Overlaying a hypergraph with a graph with bounded maximum degree

Frédéric Havet, Dorian Mazauric, Viet-Ha Nguyen, Rémi Watrigant

Overlaying a hypergraph with a graph with bounded maximum degree

Frédéric Havet[1,2][0000−0002−3447−8112], Dorian Mazauric[2][0000−0002−5589−1889], Viet-Ha Nguyen[1,2], and Rémi Watrigant[3][0000−0002−6243−5910]

[1] CNRS, I3S, France
[2] Inria, Université Côte d'Azur, France
[3] Université de Lyon, France

Abstract. Let G and H be respectively a graph and a hypergraph defined on a same set of vertices, and let F be a fixed graph. We say that G *F-overlays* a hyperedge S of H if F is a spanning subgraph of the subgraph of G induced by S, and that it *F-overlays* H if it F-overlays every hyperedge of H. Motivated by structural biology, we study the computational complexity of two problems. The first problem, $(\Delta \leq k)$ F-Overlay, consists in deciding whether there is a graph with maximum degree at most k that F-overlays a given hypergraph H. It is a particular case of the second problem Max $(\Delta \leq k)$ F-Overlay, which takes a hypergraph H and an integer s as input, and consists in deciding whether there is a graph with maximum degree at most k that F-overlays at least s hyperedges of H.
We give a complete polynomial/$\mathcal{NP}$-complete dichotomy for the Max $(\Delta \leq k)$-F-Overlay problems depending on the pairs (F, k), and establish the complexity of $(\Delta \leq k)$ F-Overlay for many pairs (F, k).

1 Introduction

A major problem in structural biology is the characterization of low resolution structures of macro-molecular assemblies [5, 20]. To attack this very difficult question, one has to determine the plausible contacts between the subunits (e.g. proteins) 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 conveniently modeled by graphs and hypergraphs. We consider the hypergraph H whose vertices represent the subunits and whose hyperedges are the complexes. We are then looking for a graph G with the same vertex set as H whose edges represent the contacts between subunits, and satisfying (i) some local properties for every complex (*i.e.* hyperedge), and (ii) some other global properties.

We first focus on the local properties. They are usually modeled by a (possibly infinite) family $\mathcal{F}$ of admissible graphs to which each complex must belong: to this end, we define the notion of *enforcement* of a hyperedge and a hypergraph. A graph G *$\mathcal{F}$-enforces* a hyperedge $S \in E(H)$ if the subgraph $G[S]$ of G induced by S belongs to $\mathcal{F}$, and it *$\mathcal{F}$-enforces* H if it $\mathcal{F}$-enforces all hyperedges of H. Very often, the considered family $\mathcal{F}$ is closed on taking edge supergraphs [1, 8]: if $F \in \mathcal{F}$, then every graph obtained from G by adding edges is also in $\mathcal{F}$. Such a family is completely defined by its set $\mathcal{M} = \mathcal{M}(\mathcal{F})$ of minimal graphs that are the elements of $\mathcal{F}$ which are not edge supergraphs of any other. In this case, a graph G $\mathcal{F}$-enforcing S is such that there is an element of $\mathcal{M}$ which is a spanning subgraph of $G[S]$. This leads to the following notion of *overlayment* when considering minimal graph families.

Definition 1. A graph G *$\mathcal{F}$-overlays* a hyperedge S if there exists $F \in \mathcal{F}$ such that F is a spanning subgraph of $G[S]$, and it *$\mathcal{F}$-overlays* H if it $\mathcal{F}$-overlays every hyperedge of H.

As said previously, the graph sought will also have to satisfy some global constraints. Since in a macro-molecular assembly the number of contacts is small, the first natural idea is to look for a graph G with the minimum number of edges. This leads to the MIN-$\mathcal{F}$-OVERLAY problem: given a hypergraph H and an integer m, decide if there exists a graph G $\mathcal{F}$-overlaying H such that $|E(G)| \leq m$.

A typical example of a family $\mathcal{F}$ is the set of all connected graphs, in which case $\mathcal{M}(\mathcal{F})$ is the set of all trees. Agarwal et al. [1] focused on MIN-$\mathcal{M}(\mathcal{F})$-OVERLAY for this particular family in the aforementioned context of structural biology. However, this problem was previously studied by several communities in other domains, as pointed out by Chen *et al.* [6]. Indeed, it is also known as SUBSET INTERCONNECTION DESIGN, MINIMUM TOPIC-CONNECTED OVERLAY or INTERCONNECTION GRAPH PROBLEM, and was considered (among others) in the design of vacuum systems [10, 11], scalable overlay networks [7, 18], and reconfigurable interconnection networks [12, 13]. Some variants have also been considered in the contexts of inferring a most likely social network [2], determining winners of combinatorial auctions [9], as well as drawing hypergraphs [4, 14].

