

COMPLEXITY OF LINEAR CIRCUITS AND GEOMETRY

FULVIO GESMUNDO, JONATHAN D. HAUNSTEIN, CHRISTIAN IKENMEYER AND J.M. LANDSBERG

ABSTRACT. We use algebraic geometry to study matrix rigidity, and more generally, the complexity of computing a matrix-vector product, continuing a study initiated in [13, 11]. In particular, we (i) exhibit many non-obvious equations testing for (border) rigidity, (ii) compute degrees of varieties associated to rigidity, (iii) describe algebraic varieties associated to families of matrices that are expected to have super-linear rigidity, and (iv) prove results about the ideals and degrees of cones that are of interest in their own right.

1. INTRODUCTION

Given an $n \times n$ matrix A , how many additions are needed to perform the map

$$(1.0.1) \quad x \mapsto Ax,$$

where x is a column vector? L. Valiant initiated a study of this question in [27]. He used the model of computation of *linear circuits* (see §1.2) and observed that for a generic linear map one requires a linear circuit of size n^2 . He posed the following problem:

Problem 1.0.2. Find an *explicit* sequence of matrices A_n needing linear circuits of size super-linear in n to compute (1.0.1).

“Explicit” has a precise meaning, see [11]. Valiant defined a notion of *rigidity* that is a measurement of the size of the best circuit of a very restricted type (see §1.2) needed to compute (1.0.1). He proved that strong lower bounds for rigidity implies super-linear lower bounds for any linear circuit computing (1.0.1), see Theorem 1.5.1 below. This article continues the use of algebraic geometry, initiated in [11] and the unpublished notes [13], to study these issues.

1.1. Why algebraic geometry? Given a polynomial P on the space of $n \times n$ matrices that vanishes on all matrices of low rigidity (complexity), and a matrix A such that $P(A) \neq 0$, one obtains a lower bound on the rigidity (complexity) of A .

For a simple example, let $\hat{\sigma}_{r,n} \subset Mat_n$ denote the variety of $n \times n$ matrices of rank at most r . (If n is understood, we write $\hat{\sigma}_r = \hat{\sigma}_{r,n}$.) Then, $\hat{\sigma}_{r,n}$ is the zero set of all minors of size $r + 1$. If one minor of size $r + 1$ does not vanish on A , we know the rank of A is at least $r + 1$.

Define the *r-rigidity* of an $n \times n$ matrix M to be the smallest s such that $M = A + B$ where $A \in \hat{\sigma}_{r,n}$ and B has exactly s nonzero entries. Write $Rig_r(M) = s$.

Define the set of *matrices of r-rigidity at most s*:

$$(1.1.1) \quad \hat{\mathcal{R}}[n, r, s]^0 := \{M \in Mat_{n \times n} \mid Rig_r(M) \leq s\}.$$

Thus if we can find a polynomial P vanishing on $\hat{\mathcal{R}}[n, r, s]^0$ and a matrix M such that $P(M) \neq 0$, we know $Rig_r(M) > s$. One says M is *maximally r-rigid* if $Rig_r(M) = (n - r)^2$, and that M is *maximally rigid* if it is maximally r -rigid for all r . (See §2.1 for justification of this terminology.)

Key words and phrases. matrix rigidity, Discrete Fourier Transform, Vandermonde matrix, MSC 68Q17.
 Haunstein supported by NSF grant DMS-1262428 and DARPA Young Faculty Award (YFA).
 Landsberg supported by NSF grant DMS-1006353.

Our study has two aspects: finding explicit polynomials vanishing on $\hat{\mathcal{R}}[n, r, s]^0$, and proving qualitative information about such polynomials. The utility of explicit polynomials has already been explained. For a simple example of a qualitative property, consider the *degree* of a polynomial. As observed in [11], for a given d , one can describe matrices that cannot be in the zero set of any polynomial of degree at most d with integer coefficients. They then give an upper bound $\Delta(n) = n^{4n^2}$ for the degrees of the polynomials generating the ideal of the polynomials vanishing on $\hat{\mathcal{R}}[n, r, s]^0$, and describe a family of matrices that do not satisfy polynomials of degree $\Delta(n)$ (but this family is not explicit in Valiant’s sense).

Following ideas in [13, 11], we not only study polynomials related to rigidity, but also to different classes of matrices of interest, such as Vandermonde matrices. As discussed in [11], one could first try to prove a general Vandermonde matrix is maximally rigid, and then afterwards try to find an explicit sequence of maximally rigid Vandermonde matrices (a problem in n variables instead of n^2 variables).

Our results are described in §1.6. We first recall basic definitions regarding linear circuits in §1.2, give brief descriptions of the relevant varieties in §1.3, establish notation in §1.4, and describe previous work in §1.5. We have attempted to make this paper readable for both computer scientists and geometers. To this end, we put off the use of algebraic geometry until §5, although we use results from it in earlier sections, and introduce a minimal amount of geometric language in §2.1. We suggest geometers read §5 immediately after §2.1. In §2.2, we present our qualitative results about equations. We give examples of explicit equations in §3. We give descriptions of several varieties of matrices in §4. In §5, after reviewing standard facts on joins in §5.1, we present generalities about the ideals of joins in §5.2, discuss degrees of cones in §5.3 and then apply them to our situation in §5.4.

1.2. Linear circuits.

Definition 1.2.1. A *linear circuit* is a directed acyclic graph \mathcal{LC} in which each directed edge is labeled by a nonzero element of \mathbb{C} . If u is a vertex with incoming edges labeled by $\lambda_1, \dots, \lambda_k$ from vertices u_1, \dots, u_k , then \mathcal{LC}_u is the expression $\lambda_1 \mathcal{LC}_{u_1} + \dots + \lambda_k \mathcal{LC}_{u_k}$.

If \mathcal{LC} has n input vertices and m output vertices, it determines a matrix $A_{\mathcal{LC}} \in \text{Mat}_{n,m}(\mathbb{C})$ by setting

$$A_i^j := \sum_{\substack{p \text{ path} \\ \text{from } i \text{ to } j}} \prod_{\substack{e \text{ edge} \\ \text{of } p}} \lambda_e,$$

and \mathcal{LC} is said to compute $A_{\mathcal{LC}}$.

The *size* of \mathcal{LC} is the number of edges in \mathcal{LC} . The *depth* of \mathcal{LC} is the length of a longest path from an input node to an output node.

Note that size is essentially counting the number of additions needed to compute $x \mapsto Ax$, so in this model, multiplication by scalars is “free.”

For example, the naïve algorithm for computing a map $A : \mathbb{C}^2 \rightarrow \mathbb{C}^3$ gives rise to the complete bipartite graph as in Figure 1. More generally, the naïve algorithm produces a linear circuit of size $O(nm)$.

If an entry in A is zero, we may delete the corresponding edge as in Figure 2.

Stacking two graphs Γ_1 and Γ_2 on top of each other and identifying the input vertices of Γ_2 with the output vertices of Γ_1 , the matrix of the resulting graph is just the matrix product of the matrices of Γ_1 and Γ_2 . So, if $\text{rank}(A) = 1$, we may write A as a product $A = A_1 A_2$ where $A_1 : \mathbb{C}^2 \rightarrow \mathbb{C}^1$ and $A_2 : \mathbb{C}^1 \rightarrow \mathbb{C}^3$ and concatenate the two complete graphs as in Figure 3.

Given two directed acyclic graphs, Γ_1 and Γ_2 , whose vertex sets are disjoint, with an ordered list of n input nodes and an ordered list of m output nodes each, we define the sum $\Gamma_1 + \Gamma_2$ to

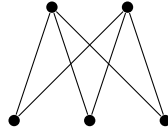


FIGURE 1. naïve linear circuit for $A \in Mat_{2 \times 3}$

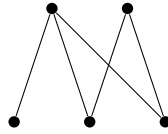


FIGURE 2. linear circuit for $A \in Mat_{2 \times 3}$ with $a_1^2 = 0$

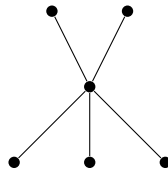


FIGURE 3. linear circuit for rank one $A \in Mat_{2 \times 3}$

be the directed graph resulting from (1) identifying the input nodes of Γ_1 with the input nodes of Γ_2 , (2) doing the same for the output nodes, and (3) summing up their adjacency matrices, see Figure 4 for an example.

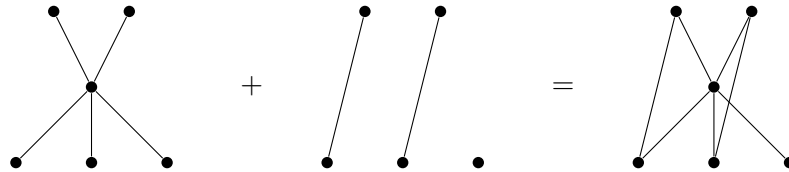


FIGURE 4. The sum of two graphs.

In what follows, for simplicity of discussion, we restrict to the case $n = m$.

With these descriptions in mind, we see rigidity is a measure of the complexity of a restricted depth two circuit computing (1.0.1). Namely the circuit is the sum of two graphs, one of depth one which has s edges and the other is depth two with r vertices at the middle level. The motivation for the restriction to such circuits is Theorem 1.5.1.

1.3. The varieties we study. Define $\hat{\mathcal{R}}[n, r, s] := \overline{\hat{\mathcal{R}}[n, r, s]^0}$, the *variety of matrices of r -border rigidity at most s* , where the overline denotes the common zero set of all polynomials vanishing on $\hat{\mathcal{R}}[n, r, s]^0$, called the *Zariski closure*. This equals the closure of $\hat{\mathcal{R}}[n, r, s]^0$ in the classical topology obtained by taking limits, see [20, Thm 2.33]. If $M \in \hat{\mathcal{R}}[n, r, s]$ we write $\underline{\mathcal{R}ig}_r(M) \leq s$, and say M has *r -border rigidity at most s* . By definition, $\underline{\mathcal{R}ig}_r(M) \leq \mathcal{R}ig_r(M)$. As pointed out in [11], strict inequality can occur. For example, when $s = 1$, one obtains points in the tangent cone as in Proposition 5.1.1(4).

It is generally expected that there are super-linear lower bounds for the size of a linear circuit computing the linear map $x_n \mapsto A_n x_n$ for the following sequences of matrices $A_n = (y_j^i)$, $1 \leq i, j \leq n$, where y_j^i is the entry of A in row i and column j :

Discrete Fourier Transform (DFT) matrix: let ω be a primitive n -th root of unity. Define the size n DFT matrix by $y_j^i = \omega^{(i-1)(j-1)}$.

Cauchy matrix: Let x^i, z_j be variables $1 \leq i, j \leq n$, and define $y_j^i = \frac{1}{x^i + z_j}$. (Here and in the next example, one means super linear lower bounds for a sufficiently general assignment of the variables.)

Vandermonde matrix: Let x_i , $1 \leq i \leq n$, be variables, define $y_j^i := (x_j)^{i-1}$.

Sylvester matrix: $Syl_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $Syl_k = \begin{pmatrix} S_{k-1} & S_{k-1} \\ S_{k-1} & -S_{k-1} \end{pmatrix}$.

We describe algebraic varieties associated to classes of matrices generalizing these examples, describe their ideals and make basic observations about their rigidity.

To each directed acyclic graph Γ with n inputs and outputs, or sums of such, we may associate a variety $\Sigma_\Gamma \subset Mat_n$ consisting of the closure of all matrices A such that (1.0.1) is computable by Γ . For example, to the graph in Figure 5 we associate the variety $\Sigma_\Gamma := \hat{\sigma}_{2,4}$ since any 4×4 matrix of rank at most 2 can be written a product of a 4×2 matrix and a 2×4 matrix.

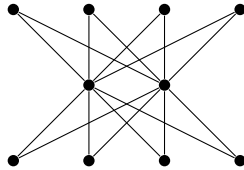


FIGURE 5. linear circuit for rank two $A \in Mat_4$

Note that the number of edges of Γ gives an upper bound of the dimension of Σ_Γ , but the actual dimension is often less, for example $\dim \hat{\sigma}_{2,4} = 12$ but Γ has 16 edges. This is because there are four parameters of choices for expressing a rank two matrix as a sum of two rank one matrices.

1.4. Notation and conventions. Since this article is for geometers and computer scientists, here and throughout, we include a substantial amount of material that is not usually mentioned.

We work exclusively over the complex numbers \mathbb{C} .

For simplicity of exposition, we generally restrict to square matrices, although most results carry over to rectangular matrices as well.

Throughout, V denotes a complex vector space, $\mathbb{P}V$ is the associated projective space of lines through the origin in V , $S^d V^*$ denotes the space of homogenous polynomials of degree d on V , and $Sym(V^*) = \bigoplus_d S^d V^*$ denotes the symmetric algebra, i.e., the ring of polynomials on V , i.e., after a choice of basis, the ring of polynomials in $\dim V$ variables. We work with projective space because the objects of interest are invariant under rescaling and to take advantage of results in projective algebraic geometry, e.g., Proposition 5.3.1. For a subset $Z \subset \mathbb{P}V$, $\hat{Z} := \pi^{-1}(Z) \cup \{0\} \subset V$ is called the *affine cone over Z* .

Let $Z \subset \mathbb{P}V$ be a *projective variety*, the zero set of a collection of homogeneous polynomials on V projected to $\mathbb{P}V$. The ideal of Z , denoted $\mathcal{I}(Z)$, is the ideal in $Sym(V^*)$ of all polynomials vanishing on \hat{Z} . Let $\mathcal{I}_d(Z) \subset S^d V^*$ denote the degree d component of the ideal of Z . The *codimension* of Z is the smallest non-negative integer c such that every linear $\mathbb{P}^c \subset \mathbb{P}V$ intersects Z and its *dimension* is $\dim \mathbb{P}V - c$. The *degree* of Z is the number of points of intersection with a

general linear space of dimension c . Here and throughout, a *general point* or *general linear space* is a point (or linear space) that does not satisfy certain (specific to the problem) polynomials, so the set of general points is of full measure, and one may simply view a general point or linear space as one that has been randomly chosen. A codimension 1 variety is called a *hypersurface* and is defined by a single equation. The degree of a hypersurface is the degree of its defining equation.

For a linear subspace $U \subset V$, its annihilator in the dual space is denoted $U^\perp \subset V^*$, and we abuse notation and write $(\mathbb{P}U)^\perp \subset V^*$ for the annihilator of U as well. The group of invertible endomorphisms of V is denoted $GL(V)$. If $G \subset GL(V)$ is a subgroup and $Z \subset \mathbb{P}V$ is a subvariety such that $g \cdot z \in Z$ for all $z \in Z$ and all $g \in G$, we say Z is a G -variety. The group of permutations on d elements is denoted \mathfrak{S}_d .

