A general framework for path convexities

João Vinicius C. Thompson[1] · Loana T. Nogueira[2] · Fábio Protti[2] ·
Raquel S. F. Bravo[2] · Mitre C. Dourado[3] · Uéverton S. Souza[2]

Abstract

In this work we deal with the so-called *path convexities*, defined over special collections
of paths. For example, the collection of the shortest paths in a graph is associated
with the well-known *geodesic convexity*, while the collection of the induced paths is
associated with the *monophonic convexity*; and there are many other examples. Besides
reviewing the path convexities in the literature, we propose a general path convexity
framework, of which most existing path convexities can be viewed as particular cases.
Some benefits of the proposed framework are the systematization of the algorithmic
study of related problems and the possibility of defining new convexities not yet
investigated.

Keywords Algorithmic complexity · Graph convexity · Path convexity

✉ Uéverton S. Souza
usouza@ic.uf.br

João Vinicius C. Thompson
joao.thompson@cefet-rj.br

Loana T. Nogueira
loana@ic.uff.br

Fábio Protti
fabio@ic.uff.br

Raquel S. F. Bravo
raquel@ic.uff.br

Mitre C. Dourado
mitre@nce.ufrj.br

[1] CEFET/Petrópolis, Petrópolis, Brazil

[2] Universidade Federal Fluminense, Niterói, Brazil

[3] Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil

1 Introduction

A *finite convexity space* is a pair $(V, \mathcal{C})$ consisting of a finite set V and a family $\mathcal{C}$ of subsets of V such that $\emptyset \in \mathcal{C}$, $V \in \mathcal{C}$, and $\mathcal{C}$ is closed under intersection. Members of $\mathcal{C}$ are called *convex sets*.

Let $\mathscr{P}$ be a collection of paths of a graph G, and let $I_{\mathscr{P}} : 2^{V(G)} \to 2^{V(G)}$ be a function (called *interval function*) such that

$$I_{\mathscr{P}}(S) = S \cup \{z \notin S \mid \exists\, u, v \in S \text{ such that } z \text{ lies in an } uv\text{-path } P \in \mathscr{P}\}.$$

Distinct choices of $\mathscr{P}$ lead to interval functions of quite different behavior. Such functions, in turn, are naturally associated with special convexity spaces (the so-called *path convexities*). For instance, if $\mathscr{P}$ contains precisely all the shortest paths in a graph then the corresponding interval function is naturally associated with the well-known *geodesic convexity*; if $\mathscr{P}$ is the collection of induced paths then the corresponding interval function is associated with the *monophonic convexity*; and there are many other examples in the literature.

In this work we propose a general path convexity framework, of which most path convexities in the literature can be viewed as particular cases. Some benefits of the proposed framework are the systematization of the algorithmic study of related problems and the possibility of defining new path convexities not yet investigated.

Our contributions are concentrated mainly in Sect. 3, where we describe in detail our framework. The idea is to control the length of the paths in $\mathscr{P}$, as well as the types of chords allowed to exist in such paths. Such control can be done by means of four matrices that specify, for each pair (u, v) of vertices, the minimum/maximum length and minimum/maximum chord length in all uv-paths of $\mathscr{P}$. We prove hardness results for the more general approach, where the matrices are part of the input of the related computational problems. We also describe some polynomial cases by restricting the usage of such matrices, including linear-time methods for bounded treewidth graphs. In addition, we show how to define most existing path convexities in the literature within the proposed framework. In Sect. 4 we provide examples of new interesting convexities and discuss future algorithmic developments.

2 Preliminaries

In this section we first provide all the necessary background. Next, we briefly review the main path convexities in the literature and list six fundamental computational problems in graph convexity that will be considered in this work. Finally, we prove two useful propositions.

All graphs are finite, simple, nonempty, and connected. Let G denote a graph with n vertices and m edges. The *length* of a path P in G, denoted by $|P|$, is its number of edges. A path P in G with endpoints u and v is an uv-*path*. An uv-path P in G is *shortest* if there is no uv-path P' in G such that $|P'| < |P|$. If an uv-path P is shortest then $|P|$ is the *distance* between u and v in G, and we write $|P| = dist_G(u, v)$. A

Table 1 Some path convexities studied in the literature

Convexity name	Collection of paths $\mathscr{P}$ considered
Geodesic (Bueno et al. 2018; Harary and Nieminen 1981; Harary et al. 1993; Nebeský 1994)	Shortest paths
Monophonic (Changat and Mathew 2004; Duchet 1988; Farber and Jamison 1986)	Induced paths
g^3 (Nielsen and Oellermann 2011)	Shortest paths of length at least three
m^3 (Cáceres et al. 2012; Dragan et al. 1999)	Induced paths of length at least three
g_k (Farber and Jamison 1987)	Shortest paths of length at most k
P_3 (Cappelle et al. 2019; Centeno et al. 2010; Dourado et al. 2012; Duarte et al. 2017; Parker et al. 2006; Penso et al. 2015)	Paths of length two
P_3^* (Araujo et al. 2013)	Induced paths of length two
Triangle-path (Changat and Mathew 1999; Changat et al. 2005, 2009)	Paths allowing only chords of length two
Total (Chepoi 1997	Paths allowing only chords of length at least three
Detour (Changat et al. 2010; Chartrand et al. 2003, 2005)	Longest paths
All-path (Changat et al. 2001; Gutin and Yeo 2009; Sampathkumar 1984)	All paths

chord of length $l \geq 2$ in a path $P = (v_0, v_1, \ldots, v_{|P|})$ is an edge $v_i v_j \in E(G)$ such that $i, j \in \{0, \ldots, |P|\}$ and $|i - j| = l \geq 2$.