Cohen et al. [8] presented a dichotomy regarding the polynomial vs. $\mathcal{NP}$-hard status of the problem MIN-$\mathcal{F}$-OVERLAY with respect to the considered family $\mathcal{F}$. Roughly speaking, they showed 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-time solvable problem: all others are $\mathcal{NP}$-complete. They also considered the FPT/W[1]-hard dichotomy for several families $\mathcal{F}$.

In this paper, we consider the variant in which the additional constraint is that G must have a bounded maximum degree: this constraint is motivated by the context of structural biology, since a subunit (e.g. a protein) cannot be connected to many other subunits. This yields the following problem for any family $\mathcal{F}$ of graphs and an integer k.

($\Delta \leq k$)-$\mathcal{F}$-OVERLAY
Input: A hypergraph H.
Question: Does there exist a graph G $\mathcal{F}$-overlaying H such that $\Delta(G) \leq k$?

We denote by $over_{\mathcal{F}}(H,G)$ the number of hyperedges of H that are $\mathcal{F}$-overlaid by G. A natural generalization is to find $over_{\mathcal{F}}(H,k)$, the maximum number of hyperedges $\mathcal{F}$-overlaid by a graph with maximum degree at most k.

Max $(\Delta \leq k)$-$\mathcal{F}$-overlay
Input: A hypergraph H and a positive integer s.
Question: Does there exist a graph G such that $\Delta(G) \leq k$ and $over_{\mathcal{F}}(H,G) \geq s$?

Observe that there is an obvious reduction from $(\Delta \leq k)$-$\mathcal{F}$-overlay to Max $(\Delta \leq k)$-$\mathcal{F}$-overlay (by setting $s = |E(H)|$).

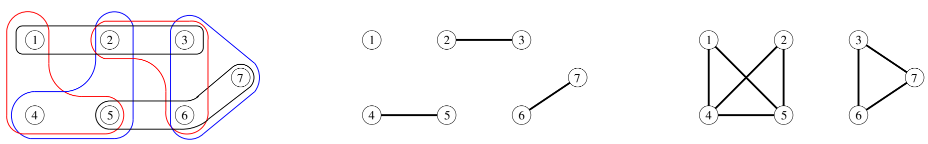

Fig. 1: Example of $(\Delta \leq k)$-$\mathcal{F}$-overlay and max $(\Delta \leq k)$-$\mathcal{F}$-overlay. In the figure, an instance H (left), a graph G with $\Delta(G) \leq 1$ that O_3-overlays H (with O_3 being the graph with three vertices and one edge) (center), and a solution to Max $(\Delta \leq 3)$-C_3-overlay (with C_3 being the cycle on three vertices) (right).

In this paper, we mainly consider the case when the family $\mathcal{F}$ contains a unique graph F. We abbreviate $(\Delta \leq k)$-$\{F\}$-Overlay and Max $(\Delta \leq k)$-$\{F\}$-Overlay as $(\Delta \leq k)$-F-Overlay and Max $(\Delta \leq k)$-F-Overlay, respectively. By definition those two problems really make sense only for $|F|$-uniform hypergraphs *i.e.* hypergraphs whose hyperedges are of size $|F|$. Therefore, we always assume the hypergraph to be $|F|$-uniform.

If F is a graph with maximum degree greater than k, then solving $(\Delta \leq k)$-F-Overlay or Max $(\Delta \leq k)$-F-Overlay is trivial as the answer is always 'No'. So we only study the problems when $\Delta(F) \leq k$.

If F is an empty graph, then Max $(\Delta \leq k)$-F-Overlay is also trivial, because for any hypergraph H, the empty graph on $V(H)$ vertices F-overlays H. Hence the first natural interesting cases are the graphs with one edge. For every integer $p \geq 2$, we denote by O_p the graph with p vertices and one edge. In Section 2, we prove the following dichotomy theorem.

Theorem 1. *Let $k \geq 1$ and $p \geq 2$ be integers. If $p = 2$ or if $k = 1$ and $p = 3$, then* Max $(\Delta \leq k)$-O_p-Overlay *and* $(\Delta \leq k)$-O_p-Overlay *are polynomial-time solvable. Otherwise, they are $\mathcal{NP}$-complete.*

Then, in Section 3, we give a complete polynomial/$\mathcal{NP}$-complete dichotomy for the Max $(\Delta \leq k)$-F-Overlay problems.

Theorem 2. Max $(\Delta \leq k)$-F-Overlay *is polynomial-time solvable if either $\Delta(F) > k$, or F is an empty graph, or $F = O_2$, or $k = 1$ and $F = O_3$. Otherwise it is $\mathcal{NP}$-complete.*

In Section 4, we investigate the complexity of $(\Delta \leq k)$-F-Overlay problems. We believe that each such problem is either polynomial-time solvable or $\mathcal{NP}$-complete. However the dichotomy seems to be more complicated than the one for Max $(\Delta \leq k)$-F-Overlay. We exhibit several pairs (F, k) such that $(\Delta \leq k)$-F-Overlay is polynomial-time solvable, while Max $(\Delta \leq k)$-F-Overlay is $\mathcal{NP}$-complete. This is in particular the case when F is a complete graph (Proposition 3), F is connected k-regular (Proposition 4), F is a path and $k = 2$ (Theorem 8), and when F is the cycle on 4 vertices and $k \leq 3$ (Theorem 7).

