Complexity Dichotomies for the Minimum F-Overlay Problem

Nathann Cohen, Frédéric Havet, Dorian Mazauric, Ignasi Sau, Rémi Watrigant

Complexity Dichotomies for the Minimum $\mathcal{F}$-Overlay Problem*

Nathann Cohen[1], Frédéric Havet[2], Dorian Mazauric[3], Ignasi Sau[4,5] and Rémi Watrigant[3]

[1] Université Paris-Sud, LRI, CNRS, Orsay, France
[2] Université Côte d'Azur, CNRS, I3S, Inria, Sophia Antipolis, France.
[3] Université Côte d'Azur, Inria, Sophia Antipolis, France.
[4] CNRS, LIRMM, Université de Montpellier, Montpellier, France.
[5] Departamento de Matemática, Universidade Federal do Ceará, Fortaleza, Brazil.

Abstract. For a (possibly infinite) fixed family of graphs $\mathcal{F}$, we say that a graph G *overlays* $\mathcal{F}$ on a hypergraph H if $V(H)$ is equal to $V(G)$ and the subgraph of G induced by every hyperedge of H contains some member of $\mathcal{F}$ as a spanning subgraph. While it is easy to see that the complete graph on $|V(H)|$ overlays $\mathcal{F}$ on a hypergraph H whenever the problem admits a solution, the Minimum $\mathcal{F}$-Overlay problem asks for such a graph with the minimum number of edges. This problem allows to generalize some natural problems which may arise in practice. For instance, if the family $\mathcal{F}$ contains all connected graphs, then Minimum $\mathcal{F}$-Overlay corresponds to the Minimum Connectivity Inference problem (also known as Subset Interconnection Design problem) introduced for the low-resolution reconstruction of macro-molecular assembly in structural biology, or for the design of networks.

Our main contribution is a strong dichotomy result regarding the polynomial vs. NP-hard status with respect to the considered family $\mathcal{F}$. Roughly speaking, we show that the easy cases one can think of (*e.g.* when edgeless graphs of the right sizes are in $\mathcal{F}$, or if $\mathcal{F}$ contains only cliques) are the only families giving rise to a polynomial problem: all others are NP-complete. We then investigate the parameterized complexity of the problem and give similar sufficient conditions on $\mathcal{F}$ that give rise to W[1]-hard, W[2]-hard or FPT problems when the parameter is the size of the solution. This yields an FPT/W[1]-hard dichotomy for a relaxed problem, where every hyperedge of H must contain some member of $\mathcal{F}$ as a (non necessarily spanning) subgraph.

Keywords: Hypergraph, Minimum $\mathcal{F}$-Overlay Problem, NP-completeness, Fixed-parameter tractability.

* This work was partially funded by 'Projet de Recherche Exploratoire', Inria, *Improving inference algorithms for macromolecular structure determination* and ANR under contract STINT ANR-13-BS02-0007.

1 Introduction

1.1 Notation

Most notations of this paper are standard. We now recall some of them, and we refer the reader to [8] for any undefined terminology. For a graph G, we denote by $V(G)$ and $E(G)$ its respective sets of *vertices* and *edges*. The *order* of a graph G is $|V(G)|$, while its *size* is $|E(G)|$. By extension, for a hypergraph H, we denote by $V(H)$ and $E(H)$ its respective sets of *vertices* and *hyperedges*. For $p \in \mathbb{N}$, a *p-uniform* hypergraph H is a hypergraph such that $|S| = p$ for every $S \in E(H)$. Given a graph G, we say that a graph G' is a *subgraph* of G if $V(G') \subseteq V(G)$ and $E(G') \subseteq E(G)$. We say that G' is a *spanning subgraph* of G if it is a subgraph of G such that $V(G') = V(G)$. Given $S \subseteq V(G)$, we denote by $G[S]$ the graph with vertex set S and edge set $\{uv \in E(G) \mid u, v \in S\}$. We say that a graph G' is an *induced subgraph* of G if there exists $S \subseteq V(G)$ such that $G' = G[S]$. Given $S \subseteq V(G)$, we say that an edge $uv \in E(G)$ is *covered* by S if $u \in S$ or $v \in S$, and we say that $uv \in E(G)$ is *induced* by S if $\{u, v\} \subseteq S$. An *isolated* vertex of a graph is a vertex of degree 0. Finally, for a positive integer p, let $[p] = \{1, \ldots, p\}$.

1.2 Definition of the Minimum $\mathcal{F}$-Overlay problem

We define the problem investigated in this paper: MINIMUM $\mathcal{F}$-OVERLAY. Given a fixed family of graphs $\mathcal{F}$ and an input hypergraph H, we say that a graph G *overlays* $\mathcal{F}$ on H if $V(G) = V(H)$ and for every hyperedge $S \in E(H)$, the subgraph of G induced by S, $G[S]$, has a spanning subgraph in $\mathcal{F}$.

Observe that if a graph G overlays $\mathcal{F}$ on H, then the graph G with any additional edges overlays $\mathcal{F}$ on H. Thus, there exists a graph G overlaying $\mathcal{F}$ on H if and only if the complete graph on $|V(H)|$ vertices overlays $\mathcal{F}$ on H. Note that the complete graph on $|V(H)|$ vertices overlays $\mathcal{F}$ on H if and only if for every hyperedge $S \in E(H)$, there exists a graph in $\mathcal{F}$ with exactly $|S|$ vertices. It implies that deciding whether there exists a graph G overlaying $\mathcal{F}$ on H can be done in polynomial time. Hence, otherwise stated, we will always assume that there exists a graph overlaying $\mathcal{F}$ on our input hypergraph H. We thus focus on minimizing the number of edges of a graph overlaying $\mathcal{F}$ on H.

