Lower Bounds for Protrusion Replacement by Counting Equivalence Classes[*][†]

Bart M. P. Jansen[1] and Jules J. H. M. Wulms[2]

1   Eindhoven University of Technology, Eindhoven, The Netherlands
    b.m.p.jansen@tue.nl
2   Eindhoven University of Technology, Eindhoven, The Netherlands
    j.j.h.m.wulms@tue.nl

— Abstract —

Garnero et al. [SIAM J. Discrete Math. 2015, 29(4):1864–1894] recently introduced a framework based on dynamic programming to make applications of the *protrusion replacement* technique constructive and to obtain explicit upper bounds on the involved constants. They show that for several graph problems, for every boundary size t one can find an explicit set $\mathcal{R}_t$ of *representatives*. Any subgraph H with a boundary of size t can be replaced with a representative $H' \in \mathcal{R}_t$ such that the effect of this replacement on the optimum can be deduced from H and H' alone. Their upper bounds on the size of the graphs in $\mathcal{R}_t$ grow triple-exponentially with t. In this paper we complement their results by lower bounds on the sizes of representatives, in terms of the boundary size t. For example, we show that each set of planar representatives $\mathcal{R}_t$ for the INDEPENDENT SET problem contains a graph with $\Omega(2^t/\sqrt{4t})$ vertices. This lower bound even holds for sets that only represent the planar subgraphs of bounded pathwidth. To obtain our results we provide a lower bound on the number of equivalence classes of the canonical equivalence relation for INDEPENDENT SET on t-boundaried graphs. We also find an elegant characterization of the number of equivalence classes in general graphs, in terms of the number of monotone functions of a certain kind. Our results show that the number of equivalence classes is at most 2^{2^t}, improving on earlier bounds of the form $(t+1)^{2^t}$.

1   Introduction

Protrusion replacement is a versatile tool for attacking optimization problems on graphs. When applied to solve an optimization problem on a graph G, the main idea is the following: repeatedly replace a *protrusion* subgraph $H \subseteq G$ that interacts with the rest of G through a small boundary, by a smaller *representative* subgraph H'. Suppose that we can ensure that (i) the change Δ in the optimum caused by this replacement only depends on H and H', and that (ii) we can efficiently analyze H to find a suitable replacement H' and the corresponding Δ. Then we can solve the problem on G by solving it on the smaller graph and adding Δ to the final result. In recent years, protrusion replacement has been applied to obtain approximation

---

algorithms [7, 8], kernelization algorithms [1, 7, 8, 10, 15], and fixed-parameter tractable algorithms [8, 15]. The generality of protrusion replacement comes at a price: it often results in proofs that efficient algorithms of a certain type *exist*, without showing explicitly how such algorithms can be *constructed* and without giving any explicit bounds on the constant factors involved in the analysis. This non-constructivity stems from the use of a property called *finite integer index* (FII, defined below). It is used to argue that for every constant boundary size t, there is a finite set of representatives $\mathcal{R}_t$ such that any t-boundaried subgraph H can safely be replaced by some representative $H' \in \mathcal{R}_t$, as described above. The key issue is that FII only guarantees that a finite set of representatives exist, without showing how to find it, how large the set is, or how many vertices the representative subgraphs have.

To deal with the issue of non-constructivity, Garnero et al. [12] introduced a framework based on dynamic programming. They showed that explicit bounds for the sizes of representatives can be obtained by analyzing the number of states required to solve the problem on graphs of bounded treewidth. By presenting explicit dynamic programming algorithms for problems such as r-INDEPENDENT SET and r-DOMINATING SET, they were able to derive upper bounds on the size of representatives in terms of the boundary size t. These upper bounds grow very quickly with t, in some cases triple-exponentially. Garnero et al. [12, §7] suggest to examine to what extent this exponential dependance is unavoidable. We pursue this direction by presenting lower bounds.

Boundaried graphs and equivalence.   To state our results we have to introduce some terminology.[1] We only consider undirected, finite, simple graphs. Let t be a positive integer. A t-boundaried graph G consists of a vertex set $V(G)$, an edge set $E(G) \subseteq \binom{V(G)}{2}$, and an injective labeling $\lambda_G \colon \{1, \dots, t\} \to V(G)$ that identifies t distinct *boundary vertices* in the graph. The *boundary* of the graph is the set $B_G := \{\lambda_G(1), \dots, \lambda_G(t)\}$. Two t-boundaried graphs G and H can be *glued* together on their boundary, resulting in the boundaried graph $G \oplus H$ that is obtained from the disjoint union of G and H by identifying corresponding boundary vertices and removing any parallel edges that are introduced. That is, we merge $\lambda_G(i)$ with $\lambda_H(i)$ for each $i \in [t]$. An *optimization problem* Π on graphs assigns to every (unboundaried) graph G an optimal solution value $\Pi(G) \in \mathbb{Z}$. We will also write $\Pi(G)$ for a boundaried graph G to denote the optimum of the underlying unboundaried graph. Two t-boundaried graphs G and H are *equivalent* with respect to Π, denoted $G \equiv_{\Pi,t} H$, if there exists a *transposition constant* $\Delta \in \mathbb{Z}$ such that for every t-boundaried graph F:

$$\Pi(G \oplus F) = \Pi(H \oplus F) + \Delta. \tag{1}$$

It is easy to see that $\equiv_{\Pi,t}$ is an equivalence relation. Problem Π has *finite integer index* if $\equiv_{\Pi,t}$ has a finite number of equivalence classes for each fixed t. In the remainder, we omit the subscript t when it is clear from the context. Observe that these notions formalize the idea behind protrusion replacement sketched above: if $G \equiv_{\Pi,t} H$, then replacing G by H changes the optimum by exactly Δ.

Our results.   We analyze the canonical equivalence relation $\equiv_{\mathrm{IS},t}$ on t-boundaried graphs for the INDEPENDENT SET (IS) problem, which asks for the maximum size of an independent

---

[1] To avoid an abundance of cumbersome definitions, our terminology differs slightly from that in earlier work (cf. [2, 3], [5, §2]). In particular, we do not allow t-boundaried graphs with fewer than t boundary vertices. The fact that we consider optimization problems as in [5], rather than decision problems as in [1, 12], forms no essential difference; our lower bounds also apply to those settings.

set of pairwise non-adjacent vertices. We focus on INDEPENDENT SET due to its simple combinatorial structure, but our techniques carry over to DOMINATING SET, as explained in §8. Define a *set of representatives* for $\equiv_{\mathrm{IS},t}$ to be a set $\mathcal{R}_t$ of t-boundaried graphs, such that for every t-boundaried graph G there exists $H \in \mathcal{R}_t$ with $G \equiv_{\mathrm{IS},t} H$. Let the *critical size* of a set of representatives be the number of vertices of its largest graph. We aim to give a lower bound on the critical size of any set of representatives for INDEPENDENT SET in terms of t. Our approach consists of two steps. First, we construct a large set of pairwise nonequivalent graphs to give a lower bound on the number of equivalence classes of $\equiv_{\mathrm{IS},t}$. Then we use a counting argument to leverage this into a lower bound on the critical size. Observe that each equivalence class must be represented by a different graph. It follows that if the number of distinct t-boundaried graphs with at most s vertices is smaller than the number of equivalence classes, then the critical size of any set of representatives must be larger than s to give each class a distinct representative. By relating the number of small graphs to the number of equivalence classes, we therefore obtain the desired lower bounds.

Protrusion replacement is often applied in the context of restricted graph classes, where the protrusions to be replaced are known to have bounded treewidth and may even belong to a family of embeddable graphs such as planar graphs. With these application areas in mind, we develop our lower bounds to apply even when we wish only to have a representative for each equivalence class that contains a planar graph whose treewidth is $t + \mathcal{O}(1)$, for boundary size t. To find a large set of nonequivalent graphs we adapt a construction of Lokshtanov et al. [16], which they used to prove that INDEPENDENT SET on graphs of treewidth w cannot be solved in time $\mathcal{O}^*((2 - \varepsilon)^w)$ for any $\varepsilon > 0$ unless the Strong Exponential Time Hypothesis fails. We show that the graphs they construct can be made planar while increasing the treewidth (and in fact the pathwidth) by only a small additive term. More importantly, we show how to use this adapted construction to build a set of $M(t) - 2$ planar graphs of small treewidth which are pairwise nonequivalent under $\equiv_{\mathrm{IS},t}$, for all t. The term $M(t) \geq 2^{\binom{t}{\lfloor t/2 \rfloor}} \geq 2^{2^t/\sqrt{4t}}$ denotes the t-th *Dedekind number*, which counts the number of *monotone Boolean functions* of t variables. The number of equivalence classes therefore grows double-exponentially with t. Using the counting argument above, this allows us to give a lower bound of $\Omega(\log M(t)) \geq \Omega(2^t/\sqrt{4t})$ on the critical size of any set of planar representatives for the equivalence classes of $\equiv_{\mathrm{IS},t}$ that contain a planar graph of bounded pathwidth.

While developing a lower bound on the number of equivalence classes for planar graphs of bounded pathwidth, we also found an exact characterization of the number of equivalence classes of $\equiv_{\mathrm{IS},t}$ in general. We define a natural class of functions from $\{0,1\}^t$ to $\mathbb{N}$ that we call t-*representative functions*. We give a bijection between the t-representative functions and the equivalence classes of $\equiv_{\mathrm{IS},t}$ for t-boundaried graphs. As we will show that all monotone Boolean functions which are not constantly zero yield a distinct t-representative function, this gives a lower bound of $M(t) - 1$ on the number of equivalence classes of $\equiv_{\mathrm{IS},t}$. On the other hand, we show that the number of such functions is at most 2^{2^t-1}. The double-exponential lower bound for the number of equivalence classes containing a bounded-pathwidth planar graph is therefore not far off from the upper bound of 2^{2^t-1} in general graphs. The fact that the base of the double-exponential in this expression is independent of t is noteworthy. The naive way to bound the number of equivalence classes is to associate a table to each t-boundaried graph. For each subset S of the boundary vertices B, the table stores the maximum size of an independent set containing no vertex of $B \setminus S$. There are at most $t + 1$ distinct values in such a table, and two boundaried graphs whose tables differ in the same universal constant in all positions are easily shown to be equivalent. As there are 2^t entries in the table, and $t + 1$ different options per entry, this gives an upper bound of $(t + 1)^{2^t}$ on the