Due to space constraints, some proofs (marked with a $\star$) were omitted.

Most notations of this paper are standard. We now recall some of them, and we refer the reader to [3] for any undefined terminology. For a positive integer p, let $[p] = \{1, \dots, p\}$.

Given $S \subseteq V(G)$, we denote by $G[S]$ the subgraph induced by S, that is the subgraph with vertex set S and edge set $\{uv \in E(G) \mid u, v \in S\}$. We denote by E_k the *edgeless graph* on k vertices, that is the graph with k vertices and no edges. The *disjoint union* of two graphs F and G is denoted by $F + G$.

Let H be a hypergraph. Two hyperedges are *adjacent* if their intersection has size at least 2. A hypergraph is *neat* if any two distinct hyperedges intersect in at most one vertex. In other words, a hypergraph is neat if there is no pair of adjacent hyperedges. We denote by $K(H)$, the graph obtained by replacing each hyperedge by a complete graph. In other words, $V(K(H)) = V(H)$ and $E(K(H)) = \{xy \mid \exists S \in E(H), \{x, y\} \subseteq S\}$. The *edge-weight function induced by* H on $K(H)$, denoted by w_H, is defined by $w_H(e) = |\{S \in E(H) \mid e \subseteq S\}|$. In words, $w_H(e)$ is the number of hyperedges of H containing e. A hypergraph H is *connected* if $K(H)$ is connected, and the *connected components* of a hypergraph H are the connected components of $K(H)$. Finally, a graph G $\mathcal{F}$-overlaying H with maximum degree at most k is called an $(\mathcal{F}, H, k)$*-graph*.

2 The graphs with one edge

In this section, we establish Theorem 1. Let $p \geq 2$, and H be a p-uniform hypergraph. Consider the edge-weighted graph $(K(H), w_H)$. For every matching M of this graph, let $G_M = (V(H), M)$. Every hyperedge O_p-overlaid by G_M contains at least one edge of M and at most $\lfloor \frac{p}{2} \rfloor$ edges of M. We thus have the following:

Observation 3 *For every matching M of $K(H)$, we have:*

$$\frac{1}{\lfloor \frac{p}{2} \rfloor} w_H(M) \leq over_{O_p}(H, G_M) \leq w_H(M), \tag{1}$$

where $w_H(M) = \sum_{e \in M} w_H(e)$.

Consider first the case when $p = 2$. Let H be a 2-uniform hypergraph. Every hyperedge is an edge, so $K(H) = H$. Moreover, a (hyper)edge of H is O_2-overlaid by G if and only if it is in $E(G)$. Hence Max $(\Delta \leq k)$-O_2-Overlay is equivalent to finding a maximum k-matching (that is a subgraph with maximum degree at most k) in $K(H)$. This problem is polynomial-time solvable, see [19, Chap. 31], hence:

Proposition 1. Max $(\Delta \leq k)$-O_2-Overlay *is polynomial-time solvable for all positive integer* k*.*

If $p = 3$, Inequalities (1) are equivalent to $over_{O_3}(H, G_M) = w_H(M)$. Since the edge set of a graph with maximum degree 1 is a matching, Max $(\Delta \leq 1)$-O_3-Overlay is equivalent to finding a maximum-weight matching in the edge-weighted graph $(K(H), w_H)$. This can be done in polynomial-time, see [15, Chap. 14].

Proposition 2. Max $(\Delta \leq 1)$-O_3-Overlay *is polynomial-time solvable.*

We shall now prove that if $p \geq 4$, or $p = 3$ and $k \geq 2$, then Max $(\Delta \leq k)$-O_p-Overlay is $\mathcal{NP}$-complete. We prove it by a double induction on k and p. Theorems 4 and 5 first prove the base cases of the induction and Lemma 1 corresponds to the inductive steps.

Theorem 4 ($\star$). $(\Delta \leq 1)$-O_4-Overlay *is* $\mathcal{NP}$*-complete.*

Theorem 5 ($\star$). $(\Delta \leq 2)$-O_3-Overlay *is* $\mathcal{NP}$*-complete.*

Lemma 1 ($\star$). *If* $(\Delta \leq k)$-O_p-Overlay *is* $\mathcal{NP}$*-complete, then* $(\Delta \leq k)$-O_{p+1}-Overlay *and* $(\Delta \leq k+1)$-O_p-Overlay *are* $\mathcal{NP}$*-complete.*

Propositions 1 and 2, Theorems 4 and 5, and Lemma 1 imply Theorem 1.

3 Complexity of Max $(\Delta \leq k)$-F-Overlay

The aim of this section is to establish Theorem 2 that gives the polynomial/$\mathcal{NP}$-complete dichotomy for the Max $(\Delta \leq k)$-F-Overlay problems.

