TYPICAL AND GENERIC RANKS IN MATRIX COMPLETION

DANIEL IRVING BERNSTEIN, GRIGORIY BLEKHERMAN, AND RAINER SINN

ABSTRACT. We consider the problem of exact low-rank matrix completion from a geometric viewpoint: given a partially filled matrix M, we keep the positions of specified and unspecified entries fixed, and study the minimal completion rank. If the entries of the matrix are complex and the known entries are chosen randomly according to a continuous distribution, then for a fixed pattern of locations of specified and unspecified entries, there is a unique minimum completion rank which occurs with probability one. We call this rank the generic completion rank. Over the real numbers there can be multiple ranks that occur with positive probability; we call them typical completion ranks. We introduce these notions formally, and provide a number of inequalities and exact results on typical and generic ranks for different families of patterns of known and unknown entries.

1. INTRODUCTION

The problem of low-rank matrix completion received a tremendous amount of attention recently [10, 11, 32, 26, 27], especially as far as efficient algorithms are concerned. Applications that have driven much of the research in this area include collaborative filtering [15], global positioning, [34, 4, 35], and the structure-from-motion problem in computer vision [36, 12].

We study the problem of *exact* low-rank matrix completion for *generic data*. Concretely, we start with a partially-filled $m \times n$ matrix M, with real or complex entries, with the goal of finding the unspecified entries (completing M) in such a way that the completed matrix has the lowest possible rank, called the *completion rank* of M. We study how the completion rank depends on the known entries, while keeping the locations of specified and unspecified entries fixed. Generic data means that we only consider partial fillings of M where a small perturbation of the entries does not change the completion rank of M. It is well known that in case of complex entries, outside of a lower dimensional set, all partially-filled $m \times n$ matrices with the same locations of known and unknown entries have the same completion rank (see [27]), which we call the *generic (completion) rank* for this given pattern. If we restrict the entries of the partially-filled matrix and its completions to real numbers, the situation becomes more complicated: there can be several full-dimensional semi-algebraic subsets in the real vector space of partially-filled matrices on which the completion ranks are different. In analogy with tensor rank, we call such ranks *typical (completion) ranks*. In this paper, we present fundamental results about generic and typical ranks, provide first techniques to study these notions, and present case studies of generic, typical, and maximal completion ranks for various families of patterns.

We encode the locations of specified and unspecified entries in a partially-filled $m \times n$ matrix by a bipartite graph G, with parts of size m and n, corresponding to rows and

columns, such that (i, j) is an edge of G if the entry (i, j) is specified. Similarly, locations of specified and unspecified entries in a partially-filled $n \times n$ *symmetric* matrix can be encoded by a semisimple graph (that is loops, but no multiple edges, allowed) where $\{i, j\}$ is an edge of G if the (i, j) and (j, i) entries are specified. We will refer to the set of matrices with the pattern of specified and unspecified entries given by G as $\mathcal{M}_G$. The number of known entries of M is $|E|$, the number of edges of G. The G-partial matrices form a vector space of dimension $|E|$.

In many applications one assumes that a given partial matrix is either exactly or approximately completable to a matrix of rank significantly below the generic completion rank, and the main question is finding this completion algorithmically. A popular approach is to relax the non-convex rank minimization problem into a convex problem of minimizing the *nuclear norm* of a completion [33]. In the present paper we only consider exact completion, and we do not treat completion to a rank below the generic rank, since this occurs on a low-dimensional subset of the vector space of G-partially filled matrices.

We now present a brief summary of the literature on generic and typical completion ranks: It is relatively easy to show that the generic completion rank of G is 1 if and only if G is a tree [19, 35]. In this case, 1 is also the unique typical completion rank of G. In [2], graphs with generic completion rank 2 were classified using techniques of tropical geometry. Generic ranks were also examined by Kalai, Nevo, and Novik [25] under the name of *bipartite rigidity*. However, not much is known beyond generic rank 2, and typical completion ranks have not been examined. A related property of bipartite graphs called rank determinacy was studied in [13, 38] (also see [30] for a survey on matrix completion problems). For the symmetric low-rank completion problem, Uhler showed that generic completion rank can be used to certify existence of the maximum likelihood estimator of a Gaussian graphical model [37, Theorem 3.3]. Bounds on symmetric generic completion rank were further analyzed in [18] and [5].

1.1. **Main Results in Detail.** In Section 2 we prove some elementary but foundational results on generic and typical completion ranks. First we show that in the case of complex entries the generic completion rank exists. This was previously observed in [27], and we provide an elementary proof in Proposition 2.2. We sketch a simple algorithm that determines the generic completion rank of a bipartite graph with probability one (see Algorithm 3.2). We then prove a simple but important result on the behavior of typical and generic ranks:

Proposition (Proposition 2.8). *Let G be a bipartite graph. The minimal typical rank of G is equal to the generic completion rank of G. Furthermore, all ranks between the minimal typical rank and the maximal typical rank of G are typical.*

See [1] and [6] for the analogous results for the rank with respect to a variety. We also prove an interesting inequality on the maximal typical and generic ranks:

Theorem (Theorem 4.6). *Let G be a bipartite graph with generic completion rank r. Then the maximal typical completion rank of G is at most $2r - 1$.*

There are two easy lower bounds on the generic completion rank of G. Recall that $K_{r,r}$ denotes the complete bipartite graph on two parts, each of size r.

Proposition (Proposition 2.5). *Let G be a bipartite graph with parts of size m and n and edge set E. Then the following are lower bounds on the generic completion rank of G:*

(1) the smallest k such that $k(m + n) - k^2 \geq |E|$
(2) the largest r such that G contains $K_{r,r}$ as a subgraph.

If the first bound is sharp, we say that the generic completion rank of G is *predicted by the dimension count*. Observe that a $K_{r,r}$ subgraph corresponds to a fully specified $r \times r$ submatrix. Therefore if the second bound above is sharp, we say that the generic completion rank of G is *predicted by the maximal specified submatrix*.

We show that while the behavior of the generic completion rank and typical ranks is quite complicated in general, the above bounds are actually sharp for several large classes of graphs. One example is the class of *bipartite chordal graphs*; see Subsection 5.2 and [16] as a general reference. We show the following:

Theorem (Theorem 5.3, parts (a) and (b)). *Let G be a bipartite chordal graph. Then the generic completion rank and the maximal typical rank of G are predicted by the maximal specified submatrix. In particular, there is only one typical completion rank in $\mathcal{M}_G$, which is equal to the generic completion rank.*

Furthermore, one may ask which partial matrices are completable to the generic completion rank. The exceptional set of partial matrices, which have completions of rank smaller than the generic completion rank or no completion of rank equal to the generic completion rank, is lower-dimensional in the complex setting, since it is contained in a Zariski-closed set. Nevertheless, finding the exceptional set exactly is often difficult. To illustrate this, in Section 3 we completely describe the behavior of generic and typical ranks for the case of 4×4 matrices with unknown diagonal. For bipartite chordal graphs, we show that if all fully specified minors of a partially specified matrix in $\mathcal{M}_G$ are non-zero, then the matrix is completable to the generic completion rank.

Theorem (Theorem 5.3 Part (c)). *Let G be a bipartite chordal graph. Every G-partial matrix M whose completely specified minors are non-vanishing can be completed to rank $\mathrm{gcr}(G)$.*

For relations to rank determinacy, see [13, 38]. We also derive a sufficient condition for a graph to have generic completion rank predicted by the dimension count (Lemma 4.13). We use this lemma to prove that a certain subclass of bipartite circulant graphs (as defined in [31]) have generic completion rank predicted by the dimension count (Propositions 5.4 and 5.7). One of our motivations for looking at this class of graphs is that none of our methods rule out the possibility that they exhibit more than one typical rank. We currently know that one of them, the graph of the 3-cube, exhibits two typical ranks, but beyond that the existence of multiple typical ranks for this class of graphs is completely open.

We prove several more "advanced" inequalities on the generic and typical completion ranks. The following inequality uses the notion of the k-core of a graph (see Definition 4.4). See also [18] for a related result in the symmetric setting.

Corollary (Corollary 4.5). *Let r be the smallest integer such that the r-core of G is empty. Then the maximal typical rank of G is at most $r - 1$.*

A *bipartite clique sum* of bipartite graphs G and H is a graph obtained by gluing G and H along a common complete bipartite subgraph. We also show that the generic and typical completion ranks behave well under the operation of bipartite clique sum (see [5] for a related result).

Theorem (Theorem 4.8). *Let $G = G_1 \cup G_2$ be a bipartite clique sum of bipartite graphs G_1, G_2 along a complete bipartite graph $K_{m,n}$. The maximal typical rank of G is the maximum $\max\{\mathrm{mtr}(G_1), \mathrm{mtr}(G_2)\}$ between the maximal typical ranks of the summands, given that this number is at least $\max\{m, n\}$.*

The above theorem allows us to find more examples of graphs with more than one typical rank; see Example 4.9. This is because it also holds with "generic completion rank" substituted for "maximal typical rank."

Finally, we briefly examine the symmetric completion problem, where both the partial matrices as well as the completions are constrained to be symmetric (see [5, 18, 37] for connections to algebraic statistics and Gaussian graphical models). In this case, patterns of known and unknown entries are encoded by semisimple graphs (i.e. loops allowed but no multiple edges) and typical ranks are defined analogously. It is possible, unlike the non-symmetric case, that a graph on n vertices has n as a typical rank. We call such graphs *full-rank typical* and prove several results about their properties. We use these results to construct a family of semisimple graphs with no upper bound on the number of typical ranks exhibited by its members.

Theorem (Theorem 6.13). *Let $\mathcal{M}$ be the collection of $2n \times 2n$ symmetric matrices with unspecified antidiagonal. Then $2n$ is a typical symmetric completion rank of $\mathcal{M}$, i.e. $\mathcal{M}$ is full-rank typical. The generic symmetric completion rank of $\mathcal{M}$ is $2n - \left\lfloor \frac{1}{2}\left(\sqrt{1 + 8n} - 1\right)\right\rfloor$.*

We immediately obtain the following Corollary:

Corollary (Corollary 6.14). *Let $\mathcal{M}$ be the collection of $2n \times 2n$ symmetric matrices with unspecified antidiagonal. Then $\mathcal{M}$ has*

$$1 + \left\lfloor \frac{1}{2}\left(\sqrt{1 + 8n} - 1\right)\right\rfloor$$

typical symmetric completion ranks.

1.2. **Open problems, and conjectures.** We end this section by a list of open problems, questions, and conjectures. Wherever specialized notation is used, we refer the reader to the section where it is introduced.

Typical Ranks: An important and mostly unexplored research direction is to find examples of graphs exhibiting multiple typical ranks.

Problem 1.1. *Find a family of bipartite graphs with an increasing number of typical ranks. Concretely, we ask: do $n \times n$ matrices with unspecified diagonal have an unbounded number of typical ranks as n grows?*

Dressler and Krone recently classified bipartite graphs with typical rank $n - 1$ [28]. It follows from their result that $n - 1$ is a typical rank for matrices with unspecified diagonal if and only if $n \leq 4$.

At present we do not have any examples of bipartite graphs with 3 typical ranks, so as a first step toward Problem 1.1 we can ask for 3 typical ranks:

Problem 1.2. *Find a bipartite graph that has three or more typical ranks. Concretely, we conjecture that the graph $G(8,6)$ exhibits three typical ranks (see Section 5).*

It is known that all planar bipartite graphs have generic completion rank 2 and it follows from Theorem 4.6 that planar bipartite graphs have maximal typical rank 3. However, we do not know which planar bipartite graphs have 3 as a typical rank.

Problem 1.3. *Characterize the planar bipartite graphs that have 3 as a typical rank. More generally, characterize bipartite graphs with generic completion rank 2 that have 3 as a typical rank.*

An answer to Problem 1.3 would be implied by the following conjecture:

Conjecture 1.4. *Let G be a graph with non-empty 3-core (see Definition 4.4). Then the maximal typical rank of G is at least 3.*

It is also reasonable to assume in addition (in Conjecture 1.4) that G is planar. This would still lead to the resolution of Problem 1.3.

