Higher connectivity of fiber graphs of Gröbner bases

Samu Potka

Abstract

Fiber graphs of Gröbner bases from contingency tables are important in statistical hypothesis testing, where one studies random walks on these graphs using the Metropolis-Hastings algorithm. The connectivity of the graphs has implications on how fast the algorithm converges. In this paper, we study a class of fiber graphs with elementary combinatorial techniques and provide results that support a new, interesting conjecture: the connectivity is given by the minimum vertex degree.

1 Introduction

We will study a class of graphs coming from Gröbner bases related to the two-way $n\times n$ contingency tables with equal row and column sums. By summing the entries of the tables both row-wise and column-wise, it is easy to see that the $n\times n$ tables are the only ones that can satisfy this property. Let $G(n,r)$ be a graph whose vertices are the $n\times n$-matrices of non-negative integers with all row and column sums r. Two vertices are adjacent if one can move between the corresponding matrices by adding one to two entries and subtracting one from two others. The graph $G(n,r)$ is a *fiber graph* of a Gröbner basis. See the foundational paper [1] or the textbook [2] for an introduction to algebraic statistics and how to construct fiber graphs from other Gröbner bases. As an example, consider the graph $G(3,2)$, drawn in Figure 1.1. The vertices are the 3×3-matrices of non-negative integers with row and column sums two.

To state our main result, we need to mention some definitions. The *degree* $d(v)$ of a vertex v in G is the number of edges at v. The *minimum degree* $\delta(G)$ of a graph G is the smallest of the degrees in the graph. A graph G is *k-connected*, $k \in \mathbb{N}$, if $|G| > k$ and $G - X$ is connected for every set $X \subseteq V(G)$ with $|X| < k$. The *connectivity* $\kappa(G)$ of a graph G is the largest k such that G is k-connected.

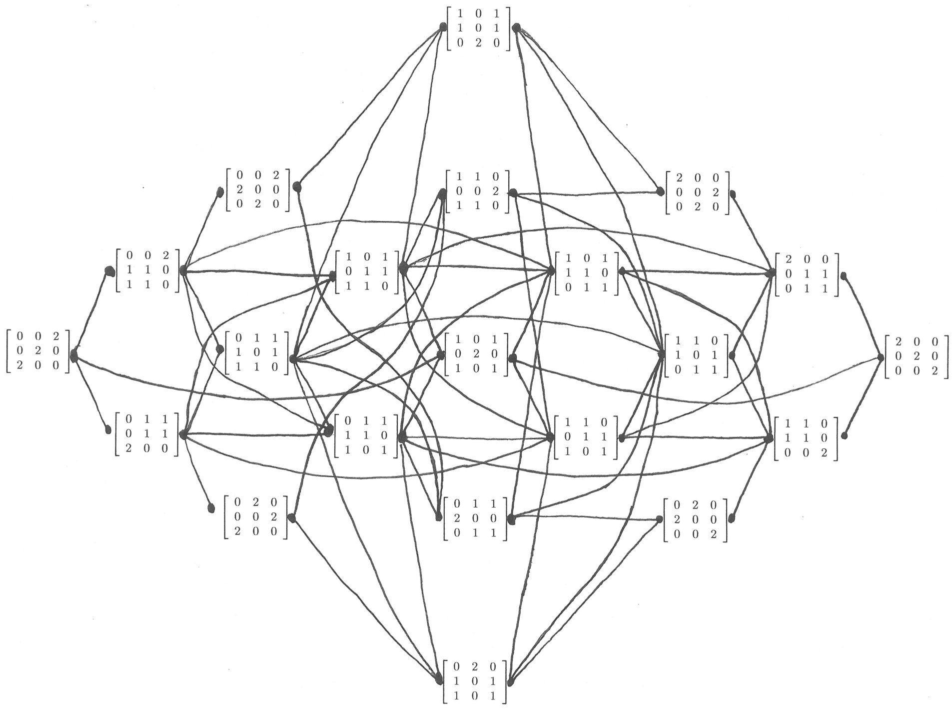

Figure 1.1: The graph $G(3, 2)$.

The *Metropolis-Hastings algorithm* can be used for statistical tests for contingency tables. The algorithm performs a random walk on the fiber graph containing the contingency table we want to study [2]. The connectivity of the fiber graphs affects the convergence of the algorithm: typically, the lower the connectivity, the slower the convergence. Our main result is:

Theorem 2.11. *The connectivity* $\kappa(G(n, r)) = \binom{n}{2}$ *for* $r > 2$.

