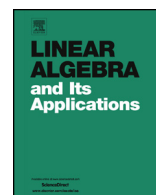




Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa



Symmetric nonnegative trifactorization of pattern matrices

Damjana Kokol Bukovšek^{a,b,*}, Helena Šmigoc^c

^a University of Ljubljana, School of Economics and Business, Slovenia

^b Institute of Mathematics, Physics and Mechanics, Ljubljana, Slovenia

^c School of Mathematics and Statistics, University College Dublin, Ireland

ARTICLE INFO

Article history:

Received 31 October 2023

Received in revised form 13 March 2024

Accepted 23 May 2024

Available online 28 May 2024

Submitted by E.C. Carson

MSC:

15A23

15B48

Keywords:

Nonnegative matrix factorization

Nonnegative symmetric matrices

Symmetric nonnegative

trifactorization

Pattern matrices

ABSTRACT

A factorization of an $n \times n$ nonnegative symmetric matrix A of the form BCB^T , where C is a $k \times k$ symmetric matrix, and both B and C are required to be nonnegative, is called the Symmetric Nonnegative Matrix Trifactorization (SN-Trifactorization). The SNT-rank of A is the minimal k for which such factorization exists. The SNT-rank of a simple graph G that allows loops is defined to be the minimal possible SNT-rank of all symmetric nonnegative matrices whose zero-nonzero pattern is prescribed by the graph G .

We define set-join covers of graphs, and show that finding the SNT-rank of G is equivalent to finding the minimal order of a set-join cover of G . Using this insight we develop basic properties of the SNT-rank for graphs and compute it for trees and cycles without loops. We show the equivalence between the SNT-rank for complete graphs and the Katona problem, and discuss uniqueness of patterns of matrices in the factorization.

© 2024 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

* Corresponding author.

E-mail addresses: damjana.kokol.bukovsek@ef.uni-lj.si (D. Kokol Bukovšek), helena.smigoc@ucd.ie (H. Šmigoc).

1. Introduction and notation

Factorizations of matrices where the factors are required to be entry-wise nonnegative provide a powerful tool in analysing nonnegative data. In this paper we consider a factorization of nonnegative symmetric matrices, which takes into account symmetry, nonnegativity and low rank of a matrix.

Denote by \mathbb{R}_+ the set of nonnegative real numbers, by $\mathbb{R}^{n \times m}$ the set of $n \times m$ real matrices, and by $\mathbb{R}_+^{n \times m}$ the set of $n \times m$ entry-wise nonnegative matrices. Furthermore, we denote

$$\mathcal{S}_n^+ = \{A \in \mathbb{R}_+^{n \times n}; A = A^T\}.$$

Definition 1.1. A factorization of $A \in \mathcal{S}_n^+$ of the form BCB^T , where $B \in \mathbb{R}_+^{n \times k}$ and $C \in \mathcal{S}_k^+$, is called *Symmetric Nonnegative Trifactorization* of A (*SN-Trifactorization* for short). Minimal possible k in such factorization is called the *SNT-rank* of A , and is denoted by $\text{st}_+(A)$.

The SN-Trifactorization was studied in [7], and is closely related to two other well known factorizations that feature nonnegative factors: Nonnegative Matrix Factorization and Completely Positive Factorization. We refer the reader to [4] for the background on the Nonnegative Matrix Factorization and to [1] for the background on the Completely Positive Factorization.

The zero-nonzero pattern of a nonnegative matrix $A \in \mathcal{S}_n^+$ poses constraints on the zero-nonzero pattern of nonnegative matrices $B \in \mathbb{R}_+^{n \times k}$ and $C \in \mathcal{S}_k^+$ satisfying $A = BCB^T$. The aim of this paper is to better understand those restrictions. To this end we define the SNT-rank of a graph G as the minimal SNT-rank that can be achieved by the matrices with the pattern prescribed by G (see Definition 2.3).

In the context of matrix patterns, graphs arise naturally. In this work, a graph $G = (V(G), E(G))$ will always be a simple undirected graph that allows loops. Hence, $E(G) \subseteq \{\{i, j\}; i, j \in V(G)\}$, where $\{i\} \in E(G)$ corresponds to a loop on the vertex i in G . A vertex $i \in V(G)$ is *isolated*, if $\{i, j\} \notin E(G)$ for all $j \in V(G)$. In particular, the vertex i does not have a loop. The vertex set $V(G)$ will often be $[n] := \{1, 2, \dots, n\}$.

Let G and H be two simple graphs that allow loops. We denote by $G \cup H$ their disjoint union, by tG the union of t copies of G and by $G \vee H$ the join of G and H , i.e. the graph with $V(G \vee H) = V(G) \cup V(H)$ and $E(G \vee H) = E(G) \cup E(H) \cup \{\{i, j\} \mid i \in V(G), j \in V(H)\}$. We use standard notation for graphs without any loops: we denote by K_n the complete graph on n vertices, by $K_{m,n}$ the bipartite graph on $m + n$ vertices, by P_n the path on n vertices, and by C_n the cycle on n vertices. We denote by K_n^ℓ the complete graph on n vertices with all possible loops.

We denote matrices by capital letters, $A \in \mathbb{R}^{n \times m}$, vectors by bold letters, $\mathbf{a} \in \mathbb{R}^n$, the zero matrix in $\mathbb{R}^{n \times m}$ by $0_{n \times m}$ and the zero vector in \mathbb{R}^n by $\mathbf{0}_n$.

Let $A, B \in \mathbb{R}^{n \times m}$. Then $A > B$ means that $A - B$ is entry-wise positive, and $A \geq B$ denotes entry-wise nonnegativity of $A - B$. The *support* of a vector $\mathbf{a} \in \mathbb{R}^n$ is the set

of all indices in $i \in [n]$ for which $a_i \neq 0$. For $A \in \mathbb{R}^{n \times m}$ and $\mathcal{S} \subset [n], \mathcal{T} \subset [m]$, we denote by $A[\mathcal{S}, \mathcal{T}]$ the submatrix of A containing entries a_{ij} for all $i \in \mathcal{S}, j \in \mathcal{T}$. For square matrices A and B we denote by $A \oplus B$ the block diagonal matrix with A and B as diagonal blocks.

After a brief introduction and development of notation in Section 1, we dedicate Section 2 to the introduction of the two main concepts of this paper: the SNT-rank and set-join covers of simple graphs that allow loops. We prove that finding the SNT-rank of a graph G is equivalent to determining minimal order of the set-join cover of G . Using this insight we develop basic properties of SNT-rank of graphs. We conclude Section 2 by defining uniqueness of optimal set-join covers of graphs. In Section 3 we find the SNT-rank of trees without loops and consider the unicyclic graphs without loops. Section 4 is dedicated to the SNT-rank and optimal set-join covers of complete graphs without loops. Section 5 offers a selection of possible applications and further research directions motivated by the work in this paper.

2. Set-join covers and $\text{st}_+(G)$

In this work we will for the most part take aside the actual values of matrices and only consider their patterns. Our main focus will be the question, how small can $\text{st}_+(A)$ be among all matrices A with a given zero-nonzero pattern.

2.1. Setup

We begin by defining $\text{st}_+(G)$ for a graph G , and developing a combinatorial question that is equivalent to finding $\text{st}_+(G)$.

Definition 2.1. Let $A \in \mathbb{R}_+^{n \times m}$. The *pattern matrix* of A is the matrix $\text{sign}(A) \in \{0, 1\}^{n \times m}$ defined by

$$\text{sign}(A)_{ij} = \begin{cases} 1; & \text{if } a_{ij} > 0, \\ 0; & \text{if } a_{ij} = 0. \end{cases}$$

(The pattern matrix is in some literature called *the derangement matrix*.)

Definition 2.2. The *pattern graph* $G(A) = (V(G), E(G))$ of a matrix $A \in \mathcal{S}_n^+$ is defined by $V(G) = [n]$, and $\{i, j\} \in E(G)$ precisely when $a_{ij} > 0$.

Let G be a graph with $|V(G)| = n$. By $\mathcal{S}^+(G)$ we denote the set of all matrices in \mathcal{S}_n^+ with the pattern graph G , and ask the question how small can the SNT-rank be on this set.

Definition 2.3. Let $G = (V(G), E(G))$ be a simple graph with loops. We define: $\text{st}_+(G) := \min\{\text{st}_+(A); A \in \mathcal{S}^+(G)\}$.

Remark 2.4. A factorization of $A \in \mathbb{R}_+^{n \times m}$ of the form UV^T , where $U \in \mathbb{R}_+^{n \times k}$ and $V \in \mathbb{R}_+^{m \times k}$, is called Nonnegative Matrix Factorization of A . Minimal possible k in such factorization is called the NMF-rank of A , and is denoted by $\text{rk}_+(A)$ (see [4]). In this context, the Boolean rank of a matrix is studied. It is defined as

$$\text{rk}_{01}(A) = \min\{\text{rk}_+(B); B \in \mathbb{R}_+^{n \times m}, \text{sign}(B) = \text{sign}(A)\}.$$

It is easy to see that Boolean rank is equal to rectangle covering bound, i.e. the minimum number of rectangles needed to cover all nonzero entries in A (see [4]).

The next result connects the patterns of nonnegative matrices A , B and C , provided $A = BCB^T$.

Proposition 2.5. Let $A = BCB^T$ be SN-Trifactorization of $A = (a_{ij}) \in \mathcal{S}_n^+$. For $i \in [n]$ let $r_i(B)^T \in \mathbb{R}^{1 \times k}$ be the i -th row of B , and $\mathcal{R}_i \subseteq \{1, \dots, k\}$ the support of $r_i(B)$, i.e. $\mathcal{R}_i := \{j; (r_i(B))_j \neq 0\}$. Then $C[\mathcal{R}_i, \mathcal{R}_j] = 0$ if and only if $a_{ij} = 0$.

Proof. Since $a_{ij} = r_i(B)^T C r_j(B) = (r_i(B)[\mathcal{R}_i])^T C[\mathcal{R}_i, \mathcal{R}_j] (r_j(B)[\mathcal{R}_j])$, and $r_s(B)[\mathcal{R}_s] > 0$ by the definition of \mathcal{R}_s , the conclusion follows. \square

Note that $\text{sign}(B)$ in Proposition 2.5 is the incidence matrix of $\mathbf{R} = (\mathcal{R}_1, \dots, \mathcal{R}_n)$. We recall the definition of the incidence matrix below.

Definition 2.6. Let $\mathbf{R} = (\mathcal{R}_1, \dots, \mathcal{R}_n)$ be a list of sets $\mathcal{R}_i \subseteq \{1, \dots, k\}$. The incidence matrix of \mathbf{R} is an $n \times k$ matrix $\iota(\mathbf{R})$ with

$$\iota(\mathbf{R})_{ij} = \begin{cases} 1; & j \in \mathcal{R}_i, \\ 0; & j \notin \mathcal{R}_i. \end{cases}$$

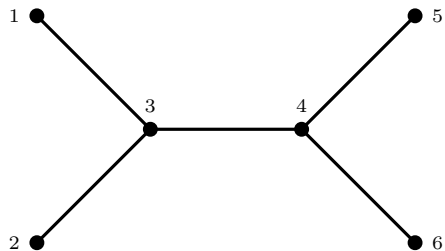
Since $\text{st}_+(G)$ is a parameter of the graph G , we would like to determine it directly from G . For this we need to introduce the set-join and the set-join cover.

Definition 2.7. Let \mathcal{S} be a finite set and \mathcal{K}, \mathcal{L} two nonempty subsets of \mathcal{S} with possibly nonempty intersection. We define the set-join of \mathcal{K} and \mathcal{L} on \mathcal{S} , denoted by $\mathcal{K} \vee_{\mathcal{S}} \mathcal{L}$, to be the graph with $V(\mathcal{K} \vee_{\mathcal{S}} \mathcal{L}) = \mathcal{S}$, and $E(\mathcal{K} \vee_{\mathcal{S}} \mathcal{L}) = \{\{i, j\}; i \in \mathcal{K}, j \in \mathcal{L}\}$.

Note that $\mathcal{K} \vee_{\mathcal{S}} \mathcal{L}$ has a loop on vertex i precisely when $i \in \mathcal{K} \cap \mathcal{L}$, and that if $\mathcal{K} \cup \mathcal{L} \neq \mathcal{S}$, then $\mathcal{K} \vee_{\mathcal{S}} \mathcal{L}$ contains some isolated vertices without loops.

Definition 2.8. Let G be a graph and $\mathcal{K}_i, \mathcal{L}_i \subseteq V(G)$. We say that

$$\mathcal{C} = \{\mathcal{K}_i \vee_{V(G)} \mathcal{L}_i, i \in [t]\}$$

Fig. 1. The graph G from Example 2.10.

is a *set-join cover* of G if $E(G) = \bigcup_{i=1}^t E(\mathcal{K}_i \vee_{V(G)} \mathcal{L}_i)$.

