Complexity aspects of the computation of the rank of a graph

Igor da Fonseca Ramos, Vinícius F. Dos Santos, Jayme L. Szwarcfiter

Complexity aspects of the computation of the rank of a graph

Igor da Fonseca Ramos[1][*] Vinícius F. dos Santos[2][†] Jayme L. Szwarcfiter[3][‡]

[1]*DCC/IM, PESC/COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*
[2]*Departamento de Matemática Aplicada, IME, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil*
[3]*DCC/IM, NCE, PESC/COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil*

We consider the P_3-convexity on simple undirected graphs, in which a set of vertices S is convex if no vertex outside S has two or more neighbors in S. The convex hull $H(S)$ of a set S is the smallest convex set containing S as a subset. A set S is a convexly independent set if $v \notin H(S \setminus \{v\})$ for all v in S. The rank $rk(G)$ of a graph is the size of the largest convexly independent set.

In this paper we consider the complexity of determining $rk(G)$. We show that the problem is NP-complete even for split or bipartite graphs with small diameter. We also show how to determine $rk(G)$ in polynomial time for the well structured classes of graphs of trees and threshold graphs. Finally, we give a tight upper bound for $rk(G)$, which in turn gives a tight upper bound for the Radon number as byproduct, which is the same obtained before by Henning, Rautenbach and Schäfer. Additionally, we briefly show that the problem is NP-complete also in the monophonic convexity.

Keywords: algorithms, complexity, graph convexity, P_3-convexity, rank

1 Introduction

When considering the spread of an infection, information or advertisement over a network, a natural model consists of starting with an initial set of nodes to which new ones are added whenever enough neighbors are already inside the set [JR09]. Such a model is strongly related with abstract convexities [vdV93]. In particular, if a vertex gets infected when at least two of its neighbors already are, we have a strong connection with the so-called P_3-convexity.

We consider the problem of finding the largest set such that none of the elements is infected by the others. More formally, for a graph G and a subset $S \subseteq V(G)$, the *hull* of S, denoted $H(S)$ is the smallest convex set containing S. We say that S is *convexly independent* if $v \notin H(S \setminus \{v\})$, for each $v \in S$, and *convexly dependent* otherwise. The *rank* of G, denoted by $rk(G)$ is the cardinality of the largest convexly independent set of G.

Recently, complexity aspects of the P_3-convexity have been considered extensively for many convexity parameters such as the Hull number [BRdSS13, CDP+11, CPRPdS13], the Radon number [DRdS+13b, DRdS+12, DRdSS12, HRS13] and the Carathéodory number [BCD+12, DRdS+13a] as well as for the partition problem [CDD+10]. In this paper we are interested in the complexity of finding the rank of G, in

the P_3-convexity. We start by giving some definitions in Section 2. In Section 3 we show that the problem is NP-complete even if the input graph is restricted to some well studied graph classes. Still, we are able to show how to find the rank in polynomial time for some restricted cases, as shown in Section 4. Finally, motivated by the hardness of the problem, we give a tight upper bound for the general case in Section 5. Additionally, in Section 6 we show that the problem is NP-complete also in the context of the monophonic convexity.

2   Definitions

Before we proceed with the results, we begin by recalling some definitions. We consider finite, simple and undirected graphs. For any graph G, let $V(G)$ denote its vertex set and $E(G)$ its edge set. For a vertex $v \in V(G)$, we call $d(v)$ the *degree* of v, i.e., the number of edges adjacent to v. Also, for a pair of vertices $u, v \in V(G)$, we denote their distance by $dist(u, v)$. The *diameter* of a graph G is the greatest distance between two of its vertices. $\delta(G)$ and $\Delta(G)$ represent, respectively the smallest and the greatest degree of a vertex that belongs to G.

We are mostly concerned with the convexity of paths of order three on simple undirected graphs, unless we clearly state otherwise. The convex hull of a set S is the smallest convex set that contains S and is represented by $H(S)$. We can think about the computation of the hull iteratively. Initially $H(S) = S$ and, on each step, a vertex with at least two neighbors inside $H(S)$ is added to $H(S)$. This computation stops when no such vertex exists. Because of this, sometimes we say that a vertex that belongs to $H(S)$ is *generated* by S.

A set R is called a *Radon set* if it can be partitioned in two subsets $R = R_1 \cup R_2$ with intersecting convex hulls, that is, $H(R_1) \cap H(R_2) \neq \emptyset$. The cardinality of the largest Radon set of a graph G is called its *Radon number* and is denoted by $r(G)$. Also, any set $S \subseteq V(G)$ that generates the whole graph, *i.e*, $H(S) = V(G)$ is called a *hull set* and we call *hull number* the size of the smallest set that satisfies such property. We represent the hull number of a graph G by $h(G)$.

A *convexly independent* set $S \subseteq V(G)$ is a set such that, for all $v \in S$, $v \notin H(S \setminus \{v\})$. Otherwise, S is called *convexly dependent*. We can say that a set is convexly independent if none of its vertices is generated by the others. The *rank* of a graph is the cardinality of its maximum convexly independent set, denoted by $rk(G)$. We call the problem of determining the rank of a graph the MAXIMUM CONVEXLY INDEPENDENT SET problem. Its decision version is the main object of study of this paper and is given below.

PROBLEM: MAXIMUM CONVEXLY INDEPENDENT SET
INPUT: A graph G and an integer k
QUESTION: Is $rk(G) \geq k$?