For the first time, the following conjecture is confirmed for a large class of fiber graphs of an important and common class of Gröbner bases.

Conjecture (Engström '12, [4][5]). *The connectivity of a fiber graph of a reduced Gröbner basis of a lattice ideal is given by the minimum vertex degree of the fiber graph.*

We also prove several other statements regarding $G(n, r)$.

1.1 Basic notation

Next, we define a number of basic notions for graphs following those in [3]. Let G be a graph, $V(G)$ be the *vertex set* of G and $|G| = |V(G)|$. The *degree* $d(v)$ of a vertex v in G is the number of edges at v. The *minimum degree* $\delta(G)$ of a graph G is the smallest of the degrees in the graph and the *maximum degree* $\Delta(G)$ the largest. We call a graph G *k-connected*, $k \in \mathbb{N}$, if $|G| > k$ and $G - X$ is connected for every set $X \subseteq V(G)$ with $|X| < k$. By Menger's Theorem [3, p. 71], a graph is k-connected if and only if it contains k independent (in other words, vertex-disjoint) paths between any two vertices. We will use disjoint as a synonym of independent. The *connectivity* $\kappa(G)$ of a graph G is the largest k such that G is k-connected, the *distance* $d_G(u, v)$ between two vertices u and v of G is the number of edges in a shortest $u - v$ path in G, and the *diameter* $\operatorname{diam}(G)$ of G is defined as the largest distance in G. If G is *r-regular*, all its vertices have the same degree r. If $V(G)$ admits a partition into two classes such that the vertices in the same class are not adjacent, G is called a *bipartite* graph. A *matching* M in G is a set of independent edges and it is called *perfect* if every vertex of G is incident to exactly one edge in M. A *multigraph* is a pair (V, E) of disjoint sets together with a map $E \mapsto [V]^2$ that assigns two vertices to each edge. Here E denotes the set of edges. A multigraph differs from an ordinary graph by allowing several edges between the same two vertices. As opposed to the definition in [3], our definition does not allow self-loops, edges that start from and end to the same vertex. The entry a_{ij} of the *adjacency matrix* A of a multigraph is the number of edges from the vertex i to the vertex j. We define the *biadjacency matrix* of a bipartite multigraph as the submatrix of the adjacency matrix, where the columns correspond to the vertices in a bipartition class of the vertex set and rows to the vertices in the other class.

2 The fiber graphs

The first results are on the degree of the vertices of $G(n, r)$.

Lemma 2.1. *If $v \in V(G(n, r))$ has r as its only positive entries, then $d(v) = \binom{n}{2}$.*

Proof. We count the number of ways we can add 1 to two entries of v and subtract 1 from two others. Since there are exactly n nonzero entries in v, there are $\binom{n}{2}$ pairs of entries from which we can subtract 1. For each such choice (i_1, j_1) and (i_2, j_2), we must add 1 to the entries (i_1, j_2) and (i_2, j_1) to keep the column and row sums of v equal to r. Thus, every pair of nonzero entries in v corresponds to exactly one edge at v and $d(v) = \binom{n}{2}$. ♫

Lemma 2.2. *If $v \in V(G(n, r))$ does not have r as its only positive entries, then $d(v) \geq \frac{(n+2)(n-1)}{2} = \binom{n}{2} + n - 1$.*

Proof. To get to this type of a vertex v from a vertex that has r as its only positive entries, and therefore n positive entries, we subtract 1 from two entries, and add 1 to two other entries. Now, because we must have $r > 1$, the number of positive entries must grow by at least 2 in the process. Therefore, there are at least $n+2$ ways to select the first entry from which to subtract. The other entry cannot be from the same row or column, as we cannot add 2 to a column sum, and 1 to two row sums (or the other way) simultaneously. However, there are $n-1$ rows, and columns to select an entry from. The $n-1$ rows must together contain at least $n-1$ positive entries such that they are not in the same column as the first one selected, because the first one is at least 1, so there cannot be any r-entries in the same column or row, which leaves 1-entries to each of the other columns and rows. Any of those entries can be chosen, since one can always do the corresponding additions by the previous remarks. All this implies that there are at least $\frac{(n+2)(n-1)}{2}$ ways to do the selections, because we do not want the take the order of the selections into account. Each selection of two subtractions corresponds to exactly one edge at v, and thus $d(v) \geq \frac{(n+2)(n-1)}{2}$. ♫

By Lemma 2.1 and Lemma 2.2, we have:

Corollary 2.3. *If $v \in V(G(n,r))$ has r as its only positive entries, then $d(v) = \delta(G(n,r)) = \binom{n}{2}$.*

Using the definition of connectivity, we get the following result as an immediate implication of Corollary 2.3.