Generic Ranks: We conjecture that all bipartite circulant graphs of the form $G(n,l)$ (see Section 5) have generic completion rank predicted by the dimension count:
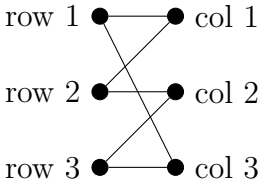
Conjecture 1.5. *All graphs $G(n,l)$ have generic completion rank predicted by the dimension count (cf. Proposition 5.4).*

2. Preliminaries

In this section, we introduce notions that are well known in the geometry of tensors in the context of matrix completion. Our setup here is the following: Let $G = (R \cup C, E)$ be a bipartite graph on parts R and C and edges $E \subset R \times C$. Throughout, we let m be the cardinality of R and n the cardinality of C. Let $\mathrm{M}^{m \times n}(K)$ be the space of $m \times n$ matrices with entries in a field K, which for us is usually the field $\mathbb{R}$ of real numbers or the field $\mathbb{C}$ of complex numbers, whose rows are indexed by elements of R and whose columns are indexed by elements of C. We let $\mathrm{M}_r^{m \times n}(K) \subseteq \mathrm{M}^{m \times n}(K)$ denote the variety consisting of $m \times n$ K-matrices with rank at most r. When the base field is clear from context, we may drop the K from our notation. We write π_G for the coordinate projection from $\mathrm{M}^{m \times n}(K)$ to K^E that takes a matrix (a_{ij}) to the vector $(a_{ij} \colon (i,j) \in E)$. Elements of K^E will be called *G-partial matrices* We think of elements in K^E as partially specified matrices. For example,

the following matrix

$$
\begin{pmatrix}
a_{11} & ? & a_{13} \\
a_{21} & a_{22} & ? \\
? & a_{32} & a_{33}
\end{pmatrix}
$$

row 1 •———• col 1

row 2 •———• col 2

row 3 •———• col 3

is a representation as a G-partial matrix of an element of K^E for the 6-cycle, which is bipartite on two parts, each of size 3.

Definition 2.1. Let $G = (R \cup C, E)$ be a bipartite graph. The $(K\text{-})completion\ rank$ of a G-partial matrix $M \in K^E$ is the smallest rank among all completions of M with entries in K, i.e. all matrices $A \in \mathrm{M}^{m \times n}(K)$ such that $\pi_G(A) = M$. For $K = \mathbb{C}$, we usually say complex completion rank. Analogously, we say real completion rank in the case $K = \mathbb{R}$.

Completion rank is not rank with respect to variety, as investigated in [1, 6, 29]. Although it may be tempting to believe that the completion rank of a G-partial matrix X is the same as its rank with respect to the projection of the variety of rank 1 matrices onto the coordinates indexed by G, this is not always the case - see Example 2.10. Nevertheless, the notions of generic rank over $\mathbb{C}$ and typical ranks over $\mathbb{R}$ apply in this context. Some elementary general results on generic and typical ranks carry over to our situation as well.

Proposition 2.2 ([27, Lemma 8]). *Let $G = (R \cup C, E)$ be a bipartite graph and suppose that K is algebraically closed (e.g. $K = \mathbb{C}$). Then there exists a unique integer r that is the completion rank of almost all G-partial matrices. Here, "almost all" means all G-partial matrices in the complement of a certain hypersurface in K^E.*

Proof. The projection π_G restricted to $\mathrm{M}_j^{m \times n}$ gives a morphism from $\mathrm{M}_j^{m \times n}$ to K^E. So the image $\pi_G(\mathrm{M}_j^{m \times n})$ is a constructible subset of K^E by Chevalley's Theorem [22, Exercise 3.19]. If it is Zariski-dense in the image, it contains a Zariski-open set [22, Exercise 3.18]. Since K^E is irreducible, the image of $\mathrm{M}_j^{m \times n}$ under π_G is either of dimension less than $\#E$ or it is Zariski-dense in K^E. So the smallest j such that $\pi_G(\mathrm{M}_j^{m \times n})$ is Zariski-dense in K^E is the integer r that we are looking for. $\qquad\square$

Remark 2.3. By generic smoothness of algebraic morphisms, the generic rank of the differential is equal to the dimension of the image of the morphism. Applied to our situation, this means that the generic completion rank of a bipartite graph $G = ([m] \times [n], E)$ is the smallest r such that the projection π_G of the tangent space to the variety of $m \times n$ matrices of rank at most r at a generic point A is surjective.

Definition 2.4. We call the integer r for $K = \mathbb{C}$ of the previous proposition 2.2 the *generic completion rank* of the bipartite graph G. We write $\mathrm{gcr}(G)$ for the generic completion rank.

Proposition 2.5. *Let $G = (R \cup C, E)$ be a bipartite graph. Then the following are both lower bounds for the generic completion rank of G:*

(1) the smallest k such that $k(m + n) - k^2 \geq \#E$
(2) the largest r such that G has $K_{r,r}$ as a subgraph.

Proof. The first lower bound follows from dimension theory in algebraic geometry. The dimension of $\mathrm{M}_k^{m\times n}$, i.e. the set of matrices of rank at most k as before, is $k(m+n) - k^2$ (see e.g. [21, Proposition 12.2]) and therefore, the dimension of the image of $\mathrm{M}_r^{m\times n}$ under π_G has dimension at most $k(m+n) - k^2$. In order for it to be dense in K^E, we need $\dim(K^E) = \#E = \dim(\pi_G(\mathrm{M}_r^{m\times n})) \leq k(m+n) - k^2$ by [21, Theorem 11.12].

The second lower bound follows by noting that a $K_{r,r}$ subgraph of G corresponds to a completely specified $r \times r$ submatrix of any G-partial matrix. $\square$

A phenomenon that is specific to the field of real numbers is the existence of typical ranks.

Definition 2.6. We call r a *typical completion rank* of a bipartite graph G if the set of points in $\mathbb{R}^E$ that have real completion rank r has non-empty interior in the Euclidean topology.

We will see examples below showing that a bipartite graph can have several typical completion ranks. The difference compared to the generic rank in the complex case is caused by the fact that Chevalley's Theorem does not hold for real algebraic sets. It must be substituted by Tarski's quantifier elimination.

Remark 2.7. We may reinterpret typical ranks from a probabilistic point of view. If we fix a "nice" probability measure on $\mathbb{R}^E$ (e.g. measures that have a continuous and positive density with respect to the Lebesgue measure), then the typical ranks of $G = (R \cup C, E)$ are exactly the real completion ranks that occur with positive probability.

Since the rank of a matrix is invariant under scaling, we can also consider probability distributions on the unit sphere in $\mathbb{R}^E$, which is compact. Again, the typical ranks are exactly the ranks that occur with positive probability for measures that have a continuous and positive density with respect to the Haar measure on the unit sphere.

The analogue of the following statement in the context of ranks in projective geometry was proved in [1, Theorem 1.1].

Proposition 2.8. *Let $G = (R \cup C, E)$ be a bipartite graph.*
(a) The smallest typical completion rank of G is the generic completion rank of G.
(b) If $r_1 < r_2$ are typical completion ranks of G, then so is every r such that $r_1 \leq r \leq r_2$.

Proof. The dimension of $\pi_G(\mathrm{M}_r^{m\times n}(\mathbb{R}))$ is equal to the dimension of its Zariski-closure in $\mathbb{C}^E$ [7, Proposition 2.8.2]. Thus part (a) follows.

To show part (b), let $r_1 \leq r < r_2$ and assume $r+1$ is not typical. Then there exists a matrix $A' \in \mathrm{M}^{m\times n}(\mathbb{R})$ of rank r with $\pi_G(A) = \pi_G(A')$ for a generic real matrix $A \in \mathrm{M}^{m\times n}(\mathbb{R})$ of rank $r+1$ because $\dim(\pi_G(\mathrm{M}_{r+1}^{m\times n}) \setminus \pi_G(\mathrm{M}_r^{m\times n})) < \dim(\pi_G(\mathrm{M}_{r+1}^{m\times n}))$. Since a generic matrix $A \in \mathrm{M}^{m\times n}(\mathbb{R})$ of rank $r+m$ can be written as $A = A_1 + A_2$ with generic matrices $A_1, A_2 \in \mathrm{M}^{m\times n}(\mathbb{R})$ satisfying $\mathrm{rk}(A_1) = r+m-1$, and $\mathrm{rk}(A_2) = 1$, we can proceed by induction on the rank to show that $r+m$ is not typical for any $m \geq 1$, which contradicts the fact that $r_2 > r$ is typical. $\square$

The *maximal rank* of a bipartite graph G is the maximum completion rank of any G-partial matrix (which usually occurs in a Zariski-thin set). Proposition 2.9 says that this is bounded above by twice the generic completion rank. At first glance, this seems like a special case of Theorem 1 in [6], but it is not quite because coordinate projections of the variety of rank-one

matrices may not be Zariski closed, see Example 2.10. However, the proof is the same simple geometric argument that applies for both the complex and real case.

Proposition 2.9. *Let* $G = (R \cup C, E)$ *be a bipartite graph.*

(a) The maximal complex completion rank of a G-partial matrix in $\mathbb{C}^E$ is at most twice the generic completion rank of G.

(b) The maximal real completion rank of a G-partial matrix in $\mathbb{R}^E$ is at most twice the minimal typical rank of G.

Proof. The argument for both cases is essentially the same. Let M be a G-partial matrix, with real or complex entries. Choose an interior point M' in the set of G-partial matrices with complex or real completion rank equal to $\mathrm{gcr}(G)$ and write $r = \mathrm{gcr}(G)$. In the complex case, open refers to the Zariski topology and existence is guaranteed by Proposition 2.2. In the real case, we use the Euclidean topology and existence is guaranteed by definition of typical rank. Consider the line L spanned by M and M'. Then this line has a spanning set of two points M_1 and M_2 with completion rank r because the intersection of the line with the set of points of completion rank r is a subset of L with non-empty interior. Fix completions A_1 and A_2 of M_1 and M_2 of rank r. Then the appropriate linear combination of A_1 and A_2 is a completion of M and has rank at most $2r$. $\square$

Example 2.10. Let T_n denote the bipartite graph corresponding to partial matrices where the known entries are precisely those on and below the diagonal. Theorem 5.3 implies that $\mathrm{gcr}(T_n) = \lceil \frac{n}{2} \rceil$ (see also [39, Theorem 2.2]), which is the maximal size of a specified submatrix in a T_n-partial matrix. Consider the T_n-partial matrix M_n

$$M_n = \begin{pmatrix} 1 & ? & ? & \ldots & ? & ? \\ 0 & 1 & ? & \ldots & ? & ? \\ 0 & 0 & 1 & \ldots & ? & ? \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}$$

with known entries corresponding to T_n. Any completion of M_n will have determinant equal to 1. Therefore, the maximum completion rank of T_n is n. This example shows that the bound provided in Proposition 2.9 for the maximum completion rank to be twice the generic completion rank is sharp. Although the completion rank of M is n, it is the limit of a sequence of T_n-partial matrices with generic completion rank 1. Explicitly, consider the sequence $(A_k)_{k \in \mathbb{N}}$ of T_n-partial matrix where the (i, j)th entry of A_k is $(2^{j-i})^k$ (where $i \leq j$). This sequence of T_n-partial matrices converges to M_n and each one has a completion of rank 1, namely the complete matrix $((2^{j-i})^k)_{i,j}$ – the sequence of completed matrices is, of course, not convergent. This argument shows that the rank of M_n with respect to the Zariski closure of the projection of rank 1 matrices is 1.
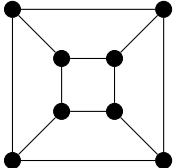
3. Example and Computational Tools

In this section we discuss an elementary example of 4×4 matrices with unspecified diagonal. We use it to showcase some of the computational tools and the difficulties in proving existence of several typical ranks. We also present a randomized algorithm for computing generic rank

(Algorithm 3.2). Let $G = ([4], [4], E)$ be the bipartite graph obtained by removing a perfect matching from $K_{4,4}$. Up to relabeling of rows and columns, the unknown entries of the corresponding partial matrices are the diagonal entries.

