Journal of Pure and Applied Algebra 224 (2020) 106351

Contents lists available at ScienceDirect

Journal of Pure and Applied Algebra

www.elsevier.com/locate/jpaa

The algebraic matroid of the finite unit norm tight frame (funtf) variety

Daniel Irving Bernstein [a], Cameron Farnsworth [b,*], Jose Israel Rodriguez [c]

[a] *Institute for Data, Systems, and Society, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA*
[b] *Department of Mathematics, Texas State University, 601 University Drive, San Marcos, TX 78666-4684, USA*
[c] *Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, WI 53706-1388, USA*

ARTICLE INFO

Article history:
Received 30 January 2019
Received in revised form 10 January 2020
Available online 11 February 2020
Communicated by G.G. Smith

MSC:
05B35

Keywords:
Algebraic matroid
Frame completion

ABSTRACT

A finite unit norm tight frame is a collection of r vectors in $\mathbb{R}^n$ that generalizes the notion of orthonormal bases. The affine finite unit norm tight frame variety is the Zariski closure of the set of finite unit norm tight frames. Determining the fiber of a projection of this variety onto a set of coordinates is called the algebraic finite unit norm tight frame completion problem. Our techniques involve the algebraic matroid of an algebraic variety, which encodes the dimensions of fibers of coordinate projections. This work characterizes the bases of the algebraic matroid underlying the variety of finite unit norm tight frames in $\mathbb{R}^3$. Partial results towards similar characterizations for finite unit norm tight frames in $\mathbb{R}^n$ with $n \geq 4$ are also given. We provide a method to bound the degree of the projections based off of combinatorial data.

1. Introduction

A set of unit vectors $\{w_1, \ldots, w_r\}$ in $\mathbb{R}^n$ is said to be a *finite unit norm tight frame* (or *funtf*) if the matrix $W := [\, w_1 \; w_2 \; \cdots \; w_r \,]$ satisfies $\frac{n}{r} WW^T = \mathrm{Id}_n$, where Id_n is the $n \times n$ identity matrix. This generalizes the notion of orthonormal bases for $\mathbb{R}^n$. The Zariski closure (over $\mathbb{C}$) of the set of $n \times r$ matrices that are finite unit norm tight frames forms an algebraic variety in $\mathbb{C}^{n \times r}$, which we denote by $X_{n,r}$. The variety $X_{n,r}$ can be expressed as an algebraic set (see [12, p. 4] and [35, Equation 4]) as follows:

$$X_{n,r} = \{W \in \mathbb{C}^{n \times r} : WW^T = \frac{r}{n} \mathrm{Id}_n, \, \mathrm{diag}(W^T W) = \mathrm{diag}(\mathrm{Id}_r)\}, \tag{1.1}$$

* Corresponding author.
E-mail addresses: dibernst@mit.edu (D.I. Bernstein), clf129@txstate.edu (C. Farnsworth), jose@math.wisc.edu (J.I. Rodriguez).
URLs: https://dibernstein.github.io (D.I. Bernstein), http://www.math.wisc.edu/~jose/ (J.I. Rodriguez).

https://doi.org/10.1016/j.jpaa.2020.106351
0022-4049/

where diag(M) denotes the diagonal entries of a matrix M. In this paper, we characterize the algebraic matroid underlying $X_{3,r}$ and provide additional partial results for $X_{n,r}$.

The remainder of this section describes the applied motivation for our work here and gives a literature review of the area. In Section 2, we provide the minimum necessary background on algebraic matroids. In Section 3 we set up our notation and collect previous results we will need about $X_{n,r}$. In particular, we discuss how matroids arise in connection with algebraic frame completion problems. In Section 4, we provide our main results on the algebraic matroid underlying $X_{n,r}$ (Theorem 4.2). Section 5 gives a recursive formula for computing the degree of a finite-to-one coordinate projection of $X_{n,r}$ (Theorem 5.2) which we then use to completely characterize the degrees of projection onto a basis of $X_{3,r}$.

1.1. Algebraic frame theory

Frames generalize the notion of a basis of a vector space and have found use in numerous fields of science and engineering. Given a Hilbert space H, a *frame* is a set of elements $\{f_k\}_{k\in I} \subset H$ such that there exist real numbers A, B such that $0 < A \leq B < \infty$ and for every $h \in H$

$$A\|h\|^2 \leq \sum_{k\in I} |\langle h, f_k\rangle|^2 \leq B\|h\|^2.$$

These frame conditions are given by Duffin and Schaeffer in [16, Equation (4)]. If $A = B$, then the frame is called *tight*. If H is n-dimensional, then any frame has at least n elements. A frame where each element has norm one is said to be a *unit norm frame*. In the literature unit norm frames are also known as normalized frames, uniform frames, and spherical frames. Throughout this text, we identify the sequence of vectors in a frame with the matrix whose columns are the vectors in the frame. The frames we are interested in are frames of the Hilbert space $\mathbb{R}^n$; however, it is important to keep in mind that we are making a relaxation to a question regarding an algebraic variety in $\mathbb{C}^{n\times r}$.

A finite frame which is both tight and unit norm is also called a *finite unit norm tight frame* and is commonly abbreviated in the literature as *funtf*. Such frames are the focus of much research because they minimize various measures of error in signal reconstruction [15,22,24,27]. Algebraic frame theory uses the powerful tools of computational algebraic geometry to solve problems involving finite frame varieties. Such approaches have found success in [11,17,35,40].

Given an $n \times r$ matrix where only a subset of the entries is observed, the *finite unit norm tight frame completion problem* asks for values of the missing entries such that the resulting completed matrix is a funtf. The jumping off point for this work is the relaxation of this problem that allows for the missing entries to take on complex values. We call this relaxation the *algebraic finite unit norm tight frame completion problem*.

Complex frames are also studied where a Hermitian inner product is used, but that is not the focus of this article. Studying the variety $X_{n,r}$ in place of the set of finite unit norm tight frames gives one access to tools from algebraic geometry, and results about $X_{n,r}$ can lead to insight about the set of finite unit norm tight frames—see for example [12]. Many works have studied the properties of various sets of frames considered as varieties. For example, dimensions of (μ, S)-frame varieties, which are spaces of matrices $W = [w_1 \cdots w_r]$, real or complex, satisfying $WW^* = S$ for some Hermitian (symmetric) positive definite matrix S such that $\|w_k\| = \mu_k$, were considered in [39]. Finite unit norm tight frames are a special case of these (μ, S)-frames where $\mu_i = 1$ and S is a scalar multiple of the identity matrix. Along with the fundamental groups, the dimensions of finite unit norm tight frame varieties were derived in [17].

In [40], nonsingular points of (μ, S)-frames are characterized along with the tangent spaces at these nonsingular points on these varieties. The connectivity of the finite unit norm tight frame variety along with its irreducibility are studied in [11]. In [26], the polytope of eigensteps of finite equal norm tight frames

is studied. These *eigensteps* are sequences of interlacing spectra used by [10] to construct finite frames of prescribed norms and the dimension of finite unit norm tight frame varieties is noted to be related to the dimension of these polytopes.

1.2. Algebraic finite unit norm tight frame completion

Due to their robustness to erasures and additive noise, unit norm tight frames play an important role in signal processing. Explicit constructions for unit norm tight frames are quite recent despite theoretical work regarding existence being quite classical. The *Schur-Horn Theorem* [28,38] characterizes the pairs (λ, μ) such that there exists a frame whose frame operator has spectrum λ and lengths μ. However, explicit constructions for these frames have remained scarce. In [23] the authors give a constructive characterization for all unit norm tight frames in $\mathbb{R}^2$ and provide a construction technique known as *harmonic frames* for unit norm tight frames in $\mathbb{R}^n$. An alternative constructive technique called *spectral tetris* is given in [14]. An explicit construction of every unit norm tight frame was finally given by [10,21].