Let $\mathscr{P}$ be a collection of paths of a graph G, and let $I_{\mathscr{P}} : 2^{V(G)} \to 2^{V(G)}$ be the interval function associated with $\mathscr{P}$, i.e.,

$$I_{\mathscr{P}}(S) = S \cup \{z \notin S \mid \exists\, u, v \in S \text{ such that } z \text{ lies in an } uv\text{-path } P \in \mathscr{P}\}. \quad (1)$$

Define $\mathcal{C}_{\mathscr{P}}$ as the family of subsets of $V(G)$ such that $S \in \mathcal{C}_{\mathscr{P}}$ if and only if $I_{\mathscr{P}}(S) = S$. Then it is easy to see that $(V(G), \mathcal{C}_{\mathscr{P}})$ is a finite convexity space, whose convex sets are precisely the fixed points of $I_{\mathscr{P}}$.

Proposition 1 (van de Vel 1993) $(V(G), \mathcal{C}_{\mathscr{P}})$ *is a finite convexity space.*

In order to ease the notation, we omit the subscript $\mathscr{P}$ whenever it is clear from the context.

2.1 Path convexities in the literature

By varying the choice of the collection $\mathscr{P}$, interval functions of different behavior can be defined using Eq. (1). The convexity spaces associated with such functions are called *path convexities*.

In Table 1 we list the main path convexities that appear in the literature. In the table, each convexity is defined by the collection of paths $\mathscr{P}$ considered.

It is important to note that although most of the graph convexities so far studied are based on paths. There are graph convexities that are not based on paths such as the And/or-convexity, a graph convexity based on processes, and deadlock models, see Lima et al. (2018). Our focus in this work is only on path convexities.

2.2 Computational problems

In this work we focus on six computational problems that are usually studied in the field of convexity in graphs. The list, of course, is not complete and other important problems could also be considered.

We need some additional definitions. Let $S \subseteq V(G)$. If $I(S) = V(G)$ then S is an *interval set*. The *convex hull $H(S)$* of S is the smallest convex set containing S. Write $I^0(S) = S$ and define $I^{i+1}(S) = I(I^i(S))$ for $i \geq 0$. Note that $I(S) = I^1(S)$ and there exists an index i for which $H(S) = I^i(S)$. If $H(S) = V(G)$ then S is a *hull set*. The *convexity number $c(G)$* of G is the size of a maximum convex set $S \neq V(G)$. The *interval number $i(G)$* of G is the size of a smallest interval set of G. The *hull number $h(G)$* of G is the size of a smallest hull set of G. Now we are in position to state the six problems dealt with in this work:

CONVEX SET - CS
Input: A graph G and a set $S \subseteq V(G)$.
Question: Is S convex?

INTERVAL DETERMINATION - ID
Input: A graph G, a set $S \subseteq V(G)$, and a vertex $z \in V(G)$.
Question: Does z belong to $I(S)$?

CONVEX HULL DETERMINATION - CHD
Input: A graph G, a set $S \subseteq V(G)$, and a vertex $z \in V(G)$.
Question: Does z belong to $H(S)$?

CONVEXITY NUMBER - CN
Input: A graph G and a positive integer r.
Question: Is $c(G) \geq r$?

INTERVAL NUMBER - IN
Input: A graph G and a positive integer r.
Question: Is $i(G) \leq r$?

HULL NUMBER - HN
Input: A graph G and a positive integer r.
Question: Is $h(G) \leq r$?

2.3 Existing complexity results

The table below shows the complexity of the six problems listed in the preceding subsection for some convexity spaces. All the entries of the table correspond to results found in the literature, or to trivial results (indicated by '[t]').

Table 2 Problems versus convexities: complexity results

	Geodesic		Monophonic		P_3		P_3^*		Triangle-path	
CS	P	[t]	P	Dourado et al. (2010)	P	[t]	P	Araujo et al. (2013)	P	Dourado and Sampaio (2016)
ID	P	[t]	NPc	Dourado et al. (2010)	P	[t]	P	Araujo et al. (2013)	NPc	Dourado and Sampaio (2016)
CHD	P	[t]	P	Dourado et al. (2010)	P	[t]	P	Araujo et al. (2013)	P	Dourado and Sampaio (2016)
CN	NPc	Gimbel (2003)	NPc	Dourado et al. (2010)	NPc	Centeno et al. (2009)	NPc	Araujo et al. (2013)	P	Dourado and Sampaio (2016)
IN	NPc	Atici (2002)	NPc	Dourado et al. (2010)	NPc	Centeno et al. (2011)	NPc	Araujo et al. (2013)	NPc	Dourado and Sampaio (2016)
HN	NPc	Dourado et al. (2009)	P	Dourado et al. (2010)	NPc	Centeno et al. (2009)	NPc	Araujo et al. (2013)	P	Dourado and Sampaio (2016)

2.4 Two useful facts

The next two propositions are useful, which say that if INTERVAL DETERMINATION or CONVEX SET can be solved in polynomial time for some convexity space then some other problems listed in Sect. 2.2 can also be solved in polynomial time, for the same convexity space.

Proposition 2 *Let* $(V(G), \mathcal{C})$ *be any convexity space. If* INTERVAL DETERMINATION *can be solved in polynomial time for* $(V(G), \mathcal{C})$ *then* CONVEX SET *and* CONVEX HULL DETERMINATION *can also be solved respectively, in polynomial time for* $(V(G), \mathcal{C})$.

Proof Let $S \subseteq V(G)$. Since INTERVAL DETERMINATION is in P for $(V(G), \mathcal{C})$, $I(S)$ can be computed in polynomial time. Let i be the smallest index such $I^{i+1}(S) = I^i(S)$. Note that determining such an index i as well as $I^i(S)$ can also be done in polynomial time, since $i = O(n)$. Therefore:

– If $I(S) = S$ then S is a convex set, otherwise S is not convex.
– If $z \in I^i(S)$ then $z \in H(S)$, otherwise $z \notin H(S)$.

