Overlaying a hypergraph with a graph with bounded maximum degree

Frédéric Havet[1,2], Dorian Mazauric[2], Viet-Ha Nguyen[1,2] and Rémi Watrigant[3]

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

RESEARCH REPORT
N° 9258
February, 2019
Project-Team ABS, COATI

ISSN 0249-6399 ISRN INRIA/RR--9258--FR+ENG

Overlaying a hypergraph with a graph with bounded maximum degree

Frédéric Havet[1,2], Dorian Mazauric[2], Viet-Ha Nguyen[1,2] and Rémi Watrigant[3]

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

Project-Team ABS, COATI

Research Report n° 9258 — February, 2019 — 19 pages

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).

Key-words: Hypergraph, graph, algorithm, complexity, computational structural biology

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Couvrant d'un hypergraphe avec un graphique à degré maximal borné

Résumé : Soit G et H respectivement un graphe et un hypergraphe défini sur un même ensemble de sommets, et F un graphe fixé. Nous disons que G F-*couvre* un hyperedge S de H si F est un sous-graphe couvrant du sous-graphe de G induit par S, et qu'il le F- *couvre* H si F recouvre chaque hyperedge de H. Motivés par la biologie structurale, nous étudions la complexité algorithmique de deux problèmes. Le premier problème, $(\Delta \leq k)$ F-Overlay, consiste à décider s'il existe ou non un graphe avec un degré maximum égal à k qui F couvre un hypergraphe H. C'est un cas particulier du deuxième problème Max $(\Delta \leq k)$ F-Overlay, qui prend en entrée un hypergraphe H et un entier s, et consiste à décider s'il y a un graphe avec degré maximum d'au plus k que F-couvre au moins s hyperges de H. Nous donnons une dichotomie complète polynomiale/$\mathcal{NP}$-complète pour les problemes Max $(\Delta \leq k)$-F-Overlay en fonction des paires (F, k), et la complexité de $(\Delta \leq k)$ F-Overlay pour plusieurs paires (F, k).

Mots-clés : Hypergraphe, graphe, algorithme, complexité, biologie structurale computationnelle

Contents

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.

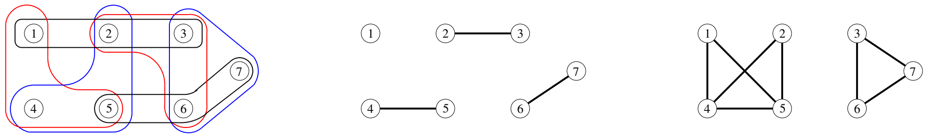
As said previously, the graph sought will also have to satisfy some global constraints. Since in a macro-molecular assembly the number of contacts is small, the first natural idea is to look for a graph G with the minimum number of edges. This leads to the MIN-$\mathcal{F}$-OVERLAY problem: given a hypergraph 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,

Figure 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).



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 $\text{over}_{\mathcal{F}}(H, G)$ the number of hyperedges of H that are $\mathcal{F}$-overlaid by G. A natural generalization is to find $\text{over}_{\mathcal{F}}(H, k)$, 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 $\text{over}_{\mathcal{F}}(H, G) \geq s$?

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

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 11), F is connected k-regular (Proposition 12), F is a path and $k = 2$ (Theorem 14), and when F is the cycle on 4 vertices and $k \leq 3$ (Theorem 13).

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 $[p] = \{1, \ldots, p\}$.

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:*

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

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

If $p = 3$, Inequalities (1) are equivalent to $\text{over}_{O_3}(H, G_M) = w_H(M)$. 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 5. 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 6 and 7 first prove the base cases of the induction and Lemma 8 corresponds to the inductive steps.

Theorem 6 ($\star$)**.** $(\Delta \leq 1)$-O_4-Overlay *is* $\mathcal{NP}$*-complete.*

Theorem 7 ($\star$)**.** $(\Delta \leq 2)$-O_3-Overlay *is* $\mathcal{NP}$*-complete.*

Lemma 8 ($\star$)**.** *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 4 and 5, Theorems 6 and 7, and Lemma 8 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 4 and 5, 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 9. *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 $n = |F|$, $a_1, \dots, a_n$ 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.

<u>Case 1</u>: $d(a_2) > \beta$. 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 $\text{over}_F(H,k) = (\gamma_1 + \gamma_2 + \gamma_3)|V(\Gamma)| + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$, 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:

$$\text{over}_F(H,k) \leq (\gamma_1 + \gamma_2 + \gamma_3)|V(\Gamma)| + (\gamma+1)|E(\Gamma)| + \alpha(\Gamma)$$

Claim 1 ($\star$)**.** *There is a graph G with $\Delta(G) \leq k$ that F-overlays* $\text{over}_F(H,k)$ *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 $\text{over}_F(H,k) \geq (\gamma_1 + \gamma_2 + \gamma_3)|V(\Gamma)| \ + (\gamma + 1)|E(\Gamma)| + \alpha(\Gamma)$.