For a set-join cover \mathcal{C} we define:

- the *component set* of \mathcal{C} to be:

$$V(\mathcal{C}) = \{\mathcal{K}_i; i \in [t]\} \cup \{\mathcal{L}_i; i \in [t]\},$$

- the *order* of \mathcal{C} to be $|\mathcal{C}| = |V(\mathcal{C})|$,
- the *graph* of \mathcal{C} denoted by $G(\mathcal{C})$ to be a graph with $V(G(\mathcal{C})) = V(\mathcal{C})$ and $\{\mathcal{K}_i, \mathcal{K}_j\} \in E(G(\mathcal{C}))$ if and only if $\mathcal{K}_i \vee_{V(G)} \mathcal{K}_j \in \mathcal{C}$.

Note that in the above definition we allow $\mathcal{K}_i = \mathcal{K}_j$ or $\mathcal{K}_i = \mathcal{L}_j$, so $|\mathcal{C}|$ can be smaller than $2t$, and can in the extreme case even be equal to 1.

Definition 2.9. Let G be graph with a set-join cover \mathcal{C} , and $S \subseteq V(G)$. The *restriction of \mathcal{C} to S* is the set

$$\mathcal{C}[S] = \{(\mathcal{K} \cap S) \vee_{V(G) \cap S} (\mathcal{L} \cap S); \mathcal{K} \vee_{V(G)} \mathcal{L} \in \mathcal{C}, \mathcal{K} \cap S \neq \emptyset, \mathcal{L} \cap S \neq \emptyset\}.$$

In the definition above, $\mathcal{C}[S]$ is a set-join cover of a subgraph of G induced on $V(G) \cap S$ that satisfies $|\mathcal{C}[S]| \leq |\mathcal{C}|$.

Example 2.10. Consider the graph G depicted in Fig. 1 with $V(G) = [6]$. Using the components $\mathcal{K}_1 = \{1, 2, 4\}$, $\mathcal{K}_2 = \{3\}$, $\mathcal{K}_3 = \{4\}$, $\mathcal{K}_4 = \{3, 5, 6\}$ we can define two set-join covers of G of order 4:

$$\mathcal{C}_1 = \{\mathcal{K}_1 \vee_{[6]} \mathcal{K}_2, \mathcal{K}_3 \vee_{[6]} \mathcal{K}_4\},$$

$$\mathcal{C}_2 = \{\mathcal{K}_1 \vee_{[6]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[6]} \mathcal{K}_3, \mathcal{K}_3 \vee_{[6]} \mathcal{K}_4\}.$$

The set-join covers \mathcal{C}_1 and \mathcal{C}_2 have the same component set: $V(\mathcal{C}_1) = V(\mathcal{C}_2) = \{\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \mathcal{K}_4\}$, but they have different graphs: $G(\mathcal{C}_1) = 2K_2$ and $G(\mathcal{C}_2) = P_4$. Note the edge $\{3, 4\} \in E(G)$ is covered twice by \mathcal{C}_1 and three times by \mathcal{C}_2 .

Using $\mathcal{K}'_1 = \{1, 2\}$, $\mathcal{K}'_2 = \{3\}$, $\mathcal{K}'_3 = \{4\}$, $\mathcal{K}'_4 = \{5, 6\}$ we can define another set-join cover of G of order 4 and graph P_4 :

$$\mathcal{C}_3 = \{\mathcal{K}'_1 \vee_{[6]} \mathcal{K}'_2, \mathcal{K}'_2 \vee_{[6]} \mathcal{K}'_3, \mathcal{K}'_3 \vee_{[6]} \mathcal{K}'_4\}.$$

This set-join cover covers every edge in the graph G only once.

Remark 2.11. A set-join cover \mathcal{C} of a graph can be interpreted in the following way. Consider a set of items V that are either required or forbidden to interact. The interactions are organised by meetings of certain subgroups of V . If two subgroups $V_1, V_2 \subseteq V$ meet, then all the items from V_1 interact with all the items from V_2 . Hence, if $i, j \in V$ are forbidden to interact, and V_1 and V_2 are two subgroups that meet, then $i \in V_1$ and $j \in V_2$ is not allowed. The desired interactions can clearly be organised by meetings of singletons, and we are asking what is the minimal number of subgroups that need to be formed, to be able to organise the desired interactions in such a way that no forbidden interactions occur.

Let G be a graph that records which interactions are required and which are forbidden: $V(G) = V$, $\{i, j\} \in E(G)$ if and only if $\{i, j\}$ are required to interact. Any set-join cover \mathcal{C} of G gives us a possible way of organising required interactions, and the minimal order of all possible set-join covers of G is the minimal number of groups that need to be formed.

2.2. Connecting SN-trifactorizations and set-join covers

Our main motivation to introduce set-join covers of graphs is the following theorem.

Theorem 2.12. *Let G be a graph. Then $\text{st}_+(G) = \min\{|\mathcal{C}|; \mathcal{C} \text{ a set-join cover of } G\}$.*

Proof. Let G be a graph with $V(G) = [n]$, and $A \in \mathcal{S}^+(G)$ with SN-Trifactorization $A = BCB^T$, where $C \in \mathcal{S}_k^+$. For $i \in [k]$, let $c_i(B)$ be the i -th column of B , and \mathcal{L}_i the support of the $c_i(B)$. Clearly, $c_i(B)c_j(B)^T + c_j(B)c_i(B)^T \in \mathcal{S}^+(\mathcal{L}_i \vee_{[n]} \mathcal{L}_j)$, and

$$\text{sign}(A) = \text{sign} \left\{ \sum_{\{i,j\} \in E(G(C))} (c_i(B)c_j(B)^T + c_j(B)c_i(B)^T) \right\}.$$

We deduce that $\mathcal{C} = \{\mathcal{L}_i \vee_{[n]} \mathcal{L}_j; \{i, j\} \in E(G(C))\}$ is a set-join cover of G with $|\mathcal{C}| = k$.

Conversely, let $\mathcal{C} = \{\mathcal{L}_i \vee_{[n]} \mathcal{K}_i; i \in [t]\}$ be a set-join cover of G , and let us list all the elements in $V(\mathcal{C})$ in some fixed order: $(\widehat{\mathcal{L}}_1, \widehat{\mathcal{L}}_2, \dots, \widehat{\mathcal{L}}_{|\mathcal{C}|})$. We define C to be a $|\mathcal{C}| \times |\mathcal{C}|$ matrix with:

$$c_{ij} = \begin{cases} 1; & \text{if } \widehat{\mathcal{L}}_i \vee_{[n]} \widehat{\mathcal{L}}_j \in \mathcal{C}, \\ 0; & \text{otherwise,} \end{cases}$$

and B to be an $n \times |\mathcal{C}|$ matrix with

$$b_{ij} = \begin{cases} 1, & \text{if } i \in \widehat{\mathcal{L}}_j \\ 0, & \text{otherwise.} \end{cases}$$

In other words, C is the zero-one matrix in $\mathcal{S}^+(G(\mathcal{C}))$, and B^T is the incidence matrix of $(\widehat{\mathcal{L}}_1, \widehat{\mathcal{L}}_2, \dots, \widehat{\mathcal{L}}_{|\mathcal{C}|})$. The proof is completed by noting that $BCB^T \in \mathcal{S}^+(G)$. \square

We illustrate Theorem 2.12 by the following examples.

Example 2.13. Let

$$B = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{pmatrix}.$$

The matrix $A = BCB^T$ has $G(A) = K_6$. Since C is a 5×5 matrix this proves $\text{st}_+(K_6) \leq 5$. (In Section 4 we will show that $\text{st}_+(K_6) = 5$.) The patterns of B and C determine a set-join cover \mathcal{C} of K_6 with the components:

$$\mathcal{K}_1 = \{1, 2, 3\}, \mathcal{K}_2 = \{4, 5, 6\}, \mathcal{K}_3 = \{1, 4\}, \mathcal{K}_4 = \{2, 5\}, \mathcal{K}_5 = \{3, 6\},$$

and $\mathcal{C} = \{\mathcal{K}_1 \vee_{[6]} \mathcal{K}_2, \mathcal{K}_3 \vee_{[6]} \mathcal{K}_4, \mathcal{K}_4 \vee_{[6]} \mathcal{K}_5, \mathcal{K}_5 \vee_{[6]} \mathcal{K}_3\}$. (The component \mathcal{K}_i contains the support of the i -th column of B and \mathcal{C} connects the components following the pattern given by the matrix C .) Notice that $G(\mathcal{C}) = G(C) = K_2 \cup K_3$.

Example 2.14. Let G be a graph with the adjacency matrix

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

We claim that $\text{st}_+(G) = 4$. Indeed, any matrix X with pattern matrix $\text{sign}(X) = A$ has rank at least 4 as the first three rows of X clearly have to be linearly independent, and the last row cannot be written as a linear combination of the first three. Since $\text{st}_+(X) \geq \text{rk}(X)$ for any $X \in \mathcal{S}^+(G)$, we have $\text{st}_+(G) \geq 4$. Let

$$\mathcal{K}_1 = \{1, 2\}, \mathcal{K}_2 = \{2, 3\}, \mathcal{K}_3 = \{3, 4\}, \mathcal{K}_4 = \{4, 5\},$$

and observe that $\mathcal{C} = \{\mathcal{K}_1 \vee_{[5]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[5]} \mathcal{K}_3, \mathcal{K}_3 \vee_{[5]} \mathcal{K}_4\}$ a set-join cover of G with $|\mathcal{C}| = 4$, proving $\text{st}_+(G) \leq 4$.

Let $B^T = \iota(\mathbf{R})$ be the incidence matrix of the list $\mathbf{R} = (\mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \mathcal{K}_4)$, C the adjacency matrix of the path on 4 vertices, and $X = BCB^T$. Then $X \in \mathcal{S}^+(G)$ and $\text{st}_+(X) = 4$.

Remark 2.15. A matrix $A \in \mathcal{S}_n^+$ is completely positive, if it can be written as BB^T for some matrix $B \in \mathbb{R}_+^{n \times k}$. Such factorization is called CP-Factorization of A . The minimal possible k in CP-Factorization as above is called the CP-rank of A , and is denoted by $\text{cp}(A)$, see [1]. The set $\mathcal{S}_n^+(G)$ contains matrices with CP-rank k if and only if there exists a set-join cover of G of the form $\mathcal{C} = \{\mathcal{K}_i \vee_{V(G)} \mathcal{K}_i, i = 1 \dots, k\}$, since any CP-factorization $A = BB^T$ can be viewed as an SN-Trifactorization $A = BCB^T$, with the middle matrix C equal to the identity. So the graph $G(\mathcal{C})$ is equal to tK_1^ℓ . Hence, all the edges of G can be covered with k complete graphs with all the loops. The lowest CP-rank that a nonnegative symmetric matrix with prescribed zero-nonzero pattern can have is equal to the clique cover number of the pattern graph.

2.3. Optimal set-join covers

Now that we know that $\text{st}_+(G)$ equals to the minimal possible order of a set-join cover of G , we are interested in set-join covers of a given graph G with the minimal possible number of components.

Definition 2.16. If \mathcal{C} is a set-join cover of G with of order $\text{st}_+(G)$, then we say that \mathcal{C} is an *optimal set-join cover* of G . We will use the abbreviation *OSJ cover* of G .

Proposition 2.17. An OSJ cover \mathcal{C} of G satisfies

$$\text{st}_+(G(\mathcal{C})) = |\mathcal{C}| = |V(G(\mathcal{C}))|.$$

Proof. Let \mathcal{C} be an OSJ cover of G , $C \in \mathcal{S}^+(G(\mathcal{C}))$ and $B = \iota(V(\mathcal{C}))^T$ the transposed incidence matrix of $V(\mathcal{C})$. Then $A = BCB^T \in \mathcal{S}^+(G)$. If $\text{st}_+(G(\mathcal{C})) < |\mathcal{C}|$, then the matrix $C \in \mathcal{S}^+(G(\mathcal{C}))$ can be chosen so that $\text{st}_+(C) < |\mathcal{C}|$. Hence, $C = B_1 C_1 B_1^T$ where $C_1 \in \mathcal{S}_k^+$ with $k < |\mathcal{C}|$. From $A = (BB_1)C_1(BB_1)^T$ we conclude $\text{st}_+(A) \leq k < |\mathcal{C}| = \text{st}_+(G)$, a contradiction. \square

Lemma 2.18. Let \mathcal{C} be a set-join cover for a graph G with $V(G) = [n]$. Assume there exist:

- $V' \subseteq V(\mathcal{C})$ with $|V'| = t$, and
- $V'' = \{\mathcal{L}_i \subseteq [n]; i \in [s]\}$,

so that $s < t$ and every element of V' can be written as the union of some elements of V'' . Then \mathcal{C} is not an OSJ cover of G .