An Open Packing of a graph G is a set $S \subseteq V(G)$ such that no pair of its vertices have intersecting open neighborhoods. This concept was introduced and proved to be NP-hard to determine even on chordal graphs by Henning and Slater [HS99]. The cardinality of the largest set with this property is called the open packing number and is represented by $\rho_o(G)$. The decision version of the OPEN PACKING NUMBER problem is stated as follows.

PROBLEM: OPEN PACKING NUMBER
INPUT: A graph G and an integer k
QUESTION: Is $\rho_o(G) \geq k$?

A *split graph* G is one whose vertex set admits a partition $V(G) = C \cup I$ into a clique C and an independent set I.

A *clique separator* is a subset of vertices that induces a complete graph whose removal disconnects the graph. An *atom* G is a graph without a clique separator.

Finally, a *bipartite graph* G is one whose vertex set can be partitioned in two sets A and B such that $A \cup B = V(G)$, $A \cap B = \emptyset$ and there is no edge from a vertex of A to a vertex of B.

3   Hardness Results

3.1   *Split graphs*

We show that computing the rank of a split graph with $\delta(G) \geq 2$, in the context of the P_3-convexity, is already NP-hard.

The following simple result is employed.

Lemma 1 *Let G be a graph, C a complete subset of it, and v_1, v_2 distinct vertices of C. Then $H(v_1, v_2) \supseteq C$.*

We can now state the NP-completeness.

Theorem 1 MAXIMUM CONVEXLY INDEPENDENT SET *is NP-complete, even for split graphs with $\delta(G) \geq 2$.*

Proof: Since the hull of a subset of vertices can be computed in polynomial time for any graph, the problem clearly lies in NP. For the reduction, we employ the MAXIMUM SUBSET PACKING problem, which is known to be NP-complete [Kar72]. The latter problem has as input a family $\mathcal{S}'$ of subsets $S_i' \in \mathcal{S}'$ of some ground set, together with an integer k'. The question is whether $\mathcal{S}'$ contains k' or more mutually disjoint subsets S_i'. Given $\mathcal{S}'$ and k', we construct the following instance of MAXIMUM CONVEXLY INDEPENDENT SET. The elements of the ground set $\cup S_i'$ of the subsets $S_i' \in \mathcal{S}'$ are all vertices of the graph G. Besides, G contains also a pair of new distinguished vertices w_i, z_i, for each subset $S_i' \in \mathcal{S}'$. The edges of G are as follows. The set of vertices of G corresponding to the ground set of $\mathcal{S}'$, together with the set of all vertices z_i form a clique C of G. In addition, for each of the distinguished vertices w_i, add an edge $w_i z_i$ and edges $w_i v$, for each $v \in C$ that corresponds to an element of the subset $S_i' \in \mathcal{S}'$. Finally, define $k = k'$. The construction of the input to the MAXIMUM CONVEXLY INDEPENDENT SET problem is completed. Observe that the set of vertices $w_i \in V(G)$ form an independent set I of G, and therefore G is a split graph with bipartition $V(G) = C \cup I$. Finally, without loss of generality, we restrict to values $k = k' \geq 3$.

We prove that $\mathcal{S}'$ contains k' mutually disjoint subsets if and only if G has a convexly independent set of size k,

Suppose that $\mathcal{S}'$ contains k' mutually disjoint subsets $S_1', S_2', \ldots, S_{k'}'$. We show that the subset $\{w_1, w_2, \ldots, w_k\}$ of $V(G)$ is a convexly independent set of G. Because $S_i' \cap S_j' = \emptyset$, for $i \neq j$, it follows that $N_G[w_i] \cap N_G[w_j] = \emptyset$, for $i \neq j$. Consequently, $dist_G(w_i, w_j) > 2$, implying that $H(\{w_i, w_j\}) = \{w_i, w_j\}$, for any $1 \leq i < j \leq k$. The latter means $\{w_1, \ldots, w_k\}$ is convexly independent.

Conversely, by hypothesis G has a convexly independent set S of size k. First, we show that $S \subseteq I$. If S contains three vertices $v_1, v_2, v_3 \in C$, by Lemma 1, $H(\{v_1, v_2\}) \supseteq C$ and therefore $H(S) = H(S \setminus \{v_3\})$, contradiction. If S contains exactly two vertices of C then $|S| \geq 3$ implies that there is some $w_i \in S \cap I$. Again, Lemma 1 and the fact that $|N(w_i)| \geq 2$ imply that $w_i \in H(\{v_1, v_2\})$, hence S must be convexly dependent. Finally, if S contains exactly one vertex $v_1 \in C$ then there are at least two distinct vertices $w_i, w_j \in S \cap I$. In this case, because $N(w_i) \geq 2$, we conclude that w_i has a neighbor v_2, distinct from v_1. That is, $v_2 \in H(\{v_1, w_1\})$, and using Lemma 1, we conclude that $w_2 \in H(\{v_1, w_i\})$, implying that $\{v_1, w_1, w_2\}$ is already convexly dependent, a contradiction. Consequently, it is true that $S \subseteq I$.

