Thin Lehman Matrices and Their Graphs

Jonathan Wang

Department of Mathematics
Harvard University, Cambridge, MA 02138, USA

`jpwang@fas.harvard.edu`

Abstract

Two square $0, 1$ matrices A, B are a pair of Lehman matrices if $AB^T = J + dI$, where J is the matrix of all 1s and d is a positive integer. It is known that there are infinitely many such matrices when $d = 1$, and these matrices are called thin Lehman matrices. An induced subgraph of the Johnson graph may be defined given any Lehman matrix, where the vertices of the graph correspond to rows of the matrix. These graphs are used to study thin Lehman matrices. We show that any connected component of such a graph determines the corresponding rows of the matrix up to permutations of the columns. We also provide a sharp bound on the maximum clique size of such graphs and give a complete classification of Lehman matrices whose graphs have at most two connected components. Some constraints on when a circulant matrix can be Lehman are also provided. Many general classes of thin Lehman matrices are constructed in the paper.

1 Introduction

Lehman matrices were defined by Lütolf and Margot [7] to aid in the classification of minimally nonideal matrices, which are a key tool for understanding when the set covering problem can be solved using linear programming (we refer the reader to [2] for more information on minimally nonideal matrices). Lehman matrices lie at the heart of Lehman's central theorem on minimally nonideal matrices [5, 6]. He showed that for $m \geq n$ almost every $m \times n$ minimally nonideal matrix contains a unique $n \times n$ Lehman matrix. Bridges and Ryser [1] showed that every Lehman matrix is r-regular for some integer $r \geq 2$, i.e., each row and column sums to r. Two infinite families of Lehman matrices are known: the point-line incidence matrices of finite nondegenerate projective planes, a widely studied topic [4], and thin Lehman matrices. Thin Lehman matrices were defined and studied by Cornuéjols et al. [3].

Two square $n \times n$ matrices A, B form a pair of *Lehman matrices* if each matrix has only $0, 1$ as entries, and $AB^T = J + dI$ for some positive integer d (where J is the matrix of all ones). Lütolf and Margot enumerated all Lehman matrices with $n \leq 11$. If $A = B$, then $AA^T = J + dI$, and A is by definition the point-line incidence matrix of a nondegenerate projective plane of order d. The classification of finite nondegenerate projective planes is an open problem, and the only known orders are prime powers [4]. A Lehman matrix is called *thin* in the case $d = 1$. A matrix is *circulant* if each row is a right 1-cyclic shift of the previous row. Given integers $r, s \geq 2$ and $n = rs - 1$, let the circulant matrix C_n^r be the $n \times n$ matrix with columns indexed by $\mathbb{Z}/n\mathbb{Z}$ and its ith row equal to the incidence vector of $\{i, i+1, \ldots, i+r-1\}$, i.e., the $0, 1$ vector that has 1s in the specified columns, for $i \in \mathbb{Z}/n\mathbb{Z}$. Also define the $n \times n$ circulant matrix D_n^s in the same way except with rows equal to the incidence vectors of $\{i, i+r-1, i+2r-1, \ldots, i+(s-1)r-1\}$. Cornuéjols et al. [3] noted that C_n^r, D_n^s form a thin Lehman pair, which shows that there are infinitely many thin Lehman matrices. Given a Lehman matrix A, Cornuéjols et al. introduced a graph G_A, which we call the *Johnson subgraph induced by* A, to study properties of the matrix A. The graph G_A has the rows of A as vertices, and two rows are adjacent if each row has all but one 1 in the same column as the other row.

In this paper, we continue the study of thin Lehman matrices. We investigate the Johnson subgraphs associated to thin Lehman matrices, which have particularly simple structures. In Section 3, we show that the structures of a Lehman matrix and its graph are closely related. Our main result shows that any connected component of the graph G_A determines the corresponding rows of A up to permutations of the columns. Bounds on the maximum clique size and maximum degree of a Johnson subgraph of a thin Lehman matrix are given in Section 4. We also prove that some of the bounds given are sharp. We believe the new restrictions we impose on thin Lehman matrices will make it easier to enumerate them. In Section 5, all Lehman matrices with graphs containing at most two connected components are classified. Lastly, the induced Johnson subgraph is used to provide constraints on when a circulant matrix is Lehman in Section 6. A complete classification of all Lehman matrices is, however, still lacking. A Lehman matrix may not be determined by its graph once the graph has more than two connected components, which reveals one limitation of the induced Johnson subgraph.

2   Preliminaries

A $0, 1$ matrix is r-regular if every row and column has exactly r ones. We restate the theorem of Bridges and Ryser [1] on regularity of Lehman matrices.

Theorem 2.1 ([1], Theorem 1.2). *Let A, B be a Lehman pair. Then there exist integers $r, s \geq 2$ such that A is r-regular, B is s-regular, and $rs = n + d$. Moreover, $B^T A = AB^T = J + dI$.*

Throughout this paper, we will use A, B to denote a Lehman pair of $n \times n$ matrices with $AB^T = J + dI$, where A is r-regular, B is s-regular, and $rs = n + d$. Observe that

$$A(B^T - \tfrac{1}{r}J) = dI \implies B^T = dA^{-1} + \tfrac{1}{r}J,$$

which shows that d and B are unique given A, since B must be $0, 1$. The matrix B is called the *Lehman dual* of A. We also see that A and B are invertible.

Two matrices A_1 and A_2 are isomorphic, denoted $A_1 \simeq A_2$, if one can be obtained from the other by permutations of rows and/or columns. Equivalently, there exist permutation matrices P and Q such that $PA_1Q = A_2$. If $A_1B_1^T = J + dI$, then

$$(PA_1Q)(PB_1Q)^T = PA_1QQ^TB_1^TP^T = P(J + dI)P^T = J + dI,$$

so A_2 is also a Lehman matrix.

Cornuéjols et al. [3] noted that if an $n \times n$ Lehman matrix A is 2-regular, then $n \geq 3$ odd, and $A \simeq C_n^2$. Therefore, we will assume $r, s > 2$ in the paper.

Let $\mathbb{Z}_{\geq 0}$ denote the set of nonnegative integers. We also define the intervals of integers $[a, b] := \{c \in \mathbb{Z} \mid a \leq c \leq b\}$, $[a, b) := [a, b] \setminus \{b\}$, $(a, b] := [a, b] \setminus \{a\}$, $(a, b) := [a, b] \setminus \{a, b\}$, and $[a] := [1, a]$. Unless otherwise specified, we index rows and columns of an $n \times n$ matrix by $[n]$. Since we are working with $0, 1$ matrices, we may identify the rows and columns of a matrix with subsets of $[n]$. Given an $n \times n$ $0, 1$ matrix A and $i \in [n]$, define

$$\operatorname{row}_i(A) = \{j \in [n] \mid a_{ij} = 1\} \subset [n]$$

to be the set of column indices where row i has a 1. Define $\operatorname{col}_i(A)$ analogously.

We provide some important observations that will be used in later proofs.

Remark 2.2. Observe that $AB^T = J + dI$ is equivalent to $|\operatorname{row}_i(A) \cap \operatorname{row}_i(B)| = d + 1$ and $|\operatorname{row}_i(A) \cap \operatorname{row}_j(B)| = 1$ for $i \neq j$. By Theorem 2.1, $AB^T = J + dI$ implies $B^TA = J + dI$, so we also have $|\operatorname{col}_i(A) \cap \operatorname{col}_i(B)| = d + 1$ and $|\operatorname{col}_i(A) \cap \operatorname{col}_j(B)| = 1$ for $i \neq j$. We therefore deduce that for $i \neq j$ and any k,

$$|\operatorname{row}_i(A) \cap \operatorname{row}_j(A) \cap \operatorname{row}_k(B)| \leq 1. \tag{1}$$

Since A is invertible, the row vectors of A must be linearly independent. Therefore,

$$\operatorname{row}_i(A) \neq \operatorname{row}_j(A) \tag{2}$$

for $i \neq j$. Note that (1) and (2) also hold with A and B switched or with rows replaced by columns.

We now define the Johnson subgraph G_A induced by an r-regular $0, 1$ matrix A. The vertices $V(G_A)$ are the rows $[n]$ of A. Two rows i and j are adjacent in G_A if

$$|\operatorname{row}_i(A) \cap \operatorname{row}_j(A)| = r - 1.$$

The vertices of the *Johnson graph* $J(n, r)$ are the size r subsets of $[n]$, and two vertices are adjacent if their intersection has size $r - 1$. Thus, G_A is the subgraph of the Johnson graph induced by the rows of A. If A is a Lehman matrix with $d > 1$, then Remark 2.2 implies that G_A has no edges. We will therefore mainly use the graph G_A to study A when A is a thin Lehman matrix.

Example. The Johnson subgraph induced by C_n^r is a single cycle with n vertices.

3 Structure of graphs

In this section we explore the relation between the structure of a thin Lehman matrix A and the Johnson subgraph G_A. We show that the structures of interest in these graphs are paths and cliques. At the end of the section we prove the following theorem.

Theorem 3.1. *Suppose A is a thin Lehman matrix. Let $W \subset V(G_A)$ be the vertices of a connected component of G_A. Then each $row_i(A)$ for $i \in W$ is determined by G_A up to permutations of the columns.*

We believe that the structure of the induced Johnson subgraphs will aid in the enumeration of all nonisomorphic thin Lehman matrices.

Note that if $A_1 \simeq A_2$, then $G_{A_1} \simeq G_{A_2}$ since permuting rows and columns does not affect the size of row intersections. Unfortunately, the converse does not hold. We provide a counterexample below.

