Overlap Number of Graphs

Daniel W. Cranston[*], Nitish Korula[†], Timothy D. LeSaulnier[‡], Kevin G. Milans[§],
Christopher Stocker[¶], Jennifer Vandenbussche[‖], Douglas B. West[**]

December 5, 2009

Abstract

An *overlap representation* of a graph G assigns sets to vertices so that vertices are adjacent if and only if their assigned sets intersect with neither containing the other. The *overlap number* $\varphi(G)$ (introduced by Rosgen) is the minimum size of the union of the sets in such a representation. We prove the following: (1) An optimal overlap representation of a tree can be produced in linear time, and its size is the number of vertices in the largest subtree in which the neighbor of any leaf has degree 2. (2) If $\delta(G) \geq 2$ and $G \neq K_3$, then $\varphi(G) \leq |E(G)| - 1$, with equality when G is connected and triangle-free and has no star-cutset. (3) If G is an n-vertex plane graph with $n \geq 5$, then $\varphi(G) \leq 2n - 5$, with equality when every face has length 4 and there is no star-cutset. (4) If G is an n-vertex graph with $n \geq 14$, then $\varphi(G) \leq n^2/4 - n/2 - 1$, with equality for even n when G arises from $K_{n/2,n/2}$ by deleting a perfect matching.

1   Introduction

Intersection representations of graphs have been studied for many years. An *intersection representation* of a graph is a family of sets corresponding to the vertices so that vertices are adjacent if and only if their assigned sets intersect. The first such model was that of *interval graphs*, in which the assigned sets are intervals on the real line.

---

[*]Virginia Commonwealth University, dcranston@vcu.edu

[†]University of Illinois, nkorula2@illinois.edu

[‡]University of Illinois, tlesaul2@illinois.edu

[§]University of Illinois, milans@math.uiuc.edu

[¶]University of Illinois, stocker2@uiuc.edu

[‖]Southern Polytechnic State University, jvandenb@spsu.edu

[**]University of Illinois, west@math.uiuc.edu. Work supported in part by the NSA under award H98230-06-1-0065.

Intersection representations may use various types of sets. Erdős, Goodman, and Pósa [3] introduced intersection representations using finite sets. The *intersection number* $\theta_1(G)$ is the minimum size of the union of the sets in an intersection representation of G by finite sets ([1] and [2] use this notation). In [3], it was shown that $\theta_1(G)$ also equals the minimum number of complete subgraphs needed to cover $E(G)$.

The "overlap" model for graph representations arose much later and is less well studied. A set *overlaps* another set if they intersect but neither contains the other. An *overlap representation* of a graph G is an assignment f of sets to the vertices of G so that $uv \in E(G)$ if and only if $f(u)$ and $f(v)$ overlap.

Just as intersection representations were first studied using intervals, so too an *overlap graph* was defined to be a graph having an overlap representation using intervals. The concept appears in the classic book by Golumbic [4], noting that a graph is an overlap graph if and only if it has an intersection representation using chords of a circle. MathSciNet returns less than 50 items for "overlap graph" and more than 600 for "interval graph", though it should be noted that overlap graphs are also discussed under equivalent terms like "circle graph".

Rosgen [6] studied overlap representations using finite sets. Under any adjacency rule for assigned sets (such as intersection, containment, or overlap), a *finite representation* of a graph G is a representation in which the assigned sets are finite. The *size* of a finite representation f of G, denoted $|f|$, is the size of the union of the assigned sets. The *overlap number* $\varphi(G)$ is the minimum size of a finite overlap representation of G.

Throughout this paper, we take n to be the number of vertices of a graph G whose overlap number is being studied. Rosgen [6] obtained upper bounds on $\varphi(G)$ for trees $(n+1)$, chordal graphs $(2n)$, planar graphs $(\frac{10}{3}n-6)$, and arbitrary graphs $(\theta_1(G)+n$, which yields $\varphi(G) \leq \lfloor n^2/4 \rfloor + n)$. He observed that $\varphi(K_n)$ is the minimum t such that a t-set contains n pairwise incomparable sets, that $\varphi(C_n) = n - 1$, and that the overlap number of any caterpillar (with $n > 2$) is the number of vertices in the longest path. He asked for the maximum value of $\varphi(G)$ in terms of n for trees, chordal graphs, planar graphs, and arbitrary n-vertex graphs, and also for the complexity of computing φ on trees and on general graphs.

We answer Rosgen's questions about trees using a special subtree. A *skeleton* is a tree in which the neighbor of any leaf vertex has degree 2. The largest skeleton in a tree T is unique up to isomorphism, obtained by deleting all leaves (yielding the *derived tree T'*) and then restoring one leaf neighbor of each leaf of T'. Hence we call this *the* skeleton of the tree. For $n \geq 3$, we prove that the overlap number of a tree is the number of vertices in its skeleton, using an algorithm that produces an overlap representation of this size in linear time.

In Section 2 we give the algorithm and formula for φ on n-vertex trees. Section 3 presents bounds in terms of the number of edges; we prove that $\varphi(G) \leq |E(G)| - 1$ when $\delta(G) \geq 2$ and $G \neq K_3$. Furthermore, equality holds when G is connected, triangle-free, and has no

star-cutset, where a *star-cutset* is a separating set S having a vertex x adjacent to all of $S - \{x\}$. These results are applied to n-vertex planar graphs in Section 4 and to the family of all n-vertex graphs in Section 5.

In particular, if G is an n-vertex plane graph with $n \geq 5$, then $\varphi(G) \leq 2n - 5$, with equality when every face is a 4-cycle and there is no star-cutset. When G is any n-vertex graph with $n \geq 14$, we prove that $\varphi(G) \leq n^2/4 - n/2 - 1$. The bound is sharp when n is even, with equality holding when G arises from $K_{n/2,n/2}$ by deleting a perfect matching. When n is odd, a graph with overlap number $\lfloor n^2/4 - n \rfloor$ is formed by duplicating a vertex in the previous construction for $n - 1$ vertices.

Henderson [5] independently obtained results on these and related problems. He obtained constant-factor approximation algorithms for computing the overlap number on trees and on planar graphs, and he proved that the maximum overlap number grows quadratically in the number of vertices for a class of bipartite graphs. It remains open whether finding the overlap number is NP-hard in general.

Some of our results use a related model. A *pure overlap representation* of G is an overlap representation in which no assigned set contains another. The *pure overlap number* $\Phi(G)$ is the minimum size of a finite pure overlap representation of G. (Rosgen used the term "containment-free overlap representation" for this model.) Note that a pure overlap representation of G is both an overlap representation and an intersection representation of G; thus always $\varphi(G) \leq \Phi(G)$ and $\theta_1(G) \leq \Phi(G)$. For this reason, $\Phi(G)$ is helpful in proving upper bounds. Note also that $\rho(H) \leq \rho(G)$ when $\rho \in \{\varphi, \Phi, \theta_1\}$ and H is an induced subgraph of G, since a representation of G restricts to a representation of H.

We say that the vertices adjacent to a vertex v in G are its *neighbors*. The number of neighbors is the *degree* of v, denoted $d_G(v)$ or simply $d(v)$. The set of neighbors is the *neighborhood* of v, denoted $N_G(v)$ or simply $N(v)$. The *closed neighborhood* of v, denoted $N[v]$, is $N(v) \cup \{v\}$. The minimum vertex degree is $\delta(G)$. A vertex of degree 1 is a *leaf*. A graph is *nontrivial* if it has at least one edge.

Before beginning the discussion of trees, we prove a lemma used in the lower bound arguments. It restricts the form of overlap representations. The idea is due to Rosgen [6].

Lemma 1.1. *Let f be an overlap representation of a graph G. If $v \in V(G)$ and H is a nontrivial component of $G - N[v]$, then either $f(v)$ properly contains all sets assigned to the vertices of H or $f(v)$ is disjoint from all sets assigned to vertices of H.*

Proof. Since no sets used in H overlap $f(v)$, and H is connected, it suffices to show that if $f(v) \supseteq f(u)$ for some $u \in V(H)$, and $x \in N(u)$, then $f(v) \supset f(x)$.

Since $f(u)$ and $f(x)$ overlap, $f(v) \cap f(x) \neq \varnothing$. Since $x \notin N(v)$, we have $f(v)$ and $f(x)$ ordered by inclusion. Since $x \in N(u)$ forbids $f(x) \supseteq f(v) \supseteq f(u)$, we have $f(v) \supset f(x)$. $\square$

2   The overlap number of trees

Rosgen [6] proved that $\varphi(T) \leq n + 1$ when T is a tree. In fact, this bound is sharp only for K_2. We provide a linear-time algorithm for producing an overlap representation of a tree. We then prove that this representation is optimal.