Finally, choose a pair of arbitrary distinct vertices $w_i, w_j \in S$, and examine $N(w_i) \cap N(w_j)$. We show that they are disjoint. Suppose the contrary, and let $v_1 \in N(w_i) \cap N(w_j)$. Because $|S| \geq 3$, there is

$w_p \in S$, such that $w_p \neq w_i, w_j$. Then $v_1 \in H(\{w_i, w_j\})$. Consequently $z_i \in H(\{w_i, w_j\})$, meaning that $C \subseteq H(\{w_i, w_j\})$ and finally $w_p \in H(\{w_i, w_j\})$. Consequently, $N(w_i) \cap N(w_j) = \emptyset$, for all distinct $w_i, w_j \in S$. Since $N(w_i)$ and $N(w_j)$ respectively, are exactly the subsets S_i' and S_j' of $\mathcal{S}'$, we conclude that $\mathcal{S}'$ contains k' mutually disjoint subsets, completing the proof. □

Corollary 1 *The* OPEN PACKING NUMBER *problem is NP-complete for split graphs with $\delta(G) \geq 2$.*

The above corollary strengthens a result of Henning and Slater [HS99], which stated that the OPEN PACKING NUMBER was NP-complete for chordal graphs, and is simpler to understand with the aid of the following lemmas.

Lemma 2 *Let G be a split graph with $\delta(G) \geq 2$ and split partition $V(G) = C \cup I$. If a set $S \subsetneq V(G)$ with $|S| \geq 3$ is convexly independent, then $S \subseteq I$.*

Proof: Suppose $S \subsetneq V(G)$ is a convexly independent set with $|S| \geq 3$ and that there is some $v \in S \cap C$. Consider any $u \in S$ such that $u \neq v$. If $u \in C$, all other vertices in C and, therefore, the whole set of vertices of G is generated by u and v alone. If $u \in I$, u and v have a common neighbor in C that is added to the convex hull and, together with v, also generates the whole graph. In both cases a third vertex of S is generated, which means S is convexly dependent. □

Lemma 3 *Let G be a split graph with $\delta(G) \geq 2$. A set $S \subsetneq V(G)$ sucht that $|S| \geq 3$ is convexly independent if and only if $H(S) = S$. Moreover, $V(G)$ is not convexly independent.*

Proof: Let G be a split graph with $\delta(G) \geq 2$ and split partition $V(G) = C \cup I$ in which C is a clique and I is an independent set. Since, by Lemma 2, no vertex of C can be in a convexly independent set, $V(G)$ is convexly dependent.

Now let $S \subsetneq V(G)$ be a convexly independent set with $|S| \geq 3$. Suppose that vertices $u, v \in S$ generate $w \notin S$. Since $S \subseteq I$ and I is stable, w must be in C, as it is a common neighbor of u and v. Since either u or v, together with w can generate the whole graph, a third vertex of S is present in $H(\{u, v\})$, hence S is not convexly dependent.

Conversely, suppose there is some $S \subsetneq V(G)$ such that $H(S) = S$ and S is convexly dependent. Since S is not convexly independent, there is some $v \in S$ that has at least two neighbors in S. That implies the existence of $u \in C \cap S$, for I is stable. This is a contradiction with Lemma 2 and, therefore, no such set can exist. □

We can now easily prove Corollary 1.

Proof: Let G be a split graph with $\delta(G) \geq 2$. Lemma 3 shows that, in G, every convexly independent set is also an open packing. Besides, every open packing is clearly a convexly independent set. This means that knowing the open packing number of G is the same as knowing the rank of said graph, thus the OPEN PACKING NUMBER problem is NP-complete for split graphs with $\delta(G) \geq 2$. □

3.2 Bipartite graphs with diameter at most 3

We prove that computing the rank of a bipartite graph with diameter at most 3, in the context of the P_3-convexity, is NP-hard.

Theorem 2 MAXIMUM CONVEXLY INDEPENDENT SET *is NP-complete, even for bipartite graphs with diameter at most three.*

Proof: As stated in the proof of Theorem 1, the problem lies in NP. We present a reduction from the MAXIMUM CONVEXLY INDEPENDENT SET problem for split graphs with minimum degree at least 2, which we have already proved to be NP-complete. As stated before, this problem receives a graph G and an integer k as input and decides whether or not $rk(G)$ is greater than k.

Given a split graph G' with $\delta(G') \geq 2$ and $k' \geq 4$, let $k = k'$. We will construct a bipartite graph G as follows. Let $V(G') = C' \cup I'$ be a split partition of the vertices in $V(G')$ in which C' is a clique and I' is an independent set. Now let $V(G) = C \cup I$, with $C = C' \cup \{y\}$ and $I = I' \cup \{x_1, x_2\}$. For every two adjacent vertices of G' that are not both in C', add an edge between them in G. Also add edges from x_1 and x_2 to all vertices in C and from y to the vertices in I. For simplicity, we will call x_1, x_2 and y *universal vertices*. Notice that G is bipartite and that the distance between any two of its vertices is not greater than 3, since it is always possible to use the universal vertices to form a path with at most 3 edges from any vertex to any other.

We prove that G' has a convexly independent set of size greater or equal to k' if and only if G has a convexly independent set of size greater or equal to k.

Suppose there is a convexly independent set $S' \subseteq V(G')$ with $|S'| \geq k'$. Let $S = S'$. Since $|S| = |S'|$ and $k = k'$, we know that $|S| \geq k$. By Lemma 2, we know that $S \subseteq I$ and, by Lemma 3, that no vertex in C but y is adjacent to more than one vertex in S. From that we can conclude that $H(S) = S \cup \{y\}$. Since all vertices in S only have neighbors in C and there is only one vertex of C in $H(S)$, no $v \in S$ is generated by a subset of S and, therefore, S is a convexly independent set with $|S| \geq k$.

Conversely, suppose that $S \subseteq V(G)$ is a convexly independent set of G with $|S| \geq k$ and let $S' = S$. We show that a S' is convexly independent in G'. In order to prove this, we show that $S \subseteq I$, that $x_1, x_2 \notin S$ and that no vertex in C but y is generated by S, meaning that $H(S') = S'$, which suffices to prove that S' is convexly independent as stated on Lemma 3.

