Covering the edges of a graph by a prescribed tree with minimum overlap

Noga Alon * Yair Caro † Raphael Yuster ‡

Abstract

Let $H = (V_H, E_H)$ be a graph, and let k be a positive integer. A graph $G = (V_G, E_G)$ is *H-coverable with overlap k* if there is a covering of the edges of G by copies of H such that no edge of G is covered more than k times. Denote by *overlap*(H, G) the minimum k for which G is H-coverable with overlap k. The *redundancy* of a covering that uses t copies of H is $(t|E_H| - |E_G|)/|E_G|$. Our main result is the following: If H is a tree on h vertices and G is a graph with minimum degree $\delta(G) \geq (2h)^{10} + C$, where C is an absolute constant, then *overlap*$(H, G) \leq 2$. Furthermore, one can find such a covering with overlap 2 and redundancy at most $1.5/\delta(G)^{0.1}$. This result is tight in the sense that for every tree H on $h \geq 4$ vertices and for every function f, the problem of deciding if a graph with $\delta(G) \geq f(h)$ has *overlap*$(H, G) = 1$ is NP-Complete.

1 Introduction

All graphs considered here are finite, undirected and simple, unless otherwise noted. For the standard graph-theoretic notations the reader is referred to [2]. Let H be a graph, and let k be a positive integer. A graph $G = (V, E)$ is *H-coverable with overlap k* if there is a set $L = \{G_1, \ldots G_t\}$ of subgraphs of G such that each G_i is isomorphic to H and every edge $e \in E$ appears in at least one member of L but in no more than k members of L. Denote by *overlap*(H, G) the minimum k for which G is H-coverable with overlap k. Clearly, *overlap*$(H, G) = 1$ if and only if there is a *decomposition* of G into H. Also, if there is an edge of G which appears in no subgraph of G which is isomorphic to H, we put *overlap*$(H, G) = \infty$. Clearly, if *overlap*(H, G) is finite then *overlap*$(H, G) \leq |E(G)| - |E(H)| + 1$. This upper bound is realized by many pairs of graphs. For

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel. Research supported in part by the Fund for Basic Research administered by the Israel Academy of Sciences.

†School of Education, Department of Mathematics, University of Haifa-ORANIM, Tivon 36006, Israel.

‡School of Education, Department of Mathematics, University of Haifa-ORANIM, Tivon 36006, Israel.

example, let H_n be the star on n vertices to which an edge has been added between two leaves. In this case we have $overlap(H_4, H_n) = n - 3$.

It has been shown by Dor and Tarsi [3] that for every fixed graph H having a connected component with at least 3 edges, the problem of deciding for a given input graph G on n vertices whether $overlap(H, G) = 1$ is NP-Complete. Thus, even if H is a tree on 4 vertices, this problem is difficult. If the minimum degree of G is very large, that is, $\delta(G) \geq (1 - \epsilon(H))n$, this decomposition problem can be solved in polynomial time, by the results of Wilson and Gustavsson [6, 5]. On the other hand, we show in Theorem 1.2 that this problem remains NP-Complete for every tree H on $h \geq 4$ or more vertices, even if $\delta(G) \geq n^{0.499}$. Hence, there is no function $f(H)$ for which we can recognize efficiently the class of graphs G having $\delta(G) \geq f(H)$ and which have $overlap(H, G) = 1$, unless P=NP. The main result in this paper is to show that such a function does exist if we allow some edges to be covered twice. In fact, this function is only a moderate polynomial function of h, and only a small fraction of the edges are covered twice. This result is summarized in the following theorem.

Theorem 1.1 *Let H be a tree on h vertices, and let $G = (V, E)$ be a graph with $\delta(G) > (2h)^{10} + 114^{10}$, then $overlap(H, G) \leq 2$. Furthermore, there exists a covering with overlap 2, where at most $1.5|E|/\delta(G)^{0.1}$ edges are covered twice.*

The overlap obtained in this result is clearly best possible in a combinatorial sense, since an exact decomposition requires additional divisibility constraints which cannot be expressed in terms of the minimal degree of G. It is also best possible in an algorithmic sense (unless P=NP), even if we significantly increase the minimum degree requirement:

Theorem 1.2 *Let $\alpha < 0.5$ be fixed, and let H be any graph having a connected component with three or more edges, and having a vertex of degree one. Deciding whether a graph G with $\delta(G) > n^\alpha$ has $overlap(H, G) = 1$ is NP-Complete.*

Note that Theorem 1.2 applies to any tree H with 4 or more vertices.

A minimum degree requirement in Theorem 1.1 is mandatory. For any tree H on $h \geq 4$ vertices, let G be the graph obtained by joining two vertex-disjoint cliques of order $h - 1$ with one edge. Clearly, $\delta(G) = h - 2$, every edge of G is on some copy of H (unless $H = K_{1,h-1}$ in which case $overlap(H, G) = \infty$), and thus $overlap(H, G)$ is finite, but every copy of H in G passes through the unique bridge. Thus, the overlap is at least $\lceil ((h-1)(h-2)+1)/(h-1) \rceil \geq h - 1 \geq 3$. The minimum degree bound of $O(h^{10})$ in Theorem 1.1 is not best possible. With some more effort we can reduce the power to a single digit number, but this is still far from the obvious lower bound of $h - 1$ described above. Furthermore, Theorem 1.1 also shows that only a small fraction of the edges are covered twice. In fact, if $\delta(G) = w(n)$ tends to infinity arbitrarily slow, then only $o(E)$ edges are covered twice. For some trees, however, we do know that a minimal degree of $h-1$ guarantees an overlap of 2:

Theorem 1.3 *Let $k > 1$ be an integer. Let G be a graph such that every edge of G has an endpoint whose degree is at least k. Then $overlap(K_{1,k}, G) \leq 2$. Consequently, if $\delta(G) \geq k$ then $overlap(K_{1,k}, G) \leq 2$.*

Note that this simply means that if a graph G is $K_{1,k}$-coverable with any overlap, then it is also $K_{1,k}$-coverable with overlap 2.

Theorem 1.4 *If $\delta(G) \geq 3$ then $overlap(P_4, G) \leq 2$, where P_4 is the path with four vertices.*

Theorem 1.3 implies that given a graph G, deciding whether $overlap(K_{1,k}, G) \leq 2$ can be done in polynomial time, for every k. This is quite different from the corresponding decomposition problem for stars. The result of Dor and Tarsi (as well as the previously known results on this question) imply that for $k \geq 3$, deciding whether $overlap(K_{1,k}, G) = 1$ is NP-Complete. However, we can still show the following:

Theorem 1.5 *There are infinitely many (fixed) trees H for which, given a graph G, deciding whether $overlap(H, G) \leq 2$ is NP-Complete.*

The rest of this paper is organized as follows. Section 2 contains the necessary lemmas needed for the proof of Theorem 1.1, and the proof itself. In Section 3 we prove the exact results for the stars $K_{1,k}$ and the path P_4, namely Theorems 1.3 and 1.4. In Section 4 we prove the NP-Completeness results stated in Theorems 1.2 and 1.5. Concluding remarks and open problems appear in Section 5.