The *$\mathcal{F}$-overlay number* of a hypergraph H, denoted $\mathrm{over}_{\mathcal{F}}(H)$, is the smallest size (*i.e.,* number of edges) of a graph overlaying $\mathcal{F}$ on H.

MINIMUM $\mathcal{F}$-OVERLAY

Input:	A hypergraph H, and an integer k.
Question:	$\mathrm{over}_{\mathcal{F}}(H) \leq k$?

We also investigate a relaxed version of the problem, called MINIMUM $\mathcal{F}$-ENCOMPASS where we ask for a graph G such that for every hyperedge $S \in E(H)$, the graph $G[S]$ contains a (non necessarily spanning) subgraph in $\mathcal{F}$. In an analogous way, we define the *$\mathcal{F}$-encompass number*, denoted $\mathrm{encomp}_{\mathcal{F}}(H)$, of a hypergraph H.

Minimum $\mathcal{F}$-Encompass

Input:	A hypergraph H, and an integer k.
Question:	$\text{encomp}_{\mathcal{F}}(H) \leq k$?

Observe that the Minimum $\mathcal{F}$-Encompass problems are particular cases of Minimum $\mathcal{F}$-Overlay problems. Indeed, for a family $\mathcal{F}$ of graphs, let $\tilde{\mathcal{F}}$ be the family of graphs containing an element of $\mathcal{F}$ as a subgraph. Then Minimum $\mathcal{F}$-Encompass is exactly Minimum $\tilde{\mathcal{F}}$-Overlay.

Throughout the paper, we will only consider graph families $\mathcal{F}$ whose $\mathcal{F}$-Recognition problem[6] is in NP. This assumption implies that Minimum $\mathcal{F}$-Overlay and Minimum $\mathcal{F}$-Encompass are in NP as well (indeed, a certificate for both problems is simply a certificate of the recognition problem for every hyperedge). In particular, it is not necessary for the recognition problem to be in P as it can be observed from the family $\mathcal{F}_{Ham}$ of Hamiltonian graphs: the $\mathcal{F}$-Recognition problem is NP-hard, but providing a spanning cycle for every hyperedge is a polynomial certificate and thus belongs to NP.

1.3 Related work and applications

Minimum $\mathcal{F}$-Overlay allows us to model lots of interesting combinatorial optimization problems of practical interest, as we proceed to discuss.

Common graph families $\mathcal{F}$ are the following: connected graphs (and more generally, ℓ-connected graphs), Hamiltonian graphs, graphs having a universal vertex (*i.e.,* having a vertex adjacent to every other vertex). When the family is the set of all connected graphs, then the problem is known as Subset Interconnection Design, Minimum Topic-Connected Overlay or Interconnection Graph Problem. It has been studied by several communities in the context of designing vacuum systems [10, 11], scalable overlay networks [5, 14, 18], reconfigurable interconnection networks [12, 13], and, in variants, in the context of inferring a most likely social network [2], determining winners of combinatorial auctions [7], as well as drawing hypergraphs [3, 16, 15, 17].

As an illustration, we explain in detail the importance of such inference problems for fundamental questions on structural biology [1]. A major problem is the characterization of low resolution structures of macro-molecular assemblies. To attack this very difficult question, one has to determine the plausible contacts between the subunits of an assembly, given the lists of subunits involved in all the complexes. We assume that the composition, in terms of individual subunits, of selected complexes is known. Indeed, a given assembly can be chemically split into complexes by manipulating chemical conditions. This problem can be formulated as a Minimum $\mathcal{F}$-Overlay problem, where vertices represent the subunits and hyperedges are the complexes. In this setting, an edge between two vertices represents a contact between two subunits.

Hence, the considered family $\mathcal{F}$ is the family of all trees: we want the complexes to be connected. Note that the minimal connectivity assumption avoids

[6] The $\mathcal{F}$-Recognition problem asks, given a graph F, whether $F \in \mathcal{F}$.

speculating on the exact (unknown) number of contacts. Indeed, due to volume exclusion constraints, a given subunit cannot contact many others.

1.4 Our contributions

In Section 2, we prove a strong dichotomy result regarding the polynomial vs. NP-hard status with respect to the considered family $\mathcal{F}$. Roughly speaking, we show that the easy cases one can think of (*e.g.* containing only edgeless and complete graphs) are the only families giving rise to a polynomial problem: all others are NP-complete. In particular, it implies that the Minimum Connectivity Inference problem is NP-hard in p-uniform hypergraphs, which generalizes previous results. In Section 3, we then investigate the parameterized complexity of the problem and give similar sufficient conditions on $\mathcal{F}$ that gives rise to W[1]-hard, W[2]-hard or FPT problems. This yields an FPT/W[1]-hard dichotomy for Minimum $\mathcal{F}$-Encompass.

Due to space restrictions, proofs of results marked by ($\star$) can be found in the long version of the paper [6].

2 Complexity dichotomy

In this section, we prove a dichotomy between families of graphs $\mathcal{F}$ such that Minimum $\mathcal{F}$-Overlay is polynomial-time solvable, and families of graphs $\mathcal{F}$ such that Minimum $\mathcal{F}$-Overlay is NP-complete.

