Computing metric hulls in graphs

Kolja Knauer, Nicolas Nisse

Discrete Mathematics and Theoretical Computer Science DMTCS vol. **21**:1, 2019, #11

Computing metric hulls in graphs

Kolja Knauer[1*] Nicolas Nisse[2†]

[1] *Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France*
[2] *Université Côte d'Azur, Inria, CNRS, I3S, France*

received 30th July 2018, *revised 12th Apr. 2019*, *accepted 15th May 2019*.

We prove that, given a closure function the smallest preimage of a closed set can be calculated in polynomial time in the number of closed sets. This implies that there is a polynomial time algorithm to compute the convex hull number of a graph, when all its convex subgraphs are given as input. We then show that deciding if the smallest preimage of a closed set is logarithmic in the size of the ground set is LOGSNP-hard if only the ground set is given. A special instance of this problem is to compute the dimension of a poset given its linear extension graph, that is conjectured to be in P.

The intent to show that the latter problem is LOGSNP-complete leads to several interesting questions and to the definition of the isometric hull, i.e., a smallest isometric subgraph containing a given set of vertices S. While for $|S| = 2$ an isometric hull is just a shortest path, we show that computing the isometric hull of a set of vertices is NP-complete even if $|S| = 3$. Finally, we consider the problem of computing the isometric hull number of a graph and show that computing it is Σ_2^P complete.

Keywords: Convex hull in graphs, complexity

1 Introduction

In the present paper we focus on two classical notions of metric subgraphs: convex subgraphs and isometric subgraphs. In both settings we study the complexity of computing a hull, i.e., a smallest metric subgraph containing a set of vertices and the hull number, i.e., a smallest set of vertices whose hull is the entire graph. While convex hulls and the convex hull-number are well studied and a subject of recent research (CHZ00; HJM$^+$05; CCG06; CHM$^+$10; DPRS10; CPRPdS13; DGK$^+$09; ACG$^+$13; AK16), isometric hulls and the corresponding hull-number seem to be as novel as it is natural.

We start introducing results concerning convexity. Let $G = (V, E)$ be a graph. A set $S \subseteq V$ is *convex* if, for any $u, v \in S$, any (u, v)-shortest path in G is included in S. The *(convex) hull* $\mathrm{conv}(S)$ of a set $S \subseteq V$ is the smallest convex set containing S. Note that since convex sets are closed under intersection, the convex hull is indeed unique. A *(convex) hull set* of G is a set $S \subseteq V$ such that $\mathrm{conv}(S) = V$. The *(convex) hull number* $\mathrm{hn}(G)$ of G is the size of a minimum hull set of G. The hull number was introduced in (ES85),

*Supported by grants DISTANCIA: ANR-17-CE40-0015 and GATO: ANR-16-CE40-0009-01

†Supported by ANR program "Investments for the Future" under reference ANR-11-LABX-0031-01, ANR MultiMod under reference ANR-17-CE22-0016 and the associated Inria team AlDyNet

ISSN 1365–8050

and since then has been the object of numerous papers. Application of graph convexity and the associated parameters include problems in contexts that require disseminating processes, such as contamination (BP98), marketing strategies (KKT15), spread of opinion (jDR09), spread of influence (KSZ14) and distributed computing (HP01). Most of the results on the hull number are about computing good bounds for specific graph classes, see e.g. (CHZ00; HJM$^+$05; CCG06; CHM$^+$10; DPRS10; CPRPdS13). While computing the convex hull of a set vertices can be done in polynomial time, computing the hull number of a graph G is known to be NP-complete (DGK$^+$09) and remains so if G is a bipartite graph (ACG$^+$13) and even if G moreover is a partial cube (AK16), i.e., an isometric subgraph of a hypercube.

In the context of the present paper it will be useful to see graph convexity from a more general perspective. Indeed, the function $\mathrm{conv} : 2^V \to 2^V$ sending vertex sets to their convex hulls belongs to a wider family of functions. Given a set A, a function $cl : 2^A \to 2^A$ is a *closure* if it satisfies the following conditions for all sets $X, Y \subseteq A$:

- $X \subseteq cl(X)$ (*extensive*)
- $X \subseteq Y \implies cl(X) \subseteq cl(Y)$ (*increasing*)
- $cl(cl(X)) = cl(X)$ (*idempotent*)

Clearly, conv is a closure. This leads to the following generalization of the hull number of a graph, that we call Minimum Generator Set (`MGS`):

> Given a set A, a polynomial time computable closure $cl : 2^A \to 2^A$ and an integer k, is there a set $X \subseteq A$ with $|X| \leq k$ such that $cl(X) = A$?

Since conv is a polynomial time computable closure, together with the results about the hull number it follows that `MGS` is NP-complete. However, in (AK16), it was conjectured that `MGS` is solvable in polynomial time if the set of images $\mathrm{Im}(cl)$ of cl is part of the input. We call this setting *Generating with large input* in Subsection 2.1. Our first result is to answer the conjecture from (AK16) in the positive for a larger class of functions. We call a function $f : 2^A \to 2^A$ a *pseudo-closure* if f satisfies $f(X \cup Y) = f(f(X) \cup f(Y))$ for any $X, Y \subseteq A$. The fact that closures are pseudo-closures is shown in Lemma 2.1. In Theorem 2.3 we devise an algorithm that takes a pseudo-closure f and finds minimum generating sets for all images of f in polynomial time (in $|A|$, $|\mathrm{Im}(f)|$ and the computation time of f). Thus, in particular we solve the conjecture of (AK16).

Pseudo-closures embody a large class of objects, most importantly closures. The latter are essentially the same as lattices (see Subsection 2.2.1). This interpretation yields that it can be decided in polynomial time if an element ℓ in a lattice L can be written as join of k join-irreducibles. Lattices encode many combinatorial objects and our algorithm can be applied there. For instance using lattices given in (Her94) and (SO96) our results yield polynomial-time algorithms to decide whether:

- a finite metric space has a hull set of size k, if all convex sets are given as input,
- a (semi)group can be generated by k elements, if all sub(semi)groups are given.

Then, in Section 2.2, we return to `MGS` in its original setting, i.e., with input (only) A, with the extra restriction that we consider only *atomistic* closures, i.e., closures $cl : 2^A \to 2^A$ that satisfy $cl(\{x\}) = \{x\}$ for all $x \in A$. We show that `MGS` is W[2]-hard (Corollary 2.6) already in this setting and its log-variant `LOGMGS`

> Given a set A, a polynomial time computable atomistic closure $cl : 2^A \to 2^A$ and an integer $k \leq \log(|A|)$, is there a set $X \subseteq A$ with $|X| \leq k$ such that $cl(X) = A$?

is LOGSNP-hard (Cor. 2.8).

It is then natural to restrict our attention again to the polynomial time computable atomistic closure conv. In this setting `LOGMGS` specializes to `LOGHULL NUMBER`(i). However, the complexity of `LOGHULL NUMBER` is open and might actually be polynomial time. Indeed, in (AK16), it was conjectured that, given a poset P together with all its linear extensions, the dimension of P can be computed in polynomial time. Moreover, the latter problem turns out to be an instance of `LOGHULL NUMBER` restricted to partial cubes, i.e., isometric subgraphs of hypercubes.

We study possible strategies to reduce to the hull number problem for partial cubes, in order to show that `LOGHULL NUMBER` for partial cubes is LOGSNP-complete. This takes us naturally to the second metric hull problem of concern in this paper. In order to introduce it, let again $G = (V, E)$ be a graph. A set $S \subseteq V$ is *isometric* if, for any $u, v \in S$, some (u, v)-shortest path of G is included in S. An *isometric hull* $\mathrm{iso}(S)$ of a set $S \subseteq V$ is a smallest isometric set containing S. Thus, this problem can be seen as a variation of the Steiner Tree problems. In particular, an isometric hull of two vertices is simply a shortest path and, already in this case, clearly there is no unique isometric hull. While computing a shortest path is easy, in Section 3 we show that computing an isometric hull of a set of vertices S is NP-complete even if $|S| = 3$ (Theorem 3.1). Note that above we have already mentioned partial cubes that are exactly isometric subgraphs of hypercubes. To illustrate the difficulty of the isometric hull computation, we present a set X of vertices of the hypercube Q_d for which we do not know if there is a partial cube containing it of size polynomial in $|X|$ and d. See Question 2.11 and the discussion thereafter.

Analogously to the convex hull set, an *isometric hull set* of G is a set $S \subseteq V$ such that $\mathrm{iso}(S) = V$. Similarly to the convex hull number, the *isometric hull number* $\mathrm{ihn}(G)$ of G is the size of a minimum isometric hull set of G. Clearly, since already computing isometric hulls is hard one can expect a higher complexity for computing the isometric hull number. Indeed, we show that computing the isometric hull number of a graph is Σ_2^P complete(ii) (Theorem 3.2).

2 Minimum generators of pseudo-closures

Let A be any set and $f : 2^A \to 2^A$ a *pseudo-closure*. Note, that setting $Y = X$ in $f(X \cup Y) = f(f(X) \cup f(Y))$ one obtains, that $f(X) = f(f(X))$ for all $X \subseteq A$, i.e., pseudo-closures are idempotent. In particular, pseudo-closures generalize closures in a different way than *preclosures*, which are not required to be idempotent. Finally, f is said *size-increasing* if $X \subseteq Y \implies |f(X)| < |f(Y)|$ or $f(X) = f(Y)$ for all $X, Y \subseteq A$. Let us first argue that in a way pseudo-closures are closures without the property of being extensive.