A *caterpillar* is a tree in which all edges are incident to a single path. Rosgen [6] proved that the overlap number of any caterpillar equals the number of vertices in a longest path. For a caterpillar, this path is the skeleton. We will need this result along with a technical property of the representation, because our procedure for extending a representation along an added caterpillar differs from the representation for the initial caterpillar.

Definition 2.1. For an overlap representation f of a graph G, the *associated poset* P_f is the inclusion order on $\{f(v) \colon v \in V(G)\}$. A vertex v is *minimal* in f if $f(v)$ is a minimal element of P_f, and v is *a-minimal* if $f(v)$ is a minimal element of the subposet of P_f consisting of the elements that contain a. In the same way that $\supseteq$ means "contains", we use $\leftrightarrow$ to mean "overlaps" and "$\|$" to mean "does not intersect".

Lemma 2.2. *Let T be a caterpillar whose longest path has vertices $v_1, \ldots, v_l$ in order. If $l \geq 3$, then T has an overlap representation f of size l. Furthermore, with $\{a_1, \ldots, a_l\}$ being the union of the assigned sets, f may be chosen so that v_i is a_i-minimal in $1 \leq i \leq l - 1$.*

Proof. Let $f(v_i) = \{a_i, a_{i+1}\}$ for $1 \leq i \leq l - 1$. All leaves (including v_l) have a neighbor in $\{v_2, \ldots, v_{l-1}\}$. For each leaf neighbor x of v_i, let $f(x) = \{a_1, \ldots, a_i\}$.

By construction, $f(v_{i-1}) \leftrightarrow f(v_i)$ for $2 \leq i \leq l - 1$, and nonconsecutive sets in that list are disjoint. If x is a leaf neighbor of v_i, then $f(x) \leftrightarrow f(v_i)$, $f(x) \supseteq f(v_j)$ for $j < i$, and $f(x) \| f(v_j)$ for $j > i$. Also the sets assigned to leaves form a chain by inclusion. Hence f is an overlap representation of T. Since no assigned sets are singletons, v_i is a_i-minimal. $\square$

Observation 2.3. If A and B are sets such that $A \supseteq B$ or $A \leftrightarrow B$, then adding an element not in $A \cup B$ to A or to both A and B preserves the relation. If $A \| B$, then the relation is preserved when the element is added to just one of $\{A, B\}$. $\square$

Lemma 2.4. *Let G be the union of a graph H and a caterpillar T such that $H \cap T$ consists of one vertex v that is not isolated in H and is an endpoint of a longest path in T. Let H have an overlap representation f of size m, and let $w_0, \ldots, w_l$ be the vertices along a longest path in T, with $v = w_0$. If v is a-minimal in f for some $a \in f(v)$, then G has an overlap representation f' of size $m + l$, with added elements $b_1, \ldots, b_l$, such that w_i is b_i-minimal in f' for $1 \leq i \leq l$, and any vertex of H that is c-minimal in f is also c-minimal in f'.*

Proof. Let $B = \{b_1, \ldots, b_l\}$, and let $b_0 = a$. Let $f'(v) = f(v)$, and let $f'(w_i) = \{b_{i-1}, b_i\}$ for $1 \leq i \leq l - 1$. Each remaining vertex of T is a leaf with neighbor in $\{w_1, \ldots, w_{l-1}\}$. For each

4

leaf x in T with neighbor w_i, let $f'(x) = \{b_i, \ldots, b_l\}$. For $u \in V(H) - \{v\}$, let $f'(u) = f(u)$ if $a \notin f(u)$; otherwise, let $f'(u) = f(u) \cup B$.

By construction, f' generates a path on $w_0, \ldots, w_l$, since $d_H(v) \geq 1$ requires $f(v) \neq \{a\}$. If x is a leaf in T adjacent to w_j, then $f'(x)$ contains the sets assigned to w_i and its leaf neighbors for $i > j$. Also $f'(x) \leftrightarrow f'(w_j)$, and $f'(x) \| f'(w_i)$ for $i < j$.

If $u \in V(H) - \{v\}$ and $y \in V(T) - \{v\}$, either $f'(u) \| f'(y)$ or $f'(u) \supseteq f'(y)$, depending on whether $f'(u)$ acquires B. Since $B \subseteq f'(u)$ if and only if $a \in f(u)$, by Observation 2.3 the relation between sets assigned to vertices of $V(H) - \{v\}$ under f' and f is the same.

Since $f'(v) \| B$, among the sets assigned by f' to $V(T) - \{v\}$ only $f'(w_1)$ overlaps $f'(v)$. Now compare v with a vertex $u \in V(H) - \{v\}$. Since v is a-minimal, $f(u) \subset f(v) = f'(v)$ implies $f'(u) = f(u)$. Otherwise, Observation 2.3 implies that $f'(u)$ and $f'(v)$ have the same relation as $f(u)$ and $f(v)$. We have shown that f' is an overlap representation of G.

Note that as in Lemma 2.2, each w_i is b_i-minimal in f'. If u is c-minimal in f, then for every y with $c \in f(y)$, Observation 2.3 implies in all cases that u is c-minimal in f'. $\qquad\square$

Theorem 2.5. *Every tree other than K_2 has an overlap representation whose size is the number of vertices in its skeleton.*

Proof. We grow a tree T by successive addition of appropriate caterpillars. The first caterpillar, T_0, consists of any maximal subtree of T that is a caterpillar whose leaves are also leaves of T. The maximality guarantees that the ends of a longest path in T_0 are leaves of T that are also leaves in the skeleton.

When the subtree absorbed so far is T_i, the next caterpillar T' is a maximal caterpillar contained in T such that an endpoint x of some longest path of T' (and no other vertex of T') is in T_i, and all leaves of T' are leaves in T. Let $T_{i+1} = T_i \cup T'$. The end opposite x of a longest path in T' is a leaf of T that is preserved in the skeleton. Thus the maximality conditions guarantee that the subtree of T formed by the union of the longest paths in the chosen caterpillars is the skeleton of T.

By Lemma 2.2, the initial caterpillar has an overlap representation of the desired size, with all non-leaf vertices being c-minimal for distinct choices of c. By Lemma 2.4, the process continues with the b-minimality conditions on non-leaf vertices preserved and the desired number of elements being added at each step. (In fact, in the final overlap representation f, only one vertex of the skeleton is not c-minimal for any c; it is a leaf of T_0.) $\qquad\square$

The skeleton of any tree G is an induced subgraph of G. Therefore, to prove that the overlap representation produced in Theorem 2.5 is optimal for every tree with $n \geq 3$, it suffices to show that if T is a skeleton with n vertices, then $\varphi(T) = n$.

The idea of the proof is inductive. Given an overlap representation f for a skeleton T, we seek one or two vertices in T (a leaf or a leaf and its neighbor) whose deletion yields a

5

smaller skeleton T' for which we can obtain an overlap representation by deleting one or two elements from f. The lower bound then follows inductively. To do this, we need to know when elements can be deleted from an overlap representation f or from a restriction of f to a subgraph. We write $f - S$ for the result of subtracting S from each set assigned under f.

Lemma 2.6. *If f is an overlap representation of a graph G, then $f - S$ is an overlap representation of G if and only if S does not contain the intersection or difference of the sets assigned to any two adjacent vertices of G.*

Proof. Necessity: Deleting a set containing the intersection or difference of the sets for adjacent vertices would delete that edge from the corresponding overlap graph.

   Sufficiency: Deleting a set S satisfying the stated condition maintains the overlap condition for any pair of overlapping sets. Deletions from disjoint sets maintain disjointness, and containments are preserved because $A \subseteq B$ implies $A - S \subseteq B - S$. $\quad\square$

Definition 2.7. Let f be an assignment of sets to $V(G)$. A set S of elements is f-*uniform* if every assigned set $f(v)$ contains all or none of S.

Observation 2.8. If f is an overlap representation of a graph G, then every proper subset of an f-uniform set is deletable from f. Hence an overlap representation having a uniform set of size 2 is not optimal. $\quad\square$

   Our next lemma is the key tool in proving the lower bound for trees. It strengthens Observation 2.8, allowing us to reduce the size of an overlap representation when it has a set that is uniform except at one vertex.

Lemma 2.9. *Let v be a vertex in a graph G such that $N(v)$ is independent and contains no leaves. Let f be an overlap representation of G, and let f' be its restriction to $G - v$. If $\{a, b\}$ is f'-uniform, then $f - \{a\}$ or $f - \{b\}$ is an overlap representation of G.*

Proof. By Observation 2.8, the claim follows unless exactly one element of $\{a, b\}$ is in $f(v)$. Hence we may assume that $a \notin f(v)$ and $b \in f(v)$.