First, notice that any two vertices of C generate the whole graph, because they add x_1 and x_2 to the convex hull which, in turn, generate all other vertices in C. That leads to the convex hull being $V(G)$. Since $|S| \geq 4$, there is some vertex in S that is generated by these two and, therefore, S would not be convexly independent. The same applies if there is exactly one vertex from C in S. Any pair of vertices of I generates y, thus y cannot be in S, and y, together with any other vertex in C, generates the whole graph which means the set is not convexly independent.

Now suppose that x_1 is in S. Let $v \in I$ be any other vertex from S. Since x_1 is an universal vertex, all the neighborhood of v is generated and, since $\delta(G') \geq 2$ and the neighborhood of the vertices in I is preserved, we now have 2 vertices in C which, in turn, generate the whole graph. S could not be convexly independent if that was the case and, therefore, x_1 is not in S. The same argument applies to x_2.

We must show, at last, that no vertex in C but y is generated. It is clear that it could not happen, though, for it would lead to $H(S) = V(G)$. Since the adjacency of the vertices in I' is preserved when we create I and C, we know that $H(S') = S'$ and, by Lemma 3, S' is convexly independent and $|S'| \geq k'$, which concludes the proof. □

4 Polynomial time algorithms

4.1 Threshold graphs

Since the problem of finding the rank of a graph G is NP-hard even for split graphs, we may ask about the complexity of calculating this parameter for a threshold graph.

Theorem 3 *If G is a connected threshold graph with $|V(G)| \geq 3$ and $D \subseteq V(G)$ is the set with all $v \in V(G)$ such that $d(v) = \delta(G)$, then:*

(i) *if G is a star, then $rk(G) = |V(G)| - 1$;*

(ii) *otherwise, if $\delta(G) = 1$, then $rk(G) = |D| + 1$;*

(iii) *otherwise, $rk(G) = 2$.*

Proof: In (i), it is clear that D itself is a maximum convexly independent set.

Moreover, (iii) is the case in which $d(v) > 1$ for all $v \in D$. That means that, no matter the choice of vertices, the nested neighborhood of vertices that characterizes threshold graphs implies that at least two vertices of the clique will be in the convex hull of any $S \subsetneq V(G)$ such that $|S| = 2$. The two vertices in the clique, in turn, will generate the whole graph. This implies that there is no third vertex that we could add to S without making it convexly dependent and, therefore, $rk(G) = 2$.

In (ii), if we take two vertices $u, v \in V(G) \setminus D$ and define $S = \{u, v\}$, we will have $H(S) = V(G) \setminus D$. The only candidates to enter S without making it convexly dependent, then, are the vertices in D, but any one of them, together with one of u or v, would generate the other vertex in S; therefore this set cannot be made any larger. From this we can understand that no convexly independent set of cardinality greater than two can be made with two or more vertices from $V(G) \setminus D$. If we take all vertices in D into a set S, since G is a threshold graph, we know that only one vertex of G, besides those in D, is in $H(S)$. We can add to the set, then, any other vertex in the graph, for it was not yet generated by the others and no vertex in D can be generated by the remaining elements of S. Since this addition makes $H(S) = V(G)$ and $|D| + 1 \geq 2$, we know that $rk(G) = |D| + 1$. $\square$

Since we can determine the degree of all vertices in $O(|E|)$ time and $|D|$ can be computed in $O(|V|)$ time given we have the vertices degrees, it is possible to solve the MAXIMUM CONVEXLY INDEPENDENT SET in $O(|E| + |V|)$ time for threshold graphs.

4.2 Trees

Let T be a tree rooted in $r \in V(T)$. We denote the subtree with root in a vertex $v \in V(T)$ as T_v.

We use a dynamic programming algorithm to find a maximum convexly independent set S^* in polynomial time for T.

We say that a vertex $u \in V(T)$ sends one unity of charge to $v \in V(T)$ if and only if $u \in H_{T-v}(S \setminus \{v\})$ and $v \in N(u)$, i.e., although u and v are neighbors, u does not depend on v to be in the convex hull of $S \setminus \{v\}$. The total amount of charge received by v with respect to S is represented as $ch(v)$, that is, $ch(v) = |N(v) \cap H_{T-v}(S \setminus \{v\})|$.

Lemma 4 *Let $S \subseteq V(T)$ be a convexly independent set in T. $v \in V(T)$ is generated by $S \setminus \{v\}$ if and only if $ch(v) \geq 2$.*

Proof: If $ch(v) \geq 2$, then v has at least two neighbors, say u_1 and u_2, such that $u_1, u_2 \in H_{T-v}(S \setminus \{v\})$. Since adding a vertex to a graph does not remove any other from the convex hull of a set in the P_3-convexity, we can put v back into $T - v$ and still have $u_1, u_2 \in H_T(S \setminus \{v\})$. Therefore, v is generated by $S \setminus \{v\}$.