Lemma 2.1 *Let $f : 2^A \to 2^A$. The following are equivalent:*

(i) f is a closure,

(ii) f is an extensive pseudo-closure,

(iii) f is an extensive and size-increasing pseudo-closure.

(i) The `LOGHULL NUMBER` Problem takes a graph $G = (V, E)$ and $k \leq \log |V|$ as inputs and asks whether $\mathrm{hn}(G) \leq k$.

(ii) See, e.g., (PY96; DF12; FG06) for the definitions of the complexity classes LOGSNP, Σ_2^P and $W[2]$, respectively.

Proof: (i) $\Longrightarrow$ (iii): Let f be a closure, then it is extensive and increasing by definition and clearly also size-increasing. Since $f(X) \subseteq f(X \cup Y)$ and $f(Y) \subseteq f(X \cup Y)$, then $f(X) \cup f(Y) \subseteq f(X \cup Y)$. Hence, $f(f(X) \cup f(Y)) \subseteq f(f(X \cup Y)) = f(X \cup Y)$. Moreover, $X \subseteq f(X)$ and $Y \subseteq f(Y)$, therefore, $f(X \cup Y) \subseteq f(f(X) \cup f(Y))$. Therefore f is an extensive and size-increasing pseudo-closure.

(iii) $\Longrightarrow$ (ii): trivial.

(ii) $\Longrightarrow$ (i): Let f be an extensive pseudo-closure. As argued above f is idempotent. It remains to prove that f is increasing. Let $X \subseteq Y$. We have

$$f(X) \subseteq f(X) \cup f(Y) \subseteq f(f(X) \cup f(Y)) = f(X \cup Y) = f(Y),$$

where the second inclusion uses that f is extensive and the first equality uses that f is a pseudo-closure. $\square$

Just to give an example of a pseudo-closure, that is not a closure, consider:

Proposition 2.2 *Let $cl : 2^A \to 2^A$ be a closure and $\emptyset \neq X' \subseteq X \subseteq A$. Then $f(Y) := cl(Y \cup X) \setminus X'$ is an increasing pseudo-closure that is not extensive.*

Proof: To see that f is a pseudo-closure, we transform $f(f(Y) \cup f(Z)) = cl(cl(Y \cup X) \setminus X' \cup cl(Z \cup X) \setminus X' \cup X) \setminus X'$ which by $X' \subseteq X$ equals $cl(cl(Y \cup X) \cup cl(Z \cup X) \cup X) \setminus X'$ which since cl is extensive equals $cl(cl(Y \cup X) \cup cl(Z \cup X)) \setminus X'$. Now, since by Lemma 2.1 cl is a pseudo-closure, we can transform to $cl((Y \cup X) \cup (Z \cup X)) \setminus X'$ which equals $cl(Y \cup Z \cup X) \setminus X' = f(Y \cup Z)$.

It is easy to see that f is increasing and since $X' \not\subseteq f(X')$ it is not extensive. $\square$

We now turn our attention to the problem of generating images of a pseudo-closure. A set $X \subseteq A$ *generates* $f(X)$, and X is *minimum* (for f) if there is no set $Y \subseteq A$ such that $|Y| < |X|$ and $f(X) = f(Y)$.

2.1 Generating with large input

In this section, we design a dynamic programming algorithm that computes a minimum generator of any $H \in \mathrm{Im}(f) = \{Y \subseteq A \mid \exists X \subseteq A, Y = f(X)\}$. We assume that, for any $X' = X \cup \{w\} \subseteq A$ and given $f(X)$, determining $f(X')$ can be done in time c_f. A similar algorithm has been published previously in a different language and restricted to closure functions (NVRG05). Moreover, the approach in (NVRG05) is incremental which leads to a time-complexity of $O(c_f|A||\mathrm{Im}(f)|^2)$ (while no runtime analysis is presented there). We include our algorithm here to be self-contained but also because the complexity $O(c_f|A||\mathrm{Im}(f)|)$ of our algorithm is slightly better and we think that our presentation might be more accessible to our community.

Let us describe the algorithm informally. Every set $S \in \mathrm{Im}(f)$ is assigned to one of its generators stored in the variable $label(S)$. Initially, $label(S)$ may be any generator of S (for instance, S itself). The algorithm considers the sets in $\mathrm{Im}(f)$ in non decreasing order of their size and aims at refining their labels. More precisely, from a set $Y \in \mathrm{Im}(f)$ with generator $label(Y)$, the algorithm considers every set $f(R)$ generated by $R = label(Y) \cup \{z\}$ for some $z \in A$. If R is smaller than $label(f(R))$ then R becomes the new label of $f(R)$.

Theorem 2.3 *Algorithm $MinGen(A, f)$ computes a minimum generator of any $H \in \mathrm{Im}(f)$ in time $O(c_f(|A||\mathrm{Im}(f)|)^2)$.*

Moreover, if f is size-increasing, its time-complexity is $O(c_f|A||\mathrm{Im}(f)|)$.

Algorithm 1 $MinGen(A, f)$.

Require: A set A, a pseudo-closure $f : 2^A \to 2^A$, and the set $\mathrm{Im}(f)$.

```
1: For any H ∈ Im(f) \ {f(∅)}, set label(H) ← H and set label(f(∅)) ← ∅
2: Set Continue ← True
3: while Continue do
4:   Set Continue ← False
5:   for i = 1 to |A| do
6:     for Y ∈ Im(f), |Y| = i do
7:       for z ∈ A \ label(Y) do
8:         Set R ← {z} ∪ label(Y)
9:         Set H ← label(f(R))
10:        if |R| < |H| then
11:          label(f(R)) ← R and Continue ← True
12: return {label(Y) | Y ∈ Im(f)}
```

Proof: Let us first show that, at the end of the execution of the algorithm, $label(Y)$ is a minimum generator for every $Y \in \mathrm{Im}(f)$.

Clearly, $label(Y)$ is initially a generator of Y (Line 1). Moreover, $label(Y)$ can only be modified when it is replaced by R such that $f(R) = Y$ (Line 11). Let us show that $label(Y)$ is minimum.

For purpose of contradiction, let $Y \in \mathrm{Im}(f)$ such that the value L of $label(Y)$ at the end of the algorithm is not a minimum generator of Y. Let us furthermore assume that the size of a minimum generator Z of Y is minimum among all counterexamples. Hence, there is $Z \subseteq A$ with $|Z| < |L|$ and $f(Z) = f(L) = Y$ and Z is a minimum generator for Y. By line 1, we know that $Z \neq \emptyset$. Hence, let $w \in Z$ and $X = f(Z \setminus \{w\})$. Any minimum generator of X has size at most $|Z| - 1$. Therefore, by minimality of the size of a minimum generator of our counterexample, $label(X)$ is a minimum generator of X. In particular, $f(label(X)) = X = f(Z \setminus \{w\})$.

First, let us show that $w \notin label(X)$. Indeed, otherwise, $X = f(label(X)) = f(label(X) \cup \{w\}) = f(f(label(X)) \cup f(w)) = f(f(Z \setminus \{w\}) \cup f(w)) = f(Z) = Y$. Therefore, $f(Z \setminus \{w\}) = X = Y$, contradicting the fact that Z is a minimum generator for Y.

Consider the step when $label(X)$ receives its final value. After this step, $Continue$ must equal $True$. Therefore, there is another iteration of the $While$-loop. During this next iteration, there must be an iteration of the For-loop (Line 6) that considers $X \in \mathrm{Im}(f)$ and an iteration of the For-loop (Line 7) that considers $w \notin label(X)$. At this iteration, we set $H = f(label(X) \cup \{w\}) = f(f(label(X)) \cup f(w)) = f(f(Z \setminus \{w\}) \cup f(w)) = f(Z) = Y$. Because the size of the set $label(Y)$ is non increasing during the execution, the value L' of $label(Y)$ at this step is such that $|L| \leq |L'|$. In particular, $|label(X) \cup \{w\}| \leq |Z| < |L| \leq |L'|$. Therefore, during this execution (Line 11), $label(Y)$ should become equal to $label(X) \cup \{w\}$. Since, again, the size of the set $label(Y)$ is non increasing, it contradicts the fact that $label(Y) = L$ at the end of the algorithm.

First, note that since f is idempotent and $H \in \mathrm{Im}(f)$ in Line 1 we can set $label(H) \leftarrow H$, i.e., this can be done in constant time. Each iteration of the $While$-loop takes time $O(c_f|A||\mathrm{Im}(f)|)$. Moreover, each new iteration of this loop comes after a modification of some label in the previous iteration (Line 11, because $Continue$ is set to $True$). Since there are $|\mathrm{Im}(f)|$ labels and each of them will receive at most $|A|$ values (because the size of a label is not increasing), the time-complexity of the algorithm is

$O(c_f(|A||\mathrm{Im}(f)|)^2)$.

In case when f is size-increasing, we prove that each label contains its final value after the first iteration of the $While$-loop. So, there are exactly 2 iterations of this loop in that case and the time-complexity is $O(c_f|A||\mathrm{Im}(f)|)$ when f is size-increasing.