Example 3.1 (The 4×4 missing diagonal)**.** We focus on the 4×4 case with unspecified diagonal, i.e. partial matrices of the following form, corresponding to the graph of the cube

$$\begin{pmatrix} ? & a_{12} & a_{13} & a_{14} \\ a_{21} & ? & a_{23} & a_{24} \\ a_{31} & a_{32} & ? & a_{34} \\ a_{41} & a_{42} & a_{43} & ? \end{pmatrix}$$

Since the determinant is multilinear and not identically constant in the unknown entries, we see that every such partial matrix has a completion of rank at most 3. So we describe exactly which partial matrices have complex completion rank 1, 2, and 3. Then we discuss the typical rank behavior over the reals. We start with (complex) rank 1 (see also [24] for the general rank 1 case). The Zariski closure of the projection of the variety $M_1^{4 \times 4}$ of matrices of rank 1 onto the entries specified by G is defined by the elimination ideal, which is generated by the polynomials

$$a_{13}a_{42} - a_{12}a_{43}, \quad a_{32}a_{41} - a_{31}a_{42}, \quad a_{23}a_{41} - a_{21}a_{43},$$

$$a_{14}a_{32} - a_{12}a_{34}, \quad a_{24}a_{31} - a_{21}a_{34}, \quad a_{14}a_{23} - a_{13}a_{24}, \quad a_{23}a_{34}a_{42} - a_{24}a_{32}a_{43},$$

$$a_{13}a_{34}a_{41} - a_{14}a_{31}a_{43}, \quad a_{12}a_{24}a_{41} - a_{14}a_{21}a_{42}, \quad a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32}.$$

The first six generators are the completely specified 2×2 minors and the other four cubic generators express the condition that the 2×2 minors of the completion involving a diagonal entry vanish simultaneously. However, the image of $M_1^{4 \times 4}$ under this projection is not closed, it is only a constructible set. The elimination ideal defines its Zariski closure.

To compute the Zariski closure of the set $cl(\pi_G(M_1^{4 \times 4})) \setminus \pi_G(M_1^{4 \times 4})$ - which we call the exceptional locus - we use the Extension Theorem [14, Chapter 3, Theorem 3] (see also [14], Chapter 3, Paragraph 2 and Chapter 8, Paragraph 5 for a more detailed description of the necessary computations). We compute a Gröbner basis of $I(M_1^{4 \times 4})$ with respect to an elimination order for the diagonal entries $a_{11}, a_{22}, a_{33}, a_{44}$. We then look at the leading coefficients. Since we eliminate several variables at the same time, we need to use the Extension Theorem iteratively, one variable at a time, or use an ad-hoc argument, which is easier in this case, because the leading coefficients turn out to be a_{ij}. By the Extension Theorem, a partial matrix satisfying the equations in the elimination ideal might not have a completion of rank 1 only if one of its entries is equal to 0. If this is the case, the entire row or column of the partial matrix must be zero in order to have a completion of rank 1, which follows from the usual parameterization of $M_1^{4 \times 4}$ as $M_1^{4 \times 4} = \{vw^t \colon v, w \in \mathbb{C}^4\}$.

Up to permutation of rows and columns, there is only one case: We can assume that $a_{12} = 0$. Adding this ideal to the elimination ideal leads to four irreducible components. Two of them correspond to what we expect, namely the first row or the second column being

zero. The other two, however, are linear spaces corresponding to partial matrices of the form

$$
\begin{pmatrix}
? & 0 & 0 & a_{14} \\
0 & ? & 0 & a_{24} \\
0 & 0 & ? & a_{34} \\
a_{41} & a_{42} & a_{43} & ?
\end{pmatrix}
\quad \text{or} \quad
\begin{pmatrix}
? & 0 & a_{13} & 0 \\
0 & ? & a_{23} & 0 \\
a_{31} & a_{32} & ? & a_{34} \\
0 & 0 & a_{43} & ?
\end{pmatrix}.
$$

A generic matrix from either linear space satisfies the equations in the elimination ideal but they do not have completions of rank 1. So these linear spaces are irreducible components of the exceptional locus. In total, there are four such linear spaces (one for each diagonal entry).

All these matrices in the exceptional locus have completions of rank at most 2. For this, consider a partial matrix as on the left. Let A denote the completion where we set $a_{11} = a_{22} = a_{33} = a_{44} = 0$. Then the resulting matrix equation $Ax = 0$ imposes only one linear condition on the three-dimensional vector space $(x_1, x_2, x_3, 0)$. This shows that A has a kernel of dimension at least 2.

The generic completion rank of this completion problem is 2. In other words, the Zariski closure of the projection of the variety $M_2^{4 \times 4}$ of matrices of rank at most 2 is equal to the image space, which is equivalent to saying that the elimination ideal is the unit ideal. Using the Extension Theorem iteratively this time, we compute the Zariski closure of the complement of the image. Looking at a Gröbner basis of the ideal of 3×3 minors with respect to an elimination order of the diagonal entries with a_{44} being the last, there is only one leading coefficient of the last variable a_{44}, namely the cubic c_4 below. In fact, the Gröbner basis contains four polynomials whose leading monomial involves only one diagonal entry a_{ii} and each coefficient is a cubic c_i. The four cubics are

$$
\begin{aligned}
c_1 &= a_{12}a_{24}a_{41} - a_{14}a_{21}a_{42} \\
c_2 &= a_{13}a_{34}a_{41} - a_{14}a_{31}a_{43} \\
c_3 &= a_{23}a_{34}a_{42} - a_{24}a_{32}a_{43} \\
c_4 &= a_{12}a_{23}a_{31} - a_{13}a_{21}a_{32}.
\end{aligned}
$$

By the Extension Theorem, we cannot lift with respect to a_{44} only if c_4 vanishes. Vanishing of this cubic does not yet describe an irreducible component of the exceptional locus. Now there are two possibilities: Either such an irreducible component comes from not being able to fill in a_{44} and another diagonal entry consistently or such an irreducible component comes from further restrictions on the given entries causing problems for the diagonal entry a_{44}. To analyze the second case, we compute a Gröbner basis of the ideal of 3×3 minors plus the cubic c_4 with respect to an elimination order. We find a new element of the Gröbner basis with leading term $g_4 a_{44}$ for a quartic polynomial g_4 in the off-diagonal entries. Now we can check that the prime ideal $\langle c_4, g_4 \rangle$ defines an irreducible component of the exceptional locus. The polynomial g_4 is

$$
a_{13}a_{24}a_{32}a_{41} + a_{12}a_{23}a_{34}a_{41} - a_{14}a_{23}a_{31}a_{42} - a_{13}a_{21}a_{34}a_{42} + a_{12}a_{24}a_{31}a_{43} - a_{14}a_{21}a_{32}a_{43}.
$$

By symmetry, for the other three diagonal entries, we find three more irreducible components with defining prime ideal $\langle c_i, g_i \rangle$. They have codimension 2 and degree 12.

To analyze the first case, looking directly at leading forms in the initial Gröbner basis, another cubic coefficient c_i has to vanish. We check that this leads to another irreducible component of the exceptional locus. So by symmetry, we find six irreducible components, one for each of the six pairs c_i, c_j of cubics. The variety cut out by such a pair is reducible with two irreducible components, both of codimension 2. The relevant irreducible component, which is contained in the exceptional locus, is defined by the two cubics and a quartic. For example, the following ideal defines an irreducible component of the exceptional locus:

$$\langle c_1, c_2, a_{12}a_{24}a_{31}a_{43} - a_{13}a_{21}a_{34}a_{42} \rangle.$$

These six irreducible components have codimension 2 and degree 8.

In total, we find that the exceptional locus for rank 2 has ten irreducible components, which come in two types. One type contributes six irreducible components, one for every two out of the four cubics c_1, c_2, c_3, c_4. The other contributes four irreducible components whose prime ideal is generated by a cubic and a quartic.

We now discuss the typical ranks of G. Since $\mathrm{gcr}(G) = 2$, Proposition 2.8 tells us that 2 is a typical rank of G. Theorem 4.6 shows that 4 is not a typical rank of G. However, 3 is a typical rank of G as we now show. If we consider the projection of the variety of rank 2 matrices on the coordinates

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & ? & a_{23} & a_{24} \\ a_{31} & a_{32} & ? & a_{34} \\ a_{41} & a_{42} & a_{34} & ? \end{pmatrix}$$

i.e. we project away the three diagonal entries a_{22}, a_{33}, a_{44}, then the image will be a hypersurface whose equation has degree 2 in a_{11}. This polynomial is most compactly expressed via determinants as follows

$$\det \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix} \det \begin{pmatrix} a_{11} & a_{12} & a_{14} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & 0 \end{pmatrix} - \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} \det \begin{pmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{41} & a_{43} & 0 \end{pmatrix}.$$

To see where this polynomial comes from, consider the 3×3 minor of the matrix (a_{ij}) obtained by removing the second row and third column, and the matrix obtained by removing the second column and third row. Both must vanish and thus can be rearranged to give an expression for a_{44}. The difference of these expressions must vanish. Clearing denominators in this difference gives the polynomial above. The discriminant of this polynomial with respect to a_{11} is written out below. When it is negative, the corresponding partial matrix has no real completion to rank 2.

$$(a_{13}a_{24}a_{32}a_{41} - a_{12}a_{23}a_{34}a_{41} - a_{14}a_{23}a_{31}a_{42} - a_{13}a_{21}a_{34}a_{42} + a_{12}a_{24}a_{31}a_{43} + a_{14}a_{21}a_{32}a_{43})^2$$
$$- 4(a_{23}a_{34}a_{42} - a_{24}a_{32}a_{43})(a_{12}a_{14}a_{23}a_{31}a_{41} - a_{12}a_{13}a_{24}a_{31}a_{41} - a_{13}a_{14}a_{21}a_{32}a_{41}$$
$$+ a_{12}a_{13}a_{21}a_{34}a_{41} + a_{13}a_{14}a_{21}a_{31}a_{42} - a_{12}a_{14}a_{21}a_{31}a_{43})$$

To see that this polynomial can in fact be negative, plug in the entries of the following matrix.

$$A = \begin{pmatrix} ? & -\frac{3}{2} & -1 & 1 \\ -5 & ? & 1 & -2 \\ -2 & 1 & ? & -1 \\ 1 & -1 & -1 & ? \end{pmatrix}.$$

Note that this implies that A and any real G-partial matrix in a sufficiently small neighborhood around A can be completed to rank 3 over $\mathbb{R}$, but not rank 2. For instance, if we specify all variables in the above discriminant as given in the matrix A except for a_{12} and a_{21}, we get an indefinite conic in the (a_{12}, a_{21})-plane that has the topology of the hyperbola. The point $(a_{12}, a_{21}) = (-3/2, -5)$ lies in a connected component where this quadratic polynomial is negative.

By doing the above computation for the other diagonal entries a_{22}, a_{33}, and a_{44} instead of a_{11}, we obtain similar polynomials, derived from the other pairs of 3×3-minors involving diagonal entries. The algebraic boundary separating the G-partial matrices of real completion ranks 2 and 3 is defined by the vanishing of the product of the discriminant conditions that we get this way.

There are software packages that compute an algebraic description of the image of such a projection, i.e. of the constructible set $\pi_G(\mathrm{M}_r^{m \times n}) \subset \mathcal{M}_G = \mathbb{C}^E$. One recent example implemented in `Macaulay2` [17] is `TotalImage` developed by Harris, Michałek, and Sertöz [20]. This is an exact algorithm based on similar ideas as discussed in Example 3.1.

To compute typical ranks over the reals, the general purpose algorithm is Tarski's quantifier elimination [7, Corollary 1.4.7]. This algorithm is implemented in various computer algebra systems. However, even the example of 4×4 matrices with missing diagonal is too complex for non-custom implementations of quantifier elimination that we have tried. Given the complexity of the output in the complex case described in the previous paragraph, it seems reasonable to suspect, that applying quantifier elimination algorithm in this case is currently not feasible.

We now give a probability-1 algorithm for computing generic completion ranks. The main idea is to check that π_G is surjective when restricted to a generic tangent space of $\mathrm{M}_r^{m \times n}$. Similar algorithms have already been proposed in e.g. [35, 27]. In the theory of algebraic matroids, this is known as linearization.