Conversely, if v is generated by $S \setminus \{v\}$, that means that there are $u_1, u_2 \in N(v)$ such that $u_1, u_2 \in H_T(S \setminus \{v\})$ and that this does not depend on v also being in $H_T(S \setminus \{v\})$. If, by removing v from T, u_i is not in $H_{T-v}(S \setminus \{v\})$, that means that u_i has exactly two neighbors in T that are also in $H_T(S \setminus \{v\})$ and v is necessarily one of them, which is a contradiction. We know, then, that $u_1, u_2 \in H_{T-v}(S \setminus \{v\})$ and, since $u_1, u_2 \in N(v)$, $ch(v) \geq 2$, which completes the proof. $\square$

Corollary 2 *S is convexly independent if and only if there is no $v \in S$ such that $ch(v) \geq 2$.*

In order to describe the algorithm, we must introduce the recurrence relation. Let us define $P_v(i, j, k)$, the *contribution* of v, as the size of the maximum convexly independent set using only vertices from the subtree rooted in v in the state defined by i, j and k as stated below. In case $P_v(i, j, k)$ is not well defined, we say that v's contribution is $-\infty$, so this value cannot be misused in other calculations.

- $i = 1$ means that v receives 1 unity of charge from its parent, while $i = 0$ means it does not.

- $j = 1$ means that v is part of the convexly independent set, while $j = 0$ means the opposite.

- k is the amount of charge that v receives from its children.

Notice that, when computing a certain $P_v(i, j, k)$, $ch(v) = i + k$.

Say that p_v is the parent of v and that $N'(v) = N(v) \setminus \{p_v\}$. Also, let us define the following functions:

$$f(v, i) = \max\{P_v(i, 0, 0), P_v(i, 0, 1)\} \tag{1}$$
$$h(v, i) = \max\{\max_{2 \le k < d(v)}\{P(i, 0, k)\}, \max_{0 \le k \le d(v)} P_v(i, 1, k)\} \tag{2}$$
$$g(v, i_1, i_2) = h(v, i_1) - f(v, i_2) \tag{3}$$

Since, in (1), $ch(v) < 2$ or, else, $ch(v) = 2$ and v depends on p_v to be in $H(S)$, we can be sure that v sends no charge to its parent. If $j = 1$ or $k \ge 2$, v will always be in $H(S)$ independently of the charges that may come from p_v. The function $f(v, i)$, then, determines the maximum contribution of v while sending no charge to its parent.

The function $h(v, i)$, comprising all the possibilities for j and k not covered by f, represents the maximum contribution of v when it must send a unity of charge to its parent.

Moreover, the function $g(v, i_1, i_2)$ is the gain, in terms of the contribution, of forcing a vertex v that is not sending charge to its parent to do so.

We can now determine the value of $P_v(i, j, k)$ for any vertex $v \in V(T)$.

$$P_v(0, 0, 0) = \sum_{u \in N'(v)} f(u, 0); \tag{4}$$

$$P_v(0, 0, 1) = \begin{cases} -\infty, & \text{if } v \text{ has no child,} \\ \sum_{u \in N'(v)} f(u, 0) + \max_{u \in N'(v)} g(u, 0, 0), & \text{otherwise;} \end{cases} \tag{5}$$

$$P_v(0, 0, 2) = \begin{cases} -\infty, & \text{if } v \text{ has less than 2 children,} \\ \sum_{u \in N'(v)} f(u, 1) + \max_{\substack{\forall X \subseteq N'(v) \\ |X| = 2}} \sum_{u \in X} g(u, 0, 1), & \text{otherwise;} \end{cases} \tag{6}$$

$$P_v(0,0,k) = \begin{cases} -\infty, & \text{if } v \text{ has less than } k \text{ children,} \\ \displaystyle\sum_{u \in N'(v)} f(u,1) + \max_{\substack{\forall X \subseteq N'(v) \\ |X|=k}} \sum_{u \in X} g(u,1,1), & \text{otherwise;} \end{cases} \tag{7}$$

$$P_v(0,1,0) = \sum_{u \in N'(v)} f(u,1) + 1; \tag{8}$$

$$P_v(0,1,1) = \begin{cases} -\infty, & \text{if } v \text{ has no child,} \\ \displaystyle\sum_{u \in N'(v)} f(u,1) + \max_{u \in N'(v)} g(u,1,1) + 1, & \text{otherwise;} \end{cases} \tag{9}$$

$$P_v(0,1,k) = -\infty; \tag{10}$$

$$P_v(1,0,0) = \begin{cases} -\infty, & \text{if } v = r, \\ \displaystyle\sum_{u \in N'(v)} f(u,0), & \text{otherwise;} \end{cases} \tag{11}$$

$$P_v(1,0,1) = \begin{cases} -\infty, & \text{if } v \text{ has no child or } v = r, \\ \displaystyle\sum_{u \in N'(v)} f(u,1) + \max_{u \in N'(v)} g(u,0,1), & \text{otherwise;} \end{cases} \tag{12}$$

$$P_v(1,0,k) = \begin{cases} -\infty, & \text{if } v \text{ has less than } k \text{ children or } v = r, \\ \displaystyle\sum_{u \in N'(v)} f(u,1) + \max_{\substack{\forall S \subseteq N'(v) \\ |S|=k}} \sum_{u \in S} g(u,1,1), & \text{otherwise;} \end{cases}$$

$$\tag{13}$$

$$P_v(1,1,0) = \begin{cases} -\infty, & \text{if } v = r, \\ \displaystyle\sum_{u \in N'(v)} f(u,1) + 1, & \text{otherwise;} \end{cases} \tag{14}$$

$$P_v(1,1,k) = -\infty. \tag{15}$$

Theorem 4 *The previous recurrence relation computes $P_v(i, j, k)$ correctly and this calculation can be done in polynomial time.*

Proof: We prove the correctness of the recurrence relation above by induction. Let us consider a vertex $v \in V(T)$ and all the possibilities for i, j and k.