Proof. We will prove the lemma by constructing a set-join cover $\widehat{\mathcal{C}}$ of G with $V(\widehat{\mathcal{C}}) = (V(\mathcal{C}) \setminus V') \cup V''$. This will prove our claim as $|(V(\mathcal{C}) \setminus V') \cup V''| \leq |\mathcal{C}| - t + s < |\mathcal{C}|$. We define $\widehat{\mathcal{C}}$ as the union of the following three sets:

- $\mathcal{C}_1 := \{\mathcal{K} \vee_{[n]} \mathcal{K}'; \mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C} \text{ and } \mathcal{K}, \mathcal{K}' \in (V(\mathcal{C}) \setminus V')\}$.
- $\mathcal{C}_2 := \{\mathcal{K} \vee_{[n]} \mathcal{L}; \mathcal{K} \in (V(\mathcal{C}) \setminus V'), \mathcal{L} \in V'' \text{ and there exists } \mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C} \text{ with } \mathcal{L} \subseteq \mathcal{K}'\}$,
- $\mathcal{C}_3 := \{\mathcal{L} \vee_{[n]} \mathcal{L}'; \mathcal{L}, \mathcal{L}' \in V'' \text{ and there exists } \mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C} \text{ with } \mathcal{L} \subseteq \mathcal{K}, \text{ and } \mathcal{L}' \subseteq \mathcal{K}'\}$.

To prove that $\widehat{\mathcal{C}}$ is a set-join cover of G we need to show that $\widehat{\mathcal{C}}$ covers all the edges of G , and that it does not cover any edges that are not in G . Since \mathcal{C} is a cover of G , it is clear that \mathcal{C}_1 does not cover any edges that are not in $E(G)$. If $\mathcal{K} \in (V(\mathcal{C}) \setminus V'), \mathcal{L} \in V''$ and there exists $\mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C}$ with $\mathcal{L} \subseteq \mathcal{K}'$, then $E(\mathcal{K} \vee_{[n]} \mathcal{L}) \subseteq E(\mathcal{K} \vee_{[n]} \mathcal{K}') \subseteq E(G)$. Hence, \mathcal{C}_2 does not cover any edges not in $E(G)$. The claim for \mathcal{C}_3 is proved in a similar way.

Now let $e \in E(G)$ and $\mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C}$ with $e \in E(\mathcal{K} \vee_{[n]} \mathcal{K}')$. If $\mathcal{K}, \mathcal{K}' \in V(\mathcal{C}) \setminus V'$, then $\mathcal{K} \vee_{[n]} \mathcal{K}' \in \mathcal{C}_1$. If $\mathcal{K} \in V(\mathcal{C}) \setminus V'$ and $\mathcal{K}' \in V'$, then \mathcal{K}' is the union of some elements from V'' , hence there exists $\mathcal{L} \in V''$ so that $\mathcal{L} \subseteq \mathcal{K}'$ and $e \in \mathcal{K} \vee_{[n]} \mathcal{L}$, hence e is covered by \mathcal{C}_2 . The case when $\mathcal{K}, \mathcal{K}' \in V'$ is proved in a similar way. \square

Let \mathcal{C} be a set-join cover of a graph G with the component set $V(\mathcal{C}) = \{\mathcal{K}_i, i \in [t]\}$. Recall that a *system of distinct representatives* (SDR for short) of $V(\mathcal{C})$ is a set $\{x_i, i \in [t]\}$ with the property that $x_i \in \mathcal{K}_i$ and x_i are distinct, [2, Section 1.2].

Lemma 2.19. *Let \mathcal{C} be an optimal set-join cover of a graph G . Then $V(\mathcal{C})$ has a system of distinct representatives.*

Proof. It is well known that a family of subsets $V(\mathcal{C})$ has an SDR if and only if the permanent of the incidence matrix $\iota(V(\mathcal{C}))$ is not zero, [2, Section 7.5]. Assume then that \mathcal{C} is an optimal cover of G with $\text{per}(\iota(V(\mathcal{C}))) = 0$. Since $\text{st}_+(G) \leq n$, this is equivalent to $\iota(V(\mathcal{C}))$ having an $w \times t$ zero submatrix, where $w + t = n + 1$, by Frobenius-König Theorem, see for example [2, Section 1.2]. Hence, there exists $V' \subset V(\mathcal{C})$ with $|V'| = t$ so that $|\cup_{\mathcal{K}_i \in V'} \mathcal{K}_i| \leq n - w = t - 1$. Let $V = \{\{x\}, x \in \cup_{\mathcal{K}_i \in V'} \mathcal{K}_i\}$ be the set of singletons from $\cup_{\mathcal{K}_i \in V'} \mathcal{K}_i$. Then V' and V satisfy the conditions of Lemma 2.18, thus \mathcal{C} is not optimal, a contradiction. \square

Theorem 2.20. *Let \mathcal{C} be an optimal set-join cover of a graph G . Then G contains a subgraph that is isomorphic to $G(\mathcal{C})$.*

Proof. Let \mathcal{C} be an optimal cover for G and $\mathcal{S} = \{x_i, i = 1 \dots, |\mathcal{C}|\}$ an SDR for $V(\mathcal{C})$ that exists by Lemma 2.19. Clearly, there exists a (not necessarily induced) subgraph of G on vertices from \mathcal{S} that is isomorphic to $G(\mathcal{C})$. \square

Remark 2.21. A graph property P is *monotone* if every subgraph of a graph with property P also has property P . Theorem 2.20 shows that if G has a monotone graph property and \mathcal{C} is an OSJ cover of G , then $G(\mathcal{C})$ has it also. For example:

- If G is a forest, then $G(\mathcal{C})$ is a forest.
- If G is triangle free, then $G(\mathcal{C})$ is also.
- If G is bipartite, then $G(\mathcal{C})$ is also.

This insight can help us in finding OSJ covers of a given graph G .

2.4. Uniqueness

After we establish $\text{st}_+(G)$, we can ask, if the set-join cover of G of order $\text{st}_+(G)$ is unique. In addition to mathematical curiosity, uniqueness can inform finding SN-Trifactorizations of matrices $A \in S_n^+$. For example, if the matrix A has a pattern that corresponds to a graph with a unique OSJ cover, then this uniquely determines the patterns of matrices B and C in any SN-Trifactorization $A = BCB^T$ that achieves the SNT-rank of G . Uniqueness can help in deciding the SNT-rank of matrices also in more subtle ways, see Example 5.1.

We will consider three different types of uniqueness as described in the definition below.

Definition 2.22. A graph G has *unique optimal set-join cover* (*unique OSJ cover* for short), if the OSJ cover of G of order $\text{st}_+(G)$ is unique. A graph G has *essentially unique OSJ cover*, if for any two covers \mathcal{C} and \mathcal{C}' of G satisfying $|\mathcal{C}| = |\mathcal{C}'| = \text{st}_+(G)$, there exists an automorphism $\sigma : V(G) \rightarrow V(G)$ of G so that $\sigma(\mathcal{C}) = \mathcal{C}'$. (For a cover $\mathcal{C} = \{\mathcal{K}_i \vee_{V(G)} \mathcal{L}_i, i = 1 \dots, \text{st}_+(G)\}$, we define $\sigma(\mathcal{C})$ to be the cover $\{\sigma(\mathcal{K}_i) \vee_{V(G)} \sigma(\mathcal{L}_i), i = 1 \dots, \text{st}_+(G)\}$.)

A graph G has the *unique OSJ cover graph*, if all covers \mathcal{C} of G of order $\text{st}_+(G)$ have the same $G(\mathcal{C})$ up to isomorphism of graphs.

Let $A \in S^+(G)$ with $\text{st}_+(A) = \text{st}_+(G)$, and let $A = BCB^T$ be an SN-Trifactorization of A achieving $\text{st}_+(A)$. If G has unique OSJ cover, then the patterns of B and C are uniquely determined, if G has essentially unique OSJ cover, then the patterns of matrices of B and C are determined up to permutation (of columns of B and permutation similarity on C). Finally, if G has unique OSJ graph then the pattern of C (but not necessarily of B) is uniquely determined. We illustrate different forms of uniqueness in the example below.

Example 2.23. Let $G_1 = (3K_1^\ell) \vee K_1^\ell$ be the star graph on 4 vertices with all the loops, and denote the vertices of $3K_1^\ell$ by 1, 2, 3 and the central vertex by 4. It is not difficult to check that $\text{st}_+(G_1) = 3$ and G_1 has the unique OSJ cover \mathcal{C}_1 , with the components:

$$\mathcal{K}_1 = \{1, 4\}, \mathcal{K}_2 = \{2, 4\}, \mathcal{K}_3 = \{3, 4\},$$

$$\mathcal{C}_1 = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_1, \mathcal{K}_2 \vee_{[4]} \mathcal{K}_2, \mathcal{K}_3 \vee_{[4]} \mathcal{K}_3\} \text{ and } G(\mathcal{C}_1) = 3K_1^\ell.$$

Next, let G_2 be defined by $V(G_2) = [4]$ and adjacency matrix

$$A_2 = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Again $\text{st}_+(G_2) = 3$, but G_2 does not even have the unique OSJ cover graph. Indeed, G_2 has two OSJ covers. Both covers have components

$$\mathcal{K}_1 = \{1, 2\}, \mathcal{K}_2 = \{1, 4\}, \mathcal{K}_3 = \{1, 2, 3\},$$

$\mathcal{C}_2 = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_1, \mathcal{K}_1 \vee_{[4]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[4]} \mathcal{K}_3\}$, and $\mathcal{C}_3 = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_1, \mathcal{K}_2 \vee_{[4]} \mathcal{K}_3\}$. Note that $G(\mathcal{C}_2)$ is a path P_3 with a loop on one pendant vertex, and $G(\mathcal{C}_3)$ is $K_1^\ell \cup P_2$.

Finally, let $G_3 = K_3 \vee K_1^\ell$ be a graph on 4 vertices. Denote the vertices of K_3 by 1, 2, 3 and the vertex with the loop by 4. Again, short case analysis shows that $\text{st}_+(G_3) = 3$. This time G_3 has the unique OSJ cover graph K_3 . This can be seen by first noting that the subgraph K_3 of G_3 has the unique OSJ cover graph K_3 . In this case OSJ cover is not unique nor essentially unique. Let

$$\mathcal{K}_1 = \{1, 4\}, \mathcal{K}_2 = \{2, 4\}, \mathcal{K}_3 = \{3, 4\}, \mathcal{K}'_3 = \{3\}.$$

The covers $\mathcal{C}_4 = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[4]} \mathcal{K}_3, \mathcal{K}_3 \vee_{[4]} \mathcal{K}_1\}$ and $\mathcal{C}_5 = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[4]} \mathcal{K}'_3, \mathcal{K}'_3 \vee_{[4]} \mathcal{K}_1\}$ are not isomorphic since the cardinalities of their components do not match.

In Example 4.10 we will see that K_6 has essentially unique but not unique OSJ cover.

2.5. Operations on graphs and SNT-rank

We proceed to develop basic properties of $\text{st}_+(G)$. We start with a simple proposition which follows from the property that for $A \in \mathcal{S}_n^+$, $B \in \mathcal{S}_m^+$ we have $\text{st}_+(A \oplus B) = \text{st}_+(A) + \text{st}_+(B)$, see [7].

Proposition 2.24. *For any graphs G, H we have $\text{st}_+(G \cup H) = \text{st}_+(G) + \text{st}_+(H)$. In particular, $\text{st}_+(G \cup K_1) = \text{st}_+(G)$ as $\text{st}_+(K_1) = 0$.*

Observations so far make it clear, that if G has two vertices with the same set of neighbours, then removing one of those vertices will not change $\text{st}_+(G)$. This gives us an operation on graphs, whose effect on SNT-rank is easily understood. To make this observation more precise, we need the definition below.

Definition 2.25. Let G be a graph and $v \in V(G)$. Then

$$N_G(v) := \{w; \{v, w\} \in E(G)\}$$

is called the *neighbourhood* of v .

Note that $v \in N_G(v)$ precisely when v has a loop in G .

Definition 2.26. Two vertices $v, w \in V(G)$ are called *twins* if $N_G(v) = N_G(w)$. A graph G is *twin-free*, if no two pairs of vertices in $V(G)$ are twins.

By $F_{tw}(G)$ we denote the biggest twin free sub-graph of G . Note that $F_{tw}(G)$ is obtained from G by removing all but one vertex from every set of twins in G .

Definition 2.27. Let $A \in \mathcal{S}_n^+$. The graph $F_{tw}(G(A))$ is called *the twin-free graph* of A . The *twin-free pattern matrix* of A , denoted by $F_{tw}(\text{sign}(A))$, is the adjacency matrix of $F_{tw}(G(A))$ and can be obtained from $\text{sign}(A)$ by removing any duplicate rows and columns from $\text{sign}(A)$.

With this new terminology we can restate our earlier observation.

Proposition 2.28. Let G be a graph. Then $\text{st}_+(G) = \text{st}_+(F_{tw}(G))$. Moreover, G has unique OSJ cover if and only if $F_{tw}(G)$ has.