We write \log to mean \log_2 .

Let $f, g : \mathbb{R} \rightarrow \mathbb{R}$ be functions. Write $f = \Omega(g)$ (resp. $f = O(g)$) if and only if there exists $C > 0$ and x_0 such that $|f(x)| \geq C|g(x)|$ (resp. $|f(x)| \leq C|g(x)|$) for all $x \geq x_0$. Write $f = \omega(g)$ (resp. $f = o(g)$) if and only if for all $C > 0$ there exists x_0 such that $|f(x)| \geq C|g(x)|$ (resp. $|f(x)| \leq C|g(x)|$) for all $x \geq x_0$. These definitions are used for any ordered range and domain, in particular \mathbb{Z} . In particular, for a function $f(n)$, $f = \omega(1)$ means f goes to infinity as $n \rightarrow \infty$.

For $I, J \subset [n] := \{1, 2, \dots, n\}$ of size $r + 1$, let M_J^I be the determinant of the size $r + 1$ submatrix defined by (I, J) . Set $\Delta_J^I := M_{J^c}^{I^c}$, where I^c and J^c denote the complementary index set to I and J , respectively. We often use x_j^i to denote coordinates on the space of $n \times n$ matrices. Write $\{x_j^i\} := \{x_j^i : i, j \in [n]\}$.

1.5. Previous Work. The starting point is the following theorem of L. Valiant:

Theorem 1.5.1. [27, Thm. 6.1, Prop. 6.2] *Suppose that a sequence $A_n \in Mat_n$ admits a sequence of linear circuits of size $\Sigma = \Sigma(n)$ and depth $d = d(n)$ where each gate has fan-in two. Then for any $t > 1$,*

$$\mathcal{R}ig_{\frac{\Sigma \log(t)}{\log(d)}}(A_n) \leq 2^{O(d/t)} n.$$

In particular, if there exist $\epsilon, \delta > 0$ such that $\mathcal{R}ig_{\epsilon n}(A_n) = \Omega(n^{1+\delta})$, then any sequence of linear circuits of logarithmic (in n) depth computing $\{A_n\}$ must have size $\Omega(n \log(\log n))$.

Proposition 1.5.2. ([6] for finite fields and [24] for the general case) *Let $r \geq (\log n)^2$, and let $A \in Mat_{n \times n}$ be such that all minors of size r of A are nonzero. Then, for all $s < \frac{n^2}{4(r+1)} \log(\frac{n}{r})$, $A \notin \hat{\mathcal{R}}[n, r, s]^0$.*

Note that if one sets $r = \epsilon n$, for n sufficiently large, the above result says $A \notin \hat{\mathcal{R}}[n, \epsilon n, \frac{1}{2} \log(\frac{1}{\epsilon})]^0$ which is far from what would be needed to apply Theorem 1.5.1, as s does not grow with n .

The matrices DFT_p with p prime, general Cauchy matrix, general Vandermonde matrix, general Sylvester matrix are such that all minors of all sizes are nonzero (see [18, §2.2] and the references therein). Thus Proposition 1.5.2 implies:

Corollary 1.5.3. [6, 24, 10] *The matrices of the following types: DFT_p with p prime, Cauchy, Vandermonde, and Sylvester, are such that for all $s < \frac{n^2}{4(r+1)} \log(\frac{n}{r})$, $A \notin \hat{\mathcal{R}}[n, r, s]^0$.*

The following theorem is proved via a theorem in graph theory from [22]:

Theorem 1.5.4. (attributed to Strassen in [27], also see [18, §2.2]) *For all $\epsilon > 0$, there exist $n \times n$ matrices A with integer entries, all of whose minors of all sizes are nonzero such that $A \in \hat{\mathcal{R}}[n, \epsilon n, n^{1+o(1)}]^0$.*

In [11], they approach the rigidity problem from the perspective of algebraic geometry. In particular, they use the effective Nullstellensatz to obtain bounds on the degrees of the hypersurfaces of maximally border rigid matrices. They show the following.

Theorem 1.5.5. [11, Thm. 7] *Let $p_{k,j} > n^{4n^2}$ be distinct primes for $1 \leq k, j \leq n$. Let A_n have entries $a_j^k = e^{2\pi i/p_{k,j}}$. Then A_n is maximally r -border rigid for all $1 \leq r \leq n - 2$.*

See Remark 1.6.6 for a small improvement of this result.

In [11], they do not restrict their field to be \mathbb{C} .

Additional references for matrix rigidity are [3, 16, 17, 4, 5, 23, 15, 19].

1.6. Our results. Previous to our work, to our knowledge, there were no explicit equations for irreducible components of $\hat{\mathcal{R}}[n, r, s]$ known other than the minors of size $r + 1$. The irreducible components of $\hat{\mathcal{R}}[n, r, s]$ are determined (non-uniquely) by cardinality s subsets $S \subset \{x_j^i\}$ corresponding to the entries one is allowed to change. Recall that x_j^i are coordinates on the space of $n \times n$ matrices. We find equations for the cases $r = 1$, (Proposition 3.2.5), $r = n - 2$ (Theorem 3.4.1), and the cases $s = 1, 2, 3$ (see §3.1). We also obtain qualitative information about the equations. Here are some sample results:

Proposition 1.6.1. *Each irreducible component of $\hat{\mathcal{R}}[n, r, s]$, described by some set $S \subset \{x_j^i, 1 \leq i, j \leq n\}$, has ideal generated by polynomials with the following property; if P is a generator of degree d , then no entries of S appear in P and P is a sum of terms of the form ΔQ where Δ is a minor of size $r + 1$ and $\deg(Q) = d - r - 1$. In particular, there are no equations of degree less than $r + 1$ in the ideal.*

Conversely any polynomial P of degree d such that no entries of S appear in P and P is a sum of terms ΔQ where Δ is a minor of size $r + 1$ is in the ideal of the component of $\hat{\mathcal{R}}[n, r, s]$ determined by S .

See §2.2 for more precise statements. These results are consequences of more general results about cones in §5.1.

We remind the reader that Δ_J^I is the determinant of the submatrix obtained by deleting the rows of I and the columns of J .

Theorem 1.6.2. *There are two types of components of the hypersurface $\hat{\mathcal{R}}[n, n - 2, 3]$:*

- (1) *Those corresponding to a configuration S where the three entries are all in distinct rows and columns, where if $S = \{x_{j_1}^{i_1}, x_{j_2}^{i_2}, x_{j_3}^{i_3}\}$ the hypersurface is of degree $2n - 3$ with equation*

$$\Delta_{j_2}^{i_3} \Delta_{j_1, j_3}^{i_1, i_2} - \Delta_{j_3}^{i_2} \Delta_{j_1, j_2}^{i_1, i_3} = 0.$$

- (2) *Those corresponding to a configuration where there are two elements of S in the same row and one in a different column from those two, or such that one element shares a row with one and a column with the other. In these cases, the equation is the unique size $(n - 1)$ minor that has no elements of S .*

If all three elements of S lie on a row or column, then one does not obtain a hypersurface.

We give numerous examples of equations in other special cases in §3. Our main tool for finding these equations are the results presented in §2.2, which follow from more general results regarding joins of projective varieties that we prove in §5.2.

If one holds not just s fixed, but moreover fixes the specific entries of the matrix that one is allowed to change, and allows the matrix to grow (i.e., the subset S is required to be contained in some $n_0 \times n_0$ submatrix of $A \in \text{Mat}_n$), there is a *propagation* result (Proposition 2.2.5), that enables one to deduce the equations in the $n \times n$ case from the $n_0 \times n_0$ case.

When one takes a cone (in the sense of Definition 2.1.1 below, not to be confused with the affine cone over a variety) over a variety with respect to a general linear space, there is a dramatic increase in the degree of the generators of the ideal because the equations of the cone are obtained using elimination theory. For example, a general cone over a codimension two complete intersection, whose ideal is generated in degrees d_1, d_2 will have defining equation in degree $d_1 d_2$. However, we are taking cones over very singular points of varieties that initially are not complete intersections, so the increase in degree is significantly less. We conjecture:

Conjecture 1.6.3. *Fix $0 < \epsilon < 1$ and $0 < \delta < 1$. Set $r = \epsilon n$ and $s = n^{1+\delta}$. Then the minimal degree of a polynomial in the ideal of each irreducible component of $\hat{\mathcal{R}}[n, r, s]$ grows like a polynomial in n .*

Although it would not immediately solve Valiant's problem, an affirmative answer to Conjecture 1.6.3 would drastically simplify the study.

While it is difficult to get direct information about the degrees of defining equations of the irreducible components of $\hat{\mathcal{R}}[n, r, s]$, as naively one needs to use elimination theory, one can use general results from algebraic geometry to get information about the degrees of the varieties.

Let $d_{n,r,s}$ denote the maximum degree of an irreducible component of $\hat{\mathcal{R}}[n, r, s]$. It will be useful to set $k = n - r$. Then (see e.g., [2, p95] for the first equality and e.g. [8, p. 50,78] for the fourth and fifth)

$$(1.6.4) \quad \begin{aligned} d_{n,r,0} &= \prod_{i=0}^{n-r-1} \frac{(n+i)!i!}{(r+i)!(n-r+i)!} \\ &= \frac{B(r)B(2n-r)B(n-r)^2}{B(n)^2 B(2n-2r)} \\ &= \frac{B(n-k)B(n+k)B(k)^2}{B(n)^2 B(2k)} \\ &= \dim S_{k^k} \mathbb{C}^n \end{aligned}$$

$$(1.6.5) \quad = \frac{\dim[k^k] B(n-k)B(n+k)}{k^2! B(n)^2}$$

Here $B(k) := G(k+1)$, where $G(m) = \prod_{i=1}^{m-2} i!$ is the Barnes G -function, $S_{k^k} \mathbb{C}^n$ denotes the irreducible GL_n -representation of type (k, k, \dots, k) , and $[k^k]$ denotes the irreducible \mathfrak{S}_{k^2} -module corresponding to the partition (k, \dots, k) .

Remark 1.6.6. The shifted Barnes G -function B has the following asymptotic expansion

$$B(z) = \left(\frac{z}{e^{\frac{3}{2}}} \right)^{\frac{z^2}{2}} O(2.51^z)$$

(see e.g. en.wikipedia.org/wiki/Barnes_G-function). Since the degree of a variety cannot increase when taking a cone over it, one can replace the n^{4n^2} upper bound for $\Delta(n)$ in Theorem 1.5.5 with roughly $n^{\epsilon n^2}$ because, setting $r = \epsilon n$, for some constant C ,

$$\frac{B(\epsilon n)B((2-\epsilon)n)B((1-\epsilon)n)^2}{B(n)^2 B(2(1-\epsilon)n)} \leq n^{n^2[\frac{\epsilon^2}{2} + \frac{(2-\epsilon)^2}{2} + (1-\epsilon)^2 - 1 - 2(1-\epsilon)^2]} C^{n^2+n} = n^{\epsilon n^2} C^{n^2+n}$$

Remark 1.6.7. A geometric interpretation of the equality between $\deg d_{n,r,0}$ and the dimension of an irreducible GL_n -module is discussed in [26].

We prove several results about the degrees $d_{n,r,s}$. For example:

Theorem 1.6.8. *Let $s \leq n$, Then,*

$$(1.6.9) \quad d_{n,r,s} \leq d_{n,r,0} - \sum_{j=1}^s d_{n-1,r-1,s-j}$$

In an earlier version of this paper, we conjectured that equality held in (1.6.9). After we submitted the paper for publication, our conjecture was answered affirmatively in [1]:

Theorem 1.6.10. [1] *Let $s \leq n$, Then,*

$$(1.6.11) \quad d_{n,r,s} = d_{n,r,0} - \sum_{j=1}^s d_{n-1,r-1,s-j}$$

In the previous version of this paper, the following theorem was stated with the hypothesis that equality holds in (1.6.9) for all $(r', n', s') \leq (r, n, s)$ and $s \leq n$. Theorem 1.6.10 renders it to the present unconditional form:

Theorem 1.6.12. *Each irreducible component of $\hat{\mathcal{R}}[n, n-k, s]$ has degree at most*

$$(1.6.13) \quad \sum_{m=0}^s \binom{s}{m} (-1)^m d_{r-m, n-m, 0}$$

with equality holding if no two elements of S lie in the same row or column, e.g., if the elements of S appear on the diagonal.

Moreover, if we set $r = n - k$ and $s = k^2 - u$ and consider the degree $D(n, k, u)$ as a function of n, k, u , then, fixing k, u and considering $D_{k,u}(n) = D(n, k, u)$ as a function of n , it is of the form

$$D_{k,u}(n) = (k^2)! \frac{B(k)^2}{B(2k)} p(n)$$

where $p(n) = \frac{n^u}{u!} - \frac{k^2-u}{2(u-1)!} n^{u-1} + O(n^{u-2})$ is a polynomial of degree u .

For example, when $u = 1$, $D(n, k, 1) = (k^2)! \frac{B(k)^2}{B(2k)} (n - \frac{1}{2}(k^2 - 1))$.

Remark 1.6.14. Note that $D_{k,u}(n) = \dim[k^k]p(n)$. It would be nice to have a geometric or representation-theoretic explanation of this equality.

Remark 1.6.15. In our earlier version of this paper, we realized that the use of *intersection theory* (see, e.g. [7]), could render Theorem 1.6.8 unconditional, so we contacted P. Aluffi, an expert in the subject. Not only was he able to render the theorem unconditional, but he determined the degrees in additional cases. We are delighted that such beautiful geometry can be of use to computer science, and look forward to further progress on these questions. We expect a substantial reduction in degree when $r = \epsilon n$ and $s = (n - r)^2 - 1$.

We define varieties modeled on different classes of families of matrices as mentioned above. We show that a general Cauchy matrix, or a general Vandermonde matrix is maximally 1-rigid and maximally $(n - 2)$ -rigid (Propositions 4.2.3 and 4.3.2). One way to understand the DFT algorithm is to factor the discrete Fourier transform matrix as a product (set $n = 2^k$) $DFT_{2^k} = S_1 \cdots S_k$ where each S_k has only $2n$ nonzero entries. Then, these sparse matrices can all be multiplied via a linear circuit of size $2n \log n$ (and depth $\log n$). We define the variety of *factorizable* or *butterfly* matrices FM_n to be the closure of the set of matrices admitting such a description as a product of sparse matrices, all of which admit a linear circuit of size $2n \log n$, and show (Proposition 4.6.1):

Proposition 1.6.16. *A general butterfly matrix admits a linear circuit of size $2n \log n$, but does not admit a linear circuit of size $n(\log n + 1) - 1$.*

1.7. Future work. Proposition 1.6.1 gives qualitative information about the ideals and we give numerous examples of equations for the relevant varieties. It would be useful to continue the study of the equations both qualitatively and by computing further explicit examples, with the hope of eventually getting equations in the Valiant range. In a different direction, an analysis of the degrees of the hypersurface cases in the range $r = \epsilon n$ could lead to a substantial reduction of the known degree bounds.