Proposition 2.4. *The connectivity of $G(n,r)$ satisfies $\kappa(G(n,r)) \leq \binom{n}{2}$.*

Proposition 2.5. *If $V(G(n,r))$ contains a vertex v that has 1 as its only positive entries, then $\Delta(G(n,r)) = d(v) = \frac{nr(nr-2r+1)}{2}$.*

Proof. The first one to subtract from can be chosen in nr ways, because there are nr 1-entries. The second one cannot be in the same row or column as the first one, and thus there are $nr - 2r + 1$ choices left. The addition can then be done in only one way, which always works, because there are no r-entries. Therefore, there are $\frac{nr(nr-2r+1)}{2}$ ways to do the selections when we do not take the order selection into account. Now, by the definition of $G(n,r)$, $d(v) = \frac{nr(nr-2r+1)}{2}$. Since other type of vertices have a lower number of positive entries to choose from, $\Delta(G(n,r)) = \frac{nr(nr-2r+1)}{2}$. Thus, $d(v) = \Delta(G(n,r)) = \frac{nr(nr-2r+1)}{2}$. ♫

Even though the maximum degree is not exactly that in some graphs ($G(n,r)$, where $n < r$), it is an upper bound. Thus, we know that, $\binom{n}{2} \leq$

$d(G) \leq \frac{nr(nr-2r+1)}{2}$. Now, having information on how the degree of the vertices of G behaves, we try to find the connectivity $\kappa(G(n,r))$. First, we will introduce a couple of auxiliary results:

Lemma 2.6. *The number of common choices for a move M from $u, v \in V(G(n,r))$ with $d_G(u,v) \leq 2$ is at least $\binom{n}{2}$ for $r > 2$.*

Proof. Because $d_G(u,v) \leq 2$ and $r > 2$, the matrices u and v have at least n positive entries in common. We want to know whether all pairs are selectable for subtraction. The positive entries only u has have to be 1 or 2. Then, because $r > 2$, there has to be entries e_i satisfying $1 \leq e_i \leq r-1$, at least one in the same column and one in the same row as such an entry. In general, each of the columns not containing an e_i has to contain a positive entry as well. Having a positive entry in a particular column means that there cannot be an r-entry in the same row. Thus, there is a positive entry not in this row in each of the other columns. We can choose two subtractions in total in $\binom{n}{2}$ ways by first selecting one of the n columns and then one of the $(n-1)$ columns. ♫

Theorem 2.7 (Kőnig, [6])**.** *Every r-regular bipartite multigraph decomposes into r perfect matchings.*

Let $E_n(i,j)$ be the $n \times n$-matrix with all entries 0, except for that position (i,j) is 1.

Lemma 2.8. *Let u be a vertex of $G(n,r)$ and $(i_1, j_1), ..., (i_k, j_k)$ positions in an $n \times n$-matrix such that $u \geq E_n(i_1, j_1) + ... + E_n(i_k, j_k)$, and $k \leq r$. Then there is a decomposition of u into a sum of matrices $u_1 + ... + u_r$ that are vertices of $G(n,1)$ such that $u_1 + ... + u_l \geq E_n(i_1, j_1) + ... + E_n(i_l, j_l)$ for all $1 \leq l \leq k$.*

Proof. The proof is by induction on k. For $k = 0$ we are done. According to Theorem 2.7, every r-regular bipartite multigraph decomposes into r perfect matchings. Interpreting u as the biadjacency matrix of an r-regular bipartite multigraph, we get a decomposition into matrices $u_1 + ... + u_r$ with row and column sum 1. Assume that we have indexed the matrices such that $(u_1)_{i_1, j_1} > 0$. Let L be a maximal subset of $\{1, 2, .., k\}$ with 1, such that $u_1 \geq \sum_{l \in L} E_n(i_l, j_l)$. By induction we can find a decomposition of $u - u_1$ admitting the conditions for $\{(i_l, j_l) \mid l \in \{1, 2, .., k\} \setminus L\}$, and then we extend it. ♫