By the above observations, the problems CONVEX SET and CONVEX HULL DETERMINATION can be solved in polynomial time for $(V(G), \mathcal{C})$. □

Let $S \subseteq V(G)$. If S is not convex then an *augmenting set* of S is any set S' such that $S \subset S' \subseteq H(S)$ (where the symbol $\subset$ stands for proper inclusion).

Proposition 3 *Let* $(V(G), \mathcal{C})$ *be a convexity space. If there is a polynomial-time certification algorithm to solve* CONVEX SET *for* $(V(G), \mathcal{C})$ *that outputs an augmenting set when the problem has a negative answer then* CONVEX HULL DETERMINATION *can also be solved in polynomial time for* $(V(G), \mathcal{C})$.

Proof Let $S \subseteq V(G)$ and $z \in V(G)$. Since CONVEX SET can be solved in polynomial time for $(V(G), \mathcal{C})$, we can apply a convexity test to S in polynomial time. If the test succeeds then S is a convex set and, consequently, the convex hull of S is S itself; otherwise, the test fails and there exists a set S_1 (a certificate outputted by the test) that can be used to augment the original set S. Recall that S is properly contained in S_1.

In order to establish the convex hull of S, we are going to apply the convexity test algorithm successively, always obtaining a set S_{i+1} that augments S_i, until it cannot be augmented anymore. Let j be the smallest index such that S_j results in a convex set. Observe that $j \leq |V(G) \setminus S|$, since at least one vertex is added in each iteration. Moreover, each convexity test can be done in polynomial time. Thus, the entire process still remains polynomial.

If a vertex $z \in V(G)$ is such that $z \in S_j$ then $z \in H(S)$. Therefore, CONVEX HULL DETERMINATION is in P for $(V(G), \mathcal{C})$. □

Note that Propositions 2 and 3 can be used to fill some entries of Table 2. For example, since INTERVAL DETERMINATION is in P for the geodesic convexity, by Proposition 2 the problems CONVEX SET and CONVEX HULL DETERMINATION are also in P for such convexity. The same applies to the P_3- and P_3^*- convexities. On the other hand, Proposition 3 implies that CONVEX HULL DETERMINATION is in P for the monophonic convexity.

3 A general framework for path convexities

In this section, we propose a general framework for the study of path convexities.

From now on, we assume that every n-vertex graph G has vertices labeled $1, 2, \ldots, n$. A *length matrix* is a symmetric $n \times n$ matrix M such that each entry $M(i, j)$, for $i, j \in V(G)$, is a natural number; in addition, all diagonal entries of M are zero.

Let A, B, C, D be four $n \times n$ length matrices. Suppose that $\mathscr{P}$ is the family of paths of G such that an ij-path P of G is a member of $\mathscr{P}$ if and only if:

(1) $|P| \geq A(i, j)$;
(2) $|P| \leq B(i, j)$;
(3) all the chords in P are of length at least $C(i, j)$;
(4) all the chords in P are of length at most $D(i, j)$.

Let $I_{\mathscr{P}} : 2^{V(G)} \to 2^{V(G)}$ be the interval function associated with $\mathscr{P}$, and let $\mathcal{C}_{\mathscr{P}}$ be the family of subsets of $V(G)$ such that $S \in \mathcal{C}_{\mathscr{P}}$ if and only if $I_{\mathscr{P}}(S) = S$. Since $\mathscr{P}$ is a particular collection of paths of G, by Proposition 1, we have that $(V(G), \mathcal{C}_{\mathscr{P}})$ is a finite convexity space, equipped with interval function $I_{\mathscr{P}}$. Let us say that such a convexity space defines a *matrix path convexity*.

Again, we omit the subscript $\mathscr{P}$ when it is clear from the context.

Say that an ij-path P *satisfies* matrices A, B, C, D if all the conditions (1) to (4) above are satisfied by P.

3.1 Putting the matrices as part of the input

In the six problems listed in Sect. 2.2, the graph G is always part of the input; however, the rule that determines which collection of paths of G must be considered is *not* part of the input. More general versions of such problems are possible when the desired convexity space, expressed as a graph G together with a set of four length matrices, is part of the input. For example, consider the following version of CONVEX SET: MATRIX CONVEX SET

Input: A graph G, four $n \times n$ length matrices A, B, C, D, and $S \subseteq V(G)$.
Question: Is S convex under the matrix path convexity ruled by A, B, C, D?

All the remaining problems listed in Sect. 2.2 can be restated analogously.

The next theorems say that such "matrix problems" are all hard. However, we shall see that restrictions on the matrices A, B, C, D lead to interesting cases. In this regard, some types of length matrices are of special interest. For a graph G, the *distance matrix* of G is the length matrix M_{dist} with entries $M_{dist}(i, j) = dist_G(i, j)$, for $i, j \in V(G)$. For a positive integer constant k, the $(n - k)$-*matrix* and the k-*matrix* are the length matrices M_{n-k} and M_k with off-diagonal entries all equal to, respectively, $n - k$ and k.

Theorem 1 MATRIX CONVEX SET *is Co-NP-complete.*

Proof A certificate for a negative answer to MATRIX CONVEX SET is a triple i, j, z (with $i, j \in S$ and $z \notin S$) and an ij-path P in G containing z such that P satisfies A to

D. Such a certificate can be clearly checked in polynomial time. Therefore, MATRIX CONVEX SET is in CoNP.

To prove that MATRIX CONVEX SET is Co-NP-complete, we show a reduction from the following NP-complete problem (Haas and Hoffmann 2006): given three distinct vertices i, j, z in a graph H, decide whether there is a chordless ij-path passing through z.