Algorithm 3.2. To determine the generic completion rank of a given bipartite graph G, we propose the following probabilistic algorithm that only uses linear algebra and correctly determines the generic completion rank with probability 1. The key is to compute the rank of the projection on the tangent space at a random point as explained in Remark 2.3. To make this a linear algebra computation, pick a random $(m - r)$-dimensional vector space V in $\mathbb{C}^m$ with basis $v_1, v_2, \ldots, v_{m-r}$, and a random $(n - r)$-dimensional vector space W in $(\mathbb{C}^n)^*$ with basis $w_1, w_2, \ldots, w_{n-r}$. Then the tangent space to $\mathrm{M}_r^{m \times n}$ at a matrix $M : \mathbb{C}^m \to \mathbb{C}^n$ of rank r with kernel V and image $W^\perp$ is the set of all $m \times n$ matrices A with $w_i A v_j = 0$ for all i, j (see [21, page 185]). These conditions are linear in the entries in A. If the linear map $\pi_G|_{T_M \mathrm{M}_r^{m \times n}} : T_M \mathrm{M}_r^{m \times n} \to \mathbb{C}^E$ is surjective, the generic completion rank of G is at most r, see Remark 2.3. If it is not, then genericity of M implies that the generic completion rank of G is strictly less than r.

4. Bounding the typical ranks of a bipartite graph

This section gives various bounds on typical ranks of a bipartite graph. This is done by studying how deleting single vertices and taking bipartite clique sums affects typical ranks (Corollary 4.2 and Theorem 4.8). One consequence is that if the r-core of a bipartite graph is empty, then its maximum typical rank is at most $r - 1$ (Corollary 4.5), which in turn implies that the maximum typical rank of a graph G is at most $2 \operatorname{gcr}(G) - 1$ (Theorem 4.6). We end this section by giving two useful lemmas. One of them, Lemma 4.13, gives a sufficient condition on a graph G for sharpness of the lower bound on $\operatorname{gcr}(G)$ given by a dimension count. We begin with the simple observation of how generic completion rank behaves with respect to adding a vertex to a given bipartite graph.

Lemma 4.1. *Let $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Let M' be an $m \times n$ matrix of rank r with entries in $\mathbb{K}$. Consider the partial $m \times (n + 1)$ matrix*

$$M = \begin{pmatrix} M' & v \end{pmatrix}$$

obtained by adding a new column to M' and suppose that v is a partially specified vector with k specified entries.

   (1) If $k \leq r$ and all $r \times r$ minors of M' are non-zero, then M has a completion of rank r with entries in $\mathbb{K}$.

   (2) If $k > r$, then generically any completion of M with entries in $\mathbb{K}$ has rank $r + 1$. More precisely, the set of pairs (M', v), for which the conclusion fails is contained in a proper Zariski closed set.

Proof. We assume after permutation of the rows that the first k entries of v are specified. Since M' is generic, we can assume that the top left $r \times n$ block of M' has full rank r. In other words, the first r rows of M' form a basis of the rowspace of M'.

With the assumption that $k \leq r$, we can choose the first r rows of M to be a basis of the rowspace of M by filling in the appropriate entries in v to keep the linear relations given by the rows of M'. This shows that M has a completion of rank r, which implies (1).

If $k > r$, then the assumption that M' and the specified entries of v are generic, implies that the row relations among the rows of M' will be violated for M, so the rank of any completion must be larger than the rank of M. Since we are adding only one more column, the rank cannot increase by more than 1. $\square$

Corollary 4.2. *Let G be a bipartite graph and let v be a vertex of G of degree k. Let G' be the graph obtained from G by deletion of v.*

   (1) If the generic completion rank of G' is greater than or equal to k, then the generic completion rank of G is equal to the generic completion rank of G'.

   (2) If the maximal typical completion rank over $\mathbb{R}$ of G' is greater than or equal to k, then the maximal typical completion rank of G is equal to the maximal typical completion rank of G'.

Proof. A G-partial matrix is a G'-partial matrix with an additional column (after possibly transposing the matrix). So the result follows from the previous Lemma 4.1. $\square$

Example 4.3. The genericity assumptions in Lemma 4.1 are important. Consider the partial matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & ? \end{pmatrix}.$$

The rank of the left 3×2 block M' is 2 and the last column v only has 2 specified entries, yet the matrix does not have a completion of rank 2 because the first two rows of M' are equal but the first two entries of v are different.

We now recall the notion of k-core from graph theory. For more on k-cores, see [9, 23, 8].

Definition 4.4. The *k-core* of a graph G is the maximal subgraph of G such that all vertices have degree at least k. Equivalently, the k-core of G is the graph obtained from G by iteratively deleting verices of degree less than k.

Corollary 4.5. *If the k-core of a bipartite graph G is empty, then the maximal typical completion rank of G is less than k.*

Proof. Saying that the k-core is empty is the same as saying that we can build the graph G by adding a vertex of degree less than k at a time. So the claim follows from Corollary 4.2. □

Theorem 4.6. *Let G be a bipartite graph with generic completion rank r. Then the maximum typical rank of G is at most $2r - 1$.*

Proof. Let k be the minimum degree of G and let m, n denote the sizes of the bipartite parts of G. Since there are at most $r(m + n - r)$ edges of G (Proposition 2.5(a)), we must have $k(m + n) \leq 2r(m + n - r)$, i.e.

$$k \leq 2 \left(r - \frac{r^2}{m + n} \right).$$

Therefore the minimum vertex degree of G is at most $2r - 1$. By induction on the number of vertices of G, this implies that the $2r$-core of G is empty and so Corollary 4.5 implies that the maximum typical rank of G is at most $2r - 1$. □

Lemma 4.7. *Let r be greater than or equal to the generic completion rank of the bipartite graph $G = ([m], [n], E)$. The set of G-partial matrices M that have a completion of rank r and such that for every completion of M some $r \times r$ minor vanishes is contained in a proper Zariski closed subset of $\mathbb{K}^E$.*

Proof. We first consider the case $\mathbb{K} = \mathbb{C}$. Let U be the set of $m \times n$ matrices of rank r such that all $r \times r$ minors are non-zero. Then U is a Zariski-open subset of the irreducible variety V_r of $m \times n$ matrices of rank at most r. So the claim follows by continuity of π_G.

For $\mathbb{K} = \mathbb{R}$, essentially the same argument can be applied to the real points of V_r, because every real matrix of rank r has full local dimension in V_r and $V_r(\mathbb{R})$ is the closure (in the Euclidean topology) of the set of real matrices of rank r. □

A bipartite clique sum of two bipartite graphs G and H is a graph obtained by gluing G and H together along a common complete bipartite subgraph. Let K be either $\mathbb{R}$ or $\mathbb{C}$. Define $\mathrm{mtr}_K(G)$ to be the maximum typical rank of G over K. Theorem 4.8 below says that the maximum typical rank of a bipartite clique sum is the larger of the maximum typical

ranks of the two pieces when the bipartite clique sum is taken along a sufficiently small common bipartite clique. The case $K = \mathbb{C}$ is easily implied by the "Gluing Lemma" in [25, Lemma 3.9(2)] but the case $K = \mathbb{R}$ requires a different proof, which we now provide. Since $\mathrm{mtr}_{\mathbb{C}}(G) = \mathrm{gcr}(G)$, the Theorem below applies to generic completion rank as well as real typical rank.

Theorem 4.8. *Let $G = G_1 \cup G_2$ be a bipartite clique sum of bipartite graphs G_1, G_2 along a complete bipartite graph $K_{m,n}$. The maximal typical rank of G over a field K is the maximum $\max\{\mathrm{mtr}_K(G_1), \mathrm{mtr}_K(G_2)\}$ of the maximal typical ranks of the summands, given that this number is at least $\max\{m, n\}$.*

Proof. Since the summands G_i are subgraphs of G, the maximal typical rank of G is at least the maximal typical ranks of G_i for $i = 1, 2$. To show the reverse inequality, let M be a G-partial matrix, which can be written, after permutation of rows and columns, in the following form

$$M = \begin{pmatrix} M_1 & M_1' & ? \\ M_1'' & K & M_2' \\ ? & M_2'' & M_2 \end{pmatrix},$$

where the upper left block is a G_1-partial matrix and the bottom right block is a G_2-partial matrix, and K is a completely specified submatrix corresponding to the bipartite clique $K_{m,n} = G_1 \cap G_2$. Suppose that M is generic. First we complete the blocks using Lemma 4.7 to get a matrix

$$A = \begin{pmatrix} A_1 & B_1 & ? \\ C_1 & K & B_2 \\ ? & C_2 & A_2 \end{pmatrix},$$

where the top left matrix is a matrix of rank at most $\mathrm{mtr}(G_1)$ and the bottom right block a matrix of rank at most $\mathrm{mtr}(G_2)$, both with minors non-vanishing according to Lemma 4.7. Set $r_i = \mathrm{mtr}(G_i)$ and assume $r_1 \geq r_2$. Our goal is to complete A to rank r_1 over $\mathbb{R}$. First, we consider the case that r_1 is less than or equal to the number of rows m_2 of the bottom right block. We claim that we can complete the $r_1 - m$ rows below C_1 (m is the number of rows of K) to give us the following partial matrix

$$A = \begin{pmatrix} A_1 & B_1 & ? \\ C_1 & K & B_2 \\ E & C_2' & A_2' \\ ? & C_2'' & A_2'' \end{pmatrix} \qquad \text{where} \qquad C_2 = \begin{pmatrix} C_2' \\ C_2'' \end{pmatrix}, \qquad A_2 = \begin{pmatrix} A_2' \\ A_2'' \end{pmatrix}$$

so that each row of $\begin{pmatrix} A_1 & B_1 \end{pmatrix}$ is a linear combination of the rows of the $r_1 \times n_1$ matrix

$$F := \begin{pmatrix} C_1 & K \\ E & C_2' \end{pmatrix}.$$

Then each such linear combination can be extended to a linear combination of the rows of

$$H := \begin{pmatrix} C_1 & K & B_2 \\ E & C_2' & A_2' \end{pmatrix}$$

in order to complete the unknown entries in the upper-right corner.

We now prove the claim. For each $i = 1, \ldots, r_1 - m$, let c_i be i^{th} row of C_2 and let b_i be the ith row of B_1. These are all well-defined since $r_1 - m \leq m_2 - m$ and $m_1 - m$. By genericity, any $(n-1) \times n$ submatrix K' of K gives rise to a linear dependence of the form $x'K' + \lambda b_i = c_i$ with $\lambda \neq 0$. By padding with zeros, we extend x' to a row vector x of size m. Let a_i be the ith row of A_1. Then we set the ith row of E equal to $xC_1 + \lambda a_i$. This construction ensures that since

$$J := \begin{pmatrix} A_1 & B_1 \\ C_1 & K \end{pmatrix}$$

is a generic $m_1 \times n_1$ matrix of rank r_1, the rows of F form a basis of the row space of J. This proves our claim.

Since $r_1 \geq r_2$, each row of $\begin{pmatrix} C_2''' & A_2'' \end{pmatrix}$ can be written as a linear combination of the rows of the $r_1 \times n_2$ matrix

$$\begin{pmatrix} K & B_2 \\ C_2' & A_2' \end{pmatrix}.$$

As before, such a linear combination can be extended to a linear combination of the rows of H in order to complete the missing entries in the lower-left corner. This gives us a completion of any generic G-partial matrix to rank r_1.

If $r_1 > m_2$, we may add $r_1 - m_2$ new vertices in G_2 corresponding to $r_1 - m_2$ fully specified rows. This increases the generic completion rank of G_2 to at most r_1 and sets $m_2 = r_1$ thus bringing us back to the case where $r_1 \leq m_2$. $\qquad\square$

Example 4.9. Theorem 4.8 can be used to construct examples of graphs exhibiting multiple typical ranks. Let $G = ([4], [4], E)$ be the graph of the cube. The corresponding pattern of missing entries in a 4×4 partial matrix has all entries known aside from the diagonals. We saw in Example 3.1 that G exhibits 2 and 3 as typical ranks over the reals. Let H be a bipartite graph with generic completion rank one (i.e. a tree) or two (see [2] for a classification). Now if we glue G and H together along a $K_{2,2}$, a path of length three, a single edge, a single vertex, or the empty graph, then Theorem 4.8 tells us that the resulting graph has 2 and 3 as typical completion ranks. See Figure 1 for two examples where $H = G$ (note that $\text{gcr}(G) = 2$).