2 Covering graphs by trees with overlap 2

The graph G in Theorem 1.1 is assumed to have a minimum degree bound, but may otherwise be highly irregular. Our proof methods require, however, that the degrees of all vertices are bounded. We can overcome this problem using the fact that any graph with a large-enough minimum degree is homeomorphic in the following strong sense to an almost-regular graph with a quadratically smaller minimum degree.

Lemma 2.1 *Let $G = (V, E)$ be a graph, $\delta(G) \geq d(d-1)$. There exists a graph $G' = (V', E')$ and a function $f : V' \to V$ such that the following hold:*

1. *For each $(u, v) \in E$ there exists exactly one edge $(x, y) \in E'$ with $f(x) = u$ and $f(y) = v$.*
2. *$(x, y) \in E'$ implies $(f(x), f(y)) \in E$.*

3. *If $x, y \in V'$, $x \neq y$ and $f(x) = f(y)$ then x and y are at distance at least 3 (in G').*

4. *The degree of every vertex of G' is either d or $d+1$.*

Proof: Let $V = \{1, \ldots, n\}$. Let d_i denote the degree of i in G. Since $d_i \geq d(d-1)$, we may partition $N(i)$, the neighbor set of i, into $s_i = \lfloor d_i/d \rfloor$ disjoint subsets $N(i,1), \ldots, N(i, s_i)$ such that $d+1 \geq |N(i,j)| \geq d$. We define the graph G' as follows. Let $V_i = \{v_{i,1}, \ldots, v_{i,s_i}\}$, $V' = \cup_{i=1}^{n} V_i$. The function f is defined as $f(v_{i,j}) = i$, $j = 1, \ldots, s_i$. In order to define E' we do the following. For each $(i,j) \in E$, we have that $j \in N(i,r)$ for some r and $i \in N(j,t)$ for some t. We therefore make $(v_{i,r}, v_{j,t})$ an edge of G'. It is easy to check that the four conditions in the lemma are satisfied by G'. $\square$

A *strong coloring* f of a multigraph is defined as a proper vertex-coloring, where two vertices of the same color do not share a common neighbor. Note that the function f in Lemma 2.1 is a strong coloring of the vertices of G'. A *simple* subgraph H of a multigraph G' with a strong-coloring f is called *colorful* with respect to f if all its vertices have different colors.

Corollary 2.2 *If G and G' are graphs as in Lemma 2.1, and G' is H-coverable with overlap k such that every copy of H in the covering is colorful with respect to the coloring function f of Lemma 2.1, then overlap$(H, G) \leq k$.* $\square$

Our proof of Theorem 1.1 is essentially divided into three stages. Given the graph G we initially create the graph G' as in Lemma 2.1. In the second stage we embed in G' a set of edge-disjoint colorful copies of the tree H, such that for every vertex of G', only a small fraction of the edges adjacent to it are non-covered. In the third stage, we embed in G' a set of edge-disjoint colorful copies of H, such that every edge that was not covered in the second stage is now covered. Note that every edge of G' is covered at most twice (at most once in stage 2 and at most once in stage 3), and thus Theorem 1.1 follows from Corollary 2.2. Lemmas 2.3 and 2.4 will provide us with stages 2 and 3 respectively. However, before we state them, we need some preparations.

Let H be a tree with $h \geq 2$ vertices. Every vertex $v \in H$ defines a unique rooted-orientation of H, denoted by $H(v)$, which results from a breadth-first search (BFS) beginning at v. The vertex v is called the *root* of such an orientation, and every vertex u of $H(v)$, except v, has a unique *parent* which is the source of the unique incoming edge into u. Given an orientation $H(v)$, let $(e_1, \ldots, e_{h-1})$ denote the *edge-addition sequence* of the BFS. Let $H^i(v)$, for $i = 1, \ldots, h-1$ denote the directed subtree of $H(v)$ on the edge-set $(e_1, \ldots, e_i)$. Note that $H^i(v)$ is obtained from $H^{i-1}(v)$ by adding a new vertex (a leaf) and directing an edge from its parent to it. We may assume that the chosen root v is a leaf of H. With this assumption, we may define, for $i = 2, \ldots, h-1$ the *parent* of the edge e_i of $H(v)$ to be the unique incoming edge of the source of e_i. The edge e_1 does not have a parent. Note that if e_j is the parent of e_i then $j < i$.

Let G' be a graph. A well-known consequence of Euler's Theorem (cf., e.g., [2]) is that the edges of G' can be oriented so that for every vertex v, $|d^+(v) - d^-(v)| \le 1$, where $d^+(v)$ and $d^-(v)$ denote the outdegree and indegree (respectively) of v in the oriented G'. We call such an orientation *balanced*. We use the notations $\Delta^+(G'), \Delta^-(G'), \delta^+(G'), \delta^-(G')$ to denote the maximum-outdegree, maximum-indegree, minimum-outdegree and minimum-indegree (resp.) of G'.

Lemma 2.3 *Let H be a tree on $h \ge 30$ vertices. Let $G' = (V', E')$ be a graph with a strong coloring f. Suppose that $32h^5 \ge d \ge 31h^5$ and $d \le \delta(G') \le \Delta(G') \le d+1$. Furthermore, assume that $2(h-1)x = d+2$ where x is a perfect square. Then there is a set L of edge-disjoint colorful subgraphs of G', each isomorphic to H, such that every vertex of G' has at most $2(h-1)\sqrt{x}$ edges adjacent to it among those not covered by members of L.*

Proof: We begin by coloring the edges of G' with the colors $\{1, \ldots, d+2\}$ such that no two adjacent edges receive the same color. This can be done by Vising's Theorem (cf. [2]). Since $h-1$ divides $d+2$ we can partition the colors into $h-1$ subsets $C_1, \ldots, C_{h-1}$ each consisting of $2x$ colors. Let E_i be the set of edges colored with a color from C_i, and put $G_i = (V', E_i)$ for $i = 1, \ldots, h-1$. Note that $\delta(G_i) \ge 2x-2$ and $\Delta(G_i) \le 2x$. We now orient the edges of each E_i such that the orientations are balanced. Thus, in these orientations, $\Delta^-(G_i), \Delta^+(G_i) \le x$ and $\delta^-(G_i), \delta^+(G_i) \ge x-1$. Consider the oriented graph G_i. By adding a perfect (directed) matching F_i from the vertices with out-degree $x-1$ to the vertices with in-degree $x-1$ (these sets have equal sizes) we obtain a regular directed multigraph $G_i^* = (V', E_i \cup F_i)$ with in-degree and out-degree x. Note that some edges of F_i may be loops or parallel to some edge of E_i. Let $G^* = (V', E_1 \cup F_1 \cup \ldots \cup E_{h-1} \cup F_{h-1})$. Note that G^* is a directed multigraph with $|V'|x(h-1)$ edges. Also, the maximum degree of a vertex in G^*, considered as an undirected multigraph, is $d+2$.