As noticed in the introduction, if $\Delta(F) > k$ or F is an empty graph then Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay is trivially polynomial-time solvable. Moreover, by Propositions 1 and 2, Max $(\Delta \leq 1)$-O_3-Overlay as well as Max $(\Delta \leq k)$-O_2-Overlay (for all positive integers k) are also polynomial-time solvable.

We shall now prove that if we are not in one of the above cases, then Max $(\Delta \leq k)$-F-Overlay is $\mathcal{NP}$-complete. We first establish the $\mathcal{NP}$-completeness when F has no isolated vertices.

Theorem 6. *Let* F *be a graph on at least three vertices with no isolated vertices. If* $k \geq \Delta(F)$*, then* Max $(\Delta \leq k)$-F-Overlay *is* $\mathcal{NP}$*-complete on neat hypergraphs.*

Proof. Assume $k \geq \Delta(F)$. Let $n = |F|$, $a_1, \ldots, a_n$ be an ordering of the vertices of F such that $\delta(F) = d(a_1) \leq d(a_2) \leq \cdots \leq d(a_n) = \Delta(F)$.
Let $\gamma = \lfloor k/\delta(F) \rfloor - 1$, $\beta = k - \gamma\delta(F)$. Observe that $\delta(F) \leq \beta \leq 2\delta(F) - 1$.

We shall give a reduction from Independent Set which is a well-known $\mathcal{NP}$-complete problem even for cubic graphs (see [16].) We distinguish two cases depending on whether $d(a_2) > \beta$ or not. The two reductions are very similar.

Case 1: $d(a_2) > \beta$. Set $\gamma_1 = \gamma_2 = \lfloor (k - d(a_2))/\delta(F) \rfloor$ and $\gamma_3 = \lfloor (k - d(a_3))/\delta(F) \rfloor$.

Let Γ be a cubic graph. For each vertex $v \in V(\Gamma)$, let $(e_1(v), e_2(v), e_3(v))$ be an ordering of the edges incident to v. We shall construct the neat hypergraph $H = H(\Gamma)$ as follows.

- For each vertex $v \in \Gamma$, we create a hyperedge $S_v = \{a_1^v, \dots, a_n^v\}$. Then, for $1 \leq i \leq 3$, we add γ_i a_i^v-*leaves*, that are hyperedges containing a_i^v and $n-1$ new vertices.
- For each edge $e = uv \in \Gamma$, let i and j be the indices such that $e = e_i(u) = e_j(v)$. We create a new vertex z_e and hyperedges S_u^e (S_v^e) containing z_e, a_i^u (a_j^v), and $n-2$ new vertices, respectively. Then, we add γ z_e-*leaves*, that are hyperedges containing z_e and $n-1$ new vertices.

We shall prove that $over_F(H,k) = (\gamma_1+\gamma_2+\gamma_3)|V(\Gamma)| + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$, where $\alpha(\Gamma)$ denotes the cardinality of a maximum independent set in Γ.

The following claim shows that there are optimal solutions with specific structure. This leads to the inequality:

$$over_F(H,k) \leq (\gamma_1+\gamma_2+\gamma_3)|V(\Gamma)| \; + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$$

Claim 1 (⋆) *There is a graph G with $\Delta(G) \leq k$ that F-overlays $over_F(H,k)$ hyperedges of H such that:*

(a) *each x-leaf L is F-overlaid and x is incident to $\delta(F)$ edges in $G[L]$ (with $x = a_i^v$ or $x = z_e$).*

(b) *for each edge $e = uv \in E(\Gamma)$, exactly one of the two hyperedges S_u^e and S_v^e is F-overlaid. Moreover if S_u^e (S_v^e) is F-overlaid, then a_i^u (a_j^v) is incident to $d(a_2)$ edges in S_u^e (S_v^e), respectively.*

(c) *the set of vertices v such that S_v is F-overlaid is an independent set in Γ.*

Conversely, consider W a maximum independent set of Γ.

Let G be the graph with vertex $V(H)$ which is the union of the following subgraphs :

- for each x-leaf L, we add a copy of F on L in which x has degree $\delta(F)$;
- for each vertex $v \in W$, we add a copy of F on S_v in which a_i^v has degree $d(a_i)$ for all $1 \leq i \leq n$.
- for each edge $e \in E(\Gamma)$, we choose an endvertex u of e such that $u \notin W$, and add a copy of F in which z_e has degree $d(a_1)$ and a_i^u has degree $d(a_2)$ (with i the index such that $e_i(u) = e$).

It is simple matter to check that $\Delta(G) \leq k$ and that G F-overlays $(\gamma_1+\gamma_2+\gamma_3)|V(\Gamma)| + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$ hyperedges of H. Thus $over_F(H,k) \geq (\gamma_1+\gamma_2+\gamma_3)|V(\Gamma)| \; + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$.