Proposition 2.9. *The graph $G(n,r)$ is connected.*

Proof. The graph $G(n, r)$ is a fiber graph of a Gröbner basis, and therefore connected [2]. ♫

Lemma 2.10 (Liu's criterion, [7])**.** *Let G be a connected graph and $|V(G)| > k$. If for any two vertices u and v of G with distance $d_G(u, v) = 2$ there are k disjoint $u - v$ paths in G, then G is k-connected.*

A proof of Lemma 2.10 can be found in [8]. With these tools, we can set out to prove our main result:

Theorem 2.11. *The connectivity $\kappa(G(n, r)) = \binom{n}{2}$ for $r > 2$.*

Proof. By Proposition 2.4, $\kappa(G(n, r)) \leq \binom{n}{2}$. Therefore, our goal is to show that $G(n, r)$ is $\binom{n}{2}$-connected. We aim to achieve this by applying Proposition 2.9 and Lemma 2.10 as well as a technique of building a large number of paths. We need to show that using the technique, we will in every case get at least $\binom{n}{2}$ independent paths. It turns out that the technique used will not work in the cases $r < 3$. If $n = 2$, $\binom{n}{2} = 1$. By Proposition 2.9, $G(n, r)$ is connected and the case $n = 2$ is done. Thus, we assume from now on that $n \geq 3$.

We will start by setting up the machinery. By Proposition 2.9, we can apply Lemma 2.10. Let $u, v \in V(G(n, r))$ with $d_G(u, v) = 2$. Then there are proper moves Δ_1 and Δ_2 such that $u + \Delta_1 + \Delta_2 = v$. Because $d_G(u, v) = 2$, $\Delta_1 + \Delta_2$ does not correspond to a single move. Now, let us consider the sequences $M, \Delta_1, \Delta_2, -M$, where M is an additional move, such that $u + M + \Delta_1 + \Delta_2 - M = v$, as depicted in Figure 2.1. We want to show that c_M, the number of ways to select M so that we get disjoint paths is at least $\binom{n}{2} - 1$. Then we would have in total $\binom{n}{2}$ disjoint paths between u and v when we count the original path of length two as well.

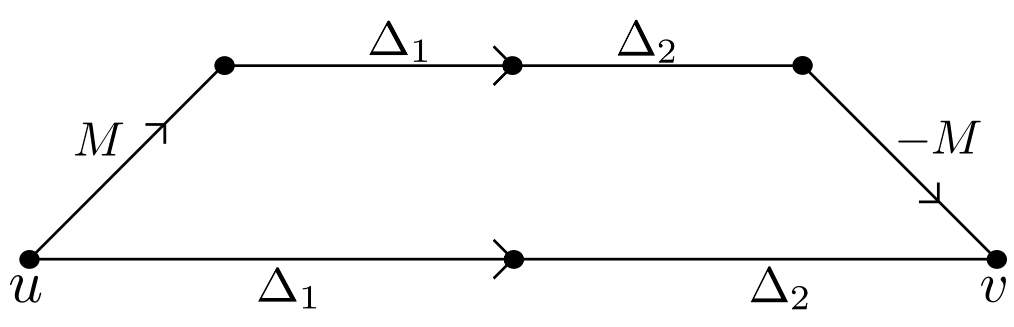

Figure 2.1: The types of paths considered in the proof with the directions corresponding to the signs of the moves.

There are some remarks to be made:

- We must have $M \neq \Delta_1$, because $M = \Delta_1$ would lead to an intersection. For the same reason, we need $M \neq -\Delta_2$.

- If we can use $M = \Delta_2$ or $M = -\Delta_1$, we have $u + \Delta_2 + \Delta_1 = v$. However, if we have both of them possible, we have to subtract one from c_M.

- On the other hand, if the selection $M = \Delta_2$ is not possible, the path using $M = -\Delta_1$ does not connect u and v.

- If $r = 1$, the entries Δ_1 subtracts from are not usable by M. By Corollary 2.3, in that case each of the vertices have the degree $\binom{n}{2}$, and thus this method does not apply, because we will not get enough ways of choosing M.

- If $r = 2, M, \Delta_1$ and Δ_2 cannot have even one same entry where they subtract from, again problematic in the cases where we start from a vertex with the degree $\binom{n}{2}$. Then we cannot get the desired result using solely this procedure. For simplicity, assume $r \geq 3$.