Proof. Let G be a graph, $v \in V(G)$, and \widehat{G} a graph obtained from G by duplicating the vertex v into vertices v_1 and v_2 . Hence, $V(\widehat{G}) = (V(G) \setminus \{v\}) \cup \{v_1, v_2\}$, and $N_{\widehat{G}}(v_i)$, $i = 1, 2$, is equal to $N_G(v)$ if $v \notin N_G(v)$, and is equal to $(N_G(v) \setminus \{v\}) \cup \{v_1, v_2\}$ otherwise.

A set-join cover of \widehat{G} can be obtained from a set-join cover \mathcal{C} of G by replacing v in each of the components of \mathcal{C} by v_1 and v_2 . Note that two distinct set-join covers of G result two distinct set-join covers of \widehat{G} in this process. Hence, $\text{st}_+(G) \geq \text{st}_+(\widehat{G})$ and unique OSJ cover of \widehat{G} implies unique OSJ cover of G .

In order to prove that unique OSJ cover of G also implies unique OSJ cover of \widehat{G} , let us assume that G has unique OSJ cover, and let $\widehat{\mathcal{C}}$ be some OSJ cover of \widehat{G} . We claim that for $\mathcal{K} \in V(\widehat{\mathcal{C}})$, $v_1 \in \mathcal{K}$ if and only if $v_2 \in \mathcal{K}$. If this is not true, then one set-join cover \mathcal{C}_1 of G can be obtained from $\widehat{\mathcal{C}}$ by removing v_2 and replacing v_1 in each of the components of $\widehat{\mathcal{C}}$ by v , and a different set-join cover \mathcal{C}_2 of G can be obtained from $\widehat{\mathcal{C}}$ by removing v_1 and replacing v_2 in each of the components of $\widehat{\mathcal{C}}$ by v . In particular, this proves that every OSJ cover of \widehat{G} can be obtained from some OSJ cover of G by

the process outlined above. Since the OSJ cover of G is by our assumption unique this establishes the uniqueness of OSJ cover of \widehat{G} . \square

Lemma 2.29. *If \mathcal{C} is an OSJ cover for a graph G , then $G(\mathcal{C})$ is twin-free.*

Proof. The assertion follows directly from Propositions 2.17 and 2.28. \square

Remark 2.30. Suppose that $G(\mathcal{C})$ contains twin vertices \mathcal{K}_1 and \mathcal{K}_2 . Let $\mathcal{K}_3 = \mathcal{K}_1 \cup \mathcal{K}_2$ and

$$\begin{aligned}\mathcal{C}_1 &= \{\mathcal{K} \vee_{V(G)} \mathcal{K}'; \mathcal{K}, \mathcal{K}' \in (V(\mathcal{C}) \setminus \{\mathcal{K}_1, \mathcal{K}_2\}), \mathcal{K} \vee_{V(G)} \mathcal{K}' \in \mathcal{C}\}, \\ \mathcal{C}_2 &= \{\mathcal{K}_3 \vee_{V(G)} \mathcal{K}'; \mathcal{K}' \in (V(\mathcal{C}) \setminus \{\mathcal{K}_1, \mathcal{K}_2\}), \mathcal{K}_1 \vee_{V(G)} \mathcal{K}' \in \mathcal{C}\}, \\ \mathcal{C}_3 &= \{\mathcal{K}_3 \vee_{V(G)} \mathcal{K}_3\}.\end{aligned}$$

If $\mathcal{K}_1 \vee_{V(G)} \mathcal{K}_2 \in \mathcal{C}$ let $\widehat{\mathcal{C}} = \mathcal{C}_1 \cup \mathcal{C}_2 \cup \mathcal{C}_3$, otherwise let $\widehat{\mathcal{C}} = \mathcal{C}_1 \cup \mathcal{C}_2$. It is straightforward to check that $\widehat{\mathcal{C}}$ is set-join cover of G with $V(\widehat{\mathcal{C}}) = (V(\mathcal{C}) \setminus \{\mathcal{K}_1, \mathcal{K}_2\}) \cup \{\mathcal{K}_3\}$, so \mathcal{C} is not an optimal set-join cover of G .

Proposition 2.31. *Let G be a graph and let $\widehat{G} = G \vee K_1^\ell$. Then:*

$$\text{st}_+(\widehat{G}) = \begin{cases} \text{st}_+(G); & \text{if } |N_G(v)| \geq 1 \text{ for all } v \in V(G), \\ \text{st}_+(H) + 2; & \text{if } G = H \cup tK_1. \end{cases}$$

Moreover, if G is a graph without isolated vertices, then G has unique OSJ cover graph if and only if \widehat{G} has one.

Proof. Since any $A \in \mathcal{S}_+(G \vee K_1^\ell)$ has $A_1 \in \mathcal{S}_+(G)$ as a principal sub-matrix $\text{st}_+(G) \leq \text{st}_+(G \vee K_1^\ell)$ holds.

Assume that G has no isolated vertices. Let $V(G) = [n]$ and $V(\widehat{G}) = [n+1]$. Replacing every component \mathcal{K} in a set-join cover of G by $\widehat{\mathcal{K}} := \mathcal{K} \cup \{n+1\}$ produces a set-join cover of \widehat{G} . Conversely, replacing every component \mathcal{K} in a set-join cover of \widehat{G} by $\mathcal{K}' := \mathcal{K} \setminus \{n+1\}$ produces a set-join cover of G . Observe that $\{n+1\}$ is not a component of any OSJ cover of \widehat{G} . Indeed, if a cover $\widehat{\mathcal{C}}$ of \widehat{G} contains $\{n+1\}$ as a component, then we can construct a set join cover $\widehat{\mathcal{C}}'$ of \widehat{G} satisfying $|\widehat{\mathcal{C}}'| = |\widehat{\mathcal{C}}| - 1$ by removing any set-joins involving $\{n+1\}$ and adding $n+1$ to all other components of \mathcal{C} . Since the process outlined above preserves the order and the graphs of the corresponding OSJ covers, the statement for graphs without isolated vertices follows. (We note in passing that G may have unique OSJ cover, but \widehat{G} doesn't, since the vertex $n+1$ is not necessarily contained in all the components of OSJ cover of $G \vee K_1^\ell$.)

Now, let $G = H \cup tK_1$, where we assume that H does not contain any isolated vertices (without loops). Let $V(G) = V(H) \cup \mathcal{I} = [n]$, and $V(\widehat{G}) = [n+1]$ as before. Defining $\widehat{\mathcal{K}}$ as above, produces from a set-join cover \mathcal{C} of G a set-join cover $\widehat{\mathcal{C}}$ of $(G \vee K_1^\ell) \cup tK_1$,

since the elements of \mathcal{I} are not contained in any component in this construction. Adding $\mathcal{I} \vee_{[n+1]} \{n+1\}$ to $\widehat{\mathcal{G}}$ results in a set-join cover of \widehat{G} , and introduces 2 new components. We conclude that $\text{st}_+(\widehat{G}) \leq \text{st}_+(H) + 2$.

Finally, let \mathcal{C}' be an OSJ cover of \widehat{G} . Since the elements of \mathcal{I} are connected only to $\{n+1\}$ in \widehat{G} , we observe that \mathcal{C}' has to contain $\{\mathcal{K}_0 \vee_{[n+1]} \{n+1\}\}$, where $\mathcal{I} \subseteq \mathcal{K}_0$, and \mathcal{K}_0 is not a component of any other set-join in \mathcal{C}' . Let \mathcal{C}'' be obtained from \mathcal{C}' by removing all set-joins with a component $\{n+1\}$. Note that $|\mathcal{C}''| \leq |\mathcal{C}'| - 2$, since $V(\mathcal{C}'')$ does not contain $\{n+1\}$ nor \mathcal{K}_0 . Replacing every $\mathcal{K} \vee_{V(\widehat{G})} \mathcal{L} \in \mathcal{C}''$ by $\mathcal{K}' \vee_{V(H)} \mathcal{L}'$, where $\mathcal{K}' := \mathcal{K} \setminus \{n+1\}$ as above, gives us a set-join cover of H , proving $\text{st}_+(\widehat{G}) = \text{st}_+(H) + 2$. \square

Remark 2.32. Note that $G = K_2 \cup K_1$ has unique OSJ cover graph, but $\widehat{G} = G \vee K_1^\ell$ doesn't.

From Propositions 2.24, 2.28 and 2.31 we see that in order to understand SNT-rank of graphs, we can from now on consider only connected twin-free graphs that are not of the form $G \vee K_1^\ell$.

Example 2.33. Threshold graphs is a family of graphs (without loops) that can be constructed by repeating two operations, adding an isolated vertex ($G \cup K_1$), and joining a vertex ($G \vee K_1$). Here we extend this definition to *threshold graphs with loops* to be all graphs that can be constructed by repeating the following two operations on G :

- $G \cup K_1$
- $G \vee K_1^\ell$

To obtain a twin free graph the two operations have to alternate. Considering only connected twin-free graphs resulting from this process, we get the following sequence of graphs:

$$T_1 := K_1 \vee K_1^\ell, T_{i+1} := (T_i \cup K_1) \vee K_1^\ell.$$

We have $\text{st}_+(T_i) = 2i$ by Proposition 2.31. This is not surprising, as it is not difficult to see that every matrix $A \in \mathcal{S}^+(T_i)$ has $\text{rk}(A) = 2i$.

Remark 2.34. Let G be a graph with a cut edge $\{u_1, u_2\}$, so that G with this edge removed is equal to $G_1 \cup G_2$, where $u_i \in V(G_i)$, $i = 1, 2$. Then:

$$\text{st}_+(G_1) + \text{st}_+(G_2) \leq \text{st}_+(G) \leq \text{st}_+(G_1) + \text{st}_+(G_2) + 2.$$

Moreover, $\text{st}_+(G) = \text{st}_+(G_1) + \text{st}_+(G_2)$ if and only if for $i = 1, 2$ there exist OSJ covers \mathcal{C}_i with $\{u_i\} \in V(\mathcal{C}_i)$, and $\text{st}_+(G) = \text{st}_+(G_1) + \text{st}_+(G_2) + 1$ if such cover exists for either $i = 1$ or $i = 2$, but not both.

3. Trees and cycles without loops

Let G be a graph without loops that does not contain any four cycles. Then a set-join cover of G can contain only elements of the form $\mathcal{K} \vee_{V(G)} \{v\}$ for some $v \in V(G)$, since all set-joins $\mathcal{K} \vee_{V(G)} \mathcal{L}$ with $|\mathcal{K}| \geq 2$ and $|\mathcal{L}| \geq 2$ contain a four cycle. Hence, for graphs without loops and four cycles a set-join cover is equivalent to an edge star cover, as defined below.

Definition 3.1. Let G be a simple graph without loops. A family of simple stars $\{S_1, S_2, \dots, S_k\}$ is an *edge star cover* of G if $E(G) = \cup_{i=1}^k E(S_i)$. The *edge star cover number* $\text{star}(G)$ of G is the minimal number of stars in any edge star cover of G .

Clearly, $\text{st}_+(G) \leq 2\text{star}(G)$ for any graph G , and we will show that this is an equality for trees. For any tree T it is known that $\text{rk}(A)$ equals twice the matching number of T for all $A \in \mathcal{S}^+(T)$, [5]. Since for a tree T the matching number equals $\text{star}(T)$, the inequality $\text{st}_+(T) \geq 2\text{star}(T)$ clearly holds. In this work we use the edge star cover number due to its immediate connection to set-join covers. We summarise this observation in the next proposition.

Proposition 3.2. Let T be a forest (without loops) and $A \in \mathcal{S}^+(T)$. Then $\text{rk}(A) = \text{st}_+(A) = 2\text{star}(T)$. In particular, $\text{st}_+(T) = 2\text{star}(T)$.

In the next lemma we see that if a graph G contains a leaf, then any set-join cover has to contain at least one element of the form $\mathcal{K} \vee \{v\}$ for some $v \in V(G)$, i.e. it contains at least one star.

Lemma 3.3. Let G be a graph and let $L(G) \subset V(G)$ be the set of all leaves without a loop in G . Let $\ell \in L(G)$, $w \in V(G)$ its unique neighbour, and G' the graph obtained from G by deleting all edges $\{w, v'\}$ with $v' \in N_G(w)$ (and all singletons that result after this deletion). Then $\text{st}_+(G) = \text{st}_+(G') + 2$ and for any optimal set-join cover \mathcal{C} of G , we have $\{w\} \vee_{V(G)} \mathcal{N} \in \mathcal{C}$, where $N_G(w) \cap L(G) \subseteq \mathcal{N} \subseteq N_G(w)$.