Given a family of graphs $\mathcal{F}$ and a positive integer p, let $\mathcal{F}_p = \{F \in \mathcal{F} : |V(F)| = p\}$. We denote by K_p the complete graph on p vertices, and by $\overline{K_p}$ the edgeless graph on p vertices.

Theorem 1. *Let $\mathcal{F}$ be a family of graphs. If, for every $p > 0$, either $\mathcal{F}_p = \emptyset$ or $\mathcal{F}_p = \{K_p\}$ or $\overline{K_p} \in \mathcal{F}_p$, then* Minimum $\mathcal{F}$-Overlay *is polynomial-time solvable. Otherwise, it is* NP*-complete.*

The first part of this theorem roughly consists in analyzing the sizes of the hyperedges, and adding cliques when necessary.

Theorem 2 ($\star$). *Let $\mathcal{F}$ be a set of graphs. If, for every $p > 0$, either $\mathcal{F}_p = \emptyset$ or $\mathcal{F}_p = \{K_p\}$ or $\overline{K_p} \in \mathcal{F}_p$, then* Minimum $\mathcal{F}$-Overlay *is polynomial-time solvable.*

The NP-complete part requires more work. We need to prove that if there exists $p > 0$ such that $\mathcal{F}_p \neq \emptyset$, $\mathcal{F}_p \neq \{K_p\}$, and $\overline{K_p} \notin \mathcal{F}_p$, then Minimum $\mathcal{F}$-Overlay is NP-complete. Actually, it is sufficient to prove the following:

Theorem 3. *Let $p > 0$, and $\mathcal{F}_p$ be a non-empty set of graphs with p vertices such that $\mathcal{F}_p \neq \{K_p\}$ and $\overline{K_p} \notin \mathcal{F}_p$. Then* Minimum $\mathcal{F}_p$-Overlay *is* NP*-complete (when restricted to p-uniform hypergraphs).*

2.1 Prescribing some edges

A natural generalization of Minimum $\mathcal{F}$-Overlay is to prescribe a set E of edges to be in the graph overlaying $\mathcal{F}$ on H. We denote by $\mathrm{over}_{\mathcal{F}}(H; E)$ the minimum number of edges of a graph G overlaying $\mathcal{F}$ on H with $E \subseteq E(G)$.

Prescribed Minimum $\mathcal{F}$-Overlay
<u>Input:</u> A hypergraph H, an integer k, and a set $E \subseteq \binom{V(H)}{2}$.
<u>Question:</u> $\mathrm{over}_{\mathcal{F}}(H; E) \leq k$?

In fact, in terms of computational complexity, the two problems Minimum $\mathcal{F}$-Overlay and Prescribed Minimum $\mathcal{F}$-Overlay are equivalent.

Theorem 4 ($\star$). *Let $\mathcal{F}$ be a (possibly infinite) class of graphs. Then* Minimum $\mathcal{F}$-Overlay *and* Prescribed Minimum $\mathcal{F}$-Overlay *are polynomially equivalent.*

2.2 Hard sets

A set $\mathcal{F}_p$ of graphs of order p is *hard* if there is a graph J of order p and two distinct non-edges e_1, e_2 of J such that

- no subgraph of J is in $\mathcal{F}_p$ (including J itself),
- $J \cup e_1$ has a subgraph in $\mathcal{F}_p$ and $J \cup e_2$ has a subgraph in $\mathcal{F}_p$.

The graph J is called the *hyperedge graph of* $\mathcal{F}_p$ and e_1 and e_2 are its two *shifting non-edges.*

For example, the set $\mathcal{F}_3 = \{P_3\}$, where P_3 is the graph with three vertices and two edges, is hard. Indeed, the graph O_3 with three vertices and one edge has no subgraph in $\mathcal{F}_3$, but adding any of the two non-edges of O_3 results in a graph isomorphic to P_3.

Lemma 1. *Let $p \geq 3$ and $\mathcal{F}_p$ be a set of graphs of order p. If $\mathcal{F}_p$ is hard, then* Prescribed Minimum $\mathcal{F}_p$-Overlay *is* NP*-complete.*

Proof. We present a reduction from Vertex Cover. Let J be the hyperedge graph of $\mathcal{F}_p$ and e_1, e_2 its shifting non-edges. We distinguish two cases depending on whether e_1 and e_2 are disjoint or not. The proofs of both cases are very similar, we thus omit the second case which can be found in the long version of the paper.

<u>Case 1:</u> e_1 and e_2 intersect. Let G be a graph. Let H_G be the hypergraph constructed as follows.

- For every vertex $v \in V(G)$ add two vertices x_v, y_v.
- For every edge $e = uv$, add a vertex z_e and three disjoint sets Z_e, Y^e_u, and Y^e_v of size $p - 3$.
- For every edge $e = uv$, create three hyperedges $Z_e \cup \{z_e, y_u, y_v\}$, $Y^e_u \cup \{x_u, y_u, z_e\}$, and $Y^e_v \cup \{x_v, y_v, z_e\}$.

We select forced edges as follows: for every edge $e = uv \in E(G)$, we force the edges of a copy of J on $Z_e \cup \{z_e, y_u, y_v\}$ with shifting non-edges $z_e y_u$ and $z_e y_v$, we force the edges of a copy of J on $Y_u^e \cup \{z_e, y_u, x_u\}$ with shifting non-edges $y_u z_e$ and $y_u x_u$, and we force the edges of a copy of J on $Y_v^e \cup \{z_e, y_v, x_v\}$ with shifting non-edges $y_v z_e$ and $y_v x_v$.