Independent of complexity theory, several interesting questions relating the differential geometry and scheme structure of tangent cones are posed in §5.

Acknowledgements. We thank the anonymous referees for very careful reading and numerous useful suggestions.

2. GEOMETRIC FORMULATION

2.1. Border rigidity.

Definition 2.1.1. For varieties $X, Y \subset \mathbb{P}V$, let

$$\mathbf{J}^0(X, Y) := \bigcup_{x \in X, y \in Y, x \neq y} \langle x, y \rangle$$

and define the *join of X and Y* as $\mathbf{J}(X, Y) := \overline{\mathbf{J}^0(X, Y)}$ with closure using either the Zariski or the classical topology. Here, $\langle x, y \rangle$ is the (projective) linear span of the points x, y . If $Y = L$ is a linear space $\mathbf{J}(X, L)$ is called the *cone* over X with *vertex* L . (Algebraic geometers refer to L as the vertex even when it is not just a point \mathbb{P}^0 .)

Let $\sigma_r = \sigma_{r,n} = \sigma_r(\text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})) \subset \mathbb{P}(\mathbb{C}^n \otimes \mathbb{C}^n)$ denote the variety of up to scale $n \times n$ matrices of rank at most r , called the *r -th secant variety of the Segre variety*. For those not familiar with this variety, the Segre variety itself, $\text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})$, is the projectivization of the rank one matrices, one may think of the first \mathbb{P}^{n-1} as parametrizing column vectors and the second as parametrizing row vectors and the corresponding (up to scale) rank one matrix as their product. The rank at most r -matrices are those appearing in some secant \mathbb{P}^{r-1} to the Segre variety.

Let $\mathbb{C}^n \otimes \mathbb{C}^n$ be furnished with linear coordinates x_j^i , $1 \leq i, j \leq n$. Let $S \subset \{x_j^i\}$ be a subset of cardinality s and let $L^S := \text{span } S$. We may rephrase (1.1.1) as

$$\hat{\mathcal{R}}[n, r, s]^0 = \bigcup_{S \subset \{x_j^i\}, |S|=s} \hat{\mathbf{J}}^0(\sigma_r, L^S)$$

The dimension of σ_r is $r(2n - r) - 1$ and $\dim \hat{\mathbf{J}}(\sigma_r, L^S) \leq \min\{r(2n - r) + s, n^2\}$ (see Proposition 5.1.1(3)). We say the dimension is the *expected dimension* if equality holds.

The variety $\hat{\mathcal{R}}[n, r, s]$, as long as $s > 0$ and it is not the ambient space, is reducible, with at most $\binom{n^2}{s}$ components, all of the same dimension $r(2n - r) + s$. (This had been observed in [11, Thm. 3.8] and [13].) To see the equidimensionality, notice that if $|S_j| = j$ and $S_j \subseteq S_{j+1}$, then the sequence of joins $\mathbf{J}_j = \mathbf{J}(\sigma_r, L^{S_j})$ eventually fills the ambient space. Moreover, $\dim \mathbf{J}_{j+1} \leq \dim \mathbf{J}_j + 1$, so the only possibilities are $\mathbf{J}_{j+1} = \mathbf{J}_j$ or $\dim \mathbf{J}_{j+1} = \dim \mathbf{J}_j + 1$. In particular, this shows that for any j there exists a suitable choice of S_j such that \mathbf{J}_j has the expected dimension. Now, suppose that $\mathbf{J}(\sigma_r, L^S)$ does not have the expected dimension, so its dimension is $r(2n - r) + s'$ for some $s' < s$. Let $S' \subseteq S$ be such that $|S'| = s'$ and $\mathbf{J}(\sigma_r, L^{S'})$

has the expected dimension. Then $\mathbf{J}(\sigma_r, L^{S'}) = \mathbf{J}(\sigma_r, L^S)$. Now let R be such that $|R| = s$, $S' \subseteq R$ and $\mathbf{J}(\sigma_r, L^R)$ has the expected dimension. Then $\mathbf{J}(\sigma_r, L^R)$ is an irreducible component of $\hat{\mathcal{R}}[n, r, s]$ that contains $\mathbf{J}(\sigma_r, L^S)$, showing that the irreducible components of $\hat{\mathcal{R}}[n, r, s]$ have dimension equal to the expected dimension of $\mathbf{J}(\sigma_r, L^S)$. Therefore,

$$(2.1.2) \quad \dim \hat{\mathcal{R}}[n, r, s] = \min\{r(2n - r) + s, n^2\},$$

and $\hat{\mathcal{R}}[n, r, s]$ is a hypersurface if and only if

$$(2.1.3) \quad s = (n - r)^2 - 1.$$

We say a matrix M is *maximally r -border rigid* if $M \notin \hat{\mathcal{R}}[n, r, (n - r)^2 - 1]$, and that M is *maximally border rigid* if $M \notin \hat{\mathcal{R}}[n, r, (n - r)^2 - 1]$ for all $r = 1, \dots, n - 2$. Throughout we assume $r \leq n - 2$ to avoid trivialities.

The set of maximally rigid matrices is of full measure (in any reasonable measure e.g. absolutely continuous with respect to Lebesgue measure) on the space of $n \times n$ matrices. In particular, a “random” matrix will be maximally rigid.

2.2. On the ideal of $\mathbf{J}(\sigma_r, L^S)$. Write $S^c = \{x_j^i\} \setminus S$ for the complement of S . The following is a consequence of Proposition 5.2.1:

Proposition 2.2.1. *Fix $L = L^S$. Generators for the ideal of $\mathbf{J}(\sigma_r, L)$ may be obtained from polynomials of the form $P = \sum_{I, J} q_I^J M_J^I$, where the q_I^J are arbitrary homogeneous polynomials all of the same degree and:*

- (1) M_J^I is the (determinant of the) size $r + 1$ minor defined by the index sets I, J (i.e., $I, J \subset [n]$, $|I| = |J| = r + 1$), and
- (2) only the variables of S^c appear in P .

Conversely, any polynomial of the form $P = \sum_{I, J} q_I^J M_J^I$, where the M_J^I are minors of size $r + 1$ and only the variables of S^c appear in P , is in $\mathcal{I}(\mathbf{J}(\sigma_r, L))$.

Let $E, F = \mathbb{C}^n$. The irreducible polynomial representations of $GL(E)$ are indexed by partitions π with at most $\dim E$ parts. Let $\ell(\pi)$ denote the number of parts of π , and let $S_\pi E$ denote the irreducible $GL(E)$ -module corresponding to π . We have the $GL(E) \times GL(F)$ -decomposition

$$S^d(E \otimes F) = \bigoplus_{|\pi|=d, \ell(\pi) \leq n} S_\pi E \otimes S_\pi F.$$

Let $T_E \subset GL(E)$ denote the torus (the invertible diagonal matrices). A vector $e \in E$ is said to be a *weight vector* if $[t \cdot e] = [e]$ for all $t \in T_E$.

Proposition 2.2.2. *Write $Mat_n = E \otimes F$. For all $S \subset \{x_j^i\}$, $\mathbf{J}(\sigma_r, L^S)$ is a $T_E \times T_F$ -variety.*

Thus a set of generators of $\mathcal{I}(\mathbf{J}(\sigma_r, L))$ may be taken from $GL(E) \times GL(F)$ -weight vectors and these weight vectors must be sums of vectors in modules $S_\pi E \otimes S_\pi F$ where $\ell(\pi) \geq r + 1$.

The length requirement follows from Proposition 2.2.1(1).

Proposition 1.6.1(1) is Proposition 2.2.2 expressed in coordinates. For many examples, the generators have nonzero projections onto all the modules $S_\pi E \otimes S_\pi F$ with $\ell(\pi) \geq r + 1$.

Recall the notation $\Delta_J^I = M_{J^c}^{I^c}$, where I^c denotes the complementary index set to I . This will allow us to work independently of the size of our matrices.

Let $S \subset \{x_j^i\}_{1 \leq i, j \leq n}$ and let $P \in \mathcal{I}_d(\sigma_r) \cap S^d(L^S)^\perp$, so $P \in \mathcal{I}_d(\mathbf{J}(\sigma_r, L^S))$, and require further that P be a $T_E \times T_F$ weight vector. Write

$$(2.2.3) \quad P = \sum_v \Delta_{J_1^v}^{I_1^v} \cdots \Delta_{J_f^v}^{I_f^v}$$

where $|I_\alpha^1| = |J_\alpha^1| =: \delta_\alpha$, so $d = fn - \sum_\alpha \delta_\alpha$. Since we allow $\delta_\alpha = 1$, any polynomial that is a weight vector may be written in this way. Write $x_j^i = \bar{e}_i \otimes \bar{f}_j$. The T_E -weight of P is $(1^{\lambda_1}, \dots, n^{\lambda_n})$ where \bar{e}_j appears λ_j times in the union of the $(I_\alpha^v)^{c_s}$ s, and the T_F -weight of P is $(1^{\mu_1}, \dots, n^{\mu_n})$ where \bar{f}_j appears μ_j times in the union of the $(J_\alpha^v)^{c_s}$ s.

Define

$$(2.2.4) \quad P_q \in S^{d+qf} \text{Mat}_{n+q}^*$$

by (2.2.3) only considered as a polynomial on Mat_{n+q} , and note that $P = P_0$.

Proposition 2.2.5. *For $P \in \mathcal{I}_d(\sigma_{r,n}) \cap S^d(L^S)^\perp$ as in (2.2.3),*

$$P_q \in \mathcal{I}_{d+fq}(\sigma_{r,n+q}) \cap S^{d+fq}(L^S)^\perp$$

where S is the same for Mat_n and Mat_{n+q} . In particular, $P_q \in \mathcal{I}_{d+fq}(\mathbf{J}(\sigma_{r+q,n+q}, L^S))$.

Proof. It is clear $P_q \in \mathcal{I}_{d+fq}(\sigma_{r,n+q})$, so it remains to show it is in $S^{d+fq}(L^S)^\perp$. By induction, it will be sufficient to prove the case $q = 1$. Say in some term, say $v = 1$, in the summation of P in (2.2.3) a monomial in S appears as a factor, some $x_{t_1}^{s_1} \cdots x_{t_g}^{s_g} Q$. Then, by Laplace expansions, we may write $Q = \tilde{\Delta}_{J_1^1}^{I_1^1} \cdots \tilde{\Delta}_{J_f^1}^{I_f^1}$, for some minors (smaller than or equal to the originals). Since this term is erased we must have, after re-ordering terms, for $v = 2, \dots, h$ (for some h),

$$x_{t_1}^{s_1} \cdots x_{t_g}^{s_g} (\tilde{\Delta}_{J_1^1}^{I_1^1} \cdots \tilde{\Delta}_{J_f^1}^{I_f^1} + \cdots + \tilde{\Delta}_{J_1^h}^{I_1^h} \cdots \tilde{\Delta}_{J_f^h}^{I_f^h}) = 0$$

that is,

$$(2.2.6) \quad \tilde{\Delta}_{J_1^1}^{I_1^1} \cdots \tilde{\Delta}_{J_f^1}^{I_f^1} + \cdots + \tilde{\Delta}_{J_1^h}^{I_1^h} \cdots \tilde{\Delta}_{J_f^h}^{I_f^h} = 0.$$

Now consider the same monomial's appearance in P_1 (only the monomials of S appearing in the summands of P could possibly appear in P_1). In the $v = 1$ term it will appear with \tilde{Q} where \tilde{Q} is a sum of terms, the first of which is $(x_{n+1}^{n+1})^f \tilde{\Delta}_{J_1^1, n+1}^{I_1^1, n+1} \cdots \tilde{\Delta}_{J_f^1, n+1}^{I_f^1, n+1}$ and each appearance will have such a term, so these add to zero because $\tilde{\Delta}_{J_1^1, n+1}^{I_1^1, n+1}$ in Mat_{n+1} , is the same minor as $\tilde{\Delta}_{J_1^1}^{I_1^1}$ in Mat_n . Next is a term say $(x_{n+1}^{n+1})^{f-1} x_{n+1}^1 \tilde{\Delta}_{J_1^1, n+1}^{I_1^1, n+1} \cdots \tilde{\Delta}_{J_f^1, n+1}^{I_f^1, n+1}$, but then there must be corresponding terms $(x_{n+1}^{n+1})^{f-1} x_{n+1}^1 \tilde{\Delta}_{J_1^\mu, n+1}^{I_1^\mu, n+1} \cdots \tilde{\Delta}_{J_f^\mu, n+1}^{I_f^\mu, n+1}$ for each $2 \leq \mu \leq h$. But these must also sum to zero because it is an identity among minors of the same form as the original. One continues in this fashion to show all terms in S in the expression of P_1 indeed cancel. \square

Corollary 2.2.7. *Fix $k = n - r$ and S with $|S| = k^2 - 1$, and allow n to grow. Then the degrees of the hypersurfaces $\mathbf{J}(\sigma_{n-k,n}, L^S)$ grow at most linearly with respect to n .*

Proof. If we are in the hypersurface case and $P \in \mathcal{I}_d(\mathbf{J}(\sigma_{r,n}, L^S))$, then even in the worst possible case where all factors $\tilde{\Delta}_{J_s^v}^{I_s^v}$ in the expression (2.2.3) but the first have degree one, the ideal of the hypersurface $\mathbf{J}(\sigma_{r+u,n+u}, L^S)$ is nonempty in degree $(d - r)u$. \square

Definition 2.2.8. Let P be a generator of $\mathcal{I}(\mathbf{J}(\sigma_{r,n}, L^S))$ with a presentation of the form (2.2.3). We say P is *well presented* if P_q constructed as in (2.2.4) is a generator of $\mathcal{I}(\mathbf{J}(\sigma_{r+q,n+q}, L^S))$ for all q .

Conjecture 2.2.9. *For all r, n, S , there exists a set of generators P^1, \dots, P^μ of $\mathcal{I}(\mathbf{J}(\sigma_{r,n}, L^S))$ that can be well presented.*

Remark 2.2.10. Well presented expressions are far from unique because of the various Laplace expansions.

Remark 2.2.11. $\mathcal{I}(\mathbf{J}(\sigma_{r+q,n+q}, L^S))$ may require additional generators beyond the P_q^1, \dots, P_q^μ .

3. EXAMPLES OF EQUATIONS FOR $\mathbf{J}(\sigma_r, L^S)$

3.1. First examples. The simplest equations for $\mathbf{J}(\sigma_r, L^S)$ occur when S omits a submatrix of size $r + 1$, and one simply takes the corresponding size $r + 1$ minor. The proofs of the following propositions are immediate consequences of Proposition 2.2.1, as when one expands each expression, the elements of S cancel.