Let $H(v)$ be a rooted orientation of H, where v is a leaf of H. Let $(e_1, \ldots, e_{h-1})$ be the edge-addition sequence of $H(v)$. For each vertex $w \in V'$ and for each $i = 2, \ldots, h-1$ we select a matching $\pi_{i,w}$ between its x incoming edges belonging to $E_j \cup F_j$ and its x outgoing edges belonging to $E_i \cup F_i$, where e_j is the parent of e_i in $H(v)$. Each matching is selected randomly, and uniformly among the $x!$ possible matchings. All matchings are independent.

We now construct a set L' of $|V'|x$ edge-disjoint subgraphs of G^*, each consisting of $h-1$ edges (hence, every edge appears in exactly one member of L'). The construction is done according to $H(v)$ and the matchings $\pi_{i,w}$ in the following inductive manner: We initially define the set L_1 to be the single-edge graphs which are the edges of $E_1 \cup F_1$. Note that L_1 has $|V'|x$ elements. We assume by induction that we have constructed L_{i-1}, which is a set of $|V'|x$ edge-disjoint subgraphs of G^*, each containing $i-1$ edges, one from each $E_k \cup F_k$, $k = 1, \ldots, i-1$. We show how to construct L_i. Let e_j be the parent of e_i in $H(v)$. Note that $1 \le j \le i-1$. Consider a copy in L_{i-1}. This copy contains exactly one (directed) edge (u, w) of $E_j \cup F_j$. We extend the copy to a copy of L_i by adding to it the edge $\pi_{i,w}((u, w))$. Clearly, this edge belongs to $E_i \cup F_i$, the new copy has i

edges, and all the copies of L_i remain edge-disjoint. Finally note that by putting $L' = L_{h-1}$ we obtain the desired construction. Note that our construction implies that each colorful member of L' is, in fact, isomorphic to $H(v)$. In particular, every colorful member of L' which contains no edge belonging to $F_1 \cup \ldots \cup F_{h-1}$ uniquely defines a colorful copy of H in G'. We therefore call a member of L' *good* if it is colorful and contains no edge from $F_1 \cup \ldots \cup F_{h-1}$, otherwise it is called *bad*. Let $L \subset L'$ be the set of good copies. Our aim is to show that, with positive probability, L satisfies the statement of the lemma.

For $e \in E_1 \cup \ldots \cup E_{h-1}$ let $L'(e)$ denote the member of L' containing e, and let $L'(e, i)$ be the edge of $L'(e)$ belonging to $E_i \cup F_i$. An edge $L'(e, i) = (u, w)$ is called *bad* if it belongs to F_i or if w's color already appears in $L'(e)$, that is $L'(e, j)$ has an endpoint colored by the same color as w where $j < i$. Let $A_{e,i}$ be the event that $L'(e, i)$ is bad and let A_e be the event that $L'(e)$ is bad. It is not difficult to see that

$$\text{Prob}[A_e] \leq \sum_{i=1}^{h-1} \text{Prob}[A_{e,i}] \leq \frac{1}{x} + \frac{1}{x} + \frac{2}{x} + \ldots + \frac{h-2}{x} \leq \frac{h^2}{2x}.$$

Let $U = \{(u_1, w), \ldots, (u_k, w)\}$ be a k-subset of the edges of E_i (for some i) that enter a vertex w. Assume that $k \leq x/2$ and let A_U be the event that $L'((u_j, w))$ is bad for all $j = 1, \ldots, k$. Clearly,

$$\text{Prob}[A_U] = \prod_{j=1}^{k} \text{Prob}[A_{(u_j,w)} | A_{(u_1,w)}, \ldots, A_{(u_{j-1},w)}].$$

On the other hand,

$$\text{Prob}[A_{(u_j,w)} | A_{(u_1,w)}, \ldots, A_{(u_{j-1},w)}] \leq \sum_{t=1}^{h-1} \text{Prob}[A_{(u_j,w),t} | A_{(u_1,w)}, \ldots, A_{(u_{j-1},w)}] \leq$$

$$\frac{1}{x-(j-1)} + \frac{1}{x-(j-1)} + \frac{2}{x-(j-1)} + \ldots + \frac{h-2}{x-(j-1)} \leq \frac{h^2}{2(x-(j-1))} \leq \frac{h^2}{2(x-(k-1))} \leq \frac{h^2}{x}.$$

Consequently,

$$\text{Prob}[A_U] \leq (\frac{h^2}{x})^k.$$

Note that exactly the same computation holds if we replace U by a set of k edges emanating from w. Let $k \leq x/2$ be fixed (we shall choose its exact value later). For $w \in V'$ and $i = 1, \ldots, h-1$ let $B_{w,i}$ be the event that there exist k edges of E_i entering w which belong to bad copies, or that there exist k edges emanating from w which belong to bad copies. We have thus shown that

$$\text{Prob}[B_{w,i}] \leq 2\binom{x}{k}(h^2/x)^k.$$

The event $B_{w,i}$ is independent of the event $B_{u,j}$ if the distance between w and u in G^*, considered as an undirected multigraph, is at least twice the height of $H(v)$. This is true since a copy of $H(v)$

in G^* which contains u cannot share an edge with a copy of $H(v)$ in G^* which contains w. The height of $H(v)$ is at most $h-1$. The number of vertices v at distance at most $2h-3$ from w is therefore bounded by

$$(d+2)+(d+1)(d+2)+(d+1)^2(d+2)+\ldots+(d+1)^{2h-4}(d+2) \le (d+1)^{2h-3}(2h-3).$$