We shall prove that $\text{over}_{\mathcal{F}_p}(H_G) = |E| + \text{vc}(G) + |E(G)|$, which yields the result. Here, $\text{vc}(G)$ denotes the size of a minimum vertex cover of G.

Consider first a minimum vertex cover C of G. For every edge $e \in E(G)$, let s_e be an endvertex of e that is not in C if such vertex exists, or any endvertex of e otherwise. Set $E_G = E \cup \{x_v y_v \mid v \in C\} \cup \{z_e y_{s_e} \mid e \in E(G)\}$. One can easily check that $(V_G, E \cup E_G)$ overlays $\mathcal{F}_p$ on H_G. Indeed, for every hyperedge S of H_G, at least one of the shifting non-edges of its forced copy of J is an edge of $E \cup E_G$. Therefore $\text{over}_{\mathcal{F}_p}(H_G) \leq |E| + |E_G| = |E| + \text{vc}(G) + |E(G)|$.

Now, consider a minimum-size graph $(V_G, E \cup E_G)$ overlaying $\mathcal{F}_p$ on H_G and maximizing the edges of the form $x_u y_u$. Let $e = uv \in E(G)$. Observe that the edge $y_u y_v$ is contained in a unique hyperedge, namely $Z_e \cup \{z_e, y_u, y_v\}$. Therefore, free to replace it (if it is not in E) by $z_e y_v$, we may assume that $y_u y_v \notin E_G$. Similarly, we may assume that the edges $x_u z_e$ and $x_v z_e$ are not in E_G, and that no edge with an endvertex in $Y_u^e \cup Y_v^e \cup Z_e$ is in E_G. Furthermore, one of $x_u y_u$ and $x_v y_v$ is in E_G. Indeed, if $\{x_u y_u, x_v y_v\} \cap E_G = \emptyset$, then $\{y_u z_e, y_v z_e\} \subseteq E_G$ because E_G contains an edge included in every hyperedge. Thus replacing $y_u z_e$ by $x_u y_u$ results in another graph overlaying $\mathcal{F}_p$ on H_G with one more edge of type $x_u y_u$ than the chosen one, a contradiction.

Let $C = \{u \mid x_u y_u \in E_G\}$. By the above property, C is a vertex cover of G, so $|C| \geq \text{vc}(G)$. Moreover, E_G contains an edge in every hyperedge $Z_e \cup \{z_e, y_u, y_v\}$, and those $|E(G)|$ edges are not in $\{x_u y_u \mid u \in V(G)\}$. Therefore $|E_G| \geq |C| + |E(G)| \geq \text{vc}(G) + |E(G)|$. ☐

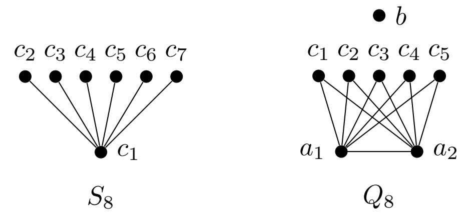
Let $\mathcal{F}_p$ be a set of graphs of order p. It is *free* if there are no two distinct elements of $\mathcal{F}_p$ such that one is a subgraph of the other. The *core* of $\mathcal{F}_p$ is the free set of graphs F having no proper subgraphs in $\mathcal{F}_p$. It is easy to see that $\mathcal{F}_p$ is overlayed by a hypergraph if and only if its core does. Henceforth, we may restrict our attention to free sets of graphs.

Lemma 2. *Let $\mathcal{F}_p$ be a free set of graphs of order p. If a graph F in $\mathcal{F}_p$ has an isolated vertex and a vertex of degree 1, then $\mathcal{F}_p$ is hard.*

Proof. Let z be an isolated vertex of F, y a vertex of degree 1, and x the neighbor of y in F. The graph $J = F \setminus xy$ contains no element of $\mathcal{F}_p$ because $\mathcal{F}_p$ is free. Moreover $J \cup xy$ and $J \cup yz$ are isomorphic to F. Hence J is a hyperedge graph of $\mathcal{F}_p$. Thus, by Lemma 1, Prescribed Minimum $\mathcal{F}_p$-Overlay is NP-complete. ☐

The *star of order p*, denoted by S_p, is the graph of order p with $p-1$ edges incident to a same vertex.

Lemma 3. *Let $p \geq 3$ and let $\mathcal{F}_p$ be a free set of graphs of order p containing a subgraph of the star S_p different from $\overline{K}_p$. Then $\mathcal{F}_p$ is hard.*



Proof. Let S be the non-empty subgraph of S_p in $\mathcal{F}_p$. If $S \neq S_p$, then S has an isolated vertex and a vertex of degree 1, and so $\mathcal{F}_p$ is hard by Lemma 2. We may assume henceforth that $S_p \in \mathcal{F}_p$.

Let Q_p be the graph with p vertices $\{a_1, a_2, b, c_1, \ldots, c_{p-3}\}$ and edge set $\{a_1a_2\} \cup \{a_ic_j \mid 1 \leq i \leq 2, 1 \leq j \leq p-3\}$. Observe that Q_p does not contain S_p but $Q_p \cup a_1b$ and $Q_p \cup a_2b$ do. If $\mathcal{F}_p$ contains no subgraph of Q_p, then $\mathcal{F}_p$ is hard. So we may assume that $\mathcal{F}_p$ contains a subgraph of Q_p.