The previous paragraph covers results on explicitly constructing frames with prescribed spectrum and whose vectors' lengths are prescribed. However, what if you have specific vectors you want included in your frame? How do you complete this partial set of vectors into a tight frame? The work [19] answers how many vectors must be added to complete your set of vectors into a tight frame, and in the case when all vectors are unit norm, they also provide a lower bound (which is not sharp) for the number of vectors required to complete the set of vectors into a tight frame. The minimum number of vectors needed to add to your set of vectors to complete it to a frame when their norms are prescribed is provided by [33]. In both papers, it is assumed that you start with a set of vectors. See also [20], in which the authors characterize the spectra of all frame operators of frames completed from the addition of extra vectors to another frame and discuss completion in such a way as to produce a frame minimizing the condition number, the mean squared reconstruction error, and the frame potential.

In this paper, we take a different approach than Feng, Wang, and Wang, Massey and Ruiz, or Fickus, Marks, and Poteet. Instead of starting with a set of vectors and asking how many more vectors are needed to have a tight frame, we have the following generalization of the problem.

Problem 1.1. Given a partially observed $n \times r$ matrix, determine if the missing entries can be completed such that the columns form a finite unit norm tight frame.

A generalization of Problem 1.1 is to ask how many completions there are.

Problem 1.2. Given some known entries of an $n \times r$ matrix, find the cardinality of the number of completions such that the columns form a finite unit norm tight frame.

When studying the algebraic geometry of low-rank matrix completion, one often studies the analogue of Problems 1.1 and 1.2, where instead of completing to a finite unit norm tight frame, one completes to a matrix of a particular specified low rank [31,7,3,5,6].

Problem 1.2 can be studied using algebraic geometry by considering the known entries as a set of defining equations of an affine linear space $\mathcal{L}$ in $\mathbb{C}^{n \times r}$ and then studying the degree of $X_{n,r} \cap \mathcal{L}$. When $X_{n,r} \cap \mathcal{L}$ is a finite set of points, the degree of $X_{n,r} \cap \mathcal{L}$ is the number of points, counted with multiplicity, and is an upper bound for the number of completions. The degree of $X_{n,r} \cap \mathcal{L}$ need not coincide with the degree of $X_{n,r}$, but when the linear space $\mathcal{L}$ is generic such that $X_{n,r} \cap \mathcal{L} \neq \emptyset$, the degree of $X_{n,r} \cap \mathcal{L}$ and $X_{n,r}$ are the same. For $n = r$, the variety $X_{n,r}$ is the orthogonal group and the degree of this variety was determined in [9] using representation theory. An analogous question from rigidity theory asks for the number of realizations of a rigid graph. This same algebraic approach for getting an upper bound works in this context, this time by computing or bounding the degree of (coordinate projections of) the Cayley-Menger variety [30,18,13,8].

After solving Problem 1.1, the most natural next problem is to actually compute such a completion, if it exists. One can attempt to do this using Gröbner bases algorithms, but this is likely to be prohibitively slow. A faster way to do this would be using numerical algebraic geometry using e.g. `Bertini` [2], `PHCPack` [41], or `Macaulay2` [32,25]. Knowing the degree of coordinate projections of $X_{n,r}$ could be helpful in this approach. One could also try using the combinatorics of the set of known entries to derive polynomials in the missing entries that could be solved to complete the finite unit norm tight frame. In the language of algebraic matroids, this is the problem of finding combinatorial descriptions of the circuit polynomials of $X_{n,r}$. This would be an interesting future direction.

2. The basics of algebraic matroids

We now take a detour to introduce the minimum necessary background on algebraic matroids. Since the only matroids considered in this paper will be algebraic, we will not discuss or define abstract matroids. Moreover, our study will be limited to those that are algebraic over $\mathbb{R}$ or $\mathbb{C}$. The reader who is interested in learning about more general (algebraic) matroids is advised to consult the textbook [36].

Let $\mathbb{K}$ be a field. Given a finite set E, we let $\mathbb{K}^E$ denote the vector space whose coordinates are indexed by the elements of E. Each subset $S \subseteq E$ of coordinates is associated with the linear projection $\pi_S : \mathbb{K}^E \to \mathbb{K}^S$ that sends each point $(x_e)_{e\in E}$ to $(x_e)_{e\in S}$. The ring of polynomials with coefficients in $\mathbb{K}$ and indeterminants indexed by E will be denoted $\mathbb{K}[x_e : e \in E]$ and the corresponding field of rational functions will be denoted $\mathbb{K}(x_e : e \in E)$. The ideal in $\mathbb{K}[x_e : e \in E]$ generated by a finite set of polynomials $f_1, \ldots, f_k \in \mathbb{K}[x_e : e \in E]$ will be notated as $(f_1, \ldots, f_k)$. Given a set $X \subseteq \mathbb{K}^E$, we let $I(X)$ denote the ideal of all polynomial functions that vanish on X.

Definition 2.1. Let E be a finite set, let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$, and let $X \subseteq \mathbb{K}^E$ be an irreducible variety. A subset of coordinates $S \subseteq E$ is

(1) *independent* in X if $I(\pi_S(X))$ is the zero ideal,
(2) *spanning* in X if $\dim(\pi_S(X)) = \dim(X)$,
(3) a *basis* of X if S is both independent and spanning.

Any one of the three set systems consisting of the independent sets, the spanning sets, or the bases of an irreducible variety determines the other two. The combinatorial structure specified by any one of these set systems is called the *algebraic matroid underlying* X.

Example 2.2. Let $X \subset \mathbb{R}^{[4]}$ be the linear variety defined by the vanishing of the linear forms $x_1 - 5x_2 = 0$ and $x_3 + 2x_4 = 0$. The bases of X are

$$\{1,3\}, \quad \{1,4\}, \quad \{2,3\}, \quad \{2,4\}.$$

The independent sets of X are the subsets of the bases, and the spanning sets are the supersets. Moreover, all the bases have cardinality two, which is also the dimension of X. This is not a coincidence—see Proposition 2.3 below.

We now describe the intuition behind the algebraic matroid underlying an irreducible variety $X \subseteq \mathbb{K}^E$. When $S \subseteq E$ is independent, the coordinates $(x_e)_{e\in S}$ can be given arbitrary generic values, and the resulting vector can be completed to a point in X. When $S \subseteq E$ is spanning and $x \in X$ is generic, then the coordinates $(x_e)_{e\notin S}$ can be determined by solving a zero-dimensional system of polynomials whose coefficients are polynomials in $(x_e)_{e\in S}$. In other words, the set $\pi_S^{-1}(\pi_S(x)) \cap X$ is generically finite.

It's important to recall Definition 2.1 requires X to be an irreducible variety. This ensures that the algebraic matroid underlying X is indeed a matroid (see for example [4, Proposition 1.2.9]). Proposition 2.3 below then follows from the fact that all bases of a matroid have the same size [36, Chapter 1].

Proposition 2.3. *For any basis $B \subseteq E$ of an irreducible variety $X \subseteq \mathbb{K}^E$, the cardinality of the set B equals* $\dim(X)$.

Given finite sets A and B and a field $\mathbb{K}$, we let $\mathbb{K}^{A\times B}$ denote the set of matrices with entries in $\mathbb{K}$ whose rows are indexed by elements of A and whose columns are indexed by elements of B. Given polynomials $f_1, \ldots, f_k \in \mathbb{K}[x_e : e \in E]$ the *Jacobian matrix of* $f_1, \ldots, f_k$ is the matrix $J(f_1, \ldots, f_k) \in (\mathbb{K}(x_e : e \in E))^{[k]\times E}$ whose (i, e) entry is the partial derivative $\frac{\partial f_i}{\partial x_e}$. In Section 4, we will often work with submatrices of a Jacobian matrix. For this reason, we introduce the following notation.

Notation 2.4. Let M denote a matrix whose columns are indexed by a set E. The submatrix of a given M with columns corresponding to the elements of a subset S of E is denoted M_S.

The following proposition is useful for computing the bases of the algebraic matroid underlying a given irreducible variety. It is well known, and usually stated in terms of matroid duals. We state it here in more elementary terms for the purposes of keeping the necessary matroid theory background at a minimum. For more details, see [37, Section 2.2] and [36, Proposition 6.7.10].