Hence, $B_{w,i}$ is independent of all other events $B_{u,j}$ but at most

$$(h-1)(2h-3)(d+1)^{2h-3}+(h-2).$$

Our aim is to show that with positive probability, none of the events $B_{w,i}$ hold. In other words, we need to show that

$$\text{Prob}[\cap_{w\in V'} \cap_{i=1}^{h-1} \overline{B}_{w,i}] > 0.$$

According to the Lovász Local Lemma (cf., e.g., [1]), it suffices to show that

$$e \cdot 2\binom{x}{k}(h^2/x)^k \cdot (h-1)((2h-3)(d+1)^{2h-3}+1) < 1 \tag{1}$$

holds. To see this, note that the following inequality holds:

$$e \cdot 2\binom{x}{k}(h^2/x)^k \cdot (h-1)((2h-3)(d+1)^{2h-3}+1) \le \tag{2}$$

$$\le 4eh^2(d+1)^{2h-3}\binom{x}{k}(h^2/x)^k < (d+1)^{2h}\binom{x}{k}(h^2/x)^k.$$

Choosing $k=\sqrt{x}$ and using the fact that $\binom{x}{\sqrt{x}} < (e\sqrt{x})^{\sqrt{x}}$ it follows from (2) that in order to prove (1) it suffices to show that

$$(\frac{h^2e}{\sqrt{x}})^{\sqrt{x}}(d+1)^{2h} < 1.$$

Recall that $x=(d+2)/(2h-2) > d/2h \ge 15.5h^4$. Hence,

$$(\frac{h^2e}{\sqrt{x}})^{\sqrt{x}}d^{2h} < (\frac{e}{3.9})^{3.9h^2}(32h^5+1)^{2h} < 1$$

where the rightmost inequality holds for $h \ge 30$. We have proved that with positive probability, none of the events $B_{w,i}$ hold. This means that there exists a set of permutations $\pi_{i,w}$ such that every vertex is adjacent to at most $2(h-1)\sqrt{x}$ bad edges. Thus L is a set of edge-disjoint colorful subgraphs of G', each one isomorphic to H, such that every vertex of G' has at most $2(h-1)\sqrt{x}$ adjacent edges which are not covered by members of L. □

Lemma 2.4 *Let H be a tree on $h \ge 2$ vertices. Let $G'=(V',E')$ be a graph with a strong coloring f. Let $G_1=(V',E_1)$ be a spanning subgraph of G' with $\Delta(G_1) \le 2s$. Furthermore, suppose that $d \ge sh^2+h^3$ and $d \le \delta(G') \le \Delta(G') \le d+1$. Then there are edge-disjoint colorful subgraphs of G', each one isomorphic to H, such that their edge-union contains the edges of G_1.*

Proof: Let $G^* = (V', E^*)$ where $E^* = E' \setminus E_1$. Clearly, $\Delta(G^*) \leq d+1$, and $\delta(G^*) \geq d-2s$. As in Lemma 2.3, we color the edges of E^* with the colors $\{1, \ldots, d+2\}$ such that no two adjacent edges receive the same color. We may partition the colors into $h-1$ disjoint sets, $C_2, \ldots, C_h$ where C_i contains exactly $2(is + \binom{i}{2})$ colors, for $i = 2, \ldots, h-1$. C_h contains the rest of the colors, if there are any. This can be done since

$$\sum_{i=2}^{h-1} 2(is + \binom{i}{2}) \leq sh^2 + h^3 < d+2.$$

Let E_i be the set of edges of E^* whose color belong to C_i, and $G_i = (V', E_i)$, $i = 2, \ldots, h$. Thus $E' = E_1 \cup \ldots \cup E_h$. Note that the property of our coloring and the degree bounds of G^* imply that $\delta(G_i) \geq 2(is + \binom{i}{2}) - (2s+2)$ and $\Delta(G_i) \leq 2(is + \binom{i}{2})$, for $i = 2, \ldots, h-1$. We now orient the edges of E_i for $i = 1, \ldots, h$ such that the orientations are balanced. Thus, in these orientations, $\Delta^-(G_1), \Delta^+(G_1) \leq s$, and for $i = 2, \ldots, h-1$ we have $\Delta^-(G_i), \Delta^+(G_i) \leq is + \binom{i}{2}$ and $\delta^-(G_i), \delta^+(G_i) \geq is + \binom{i}{2} - (s+1)$. (We claim nothing on the degrees of the oriented G_h. In fact, we will ignore the edges of E_h). Note that we have oriented every edge of G', and we may now consider it as a directed graph. Let $H(v)$ be a rooted orientation of H, where v is a leaf of H. Let $(e_1, \ldots, e_{h-1})$ be the edge-addition sequence of $H(v)$. We will create $|E_1|$ edge-disjoint colorful subgraphs of (the directed) G', each isomorphic to $H(v)$, such that the edge corresponding to e_i in each copy belongs to E_i for $i = 1, \ldots, h-1$. We do this in $h-1$ stages where after stage i we shall have $|E_1|$ edge-disjoint colorful subgraphs isomorphic to $H^i(v)$. For $i = 1$ we simply take every directed edge of E_1 as a subgraph, which is trivially isomorphic to $H^1(v)$. Note that we have already guaranteed that all the edges of G_1 are covered. All these subgraphs are colorful since the coloring f is proper. Suppose we have already constructed $|E_1|$ edge-disjoint colorful copies of $H^i(v)$, so that in each copy the edge playing the role of e_j is taken from E_j, $j = 1, \ldots, i$. We show how to extend these copies to edge-disjoint colorful copies of $H^{i+1}(v)$, only by using edges from E_{i+1}. Let e_j be the parent of e_{i+1} in $H(v)$. Note that $j \leq i$. Let $w \in G'$, and consider all the copies of $H^i(v)$ where w plays the role of the target of e_j (and thus should become the source of e_{i+1} after the extension). By our assumption, there is a one-to-one correspondence between these copies and some of the edges of E_j whose target is w (there may be other edges of E_j whose target is w that were not covered). Thus, the number of these copies is at most $js + \binom{j}{2}$ (note that this also holds if $j = 1$). Each such copy must be extended to a copy of $H^{i+1}(v)$ by an edge of E_{i+1} whose source is w. Thus, each copy must select an edge $(w, u) \in E_{i+1}$ such that all the selections are distinct, and such that u is not colored by any of the $i+1$ colors of the vertices of the copy of $H^i(v)$. In fact, for each copy we may only worry about $i-1$ forbidden colors, since u is already guaranteed not to have the color of w nor the color of the source of the edge playing the role of e_j in the copy (recall that the coloring is strong). This can be done if we can show that $\delta^+(G_{i+1}) \geq js + \binom{j}{2} + (i-1)$.

