{R}^{2\times(n-k)}$ are both of full row-rank. In these coordinates, we get $M(A_J) = M(\hat{A})$ and $M(A/J) = M(B)$.

Now, let us return to our generic point $(p, 0) \in \mathcal{L}_J^{-1}$. The partial specialization $g_A(p, \mathbf{x}_{J^c})$ is a polynomial in $\mathbb{C}[x_{k+1}, x_{k+2}, \ldots, x_n]$. It is non-homogeneous and its terms of lowest total degree come from those bases $I$ of $M(A)$ for which $|I \cap J| = d - 2$. For any such $I$, we have

$$\det(A_I) = \det(\hat{A}_{I\cap J}) \cdot \det(B_{I\cap J^c}).$$

From this we see that the initial form of $g_A(p, \mathbf{x}_{J^c})$ of lowest degree terms can be written as

$$\mathrm{in}_{-\mathbf{1}}(g_A(p, \mathbf{x}_{J^c})) = g_{\hat{A}}(p) \cdot g_B(\mathbf{x}_{J^c}). \tag{45}$$

From the results of Section 6 we know that $\{g_{\hat{A}} = 0\} \cap \mathcal{L}_J^{-1}$ has codimension 1 in $\mathcal{L}_J^{-1}$. This implies $g_{\hat{A}}(p) \neq 0$ because the point $(p, 0)$ was chosen to be generic in $\mathcal{L}_J^{-1}$.

In order to proceed, we need to represent the ramification locus around $p$ by a single polynomial, rather than as a subvariety of $\mathcal{L}^{-1}$. To do this, we rationally parametrize the points $\mathbf{x}_{J^c}$ for which $(p, \mathbf{x}_{J^c})$ lies in $\mathcal{L}^{-1}$ using the matrix $B$. First, note that the intersection of the linear space $\mathcal{L}$ with $\{p\} \times \mathbb{C}^{J^c}$ gives an affine linear space in $\mathbb{C}^{J^c}$ of the form $v + \mathrm{rowspan}(B)$ for some vector $v$ in $\mathbb{C}^{J^c}$. We can parametrize this space by $v + \mathbf{z}B$ where $\mathbf{z} = (z_1, z_2)$. This gives the rational parametrization $(p, (v + \mathbf{z}B)^{-1})$ of the intersection of $\mathcal{L}^{-1}$ with $\{p\} \times \mathbb{C}^{J^c}$.

Now we plug this parametrization into $g_A(p, \mathbf{x}_{J^c})$ and clear denominators to get a polynomial in $\mathbb{C}[z_1, z_2]$. Define $g(\mathbf{z}) \in \mathbb{C}[z_1, z_2]$ to be this polynomial,

$$\begin{aligned} g(\mathbf{z}) &= g_A(p, (v + \mathbf{z}B)^{-1}) \prod_{i \in J^c} (v_i + \mathbf{z}B_i)^2 \\ &= \sum_{I \in \binom{[n]}{d}} \det(A_I)^2 \prod_{i \in I \cap J} p_i^2 \prod_{j \in J^c \setminus I} (v_j + \mathbf{z}B_j)^2. \end{aligned} \tag{46}$$

If $\mathbf{z}$ is a solution to $g(\mathbf{z}) = 0$ for which each coordinate of $v + \mathbf{z}B$ is non-zero, then the point $(p, (v + \mathbf{z}B)^{-1})$ lies in the ramification locus $\mathcal{R}_A$.