   Suppose that $f - \{a\}$ is not an overlap representation of G. Since Observation 2.8 implies that $f - \{a\}$ is an overlap representation of $G - v$, Lemma 2.6 implies that some edge incident to v is lost when a is deleted from f. Let vw_1 be such an edge. Because $a \notin f(v)$, we conclude that $f(w_1) - f(v) = \{a\}$. Because $\{a, b\}$ is f'-uniform, also $b \in f(w_1)$.

   If $f - \{b\}$ also is not a representation, then deleting b also destroys some edge vw_2 incident to v. Since $b \in f(v)$, either $f(v) \cap f(w_2) = \{b\}$ or $f(v) - f(w_2) = \{b\}$. We obtain a contradiction from each case. Note first that since each w_i has a neighbor other than v, and $\{a, b\}$ is f'-uniform, each $f(w_i)$ contains an element outside $\{a, b\}$.

6

Case 1: $f(v) \cap f(w_2) = \{b\}$. Since $\{a, b\}$ is f'-uniform, also $a \in f(w_2)$. Thus $\{a, b\} \subseteq f(w_1) \cap f(w_2)$. Since $w_1 w_2 \notin E(G)$, the sets $f(w_1)$ and $f(w_2)$ cannot each have an element outside the other. However, each has an element outside $\{a, b\}$, so they share another element c. Now $f(w_1) - f(v) = \{a\}$ yields $c \in f(v)$, while $f(v) \cap f(w_2) = \{b\}$ yields $c \notin f(v)$.

Case 2: $f(v) - f(w_2) = \{b\}$. Since $\{a, b\}$ is f'-uniform, $f(w_2) \cap \{a, b\} = \varnothing$. Since $f(w_1) - f(v) = \{a\}$, the guaranteed element c in $f(w_1) - \{a, b\}$ must lie in $f(v)$. Since $f(v) - f(w_2) = \{b\}$, also $c \in f(w_2)$. Meanwhile, $f(v) \leftrightarrow f(w_2)$ requires an element d in $f(w_2) - f(v)$. Since $f(w_1) - f(v) = \{a\}$, we have $d \notin f(w_1)$. Now $f(w_1)$ and $f(w_2)$ share c and overlap, contradicting $w_1 w_2 \notin E(G)$. $\qquad\square$

In a skeleton, the neighbor of a leaf vertex has degree 2.

Definition 2.10. In an overlap representation f of a skeleton T, a leaf l is *doubly-minimal* if both l and the neighbor of l are minimal in f.

Lemma 2.11. *In an overlap representation f of any skeleton T, there is at most one non-minimal leaf. If $T \neq P_4$, then there is a doubly-minimal leaf.*

Proof. If $T = P_3$, then there is only one non-edge, so only for the two leaves can one set properly contain another. Thus a leaf and the center are minimal in f. Henceforth assume $T \neq P_3$. In a skeleton other than P_3, no two neighborhoods are equal. Hence also no two assigned sets are equal.

Since the neighbor of any leaf x has degree 2, $T - N[x]$ is connected. If x is nonminimal, then by Lemma 1.1 $f(x)$ properly contains the sets assigned to all vertices other than its neighbor, including the other leaves. Hence only one leaf can be nonminimal.

Let A be the set of neighbors of minimal leaves; we have shown that $A \neq \varnothing$. Choose $v \in A$ such that $f(v)$ is minimal in $\{f(y) \colon y \in A\}$. Let $N(v) = \{x, u\}$, with x being the leaf. Since $T \neq P_3$, u is not a leaf.

When $T = P_4$, the claim fails when the sets in f are ab, $bceg$, $abde$, and eg.

If $T \neq P_4$, then u has no leaf neighbor, and each component of $T - N[v]$ is nontrivial. If x is not doubly-minimal, then $f(v)$ properly contains the sets for all vertices in some component T' of $T - N[v]$, by Lemma 1.1. Let x' be a leaf of T contained in T', and let v' be its neighbor, also in T'. Since $f(v)$ contains $f(v')$, the choice of v from A requires x' to be nonminimal. As observed earlier, this yields $f(v) \subset f(x')$, contradicting $f(v) \supset f(x')$. $\qquad\square$

Theorem 2.12. *If T is a skeleton with n vertices, where $n \geq 3$, then $\varphi(T) \geq n$.*

Proof. We first note a tool that allows us to apply Lemma 2.9 when a leaf is a minimal vertex. Recall that $N[v]$ denotes $N(v) \cup \{v\}$.

(*) If x is a minimal vertex in an overlap representation f of a graph G, and f' is the restriction of f to $G - N[x]$, then $f(x)$ is f'-uniform.

To prove (*), note that $v \in V(G) - N[x]$ implies $xv \notin E(G)$. Hence $f(x) \| f(v)$ or $f(x) \subseteq f(v)$ or $f(x) \supset f(v)$. Minimality of x excludes the last, so $f(v)$ contains all or none of $f(x)$.

We prove the lower bound on $\varphi(T)$ by induction on n. Since P_3 has an edge, $\varphi(P_3) \geq 3$, so we may assume $n \geq 4$. Let f be an optimal overlap representation of T; Lemma 2.11 yields a leaf x of T that is minimal in f. Let v be the neighbor of x, and let u be the other neighbor of v; note that $d(u) \geq 2$. Since $xv \in E(G)$, there exist $a \in f(x) - f(v)$, $b \in f(x) \cap f(v)$, and $c \in f(v) - f(x)$.

Let $T' = T - x$ and $T'' = T - x - v$. Let f' and f'' be the restrictions of f to T' and T'', respectively. We consider two cases, depending upon $d(u)$.

If $d(u) = 2$, then T' is a skeleton. Since x is minimal, (*) implies that $f(x)$ (and therefore $\{a, b\}$) is f''-uniform. Since $d(u) = 2$, the neighborhood of v in T' is independent and contains no leaves. Hence Lemma 2.9 applies, and $f' - \{a\}$ or $f' - \{b\}$ is an overlap representation of T'. By the induction hypothesis, $|f'| \geq n - 1$, so $|f| \geq n$.

If $d(u) > 2$, then $T \neq P_4$. Now Lemma 2.11 allows us to choose x to be doubly-minimal in f. Since $d(u) \geq 3$, deleting x and v from T does not create any new leaves, so T'' is a skeleton. Since x is minimal, (*) implies that $f(x)$ (and therefore $\{a, b\}$) is f''-uniform. Thus $f'' - \{a\}$ is an overlap representation of T''. Let $g = f'' - \{a\}$.

Since x is doubly-minimal, v is minimal in f, and thus (*) implies that $f(v)$ (and therefore $\{b, c\}$) is g-uniform. We now apply Lemma 2.9 to the vertex u, graph T'', and overlap representation g of T''. Let g' be the restriction of g to $T'' - u$. Since $\{b, c\}$ is g'-uniform, and $d(u) \geq 3$ implies that u has no leaf neighbors in the skeleton T'', Lemma 2.9 implies that $g - \{b\}$ or $g - \{c\}$ is an overlap representation of T''. By the induction hypothesis, $|g| \geq n - 1$ and $|f| \geq n$. $\qquad\square$

We have proved the following conclusion.

Theorem 2.13. *If T is a tree, then $\varphi(T)$ is the number of vertices in the skeleton of T. Furthermore, there is a linear-time algorithm to produce an optimal overlap representation.*

3   Bounds from the Number of Edges

As mentioned earlier, Erdős, Goodman, and Pósa [3] observed that finite intersection representations of a graph G correspond to families of complete subgraphs covering $E(G)$. In cases where the intersection representation arising from a decomposition into complete subgraphs is also an overlap representation, its size must be at least the pure overlap number $\Phi(G)$, defined in Section 1. On the other hand, always $\varphi(G) \leq \Phi(G)$.

Several upper bounds will be used repeatedly in the remainder of the paper; we give them names to improve readability. A *decomposition* of a graph G is a family $\mathcal{F}$ of pairwise edge-disjoint subgraphs whose union is G.

Lemma 3.1 (Decomposition Bound). *Let $\mathcal{F}$ be a decomposition of a graph G into complete subgraphs of order at most k, where $k \geq 2$. If $\delta(G) \geq k$, then $\Phi(G) \leq |\mathcal{F}|$. In particular, $\delta(G) \geq 2$ implies $\Phi(G) \leq |E(G)|$.*

Proof. For each $v \in V(G)$, let $f(v)$ be the set of all members of $\mathcal{F}$ that contain v. Each edge lies in some member of $\mathcal{F}$, so f is an intersection representation.

A vertex has at most $k-1$ neighbors in a complete subgraph of order k. Since $\delta(G) \geq k$, each $|f(v)|$ is at least 2. Each edge is covered only once, so $|f(v) \cap f(u)| \leq 1$. Hence no assigned set contains another, and f is a pure overlap representation. $\qquad\square$

Corollary 3.2. *If G is triangle-free, then $\Phi(G) \geq |E(G)|$, with equality when $\delta(G) \geq 2$.*

Proof. A pure overlap representation is also an intersection representation. $\qquad\square$

Lemma 3.1 provides an upper bound when $\delta(G) \geq 2$, and we will also apply it with $k = 3$ for decompositions into edges and triangles. Hence we want vertices of degree at most 2 to contribute little to $\Phi(G)$.