Case 2: $d(a_2) \leq \beta$. The proof is very similar to Case 1. The main difference is the definition of the γ_i. In this case, we set $\gamma_i = \lfloor (k - d(a_i))/\delta(F) \rfloor$ for $1 \leq i \leq 3$, and we can adapt the proof of Claim 1.

Conversely, if we have W a maximum independent set of Γ, then we construct graph G, union of the subgraphs as Case 1 except the subgraphs for hyperedges $S_u^{e_1(v)}$, that we add a copy of F in which $d(z_e) = d(a_2)$ and $d(a_1^u) = d(a_1)$.

We then establish the following lemma, which allows to derive the $\mathcal{NP}$-completeness of Max $(\Delta \leq k)$-F-Overlay when F has isolated vertices.

Lemma 2 ($\star$). *Let k be a positive integer, let F be a graph with $\delta(F) \geq 1$, and let q be a non-negative integer. If* Max $(\Delta \leq k)$-$(F + E_q)$-Overlay *is $\mathcal{NP}$-complete, then* Max $(\Delta \leq k)$-$(F + E_{q+1})$-Overlay *is also $\mathcal{NP}$-complete.*

Now we can prove Theorem 2. As explained in the beginning of the section, it suffices to prove that Max $(\Delta \leq k)$-F-Overlay remains $\mathcal{NP}$-complete when $\Delta(F) \leq k$, $F \neq E_{|F|}$, $|F| \geq 3$ and $(F, k) \neq (O_3, 1)$. Assume that the above conditions are satisfied. Let F' be the graph induced by the non-isolated vertices of F. Then $F = F' + E_q$ with $q = |F| - |F'|$. If $|F'| = 2$, then $F = O_{|F|}$, and we have the result by Theorem 1. If $|F'| \geq 3$, then the result follows from Theorem 6, Lemma 2, and an immediate induction.

4 Complexity of $(\Delta \leq k)$-$\mathcal{F}$-Overlay

4.1 Regular graphs

Proposition 3. *For every complete graph K and every positive integer k, $(\Delta \leq k)$-K-*Overlay *is polynomial-time solvable.*

Proof. Observe that a $|V(K)|$-uniform hypergraph H is a positive instance of $(\Delta \leq k)$-K-Overlay if and only if $K(H)$ is a (K, H, k)-graph.

Proposition 4. *For every connected k-regular graph F, $(\Delta \leq k)$-F-*Overlay *is polynomial-time solvable.*

Proof. One easily sees that a $|V(F)|$-uniform hypergraph H admits an (F, H, k)-graph if and only if the hyperedges of H are pairwise non-intersecting.

Let C_4 denote the cycle on 4 vertices. Proposition 4 implies that $(\Delta \leq 2)$-C_4-Overlay is polynomial-time solvable. We now show that $(\Delta \leq 3)$-C_4-Overlay is also polynomial-time solvable.

Theorem 7. $(\Delta \leq 3)$-C_4-Overlay *is polynomial-time solvable.*

Proof. Let H be a 4-uniform hypergraph.

Let us describe an algorithm to decide whether there is a $(C_4, H, 3)$-graph. It is sufficient to do it when H is connected since the disjoint union of the $(C_4, K, 3)$-graphs for connected components K of H is a $(C_4, H, 3)$-graph.

Observe first that if two hyperedges of H intersect in exactly one vertex u, then no such graph exists, since u must have degree 2 in each of the hyperedges if they are C_4-overlaid, and thus degree 4 in total. Therefore if there are two such hyperedges, we return 'No'. At this point we may assume that $|E(H)| \geq 2$ for otherwise we return 'Yes'.

From now on we may assume that two hyperedges either do not intersect, or are adjacent (intersect on at least two vertices).

Claim 2 *If two hyperedges S_1 and S_2 intersect on three vertices and there is a $(C_4, H, 3)$-graph G, then $|V(H)| \leq 6$.*

Proof of claim: Assume $S_1 = \{a_1, b, c, d\}$ and $S_2 = \{a_2, b, c, d\}$. Let G be a $(C_4, H, 3)$-graph. In G, a_1 and a_2 have the same two neighbours in $\{b, c, d\}$ and the third vertex of $\{b, c, d\}$ is also adjacent to those two. Consider a hyperedge S_3 intersecting $S_1 \cup S_2$. Since it is C_4-overlaid by G, at least two edges connect $S_3 \cap (S_1 \cup S_2)$ to $S_3 \setminus (S_1 \cup S_2)$. The endvertices of those edges in $S_1 \cup S_2$ must have degree 2 in $G[S_1 \cup S_2]$. Hence, without loss of generality, either $S_3 = \{a_1, a_2, b, e\}$, or $S_3 = \{a_1, b, c, e\}$ for some vertex e not in $S_1 \cup S_2$. Now no hyperedge can both intersect $S_1 \cup S_2 \cup S_3$ and contain a vertex not in $S_1 \cup S_2 \cup S_3$, for such a hyperedge must contain either the vertices c, e or a_2, e which are at distance 3 in $G[S_1 \cup S_2 \cup S_3]$. (However there can be more hyperedges contained in $S_1 \cup S_2 \cup S_3$.) Hence $|V(H)| \leq 6$. $\triangleleft$

In view of Claim 2, if there are two hyperedges with three vertices in common, either we return 'No' if $|V(H)| > 6$, or we check all possibilities (or follow the proof of the above claim) to return the correct answer otherwise. Henceforth, we may assume that any two adjacent hyperedges intersect in exactly two vertices.

