# Covering a hypergraph of subgraphs

Noga Alon *

*Dedicated to Dan Kleitman, for his* $65^{th}$*-birthday.*

**Abstract**

Let $G$ be a tree and let $\mathcal{H}$ be a collection of subgraphs of $G$, each having at most $d$ connected components. Let $\nu(\mathcal{H})$ denote the maximum number of members of $\mathcal{H}$ no two of which share a common vertex, and let $\tau(\mathcal{H})$ denote the minimum cardinality of a set of vertices of $G$ that intersects all members of $\mathcal{H}$. It is shown that $\tau(\mathcal{H}) \leq 2d^2\nu(\mathcal{H})$. A similar, more general result is proved replacing the assumption that $G$ is a tree by the assumption that it has a bounded tree-width. These improve and extend results of various researchers.

## 1 Introduction

Let $\mathcal{H}$ be a finite collection of subgraphs of a finite graph $G$. The *covering number* (or *piercing number*) $\tau(\mathcal{H})$ of $\mathcal{H}$ is the minimum cardinality of a set of vertices of $G$ that intersects every member of $\mathcal{H}$. The *matching number* $\nu(\mathcal{H})$ of $\mathcal{H}$ is the maximum number of pairwise vertex disjoint members of $\mathcal{H}$. Clearly $\tau(\mathcal{H}) \geq \nu(\mathcal{H})$. In general, $\tau(\mathcal{H})$ cannot be bounded from above by a function of $\nu(\mathcal{H})$, as shown, for example, by all induced subgraphs on $n$ vertices of an arbitrary graph on $2n-1$ vertices, where $\nu = 1$ and $\tau = n$. If, however, the graph $G$ is a tree and each member of $\mathcal{H}$ has at most $d$ connected components, then $\tau$ can be bounded by a function of $\nu$ and $d$.

Gallai noticed that if $G$ is a path and $d = 1$ then $\tau = \nu$. More generally, Surányi (see [4]) proved that the intersection graph of subtrees of a tree is chordal, implying that if $G$ is any tree and $d = 1$ then $\nu = \tau$. Gyárfás and Lehel [4] proved that for $d = 2$, if $\nu = 1$ then $\tau \leq 3$, and that if $G$ is a path then for general $d$, $\tau \leq O(\nu^{d!})$. They also mentioned that $\tau$ can be bounded by a (similarly fast growing) function of $\nu$ and $d$ for general trees using related ideas. For $G$ being a path and general $d$, Kaiser [5] proved that $\tau \leq (d^2 - d + 1)\nu$. His proof is topological, applies the Borsuk-Ulam theorem and extends and simplifies a result of Tardos [9]. A short proof of the slightly weaker estimate that in this case $\tau \leq 2d^2\nu$ is described in [1]. This proof is based on the ideas of [3]. See also [10] for a short survey.

Here we prove the following result, extending and improving some of the above mentioned ones.

*Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Research supported in part by a USA-Israeli BSF grant, by the Israel Science Foundation and by the Hermann Minkowski Minerva Center for Geometry at Tel Aviv University. Email: noga@math.tau.ac.il.

**Theorem 1.1** *Let $G$ be an arbitrary tree and let $\mathcal{H}$ be a collection of subgraphs of $G$, each having at most $d$ connected components. Then $\tau(\mathcal{H}) \leq 2d^2\nu(\mathcal{H})$.*

We also prove a more general result, for graphs with bounded tree-width (see Section 4 for the relevant definitions).

**Theorem 1.2** *Let $G$ be an arbitrary graph of tree-width at most $b$ and let $\mathcal{H}$ be a collection of subgraphs of $G$, each having at most $d$ connected components. Then $\tau(\mathcal{H}) \leq 2(b+1)d^2\nu(\mathcal{H})$.*

The proofs are based on the method of [3] (see also [2]) but require some additional ideas for dealing with subgraphs of trees or subgraphs of graphs with bounded tree-width. We first obtain an upper bound for the fractional covering number $\tau^*(\mathcal{H})$ $(= \nu^*(\mathcal{H}))$ in terms of $\nu(\mathcal{H})$ and then bound $\tau(\mathcal{H})$ in terms of $\tau^*(\mathcal{H})$.

The term *piercing* is used in the study of these questions in the geometric context (see, e.g., [3]), where, for a family of planar sets $\mathcal{H}$, the parameter $\tau(\mathcal{H})$ is the minimum number of needles needed to pierce all the members of the family. Since here we are dealing with a graph theoretic variant, we prefer to call $\tau(\mathcal{H})$ the covering number of $\mathcal{H}$, as usual.