More precisely, we show that the label of $Y \in \mathrm{Im}(f)$ contains its final value just before Y is considered in the For-loop (line 6) of the first iteration. The proof is similar to the one of the correctness of the algorithm.

For purpose of contradiction, let $Y \in \mathrm{Im}(f)$ such that the value L of $label(Y)$ just before Y is considered in the For-loop (line 6) of the first iteration is not a minimum generator of Y. Moreover, let us consider such a counter example such that $|Y|$ is minimum. Hence, there is $Z \subseteq A$ with $|Z| < |L|$ and $f(Z) = f(L) = Y$. Let $w \in Z$ and $X = f(Z \setminus \{w\})$. Any minimum generator of X has size at most $|Z| - 1$. Moreover, because f is size-increasing, $|X| = |f(Z \setminus \{w\})| < |f(Z)| = |Y|$ (because $X \neq Y$ since their minimum generators have different sizes). Therefore, by minimality of the counterexample, $label(X)$ is a minimum generator of X just before X is considered in the For-loop of the first iteration, and moreover, X is considered before Y. In particular, $f(label(X)) = X = f(Z \setminus \{w\})$.

Similarly as before, $w \notin label(X)$. Hence, during the iteration (of the For-loops) that considers X and w, either $label(Y)$ must become $label(X) \cup \{w\}$ or $|label(Y)| \leq |label(X) \cup \{w\}| \leq |Z|$. In both cases, it is a contradiction since X is considered before Y and $|label(X) \cup \{w\}| < |L|$. □

From Th. 2.3, Lemma 2.1, and the preceding discussion we immediately get:

Corollary 2.4 *Let $cl : 2^A \to 2^A$ be a closure.* `MGS` *can be solved in $O(c_{cl}|A||\mathrm{Im}(cl)|)$ time.*

This confirms a conjecture of (AK16) (which could have probably also been extracted from (NVRG05)) and slightly improves the time-complexity of (NVRG05) in the case of a closure. Furthermore, it is well known and easy to see that for a closure $cl(X) = \bigcap_{X \subseteq Y \in \mathrm{Im}(cl)} Y$. Thus, c_{cl} is in $O(|\mathrm{Im}(cl)|)$ yielding a uniform bound of $O(|A||\mathrm{Im}(cl)|^2)$.

2.2 Generating with small input

In this section we show that for an atomistic closure $cl : 2^A \to 2^A$, the problem `MGS` is W[2]-hard with respect to the size of the solution, when only A is the input. Furthermore, `LOGMGS` is LOGSNP-hard. We then introduce the problem `COORDINATE REVERSAL`, show an equivalence with `HITTING SET`, and finally relate it to the hull number problem in partial cubes.

First, recall that a set $X \subseteq A$ is *closed* for a closure cl if $cl(X) = X$. For instance, the closed sets for the closure conv are exactly the convex sets.

We will present a reduction from the `HITTING SET` Problem that takes a ground set U and a set $\mathcal{X} \subseteq 2^U$ of subsets of U and an integer k as inputs and aims at deciding if there exists $K \subseteq U$ of size at most k such that $K \cap X \neq \emptyset$ for every $X \in \mathcal{X}$.

Proposition 2.5 `HITTING SET` *is L-reducible to* `MGS`*, i.e., there is a polynomial time reduction that preserves the size of optimal solutions.*

Proof: Let $(U, \mathcal{X})$ be an instance of `HITTING SET`, where we assume without loss of generality that for any two elements $u, v \in U$ there is a set $X \in \mathcal{X}$ such that $u \in X$ but $v \notin X$. We now define the following function $cl : 2^U \to 2^U$ by mapping $S \subseteq U$ to the minimal set $cl(S)$ that contains S and is the intersection of complements of elements of $\mathcal{X}$. Clearly, $cl(S)$ can be computed in polynomial time, since it is formed by passing through all members of $\mathcal{X}$ and including their complement into the intersection if necessary. Furthermore, it is easy to check, that cl indeed satisfies the axioms of a closure. Since for any two $u, v \in U$ there is a set $X \in \mathcal{X}$ such that $u \in X$ but $v \notin X$, we moreover get that cl is atomistic. Finally, note that for a subset $K \subset U$ we have that K hits every element of $\mathcal{X}$ if and only if for every $U \setminus X$ with $X \in \mathcal{X}$, there is an element $u \in K$ with $u \notin U \setminus X$, which is equivalent with K not being contained in any set $U \setminus X$, i.e., $cl(K) = U$. This concludes the proof. □

It was already known that the problem of determining the hull number is NP-complete in general graphs (DGK$^+$09). This result has been proved for bipartite graphs in (ACG$^+$13) and in even partial cubes in (AK16). However, all known reductions reduce variants of 3-SAT to the decision version of hull number. Here, we have shown reduction from a a combinatorial optimization problems are equivalent in the stronger sense of L-reductions, i.e., sizes of solutions are preserved. This has some immediate consequences.

Since `HITTING SET` is W[2]-complete (DF12), Proposition 2.5 gives:

Corollary 2.6 *`MGS` is W[2]-hard.*

Using results of Dinur and Steurer (DS14), by Proposition 2.5 we get:

Corollary 2.7 *`MGS` cannot be approximated to* $\big(1 - o(1)\big) \cdot \ln n$ *unless P=NP.*

In (FG06, Theorem 15.59) it is shown that `LOGHITTING SET` is LOGSNP (aka LOG[2]) complete, which with Proposition 2.5 gives:

Corollary 2.8 *`LOGMGS` is LOGSNP-hard.*

Concerning Proposition 2.5 we are not aware of a reduction the other way around and state here as an open question the complexity status of `LOGMGS`.

Since the conv-operator for graphs – sending sets of vertices to their convex hull – is an atomistic closure, we wonder if similar results can be proved for the hull number problem, or if this problem is essentially easier. For instance in (AMS$^+$16), a fixed parameter tractable algorithm to compute the hull number of any graph was obtained. But there the parameter is the size of a vertex cover. How about the complexity when parameterized by the size of a solution?

2.2.1 Lattices

Let us discuss the above from a lattice-theoretic point of view. A *lattice* is a poset $L = (X, \leq)$ such that for any two elements $x, y \in X$ there is a unique smallest element $x \vee y \in L$ such that $x, y \leq x \vee y$ (the *join* of x and y) and a unique largest element $x \wedge y \in L$ such that $x, y \geq x \wedge y$ (the *meet* of x and y). It is well known that closures correspond to lattices in the following way: Given a closure cl define the inclusion order on the closed sets, i.e., $L_{cl}(\mathrm{Im}(cl), \subseteq)$. Indeed, since L_{cl} the unique maximal element A and the intersection of closed sets is closed, it is easy to see that lattice L_{cl} where the meet of two closed sets is their intersection and their join is the closure of their union. On the other hand given a lattice L, an element $j \in L$ is called *join-irreducible* if $j = x_1 \vee \ldots \vee x_k \implies j \in \{x_1, \ldots, x_k\}$ for all $x_1, \ldots, x_k \in L$. If J is the set of join-irreducibles of L, we associate to every $\ell \in L$ the set $\downarrow\ell = \{j \in J \mid j \leq \ell\}$ of join-irreducibles below ℓ. One can see that L is isomorphic to the inclusion order $(\{\downarrow\ell \mid \ell \in L\}, \subseteq)$. The function $cl : 2^J \to 2^J$ with $cl(J') =\downarrow\bigvee J'$, where $\bigvee J' \in L$ denotes the join of J', is a closure and $(\{\downarrow\ell \mid \ell \in L\}, \subseteq)$ corresponds to the closed sets of cl.

The above correspondence specializes to atomistic closures. Indeed, a lattice is *atomistic* if its join-irreducibles are exactly its *atoms*, i.e., those elements directly above the minimum. The corresponding class of closures are the atomistic closures. Now we can state the following :

Question 2.9 *What are the atomistic lattices that come from the convex subgraphs of a graph?*

Clearly, these lattices are quite special, in particular any such lattice is entirely determined by its first two levels, since these correspond to vertices and edges of the graph. On the other hand, it is not clear what other properties such lattices enjoy. For instance, the graph in Figure 1 shows that convexity lattices of graphs are not *ranked* in general, i.e., not all maximal chains are of the same length.

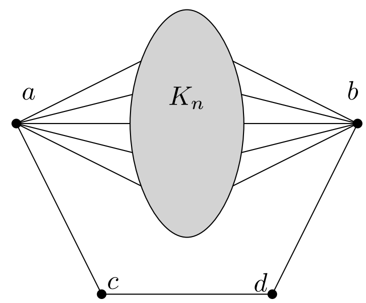

Fig. 1: A family of graphs with lattice of convex subgraphs being arbitrary far from ranked.

For example, *Ptolemaic* graphs are exactly those graphs whose lattice of convex subgraphs is lower locally distributive (FJ86). Also, in (AK16), the lattices of convex subgraphs of partial cubes were characterized. However, we do not know how to make use of these characterizations.

Let us now approach the hull number problem in partial cubes via a different reduction. Given a vertex $x \in Q_d$ of the hypercube of dimension d, let us consider x as a binary word with d bits (and two vertices are adjacent if they differ by exactly one bit). Given a coordinate $e \leq d$, let us denote by x_e the e^{th} bit of x. We call `COORDINATE REVERSAL` the following problem:

> Given a set X of vertices of the hypercube Q_d and an integer k, is there a subset $X' \subseteq X$, with $|X'| \leq k$ and such that for every coordinate e of Q_d there are vertices $x, y \in X'$ with $x_e \neq y_e$.