Since $J$ is a flat, the $n - k$ linear forms $\mathbf{z}B_i$ are non-zero for all indices $i$. This implies that $x_i = 1/(v_i + \mathbf{z}B_i)$ has degree $-1$. Thus the terms of highest degree in $g(\mathbf{z})$ correspond exactly to the terms of lowest degree in $g_A(p, \mathbf{x}_{J^c})$. From (45), we see that the leading form of $g(\mathbf{z})$ is

$$\mathrm{in}_{\mathbf{1}}(g(\mathbf{z})) = g_{\hat{A}}(p) \cdot g_B((\mathbf{z}B)^{-1}) \cdot \prod_{i \in J^c} (\mathbf{z}B_i)^2.$$

Our next step is to find a solution to the initial equation $\mathrm{in}_{\mathbf{1}} g(\mathbf{z}) = 0$ and to then extend it to the desired sequence of points in $\mathcal{R}_A$. As the matroid $M(A/J) = M(B)$ is non-basic, it follows from Corollary 5 that the ramification $\mathcal{R}_{A/J}$ is nonempty. Hence there is a point $q \in \mathcal{L}_{A/J} \cap (\mathbb{C}^*)^{n-k}$ such that $g_{A/J}(q^{-1}) = g_B(q^{-1}) = 0$. Let $z$ be the unique vector such that $zB = q$. We may assume that $B$ has the form $(\,\mathrm{Id}_2 \;\; B'\,)$. Thus implying that $z_i = q_i \neq 0$ for $i = 1, 2$.

By Lemma 40, we can extend this solution $z \in (\mathbb{C}^*)^2$ to a solution $Z = Z(\epsilon)$ of $g(\mathbf{z})$, where the coordinates of $Z = (Z_1, Z_2)$ lie in the field $\mathbb{C}\{\{\epsilon\}\}$ of Puiseux series:

$$Z_i = z_i \frac{1}{\epsilon} + \text{higher order terms} \in \mathbb{C}\{\{\epsilon\}\} \quad \text{for } i = 1, 2.$$

Moreover, by Lemmas 39 and 40, these series converge in a neighborhood of zero in $\mathbb{R}_{>0}$.

Now consider the point $Q = Q(\epsilon) = v + ZB$ with coordinates $Q_i = q_i \frac{1}{\epsilon} + \cdots$ in $\mathbb{C}\{\{\epsilon\}\}$. We can invert $Q_i$ in the field of Puiseux series to get

$$Q_i^{-1} = q_i^{-1}\epsilon + \text{higher order terms} \in \mathbb{C}\{\{\epsilon\}\} \quad \text{for } i = 1, \ldots, n-k,$$

and these series converge for real $\epsilon$ in an open segment $(0, \epsilon_0)$ near zero (see Lemma 39).

Then, by (46), the point $(p, Q^{-1})$ in $\mathcal{L}^{-1} \otimes_{\mathbb{C}} \mathbb{C}\{\{\epsilon\}\}$ is a zero of the polynomial $g_A(\mathbf{x})$. Specializing to sufficiently small $\epsilon \in \mathbb{R}_{>0}$, gives a point $(p, Q(\epsilon)^{-1}) \in (\mathbb{C}^*)^n$ that belongs to the ramification locus $\mathcal{R}_A$. Furthermore, as $\epsilon$ approaches 0, the limit of the points $(p, Q^{-1}(\epsilon))$ is $(p, 0)$ in $\mathcal{L}_J^{-1} \times \{0\}$. This shows $\mathcal{L}_J^{-1} \subseteq \mathcal{R}_A$ and consequently $\mathrm{Sing}(\mathcal{L}^{-1}) \subseteq \mathcal{R}_A$. $\square$

Before Lemma 40, we need a short lemma on the convergence of reciprocals of Puiseux series.

**Lemma 39.** *If $x(\epsilon)$ is a nonzero Puiseux series that converges for $\epsilon > 0$ in a neighborhood of 0, then its inverse $x(\epsilon)^{-1}$ in $\mathbb{C}\{\{\epsilon\}\}$ also converges for real $\epsilon$ in an open segment $(0, \epsilon_0)$.*

**Proof.** Suppose $x(\epsilon) = u\epsilon^k +$ higher order terms. We can write the field of Puiseux series as the union of $\mathbb{C}((\epsilon^{1/m}))$ over $m \in \mathbb{Z}_+$. Thus for some $m \in \mathbb{Z}_+$, replacing $\epsilon$ with $\epsilon^m$ yields a Laurent series $x(\epsilon^m)$, which also converges in a neighborhood of 0. In particular, $\epsilon^{-mk}x(\epsilon^m)$ is a convergent power series with constant term $u$ and has an inverse $y(\epsilon)$ in the ring of convergent power series (see [8, Section 6.4]). Then $y(\epsilon) = 1/u + \cdots$ satisfies $\epsilon^{-mk}x(\epsilon^m)y(\epsilon) = 1$.