The basic case. Let us first assume that M can use the same entries as Δ_1 and Δ_2. Consider $\Delta_1 = \Delta_2$.

- We have at least $\binom{n}{2} - 1$ ways of choosing M such that $M \neq \Delta_1$, because $\mathrm{d}(u) \geq \binom{n}{2}$ by Corollary 2.3.

- If it is even possible to select $M = -\Delta_1$, some nonzero-entries of U are not r, and by Lemma 2.2, the degree of the vertex we are at is at least $\binom{n}{2} + n - 1$. Therefore, after subtracting the disallowed moves $M = \Delta_1$ and $M = -\Delta_1$, we have $c_M \geq \binom{n}{2} + n - 3 \geq \binom{n}{2}$ in this case, because $n \geq 3$.

However, we also have to take the case $\Delta_1 \neq \Delta_2$ into account.

- If $M = \Delta_2$ is possible, but $d(u) = \binom{n}{2}$, $M = -\Delta_2$ is not possible. Therefore, the previous results hold in this case as well.

- If also $M = -\Delta_1$ is possible as well as $M = -\Delta_2$, by the earlier analysis we get $c_M \geq \binom{n}{2} - 1$, because with our assumption $n \geq 3$, $\binom{n}{2} + n - 4 \geq \binom{n}{2} - 1$.

- If on the other hand $M = \Delta_2$ is not possible, we want to know whether the possibility $M = -\Delta_2$ is included in $d(u)$. If the number of problematic entries in Δ_2 is at least two, Δ_1 must be $-\Delta_2$, because Δ_2 will then subtract from entries zero in u Δ_1 adds to. However, there is no point in this. Therefore, consider that Δ_2 has only one problematic entry. Then Δ_2 subtracts from an entry zero in u, which implies that Δ_1 has to add to that entry, but then $-\Delta_1$ would subtract from the entry. Thus, $M = -\Delta_1$ is not included in $d(u)$. If $M = -\Delta_2$ is to be possible, by Lemma 2.2, we need to be at u with $d(u) \geq \binom{n}{2} + n - 1$, because otherwise we would subtract from an r-entry with Δ_2, but then we would add to a zero-entry. Then $c_M \geq \binom{n}{2} + n - 3 \geq \binom{n}{2}$. Otherwise we only need to avoid $M = \Delta_1$ and have $c_M \geq \binom{n}{2} - 1$, because $d(u) \geq \binom{n}{2}$ by Lemma 2.1.

Problematic entries. Let us now move on to the cases where M cannot use all the entries Δ_1 and Δ_2 use. Then, the moves Δ_1 and Δ_2 subtract from entries smaller than two in u. The number of problematic entries can range from one to four. By Lemma 2.6, u and v must have at least $\binom{n}{2}$ common choices for M.

- First, say that $\Delta_1 + \Delta_2$ subtracts from either four or three one-entries or two or one two-entry. Then the choices at v do not include Δ_1 or Δ_2. We have to avoid $-\Delta_2$, and thus $c_M \geq \binom{n}{2} - 1$.

- If Δ_1 and Δ_2 subtract from three one-entries in total, but the sum $\Delta_1 + \Delta_2$ does not, there are six different cases: either Δ_1 or Δ_2 subtracts from two one-entries, and Δ_1, Δ_2 or both add to an entry the other subtracts from. If Δ_1 subtracts from two one-entries, the moves Δ_1 and Δ_2 are clearly not possible at v. Then we have $c_M \geq \binom{n}{2} - 1$. The same thing happens when Δ_2 subtracts from two one-entries and Δ_2 does not add to an entry Δ_1 subtracts from. In the two cases left, we cannot rely on Lemma 2.6.

- If Δ_1 and Δ_2 subtract from a total number of two one-entries, we either have the other one subtracting from two or both subtracting from one. In the latter case, if neither of them or only Δ_1 adds to an entry the other subtracts from, Δ_1 and Δ_2 are not possible at v. Hence, in this case as well, we have $c_M \geq \binom{n}{2} - 1$.

The cases left are: only Δ_1 or Δ_2 subtracts from one-entries; Δ_1 subtracts from one one-entry, while Δ_2 subtracts from at least one different one-entry but adds to the one-entry Δ_1 subtracts from.