FIGURE 1. Two copies of a cube glued along a four-cycle, and along a single edge. Theorem 4.8 implies that both have two and three as their typical ranks.

Example 4.10. The assumption that $\max\{r_1, r_2\} \geq \max\{m, n\}$ in Theorem 4.8 is important. Let G be the clique sum of $K_{2,4}$ and $K_{3,4}$ with two edges removed, along a $K_{1,3}$ such that a

G-partial matrix looks like this

$$\begin{pmatrix} * & * & * & * & ? \\ * & * & * & * & * \\ ? & * & * & * & ? \\ ? & * & * & ? & * \end{pmatrix}.$$

Both blocks have generic completion rank 2, but the clique sum has generic completion rank 3, since G contains a $K_{3,3}$.

The following lemma follows quickly from the "Cone Lemma" in [25]. However, we provide an elementary proof using arguments from linear algebra directly. We will make use of it in the next section.

Lemma 4.11. *Let* $G = ([n] \sqcup [n], E)$ *be a bipartite graph and let* r *be an integer such that the differential of the projection* π_G *restricted to the variety of* $n \times n$ *matrices of rank at most* r *is generically injective, i.e. injective outside of a proper Zariski closed set. Define* $G' = ([n+1] \sqcup [n+1], E \cup ([n+1] \times \{n+1\}) \cup (\{n+1\} \times [n]))$, *the bipartite graph obtained from* G *by adding a fully connected vertex to each part. Then the following statements hold.*
(a) *The differential of* $\pi_{G'}$ *restricted to the variety of* $(n+1) \times (n+1)$ *matrices of rank at most* $r+1$ *is generically injective.*
(b) *If* G *is maximal (in the partial order given by containment of edge sets) among all bipartite graphs on* $[n] \sqcup [n]$ *of generic completion rank* r, *then* G' *is maximal among all bipartite graphs on* $[n+1] \sqcup [n+1]$ *of generic completion rank* $r+1$.

Before proving Lemma 4.11 we pause to recall a basic fact from linear algebra.

Lemma 4.12. *Let* I, K *be linear subspaces of* $\mathbb{K}^n$ *of dimensions* d *and* $n - d$ *respectively. Then there exists a matrix with image* I *and kernel* K.

Proof. Let $e_1, \ldots, e_d$ be a basis of I and $f_1, \ldots, f_{n-d}$ be a basis of K. Let F be the $n \times (n-d)$ matrix with columns $f_1, \ldots, f_{n-d}$. Let B be the $(n-d) \times (n-d)$ matrix consisting of the bottom $n - d$ rows of F, and T be the $d \times (n-d)$ matrix consisting of the top d rows of F. Reordering rows of F if necessary, we may without loss of generality assume that B is nonsingular. Now we construct our matrix M with image I and kernel K. Set the first d columns of M equal to $e_1, \ldots, e_d$. Let $\mathbf{v}^i$ denote the row vector consisting of the d specified entries of the i-th row of M. Set the unspecified entries of the i-th row of M to be the unique solution $\mathbf{x}$ to the linear equation $\mathbf{v}^i T + \mathbf{x} B = 0$. □

Proof of Lemma 4.11. Let M be a generic $(n+1) \times (n+1)$ matrix of rank $r+1$ and set $k = n - r$. Choose linearly independent linear functionals $c_1, \ldots, c_k \in (\mathbb{K}^{n+1})^*$, whose vanishing defines the column span of M and choose a basis $v_1, \ldots, v_k \in \mathbb{K}^{n+1}$ of the kernel of M. Now let $A \in T_M V_{r+1}^{n+1}$ be an element of the tangent space to the variety V_{r+1}^{n+1} of $(n+1) \times (n+1)$ matrices of rank at most $r+1$ satisfying $\pi_{G'}(A) = 0$. We need to show that $A = 0$. The condition $A \in T_M V_{r+1}^{n+1}$ is equivalent to $c_i A v_j = 0$ for all i, j. Since the last row and last column of a G'-partial matrix are completely specified by construction of G', the last row and column of A must be zero. Let us denote by c_i' and v_j' the vectors obtained from c_i and v_j, respectively, by removing their last coordinate. Let A' be the matrix obtained by removing the last row and column of A. Since $\pi_{G'}(A) = 0$, the condition that $A \in T_M V_{r+1}^{n+1}$

is equivalent to $c_i' A' v_j' = 0$ for all i, j. If $r = 0$, the proposition is clearly true so we assume $r \geq 1$. It therefore follows from genericity of M that $\{c_1', \ldots, c_k'\}$ and $\{v_1', \ldots, v_k'\}$ are both linearly independent sets. So Lemma 4.12 implies the existence of an $n \times n$ matrix M' of rank r, whose column span is cut out by $c_1', \ldots, c_k'$ and whose kernel is spanned by $v_1', \ldots, v_k'$. The condition that $c_i' A' v_j' = 0$ for all i, j is equivalent to the condition that A' is in the tangent space to V_r^n at M', the variety of $n \times n$ matrices of rank at most r. Since M' has rank r, it is a smooth point of V_r^n. Genericity of M in V_{r+1}^{n+1} implies genericity of M' in V_r^n because every $n \times n$ matrix of rank r can be obtained from an $(n+1) \times (n+1)$ matrix of rank $r+1$ by deletion of the last row and last column. So $\pi_E(A') = 0$ implies $A' = 0$ by our injectivity assumption on the differential of π_G restricted to V_r^n. Since $A' = 0$ implies $A = 0$, we have proved part (a).

The second statement follows from the first by a dimension count. If the differential of π_G restricted to V_r^n is not injective, then add an edge to G in such a way that the rank of the differential increases. The maximality of G with respect to the generic completion rank implies that the differential is in fact a bijection. The injectivity from part (a) together with a dimension count shows in fact that the differential of $\pi_{G'}$ restricted to V_{r+1}^{n+1} is also bijective. Indeed, the dimension of V_{r+1}^{n+1} turns out to be $\dim(V_r^n) + 2n + 1$, i.e. the number of edges of G'. This proves that the generic completion rank of G' is $r+1$ and if we add an edge to G', the generic completion rank will necessarily increase, see Remark 2.3. $\square$

We now give a lemma that is useful for showing that a given bipartite graph $([m], [n], E)$ satisfying $\#E = r(m + n - 1)$ has generic completion rank r. Given a bipartite graph G and subsets A and B of the two parts of G, denote by $G_{A,B}$ the induced subgraph on $A \cup B$. We use it in the next section to show that certain bipartite circulant graphs have generic completion rank predicted by the dimension count.

Lemma 4.13. *Let m, n, r be integers with $r \leq m, n$ and let $G = ([m], [n], E)$ be a bipartite graph such that $|E| = r(m+n-r)$. Then $\mathrm{gcr}(G) = r$ if there exist set partitions $P_1, \ldots, P_{m-r}$ of $[m]$ and $Q_1, \ldots, Q_{n-r}$ of $[n]$ such that each bipartite subgraph G_{P_i, Q_j} on parts P_i and Q_j contains exactly one non-edge of G, and every non-edge of G lies in some such G_{P_i, Q_j}.*

Example 4.14. Before proving Lemma 4.13, it will be helpful to have an example illustrating the statement. Let $m = n = 4$, let $r = 2$, and let $G = ([4], [4], E)$ be the bipartite graph with $E = \{(i, j) : i \neq j\}$. Then define $P_1 := \{1, 2\}, P_2 := \{3, 4\}, Q_1 := \{1, 3\}$, and $Q_2 := \{2, 4\}$. The unique non-edge of G_{P_1, Q_1} is $(1, 1)$, the unique non-edge of G_{P_1, Q_2} is $(2, 2)$, the unique non-edge of G_{P_2, Q_1} is $(3, 3)$, and the unique non-edge of G_{P_2, Q_2} is $(4, 4)$. Note that all non-edges of G are accounted for. According to Lemma 4.13, this implies $\mathrm{gcr}(G) = 2$.

Proof of Lemma 4.13. Since G has $r(m + n - r)$ edges, the lemma is equivalent to the statement that $\mathrm{M}_r^{m \times n}$ has the same dimension as V, the Zariski closure of its image under the projection π_G. We proceed by showing that $\mathrm{d}\,\pi_G : T_M \mathrm{M}_r^{m \times n} \to T_{\pi_G(M)} V$, the differential of $\pi_G|_{\mathrm{M}_r^{m \times n}} \mathrm{M}_r^{m \times n} \to \mathbb{C}^E$, is one-to-one for a particular choice of smooth point M. For now, we leave M unspecified. Since π is linear, we abuse notation by writing π instead of $\mathrm{d}\,\pi$.

Let $A \in T_M \mathrm{M}_r^{m \times n}$ and let $B \in (\mathrm{d}\,\pi)^{-1}(\pi(A))$. Our goal is to show that for a particular choice of M, B must equal A. By [22, Example 14.16], a matrix C is in $T_M \mathrm{M}_r^{m \times n}$ if and only if C maps the kernel of M into the image of M. So let $v_1, \ldots, v_{n-r} \in \mathbb{C}^n$ span the

kernel of M and let $c_1, \ldots, c_{m-r} \in (\mathbb{C}^m)^*$ be linear functionals whose vanishing cuts out the column span of M. The entries of B must satisfy the $(n-r)(m-r)$ linear equations given by $c_i B v_j = 0$. If we plug in entries of B corresponding to the edges in G, this gives us a system of $(n-r)(m-r)$ affine-linear equations that must be satisfied by the $(n-r)(m-r)$ entries of B corresponding to the non-edges of G. We denote the $(n-r)(m-r) \times (n-r)(m-r)$ coefficient matrix of this linear system by C. The constant terms in this linear system are determined by v_i, c_j, and the entries of B corresponding to edges in G. Note that the entries in these positions are the same in B and A. Therefore, if C is nonsingular, then $B = A$. We now finish the proof by constructing a smooth point M of $\mathrm{M}_r^{m \times n}$ such that the corresponding C is nonsingular.

For $i = 1, \ldots, n - r$, define c_i to be the characteristic row vector of P_i and for $j = 1, \ldots, m - r$, define v_j to be the characteristic column vector of Q_j. Note that each set $\{c_i\}, \{v_j\}$ is linearly independent. So Lemma 4.12 implies that there exists an $m \times n$ matrix M of rank r whose span is cut out by the c_is and whose kernel is spanned by the v_js (see below for an example). Moreover M is a smooth point of $\mathrm{M}_r^{m \times n}$ [22, Example 14.16]. Our hypotheses imply that the coefficient matrix C corresponding to $\{c_i\}$ and $\{v_j\}$ is a permutation matrix. In particular, C is nonsingular. $\qquad\square$

Example 4.15. We illustrate the construction M from the proof of Lemma 4.13 for the situation given in Example 4.14. Here we would have $c_1 = \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix}$, $c_2 = \begin{pmatrix} 0 & 0 & 1 & 1 \end{pmatrix}$, $v_1 = \begin{pmatrix} 1 & 0 & 1 & 0 \end{pmatrix}^T$, and $v_2 = \begin{pmatrix} 0 & 1 & 0 & 1 \end{pmatrix}^T$. Any matrix M whose column span is orthogonal to c_1 and c_2 and whose kernel is spanned by v_1 and v_2 is an acceptable choice. One such example is

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

5. Planar bipartite, bipartite chordal, and circulant graphs

In this section we discuss three classes of bipartite graphs with respect to their typical and generic completion ranks. First we make a brief note about bipartite planar graphs. Then we discuss bipartite chordal graphs, which have a long history in the subject of matrix completion [30, 38, 39]. As we will see, they all have generic completion rank predicted by the maximal specified submatrix. Moreover, they have only one typical rank. We then turn to bipartite circulant graphs. We identify a subset of them whose generic completion ranks are predicted by a dimension count and conjecture that this is the case for all such graphs. We also discuss why this class of graphs might offer a promising direction in the search for bipartite graphs exhibiting many typical ranks.