Replacing $\epsilon$ with $\epsilon^{1/m}$, we see that $\epsilon^{-k}y(\epsilon^{1/m})$ is an inverse for $x(\epsilon)$. Furthermore, since $y(\epsilon)$ and $y(\epsilon^{1/m})$ converge in a neighborhood of zero, $x(\epsilon)^{-1} = \epsilon^{-k}y(\epsilon^{1/m})$ also converges for $\epsilon > 0$ in a neighborhood of zero. $\square$

Now all that remains is to lift roots of initial forms to solutions over $\mathbb{C}\{\{\epsilon\}\}$.

**Lemma 40.** *Let $g(z_1, z_2)$ be a polynomial with complex coefficients and initial form $\text{in}_{\mathbf{1}}(g)$, consisting of the highest terms with respect to total degree. Let $u = (u_1, u_2) \in (\mathbb{C}^*)^2$ be any solution to the equation $\text{in}_{\mathbf{1}}(g)(u_1, u_2) = 0$. Then there exists a vector $v(\epsilon)$ that satisfies $g(v(\epsilon)) = 0$ and whose coordinates are Puiseux series of the form*

$$v_i(\epsilon) = u_i \frac{1}{\epsilon} + \textit{higher order terms in } \epsilon, \quad \textit{for } i = 1, 2,$$

*that converge for $\epsilon$ in some neighborhood $(0, \epsilon_0)$ of zero.*

**Proof.** We invert the variables $z_i$ and work with the polynomial

$$\overline{g}(z) = z_1^{\deg(g)} \cdot z_2^{\deg(g)} \cdot g(z_1^{-1}, z_2^{-1}).$$

The highest-degree terms of $g$ then correspond to the lowest-degree terms of $\overline{g}$. Furthermore, the point $u^{-1} = (1/u_1, 1/u_2)$ is a solution of $\text{in}_{-\mathbf{1}}(\overline{g})$.

Our hypothesis states that the Newton polygon of $\overline{g}(z_1, z_2)$ has an edge of slope $-1$, and $(1/u_1, 1/u_2)$ is a root of the corresponding binary form $\text{in}_{-\mathbf{1}}(\overline{g})(z_1, z_2)$. Using the classical Newton–Puiseux algorithm, we can construct a power series expansion of $z_2$ in terms of $z_1 = \frac{1}{u_1}\epsilon$, having the form $z_2 = \frac{1}{u_2}\epsilon + \cdots$. The resulting series in $\epsilon$ converges by the arguments in [8, Section 7.11].

This solution has an inverse in the field of Puiseux series, and this inverse will be our desired solution $(v_1(\epsilon), v_2(\epsilon))$ of $g(z_1, z_2) = 0$. Namely, if $w(\epsilon) = (w_1(\epsilon), w_2(\epsilon)) \in \mathbb{C}\{\{\epsilon\}\}^2$ is the solution to $\overline{g}(z_1, z_2)$ found in the paragraph above, then $v_i(\epsilon) = w_i(\epsilon)^{-1}$ is a solution to $g(z_1, z_2)$. By Lemma 39, the Puiseux series $v_i(\epsilon)$ converge in a neighborhood $(0, \epsilon_0)$ of the origin in $\mathbb{R}_{>0}$. $\square$

This concludes our study of the entropic discriminant. In spite of the progress that has been achieved, there are still many unresolved problems concerning $H_A(\mathbf{b})$. We list five open questions:

*Open Questions*:

(1) *Is the entropic discriminant $H_A(\mathbf{b})$ always a sum of squares*?