Let Q be the subgraph of Q_p in $\mathcal{F}_p$ that has the minimum number of triangles. If Q has a degree 1 vertex, then $\mathcal{F}_p$ is hard by Lemma 2. Henceforth we may assume that Q has no vertex of degree 1. So, without loss of generality, there exists q such that $E(Q) = \{a_1a_2\} \cup \{a_ic_j \mid 1 \leq i \leq 2, 1 \leq j \leq q\}$.

Let $R = (Q \setminus a_1c_1) \cup a_2b$. Observe that $R \cup a_1c_1$ and $R \cup a_1b$ contain Q. If $\mathcal{F}_p$ contains no subgraph of R, then $\mathcal{F}_p$ is hard. So we may assume that $\mathcal{F}_p$ contains a subgraph R' of R. But $\mathcal{F}_p$ contains no subgraph of Q because it is free, so both a_2c_1 and a_2b are in R'. In particular, c_1 and b have degree 1 in R'.

Let $T = (Q \setminus a_1c_1)$. It is a proper subgraph of Q, so $\mathcal{F}_p$ contains no subgraph of T, because $\mathcal{F}_p$ is free. Moreover $T \cup a_1c_1 = Q$ is in $\mathcal{F}_p$ and $T \cup a_2b = R$ contains $R' \in \mathcal{F}_p$. Hence $\mathcal{F}_p$ is hard. □

2.3 Proof of Theorem 3

For convenience, instead of proving Theorem 3, we prove the following statement, which is equivalent by Theorem 4.

Theorem 5. *Let p be a positive integer. Let $\mathcal{F}_p$ be a non-empty set of graphs of order p.* Prescribed Minimum $\mathcal{F}_p$-Overlay *is* NP*-complete unless $\overline{K}_p \in \mathcal{F}_p$ or $\mathcal{F}_p = \{K_p\}$.*

Proof. We proceed by induction on p, the result holding trivially when $p = 1$ and $p = 2$. Assume now that $p \geq 3$. Without loss of generality, we may assume that $\mathcal{F}_p$ is a free set of graphs.

A *hypograph* of a graph G is an induced subgraph of G of order $|G| - 1$. In other words, it is a subgraph obtained by removing a vertex from G. Let $\mathcal{F}^-$ be the set of hypographs of elements of $\mathcal{F}_p$.

If $\mathcal{F}^- = \{K_{p-1}\}$, then necessarily $\mathcal{F}_p = \{K_p\}$, and Prescribed Minimum $\mathcal{F}_p$-Overlay is trivially polynomial-time solvable.

If $\mathcal{F}^- \neq \{K_{p-1}\}$ and $\overline{K}_{p-1} \notin \mathcal{F}^-$, then Prescribed Minimum $\mathcal{F}^-$-Overlay is NP-complete by the induction hypothesis. We shall now reduce this problem to Prescribed Minimum $\mathcal{F}_p$-Overlay. Let (H^-, k^-, E^-) be an instance of Prescribed Minimum $\mathcal{F}^-$-Overlay. For every hyperedge S of H^-, we create a new vertex x_S and the hyperedge $X_S = S \cup \{x_S\}$. Let H be the hypergraph defined by $V(H) = V(H^-) \cup \bigcup_{S \in E(H^-)} x_S$ and $E(H) = \{X_S \mid S \in E(H^-)\}$. We set $E = E^- \cup \bigcup_{S \in E(H^-)} \{x_S v \mid v \in S\}$.

Let us prove that $\text{over}_{\mathcal{F}_p}(H; E) = \text{over}_{\mathcal{F}^-}(H^-; E^-) + (p-1) \cdot |S|$. Clearly, if $G^- = (V(H^-), F^-)$ overlays $\mathcal{F}^-$, then $G = (V(H), F^- \cup \bigcup_{S \in E(H^-)} \{x_S v \mid v \in S\})$ overlays $\mathcal{F}_p$. Hence $\text{over}_{\mathcal{F}_p}(H; E) \leq \text{over}_{\mathcal{F}^-}(H^-; E^-) + (p-1) \cdot |S|$. Reciprocally, assume that G overlays $\mathcal{F}_p$. Then for each hyperedge S of H^-, the graph $G[X_S] \in \mathcal{F}_p$, and so $G[S] \in \mathcal{F}^-$. Therefore, setting the graph $G^- = G[V(H^-)]$ overlays $\mathcal{F}^-$. Moreover $E(G) \setminus E(G^-) = \bigcup_{S \in E(H^-)} \{x_S v \mid v \in S\}$. Hence $\text{over}_{\mathcal{F}_p}(H; E) \geq \text{over}_{\mathcal{F}^-}(H^-; E^-) + (p-1) \cdot |S|$.

Assume now that $\overline{K}_{p-1} \in \mathcal{F}^-$. Then $\mathcal{F}_p$ contains a subgraph of the star S_p. If $\mathcal{F}_p$ contains $\overline{K}_p$, then Prescribed Minimum $\mathcal{F}_p$-Overlay is trivially polynomial-time solvable. Henceforth, we may assume that $\mathcal{F}_p$ contains a non-empty subgraph of S_p. Thus, by Lemma 3, $\mathcal{F}_p$ is hard, and so by Lemma 1, Prescribed Minimum $\mathcal{F}_p$-Overlay is NP-complete. □

3 Parameterized analysis

We now focus on the parameterized complexity of our problems. A *parameterization* of a decision problem Q is a computable function κ that assigns an integer $\kappa(I)$ to every instance I of the problem. We say that (Q, κ) is *fixed-parameter tractable* (FPT) if every instance I can be solved in time $O(f(\kappa(I))|I|^c)$, where f is some computable function, $|I|$ is the encoding size of I, and c is some constant independent of I (we will sometimes use the $O^*(\cdot)$ notation that removes polynomial factors and additive terms). Finally, the W[i]-hierarchy of parameterized problems is typically used to rule out the existence of FPT algorithms, under the widely believed assumption that FPT $\neq$ W[1]. For more details about fixed-parameter tractability, we refer the reader to the monograph of Downey and Fellows [9].