Proposition 2.10 *`COORDINATE REVERSAL` is L-equivalent to `HITTING SET`.*

Proof: Let $(U, \mathcal{X})$ be an instance of hitting set with $\mathcal{X} = \{X_1, \ldots, X_d\}$. Let us furthermore assume that $U \in \mathcal{X}$ and that for any two distinct $u, v \in U$ there is an $X \in \mathcal{X}$ such that $u \in X$ but $v \notin X$. Clearly, these assumptions do not change the complexity of `HITTING SET`.

Now, define a new instance $(U^+, \mathcal{X}^+)$, where $U^+ = U \cup \{x\}$ is U extended with one vertex x and $\mathcal{X}^+ = \{X_i, U^+ \setminus X_i \mid 1 \leq i \leq d\}$ is $\mathcal{X}$ together with the set of complements with respect to the new ground set.

If H is a hitting set of size k of $(U, \mathcal{X})$, then $H \cup \{x\}$ is a hitting set of size $k+1$ of $(U^+, \mathcal{X}^+)$. Conversely, if I is a hitting set of size $k+1$ of $(U^+, \mathcal{X}^+)$, then since $U \in \mathcal{X}$ we have $\{x\} \in \mathcal{X}^+$ and therefore $x \in I$. Hence $I \setminus \{x\}$ is a hitting set of size k of $(U, \mathcal{X})$. Thus, the new instance $(U^+, \mathcal{X}^+)$ has a hitting set of size $k+1$ if and only if the old one has one of size k.

Let us now show how to interpret the hitting set instance $(U^+, \mathcal{X}^+)$ as an equivalent instance of COORDINATE REVERSAL. We encode U^+ as a subset of vertices of the hypercube of dimension $d = \frac{|\mathcal{X}^+|}{2}$ by associating to every $v \in U^+$ a vector $\overline{v}$ with coordinates $\overline{v}_i = \begin{cases} 1, & v \in X_i, \\ -1, & v \in U^+ \setminus X_i. \end{cases}$ Note that by the assumption that for any two distinct $u, v \in U$ there is an $X \in \mathcal{X}$ such that $u \in X$ but $v \notin X$, the mapping $v \mapsto \overline{v}$ is a bijection. Now, under this mapping a hitting set I of $(U^+, \mathcal{X}^+)$ corresponds to a solution of COORDINATE REVERSAL, i.e., a subset $\overline{I} \subseteq \overline{U}^+$ such that each coordinate is reversed, i.e., appears once positively and once negatively. Conversely, a solution to COORDINATE REVERSAL corresponds to a subset of U^+ that for any coordinate i contains an element $v \in X_i$ and an element $w \in (U^+ \setminus X_i)$. Thus, it is a hitting set of $(U^+, \mathcal{X}^+)$.

Now, take conversely an instance (X, k) of COORDINATE REVERSAL in Q_d. We construct an instance $(X, \mathcal{X})$ of HITTING SET. Define for $1 \leq i \leq d$ the set $X_i := \{x \in X \mid x_i = +\}$ and let $\mathcal{X} = \{X_i, X \setminus X_i \mid 1 \leq i \leq d\}$. As in the previous paragraph solution of size k of COORDINATE REVERSAL are in bijection to solutions of size k of HITTING SET in $(X, \mathcal{X})$. □

Now, in (AK16), it is shown that, in a partial cube $G = (V, E) \subseteq Q_d$, HULL NUMBER coincides with COORDINATE REVERSAL for V and Q_d, therefore, HULL NUMBER in partial cubes is a special case of COORDINATE REVERSAL. In order to L-reduce COORDINATE REVERSAL to partial cube hull number along the lines of Proposition 2.10, it would be interesting to check if, given a subset $V' \subseteq Q_d$, a smallest partial cube containing V' has to be polynomial in $|V'| + d$. Moreover it is important to maintain the same solution size with respect to COORDINATE REVERSAL. In (FG06, Theorem 15.59) it is shown that LOGHITTING SET is LOGSNP (aka LOG[2]) complete. Hence, this would show LOGSNP-completeness of LOGHULL NUMBER for partial cubes, one instance of which is calculating the dimension of a poset given its linear extensions, see (AK16). So as a first step we wonder:

Question 2.11 *Let X be a set of vertices of the hypercube Q_d, does there exist an isometric subgraph G of Q_d, containing X, such that $|G|$ is polynomial in $|X| + d$?*

Let M_k be a $(0,1)$-matrix whose columns are all the $(0,1)$-vectors of length k. Now, $X_k \subseteq Q_{2^k}$ is defined as the set of rows of M_k. We do not know the answer to Question 2.11 for the set X_k.

These questions lead to the problem of computing a small isometric subgraph containing a given set of vertices, which is the subject of the next section.

3 Isometric hull

We recall the definitions related to the isometric hull from the introduction. Let $G = (V, E)$ be a graph. For any $v, u \in V$, let $\mathrm{dist}_G(u, v)$ denote the *distance* between u and v, i.e., the minimum number of edges of a path between u and v in G. A subgraph $H = (V', E')$ (i.e., $V' \subseteq V$ and $E' \subseteq (V' \times V') \cap E)$) of G is *isometric* if $\mathrm{dist}_G(u, v) = \mathrm{dist}_H(u, v)$ for any $u, v \in V'$. Given $S \subseteq V$, an *isometric hull* of S is any subgraph $H = (V', E')$ of G such that $S \subseteq V'$ and H is isometric. An isometric hull $H = (V', E')$ of S is *minimum* if $|V'|$ is minimum, i.e., there are no isometric hulls of S with strictly less vertices. Note that a set S may have several minimum isometric hulls. As an example, consider the 4-node cycle $C_4 = (a, b, c, d)$: the subgraphs induced by $\{a, b, c\}$ and $\{a, d, c\}$ are minimum isometric hulls of

$S = \{a, c\}$. More generally, for any $S = \{u, v\}$, inclusion-minimal isometric hulls of S are any shortest path between u and v. So, if $|S| = 2$, all minimal isometric supergraphs of S are of the same size and computing a minimum isometric hull of S is easy. For $|S| > 2$ it is easy to find examples with minimal isometric supergraphs that are not minimum. We show below that computing a minimum isometric hull is NP-complete if $|S| > 2$. An *isometric hull set* $S \subseteq V$ of G is any subset of the vertices such that G is the (unique) minimum isometric hull of S.

This section is devoted to prove the following theorems.

Theorem 3.1 *Given an n-node bipartite graph $G = (V, E)$, $S \subseteq V$ and $k \in \mathbb{N}$, deciding whether there exists an isometric hull H of S with $|V(H)| \leq k$ is NP-complete, even if $|S| = 3$ and $k = n - 1$. In particular, deciding whether a set S of vertices is an isometric hull set of a graph is co-NP-complete.*

Theorem 3.2 *Given a graph G and $k \in \mathbb{N}$, the problem of deciding whether G admits an isometric hull set of size at most k is Σ_2^P-complete.*

Let us start with an easier result where neither the size of the input set S nor the size k of the isometric hull are constrained, but where the class of bipartite graphs is restricted to have diameter 3.

Lemma 3.3 *Given $G = (V, E)$ bipartite with diameter 3, $S \subseteq V$ and $k \in \mathbb{N}$, deciding if there exists an isometric hull I of S with $|V(I)| \leq k$ is NP-complete.*

Proof: The problem is clearly in NP since testing whether a subgraph is isometric can be done in polynomial-time.

To prove that the problem is NP-hard, let us present a reduction from the `HITTING SET` Problem that takes a ground set $U = \{u_1, \cdots, u_n\}$ and a set $\mathcal{X} = \{X_1, \cdots, X_m\} \subseteq 2^U$ of subsets of U and an integer k as inputs and aims at deciding if there exists $K \subseteq U$ of size at most k such that $K \cap X_j \neq \emptyset$ for every $j \leq m$. Note that we may assume that at least two sets of $\mathcal{S}$ are disjoint (up to adding a dummy vertex in U and a set restricted to this vertex).

Let us build the graph G as follows. We start with the incidence graph of $(U, \mathcal{X})$, i.e., the graph with vertices $U \cup \mathcal{X} = \{u_1, \cdots, u_n, X_1, \cdots, X_m\}$ and edges $\{u_i, X_j\}$ for every $i \leq n$, $j \leq m$ such that $u_i \in X_j$. Then add a vertex x adjacent to every vertex in U and a vertex y adjacent to every vertex in $\mathcal{X}$. Note that G has diameter 3. Finally, let $S = \{x\} \cup \mathcal{X}$.

We show that $(U, \mathcal{X})$ admits a hitting set of size k if and only if S has an isometric hull of size $k+m+2$. Note that, because at least two sets are disjoint, y must be in any isometric hull of S (to ensure that these sets are at distance two). Moreover, for every set containing (at least) x, y and $\mathcal{X}$, all distances are preserved but possibly the ones between x and some vertices of $\mathcal{X} \cup \{y\}$. We show that I is an isometric hull of S if and only if $K = V(I) \setminus (S \cup \{y\})$ is a hitting set of $(U, \mathcal{X})$. Indeed, for every $j \leq m$, the distance between X_j and x equals 2 in I if and only if K contains a vertex u_i adjacent to X_j, i.e., $K \cap X_j \neq \emptyset$ for every $j \leq m$. □

Now, let us consider a restriction of Theorem 3.1 in the case $k = n - 1$ (without constraint on $|S|$). For this purpose, we present a reduction from 3-SAT.