We know that the answer is yes for $n = d+1$ and for $d = 2$, but even the case $d = 3$ of plane curves is open. It would be especially nice to write $H_A(\mathbf{b})$ as sum of squares in the maximal minors of the matrix $(A, \mathbf{b})$, as we did in (16) and (17) for $(d, n) = (2, 3),\ (2, 4)$.

(2) *What is the Newton polytope of the entropic discriminant $H_A(\mathbf{b})$*?

For instance, when $A$ is the matrix in (20) then the table (26) suggests that the Newton polytope of $H_A(\mathbf{b})$ is the standard permutohedron, scaled by a factor of two.

(3) *Find ideal generators for the ramification scheme*.
Here is a concrete conjecture about minimal generators of the ideal $J_A$ in (41). Fix $n \geq d+2$ and a $d \times n$-matrix $A$ whose matroid is uniform. We know from Example 23 that $I(\mathcal{L}^{-1})$ is minimally generated by $\binom{n-1}{d}$ polynomials of degree $d$. We conjecture that $J_A$ has precisely $\binom{d+1}{2}$ additional minimal generators of degree $2d-2$, namely, the restrictions to $\mathcal{L}^{-1}$ of the rational functions $g_A(\mathbf{x})/x_i x_j$ for some $i, j \in [n]$. We can show that these rational functions are polynomials on $\mathcal{L}^{-1}$ and that they vanish on $\mathrm{Sing}(\mathcal{L}^{-1})$. Do they generate our ideal?
(4) *How is the entropic discriminant related to the Gauss curve of the central curve*?
The degree formula for the Gauss curve in [3, Section 5] is essentially the same as the degree formula we derived for $H_A(\mathbf{b})$. What is the most natural geometric explanation for this?
(5) *How does the entropic discriminant depend on the choice of monomial to be maximized*?
In light of Varchenko's work [27], it is natural to replace $x_1 x_2 \cdots x_n$ in (1) by a monomial $\mathbf{x}^{\mathbf{u}} = x_1^{u_1} x_2^{u_2} \cdots x_n^{u_n}$ with indeterminate exponents. This would lead to a refined discriminant that is a bihomogeneous polynomial in $(\mathbf{b}, \mathbf{u})$. What is the bidegree of that polynomial?

## Acknowledgments

We thank Anders Björner, Igor Dolgachev, Chris Hillar, Daniel Plaumann, Frank Sottile and David Speyer for helpful discussions. We are especially grateful to David Speyer for helping us to a proof of Theorem 35. We also thank the anonymous referee for thoughtful comments and suggestions. Raman Sanyal was supported by a Miller Postdoctoral Research Fellowship at UC Berkeley. Bernd Sturmfels and Cynthia Vinzant were partially supported by the US National Science Foundation (DMS-0757207 and DMS-0968882). Bernd Sturmfels also thanks the Mittag-Leffler-Institute and MATHEON Berlin for their hospitality during this project.