<u>Case 2</u>: $d(a_2) \leq \beta$. 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 10 ($\star$)**.** *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 9, Lemma 10, and an immediate induction.

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

4.1 Regular graphs

Proposition 11. *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 12. *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 12 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 13. $(\Delta \leq 3)$-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 $|V(H)| \leq 6$.*

Subproof. 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 $|V(H)| \leq 6$. ♢

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 $|V(H)| \leq 8$.*

Subproof. 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, $|V(H)| = 8$. ◊

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 14. $(\Delta \leq 2)$-$\mathcal{P}$-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:

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

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 $\mathbb{C}_n$.*

Subproof. 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_iv_{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 $E(C[S]) = \{v_iv_{i+1} \mid [i, i+1] \subseteq A_S\}$. $\Diamond$

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'.

$\square$

***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 Macromolecular 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.

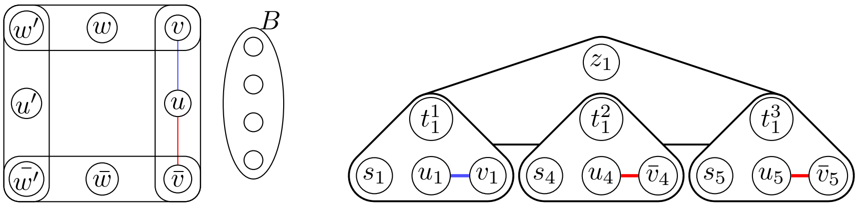
[17] J. Köbler, S. Kuhnert, and O. Watanabe. Interval graph representation with given interval and intersection lengths. *J. Discrete Algorithms*, 34:108–117, 2015.

[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.

Figure 2 The construction of a variable gadget (left) and an example of clause gadget of $C_1 = (x_1 \vee \bar{x_4} \vee \bar{x_5})$ (right).



A Proof of Theorem 6

We reduce 3-SAT to $(\Delta \leq 1)$-O_4-Overlay.

We shall need the following gadget hypergraph $J = J(A, B)$. $V(J) = A \cup B$ with $|A| = 3$ and $|B| = 4$ and $E(J) = \{A \cup \{b\} \mid b \in B\}$.

Claim 5. *Every $(O_4, J, 1)$-graph contains exactly one edge with both ends in A.*

Subproof. Let G be an $(O_4, J, 1)$-graph. Each hyperedge of J contains an edge of G, which necessarily has an endvertex in A. Since $\Delta(G) \leq 1$, G has at most three edges with an endvertex in A. Since J has four hyperedges, there is an edge of G contained in two hyperedges of J. The two endvertices of this edge are in A. ♢

Let Φ be an instance of 3-SAT with m clauses $C_1, \dots, C_m$ on n variables $x_1, \dots, x_n$. Let $H = H(\Phi)$ be the 4-uniform hypergraph constructed as follows (see Figure 2).

- For each variable x_i, we create the variable gadget VG_i as follows.
  Its vertex set is $\{u_i, v_i, \bar{v}_i, w_i, \bar{w}_i, w'_i, \bar{w}'_i, u'_i\} \cup B_i$, with $|B_i| = 4$. We add the hyperedges of $J(\{u_i, v_i, \bar{v}_i\}, B_i)$, $J(\{v_i, w_i, w'_i\}, B_i)$, $J(\{\bar{v}_i, \bar{w}_i, \bar{w}'_i\}, B_i)$, and $J(\{u'_i, w'_i, \bar{w}'_i\}, B_i)$.
- For each clause $C_j = \ell_1 \vee \ell_2 \vee \ell_3$, we create the clause hypergraph CG_j. Its vertex set is $\{z_j\} \cup \bigcup_{k=1}^{3} \{s_j^k, t_j^k, u_j^k, v_j^k\}$. Set $Z_j = \{t_j^1, t_j^2, t_j^3, z_j\}$ and for $k \in \{1, 2, 3\}$, set $T_j^k = \{s_j^k, t_j^k, u_j^k, v_j^k\}$. The hyperedge set of CG_j is $\{Z_j, T_j^1, T_j^2, T_j^3\}$. Then for each $k \in \{1, 2, 3\}$, we identify u_j^k and v_j^k with u_i and v_i if $\ell_k = x_i$ and with u_i and $\bar{v}_i$ if $\ell_k = \bar{x}_i$.

Let us prove that there is an $(O_4, H, 1)$-graph of if and only if Φ is satisfiable.

Assume first that there is an $(O_4, H, 1)$-graph G.