Proposition 2.5. *Let E be a finite set, let $\mathbb{K}$ be $\mathbb{R}$ or $\mathbb{C}$ and let $X \subseteq \mathbb{K}^E$ be an irreducible variety of dimension d such that $I(X) = (f_1, \ldots, f_k)$. A subset $S \subseteq E$ of size d is a basis of X if and only if the rank of $J(f_1, \ldots, f_k)_{E\setminus S}$ is $|E| - d$.*

3. Algebraic matroids to algebraic funtf completion

Recall from equation (1.1) that the Zariski closure in $\mathbb{C}^{n\times r}$ of the set of $n \times r$ matrices such that the columns form a finite unit norm tight frame is denoted by $X_{n,r}$. In other words, this paper studies the following algebraic relaxation of Problem 1.1.

Problem 3.1 *(The algebraic frame completion problem).* Given some known entries of an $n \times r$ matrix, determine if the matrix can be completed to an element of $X_{n,r}$.

We say $X_{n,r}$ is an *affine finite unit norm tight frame variety* and call a matrix in $X_{n,r}$ a *finite unit norm tight frame (funtf) matrix*. The $\binom{n+1}{2} + r$ scalar equations defining $X_{n,r}$ shown in (1.1) were found in [12, p. 4] and [35, Equation 4]. We will express the polynomials defining the affine finite unit norm tight frame variety in the ring $\mathbb{K}[x_{ij} : 1 \le i \le n, 1 \le j \le r]$ where x_{ij} will represent the ij entry of a matrix. Indeed, the column norm constraints on W can be expressed as the following r polynomials set to zero:

$$(g_1, \ldots, g_r) := \operatorname{diag}(W^T W - \operatorname{Id}_r), \tag{3.1}$$

while the orthogonal row constraints on W can be expressed as the following $\binom{n+1}{2}$ polynomials f_{ij}, $i \le j$:

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{12} & f_{22} & \cdots & f_{2n} \\ \vdots & & \ddots & \vdots \\ f_{1n} & f_{2n} & \cdots & f_{nn} \end{bmatrix} := WW^T - \frac{r}{n}\operatorname{Id}_n . \tag{3.2}$$

Elementary algebra shows that these polynomials satisfy the following relation

$$\sum_{i=1}^{n} f_{i,i} = \sum_{j=1}^{r} g_j = ||W||_F^2 - r, \tag{3.3}$$

where $||\cdot||_F$ denotes the Frobenius norm.

The problem of algebraic finite unit norm tight frame completion can be cast as the problem of projecting an affine finite unit norm tight frame variety to a subset of coordinates. Let $E \subseteq [n] \times [r]$ denote a subset of coordinates of $\mathbb{C}^{n\times r}$. We will think of E as indexing "known" entries, and the algebraic finite unit norm tight frame completion problem is to determine the remaining "unknown" entries so that the completed matrix is a finite unit norm tight frame. Let π_E denote the respective coordinate projection. The algebraic finite unit norm tight frame completions of a given $M \in \mathbb{C}^E$ are the elements of the fiber $\pi_E^{-1}(M)$. It follows that E is independent in $X_{n,r}$ if and only if every generic $M \in \mathbb{C}^E$ has an algebraic finite unit norm tight frame completion and that E is spanning in $X_{n,r}$ if and only if each nonempty fiber $\pi^{-1}(\pi_E(M))$ is finite when M is generic. Thus in the generic case, Problem 3.1 is equivalent to the following.

Problem 3.2. Find a combinatorial description of the algebraic matroid underlying $X_{n,r}$.

The first steps towards solving Problem 3.2 are determining the irreducibility and dimension of $X_{n,r}$. Fortunately, this was done in [17, Theorem 4.3(ii)], [40, Corollary 3.5], and [11, Theorem 1.4] which we summarize in the following theorem.

Theorem 3.3. *The dimension of the affine finite unit norm tight frame variety $X_{n,r}$ is*

$$\dim(X_{n,r}) = nr - \binom{n+1}{2} - r + 1 \qquad \text{provided } r > n \geq 2. \tag{3.4}$$

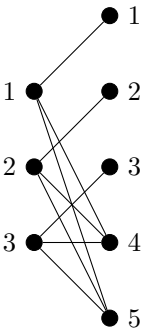
It is irreducible when $r \geq n + 2 > 4$.

In our work, we look to determine each basis (Definition 2.1, item 3) of $X_{n,r}$. We restrict our study to $r \geq n + 2 > 4$ so that $X_{n,r}$ is irreducible and thus gives a matroid. We seek a combinatorial description using bipartite graphs. Bipartite graphs provide a natural language for attacking Problem 3.2. Given finite sets A and B and a subset $S \subseteq A \times B$, we let (A, B, S) denote the *bipartite graph* with partite vertex sets A and B and edge set S. We call two bipartite graphs (A_1, B_1, S_1) and (A_2, B_2, S_2) *bipartite isomorphic* if there exists a graph isomorphism $\phi : A_1 \cup B_1 \to A_2 \cup B_2$ such that $\phi(A_1) = A_2$ and $\phi(B_1) = B_2$.

Every subset E of entries of an $n \times r$ matrix can be identified with the bipartite graph $([n], [r], [n]\times[r]\setminus E)$, which we denote by G_E (Fig. 1). The edges of G_E are in bijection with the complement of E and not E itself. This stands in contrast to what is often done in the algebraic matrix completion literature, but will make our results much cleaner to state. Neither row-swapping nor column-swapping affects whether a given subset E of entries of an $n \times r$ matrix is an independent set (or a basis, or spanning set) of $X_{n,r}$. Therefore, whether a given subset E of entries is independent (or a basis or spanning) in $X_{n,r}$ only depends on the bipartite isomorphism equivalence class of G_E. The (non-bipartite) graph isomorphism class of G_E may not be sufficient to determine whether E is independent (or a basis or spanning) in $X_{n,r}$ because the transpose of a finite unit norm tight frame matrix W may not be funtf. So from now on, we will only consider bipartite graphs up to their bipartite isomorphism classes. We may now phrase Problem 3.2 more concretely as follows.

Problem 3.4. For which (bipartite isomorphism classes of) bipartite graphs G_E is E a basis of $X_{n,r}$?

We will sometimes find it useful to represent a subset $E \subseteq [n] \times [r]$ as the $\{0, 1\}$-matrix whose ij entry is 1 if $(i, j) \in E$ and 0 otherwise (Fig. 1). Such a representation will be called a *matrix entry representation*.



$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Fig. 1. Depicting $E := \{(1,2),(1,3),(2,1),(2,3),(3,1),(3,2)\} \subset [3] \times [5]$ as the bipartite graph G_E and as a $\{0,1\}$-matrix.

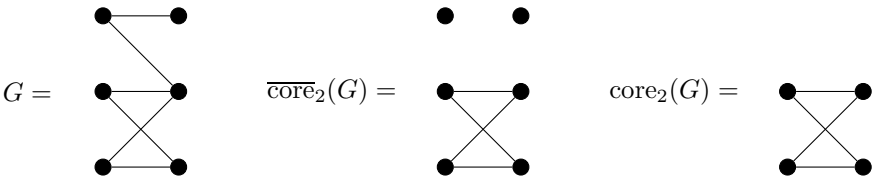


Fig. 2. A bipartite graph G alongside its greater 2-core and its 2-core.

4. The algebraic matroid underlying the finite unit norm tight frame variety

In this section we give combinatorial criteria on the bases of $X_{n,r}$. First we show that if E is a basis of $X_{n,r}$, then the graph G_E is connected. Moreover, when $n = 3$ the converse is true as well. Second, we show that whether or not E is a basis of $X_{n,r}$ only depends on the 2-core of G_E. This allows us to determine a combinatorial criterion for every r after fixing n.

4.1. Graph connectivity

We begin with some graph theoretic definitions. Let $G = (A, B, S)$ be a bipartite graph. The *greater 2-core of* G, denoted $\overline{\text{core}}_2(G)$, is the graph obtained from G by iteratively removing all edges that are incident to a vertex of degree one. The 2-*core of* G, denoted $\text{core}_2(G)$, is the graph obtained by deleting the isolated vertices from $\overline{\text{core}}_2(G)$. Fig. 2 shows a graph alongside its greater 2-core and its 2-core.