Lemma 3.3 (Deletion Bound). *If v is a vertex of degree at most 2 in a graph G with at least three vertices, then $\Phi(G) \leq \Phi(G - v) + 2$, with strict inequality when $d(v) = 0$.*

Proof. If $d(v) = 0$, then to avoid overlap and containment with all other assigned sets, we must assign v an element not assigned to any other vertex. Thus $\Phi(G) = \Phi(G - v) + 1$.

For $d(v) \in \{1, 2\}$, let f be an optimal pure overlap representation of $G - v$. Introduce new labels a and b. Let $f'(v) = \{a, b\}$. Let $f'(x) = f(x)$ for $x \notin N[v]$. For $x \in N(v)$, let $f'(x) = f(x) \cup \{c\}$, where c is one of the new labels, each used once when $d(v) = 2$.

Changing from f to f' creates no new intersections except to establish the edge(s) incident to v. Adding a new element to its neighbor(s) does not create containments, and there is no containment involving $f(v)$ and a set assigned to a neighbor, since the neighbors also receive an old label (even if isolated in $G - v$). $\qquad\square$

Next we discuss bounds on φ in terms of the number of edges. In contrast to $\Phi(G)$, generally $\varphi(G) < |E(G)|$ (though not for skeletons, as we have seen). An easy reduction allows us to forbid repeated vertex neighborhoods and isolated vertices.

Observation 3.4. *If a graph G has a vertex v such that $N(v)$ is empty or equals another vertex neighborhood, then $\varphi(G) = \varphi(G - v)$. In the first case, we extend an overlap representation f of $G - v$ by assigning v the set $\bigcup_{u \in V(G-v)} f(u)$. In the second case, we assign v the same set as w, where $N(v) = N(w)$.* $\qquad\square$

Let B_n denote the graph that is the union of $n-2$ triangles having a common edge; this graph is sometimes called the *n-book*.

Lemma 3.5. *The overlap number of the n-book B_n is 3.*

Proof. Since all the vertices besides the two vertices of degree $n-1$ have identical neighborhoods, Observation 3.4 allows us to remove them without changing the overlap number until we are left with a triangle, which has overlap number 3. □

Lemma 3.6 (Edge Bound). *Let G be an n-vertex graph other than the book B_n. If $\delta(G) \geq 2$ and $uv \in E(G)$, then G has an overlap representation f with size $|E(G)|-1$ such that neither $f(u)$ nor $f(v)$ is properly contained in the set assigned to any other vertex. In particular, $\varphi(G) \leq |E(G)| - 1$ when $\delta(G) \geq 2$ unless G is the 3-book K_3.*

Proof. We define an explicit representation using a label for each edge *other than uv*. For $w \notin \{u, v\}$, let $f(w)$ be the set of labels for edges incident to w. For $w \in \{u, v\}$, let $f(w)$ be the set of labels for edges *not* incident to w. The restriction of f to $G-u-v$ is a pure overlap representation of $G - u - v$ (labels for edges to u or v can establish non-containment).

By construction, $f(u) \supseteq f(w)$ when w is a nonneighbor of u. This establishes nonadjacency and prohibits $f(u)$ from proper containment in another assigned set. Similarly for v. (However, $f(u) = f(w)$ when $G = K_{2,n-2}$ and $\{u, w\}$ is a partite set of size 2.)

For $w \in N(u) - \{v\}$, the label for uw is in $f(w) - f(u)$. Since $d(w) \geq 2$, the label for some other edge incident to w lies in $f(u) \cap f(w)$. To establish $f(u) - f(w) \neq \varnothing$, it suffices to have an edge incident to neither w nor u. If every edge is incident to w or u, then $G = B_n$, and $\varphi(G) = 3$. The same argument applies to edges at v.

For the edge uv itself, $f(u) - f(v)$ contains the label for an edge other than uv incident to v. Similarly, $f(v) - f(u) \neq \varnothing$. To ensure that $f(u) \cap f(v) \neq \varnothing$, we need an edge incident to neither u nor v. As above, this exists unless $G = B_n$. □

In Theorem 3.10, we will prove equality in the upper bound $\varphi(G) \leq |E(G)| - 1$ for a special family of graphs with $\delta(G) \geq 2$. For this we will need a definition and two lemmas.

Definition 3.7. A *star-cutset* in a graph G is a separating set S containing a vertex x adjacent to all of $S - x$. If G has no star-cutset, then it is *star-cutset-free*.

Lemma 3.8. *If f is an overlap representation of a connected graph G with no star-cutset, then any two vertices that are not minimal in f are adjacent.*

Proof. Let u and v be such vertices. If $v \notin N(u)$, then v remains in $G - N[u]$. Since G has no star-cutset, $G - N[u]$ is connected. Since u is not minimal, $f(u)$ properly contains the set assigned to some vertex of $G - N[u]$. By Lemma 1.1, $f(u)$ properly contains $f(v)$. The same argument yields $f(v) \supset f(u)$, a contradiction. □

Lemma 3.9. *If G is an n-vertex triangle-free graph with no star-cutset, then G does not have distinct vertices with the same neighborhood, unless $G = K_{2,n-2}$ with $n \leq 4$.*

Proof. If $N(u) = N(v)$, then v is isolated in $G - N[u]$. Since G has no star-cutset, $G - N[u]$ contains only v. Thus $V(G) = \{u, v\} \cup N(u)$. Also $N(u)$ induces no edges, since G is triangle-free. Thus $G = K_{2,n-2}$. Also, $n - 2 \leq 2$, since otherwise deleting $N[x]$ for some $x \in N(u)$ disconnects G, contradicting the absence of star-cutsets. $\qquad \square$

Theorem 3.10. *If G is a triangle-free graph with no star-cutset, then $\varphi(G) \geq |E(G)| - 1$.*

Proof. Let f be an overlap representation of G. If two assigned sets are equal, then those vertices have the same neighborhood, and Lemma 3.9 yields $G = K_{2,n-2}$ with $n \leq 4$. By Observation 3.4, $\varphi(K_{2,0}) = 1$ and $\varphi(K_{2,1}) = \varphi(K_{2,2}) = 3$. In each case, $\varphi(G) \geq |E(G)| - 1$.

Hence we may assume that no two sets assigned by f are equal. Since G is triangle-free, by Lemma 3.8 at most two vertices are non-minimal. We consider three cases.

Case 0: Every vertex is minimal in f. In this case, f is a pure overlap representation, and Corollary 3.2 yields $|f| \geq |E(G)|$.

Case 1: One vertex u is nonminimal in f. Since $G - N[u]$ is connected and $f(u)$ contains some other assigned set, $f(u)$ contains all elements assigned to the nonneighbors of u. Also $N(u)$ is independent, so every edge of $G - u$ has an endpoint outside $N[u]$.

All containments involve $f(u)$. Hence f restricts to a pure overlap representation and thus an intersection representation on $G - u$. Since G is triangle-free, for each edge e of $G - u$ there is an element assigned by f to the endpoints of e. It also lies in $f(u)$, since e has an endpoint outside $N[u]$. Let S be this set of elements.

Since $S \subseteq f(u)$, we still must make $f(w) - f(u)$ nonempty for $w \in N(u)$. Since $N(u)$ is independent and u is the only nonminimal vertex, the sets assigned to $N(u)$ are pairwise disjoint. Hence $N(u)$ requires distinct additional elements, yielding $|f| \geq |E(G)|$.

Case 2: Two vertices, u and v, are nonminimal in f. By Lemma 3.8, $uv \in E(G)$. As above, f restricts to an intersection representation on $G - u - v$ with an element for each edge; let S be this set of elements. Since u is nonminimal and $G - N[u]$ is connected, $f(u)$ contains all elements assigned to vertices outside $N[u]$.

As above, each $w \in N(u) - \{v\}$ needs an element not in $f(u)$, and these elements are distinct since G is triangle-free. The same holds for $N(v)$. We thus have $|f| \geq |E(G)| - 1$ unless there exist $x \in N(u) - \{v\}$ and $y \in N(v) - \{u\}$ with $f(x)$ and $f(y)$ having a common element outside $f(u) \cup f(v)$. Since G is triangle-free, u and v have no common neighbors, so $f(u) \supset f(y)$ and $f(v) \supset f(x)$. Hence the elements establishing the edges between $\{u, v\}$ and their neighbors are distinct, and $|f| \geq |E(G)| - 1$. $\qquad \square$

The proof of Theorem 3.10 shows that for a connected triangle-free graph G with $\delta(G) \geq 2$ and no star-cutset, the *only* way to form an overlap representation with fewer than $|E(G)|$ el-

ements is as in the proof of the Edge Bound (Lemma 3.6). Theorem 3.10 and Observation 3.4 determine the overlap number for several constructions of interest.