Claim 6. *For each $1 \leq i \leq n$, G contains either the edge $u_i v_i$ or the edge $u_i \bar{v}_i$. Moreover, all vertices of $\{u_i, v_i, \bar{v}_i\}$ are adjacent in G to a vertex in VG_i.*

Subproof. Since VG_i contains $J(\{u'_i, w'_i, \bar{w}'_i\}, B_i)$, by Claim 5, G contains an edge e'' with both endvertices in $\{u'_i, w'_i, \bar{w}'_i\}$. Assume that e'' is incident to w'_i. Now VG_i contains $J(\{v_i, w_i, w'_i\}, B_i)$. So, by Claim 5, G contains an edge e' with both endvertices in $\{v_i, w_i, w'_i\}$. Since $\Delta(G) \leq 1$, e' is not incident to w'_i, so $e' = v_i w_i$. Now VG_i contains $J(\{u_i, v_i, \bar{v}_i\}, B_i)$. So, by Claim 5, G contains an edge e with both endvertices in $\{u_i, v_i, \bar{v}_i\}$. This edge is not incident to v_i, so $e = u_i \bar{v}_i$. Similarly, we get that G contains $\bar{v}_i \bar{w}_i$ and $u_i v_i$ if e'' is incident to $\bar{w}'_i$. ♢

In view of this claim, one defines the truth assignment ϕ by $\phi(x_i) = true$ if $u_i v_i \in E(G)$ and $\phi(x_i) = false$ if $u_i \bar{v}_i \in E(G)$. Let us prove that it satisfies Φ. Consider a clause $C_j = \ell_1 \vee \ell_2 \vee \ell_3$ and its corresponding clause gadget CG_j. If there is none of $u_j^k v_j^k$ in $E(G)$ (for all $k \in \{1, 2, 3\}$), that means both u_j^k and v_j^k are incident to an edge adjacent to a vertex in the variable gadget it they belongs to, T_j^k must be O_4-overlaid by the edge $s_j^k t_j^k$. So, t_j^k (for all $k \in \{1,2,3\}$) has degree one in G. However, the hyperedge Z_j is O_4-overlaid by $z_j t_j^k$ or $t_j^k t_j^{k'}$ for $k' \neq k$ and $k, k' \in \{1,2,3\}$, which implies that there is one of vertices t_j^k which is of degree one in $G[Z_j]$. It is a contradiction, thus there is an edge $u_j^k v_j^k \in E(G)$ (for some $k \in \{1,2,3\}$). Hence $\phi(\ell_k) = true$. Hence ϕ satisfies Φ.

Conversely, assume ϕ is a truth assignment satisfying Φ. Let G be the graph with vertex set $V(H)$ and whose edge set is constructed as follows. For every $1 \leq i \leq n$, add $\{u_i v_i, w_i w_i', u_i' \bar{w}_i', \bar{w}_i \bar{v}_i\}$ to $E(G)$ if $\phi(x_i) = true$ and $\{u_i \bar{v}_i, \bar{w}_i \bar{w}_i', u_i' w_i', w_i v_i\}$ to $E(G)$ if $\phi(x_i) = false$. For every clause $C_j = \ell_1 \vee \ell_2 \vee \ell_3$, there exists $k \in \{1,2,3\}$ such that $\phi(\ell_k) = true$. Add the edges $z_j t_j^k$, $s_j^{k+1} t_j^{k+1}$ and $s_j^{k+2} t_j^{k+2}$ (the superscript are modulo 3) to $E(G)$. One easily checks that G is an $(O_4, H, 1)$-graph.

B Proof of Theorem 7

In order to simplify the proof, we will prove that the problem $(\Delta \leq 2)$-O_3-Overlay with handicap is $\mathcal{NP}$-complete, where, in addition to a hypergraph H, we are given a subset of vertices $A \subseteq V(H)$ (called *handicap vertices*), and we ask for a graph G that O_3-overlays H with the additional constraint that for every $v \in A$, the degree of v is at most 1. Observe that we can easily reduce this problem to $(\Delta \leq 2)$-O_3-Overlay: for every $v \in A$, add six vertices $t_v^1, \ldots, t_v^6$, and hyperedges $\{v, t_v^1, t_v^p\}$ for every $p \in \{2,3,4,5,6\}$. Let H' be the obtained hypergraph. Observe that an $(O_3, H, 2)$-graph G in which every $v \in A$ has degree at most 1 can be extended into an $(O_3, H', 2)$-graph by adding the edge vt_v^1 (and all vertices of $V(H') \setminus V(H)$). Conversely, $(O_3, H', 2)$-graph G' must contain the edge vt_v^1, and thus v is of degree at most 1 in $G'[V(H)]$.

We reduce 3-SAT to $(\Delta \leq 2)$-O_3-Overlay with handicap. Let Φ be an instance of 3-SAT with m clauses $C_1, \ldots, C_m$ on n variables $x_1, \ldots, x_n$. We construct a 3-uniform hypergraph H as follows (see Figure 3 for an example):