Indeed,

$$\delta^+(G_{i+1}) \geq (i+1)s + \binom{i+1}{2} - (s+1) = is + \binom{i}{2} + (i-1) \geq js + \binom{j}{2} + (i-1).$$

□

Proof of Theorem 1.1 We shall prove that if H is a tree with $h \geq 2$ vertices and G is a graph with $\delta(G) > (2h)^{10} + 114^{10}$, then $overlap(H,G) \leq 2$. Let h_0 be the maximal integer such that $\delta(G) > (2h_0)^{10}$ and $h-1$ divides $h_0 - 1$. Note that $h_0 \geq \max\{30, h\}$. It is very easy to construct a tree H_0 on h_0 vertices which has a decomposition into $(h_0-1)/(h-1)$ copies of H. Hence, it suffices to show that $overlap(H_0, G) \leq 2$. Let d be an integer satisfying $32h_0^5 \geq d \geq 31h_0^5$, such that $(d+2)/(2h_0-2) = x$ is a perfect square. Such a d certainly exists. Note that $\delta(G) \geq d(d-1)$, so we can construct the graph G' and the strong coloring function f, as guaranteed by Lemma 2.1. We can now apply Lemma 2.3 to G' and obtain a set L of edge-disjoint colorful subgraphs of G' which are isomorphic to H_0, where every vertex $w \in G'$ is adjacent to at most $2(h_0-1)\sqrt{x}$ non-covered edges. Let $s = (h_0-1)\sqrt{x}$, and let $G_1 = (V', E_1)$ be the spanning subgraph of G' where E_1 is the set of the non-covered edges. Note that $\Delta(G_1) \leq 2s$. Furthermore,

$$d \geq h_0^3\sqrt{x} \geq (h_0-1)\sqrt{x}h_0^2 + h_0^3 \geq sh_0^2 + h_0^3.$$

Hence, according to Lemma 2.4, there is a set M of edge-disjoint colorful subgraphs of G' which are isomorphic to H_0, such that every edge of E_1 is covered. Now $L \cup M$ is a covering of G' with colorful copies of H_0 such that every edge is covered at most twice. By Corollary 2.2, we have $overlap(H_0, G) \leq 2$. Lemma 2.4 and its proof imply that at most $|E_1|(h_0-2)$ edges are covered twice. Note that

$$|E_1|(h_0-2) \leq \frac{2(h_0-1)\sqrt{x}|E|}{d}(h_0-2) = \frac{2(h_0-1)(h_0-2)\sqrt{(d+2)/(2h_0-2)}}{d}|E| \leq$$

$$|E| \cdot 2\sqrt{h_0^3/d} \leq 0.36|E|/h_0 \leq 0.36|E|/(0.25\delta(G)^{0.1}) \leq 1.5|E|/\delta(G)^{0.1}.$$

□

3 Covering graphs by $K_{1,k}$ or P_4 with overlap 2

Proof of Theorem 1.3 Let $G = (V, E)$ be a graph such that if $(a,b) \in E$ then either $d(a) \geq k$ or $d(b) \geq k$, where $d(v)$ denotes the degree of v. We must find a set L of edge-disjoint subgraphs of G which are isomorphic to $K_{1,k}$ such that every edge of G appears in a member of L, but in no more than two members of L. Let $V' = \{v_1, \ldots, v_s\}$ be the set of vertices of G with degree at least k. We initially mark all edges of G as uncovered, and put $L = \emptyset$. We add elements to L by performing the following process for every $v_i \in V'$, where $i = 1, \ldots, s$. Let E_i be the uncovered edges adjacent to

v_i. We can create $\lfloor |E_i|/k \rfloor$ edge-disjoint copies of $K_{1,k}$ whose roots are v_i and whose edges belong to E_i. We add these copies to L, and mark the $k\lfloor |E_i|/k \rfloor$ edges of these copies as covered once. Now v_i only has $F_i \subset E_i$ non-covered adjacent edges, where $0 \le |F_i| < k$. If $|F_i| = 0$, we are done with v_i. Otherwise, $|F_i| > 0$, and we create another copy of $K_{1,k}$ whose root is v_i as follows. The copy uses the edges of F_i, but still requires $k - |F_i|$ more edges. If there is a set D_i of $k - |F_i|$ edges adjacent to v_i which are covered only once, we may use the edges of D_i for the copy, add the copy to L, mark the edges of F_i as covered once, and the edges of D_i as covered twice. Otherwise, let (v_i, u) be any edge that is covered twice. The two elements of L that use (v_i, u) have u as their root. Assume they are S_1 and S_2. If every edge of S_1 is covered twice, we delete S_1 from L, and all the edges of S_1 are marked as covered once, in particular (v_i, u) is covered once. If this is not the case, there is some edge, say (u, a) of S_1 which is covered once. In this case, we delete S_2 from L and replace it with the star obtained from S_2 by deleting the edge (u, v_i) and adding the edge (u, a). Note that now (u, a) is covered twice, but (v_i, u) is covered once. This process can be performed on any edge adjacent to v_i that is covered twice until we have $k - |F_i|$ edges adjacent to v_i that are covered once.

Our process has the property that at any stage no edge is covered more than twice, and after stage i, all edges adjacent to v_i are covered at least once. Thus, after the final stage L is a covering with overlap at most 2. $\Box$

Note that the proof of Theorem 1.3 is algorithmic, and can be performed in $O(V + E)$ time. Furthermore, a graph that does not satisfy the requirements of Theorem 1.3 has $overlap(K_{1,k}, G) = \infty$, as there is an edge (a, b) with $d(a), d(b) < k$, and this edge cannot belong to a $K_{1,k}$. This degree requirement is also detectable in polynomial time, so given a graph G we can decide if $overlap(K_{1,k}, G) \le 2$ in polynomial time, for every k.

Proof of Theorem 1.4 Let $G = (V, E)$ be a graph with $\delta(G) \ge 3$. Let L be a maximal set of edge-disjoint paths of length 3 of G (with respect to containment). Let E_1 be the set of edges of all the members of L, and put $E_2 = E \setminus E_1$. The maximality of L implies that $G_2 = (V, E_2)$ is a spanning subgraph of G whose connected components are either stars, or triangles or isolated vertices. Denote the connected components which are not isolated vertices by $S_1, \ldots, S_t$. We now perform the following process, which creates a set M of edge-disjoint paths of length 3, and shrinks $S_1, \ldots, S_t$ into connected subgraphs $T_1, \ldots T_t$, respectively. Initially, M is empty, and $T_i = S_i$ for all $i = 1, \ldots, t$. At any point in this process, the edges of $S_i \setminus T_i$ are the edges of S_i that appear in M. Furthermore, any edge of S_i that appears in M is not the central edge of the member of M in which it appears. Note that these properties hold initially.