Corollary 3.11. *For even n with $n \geq 6$, if G_n is the n-vertex graph obtained by deleting from $K_{n/2,n/2}$ a matching of size $n/2$, then $\varphi(G_n) = n^2/4 - n/2 - 1$. For odd n with $n \geq 7$, if G_n is obtained from G_{n-1} by adding one vertex with the same neighborhood as some vertex of G_{n-1}, then $\varphi(G_n) = \varphi(G_{n-1}) = \lfloor n^2/4 - n \rfloor$.* $\qquad\square$

Corollary 3.12. *If G is a triangle-free plane graph in which every face has length 4, and G has no star-cutset, then $\varphi(G) = 2n - 5$.* $\qquad\square$

In Section 5, we will show that the graphs of Corollary 3.11 are extremal when n is even and are within $n/2$ of the extreme when n is odd. In Section 4, we will show that graphs having the properties in Corollary 3.12 are extremal in the class of n-vertex planar graphs. We next give explicit examples of such graphs.

Example 3.13. Graphs as described in Corollary 3.12 exist for $n \geq 10$ (also for $n = 4$ and $n = 8$). When $n \equiv 0 \mod 4$, the cartesian product of $P_{n/4}$ and C_4 suffices. When $n \equiv 0 \mod 2$, one can start with an even cycle C and add a vertex inside adjacent to the even-indexed vertices on C and a vertex outside adjacent to the odd-indexed vertices on C.

For odd n, take such a graph G with $n-1$ vertices embedded in the plane, let x be a vertex of degree at least 4 in G (x exists if $|V(G)| \geq 10$), and let u and v be nonconsecutive neighbors of x in the embedding. Form G' by replacing x with nonadjacent vertices x' and x'' whose neighborhoods in the new graph G' partition $N_G(x)$, except that $N_{G'}(x') \cap N_{G'}(x'') = \{u, v\}$. The vertices $\{x', u, x'', v\}$ form a new face surrounding the former edges xu and xv, and the other edges at x attach instead to x' and x''. $\qquad\square$

As we did with pure overlap number, for overlap number we will want to accommodate vertices of degree less than 2 without much cost. By Observation 3.4, we may assume that there are no isolated vertices and that vertex neighborhoods are distinct.

The corresponding result for vertices of degree 1 is a special case of a more general result (proved in the same way) that permits saving labels when overlap representations of subgraphs are combined at a cut-vertex. For clarity, we present only the result that we use to obtain our extremal results in the subsequent sections.

Lemma 3.14. *If v is a leaf in a graph G and $G - v$ is nontrivial, then $\varphi(G) \leq \varphi(G - v) + 2$.*

Proof. Let u be the neighbor of v. If uv is an isolated edge and $G - v$ is nontrivial, then let $R = \bigcup_{x \in V(G - \{u,v\})} f(x)$, where f is an overlap representation of $G - \{u, v\}$. Note that $|R| \geq 1$. Modify f by assigning $R \cup \{a\}$ to u and $R \cup \{b\}$ to v, where $a, b \notin R$. This produces an overlap representation of G, so $\varphi(G) \leq \varphi(G - \{u, v\}) + 2 \leq \varphi(G - v) + 2$.

Hence we may assume that $d_G(u) \geq 2$. Let f be an optimal overlap representation of $G - v$. Let $W = V(G) - \{u, v\}$. Define f' on $V(G)$ as follows. Let $f'(v) = S = \{a, b\}$, where $a, b \notin \bigcup_{x \in W} f(x)$. Let $f'(u) = f(u) \cup \{b\}$. For $x \in W$, let $f'(x) = f(x) \cup S$ if $f(x) \supseteq f(u)$; otherwise, let $f'(x) = f(x)$. Note that $|f'| = |f| + 2$.

We check that f' is an overlap representation of G. Since $|f(u)| \geq 2$, we have $f'(u) \leftrightarrow f'(v)$. For $x \in W$, we have $f'(x) \| f'(v)$ or $f'(x) \supseteq f'(v)$, depending on whether $f'(x)$ acquires S, so v receives no other edges.

For $x, y \in W$, the assigned sets acquire S if and only if they contain $f(u)$. By Observation 2.3, the relation between $f'(x)$ and $f'(y)$ is the same as between $f(x)$ and $f(y)$. If $f'(x) = f(x)$, then $f(x) \not\supseteq f(u)$, and the relation between $f(x)$ and $f(u)$ is preserved. If $f(x) \supseteq f(u)$, then $f'(x) = f(x) \cup S$ and again the relation is preserved. $\qquad \square$

4 Overlap Number of Planar Graphs

In order to apply the Decomposition Bound for planar graphs that may contain triangles, we need an efficient decomposition into small complete subgraphs. By Euler's Formula, a triangle-free planar graph has as most $2n - 4$ edges, with equality only if every face is a 4-cycle.

Lemma 4.1. *If G is an n-vertex plane graph, and $n \geq 3$, then G decomposes into at most $2n - 5$ edges and facial triangles unless:*
(a) every face is a 4-cycle, in which case G decomposes into $2n - 4$ edges, or
(b) G is K_4, which decomposes into three edges and one facial triangle.

Proof. We use induction on the number of facial triangles in G. If there are none, then Euler's Formula suffices. If G has a facial triangle T, then form a plane graph G' from G by deleting $E(T)$ and introducing a new vertex v adjacent to $V(T)$. Since v belongs to no triangle, G' has fewer facial triangles than G.

Suppose first that G' has a facial cycle that is not a 4-cycle. By the induction hypothesis, G' decomposes into at most $2n - 3$ triangles and edges (G' has $n + 1$ vertices). Since v is in no triangle, the three edges incident to v are edges in the decomposition. Replacing them with T yields the desired decomposition of G.

If every face in G' is a 4-cycle, then each edge of T lies in another facial triangle in G. By the induction hypothesis, G' decomposes into $2n - 2$ edges. If the faces incident to v in G' have no shared edges not incident to v, then their nine edges G' can be replaced with three triangles to decompose G into $2n - 8$ edges and facial triangles (Figure 1a).

If two of these faces share an edge, then the eight distinct edges can be replaced with two triangles and two edges to decompose G into $2n - 6$ edges and facial triangles (Figure 1b).
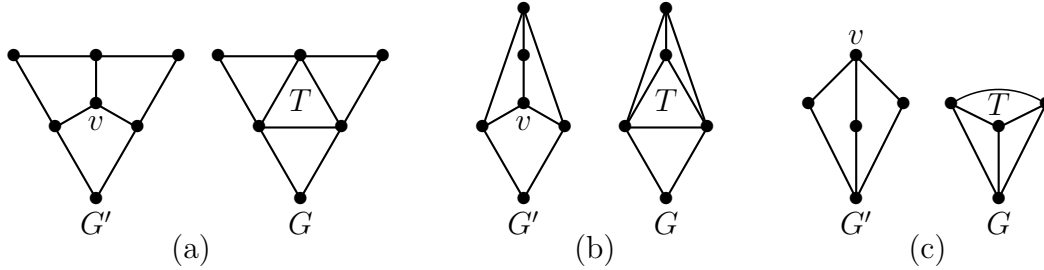
13

Figure 1: Three cases for G' and G in Lemma 4.1.

Finally, the additional edges may be shared in pairs (Figure 1c). Now the component of G' containing v is $K_{2,3}$, and the component of G containing T is K_4. Form G'' from G by deleting this component; G'' has fewer facial triangles than G. If G'' has at least three vertices, then it decomposes into at most $2n - 12$ edges and facial triangles, yielding a decomposition of size at most $2n - 8$ for G. If G'' has at most two vertices, then it is K_1, $2K_1$ or K_2; in each case, G decomposes into at most $2n - 6$ edges and facial triangles. $\square$

Corollary 4.2. *If G is an n-vertex plane graph with $\delta(G) \geq 3$, then $\Phi(G) \leq 2n - 4$, with equality only if every facial cycle is a 4-cycle or $G = K_4$. The same holds for $\theta_1(G)$.*

Proof. Immediate from Lemma 4.1 and the Decomposition Bound (Lemma 3.1). $\square$

In many cases, we will obtain bounds on $\varphi(G)$ from bounds on $\Phi(G)$. The next several remarks and computations facilitate characterization of the n-vertex planar graphs and the n-vertex graphs with largest pure overlap number.

Observation 4.3. With the convention that $\Phi(K_1) = 1$, pure overlap number and intersection number are additive under disjoint union; in particular, $\Phi(2K_2) = 6$. Overlap number is not additive under disjoint union: $\varphi(K_2) = 3$, but $\varphi(2K_2) = 5$. $\square$

We observed in Corollary 3.2 that $\Phi(G) = |E(G)|$ when G is triangle-free and $\delta(G) = 2$. When $\delta(G) = 1$, more labels may be needed.