In case v is a leaf of T_r, k must be 0, otherwise the contribution is said to be $-\infty$. There are four cases in which this occurs: $P_v(0, 0, 0)$, $P_v(0, 1, 0)$, $P_v(1, 0, 0)$ and $P_v(1, 1, 0)$. Since v is a leaf, we know that its only neighbor is p_v, and that means that $P_v(i, j, 0)$ is always a valid contribution. In this case, it can be seen in (4), (8), (11) and (14) that $P_v(i, 0, 0) = \sum_{u \in N'(v)} f(u, 0)$ and $P_v(i, 1, 0) = \sum_{u \in N'(v)} f(u, 1) + 1$. Since all sums amount to zero, we have $P_v(i, 0, 0) = 0$ and $P_v(i, 1, 0) = 1$, which is correct, as the maximum contribution of T_v can only be 0 when $v \notin S$ and 1 otherwise.

For all other vertices, we must make some further analysis.

If $i = 0$, $j = 0$ and $k = 0$, we want the maximum number of nodes in a convexly independent set that are also in T_v, given that v receives no charge either from p_v or from its children. For that, each $u \in N'(v)$ must give the maximum contribution without sending any charges to v, that is,

$$P_v(0, 0, 0) = \sum_{u \in N'(v)} f(u, 0).$$

If we change i to 1 while the values of j and k remain the same, the idea is absolutely the same, though we must consider the case in which v is the root and, therefore, has no parent that could send it any charge.

When we consider $i = 0$, $j = 0$ and $k = 1$, it is necessary to take into account that some $u \in N'(v)$ must send a charge to v and, therefore, its maximum contribution is given by $h(u, 0)$ instead of $f(u, 0)$. Since v is not in $H(S)$ and since we must choose u to be the child that maximizes the overall contribution of v, we can write

$$P_v(0, 0, 1) = \max_{u \in N'(v)} \left(\sum_{w \in N'(v) \setminus \{u\}} f(w, 0) + h(u, 0) \right).$$

We can, however, add $f(u, 0) - f(u, 0)$ to $P_v(0, 0, 1)$ without changing the result, which ultimately leads us to

$$P_v(0, 0, 1) = \sum_{u \in N'(v)} f(u, 0) + \max_{u \in N'(v)} g(u, 0, 0).$$

Considering any other case in which $j = 0$ and $k > 0$ we have the same situation. We must only be aware that the value of $ch(v)$ influences the amount of charge v sends to its children and, therefore, the parameter i given to functions f and g. If $ch(v) < 2$, then v does not send any charge to its children. However, if $ch(v) = 2$, then v sends charge to all its children, except for the ones that send a unity of charge to v. This implies the use of f with $i = 1$ and the use of g with $i_1 = 0$ and $i_2 = 1$. Furthermore, if $ch(v) > 2$, then v does not depend specifically on any of its neighbors to be in $H(S)$ and, therefore, both f must be used with $i = 1$ and g with $i_1 = i_2 = 1$. To choose which of v's children will send charges to v, we can also use the same argument and simply add the k greatest values of g to $P_v(i, 0, 0)$.

At last, when $j = 1$ and $k > 0$, there are two possibilities. If $ch(v) \geq 2$, we know that we cannot make a convexly independent set and, therefore, $P_v(i, 1, k) = -\infty$ whenever $i + k \geq 2$. If $ch(v) < 2$, however, v is allowed into S, though it will never depend on any neighbor to be in $H(S)$ and will always send charge to them. Also, it must be added to the count of the vertices in S that belong to T_v. This means that we can still use the same formulas above, but every time f or g are used i or i_2 must be 1 and we must add 1 to each $P_v(i, 1, k)$.

Considering the time to compute $P_v(i, j, k)$, in the worst case, it will be necessary to calculate f and g for all vertices in $V(T)$. We may assume we must compute $f(v, 0)$, $f(v, 1)$, $h(v, 0)$ and $h(v, 1)$ for all $v \in V(T)$. The function f can be evaluated in $O(1)$ given we have the corresponding $P_v(i, 0, 0)$ and $P_v(i, 0, 1)$. The function h requires $O(\Delta(G))$ for each vertex, but, since the graph is a tree, the sum of the degrees of all vertices is $O(|V|)$, so the time required to evaluate $h(v, 0)$ and $h(v, 1)$ for all $v \in V(T)$ if we have all required $P_v(i, j, k)$ is $O(|V|)$. To evaluate each $P_v(i, j, k)$ it may also be necessary to choose the k greatest values of $g(u, i_1, i_2)$ for all $u \in N'(v)$. For each $v \in V(T)$, there are $O(d(v))$ different values for k and, by sorting the values of g in non-increasing order and using accumulated sums, this can be done in $O(d(v) \log d(v)) \leq O(d(v) \log \Delta(T))$. The total time complexity for all vertices is, thus, $\sum_{v \in V(T)} O(d(v) \log \Delta(T))$, which, in turn, can be rewritten as $O(\log \Delta(T)) \sum_{v \in V(T)} O(d(v))$. Replacing the sum of the degrees of all vertices of T by $O(|V|)$, the total time to compute all $P_v(i, j, k)$ is $O(|V| \log \Delta(T))$.   □

Corollary 3 $|S^*|$ *is given by*

$$\max_{j \in \{0,1\}, 0 \leq k \leq d(r)} P_r(0, j, k).$$

Moreover, it can be obtained in $O(|V| \log \Delta(T))$ time.