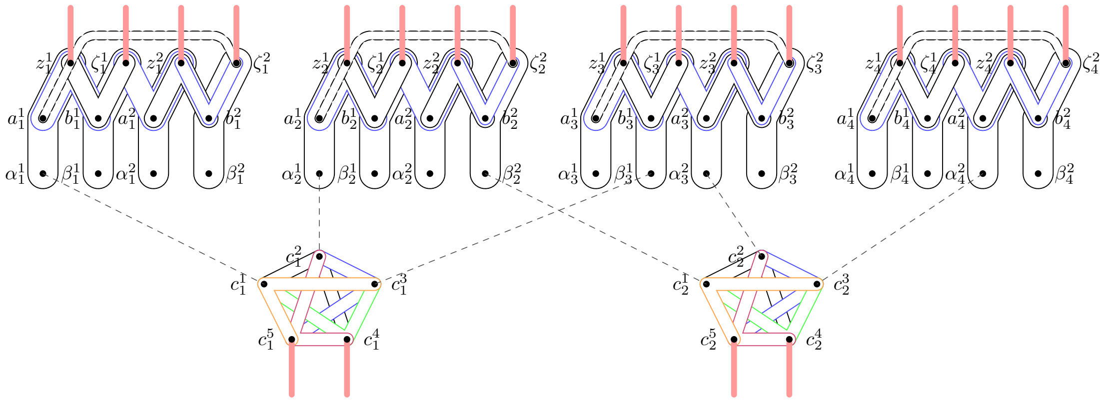
- for every variable x_i, we create a variable gadget VG_i composed of vertices $a_i^1, \ldots, a_i^{p_i}$, $\alpha_i^1, \ldots, \alpha_i^{p_i}$, $b_i^1, \ldots, b_i^{p_i}$, $\beta_i^1, \ldots, \beta_i^{p_i}$, $z_i^1, \ldots, z_i^{p_i}$, and $\zeta_i^1, \ldots, \zeta_i^{p_i}$ where p_i is the number of clauses where x_i appears. Then, for every $j \in [p_i]$, we create the following hyperedges (additions in superscripts are modulo p_i, and we write p_i instead of 0): $h_{i,j}^1 = \{z_i^j, a_i^j, \alpha_i^j\}$, $h_{i,j}^2 = \{z_i^j, b_i^j, \beta_i^j\}$, $h_{i,j}^3 = \{a_i^j, z_i^j, b_i^j\}$, $h_{i,j}^4 = \{z_i^j, b_i^j, \zeta_i^j\}$, $h_{i,j}^5 = \{b_i^j, \zeta_i^j, a_i^{j+1}\}$, $h_{i,j}^6 = \{\zeta_i^j, a_i^{j+1}, z_i^{j+1}\}$.
  Moreover, all vertices $z_i^1, \ldots, z_i^{p_i}$, and $\zeta_i^1, \ldots, \zeta_i^{p_i}$ are handicap vertices.
- for every clause C_j, we create a clause gadget CG_j composed of vertices $c_j^1, \ldots, c_j^5$, and, for every $p \in [5]$, of the hyperedge $\{c_j^p, c_j^{p+1}, c_j^{p+3}\}$ (additions are modulo 5, and we write 5 instead of 0). The vertices c_j^4 and c_j^5 are handicap vertices.
- finally, for every clause $C_j = \ell_1 \vee \ell_2 \vee \ell_3$, if $\ell_p = x_i$ ($\ell_p = \bar{x}_i$) for some $p \in \{1,2,3\}$, then we identify the vertex c_j^p of CG_j with the vertex α_i^j (β_i^j) of VG_i, respectively.

Suppose first that $\phi : \{x_1, \ldots, x_n\} \to \{true, false\}$ is a truth assignment satisfying Φ.

Let us first select a set of edges in the variable gadgets. For every $i \in [n]$, if $\phi(x_i) = true$, then select edges $\{z_i^j, a_i^j\}$, $\{b_i^j, \beta_i^j\}$ and $\{\zeta_i^j, b_i^j\}$ for all $j \in [p_i]$. If $\phi(x_i) = false$, then select edges $\{z_i^j, b_i^j\}$, $\{a_i^j, \alpha_i^j\}$ and $\{\zeta_i^j, a_i^{j+1}\}$ for all $j \in [p_i]$ (additions are modulo p_i). Let E_1 be the set of selected edges, and $G_1 = (V(H), E_1)$. Observe that G_1 O_3-overlays the hyperedges of all variable gadgets: for every $i \in [n]$ and every $j \in [p_i]$,

- $h_{i,j}^1$ contains the edge $\{z_i^j, a_i^j\}$ if $\phi(x_i) = true$, and the edge $\{a_i^j, \alpha_i^j\}$ otherwise.

Figure 3 An example of the reduction for $\Phi = (x_1 \vee x_2 \vee \bar{x_3}) \wedge (\bar{x_2} \vee x_3 \vee x_4)$. The dotted lines join two vertices that are identified. The thick red lines indicate the handicap vertices.



- $h^2_{i,j}$ contains the edge $\{b^j_i, \beta^j_i\}$ if $\phi(x_i) = true$, and the edge $\{z^j_i, b^j_i\}$ otherwise.
- $h^3_{i,j}$ contains the edge $\{z^j_i, a^j_i\}$ if $\phi(x_i) = true$, and the edge $\{z^j_i, b^j_i\}$ otherwise.
- $h^4_{i,j}$ contains the edge $\{\zeta^j_i, b^j_i\}$ if $\phi(x_i) = true$, and the edge $\{z^j_i, b^j_i\}$ otherwise.
- $h^5_{i,j}$ contains the edge $\{\zeta^j_i, b^j_i\}$ if $\phi(x_i) = true$, and the edge $\{\zeta^j_i, a^{j+1}_i\}$ otherwise.
- $h^6_{i,j}$ contains the edge $\{z^{j+1}_i, a^{j+1}_i\}$ if $\phi(x_i) = true$, and the edge $\{\zeta^j_i, a^{j+1}_i\}$ otherwise.