Proof. Let \mathcal{C} be an optimal set-join cover of G . Since $N_G(\ell) = \{w\}$, there exists $\mathcal{N} \subseteq N_G(w)$ with $\ell \in \mathcal{N}$, so that $\{w\} \vee_{V(G)} \mathcal{N} \in \mathcal{C}$. If there exists $\ell' \in (N_G(w) \cap L(G)) \setminus \mathcal{N}$, then we also have $\{w\} \vee_{V(G)} \mathcal{N}' \in \mathcal{C}$ for some \mathcal{N}' with $\ell' \in \mathcal{N}'$. Since $\mathcal{C}_1 := (\mathcal{C} \cup (\{w\} \vee_{V(G)} N_G(w))) \setminus \{\{w\} \vee_{V(G)} \mathcal{N}, \{w\} \vee_{V(G)} \mathcal{N}'\}$ is a set-join cover of G with $|\mathcal{C}_1| \leq |\mathcal{C}| - 1$, we get a contradiction with the assumption that \mathcal{C} is optimal. Hence, $N_G(w) \cap L(G) \subseteq \mathcal{N}$.

Observe that $\mathcal{C}' := \mathcal{C} \setminus \{\{w\} \vee_{V(G)} \mathcal{N}\}$ is a set-join cover for $G' \cup tK_1$ for some $t \in \mathbb{N}$. Since, $\text{st}_+(G') = \text{st}_+(G' \cup tK_1)$, this implies $\text{st}_+(G') \leq \text{st}_+(G) + 2$. On the other hand, $\mathcal{C} = \{\{w\} \vee_{V(G)} N_G(w)\} \cup \mathcal{C}'$ is a set-join cover of G for any set-join cover \mathcal{C}' of G' . Hence, $\text{st}_+(G) \leq \text{st}_+(G') + 2$, as desired. \square

Example 3.4. For paths we have $\text{st}_+(P_{2k}) = \text{st}_+(P_{2k+1}) = 2k$ by inductive application of Lemma 3.3.

In the result below we resolve the question of uniqueness of OSJ covers for trees without loops.

Theorem 3.5. *Let T be a tree with $|V(T)| \geq 3$. Then T has unique OSJ cover if and only if the distance between any two leaves in T is even.*

Proof. Assume the distance between any two leaves in T is even, and let V_0 be the set of vertices in T at an odd distance to all leaves. An inductive application of Lemma 3.3 shows that

$$\mathcal{C}_0 := \{\{v\} \vee N_T(v); v \in V_0\}$$

is an OSJ cover of T . Let \mathcal{C}' be an OSJ cover of T and V' the set of all central vertices of stars in \mathcal{C}' . Then $|\mathcal{C}'| = 2|V'| = 2|V_0|$. Since for any $v \in V(T) \setminus V'$ there exists $w \in V' \cap N_T(v)$, we necessarily have $V' = V_0$. Since no two vertices in V_0 are connected in T , \mathcal{C}_0 is the only collection of stars with central vertices V_0 that covers T . We conclude that the cover is unique. We further remark, that the unique OSJ cover graph of T is $|V_0|P_2$.

Assume now, that there are two leaves ℓ_1 and ℓ_2 in a tree T at odd distance d , $d \geq 3$. Observe that in any set-join cover \mathcal{C} of T at least every second vertex of the path between ℓ_1 and ℓ_2 is a central vertex of a star in that cover, hence a component in \mathcal{C} . Since d is odd, we have $\{v_1\} \vee \mathcal{S}_1$ and $\{v_2\} \vee \mathcal{S}_2$ in \mathcal{C} , such that v_1 and v_2 are neighbours in T , and at least one of the inclusions $v_1 \in \mathcal{S}_2$ and $v_2 \in \mathcal{S}_1$ holds. Assume now that \mathcal{C} is an OSJ cover, and without loss of generality that $v_1 \in \mathcal{S}_2$. Then we get a different OSJ cover of T by replacing $\{v_1\} \vee \mathcal{S}_1$ and $\{v_2\} \vee \mathcal{S}_2$ by $\{v_1\} \vee (\mathcal{S}_1 \cup \{v_2\})$ and $\{v_2\} \vee (\mathcal{S}_2 \setminus \{v_1\})$. \square

Example 3.6. Note that the only tree with $|V(T)| = 2$ is a path P_2 and it has the unique OSJ cover. Moreover, a path P_n has unique OSJ cover if and only if n is odd or $n = 2$, see Theorem 3.5.

For integers $t \geq 3$ and $k_i \geq 1$, $i \in [t]$, we denote by $\text{star}(k_1, k_2, \dots, k_t)$ the graph with the central vertex v and t arms, where each arm is a path on $k_i + 1$ vertices and one of its end-vertices is v . Hence, $\text{star}(k_1, k_2, \dots, k_t)$ is a generalised star with $\sum_{i=1}^t k_i + 1$ vertices. If $k_i = 1$ for all i , then $\text{star}(1, 1, \dots, 1) = K_{1,t}$ is a star graph with t leaves.

Example 3.7. Let $G = \text{star}(k_1, k_2, \dots, k_t)$, $t \geq 3$, be a generalised star. Then

$$\text{st}_+(G) = \begin{cases} |V(G)| + 1 - |\{k_i; k_i \text{ odd}\}|; & \text{if at least one } k_i \text{ is odd,} \\ |V(G)| - 1; & \text{if all } k_i \text{ are even.} \end{cases}$$

We will prove this claim by induction on

$$\psi(G) := \sum_{i=1}^t k_i - t.$$

We first assume $\psi(G) = 0$. This implies $G = G_0(t) := \text{star}(1, 1, \dots, 1)$, i.e. a star with t leaves. Clearly, $\text{st}_+(G_0(t)) = 2$, and the claim holds. For $\psi(G) = 1$, we have $G = G_1(t) := \text{star}(2, 1, \dots, 1)$, i.e. a generalised star with t arms. Applying Lemma 3.3 for the leaf on the long arm, we get:

$$\text{st}_+(G_1(3)) = \text{st}_+(P_3) + 2 = 4$$

by Example 3.4, and for $t > 3$:

$$\text{st}_+(G_1(t)) = \text{st}_+(G_0(t-1)) + 2 = 4.$$

Since $G_1(t)$ has $t+2$ vertices and $t-1$ arms of odd length, this establishes the base of induction.

Now assume the claim holds for all generalised stars $G' = \text{star}(k'_1, k'_2, \dots, k'_{t'})$ with $\psi(G') \leq \psi$, and let $G := \text{star}(k_1, k_2, \dots, k_t)$ with $\psi(G) = \psi + 1 \geq 2$. Without loss of generality we can assume $k_1 \geq 2$. By Lemma 3.3 we have $\text{st}_+(G) = \text{st}_+(G') + 2$, where:

- $G' := \text{star}(k_2, \dots, k_t)$ with $\psi(G') = \psi(G) - 1 \geq 1$ for $k_1 = 2$,
- $G' := \text{star}(k_1 - 2, k_2, \dots, k_t)$ with $\psi(G') = \psi(G) - 2$ for $k_1 \geq 3$.

If $k_1 = 2$ and $t = 3$, then $G' = P_{k_2+k_t+1}$, and the claim holds for G by Example 3.4. In all other cases G' is again a generalised star, so we can use the induction hypothesis. Noticing that $V(G) - V(G') = 2$ and that G and G' have same number of arms of odd length, we establish the claim.

The graph G has unique OSJ cover if and only if all k_i 's are odd or all even by Theorem 3.5.

A vertex v is a *cut* vertex of a connected graph G , if by removing v and all edges $\{v, w\}$, $w \in N_G(v)$, from G we get a disconnected graph that we will denote by $G \setminus \{v\}$. (We allow for a cut vertex to have a loop.)

Proposition 3.8. *Let v be a cut vertex of a connected graph G , where $k \geq 2$ and $G \setminus \{v\} = \cup_{i=1}^k G_i$ for some connected graphs G_i . Then $\text{st}_+(G) \leq 2 + \sum_{i=1}^k \text{st}_+(G_i)$.*

Proof. Let \mathcal{C}_i be a set-join cover of G_i , $i \in [k]$. Then

$$\mathcal{C} := \cup_{i=1}^k \mathcal{C}_i \cup \{\{v\} \vee_{V(G)} N_G(v)\}$$

is a set-join cover of G . Since $|\mathcal{C}| = 2 + \sum_{i=1}^k |\mathcal{C}_i|$, the claim follows. \square

Example 3.9. The inequality of Proposition 3.8 is not always an equality, but it cannot be improved in general.

1. Let $G = \text{star}(2, 2, 2)$. Then $\text{st}_+(G) = 6$ by Proposition 3.7 and $G \setminus \{v\} = 3P_2$, so $2 + \sum_{i=1}^k \text{st}_+(G_i) = 8$.
2. Let G be a star graph $G = \text{star}(3, 3, 3)$. Then $\text{st}_+(G) = 8$ by Proposition 3.7 and $G \setminus \{v\} = 3P_3$, so $2 + \sum_{i=1}^k \text{st}_+(G_i) = 8$.

We now consider SNT-rank of cycles.

Proposition 3.10. *For cycles C_n we have $\text{st}_+(C_4) = 2$ and $\text{st}_+(C_n) = n$ for $n = 3$ and all $n \geq 5$. Furthermore, the graph C_n has unique OSJ cover if and only if $n = 4$ or n is odd.*

Proof. If $n = 3$, then the determinant of any matrix $A \in \mathcal{S}^+(C_3)$ is positive, so $\text{rk}(A) = 3$, thus $\text{st}_+(C_3) = 3$. If $n = 4$, then

$$\mathcal{C} = \{\{1, 3\} \vee_{[4]} \{2, 4\}\}$$

is a set-join cover of C_4 , and it is easy to check that it is unique. Hence, $\text{st}_+(C_4) = 2$.

Now let $n \geq 5$. If the only OSJ cover of C_n has all its components equal to a singleton, then $\text{st}_+(C_n) = n$ and C_n has unique OSJ cover. So assume that \mathcal{C} is an OSJ cover of C_n that contains a component \mathcal{K}_1 , which is not a singleton. In this case $\mathcal{K}_1 = N_{C_n}(v)$ for some $v \in V(C_n)$ and $\{v\} \vee_{[n]} \mathcal{K}_1 \in \mathcal{C}$. Note that the components $\{v\}$ and \mathcal{K}_1 cannot be part of any other set-join in \mathcal{C} . This implies that $\mathcal{C}' = \mathcal{C} \setminus \{\{v\} \vee_{[n]} \mathcal{K}_1\}$ is a set-join cover of P_{n-1} , and since \mathcal{C} is optimal so is \mathcal{C}' . From $|\mathcal{C}| = 2 + |\mathcal{C}'| = 2 + \text{st}_+(P_{n-1})$ and Example 3.4, we deduce that \mathcal{C} has to be even, and $|\mathcal{C}| = n$. In particular, we have shown that for n odd, C_n has unique OSJ cover (with all components equal to singletons). For $n = 2k$ we have two additional OSJ covers:

$$\mathcal{C}_1 = \{\{2i\} \vee_{[n]} \mathcal{N}_{C_n}(2i); i = 1, 2, \dots, k\}$$

$$\mathcal{C}_2 = \{\{2i - 1\} \vee_{[n]} \mathcal{N}_{C_n}(2i - 1); i = 1, 2, \dots, k\},$$

proving that in this case OSJ cover is not unique. \square

Remark 3.11. Let G be an unicyclic graph, i.e. connected simple graph without loops containing exactly one cycle. Then G is either a cycle or it contains at least one leaf. We can find the SNT-rank of G by repeatedly applying Lemma 3.3 on a leaf. In each step we obtain a disjoint union of trees and at most one unicyclic graph. Eventually we obtain a disjoint union of trees and at most one cycle. For this graph the SNT-rank can be computed by Proposition 3.2 and Proposition 3.10.

4. Complete graphs without loops

Let \mathcal{C} be a set-join cover of K_n with $V(K_n) = [n]$. If $\mathcal{K} \vee_{[n]} \mathcal{L} \in \mathcal{C}$, then $\mathcal{K} \cap \mathcal{L} = \emptyset$, since K_n has no loops. On the other hand, for any pair $i, j \in [n]$, $i \neq j$, there exist $\mathcal{K} \vee_{[n]} \mathcal{L} \in \mathcal{C}$, so that $i \in \mathcal{K}$, $j \in \mathcal{L}$.

Definition 4.1. Let \mathbf{T} be a k -tuple of subsets of $[n]$. If for any pair $i, j \in [n]$, $i \neq j$, there exists $\mathcal{K}, \mathcal{L} \in \mathbf{T}$ so that $i \in \mathcal{K}$, $j \in \mathcal{L}$ and $\mathcal{K} \cap \mathcal{L} = \emptyset$, then \mathbf{T} is called a *separating cover* of n elements with k sets.

Hence, if \mathcal{C} is a set-join cover of K_n , then $V(\mathcal{C})$ is a separating cover of n elements with $|V(\mathcal{C})|$ sets. To determine $\text{st}_+(K_n)$ for a given n , we need to find minimal separating cover of n elements. This problem is known as Katona problem and has been solved independently by A. C. C. Yao, [9] and M. C. Cai [3], hence the problem of determining $\text{st}_+(K_n)$ is resolved. In this section we expand on the work in [9], by showing how various set-join covers of K_n can be constructed, investigate when these covers are optimal, and when they are essentially unique.