Example. We give two thin Lehman matrices A_1, A_2 with $n = 14, r = 3$ such that $G_{A_1} \simeq G_{A_2} \simeq P_1 \sqcup P_2 \sqcup P_3 \sqcup P_4 \sqcup P_4$, where P_k is a path with k vertices. We checked with a computer program that $A_1 \not\simeq A_2$. In the diagram, dots represent 1s and blank spaces represent 0s.
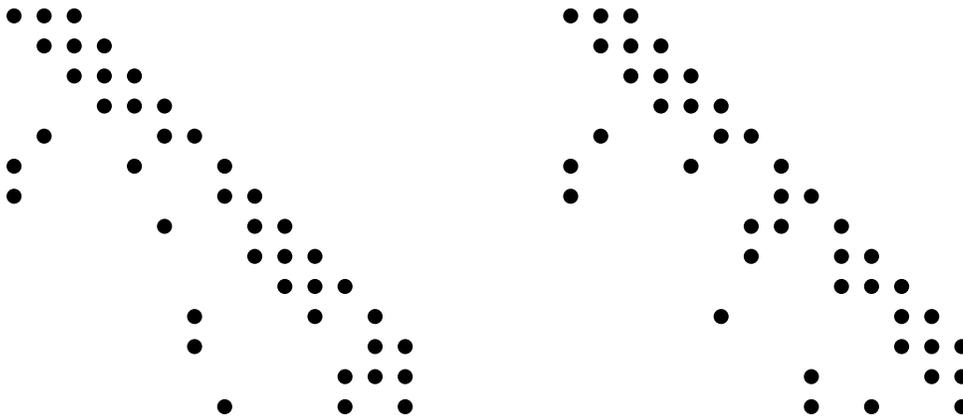
For the rest of this section, we assume A is an r-regular thin Lehman matrix, B is the s-regular dual, and $n = rs - 1$.

3.1 Paths

We first build up some machinery to prove the following key lemma on the structure of the rows in A corresponding to a subpath of G_A.

Lemma 3.2. *Let $[k]$ be the vertices of a subpath of G_A such that $i, i+1$ are adjacent for $i < k$, but $i, i+2$ are not adjacent for any $i < k-1$. Then either $A \simeq C_n^r$ or the columns of A can be permuted such that $\text{row}_i(A) = [i, i+r)$ for $i \in [k]$.*

If rows $[k]$ of A satisfy $\text{row}_i(A) = [i, i+r)$ for $i \in [k]$, then we say these rows have a *cascading structure*. We show that the cascading structure of rows in A determines part of the dual matrix B.

Lemma 3.3. *Suppose* $\text{row}_i(A) = [i, i+r)$ *for* $i \in [k]$. *Then there exists a permutation matrix* P *such that* $\text{row}_i(A) = \text{row}_i(AP)$ *and*

$$\text{row}_i(BP) \cap [k+r-1] = \{i - r\ell, i + (r-1) + r\ell \mid \ell \in \mathbb{Z}_{\geq 0}\} \cap [k+r-1]$$

for $i \in [k]$. *That is, we can simultaneously permute the columns of* A *and* B *such that* B *has the above form without changing the first* k *rows of* A.

Proof. The claim is clear for $k = 1$, so assume otherwise. For $i \in [2, k)$,

$$\text{row}_{i-1}(A) \cap \text{row}_i(A) = [i, i+r-2] \quad \text{and} \quad \text{row}_i(A) \cap \text{row}_{i+1}(A) = [i+1, i+r-1],$$

so $\{i, i+r-1\} = \text{row}_i(A) \cap \text{row}_i(B)$ by (1). Since $\text{row}_1(A) \cap \text{row}_1(B) = [r] \cap \text{row}_1(B)$ contains two elements and $\text{row}_1(A) \cap \text{row}_2(A) = [2, r]$, we must have $1 \in \text{row}_1(B)$ by (1). Similarly, $\text{row}_{k-1}(A) \cap \text{row}_k(A) = [k, k+r-2]$ implies $k+r-1 \in \text{row}_k(B)$.

Suppose $k < r$. By assumption, $[k] \subset \text{col}_i(A)$ for $i \in [k, r]$. The analog of (1) for columns implies that

$$|[k] \cap \text{col}_j(B)| \leq |\text{col}_k(A) \cap \text{col}_r(A) \cap \text{col}_j(B)| \leq 1 \tag{3}$$

for any j. We have shown above that $i \in \text{row}_i(B)$ for $i \in [1, k)$, which implies $i \in \text{col}_i(B)$ for $i \in [1, k)$. By (3), we have $1 \notin \text{col}_i(B)$ for $i \in [2, k)$. Therefore, $\text{row}_1(B) \cap [1, k) = \{1\}$, so $\text{row}_1(B)$ contains one element in $[k, r]$. Let $\text{row}_1(B) \cap [k, r] = \{j\}$. Since $[k] \subset \text{col}_j(A) \cap \text{col}_r(A)$, we can swap columns j, r in both A and B to get $\{1, r\} \subset \text{row}_1(B)$ while the first k rows of A stay the same. Thus, we have

$$\{i, i+r-1\} \subset \text{row}_i(B) \quad \text{for } i \in [k-1]$$

and $k+r-1 \in \text{row}_k(B)$. Observe that $i \in \text{col}_{i+r-1}(B)$ for $i \in [k]$. Now (3) implies that $k \notin \text{col}_{i+r-1}(B)$ for $i \in [k-1]$, or equivalently $\text{row}_k(B) \cap [r, k+r) = \{k+r-1\}$. Hence $\text{row}_k(B)$ contains one element in $[k, r)$. Let $\text{row}_k(B) \cap [k, r) = \{j\}$. As $[k] \subset \text{col}_k(A) \cap \text{col}_j(A)$, we can swap columns k, j in both A and B to get $\{k, k+r-1\} \subset \text{row}_k(B)$ without changing the first k rows of A. We conclude that

$$\{i, i+r-1\} \subset \text{row}_i(B) \quad \text{for } i \in [k]. \tag{4}$$

Suppose $k \geq r$. Then $[r-1] = \text{col}_{r-1}(A) \cap \text{col}_r(A)$. Since $i \in \text{col}_i(B)$ for $i \in [2, r)$, the analog of (1) for columns implies that $1 \notin \text{col}_i(B)$ for $i \in [2, r)$. Therefore, $\text{row}_1(A) \cap \text{row}_1(B) = \{1, r\}$. Similarly, $[k-r+2, k] = \text{col}_k(A) \cap \text{col}_{k+1}$, and $i \in \text{col}_{i+r-1}(B)$ for $i \in [k-r+2, k)$ implies that $k \notin \text{col}_{i+r-1}(B)$ for $i \in [k-r+2, k)$. Therefore, $\text{row}_k(A) \cap \text{row}_k(B) = \{k, k+r-1\}$. We deduce that (4) holds.

In both cases, (4) is true. Fix $i \in [k]$. Given $i - r\ell \in \text{row}_i(B)$ for $\ell \in \mathbb{Z}_{\geq 0}$ such that $i - r(\ell+1) > 0$, we have $\text{row}_{i-r(\ell+1)+1}(A) \cap \text{row}_i(B) = \{i - r\ell\}$. However, $\text{row}_i(B)$ must

intersect $\text{row}_{i-r(\ell+1)}(A)$, so $i - r(\ell+1) \in \text{row}_i(B)$. Similarly $i + (r-1) + r\ell \in \text{row}_i(B)$ for $\ell \in \mathbb{Z}_{\geq 0}$ implies $i + r - 1 + r(\ell+1) \in \text{row}_i(B)$, assuming $i + (r-1) + r(\ell+1) \leq k + r - 1$. Therefore, starting with $\ell = 0$, we have by induction that

$$\text{row}_i(B) \cap [k+r-1] = \{i - r\ell, i + (r-1) + r\ell \mid \ell \in \mathbb{Z}_{\geq 0}\} \cap [k+r-1]. \qquad \square$$

Observe that the first $n - r + 1 = r(s-1)$ rows of C_n^r have the cascading structure. We show that if the same number of rows in A have the cascading structure, then A must actually be isomorphic to C_n^r.

Lemma 3.4. *If* $\text{row}_i(A) = [i, i+r)$ *for* $i \in [r(s-1)]$, *then* $A \simeq C_n^r$.

Proof. Permute A and B to have the form described in Lemma 3.3. Since A has dimension $n = rs - 1$ and is r-regular, the size of A forces $\text{col}_1(A) = \{1\} \cup [r(s-1) + 1, rs)$ and $\text{col}_n(A) = [r(s-1), rs)$. Then by (1),

$$|\text{col}_i(B) \cap [r(s-1) + 1, rs)| \leq 1$$

for all i. For $i \in [r-1]$, Lemma 3.3 implies

$$\{i, r+i, \ldots, r(s-2) + i\} = \text{col}_i(B) \cap [r(s-1)].$$

Since B is s-regular, each $\text{col}_i(B)$ contains exactly one additional element. By permuting the last $r-1$ rows of A and B, we can assume $r(s-1) + i \in \text{col}_i(B)$. Taking $1 \leq i < j < r$, observe that $i \in \text{col}_i(A) \cap \text{col}_j(A) \cap \text{col}_i(B)$. This implies $r(s-1) + i \notin \text{col}_j(A)$ by the column analog of (1). Now using r-regularity of A, we must have

$$[r(s-1) + i, n] \subset \text{col}_i(A) \implies [i] \subset \text{row}_{r(s-1)+i}(A)$$

for $i \in [r-1]$. Starting with $i = 1$, only columns $[r(s-1) + 1, rs - i]$ and rows $[r(s-1) + 1, rs - i]$ of A do not have r ones already allocated. By r-regularity, we must have

$$\text{col}_{rs-i}(A) = (r(s-1) - i, rs - i] \quad \text{and} \quad \text{row}_{rs-i}(A) = [r-i] \cup [rs - i, rs).$$

This fills row $rs - i$ and column $rs - i$ of A. Proceeding inductively for $i = 1, \ldots, r-1$, we fill the matrix A and conclude that $A \simeq C_{rs-1}^r = C_n^r$. $\qquad \square$

The next lemma demonstrates that if A contains the first $n' - r + 2$ rows of $C_{n'}^r$ for some n', then $n = n'$ and $A \simeq C_n^r$.