Proof: Since $P_v(i, j, k)$ is the maximum number of vertices in the subtree with root in v that can be in a convexly independent set under the conditions imposed by i, j and k, when we have $r = v$, we have the size of a maximum convexly independent set for the whole graph, that is, the rank of T. Also, since we can compute all $P_v(i, j, k)$ in $O(|V| \log \Delta(T))$ time, it is possible to obtain the rank of T in $O(|V| \log \Delta(T))$ time.   □

5   A general upper bound

We remark that every minimum hull set is a convexly independent set, hence $h(G) \leq rk(G)$. A *Radon partition* of a set of vertices R is a partition of R into two disjoint sets R_1 and R_2 such that $H(R_1) \cap H(R_2) \neq \emptyset$. The set R is *Radon independent* if it has no Radon partition. Since every Radon independent set is a convexly independent set and the Radon number is exactly the size of the largest Radon independent set plus one, we also have $r(G) - 1 \leq rk(G)$. It is known [HRS13] that $r(G) - 1 \leq \frac{2n}{\delta(G)+1}$. We now show that in fact this is also an upper bound for the rank of a graph getting a simpler proof of the bound in [HRS13] as a byproduct. Note that this also shows that the rank of a graph can be used as a tighter bound for the Radon number, since $r(G) - 1 \leq rk(G) \leq \frac{2n}{\delta(G)+1}$.

Theorem 5 *Let G be a graph with minimum degree $\delta(G)$. Then $rk(G) \leq \frac{2n}{\delta(G)+1}$. Moreover, this bound is tight.*

Proof: First note that for complete graphs the equality holds, showing that the bound is tight.

Let $S \subseteq V(G)$ be a convexly independent set. We define S_i as the subset of S of vertices containing exactly i neighbors in S and R_j as the subset of vertices of $R = V(G) \setminus S$ containing exactly j neighbors in S. For notational simplicity we also define $r_i = |R_i|$ and $s_i = |S_i|$.

First note that from the independence we have $S_i = \emptyset$ for $i \geq 2$, hence $S = S_0 \cup S_1$. Moreover, note that no vertex of S_1 can have a neighbor in $\cup_{i \geq 3} R_i$ and that, if a vertex $v \in S_0$ has more than one neighbor in $\cup_{i \geq 3} R_i$, then $v \in H(S \setminus \{v\})$. Also note that the sets S_i and R_j define a partition of $V(G)$ and that $\sum_{v \in S} |N(v) \cap R| = \sum_{w \in R} |N(w) \cap S|$.

It follows from the previous observations that the optimal solution of the following linear program is an upper bound for $rk(G)$, since any convexly independent set must satisfy all the linear constraints of (P).

$$(P) \begin{cases} \max & s_0 + s_1 \\ & s_0 + s_1 + \sum_{i=0}^{\Delta(G)} r_i & = & n \\ & \sum_{i=3}^{\Delta(G)} i r_i & \leq & s_0 \\ & \sum_{v \in S_0} d(v) + \sum_{w \in S_1} (d(v) - 1) & = & \sum_{i=0}^{\Delta(G)} i r_i \\ & s_i, r_i & \in & \mathbb{N}_0. \end{cases}$$

From the last two constraints we have

$$\sum_{v \in S_0} d(v) + \sum_{w \in S_1} (d(v) - 1) \leq r_1 + 2r_2 + s_0$$

$$\sum_{v \in S_0} (d(v) - 1) + \sum_{w \in S_1} (d(v) - 1) \leq r_1 + 2r_2 \leq 2(r_1 + r_2)$$

$$\sum_{v \in S_0} (\delta(G) - 1) + \sum_{w \in S_1} (\delta(G) - 1) \leq 2(r_1 + r_2)$$

$$\frac{\delta(G) - 1}{2}(s_0 + s_1) \leq r_1 + r_2.$$

Combining the previous inequality with the first constraint of (P) we get

$$s_0 + s_1 + r_0 + r_1 + r_2 + \sum_{i=3}^{\Delta(G)} r_i = n$$

$$s_0 + s_1 + r_1 + r_2 \leq n$$

$$s_0 + s_1 + \frac{\delta(G) - 1}{2}(s_0 + s_1) \leq n$$

$$s_0 + s_1 \leq \frac{2n}{\delta(G) + 1},$$

which completes the proof. $\qquad\square$

6 Monophonic convexity

In this section we deal with the monophonic convexity instead of the P_3-convexity. A set $S \subseteq V(G)$ is convex in the monophonic convexity if, for all pairs of vertices $u, v \in S$, every vertex that lies on an induced path from u to v is also in S. We prove that the problem MAXIMUM CONVEXLY INDEPENDENT SET is also NP-complete when considering the monophonic convexity.

We use the following result as stated by Dourado et al [DPS10].

Theorem 6 *If G is an atom that is not a complete graph, then every pair of non-adjacent vertices is an m-hull set of G.*

We also introduce the following lemma.

Lemma 5 *The* MAXIMUM CLIQUE *problem is NP-complete even for graphs with no clique separator.*

Proof: The problem is clearly in NP. We present a reduction from the MAXIMUM CLIQUE problem for general graphs.

Consider a connected graph G' given as input for the MAXIMUM CLIQUE problem for general graphs. If G' is already an atom there is nothing to prove, so we must assume that G' has a clique separator. Also, we restrict $|V(G')| \geq 3$. We construct an atom G by making the disjunct union between G' and $C_{|V(G')|}$, which is a cycle with $|V(G')|$ vertices. Also, each vertex of $C_{|V(G')|}$ is adjacent to exactly one vertex of G' and each vertex of G' has exactly one neighbor in $C_{|V(G')|}$.