Moreover the vertices $z^1_i, \ldots, z^{p_i}_i$, and $\zeta^1_i, \ldots, \zeta^{p_i}_i$ have degree 1 in G_1 and the vertices α^j_i have degree 0 in G_1 if and only if $\phi(x_i) = true$ and the vertices β^j_i have degree 0 in G_1 if and only if $\phi(x_i) = false$.

We now select a set of edges in the clause gadgets for every clause $C_j = \ell_1 \vee \ell_2 \vee \ell_3$. Since ϕ satisfies Φ, there exists $p = p_j \in \{1, 2, 3\}$ such that $\phi(\ell_p) = true$. Select the edges $\{c^p_j, c^{p+2}_j\}$, $\{c^p_j, c^{p+3}_j\}$ and $\{c^{p+1}_j, c^{p+4}_j\}$. Let E_2 be the set of selected edges in clause gadgets, and $G_2 = (V(H), E_2)$. Clearly, G_2 O_3-overlays all hyperedges of the clause gadgets and the vertices c^4_j and c^5_j, $j \in [m]$, have degree 1 in G_2. Moreover, for all $j \in [m]$ the vertex $c^{p_j}_j$ is identified with the vertex α^j_i or β^j_i, which has degree 0 in G_1 since $\phi(\ell_p) = true$. Thus $G_1 \cup G_2$ has maximum degree 2, which proves that it is an $(O_3, H, 2)$-graph.

Conversely, assume that there is an $(O_3, H, 2)$-graph G such that all handicap vertices have degree at most 1.

For $i \in [n]$, let $T_i = \bigcup_{j=1}^{p_i} \{z^j_i a^j_i, b^j_i \beta^j_i, b^j_i \zeta^j_i\}$ and $F_i = \bigcup_{j=1}^{p_i} \{a^j_i \alpha^j_i, z^j_i b^j_i, \zeta^j_i a^{j+1}_i\}$.

Claim 7. *Let $i \in [n]$. We have the following:*

1. $z^j_i \zeta^j_i \notin E(G)$ *and* $\zeta^j_i z^{j+1}_i \notin E(G)$, $j \in [m]$.
2. *Either* $T_i \subseteq E(G)$ *or* $F_i \subseteq E(G)$.

Subproof. 1. Assume for a contradiction that $z^j_i \zeta^j_i \in E(G)$. Since z^j_i and ζ^j_i are handicap vertices, the hyperedges $h^2_{i,j}$,$h^3_{i,j}$, $h^5_{i,j}$ must be O_3-overlaid by the edges $b^j_i \beta^j_i$, $a^j_i b^j_i$, $b^j_i a^{j+1}_i$, respectively, contradicting the fact that b^j_i has degree 2 in G.
Similarly, if $\zeta^j_i z^{j+1}_i \in E(G)$, then the hyperedges $h^5_{i,j}$,$h^1_{i,j+1}$, $h^3_{i,j+1}$ must be O_3-overlaid by the edges $b^j_i a^{j+1}_i$, $a^{j+1}_i \alpha^j_i$, $a^{j+1}_i b^{j+1}_i$, respectively, a contradiction.

2. By 1. the hyperedge $h^4_{i,j}$ is O_3-overlaid by $z^j_i b^j_i$ or $\zeta^j_i b^j_i$.

In the first case $z_i^j b_i^j \in E(G)$, since z_i^j is a handicap vertex, the hyperedge $h_{i,j-1}^6$ must be O_3-overlaid by $\zeta_i^{j-1} a_i^j$, and so hyperedge $h_{i,j-1}^3$ is O_3-overlaid by $z_i^{j-1} b_i^{j-1}$. And so on by induction, we get that $\bigcup_{j=1}^m \{\zeta_i^j a_i^{j+1}, z_i^j b_i^j\} \subseteq E(G)$. Since the z_i^j are handicap vertices, the hyperedge $h_{i,j}^1$ is O_3-overlaid by $a_i^j \alpha_i^j$. Hence $F_i \subseteq E(G)$.

In the second case $\zeta_i^j b_i^j \in E(G)$, a symmetrical argument shows $T_i \subseteq E(G)$. $\Diamond$

We define the truth assignment ϕ by $\phi(x_i) = true$ if $T_i \subseteq E(G)$ and $\phi(x_i) = false$ if $F_i \subseteq E(G)$. This is well-defined by Claim 7 and the fact that $E(G)$ cannot contain both T_i and F_i because it has maximum degree 2.