Let G be the graph obtained from H by replacing each edge (s, z) incident to z by an sz-path containing $n - 1$ internal vertices of degree two, where $n = |V(H)|$. In other words, G is a subdivision of H obtained by subdividing each edge incident to z using $n - 1$ vertices. Set A and B as the length matrices with off-diagonal entries all equal to, respectively, $2n$ and $3n - 3$. Also, set $C = D = M_1$ (the k-matrix for $k = 1$). Finally, set $S = \{i, j\}$. Note that the collection of paths $\mathscr{P}$ defined by A, B, C, D contains the chordless paths with length at least $2n$ and at most $3n - 3$.

Suppose that there is a chordless ij-path P_H in H passing through z. Write $P_H = (s_0 = i, s_1, \ldots, s_{h-1}, s_h = z, s_{h+1}, \ldots, s_l = j)$. Then there is a chordless ij-path P_G in G obtained from P_H by subdividing edges (s_{h-1}, z) and (z, s_{h+1}) using $n - 1$ vertices of degree two for each edge. Note that $|P_G| = (l-2)+2n$. Since $2 \leq l \leq n-1$, we have $2n \leq |P_G| \leq 3n - 3$. Therefore, P_G satisfies A, B, C, D, and its existence implies that S is not convex.

Conversely, suppose that S is not convex. Then there is a chordless ij-path P_G of length at least $2n$ passing through some vertex of G lying outside S. But, by the construction of G, all the ij-paths of length at least $2n$ must necessarily pass through z. Let P_{hz} be the subpath of P_G with length n that starts at a vertex h and ends at z. Similarly, let $P_{zh'}$ be the subpath of P_G with length n that starts at z and ends at a vertex h'. By replacing P_{hz} and $P_{zh'}$ by edges (h, z) and (z, h'), we obtain a chordless ij-path in H passing through z. This completes the proof.   □

Theorem 2 MATRIX INTERVAL DETERMINATION *is NP-complete.*

Proof A certificate for a positive answer to MATRIX INTERVAL DETERMINATION is a pair i, j (with $i, j \in S$) and an ij-path P in G containing z (recall that z is part of the input) such that P satisfies A, B, C, D. Moreover, this certificate can be checked in polynomial time. Therefore, MATRIX INTERVAL DETERMINATION is in NP.

To prove that MATRIX INTERVAL DETERMINATION is NP-complete, recall from Table 2 that INTERVAL DETERMINATION is NP-complete for the monophonic convexity. If $A = M_2$, $B = M_{n-1}$, and $C = D = M_1$, the collection of paths $\mathscr{P}$ associated with such matrices is precisely the collection of induced paths of G. Then INTERVAL DETERMINATION for the monophonic convexity is a restriction of MATRIX INTERVAL DETERMINATION, i.e., the former problem is NP-complete.   □

Theorem 3 MATRIX CONVEX HULL DETERMINATION *is NP-complete.*

Proof A certificate for a positive answer to MATRIX CONVEX HULL DETERMINATION is formed by a sequence of paths $P_1, P_2, \ldots, P_r$ such that $r \leq n - 1$, $z \in V(P_r)$, and each P_k is an $i_k j_k$-path satisfying matrices A, B, C, D with $i_k, j_k \in S \cup V(P_1) \cup \cdots \cup V(P_{k-1})$. It is easy to see that such a certificate can be checked in polynomial time. Therefore, MATRIX CONVEX HULL DETERMINATION is in NP.

For the hardness proof, we use the same reduction described in Theorem 1. Again, if there is a chordless ij-path $P_H = (i, \ldots, h, z, h', \ldots, j)$ in H then there is a chordless ij-path P_G in G obtained from P_H by subdividing edges (h, z) and (z, h'), as explained in the proof of Theorem 1, such that P_G satisfies A, B, C, D. Therefore, $z \in I(S) \subseteq H(S)$.

Conversely, suppose that $z \in H(S)$. Then there is a sequence of paths $P_1, P_2, \ldots,$ P_r such that $z \in V(P_r)$ and, for $1 \leq k \leq r$, P_k is an $i_k j_k$-path satisfying matrices A, B, C, D, where $i_k, j_k \in S \cup V(P_1) \cup \cdots \cup V(P_{k-1})$.

Note that $i_1, j_1 \in S$. Thus, $\{i_1, j_1\} = \{i, j\}$. In addition, $|P_1| \geq 2n$. But, by the construction of G, all the ij-paths of length at least $2n$ must necessarily pass through z. Hence, $z \in V(P_1)$. The rest of the proof follows as in the proof of Theorem 1: let P_{hz} (resp., $P_{zh'}$) be the subpath of P_1 with length n starting at some h (resp., at z) and ending at z (resp., at some h'). Replacing P_{hz} and $P_{zh'}$ by edges (h, z) and (z, h') produces a chordless ij-path in H passing through z. □

Theorem 4 1. MATRIX CONVEXITY NUMBER *is NP-hard.*
2. MATRIX INTERVAL NUMBER *is NP-complete.*
3. MATRIX HULL NUMBER *is NP-complete.*

Proof We first prove that both MATRIX INTERVAL NUMBER and MATRIX HULL NUMBER are in NP.

A certificate for a positive answer to MATRIX INTERVAL NUMBER consists of a set $S \subseteq V(G)$ of size at most r, and a collection of paths $\{P_z \mid z \in V(G) \setminus S\}$ such that for each path P_z there exist $i_z, j_z \in S$ for which P_z is an $i_z j_z$-path containing z and satisfying A, B, C, D. Since the number of paths in the collection is $O(n)$, MATRIX INTERVAL NUMBER is in NP.