## 2 Two Lemmas

Our approach is based on the one in [3], where the key ingredients are the notions of fractional Helly theorems and weak $\epsilon$-nets, together with linear programming duality. The following lemma is a fractional Helly type result for subtrees of a tree.

**Lemma 2.1** *Let $\mathcal{H}$ be a collection of $n$ (not necessarily distinct) subtrees of a tree $G$, and suppose that there are at least $nf/2$ intersecting unordered pairs of members of $\mathcal{H}$. Then there is a vertex of $G$ contained in at least $f/2+1$ members of $\mathcal{H}$.*

**Proof.** As long as there is a subtree in the family $\mathcal{H}$ that intersects less than $f/2$ others, omit one such subtree from the family. Note that this process must terminate with a nonempty subfamily of $\mathcal{H}$, since the number of intersecting pairs decreases in each step by less than $f/2$, and hence would stay positive if the remaining family would vanish, which is impossible. Therefore, there is a nonempty subfamily $\mathcal{H}'$ of subtrees in which each member intersects at least $f/2$ others. Let $u$ be an arbitrary vertex of $G$ and consider $G$ as a tree rooted at $u$. Among all vertices $x$ for which there is a member of $\mathcal{H}'$ which is contained in the subtree rooted at $x$, let $v$ be one whose distance from $u$ is maximum. Suppose $T \in \mathcal{H}'$ is contained in the subtree rooted at $v$. Then every element of $\mathcal{H}'$ that intersects $T$ must contain the vertex $v$, and since there are at least $f/2$ such elements besides $T$ itself, the desired result follows. $\square$

The next lemma is applied in Section 3 to construct the weak $\epsilon$-net suitable for our purpose here.

**Lemma 2.2** *For two positive integers $m$ and $r$, let $R$ be an arbitrary multi-set of at most $rm$ vertices in a tree $G$. Then, there is a set $S$ of at most $m-1$ vertices of $G$ so that each connected component of $G-S$ contains at most $r$ members of $R$.*

**Proof.** We apply induction on $m$, the result being trivial for $m = 1$. Assuming it holds for $m-1$, we prove it for $m$ ( $\geq 2$). Let $u$ be an arbitrary vertex of $G$ and consider $G$ as a tree rooted at $u$. Among all vertices $x$ for which the total number of members of $R$ in the subtree rooted at $x$ is at least $r$, let $v$ be one whose distance from $u$ is maximum. Then the number of vertices of $R$ in each connected component of $G - v$ besides the one containing the root $u$ is less than $r$. Let $G'$ be the tree obtained from $G$ by removing the subtree rooted at $v$ (including $v$). Note that $G'$ contains at most $r(m-1)$ members of $R$. By the induction hypothesis there is a set $S'$ in $G'$ such that each connected component of $G' - S'$ contains at most $r$ members of $R$. The set $S = S' \cup \{v\}$ clearly satisfies the required assertion, completing the proof. $\square$

## 3 Trees

Let $\mathcal{H}$ be a collection of subgraphs of a finite graph $G = (V, E)$. The *fractional matching number* $\nu^*(\mathcal{H})$ of $\mathcal{H}$ is the maximum possible value of the sum $\sum_{T \in \mathcal{H}} g(T)$, where the maximum is taken over all real-valued functions $g : \mathcal{H} \mapsto [0, 1]$ satisfying $\sum_{T: v \in T \in \mathcal{H}} g(T) \leq 1$ for every vertex $v$ of $G$. Note that this maximum is obtained for a function attaining rational values. Note also that if we let $g : \mathcal{H} \mapsto \{0, 1\}$ instead, this integer program now defines $\nu(\mathcal{H})$. The *fractional covering number* $\tau^*(\mathcal{H})$ of $\mathcal{H}$ is the minimum possible value of the sum $\sum_{v \in V} h(v)$, where the minimum is taken over all real valued functions $h : V \mapsto [0, 1]$ satisfying $\sum_{v \in V: v \in T} h(v) \geq 1$ for every $T \in \mathcal{H}$. Here, too, the minimum is obtained for a function attaining rational values. By the duality theorem of linear programming we have $\nu^*(\mathcal{H}) = \tau^*(\mathcal{H})$, and by definition $\nu(\mathcal{H}) \leq \nu^*(\mathcal{H})$ and $\tau^*(\mathcal{H}) \leq \tau(\mathcal{H})$. We next show that if $\mathcal{H}$ is nonempty, $G$ is a tree, and each member of $\mathcal{H}$ has at most $d$ components, then

$$\tau^*(\mathcal{H}) = \nu^*(\mathcal{H}) < 2d\nu(\mathcal{H})$$

and

$$\tau(\mathcal{H}) \leq d\tau^*(\mathcal{H}).$$

This clearly implies the assertion of Theorem 1.1.