Since the graph G need not be connected, there may not exist a spanning tree for G. Instead, consider a *spanning forest* for G, which is a maximal acyclic subgraph of G, or equivalently a subgraph consisting of a spanning tree in each connected component. A *circuit* is a nonempty trail in G such that the first and last vertex coincide, or equivalently, a non-empty sequence $(e_1, \ldots, e_k)$ of edges in G for which there is a sequence of vertices $(v_1, \ldots, v_k, v_1)$ such that for $i = 1, \ldots, k-1$ the vertices of e_i are (v_i, v_{i+1}) and the vertices of e_k are (v_k, v_1). A cycle is an example of a circuit. Given a spanning forest F of G and an edge e of G not appearing in F, the graph $F \cup \{e\}$ has exactly one cycle which must contain e. This cycle is called *the fundamental circuit of* e *with respect to* F.

Given a subset $S' \subseteq S$ of the edge set of G, the *characteristic vector of* S' is the vector in $\{0,1\}^S$ that has ones at entries corresponding to elements of S' and zeros at all other entries. The *incidence matrix of* G is the matrix whose rows are indexed by the vertices of G, and the row corresponding to a vertex v is the characteristic vector of the set of edges that are incident to v. The columns of the incidence matrix of G are indexed by the edges of G.

Example 4.1. We use the notation $K_{a,b}$ to denote the complete bipartite graph on partite sets of size a and b. The incidence matrix of $K_{3,5}$ is given by the 8×15 matrix below

$$\left(\begin{array}{cccccccccccccc} 1&1&1&1&1&0&0&0&0&0&0&0&0&0\\ 0&0&0&0&0&1&1&1&1&1&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&1&1&1\\ \hline 1&0&0&0&0&1&0&0&0&0&1&0&0&0\\ 0&1&0&0&0&0&1&0&0&0&0&1&0&0\\ 0&0&1&0&0&0&0&1&0&0&0&0&1&0\\ 0&0&0&1&0&0&0&0&1&0&0&0&0&1\\ 0&0&0&0&1&0&0&0&0&1&0&0&0&0 \end{array}\right).$$

This matrix is naturally partitioned via the vertices in each partite of the graph $K_{3,5}$.

Theorem 4.2. *Assume $r \geq n+2 > 4$ and let $E \subseteq [n] \times [r]$ have cardinality $nr - \binom{n+1}{2} - r + 1$. If E is a basis of $X_{n,r}$, then G_E is connected. When $n = 3$, the converse is true as well.*

Proof. Let g_i and f_{ij} denote the polynomials as in (3.1) and (3.2); this set of polynomials generates the ideal of $X_{n,r}$. Let J be the $(r + \binom{n+1}{2}) \times nr$ Jacobian matrix

$$J := J(g_1, \ldots, g_r, f_{11}, f_{12} \ldots, f_{nn}).$$

Proposition 2.5 implies that E is a basis of $X_{n,r}$ if and only if the $(r + \binom{n+1}{2}) \times (r + \binom{n+1}{2} - 1)$ matrix $J_{[n]\times[r]\setminus E}$ has full rank, or equivalently,

$$\operatorname{rank}(J_{[n]\times[r]\setminus E}) = r + \binom{n+1}{2} - 1.$$

Let J' and J'' denote the following matrices of size $(r + \binom{n+1}{2}) \times nr$ and $(r+n) \times nr$, respectively

$$J' := J \cdot \operatorname{diag}(1/x_{11}, 1/x_{12} \ldots, 1/x_{nr}),$$
$$J'' := J(g_1, \ldots, g_r, f_{11}, f_{22}, \ldots, f_{ii}, \ldots, f_{nn}) \cdot \operatorname{diag}(1/x_{11}, 1/x_{12} \ldots, 1/x_{nr}),$$

where $\operatorname{diag}(w)$ denotes the matrix with the vector w along its diagonal. The matrix J'' is twice the incidence matrix of $K_{n,r}$ from Example 4.1, and $J''_{[n]\times[r]\setminus E}$ is twice the incidence matrix of G_E. Thus if G_E has c connected components, then

$$\operatorname{rank}(J''_{[n]\times[r]\setminus E}) = r + n - c.$$

On the other hand, since J' can be obtained from J'' by including $\binom{n}{2}$ additional rows, we have

$$\operatorname{rank}(J'_{[n]\times[r]\setminus E}) \leq \operatorname{rank}(J''_{[n]\times[r]\setminus E}) + \binom{n}{2} = r + \binom{n+1}{2} - c.$$

Since $\operatorname{rank}(J_{[n]\times[r]\setminus E}) = \operatorname{rank}(J'_{[n]\times[r]\setminus E})$, we have that if G_E is disconnected, then $J_{[n]\times[r]\setminus E}$ is rank deficient and thus E is not a basis.

Now, having proved that E being a basis implies connectivity of G_E, we assume that $n = 3$ and prove the converse. Further assume that G_E is connected with $\binom{n+1}{2} + r - 1 = r + 2$ edges. We will show that E is a basis of $X_{n,r}$ by showing that $J'_{[n]\times[r]\setminus E}$ has full rank. This is done by splitting $J'_{[n]\times[r]\setminus E}$ into two row submatrices whose kernels intersect trivially.

Twice the incidence matrix of the complete bipartite graph $K_{3,r}$ is a row-submatrix of J'. Therefore, any linear relation among the columns of J' must lie in the linear space

$$\mathbb{C}\{v_C : C \text{ is a circuit of } K_{3,r}\}$$

where $v_C \in \mathbb{C}^{[3]\times[r]}$ is the $\{1,-1,0\}$-vector obtained from the characteristic vector of C by giving adjacent edges opposite signs. The row of J' corresponding to the constraint $f_{ab} = 0$ with $a \neq b$ has x_{bi}/x_{ai} at the column corresponding to x_{ai} and x_{ai}/x_{bi} at the column corresponding to x_{bi}. For ease of notation, we introduce the change of variables

$$t_{1i} := x_{2i}/x_{1i} \qquad t_{2i} := x_{3i}/x_{1i}.$$

With this change of variables, the rows of J' corresponding to the constraints $f_{12} = f_{13} = f_{23} = 0$ form the matrix K shown below

$$K := \begin{array}{c} \\ f_{12} \\ f_{13} \\ f_{23} \end{array} \begin{array}{c} \begin{array}{ccccccccc} x_{11} & \dots & x_{1r} & x_{21} & \dots & x_{2r} & x_{31} & \dots & x_{3r} \end{array} \\ \begin{bmatrix} t_{11} & \dots & t_{1r} & t_{11}^{-1} & \dots & t_{1r}^{-1} & 0 & \dots & 0 \\ t_{21} & \dots & t_{2r} & 0 & \dots & 0 & t_{21}^{-1} & \dots & t_{2r}^{-1} \\ 0 & \dots & 0 & t_{21}t_{11}^{-1} & \dots & t_{2r}t_{1r}^{-1} & t_{11}t_{21}^{-1} & \dots & t_{1r}t_{2r}^{-1} \end{bmatrix} \end{array}.$$

Fix a spanning tree T of G_E and let e_1, e_2, e_3 denote the three edges of G_E that are not contained in T. Let C_i denote the fundamental circuit of e_i with respect to T. The space of linear relations among the columns of J' corresponding to the edges of G_E lies within the three-dimensional subspace $\mathbb{C}\{v_{C_i} : i = 1,2,3\}$. We now show that no nonzero element of $\mathbb{C}\{v_{C_i} : i = 1,2,3\}$ lies in the kernel of K. It will then follow that the column-submatrix of J' corresponding to the edges of G_E has maximum rank.