Now, for any $j \in [m]$, there exists $p \in \{1, 2, 3\}$ such that c_j^p has degree 2 in $G[CG_j]$ (If not, then there are at most 2 edges. Hence there is a vertex c_j^p of degree 0, and thus there is a hyperedge containing c_j^p which is not overlaid). Therefore, it has degree 0 in the variable gadget it belongs to. By our definition of ϕ, we get that $\phi(\ell_p) = true$. Hence ϕ satisfies Φ.

C Proof of Lemma 8

Let k and p be two positive integers with $p \geq 2$, let $H_{k,p}$ be the p-uniform hypergraph on $(k+1)(p-1)$ vertices whose hyperedge set contains all possible p-subsets of vertices.

We first need the following claim in order to prove Lemma 8.

Claim 8. *Every $(O_p, H_{k,p}, k)$-graph is k-regular. Moreover, there exists an $(O_p, H_{k,p}, k)$-graph.*

Subproof. Let G be an $(O_p, H_{k,p}, k)$-graph. Assume for a contradiction that it has a vertex v of degree at most $k-1$. Then, observe that $|V(G) \setminus N[v]| \geq (k+1)(p-2)+1$. Now, by removing successively the closed neighbourhood of a vertex of the graph (chosen arbitrarily) until the graph is empty, we can construct, together with v, an independent set S of G of size p. This is impossible, since S is a hyperedge of $H_{k,p}$ which is not O_p-overlaid by G.
Finally, observe that the disjoint union of $p-1$ cliques of size $k+1$ is a k-regular graph which O_p-overlays $H_{k,p}$. $\Diamond$

Proof. First, we show the $\mathcal{NP}$-completeness of $(\Delta \leq k)$-O_{p+1}-Overlay by providing a polynomial time many one reduction from $(\Delta \leq k)$-O_p-Overlay. Let H be a p-uniform hypergraph. We construct a $(p+1)$-uniform hypergraph H^* from a copy of H and adding, for every hyperedge S, a new vertex v_s. We also add, for every hyperedge S, a set of $(k+1)p-1$ new vertices N_S, and form, together with v_S, a hypergraph H_S isomorphic to $H_{k,p+1}$.

Let G be an (O_p, H, k)-graph. Construct G^* from G by adding, for every hyperedge S of H, a disjoint union of p cliques of size $k+1$ on the vertices $N_S \cup \{v_S\}$. Clearly, G^* is an (O_{p+1}, H^*, k)-graph.

Conversely, if G^* is an (O_{p+1}, H^*, k)-graph, then, by Claim 8, v_S must have degree k in N_S for every hyperedge S of H. Thus, the restriction of G^* to the vertices of H is an (O_p, H, k)-graph.

Now, we reduce $(\Delta \leq k)$-O_p-Overlay to $(\Delta \leq k+1)$-O_p-Overlay.

Let H be a p-uniform hypergraph. Let H' be the p-uniform hypergraph obtained from H by doing the following for each vertex $v \in V(H)$:

- add the vertices $x^v, y_1^v, \ldots, y_{p-2}^v$ and the hyperedge $S_v = \{v, x^v, y_1^v, \ldots, y_{p-2}^v\}$.
- for any $i = \{1, \ldots, p-2\}$, add a copy H_i^v of $H_{k+1,p}$ whose vertices are y_i^v and $(p-1)(k+2)-1$ new vertices.

Assume that G' is an $(O_p, H', k+1)$-graph. By Claim 8, the graph induced by G' on $V(H_i^v)$ is $(k+1)$-regular, for all $v \in V(H)$, $1 \leq i \leq p-2$. Therefore each y_i^v has $k+1$-neighbours outside the hyperedge S_v. Hence G' must contain the edge vx^v because it O_p-overlays S_v. Thus the subgraph $G = G'[V(H)]$ has maximum degree at most $\Delta(G') - 1 \leq k$ and it O_p-overlays H. In other words, G is an (O_p, H, k)-graph.

Conversely, assume G is an (O_p, H, k)-graph. Let G' be the graph with vertex set $V(H')$ and whose edge set is the union of $E(G)$, $\{vx^v \mid v \in V(H)\}$, and the edge set of an $(O_p, H_v^i, k+1)$-graph (which exists by Claim 8), for all $v \in V(H), 1 \leq i \leq p-2$. One easily checks that G' is $(O_p, H', k+1)$-graph.

Therefore, there is an (O_p, H, k)-graph if and only if there is an $(O_p, H', k+1)$-graph. □

D Proof of Lemma 10

We shall reduce MAX $(\Delta \leq k)$-$(F + E_q)$-OVERLAY to MAX $(\Delta \leq k)$-$(F + E_{q+1})$-OVERLAY. Set $p = |V(F)|$ and $r = \lfloor k/\delta(F) \rfloor$. Note that $r \geq 1$, for otherwise MAX $(\Delta \leq k)$-$(F + E_q)$-OVERLAY would trivially be polynomial-time solvable.