Consider the example $n = 3$, $r = 1$, $S = \{x_1^1, x_2^2, x_3^3\}$, and $r = 1$. Then

$$(3.1.1) \quad x_1^2 x_2^3 x_3^1 - x_2^1 x_3^2 x_1^3 = M_{12}^{23} x_3^1 - M_{23}^{12} x_1^3 \in \mathcal{I}_3(\mathbf{J}(\sigma_1, L^S)).$$

This example generalizes in the following two ways. First, Proposition 2.2.5 implies:

Proposition 3.1.2. *If there are two size $r + 1$ submatrices of Mat_n , say respectively indexed by (I, J) and (K, L) , that each contain some $x_{j_0}^{i_0} \in S$ but no other point of S , then setting $I' = I \setminus i_0$, $J' = J \setminus j_0$, $K' = K \setminus i_0$, $L' = L \setminus j_0$, the degree $2r + 1$ equations*

$$(3.1.3) \quad M_J^I M_{L'}^{K'} - M_L^K M_{J'}^{I'}$$

are in the ideal of $J(\sigma_r, L^S)$.

By Proposition 2.2.5, (3.1.1) also generalizes to:

Proposition 3.1.4. *Suppose that there exists two size $r + 2$ submatrices of S , indexed by $(I, J), (K, L)$, such that*

- (1) *there are only three elements of S appearing in them, say $x_{j_1}^{i_1}, x_{j_2}^{i_2}, x_{j_3}^{i_3}$ with both i_1, i_2, i_3 and j_1, j_2, j_3 distinct, and*
- (2) *each element appears in exactly two of the minors.*

Then the degree $2r + 1$ equations

$$(3.1.5) \quad M_{J \setminus j_1}^{I \setminus i_1} M_{L \setminus i_2, i_3}^{K \setminus i_2, i_3} - M_{J \setminus j_2}^{I \setminus i_2} M_{L \setminus i_1, i_3}^{K \setminus i_1, i_3}$$

are in the ideal of $J(\sigma_r, L^S)$.

For example, when $S = \{x_1^1, x_2^2, x_3^3\}$, equation (3.1.5) may be written

$$(3.1.6) \quad \Delta_2^3 \Delta_{13}^{12} - \Delta_3^2 \Delta_{12}^{13}.$$

Now consider the case $n = 4$, $r = 1$ and $S = \{x_3^1, x_4^1, x_1^2, x_4^2, x_1^3, x_2^3, x_2^4, x_3^4\}$. Proposition 3.1.4 cannot be applied. Instead we have the equation

$$x_1^1 x_2^2 x_3^3 x_4^4 - x_2^1 x_3^2 x_4^3 x_1^4 = M_{12}^{12} x_3^3 x_4^4 + M_{13}^{23} x_2^1 x_4^4 + M_{14}^{34} x_3^2 x_2^1.$$

This case generalizes to

Proposition 3.1.7. *If there are three size $r + 1$ submatrices of $Mat_{n \times n}$, indexed by $(I, J), (K, L), (P, Q)$, such that*

- (1) *the first two contain one element of S each, say the elements are $x_{j_1}^{i_1}$ for (I, J) and $x_{j_1}^{i_2}$ for (K, L) ,*
- (2) *these two elements lie in the same column (or row), and*
- (3) *the third submatrix contains $x_{j_1}^{i_1}, x_{j_1}^{i_2}$ and no other element of S ,*

then the degree $3r + 1$ equations

$$(3.1.8) \quad M_J^I M_{L'}^{K'} M_{Q'}^{P'} + M_L^K M_{J'}^{I'} M_{Q''}^{P''} + M_Q^P M_{J'}^{I'} M_{L'}^{K'},$$

where $I' = I \setminus i_0$, $J' = J \setminus j_0$, $K' = K \setminus i_0$, $L' = L \setminus j_0$, $P' = P \setminus i_1$, $Q' = Q \setminus j_1$, and $P'' = P \setminus i_2$, are in the ideal of $J(\sigma_r, L^S)$.

Lemma 3.1.9. *Let $S' \subsetneq S$ with $S' = \{x_1^1, \dots, x_1^{n-r}\}$ and $S = S' \cup \{x_1^{n-r+1}, \dots, x_1^n\}$. Then $\mathbf{J}(\sigma_r, L^S) = \mathbf{J}(\sigma_r, L^{S'})$.*

Proof. Clearly $\mathbf{J}(\sigma_r, L^S) \supseteq \mathbf{J}(\sigma_r, L^{S'})$. To prove the other inclusion, let $A = (a_j^i) \in \hat{\mathbf{J}}(\sigma_r, L^S)$ be general. Let \tilde{A} be the $r \times r$ submatrix of A given by the last r rows and the last r columns. By generality assumptions, \tilde{A} is non-singular. Therefore, there exist $c_{n-r+1}, \dots, c_n \in \mathbb{C}$ such that

$$\begin{pmatrix} a_1^{n-r+1} \\ \vdots \\ a_1^n \end{pmatrix} = \sum_{j=n-r+1}^n c_j \begin{pmatrix} a_j^{n-r+1} \\ \vdots \\ a_j^n \end{pmatrix}.$$

Let $B = (b_j^i)$ be an $n \times n$ matrix such that $b_j^i = a_j^i$ if $j \geq 2$ and $b_1^i = \sum_{j=n-r+1}^n c_j b_j^i$. Then $B \in \hat{\sigma}_r$ and $A \in \hat{\mathbf{J}}([B], L^{S'}) \subset \hat{\mathbf{J}}(\sigma_r, L^{S'})$. \square

It will be useful to represent various S pictorially. We will use black diamonds \blacklozenge for entries in S and white diamonds \lozenge for entries omitted by S . For example, $S = \{x_1^1, x_2^2, \dots, x_5^5\}$ is represented by

$$\begin{pmatrix} \blacklozenge & & & & \\ & \blacklozenge & & & \\ & & \blacklozenge & & \\ & & & \blacklozenge & \\ & & & & \blacklozenge \end{pmatrix}$$

while $S = \{x_j^i\}_{i,j} \setminus \{x_1^1, x_2^2, \dots, x_5^5\}$ is represented by

$$\begin{pmatrix} \lozenge & & & & \\ & \lozenge & & & \\ & & \lozenge & & \\ & & & \lozenge & \\ & & & & \lozenge \end{pmatrix}.$$

3.2. Case $r = 1$.

Lemma 3.2.1. *Let S be a configuration omitting $x_1^1, \dots, x_k^k, x_2^1, \dots, x_k^{k-1}$ and x_1^k for some $k \geq 2$. Then $\mathcal{I}_k(\mathbf{J}(\sigma_1, L^S))$ contains the binomial*

$$x_1^1 \cdots x_k^k - x_2^1 \cdots x_1^k$$

Moreover, if the complement $S^c = \{x_1^1, \dots, x_k^k, x_2^1, \dots, x_k^{k-1}, x_1^k\}$, then $\mathbf{J}(\sigma_1, L^S)$ is a hypersurface.

Proof. Let $f = x_1^1 \cdots x_k^k - x_2^1 \cdots x_1^k$. It suffices to show that $f \in \mathcal{I}(\sigma_1)$, namely that it can be generated by 2×2 minors. If $k = 2$, then f is the 2×2 minor M_{12}^{12} . Suppose $k \geq 3$.

Define $f_2 = x_1^1 x_2^2 - x_2^1 x_1^2 = M_{12}^{12}$. For any $j = 3, \dots, k$ define

$$f_j = x_j^j f_{j-1} - x_2^1 \cdots x_{j-1}^{j-2} M_{1,j}^{j-1,j}.$$

Thus $f_j = x_1^1 \cdots x_j^j - x_2^1 \cdots x_1^j \in \mathcal{I}(\sigma_1)$ for all $j = 3, \dots, k$ and $f_k = f$.

The last assertion follows because for any $S' \supseteq S$, iterated applications of Lemma 3.1.9 implies $\mathbf{J}(\sigma_1, L^{S'}) = \text{Mat}_{n \times n}$. \square

Lemma 3.2.2. *Let S be a configuration omitting at least two entries in each row and in each column. Then there exists $k \geq 2$ such that, up to a permutation of rows and columns, $S^c \supseteq \{x_1^1, \dots, x_k^k, x_2^1, \dots, x_1^k\}$.*

Proof. After a permutation, we may assume $x_1^1 \in S^c$, and, since S omits at least another entry in the first column, $x_1^2 \in S^c$. Since S omits at least 2 entries in the second row, assume $x_2^2 \in S^c$. S omits at least another entry in the second column: if that entry is x_2^1 , then $k = 2$ and S omits a 2×2 minor; otherwise we may assume $x_2^3 \in S^c$. Again S omits another entry on the third row: if that entry is x_1^3 (resp. x_2^3), then $k = 3$ (resp. $k = 2$) and S omits a set of the desired form. After at most $2n$ steps, this procedure terminates, giving a $k \times k$ submatrix K with one of the following configurations, one the transpose of the other:

$$\begin{bmatrix} \diamond & \diamond & & & \\ & \ddots & \ddots & & \\ & & \ddots & \diamond & \\ \diamond & & & & \diamond \end{bmatrix}, \quad \begin{bmatrix} \diamond & & & & \diamond \\ \diamond & \ddots & & & \\ & \ddots & \ddots & & \\ & & \ddots & \diamond & \\ & & & & \diamond \end{bmatrix}.$$

K and its transpose are equivalent under permutations of rows and columns because $K^T = PKP$ where P is the $k \times k$ permutation matrix having 1 on the anti-diagonal and 0 elsewhere. \square

Lemma 3.2.3. *Let S be a configuration of $n^2 - 2n$ entries. Then there exist $k \in [n]$ and a $k \times k$ submatrix K such that, up to a permutation of rows and columns, at least $2k$ entries of the complement S^c of S lie in K in the following configuration*

$$(3.2.4) \quad \left(\begin{array}{cccc|cccc} \diamond & \diamond & & & & & & \\ & \ddots & \ddots & & & & & \\ & & \ddots & \diamond & & & & \\ \diamond & & & & \diamond & & & \\ \hline & & & & & & & \end{array} \right).$$

Moreover, if $\mathbf{J}(\sigma_1, L^S)$ is a hypersurface then these are the only omitted entries in K and the ideal of $\mathbf{J}(\sigma_1, L^S)$ is generated by $(x_1^1 \cdots x_k^k - x_2^1 \cdots x_1^k)$.

Proof. To prove the first assertion, we proceed by induction on n . The case $n = 2$ provides $s = 0$ and $k = 2$ trivially satisfies the statement.

If S omits at least (and therefore exactly) 2 entries in each row and in each column, then we conclude by Lemma 3.2.2.

Suppose that S contains an entire row (or an entire column). Then S^c is concentrated in a $(n - 1) \times n$ submatrix. In this case we may consider an $(n - 1) \times (n - 1)$ submatrix obtained by removing a column that contains at most 2 entries of S^c . Thus, up to reduction to a smaller matrix, we may always assume that S omits at least one entry in every row (or at least one entry in every column).

After a permutation, we may assume that the first row omits x_1^1 . If the first column omits at most one more entry, then S omits at least $2(n - 1)$ entries in the the submatrix obtained by removing the first row and the first column. We conclude by induction that there exists a k and a $k \times k$ submatrix with the desired configuration in the submatrix.

If the first column omits at least 2 entries other than x_1^1 , then there is another column omitting only 1 entry. Consider the submatrix obtained by removing this column and the first row: S omits exactly $2(n-1)$ entries in this submatrix, and again we conclude by induction.

To prove the last assertion, if other omitted entries lie in K , then they provide another equation for $\mathbf{J}(\sigma_1, L^S)$. \square

Theorem 3.2.5. *The number of irreducible components of $\mathcal{R}[n, 1, n^2 - 2n]$ coincides with the number of cycles of the complete bipartite graph $K_{n,n}$. Moreover, every ideal of an irreducible component is generated by a binomial of the form*

$$x_{j_1}^{i_1} \cdots x_{j_k}^{i_k} - x_{j_{\tau(1)}}^{i_1} \cdots x_{j_{\tau(k)}}^{i_k},$$

for some k , where $\tau \in \mathfrak{S}_k$ is a cycle and $I, J \subset [n]$ have size k .

Proof. $\mathcal{R}[n, 1, n^2 - 2n]$ is equidimensional and its irreducible components are $\mathbf{J}(\sigma_1, L^S)$ where S is a configuration of entries providing a join of expected dimension.

By Lemma 3.2.3 there exists a k such that, up to a permutation of rows and columns, S omits the entries $x_1^1, \dots, x_k^k, x_2^1, \dots, x_1^k$ and the equation of $\mathbf{J}(\sigma_1, L^S)$ is $x_1^1 \cdots x_k^k - x_2^1 \cdots x_1^k = 0$. In particular, entries in S^c that do not lie in the submatrix K are free to vary. Let F be the set of entries whose complement is $\{x_1^1, \dots, x_k^k, x_2^1, \dots, x_1^k\}$; we obtain $\mathbf{J}(\sigma_1, L^S) = \mathbf{J}(\sigma_1, L^F)$. This shows that the irreducible components are determined by the choice of a $k \times k$ submatrix and by the choice, in this submatrix, of a configuration of $2k$ entries such that, after a permutation of rows and columns, it has the form of (3.2.4).

Every configuration of this type, viewed as the adjacency matrix of a (n, n) -bipartite graph, determines a cycle in the complete bipartite graph $K_{n,n}$. This shows that the number of irreducible components of $\mathcal{R}[n, 1, n^2 - 2n]$ is the number of such cycles \square

Remark 3.2.6. The number of cycles in $K_{n,n}$ is

$$\sum_{k=2}^n \binom{n}{k}^2 \frac{k!(k-1)!}{2}.$$

Example 3.2.7. Examples of generators of ideals of $\mathbf{J}(\sigma_1, L^S)$:

- (1) If $s = 1$, the ideal is generated by the 2×2 minors not including the element of S and the degree is $\deg(\sigma_1) - 1 = \frac{(2n-2)!}{[(n-1)!]^2} - 1$.
- (2) If $s = 2$, the ideal is generated by the 2×2 minors not including the elements of S . If the elements of S lie in the same column or row, the degree is $\deg(\sigma_1) - (n-1) = \frac{(2n-2)!}{[(n-1)!]^2} - (n-1)$ and otherwise it is $\deg(\sigma_1) - 2 = \frac{(2n-2)!}{[(n-1)!]^2} - 2$.
- (3) If $s = 3$ and there are no entries of S in the same row or column, the ideal is generated in degrees two and three by the 2×2 minors not including the elements of S and the difference of the two terms in the 3×3 minor containing all three elements of S and the degree is $\frac{(2n-2)!}{[(n-1)!]^2} - 3$.
- (4) If $s = 3$ and there are two entries of S in the same row or column, the ideal is generated in degree two by the 2×2 minors not including the elements of S .
- (5) If $s \leq n$ and there are no elements of S in the same row or column, then $\deg(\mathbf{J}(\sigma_1, L^S)) = \frac{(2n-2)!}{[(n-1)!]^2} - s$. (See Theorem 5.4.7.)