Proposition 4.4. *If $n \geq 2$, then $\Phi(P_n) = n + 1$. If $n \geq 3$, then $\Phi(C_n) = n$.*

Proof. A representation using the sets $\{i, i + 1\}$ for $1 \leq i \leq n$ provides the upper bound. Since a pure overlap representation is an intersection representation and P_n is triangle-free, each label is used at most twice. The endpoints of an edge have a common label used at no other vertex. Also, each endpoint of P_n has a label used nowhere else. Hence $|f|$ must exceed the number of edges by 2.

If $n \geq 4$, then C_n is triangle-free and $\delta(C_n) = 2$, so $\Phi(C_n) = |E(G)| = n$. For C_3, using the three 2-sets in $\{1, 2, 3\}$ yields $\Phi(C_n) = n$ again. $\square$

14

Proposition 4.5. *If $m \geq 2$, then $\Phi(K_{1,m}) = 2m$. If $G - v = K_{1,m}$ and $d_G(v) \leq 2$, then $\Phi(G) \leq 2m + 1$.*

Proof. Since a pure overlap representation is an intersection representation, the sets for leaves of $K_{1,m}$ are disjoint. They must overlap the central set, so each has size at least 2. Equality holds using two labels at each leaf and putting one label from each leaf in the central set.

If $G - v = K_{1,m}$ and $G \neq K_{1,m+1}$, then G is obtained from C_3, C_4, or P_4 by appending leaves at one vertex of degree 2. By the Deletion Bound (Lemma 3.3), each leaf costs at most two new labels; combined with Proposition 4.4, this yields $\Phi(G) \leq 2m + 1$. $\qquad\square$

A graph G is *k-degenerate* if every subgraph of G has a vertex of degree at most k.

Theorem 4.6. *If G is an n-vertex planar graph with $n \geq 3$, then $\Phi(G) \leq 2n - 2$, with equality if and only if $G \in \{K_{1,n-1}, 2K_2, K_2 + K_1\}$. Furthermore, if G is not 2-degenerate, then $\Phi(G) \leq 2n - 4$, with equality when $\delta(G) \geq 2$ only if G has $2n - 4$ edges or is K_4.*

Proof. Observation 4.3 and Propositions 4.4–4.5 take care of the case $n = 3$ and confirm equality for $K_{1,n-1}$. This provides a basis for induction. Suppose that $n \geq 4$. If $\delta(G) \geq 3$, then Corollary 4.2 yields $\Phi(G) \leq 2n-4$, with equality only when G has $2n-4$ edges or is K_4.

Hence we may assume that $\delta(G) \leq 2$. Let v be a vertex of minimum degree. By the Deletion Bound and the induction hypothesis, $\Phi(G) \leq \Phi(G - v) + 2 \leq 2n - 2$. Equality requires $\Phi(G - v) = 2n - 4$ and $d(v) \in \{1, 2\}$ and $G - v \in \{K_{1,n-2}, 2K_2, K_2 + K_1\}$.

If $G - v = K_2 + K_1$, then $G \in \{2K_2, P_3 + P_1, C_3 + P_1, P_4\}$, and $\Phi(G)$ is $6, 5, 4, 5$, respectively, using Observation 4.3 and Proposition 4.4. If $G - v = 2K_2$, then $G \in \{P_3 + P_2, P_5, C_3 + P_2\}$, and $\Phi(G)$ is $7, 6, 6$, respectively, using the same facts. If $G - v = K_{1,n-2}$ and $G \neq K_{1,n-1}$, then Proposition 4.5 states that $\Phi(G) \leq 2n - 3$.

For the final statement, suppose that G is not 2-degenerate. Now $G - v$ has a subgraph with minimum degree at least 3. Hence $G - v$ is not 2-degenerate, and the bound improves to $\Phi(G) \leq \Phi(G - v) + 2 \leq 2n - 4$, with equality only if $\Phi(G - v) = 2n - 6$.

When also $\delta(G) = 2$, we have $G \neq K_4$ and need to show $|E(G)| = 2n - 4$. If $\delta(G - v) \geq 2$, then by the induction hypothesis, $G - v$ has $2n - 6$ edges or is K_4. In the former case $|E(G)| = 2n - 4$, since $d(v) = \delta(G) = 2$. In the latter case, $\Phi(G) \leq 5 = 2n - 5$ by using the five sets $123, 14, 24, 34, 15$.

Since $\delta(G) = 2$ prohibits $\delta(G - v) = 0$, the remaining case is $\delta(G - v) = 1$, and a leaf u in $G - v$ must be a neighbor of v in G. Consider a pure overlap representation f of $G - v$ using $2n - 6$ elements; since u is a leaf, $f(u)$ must have an element a assigned to no other vertex. Assign a and a new element b to v, and add b to the set for the other neighbor of v. Since no assigned set contains another, this use of b causes no trouble. We have extended f to a pure overlap representation of G with only one new label, so $\Phi(G) = 2n - 5$. $\qquad\square$

Corollary 4.7. *If G is a planar n-vertex graph with $n \geq 3$, then $\varphi(G) \leq 2n - 2$.*

The upper bounds on pure overlap number simplify some cases in proving the best upper bound on overlap number, which in general is smaller by 1.

Theorem 4.8. *If G is a planar n-vertex graph and $n \geq 5$, then $\varphi(G) \leq 2n - 5$, which is sharp for $n = 8$ and $n \geq 10$.*

Proof. Example 3.13 establishes sharpness for $n = 8$ and $n \geq 10$ (and $n = 4$). To prove the bound, we use induction on n, postponing the base case to Proposition 4.9 below. For $n > 5$, we may assume $\delta(G) \geq 2$ by Observation 3.4, Lemma 3.14, and the induction hypothesis.

If G is 2-degenerate, then $|E(G)| \leq 2n - 3$. If $|E(G)| < 2n - 3$, then the Edge Bound (Lemma 3.6) yields $\varphi(G) \leq 2n - 5$, so we may assume equality. By Euler's Formula, G contains a triangle T. If every vertex of T has a neighbor outside T, then each vertex of G is incident to at least two subgraphs in the decomposition of G consisting of T and $2n - 6$ individual edges. The Decomposition Bound (Lemma 3.1) now yields $\Phi(G) \leq 2n - 5$, and hence $\varphi(G) \leq 2n - 5$.

If $|E(G)| = 2n - 3$ and every triangle has a vertex v with $d_G(v) = 2$, then $G - v$ has the same property, and by induction G is the book B_n and $\varphi(G) = 3$ (Lemma 3.5).

In the remaining case, G is not 2-degenerate and $\delta(G) \geq 2$. By Theorem 4.6, either $\varphi(G) \leq \Phi(G) \leq 2n - 5$, or $|E(G)| = 2n - 4$ and the Edge Bound applies. $\square$

When the complement of a graph G is edge-transitive, G^+ denotes the graph obtained by adding any edge of the complement to G.

Proposition 4.9. *If G is an n-vertex graph, where $n \in \{4, 5\}$, then $\varphi(G) \leq 2n - 5$, except that $\varphi(G) = 4$ for $G \in \{P_4, K_4, K_{1,3}^+\}$.*

Proof. If G is a forest, then Theorem 2.13 suffices. Note also that $\varphi(K_3) = 3$, and we may assume that G has no isolated vertex or repeated neighborhood, by Observation 3.4.

If $n = 4$ and G is not a forest and has no isolated vertex, then $G \in \{C_4, C_4^+, K_{1,3}^+, K_4\}$. Each graph has an edge, so $\varphi(G) \geq 3$. For C_4 and C_4^+, the repeated neighborhoods let three elements suffice. For K_4, we need an intersecting family of four incomparable sets, which does not exist in $\{1, 2, 3\}$, but $\{123, 41, 42, 43\}$ suffices. For $K_{1,3}^+$, the triangle can only be represented in subsets of $\{1, 2, 3\}$ using $\{12, 23, 13\}$, and no fourth subset intersects just one of these. Hence four elements are needed, and $\{123, 124, 13, 23\}$ is an overlap representation.

For $n = 5$, if G has a vertex v of degree 1 such that $\varphi(G - v) \leq 3$, then Lemma 3.14 applies. With no repeated neighborhood, a vertex of degree 1 now restricts G to be K_4 plus one pendant edge, K_3 plus pendant edges at two distinct vertices, or K_3 plus a pendant path

16

of length two at one vertex. These three graphs are represented by $\{145, 245, 345, 1234, 45\}$, $\{12, 23, 34, 45, 1245\}$, and $\{12, 23, 34, 45, 1235\}$, respectively.