Let H' be a $(p+q)$-uniform hypergraph. Let H be the $(p+q+1)$-uniform hypergraph obtained from H' as follows. Its vertex set is partitioned into $V(H')$, $P_1, \dots, P_r$, K, and $\{z\}$ with $|P_i| = p-1$ for all $1 \leq i \leq r$ and $|K| = (m+1)(k+1) + kp$ with $m = \max\{|E(H')|, q+1\}$. The hyperedge set of H contains $Q \cup P_i \cup \{z\}$ for all $(q+1)$-subsets Q of K and all $1 \leq i \leq r$, and $S \cup z$ for all $S \in E(H')$.

The reduction can be performed in polynomial time, because for each hyperedge in H' we construct a polynomial number of hyperedges in H (recall that q is fixed), and each of them is straightforward to obtain. We shall now prove that $\text{over}_{F+E_{q+1}}(H,k) = \text{over}_{F+E_q}(H',k) + r\binom{|K|}{q+1}$.

Let G_0' be a graph with $V(G_0') = V(H')$ and $\Delta(G_0') \leq k$ that $(F+E_q)$-overlays $\text{over}_{F+E_q}(H',k)$ hyperedges of H'. For $1 \leq i \leq r$, let F_i be a copy of F on $P_i \cup \{z\}$ in which z has degree $\delta(F)$. Let G_0 be the graph defined by $V(G_0) = V(H)$ and $E(G_0') \cup \bigcup_{i=1}^r E(F_i)$. Clearly, $\Delta(G_0) \leq k$ and G_0 F-overlays all hyperedges of H except the $S \cup z$ for all $S \in E(H')$. So $\text{over}_{F+E_{q+1}}(H,k) \geq \text{over}_{F+E_q}(H',k) + r\binom{|K|}{q+1}$.

Conversely, let G be a graph with $V(G) = V(H)$ and $\Delta(G) \leq k$ that maximizes the number of $(F+E_{q+1})$-overlaid hyperedges.

Claim 9. *In G, vertex z has at least $\delta(F)$ neighbours in each P_i, $1 \leq i \leq r$.*

Subproof. Suppose for a contradiction that z has less than $\delta(F)$ neighbours in P_i. Since $\Delta(G) \leq k$, at most kp vertices of K are incident to an edge with an endvertex in $P_i \cup \{z\}$. Thus there is a subset R of K of $(m+1)(k+1)$ vertices non-adjacent to all vertices of $P_i \cup \{z\}$. Now $G[R]$ has maximum degree k and is thus $(k+1)$-colourable. Hence $G[R]$ has an independent set S of size $m+1$. Observe that for each $(q+1)$-subset Q of S, the subgraph $G[Q \cup P_i \cup \{z\}]$ has $q+2$ vertices of degree less than $\delta(F)$ (namely the ones of $Q \cup \{z\}$), and so the hyperedge $Q \cup P_i \cup \{z\}$ is not $(F+E_{q+1})$-overlaid. Thus there are at least $m+1 \geq |E(H')|+1$ hyperedges not $(F+E_{q+1})$-overlaid by G. This contradicts the maximality of G as G_0 overlays more hyperedges of H. ◊

Now z has degree at least $r\delta(F)$ in $\bigcup_{i=1}^r P_i$ and so less than $\delta(F)$ in each hyperedge $S \cup z$ for $S \in E(H')$. Hence if G $(F+E_{q+1})$-overlays $S \cup z$, it must also $(F+E_q)$-overlay S. Thus $\text{over}_{F+E_{q+1}}(H,k) \leq \text{over}_{F+E_q}(H',k) + r\binom{|K|}{q+1}$.

E Proof of Claim 1

Let G be a graph with $\Delta(G) \leq k$ that F-overlays $\mathrm{over}_F(H,k)$ hyperedges of H. Observe first that since H is a neat hypergraph, we may assume that, for every hyperedge S of H, $G[S] = F$ if S is F-overlaid and $G[S]$ has no edge otherwise.

Therefore if an x-leaf L is F-overlaid, then replacing $G[L]$ by a copy of F in which x has degree $\delta(F)$, we obtain another graph with maximum degree at most k that F-overlays $\mathrm{over}_F(H,k)$ hyperedges of H. Henceforth we may assume that if an x-leaf is F-overlaid then x is incident to $\delta(F)$ edges included in it.

Now assume that the z_e-leaf L, for $e \in E(\Gamma)$, is not F-overlaid by G. We have $d_G(z_e) > k - \delta(F)$ for otherwise adding on L the edges of a copy of F in which z_e has degree $\delta(F)$, we obtain a graph F-overlaying more hyperedges than G. But since there are γ z_e-leaves and z_e is incident to $\delta(F)$ edges in each of them, at least one of the hyperedges S^e_u and S^e_v is F-overlaid. Now replacing the edges of G in this hyperedge by those of a copy of F in L, in which z_e has degree $\delta(F)$, we obtain another graph with maximum degree at most k that F-overlays $\mathrm{over}_F(H,k)$ hyperedges of H. Henceforth we may assume that every z_e-leaf is F-overlaid.

A similar reasoning shows that every a^v_i-leaf is F-overlaid, for all $v \in V(\Gamma)$, $i \in \{1,2,3\}$. This proves (a).