Theorem 4.2 ([9]). Let $n \in \mathbb{N}$, and let $s(n) := \min\{k: \text{there is a separating cover on } n \text{ elements with } k \text{ sets}\}$. For all $n \geq 2$,

$$s(n) = \begin{cases} 3i & \text{if } 2 \cdot 3^{i-1} < n \leq 3^i \\ 3i + 1 & \text{if } 3^i < n \leq 4 \cdot 3^{i-1} \\ 3i + 2 & \text{if } 4 \cdot 3^{i-1} < n \leq 2 \cdot 3^i, \end{cases} \quad (1)$$

while $s(1) = 0$.

Corollary 4.3. The SNT-rank of a complete graph is $\text{st}_+(K_n) = s(n)$, where $s(n)$ is defined in (1).

To give some intuition on how a set-join cover for K_n can be constructed, we offer the following example.

Example 4.4. Suppose that $n \in \mathbb{N}$ is not prime and $n = p \cdot q$ with $p, q \geq 2$. Let $V(K_n) = [n]$, and define:

$$\begin{aligned} \mathcal{L}_i &= \{(i-1)p + s; s = 1, \dots, p\} \text{ for every } i \in [q], \\ \mathcal{K}_j &= \{j + tp; t = 0, \dots, q-1\}, \text{ for every } j \in [p]. \end{aligned}$$

Then $\mathcal{L}_i \cap \mathcal{L}_j = \emptyset$ for $i \neq j$ and $\mathcal{K}_{i'} \cap \mathcal{K}_{j'} = \emptyset$ for $i' \neq j'$, and

$$\mathcal{C} = \{\mathcal{L}_i \vee_{[n]} \mathcal{L}_j, i, j \in [q], i \neq j\} \cup \{\mathcal{K}_i \vee_{[n]} \mathcal{K}_j, i, j \in [p], i \neq j\}$$

is a set-join cover of K_n . Indeed, let $a, b \in [n]$, $a \neq b$, and let $a = (i_a - 1)p + j_a$ and $b = (i_b - 1)p + j_b$ with $i_a, i_b \in [q]$ and $j_a, j_b \in [p]$. If $j_a = j_b$, then $i_a \neq i_b$ and $\{a, b\} \in E(\mathcal{L}_{i_a} \vee \mathcal{L}_{i_b})$, and if $j_a \neq j_b$ then $\{a, b\} \in E(\mathcal{K}_{j_a} \vee \mathcal{K}_{j_b})$. This argument shows that $\text{st}_+(K_n) \leq p + q$.

A generalisation of this example leads to an OSJ cover of K_n .

Proposition 4.5. *Let $n \in \mathbb{N}$ and $q_i \in \mathbb{N}$, $i \in [t]$, satisfy $q_i \geq 2$ and $n \leq \prod_{k=1}^t q_k$. Then there exists a set-join cover \mathcal{C} of K_n with $|\mathcal{C}| = \sum_{i=1}^t q_i$ and $G(\mathcal{C}) = \cup_{i=1}^t K_{q_i}$. Furthermore,*

$$\text{st}_+(K_n) = \min \left\{ \sum_{i=1}^t q_i; n \leq \prod_{i=1}^t q_i \right\}. \quad (2)$$

Proof. We prove the first claim by induction on t . If $t = 1$, then the claim is trivial. If $n \leq m$, then K_n is a subgraph of K_m and a restriction of any set-join cover of K_m to the elements of K_n results in a set-join cover of K_n , in particular $\text{st}_+(K_n) \leq \text{st}_+(K_m)$. This observation allows us to assume $n = \prod_{k=1}^t q_k$. The claim for $t = 2$ is proved in Example 4.4 and establishes the base of induction. Let $n_1 = \prod_{k=1}^{t-1} q_k$, $V(K_n) = [n]$ and $V(K_{n_1}) = [n_1]$. Let \mathcal{C} be a set-join cover of K_{n_1} with $|\mathcal{C}| = \sum_{k=1}^{t-1} q_k$. For every $\mathcal{K} \in V(\mathcal{C})$ let

$$\mathcal{K}' = \cup_{i \in \mathcal{K}} \{i + tn_1; t = 0, \dots, q_t - 1\} \subset [n].$$

Furthermore, let

$$\mathcal{L}_i = \{(i - 1)n_1 + s; s = 1, \dots, n_1\} \text{ for every } i \in [q_t].$$

Then

$$\mathcal{C}' = \{\mathcal{K}'_1 \vee_{[n]} \mathcal{K}'_2; \mathcal{K}_1 \vee_{[n_1]} \mathcal{K}_2 \in \mathcal{C}\}$$

covers all the edges of K_n within the sets $\mathcal{L}_1, \mathcal{L}_2, \dots, \mathcal{L}_{q_t}$. So

$$\widehat{\mathcal{C}} = \mathcal{C}' \cup \{\mathcal{L}_i \vee_{[n]} \mathcal{L}_j, i, j \in [q_t], i \neq j\}$$

is a set-join cover of K_n with $|\widehat{\mathcal{C}}| = \sum_{i=1}^t q_i$.

To complete the proof, we need to show $\text{st}_+(K_n) \geq \min\{\sum_{i=1}^t q_i; n \leq \prod_{i=1}^t q_i\}$. Let us fix n and define $m = \max\{t; s(t) = s(n)\}$. From Theorem 4.2 we know that m is the largest positive integer that satisfies $\text{st}_+(K_n) = \text{st}_+(K_m)$. In particular, m has one of the forms $m = 3^i$, $m = 4 \cdot 3^{i-1}$, or $m = 2 \cdot 3^i$. In each case we can write $m = \prod_{k=1}^t q_k$, where $q_j \in \{2, 3, 4\}$ for all $j \in [t]$. It is straightforward to check that the equality $\text{st}_+(K_m) = \sum_{i=1}^t q_i$ holds in each of these cases. \square

Remark 4.6. An optimal set-join cover of K_6 is given in Example 2.13. Following the construction outlined in the proof of Proposition 4.5, we can produce examples of optimal set-join covers of K_n for every $n \geq 2$.

If $n = 2, 3, 4, 5$, then $\text{st}_+(K_n) = n$, and the set-join cover \mathcal{C} , where all the components of \mathcal{C} are singletons and $G(\mathcal{C}) = K_n$, is an OSJ cover of K_n . Note that for $n = 2$ and for $n = 3$ these OSJ covers are unique. For $n = 4$ this set-join cover is not unique, since the factorization $4 = 2^2$ gives rise to an OSJ cover \mathcal{C}' of K_4 with $G(\mathcal{C}') = 2K_2$. In the case $n = 5$ this set-join cover is also not unique, since the factorization $6 = 2 \cdot 3$ gives rise to an OSJ cover \mathcal{C}' of K_5 with $G(\mathcal{C}') = K_2 \cup K_3$.

Suppose that $n \geq 6$ and let $m = \max\{t; s(t) = s(n)\}$. An inductive construction outlined in the proof of Proposition 4.5 gives an OSJ cover \mathcal{C} of K_m , which we can restrict to K_n to obtain an OSJ cover of K_n . Let us look at some properties of an OSJ cover \mathcal{C} constructed in this way, for a few choices of n :

- For $n = 3^i$, all components of \mathcal{C} have cardinality 3^{i-1} , and for each $\mathcal{K} \in \mathcal{C}$ there exist two other components $\mathcal{K}', \mathcal{K}'' \in V(\mathcal{C})$, so that $\mathcal{K}, \mathcal{K}'$ and \mathcal{K}'' are pairwise disjoint.
- For $n = 3^i - 1$, $2i$ components have cardinality 3^{i-1} , and i components have cardinality $3^{i-1} - 1$. For each $\mathcal{K} \in \mathcal{C}$ there exist two other components $\mathcal{K}', \mathcal{K}'' \in V(\mathcal{C})$, so that $\mathcal{K}, \mathcal{K}'$ and \mathcal{K}'' are pairwise disjoint.
- For $n = 2 \cdot 3^i$, the largest two components $\mathcal{K}, \mathcal{K}' \in V(\mathcal{C})$ are disjoint and have cardinality 3^i . All other components have cardinality $2 \cdot 3^{i-1}$.
- For $n = 2 \cdot 3^i - 1$, the largest two components in $V(\mathcal{C})$ are disjoint, one has 3^i and the other $3^i - 1$ elements. There are $2i$ components of cardinality $2 \cdot 3^{i-1}$, and i components of cardinality $2 \cdot 3^{i-1} - 1$.

Having constructed an OSJ cover of K_n for each $n \in \mathbb{N}$, we want to consider what properties any OSJ cover \mathcal{C} of K_n has to have. In particular, we consider the components of minimal and maximal cardinality.

Lemma 4.7. Let $m, n \in \mathbb{N}$, $n < m$, $\text{st}_+(K_n) = \text{st}_+(K_m)$, and let \mathcal{C} be an OSJ cover of K_m . Then any component $\mathcal{K} \in V(\mathcal{C})$ satisfies $|\mathcal{K}| > m - n$.

Proof. Let \mathcal{C} be an OSJ cover of K_m , and $\mathcal{K} \in V(\mathcal{C})$. Let \mathcal{C}' be obtained from \mathcal{C} by removing \mathcal{K} , and elements from \mathcal{K} from all components of \mathcal{C} . Then \mathcal{C}' is a set-join cover of $K_{m-|\mathcal{K}|}$ with $|\mathcal{C}'| \leq |\mathcal{C}| - 1$. From $\text{st}_+(K_n) = \text{st}_+(K_m)$ we get $m - |\mathcal{K}| < n$, as required. \square

Corollary 4.8. Let \mathcal{C} be an OSJ cover of K_n and $\mathcal{K} \in V(\mathcal{C})$.

1. If $n = 3^i$, then $|\mathcal{K}| \geq 3^{i-1}$.
2. If $n = 4 \cdot 3^i$, then $|\mathcal{K}| \geq 3^i$.
3. If $n = 2 \cdot 3^i$, then $|\mathcal{K}| \geq 2 \cdot 3^{i-1}$.

Proof. Let $n = 3^i$, then $\text{st}_+(K_{3^i}) = \text{st}_+(K_{2 \cdot 3^{i-1} + 1})$ by Theorem 4.2. Hence, $|\mathcal{K}| > 3^{i-1} - 1$ by Lemma 4.7. The other two items are proved in the same way. \square

Proposition 4.9. *Suppose that $n \in \mathbb{N}$ satisfies one of the following conditions: $n \in \{4, 5, 7, 8\}$, or $3^i + 1 \leq n \leq 2 \cdot 3^i - 2$ for some $i \geq 2$, or $2 \cdot 3^i + 1 \leq n \leq 3^{i+1} - 2$ for some $i \geq 2$. Then an OSJ cover of K_n is not essentially unique.*

Proof. In Remark 4.6 we have already seen that an OSJ cover of K_n is not essentially unique for $n \in \{4, 5\}$. K_8 has two essentially different OSJ covers: one arising from the factorization $8 = 2 \cdot 4$, and the other from $8 \leq 3^2$. In the first case the cover has the graph $K_2 \cup K_4$, and in the second case the cover has the graph $2K_3$. Since $s(8) = 6$ both covers are optimal. Restrictions of those covers to the elements of K_7 gives two OSJ covers of K_7 .

Suppose that $n + 2 \leq m$ and $s(n) = s(m)$. We will show that in this case an OSJ cover of K_n is not essentially unique. Let \mathcal{C} be an OSJ cover for K_m and let $\mathcal{K} \in V(\mathcal{C})$ be a component of \mathcal{C} of minimal cardinality. Then $|\mathcal{K}| > m - n \geq 2$, by Lemma 4.7. Now, choose $m - n$ vertices in \mathcal{K} and remove them from K_m and \mathcal{C} to obtain a set-join cover \mathcal{C}' of K_n . The set-join cover \mathcal{C}' contains a component of cardinality $|\mathcal{K}| - (m - n)$. Next, choose $m - n$ vertices in K_m so that they do not all belong to the same component of \mathcal{C} . This is possible since $m - n \geq 2$, and $m \geq 9$, so $m \geq |V(\mathcal{C})| + 3$. Remove them from K_m and \mathcal{C} to obtain a set-join cover \mathcal{C}'' of K_n . The set-join cover \mathcal{C}'' does not contain a component of cardinality $|\mathcal{K}| - (m - n)$, so it is essentially different from \mathcal{C}' .

For $n = 4 \cdot 3^i = 2^2 \cdot 3^i$ the construction in Proposition 4.5 gives rise to two set-join covers \mathcal{C}_1 and \mathcal{C}_2 , one with the graph $K_4 \cup iK_3$, and one with the graph $2K_2 \cup iK_3$. Both covers are optimal, proving that in this case OSJ cover is not unique. Now let m satisfy $m < n$ and $s(m) = s(n)$. By restricting \mathcal{C}_1 and \mathcal{C}_2 to $[m]$, we get two OSJ covers of K_m with different graphs, proving that OSJ covers of K_m are not essentially unique. \square

Before we move to more general cases, we show in the example below that K_6 has essentially unique OSJ cover.