3.3. Case $r = 2$. The following propositions are straight-forward to verify with the help of a computer with explicit computations available at www.nd.edu/~jhauenst/rigidity.

Proposition 3.3.1. *Let $n = 5$, $r = 2$ and let $S = \{x_1^1, x_2^2, \dots, x_5^5\}$. Then $\mathbf{J}(\sigma_{2,5}, L^S)$ has 27 generators of degree 5 of the form (3.1.3), e.g., $M_{456}^{123} M_{12}^{45} - M_{123}^{345} M_{45}^{12}$, and 5 generators of degree 6 with 6 summands in their expression, each of the form $M_{\bullet\bullet\bullet}^{\bullet\bullet\bullet} M_{\bullet}^{\bullet} M_{\bullet}^{\bullet} M_{\bullet}^{\bullet}$:*

$$\begin{aligned} & - M_{123}^{345} M_4^1 M_5^1 M_1^2 + M_{134}^{235} M_2^1 M_5^1 M_1^4 - M_{135}^{234} M_2^1 M_4^1 M_1^5 \\ & + M_{235}^{134} M_4^1 M_1^2 M_1^5 - M_{345}^{123} M_2^1 M_4^1 M_1^5 - M_{234}^{135} M_5^1 M_1^2 M_1^4. \end{aligned}$$

Proposition 3.3.2. *Let $n = 6$, $r = 2$, $s = 15$ and let S be given by*

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \blacklozenge & \blacklozenge \\ \blacklozenge & 0 & 0 & \blacklozenge & 0 & 0 \\ 0 & \blacklozenge & \blacklozenge & 0 & 0 & 0 \\ 0 & \blacklozenge & \blacklozenge & 0 & \blacklozenge & 0 \\ \blacklozenge & 0 & 0 & \blacklozenge & 0 & \blacklozenge \\ 0 & 0 & 0 & \blacklozenge & \blacklozenge & \blacklozenge \end{pmatrix}.$$

Then $\mathbf{J}(\sigma_{2,6}, L^S)$ is a hypersurface of degree 9 whose equation is:

$$\begin{aligned} (3.3.3) \quad & - M_{235}^{235} M_{36}^{12} M_{12}^{16} M_{14}^{34} + M_{235}^{235} M_{26}^{12} M_{13}^{16} M_{14}^{34} + M_{236}^{126} M_{13}^{13} M_{25}^{25} M_{14}^{34} - M_{236}^{126} M_{12}^{13} M_{35}^{25} M_{14}^{34} \\ & + M_{235}^{126} M_{16}^{13} M_{23}^{25} M_{14}^{34} - M_{235}^{126} M_{14}^{13} M_{23}^{25} M_{16}^{34} + M_{146}^{134} M_{23}^{12} M_{23}^{25} M_{15}^{36} - M_{146}^{134} M_{15}^{13} M_{23}^{25} M_{23}^{26} \\ & - M_{136}^{136} M_{23}^{12} M_{25}^{25} M_{14}^{34} + M_{126}^{136} M_{23}^{12} M_{35}^{25} M_{14}^{34}. \end{aligned}$$

The weight of equation (3.3.3) is $(1^2, 2^2, 3^2, 4, 5, 6) \times (1^2, 2^2, 3^2, 4, 5, 6)$. (This weight is hinted at because the first, second and third columns and rows each have two elements of S in them and the fourth, fifth and sixth rows and columns each have three.)

Proposition 3.3.4. *Let $n = 6$, $r = 2$, $s = 15$, and let S be given by*

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \blacklozenge & \blacklozenge \\ 0 & \blacklozenge & 0 & 0 & \blacklozenge & 0 \\ 0 & 0 & \blacklozenge & \blacklozenge & 0 & 0 \\ \blacklozenge & 0 & 0 & \blacklozenge & \blacklozenge & 0 \\ \blacklozenge & 0 & \blacklozenge & 0 & 0 & \blacklozenge \\ 0 & \blacklozenge & 0 & \blacklozenge & 0 & \blacklozenge \end{pmatrix}.$$

Then, $\mathbf{J}(\sigma_{2,6}, L^S)$ is a hypersurface of degree 16.

We do not have a concise expression for the equation of $\mathbf{J}(\sigma_{2,6}, L^S)$. Expressed naïvely, it is the sum of 96 monomials, each with coefficient ± 1 plus two monomials with coefficient ± 2 , for a total of 100 monomials counted with multiplicity. The monomials are of weight $(1^4, 2^3, 3^3, 4^2, 5^2, 6^2) \times (1^4, 2^3, 3^3, 4^2, 5^2, 6^2)$. (This weight is hinted at because the first, second and third columns and rows each have two elements of S in them, but the first is different because in the second and third a column element equals a row element, and the fourth, fifth and sixth rows and columns each have three.)

3.4. Case $r = n - 2$. Proposition 3.1.4 implies:

Theorem 3.4.1. *In the hypersurface case $r = n - 2$, $s = 3$, there are two types of varieties up to isomorphism:*

- (1) *if no two elements of S are in the same row or column, then the hypersurface is of degree $2n - 3$ and can be represented by an equation of the form (3.1.6). There are $\binom{n}{3}^2$ such components.*

- (2) If two elements are in the same row and one in a different column from those two, or such that one element shares a row with one and a column with the other, then the equation is the unique size $(n - 1)$ minor that has no elements of S in it. There are n^2 such components.

If all three elements of S lie on a row or column, then $\mathbf{J}(\sigma_{n-2}, L^S)$ is not a hypersurface.

Corollary 3.4.2. Let M be an $n \times n$ matrix. Then M is maximally $(n - 2)$ -border rigid if and only if no size $n - 1$ minor is zero and for all index sets $\{i_1, i_2, i_3\} \subset [n]$, $\{j_1, j_2, j_3\} \subset [n]$, the equation $\Delta_{j_1}^{i_1} \Delta_{j_2 j_3}^{i_2 i_3} - \Delta_{j_2}^{i_2} \Delta_{j_1 j_3}^{i_1 i_3}$ does not vanish on M .

4. VARIETIES OF MATRICES

4.1. General remarks. Recall the construction of matrices from directed acyclic graphs in §1. To each graph Γ that is the disjoint union of directed acyclic graphs with n input gates and n output gates we associate the set $\Sigma_\Gamma^0 \subset Mat_n$ of all matrices admitting a linear circuit (see §1) with underlying graph Γ . We let $\Sigma_\Gamma := \overline{\Sigma_\Gamma^0} \subset Mat_n$, the *variety of linear circuits associated to Γ* .

For example $\mathcal{R}[n, r, s]^0 = \cup \Sigma_\Gamma^0$ where the union is over all $\Gamma = \Gamma_1 + \Gamma_2$ (addition as in Figure 4) where Γ_1 is of depth two with r vertices at the second level and is a complete bipartite graph at each level, and Γ_2 is of depth one, with s edges.

Proposition 4.1.1. Let $\Sigma \subset Mat_n$ be a variety of dimension δ . Then a general element of Σ cannot be computed by a circuit of size $\delta - 1$.

Proof. Let Γ be a fixed graph representing a family of linear circuits with γ edges. Then Γ can be used for at most a γ -dimensional family of matrices. Any variety of matrices of dimension greater than γ cannot be represented by Γ , and since there are a finite number of graphs of size at most γ , the dimension of their union is still γ . \square

4.2. Cauchy matrices. Let $1 \leq i, j, \leq n$. Consider the rational map

$$(4.2.1) \quad \text{Cau}_n : \mathbb{C}^n \times \mathbb{C}^n \dashrightarrow Mat_n$$

$$((x^i), (z_j)) \mapsto (y_j^i) := \frac{1}{x^i + z_j}$$

The variety of Cauchy matrices $\text{Cauchy}_n \subset Mat_n$ is defined to be the closure of the image of (4.2.1). It is $\mathfrak{S}_n \times \mathfrak{S}_n$ invariant and has dimension $2n - 1$. To see the dimension, note that Cauchy_n is the *Hadamard inverse* or *Cremona transform* of a linear subspace of Mat_n of dimension $2n - 1$ (that is contained in σ_2). The Cremona map is

$$\text{Crem}_N : \mathbb{C}^N \dashrightarrow \mathbb{C}^N$$

$$(w_1, \dots, w_N) \mapsto \left(\frac{1}{w_1}, \dots, \frac{1}{w_N} \right)$$

which is generically one to one. The fiber of $\text{Crem}_n \circ \text{Cau}_n$ over $(x^i + z_j)$ is $((x^i + \lambda), (z_j - \lambda))$, with $\lambda \in \mathbb{C}$.

One can obtain equations for Cauchy_n by transporting the linear equations of its Cremona transform, which are the $(n - 1)^2$ linear equations, e.g., for $i, j = 2, \dots, n$, $y_1^1 + y_j^i - y_1^i - y_j^1$. (More generally, it satisfies the equation $y_{j_1}^{i_1} + y_{j_2}^{i_1} - y_{j_1}^{i_2} - y_{j_2}^{i_2}$ for all i_1, j_1, i_2, j_2 .) Thus, taking reciprocals and clearing denominators, the Cauchy variety has cubic equations

$$y_{j_2}^{i_1} y_{j_1}^{i_2} y_{j_2}^{i_1} + y_{j_1}^{i_1} y_{j_1}^{i_2} y_{j_2}^{i_1} - y_{j_1}^{i_1} y_{j_2}^{i_1} y_{j_2}^{i_1} - y_{j_1}^{i_1} y_{j_2}^{i_1} y_{j_1}^{i_2}.$$

Alternatively, $Cauchy_n$ can be parametrized by the first row and column: let $2 \leq \rho, \sigma \leq n$, and denote the entries of A by a_j^i . Then the space is parametrized by $a_1^1, a_1^\rho, a_\sigma^1$, by setting $a_\sigma^\rho = [\frac{1}{a_1^\rho} + \frac{1}{a_\sigma^1} - \frac{1}{a_1^1}]^{-1}$.

Any square submatrix of a Cauchy matrix is a Cauchy matrix, and the determinant of a Cauchy matrix is given by

$$(4.2.2) \quad \frac{\prod_{i < j} (x^i - x^j) \prod_{i < j} (z^i - z^j)}{\prod_{i, j} (x^i + z_j)}$$

In particular, if $x^i, -z_j$ are all distinct, then all minors of the Cauchy matrix are nonzero.

Proposition 4.2.3. *A general Cauchy matrix is both maximally $r = 1$ rigid and maximally $r = n - 2$ rigid.*

Proof. For the $r = 1$ case, let σ be a k -cycle in \mathfrak{S}_k and say there were an equation

$$y_{j_1}^{i_1} \cdots y_{j_k}^{i_k} - y_{j_{\sigma(1)}}^{i_1} \cdots y_{j_{\sigma(k)}}^{i_k}$$

$Cauchy_n$ satisfied. By the $\mathfrak{S}_n \times \mathfrak{S}_n$ invariance we may assume the equation is

$$y_1^1 \cdots y_k^k - y_{\sigma(1)}^1 \cdots y_{\sigma(k)}^k$$

which may be rewritten as

$$\frac{1}{y_1^1 \cdots y_k^k} = \frac{1}{y_{\sigma(1)}^1 \cdots y_{\sigma(k)}^k}$$

The first term contains the monomial $x^1 \cdots x^{k-1} z_k$, but the second does not.

For the $r = n - 2$ case, we may assume the equation is $\Delta_2^1 \Delta_{13}^{23} - \Delta_1^2 \Delta_{23}^{13}$, because for a general Cauchy matrix all size $n - 2$ minors are nonzero and we have the $\mathfrak{S}_n \times \mathfrak{S}_n$ invariance.

Write our equation as $\frac{\Delta_2^1}{\Delta_1^2} = \frac{\Delta_{23}^{13}}{\Delta_{13}^{23}}$. Then using (4.2.2) and canceling all repeated terms we get

$$\frac{(x^2 - x^3)(z_1 - z_3)(x^1 + z_3)(x^3 + z_2)}{(x^1 - x^3)(z_2 - z_3)(x^2 + z_3)(x^3 + z_1)} = 1$$

which fails to hold for a general Cauchy matrix. \square

4.3. The Vandermonde variety. In [11, p26], they ask if a general Vandermonde matrix has maximal rigidity.

Consider the map

$$(4.3.1) \quad Van_n : \mathbb{C}^{n+1} \rightarrow Mat_n$$

$$(y_0, y_1, \dots, y_n) \mapsto \begin{pmatrix} (y_0)^{n-1} & \cdots & (y_0)^{n-1} \\ (y_0)^{n-2} y_1 & \cdots & (y_0)^{n-2} y_n \\ (y_0)^{n-3} y_1^2 & \cdots & (y_0)^{n-3} (y_n)^2 \\ \vdots & & \vdots \\ (y_1)^{n-1} & \cdots & (y_n)^{n-1} \end{pmatrix} = (x_j^i)$$

Define the *Vandermonde variety* $Vand_n$ to be the closure of the image of this map. Note that this variety contains n rational normal curves (set all y_j except y_0, y_{i_0} to zero), and is \mathfrak{S}_n -invariant (permutation of columns). The (un-normalized) Vandermonde matrices are the Zariski open subset where $y_0 \neq 0$ (set $y_0 = 1$ to obtain the usual Vandermonde matrices). Give $Mat_{n \times n}$ coordinates x_j^i . The variety $Vand_n$ is contained in the linear space $\{x_1^1 - x_2^1 = 0, \dots, x_1^1 - x_n^1 = 0\}$ and it is the zero set of these linear equations and the generators of the ideals of the rational

normal curves $\text{Van}[y_0, 0, \dots, 0, y_j, 0, \dots, 0]$. Explicitly, fix j , the generators for the rational normal curves are the 2×2 minors of

$$\begin{pmatrix} x_j^1 & x_j^2 & \cdots & x_j^{n-1} \\ x_j^2 & x_j^3 & \cdots & x_j^n \end{pmatrix}$$

see, e.g., [9, p. 14], and thus the equations for the variety are, fixing j and $i < k$, the quadratic polynomials $x_j^i x_j^k - x_j^{i+1} x_j^{k-1}$.