Let S_1 and S_2 be two adjacent hyperedges, say $S_1 = \{a, b, c, d\}$ and $S_2 = \{c, d, e, f\}$. Note that every $(C_4, H, 3)$-graph contains the edges ab, cd and ef, and that $N(c) \cup N(d) = S_1 \cup S_2$.

Claim 3 *If there is another hyperedge than S_1 and S_2 containing c or d, and there is a $(C_4, H, 3)$-graph G, then $|V(H)| \leq 8$.*

Proof of claim: Without loss of generality, we may assume that G contains the cycle (a, b, d, f, e, c, a) and the edge cd. Hence the only possible hyperedges containing c or d and a vertex not in $S_1 \cup S_2$ are $S_3 = \{a, c, e, g\}$ for some $g \notin S_1 \cup S_2$ and $S_4 = \{b, d, f, h\}$ for some $h \notin S_1 \cup S_2$.

If H contains both S_3 and S_4, then G contains the edges ag, eg, bh and hf. If G contains also gh, then $G[S_1 \cup S_2 \cup S_3 \cup S_4]$ is 3-regular, so $G = G[S_1 \cup S_2 \cup S_3 \cup S_4]$. If G does not contain gh, then the only vertices of degree 2 in $G[S_1 \cup S_2 \cup S_3 \cup S_4]$ are g and h, and they are at distance at least 3 in this graph. Thus every hyperedge intersecting $S_1 \cup S_2 \cup S_3 \cup S_4$ is contained in this set, so $|V(H)| = 8$.

Assume now that G contains only one of S_3, S_4. Without loss of generality, we may assume that this is S_3. Hence G also contains the edges ag and eg. If $V(G) \neq S_1 \cup S_2 \cup S_3$, then there is a hyperedge S that intersects $S_1 \cup S_2 \cup S_3$ and that is not contained in $S_1 \cup S_2 \cup S_3$. It does not contain c and d. Hence it must contain one of the vertices a or e, because it intersects each S_i along an edge of G or not at all. Without loss of generality, $a \in S$. Hence $S = \{a, b, i, g\}$ for some vertex i not in $S_1 \cup S_2 \cup S_3$, and G contains the edges bi and ig. Now, as previously, either i and f are adjacent and $G = G[S_1 \cup S_2 \cup S_3 \cup S]$ or they are not adjacent, and every hyperedge intersecting $S_1 \cup S_2 \cup S_3 \cup S$ is contained in this set. In both cases, $|V(H)| = 8$. $\triangleleft$

We now summarize the algorithm: if $|V(G)| \leq 8$, then we solve the instance by brute force. Otherwise, for every pair of hyperedges S_1, S_2, if their intersection is of size 1 or 3, we answer 'No'. In the remaining cases, if S_1 and S_2 have non-empty intersection, then, they must intersect on two vertices c and d, and these vertices do not belong to any other hyperedges but S_1 and S_2.

In this case, let H' be the hypergraph with vertex set $V(H) \setminus \{c, d\}$ and hyperedge set $(E(H) \cup \{\{a, b, e, f\}\}) \setminus \{S_1, S_2\}$. It is simple matter to check that there is a $(C_4, H, 3)$-graph if and only if there is a $(C_4, H', 3)$-graph. Consequently, we recursively apply the algorithm on H'.

Clearly, the above-described algorithm runs in polynomial time.

4.2 Paths

Let $\mathcal{P}$ be the set of all paths. We have the following:

Theorem 8. $(\Delta \leq 2)$-$\mathcal{P}$-Overlay *is linear-time solvable.*

Proof. Clearly, if H is not connected, it suffices to solve the problem on each of the components and to return 'No' if the answer is negative for at least one of the components, and 'Yes' otherwise. Henceforth, we shall now assume that H is connected. In such a case, a $(\mathcal{P}, H, 2)$-graph is either a path or a cycle. However, if H is $\mathcal{P}$-overlaid by a path P, then it is also $\mathcal{P}$-overlaid by the cycle obtained from P by adding an edge between its two endvertices. Thus, we focus on the case where G is a cycle.

Let $\mathcal{S}$ be a family of sets. The *intersection graph* of a set $\mathcal{S}$ is the graph $IG(\mathcal{S})$ whose vertices are the sets of $\mathcal{S}$, and in which two vertices are adjacent if the corresponding sets in S intersect.