Lemma 3.5. *If* $\text{row}_i(A) = [i, i+r)$ *for* $i \in [k]$ *and* $\text{row}_{k+1}(A) = \{1\} \cup [k+1, k+r)$, *then* $k = r(s-1)$ *and* $A \simeq C_{k+r-1}^r$.

Proof. Write $k = r(t-1) + \ell$ for $t \geq 1$ and $0 \leq \ell < r$. Since

$$[i+1, i+r) \subset \text{row}_i(A) \cap \text{row}_{i+1}(A)$$

for $i \in [k]$, we have by Remark 2.2 that $i \in \mathrm{row}_i(B)$ for $i \in [k]$, and $1 \in \mathrm{row}_{k+1}(B)$. By the cascading structure of the first k rows of A, we deduce that

$$\{1, r+1, \ldots, r(t-1)+1, k+1\} \subseteq \mathrm{col}_1(B). \tag{5}$$

Suppose $\ell > 0$. Then $r(t-1)+1$ and $k+1$ are distinct. This contradicts $|\mathrm{col}_{k+1}(A) \cap \mathrm{col}_1(B)| = 1$ since $[r(t-1)+1, k+1] \subset \mathrm{col}_{k+1}(A)$. Therefore, $\ell = 0$ and $k = r(t-1)$.

We claim that B is t-regular. Suppose that $1 \in \mathrm{row}_i(B)$ for $i > k+1$. By the cascading structure of the first k rows of A, this implies $\{1, r+1, \ldots, r(t-1)+1\} \subseteq \mathrm{row}_i(B)$. This contradicts $|\mathrm{row}_{k+1}(A) \cap \mathrm{row}_i(B)| = 1$ since $\{1, r(t-1)+1\} \subset \mathrm{row}_{k+1}(A)$. Therefore, B is t-regular, $s = t$, and $n = rs - 1 = k + r - 1$.

Lemma 3.4 implies that $A \simeq C_n^r = C_{k+r-1}^r$. $\qquad\square$

We now use the previous lemmas to present the proof of Lemma 3.2.

Proof of Lemma 3.2. We prove the lemma by induction on the rows of A. We can assume $\mathrm{row}_1(A) = [r]$ and $\mathrm{row}_2(A) = [2, r+1]$. Now suppose $\mathrm{row}_i(A) = [i, i+r)$ for all $i \in [\ell]$ and $\ell > 1$. Then we apply Lemma 3.3 to assume

$$\mathrm{row}_i(B) \cap [\ell + r - 1] = \{i - r\mathbb{Z}_{\geq 0}, i + r - 1 + r\mathbb{Z}_{\geq 0}\} \cap [\ell + r - 1].$$

By assumption, rows $\ell, \ell+1$ are adjacent in A but rows $\ell-1, \ell+1$ are not, so

$$[\ell, \ell + r - 1) \not\subset \mathrm{row}_{\ell-1}(A) \cap \mathrm{row}_\ell(A) \cap \mathrm{row}_{\ell+1}(A).$$

Therefore, $\ell + r - 1 \in \mathrm{row}_{\ell+1}(A)$. Since $\{\ell, \ell + r - 1\} \subset \mathrm{row}_\ell(B)$, $\ell \notin \mathrm{row}_{\ell+1}(A)$. By adjacency, we deduce that

$$[\ell + 1, \ell + r) \subset \mathrm{row}_{\ell+1}(A).$$

Suppose $\mathrm{row}_{\ell+1}(A) = \{i\} \cup [\ell + 1, \ell + r)$ for $i < \ell$. Then rows $[i, \ell+1]$ satisfy Lemma 3.5. Therefore, we either get a contradiction or $A \simeq C_n^r$.

Otherwise $\mathrm{row}_{\ell+1}(A) \not\subset [\ell + r - 1]$, and we can permute columns to assume $\mathrm{row}_{\ell+1}(A) = [\ell + 1, \ell + r]$. This completes the inductive step. Hence either $A \simeq C_n^r$ or we can permute the columns of A such that $\mathrm{row}_i(A) = [i, i+r)$ for $i \in [k]$. $\qquad\square$

Corollary 3.6. *Suppose G_A contains a cycle where vertices of distance 2 apart in the cycle are not adjacent in G_A. Then $A \simeq C_n^r$.*

Proof. Let rows $[k]$ correspond to the vertices of the cycle. We must have $k > 3$ in order for the assumptions to hold. Suppose $A \not\simeq C_n^r$. Then by Lemma 3.2, $\mathrm{row}_i(A) = [i, i+r)$ for $i \in [k]$. Since $|\mathrm{row}_1(A) \cap \mathrm{row}_k(A)| = \max(r - k + 1, 0) < r - 2$, rows 1 and k cannot be adjacent, which is a contradiction. $\qquad\square$

3.2 Cliques

In the previous section we considered triangle-free paths in the graph G_A. We now look at the structure of triangles, and in greater generality, cliques in G_A. In particular, we provide a lemma analogous to Lemma 3.2 for cliques.

Lemma 3.7. *If rows $[k]$ form a k-clique in G_A, then the columns of A can be permuted such that $\mathrm{row}_i(A) = [r-1] \cup \{r+i-1\}$ and $\{i, r+i-1\} \subset \mathrm{row}_i(B)$ for $i \in [k]$.*

Proof. Permute the columns so $\mathrm{row}_1(A) = [r]$ and $\{1, r\} \subset \mathrm{row}_1(B)$. Thus for $i > 1$,

$$|\mathrm{row}_i(A) \cap \{1, r\}| \leq 1.$$

Since each row $i \in (1, k]$ is adjacent to row 1, we have either

$$[r-1] \subset \mathrm{row}_i(A) \quad \text{or} \quad [2, r] \subset \mathrm{row}_i(A).$$

By possibly switching columns 1 and r, we may assume without loss of generality that $[r-1] \subset \mathrm{row}_2(A)$. Suppose $[2, r] \subset \mathrm{row}_i(A)$ for some $i > 2$. Since $\{1, 2, i\} \subset \mathrm{col}_2(A)$ and $1 \in \mathrm{col}_r(B)$, we deduce that $r \notin \mathrm{row}_i(B)$. Since rows 2 and i must be adjacent,

$$\mathrm{row}_i(A) \setminus \{r\} \subset \mathrm{row}_2(A).$$

Thus, $\mathrm{row}_i(B)$ must contain two elements in $\mathrm{row}_2(A)$, which is a contradiction. Therefore, $[r-1] \subset \mathrm{row}_i(A)$ for $i \in [k]$. No two rows of A may be equal, so we can permute columns $[r, n]$ to assume

$$\mathrm{row}_i(A) = [r-1] \cup \{r+i-1\}.$$

Since $[r-1] \subset \mathrm{row}_1(A) \cap \mathrm{row}_i(A)$, we must have $r+i-1 \in \mathrm{row}_i(B)$ by (1). Additionally, we know that

$$[k] \subset \bigcap_{i=1}^{r-1} \mathrm{col}_i(A),$$

so no column of B can have two 1s in the first k rows. We may therefore permute the first $r-1$ columns of A and B simultaneously to assume $\{i, r+i-1\} \subset \mathrm{row}_i(B)$. $\square$

Example. We give an example of the rows in A and B corresponding to a clique in G_A. Here $r = 4$ and $k = 3$. The diagram on the left shows $\mathrm{row}_i(A)$ and the diagram on the right shows $\mathrm{row}_i(B) \cap [k+r-1]$ for $i \in [k]$.

Remark 3.8. Suppose rows $[k]$ form a clique in G_A. By Lemma 3.7, we can permute columns to get $\mathrm{row}_i(A) = [r-1] \cup \{r+i-1\}$ for $i \in [k]$. Then

$$[r-1] = \mathrm{row}_1(A) \cap \mathrm{row}_2(A) = \bigcap_{i=1}^{k} \mathrm{row}_i(A). \tag{6}$$

3.3 Connected components

We define a *clique tree* as follows. Start with a tree T. Create a new graph equal to the disjoint union of $|V(T)|$ cliques of arbitrary size. For each edge $ij \in E(T)$, choose one vertex in clique i and one vertex in clique j of the new graph, and combine the two chosen vertices into one vertex. We additionally require that the new graph does not contain a vertex incident to more than two maximal cliques. We call the resulting graph a clique tree. Note that a triangle-free clique tree is a path.

In this section, we show that if $A \not\simeq C_n^r$, then the connected components of G_A must be clique trees. Moreover, connected components containing a triangle must contain fewer than r vertices. At the end of the section we prove that a connected component of G_A uniquely determines, up to permutation of the columns, the corresponding rows of A.

Lemma 3.9. *Suppose rows $1, 3, 4$ are all adjacent to row 2 in A. Then two rows in $\{1, 3, 4\}$ must be adjacent.*

Proof. Suppose rows 1 and 3 are not adjacent. Then using Lemmas 3.2 and 3.3, we can permute columns such that $\mathrm{row}_i(A) = [i, i+r)$ for $i \in [3]$ and $\{2, r+1\} \subset \mathrm{row}_2(B)$. Since rows 2 and 4 are adjacent in A, we must have either $[2, r] \subset \mathrm{row}_4(A)$ or $[3, r+1] \subset \mathrm{row}_4(A)$. Therefore, row 4 is adjacent to row 1 or 3 in A. $\quad\square$

Note that the lemma implies that the only possible trees in G_A are paths. We next prove that if a vertex is adjacent to two vertices of a clique in G_A, then it must be adjacent to every vertex in the clique.