The three fundamental circuits C_1, C_2, and C_3 all lie in $\mathrm{core}_2(G_E)$ which is a bipartite graph on partite sets of size $n' \leq 3$ and $r' \leq r$. Each vertex of $\mathrm{core}_2(G_E)$ has degree at least 2, so $n', r' \geq 2$. Since G_E is connected, $\mathrm{core}_2(G_E)$ must also be connected. Hence since C_1, C_2, and C_3 all lie in $\mathrm{core}_2(G_E)$, $\mathrm{core}_2(G_E)$ must have exactly $n'+r'+2$ edges. Since each vertex has degree at least 2, $2r' \leq n'+r'+2$ and so $r' \leq n'+2$. So thus far, we only need to consider $(n', r') = (2,2), (2,3), (2,4), (3,2), (3,3), (3,4), (3,5)$. Among these, the only (n', r')-pairs such that there even exists such a bipartite graph with the correct number of edges are $(2,4), (3,3), (3,4), (3,5)$. For these values of (n', r'), we may compute all the connected bipartite graphs on partite sets of size n' and r' with minimum degree 2 and exactly $n' + r' + 2$ edges using the `genbg` command of `Nauty and Traces` [34]. There are seven such graphs and they are displayed in Fig. 3 with vertices labeled according to which row or column they correspond to.

By relabeling vertices, we may assume that $\mathrm{core}_2(G_E)$ is supported on partite vertex sets $1, \dots, n'$ and $1, \dots, r'$. Let $A_{\mathrm{core}_2(G_E)}$ denote the incidence matrix of $\mathrm{core}_2(G_E)$ and let $M_{\mathrm{core}_2(G_E)}$ denote the matrix whose columns are a basis of the kernel of $A_{\mathrm{core}_2(G_E)}$. Then $\mathbb{C}\{v_{C_i} : i = 1,2,3\}$ is the span of $M_{\mathrm{core}_2(G_E)}$. Letting K' be the submatrix of K with columns corresponding to the edge set of $\mathrm{core}_2(G_E)$, we are done if we show that $K'M_{\mathrm{core}_2(G_E)}$ has rank 3 for the seven values of $\mathrm{core}_2(G_E)$ above. This is verified in a Mathematica script available at the following url [29].

https://dibernstein.github.io/Supplementary_materials/funtf.html

This Mathematica script computes each of the symbolic matrices $K'M_{\mathrm{core}_2(G_E)}$, then calculates their rank as matrices with entries in the appropriate function field. □

The following proposition is useful to construct examples showing that the converse of Theorem 4.2 is not true for $n \geq 4$.

Proposition 4.3. *Assume $r \geq n+2 > 4$ and let $E \subseteq [n] \times [r]$. If E is spanning in $X_{n,r}$, then at most two vertices of G_E corresponding to columns can have degree n.*

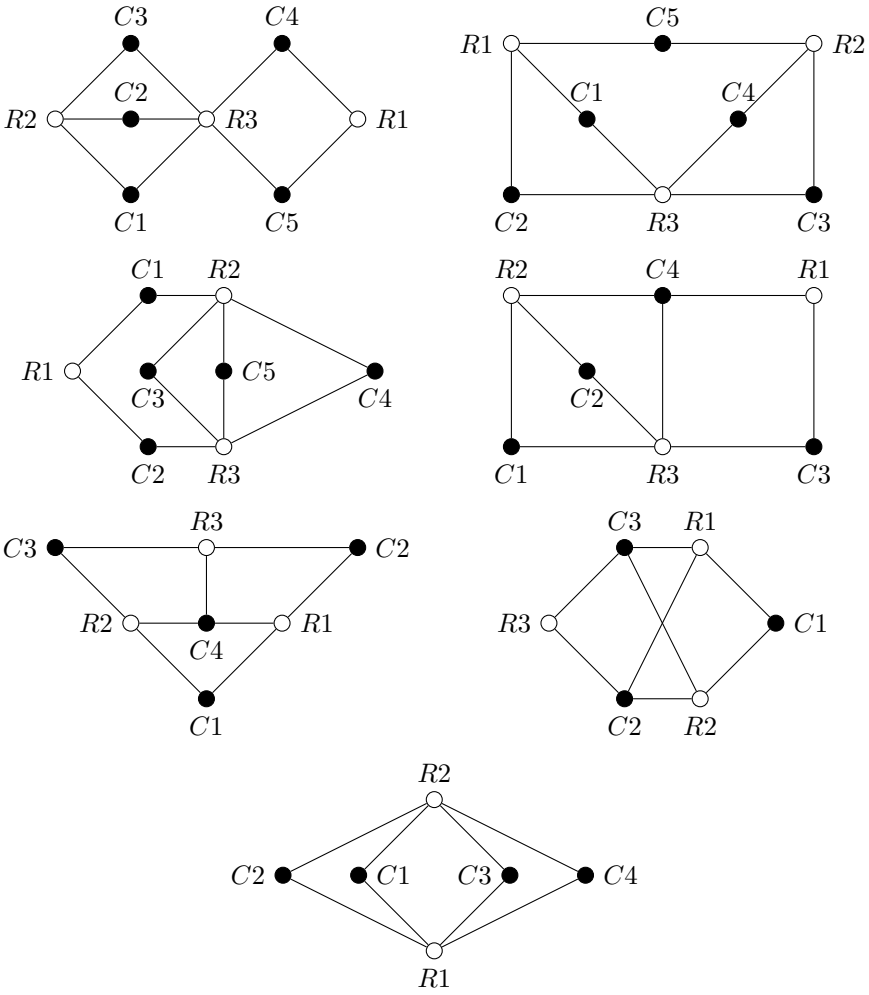


Fig. 3. The seven possibilities (up to bipartite isomorphism) for $\mathrm{core}_2(G_E)$ when E is a basis of $X_{3,r}$.

Proof. Assume G_E has k column vertices of degree n. Without loss of generality, assume they correspond to the first k columns so that $(a,i) \notin E$ for all $1 \le a \le n$, $1 \le i \le k$. Define $E' := \{(a,i) : 1 \le a \le n, k+1 \le i \le r\}$, and observe $E \subseteq E'$. We show that when $k \ge 3$, the dimension of $\pi_{E'}^{-1}(\pi_{E'}(M))$ is positive for generic M. It follows that E', and therefore E, is not spanning.

Let $M \in X_{n,r}$ be a generic finite unit norm tight frame. The (i,a) entry of M will be denoted m_{ia}. Let $\tilde{g}_a$, $\tilde{f}_{ij}$ denote the polynomials obtained from g_a and f_{ij}, as in (3.1) and (3.2), by plugging in m_{ia} for x_{ia} when $(i,a) \in E'$. The Zariski closure of $\pi_{E'}^{-1}(\pi_{E'}(M))$ can be identified with the variety in $\mathbb{C}^{n\times k}$ defined by the polynomials $\tilde{g}_1, \dots, \tilde{g}_k$ and $\tilde{f}_{ij}$ for $1 \le i \le j \le n$.

Following from Equation (3.3), we have

$$\tilde{f}_{11} + \tilde{f}_{22} + \cdots + \tilde{f}_{nn} = \tilde{g}_1 + \cdots + \tilde{g}_k \tag{4.1}$$

and so $\pi_{E'}^{-1}(\pi_{E'}(M))$ is in fact the vanishing locus of $\tilde{g}_1, \dots, \tilde{g}_{k-1}$ and $\tilde{f}_{ij}$ for $1 \le i,j \le n$. Moreover, the polynomials

$$\tilde{f}_{ij} + \delta_{ij}\frac{r}{n} - \sum_{a=k+1}^{r} m_{ia}m_{ja} \qquad \text{where} \qquad \delta_{ij} := \begin{cases} 1 & i = j \\ 0 & i \ne j \end{cases} \tag{4.2}$$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Fig. 4. Matrix entry representations of examples of the construction given in Example 4.4 for $n = 4$ and $n = 5$. These show that when $n \geq 4$, a graph on $nr - \binom{n+1}{2} - r + 1$ edges whose bipartite complement is connected may fail to be a basis of $X_{n,r}$.

parameterize the variety of $n \times n$ symmetric matrices of rank at most k, which has dimension $nk - \binom{k}{2}$ (see for example [6, Lemma 6.2]). Thus the $\binom{n+1}{2}$ polynomials $\tilde{f}_{ij}, 1 \leq i \leq j \leq n$ together contribute at most $nk - \binom{k}{2}$ to the codimension of $\pi_{E'}^{-1}(\pi_{E'}(M))$. Hence the codimension of $\pi_{E'}^{-1}(\pi_{E'}(M))$ is at most $nk - \binom{k}{2} + k - 1$, which is strictly less than nk for $k \geq 3$. □

Example 4.4. Let $n \geq 4$ and $r \geq n + 2$ and let $E' := \{(i, a) : 2 \leq i \leq n, 4 \leq a \leq r\}$. Let E be obtained from E' by removing any $\binom{n-2}{2} - 1$ elements. Then E has $nr - \binom{n+1}{2} - r + 1$ elements and G_E is connected. However, Proposition 4.3 implies that E cannot be a basis of $X_{n,r}$. Fig. 4 shows examples of this construction for $n = 4$ and $n = 5$.