To complete the proof it thus suffices to prove the above two inequalities. This is done in the following two lemmas.

**Lemma 3.1** *Let $G$ be a tree, and let $\mathcal{H}$ be a nonempty collection of subgraphs of $G$, each having at most $d$ connected components. Then $\nu^*(\mathcal{H}) < 2d\nu(\mathcal{H})$.*

**Proof.** Put $k = \nu(\mathcal{H})$ and let $g : \mathcal{H} \mapsto [0, 1]$ be a function, where $g(T)$ is rational for each $T \in \mathcal{H}$, $\sum_{T \in \mathcal{H}} g(T) = \nu^*(\mathcal{H})$, and $\sum_{T: v \in T \in \mathcal{H}} g(T) \leq 1$ for every vertex $v$ of $G$. Let $m$ be an integer for which $mg(T)$ is integral for each $T \in \mathcal{H}$ and put $M = \sum_{T \in \mathcal{H}} mg(T)$. Let $\mathcal{H}'$ be the multiset consisting of $mg(T)$ copies of $T$ for each $T \in \mathcal{H}$, and note that $|\mathcal{H}'| = M$. Let $\mathcal{H}''$ be the multiset obtained from $\mathcal{H}'$ by replacing each member of $\mathcal{H}'$ by its components. Put $n = |\mathcal{H}''|$ and note that $n \leq Md$. Since there are no $k+1$ pairwise disjoint members of $\mathcal{H}'$, Turán's Theorem implies that there are at least

$\frac{M}{2}(\frac{M}{k}-1)$ intersecting pairs of members of $\mathcal{H}'$. Thus there are at least

$$\frac{M}{2}(\frac{M}{k}-1) \geq \frac{n}{2d}(\frac{M}{k}-1) > \frac{n}{2}(\frac{M}{kd}-2)$$

intersecting pairs of members of $\mathcal{H}''$. By Lemma 2.1 this implies that there is a vertex $v$ of $G$ contained in more than $\frac{M}{2kd}$ members of $\mathcal{H}''$ (and hence of $\mathcal{H}'$). Therefore

$$1 \geq \sum_{T: v\in T \in \mathcal{H}} g(T) > \frac{1}{m}\frac{M}{2kd} = \frac{\sum_{T\in\mathcal{H}} g(T)}{2kd} = \frac{\nu^*(\mathcal{H})}{2kd},$$

implying that $\nu^*(\mathcal{H}) < 2kd$, and completing the proof. $\Box$

**Lemma 3.2** *Let $G=(V,E)$ be a tree, and let $\mathcal{H}$ be a nonempty collection of subgraphs of $G$, each having at most $d$ connected components. Then $\tau(\mathcal{H}) \leq d\tau^*(\mathcal{H})$.*

**Proof.** Let $h: V \mapsto [0,1]$ satisfy $\sum_{v\in V: v\in T} h(v) \geq 1$ for every $T \in \mathcal{H}$, where $h(v)$ is rational for all $v \in V$ and $\tau^*(\mathcal{H}) = \sum_{v\in V} h(v)$. Let $r > r'$ be two positive integers such that $(rd+r')h(v)$ is an integer for all $v$, and let $R$ be the multiset consisting of $(rd+r')h(v)$ copies of $v$, for each $v \in V$. Note that each member of $\mathcal{H}$ contains at least $rd+r'$ points of $R$, and hence it has some connected component that contains at least $r+1$ points of $R$. By Lemma 2.2 with $m = \lceil (d+\frac{r'}{r})\sum_{v\in V} h(v)\rceil$ there is a set $S$ of at most $m-1 < (d+\frac{r'}{r})\sum_{v\in V} h(v) = (d+\frac{r'}{r})\tau^*(\mathcal{H})$ vertices of $G$ such that every connected component of $G-S$ contains at most $r$ points of $R$. This means that each member of $\mathcal{H}$ contains a point of $S$, since otherwise each of its components (including the one containing more than $r$ points of $R$) would lie in a component of $G-S$, which contains at most $r$ points of $R$. Therefore, $\tau(\mathcal{H}) < (d+\frac{r'}{r})\tau^*(\mathcal{H})$, and since we can keep $r'$ fixed and choose an arbitrarily large $r$ the desired result follows. $\Box$

## 4 Bounded tree-width

In this section we observe that Theorem 1.2 follows from Theorem 1.1.

The concept of tree-width was introduced by Robertson and Seymour in their series of works on graph minors. See, e.g., [7].