Assume there exists an edge $(a_i, a_j) \in E_1$ which does not appear (yet) in a member of M such that $a_i \in T_i$ and $a_j \in T_j$ where $i \ne j$, and at least one of the following conditions holds for $k = i, j$:

1. S_k is a triangle. (Note that T_k is either a triangle or a proper subgraph of it at this stage). If T_k is a triangle, let (c_k, a_k) be any edge of this triangle. If $T_k = K_{1,2}$ and a_k is the root

in T_k, Let b_k, c_k be the leaves of T_k. By our assumption, (b_k, c_k) is the starting edge of some member of M. We may assume that c_k is the a non-endpoint of this member. If $T_k = K_{1,2}$ and a_k is not the root of T_k, let c_k be the root of T_k. If $T_k = K_{1,1}$ let c_k be the other member of T_k.

2. S_k is a star and a_k has degree 1 in T_k. (Note that since T_k is a subgraph of S_k, a_k also has degree 1 in S_k, unless a_k was the root of S_k and S_k contained at least three vertices). Let $(c_k, a_k) \in T_k$ (there is only one such edge).

The path (c_i, a_i, a_j, c_j) is a path of length 3, which is added to M. We update T_k, for $k = i, j$ by deleting the edge (c_k, a_k) from it. If either c_k or a_k becomes isolated by this deletion, it is also deleted from T_k. If T_k consisted only of a_k and c_k, we put $T_k = \emptyset$. Note that, indeed, M remains a set of edge-disjoint paths of length 3, and that the edges of S_i that appear in M, are exactly the edges of $S_i \setminus T_i$. Furthermore, any edge of S_i that appears in M is not the central edge of the member of M in which it appears.

We repeat the process described in the last paragraph until there is no such edge $(a_i, a_j) \in E_1$ with the required properties. When this process is complete we have that any edge appears at most once in L and at most once in M, but some may appear in both, namely, the middle edges of the members of M. Let $E'_1 \subset E_1$ denote the set of edges that appear in both L and M.

Consider the graph $G_3 = (V, E_3)$ where E_3 is the set of edges that do not appear in L nor in M. The non-isolated connected components of G_3 are exactly the subgraphs $T_1, \dots T_t$ for which $T_i \neq \emptyset$ at the end of the process of creating M. We may thus assume the non-isolated connected components of G_3 are $T_1, \dots T_{t'}$ where $t' \leq t$. For $i = 1, \dots, t'$, let $F_i \subset E_1 \setminus E'_1$ be defined as follows. If S_i is a star, F_i is the set of all edges of $E_1 \setminus E'_1$ adjacent to a vertex of degree 1 in T_i. If S_i is a triangle, F_i is the set of all edges of $E_1 \setminus E'_1$ adjacent to any vertex of T_i. Clearly, $F_i \cap F_j = \emptyset$ for $1 \leq i < j \leq t'$ (otherwise, M would have been extended, and the process of creating M would not have been completed). For each $i = 1, \dots, t'$ we create a set of paths of length 3 that cover all the edges of T_i, each one at most twice, and some edges of F_i, each one at most once, and some edges of $S_i \setminus T_i$, each one at most once. This will clearly conclude the proof of the theorem.

Consider T_i and F_i. We distinguish between the following cases:

1. $S_i = (a, b, c)$ is a triangle, and $T_i = S_i$. Since $\delta(G) \geq 3$ we have that F_i contains at least three edges, and every vertex of T_i is adjacent to at least one edge of F_i. Let $(a, d) \in F_i$ and $(b, e) \in F_i$ (it may be that $d = e$). The two paths (d, a, c, b) and (e, b, a, c) are the desired covering in this case.

2. $S_i = (a, b, c)$ is a triangle, and $T_i = K_{1,2}$ where a is the root of T_i. The edge (b, c) appears in a member of M as a non-middle edge. We may hence assume that b is the end-vertex of this member. This, and the fact that b has at least 3 neighbors in G, imply that $(b, d) \in F_i$ for some d. The path (d, b, a, c) is the desired covering in this case.

3. $S_i = (a, b, c)$ is a triangle, and $T_i = K_{1,1}$ consists only of a and b. The edges (a, c) and (b, c) appear in distinct members P and Q of M (respectively) as a non-middle edges. We claim that c cannot be the endpoint of both P and Q. To see this, assume that P was added to M prior to Q, and that c is the endpoint of P. At the beginning of the iteration that added Q to M, T_i was a $K_{1,2}$ where b was the root. The middle edge of Q cannot be adjacent to b, as this would cause the algorithm to select (a, b) for Q and not (b, c), as we assume. Thus, b is the endpoint of Q. Assume, therefore, that c is not the endpoint of Q (and hence, b is). This implies that $(b, d) \in F_i$ for some d. The path (d, b, a, c) is the desired covering in this case.

4. S_i is a star, and $T_i = K_{1,1}$. Let a, b be the vertices of T_i. If $S_i = K_{1,1}$ then both a and b each have two adjacent edges in F_i. Let $(a, c) \in F_i$ and $(b, d) \in F_i$ where $b \neq d$. The path (c, a, b, d) is the desired covering in this case. If $S_i \neq K_{1,1}$, assume a is the root of S_i. Let $c \neq b$ be another leaf of S_i. Since b has two adjacent edges in F_i, let $(b, d) \in F_i$ where $d \neq c$. The path (d, b, a, c) is the desired covering in this case.

5. S_i is a star, and $T_i = K_{1,k}$ where $k \geq 2$. This, and the fact that $\delta(G) \geq 3$, imply that each one of the leaves of T_i is adjacent to at least two edges of F_i, and hence $|F_i| \geq k$. Let $v_1, \ldots v_k$ be the leaves of T_i, and let v_0 be the root. Let $R_j = \{v_{2j-1}, v_{2j}\}$ for $j = 1, \ldots, \lfloor k/2 \rfloor$. Consider the bipartite graph $H = (A \cup F_i, P)$ which is defined as follows. The members of A are the subsets R_j, and an edge $p \in P$ connects $R_j \in A$ with $(a, b) \in F_i$ if $a \in R_j$ or $b \in R_j$ and $R_j \neq \{a, b\}$. We claim that H has a matching which matches all the elements of A. To see this, we show that Hall's condition applies (cf., e.g., [2]). Let $X \subset A$. Consider the set of $2|X|$ leaves that belong to the subsets that comprise X. There are at least $2|X|$ edges of F_i that are adjacent to one of these leaves. At-most X of them are non-neighbors of X in H, since any $R_j \in X$ disallows at most one edge (namely, the edge (v_{2j-1}, v_{2j}) if it exists). Thus X has at least $|X|$ neighbors in H. By Hall's condition, H has a matching which matches all the elements of A. We may assume that R_j is matched with the edge $(v_{2j}, w_j) \in F_i$. The set of paths $(v_{2j-1}, v_0, v_{2j}, w_j)$ for $j = 1, \ldots, \lfloor k/2 \rfloor$ is the desired covering in this case. The edge (v_0, v_k) may still be uncovered in case k is odd. We may cover it as follows. Let $f \in F_i$ be an edge that was not used for the matching. Such an edge exists since $|F_i| \geq k$ and only $\lfloor k/2 \rfloor$ edges have been used. If v_k is not an endpoint of f, we may assume $f = (v_j, w)$ for some $j \leq k - 1$. The path (v_k, v_0, v_j, w) completes the covering. If v_k is an endpoint of f, then $f = (v_k, w)$. Let v_j be such that $j \leq k - 1$ and $v_j \neq w$. Such a j exists since $k \geq 3$. The path (w, v_k, v_0, v_j) completes the covering in this case.