Lemma 3.10. *Suppose rows $[k]$ of A form a clique in G_A, and row $k + 1$ is adjacent to rows 1 and 2 in G_A. Then row $k + 1$ is adjacent to every row i for $i \in [k]$.*

Proof. Lemma 3.7 implies that $\mathrm{row}_i(A) = [r-1] \cup \{r + i - 1\}$ for $i \in [k]$. Observe that since rows $1, 2, k+1$ form a triangle, (6) implies that

$$[r - 1] = \mathrm{row}_1(A) \cap \mathrm{row}_2(A) \subset \mathrm{row}_{k+1}(A).$$

Therefore, row $k + 1$ is adjacent to every row i for $i \in [k]$. $\quad\square$

Observe that if two cliques share at least two vertices, Lemma 3.10 shows that their union must also be a clique. Now Lemma 3.9 implies that any vertex is incident to at most two maximal cliques. The previous two lemmas show that G_A essentially contains only paths and cliques. The next proposition will show that a connected component of G_A for $A \not\simeq C_n^r$ must indeed be a clique tree.

Lemma 3.11. *If G_A has a cycle that is not contained inside a clique, then $A \simeq C_n^r$.*

Proof. Suppose G_A contains such a cycle, and let the vertices of the cycle be $[k]$. If $k > 3$ and there exists a row $i \in [k]$ in the cycle with cyclically shifted rows $i - 1$ and $i + 1$ adjacent in G_A, consider instead the cycle with vertices $[k] \setminus \{i\}$. Repeating, we either reduce the cycle to a triangle or a cycle where vertices of distance 2 apart are not

adjacent. If the reduced cycle is a triangle, Lemma 3.10 implies that the original rows $[k]$ form a clique. Otherwise, the reduced cycle satisfies the conditions of Corollary 3.6, so $A \simeq C_n^r$. $\qquad \square$

Combining Lemmas 3.9, 3.10, and 3.11, we conclude the following theorem.

Theorem 3.12. *If $A \not\simeq C_n^r$, then each connected component of G_A is a clique tree.*

Corollary 3.13. *If G_A is triangle-free, then either $A \simeq C_n^r$ or G_A is a disjoint union of paths.*

We next give a bound on the size of connected components that do contain triangles.

Lemma 3.14. *If a connected component of G_A contains a triangle, then the component has fewer than r vertices.*

Proof. Suppose a connected component contains at least r vertices. We can then choose a subset of r vertices $W \subset V(G_A)$ such that the subgraph of G_A induced by W is connected and contains a triangle. Rearrange the rows so that $W = [r]$, and each row $j \in W \setminus \{1\}$ is adjacent to some $i < j$. Let $t_i \in W$ for $i \in [3]$ induce a triangle, with $t_1 < t_2 < t_3$.

We claim that
$$\left| \bigcap_{i=1}^{k} \mathrm{row}_i(A) \right| \geq r - (k-1)$$
for $k \in [r]$. We prove this by induction. The case $k = 1$ is clear. Now assume the claim is true for some k. Since $|\mathrm{row}_{k+1}(A) \cap \mathrm{row}_i(A)| = r - 1$ for some $i \leq k$, there is at most one element of $\mathrm{row}_i(A)$ that is not in $\mathrm{row}_{k+1}(A)$. Consequently there is at most one element of $\bigcap_{i=1}^{k} \mathrm{row}_i(A)$ that is not in $\mathrm{row}_{k+1}(A)$. Thus,
$$\left| \bigcap_{i=1}^{k+1} \mathrm{row}_i(A) \right| \geq r - (k-1) - 1 = r - ((k+1) - 1),$$
proving the claim.

Therefore, $|\bigcap_{i=1}^{r} \mathrm{row}_i(A)| \geq 1$, with equality if and only if
$$\bigcap_{i=1}^{k} \mathrm{row}_i(A) \neq \bigcap_{i=1}^{k+1} \mathrm{row}_i(A) \tag{7}$$
for every $k \in [r-1]$. Since t_1, t_2, t_3 induce a triangle in G_A, (6) implies that $\mathrm{row}_{t_1}(A) \cap \mathrm{row}_{t_2}(A) \subset \mathrm{row}_{t_3}(A)$. Thus, (7) does not hold for $k = t_3 - 1$. Therefore,
$$\left| \bigcap_{i=1}^{r} \mathrm{row}_i(A) \right| > 1.$$

Thus, there exist two columns of A with r ones in the same rows. This is a contradiction since A is invertible. $\qquad \square$

Corollary 3.15. *The graph G_A is connected if and only if $A \simeq C_n^r$.*

Proof. Suppose G_A is connected and not isomorphic to C_n^r. Since $n = rs - 1 > r$, Lemma 3.14 implies G_A is triangle-free. By Corollary 3.13, G_A is a disjoint union of paths. If G_A is a single path, Lemma 3.2 allows us to permute A such that $\mathrm{row}_i(A) = [i, i+r)$ for $i \in [n]$. Since A is a square matrix, this is impossible.

The other direction is obvious. $\qquad\square$

We now present the proof of our main result relating induced Johnson subgraphs and Lehman matrices.

Proof of Theorem 3.1. We may assume that $W = [k]$. If W induces a path, then Lemma 3.2 proves the claim. Otherwise, the connected component contains a triangle, so Lemma 3.14 implies $k < r$. From the proof of Lemma 3.14, we know that $|\bigcap_{i=1}^{k} \mathrm{row}_i(A)| \geq 2$. Therefore, there exist two columns of A that have 1s in rows $[k]$. Then (1) implies that

$$|\mathrm{col}_i(B) \cap [k]| \leq 1 \tag{8}$$

for all i. Rearrange W so that each row $j \in W \setminus \{1\}$ is adjacent to some $i < j$. We will prove inductively that for $\ell \in [k]$, the columns of A can be permuted such that

1. $\mathrm{row}_\ell(A) \subset [\ell + r - 1]$,

2. $[\ell, r) \subset \bigcap_{i=1}^{\ell} \mathrm{row}_i(A)$,

3. $\{\ell, \ell + r - 1\} = \mathrm{row}_\ell(A) \cap \mathrm{row}_\ell(B)$, and

4. rows $[\ell]$ are unique up to permutations of columns.

The claim is trivial for $\ell \leq 2$. Now assume the claim is true for all $i \leq \ell$ for some $\ell \in [2, k)$. By statement 3 of the inductive hypothesis,

$$i \in \mathrm{col}_i(B) \cap \mathrm{col}_{i+r-1}(B)$$

for $i \in [\ell]$. Therefore, (8) implies that $\ell + 1 \notin \mathrm{col}_i(B)$ for $i \in [\ell] \cup [r, r + \ell)$, or equivalently $\mathrm{row}_{\ell+1}(B)$ is disjoint from $[\ell] \cup [r, r + \ell)$. Since $[\ell, r) \subset \bigcap_{i=1}^{\ell} \mathrm{row}_i(A)$, (1) implies that

$$|\mathrm{row}_{\ell+1}(B) \cap [\ell, r)| \leq 1.$$

Thus, $\mathrm{row}_{\ell+1}(A) \cap \mathrm{row}_{\ell+1}(B) \not\subset [r + \ell - 1]$ because the intersection has size 2.

There exists $i \in [\ell]$ such that rows i and $\ell + 1$ in A are adjacent. Thus,

$$|\mathrm{row}_{\ell+1}(A) \setminus \mathrm{row}_i(A)| = 1,$$

so row $\ell + 1$ in A has exactly one 1 outside $[r + \ell - 1]$. We can permute columns to assume

$$r + \ell \in \mathrm{row}_{\ell+1}(A) \cap \mathrm{row}_{\ell+1}(B).$$

Thus, $\mathrm{row}_{\ell+1}(A) \subset [r+\ell]$, proving statement 1. Since rows i and $\ell+1$ are adjacent in A and $\{i, i+r-1\} \subset \mathrm{row}_i(A) \cap \mathrm{row}_i(B)$, we have either

$$\mathrm{row}_{\ell+1}(A) = \mathrm{row}_i(A) \cup \{r+\ell\} \setminus \{i\} \text{ or } \mathrm{row}_{\ell+1}(A) = \mathrm{row}_i(A) \cup \{r+\ell\} \setminus \{i+r-1\}. \quad (9)$$

Note that since $[\ell, r) \subset \mathrm{row}_i(A)$, in both cases $[\ell+1, r) \subset \mathrm{row}_{\ell+1}(A)$, asserting statement 2. We showed above that $\mathrm{row}_{\ell+1}(B)$ is disjoint from $[\ell] \cup [r, r+\ell]$. Since $r+\ell \in \mathrm{row}_{\ell+1}(B)$ and $\mathrm{row}_{\ell+1}(A) \subset [r+\ell]$, we deduce that

$$|\mathrm{row}_{\ell+1}(B) \cap [\ell+1, r)| = 1.$$

Since $[\ell+1, r) \subset \bigcap_{i=1}^{\ell+1} \mathrm{row}_i(A)$, we can permute the columns $[\ell+1, r)$ of A and B simultaneously to assume $\ell+1 \in \mathrm{row}_{\ell+1}(B)$, leaving the first $\ell+1$ rows of A unchanged. We now have $\{\ell+1, \ell+r\} \subset \mathrm{row}_{\ell+1}(A) \cap \mathrm{row}_{\ell+1}(B)$, proving statement 3.