Preliminaries: the triangle gadget T_γ. Let us first describe a gadget subgraph, parameterized by an odd integer γ [(iii)], for which only 3 vertices generate the whole graph. That is, we describe a graph T_γ with

[(iii)] γ is set odd only to avoid parity technicality in the computation of the distances.

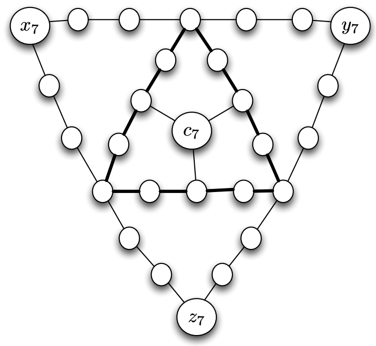

Fig. 2: Example of T_7. Edges are bold only to better distinguish the different "levels".

size $\Theta(\gamma^2)$ such that there are 3 vertices (called the corners) whose minimum isometric hull is the whole graph. Moreover, some vertex (called the center) of T_γ is "far" (at distance $\Theta(\gamma^2)$) from the corners.

Let $\gamma \in \mathbb{N}^*$ be any odd integer. Let us define recursively a γ-triangle with *corners* $\{x_\gamma, y_\gamma, z_\gamma\}$ and *center* c_γ as follows.

A 3-triangle T_3 is a $K_{1,3}$ where the big bipartition class $\{x_3, y_3, z_3\}$ are the corners and the center is the remaining vertex c_3.

Let $\gamma > 3$ be an odd integer and let $T_{\gamma-2}$ be a $(\gamma-2)$-triangle with corners $\{x_{\gamma-2}, y_{\gamma-2}, z_{\gamma-2}\}$ and center $c_{\gamma-2}$. The γ-triangle T_γ is obtained as follows. First, let U_γ be the cycle of length $3(\gamma-1)$ with vertices $x_\gamma, y_\gamma, z_\gamma$ that are pairwise at distance $\gamma-1$. For any $u, v \in \{x_\gamma, y_\gamma, z_\gamma\}$, let a_{uv} be the vertex at distance $\lfloor \gamma/2 \rfloor$ from u and v in U_γ. The graph T_γ is obtained from U_γ and $T_{\gamma-2}$ by identifying $x_{\gamma-2}, y_{\gamma-2}, z_{\gamma-2}$ with a_{x_γ,y_γ}, a_{y_γ,z_γ} and a_{z_γ,x_γ}, respectively. The corners of T_γ are x_γ, y_γ and z_γ, and the center c_γ of T_γ is the center $c_{\gamma-2}$ of $T_{\gamma-2}$. Note that the center c_γ of T_γ is the center c_3 of the "initial triangle" T_3. An example is depicted on Figure 2.

The following claim can be easily proved by induction on γ. The second statement also comes from the fact that $T_{\gamma-2}$ is an isometric subgraph of T_γ.

Claim 1 *For any odd integer $\gamma > 3$, let T_γ with corners $S = \{x_\gamma, y_\gamma, z_\gamma\}$*

- $|V(T_\gamma)| = |V(T_{\gamma-2})| + 3(\gamma-2) = \Theta(\gamma^2)$*;*
- *the (unique) isometric hull of S is T_γ;*
- *the distance between any two corners in T_γ is $\gamma - 1$;*
- *the distance between the center and any corner in T_γ is $\sum_{i=1}^{\lceil \frac{\gamma}{2} \rceil} i = \Theta(\gamma^2)$;*
- *since T_γ is planar and all faces are even, T_γ is bipartite.*

Lemma 3.4 *Given a bipartite n-node graph $G = (V, E)$, $X \subseteq V$, deciding whether there exists an isometric hull I of X with $|V(I)| < n$ is NP-complete.*

Proof: The problem is clearly in NP since testing whether a subgraph is isometric can be done in polynomial time. To prove that the problem is NP-hard, let us present a reduction from 3-SAT.

We first present a reduction proving that the problem of deciding whether there exists an isometric hull with size at most k (given as input) is NP-complete. Then, in the last paragraph, we prove the lemma (i.e., for $k = n - 1$).

Let Φ be a CNF formula with n variables $v_1, \cdots, v_n$ and m clauses $C_1, \cdots, C_m$. Let us describe a graph $G_0 = (V, E)$, $S \subseteq V$ and $k \in \mathbb{N}$ such that an isometric hull of S has size at most k if and only if Φ is satisfiable.
Let α, β and γ be three integers satisfying: α and β are even and γ is odd and

$$m << 2\alpha < 2\beta < \gamma < 2(\alpha + \beta).$$

The graph G_0 is built by combining some variable gadgets, clause-gadgets and adding some paths connecting the variable gadgets with some particular vertex r.

Variable gadget. For any $1 \leq i \leq n$, the variable gadget V^i consists of a cycle of length 4α with four particular vertices d_i, n_i, p_i, g_i such that d_i and g_i are antipodal, i.e., at distance 2α of each other, n_i and p_i are antipodal, and $\text{dist}_{V^i}(d_i, n_i) = \text{dist}_{V^i}(d_i, p_i) = \text{dist}_{V^i}(g_i, n_i) = \text{dist}_{V^i}(g_i, p_i) = \alpha$. Let P^i and N^i be the shortest path between d_i and g_i in V^i passing through p_i and n_i, respectively.
Clause gadget. For any $1 \leq j \leq m$ and clause $C_j = (\ell_i \wedge \ell_k \vee \ell_h)$ (where ℓ_i is the literal corresponding to variable v_i in clause C_j), the clause gadget C^j is a γ-triangle with corners denoted by ℓ_i, ℓ_k, ℓ_h (abusing the notation, we identify the corner-vertices and the literals they correspond to) and center denoted by c^j.
The graph G_0. The graph G_0 is obtained as follows. First, let us start with disjoint copies of V^i, for $1 \leq i \leq n$, and of C^j, for $1 \leq j \leq m$. Then, add one vertex r and, for any $1 \leq i \leq n$, add a path $P(r, d_i)$ of length β between r and d_i and a path $P(r, g_i)$ of length β between r and g_i (these $2n$ paths are vertex-disjoint except in r). Finally, for any $1 \leq j \leq m$ and any literal ℓ_i in the clause C_j, let us identify the corner ℓ_i of C^j with vertex p_i (in the variable gadget V^i) if variable v_i appears negatively in C_j (i.e., if $\ell_i = \bar{v}_i$) and identify the corner ℓ_i of C^j with vertex n_i if variable v_i appears positively in C_j (i.e., if $\ell_i = v_i$). Let us emphasize that, if variable v_i appears positively (negatively) in C_j, then a corner of C^j is identified with a vertex of the path N^i (P^i). By the parity of α, β and γ, G_0 is clearly bipartite. An example is depicted in Figure 3.
The set S. Finally, let $S = \{r\} \cup \{d_i, g_i \mid 1 \leq i \leq n\}$.

We first show that S has an isometric hull of size at most $k := n(\alpha + 2\beta) + m\gamma$ in G_0 if and only if Φ is satisfiable.

Claim 2 *S has an isometric hull of size at most $k := n(\alpha+2\beta)+m\gamma$ in G_0 if and only if Φ is satisfiable.*

Proof of the claim.
Let us start with some simple observations (following from the constraints on α, β and γ):

1. For any $1 \leq i \leq n$, $\text{dist}_{G_0}(d_i, g_i) = \text{dist}_{G_0}(p_i, n_i) = 2\alpha$ and there are exactly two shortest paths P^i and N^i between d_i and g_i. Intuitively, choosing P^i (resp., N^i) in the isometric hull will correspond to a positive (resp., negative) assignment of variable v_i.

2. For any $1 \leq i \leq n$, $\text{dist}_{G_0}(r, d_i) = \text{dist}_{G_0}(r, g_i) = \beta$ and $P(r, d_i)$ (resp., $P(r, g_i)$) is the unique shortest path between r and d_i (resp., between r and g_i). In particular, each of these paths has to be in any isometric hull of S.

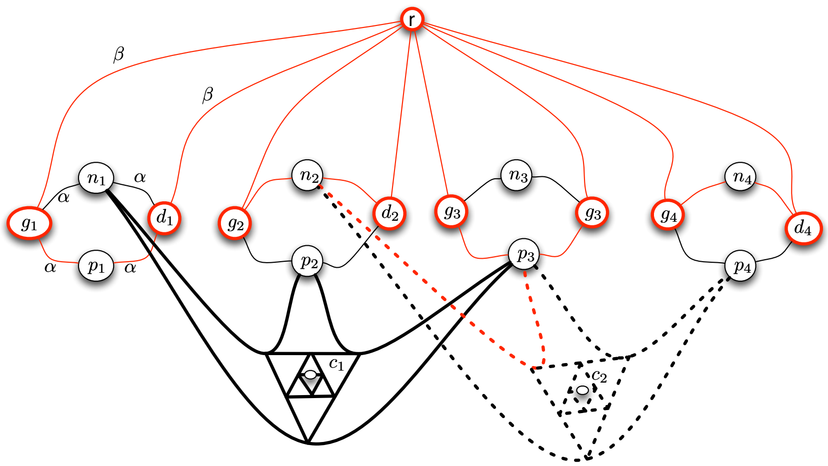

Fig. 3: An example for the graph G_0 of the reduction of Lemma 3.3 for Φ with variables v_1, v_2, v_3, v_4 and two clauses $C_1 = v_1 \vee \bar{v}_2 \vee \bar{v}_3$ and $C_2 = v_2 \vee \bar{v}_3 \vee \bar{v}_4$. The solid bold lines represent the clause gadget C^1 which is a γ-triangle (only few “levels” are depicted), and the dotted bold lines represent the clause gadget C^2. The vertices c_1 and c_2 denote the centers of C^1 and C^2 respectively. Red vertices are the ones of the set $S = \{r, d_1, g_1, d_2, g_2, d_3, g_3, d_4, g_4\}$. Finally, the red subgraph is the isometric hull of S corresponding to the truth assignment $(v_1, v_2, v_3, v_4) = (1, 0, 1, 0)$. The graph G is obtained from G_0 by adding a vertex q adjacent to c_1 and c_2.