4.2. Combinatorial criteria with fixed row size

The goal of this section is to fix n and find combinatorial criteria to determine if E is a basis of $X_{n,r}$ for any r. This is made precise in Remark 4.6.

The following theorem tells us that whether or not a given $E \subseteq [n] \times [r]$ of cardinality $nr - \binom{n+1}{2} - r + 1$ is a basis in $X_{n,r}$ depends only on $\text{core}_2(G_E)$.

Theorem 4.5. *Assume $r \geq n + 2 > 4$ and let $E \subseteq [n] \times [r]$ such that $|E| = nr - \binom{n+1}{2} - r + 1$. Then E is a basis in the algebraic matroid underlying $X_{n,r}$ if and only if the set $E' \supseteq E$ satisfying $G_{E'} = \overline{\text{core}}_2(G_E)$ is spanning in $X_{n,r}$. Moreover, for a fixed n, there are only finitely many possible graphs appearing as $\text{core}_2(G_E)$ as E ranges over all bases of $X_{n,r}$.*

Proof. Let $M \in X_{n,r}$ be a generic finite unit norm tight frame whose (i, a) entry is m_{ia}. Let $\tilde{g}_a$ and $\tilde{f}_{ij}$ denote the polynomials obtained from g_a and f_{ij} by setting $x_{ia} = m_{ia}$ when $(i, a) \in E$. Then $\pi_E^{-1}(\pi_E(M))$ can be viewed as the zero-dimensional variety in $\mathbb{C}^{[n]\times[r]\setminus E}$ defined by the vanishing of the polynomials $\tilde{g}_a$, $1 \leq a \leq r$ and $\tilde{f}_{ij}$, $1 \leq i \leq j \leq n$. Since G_E is connected, the edges of G_E that are not in $\overline{\text{core}}_2(G_E)$ can be ordered $(i_1, a_1), \ldots, (i_k, a_k)$ such that for each j, in either $\tilde{g}_{a_j}$ or $\tilde{f}_{i_j i_j}$, every variable other than $x_{i_j a_j}$ that appears is of the form $x_{i_l a_l}$ for some $l < j$. It follows that given $\pi_E(M)$, one can solve a series of quadratic equations in order to recover, up to finite ambiguity, the entries of M at positions corresponding to edges of G_E that are not in $\overline{\text{core}}_2(G_E)$. One can then solve for the remaining entries of M precisely when $\overline{\text{core}}_2(G_E)$ is spanning in $X_{n,r}$. The "moreover" clause follows by Proposition 4.7 below. □

Remark 4.6. Given a set $E \subseteq [n] \times [r]$ that is spanning in $X_{n,r}$, the set $E \cup \{(1, r + 1), \ldots, (n, r + 1)\}$ is spanning in $X_{n,r+1}$. Thus Theorem 4.5 tells us that if we fix n but allow r to vary, then the problem of determining whether or not $E \subset [n] \times [r]$ is a basis of $X_{n,r}$ is equivalent to determining whether or not $\text{core}_2(G_E)$ appears on a certain finite list. Proposition 4.7 below gives us the finiteness statement in Theorem 4.5, as well as bounds on the size of $\text{core}_2(G)$.

Proposition 4.7. *Let $r \geq n + 2 > 4$ and let $E \subseteq [n] \times [r]$ such that G_E is connected. Let α and β be the number of row- and column-vertices (respectively) in $\text{core}_2(G_E)$. If E is a basis of $X_{n,r}$, then*

(1) $\alpha = n-1$ *or* $\alpha = n$,
(2) $\alpha \leq \beta \leq \binom{n}{2} + \alpha - 1$.

Proof. Let $M \in X_{n,r}$ be a generic finite unit norm tight frame whose (i,a) entry is m_{ia}. Let $\tilde{g}_a$ and $\tilde{f}_{ij}$ denote the polynomials obtained from g_a and f_{ij} by setting $x_{ia} = m_{ia}$ when $(i,a) \in E$. Then $\pi_E^{-1}(\pi_E(M))$ can be viewed as the zero-dimensional variety in $\mathbb{C}^{[n]\times[r]\setminus E}$ defined by the vanishing of the polynomials $\tilde{g}_a$, $1 \leq a \leq r$ and $\tilde{f}_{ij}$, $1 \leq i \leq j \leq n$.

First we show $\alpha = n-1$ or $\alpha = n$. Without loss of generality, assume that $\mathrm{core}_2(G_E)$ has row-vertices $1,\ldots,\alpha$ and column vertices $1,\ldots,\beta$. Let $F := \{(i,a) : 1 \leq i \leq \alpha, 1 \leq a \leq \beta\} \setminus E$ be the edge set of $\overline{\mathrm{core}}_2(G_E)$. This set has cardinality given by $|F| = \alpha + \beta + \binom{n}{2} - 1$. The elements of F index the entries in the upper-left submatrix of M that are, in principle, allowed to vary over the fiber $\pi_E^{-1}(\pi_E(M))$. After dropping one of the redundant $\tilde{g}_a$'s via (4.1), there are exactly $\binom{\alpha+1}{2} + \beta - 1 + \alpha(n-\alpha)$ equations among the $\tilde{g}_a$'s and $\tilde{f}_{ij}$'s that involve entries in the upper-left $\alpha \times \beta$ block of M. Since $\pi_E^{-1}(\pi_E(M))$ is zero-dimensional, we must have

$$\alpha + \beta + \binom{n}{2} - 1 \leq \binom{\alpha+1}{2} + \beta - 1 + \alpha(n-\alpha)$$

which simplifies to

$$-\frac{1}{2}\alpha^2 + \left(n - \frac{1}{2}\right)\alpha - \binom{n}{2} \geq 0. \tag{4.3}$$

Let us now consider the left-hand side of the inequality (4.3) as a polynomial h in α, treating n as a constant. The only roots of h are $n-1$ and n, and $h(\alpha)$ is nonnegative if and only if $n-1 \leq \alpha \leq n$. Therefore, we must have $\alpha \in \{n-1, n\}$.

Now we show $\alpha \leq \beta$. As noted in the proof of Proposition 4.3, when (i,a) is an edge in G_E but not $\overline{\mathrm{core}}_2(G_E)$, we may solve a zero-dimensional quadratic system for x_{ia} given $\{m_{ia} : (i,a) \in E\}$. Thus we may now assume that $G_E = \overline{\mathrm{core}}_2(G_E)$ and allow E to be spanning in $X_{n,r}$ (as opposed to a basis of $X_{n,r}$).

Assume for the sake of contradiction that $\alpha > \beta$. Then, the $\binom{\alpha+1}{2}$ constraints $\tilde{f}_{ij} = 0$ where $1 \leq i \leq j \leq \alpha$ together can contribute at most $\alpha\beta - \binom{\beta}{2}$ to codimension. This is because for $1 \leq i \leq j \leq \alpha$, the polynomials from (4.2) with $k = \beta$ give the entries of an $\alpha \times \alpha$ symmetric matrix with rank at most β, and the dimension of the variety of $\alpha \times \alpha$ symmetric matrices of rank at most β is $\alpha\beta - \binom{\beta}{2}$ (see for example [6, Lemma 6.2]). Also, as before, at least one of the constraints $\tilde{g}_a = 0$, $1 \leq a \leq \beta$ is redundant. Since $\pi_E^{-1}(\pi_E(M))$ is zero-dimensional, we must have $|F| \leq \alpha\beta - \binom{\beta}{2} + \beta - 1$ and therefore

$$\alpha(\beta - 1) - \binom{\beta}{2} - \binom{n}{2} \geq 0. \tag{4.4}$$

After plugging in $n-1$ for α, (4.4) becomes $(\beta - n)^2 + n + \beta \leq 2$, which is a contradiction because $n \geq 3$. Plugging in n for $\alpha = n$ in (4.4), we get the inequality $-n - \beta \geq (\beta - n)^2$, which is a contradiction because the left hand side is strictly negative, and the right hand side is nonnegative. Hence, we have $\alpha \leq \beta$.