- If Δ_1 subtracts from one one-entry, $u \geq E_n(i, j)$ where (i, j) is the position of that particular one-entry. Following Lemma 2.8, decompose u: $u = u_1 + ... + u_r = u_1 + u'$, where $u_1 \geq E_n(i, j)$ and $u' \in V(G(n, r-1))$. The one-entry in the position (i, j) in u is now zero in u'. Because u

has one one-entry, it must have at least another. The second one-entry can either be in u_1 or u'. If it is in u_1, $d(u') \geq \binom{n}{2}$, and if it is in u', $d(u') \geq \binom{n}{2} + n - 1$ by Lemma 2.2. In the former case we get $c_M \geq \binom{n}{2} + (n-2) - 2 \geq \binom{n}{2} - 1$, where $\binom{n}{2}$ comes from the moves for u' and $(n-2)$ from the moves using the one entry not problematic in u now in u_1. In the latter case we have $c_M \geq= \binom{n}{2} + (n-1) - 2 \geq \binom{n}{2}$. We subtract two in both cases to avoid counting $M = -\Delta_1$ and $M = -\Delta_2$.

- If Δ_1 subtracts from two one-entries at positions (i_1, j_1) and (i_2, j_2), $u \geq E_n(i_1, j_1) + E_n(i_2, j_2)$, and we decompose $u = u_1 + u_2 + u'$, where $u' \in V(G(n, r-2))$ and $u_1 + u_2 \geq E_n(i_1, j_1) + E_n(i_2, j_2)$. Thus, the problematic entries are zero in u', and therefore also the move Δ_1 is not possible from u'. If $d(u') = \binom{n}{2}$, $-\Delta_1$ is not included in $d(u')$, and we have $c_M = \binom{n}{2} - 1$. Otherwise $d(u') > \binom{n}{2}$, and we get $c_M \geq \binom{n}{2} - 1$.

- If Δ_2 subtracts from one-entries some of which are also in u, the case is treated exactly the same way as the two previous ones. If the particular one-entries are not in u, M cannot use them and thus there is nothing to avoid.

- The case where Δ_1 subtracts from one one-entry and Δ_2 subtracts from one or two different one-entries, but Δ_2 adds to the one-entry Δ_1 subtracts from and at most one of the one-entries Δ_2 subtracts from is present in u already, is treated exactly same way as the previous ones, because we have to avoid one or two problematic one-entries. If there are two problematic entries both already in u, they can be avoided the same way as before. If there are three of them, all present in u at positions (i_1, j_1), (i_2, j_2) and (i_3, j_3), we have $u \geq E_n(i_1, j_1) + E_n(i_2, j_2) + E_n(i_3, j_3)$. Say that the two first are the ones used by Δ_2. They can be put in the same u_1 in the proof of Lemma 2.8. Then we have $u = u_1 + u_2 + u'$, where $u' \in V(G(n, r-2))$. The problematic entries are zero in u' and the moves Δ_1 and Δ_2 are not possible from u'. Then $d(u') \geq \binom{n}{2}$ only includes the disallowed choice $M = -\Delta_2$. Thus $c_M \geq \binom{n}{2} - 1$.

Intersections. The last question is what if different paths $M + \Delta_1 + \Delta_2 - M$ and $M' + \Delta_1 + \Delta_2 - M'$ intersect. By symmetry and straightforward calculations, the number of cases reduces to three: $M' - M = \Delta_1$; $M' - M = \Delta_2$; $M' - M = \Delta_1 + \Delta_2$. The different types are drawn in Figure 2.2:

The last case is the easiest to handle. Assume that we only have intersections of this type. An intersection can happen in two different ways.

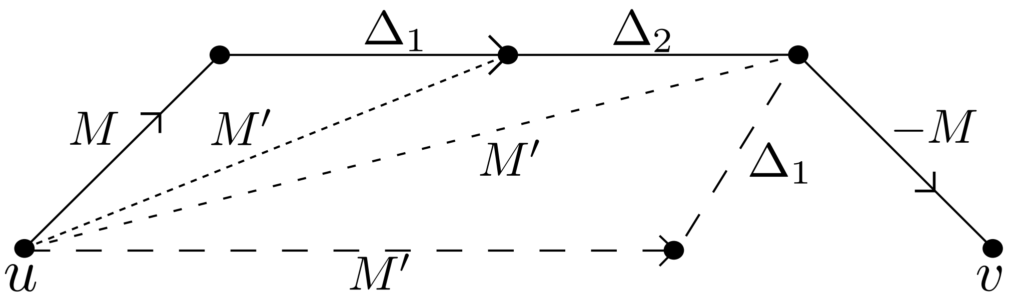

Figure 2.2: The possible types of intersection.