The *intersection graph* of a hypergraph H, denoted by $IG(H)$, is the intersection graph of its hyperedge set. We define two functions l_H and s_H as follows:

$$l_H(S) = |S| - 1 \text{ for all } S \in E(H) \text{ and } s_H(S, S') = |S \cap S'| - 1 \text{ for all } S, S' \in E(H).$$

Let $\mathbb{C}_\ell$ be the circle of circumference ℓ. We identify the points of $\mathbb{C}_\ell$ with the integer numbers (points) of the segment $[0, \ell]$, (with 0 identified with ℓ). A *circular-arc graph* is the intersection graph of a set of arcs on $\mathbb{C}_\ell$. A set $\mathcal{A}$ of arcs such that $IG(\mathcal{A}) = G$ is called an *arc representation* of G. We denote by A_v the arc corresponding to v in $\mathcal{A}$. Let G be a graph and let $l : V(G) \to \mathbb{N}$ and $s : E(G) \to \mathbb{N}$ be two functions. An arc representation $\mathcal{A}$ of G is *l-respecting* if A_v has length $l(v)$ for any $v \in V(G)$, *s-respecting* if $A_v \cap A_u$ has length $s(u, v)$ for all $uv \in E(G)$, and *(l, s)-respecting* if it is both l-respecting and s-respecting. One can easily adapt the algorithm given by Köbler et al. [17] for (l, s)-respecting interval representations to decide in linear time whether a graph admits an (l, s)-respecting arc representation in $\mathbb{C}_n$ for every integer n.

Claim 4 *Let H be a connected hypergraph on n vertices. There is a cycle $\mathcal{P}$-overlaying H if and only if $IG(H)$ admits an (l_H, s_H)-respecting arc representation into $\mathbb{C}_n$.*

Proof of claim: Assume that H is $\mathcal{P}$-overlaid by a cycle $C = (v_0, v_1, \dots, v_{n-1}, v_0)$. There is a canonical embedding of C to $\mathbb{C}_n$ in which every vertex v_i is mapped to i and every edge $v_i v_{i+1}$ to the circular arc $[i, i+1]$. For every hyperedge $S \in E(H)$, $P[S]$ is a subpath, which is mapped to the circular arc A_S of $\mathbb{C}_n$ that is the union of the circular arcs to which its edges are mapped. Clearly, $\mathcal{A} = \{A_S \mid S \in E(H)\}$ is an (l_H, s_H)-respecting interval representation of $IG(H)$.

Conversely, assume that $IG(H)$ admits an (l_H, s_H)-respecting interval representation $\mathcal{A} = \{A_S \mid S \in E(H)\}$ into $\mathbb{C}_n$. Let S_0 be a hyperedge of minimum size. Free to rotate all intervals, we may assume that A_{S_0} is $[1, |S_0|]$. Now since $\mathcal{A}$ is (l_H, s_H)-respecting and H is connected, we deduce that the extremities of A_S are integers for all $S \in E(H)$. Let v_1 be a vertex of H that belongs to the hyperedges whose corresponding arcs of $\mathcal{A}$ contain 1. Then for all $i = 2$ to $n = |V(H)|$, denote by v_i an arbitrary vertex not in $\{v_1, \dots, v_{i-1}\}$ that belongs to the hyperedges whose corresponding arcs of $\mathcal{A}$ contain i. Such a vertex exists because $\mathcal{A}$ is (l_H, s_H)-respecting. Observe that such a construction yields $S = \{v_i \mid i \in A_S\}$ for all $S \in E(H)$. Furthermore, the cycle $C = (v_1, \dots, v_n, v_1)$ $\mathcal{P}$-overlays H. Indeed, for each $S \in E(H)$, $C[S]$ is the subpath corresponding to A_S, that is $V(C[S]) = \{v_i \mid i \in A_S\}$ and $E(C[S]) = \{v_i v_{i+1} \mid [i, i+1] \subseteq A_S\}$. $\triangleleft$

The algorithm to solve $(\Delta \leq 2)$-$\mathcal{P}$-Overlay for a connected hypergraph H in linear time is thus the following:

1. Construct the intersection graph $IG(H)$ and compute the associated functions l_H and s_H.
2. Check whether graph $IG(H)$ has an (l_H, s_H)-respecting interval representation. If it is the case, return 'Yes'. If not return 'No'.

Remark 1. We can also detect in polynomial time whether a connected hypergraph H is $\mathcal{P}$-overlaid by a path. Indeed, similarly to Claim 4, one can show that there is a path $\mathcal{P}$-overlaying H if and only if $IG(H)$ admits an (l_H, s_H)-respecting interval representation.

5 Further research

Theorem 2 characterizes the complexity of Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay when $\mathcal{F}$ contains a unique graph. It would be nice to extend this characterization to families $\mathcal{F}$ of arbitrary size.

Problem 1. Characterize the pairs $(\mathcal{F}, k)$ for which Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay is polynomial-time solvable and those for which it is $\mathcal{NP}$-complete.