The final inequality $\beta \leq \binom{n}{2} + \alpha - 1$ follows from the fact that $\mathrm{core}_2(G_E)$ has $\binom{n}{2} + \alpha + \beta - 1$ edges and each of the β non-isolated column vertices has degree at least 2. □

5. Degree of projection and algebraic identifiability

Now that we have a handle on which subsets $E \subseteq [n] \times [r]$ yield projections $\pi_E : X_{n,r} \to \mathbb{C}^E$ that are generically finite-to-one, we can ask about the cardinality of a generically finite fiber. In other words, we want to solve the following problem.

Problem 5.1 *(Algebraic identifiability complexity).* Develop a combinatorial method for computing the degree of the map $\pi_E : X_{n,r} \to \mathbb{C}^E$ from G_E when E is a basis of $X_{n,r}$.

The following theorem gets us part of the way towards a solution to Problem 5.1.

Theorem 5.2. *Let $r \geq n+2 > 4$ and let $E \subseteq [n] \times [r]$. Define $F \subseteq [n] \times [r]$ such that $\overline{\text{core}}_2(G_E) = G_F$, and let k denote the number of vertices that are isolated in $\overline{\text{core}}_2(G_E)$ but not in G_E. If E is a spanning set of $X_{n,r}$, then we have*

$$\deg \pi_E = 2^k \cdot \deg \pi_F.$$

Proof. First observe, $E \subseteq F$ and so we have a projection map $h : \pi_F(X_{n,r}) \to \mathbb{C}^E$ that omits all the coordinates corresponding to elements of $F \setminus E$. Then we have $\pi_E = h \circ \pi_F$ and

$$\deg \pi_E = \deg \pi_F \cdot \deg h,$$

which can be seen as follows. The maps $h : \pi_F(X_{n,r}) \to \mathbb{C}^E$ and $\pi_E : X_{n,r} \to \mathbb{C}^E$ are each branched covers of $\mathbb{C}^E$. In other words, there exist dense Zariski open subsets U_1 and U_2 of $\mathbb{C}^E$ such that h restricted to $h^{-1}(U_1)$ and π_E restricted to $\pi_E^{-1}(U_2)$ are covering spaces. Moreover, h and π_E restricted to $U_1 \cap U_2$ are also covering spaces. Since $\pi_E = h \circ \pi_F$, we have π_F is a homomorphism of covering spaces and thus the topological degree of the restricted maps with image $U_1 \cap U_2$ satisfy $\deg \pi_E = \deg \pi_F \cdot \deg h$. As U_1 and U_2 are dense Zariski open subsets of $\mathbb{C}^E$, we have $U_1 \cap U_2$ is also a dense Zariski open subset of $\mathbb{C}^E$. Therefore, the equality above follows.

It now suffices to show $\deg h = 2^k$. Let $M \in X_{n,r}$ be generic and define $\tilde{g}_a$ and $\tilde{f}_{ij}$ as in the proofs of Theorem 4.5 and Proposition 4.7. We can order the elements of $F \setminus E$ as $(i_1, a_1), \ldots, (a_k, i_k)$ such that for each $j \in \{1, \ldots, k\}$, the only non-$x_{i_j a_j}$ variables in at least one of g_{a_j} or $f_{i_j i_j}$ will be of the form $x_{i_l a_l}$ with $l < j$. Let $\mathcal{F}$ denote the system of all such polynomials. The non-constant coordinates of $h^{-1}(\pi_E(M))$ are given by the variety defined by the vanishing of $\mathcal{F}$. By solving $\mathcal{F}$ via "back-substitution" in the order $x_{a_1 i_1}, \ldots, x_{a_k i_k}$, we see that this variety has exactly 2^k points. Thus, $|h^{-1}(f(M))| = 2^k$ and $\deg h = 2^k$. □

Let $E \subseteq [n] \times [r]$ with E a basis of $X_{n,r}$ and define $F \subseteq [n] \times [r]$ so that $G_F = \overline{\text{core}}_2(G_E)$. Let β denote the number of column vertices in $\text{core}_2(G_E)$ and let $M \in X_{n,r}$ be generic whose (i,a) entry is m_{ia}. If $r \geq n+2 > 4$ and $r \geq \beta + 1$, then the degree of the projection map $\pi_F : X_{n,r} \to \mathbb{C}^F$ only depends on $\text{core}_2(G_E)$ and not on r. This follows from the fact that if $r \geq \beta + 1$, then the set of non-constant polynomials $\tilde{f}_{ij}$ and $\tilde{g}_a$ obtainable by substituting $x_{ia} = m_{ia}$ for $(i,a) \in F$ does not depend on r. So for a graph H such that $H = \text{core}_2(G_E)$ for some basis E of $X_{n,r}$, let $\deg(H)$ denote the degree of π_F when $r \geq \beta + 1$. If $r = \beta$ ($r < \beta$ is not possible), then $\deg(\pi_F) \leq \deg(\text{core}_2(G_E))$. Thus Theorem 5.2 gives us the bound $\deg(\pi_E) \leq 2^k \deg(\text{core}_2(G_E))$.

Theorem 4.5 tells us, that for fixed n, there are only *finitely many* $\text{core}_2(G_E)$. Thus one can compute all values of $\deg(\text{core}_2(G_E))$ for a fixed n and use this to produce an algorithm that bounds the size of a finite fiber $|\pi_E^{-1}(\pi_E(M))|$ by computing the 2-core of G_E. This is done in Algorithm 1 and Example 5.3 illustrates this for the case $n = 3$.

Algorithm 1 For fixed $n \geq 3$, bounds the size of a generic fiber $\pi_E^{-1}(\pi_E(M))$ when E is a basis of $X_{n,r}$. Assumes that all possible values of $\deg(\text{core}_2(G_E))$ have been precomputed.

```
1: procedure BoundFiber(r, E)                ▷ r ≥ n + 2 and E is a basis of X_{n,r}
2:     H ← core_2(G_E)
3:     k ← number of vertices in G_E but not H
4:     d ← deg(H)                            ▷ obtain by looking up in precomputed table
5:     Return: d · 2^k
6: end procedure
```

Table 1
Each possible value of $F \subseteq [3] \times [5]$ such that $G_F = \mathrm{core}_2(G_E)$ when E is a basis of $X_{3,r}$, alongside the degree of the corresponding coordinate projection map.

F such that $G_F = \mathrm{core}_2(G_E)$	$\lvert \deg(\mathrm{core}_2(G_E)) \rvert$
$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$	32
$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$	24
$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$	96
$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$	128
$\begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$	288
$\begin{pmatrix} 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$	576
$\begin{pmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$	384

Example 5.3. When E is a basis of $X_{3,r}$, $\mathrm{core}_2(G_E)$ is one of seven graphs, displayed in Fig. 3. For each possible $\mathrm{core}_2(G_E)$, we compute the cardinality of a projection of an $X_{3,r}$ onto $\overline{\mathrm{core}}_2(G_E)$ using probability-one methods in `Bertini` [2] via `Macaulay2` [25,1]. When $\mathrm{core}_2(G_E)$ has five or more column vertices, we took $r = 5$ and $r = 6$ and observed that in both cases, the degree of projection was the same. When $\mathrm{core}_2(G_E)$ has fewer than five column vertices, we take $r = 5$. These degrees are given in Table 1. Via the above discussion, this characterizes the possible degrees of a projection of $X_{3,r}$ onto a basis. For example, if $E \subseteq [3] \times [5]$ where as a zero-one matrix,

$$E = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{pmatrix},$$

then the degree of a generic fiber of this projection is $128 = 2^2 \cdot 32$. This can be read off from Table 1 by noting that F such that $G_F = \mathrm{core}_2(G_E)$ is given in the top row of the table and the degree of the corresponding fiber is 32.