None of the clique separators of G' can disconnect the vertices of G, since it is always possible to go from a vertex to any other going through the vertices of $C_{|V(G')|}$. Also, since no vertex from $C_{|V(G')|}$ is adjacent to more than one vertex of G', they cannot be part of any clique separator. Furthermore, no clique of G is larger than the maximum clique of G', since we only add cliques of size 2 and no clique of G' can grow by the addition of a vertex from $C_{|V(G')|}$. This implies that the size of the maximum clique of G is the same as the size of the maximum clique of G' and, being G an atom, we prove that the MAXIMUM CLIQUE problem is NP-complete for atoms. □

We can now state de the NP-completeness.

Theorem 7 *The* MAXIMUM CONVEXLY INDEPENDENT SET *problem is NP-complete in the monophonic convexity even for atoms.*

Proof: We present a reduction of the MAXIMUM CLIQUE problem for atoms to the MAXIMUM CONVEXLY INDEPENDENT SET problem. Let G be a graph with no clique separator. According to Theorem 6, any two non-adjacent vertices generate the whole graph, thus no convexly independent set with size greater than two exists in which two of its vertices are not neighbors. This implies that any convexly independent set of G with size at least three must be a clique. On the other hand, a clique is always a convexly independent set in the monophonic convexity, since the only induced path between any two of its vertices is the edge that connects them and, therefore, no vertex is generated. From this it is clear that the rank of G in the monophonic convexity coincides with the size of its maximum clique, thus determining the former immediately gives us latter. From Lemma 5, we know that the MAXIMUM CLIQUE problem is NP-complete for atoms. Therefore, so is the MAXIMUM CONVEXLY INDEPENDENT SET problem in the monophonic convexity. □

References

[BCD$^+$12] Rommel M. Barbosa, Erika M. M. Coelho, Mitre C. Dourado, Dieter Rautenbach, and Jayme L. Szwarcfiter. On the Carathéodory number for the convexity of paths of order three. *SIAM Journal on Discrete Mathematics*, 26(3):929–939, 2012.

[BRdSS13] Rommel Barbosa, Dieter Rautenbach, Vinícius F. dos Santos, and Jayme L. Szwarcfiter. On minimal and minimum hull sets. *Electronic Notes in Discrete Mathematics*, 44(0):207 – 212, 2013.

[CDD$^+$10] Carmen C. Centeno, Simone Dantas, Mitre C. Dourado, Dieter Rautenbach, and Jayme L. Szwarcfiter. Convex partitions of graphs induced by paths of order three. *Discrete Mathematics & Theoretical Computer Science*, 12(5), 2010.

[CDP$^+$11] Carmen C. Centeno, Mitre C. Dourado, Lucia D. Penso, Dieter Rautenbach, and Jayme L. Szwarcfiter. Irreversible conversion of graphs. *Theoretical Computer Science*, 412(29):3693 – 3700, 2011.

[CPRPdS13] Carmen C. Centeno, Lucia Penso, Dieter Rautenbach, and Vinícius P. de Sá. Geodetic number versus hull number in P_3-convexity. *SIAM Journal on Discrete Mathematics*, 27(2):717–731, 2013.

[DPS10] Mitre C. Dourado, Fábio Protti, and Jayme L. Szwarcfiter. Complexity results related to monophonic convexity. *Discrete Applied Mathematics*, 158(12):1268 – 1274, 2010. Traces from LAGOS'07 IV Latin American Algorithms, Graphs, and Optimization Symposium Puerto Varas - 2007.

[DRdS$^+$12] Mitre C. Dourado, Dieter Rautenbach, Vinícius F. dos Santos, Philipp M. Schäfer, Jayme L. Szwarcfiter, and Alexandre Toman. An upper bound on the P_3-Radon number. *Discrete Mathematics*, 312(16):2433 – 2437, 2012.

[DRdS$^+$13a] Mitre C. Dourado, Dieter Rautenbach, Vinícius F. dos Santos, Philipp M. Schäfer, and Jayme L. Szwarcfiter. On the carathéodory number of interval and graph convexities. *Theoretical Computer Science*, 2013.

[DRdS$^+$13b] Mitre C. Dourado, Dieter Rautenbach, Vinícius F. dos Santos, Philipp M. Schäfer, Jayme L. Szwarcfiter, and Alexandre Toman. Algorithmic and structural aspects of the P_3-Radon number. *Annals of Operations Research*, 206(1):75–91, 2013.

[DRdSS12] Mitre C. Dourado, Dieter Rautenbach, Vinícius F. dos Santos, and Jayme L. Szwarcfiter. Characterization and recognition of Radon-independent sets in split graphs. *Information Processing Letters*, 112(24):948 – 952, 2012.

[HRS13] Michael A. Henning, Dieter Rautenbach, and Philipp M. Schäfer. Open packing, total domination, and the P_3-Radon number. *Discrete Mathematics*, 313(9):992 – 998, 2013.

[HS99] Michael A. Henning and Peter J. Slater. Open packing in graphs. *Journal of Combinatorial Mathematics and Combinatorial Computing*, 29:3–16, 1999.

[JR09] Paul A. Dreyer Jr. and Fred S. Roberts. Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Applied Mathematics*, 157(7):1615 – 1627, 2009.

[Kar72]   Richard M. Karp. Reducibility among combinatorial problems. In Raymond E. Miller, James W. Thatcher, and Jean D. Bohlinger, editors, *Complexity of Computer Computations*, The IBM Research Symposia Series, pages 85–103. Springer US, 1972.

[vdV93]   Marcel L. J. van de Vel. *Theory of Convex Structures*. North-Holland Mathematical Library. Elsevier Science, 1993.