A certificate for a positive answer to MATRIX HULL NUMBER consists of a set S of size at most r and a sequence of paths $P_1, P_2, \ldots, P_s$ such that:

(a) $s \leq n - 1$;
(b) $V(G) \setminus S \subseteq V(P_1) \cup \cdots \cup V(P_s)$;
(c) for $1 \leq k \leq s$, P_k is an $i_k j_k$-path satisfying matrices A, B, C, D,

with $i_k, j_k \in S \cup V(P_1) \cup \cdots \cup V(P_{k-1})$.

Since such a certificate can be checked in polynomial time, MATRIX HULL NUMBER is in NP.

Now we describe the hardness proof for the three problems. As in Theorem 2, we use a proof by restriction. Recall from Table 2 that CONVEXITY NUMBER, INTERVAL NUMBER, and HULL NUMBER are all NP-complete for the geodesic convexity. If $A = B = M_{dist}$ and $C = D = M_1$, the collection of paths $\mathscr{P}$ associated with such matrices is precisely the collection of shortest paths of G. This means that CONVEXITY NUMBER, INTERVAL NUMBER, and HULL NUMBER for the geodesic convexity are, respectively, restrictions of MATRIX CONVEXITY NUMBER, MATRIX INTERVAL NUMBER, and MATRIX HULL NUMBER. Hence, the theorem follows. □

3.2 Constant matrices: the (*a*, *b*, *c*, *d*)-path convexity

In this section, we study the case in which there are constants a, b, c, d such that $A = M_a$, $B = M_b$, $C = M_c$, and $D = M_d$. In this scenario we can assume that the matrices are *not* part of the input, because length restrictions are known in advance. This gives rise to "constant matrix versions" of the problems studied in the preceding subsection. For example, consider the following problems:

(*a*, *b*, *c*, *d*)-CONVEX SET
Input: A graph G and a set $S \subseteq V(G)$.
Question: Is S convex under the matrix path convexity ruled by A, B, C, D?
Equivalently: Is S convex under the path convexity defined by the collection $\mathscr{P}(a, b, c, d)$ of paths of G whose length is at least a and at most b, and whose chords have length at least c and at most d?

(*a*, *b*, *c*, *d*)-INTERVAL DETERMINATION
Input: A graph G, a set $S \subseteq V(G)$, and a vertex $z \in V(G)$.
Question: Does z belong to $I(S)$, where I is the interval function associated with the collection $\mathscr{P}(a, b, c, d)$ of paths of G?

The remaining matrix problems can be restated analogously.

The path convexity for which the path/chord length restrictions are ruled by four constants a, b, c, d as explained above is called (*a*, *b*, *c*, *d*)-*path convexity*.

Theorem 5 (*a*, *b*, *c*, *d*)-INTERVAL DETERMINATION *is in* P.

Proof Note that for a constant ℓ there are $O(n^{\ell+1})$ paths of length ℓ in G. Thus there are $O(\sum_{\ell=a}^{b} n^{\ell+1})$ paths in G with length at least a and at most b. Now, for each pair (i, j) of distinct vertices in S, there are $O(\sum_{\ell=a}^{b} n^{\ell-1})$ ij-paths with length at least a and at most b, because i and j are the fixed endpoints of each such path. But S contains $O(|S|^2)$ pairs of distinct vertices. This amounts to checking $O(|S|^2 \sum_{\ell=a}^{b} n^{\ell-1})$ paths. Since a path of length ℓ can have at most $O(\ell^2)$ chords, we can select all the ij-paths in $\mathscr{P}(a, b, c, d)$ with $i, j \in S$ in $O(|S|^2 \sum_{\ell=a}^{b} \ell^2 n^{\ell-1})$ time. Finally, checking whether z belongs to one of such paths can be done in $O(1)$ time per path, because the lenght of each path is bounded by b. This gives a naïve polynomial-time brute-force algorithm to check whether $z \in I(S)$. $\qquad\square$

By Proposition 2, we have:

Corollary 1 (*a*, *b*, *c*, *d*)-CONVEX SET *and* (*a*, *b*, *c*, *d*)-CONVEX HULL DETERMINA-TION *are in* P. $\qquad\square$

As for the other three problems, (*a*, *b*, *c*, *d*)-CONVEXITY/INTERVAL/HULL NUMBER, we remark that the special cases

$$(a = 2, b = 2, c = 1, d = 2) \text{ and } (a = 2, b = 2, c = 1, d = 1)$$

correspond precisely to the P_3- and P_3^*- convexities, as indicated in Table 1. For both convexities, all the three problems are NP-complete (see Table 2).

3.3 (a, b, c, d)-path convexity and bounded treewidth graphs

In this section, we investigate the complexity of the six (a, b, c, d)-path convexity problems in Sect. 3.2 when applied to bounded treewidth graphs. As we shall see, linear-time methods will be possible in this case.

Let G be a graph, T a tree, and $V = (V_t)_{t \in T}$ a family of vertex sets $V_t \subseteq V(G)$ indexed by the vertices t of T. The pair (T, V) is called a *tree-decomposition of G* if it satisfies the following three conditions (Diestel 2005):

[(T1)] $V(G) = \bigcup_{t \in T} V_t$;
[(T2)] for every $e \in E(G)$ there exists $t \in T$ such that both ends of e lie in V_t;
[(T3)] if V_{t_i} and V_{t_j} both contain a vertex v then $v \in V_{t_k}$ for all vertices t_k in the path between t_i and t_j.

The *width* of (T, V) is the number $\max\{|V_t| - 1 \mid t \in T\}$, and the *treewidth $tw(G)$* of G is the minimum width of any tree-decomposition of G.