Acknowledgements

We are very thankful for the reviewers' comments to improve this article. This work was partially supported by the grants NSF CCF-1708884 and NSA H98230-18-1-0016. Daniel Irving Bernstein was supported by an NSF Mathematical Sciences Postdoctoral Research Fellowship (DMS-1802902). He also completed some of this work while employed by the NSF-supported (DMS-1439786) Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the Fall 2018 semester program on nonlinear algebra. Cameron Farnsworth's research was supported in part by the National Research Foundation of Korea (Grant Number 20151009350). Jose Israel Rodriguez was partially supported by the College of Letters and Science, UW-Madison.

References

[1] D.J. Bates, E. Gross, A. Leykin, J.I. Rodriguez, Bertini for Macaulay2, preprint, arXiv:1310.3297, 2013.
[2] D.J. Bates, J.D. Hauenstein, A.J. Sommese, C.W. Wampler, Bertini: software for numerical algebraic geometry, available at bertini.nd.edu, with permanent doi: https://doi.org/10.7274/R0H41PB5.
[3] D.I. Bernstein, Completion of tree metrics and rank 2 matrices, Linear Algebra Appl. 533 (2017) 1–13.

[4] D.I. Bernstein, Matroids in Algebraic Statistics, PhD thesis, North Carolina State University, 2018, https://repository.lib.ncsu.edu/handle/1840.20/35009.
[5] D.I. Bernstein, G. Blekherman, K. Lee, Typical ranks in symmetric matrix completion, arXiv preprint, arXiv:1909.06593, 2019.
[6] D.I. Bernstein, G. Blekherman, R. Sinn, Typical and generic ranks in matrix completion, Linear Algebra Appl. 585 (2020) 71–104.
[7] G. Blekherman, R. Sinn, Maximum likelihood threshold and generic completion rank of graphs, Discrete Comput. Geom. 61 (2) (2019) 303–324.
[8] C. Borcea, I. Streinu, The number of embeddings of minimally rigid graphs, Discrete Comput. Geom. 31 (2) (2004) 287–303.
[9] M. Brandt, J. Bruce, T. Brysiewicz, R. Krone, E. Robeva, The degree of $\mathrm{SO}(n, \mathbb{C})$, in: Combinatorial Algebraic Geometry, in: Fields Inst. Commun., vol. 80, Fields Inst. Res. Math. Sci., Toronto, ON, 2017, pp. 229–246.
[10] J. Cahill, M. Fickus, D.G. Mixon, M.J. Poteet, N. Strawn, Constructing finite frames of a given spectrum and set of lengths, Appl. Comput. Harmon. Anal. 35 (1) (2013) 52–73.
[11] J. Cahill, D.G. Mixon, N. Strawn, Connectivity and irreducibility of algebraic varieties of finite unit norm tight frames, SIAM J. Appl. Algebra Geom. 1 (1) (2017) 38–72.
[12] J. Cahill, N. Strawn, Algebraic geometry and finite frames, in: Finite Frames, in: Appl. Numer. Harmon. Anal., Birkhäuser/Springer, New York, 2013, pp. 141–170.
[13] J. Capco, M. Gallet, G. Grasegger, C. Koutschan, N. Lubbes, J. Schicho, The number of realizations of a Laman graph, SIAM J. Appl. Algebra Geom. 2 (1) (2018) 94–125.
[14] P.G. Casazza, M. Fickus, D.G. Mixon, Y. Wang, Z. Zhou, Constructing tight fusion frames, Appl. Comput. Harmon. Anal. 30 (2) (2011) 175–187.
[15] P.G. Casazza, J. Kovačević, Equal-norm tight frames with erasures, Adv. Comput. Math. 18 (2–4) (2003) 387–430, Frames.
[16] R.J. Duffin, A.C. Schaeffer, A class of nonharmonic Fourier series, Trans. Am. Math. Soc. 72 (1952) 341–366.
[17] K. Dykema, N. Strawn, Manifold structure of spaces of spherical tight frames, Int. J. Pure Appl. Math. 28 (2) (2006) 217–256.
[18] I.Z. Emiris, E.P. Tsigaridas, A.E. Varvitsiotis, Algebraic methods for counting Euclidean embeddings of rigid graphs, in: International Symposium on Graph Drawing, Springer, 2009, pp. 195–200.
[19] D.-J. Feng, L. Wang, Y. Wang, Generation of finite tight frames by Householder transformations, Adv. Comput. Math. 24 (1–4) (2006) 297–309.
[20] M. Fickus, J.D. Marks, M.J. Poteet, A generalized Schur-Horn theorem and optimal frame completions, Appl. Comput. Harmon. Anal. 40 (3) (2016) 505–528.
[21] M. Fickus, D.G. Mixon, M.J. Poteet, N. Strawn, Constructing all self-adjoint matrices with prescribed spectrum and diagonal, Adv. Comput. Math. 39 (3–4) (2013) 585–609.
[22] V.K. Goyal, Beyond traditional transform coding, PhD thesis, University of California, Berkeley, 1998.
[23] V.K. Goyal, J. Kovačević, J.A. Kelner, Quantized frame expansions with erasures, Appl. Comput. Harmon. Anal. 10 (3) (2001) 203–233.
[24] V.K. Goyal, M. Vetterli, N.T. Thao, Quantized overcomplete expansions in $\mathbf{R}^N$: analysis, synthesis, and algorithms, IEEE Trans. Inf. Theory 44 (1) (1998) 16–31.
[25] D.R. Grayson, M.E. Stillman, Macaulay2, a software system for research in algebraic geometry, available at http://www.math.uiuc.edu/Macaulay2/.
[26] T. Haga, C. Pegel, Polytopes of eigensteps of finite equal norm tight frames, Discrete Comput. Geom. 56 (3) (2016) 727–742.
[27] R.B. Holmes, V.I. Paulsen, Optimal frames for erasures, Linear Algebra Appl. 377 (2004) 31–51.
[28] A. Horn, Doubly stochastic matrices and the diagonal of a rotation matrix, Am. J. Math. 76 (1954) 620–630.
[29] Wolfram Research, Inc., Mathematica, Version 11.3, Champaign, IL, 2018.
[30] B. Jackson, J. Owen, The number of equivalent realisations of a rigid graph, arXiv preprint, arXiv:1204.1228, 2012.
[31] F.J. Király, L. Theran, R. Tomioka, The algebraic combinatorial approach for low-rank matrix completion, J. Mach. Learn. Res. 16 (1) (2015) 1391–1436.
[32] A. Leykin, Numerical algebraic geometry, J. Softw. Algebra Geom. 3 (2011) 5–10.
[33] P.G. Massey, M.A. Ruiz, Tight frame completions with prescribed norms, Sampl. Theory Signal Image Process. 7 (1) (2008) 1–13.
[34] B.D. McKay, A. Piperno, Practical graph isomorphism, {II}, J. Symb. Comput. 60 (2014) 94–112.
[35] L. Oeding, E. Robeva, B. Sturmfels, Decomposing tensors into frames, Adv. Appl. Math. 73 (2016) 125–153.
[36] J.G. Oxley, Matroid Theory, vol. 3, Oxford University Press, USA, 2006.
[37] Z. Rosen, Computing algebraic matroids, arXiv preprint, arXiv:1403.8148, 2014.
[38] I. Schur, Über eine Klasse von Mittelbildungen mit Anwendungen auf die Determinantentheorie, Sitzungsber. Berl. Math. Ges. 22 (1923) 9–20.
[39] N. Strawn, Geometry and constructions of finite frames, Master's thesis, Texas A&M University, College Station, TX, USA, 2007.
[40] N. Strawn, Finite frame varieties: nonsingular points, tangent spaces, and explicit local parameterizations, J. Fourier Anal. Appl. 17 (5) (2011) 821–853.
[41] J. Verschelde, Algorithm 795: Phcpack: a general-purpose solver for polynomial systems by homotopy continuation, ACM Trans. Math. Softw. 25 (2) (1999) 251–276.