Since Minimum $\mathcal{F}$-Overlay is NP-hard for most non-trivial cases, it is natural to ask for the existence of FPT algorithms. In this paper, we consider the so-called *standard parameterization* of an optimization problem: the size of a solution. In the setting of our problems, this parameter corresponds to the number k of edges in a solution. Hence, the considered parameter will always be k in the remainder of this section.

Similarly to our dichotomy result stated in Theorem 1, we would like to obtain necessary and sufficient conditions on the family $\mathcal{F}$ giving rise to either an FPT or a W[1]-hard problem. One step towards such a result is the following FPT-analogue of Theorem 2.

Theorem 6. *Let $\mathcal{F}$ be a family of graphs. If there is a non-decreasing function $f : \mathbb{N} \to \mathbb{N}$ such that $\lim_{n\to+\infty} f(n) = +\infty$ and $|E(F)| \geq f(|V(F)|)$ for all $F \in \mathcal{F}$, then* Minimum $\mathcal{F}$-Overlay *is* FPT.

Proof. Let $g : \mathbb{N} \to \mathbb{N}$ be the function that maps every $k \in \mathbb{N}$ to the smallest integer ℓ such that $f(\ell) \geq k$. Since $\lim_{n\to+\infty} f(n) = +\infty$, g is well-defined. If a hyperedge S of a hypergraph H is of size at least $g(k+1)$, then since f is non-decreasing, $\text{over}_{\mathcal{F}}(H) > k$ and so the instance is negative. Therefore, we may assume that every hyperedge of H has size at most $g(k)$. Applying a simple branching algorithm (see [9]) allows us to solve the problem in time $O^*(g(k)^{O(k)})$. □

Observe that if $\mathcal{F}$ is finite, setting $N = \max\{|E(F)| \mid F \in \mathcal{F}\}$, the function f defined by $f(n) = 0$ for $n \leq N$ and $f(n) = n$ otherwise satisfies the condition of Theorem 6, and so Minimum $\mathcal{F}$-Overlay is FPT. Moreover, Theorem 6 encompasses some interesting graph families. Indeed, if $\mathcal{F}$ is the family of connected graphs (resp. Hamiltonian graphs), then $f(n) = n - 1$ (resp. $f(n) = n$) satisfies the required property. Other graph families include c-vertex-connected graphs or c-edge-connected graphs for any fixed $c \geq 1$, graphs of minimum degree at least d for any fixed $d \geq 1$. In sharp contrast, we shall see in the next subsection (Theorem 7) that if, for instance, $\mathcal{F}$ is the family of graphs containing a matching of size at least c, for any fixed $c \geq 1$, then the problem becomes W[1]-hard (note that such a graph might have an arbitrary number of isolated vertices).

3.1 Negative result

In view of Theorem 6, a natural question is to know what happens for graph families not satisfying the conditions of the theorem. Although we were not able to obtain an exact dichotomy as in the previous section, we give sufficient conditions on $\mathcal{F}$ giving rise to problems that are unlikely to be FPT (by proving W[1]-hardness or W[2]-hardness).

An interesting situation is when $\mathcal{F}$ is *closed by addition of isolated vertices*, *i.e.,* for every $F \in \mathcal{F}$, the graph obtained from F by adding an isolated vertex is also in $\mathcal{F}$. Observe that for such a family, Minimum $\mathcal{F}$-Overlay and Minimum $\mathcal{F}$-Encompass are equivalent, which is the reason that motivated us defining this relaxed version. We have the following result, which implies an FPT/W[1]-hard dichotomy for Minimum $\mathcal{F}$-Encompass.

Theorem 7. *Let $\mathcal{F}$ be a fixed family of graphs closed by addition of isolated vertices. If $\overline{K}_p \in \mathcal{F}$ for some $p \in \mathbb{N}$, then* Minimum $\mathcal{F}$-Overlay *is* FPT. *Otherwise, it is* W[1]*-hard parameterized by k.*

Proof. To prove the positive result, let p be the minimum integer such that $\overline{K}_p \in \mathcal{F}$. Observe that no matter the graph G, for every hyperedge $S \in E(H)$, $G[S]$ will contain $\overline{K}_{|S|}$ as a spanning subgraph, which is in $\mathcal{F}$ whenever $|S| \geq p$ (recall that $\mathcal{F}$ is closed by addition of isolated vertices). Then, a simple branching algorithm allows us to enumerate all graphs (with at least one edge) induced by hyperedges of size at most $p-1$ in $O^*(p^{O(k)})$ time.

To prove the negative result, we use a recent result of Chen and Lin [4] stating that any constant-approximation of the parameterized Dominating Set is $\mathsf{W}[1]$-hard, which directly transfers to Hitting Set[7]. For an input of Hitting Set, namely a finite set U (called the *universe*), and a family $\mathcal{S}$ of subsets of U, let $\tau(U,\mathcal{S})$ be the minimum size of a set $K \subseteq U$ such that $K \cap S \neq \emptyset$ for all $S \in \mathcal{S}$ (such a set is called a *hitting set*). The result of Chen and Lin implies that the following problem is $\mathsf{W}[1]$-hard parameterized by k.