To show uniqueness, we only need to show exactly one case in (9) is true. Take some row j such that rows i and j are adjacent in G_A and $j \leq \ell$ (this is possible since $\ell \geq 2$ and rows 1 and 2 are adjacent). Let

$$\{a\} = \{i, i+r-1\} \cap \mathrm{row}_j(A) = \mathrm{row}_i(A) \cap \mathrm{row}_i(B) \cap \mathrm{row}_j(A),$$

which is a one element set since $|\mathrm{row}_i(A) \cap \mathrm{row}_j(A)| = r-1$ implies $|\{i, i+r-1\} \cap \mathrm{row}_j(A)| \geq 1$ and $|\mathrm{row}_i(B) \cap \mathrm{row}_j(A)| = 1$ implies $|\{i, i+r-1\} \cap \mathrm{row}_j(A)| \leq 1$. Let

$$\{b\} = \{i, i+r-1\} \setminus \{a\}.$$

Since rows i, j are adjacent, we have $\mathrm{row}_i(A) \cap \mathrm{row}_j(A) = \mathrm{row}_i(A) \setminus \{b\}$. If $\mathrm{row}_{\ell+1}(A) = \mathrm{row}_i(A) \cup \{r+\ell\} \setminus \{b\}$, then rows $j, \ell+1$ are also adjacent. If rows $j, \ell+1$ are adjacent in G_A, then rows $i, j, \ell+1$ form a triangle, so

$$a \in \mathrm{row}_i(A) \cap \mathrm{row}_j(A) \subset \mathrm{row}_{\ell+1}(A)$$

by (6). Therefore,

$$\mathrm{row}_{\ell+1}(A) = \mathrm{row}_i(A) \cup \{r+\ell\} \setminus \{b\}$$

if and only if rows $j, \ell+1$ are adjacent in G_A.

Thus, exactly one case in (9) can be true, and we conclude that the rows $[\ell+1]$ are uniquely determined by the graph G_A up to permutations of the columns. This completes the inductive step and proves the theorem. $\qquad\square$

4 Bounds on clique size and degree

In this section, we continue to assume that A is an r-regular thin Lehman matrix of dimension $n \times n$, B is its s-regular dual, and $n = rs - 1$. In Section 3, we showed that either $A \simeq C_n^r$ or the connected components of G_A are clique trees. We are therefore interested in the maximum size of cliques in G_A, which would give us a better idea of the possible structures of the induced Johnson subgraph.

We provide a sharp upper bound on the maximum clique size of the graph G_A, and give a relation between the clique sizes of G_A and G_{B^T}. We also give an upper bound on the maximum degree of G_A.

4.1 Maximum clique size

Let the clique number $\omega(G_A)$ denote the number of vertices in a maximal clique in G_A. The equation $AB^T = J + I$ gives a condition on the size of cliques in G_A and G_{B^T}.

Lemma 4.1. *The clique numbers $\omega(G_A)$ and $\omega(G_{B^T})$ are equal.*

Proof. Let $\omega(G_A) = k$. Use Lemma 3.7 to permute rows/columns such that

$$\mathrm{row}_i(A) = [r-1] \cup \{i+r-1\} \quad \text{and} \quad \{i, i+r-1\} \subset \mathrm{row}_i(B)$$

for $i \in [k]$. Now consider any row $i > k$ with $r \in \mathrm{row}_i(B)$. Then $|\mathrm{row}_1(A) \cap \mathrm{row}_i(B)| = 1$ implies $[r-1] \cap \mathrm{row}_i(B) = \emptyset$. Now for $j \in (1, k]$, $|\mathrm{row}_j(A) \cap \mathrm{row}_i(B)| = 1$ implies $j + r - 1 \in \mathrm{row}_i(B)$. Thus, $[r, r+k) \subset \mathrm{row}_i(B)$. Since $[k] \subset \mathrm{col}_i(A)$ for $i \in [r-1]$, (1) implies that $|\mathrm{col}_r(B) \cap [k]| = 1$. By s-regularity of B, there must be $s-1$ rows $i > k$ with $r \in \mathrm{row}_i(B)$. Consequently there are $s-1$ rows of B containing columns $[r, r+k)$. These columns form a k-clique in G_{B^T}.

This shows that $\omega(G_A) \le \omega(G_{B^T})$. Since $AB^T = B^T A = J + I$ by Theorem 2.1, we get equality via symmetry. $\qquad\square$

Corollary 4.2. *The maximum clique size $\omega(G_A) \le \min(r-1, s-1)$.*

Proof. Lemma 3.14 implies $\omega(G_A) \le r-1$ and $\omega(G_{B^T}) \le s-1$. It follows that

$$\omega(G_A) = \omega(G_{B^T}) \le \min(r-1, s-1). \qquad\square$$

We will show that the previous upper bound is sharp. For $r = s$ and $n = r^2 - 1$, define the block matrix

$$\Omega_r = \begin{bmatrix} J_{r-1} & E_{11} & E_{22} & \cdots & E_{r-1,r-1} & 0 \\ 0 & J_{r-1} & E_{11} & & E_{r-2,r-2} & E_{r-1,r-1} \\ E_{r-1,r-1} & 0 & J_{r-1} & & E_{r-3,r-3} & E_{r-2,r-2} \\ \vdots & & & \ddots & & \vdots \\ E_{22} & E_{33} & E_{44} & & J_{r-1} & E_{11} \\ E_{11} & E_{22} & E_{33} & \cdots & 0 & J_{r-1} \end{bmatrix},$$

where each block is $(r-1) \times (r-1)$, J_{r-1} is the matrix of all 1s, and E_{ij} is the matrix with a single 1 in row i, column j. Observe that $G_{\Omega_r} = \bigsqcup^{r+1} K_{r-1}$, where K_{r-1} is the complete graph with $r-1$ vertices. For this reason, we will call Ω_r a *clique matrix*. In [8], it is shown that Ω_r is a thin Lehman matrix for $r \ge 3$. For completeness, we give a sketch of the proof.

Lemma 4.3 ([8], Lemma 4). *For $r \ge 3$, the clique matrix Ω_r is a thin Lehman matrix.*

Proof. Let $\pi \in S_{r+1}$ be the cycle permutation sending

$$\pi : 1 \mapsto 2, \ldots, r \mapsto r+1, r+1 \mapsto 1.$$

For $(i,j) \in [r+1] \times [r-1]$, define the $0,1$ vector y^{ij} to be the incidence vector of

$$\{(\pi^k(i)-1)(r-1)+k, (\pi^r(i)-1)(r-1)+j \mid k = 1, \ldots, r-1\}.$$

By letting B be the $n \times n$ matrix with row $(i-1)(r-1)+j$ equal to $y^{\pi^{j+1}(i),j}$, it can be checked that $\Omega_r \cdot B^T = J + I$. $\qquad\square$

Next, we consider the case when $r < s$.

Proposition 4.4. *For any $3 \le r < s$, there exists an r-regular Lehman matrix A such that $\omega(G_A) = r-1$, i.e., A contains J_{r-1} as a submatrix.*

Proof. Define the $n \times n$ matrix A_0 to be

$$A_0 = \left[\begin{array}{c|ccccccc}
J_{r-1} & E_{11} & E_{21} & E_{31} & \ldots & & E_{r-2,1} & E_{r-1,1} \\
\hline
E_{11} & J_{r-1} & E_{22} & E_{33} & \ldots & & E_{r-2,r-2} & 0 \\
E_{12} & 0 & J_{r-1} & E_{22} & & & E_{r-3,r-3} & E_{r-2,r-2} \\
E_{13} & E_{r-2,r-2} & 0 & & & & & \vdots \\
\vdots & \vdots & & & & & & E_{22} \\
E_{1,r-1} & E_{22} & \ldots & & E_{r-2,r-2} & & 0 & J_{r-1} \\
& & & & & & & & C
\end{array}\right],$$

where J_{r-1} is the $(r-1) \times (r-1)$ matrix of all 1s, E_{ij} is the $(r-1) \times (r-1)$ matrix with a single 1 in row i, column j, and C is the circulant matrix $C^r_{(s-r+1)r}$ with the first column and last row removed. The lines are placed to emphasize the matrix pattern.

Let $B_1(i,j) := (i-1)(r-1)+j$ and $B_2(i,j) := r(r-1)+r(i-1)+j$. Define the $n \times n$ matrix Σ with 1s only at

$$(B_1(2,r-1), B_2(s-r+1,r-1)), (B_1(r+1-i,r-1)), B_2(1,i)),$$
$$(B_2(1,1), B_1(r,r-1)), (B_2(s-r+1,r-i), B_1(i+1,r-1))$$

for $i \in [r-2]$. We claim that $A := A_0 + \Sigma$ is an r-regular thin Lehman matrix.

We provide the rows of B. For convenience, given a subset $S \subset [n]$, let $\mathrm{vec}(S)$ denote the incidence vector in $\{0,1\}^n$ of S. Let $e_i := \mathrm{vec}\{i\}$. Define

$$x_i = \mathrm{vec}\{B_1(2,1), \ldots, B_1(r,1), B_2(1,i), \ldots, B_2(s-r+1,i)\}$$

for $i \in [r-1]$. Let $\tau \in S_{r-1}$ be the cyclic permutation sending $1 \mapsto 2, \ldots, r-1 \mapsto 1$. For $(i,j) \in [r-1] \times [r-1]$, we define

$$y_{ij} = \mathrm{vec}\{B_1(1,i), B_1(1+\tau(i),2), \ldots, B_1(1+\tau^{r-3}(i), r-2),$$
$$B_1(1+\tau^{-1}(i),j), B_2(1,r-i), \ldots, B_2(s-r+1,r-i)\}.$$

Lastly for $i \in [r-1]$, define

$$z_i = \mathrm{vec}\{B_1(1,i), B_1(2,r-1), \ldots, B_1(r,r-1), B_2(1,r), \ldots, B_2(s-r,r)\}.$$