To see the assertion about the zero set, first consider the larger parametrized variety where instead of y_0 appearing in each column, in the j -th column, replace y_0 by a variable y_{0j} . The resulting variety is the join of n rational normal curves, each contained in a $\mathbb{P}^{n-1} \subset \mathbb{P}\text{Mat}_n$, where the \mathbb{P}^{n-1} 's are just the various columns. In general, given varieties $X_j \subset \mathbb{P}V_j$, $j = 1, \dots, q$, the join $J(X_1, \dots, X_q) \subset \mathbb{P}(V_1 \oplus \cdots \oplus V_q)$ has ideal generated by $I(X_1, \mathbb{P}V_1), \dots, I(X_q, \mathbb{P}V_q)$, see, e.g. [9, 18.17, *Calculation III*]. The second set of equations exactly describes this join. Now intersect this variety with the linear space where all entries on the first row are set equal. We obtain the Vandermonde variety.

Proposition 4.3.2. *$\text{Vand}_n \not\subset \mathcal{R}[n, 1, n^2 - 2n]$ and $\text{Vand}_n \not\subset \mathcal{R}[n, n - 2, 3]$, i.e., Vandermonde matrices are generically maximally 1-border rigid and $(n - 2)$ -border rigid.*

Proof. Say we had Vand_n in some component of $\mathcal{R}[n, 1, n^2 - 2n]$. Using the \mathfrak{S}_n -invariance, we may assume the equation it satisfies is $x_1^{i_1} \cdots x_k^{i_k} - x_{\sigma(1)}^{i_1} \cdots x_{\sigma(k)}^{i_k}$ for some k , where $\sigma \in \mathfrak{S}_k$ is a k -cycle. Assume $i_k = \max\{i_\ell\}$. Then the first monomial is divisible by $(y_k)^{i_k-1}$ but the second is not.

For the $n - 2$ -rigidity, since no minors are zero, by Corollary 3.4.2 and the \mathfrak{S}_n -invariance, it suffices to consider equations of the form $\Delta_2^j \Delta_{13}^{ik} - \Delta_3^k \Delta_{12}^{ij}$, where $S = \{x_1^i, x_2^j, x_3^k\}$. First consider the case that $2 \notin \{i, j, k\}$. The y_2 -linear coefficient of $\Delta_2^j \Delta_{13}^{ik} - \Delta_3^k \Delta_{12}^{ij}$ is $\Delta_2^j \Delta_{132}^{ik2} - \Delta_{32}^{k2} \Delta_{12}^{ij}$. This expression is nonzero, because as a polynomial in y_1 it has linear coefficient $\Delta_{21}^{j2} \Delta_{132}^{ik2}$, which is a product of minors and hence nonzero. Now for the other case let $i = 2, j \neq 2, k \neq 2$. But the y_2 -linear coefficient of $\Delta_2^j \Delta_{13}^{2k} - \Delta_3^k \Delta_{12}^{2j}$ is $-\Delta_{32}^{k2} \Delta_{12}^{2j}$, which is also nonzero. \square

4.4. The DFT matrix. The following ‘folklore result’ was communicated to us (independently) by A. Kumar and A. Wigderson:

Proposition 4.4.1. *Let A be a matrix with an eigenvalue of multiplicity $k > \sqrt{n}$. Then $A \in \hat{\mathcal{R}}[n, n - k, n]^0$.*

Proof. Let λ be the eigenvalue with multiplicity k , then $A - \lambda Id$ has rank $n - k$. To have the condition be nontrivial, we need $r(2n - r) + s = (n - k)(2n - (n - k)) + n < n^2$, i.e., $n < k^2$. \square

Equations for the variety of matrices with eigenvalues of high multiplicity can be obtained via resultants applied to the coefficients of the characteristic polynomial of a matrix.

Corollary 4.4.2. *Let $n = 2^k$, then $DFT_n \in \hat{\mathcal{R}}[n, \frac{3n}{4}, n]^0$.*

Proof. The eigenvalues of DFT_n are $\pm 1, \pm \sqrt{-1}$ with multiplicity roughly $\frac{n}{4}$ each. \square

Proposition 4.4.3. *Any matrix with \mathbb{Z}_2 symmetry (either symmetric or symmetric about the anti-diagonal) is not maximally 1-border rigid.*

Proof. Say the matrix is symmetric. Then $x_2^1 x_3^2 \cdots x_n^{n-1} x_1^n - x_1^2 x_2^3 \cdots x_{n-1}^n x_n^1$ is in the ideal of the hypersurface $\mathbf{J}(\sigma_1, L^S)$ where S is the span of all the entries not appearing in the expression. \square

4.5. The DFT curve. We define two varieties that contain the DFT matrix, the first corresponds to a curve in projective space.

Define the *DFT curve* $CDFT_n \in Mat_n$ to be the image of the map

$$(4.5.1) \quad \mathbb{C}^2 \rightarrow Mat_n$$

$$(x, w) \mapsto \begin{pmatrix} x^{n-1} & x^{n-1} & x^{n-1} & \cdots & x^{n-1} \\ x^{n-1} & x^{n-2}w & x^{n-3}w^2 & \cdots & w^{n-1} \\ & \vdots & & & \\ x^{n-1} & w^{n-1} & x^1w^{n-2} & \cdots & x^{n-2}w \end{pmatrix}$$

This curve is a subvariety of $Vand_n$ where $y_0 = y_1 = x$ and $y_j = w^{j-1}$. From this one obtains its equations.

Proposition 4.5.2. *For general w , and even for w a fifth root of unity, the matrix*

$$M(w) := \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & w & w^2 & w^3 & w^4 \\ 1 & w^2 & w^4 & w & w^3 \\ 1 & w^3 & w & w^4 & w^2 \\ 1 & w^4 & w^3 & w^2 & w \end{pmatrix}$$

satisfies $M(w) \notin \hat{\mathcal{R}}[5, 3, 2]$, $M(w) \in \mathcal{R}[5, 3, 3]^0$, $M(w) \notin \hat{\mathcal{R}}[5, 1, 12]$, and $M(w) \in \hat{\mathcal{R}}[5, 1, 13]^0$.

This is proved by explicit calculation at www.nd.edu/~jhauenst/rigidity. For a more general DFT matrix we have:

Proposition 4.5.3. *Let p be prime, then the DFT curve $CDFT_p$ satisfies, for all $A \in CDFT_p$*

$$\mathcal{R}ig_1(A) \leq (p-1)^2 + 1 - (p-1).$$

In other words, $CDFT_p \subset \hat{\mathcal{R}}[p, 1, p^2 - 3p + 3]^0$.

Proof. Change the $(1, 1)$ -entry to w^{-1} and all entries in the lower right $(p-1) \times (p-1)$ submatrix not already equal to w to w . The resulting matrix is

$$\begin{pmatrix} w^{-1} & 1 & 1 & \cdots & 1 \\ 1 & w & w & \cdots & w \\ & & \vdots & & \\ 1 & w & w & \cdots & w \end{pmatrix}$$

□

4.6. The variety of factorizable matrices/the butterfly variety. The DFT algorithm may be thought of as factorizing the size $n = 2^k$ DFT matrix into a product of k matrices S_1, \dots, S_k with each S_i having $2n$ nonzero entries.

If S_j , $1 \leq j \leq d$ are matrices with s_j nonzero entries in S_j , with $s_j = f_j n$ for some natural numbers f_j , then $S_1 S_2$ has at most $f_1 f_2 n$ nonzero entries. Consider the set of matrices A such that we may write $A = S_1 \cdots S_d$ with $s_j = f_j n$ and $f_1 \cdots f_d = n$. Then A may be computed by a linear circuit of depth d and size $(f_1 + \cdots + f_d)n$. In the DFT case we have $f_j = 2$ and $d = \log(n)$.

This space of matrices is the union of a large number of components, each component is the image of a map:

$$Bfly: \hat{L}_{s_1} \times \cdots \times \hat{L}_{s_d} \rightarrow Mat_{n \times n}$$

where $\hat{L}_{s_j} \subset Mat_{n \times n}$ is the span of some $S \subset \{x_j^i\}$ of cardinality s_j . In the most efficient configurations (those where the map has the smallest dimensional fibers), each entry y_j^i in a matrix in the image will be of the form $y_j^i = (x_1)_{j_1}^i (x_2)_{j_2}^{j_1} \cdots (x_d)_j^{j_{d-1}}$ where the j_u 's are fixed indices (no sum).

If we are not optimally efficient, then the equations for the corresponding variety become more complicated, and the dimension will drop.

From now on, for simplicity assume $n = 2^k$, $d = k$ and $s_j = 2n$ for $1 \leq j \leq k$. Let FM_n^0 denote the set of factorizable or *butterfly* matrices, the set of matrices A such that $A = S_1 \cdots S_k$ with S_j as above, and let $FM_n := \overline{FM_n^0}$ denote its Zariski closure. The term ‘‘butterfly’’ comes from the name commonly used for the corresponding circuit, e.g., see [14, §3.7]. By construction every $A \in FM_n^0$ admits a linear circuit of size $2n \log n$, see, e.g., Figure 6: the graph has 48 edges compared with 64 for a generic 8×8 matrix, and in general one has $2^{k+1}k = 2n \log n$ edges compared with $2^{2k} = n^2$ for a generic matrix.

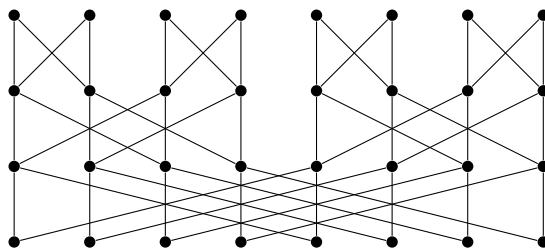


FIGURE 6. linear circuit for element of FM_8^0 , support is the ‘‘butterfly graph’’

Proposition 4.6.1. *A general factorizable matrix does not admit a linear circuit of size $n(\log n + 1) - 1$.*

Proof. We will show that a general component of FM_n has dimension $n(\log n + 1)$, so Proposition 4.1.1 applies.

First it is clear that $\dim FM_n$ is at most $n(\log n + 1)$, because $\dim(\hat{L}_1 \oplus \cdots \oplus \hat{L}_k) = 2nk$ and if D_1, \dots, D_{k-1} are diagonal matrices (with nonzero entries on the diagonal), then $Bfly(S_1 D_1, D_1^{-1} S_2 D_2, \dots, D_{k-1}^{-1} S_k) = Bfly(S_1, S_2, \dots, S_k)$, so the fiber has dimension at least $n(k - 1)$. Consider the differential of $Bfly$ at a general point:

$$d(Bfly)|_{(S_1, \dots, S_k)} : \hat{L}_1 \oplus \cdots \oplus \hat{L}_k \rightarrow Mat_{n \times n}$$

$$(Z_1, \dots, Z_k) \mapsto Z_1 S_2 \cdots S_k + S_1 Z_2 S_3 \cdots S_k + \cdots + S_1 \cdots S_{k-1} Z_k$$

The rank of this linear map is the dimension of the image of FM_n as its image is the tangent space to a general point of FM_n . We may use Z_1 to alter $2n$ entries of the image matrix $y = S_1 \cdots S_k$. Then, *a priori* we could use Z_2 to alter $2n$ entries, but n of them overlap with the entries altered by Z_1 , so Z_2 may only alter n new entries. Now think of the product of the first two matrices as fixed, then Z_3 multiplied by this product again can alter n new entries, and similarly for all Z_j . Adding up, we get $2n + (k - 1)n = n(\log n + 1)$. \square

5. GEOMETRY

5.1. Standard facts on joins. We review standard facts as well as observations in [11, 13]. Recall the notation $\mathbf{J}(X, Y)$ from Definition 2.1.1. The following are standard facts:

Proposition 5.1.1.

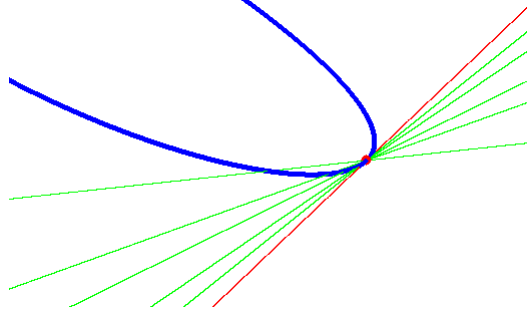


FIGURE 7. Secant lines limiting to a tangent line

- (1) If X, Y are irreducible, then $\mathbf{J}(X, Y)$ is irreducible.
- (2) Let $X, Y \subset \mathbb{P}V$ be varieties, then $\mathcal{I}(\mathbf{J}(X, Y)) \subset \mathcal{I}(X) \cap \mathcal{I}(Y)$.
- (3) (Terracini's Lemma) The dimension of $\mathbf{J}(X, Y)$ is $\dim X + \dim Y + 1 - \dim \hat{T}_x X \cap \hat{T}_y Y$, where $x \in X, y \in Y$ are general points. In particular,
 - (a) the dimension is $\dim X + \dim Y + 1$ if there exist $x \in X, y \in Y$ such that $\hat{T}_x X \cap \hat{T}_y Y = 0$. ($\dim X + \dim Y + 1$ is called the expected dimension.)
 - (b) If $Y = L$ is a linear space, $\mathbf{J}(X, L)$ will have the expected dimension if and only if there exists $x \in X$ such that $\hat{T}_x X \cap \hat{L} = 0$.
- (4) If $z \in \mathbf{J}(X, p)$ and $z \notin \langle x, p \rangle$ for some $x \in X$, then z lies on a line that is a limit of secant lines $\langle x_t, p \rangle$, for some curve x_t with $x_0 = p$.

Proof. For assertions (1), (3), (4) respectively see e.g., [9, p157], [12, p122], and [12, p118]. Assertion (2) holds because $X, Y \subset \mathbf{J}(X, Y)$. \square

To gain intuition regarding Terracini's lemma, a point on $\mathbf{J}(X, Y)$ is obtained by selecting a point $x \in X$ ($\dim X$ parameters), a point $y \in Y$ ($\dim Y$ parameters) and a point on the line joining x and y (one parameter). Usually these parameters are independent, and Terracini's lemma says that if the infinitesimal parameters are independent, the actual parameters are as well.

In the special case (3b), since Y is a linear space, it is equal to its tangent space.

To understand (4), consider Figure 5.1 where a point on a limit of secant lines lies on a tangent line.

5.2. Ideals of cones. Define the *primitive part* of the ideal of a variety $Z \subset \mathbb{P}V$ as $\mathcal{I}_{prim,d}(Z) := \mathcal{I}_d(Z) / (\mathcal{I}_{d-1}(Z) \circ V^*)$. Here, if $A \subseteq S^d V$ and $B \subseteq S^d V$, $A \circ B := \{pq \mid p \in A, q \in B\}$. Note that $\mathcal{I}_{prim,d}(Z)$ is only nonzero in the degrees that minimal generators of the ideal of Z appear and that (lifted) bases of $\mathcal{I}_{prim,d}(Z)$ for each such d furnish a set of generators of the ideal of Z .