Gap$_\rho$ Hitting Set

Input:	A finite set U, a family $\mathcal{S}$ of subsets of U, and a positive integer k.
Question:	Decide whether $\tau(U,\mathcal{S}) \leq k$ or $\tau(U,\mathcal{S}) > \rho k$.

Let F_{is} be a graph from $\mathcal{F}$ minimizing the two following criteria (in this order): number of non-isolated vertices, and minimum degree of non-isolated vertices. Let r_{is} and δ_{is} be the respective values of these criteria, $n_{is} = |V(F_{is})|$, and $m_{is} = |E(F_{is})|$. We thus have $\delta_{is} \leq r_{is}$. Let F_e be a graph in $\mathcal{F}$ with the minimum number of edges, and $n_e = |V(F_e)|$, $m_e = |E(F_e)|$.

Let $U, \mathcal{S}, k$ be an instance of Gap$_{2r_{is}}$ Hitting Set, with $U = \{u_1, \ldots, u_n\}$. We denote by H the hypergraph constructed as follows. Its vertex set is the union of:

- a set V_{is} of $r_{is}-1$ vertices;
- a set $V_U = \bigcup_{i=1}^n V^i$, where $V^i = \{v_1^i, \ldots, v_{n_{is}-r_{is}+1}^i\}$; and
- for every $u, v \in V_{is}$, $u \neq v$, a set $V_{u,v}$ of $n_e - 2$ vertices.

Then, for every $u, v \in V_{is}$, $u \neq v$, create a hyperedge $h_{u,v} = \{u, v\} \cup V_{u,v}$ and, for every set $S \in \mathcal{S}$, create the hyperedge $h_S = V_{is} \cup \bigcup_{i : u_i \in S} V^i$. Finally, let $k' = \binom{n_{is}-1}{2} m_e + k\delta_{is}$. Since $\mathcal{F}$ is fixed, k' is a function of k only.

We shall prove that if $\tau(U,\mathcal{S}) \leq k$, then $\mathrm{over}_{\mathcal{F}}(H) \leq k'$ and, conversely, if $\mathrm{over}_{\mathcal{F}}(H) \leq k'$, then $\tau(U,\mathcal{S}) \leq 2r_{is}k$.

Assume first that U has a hitting set K of size at most k. For every $u, v \in V_{is}$, $u \neq v$, add to G the edges of a copy of F_e on $h_{u,v}$ with $uv \in E(G)$. This already adds $\binom{n_{is}-1}{2} m_e$ edges to G and, obviously, $G[h_{u,v}]$ contains F_e as a subgraph. Now, for every $u_i \in K$, add all edges between v_1^i and δ_{is} arbitrarily chosen vertices in V_{is}. Observe that for every $S \in \mathcal{S}$, $G[h_S]$ contains F_{is} as a subgraph, and also $|E(G)| \leq k'$.

Conversely, let G be a solution for Minimum $\mathcal{F}$-Overlay with at most k' edges. Clearly, for all $u, v \in V_{is}$, $u \neq v$, $G[V_{u,v}]$ has at least m_e edges, hence the subgraph of G induced by $V(H) \setminus V_U$ has at least $\binom{n_{is}-1}{2} m_e$ edges, and thus the number of edges of G covered by V_u is at most $k\delta_{is}$. Let K be the

[7] Roughly speaking, each element of the universe represents a vertex of the graph, and for each vertex, create a set with the elements corresponding to its closed neighborhood.

set of non-isolated vertices of V_U in G, and $K' = \{u_i \mid v^i_j \in K$ for some $j \in \{1, \ldots, n_{is} - r_{is} + 1\}\}$. We claim that K' is a hitting set of $(U, \mathcal{S})$: indeed, for every $S \in \mathcal{S}$, $G[h_S]$ must contain some $F \in \mathcal{F}$ as a subgraph, but since V_{is} is composed of $r_{is} - 1$ vertices, and since F_{is} is a graph from $\mathcal{F}$ with the minimum number r_{is} of non-isolated vertices, there must exist $i \in \{1, \ldots, n\}$ such that $u_i \in S$, and $j \in \{1, \ldots, n_{is} - r_{is} + 1\}$ such that $v^i_j \in h_S \cap K$, and thus $S \cap K' \neq \emptyset$. Finally, observe that K is a set of non-isolated vertices covering $k\delta_{is}$ edges, and thus $|K| \leq 2k\delta_{is}$ (in the worst case, K induces a matching), hence we have $|K'| \leq |K| \leq 2k\delta_{is} \leq 2r_{is}k$, *i.e.*, $\tau(U, \mathcal{S}) \leq 2r_{is}k$, concluding the proof. □

It is worth pointing out that the idea of the proof of Theorem 7 applies to broader families of graphs. Indeed, the required property 'closed by addition of isolated vertices' forces $\mathcal{F}$ to contain all graphs $F_{is} + \overline{K}_i$ (where + denotes the disjoint union of two graphs) for every $i \in \mathbb{N}$. Actually, it would be sufficient to require the existence of a polynomial $p : \mathbb{N} \to \mathbb{N}$ such that for any $i \in \mathbb{N}$, we have $F_{is} + \overline{K}_{p(i)} \in \mathcal{F}$ (roughly speaking, for a set S of the Hitting Set instance, we would construct a hyperedge with $|V(F_{is} + \overline{K}_{p(|S|)})|$ vertices). Intuitively, most families of practical interest not satisfying such a constraint will fall into the scope of Theorem 6. Unfortunately, we were not able to obtain the dichotomy in a formal way.