## References

[1] P. Alexandersson, B. Shapiro, Discriminants, symmetrized graph monomials and sums of squares, Experiment. Math. 21 (2012) 353–361.
[2] C.W. Borchardt, Neue Eigenschaft der Gleichung, mit deren Hülfe man die seculären Störungen der Planeten bestimmt, J. Reine Angew. Math. 30 (1846) 38–45.
[3] J.A. de Loera, B. Sturmfels, C. Vinzant, The central curve in linear programming, Found. Comput. Math. 12 (2012) 509–540.
[4] A. Dimca, S. Papadima, Hypersurface complements, Milnor fibers and higher homotopy groups of arrangments, Ann. of Math. 158 (2003) 473–507.
[5] I.V. Dolgachev, Polar Cremona transformations, Michigan Math. J. 48 (2000) 191–202.
[6] I.V. Dolgachev, Classical Algebraic Geometry: A Modern View, Cambridge Univ. Press, 2012.
[7] M. Domokos, The discriminant of symmetric matrices as a sum of squares and the orthogonal group, Comm. Pure Appl. Math. 64 (2011) 443–465.
[8] G. Fischer, Plane Algebraic Curves, in: Student Mathematical Library, vol. 15, American Mathematical Society, Providence, RI, 2001.
[9] H. Flenner, L. O'Carroll, W. Vogel, Joins and Intersections, Springer Verlag, New York, 1999.
[10] C. Greene, T. Zaslavsky, On the interpretation of Whitney numbers through arrangements of hyperplanes, zonotopes, non-Radon partitions, and orientations of graphs, Trans. Amer. Math. Soc. 280 (1983) 97–126.
[11] C. Hillar, A. Wibisono, Maximum entropy distributions on graphs, arXiv:1301.3321.
[12] J. Huh, Milnor numbers of projective hypersurfaces and the chromatic polynomial of graphs, J. Amer. Math. Soc. 25 (2012) 907–927.
[13] J. Huh, E. Katz, Log-concavity of characteristic polynomials and the Bergman fan of matroids, Math. Ann. 354 (2012) 1103–1116.
[14] N.V. Ilyushechkin, The discriminant of the characteristic polynomial of a normal matrix, Mat. Zametki 51 (1992) 16–23.

[15] P.D. Lax, On the discriminant of real symmetric matrices, Comm. Pure Appl. Math. 51 (1998) 1387–1396.
[16] T. Muir, A Treatise on the Theory of Determinants, Dover Publications Inc., New York, 1960. Revised and enlarged by William H. Metzler.
[17] M. Nagata, On the purity of branch loci in regular local rings, Illinois J. Math. 3 (1959) 328–333.
[18] M.J. Newell, On identities associated with a discriminant, Proc. Edinb. Math. Soc. 18 (1972–1973) 287–291.
[19] J. Oxley, Matroid Theory, Oxford University Press, 1992.
[20] N. Proudfoot, D. Speyer, A broken circuit ring, Beitr. Algebra Geom. 47 (2006) 161–166.
[21] R. Sanyal, On the derivative cones of polyhedral cones, Adv. Geom. 13 (2013) 315–321.
[22] H. Schenck, Ş. Tohăneanu, The Orlik–Terao algebra and 2-formality, Math. Res. Lett. 16 (2009) 171–182.
[23] F. Sottile, Real Solutions to Equations from Geometry, in: University Lecture Series, American Mathematical Society, Providence, Rhode Island, 2011.
[24] R.P. Stanley, An introduction to hyperplane arrangements, in: Geometric Combinatorics, in: IAS/Park City Math. Ser., vol. 13, American Mathematical Society, Providence, Rhode Island, 2007, pp. 389–496.
[25] B. Sturmfels, Solving Systems of Polynomial Equations, in: CBMS Regional Conference Series in Mathematics, vol. 97, American Mathematical Society, Providence, Rhode Island, 2002.
[26] C. Tatakis, A. Thoma, On the universal Gröbner bases of toric ideals of graphs, J. Combin. Theory Ser. A 118 (2011) 1540–1548.
[27] A. Varchenko, Critical points of the product of powers of linear functions and families of bases of singular vectors, Compos. Math. 97 (1995) 385–401.
[28] N. White (Ed.), Combinatorial Geometries, in: Encyclopedia of Mathematics and its Applications, vol. 29, Cambridge University Press, Cambridge, 1987.
[29] J. Yu, D. Yuster, Representing tropical linear spaces by circuits, in: Proceedings of FPSAC, 2007.
[30] T. Zaslavsky, Chromatic invariants of signed graphs, Discrete Math. 42 (1982) 287–312.
[31] T. Zaslavsky, Signed graphs, Discrete Appl. Math. 4 (1982) 47–74.