Proposition 5.2.1. *Let $X \subset \mathbb{P}V$ be a variety and let $L \subset \mathbb{P}V$ be a linear space.*

- (1) Then

$$\mathcal{I}_d(X) \cap S^d L^\perp \subseteq \mathcal{I}_d(\mathbf{J}(X, L)) \subseteq \mathcal{I}_d(X) \cap (L^\perp \circ S^{d-1} V^*).$$

- (2) A set of generators of $\mathcal{I}(\mathbf{J}(X, L))$ may be taken from $\mathcal{I}(X) \cap \text{Sym}(L^\perp)$.
- (3) In particular, if $\mathcal{I}_k(X)$ is empty, then $\mathcal{I}_k(\mathbf{J}(X, L))$ is empty and

$$\mathcal{I}_{k+1}(X) \cap S^{k+1} L^\perp = \mathcal{I}_{k+1,prim}(\mathbf{J}(X, L)) = \mathcal{I}_{k+1}(\mathbf{J}(X, L)).$$

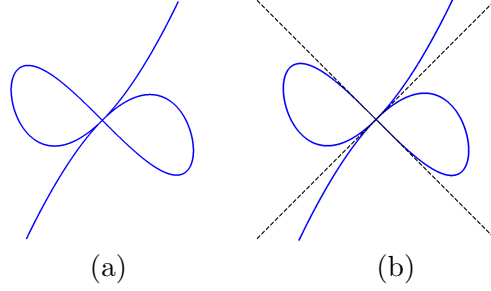


FIGURE 8. (a) is graph of $2x^5 - x^3 + x^2y + xy^2 - y^3 = 0$ and (b) is the graph with the tangent cone at the origin.

Proposition 5.2.1 says that we only need to look for polynomials in the variables of L^\perp when looking for equations of $\mathbf{J}(X, L)$.

Proof. For the first assertion, $P \in \mathcal{I}_d(\mathbf{J}(X, L))$ if and only if $P_{k,d-k}(x, \ell) = 0$ for all $[x] \in X$, $[\ell] \in L$ and $0 \leq k \leq d$ where $P_{k,d-k} \in S^k V^* \otimes S^{d-k} V^*$ is a polarization of P (in coordinates, $P_{k,d-k}$ is the coefficient of t^k in the expansion of $P(tx + y)$ in t , where x, y are independent sets of variables and t is a single variable, see [12, §7.5] for more details). Now $P \in S^d L^\perp$ implies all the terms vanish identically except for the $k = d$ term. But $P \in \mathcal{I}_d(X)$ implies that term vanishes as well. The second inclusion of the first assertion is Proposition 5.1.1(2).

For the second assertion, we can build L up by points as $\mathbf{J}(X, \langle L', L'' \rangle) = \mathbf{J}(\mathbf{J}(X, L'), L'')$, so assume $\dim L = 0$. Let $P \in \mathcal{I}_d(\mathbf{J}(X, L))$. Choose a (one-dimensional) complement W^* to L^\perp in V^* . Write $P = \sum_{j=1}^d q_j u^{d-j}$ where $q_j \in S^j L^\perp$ and $u \in W^*$. Then

$$(5.2.2) \quad P_{j,d-j}(x^j, \ell^{d-j}) = \sum_{i=0}^j \sum_{t=0}^i (q_i)_{t,i-t}(x^t, \ell^{i-t})(u^{d-i})_{j-t,d-j+t-i}(x^{j-t}, \ell^{d-j+t-i})$$

$$(5.2.3) \quad = \sum_{i=0}^j q_i(x)(u^{d-i})_{j-i,d-j}(x^{j-i}, \ell^{d-j})$$

Consider the case $j = 1$, then (5.2.3) reduces to $q_1(x)u^{d-1}(\ell) = 0$ which implies $q_1 \in \mathcal{I}_1(X) \cap L^\perp$. Now consider the case $j = 2$, since $q_1(x) = 0$, it reduces to $q_2(x)u^{d-2}(\ell)$, so we conclude $q_2(x) \in \mathcal{I}_2(X) \cap S^2 L^\perp$. Continuing, we see each $q_j \in \mathcal{I}_j(X) \cap S^j L^\perp \subset \mathcal{I}(\mathbf{J}(X, L))$ and the result follows. \square

5.3. Degrees of cones. For a projective variety $Z \subset \mathbb{P}V$ and $z \in Z$, let $\hat{TC}_z Z \subset V$ denote the *affine tangent cone* to Z at z and $TC_z Z = \mathbb{P}\hat{TC}_z Z \subset \mathbb{P}V$ the (*embedded*) *tangent cone*. Set-theoretically $\hat{TC}_z Z$ is the union of all points on all lines of the form $\lim_{t \rightarrow 0} \langle z, z(t) \rangle$ where $z(t) \subset \hat{Z}$ is a curve with $[z(0)] = z$. If Z is irreducible, then $\dim TC_z Z = \dim Z$.

To gain some intuition regarding tangent cones, we compute the tangent cone to $2x^5 - w^2x^3 + w^2x^2y + w^2xy^2 - w^2y^3 = 0$ at $[(w, x, y)] = [(1, 0, 0)]$.

Figure 8 depicts this curve in the affine space $w = 1$. The line $\{x + y = 0\}$ has multiplicity one in the tangent cone, and the line $\{x - y = 0\}$ has multiplicity two because two of the branches of the curve that go through the origin are tangent to it. We will need to keep track of these multiplicities in order to compute the degree of the tangent cone as a subscheme of the Zariski tangent space. That is, we need to keep track of its ideal, not just its zero set. To this end, let \mathfrak{m} denote the maximal ideal in $\mathcal{O}_{Z,z}$ of germs of regular functions on Z at z vanishing at z , so

the Zariski tangent space is $T_z Z = (\mathfrak{m}/\mathfrak{m}^2)^*$. Then the (abstract) tangent cone is the subscheme of $T_z Z$ whose coordinate ring is the graded ring $\bigoplus_{j=0}^{\infty} \mathfrak{m}^j / \mathfrak{m}^{j+1}$.

To compute the ideal of the tangent cone in practice, one takes a set of generators for the ideal of Z and local coordinates (w, y^α) such that $z = [(1, 0)]$, and writes, for each generator $P \in \mathcal{I}(Z)$, $P = (w - 1)^j Q(y) + O((w - 1)^{j+1})$. The generators for the ideal of the tangent cone are the lowest degree homogeneous components of the corresponding $Q(y)$. See either of [9, Ch. 20] or [20, Ch. 5] for details. The *multiplicity* of Z at z is defined to be $\text{mult}_z Z = \deg(TC_z Z)$.

In our example, $(y^1, y^2) = (x, y)$ and we take coordinates with origin at $[(1, 0, 0)]$ so let $\tilde{w} = w - 1$ to have the expansion $P = -\tilde{w}^2(x + y)(x - y)^2 - 2\tilde{w}(x + y)(x - y)^2 - (x + y)(x - y)^2 + 2x^5 = \tilde{w}^0[-(x + y)(x - y)^2 + 2x^5] + O(\tilde{w})$ and the ideal of the tangent cone is generated by $(x + y)(x - y)^2$. So in Figure 8, the multiplicity at the origin is three. We will slightly abuse notation writing $TC_z Z$ for both the abstract and embedded tangent cone. While $TC_x Z$ may have many components with multiplicities, it is equi-dimensional, see [21, p162].

Proposition 5.3.1. *Let $X \subset \mathbb{P}V$ be a variety and let $x \in X$. Assume that $\mathbf{J}(X, x) \neq X$. Let $p_x : \mathbb{P}V \setminus x \rightarrow \mathbb{P}(V/\hat{x})$ denote the projection map and let $\pi := p_x|_{X \setminus x}$. Then*

$$\deg(\mathbf{J}(X, x)) = \frac{1}{\deg \pi} [\deg(X) - \deg(TC_x X)].$$

To gain intuition for Proposition 5.3.1, assume π has degree one, which it will in our situation and let $\mathbb{P}W \subset \mathbb{P}V$ be a general linear space of complementary dimension to $\mathbf{J}(X, x)$, so it intersects $\mathbf{J}(X, x)$ in $\deg(\mathbf{J}(X, x))$ points, each of multiplicity one. Now consider the linear space spanned by $\mathbb{P}W$ and x . It intersects X in $\deg(\mathbf{J}(X, x)) + 1$ points ignoring multiplicity but it may intersect x with multiplicity greater than one. The degree of X is the number of points of intersection counted with multiplicity, so the degree of $\mathbf{J}(X, x)$ is the degree of X minus the multiplicity of the intersection at x . If $x \in X$ is a smooth point, the multiplicity will be one, in general the multiplicity will equal the degree of the tangent cone, which can be visualized by considering a horizontal line through the curve in Fig. 8(a), and moving the line upwards just a little. The three physical points of intersection become five on the moved line. Here is the formal proof:

Proof. By [20, Thm. 5.11],

$$\deg(\overline{\pi(X \setminus x)}) = \frac{1}{\deg \pi} [\deg(X) - \deg(TC_x X)].$$

Now let $H \subset \mathbb{P}V$ be a hyperplane not containing x that intersects $\mathbf{J}(X, x)$ transversely. Then $\overline{\pi(X \setminus x)} \subset \mathbb{P}(V/\hat{x})$ is isomorphic to $\mathbf{J}(X, x) \cap H \subset H$. In particular their degrees are the same. \square

Note that the only way to have $\deg(\pi) > 1$ is for every secant line through x to be at least a trisecant line.

Proposition 5.3.2. *Let $X \subset \mathbb{P}V$ be a variety, let $L \subset \mathbb{P}V$ be a linear space, and let $x \in X$. Then we have the inclusion of schemes*

$$(5.3.3) \quad \mathbf{J}(TC_x X, \tilde{L}) \subseteq TC_x \mathbf{J}(X, L)$$

where $\tilde{L} \subset T_x \mathbb{P}V$ is the image of L in the projectivized Zariski tangent space, and both are sub-schemes of $T_x \mathbb{P}V$.

Proof. Write $x = [v]$. For any variety $Y \subset \mathbb{P}V$, generators for $TC_x Y$ can be obtained from generators Q_1, \dots, Q_s of $\mathcal{I}(Y)$, see e.g., [9, Chap. 20]. The generators are $Q_1(v^{f_1}, \cdot), \dots, Q_s(v^{f_s}, \cdot)$ where f_j is the largest nonnegative integer (which is at most $\deg Q_j - 1$ since $x \in Y$) such that $Q_j(v^{f_j}, \cdot) \neq 0$. Here, if $\deg(Q_j) = d_j$, then strictly speaking $Q_j(v^{f_j}, \cdot) \in S^{d_j - f_j} T_x^* Y$, but we may

consider $T_x Y \subset T_x \mathbb{P}V$ and may ignore the additional linear equations that arise as they don't effect the proof.

Generators of $\mathbf{J}(X, L)$ can be obtained from elements of $\mathcal{I}(X) \cap \text{Sym}(L^\perp)$. Let $P_1, \dots, P_g \in \mathcal{I}(X) \cap \text{Sym}(L^\perp)$ be such a set of generators. Then, choosing the f_j as above, $P_1(v^{f_1}, \cdot), \dots, P_g(v^{f_g}, \cdot)$ generate $\mathcal{I}(TC_x(\mathbf{J}(X, L)))$.

Note that $P_1(v^{f_1}, \cdot), \dots, P_g(v^{f_g}, \cdot) \in \mathcal{I}(TC_x X) \cap \text{Sym}(L^\perp)$, so they are in $\mathcal{I}(\mathbf{J}(TC_x X, \tilde{L}))$. Thus $\mathcal{I}(TC_x(\mathbf{J}(X, L))) \subseteq \mathcal{I}(\mathbf{J}(TC_x X, \tilde{L}))$. \square

Remark 5.3.4. The inclusion (5.3.3) may be strict. For example $\mathbf{J}(TC_{[a_1 \otimes b^1]} \sigma_r, [a_1 \otimes b^2]) \neq TC_{[a_1 \otimes b^1]} \mathbf{J}(\sigma_r, [a_1 \otimes b^2])$, where $a_i \otimes b^j$ is the matrix having 1 at the entry (i, j) and 0 elsewhere. To see this, first note that $[a_1 \otimes b^2] \subset TC_{[a_1 \otimes b^1]} \sigma_r$, so as a set $\mathbf{J}(TC_{[a_1 \otimes b^1]} \sigma_r, [a_1 \otimes b^2]) = TC_{[a_1 \otimes b^1]} \sigma_r$, in particular it is of dimension one less than $\mathbf{J}(\sigma_r, [a_1 \otimes b^2])$ which has the same dimension as its tangent cone at any point.

Proposition 5.3.2 implies:

Corollary 5.3.5. *Let $X \subset \mathbb{P}V$ be a variety, let $L \subset \mathbb{P}V$ be a linear space, and let $x \in X$. Assume $TC_x \mathbf{J}(X, L)$ is reduced, irreducible, and $\dim \mathbf{J}(TC_x X, \tilde{L}) = \dim TC_x \mathbf{J}(X, L)$. Then we have the equality of schemes*

$$\mathbf{J}(TC_x X, \tilde{L}) = TC_x \mathbf{J}(X, L).$$

5.4. Degrees of the varieties $\mathbf{J}(\sigma_r, L^S)$.

Lemma 5.4.1. *Let S be such that no entries of S lie in a same column or row, and let $x \in S$. Assume $s < (n-r)^2$, and let $S' = S \setminus x$. Let $\pi : \mathbf{J}(\sigma_r, L^{S'}) \dashrightarrow \mathbb{P}^{n^2-2}$ denote the projection from $[a \otimes b]$, where $a \otimes b$ is the matrix having 1 at the entry x and 0 elsewhere. Then $\deg(\pi) = 1$.*

Proof. We need to show a general line through $[a \otimes b]$ that intersects $\mathbf{J}(\sigma_r, L^{S'})$, intersects it in a unique point. Without loss of generality, take S to be the first s diagonal entries and $x = x_1^1$. It will be sufficient to show that there exist $A \in \hat{\sigma}_r$ and $M \in \hat{L}^{S'}$ such that are no elements $B \in \hat{\sigma}_r, F \in \hat{L}^{S'}$ such that $u(A + M) + va_1 \otimes b^1 = B + F$ for some $u, v \neq 0$ other than when $[B] = [A]$. Assume A has no entries in the first row or column, so, moving F to the other side of the equation, in order that the corresponding B has rank at most r , there must be a matrix D with entries in S' , such that $A + D$ with the first row and column removed has rank at most $r - 1$.