□

Note that the proof of Theorem 1.4 is algorithmic. Given a graph G with $\delta(G) \geq 3$ we can find a P_4 covering with overlap 2 in polynomial time. Unlike Theorem 1.3, however, this is not an "if and only if" result. There are graphs containing some vertices of degree 1 or 2 which have a P_4-covering with overlap 2.

4 The hardness aspects of covering with small overlap

Proof of Theorem 1.2 Let $\alpha < 0.5$ and let H be any graph on h edges having a connected component with three or more edges, and having a vertex of degree one. The decision problem stated in the theorem clearly belongs to NP as given a graph $G = (V, E)$ and a set L of subgraphs of G, we may verify efficiently that each member of L is isomorphic to H, and that each edge of G appears exactly once in a member of L. We show that the problem is NP-Complete by reducing from the general H-decomposition problem (which is NP-Complete by [3]). Let $G = (V, E)$ be an n-vertex graph, which is an input to the general H-decomposition problem, where n is large. Let $x > 0$ be the solution to $x - 2 = n^\alpha(x^2 + 1)^\alpha$. For every $\alpha < 0.5$ such a solution exists and $x = O(n^{\alpha/(1-2\alpha)})$. Note that x is bounded by a polynomial function of n, and for all $y \geq x$ we have $y - 2 \geq n^\alpha(y^2 + 1)^\alpha$. Let $f(H)$ be an integer such that K_k has a decomposition into H, for all $k \geq f(H)$, $h|\binom{k}{2}$. Note that $f(H)$ exists by Wilson's Theorem [6]. Let y be the minimal integer such that $y \geq x$ and K_y has an H-decomposition. Clearly, $y \leq x + f(H) + h$. Note that y is polynomial in n. Let K'_y be the graph on $y + 1$ vertices obtained from K_y by deleting some edge (a, b) from K_y and adding a new vertex c and an edge (c, a). We call (c, a) the *bridge* of K_y. Clearly, the assumption that H has a vertex of degree one implies that K'_y also has an H-decomposition. We create the graph G' as follows. To each $v \in G$ we connect y copies of K'_y where v is identified with the vertex corresponding to c in each such copy. The other y vertices of each copy belong only to that copy. The graph G' has $n' = n(y^2 + 1)$ vertices, and hence G' can be constructed in polynomial time. Also, note that

$$\delta(G') \geq y - 2 \geq n^\alpha(y^2 + 1)^\alpha = n'^\alpha.$$

It remains to show that G has an H-decomposition iff G' has. Clearly, if G is H-decomposable so is G' since G' contains G as an induced subgraph, and the remaining part of G' is just a set of ny copies of K'_y which are H-decomposable. On the other hand, consider any H-decomposition of G'. The bridges that connect each attached copy of K'_y to the vertices of G imply that any copy of H in this decomposition is either entirely in an attached K'_y copy, or entirely within G. Thus, G has an H-decomposition as well. □

The requirement that $\alpha < 0.5$ in Theorem 1.2 can be replaced with the weaker requirement that $\alpha < 1$ when $H = K_{1,k}$ and $k \geq 3$, by a slightly more complicated argument which we do not include here. We conjecture, however, that for any graph H having a connected component with three or more edges, and for $\alpha < 1$, deciding whether a graph G with $\delta(G) > n^\alpha$ has $overlap(H, G) = 1$ is NP-Complete.

In order to prove Theorem 1.5 we should first define an infinite family of graphs for which the 2-overlap decision problem is NP-Complete. Consider the tree H_k which is obtained by taking k paths of length 4 where all of the paths have a common endpoint, but are otherwise edge-disjoint. H_k has $4k+1$ vertices and $4k$ edges. For $k \geq 3$ there is a unique *root* which is the vertex of degree k in H. Alternatively, one may view H_k as a 4-subdivision of the edges of $K_{1,k}$.

Proof of Theorem 1.5 We show that for each fixed $k \geq 3$, given a graph G on n vertices, deciding whether $overlap(H_k, G) \leq 2$ is NP-Complete. The problem clearly belongs to NP as one can verify, in polynomial (in n) time if a set of subgraphs forms a covering of G by copies of H_k where each edge is covered at most twice.

Our reduction will be from the general $K_{1,k}$-decomposition problem. In order to define our construction we define the tree H'_k to be the tree obtained from H_k by contracting one of the k paths of length 4 into a path of length 1. H'_k has $4k-2$ vertices and $4k-3$ edges. Also, H'_k has a unique vertex of degree one which is adjacent to the root of H'_k. Let $G=(V,E)$ be an input for the $K_{1,k}$-decomposition problem. We construct a graph G' as follows.

1. Each edge $e=(u,v)$ of G is subdivided into four edges. We denote the three new vertices on this path by e_u, e_m, e_v and the four edges are $(u,e_u),(e_u,e_m),(e_m,e_v),(e_v,v)$. This operation introduces $3|E|$ new vertices and $4|E|$ new edges instead of the original edges of G, which we call *subdivision edges.*

2. To each vertex of type e_u (that is, a vertex that was introduced when e is subdivided and is not the middle vertex in the subdivision) we attach a path of length 2 which we denote by (e_u, e'_u, e''_u). We call this path the *forcing path.* This operation introduces $4|E|$ new vertices and $4|E|$ new edges which we call *forcing edges.*

3. To each vertex of type e_m (that is, the middle vertex in the subdivision of e) we attach a copy of H'_k which we denote by $H(e)$. The attachment is done by identifying e_m with the unique degree one vertex of H'_k which is adjacent to the root of H'_k. This operation introduces $|E|(4k-3)$ new vertices and $|E|(4k-3)$ new edges which we call *forced edges.*

The new graph G' has $|V|+|E|4(k+1)$ vertices and $|E|(4k+5)$ edges, and can thus be constructed in polynomial time.