5.1. **Planar bipartite graphs.**

Theorem 5.1 ([25, Theorem 4.1]). *Let G be a planar bipartite graph. Then $\mathrm{gcr}(G) = 2$.*

Theorems 4.6 and 5.1 imply that two and three are the only possible typical ranks of a planar bipartite graph. Hence we pose the following problem.

Problem 5.2. *Characterize the planar bipartite graphs that exhibit 3 as a typical rank.*

Corollary 4.2 implies that G has three as a typical rank if and only if its 3-core has three as a typical rank. Hence to solve Problem 5.2, it suffices to restrict attention to the case of planar bipartite graphs with minimum vertex degree three.

5.2. **Bipartite chordal graphs.** A bipartite graph G is said to be *bipartite chordal* if every induced cycle has length 4. Equivalently, G can be constructed by gluing together several complete bipartite graphs along common cliques. For more on bipartite chordal graphs, see [16]. The following theorem tells us that bipartite chordal graphs are quite simple in terms of their matrix completion properties.

Theorem 5.3. *Let G be a bipartite chordal graph.*

(a) *The generic completion rank of G is the largest r such that G contains $K_{r,r}$ as an induced subgraph.*

(b) *The only typical rank of G is its generic completion rank.*

(c) *Every G-partial matrix M whose completely specified minors are nonzero can be completed to rank $\mathrm{gcr}(G)$.*

Proof. We prove (a) by induction on the number of vertices of G. The base case $G = K_{1,1}$ is trivial. For the induction step, we use that every bipartite chordal graph has a bisimplicial edge, i.e. an edge $\{v, w\}$ such that the induced graph on $N(v) \cup N(w)$, the union of the neighborhoods of v and w, is a complete bipartite graph, see [16, Corollary 5]. Let m be the degree of v and n be the degree of w and assume that $m \leq n$. Let G' be the graph obtained from G by deleting the vertex v. Then G' is also bipartite chordal and the induced subgraph on $N(v) \cup (N(w) \setminus \{v\})$ is a complete bipartite graph $K_{m,n-1}$ because $\{v, w\}$ is bisimplicial. By induction, the generic completion rank of G' is given by the largest bipartite clique in G'. Let r be the generic completion rank of G'.

Assume without loss of generality that $m \leq n$. In case that $m < n$, we know that $r \geq m$, so the claim follows by Lemma 4.1, because the degree of v is m. If $m = n$ and $r \geq m$, the same argument applies. If $m = n$ and $r < m$ then we must have $r = n - 1$. In this case, G contains $K_{m,n} = K_{r+1,r+1}$ as an induced subgraph on $N(v) \cup N(w)$ and its generic completion rank is at most $r + 1$, because the generic completion rank of G' is r and we are adding one vertex, which corresponds to a new row or column in the partial matrix. This shows (a).

To prove (b), we argue similarly by induction on the number of vertices of G applying Lemma 4.1.

Finally, to prove (c), we also proceed by induction on the number of vertices by deleting vertices contained in bisimplicial edges. The assumption, that every completely specified minor of M is non-zero, clearly transfers to submatrices. So by applying Lemma 4.1, we get a completion of M of rank at most $\mathrm{gcr}(G)$. Since G contains a clique of size $K_{r,r}$ for $r = \mathrm{gcr}(G)$, the rank of M is at least $\mathrm{gcr}(G)$ by assumption on the specified minors to be non-zero. $\square$

5.3. **Bipartite circulant graphs.** In [31], a *bipartite circulant graph* was defined to be a bipartite graph whose biadjacency matrix is a circulant matrix (see [31] for a definition of circulant matrix). We consider the subset of such graphs whose parts A and B are disjoint copies of $[n]$, where each $i \in A$ is adjacent to all vertices in B *aside from $i, i+1, \ldots, i+n-l-1$.* We denote such a graph by $G(n, l)$. Our motivation for studying this class of graphs is that

it contains the smallest bipartite graphs we know of where the possibility of multiple typical ranks is not immediately ruled out by a dimension count and Corollary 4.5. For example, we have already seen in Example 3.1 that $G(4, 3)$ exhibits both 2 and 3 as typical ranks. The graphs $G(8, 6)$ and $G(9, 5)$ are the smallest graphs we know of that could potentially have three typical ranks, although we are still unable to determine whether or not they do. The number of edges in $G(n, l)$ is nl, and so a dimension count says that its generic completion rank is at least $\lceil n - \sqrt{n^2 - nl} \rceil$. We begin by showing that this lower bound is obtained in some cases.

Proposition 5.4. *Let n be a positive integer and choose another integer k such that k divides n and n divides k^2. Define $l := n - \frac{k^2}{n}$. Then the generic completion rank of $G(n, l)$ is $n - \sqrt{n^2 - nl} = n - k$.*

Proof. Note that $G(n, l)$ has $n^2 - k^2$ edges which is equal to the dimension of $M_{n-k}^{n \times n}$. Therefore it suffices to find two partitions of $[n]$ satisfying the hypotheses of Lemma 4.13. For $i = 1, \ldots, k$, define P_i as

$$P_i := \left\{ i, i + k, \ldots, i + \left(\frac{n}{k} - 1 \right) k \right\}.$$

For $a = 1, \ldots, \frac{n}{k}$ and $b = 1, \ldots, \frac{k^2}{n}$, define Q_{ab} as

$$Q_{ab} := \left\{ (a - 1)k + b, (a - 1)k + b + \frac{k^2}{n}, \ldots, k(a - 1) + b + \left(\frac{n}{k} - 1 \right) \frac{k^2}{n} \right\}.$$

For each pair (i, ab), the edges of the graph $G(n, l)_{P_i, Q_{ab}}$ must be obtainable from the following expression, allowing p and q to range over $\{0, \ldots, \frac{n}{k} - 1\}$

$$(i + pk, k(a - 1) + b + q\frac{k^2}{n}).$$

This is a non-edge of $G(n, l)$ if and only if

$$k(a - 1) + b + q\frac{k^2}{n} - i - pk \in \{0, 1, \ldots, \frac{k^2}{n} - 1\}.$$

Since k divides n, $k > \frac{k^2}{n}$ and therefore this can happen for at most one value of (p, q).

Since there are exactly as many pairs (i, ab) as there are non-edges of $G(n, l)$, it now suffices to show that each non-edge of $G(n, l)$ appears as a non-edge of some $G(n, l)_{P_i, Q_j}$. Note that each $j \in [n]$ can be expressed as $i + pk$ for some $p \in \{0, \ldots, \frac{n}{k} - 1\}$ and $i \in \{1, \ldots, k\}$. Also, every $j \in [n]$ can be expressed as $k(a-1) + b + q\frac{k^2}{n}$ for some $q \in \{0, \ldots, \frac{n}{k} - 1\}$, $a \in \{1, \ldots, \frac{n}{k}\}$, and $b \in \{1, \ldots, \frac{k^2}{n}\}$. So for a given non-edge (u, v) of G, we can choose i such that $u \in P_i$ and ab such that $v \in Q_{ab}$ thus making (u, v) a non-edge of $G(n, l)_{P_i, Q_j}$. $\square$

It follows from Proposition 5.4 that $\text{gcr}(G(8, 6)) = 4$. Since the 6-core of $G(8, 6)$ is nonempty, Corollary 4.5 does not rule out the possibility that $G(8, 6)$ has both 5 and 6 as typical ranks. It would be interesting to know whether or not this is the case because if so, this would provide us with the first example of a bipartite graph exhibiting three or more typical ranks. Hence we ask the following question.

Question 5.5. *Does $G(8, 6)$ exhibit 6 as a typical rank?*

We believe that Proposition 5.4 holds in more generality. In particular, we make the following conjecture.

Conjecture 5.6. *Bipartite circulant graphs of the form $G(n, l)$ have the generic completion rank predicted by the dimension count. That is, $\mathrm{gcr}(G(n, l)) = n - \lfloor \sqrt{n^2 - nl} \rfloor$.*

Computations provide some evidence for Conjecture 5.6. A first step towards proving may involve relaxing the assumption that k divides n in Proposition 5.4.

In Section 6 below, we identify a family of semisimple graphs which exhibit arbitrarily many typical ranks in the symmetric matrix situation. We would like to find something similar in the non-symmetric situation. As of now, it seems that the most promising family of bipartite graphs for this goal are those of the form $G(n, n-1)$. Such graphs are sometimes called "crown graphs" and, up to row-swapping, their corresponding partial matrices have all entries known aside from the diagonals. In Proposition 5.7 below we compute their generic completion ranks.

Proposition 5.7. *The generic completion rank of $G(n, n-1)$ is $n - \lfloor \sqrt{n} \rfloor$.*

Proof. When n is a perfect square the proposition follows from Proposition 5.4 with $k = \sqrt{n}$. So let k be a positive integer and assume $n = k^2 + a$ for some $0 < a < 2k + 1$. A dimension count shows that the generic completion rank of $G(n, n-1)$ is at least $n - \lfloor \sqrt{n} \rfloor$. Let H_n denote the bipartite graph corresponding to the $n \times n$ partial matrix that is missing k^2 diagonal entries. Note that one can obtain H_n from $G(n, n-1)$ by adding a edges. Then $\mathrm{gcr}(G(n, n-1)) \leq \mathrm{gcr}(H_n)$ and Lemma 4.11 implies that $\mathrm{gcr}(H_n) = k^2 - k + a$. Then note that $n - \lfloor \sqrt{n} \rfloor = k^2 - k + a$. $\square$

6. Symmetric matrix completion

In this section, we discuss some aspects of the matrix completion problems for symmetric partially filled matrices. Instead of using bipartite graphs as in the non-symmetric case, a pattern of known entries will be encoded by a semisimple graph G (that is, loops are allowed but no multiple edges), where we put in an edge between i and j if the entries (i, j) and (j, i) are known. After preliminaries, we give some operations for constructing graphs on n vertices that have n as a typical rank. We then construct a family of graphs exhibiting an unbounded number of typical ranks.

We now introduce notation and terminology. Let $G = (V, E)$ be a semisimple graph on vertex set V, which we usually take to be $[n]$. Let $\mathcal{S}^n(K)$ denote the set of $n \times n$ symmetric matrices with entries in a field K whose rows and columns are indexed by $[n]$. Denote the variety consisting of matrices of rank at most r by $\mathcal{S}_r^n(K)$. As in the non-symmetric case, we write π_G for the coordinate projection from $\mathcal{S}_r^n(K)$ to K^E taking a matrix (a_{ij}) to the vector $(a_{ij} : (i, j) \in E)$. In analogy to the non-symmetric case, we define the *(K-)completion rank* of a G-partial matrix X to be the minimum among all ranks of symmetric K-completions of X. Many of the preliminary results for non-symmetric low-rank matrix completion problems have analogues in the symmetric case. In particular, there exists a unique integer, which we call the *generic completion rank of G* and denote by $\mathrm{sgcr}(G)$, that is the completion rank of almost all G-partial $\mathbb{C}$-matrices. This was noted in [37] for its utility in bounding the number of observations required to ensure existence of maximum likelihood estimators for

Gaussian graphical models. We summarize some other preliminary results for symmetric matrix completion in the proposition below.

Proposition 6.1. *Let $G = (V, E)$ be a semisimple graph with $|V| = n$.*

(1) *The generic completion rank of G is at least the smallest k such that $nk - \binom{k}{2} \geq \#E$.*
(2) *The smallest (real) typical rank of G is $\mathrm{sgcr}(G)$.*
(3) *If $r_1 < r_2$ are typical ranks of G, then so is every r satisfying $r_1 \leq r \leq r_2$.*
(4) *The maximal real or complex completion rank of G is at most twice $\mathrm{sgcr}(G)$.*

The proof of Proposition 6.1 is essentially the same as the analogous results for the non-symmetric case. The only missing piece is the dimension of $\mathcal{S}_r^n$, which we give in Lemma 6.2 below, along with some other basic facts about about $\mathcal{S}_r^n$ that we will use later.