Nevertheless, as explained before, this still yields an FPT/W[1]-hardness dichotomy for the Minimum $\mathcal{F}$-Encompass problem.

Corollary 1. *Let $\mathcal{F}$ be a fixed family of graphs. If $\overline{K}_p \in \mathcal{F}$ for some $p \in \mathbb{N}$, then* Minimum $\mathcal{F}$-Encompass *is* FPT. *Otherwise, it is* W[1]*-hard parameterized by* k.

We conclude this section with a stronger negative result than Theorem 7, but concerning a restricted graph family (hence both results are incomparable).

Theorem 8 ($\star$). *Let $\mathcal{F}$ be a fixed graph family such that (i) $\mathcal{F}$ is closed by addition of isolated vertices; (ii) $\overline{K}_p \notin \mathcal{F}$ for every $p \geq 0$; and (iii) all graphs in $\mathcal{F}$ have the same number of non-isolated vertices. Then* Minimum $\mathcal{F}$-Overlay *is* W[2]*-hard parameterized by* k.

4 Conclusion and future work

Naturally, the first open question is to close the gap between Theorems 6 and 7 in order to obtain a complete FPT/W[1]-hard dichotomy for any family $\mathcal{F}$.

As further work, we are also interested in a more constrained version of the problem, in the sense that we may ask for a graph G such that for every hyperedge $S \in E(H)$, the graph $G[S]$ belongs to $\mathcal{F}$ (hence, we forbid additional edges). The main difference between Minimum $\mathcal{F}$-Overlay and this problem, called Minimum $\mathcal{F}$-Enforcement, is that it is no longer trivial to test for the existence of a feasible solution (actually, it is possible to prove the NP-hardness of this existence test for very simple families, *e.g.* when $\mathcal{F}$ only contains P_3, the path on three vertices). We believe that a dichotomy result similar to Theorem 1 for Minimum $\mathcal{F}$-Enforcement is an interesting challenging question, and will need a different approach than the one used in the proof of Theorem 5.

References

1. Deepesh Agarwal, Christelle Caillouet, David Coudert, and Frédéric Cazals. Unveiling Contacts within Macro-molecular assemblies by solving Minimum Weight Connectivity Inference Problems. *Molecular and Cellular Proteomics*, 14:2274–2284, April 2015.
2. Dana Angluin, James Aspnes, and Lev Reyzin. Inferring social networks from outbreaks. In *ALT '10*, pages 104–118, 2010.
3. Ulrik Brandes, Sabine Cornelsen, Barbara Pampel, and Arnaud Sallaberry. Blocks of hypergraphs - applied to hypergraphs and outerplanarity. In *IWOCA '10*, pages 201–211, 2010.
4. Yijia Chen and Bingkai Lin. The constant inapproximability of the parameterized dominating set problem. In *FOCS 2016*, pages 505–514, 2016.
5. Gregory Chockler, Roie Melamed, Yoav Tock, and Roman Vitenberg. Constructing scalable overlays for pub-sub with many topics. In *PODC '07*, pages 109–118. ACM, 2007.
6. Nathann Cohen, Dorian Mazauric, Ignasi Sau, and Rémi Watrigant. Complexity dichotomies for the minimum f-overlay problem. *CoRR*, abs/1703.05156, 2017.
7. Vincent Conitzer, Jonathan Derryberry, and Tuomas Sandholm. Combinatorial auctions with structured item graphs. In *AAAI '04*, pages 212–218, 2004.
8. Reinhard Diestel. *Graph Theory, 4th Edition*, volume 173 of *Graduate texts in mathematics*. Springer, 2012.
9. Rod G. Downey and Michael R. Fellows. *Fundamentals of Parameterized Complexity*. Texts in Computer Science. Springer, 2013.
10. Ding Zhu Du and Dean F. Kelley. On complexity of subset interconnection designs. *Journal of Global Optimization*, 6(2):193–205, 1995.
11. Ding-Zhu Du and Zevi Miller. Matroids and subset interconnection design. *SIAM Journal on Discrete Mathematics*, 1(4):416–424, 1988.
12. Hongbing Fan, Christian Hundt, Yu-Liang Wu, and Jason Ernst. Algorithms and implementation for interconnection graph problem. In *COCOA '08*, pages 201–210, 2008.
13. Hongbing Fan and Yu-Liang Wu. Interconnection graph problem. In *FCS '08*, pages 51–55, 2008.
14. Jun Hosoda, Juraj Hromkovic, Taisuke Izumi, Hirotaka Ono, Monika Steinov, and Koichi Wada. On the approximability and hardness of minimum topic connected overlay and its special instances. *Theoretical Computer Science*, 429:144 – 154, 2012.
15. D. S. Johnson and H. O. Pollak. Hypergraph planarity and the complexity of drawing venn diagrams. *Journal of Graph Theory*, 11(3):309–325, 1987.
16. Boris Klemz, Tamara Mchedlidze, and Martin Nöllenburg. Minimum tree supports for hypergraphs and low-concurrency euler diagrams. In *SWAT '14*, pages 265–276, 2014.
17. Ephraim Korach and Michal Stern. The clustering matroid and the optimal clustering tree. *Mathematical Programming*, 98(1):385–414, 2003.
18. Melih Onus and Andréa W. Richa. Minimum maximum-degree publish-subscribe overlay network design. *IEEE/ACM Trans. Netw.*, 19(5):1331–1343, October 2011.