- The moves Δ_1 and Δ_2 share one entry the other adds to and the other subtracts from. This sum can be written in two ways, one being the original, because M and M' have to be proper moves and both have to use three of the operations in $\Delta_1 + \Delta_2$. Therefore, this case amounts to one intersection. Let us a write an example to illustrate this:

$$\Delta_1 + \Delta_2 = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix} + \begin{pmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ -1 & 1 & 0 \end{pmatrix}.$$

- The other possibility, disjoint from the previous one, is that Δ_1 and Δ_2 are on the same two rows or columns. Then there are two ways, the original and another with swapped rows, to write the sum $\Delta_1 + \Delta_2$. Also this gives one intersection. Again, let us do a basic example:

$$\Delta_1 + \Delta_2 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

To analyse how these affect the earlier calculations, we have to first note that the entries $\Delta_1 + \Delta_2$ subtracts from must be at least two, because we want to subtract from the same entries with M and M'. Only one of the two types is possible at a time.

- In the former case, if $-\Delta_2$ (or $-\Delta_1$ if the order of the moves is switched) is to be included in $d(u)$, we must have $d(u) > \binom{n}{2}+n-1$, because there has to be at least three positive entries in one column, and therefore $c_M > \binom{n}{2}+n-1-4 \geq \binom{n}{2}-2$, where -4 comes from three disallowed moves and one intersection – if not, the degree is at least $\binom{n}{2}+n-1$ by Lemma 2.2, which means we have $c_M \geq \binom{n}{2}+n-4 \geq \binom{n}{2}-1$.

- In the latter case, $d(u)$ must be at least $\binom{n}{2}+n-1$ by Lemma 2.2, and we have $c_M \geq \binom{n}{2}+n-1-4 \geq \binom{n}{2}-1$, because we must have $n \geq 4$. We subtract four, because there are at most three disallowed moves and one intersection.

In the two other cases we have M' and M sharing one row or column, which disappears in the sum $M'+(-M)$. An example is presented below:

$$\Delta = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 & -1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = M' + (-M).$$

We assume that either Δ_1 or Δ_2 causes intersections, and denote the one causing them with Δ. Let the other one be Δ'. Because M' and $-M$ share one row with Δ, Δ also adds to a positive entry, because M needs to subtract from that. Let the position of that entry be (i_1, j_1).

- Assume that at least one of the entries e_i Δ' subtracts from satisfies $1 \leq e_i \leq r-1$. Then there must be at least one positive entry in the same column and one in the same row. If they are both in the row i_1 and column j_1, e_i is in the position (i_1, j_1). Otherwise, we can find 1-entries that do not use the row i_1 and the column j_1 for each e_i Δ' subtracts from satisfying $1 \leq e_i \leq r-1$. Denote them with (i_2, j_2) and (i_3, j_3). It might be that (i_3, j_3) does not exist or $(i_2, j_2) = (i_3, j_3)$. We have $u \geq E_n(i_1, j_1) + E_n(i_2, j_2) + E_n(i_3, j_3)$. They can all be put in the same $u_1 \in V(G(n, 1))$ in the construction of the proof of Lemma 2.8. Thus, by Lemma 2.8, we have $u = u_1 + u'$, where u_1 is such that it does not contain the entries at most $r-1$ Δ or Δ' subtract from. Because $d(u_1) = \binom{n}{2}$, and the choices $M = \Delta_1$ and $M = \Delta_2$ as well as intersections are avoided in u_1, we have $c_M \geq \binom{n}{2} - 1$.

- Now, assume that both of the entries Δ' subtracts from are r. As before, decompose $u = u_1 + u'$ using Lemma 2.8. This time, we cannot avoid the entries used by Δ', but they will surely be large enough to be usable by M. Again, u_1 does not contain the entries subtracted from by Δ. Hence, we cannot have intersections of the other type occuring with moves from u_1 and have to only avoid $M = -\Delta_1$, because Δ_2 adds to zero-entries, and thus $M = -\Delta_2$ is not included in $d(u_1)$. We have $c_M \geq \binom{n}{2} - 1$.