3. For any $1 \leq i < j \leq n$, $\text{dist}_{G_0}(d_j, d_i) = \text{dist}_{G_0}(d_j, g_i) = 2\beta$ and the unique shortest path between them is the one going through r (because $2\beta < \gamma$).

4. For any $1 \leq j \leq m$ and clause $C_j = (\ell_i \vee \ell_k \vee \ell_h)$, $\text{dist}_{G_0}(\bar{\ell}_i, \bar{\ell}_h) = \text{dist}_{G_0}(\bar{\ell}_h, \bar{\ell}_k) = \text{dist}_{G_0}(\bar{\ell}_i, \bar{\ell}_k) = \gamma - 1$ (where $\bar{\ell}_i$ denotes n_i if v_i appears positively in C_j and it denotes p_i otherwise). This is because $\gamma < 2(\alpha + \beta)$ and the unique shortest path between these vertices is the one in C^j.

5. For any $1 \leq h < k \leq n$, $\ell_h \in \{n_h, p_h\}$ and $\ell_k \in \{n_k, p_k\}$ such that literals $\bar{\ell}_h$ and $\bar{\ell}_k$ do not appear in a same clause, then $\text{dist}_{G_0}(\ell_h, \ell_k) = 2(\alpha + \beta)$ (because $2\beta < \gamma$). In particular, every shortest path between ℓ_h and ℓ_k does not cross any clause gadget.

6. Let $1 \leq h < k \leq n$, $\ell_h \in \{n_h, p_h\}$ and $\ell_k \in \{n_k, p_k\}$ appearing in a clause C_j. For any vertex u in the shortest path between ℓ_h and ℓ_k in the clause gadget C^j, and for any $v \in \{d_i, g_i\}$ for some $i \notin \{h, k\}$, $\text{dist}_{G_0}(u, v) \leq \gamma/2 + \alpha + 2\beta$. In particular, any shortest path between u and v does not pass through the third corner (different from ℓ_h and ℓ_k) of C^j. This is because $\gamma > 2\beta$.

- First, let us show that, if Φ is satisfiable, there is an isometric hull of S with size at most $n(\alpha + 2\beta) + m\gamma$ in G_0. Indeed, consider a truth assignment of Φ and let H be the subgraph defined as follows. For any $1 \leq i \leq n$, the paths $P(r, d_i)$ and $P(r, g_i)$ belong to H. For any $1 \leq i \leq n$, if v_i is assigned to True, add P^i in H, and add N^i otherwise. Finally, for any $1 \leq j \leq m$, for any two corners of the clause gadget C^j, if these two corners are in H, then add to H the path of length $\gamma - 1$ between them in C^j.

Clearly, H contains all vertices in S. To show that H is isometric, let us first show that any clause gadget has at most two corners in H. Let $x \in \{n_i, p_i\}$ be a corner of a clause gadget C^j which is in H. If $x = n_i$ (resp., $x = p_i$) is in H, it implies that the path N^i (resp., P^i) has been added in H. Therefore, the variable v_i is assigned to False (resp., to True) in the assignment. On the other hand, if $x = n_i$ (resp., $x = p_i$) is a corner of C^j, it means that the variable v_i appears positively (resp., negatively) in clause C_j. Altogether, this implies that, in the assignment, Variable v_i cannot satisfy clause C_j. Since the assignment satisfies Φ, each clause must be satisfied by at least one of its variables, which implies that at least one of its corners is not in H.

To sum-up H consists of the $2n$ paths from r to the vertices d_i, g_i, $1 \leq i \leq n$, of exactly on path P^i or N^i, $1 \leq i \leq n$, and of at most one path between two corners of C^j, $1 \leq j \leq m$. Hence, H has at most $n(\alpha + 2\beta) + m\gamma$ vertices. The fact that H is isometric comes from the above observations on the shortest paths in G_0.

- To conclude, let us show that, if Φ is not satisfiable, then any isometric hull of S, in G_0, has size at least $n(\alpha + 2\beta) + \Omega(\gamma^2)$, i.e., strictly larger than $n(\alpha + 2\beta) + m\gamma$ (since $\gamma >> m$).

  As already mentioned, any isometric hull of S has to contain each of the paths $P(r, d_i)$ and $P(r, g_i)$ and at least one of the paths P^i and N^i, $1 \leq i \leq n$. This consists of at least $n(\alpha + 2\beta)$ vertices. It remains to show that, for any isometric hull H of S, there exists $j \leq m$ such that the entire clause gadget C^j belongs to H. This will consist of $\Omega(\gamma^2)$ additional vertices.

  Let H be an isometric hull of S. For any $1 \leq i \leq n$, at least P^i or N^i belongs to H. If P^i belongs to H, assign variable v_i to True and assign it to False otherwise. Since Φ is not satisfiable, there is a clause $C_j = (\ell_i \vee \ell_k \vee \ell_h)$ that is not satisfied. Let $u \in \{i, h, k\}$. If v_u appears positively (resp., negatively) in C_j, then v_u is assigned to False (resp., to True) since C_j is not satisfied. Moreover, it implies that P^u (resp., N^u) belongs to H. By construction, the corner ℓ_u of C^j belongs to P^u (resp., N^u) and so, ℓ_u belongs to H. Hence, all the three corners of C^j belong to H and it is easy to see that the entire C^j must belong to H since, recursively, all paths in C^j have to be added to preserve the fact that H is isometric.

$\diamond$

To prove Lemma 3.4, from G_0, let us build a graph G such that S (which remains unchanged) has an isometric hull of size at most $|V(G)| - 1$ if and only if Φ is satisfiable. The graph G is obtained from G_0 by adding to it a gadget (one vertex) that will ensure that if the center of one clause gadget belongs to an isometric hull (recall that, in the first part of the proof, this is the case if and only if Φ is not satisfiable), then all vertices of the graph will have to be in the isometric hull.

Let us add to G_0 one vertex q adjacent to all the centers of the clause gadgets in G_0. Note that the obtained graph G is bipartite.

- If Φ is satisfiable, consider any truth assignment and let H be the subgraph (as defined in the previous proof) that consists of the $2n$ paths from r to the vertices d_i, g_i, $1 \leq i \leq n$, of exactly on path P^i or N^i, $1 \leq i \leq n$, and of at most one path between two corners of C^j, $1 \leq j \leq m$. Because each center of a clause gadget is at distance $\Omega(\gamma^2)$ from any vertex of H, the addition of vertex q in the graph has not modified the distances between vertices in H. Therefore, H is isometric (as in the first part of the proof) and S is not an isometric hull set of G.

- If Φ is not satisfiable, then we prove that S is an isometric hull set of G.

  If Φ is not satisfiable, then it can be shown as previously that any isometric hull H of S contains at least one of the paths P^i or N^i for any $1 \leq i \leq n$, and an entire clause gadget C^j for some $j \leq m$. In particular, the center c of C^j belongs to H. Now, let $1 \leq i \leq n$ and let us assume that N^i is not in H. Note that, in this case, P^i must be in H. Therefore $p_i \in V(H)$ and $n_i \notin V(H)$. By assumption, there is a clause C_z that contains v_i positively and does not contain v_i negatively. By construction, the clause gadget C^z has n_i as a corner and p_i is not a corner of C^z. Note that $z \neq j$ since all corners of C^j belong to H. Now, any shortest path between c and d_i must go from c to q then to the center of the clause gadget C_z and then through n_i to d_i. In particular, n_i must be added to the isometric hull. It can be proved similarly that if P^i does not belong to H, then p_i has to be included into H.

  Altogether, we just proved that, for any $1 \leq i \leq n$, both n_i and p_i belong to H. It is easy to conclude that $H = G$. Indeed, in particular, any clause gadget has all its corner in H and therefore, the entire clause gadget must be included in H.

$\square$

Finally, to prove Th. 3.1, we will reduce the problems that we proved NP-complete in Lemma 3.4 to the same problems in the case $|S| = 3$. Note that, in both reductions of Lemmas 3.3 and 3.4, the distance between any pair of vertices of S is even, so both problems are NP-complete with this extra constraint.

Proof: of Th. 3.1. Let G, S, k be an instance of the problem of finding an isometric hull of S with size at most k. Let $n = |V(G)|$ and let $S = \{u_1, \cdots, u_s\}$. Moreover, let us assume that the distance between any pair of vertices of S is even.