If $r \leq \lfloor \frac{n}{2} \rfloor - 1$, take A to be the matrix $\sum_{j=1}^r a_{\lfloor \frac{n}{2} \rfloor + 1 + j} \otimes b^{j+1}$. Then the determinant of a size r submatrix in the lower left quadrant of A is always 1.

If $\lfloor \frac{n}{2} \rfloor - 1 < r \leq n - 2$, take $A = \sum_{j=1}^{\lfloor \frac{n}{2} \rfloor - 1} a_{\lfloor \frac{n}{2} \rfloor + 1 + j} \otimes b^{j+1} + \sum_{i=1}^{r - \lfloor \frac{n}{2} \rfloor - 1} a_{i+1} \otimes b^{\lfloor \frac{n}{2} \rfloor + 1 + i}$. Then the size r minor consisting of columns $\{2, 3, \dots, r+1\}$ and rows $\{\lfloor \frac{n}{2} \rfloor + 2, \lfloor \frac{n}{2} \rfloor + 3, \dots, n, 2, 3, \dots, r - \lfloor \frac{n}{2} \rfloor + 2\}$ is such that its determinant is also always ± 1 , independent of choice of D . \square

Let $A = \mathbb{C}^n$ with basis a_1, \dots, a_n , and let $A' = \langle a_2, \dots, a_n \rangle$, and similarly for $B = \mathbb{C}^n$. Let $x = [x_1^1]$. It is a standard fact (see, e.g. [9, p 257]), that

$$TC_x \sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)) = \mathbf{J}(\mathbb{P}\hat{T}_x \sigma_1(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), \sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B'))),$$

so by Proposition 5.3.2

$$TC_x(\mathbf{J}(\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), L^{S'})) \supseteq \mathbf{J}(\mathbb{P}\hat{T}_x \sigma_1(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), \mathbf{J}(\sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B')), L^{S'})).$$

Since $\mathbb{P}\hat{T}_x \sigma_1(\text{Seg}(\mathbb{P}A \times \mathbb{P}B))$ is a linear space and $\mathbf{J}(\sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B')), L^{S'})$ lies in a linear space disjoint from it,

$$\deg \mathbf{J}(\mathbb{P}\hat{T}_x \sigma_1(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), \mathbf{J}(\sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B')), L^{S'})) = \deg \mathbf{J}(\sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B')), L^{S'})$$

because if L is a linear space and Y any variety and $L \cap Y = \emptyset$, then $\deg J(Y, L) = \deg Y$.

Thus if

$$(5.4.2) \quad \dim(TC_x(\mathbf{J}(\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), L^{S'})) = \dim \mathbf{J}(TC_x \sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), L^{S'}),$$

we obtain

$$(5.4.3) \quad \deg TC_x(\mathbf{J}(\sigma_r(\text{Seg}(\mathbb{P}A \times \mathbb{P}B)), L^{S'})) \geq \deg \mathbf{J}(\sigma_{r-1}(\text{Seg}(\mathbb{P}A' \times \mathbb{P}B')), L^{S'}).$$

Recall the notation $d(n, r, s) := \deg \mathbf{J}(\sigma_r, L^S)$ where L^S is general with $|S| = s$. In particular $d(n, r, 0) = \deg(\sigma_r(\text{Seg}(\mathbb{P}^{n-1} \times \mathbb{P}^{n-1})))$.

Proposition 5.4.4. *Let S be such that no two elements of S lie in the same row or column. Then*

$$(5.4.5) \quad d_{n,r,s} \leq d_{n,r,0} - \sum_{j=1}^s d_{n-1,r-1,s-j}$$

Proof. In this situation the equality (5.4.2) holds, and Lemma 5.4.1 says the degree of π in Proposition 5.3.1 equals one, so apply it and equation (5.4.3) iteratively to obtain the inequalities $d_{n,r,t} \leq d_{n,r,t-1} - d_{n-1,r-1,t-1}$. \square

As mentioned in the introduction, P. Aluffi [1] proved that equality holds in (5.4.5). This has the following consequence, which was stated as a conjecture in an earlier version of this paper:

Proposition 5.4.6. *Let S be such that no two elements of S lie in the same row or column and let $x \in S$. Then*

$$TC_x \mathbf{J}(\sigma_r, L^S) = \mathbf{J}(TC_x \sigma_r, L^{S'}).$$

In particular $TC_x \mathbf{J}(\sigma_r, L^S)$ is reduced and irreducible.

Theorem 5.4.7. *Each irreducible component of $\hat{\mathcal{R}}[n, n-k, s]$ has degree at most*

$$(5.4.8) \quad \sum_{m=0}^s \binom{s}{m} (-1)^m d_{r-m, n-m, 0}$$

with equality holding if no two elements of S lie in the same row or column, e.g., if the elements of S appear on the diagonal.

Moreover, if we set $r = n - k$ and $s = k^2 - u$ and consider the degree $D(n, k, u)$ as a function of n, k, u , then, fixing k, u and considering $D_{k,u}(n) = D(n, k, u)$ as a function of n , it is of the form

$$D_{k,u}(n) = (k^2)! \frac{B(k)^2}{B(2k)} p(n)$$

where $p(n) = \frac{n^u}{u!} - \frac{k^2-u}{2(u-1)!} n^{u-1} + O(n^{u-2})$ is a polynomial of degree u .

For example:

$$D(n, k, 1) = \frac{(k^2)! B(k)^2}{B(2k)} \left(n - \frac{1}{2}(k^2 - 1) \right),$$

$$D(n, k, 2) = \frac{(k^2)! B(k)^2}{B(2k)} \left(\frac{1}{2} n^2 - \frac{1}{2}(k^2 - 2)n + \frac{1}{6} \left(\frac{3}{4} k^4 - \frac{11}{4} k^2 + 2 \right) \right).$$

Proof. Apply induction on all terms of (5.4.5). We get

$$\begin{aligned}
d_{n,r,s} &= d_{n,r,s-1} - d_{n-1,r-1,s-1} \\
&= (d_{n,r,s-2} - d_{n-1,r-1,s-2}) - (d_{n-1,r-1,s-2} - d_{n-2,r-2,s-2}) \\
&= d_{n,r,s-3} - 3d_{n-1,r-1,s-3} + 3d_{n-2,r-2,s-3} + d_{n-3,r-3,s-3} \\
&\vdots \\
&= \sum_{m=0}^{\ell} (-1)^m \binom{\ell}{m} d_{n-m,r-m,s-\ell},
\end{aligned}$$

for any $\ell \leq s$, in particular, for $\ell = s$.

To see the second assertion, note that

$$\sum_{m=0}^s \binom{s}{m} (-1)^m q(m) = 0$$

where $q(m)$ is any polynomial of degree less than s and

$$\begin{aligned}
\sum_{m=0}^s \binom{s}{m} (-1)^m m^s &= s! (-1)^s \\
\sum_{m=0}^s \binom{s}{m} (-1)^m m^{s+1} &= s! \binom{s-1}{2} (-1)^s.
\end{aligned}$$

(See, e.g. [25, §1.4], where the relevant function is called $S(n, k)$.)

Consider

$$\begin{aligned}
&\sum_{m=0}^s \binom{s}{m} (-1)^m d_{r-m,n-m,0} \\
&= \sum_{m=0}^s \binom{s}{m} (-1)^m \frac{B(k)^2 B(n-m+k) B(n-m-k)}{B(2k) B(n-m)^2} \\
&= \frac{B(k)^2}{B(2k)} \sum_{m=0}^s \binom{s}{m} (-1)^m \frac{B(n-m+k) B(n-m-k)}{B(n-m)^2} \\
&= \frac{B(k)^2}{B(2k)} \sum_{m=0}^s \binom{s}{m} (-1)^m (n-m)^k \prod_{t=1}^{k-1} (n-m+k-t)^t (n-m-k+t)^t
\end{aligned}$$

Write $(n-m)^k \prod_{t=1}^{k-1} (n-m+k-t)^t (n-m-k+t)^t = \sum_j c_{k,n,j} m^j$, then all values of j less than $s = k^2 - u$ contribute zero to the sum, the $j = s$ case gives $c_{k,n,k^2-u} (k^2 - u)! (-1)^{k^2-u}$. Now consider the highest power of n in c_{k,n,k^2-u} . $\sum_j c_{k,n,j} m^j$ is a product of $k+2 \binom{k}{2} = k^2$ linear forms, if we use $k^2 - u + t$ of them for the m , there will be $u - t$ to which n can contribute, so the only term with n^u can come from the case $t = 0$, in which case the coefficient of $n^u m^{k^2-u}$ is $(-1)^{k^2-u} \binom{k^2}{u}$. Putting it all together, we obtain the coefficient. The next highest power, n^{u-1} a priori could appear in two terms: c_{k,n,k^2-u} , but there the coefficient is $\binom{k^2}{u} [\sum_{t=1}^{k-1} (k-t) + \sum_{t=1}^{k-1} (-k+t)] = 0$, and c_{k,n,k^2-u+1} , where the total contribution is

$$\binom{k^2}{u-1} \binom{k^2 - u + 1}{2} (k^2 - u)! = \frac{k^2!}{(u-1)!} \frac{k^2 - u}{2}.$$

□

REFERENCES

1. Paolo Aluffi, *Degrees of projections of rank loci*, preprint arXiv:1408.1702.
2. E. Arbarello, M. Cornalba, P. A. Griffiths, and J. Harris, *Geometry of algebraic curves. Vol. I*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 267, Springer-Verlag, New York, 1985. MR 770932 (86h:14019)
3. Bruno Codenotti, *Matrix rigidity*, Linear Algebra Appl. **304** (2000), no. 1-3, 181–192. MR 1734200 (2000j:15002)
4. Ronald de Wolf, *Lower bounds on matrix rigidity via a quantum argument*, Automata, languages and programming. Part I, Lecture Notes in Comput. Sci., vol. 4051, Springer, Berlin, 2006, pp. 62–71. MR 2305515 (2007m:68094)
5. Zeev Dvir, *On matrix rigidity and locally self-correctable codes*, Comput. Complexity **20** (2011), no. 2, 367–388. MR 2822875 (2012h:68096)
6. Joel Friedman, *A note on matrix rigidity*, Combinatorica **13** (1993), no. 2, 235–239. MR 1237045 (94e:15043)
7. William Fulton, *Intersection theory*, second ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 2, Springer-Verlag, Berlin, 1998. MR 1644323 (99d:14003)
8. William Fulton and Joe Harris, *Representation theory*, Graduate Texts in Mathematics, vol. 129, Springer-Verlag, New York, 1991, A first course, Readings in Mathematics. MR 1153249 (93a:20069)
9. Joe Harris, *Algebraic geometry*, Graduate Texts in Mathematics, vol. 133, Springer-Verlag, New York, 1995, A first course, Corrected reprint of the 1992 original. MR 1416564 (97e:14001)
10. B. S. Kashin and A. A. Razborov, *New lower bounds for the stability of Hadamard matrices*, Mat. Zametki **63** (1998), no. 4, 535–540. MR 1680943 (2000c:05033)
11. Abhinav Kumar, Satyanarayana V. Lokam, Vijay M. Patankar, and Jayalal Sarma M. N., *Using elimination theory to construct rigid matrices*, Foundations of software technology and theoretical computer science—FSTTCS 2009, LIPIcs. Leibniz Int. Proc. Inform., vol. 4, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2009, pp. 299–310. MR 2870721
12. J. M. Landsberg, *Tensors: geometry and applications*, Graduate Studies in Mathematics, vol. 128, American Mathematical Society, Providence, RI, 2012. MR 2865915
13. JM Landsberg, J. Taylor, and N.K. Vishnoi, *The geometry of matrix rigidity*, preprint <https://smartech.gatech.edu/handle/1853/6514>.
14. F. Thomson Leighton, *Introduction to parallel algorithms and architectures*, Morgan Kaufmann, San Mateo, CA, 1992, Arrays, trees, hypercubes. MR 1137272 (92j:68041)
15. Satyanarayana V. Lokam, *On the rigidity of Vandermonde matrices*, Theoret. Comput. Sci. **237** (2000), no. 1-2, 477–483. MR 1756225 (2001d:15021)
16. ———, *Spectral methods for matrix rigidity with applications to size-depth trade-offs and communication complexity*, J. Comput. System Sci. **63** (2001), no. 3, 449–473. MR 1892859 (2003a:68059)
17. ———, *Quadratic lower bounds on matrix rigidity*, Theory and applications of models of computation, Lecture Notes in Comput. Sci., vol. 3959, Springer, Berlin, 2006, pp. 295–307. MR 2277251 (2007h:68087)
18. ———, *Complexity lower bounds using linear algebra*, Found. Trends Theor. Comput. Sci. **4** (2008), no. 1-2, front matter, 1–155 (2009). MR 2539154 (2011c:68060)
19. Meena Mahajan and Jayalal Sarma M. N., *Rigidity of a simple extended lower triangular matrix*, Inform. Process. Lett. **107** (2008), no. 5, 149–153. MR 2436904 (2009k:68094)
20. David Mumford, *Algebraic geometry. I*, Classics in Mathematics, Springer-Verlag, Berlin, 1995, Complex projective varieties, Reprint of the 1976 edition. MR 1344216 (96d:14001)
21. ———, *The red book of varieties and schemes*, expanded ed., Lecture Notes in Mathematics, vol. 1358, Springer-Verlag, Berlin, 1999, Includes the Michigan lectures (1974) on curves and their Jacobians, With contributions by Enrico Arbarello. MR 1748380 (2001b:14001)
22. Nicholas Pippenger, *Superconcentrators*, SIAM J. Comput. **6** (1977), no. 2, 298–304. MR 0446750 (56 #5074)
23. Pavel Pudlák and Zdeněk Vavřín, *Computation of rigidity of order n^2/r for one simple matrix*, Comment. Math. Univ. Carolin. **32** (1991), no. 2, 213–218. MR 1137781 (93b:15001)
24. M. A. Shokrollahi, D. A. Spielman, and V. Stemann, *A remark on matrix rigidity*, Inform. Process. Lett. **64** (1997), no. 6, 283–285. MR 1608240
25. Richard P. Stanley, *Enumerative combinatorics. Volume 1*, second ed., Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 2012. MR 2868112

26. Harry Tamvakis, *The connection between representation theory and Schubert calculus*, Enseign. Math. (2) **50** (2004), no. 3-4, 267–286. MR 2116717 (2006a:20086)
27. Leslie G. Valiant, *Graph-theoretic arguments in low-level complexity*, Mathematical foundations of computer science (Proc. Sixth Sympos., Tatranská Lomnica, 1977), Springer, Berlin, 1977, pp. 162–176. Lecture Notes in Comput. Sci., Vol. 53. MR 0660702 (58 #32067)

E-mail address: fulges@math.tamu.edu, hauenstein@nd.edu, ciken@math.tamu.edu, jml@math.tamu.edu