Example 4.10. Let $V(K_6) = [6]$ and let \mathcal{C} , with $V(\mathcal{C}) = \{\mathcal{K}_i; i = 1, \dots, 5\}$, be an OSJ cover of K_6 . In addition, we assume $|\mathcal{K}_1| \geq |\mathcal{K}_i|$, $i = 2, 3, 4, 5$, and $\mathcal{K}_1 \vee_{[6]} \mathcal{K}_2 \in \mathcal{C}$.

Next we show that $|\mathcal{K}_1| = 3$. Since \mathcal{K}_1 and \mathcal{K}_2 are disjoint, the set-joins of \mathcal{C} with components $\mathcal{K}_3, \mathcal{K}_4$, and \mathcal{K}_5 , need to cover all the edges between the elements of \mathcal{K}_1 . In particular, the set $V_1 = \{\mathcal{K}_3 \cap \mathcal{K}_1, \mathcal{K}_4 \cap \mathcal{K}_1, \mathcal{K}_5 \cap \mathcal{K}_1\}$ needs to contain a component set of a set-cover for a complete graph with the vertex set \mathcal{K}_1 . Since $s(4) = 4$, this excludes $|\mathcal{K}_1| \geq 4$. On the other hand, $s(6) = s(5)$ implies that $V(\mathcal{C})$ doesn't contain singletons by Lemma 4.7. If $|\mathcal{K}_i| = 2$ for all $i \in [5]$, then we may without loss of generality assume that $\mathcal{K}_1 = \{1, 2\}$ and $\mathcal{K}_2 \subset \{3, 4, 5, 6\}$. Since the edge $\{1, 2\}$ needs to be covered, we may also assume that $\mathcal{K}_3 = \{1, 3\}$ and $\mathcal{K}_4 = \{2, 4\}$. Similarly, since the edge $\{5, 6\}$ needs to be covered, we may assume $5 \in \mathcal{K}_2$, $6 \in \mathcal{K}_5$ and $\mathcal{K}_2 \cap \mathcal{K}_5 = \emptyset$. Now we have $j \in \mathcal{K}_2$ for

$j = 3$ or $j = 4$. In either case, the edge $\{j, 5\}$ cannot be covered by the components \mathcal{K}_i , $i = 1, \dots, 5$.

At this point we can assume $\mathcal{K}_1 = \{1, 2, 3\}$, and deduce that V_1 is a component set of an OSJ cover for the complete graph on \mathcal{K}_1 , hence $V_1 = \{\{1\}, \{2\}, \{3\}\}$ and $\{\mathcal{K}_3 \vee_{[6]} \mathcal{K}_4, \mathcal{K}_4 \vee_{[6]} \mathcal{K}_5, \mathcal{K}_5 \vee_{[6]} \mathcal{K}_3\} \subset \mathcal{C}$. Now we know that $\mathcal{K}_3, \mathcal{K}_4$ and \mathcal{K}_5 are pairwise disjoint, and without loss of generality we may assume $\mathcal{K}_1 \cap \mathcal{K}_3 = \{1\}$, $\mathcal{K}_1 \cap \mathcal{K}_4 = \{2\}$ and $\mathcal{K}_1 \cap \mathcal{K}_5 = \{3\}$.

Next we argue that $\mathcal{K}_2 = \{4, 5, 6\}$. Assume that $4 \notin \mathcal{K}_2$, and note that 4 can be contained in at most one of $\mathcal{K}_3, \mathcal{K}_4$ and \mathcal{K}_5 . Without loss of generality let $4 \in \mathcal{K}_3$. This leads to a contradiction, since the edge $\{1, 4\}$ clearly cannot be covered. Arguing as above, we prove that $\{\mathcal{K}_3 \cap \mathcal{K}_2, \mathcal{K}_4 \cap \mathcal{K}_2, \mathcal{K}_5 \cap \mathcal{K}_2\} = \{\{4\}, \{5\}, \{6\}\}$. We may assume without loss of generality that $\mathcal{K}_3 = \{1, 4\}$, $\mathcal{K}_4 = \{2, 5\}$, and $\mathcal{K}_5 = \{3, 6\}$. We obtain that

$$\mathcal{C} = \{\mathcal{K}_1 \vee_{[6]} \mathcal{K}_2, \mathcal{K}_3 \vee_{[6]} \mathcal{K}_4, \mathcal{K}_4 \vee_{[6]} \mathcal{K}_5, \mathcal{K}_5 \vee_{[6]} \mathcal{K}_3\}$$

and $G(\mathcal{C}) = K_2 \cup K_3$.

Notice that an OSJ cover of K_6 constructed above is not unique, since by permuting the vertices of K_6 we can obtain a different set-join cover of K_6 .

To identify K_n with essentially unique OSJ covers, we need to extend the arguments seen in Example 4.10. To do so, we depend on a selection of observations, that we do not prove here, but that can be deduced from the proof of Theorem 4.2 given in [6, Ch. 18].

Lemma 4.11. ([9, 6]) *Let \mathcal{C} be an OSJ cover of K_n for some $n \geq 4$. Then*

$$s(n) = \min \left\{ k + s \left(\left\lceil \frac{n}{k} \right\rceil \right); k = 2, 3, 4, 5 \right\}. \quad (3)$$

Let \mathcal{M}_n be the set of all $k \in \{2, 3, 4, 5\}$ for which the minimum is achieved, and \mathcal{K}_1 a component of \mathcal{C} of maximal cardinality. Then there exists $k_0 \in \mathcal{M}_n$ so that:

1. \mathcal{K}_1 is disjoint with precisely $k_0 - 1$ components $\mathcal{K}_2, \dots, \mathcal{K}_{k_0} \in V(\mathcal{C})$.
2. The restriction of \mathcal{C} to \mathcal{K}_1 is an OSJ cover of the complete graph on \mathcal{K}_1 , and the restriction of \mathcal{C} to $\mathcal{S} = [n] \setminus \cup_{i=2}^{k_0} \mathcal{K}_i$ is an OSJ cover of the complete graph on \mathcal{S} .
3. $s(|\mathcal{S}|) = s(|\mathcal{K}_1|) = s(\lceil \frac{n}{k_0} \rceil) = s(n) - k_0$.

Proposition 4.12. *Suppose that $n \in \mathbb{N}$, $n = 3^i$ or $n = 2 \cdot 3^i$ for some $i \geq 1$. Then K_n has essentially unique OSJ cover.*

Proof. We will prove the claim by induction on i . Since we have already seen that K_3 and K_6 have essentially unique OSJ covers, the base of induction is established. We assume $i \geq 2$, and that the claim holds for $n = 3^{i-1}$ and $n = 2 \cdot 3^{i-1}$. Throughout the proof we also assume that \mathcal{C} is an OSJ cover of K_n , and $\mathcal{K}_1 \in V(\mathcal{C})$ is one of its components with maximal cardinality.

First we want to determine the parameter k_0 from Lemma 4.11. Combining (1) and (3) it is straightforward to see that for $i \geq 2$ we have $\mathcal{M}_{3^i} = \{3\}$ and $\mathcal{M}_{2 \cdot 3^i} = \{2, 3\}$. Suppose that for $n = 2 \cdot 3^i$ we have $k_0 = 3$. Applying Lemma 4.11 we get $s(|\mathcal{K}_1|) = s(|\mathcal{S}|) = s(\lceil \frac{n}{3} \rceil) = s(2 \cdot 3^{i-1}) = 3i - 1$. This restricts $4 \cdot 3^{i-2} < |\mathcal{K}_1| \leq |\mathcal{S}| \leq 2 \cdot 3^{i-1}$. From

$$\begin{aligned} |\mathcal{K}_2 \cup \mathcal{K}_3| &\leq 2|\mathcal{K}_1| \leq 4 \cdot 3^{i-1}, \\ |\mathcal{K}_2 \cup \mathcal{K}_3| &= n - |\mathcal{S}| \geq 2 \cdot 3^i - 2 \cdot 3^{i-1} = 4 \cdot 3^{i-1} \end{aligned}$$

we conclude that $|\mathcal{K}_2 \cup \mathcal{K}_3| = 4 \cdot 3^{i-1}$. Hence, $\mathcal{K}_2 \cap \mathcal{K}_3 = \emptyset$, and $|\mathcal{K}_i| = 2 \cdot 3^{i-1}$ for $i = 1, 2, 3$. Note that for $i = 1, 2, 3$, $|\mathcal{C}[\mathcal{K}_i]| = |\mathcal{C}| - 3$, hence $\mathcal{C}[\mathcal{K}_i]$ is an OSJ cover of \mathcal{K}_i , and by the induction hypothesis essentially unique. By Remark 4.6, $\mathcal{C}[\mathcal{K}_i]$ has two components of cardinality 3^{i-1} , and all other components of cardinality $2 \cdot 3^{i-2}$. Let \mathcal{K}_4 be one of the components of \mathcal{C} with the largest intersection with \mathcal{K}_1 . From $|\mathcal{C}[\mathcal{K}_i]| = |\mathcal{C}| - 3$ for $i = 2, 3$, we deduce that $\mathcal{K}_4 \cap \mathcal{K}_i \neq \emptyset$ for $i = 2, 3$. Then $|\mathcal{K}_4 \cap \mathcal{K}_1| = 3^{i-1}$ and $|\mathcal{K}_4 \cap \mathcal{K}_i| \geq 2 \cdot 3^{i-2}$ for $i = 2, 3$. It follows that $|\mathcal{K}_4| \geq 3^{i-1} + 2 \cdot 3^{i-2} + 2 \cdot 3^{i-2} > 2 \cdot 3^{i-1} = |\mathcal{K}_1|$, a contradiction with maximality of $|\mathcal{K}_1|$.

Let $n = 3^i$ for $i \geq 2$, and hence $s(n) = 3i$ and $k_0 = 3$. From Lemma 4.11 we get $s(|\mathcal{K}_1|) = s(|\mathcal{S}|) = 3(i - 1)$, implying $|\mathcal{K}_1| \leq |\mathcal{S}| \leq 3^{i-1}$. Hence,

$$\begin{aligned} |\mathcal{K}_2 \cup \mathcal{K}_3| &\leq 2|\mathcal{K}_1| \leq 2 \cdot 3^{i-1}, \\ |\mathcal{K}_2 \cup \mathcal{K}_3| &= n - |\mathcal{S}| \geq 3^i - 3^{i-1} = 2 \cdot 3^{i-1}. \end{aligned}$$

As above, we conclude that \mathcal{K}_1 , \mathcal{K}_2 , and \mathcal{K}_3 are pairwise disjoint, their union is $[n]$, and they all have $\frac{n}{3} = 3^{i-1}$ elements. Since we are allowing isomorphism of graphs, \mathcal{K}_j , $j \in [3]$, can be taken to be any three subsets of $[n]$ that satisfy those conditions. Furthermore, $\mathcal{C}[\mathcal{K}_j]$, $j \in [3]$, is an OSJ cover of the complete graph on \mathcal{K}_j and by the induction hypothesis essentially unique. Since \mathcal{K}_j , $j \in [3]$, are pairwise disjoint and $\cup_{j \in [3]} \mathcal{K}_j = [n]$ this implies essential uniqueness of \mathcal{C} .

The proof for $n = 2 \cdot 3^i$, $i \geq 2$ with $s(n) = 2 + 3i$ and $k_0 = 2$ is very similar and we leave it to the reader. \square

Remark 4.13. With arguments, akin to the ones in the proof of Proposition 4.12, it can be shown that for $n = 4 \cdot 3^i$ the graph K_n has exactly two essentially different OSJ covers. In particular, the components of any OSJ cover of K_n either satisfy $|\mathcal{K}_1| = \dots = |\mathcal{K}_4| = 2 \cdot 3^i$, $|\mathcal{K}_5| = \dots = |\mathcal{K}_{3i+4}| = 4 \cdot 3^{i-1}$ or $|\mathcal{K}_1| = \dots = |\mathcal{K}_{3i}| = 4 \cdot 3^{i-1}$, $|\mathcal{K}_{3i+1}| = \dots = |\mathcal{K}_{3i+4}| = 3^i$.

Notice that in this case two possible values of $k_0 \in \mathcal{M}_n$ can be attained, namely in the first cover we have $k_0 = 2$ and in the second cover $k_0 = 3$.

The following proposition covers the remaining cases with essentially unique OSJ cover. Only a sketch of the proof is given. The proof with all the technical details included would be rather long, and would not give significant new insights to the reader.

Proposition 4.14. *Let $n \in \mathbb{N}$ be of the form $n = 3^i - 1$ for some $i \geq 3$, or of the form $n = 2 \cdot 3^i - 1$ for some $i \geq 2$. Then K_n has essentially unique OSJ cover.*

Sketch of the proof. We claim that K_n has essentially unique OSJ cover for every n in the set

$$\mathcal{N} = \{n \in \mathbb{N}, n = 3^i - 1, i \geq 3\} \cup \{n \in \mathbb{N}, n = 2 \cdot 3^i - 1, i \geq 2\}.$$

This claim can be proved by induction, where the base of induction needs to be established for $n \in \{17, 26\}$. (This is left to the reader.) Next we fix n , and assume that the claim holds for $m \in \mathcal{N}$, with $m < n$. We assume notation and definitions from Lemma 4.11.