Graphs of treewidth at most k are called *partial k-trees*. Some graph classes with bounded treewidth include: forests (treewidth 1); pseudoforests, cacti, series-parallel graphs, and outerplanar graphs (treewidth at most 2); Halin graphs and Apollonian networks (treewidth at most 3) (Bodlaender 1998; Brandstädt et al. 1999). Control flow graphs arising in the compilation of structured programs also have bounded treewidth (at most 6) (Thorup 1998).

In 1990, Courcelle (1990) stated that for any graph G with treewidth bounded by a constant k and for any graph property Π that can be formulated in CMSOL$_2$ (*Counting Monadic Second-Order Logic* where quantification over sets of vertices or edges and predicates testing the size of sets modulo constants are allowed), there is a linear-time algorithm that decides if G satisfies Π (Courcelle 1990, 1997; Courcelle and Mosbah 1993; Courcelle and Engelfriet 2011). This result has been extended a number of times (Arnborg et al. 1991; Borie et al. 1992; Courcelle et al. 2000; Hliněný et al. 2008). In particular, Arnborg et al. (1991) study optimization problems over sets definable in Counting Monadic Second-Order Logic.

By Courcelle's meta-theorems based on CMSOL$_2$ (Courcelle 1990, 1997; Courcelle and Mosbah 1993), obtaining linear-time methods to solve the six problems of Sect. 3.2 on bounded treewidth graphs amounts to showing that the related properties are expressible in CMSOL$_2$.

Theorem 6 (a, b, c, d)-INTERVAL DETERMINATION *is solvable in linear time on bounded treewidth graphs.*

Proof It is enough to show that the property "$z \in I(S)$" is CMSOL$_2$-expressible. Given G, S, and z, we construct $\varphi(G, S, z, a, b, c, d)$ such that $z \in I(S) \Leftrightarrow \varphi(G, S, z, a, b, c, d)$ as follows:

$$(z \in S) \vee (\exists u, v, P$$
$$(u, v \in S \wedge$$
$$P \text{ is an } uv\text{-path} \wedge z \text{ is in } P \wedge$$
$$Card(P) \geq a \wedge Card(P) \leq b \wedge$$

$$\forall P' \, (\, (\, P' \subseteq P \, \wedge \, Card(P') \geq 2 \, \wedge$$
$$\exists \, u', v'(P' \text{ is an } u'v'\text{-path} \, \wedge \, adj(u', v')))$$
$$\Rightarrow (Card(P') \geq c \, \wedge \, Card(P') \leq d) \,) \,)) \qquad (2)$$

In the above formula, paths are regarded as subsets of edges. Using this approach, the subformula "P is an uv-path" can be expressed in CMSOL$_2$ (see Courcelle 1997). Note that a chord is expressed as an $u'v'$-subpath P' of P with length at least c and at most d such that u' is adjacent to v'. □

Corollary 2 (a, b, c, d)-CONVEX SET *can be solved in linear time on bounded treewidth graphs.*

Proof The property "S is convex" is equivalent to "there is no z such that $z \notin S$ and $z \in I(S)$". By Theorem 6, "$z \in I(S)$" is CMSOL$_2$-expressible. Thus the result easily follows. □

Corollary 3 (a, b, c, d)-CONVEX HULL DETERMINATION *can be solved in linear time on bounded treewidth graphs.*

Proof The property "$z \in H(S)$" is equivalent to "there exists S_1 such that: (a) S_1 is convex, (b) $S \subseteq S_1$, (c) $z \in S_1$, and (d) there is no S_2 such that S_2 is convex, $S \subseteq S_2$, and S_2 is properly contained in S_1". By Corollary 2, we can use CMSOL$_2$ to say that the sets S_1 and S_2 are convex. Thus the result follows. □

For the remaining three problems ((a, b, c, d)-CONVEXITY/INTERVAL/HULL NUMBER), we consider their optimization versions (maximization in the case of CONVEXITY NUMBER, and minimization in the case of INTERVAL/HULL NUMBER). Note that the properties "S is a convex set distinct from $V(G)$", "S is an interval set", and "S is a hull set" can be expressed in CMSOL$_2$. Therefore the optimization versions ("find an optimal set satisfying the required property") are LinCMSOL$_2$ problems (Arnborg et al. 1991; Courcelle 1992), to which the following result applies:

Theorem 7 (Arnborg et al. 1991; Courcelle 1992) *Let k be a positive constant, and Π be a* LinCMSOL$_2$ *problem. Then Π can be solved in linear time on graphs of treewidth bounded by k (if the tree-decomposition is given with the input graph).*

Therefore:

Corollary 4 *The optimization versions of (a, b, c, d)-CONVEXITY NUMBER, (a, b, c, d)-INTERVAL NUMBER, and (a, b, c, d)-HULL NUMBER can be solved in linear time on bounded treewidth graphs (if the tree-decomposition is given with the input graph).*

3.4 Particular cases of the (a, b, c, d)-path convexity

In this section we show that, by extending the meaning of the parameters a, b, c, d, some path convexities in the literature can be viewed as particular cases of the

Table 3 Path convexities as particular cases of the (a, b, c, d)-path convexity

Convexity	a	b	c	d
Geodesic	σ	σ	1	1
Monophonic	2	∞	1	1
g^3	3	σ	1	1
g_k	σ	$k \mid \sigma$	1	1
m^3	3	∞	1	1
P_3	2	2	1	2
P_3^*	2	2	1	1
Triangle-path	2	∞	1	2
Total	2	∞	3	∞
Detour	ℓ	ℓ	1	∞
All-path	2	∞	1	∞

Note that putting $c = d = 1$ implies that all the paths of the considered collection $\mathscr{P}$ are chordless

(a, b, c, d)-path convexity. In Table 3 below, the symbol 'σ' (resp.,'ℓ') means that the length of the shortest (resp., longest) path between each pair of distinct vertices must be considered. The symbol '∞' stands for no length restriction. For a constant k, the symbol '$k \mid \sigma$' means that, for each pair (i, j) of distinct vertices, the minimum value between k and the length of the shortest ij-path must be considered.