Lemma 6.2. *(a) Let $L \subseteq \mathbb{K}^n$ be a linear subspace of dimension d. Then there exists an $n \times n$ symmetric matrix of rank $n - d$ whose kernel is L.*
(b) The dimension of $\mathcal{S}_r^n$ is $nr - \binom{r}{2}$.
(c) Let M be a generic $n \times n$ symmetric matrix of rank r. Then the tangent space to $\mathcal{S}_r^n$ at M is the set of $n \times n$ symmetric matrices B such that $x^T B x = 0$ for all $x \in \ker M$.

Proof. Let $A \in \mathbb{K}^{(n-d) \times n}$ be a matrix whose rows are a basis of the space of linear functionals vanishing on L. Then $A^T A$ is an $n \times n$ symmetric matrix with kernel L thus proving (a).

Part (b) is well known, but a precise reference is difficult to locate. Hence we include a proof here. If A is a generic $n \times n$ symmetric matrix of rank r, then the upper-left $r \times r$ submatrix, which we will denote A', will be a generic symmetric matrix. The lower-left $(n - r) \times r$ submatrix will also be generic. This gives us a total of $rn - \binom{r}{2}$ independent parameters. These entries uniquely determine the rest. Specifically, the upper-right $r \times (n-r)$ submatrix is the transpose of the lower-left $(n - r) \times r$ submatrix. Each remaining entry a_{ij} is uniquely determined by the equation

$$\det \begin{pmatrix} A' & b_j^T \\ b_i & a_{ij} \end{pmatrix} = 0$$

where b_i denotes the i-th row of the lower-left $(n - r) \times r$ submatrix of A.

For part (c), we differentiate the usual parameterization of the variety of $n \times n$ symmetric complex matrices of rank r given by the factorization $M = UU^T$, which shows that the tangent space to $\mathcal{S}_r^n$ at M is the set of all matrices of the form $UA^T + AU^T$. We have $x^T(UA^T + AU^T)x = 0$ for all $x \in \ker(U^T) = \ker(M)$. By part (b), it suffices to prove that the dimension of the set of symmetric matrices B such that $x^T B x = 0$ for all $x \in \ker(M)$ is $nr - \binom{r}{2}$. By diagonalizing M and possibly permuting rows and columns, we can assume that M is diagonal with the r non-zero diagonal entries at the top left. Then the set of all symmetric B with $x^T B x = 0$ for all $x \in \ker(M)$ is the set of symmetric matrices whose lower $(n - r) \times (n - r)$ block is all zeros. This set has dimension $nr - \binom{r}{2}$. □

6.1. **Full-rank typical graphs.** We say that a semisimple graph $G = (V, E)$ is *full-rank typical* if $|V|$ is a typical rank of G. This subsection explores how to construct examples of full-rank typical graphs with elementary arguments. In particular, we show that the disjoint union of a complete semi-simple graph and an isolated loop is full-rank typical. Moreover, given a full-rank typical graph G, we show that one can construct a new full-rank typical

graph by adding edges, adding a suspension vertex with a loop, or taking the join with another full-rank typical graph.

Since the determinant of a matrix is a continuous function of its entries, one can prove that a given graph G is full-rank typical by exhibiting a single real G-partial matrix X such that every real completion of X has nonzero determinant. This will be our primary tool for showing that a given graph is full-rank typical.

Proposition 6.3. *Assume $G = (V, E)$ is full-rank typical. Then E must contain all loops.*

Proof. Assume G does not contain the loop at some vertex v, i.e. the G-partial matrix is not specified at the diagonal entry corresponding to v. Let G' be the graph obtained by adding all missing edges to G except for the loop at v. Let X be a generic G-partial matrix and let X' be a generic G'-partial matrix that agrees with X at all common entries. The determinant of X' is a polynomial of degree 1 in the single missing entry. This polynomial has a real zero corresponding to a completion of X', and therefore of X, with rank at most $|V| - 1$. $\square$

Now we give three operations on graphs that preserve the property of being full-rank typical. We begin with the simplest such operation - adding edges.

Proposition 6.4. *If $G = (V, E)$ is full-rank typical, then so is any graph $H = (V, E \cup F)$ on the same set of vertices that contains G as a subgraph.*

Proof. This follows from the fact that the completion rank of a partial matrix can only go up when we specify more entries. $\square$

Given a graph $G = (V, E)$, a *suspension vertex* is a vertex $v \in V$ that is connected to every other vertex. That is, a vertex v satisfying $\{v, w\} \in E$ for all $w \in V \setminus \{v\}$. If G has a loop at a suspension vertex v, then we call v a *looped suspension vertex*. As we see in the following proposition, the operation of adding a new looped suspension vertex preserves the property of being full-rank typical.

Proposition 6.5. *If G is full-rank typical, so is the graph obtained from G by adding a looped suspension vertex.*

Proof. In terms of matrices, this means that we grow G-partial matrices by a new row and column, which are completely specified. So the completion rank of a generic G-partial matrix increases by 1. $\square$

Our last graph operation that preserves the property of being full-rank typical is binary. Let $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ be graphs on disjoint vertex sets. Define their *join* as $G_1 + G_2 = (V_1 \cup V_2, E_1 \cup E_2 \cup E)$, where E is the edge set of the complete bipartite graph on parts V_1 and V_2. So a $G_1 + G_2$-partial matrix is a block sum of G_1- and G_2-partial matrices that are also completely specified in the top right and bottom left blocks.

Proposition 6.6. *If $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ are full-rank typical, then so is $G_1 + G_2$, the join of G_1 and G_2.*

Proof. Let X_i be a generic G_i-partial matrix that is minimally completable to full rank. Let Y be the $(G_1 + G_2)$-partial matrix whose entries corresponding to edges in $V_1 \times V_2$ are all 0, and whose entries corresponding to edges in E_i agree with X_i. Then the determinant of

any completion of Y is the product of the determinants of completions of X_1 and X_2, both of which must be nonzero. $\qquad\square$

We now give an infinite family of full-rank typical graphs that are *minimal* in the sense that they cannot be built up from smaller full-rank typical graphs via the operations of adding edges, adding suspension vertices or taking joins (see Proposition 6.9). Let K_n° denote the complete semisimple graph on n vertices. Given two graphs $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ on disjoint vertex sets V_1 and V_2, let $G_1 \cup G_2$ denote their disjoint union. That is, $G_1 \cup G_2 := (V_1 \cup V_2, E_1 \cup E_2)$. Our infinite family of (minimally) full-rank typical graphs is $\{K_n^\circ \cup K_1^\circ\}_{n=1}^\infty$.

Proposition 6.7. *The typical ranks of $K_n^\circ \cup K_1^\circ$ are n and $n+1$.*

Proof. Since $K_n^\circ \cup K_1^\circ$ has a fully specified $n \times n$ minor, its generic completion rank is at least n. The determinant of a generic $K_n^\circ \cup K_1^\circ$-partial matrix is a non-constant polynomial in the n unknowns which must have a zero over the complex numbers. Hence the generic completion rank (which is the minmal typical rank) is n.

To see that $n+1$ is also a typical completion rank, we construct an explicit example of a real partial $K_n^\circ \cup K_1^\circ$-matrix with determinant bounded away from zero for all real completions. Namely, let X_n be the $K_n^\circ \cup K_1^\circ$-partial matrix whose (i, i) entry is 1 for $1 \leq i \leq n$, whose $(n+1, n+1)$ entry is -1, and whose other known entries are 0. Denote the unknown entries of X by $x_1, \ldots, x_n$. Then the determinant of X_n is $-1 - x_1^2 \cdots - x_n^2$ (this is perhaps easiest to see via cofactor expansion along the bottom row). Note that this polynomial is strictly negative for all real $x_1, \ldots, x_n$. $\qquad\square$

Now we want to show that $K_n^\circ \cup K_1^\circ$ is minimally full-rank typical with respect to our three graph operations that preserve the property of being full-rank typical. Note that since $K_n^\circ \cup K_1^\circ$ has an isolated vertex, it cannot be constructed by adding a suspension vertex to a smaller graph, nor via a graph join. So we only need to show that no subgraph obtained by edge deletion is full-rank typical. To establish this, we need Lemma 6.8, which gives us a convenient description of the generic real $(K_n^\circ \cup K_1^\circ)$-partial matrices that are completable only to full rank.

Lemma 6.8. *Let X be a generic real $K_n^\circ \cup K_1^\circ$-partial matrix. Let A denote the fully specified $n \times n$ submatrix of X corresponding to K_n° and let λ denote the diagonal entry corresponding to K_1°. Then X can only be completed to a full rank matrix if and only if A is positive definite and $\lambda < 0$ or A is negative definite and $\lambda > 0$.*

Proof. By the rank additivity of Schur complements [40, Section 0.9], the completion rank of X is $n+1$ if and only if λ is not in the range of the quadratic form defined by A. Recall that the range of a quadratic form of a matrix A is $\mathbb{R}$ if and only if A is indefinite, $\mathbb{R}_{\geq 0}$ if and only if A is positive semidefinite, and $\mathbb{R}_{\leq 0}$ if and only if A is negative semidefinite. $\qquad\square$

Proposition 6.9. *No proper subgraph of $K_n^\circ \cup K_1^\circ$ is full-rank typical.*

Proof. Let G be a proper subgraph of $K_n^\circ \cup K_1^\circ$. If G lacks any loops, then Proposition 6.3 implies that G is not full-rank typical, so assume G has all loops. Then any G-partial matrix has some unspecified off-diagonal entry in the upper-left block. Choose one such entry, and randomly specify all other unknown entries in the upper-left block. By setting

this yet unspecified entry to be sufficiently large or sufficiently negative, we can ensure that the upper-left block is neither positive definite, nor negative definite. Lemma 6.8 then implies that a completion to rank n exists. $\square$

Given the results in this section, one can construct many examples of full-rank typical graphs by starting with an instance of $K_n^\circ \cup K_1^\circ$, adding edges and looped suspension vertices, and taking joins with other similarly constructed graphs.

Problem 6.10. *Characterize semisimple graphs that are full-rank typical.*

Problem 6.10 was recently solved by Lee and the first two authors [3], who showed that G is full-rank typical if and only if the complement of G is bipartite.

6.2. **Many typical ranks.** Let $G_n := (K_1^\circ \cup K_1^\circ) + \cdots + (K_1^\circ \cup K_1^\circ)$ denote the graph obtained by taking the join of n copies of $K_1^\circ \cup K_1^\circ$. So G_n is the complete n-partite graph on n parts with two elements each and every loop. It has $2n$ vertices and $\binom{2n}{2} - n + 2n = 2n^2$ edges (including the loops). It follows from Propositions 6.7 and 6.6 that G_n is full-rank typical - i.e. its maximum typical rank is $2n$. The main result of this subsection is to show that the generic completion rank of G_n is $2n - \left\lfloor \frac{1}{2} \left(\sqrt{1 + 8n} - 1 \right) \right\rfloor$, thus establishing that G_n exhibits many typical ranks (Theorem 6.13). Note that up to permutation of rows and columns, the partial symmetric matrix corresponding to G_n has all entries known aside from the anti-diagonal.

Proposition 6.11. *The generic completion rank of $G_{(k^2+k)/2}$ is k^2.*

Proof. A dimension count shows that $\mathrm{sgcr}(G_{(k^2+k)/2}) \geq k^2$ because the codimension of the variety of $k^2 + k$ symmetric matrices of rank at most $k^2 - 1$ is $\binom{k+2}{2}$, which is greater than the dimension $\binom{k+1}{2}$ of the kernel of the projection $\pi_{G_{(k^2+k)/2}}$. To establish the reverse inequality we will show that the projection map from the tangent space of $\mathcal{S}_{k^2}^{k^2+k}$ at a particular point onto the non-anti-diagonal entries is one-to-one. By Lemma 6.2, parts (a) and (c), it suffices to find a linear subspace $L \subseteq \mathbb{C}^{k^2+k}$ of dimension k such that if a symmetric matrix B defines a quadratic form that is identically zero on L and all non-anti-diagonal entries of B are zero, then B is zero.