For $n = 3^i - 1, i \geq 3$, we have $\mathcal{M}_n = \{3\}$. Lemma 4.11 implies $s(|\mathcal{K}_1|) = s(|\mathcal{S}|) = 3(i - 1)$. From here we get:

$$|\mathcal{K}_2 \cup \mathcal{K}_3| \geq n - |\mathcal{S}| \geq 2 \cdot 3^{i-1} - 1.$$

Since $\mathcal{K}_1 \cap (\mathcal{K}_2 \cup \mathcal{K}_3) = \emptyset$ and $\mathcal{K}_1 \cup \mathcal{K}_2 \cup \mathcal{K}_3 = [3^i - 1]$, we conclude that $|\mathcal{K}_1| = 3^{i-1}$ and $|\mathcal{K}_2 \cup \mathcal{K}_3| = 2 \cdot 3^{i-1} - 1$. Note that $|\mathcal{C}[\mathcal{K}_2 \cup \mathcal{K}_3]| = |\mathcal{C}| - 1 = 3i - 1$, since $\mathcal{K}_1 \cap (\mathcal{K}_2 \cup \mathcal{K}_3) = \emptyset$. Hence $\mathcal{C}[\mathcal{K}_2 \cup \mathcal{K}_3]$ is an OSJ set-join cover of the complete graph on $\mathcal{K}_2 \cup \mathcal{K}_3$, and it is by the induction hypothesis essentially unique. By Remark 4.6 we know that its largest components are disjoint (and of order 3^{i-1} and $3^{i-1} - 1$). This implies that \mathcal{K}_2 and \mathcal{K}_3 are disjoint, $|\mathcal{K}_1| = |\mathcal{K}_2| = 3^{i-1}$, and $|\mathcal{K}_3| = 3^{i-1} - 1$. The claim is proved by noting that $\mathcal{C}[\mathcal{K}_j]$ is an essentially unique OSJ cover of the complete graph on \mathcal{K}_j for $j \in [3]$.

For $n = 2 \cdot 3^i - 1, i \geq 2$, we have $\mathcal{M}_n = \{2, 3\}$. The case $k_0 = 3$ can be excluded in a similar way, as was done in the proof of Proposition 4.12 for the case $n = 2 \cdot 3^i$. The rest of the proof of this case is very similar to the proof for the case $n = 3^i - 1$ above, and it is left to the reader.

Propositions 4.9, 4.12, and 4.14 completely decide the question of essential uniqueness of OSJ covers of K_n . The result is summarized in the theorem below.

Theorem 4.15. *Let $n \in \mathbb{N}, n \geq 2$. The graph K_n has essentially unique OSJ cover if and only if $n = 3^i$ for some $i \geq 1$, $n = 2 \cdot 3^i$ for some $i \geq 1$, $n = 3^i - 1$ for some $i \geq 3$, or $n = 2 \cdot 3^i - 1$ for some $i \geq 2$.*

Remark 4.16. Equations (1), (2) and (3) express $\text{st}_+(K_n)$ in three different ways:

$$\text{st}_+(K_n) = \begin{cases} 3i & \text{if } 2 \cdot 3^{i-1} < n \leq 3^i \\ 3i + 1 & \text{if } 3^i < n \leq 4 \cdot 3^{i-1} \\ 3i + 2 & \text{if } 4 \cdot 3^{i-1} < n \leq 2 \cdot 3^i \end{cases}$$

$$\begin{aligned}
 &= \min \left\{ \sum_{i=1}^t q_i; n \leq \prod_{i=1}^t q_i \right\} \\
 &= \min \left\{ k + \text{st}_+ \left(K_{\lceil \frac{n}{k} \rceil} \right); k = 2, 3, 4, 5 \right\}.
 \end{aligned}$$

Through those expressions we have seen how OSJ covers can be constructed recursively.

Remark 4.17. In this section, an idea of constructing set-join covers from set-join covers of smaller graphs was repeatedly used. This idea can be formalised using co-normal products of graphs. Let G and H be simple graphs (with loops). The co-normal product of G and H , denoted by $G * H$, is the graph with the vertex set $V(G) \times V(H)$, and the edge set

$$E(G * H) = \{ \{(g, h), (g', h')\}; \{g, g'\} \in E(G) \text{ or } \{h, h'\} \in E(H) \}.$$

Let \mathcal{C}_G and \mathcal{C}_H be set-join covers of G and H , respectively. Using \mathcal{C}_G and \mathcal{C}_H we can build a set-join cover \mathcal{C}' of $G * H$ as follows. For every $\mathcal{K} \in V(\mathcal{C}_G)$ define $\mathcal{K}' := \{(g, h); g \in \mathcal{K}, h \in V(H)\}$ and for every $\mathcal{L} \in V(\mathcal{C}_H)$ define $\mathcal{L}' := \{(g, h); g \in V(G), h \in \mathcal{L}\}$. Then

$$\mathcal{C}' := \{ \mathcal{K}'_1 \vee_{V(G * H)} \mathcal{K}'_2; \mathcal{K}_1 \vee_{V(G)} \mathcal{K}_2 \in \mathcal{C}_G \} \cup \{ \mathcal{L}'_1 \vee_{V(G * H)} \mathcal{L}'_2; \mathcal{L}_1 \vee_{V(H)} \mathcal{L}_2 \in \mathcal{C}_H \}$$

is a set-join cover of $G * H$. In particular,

$$\text{st}_+(G * H) \leq \text{st}_+(G) + \text{st}_+(H). \quad (4)$$

Since $K_p * K_q$ is isomorphic to K_{pq} , we see from above, that the inequality (4) can be strict (for example for $p = q = 5$), but it cannot be improved in general (it is equality for example for $p = 3^i$ and $q = 3^j$).

Remark 4.18. If $A \in \mathcal{S}_n^+$, then its Boolean rank is the NMF-rank of the pattern graph of the matrix A (see [4]):

$$\text{rk}_{01}(A) = \min \{ \text{rk}_+(B); B \in \mathcal{S}^+(G(A)) \}.$$

It follows immediately from (1) that $\text{st}_+(K_n)$ behaves asymptotically as $\text{st}_+(K_n) \sim \frac{3}{\log 3} \log n$. On the other hand, if $A \in \mathcal{S}^+(K_n)$ its Boolean rank $\text{rk}_{01}(A)$ is the minimal k such that $n \leq \binom{k}{\lfloor k/2 \rfloor}$, [4]. It follows that NMF-rank of such matrix is asymptotically bounded below by $\frac{1}{\log 2} \log n$. It means that the minimal SNT-rank and the minimal NMF-rank that can be achieved on the complete graphs have the same order of asymptotic behaviour, but the constant differs by a factor $\frac{3 \log 2}{\log 3} \approx 1.893$.

5. Concluding remarks

Finding $\text{st}_+(G)$ for a given graph G is well-defined combinatorial optimisation problem, that is of independent interest. In this section we highlight a few settings, where the theory of this paper naturally appears and provides relevant groundwork.

The most immediate application of $\text{st}_+(G)$ is the study $\text{st}_+(A)$ and SN-Trifactorizations of $A \in \mathcal{S}_n^+$. Below we offer an example how this can be done for a given matrix.

Example 5.1. The matrix

$$A = \begin{pmatrix} 4 & 1 & 1 & 4 \\ 1 & 1 & 2 & 0 \\ 1 & 2 & 0 & 3 \\ 4 & 0 & 3 & 1 \end{pmatrix}$$

has rank 3. Let $G := G(A)$, and let H be an induced subgraph of G on the set $\{2, 3, 4\}$. Since any matrix in $\mathcal{S}^+(H)$ is invertible, we have $\text{st}_+(G) \geq 3$. Note also that H has the unique OSJ cover \mathcal{C}_H , where each component in $V(\mathcal{C}_H)$ is a singleton.

The set-join cover \mathcal{C}_H can be extended to a set-join cover of G :

$$\mathcal{C} = \{\mathcal{K}_1 \vee_{[4]} \mathcal{K}_1, \mathcal{K}_1 \vee_{[4]} \mathcal{K}_2, \mathcal{K}_2 \vee_{[4]} \mathcal{K}_3, \mathcal{K}_3 \vee_{[4]} \mathcal{K}_3\},$$

where

$$2 \in \mathcal{K}_1, 3 \in \mathcal{K}_2, 4 \in \mathcal{K}_3,$$

and 1 is an element of at least two of \mathcal{K}_i , $i = 1, 2, 3$. This describes all possible set-join covers of G , and, in particular, shows that $\text{st}_+(G) = 3$.

Suppose that $\text{st}_+(A) = 3$ and let $A = BCB^T$ be an optimal SN-Trifactorization. By observations above and Theorem 2.12 the matrix B has the form

$$B = \begin{pmatrix} \mathbf{b}^T \\ D \end{pmatrix},$$

where $\mathbf{b} \in \mathbb{R}_+^3$ has at least two nonzero entries and D is an invertible nonnegative diagonal matrix. It follows that

$$A = BCB^T = \begin{pmatrix} \mathbf{b}^T \\ D \end{pmatrix} C \begin{pmatrix} \mathbf{b} & D \end{pmatrix} = \begin{pmatrix} \mathbf{b}^T C \mathbf{b} & \mathbf{b}^T C D \\ DC \mathbf{b} & DCD \end{pmatrix},$$

so

$$DCD = \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix} \text{ and } DC\mathbf{b} = \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix}.$$

Thus

$$\mathbf{b} = D(DCD)^{-1}(DC\mathbf{b}) = D \begin{pmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \\ 4 \end{pmatrix} = D \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix},$$

which means that vector \mathbf{b} has a negative entry, a contradiction. It follows that $\text{st}_+(A) = 4$.

Given an $n \times n$ nonnegative symmetric matrix A there are existing algorithms that given k find matrices $B \in \mathbb{R}_+^{n \times k}$ and $C \in \mathcal{S}_k^+$ that minimize $\|A - BCB^T\|_F$, [8]. It might be possible to extend such algorithms to find matrices $B \in \mathbb{R}_+^{n \times k}$ and $C \in \mathcal{S}_k^+$ where BCB^T has some prescribed zero-nonzero pattern. In particular, the properties proved in this paper could be used to restrict the search space.

In Remark 2.11 we have seen an interpretation of $\text{st}_+(G)$ as the minimal number of subsets of a given set V that need to be formed if we want to organize required and forbidden interactions as given by the graph G . This interpretation motivates variations of the question considered in this paper. For example, minimizing the number of meetings corresponds to finding set-join covers \mathcal{C} containing minimal number of set-joins. Posing restrictions on the size of groups would mean looking for set-join covers \mathcal{C} with restrictions on the cardinalities of the elements in $V(\mathcal{C})$. In addition to having pairs of elements that are required and pairs that are forbidden to interact, we may have pairs of elements that can (but are not required to) interact. To solve this question we would be looking for partial set-join covers of G .

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgements

Damjana Kokol Bukovšek acknowledges financial support from the ARIS (Slovenian Research and Innovation Agency, research core funding No. P1-0222).

References

- [1] Abraham Berman, Naomi Shaked-Monderer, *Completely Positive Matrices*, World Scientific Publishing Co., Inc., River Edge, NJ, 2003. MR 1986666.

- [2] Richard A. Brualdi, Herbert J. Ryser, *Combinatorial Matrix Theory*, paperback ed., *Encyclopedia of Mathematics and Its Applications*, vol. 39, Cambridge University Press, New York, 2013. MR 3469704.
- [3] Mao Cheng Cai, On a problem of Katona on minimal completely separating systems with restrictions, *Discrete Math.* 48 (1) (1984) 121–123. MR 732208.
- [4] Nicolas Gillis, *Nonnegative Matrix Factorization*, SIAM, 2020.
- [5] Frank J. Hall, Zhongshan Li, Di Wang, Symmetric sign pattern matrices that require unique inertia, *Linear Algebra Appl.* 338 (2001) 153–169. MR 1861120.
- [6] Ross Honsberger, *Mathematical Gems. III*, *The Dolciani Mathematical Expositions*, vol. 9, Mathematical Association of America, Washington, DC, 1985. MR 801393.
- [7] Damjana Kokol Bukovšek, Helena Šmigoc, Symmetric nonnegative matrix trifactorization, *Linear Algebra Appl.* 665 (2023) 36–60. MR 4547308.
- [8] Lu Lin, Zhong-Yun Liu, An alternating projected gradient algorithm for nonnegative matrix factorization, *Appl. Math. Comput.* 217 (24) (2011) 9997–10002. MR 2806386.
- [9] Andrew Chi, Chih Yao, On a problem of Katona on minimal separating systems, *Discrete Math.* 15 (2) (1976) 193–199. MR 422036.