4 Concluding remarks

In this work we described a matrix path convexity framework, where, by means of four input matrices, we can specify the types of paths that must be considered for each pair of vertices of the input graph, "customizing" the path convexity to be dealt with. Since this general approach results in the hardness of the related computational problems (at least the more studied ones), we also investigate the case of constant matrices. The latter case leads to the study of the (a, b, c, d)-path convexity, where the rule that defines the convexity is not part of the input. If a, b, c, and d are positive constants then, in such convexity, the problems of computing the interval of a set, deciding whether a set is convex, and computing the convex hull of a set, are all solvable in polynomial time. In addition, all the "(a, b, c, d)-versions" of the six problems listed in Sect. 2.2 are solvable in linear time if the input graph has bounded treewidth. We have also shown that, by extending the meaning of the parameters a, b, c, d, most path convexities considered in the literature can be viewed as particular cases of the (a, b, c, d)-path convexity. In this regard, other interesting convexities, not yet considered in the literature up to the authors' knowledge, can be defined by choosing other values for the tuple (a, b, c, d). Tables 4 and 5 describe such convexities. The symbol 'n$^-$' means that paths with length one less than the number of vertices of the input graph must be considered.

Let $\mathbb{K} = \{k \mid \sigma\}_{k \in \mathbb{N}^*}$ and $\Sigma = \mathbb{N}^* \cup \{\sigma, \infty, \ell, \text{n}^-\} \cup \mathbb{K}$, and consider the domain of tuples $(a, b, c, d) \in \Sigma^4$ (considering, of course, only meaningful cases). Such

Table 4 Some new convexities proposed in this work

Convexity name	Collection of paths $\mathscr{P}$ considered
g^k	Shortest paths of length at least k
m^k	Induced paths of length at least k
(k, l)-path	Chordless paths of length between k and l
k-path	Chordless paths with k vertices
Hamiltonian	Hamiltonian paths

Table 5 Convexities from Table 4 described according the (a, b, c, d)-path convexity framework

Convexity	a	b	c	d
g^k	k	σ	1	1
m^k	k	∞	1	1
(k, l)-path	k	l	1	1
k-path	$k - 1$	$k - 1$	1	1
Hamiltonian	n^-	n^-	1	∞

domain can be used to systematize algorithmic studies in path convexity in some ways. One example is to find complexity dichotomies (complete classifications of the complexity of a fixed computational problem Π). For example, if $\Pi =$ INTERVAL DETERMINATION, $c = 1$, and $d \in \{1, 2\}$, which values of (a, b, c, d) imply tractability (hardness) of Π under the (a, b, c, d)-path convexity? From Tables 2 and 3 we know that the cases $(\sigma, \sigma, 1, 1)$, $(2, 2, 1, 1)$, and $(2, 2, 1, 2)$ are in P, while $(2, \infty, 1, 1)$ and $(2, \infty, 1, 2)$ are NP-complete.

References

Araujo RT, Sampaio RM, Szwarcfiter JL (2013) The convexity of induced paths of order three. Discrete Math 44:109–114

Arnborg S, Lagergren J, Seese D (1991) Easy problems for tree-decomposable graphs. J Algorithms 12(2):308–340

Atici M (2002) Computational complexity of geodetic set. Int J Comput Math 79:587–591

Bodlaender HL (1998) A partial k-arboretum of graphs with bounded treewidth. Theor Comput Sci 209(1–2):1–45

Borie RB, Parker RG, Tovey CA (1992) Automatic generation of linear-time algorithms from predicate calculus descriptions of problems on recursively constructed graph families. Algorithmica 7:555–581

Brandstädt A, Le VB, Spinrad JP (1999) Graph classes: a survey, vol 3. SIAM, Philadelphia

Bueno LR, Penso LD, Protti F, Ramos VR, Rautenbach D, Souza US (2018) On the hardness of finding the geodetic number of a subcubic graph. Inf Process Lett 135:22–27

Cáceres J, Oellermann OR, Puertas ML (2012) Minimal trees and monophonic convexity. Discussiones Mathematicae Graph Theory 32(4):685–704

Cappelle MR, Coelho EMM, Coelho H, Silva BR, Protti F, Souza US (2019) P3-hull number of graphs with diameter two. In: Latin and American algorithms, graphs and optimization symposium. LAGOS 2019. Electronic notes in theoretical computer science, vol 346, pp 309–320

Centeno CC, Dourado MC, Szwarcfiter JL (2009) On the convexity of paths of length two in undirected graphs. Electron Notes Discrete Math 32:11–18

Centeno CC, Dantas S, Dourado MC, Rautenbach D, Szwarcfiter JL (2010) Convex partitions of graphs induced by paths of order three. Discrete Math 12(5):175–184

Centeno CC, Dourado MC, Penso LD, Rautenbach D, Szwarcfiter JL (2011) Irreversible interval of graphs. Theor Comput Sci 412:3693–3700

Changat M, Mathew J (1999) On triangle path convexity in graphs. Discrete Math 206:91–95

Changat M, Mathew J (2004) Induced path transit function, monotone and Peano axioms. Discrete Math 286(3):185–194

Changat M, Klavzar S, Mulder HM (2001) The all-paths transit function of a graph. Czechoslovak Math J 51(2):439–448