Now by (a) each vertex z_e is incident to $\gamma\delta(F) = k - \beta > k - 2\delta(F)$ edges in the z_e-leaves. Thus z_e is incident to at most β edges in $S^e_u \cup S^e_v$. Since $\beta \leq 2\delta(F) - 1$, at most one of S^e_u and S^e_v is F-overlaid. Moreover, since $\beta < d(a_2)$, if S^e_u is F-overlaid, then z_e must be the vertex corresponding to a_1 in $G[S^e_u]$.
Assume now that none of S^e_u, S^e_v is F-overlaid, and let $e = e_i(u)$, a^u_i has γ_i a^u_i-leaves, so a^u_i is incident to $\gamma_i\delta(F) \leq k - d(a_i)$ edges in these leaves, thus $d_G(a^u_i) \leq k - d(a_2)$. Thus, replacing the edges of G in S^e_u by those of a copy of F, in which z_e has degree $\delta(F)$ and a^u_i has degree $d(a_2)$, we obtain another graph with maximum degree at most k that F-overlays $\mathrm{over}_F(H,k)$ hyperedges of H. Henceforth, we may assume that at least one of S^e_u and S^e_v is F-overlaid. This completes the proof of (b).

Consider now a vertex $v \in V(\Gamma)$ such that S_v is F-overlaid. By (a) a^v_1 is incident to $\gamma_1\delta(F)$ edges in the a^v_1-leaves and at least $\delta(F) \in S_v$. Therefore $S^{e_1(v)}_v$ cannot be F-overlaid for otherwise, by (b), a^v_1 would have degree at least $\gamma_1\delta(F) + \delta(F) + d(a_2) > k$. Similarly, $S^{e_2(v)}_v$ cannot be F-overlaid.
Assume now for a contradiction that $S^{e_3(v)}_v$ is F-overlaid. By (a) and (b), a^v_3 is incident to $\gamma_3\delta(F)$ edges in the a^v_3-leaves and at least $d(a_2) \in S^{e_3(v)}_v$. Since $\gamma_3\delta(F) + d(a_2) + d(a_3) > k$, a^v_3 is incident to less than $d(a_3)$ edges in S_v. Thus, there is a vertex $a \in \{a^v_1, a^v_2\}$ which is incident to at least $d(a_3)$ edges in S_v and thus at least $\gamma_1\delta(F) + d(a_3)$ edges in G. But a^v_3 is adjacent to at least $\delta(F)$ edges in S_v so at least $\gamma_3\delta(F) + d(a_2) + \delta(F)$ edges in total. Thus $\gamma_3\delta(F) + d(a_2) + \delta(F) \leq k$, so $\gamma_3 \leq \gamma_1 - 1$. Thus a is adjacent to at least $(\gamma_3 + 1)\delta(F) + d(a_3) > k$ edges, a contradiction.
To summarize, if S_v is F-overlaid, then none of $S^{e_1(v)}_v$, $S^{e_2(v)}_v$, $S^{e_3(v)}_v$ is F-overlaid. Together with (b), this implies (c).

In the second case $d(a_2) \leq \beta$, as the definition of the γ_i is different to that in the first case, thus there are differences in the construction in the proof of Claim 1 which are briefly shown as follows:

- In Claim 1(a), it's the same. Precisely, x-leaf L is overlaid and we replace $G[L]$ by a copy of F in which x has degree $\delta(F)$ for any $x = a^v_i$ or $x = z_e$.
- In Claim 1(b), there is exactly one of S^e_u and S^e_v is F-overlaid. Let $e = e^u_i$, then in $G[S^e_u]$ which is F-overlaid, for $i = 1$, $d(z_e) = a_2$ and $d(a^u_1) = d(a_1)$; and for $i = \{2,3\}$ $d(z_e) = d(a_1)$ and $d(a^u_i) = d(a_2)$.

- In Claim 1(c), consider a vertex $v \in V(\Gamma)$ such that S_v is F-overlaid. By (a) a_1^v is incident to $\gamma_1\delta(F)$ edges in the a_1^v-leaves and at least $\delta(F) \in S_v$. Hence, If $S_v^{e_1(v)}$ is F-overlaid, then $d(a_i^v) = \gamma_i\delta(F) + \delta(F) + \delta(F) > k$ by (b). If $S_v^{e_2(v)}$ is F-overlaid, then $d(a_2^v) = \delta$ in $G[S_v]$. We obtain that $d(a_2^v) = \gamma_2\delta + d(a_2) + \delta \leq k$ which implies that $\gamma_2 \leq \gamma_2 - 1$, it is a contradiction. If $S_v^{e_3(v)}$ is F-overlaid, similarly to the previous proof, we obtain $\gamma_3\delta(F) + d(a_2) + \delta(F) \leq k$, so $\gamma_3 \leq \gamma_2 - 1 \leq \gamma_1 - 1$ since $\gamma_2 \leq \gamma_1$.

RESEARCH CENTRE
SOPHIA ANTIPOLIS – MÉDITERRANÉE

2004 route des Lucioles - BP 93
06902 Sophia Antipolis Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399