In the latter case, Δ' cannot cause intersections because of the assumption that Δ' subtracts from r-entries, but in the former case it could. Because the entries Δ' subtracts from are in u', the calculations hold even if intersections of the type $M' - M = \Delta'$ are assumed possible. ♫

The last result in this paper concerns the diameter of $G(n, r)$:

Proposition 2.12. *The diameter of $G(n, r)$ is $(n-1)r$.*

Proof. Every row sum is r, and each of the positive entries can be selected to be subtracted from. Therefore, r changes are enough to transform a row to any other. The n:th row must be correct at least after changing the $(n-1)$:th row, because otherwise we would have to change an already correct row to incorrect. The maximal number of changes needed is then $(n-1)r$, and $\operatorname{diam}(G(n, r)) \leq (n-1)r$.

Now, it suffices to show that $\operatorname{diam}(G(n, r)) \geq (n-1)r$. Take the diagonal matrix

$$A = \begin{pmatrix} r & 0 & \cdots & 0 \\ 0 & r & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r \end{pmatrix}.$$

The coordinates of the nonzero-entries are of the form $(i, i), i \in \mathbb{N} \cap [1, n]$. Consider permuting the rows so that $(i, i) \mapsto (i, i-1), i \neq 1$, and $(1, 1) \mapsto (1, n)$. The result is

$$A' = \begin{pmatrix} 0 & 0 & \cdots & r \\ r & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix},$$

and the permutation matrix

$$P = \begin{pmatrix} 0 & 0 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

On the other hand, $A = rI$. If p is the number of operations needed to change $\frac{1}{r}A' = P$ to I, the number of operations needed to change A' to A is clearly pr.

Consider this procedure: start from the row $i = 1$. Find the row which has its 1-entry in the column i, in this case the second row, and swap the rows. Repeat this for each of the rows except for the n:th one. Before the $(n-1)$:th row is swapped for the second time, it will have its 1-entry in the n:th column, so by interchanging it with the n:th row we will get to I.

In our procedure, each of the swaps corrects the place of one one-entry except for the last one which corrects two. However, we might be able to use more swaps that correct two positions. These kind of interchanges require pairs of one-entries to be in positions of the form (i, j) and (j, i). Say that we swap (i, j) with (i', j') to get (i', j) and (i, j'). Assume $i > i'$. If $i' < j$ and $i < j'$, $j' > i'$. Thus, the number of entries in a position of the form $(j, i), i > j$ increases by at most one with each swap. There are $n-1$ positive entries in positions of the form $(i, j), i > j$ in P. To interchange the positions of two of $n-2$ entries (the entries not in the positions $(1, n)$ and some other) correcting both, we would then need at least one extra swap. Thus, the best possible result we could get this way is still $n-1$ swaps.

Each swap consists of one operation. Thus, $p = n-1$, and therefore we need $(n-1)r$ operations to make A' from A. Hence, $\mathrm{diam}(G(n, r)) \geq (n-1)r$, but because also $\mathrm{diam}(G(n, r)) \leq (n-1)r$, $\mathrm{diam}(G(n, r)) = (n-1)r$. ♫

References

[1] Persi Diaconis and Bernd Sturmfels. Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26** (1998), no. 1, 363–397.

[2] Mathias Drton, Bernd Sturmfels and Seth Sullivant. *Lectures on Algebraic Statistics*. Oberwolfach Seminars, Vol. 39. *Birkhäuser, Basel*, 2009. 172 pp.

[3] Reinhard Diestel. *Graph theory*. Fourth edition. Graduate Texts in Mathematics, 173. *Springer, Heidelberg*, 2010. 437 pp.

[4] Alexander Engström. Private communication, June 2012.

[5] Samu Potka. Connectivity, in "Problem book for Sannäs workshop, August 9-10, 2012", edited by Alexander Engström, 2012.

[6] Dénes Kőnig. Über Graphen und ihre Anwendung auf Determinantentheorie und Mengenlehre. *Math. Ann.* **77** (1916), 453–465.

[7] Gui Zhen Liu. Proof of a conjecture on matroid base graphs. *Sci. China Ser. A* **33** (1990), no. 11, 1329–1337.

[8] Anders Björner and Kathrin Vorwerk. Connectivity of chamber graphs of buildings and related complexes. *European J. Combin.* **31** (2010), no. 8, 2149–2160.

Samu Potka
Aalto University
Department of Mathematics and Systems Analysis
PO Box 11100
FI-00076 Aalto
Finland
e-mail: samu.potka@aalto.fi