Now let the rows of B equal

$$x_i \qquad\qquad\qquad\qquad\qquad\qquad \text{for } i \in [r-1], \qquad\qquad\qquad \text{(R1)}$$

$$x_{r-1} + \sum_{j=k}^{s-r+1} \left(e_{B_2(j,r-2)} - e_{B_2(j,r-1)}\right) \qquad \text{for } k \in [2, s-r+1], \qquad\qquad \text{(R2)}$$

$$y_{ij} \qquad\qquad\qquad\qquad\qquad\qquad \text{for } (i,j) \in [r-1] \times [r-1], \qquad \text{(R3)}$$

$$y_{1,r-1} + \sum_{j=1}^{k} \left(e_{B_2(j,r)} - e_{B_2(j,r-1)}\right) \qquad \text{for } k \in [s-r], \qquad\qquad\qquad \text{(R4)}$$

$$y_{i,r-1} + \sum_{j=k}^{s-r+1} \left(e_{B_2(j,r-i-1)} - e_{B_2(j,r-i)}\right) \quad \text{for } (i,k) \in [2, r-1] \times [2, s-r+1], \text{ and} \quad \text{(R5)}$$

$$z_i \qquad\qquad\qquad\qquad\qquad\qquad \text{for } i \in [r-1]. \qquad\qquad\qquad\qquad \text{(R6)}$$

There are

$$(r-1) + (s-r) + (r-1)(r-1) + (s-r) + (r-2)(s-r) + (r-1) = rs - 1 = n$$

rows in total, so B is an $n \times n$ matrix. To avoid excessive technical details, we leave it to the reader to check that $AB^T = J + P$ for some permutation matrix P. Thus, $A(PB)^T = J + PP^T = J + I$. We conclude that A is a thin Lehman matrix. $\qquad\square$

Example. We provide an example of the matrices A and B described in Proposition 4.4 such that $AB^T = J + I$. In the example $r = 4, s = 5$, and $n = 19$.
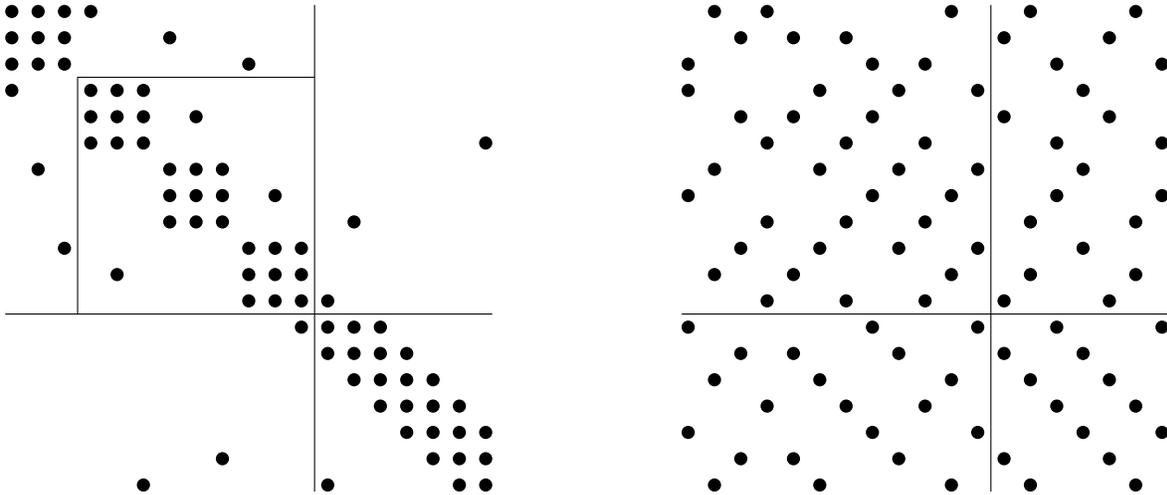
Proposition 4.4 shows that Corollary 4.2 is sharp for $r < s$. Since $\omega(G_A) = \omega(G_B^T)$ by Lemma 4.1, taking B^T from Proposition 4.4 shows the bound is sharp for $r > s$. The clique matrix Ω_r shows sharpness for $r = s$, so we conclude that $\omega(G_A) \leq \min(r-1, s-1)$ is sharp in general.

4.2 Maximum degree

We provide a relation between the maximum degree and the maximum clique size of the induced Johnson subgraph. Using bounds on maximum clique size, we give a bound on the maximum degree of G_A. We do not, however, believe this bound is sharp.

Lemma 4.5. *The maximum degree* $\Delta(G_A) \leq 2(\omega(G_A) - 1)$.

Proof. If $A \simeq C_n^r$, then $\Delta(G_A) = 2$ and $\omega(G_A) = 2$, so the lemma holds. If $A \not\simeq C_n^r$, Theorem 3.12 says that each connected component of G_A is a clique tree. Take a vertex i of G_A. By the definition of a clique tree, i is incident to at most two maximal cliques. The degree of any vertex in a maximal clique of G_A is $\omega(G_A) - 1$. Therefore, the degree of i is at most $2(\omega(G_A) - 1)$. $\qquad\square$

Observe that in the case $r = 3$, Corollary 4.2 says $\omega(G_A) \leq 2$, so the previous lemma implies that $\Delta(G_A) \leq 2$.

Proposition 4.6. *For $r > 3$, the maximum degree* $\Delta(G_A) \leq \min(r - 2, 2(s - 2))$.

Proof. Assume $\mathrm{row}_1(A) = [r]$, and suppose that the $r - 1$ rows $[2, r]$ are all adjacent to row 1. We may also assume by permuting columns that $\{1, 2\} \subset \mathrm{row}_1(B)$, which implies that $[3, r] \subset \mathrm{row}_i(A)$ for $i \in [r]$ by adjacency. Thus,

$$\mathrm{col}_3(A) = \mathrm{col}_4(A) = [r],$$

which contradicts A being invertible. Therefore, $\Delta(G_A) \leq r - 2$. It follows from Corollary 4.2 and Lemma 4.5 that $\Delta(G_A) \leq 2(s - 2)$. $\qquad\square$

5 Classification of graphs with two connected components

In Section 3, we showed that a Lehman matrix A has a connected graph G_A if and only if $A \simeq C_n^r$. We now classify all Lehman matrices such that the graph G_A has exactly two connected components. Recall that a Lehman matrix A with $d > 1$ has no edges in G_A. Therefore, we may assume A is an r-regular thin Lehman matrix, B is the s-regular dual, and $n = rs - 1$.

Theorem 5.1. *For any $r, s > 2$, a graph G with n vertices and two connected components is the Johnson subgraph G_A associated to an r-regular Lehman matrix A if and only if each component of G is a path with length greater than r and not equivalent to $-1, 0$ (mod r). Furthermore the matrix A is determined, up to isomorphism, by the graph G.*

Proof. ($\implies$) Suppose a component of G_A has length greater than $n - r > r$. Then it must be either a path or a cycle by Lemma 3.14. Lemmas 3.2 and 3.4 imply that $A \simeq C_n^r$. Since G_A has two connected components, we must have that the number of vertices in each component is at most $n - r$. Equivalently, each component has at least r vertices and must be a path by Lemma 3.14.

Claim 1. *Neither path has length equivalent to* $-1, 0 \pmod{r}$.

Let the lengths of the two paths be k and $n - k$. Since $k + (n - k) = n \equiv -1 \pmod{r}$, we can assume the length of one of the paths is

$$k = r\ell + j$$

for $\ell \geq 1$ and $0 \leq j < r/2$. By Lemmas 3.2 and 3.3, we rearrange rows and columns to get $\text{row}_i(A) = [i, i + r)$ and

$$\text{row}_i(B) \cap [k + r - 1] = \{i - r\mathbb{Z}_{\geq 0}, i + r - 1 + r\mathbb{Z}_{\geq 0}\} \cap [k + r - 1] \tag{10}$$

for $i \in [k]$. Lemma 3.2 also implies that there exists a permutation $\sigma \in S_n$ such that

$$\text{row}_{k+i}(A) = \sigma([i, i + r)) \quad \text{for} \quad i \in [n - k]. \tag{11}$$

Additionally since $n - k > r$, we may assume

$$\{\sigma(i), \sigma(i + r - 1)\} = \text{row}_{k+i}(B) \cap \sigma([i, i + r)) \tag{12}$$

for $i \in [n - k]$ from the proof of Lemma 3.3.

Observe that

$$|\text{col}_i(A) \cap [k]| = \min(i, k + r - i, r) \qquad \text{for } i \in [k + r - 1], \text{ and} \tag{13}$$
$$|\text{col}_{\sigma(i)}(A) \cap (k, n]| = \min(i, n - k + r - i, r) \qquad \text{for } i \in [n - k + r - 1]. \tag{14}$$

Since A is r-regular, we must have

$$|\text{col}_{\sigma(i)}(A) \cap [k]| + |\text{col}_{\sigma(i)}(A) \cap (k, n]| = \min(\sigma(i), k + r - \sigma(i), r) + \min(i, n - k + r - i, r) = r.$$

From this observation we deduce that

$$\sigma(i) \in \{r - i, k + i\} \tag{15}$$

for $i \in [r - 1]$. If $\sigma(1) = r - 1$, we can swap rows $i, k - i$ for $i \in [k]$ and columns $i, k + r - i$ for $i \in [k + r - 1]$ in A and B simultaneously to assume that $\sigma(1) = k + 1$ and $k + 1 \in \text{row}_{k+1}(A) \cap \text{row}_{k+1}(B)$. Due to the cascading structure of rows $[k]$ of A, we have that

$$\{k + 1 - r\mathbb{Z}_{\geq 0}\} \cap [k] \subset \text{row}_{k+1}(B),$$

so $j + 1 \in \text{row}_{k+1}(B)$. Since $\{j + 1, k + 1\} \subset \text{col}_{j+1}(B)$, Remark 2.2 says that

$$|\{j + 1, k + 1\} \cap \text{col}_i(A)| \leq 1$$

for columns $i > j + 1$. As $\text{row}_{j+1}(A) = [j + 1, j + r]$, we deduce that

$$[j + 2, j + r] \cap \text{row}_{k+1}(A) = \emptyset.$$

Noting that $\text{row}_{k+1}(A) = \sigma([1, r])$, if $\sigma(r - i) = i$ for any $i \in [j + 2, r)$, then

$$\sigma([1, r]) \cap [j + 2, r) \neq \emptyset,$$

which gives a contradiction. Therefore, (15) implies that $\sigma(r-i) = k+r-i$ for $i \in [j+2, r)$. Thus,

$$[k + 1, k + r - j - 2] \subset \text{row}_{k+1}(A).$$

Note that $\sigma(r - j - 1) \notin \text{row}_{k+1}(B)$ from (12). Therefore, $\sigma(r - j - 1) \neq j + 1$, which forces $\sigma(r-j-1) = k+r-j-1$. If $j = 0$, we have shown that $[k+1, k+r) \subset \text{row}_{k+1}(A)$, which contradicts rows k and $k + 1$ not being adjacent in G_A. Therefore, $k \not\equiv 0 \pmod{r}$ and since $n = rs - 1$, $n - k \not\equiv -1 \pmod{r}$. Switching k and $n - k$, we get $n - k \not\equiv 0 \pmod{r}$ and $k \not\equiv -1 \pmod{r}$. This proves the claim.