Theorem 1 and the results obtained in Section 4 give a first view of the complexity of $(\Delta \leq k)$-F-Overlay. A natural problem is to close the dichotomy:

Problem 2. Characterize the pairs (F, k) for which $(\Delta \leq k)$-F-Overlay is polynomial-time solvable and those for which it is $\mathcal{NP}$-complete.

It would be interesting to consider the complexity of this problem when F is k-regular but non-connected, and when F is a cycle. In order to attack Problem 2, it would be helpful to prove the following conjecture.

Conjecture 1. If $(\Delta \leq k)$-F-Overlay is $\mathcal{NP}$-complete, then $(\Delta \leq k+1)$-F-Overlay is also $\mathcal{NP}$-complete.

Furthermore, for each pair $(\mathcal{F}, k)$ such that Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay is $\mathcal{NP}$-complete and $(\Delta \leq k)$-$\mathcal{F}$-Overlay is polynomial-time solvable, it is natural to consider the parameterized complexity of Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay when parameterized by $|E(H)| - s$, because $(\Delta \leq k)$-$\mathcal{F}$-Overlay is the case $s = 0$.

Finally, it would be interesting to obtain approximation algorithms for Max $(\Delta \leq k)$-$\mathcal{F}$-Overlay when this problem is $\mathcal{NP}$-complete.

References

1. D. Agarwal, C. Caillouet, D. Coudert, and F. Cazals. Unveiling Contacts within Macro-molecular assemblies by solving Minimum Weight Connectivity Inference Problems. *Molecular and Cellular Proteomics*, 14:2274–2284, Apr. 2015.
2. D. Angluin, J. Aspnes, and L. Reyzin. Inferring social networks from outbreaks. In *ALT '10*, pages 104–118, 2010.
3. A. J. Bondy and U. S. R. Murty. *Graph theory*, volume 244 of *Graduate Texts in Mathematics*. Springer, New York, 2008.
4. U. Brandes, S. Cornelsen, B. Pampel, and A. Sallaberry. Blocks of hypergraphs - applied to hypergraphs and outerplanarity. In *IWOCA '10*, pages 201–211, 2010.
5. R. Burnley, E. Damoc, E. Denisov, A. Makarov, and A. Heck. High-sensitivity orbitrap mass analysis of intact macromolecular assemblies. *Nature methods*, 9, 10 2012.
6. J. Chen, C. Komusiewicz, R. Niedermeier, M. Sorge, O. Suchý, and M. Weller. Polynomial-time data reduction for the subset interconnection design problem. *SIAM J. Discrete Math.*, 29(1):1–25, 2015.
7. G. Chockler, R. Melamed, Y. Tock, and R. Vitenberg. Constructing scalable overlays for pub-sub with many topics. In *PODC '07*, pages 109–118. ACM, 2007.
8. N. Cohen, F. Havet, D. Mazauric, I. Sau, and R. Watrigant. Complexity dichotomies for the minimum F-overlay problem. *Journal of Discrete Algorithms*, 52:133–142, 2019.
9. V. Conitzer, J. Derryberry, and T. Sandholm. Combinatorial auctions with structured item graphs. In *16th Conference on Innovative Applications of Artificial Intelligence (AAAI 2004)*, pages 212–218, 2004.
10. D. Z. Du and D. F. Kelley. On complexity of subset interconnection designs. *Journal of Global Optimization*, 6(2):193–205, 1995.
11. D.-Z. Du and Z. Miller. Matroids and subset interconnection design. *SIAM Journal on Discrete Mathematics*, 1(4):416–424, 1988.
12. H. Fan, C. Hundt, Y.-L. Wu, and J. Ernst. Algorithms and implementation for interconnection graph problem. In *Combinatorial Optimization and Applications*, pages 201–210, 2008.
13. H. Fan and Y. Wu. Interconnection graph problem. In *Proceedings of the 2008 International Conference on Foundations of Computer Science, FCS 2008*, pages 51–55, 2008.

14. 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.
15. D. Jungnickel. *Graphs, Networks and Algorithms*, volume 5 of *Algorithms and Computation in Mathematics*. Springer-Verlag Berlin Heidelberg, 2013.
16. R. M. Karp. *Reducibility among Combinatorial Problems*. Springer US, 1972.
17. J. Köbler, S. Kuhnert, and O. Watanabe. Interval graph representation with given interval and intersection lengths. *J. Discrete Algorithms*, 34:108–117, 2015.
18. M. Onus and A. W. Richa. Minimum maximum-degree publish-subscribe overlay network design. *IEEE/ACM Trans. Netw.*, 19(5):1331–1343, Oct. 2011.
19. A. Schrijver. *Combinatorial optimization. Polyhedra and efficiency. Vol. A*, volume 24 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2003.
20. M. Sharon and C. V. Robinson. The role of mass spectrometry in structure elucidation of dynamic protein complexes. *Annual Review of Biochemistry*, 76(1):167–193, 2007. PMID: 17328674.