Let G' be obtained as follows. Start with a copy of G, a path $P = (x = v_0, v_1, w_1, v_2, w_2, \cdots, w_{s-1}, v_s, v_{s+1} = y\}$ and a vertex z. Let $n' = n$ if n even and $n' = n + 1$ otherwise. For any $1 \leq i \leq s$, add a path of length n' between v_i and u_i and add a path of length n' between z and u_i. Note that G is an isometric subgraph of G' and that G' is bipartite. Finally, let $S' = \{x, y, z\}$.

Any isometric hull H of S' has to contain the (unique) shortest path P between x and y. Hence, for any $1 \leq i \leq s$, H contains v_i and therefore must contain the (unique) shortest path P_i between v_i and z (of length $2n'$). In particular H contains u_i for any $1 \leq i \leq s$. Since G is isometric in G', then the subgraph induced by the vertices in $V(G) \cap V(H)$ is an isometric hull of S in G.

Therefore, S admits an isometric hull of size at most k in G if and only if S' admits an isometric hull of size $k + |V(G') \setminus V(G)| = k + 2sn' + s + 1$. In particular, if $k = n - 1$, then the formula gives $|V(G')| - 1$. $\square$

Note that deciding whether a set S of vertices is not an isometric hull set of an n-node graph is equivalent to decide whether S has an isometric hull of size $< n$. Therefore:

Corollary 3.5 *Deciding whether a set of vertices is an isometric hull set is co-NP-complete.*

Finally, to prove Theorem 3.2, we present a reduction from the problem of satisfiability for quantified Boolean formulas with 2 alternations of quantifiers $QSAT_2$ which is well known to be Σ_2^P-complete (Pap07). The reduction is an adaptation of the one presented in the proof of Theorem 3.1.

In the proof below, we will use the following easy claim to force some vertices to belong to any isometric hull set.

Claim 3 *For any graph $G = (V, E)$ and any vertex $v \in V$ such that $G \setminus v$ is isometric, we have that v has to belong to any isometric hull set of G. In particular, any one-degree vertex of G has to belong to any isometric hull set of G.*

Proof: of Theorem 3.2. First, the problem is in Σ_2^P. Indeed, by Theorem 3.1, a certificate S (i.e., a set of vertices which is supposed to be an isometric hull set of G) can be checked using an NP oracle.

To prove that it is hard for Σ_2^P, let us give a reduction from $QSAT_2$ where the input is a Boolean formula Φ on two sets $X = \{x_1, \cdots, x_{n_x}\}$ and $Y = \{y_1, \cdots, y_{n_y}\}$ of variables and the question is to decide whether $\exists X, \forall Y, \Phi(X, Y)$. We moreover may assume that Φ is 3-DNF formula, i.e., the disjunction of conjunctive clauses $C_1, \cdots, C_m$ with 3 variables each. We also assume that, for each variable, some clause contains it positively and some clause contains it negatively, and that no variable appears positively and negatively in some clause.

Let us describe a graph $G = (V, E)$ and $k \in \mathbb{N}$ such that there exists an isometric hull set S of size at most k if and only if $\exists X, \forall Y, \Phi(X, Y)$.

Let α, β and γ be three integers satisfying: α and β are even and γ is odd and

$$m << 2\alpha < 2\beta < \gamma < 2(\alpha + \beta).$$

The graph G is built by combining some variable gadgets, clause gadgets and adding some paths connecting the variable gadgets with some particular vertex r. We emphasize the differences with the graph proposed in previous subsection.

Variable gadget. For any $1 \leq i \leq n_y$, the variable gadget Y^i consists of a cycle of length 4α with four particular vertices $d_i^y, n_i^y, p_i^y, g_i^y$ such that d_i^y and g_i^y are antipodal, i.e., at distance 2α, n_i^y and p_i^y are antipodal, and $\text{dist}_{Y^i}(d_i^y, n_i^y) = \text{dist}_{Y^i}(d_i^y, p_i^y) = \text{dist}_{Y^i}(g_i^y, n_i^y) = \text{dist}_{Y^i}(g_i^y, p_i^y) = \alpha$. Let P_y^i (resp., N_y^i) be the shortest path between d_i^y and g_i^y in Y^i passing through p_i^y and n_i^y, respectively.

Moreover, let us add a one-degree vertex dd_i^y adjacent to d_i^y and a one-degree vertex gg_i^y adjacent to g_i^y (This is the first difference with the previous section). By the above claim both vertices dd_i^y and gg_i^y have to belong to any isometric hull set of G.

For any $1 \leq i \leq n_x$, the variable gadget X^i, the vertices $d_i^x, n_i^x, p_i^x, g_i^x, dd_i^x, gg_i^x$ and the paths P_x^i and N_x^i are defined similarly.

Clause gadget. For any $1 \leq j \leq m$ and clause $C_j = (\ell_i \wedge \ell_k \wedge \ell_h)$, the clause gadget C^j is a γ-triangle with corners denoted by ℓ_i, ℓ_k, ℓ_h (abusing the notation, we identify the corner-vertices and the literals they correspond to) and center denoted by c^j.

The graph G. The graph G is obtained as follows. First, let us start with disjoint copies of X^i, for $1 \leq i \leq n_x$, of Y^i for $1 \leq i \leq n_y$, and of C^j, for $1 \leq j \leq m$. Then, add one vertex r and, for any $1 \leq i \leq n_x$, add a path $P(r, d_i^x)$ of length β between r and d_i^x and a path $P(r, g_i^x)$ of length β between r and g_i^x (these $2n_x$ paths are vertex-disjoint except in r). Similarly, for any $1 \leq i \leq n_y$, add a path $P(r, d_i^y)$ of length β between r and d_i^y and a path $P(r, g_i^y)$ of length β between r and g_i^y (these $2n_y$ paths are vertex-disjoint except in r).

Then, add a one-degree vertex r' adjacent to r (This is another difference with the previous section). Again, by the above claim, vertex r' has to belong to any isometric hull set of G.

A main difference with the construction in the previous section is the way the clause gadgets are connected to the variable gadgets. Intuitively, this is because we consider now a DNF formula while previously it was a CNF formula.

For any $1 \leq j \leq m$ and any literal ℓ_i in the clause C_j (corresponding to some variable $v_i \in X \cup Y$), let us identify the corner ℓ_i of C^j with vertex p_i (in the variable gadget of variable v_i) if variable v_i appears positively in C_j and identify the corner ℓ_i of C^j with vertex n_i if variable v_i appears negatively in C_j. Let us emphasis that, contrary to the previous section, if variable v_i appears positively (resp., negatively) in C_j, then a corner of C^j is identified with a vertex of the path P^i (resp., N^i).

Finally, add a vertex q adjacent to all centers of the clause gadgets.

The last touch. Let δ be any odd integer larger than the diameter of the graph built so far. For any $1 \leq i \leq n_x$, let us add a path H^i of length δ between p_i^x and n_i^x.

The key point is that any isometric hull set of G has to contain at least one internal vertex of each path H^i. Indeed, by the choice of δ, for any $1 \leq i \leq n_x$, the graph obtained from G by removing the internal vertices of H^i is isometric in G.

Another important remark is that, since δ is odd, each vertex in H^i is either closer to p_i^x than to n_i^x or vice-versa (no vertex is at equal distance from both). For any $1 \leq i \leq n_x$, let $\{h_i^p, h_i^n\}$ be the middle edge of H^i where h_i^p is closer than p_i^x and h_i^n is closer than n_i^x

As we have already said, any isometric hull set of G must contain all vertices in $I = \{dd_i^x, gg_i^x dd_j^y, gg_j^y \mid 1 \leq i \leq n_x, 1 \leq j \leq n_y\} \cup \{r'\}$ and at least one internal vertex in H^i for each $1 \leq i \leq n_x$. That is, any isometric hull set of G has at least $3n_x + 2n_y + 1$ vertices.

We show that G has an isometric hull set of size $3n_x + 2n_y + 1$ if and only if $\exists X, \forall Y, \Phi(X, Y)$.

- First, assume that there exists an assignment X^* of X such that every assignment of Y satisfies $\Phi(X, Y)$. For any $1 \leq i \leq n_x$, let s_i denote the vertex h_i^p if variable x_i is set to True, and s_i denote h_i^n otherwise.

  We prove that $S = I \cup \{s_1, \cdots, s_{n_x}\}$ is an isometric hull set of G, i.e., G is the unique isometric hull of S.

  If $s_i = h_i^p$ then the path P_x^i and the shortest path from p_i^x to h_i^p (i.e., the subpath of H^i) must belong to any isometric hull of S. Symmetrically, if $s_i = n_i^p$ then the path N_x^i and the shortest path from n_i^x to h_i^n (i.e., the subpath of H^i) must belong to any isometric hull of S.

  Moreover, for any $1 \leq i \leq n_y$, any isometric hull of S must contain either P_y^i or N_y^i.

  Let us consider any isometric hull H of S and, for any $1 \leq i \leq n_y$, let $L_i \in \{P_y^i, N_y^i\}$ be a path contained in H.