We claim that G has a $K_{1,k}$-decomposition iff G' has $overlap(H_k, G) \leq 2$. Consider first a decomposition of G. Let G'' be the subgraph of G' obtained from the subdivision edges. G'' is simply a 4-subdivision of G. However, H_k is also a 4-subdivision of $K_{1,k}$, and hence G'' has an H_k decomposition. We still need to cover the forcing edges and the forced edges of G'. Consider a two-path (e_u, e'_u, e''_u) of forcing edges. There is exactly one copy of H_k in G' which covers the edge (e''_u, e'_u). This copy contains the edges of $H(e)$, the edge (e_u, e_m) and the edges (e_u, e'_u) and (e'_u, e''_u). Hence this copy of H_k which we denote by $H(e,u)$ must be in the covering. Taking $H(e,u)$ and

$H(e,v)$ for all $e=(u,v) \in E$, we obtain a covering of G' where the forcing edges are covered once, the forced edges are covered twice, half of the subdivision edges are covered twice (the middle edges in every subdivision), and half of the subdivision edges are covered once (the side edges in every subdivision). Consider now an H_k covering of G' with overlap at most 2. Denote this covering by L. As before, we must have that $H(e,u)$ and $H(e,v)$ are members of L for each $e=(u,v) \in E$. This already implies that the forced edges are covered twice and the other members of L do not include them. Put $L' = L \setminus \{H(e,u), H(e,v) \mid e=(u,v) \in E\}$. The members of L' only contain subdivision edges and forcing edges. We claim that every $H \in L'$ only uses subdivision edges. Indeed H has a unique vertex of degree $k \geq 3$, the root of H. The root cannot be of type e_m since e_m has degree 3, but one of its adjacent edges is a forced edge. The root cannot be of type e_u since e_u has degree 3, but it is an endpoint of a forcing path, which only has length 2, which is smaller than 4. Hence, the root of H must be an original vertex $u \in V$. Consider a path of length 4 in H which begins at u. Since it cannot use forced edges, and since forcing paths are too short, this path only uses subdivision edges. Hence, $H \in L'$ only uses subdivision edges, and every 4-path of H which begins in the root maps to a subdivision of a single edge $e \in E$. We now claim that if $H \in L'$ and $H' \in L'$ then H and H' are edge-disjoint. Indeed, if this were not the case, we would have that H and H' use a common subdivision edge, of some edge $e=(u,v) \in E$, and thus use all the 4 subdivision edges that correspond to e. In particular, they both use the edge (e_m, e_u). But (e_m, e_u) is also used by $H(e,u)$, contradicting the fact that L is a covering with overlap at most 2. We have shown that each member $H \in L'$ corresponds to k edges of E with a common endpoint, that is, to a $K_{1,k}$ in G. No two $K_{1,k}$'s share an edge since the members of L' are edge-disjoint. Furthermore every $e=(u,v) \in E$ belongs to one of these $K_{1,k}$'s since the edge (u, e_u) must be covered by a member of L'. We have thus shown that G has a $K_{1,k}$-decomposition. □

There are many other trees for which we can deduce an NP-Completeness result. Let H be any tree containing a vertex of degree 3. Let H' be obtained from H by an r-subdivision, where $r \geq 4$ is even. A similar construction to the one described in Theorem 1.4 shows that deciding whether $overlap(H', G) \leq 2$ is NP-Complete. The result can also be extended to many other graphs H, which are non-trees.

5 Concluding remarks and open problems

1. As mentioned in the introduction, the minimum degree bound in Theorem 1.1 is not best possible. By modifying (and significantly complicating) the proofs to allow more flexibility in the degrees of the graph G' one can obtain a bound which is $O(h^6)$. This is done by allowing the degrees of G' to vary between d and, say, $d + o(d/h)$ instead of d and $d+1$ and by modifying Lemma 2.3 accordingly. However, this is still far from the obvious lower bound of $h-1$ described in the introduction. We thus conjecture the following:

Conjecture 5.1 *For every tree H on h vertices, any graph G with $\delta(G) \geq h-1$ has overlap$(H,G) \leq 2$.*

Note that Theorems 1.3 and 1.4 show that Conjecture 5.1 holds for stars and for P_4.

2. Conjecture 5.1, if true, does not imply Theorem 1.1, as Theorem 1.1 also guarantees that a small fraction, of $O(\delta(G)^{-0.1})$, of the edges of G are covered twice. This near-packing result does not hold for graphs with minimum degree $h-1$. Consider a covering of K_h with $K_{1,h-1}$ having overlap 2. Such a covering must contain at least $h-1$ members, and hence all but at most $h-1$ edges are covered twice.

3. An H-covering of G is *k-intersecting* if every two elements in the covering share at most k edges. Clearly, if *overlap*$(H,G) > 1$ then any H-covering of G is at least 1-intersecting. It is quite easy to modify the proof of Lemma 2.4 such that when we create the copies of H, we maintain a 1-intersection property as-well. Each time we extend a subtree H' of H on i vertices by adding to it a new edge, we choose an edge that does not belong to any of the copies that already intersect H'. At-most $i-1$ copies intersect it, and they each have no more than i edges, thus we should avoid less than i^2 edges. The lemma still holds if, say, $d \geq sh^2 + 2h^3$. Thus we can strengthen Theorem 1.1 to include a 1-intersection requirement if the minimal degree is, say, $(200h)^{10}$. Conjecture 5.1 may also be strengthened to include a 1-intersection requirement.

4. Theorem 1.3 implies that given a graph G, deciding whether *overlap*$(K_{1,k}, G) \leq 2$ can be done in polynomial time, for every k. On the other hand, Theorem 1.5 shows that there are infinitely many (fixed) trees for which this decision problem is NP-Complete. The smallest tree for which we have an NP-Completeness result is the tree H_3, defined in Section 4, which contains 12 edges. A challenging open problem is to characterize all graphs (or, alternatively, all trees) for which the 2-overlap problem is NP-Complete, and to characterize all trees for which the 2-overlap problem is polynomial.

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley and Sons Inc., New York, 1991.

[2] J. A. Bondy and U.S. R. Murty, *Graph Theory with Applications*, Macmillan Press, London, 1976.

[3] D. Dor and M. Tarsi, *Graph decomposition is NPC - A complete proof of Holyer's conjecture*, Proc. 20th ACM STOC, ACM Press (1992), 252-263.

[4] M. R. Garey and D. S. Johnson, *Computers and Intractability, A Guide to the Theory of NP-Completeness*, W. H. Freeman and Company, New York, 1979.

[5] T. Gustavsson, *Decompositions of large graphs and digraphs with high minimum degree*, Doctoral Dissertation, Dept. of Mathematics, Univ. of Stockholm, 1991.

[6] R. M. Wilson, *Decomposition of complete graphs into subgraphs isomorphic to a given graph*, Congressus Numerantium XV (1975), 647-659.