Changat M, Mulder HM, Sierksma G (2005) Convexities related to path properties on graphs. Discrete Math 290(2–3):117–131

Changat M, Narasimha-Shenoi PG, Mathews J (2009) Triangle path transit functions, betweenness and pseudo-modular graphs. Discrete Math 309(6):1575–1583

Changat M, Narasimha-Shenoi PG, Pelayo I (2010) The longest path transit function of a graph and betweenness. Utilitas Math 82:111–127

Chartrand G, Garry L, Zhang P (2003) The detour number of a graph. Utilitas Math 64:97–113

Chartrand G, Escudaro H, Zhang P (2005) Detour distance in graphs. J Combin Math Combin Comput 52:75–94

Chepoi V (1997) Peakless functions on graphs. Discrete Appl Math 73(2):175–189

Courcelle B (1997) The expression of graph properties and graph transformations in monadic second-order logic. In: Handbook of graph grammars and computing by graph transformations, vol 1, pp 313–400

Courcelle B (1990) The monadic second-order logic of graphs I. Recognizable sets of finite graphs. Inf Comput 25(1):12–75

Courcelle B (1992) The monadic second-order logic of graphs III: tree-decompositions, minor and complexity issues. Informatique Théorique et Applications 26:257–286

Courcelle B, Engelfriet J (2011) Graph structure and monadic second-order logic. Cambridge University Press, Cambridge

Courcelle B, Mosbah M (1993) Monadic second-order evaluations on tree-decomposable graphs. Theor Comput Sci 109(1):49–82

Courcelle B, Makowsky JA, Rotics U (2000) Linear time solvable optimization problems on graphs of bounded clique-width. Theory Comput Syst 33(2):125–150

Diestel R (2005) Graph theory, 3rd edn. Springer, Berlin

Dourado MC, Sampaio RM (2016) Complexity aspects of the triangle path convexity. Discrete Appl Math 206:39–47

Dourado MC, Gimbel JG, Kratochvil J, Protti F, Szwarcfiter JL (2009) On the computation of the hull number of a graph. Discrete Math 309:5668–5674

Dourado MC, Protti F, Szwarcfiter JL (2010) Complexity results related to monophonic convexity. Discrete Math 158:1268–1274

Dourado MC, Rautenbach D, dos Santos VF, Schäfer PM, Szwarcfiter JL, Toman A (2012) An upper bound on the P_3-Radon number. Discrete Math 312(16):2433–2437

Dragan FF, Nicolai F, Brandstädt A (1999) Convexity and HHD-free graphs. SIAM J Discrete Math 12:119–135

Duarte MA, Penso LD, Rautenbach D, Souza US (2017) Complexity properties of complementary prisms. J Combin Optim 33:365–372

Duchet P (1988) Convex sets in graphs, II. Minimal path convexity. J Combin Theory Ser B 44:307–316

Farber M, Jamison RE (1986) Convexity in graphs and hypergraphs. SIAM J Algebr Discrete Methods 7(3):433–444

Farber M, Jamison RE (1987) On local convexity in graphs. Discrete Math 66:231–247

Gimbel JG (2003) Some remarks on the convexity number of a graph. Graphs Combin 19:357–361

Gutin G, Yeo A (2009) On the number of connected convex subgraphs of a connected acyclic digraph. Discrete Appl Math 157(7):1660–1662

Haas R, Hoffmann M (2006) Chordless path through three vertices. Theor Comput Sci 351:360–371

Harary F (1984) Convexity in graphs: achievement and avoidance games. Ann Discrete Math 20:323

Harary F, Nieminen J (1981) Convexity in graphs. J Differ Geom 16:185–190

Harary F, Loukakis E, Tsouros C (1993) The geodetic number of a graph. Math Comput Model 17(11):89–95

Hlíněný P, Oum S, Seese D, Gottlob G (2008) Width parameters beyond tree-width and their applications. Comput J 51(3):326–362

Kanté MM, Nourine L (2013) Polynomial time algorithms for computing a minimum hull set in distance-hereditary and chordal graphs. Theory and practice of computer science. Lect Notes Comput Sci 7741:268–279

Lima CVGC, Protti F, Rautenbach D, Souza US, Szwarcfiter JL (2018) And/or-convexity: a graph convexity based on processes and deadlock models. Ann Oper Res 264:267–286

Nebeský L (1994) A characterization of the interval function of a connected graph. Czechoslov Math J 44(1):173–178

Nielsen MH, Oellermann OR (2011) Steiner trees and convex geometries. SIAM J Discrete Math 23(2):680–693

Parker DB, Westhoff RF, Wolf MJ (2006) Two-path convexity in clone-free regular multipartite tournaments. Australas J Combin 36:177–196

Parvathy KS (1995) Studies on convex structures with emphasis on convexity in graphs. Ph.D. Thesis, Cochin University, Kochi

Pelayo IM (2013) Geodesic convexity in graphs. Springer, New York

Penso LD, Protti F, Rautenbach D, Souza US (2015) Complexity analysis of P3-convexity problems on bounded-degree and planar graphs. Theor Comput Sci 607:83–95

Peterin I (2012) The pre-hull number and lexicographic product. Discrete Appl Math 312:2153–2157

Sampathkumar E (1984) Convex sets in a graph. Indian J Pure Appl Math 15(10):1065–1071

Thompson JVC, Nogueira LT, Protti F, Bravo RSF, Dourado MC, Souza US (2019) A general framework for path convexities. Algorithmic aspects in information and management. AAIM 2019. Lect Notes Comput Sci 11640:272–283

Thorup M (1998) All structured programs have small tree width and good register allocation. Inf Comput 142(2):159–181

van de Vel MLJ (1993) Theory of convex structures. Elsevier, North Holland