We are left with $n = 5$ and $\delta(G) \geq 2$. By Lemma 3.6, we may assume that $|E(G)| \geq 7$. The remaining 5-vertex graphs with at least seven edges are listed below with overlap representations ("+" denotes disjoint union).

$$
\begin{array}{llll}
K_5 & \{123, 234, 345, 451, 512\} & K_{2,2,1} & \{12, 34, 14, 23, 13\} \\
\overline{P_2 + 3K_1} & \{123, 234, 345, 14, 25\} & K_{3,1,1} & \{12, 34, 1234, 513, 524\} \qquad \square \\
\overline{P_3 + 2K_1} & \{123, 345, 14, 25, 1245\} & \overline{P_4 + K_1} & \{12, 23, 34, 45, 135\}
\end{array}
$$

5 Extremal values for n-vertex graphs

In this section we study the maximum values for $\Phi(G)$ and $\varphi(G)$ over n-vertex graphs. As usual, the problem is easier for $\Phi(G)$, and solving it simplifies the analysis for $\varphi(G)$. In addition to our earlier computations, we need one more special family.

Proposition 5.1. *If* $n \geq 1$ *and* $\binom{2k-1}{k} \geq n$, *then* $\Phi(K_n) \leq 2k - 1$.

Proof. The k-element subsets of a $(2k-1)$-set are pairwise intersecting and incomparable. $\square$

Lemma 5.2. *Let* T *be the vertex set of a triangle in an* n-vertex graph G. *If* n' *is the number of vertices of degree 1 in* $G - T$, *then* $\Phi(G) \leq \Phi(G') + n - n'$.

Proof. Let f be an optimal pure overlap representation of $G - T$. Add three new labels, two assigned to each vertex, to represent the triangle. Consider each vertex x outside T. If $d(x) = 1$, then $f(x)$ has at least two labels, with one appearing on no other vertex; add that label to the sets for the neighbors of x in T. If $d(x) \neq 1$, then introduce a new label assigned to x and its neighbors in T. The total number of labels used is at most $\Phi(G - T) + 3 + n - 3 - n'$. $\square$

Theorem 5.3. *Let* G *be an* n-vertex graph. *If* $3 \leq n \leq 5$, *then* $\Phi(G) \leq 2n - 3$ *unless* $G \in \{K_{1,n-1}, 2K_2, K_2 + K_1\}$. *If* $n \geq 6$ *and* $G \neq K_{1,5}$, *then* $\Phi(G) \leq \lfloor n^2/4 \rfloor$. *If* $n \geq 7$, *then equality holds only when* $G = K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Proof. The first statement was proved in Theorem 4.6 except for the nonplanar graph K_5, and $\Phi(K_5) \leq 5$, by Proposition 5.1. For $n \geq 6$, we proceed inductively.

Case 1: There exists $v \in V(G)$ *with* $d(v) \leq 2$. By the Deletion Bound (Lemma 3.3), $\Phi(G) \leq \Phi(G - v) + 2$. If $n = 6$, then $\Phi(G - v) \leq 7$ unless $G - v = K_{1,4}$; in either case, $\Phi(G) \leq 9 = 6^2/4$ (Proposition 4.5). If $n \geq 7$, then $\Phi(G) \leq \lfloor (n-1)^2/4 \rfloor + 2 < \lfloor n^2/4 \rfloor$ unless $G - v = K_{1,5}$, in which case $\Phi(G) \leq 11 < \lfloor 7^2/4 \rfloor$ (Proposition 4.5).

Case 2: G is triangle-free and $\delta(G) \geq 3$. By the Decomposition Bound and Corollary 3.2 and the well-known fact that $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$ is the unique triangle-free n-vertex graph with the most edges, $\Phi(G) \leq \lfloor n^2/4 \rfloor$, with equality only for $K_{\lfloor n/2 \rfloor, \lceil n/2 \rceil}$.

Case 3: G has a triangle and $\delta(G) \geq 3$. Let T be the vertex set of a triangle. By Lemma 5.2, $\Phi(G) \leq \Phi(G - T) + n - n'$, where n' is the number of vertices of degree 1 in $G - T$. For $n \geq 9$, we have $\Phi(G) \leq \lfloor (n-3)^2/4 \rfloor + n < \lfloor n^2/4 \rfloor$ unless $G - T = K_{1,5}$, in which case $n' = 5$ and $\Phi(G) \leq 10 + 1 = 11$.

For $6 \leq n \leq 8$, we have $\Phi(G - T) + n \leq 2n - 9 + n \leq \lfloor n^2/4 \rfloor$ unless $G - T \in \{K_{1,n-1}, 2K_2, K_2 + K_1\}$. In those cases, $\Phi(G - T) = 2n - 8$ and $n' \geq 2$, so $\Phi(G - T) < \lfloor n^2/4 \rfloor$. In the remaining cases with $n = 8$, we have $\Phi(G) \leq 3n - 9 < \lfloor n^2/4 \rfloor$.

If $n = 7$, then $\lfloor n^2/4 \rfloor = 12$, so it remains only to prove $\Phi(G) \leq 11$ when $\Phi(G - T) = 2n - 9 = 5$ and $\delta(G - T) \geq 2$. From $\delta(G - T) = 2$, we have $G - T \in \{C_4, C_4^+, K_4\}$. We have shown $\Phi(C_4) = 4$, and also $\Phi(C_4^+) = 4$ by adding a set with one new label and one old label to the pure overlap representation of C_3 using three labels. Hence only $G - T = K_4$ remains. Since this must hold for every triangle in G, we have $G = K_7$, but $\Phi(K_7) = 5$. $\square$

It remains to study the maximum of $\varphi(G)$ over n-vertex graphs. Rosgen [6] showed $\varphi(G) \leq n^2/4$. In Corollary 3.11, we constructed for even n with $n \geq 6$ a graph G_n with $\varphi(G) = n^2/4 - n/2 - 1$. By improving the upper bound from $n^2/4$ to $n^2/4 - n/2 - 1$ for $n \geq 14$, we show that this construction is extremal and that the construction in Corollary 3.11 for odd n is within $n/2$ of the extreme. We consider the main cases in separate lemmas: bipartite graphs, triangle-free non-bipartite graphs, and graphs containing a triangle.

Lemma 5.4. *Let G be an n-vertex bipartite graph in which no two vertices have the same neighborhood. If $n \geq 7$ and $\delta(G) \geq 2$, then $\varphi(G) \leq n^2/4 - n/2 - 1$.*

Proof. By the Edge Bound (Lemma 3.6), $\varphi(G) \leq |E(G)| - 1$, so we may assume that $|E(G)| > n^2/4 - n/2$. Let X and Y be the partite sets, with $k = |X| \leq |Y|$. If some neighborhood is repeated, then by Observation 3.4 we can reduce the problem to smaller n. Otherwise, at most one vertex of Y has degree k, and vertices of degree $k - 1$ have distinct nonneighbors in X. Summing the vertex degrees in Y then yields $|E(G)| \leq (k-1)(n-k)+1$. If $k \leq (n - 1)/2$, then the bound is at most $n^2/4 - n - 1/2$. Hence we may assume that $k = n/2$ and $n \geq 8$. Now $|E(G)| \leq n^2/4 - n/2 + 1$, so equality holds, and G arises from $K_{n/2,n/2}$ by deleting a matching of size $n/2 - 1$.

Let y be the vertex in Y having degree $n/2$, and let $G' = G - y$. Since $n \geq 8$, we have $\delta(G') \geq 2$. By the Decomposition Bound (Lemma 3.1), $\Phi(G') \leq |E(G')| = n^2/4 - n + 1$. Choose $y' \in Y - \{y\}$, and let x' be the nonneighbor of y' in X. Let f' be a pure overlap representation of G' using one label for each edge. Define f as follows: Put $f(y) = f'(y') \cup \{a\}$, where a is a new label, and let $f(x') = f'(x') \cup \{a\}$. For $v \notin \{y, x'\}$, let $f(v) = f'(v)$.

Since the only vertex of G' receiving a is x', overlaps and disjointness are preserved within the sets assigned to $V(G')$ Hence it suffices to check pairs involving y. We have $f(y) \supset f(y')$ and $f(y) \| f(z)$ for $z \in Y - \{y'\}$. For $x \in X$, we have $f(y) \leftrightarrow f(x)$. Thus, f is an overlap representation of G with $n^2/4 - n + 2$ labels. Since $n \geq 8$, the desired bound holds. $\square$

Lemma 5.5. *If G is an n-vertex bipartite graph, then $\varphi(G) \leq \max\{2n, n^2/4 - n/2 - 1\}$.*

Proof. The claim is $\varphi(G) \leq 2n$ for $n \leq 10$ and $\varphi(G) \leq n^2/4 - n/2 - 1$ for $n > 10$. Since $\varphi(G) \leq \Phi(G)$, Theorem 5.3 implies the claim when $n \leq 8$, using $2n - 2 \leq 2n$ always and $\lfloor n^2/4 \rfloor \leq 2n$ for $n \leq 8$.