  Consider the assignment Y^* of Y defined by H as follows: if $L_i = P_y^i$ then variable y_i is set to true, and it is set to False otherwise (i.e., if $L_i = N_y^i$). Since the formula is true for any assignment of Y, then $\Phi(X^*, Y^*)$ is true. In particular, there is a clause C_j satisfied by all its variables (recall that Φ is disjunctive). By definition of X^*, Y^* and H, this implies that all its three corners belong to H and, as in the proof of Lemma 3.4, this implies that the entire clause gadget C^j is in H. Therefore, using vertex q as in proof of Lemma 3.4, this implies that all vertices p_x^i, n_x^i for $1 \leq i \leq n_x$ compared and all vertices p_y^i, n_y^i for $1 \leq i \leq n_y$ belong to H. From there, it is easy to conclude that all vertices of G belong to H. Therefore, G is the unique isometric hull of S and S is an isometric hull set of the desired size.

- To conclude, we prove that, if for any assignment X^* of X there exists an assignment Y^* of Y such that $\Phi(X^*, Y^*)$ is False, then no set of at most $3n_x + 2n_y + 1$ vertices is an isometric hull set of G.

Let S be a set of at most $3n_x + 2n_y + 1$ vertices. As already said, to be an isometric hull set, S must be equal to $I \cup \{s_1, \cdots, s_{n_x}\}$ where, for any $1 \leq i \leq n_x$, vertex s_i is an internal vertex of the path H^i.

Let X^* be the assignment of X defined as follows: for any $1 \leq i \leq n_x$, variable x_i is set to True if s_i is closer to h_i^p and x_i is set to False otherwise.

By assumption, there is an assignment Y^* of Y such that $\Phi(X^*, Y^*)$ is False.

Let H be the subgraph of G built as follows. First, H contains S and all paths $P(r, d_i^x)$ and $P(r, g_x^i)$ for $1 \leq i \leq n_x$ and H contains all paths $P(r, d_i^y)$ and $P(r, g_y^i)$ for $1 \leq i \leq n_y$. For any $1 \leq i \leq n_x$, H contains P_x^i and the shortest path between s_i and p_x^i if x_i is assigned to True, and H contains N_x^i and the shortest path between s_i and n_x^i if x_i is assigned to False. For any $1 \leq i \leq n_y$, H contains P_y^i if y_i is assigned to True, and H contains N_y^i if y_i is assigned to False.

As in the proof of Lemma 3.4, because no clause is satisfied by $X^* \cup Y^*$, it can be proved that each clause gadget has at most two corners in the current graph H.

Finally, for any clause gadget C^j that has exactly two corners in H, add to H the shortest path (in C^j) between these two corners.

Similar arguments as those in the proof of Lemma 3.4 give that H is a proper isometric subgraph of G and contains S. Therefore, S is not an isometric hull set of G.

□

4 Further work

We have devised a polynomial time algorithm for MGS when all images of a pseudo-closure are given as an input. While pseudo-closures generalize closures, they do not capture other generalizations from the literature such as preclosures (APG90) (since they are not idempotent) or closure functions of greedoids (BZ92) (since they are extensive). Can similar algorithms be provided for these classes?

An open problem with respect to closures is Question 2.9, i.e., find a characterization of those closures coming from the convex subgraphs of a graph. The corresponding question for (finite) metric spaces is also open, see (Her94). Moreover, we wonder about the complexity of `LOGMGS` and `LOGHULL NUMBER`, where the latter even for partial cubes is interesting. A particular question arising in this context is, whether `HULL NUMBER` admits an FPT algorithm parameterized by solution size k. Finally, we would like to recall Question 2.11, i.e., is there a subset X of the hypercube Q_d such that a smallest partial cube in Q_d containing X is not of polynomial size in $d + |X|$?

References

[ACG+13] J. Araujo, V. Campos, F. Giroire, N. Nisse, L. Sampaio, and R. Soares. On the hull number of some graph classes. *Theoret. Comput. Sci.*, 475:1–12, 2013.

[AK16] M. Albenque and K. Knauer. Convexity in partial cubes: the hull number. *Discrete Math.*, 339(2):866–876, 2016.

[AMS+16] J. Araújo, G. Morel, L. Sampaio, R. P. Soares, and V. Weber. Hull number: P_5-free graphs and reduction rules. *Discrete Applied Mathematics*, 210:171–175, 2016.

[APG90] A.V. Arkhangel'skij, L.S. Pontryagin, and R.V. Gamkrelidze, editors. *General topology. I. Basic concepts and constructions. Dimension theory.* Berlin etc.: Springer-Verlag, 1990.

[BP98] József Balogh and Gábor Pete. Random disease on the square grid. *Random Struct. Algorithms*, 13(3-4):409–422, 1998.

[BZ92] Anders Björner and Günter M. Ziegler. Introduction to greedoids. In *Matroid applications*, pages 284–357. Cambridge: Cambridge University Press, 1992.

[CC05] S. R. Canoy, Jr. and G. B. Cagaanan. On the hull number of the composition of graphs. *Ars Combin.*, 75:113–119, 2005.

[CCG06] S. R. Canoy, Jr., G. B. Cagaanan, and S. V. Gervacio. Convexity, geodetic, and hull numbers of the join of graphs. *Util. Math.*, 71:143–159, 2006.

[CHM+10] J. Cáceres, C. Hernando, M. Mora, I. M. Pelayo, and M. L. Puertas. On the geodetic and the hull numbers in strong product graphs. *Comput. Math. Appl.*, 60(11):3020–3031, 2010.

[CHZ00] G. Chartrand, F. Harary, and P. Zhang. On the hull number of a graph. *Ars Combin.*, 57:129–138, 2000.

[CPRPdS13] C. C. Centeno, L. D. Penso, D. Rautenbach, and V. G. Pereira de Sá. Geodetic Number versus Hull Number in P_3-Convexity. *SIAM J. Discrete Math.*, 27(2):717–731, 2013.

[CZ01] G. Chartrand and P. Zhang. On graphs with a unique minimum hull set. *Discuss. Math. Graph Theory*, 21(1):31–42, 2001.

[DF12] Rodney G. Downey and M. R. Fellows. *Parameterized Complexity*. Springer Publishing Company, Incorporated, 2012.

[DGK+09] M. C. Dourado, J. G. Gimbel, J. Kratochvíl, F. Protti, and J. L. Szwarcfiter. On the computation of the hull number of a graph. *Disc. Maths.*, 309(18):5668–5674, 2009.

[DPRS10] M. C. Dourado, F. Protti, D. Rautenbach, and J. L. Szwarcfiter. On the hull number of triangle-free graphs. *SIAM J. Discrete Math.*, 23:2163–2172, 2009/10.

[DS14] I. Dinur and D. Steurer. Analytical approach to parallel repetition. In *46th annual ACM symposium on theory of computing (STOC '14)*, pages 624–633. ACM, 2014.

[ES85] M. G. Everett and S. B. Seidman. The hull number of a graph. *Discrete Math.*, 57(3):217–223, 1985.

[FG06] J. Flum and M. Grohe. *Parameterized Complexity Theory (Texts in Theoretical Computer Science. An EATCS Series)*. Springer-Verlag, 2006.

[FJ86] Martin Farber and Robert E. Jamison. Convexity in graphs and hypergraphs. *SIAM J. Algebraic Discrete Methods*, 7:433–444, 1986.

[FLL+04] Paola Flocchini, Elena Lodi, Fabrizio Luccio, Linda Pagli, and Nicola Santoro. Dynamic monopolies in tori. *Discrete Appl. Math.*, 137(2):197–212, 2004.

[Her94] Eike Hertel. Convexity in finite metric spaces. *Geom. Dedicata*, 52(3):215–220, 1994.

[HJM+05] C. Hernando, T. Jiang, M. Mora, I. M. Pelayo, and C. Seara. On the Steiner, geodetic and hull numbers of graphs. *Discrete Math.*, 293(1-3):139–154, 2005.

[HLT93] Frank Harary, Emmanuel Loukakis, and Constantine Tsouros. The geodetic number of a graph. *Math. Comput. Modelling*, 17(11):89–95, 1993.

[HP01] Yehuda Hassin and David Peleg. Distributed probabilistic polling and applications to proportionate agreement. *Inf. Comput.*, 171(2):248–268, 2001.

[jDR09] Paul A. jun. Dreyer and Fred S. Roberts. Irreversible k-threshold processes: Graph-theoretical threshold models of the spread of disease and of opinion. *Discrete Appl. Math.*, 157(7):1615–1627, 2009.

[KKT15] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. *Theory Comput.*, 11:105–147, 2015.

[KSZ14] Kaveh Khoshkhah, Hossein Soltani, and Manouchehr Zaker. Dynamic monopolies in directed graphs: the spread of unilateral influence in social networks. *Discrete Appl. Math.*, 171:81–89, 2014.

[NVRG05] K. Nehmé, P. Valtchev, M. H. Rouane, and R. Godin. On computing the minimal generator family for concept lattices and icebergs. In 3^{rd} *int. conf. on Formal concept analysis (ICFCA'05)*, pages 192–207. Berlin: Springer, 2005.

[Pap07] Christos H. Papadimitriou. *Computational complexity*. Academic Internet Publ., 2007.

[PY96] C. H. Papadimitriou and M. Yannakakis. On limited nondeterminism and the complexity of the V-C dimension. *J. Comput. Syst. Sci.*, 53(2):161–170, 1996.

[SO96] L.N. Shevrin and A.Ya. Ovsyannikov. *Semigroups and their subsemigroup lattices.* Dordrecht: Kluwer Academic Publishers, 1996.