We index the coordinates of $\mathbb{C}^{k^2+k}$ by two disjoint copies of $[k] \sqcup \binom{[k]}{2}$, which is possible because $2\binom{k+1}{2} = k^2 + k$. We write x_i, x_{ij}, x_{i^*}, and x_{ij^*} for these coordinates, which we order essentially lexicographically as

$$x_1, \ldots, x_k, x_{12}, \ldots, x_{(k-1)k}, x_{1^*}, \ldots, x_{k^*}, x_{12^*}, \ldots, x_{(k-1)k^*}$$

Let L be the subspace defined by $x_i = x_{i^*}$, $x_{ij} = x_{ij^*}$ and $x_{ij} = x_i + x_j$. Then L is k-dimensional subspace of $\mathbb{C}^{k^2+k}$. Let B be a symmetric matrix whose entries off the anti-diagonal are all zero. Then the quadratic form that B defines with our labels is

$$x^T B x = \sum_i b_i x_i x_{i^*} + \sum_{ij} b_{ij} x_{ij} x_{ij^*}.$$

Restricting to L, we have

$$x^T B x = \sum_i c_i x_i^2 + \sum_{ij} c_{ij} x_i x_j$$

where $c_{ij} = 2b_{ij}$ and $c_i = b_i + \sum_j b_{ij}$. Therefore $x^T B x = 0$ for all $x \in L$ implies that c_i, c_{ij} are all zero. Since the map from the b's to the c's is invertible, this implies that the bs are all zero. So B is the zero matrix. $\qquad\square$

The general case will follow easily from the lemma below. Its proof is very similar to that of its nonsymmetric analogue (Lemma 4.11).

Lemma 6.12. *Let $G = ([n], E)$ be a semisimple graph such that the differential of the projection π_G restricted to variety of symmetric matrices of rank at most r is generically injective. Define $G' = ([n + 1], E \cup [n] \times \{n + 1\})$ by adding a suspension vertex and its corresponding loop edge. Then the following statements hold.*

(1) *The differential of the projection $\pi_{G'}$ restricted to the variety of $(n + 1) \times (n + 1)$ symmetric matrices of rank at most $r + 1$ is generically injective.*

(2) *If G is maximal among all semisimple graphs on n vertices of generic completion rank r then G' is maximal among all semisimple graphs on $n + 1$ vertices of generic completion rank $r + 1$.*

Proof. Let $M \in \mathcal{S}_{r+1}^{n+1}$ be a generic $(n + 1)$ symmetric matrix of rank $r + 1$. Let $A \in T_M \mathcal{S}_{r+1}^{n+1}$ be a tangent vector to the variety of symmetric matrices of rank at most $r + 1$ at M such that $\pi_{G'}(A) = 0$. We will show that this implies $A = 0$. Let $v_1, \ldots, v_{n-r}$ be a generic basis of $\ker M$. By Lemma 6.2(c), $A \in T_M \mathcal{S}_{r+1}^{n+1}$ is equivalent to the condition that $v_i^T A v_j = 0$ for all $1 \le i, j \le n - r$. Since the last row and column of a G'-partial matrix are completely specified by construction of G', $\pi_{G'}(A) = 0$ implies that the last row and column of A are zero. Let v_i' denote the vector obtained by deleting the last entry from v_i and let A' denote the symmetric matrix obtained by deleting the last row and column from A. Then $A \in T_M \mathcal{S}_{r+1}^{n+1}$ is equivalent to $(v_i')^T A' v_j' = 0$ for all $1 \le i, j \le n - r$. So $A' \in T_N \mathcal{S}_r^n$ for some symmetric matrix N whose kernel is spanned by $v_1', \ldots, v_{n-r}'$ which exists by Lemma 6.2(a). Note that $\pi_G(A') = 0$. Since N is a generic matrix of rank r, the differential of π_G restricted to $\mathcal{S}_r^n$ is injective on $T_N \mathcal{S}_r^n$, which shows $A' = 0$ and therefore $A = 0$.

Again, the second statement follows from the first by a dimension count as in the proof of Lemma 4.11. $\qquad\square$

Theorem 6.13. *Let G_n be the graph obtained as a join of n copies of $K_1^\circ \cup K_1^\circ$. The generic completion rank of G_n is*

$$2n - \left\lfloor \frac{1}{2}\left(\sqrt{1 + 8n} - 1\right)\right\rfloor$$

or equivalently, $\mathrm{sgcr}(G_n) = 2n - k$ where k is the least integer such that $\binom{k+1}{2} \ge n$. Moreover, every integer between $\mathrm{sgcr}(G_n)$ and $2n$ is a typical rank of G_n.

Proof. The "moreover" clause follows from Propositions 6.1, 6.6, and 6.7. The case $n = (k^2 + k)/2$ is Proposition 6.11. So assume $n = (k^2 + k)/2 + a$ with $1 \le a \le k$. Note that the proposed value for $\mathrm{sgcr}(G_n)$ is a lower bound by a dimension count (Lemma 6.2(b)). We show that it is also an upper bound. Let G_n' denote the graph obtained by adding $2a$ suspension vertices (with loops) to G_n. Since $G_{(k^2+k)/2}$ is maximal of generic completion rank k^2, Lemma 6.12 implies that G_n' is maximal of generic completion rank $k^2 + 2a$. Since

$\mathrm{sgcr}(G_n) \leq \mathrm{sgcr}(G'_n)$, we now have $\mathrm{sgcr}(G_n) \leq k^2 + 2a$. Plugging in $n = (k^2 + k)/2 + a$ to the proposed expression for $\mathrm{sgcr}(G_n)$ gives

$$k^2 + k + 2a - \left\lfloor \frac{1}{2} \left(\sqrt{(2k+1)^2 + 8a} - 1 \right) \right\rfloor.$$

Note that the expression inside the $\lfloor \cdot \rfloor$ is k for all $1 \leq a \leq k$. Hence the proposed value for $\mathrm{sgcr}(G_n)$ is equal to $k^2 + 2a$. $\qquad \square$

Corollary 6.14. *The graph G_n exhibits $1 + \left\lfloor \frac{1}{2} \left(\sqrt{1 + 8n} - 1 \right) \right\rfloor$ typical ranks.*

References

[1] A. Bernardi, G. Blekherman, and G. Ottaviani. On real typical ranks. *Bollettino dell'Unione Matematica Italiana*, 11(3):293–307, Sep 2018.

[2] D. I. Bernstein. Completion of tree metrics and rank 2 matrices. *Linear Algebra Appl.*, 533:1–13, 2017.

[3] D. I. Bernstein, G. Blekherman, and K. Lee. Typical ranks in symmetric matrix completion. *ArXiv e-prints arXiv:1909.06593*, 2019.

[4] P. Biswas, T.-C. Lian, T.-C. Wang, and Y. Ye. Semidefinite programming based algorithms for sensor network localization. *ACM Transactions on Sensor Networks (TOSN)*, 2(2):188–220, 2006.

[5] G. Blekherman and R. Sinn. Maximum likelihood threshold and generic completion rank of graphs. *Discrete Comp. Geom.*, 61(2):303–324, 2019.

[6] G. Blekherman and Z. Teitler. On maximum, typical and generic ranks. *Math. Ann.*, 362(3-4):1021–1031, 2015.

[7] J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.

[8] B. Bollobás. The evolution of sparse graphs. In *Graph theory and combinatorics (Cambridge, 1983)*, pages 35–57. Academic Press, London, 1984.

[9] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.

[10] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Found. Comput. Math.*, 9(6):717–772, 2009.

[11] E. J. Candès and T. Tao. The power of convex relaxation: near-optimal matrix completion. *IEEE Trans. Inform. Theory*, 56(5):2053–2080, 2010.

[12] P. Chen and D. Suter. Recovering the missing components in a large noisy low-rank matrix: Application to sfm. *IEEE transactions on pattern analysis and machine intelligence*, 26(8):1051–1063, 2004.

[13] N. Cohen, C. R. Johnson, L. Rodman, and H. J. Woerdeman. Ranks of completions of partial matrices. In *The Gohberg anniversary collection, Vol. I (Calgary, AB, 1988)*, volume 40 of *Oper. Theory Adv. Appl.*, pages 165–185. Birkhäuser, Basel, 1989.

[14] D. A. Cox, J. Little, and D. O'Shea. *Ideals, varieties, and algorithms*. Undergraduate Texts in Mathematics. Springer, Cham, fourth edition, 2015. An introduction to computational algebraic geometry and commutative algebra.

[15] D. Goldberg, D. Nichols, B. M. Oki, and D. Terry. Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35(12):61–70, 1992.

[16] M. C. Golumbic and C. F. Goss. Perfect elimination and chordal bipartite graphs. *J. Graph Theory*, 2(2):155–163, 1978.

[17] D. R. Grayson and M. E. Stillman. Macaulay2, a software system for research in algebraic geometry. Available at `https://faculty.math.illinois.edu/Macaulay2/`.

[18] E. Gross and S. Sullivant. The maximum likelihood threshold of a graph. *Bernoulli*, 24(1):386–407, 2018.

[19] D. Hadwin, K. Harrison, and J. Ward. Rank-one completions of partial matrices and completely rank-nonincreasing linear functionals. *Proceedings of the American Mathematical Society*, 134(8):2169–2178, 2006.

[20] C. Harris, M. Michałek, and E. C. Sertöz. Computing images of polynomial maps. to appear in *Adv. Comput. Math.*, 2019.

[21] J. Harris. *Algebraic geometry*, volume 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. A first course, Corrected reprint of the 1992 original.

[22] R. Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[23] S. Janson, T. Łuczak, and A. Rucinski. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.

[24] T. Kahle, K. Kubjas, M. Kummer, and Z. Rosen. The geometry of rank-one tensor completion. *SIAM Journal on Applied Algebra and Geometry*, 1(1):200–221, 2017.

[25] G. Kalai, E. Nevo, and I. Novik. Bipartite rigidity. *Trans. Amer. Math. Soc.*, 368(8):5515–5545, 2016.

[26] R. H. Keshavan, A. Montanari, and S. Oh. Matrix completion from a few entries. *IEEE Trans. Inform. Theory*, 56(6):2980–2998, 2010.

[27] F. J. Király, L. Theran, and R. Tomioka. The algebraic combinatorial approach for low-rank matrix completion. *J. Mach. Learn. Res.*, 16:1391–1436, 2015.

[28] R. Krone and M. Dressler. personal communication.

[29] J. M. Landsberg. *Tensors: geometry and applications*, volume 128 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.

[30] M. Laurent. Matrix completion problems. In C. Floudas and P. Pardalos, editors, *Encyclopedia of Optimization*, pages 1967–1975. Springer, Boston, MA, 2009.

[31] S. A. Meyer. Zero forcing sets and bipartite circulants. *Linear Algebra and its Applications*, 436(4):888–900, 2012.

[32] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[33] B. Recht, M. Fazel, and P. A. Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.*, 52(3):471–501, 2010.

[34] A. Singer. A remark on global positioning from local distances. *Proceedings of the National Academy of Sciences*, 105(28):9507–9511, 2008.

[35] A. Singer and M. Cucuringu. Uniqueness of low-rank matrix completion by rigidity theory. *SIAM Journal on Matrix Analysis and Applications*, 31(4):1621–1641, 2010.

[36] C. Tomasi and T. Kanade. Shape and motion from image streams under orthography: a factorization method. *International Journal of Computer Vision*, 9(2):137–154, 1992.

[37] C. Uhler. Geometry of maximum likelihood estimation in gaussian graphical models. *The Annals of Statistics*, 40(1):238–261, 2012.

[38] H. J. Woerdeman. The lower order of lower triangular operators and minimal rank extensions. *Integral Equations and Operator Theory*, 10(6):859–879, 1987.

[39] H. J. Woerdeman. A matrix and its inverse: revisiting minimal rank completions. In *Recent advances in matrix and operator theory*, volume 179 of *Oper. Theory Adv. Appl.*, pages 329–338. Birkhäuser, Basel, 2008.

[40] F. Zhang, editor. *The Schur complement and its applications*, volume 4 of *Numerical Methods and Algorithms*. Springer-Verlag, New York, 2005.