We proceed inductively. The desired bound exceeds the desired bound for $(n-1)$-vertex graphs by at least 2. Consider $v, w \in V(G)$. If $d(v) = 0$, then $\varphi(G) \leq \varphi(G - v) + 1$. If $d(v) = 1$, then $\varphi(G) \leq \varphi(G-v)+2$ (by Lemma 3.14). If $N(v) = N(w)$, then Observation 3.4 yields $\varphi(G) = \varphi(G - v)$. If $\delta(G) \geq 2$ and no two vertices have the same neighborhood, then Lemma 5.4 yields $\varphi(G) \leq n^2/4 - n/2 - 1$. Thus the bound holds in all cases. $\square$

Lemma 5.6. *If G is an n-vertex triangle-free graph that is not bipartite, then $\varphi(G) \leq \max\{2n + 7, n^2/4 - n/2 - 1\}$.*

Proof. The claim is $\varphi(G) \leq 2n + 7$ for $n \leq 12$ and $\varphi(G) \leq n^2/4 - n/2 - 1$ for $n > 12$. Theorem 5.3 suffices when $n \leq 10$. As above, we proceed inductively; the desired bound increases by at least 2 per step, and we may assume that $\delta(G) \geq 2$ and that neighborhoods are distinct. By the Edge Bound (Lemma 3.6), $\varphi(G) \leq |E(G)| - 1$, so we seek $|E(G)| \leq \max\{2n + 8, n^2/4 - n/2\}$ for a triangle-free graph G with no repeated neighborhood.

Let C be a shortest odd cycle in G, with length $2k + 1$, and let $G' = G - V(C)$. Since C has no chords, $V(C)$ induces $2k+1$ edges. Since G has no triangle, each vertex not on C has at most k neighbors on C. Since G' is triangle-free, $|E(G')| \leq (n - 2k - 1)^2/4$. Summing the bounds $2k+1$, $k(n-2k-1)$, and $(n-2k-1)^2/4$ yields $|E(G)| \leq n^2/4 - n/2 - (k^2 - 2k - 5/4)$. If $k \geq 3$, then $k^2 - 2k > 5/4$, so we may assume $k = 2$.

If G' is not bipartite, then let C' be a shortest odd cycle in G', with length $2l + 1$. With $|V(G') - V(C')| = n - 2l - 6$, we have

$$|E(G)| \leq 5 + 2(n-5) + (2l+1) + l(n-2l-6) + (n-2l-6)^2/4 \leq n^2/4 - n + 5 - l(l-2).$$

Since $l \geq 2$, and $n - 5 \geq n/2$ when $n \geq 10$, this is small enough.

Finally, suppose that G' is bipartite. Since $k = 2$, we have $|E(G)| \leq n^2/4 - n/2 + 5/4$. Call a vertex of G' *full* if it has two (nonadjacent) neighbors in C and is adjacent to all vertices in the other partite set of G'. Each pair of nonadjacent vertices in C is adjacent to at most one full vertex, since otherwise G has a triangle or a repeated neighborhood. Thus, at most five vertices of G' are full, so at least $(n - 5) - 5$ vertices are not. This yields $|E(G)| \leq n^2/4 - n/2 + 5/4 - (n - 10)/2$, which suffices. $\square$

19

Lemma 5.7. *If G is an n-vertex graph with a triangle T, and $n \geq 14$, then $\varphi(G) \leq n^2/4 - n/2 - 1$.*

Proof. View T as a triple of pairwise adjacent vertices, and let $G' = G - T$. In most cases we show $\Phi(G) \leq n^2/4 - n/2 - 1$, which suffices.

Case 1: G' has a triangle T'. By Theorem 5.3, $\Phi(G' - T') \leq (n-6)^2/4$ when $n \geq 13$. By Lemma 5.2 (twice), $\Phi(G) \leq \Phi(G' - T') + 2n - 3$. If $n \geq 14$, then $(n-6)^2/4 + 2n - 3 \leq n^2/4 - n/2 - 1$.

Case 2: $\delta(G') \leq 1$. By Theorem 5.3, $\Phi(G' - v) \leq (n-4)^2/4$ when $n \geq 11$. By the Deletion Bound and then Lemma 5.2, $\Phi(G) \leq \Phi(G' - v) + n + 1$. Hence $\Phi(G) \leq n^2/4 - n + 5$. Since $n - 5 \geq n/2 + 1$ when $n \geq 12$, we have $\Phi(G) \leq n^2/4 - n/2 - 1$.

Case 3: $\delta(G') \geq 2$ and G' is triangle-free but not bipartite. The argument in the second paragraph of Lemma 5.6 yields $|E(G')| \leq (n-3)^2/4 - (n-3)/2 + 5/4 = n^2/4 - 2n + 5$. By the Decomposition Bound (Lemma 3.1), $\Phi(G') \leq |E(G')|$. By Lemma 5.2, $\Phi(G) \leq \Phi(G') + n \leq n^2/4 - n + 5$. As in Case 2, the claim holds when $n \geq 12$.

Case 4: $\delta(G') \geq 2$ and G' is bipartite. Suppose first that $T \subseteq N_G(v)$ for some $v \in V(G')$. Since K_4 has a pure overlap representation using $\{123, 41, 42, 43\}$, the method of Lemma 5.2 yields $\Phi(G) \leq \Phi(G' - v) + n \leq (n-4)^2/4 + n = n^2/4 - n + 4$. This suffices when $n \geq 11$.

Thus, we may assume that each vertex of G' has at most two neighbors in T. Using Lemma 5.2, our present bound on $\Phi(G)$ is $|E(G')| + n$, and $|E(G')| \leq (n-3)^2/4$, so $\Phi(G) \leq n^2/4 - n/2 + 9/4$, and we only need to reduce this by $13/4$.

Call a vertex of G' *full* if it has at least one neighbor in T and is adjacent to all vertices in the other partite set of G'. Each vertex that is not full reduces the added number of labels in the construction of Lemma 5.2 by 1 or reduces the degree-sum in G' by 1. Hence when $n \geq 13$ it suffices to show that there are at most six full vertices. We prove this for $n \geq 14$.

Since the neighborhood in T of a full vertex has size 1 or 2, there are only six possible such neighborhoods. Nonadjacent full vertices with the same neighborhood in T would have the same neighborhood in G, in which case Observation 3.4 completes the proof (using the inductive bound on $\varphi(G)$, not on $\Phi(G)$). Hence having seven full vertices requires full vertices u and v that are adjacent in G' and have the same neighborhood S in T. We argue that this leads to two disjoint triangles in G, which allows Case 1 to complete the proof.

Since nonadjacent full vertices cannot have the same neighborhood in T, no other full vertex has neighborhood S in T. If two other adjacent full vertices x and y have neighborhood S' in T, then $S \neq S'$, so there exist distinct vertices $s \in S$ and $s' \in S'$, and $\{s, u, v\}$ and $\{s', x, y\}$ are disjoint triangles.

Otherwise, for each $S' \subseteq T$ with $|S'| \leq 2$, some full vertex has neighborhood S' in T. For $s \in S$, the triangle $\{s, u, v\}$ is disjoint from the triangle formed by $T - \{s\}$ and the full vertex having neighborhood $T - \{s\}$ in T. $\qquad\square$

Theorem 5.8. *If $n(G) \geq 14$, then $\varphi(G) \leq n^2/4 - n/2 - 1$, with equality for G_n of Corollary 3.11 when n is even.*

Proof. The claim follows immediately from Lemmas 5.5, 5.6, and 5.7. □

We believe that the bound of Theorem 5.8 in fact holds for $n \geq 8$ and is sharp only for G_n when n is even. We also think that when n is odd the alternative construction of G_n in Corollary 3.11 is sharp, meaning that for odd n the upper bound can be strengthened slightly to match it. However, proving any of these statements seems likely to require substantial case analysis.

References

[1] M.-S. Chung and D. B. West, The p-intersection number of a complete bipartite graph and orthogonal double coverings of a clique. *Combinatorica* 14(1994), 453–461.

[2] N. Eaton, Intersection representation of complete unbalanced bipartite graphs. *J. Combin. Theory Ser. B* 71 (1997), 123–129.

[3] P. Erdős, A. W. Goodman, and L. Pósa, The representation of a graph by set intersections. *Canad. J. Math.* 18 (1966), 106–112.

[4] M. C. Golumbic, *Algorithmic graph theory and perfect graphs.* Academic Press, 1980.

[5] P. Henderson, Lower bounds on the overlap number of graphs, manuscript.

[6] B. Rosgen, Set Representations of Graphs. M.Sc. Thesis, University of Alberta, 2005.