A *tree-decomposition* of a graph $G=(V,E)$ is a pair $(X,T)$ where $T=(I,F)$ is a tree and $X=\{X_i : i \in I\}$ is a family of subsets of $V$ such that (i) $\cup_{i\in I} X_i = V$; (ii) for every edge $(u,v) \in E$, there exists an $i \in I$ such that $u, v \in X_i$; and (iii) if $i,j,k \in I$ and $j$ is on the path from $i$ to $k$ in $T$, then $X_i \cap X_k \subseteq X_j$. The *tree-width* of the tree-decomposition $(X,T)$ is $\max_{i\in I}|X_i|-1$. The tree-width of a graph $G$ is the minimum tree-width over all possible tree-decompositions of $G$. Graphs with tree-width at most $b$ are also called *partial $b$-trees*. In particular, a connected graph has tree-width 1 if and only if it is a tree.

**Proof of Theorem 1.2** Fix a tree-decomposition of $(X,T)$ of $G$, where $T=(I,F)$, $X=\{X_i : i \in I\}$ and $|X_i| \leq b+1$ for each $i \in I$. For each subgraph $H \in \mathcal{H}$ let $H'$ be the subgraph of $T$ induced on all vertices $i \in I$ for which $X_i$ contains a vertex of $H$. Let $\mathcal{H}'$ denote the set of all subgraphs $H'$ of

$T$ obtained in this way. It is not difficult to check that each member of $\mathcal{H}'$ has at most $d$ connected components, and that $\nu(\mathcal{H}') \leq \nu(\mathcal{H})$. Therefore, by Theorem 1.1 there is a set $S' \subset I$ of at most $2d^2\nu(\mathcal{H}') \leq 2d^2\nu(\mathcal{H})$ vertices of $T$ that intersects each member $H'$ of $\mathcal{H}'$. The set $S = \cup_{i \in S'} X_i$ is thus a set of size at most $2(b+1)d^2\nu(\mathcal{H})$ that intersects all members of $\mathcal{H}$, completing the proof. □

## 5 Concluding remarks and open problems

- The assumption that $G$ has a bounded tree-width is necessary in Theorem 1.2. Indeed, for every integer $c$ there exists a $b = b(c)$ such that **every** graph $G$ with tree-width at least $b$ contains a collection $\mathcal{H}$ of subtrees such that $\nu(\mathcal{H}) = 1$ and $\tau(\mathcal{H}) \geq c$. This is because any $G$ with a sufficiently large tree-width contains a large grid minor (see [8]), and by considering the collection of all subgraphs of that grid consisting of a union of a horizontal path and a vertical path in it, we obtain the desired family.
- Very recently, J. Matoušek [6] applied a construction of J. Sgall and proved that even when the graph $G$ is a path, the quadratic dependence on the number of components $d$ in Theorems 1.1 and 1.2 is optimal, up to a logarthmic factor. It would be interesting to decide if this logarithmic factor is indeed necessary. Simple examples show that a better than linear dependence on $b$ in Theorem 1.2 does not hold.

## References


[1] N. Alon, Piercing $d$-intervals, Discrete and Computational Geometry 19 (1998), 333-334.

[2] N. Alon and G. Kalai, Bounding the piercing number, Discrete and Computational Geometry 13 (1995), 245-256.

[3] N. Alon and D. J. Kleitman, Piercing convex sets and the Hadwiger Debrunner $(p, q)$-problem, Advances in Mathematics 96 (1992), 103-112.

[4] A. Gyárfás and J. Lehel, A Helly-type problem in trees, in *Combinatorial Theory and its Applications*, Eds. P. Erdős, A. Rényi and V.T. Sós, North-Holland, Amsterdam 1970, 571–584.

[5] T. Kaiser, Transversals of $d$-intervals, Discrete and Computational Geometry 18 (1997), 195-203.

[6] J. Matoušek, Lower bounds on the transversal numbers of $d$-intervals, to appear.

[7] N. Robertson and P. Seymour, Graph minors. II. Algorithmic aspects of tree-width, Journal of Algorithms 7 (1986), 309–322.

[8] N. Robertson, P. Seymour and R. Thomas, Quickly excluding a planar graph, J. Combinatorial Theory Ser. B 62 (1994), 323–348.

[9] G. Tardos, Transversals of 2-intervals, a topological approach, Combinatorica 15 (1995), 123–134.

[10] G. Tardos, Transversals of $d$-intervals—comparing three approaches, European Congress of Mathematics, Vol. II (Budapest, 1996), Progr. Math., 169, Birkhauser, Basel, 1998, pp. 234–243.