Claim 2. *If there exists a Lehman matrix A with $G_A = P_k \sqcup P_{n-k}$, then A is unique up to isomorphism.*

We continue the discussion from the proof of Claim 1. We previously concluded that

$$[k + 1, k + r - j) \subset \text{row}_{k+1}(A).$$

Suppose $k + r - j \in \text{row}_{k+1}(A)$. Then $\sigma(r - j) = k + r - j$ by (11) and (15), so

$$k + r - j \in \text{col}_{k+r-j}(B)$$

by (12). Since $k + r - j = r(\ell + 1)$, we also have $k + r - j \in \text{row}_1(B)$ by (10). Thus,

$$\text{col}_{k+r-j}(B) \supset \{1, k + r - j\}.$$

If $i \in \text{row}_{k+1}(A)$ for $i \in [j]$, then $\sigma(r - i) = i$ and

$$\text{col}_i(A) = [i] \cup [k + 1, k + r - i],$$

which contains $\{1, k+r - j\}$, giving a contradiction. Therefore, $\text{row}_{k+1}(A) \cap [j] = \emptyset$, and $\sigma(r - i) = k + r - i$ for $i \in [j]$. This implies

$$[k + 1, k + r) \subset \text{row}_{k+1}(A)$$

which is another contradiction because rows k and $k+1$ are not adjacent. Hence $k+r-j \notin \text{row}_{k+1}(A)$, so $\sigma(r - j) = j$.

Observe that $j \in \text{row}_k(B)$ by (10) since $k = r\ell + j$. Thus, (12) implies that

$$\text{col}_j(B) \supset \{k, k + r - j\}.$$

If $k + r - i \in \text{row}_{k+1}(A)$ for $i \in [j]$, then $\sigma(r - i) = k + r - i$ and

$$\text{col}_{k+r-i}(A) = (k - i, k + r - i],$$

which contains $\{k, k+r-j\}$, giving a contradiction. We conclude that

$$[k+r-j, k+r) \cap \mathrm{row}_{k+1}(A) = \emptyset.$$

Now by (15), we conclude that

$$[j] \cup [k+1, k+r-j) \subset \mathrm{row}_{k+1}(A) \text{ and}$$

$$\sigma(i) = \begin{cases} k+i & \text{if } i \in [r-j-1] \\ r-i & \text{if } i \in [r-j, r) \end{cases}.$$

We further deduce from (13) and (14) that

$$\sigma(n-k+i) = \begin{cases} k+r-i & \text{if } i \in [j] \\ i & \text{if } i \in [j+1, r) \end{cases}.$$

We have thus determined the columns $[k+r-1]$ in A. This determines A up to a choice of $\sigma([r, n-k])$. Since all 1s in rows $[k]$ of A are contained in columns $[k+r-1]$, we can permute the rest of the columns without affecting the first k rows. Therefore, we have shown that if a thin Lehman matrix A exists with $G_A = P_k \sqcup P_{n-k}$, then A is unique up to isomorphism (P_k is a path of length k).

($\impliedby$) Let $r < k < n-r$ and $k = r\ell + j$ for $0 < j < r/2$. We will construct an $n \times n$ r-regular thin Lehman matrix A such that $G_A = P_k \sqcup P_{n-k}$. First we divide C_n^r and D_n^s into blocks. Write

$$C_n^r = \begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad D_n^s = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

where A_{11} and B_{11} are $k \times (k+r-1)$ matrices, and A_{22} and B_{22} are $(n-k) \times (n-k-r+1)$ matrices. Let P be the $(k+r-1) \times (k+r-1)$ permutation matrix such that $A_{21}P$ switches columns

$$i \leftrightarrow k+r-i \quad \text{for} \quad i \notin (rt+j, r(t+1)), t \in [0, \ell]$$

and keeps all other columns fixed. Let Q be the $(n-k-r+1) \times (n-k-r+1)$ permutation matrix such that $B_{12}Q$ switches columns

$$rt-j+i \leftrightarrow rt-i+1 \quad \text{for} \quad i \in [j], t \in [1, s-\ell-2]$$

and keeps all other columns fixed.

Note that if $\ell = s-2$, then $n-k-r+1 = r(s-\ell-1)-j = r-j$, so Q is the identity matrix. If $\ell = s-3$, then $(r-j, r] = (r(s-\ell-2)-j, r(s-\ell-2)]$, so Q only switches one block of columns.

Define A and B by

$$A = \begin{bmatrix} A_{11} & 0 \\ A_{21}P & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12}Q \\ B_{21}P & B_{22} \end{bmatrix}.$$

Observe that A is r-regular with rows $[k]$ and $(k, n]$ forming two disjoint paths in G_A. To avoid excessive technical details, we leave it to the reader to confirm that $AB^T = J + I$, i.e., A and B form a thin Lehman pair. This proves the theorem. $\square$

Example. We give an example of A from Theorem 5.1 with $G_A = P_k \sqcup P_{n-k}$. We also provide the dual B. In the example $r = 5$, $s = 5$, $n = 24$, and $k = 7$.
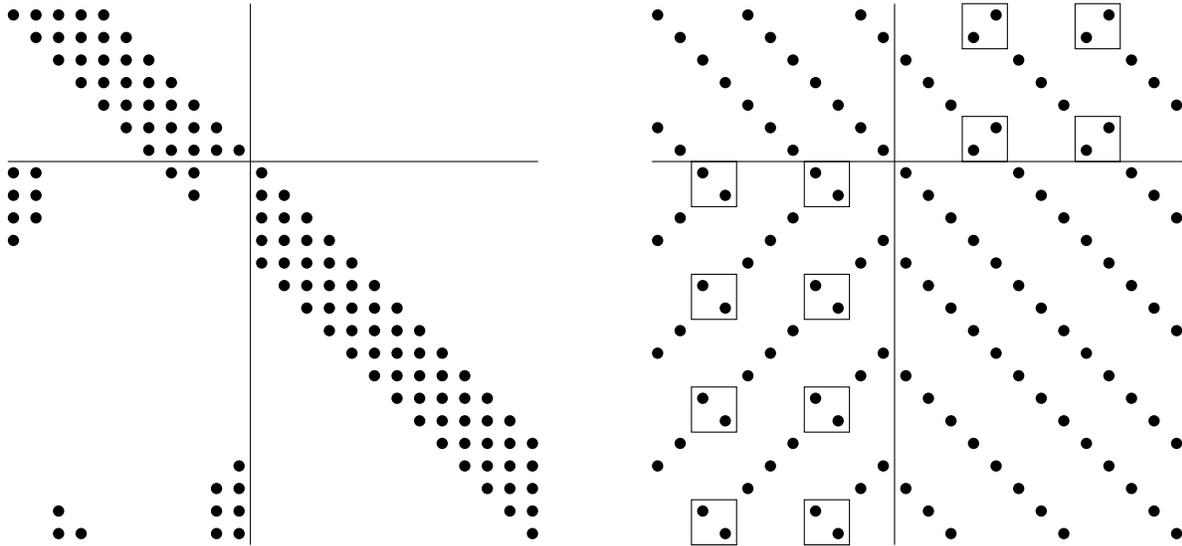
6   Circulant matrices

Let $S \subset \mathbb{Z}/n\mathbb{Z}$. We define the circulant matrix C_S to be the $n \times n$ matrix with columns indexed by $\mathbb{Z}/n\mathbb{Z}$ and rows equal to the incidence vectors of

$$i + S := \{i + s \mid s \in S\}$$

for $i \in \mathbb{Z}/n\mathbb{Z}$. By definition, any circulant matrix can be written in this form. We study when a circulant matrix is also a Lehman matrix. We show that if C_S is a Lehman matrix, then either $C_S \simeq C_n^r$ or the Johnson subgraph G_{C_S} has no edges. We then provide some general constructions of circulant Lehman matrices C_S such that the Johnson subgraph indeed has no edges.

Note that if C_S is a Lehman matrix, then its dual must also be circulant, since a translation of $\mathbb{Z}/n\mathbb{Z}$ is a bijection and the dual is unique. If C_S is Lehman for $d > 1$, then G_{C_S} cannot have any edges. We therefore again only consider thin Lehman matrices.

Proposition 6.1. *Let $r = |S|$. If C_S is a thin Lehman matrix, then either $C_S \simeq C_n^r$ or G_{C_S} has no edges.*

Proof. Note that since C_S is circulant, the graph G_{C_S} is vertex transitive (translating the rows in $\mathbb{Z}/n\mathbb{Z}$ is a graph automorphism). Suppose G_{C_S} contains an edge. Then there must exist $t < n$ such that

$$|S \cap (t + S)| = r - 1.$$

Let $\{b\} = S \setminus (t + S)$. There must exist $k \geq 0$ such that

$$b, b + t, \ldots, b + kt \in S \quad \text{but} \quad b + (k+1)t \notin S,$$

since otherwise $b + nt = b$ contradicts $b \notin t + S$. Therefore, $\{b + (k+1)t\} = (t + S) \setminus S$. Define

$$X = S \setminus \{b, b + t, \ldots, b + kt\}.$$

We claim that $t + X = X$. Take $a \in X$ and suppose $a + t \notin X$. Then either $a + t \notin S$ or $a + t \in \{b, b + t, \ldots, b + kt\}$. If $a + t \notin S$, then $a + t \in (t + S) \setminus S$, so $a + t = b + (k+1)t$. Hence $a = b + kt \notin X$, a contradiction. Now suppose

$$a + t \in \{b, b + t, \ldots, b + kt\}.$$

We cannot have $a + t = b$ since $b \notin t + S$. Therefore, $a + t = b + \ell t$ for $\ell \geq 1$, so $a = b + (\ell - 1)t \notin X$, which is another contradiction. We conclude that $t + X = X$.

Suppose k is positive. Since $X - t = X + nt - t = X + (n-1)t$, we have $X - t = X$. Therefore,

$$S - t = \{b - t, b, \ldots, b + (k-1)t\} \cup (X - t) = \{b - t, b, \ldots, b + (k-1)t\} \cup X.$$

By assumption $b + (k+1)t \notin S$, so $b + (k+1)t \neq b$. Then $b - t \neq b + kt$, and

$$|S \cap (S - t)| = r - 1.$$

Observe that

$$t + S = \{b + t, b + 2t, \ldots, b + kt, b + (k+1)t\} \cup X.$$

Since $b \neq b + (k+1)t$, we also have $|S \cap (t + S)| = r - 1$. Note that

$$S \cap (S - t) = \{b, \ldots, b + (k-1)t\} \cup X \quad \text{and} \quad S \cap (t + S) = \{b + t, \ldots, b + kt\} \cup X.$$

Therefore, rows $-t, 0, t$ of C_S cannot form a clique in G_{C_S} because $S \cap (S - t) \neq S \cap (t + S)$, by (6). If C_S is not isomorphic to C_n^r, then G_{C_S} is a union of clique trees. As G_{C_S} is vertex transitive and thus regular, the only possible clique trees are single cliques. Since rows $-t$ and 0 and rows 0 and t are adjacent, but rows $-t$ and t are not, this is a contradiction. Therefore, we must have $C_S \simeq C_n^r$.

Now suppose $k = 0$. Then $S = \{b\} \cup X$ and row t is the incidence vector of

$$t + S = \{b + t\} \cup X.$$

Let B be the Lehman dual of C_S. Take some row $i \in \mathrm{col}_b(B)$ distinct from 0 or t, so $b \in \mathrm{row}_i(B)$. Since $|S \cap \mathrm{row}_i(B)| = 1$, $|(t + S) \cap \mathrm{row}_i(B)| = 1$, and $S \cap (t + S) = X$, we must have

$$b + t \in \mathrm{row}_i(B).$$

Observe that $|X| = r - 1 \geq 2$, and pick distinct $a_1, a_2 \in X$. For $j \in \{1, 2\}$, we have $\{a_j, a_j + t\} \subset X$, so

$$(b - a_j) + S \supset \{b, b + t\}.$$

Thus, the distinct rows $b - a_1$ and $b - a_2$ of C_S both intersect row i of B in at least 2 columns, which is a contradiction. In this case, C_S cannot be a thin Lehman matrix. $\square$

The previous lemma motivates the question of when a circulant Lehman matrix can have a Johnson subgraph with no edges. We show that for composite r and s, there do exist circulant matrices C_S that are thin Lehman matrices with edgeless graphs. To simplify our expressions, we use

$$\mathrm{AP}(a, k, \delta) := \{a, a + \delta, \ldots, a + (k-1)\delta\} \subset \mathbb{Z}/n\mathbb{Z}$$

to denote arithmetic progressions of length k and difference δ starting with element a.

Given two subsets $X, Y \subset \mathbb{Z}/n\mathbb{Z}$, let

$$X + Y := \{x + y \mid x \in X, y \in Y\}.$$

Proposition 6.2. *Suppose $r = r_1 r_2$ and $s = s_1 s_2$ are composite integers, with $r_i, s_i \geq 2$. Define*

$$S_A = \mathrm{AP}(0, r_1, 1) + \mathrm{AP}(0, r_2, r_1 s_1) \ and$$
$$-S_B = \mathrm{AP}(0, s_1, r_1) + \mathrm{AP}(0, s_2, r s_1).$$

Then C_{S_A}, C_{S_B} form a thin Lehman pair. Moreover $G_{C_{S_A}}, G_{C_{S_B}}$ are edgeless graphs.

Proof. Let P be the circulant matrix $C_{\{1\}}$ of order n, where $\{1\} \subset \mathbb{Z}/n\mathbb{Z}$. Note that P is a permutation matrix. Then we may write

$$C_{S_A} = \sum_{i \in S_A} P^i \quad \text{and} \quad C_{S_B} = \sum_{i \in S_B} P^i.$$

Since $P^T = P^{-1}$, we have $(C_{S_B})^T = C_{-S_B}$. We can thus express

$$C_{S_A}(C_{S_B})^T = \sum_{(i,j) \in S_A \times S_B} P^{i-j}.$$

Since $r_1 < r_1 s_1$ and $r_1 s_1 < r s_1$, we see that $|S_A| = r_1 r_2 = r$ and $|S_B| = s_1 s_2 = s$. Observe that $0 \in S_A \cap S_B$ and

$$(r_1 - 1) + (r_2 - 1)r_1 s_1 = n - (s_1 - 1)r_1 - (s_2 - 1)r s_1 \in S_A \cap S_B.$$

Thus, there are two pairs $(i, j) \in S_A \times S_B$ with $i - j = 0$. It is easy to check that

$$S_A - S_B = \mathbb{Z}/n\mathbb{Z}.$$

Since $n = rs - 1$, we conclude that

$$C_{S_A}(C_{S_B})^T = P^0 + \sum_{i \in \mathbb{Z}/n\mathbb{Z}} P^i = J + I.$$

Since $r_2 \geq 2$ and $r_1 + r_2(r_1 s_1) < n$, we can deduce that $G_{C_{S_A}}$ has no edges. Then $\omega(G_{C_{S_A}}) = 0$ implies $\omega(G_{C_{S_B}^T}) = 0$ by Lemma 4.1. Since $C_{S_B}^T = C_{-S_B}$, this implies $G_{C_{S_B}}$ also has no edges. □

Note that some of the matrices C_{S_A} given above may be isomorphic for different choices of r_1, r_2, s_1, s_2. We believe, however, that the previously mentioned matrices are the only possible circulant thin Lehman matrices, up to isomorphism.

Conjecture 6.3. *If C_S is a thin Lehman matrix, then C_S is isomorphic to C_n^r or one of the matrices in Proposition 6.2.*

We used a computer program to confirm that the conjecture is true when $n < 48$.

7 Open problems

The classification of Lehman matrices, or even only thin Lehman matrices, is still an open problem. We have used the Johnson subgraph to give structural results for thin Lehman matrices, which we believe will make it easier to enumerate them. We showed that a connected component of the graph uniquely determines the corresponding rows in the Lehman matrix, and we completely classified matrices where their graphs have two connected components. Lehman matrices with graphs containing more connected components have not been classified, however. Some of the constraints we provide for the Johnson subgraph may also be improved upon.

If two matrices are isomorphic, then their Johnson subgraphs are also isomorphic. We noted that the converse is false. It would be highly useful to find some simple structure associated to a Lehman matrix such that the matrix is uniquely determined up to isomorphism. One possibility may be to use colored edges to combine the graphs G_A and G_B into a single graph. Another possibility is to investigate when row intersections have a fixed size other than $r - 1$.

We showed that a circulant Lehman matrix must either be isomorphic to C_n^r or have an edgeless Johnson graph. We propose a conjecture on the classification of all circulant thin Lehman matrices. In particular, it would be interesting to prove that if r or s is prime, then any $n \times n$ circulant thin Lehman matrix is isomorphic to C_n^r, for $n = rs - 1$.

The only known infinite families of Lehman matrices are the thin Lehman matrices and the point-line incidence matrices of nondegenerate finite projective planes. We would like to know if there are any other infinite families of Lehman matrices.

8 Acknowledgments

References

[1] W. G. Bridges and H. J. Ryser. Combinatorial designs and related systems. *J. Algebra*, 13:432–446, 1969.

[2] G. Cornuéjols. *Combinatorial optimization*, volume 74 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

[3] G. Cornuéjols, B. Guenin, and L. Tunçel. Lehman matrices. *J. Combin. Theory Ser. B*, 99(3):531–556, 2009.

[4] D. R. Hughes and F. C. Piper. *Projective planes*. Springer-Verlag, New York, 1973. Graduate Texts in Mathematics, Vol. 6.

[5] A. Lehman. On the width-length inequality. *Math. Programming*, 16(2):245–259, 1979.

[6] A. Lehman. The width-length inequality and degenerate projective planes. In *Polyhedral combinatorics (Morristown, NJ, 1989)*, volume 1 of *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, pages 101–105. Amer. Math. Soc., Providence, RI, 1990.

[7] C. Lütolf and F. Margot. A catalog of minimally nonideal matrices. *Math. Methods Oper. Res.*, 47(2):221–241, 1998.

[8] J. Wang. A new infinite family of minimally nonideal matrices. *J. Combin. Theory Ser. A*, 118(2):365-372, 2011.