number of equivalence classes. Garnero et al. [12, Lemma 3.7] obtain the same bound using a subtly different definition for the table. Our result of 2^{2^t-1} yields a slight improvement.

2 Preliminaries

We use $\mathbb{N}$ to denote the natural numbers, including 0. For a positive integer n and a set X we use $\binom{X}{n}$ to denote the collection of all subsets of X of size n. The *power set* of X is denoted 2^X. The set $\{1, \ldots, n\}$ is abbreviated as $[n]$. A *Boolean function* is a function of the form $f: \{0,1\}^n \to \{0,1\}$. We sometimes use the equivalent view that a Boolean function assigns a 0/1-value to every subset $S \subseteq [n]$, which is the value of f when the arguments whose index is in S are set to 1 and the remaining arguments are set to 0. A Boolean function $f: 2^{[n]} \to \{0,1\}$ is *monotone* if $f(S') \le f(S)$ whenever $S' \subseteq S \subseteq [n]$. We will call Boolean functions in this form set-functions, and may replace $[n]$ by other finite sets of ordered elements. A formula in conjunctive normal form (CNF) is *monotone* if no literal appears negated. Proofs for statements marked (★) can be found in the full version [14].

▶ **Proposition 1** (★). *For every non-constant monotone Boolean set-function $f: 2^{[n]} \to \{0,1\}$ there is a monotone CNF formula ϕ such that for all $x_1, \ldots, x_n \in \{0,1\}^n$ we have $\phi(x_1, \ldots, x_n) = 1$ if and only if $f(\{i \mid x_i = 1\}) = 1$.*

Graphs. We will denote the treewidth of a graph G by $\mathbf{tw}(G)$ and its pathwidth by $\mathbf{pw}(G)$. It is well-known that $\mathbf{pw}(G) \ge \mathbf{tw}(G)$; refer to a textbook for further details [4, §7]. We use the following consequence of the gluing operation.

▶ **Proposition 2** (★). *Let G and H be t-boundaried graphs that share the same set of boundary vertices $B = \{v_1, \ldots, v_t\}$ but are otherwise vertex-disjoint. Then a vertex set $X \subseteq V(G \oplus H)$ is independent in $G \oplus H$ if and only if $X \cap V(G)$ is independent in G and $X \cap V(H)$ is independent in H.*

3 Characterizing equivalence classes for Independent Set

In this section we derive several tools to analyze the equivalence classes of $\equiv_{\mathrm{IS}}$. For each t-boundaried graph G we define a function that captures the interaction of optimal independent sets with its boundary. These will be useful to reason about the (non)equivalence of pairs of graphs with respect to $\equiv_{\mathrm{IS}}$.

▶ **Definition 3.** Let G be a t-boundaried graph with boundary $B = \{v_1, \ldots, v_t\}$. The function $\mathfrak{s}_G: 2^B \to \mathbb{N}$ expresses the size of a maximum independent set in G whose intersection with the boundary is a *subset* of a given set:

$$\mathfrak{s}_G(S) := \max\big\{|X| \,\big|\, X \text{ is an independent set in } G \text{ with } X \cap B \subseteq S\big\}. \tag{2}$$

We will see that equivalence classes can be characterized by the functions $\mathfrak{s}_G$ of the graphs G in that class. The next lemma shows that when gluing two boundaried graphs G and H together, the optimum of the resulting graph $G \oplus H$ can be deduced from $\mathfrak{s}_G$ and $\mathfrak{s}_H$. The identity we prove is reminiscent of the recurrence that is used for JOIN nodes when solving INDEPENDENT SET on graphs of bounded treewidth [4, §7.3.1].

▶ **Lemma 4** (★). *Let G and H be t-boundaried graphs for some t. The following holds:*

$$\max_{S \subseteq B}\{\mathfrak{s}_G(S) + \mathfrak{s}_H(S) - |S|\} = \mathrm{OPT}_{\mathrm{IS}}(G \oplus H).$$

To relate the equivalence of graphs to properties of the corresponding functions $\mathfrak{s}$, the following indicator graphs will be convenient.

▶ **Definition 5.** Let t be a positive integer and $B = \{v_1, \ldots, v_t\}$. For each $S \subseteq B$ define the t-boundaried *indicator* graph I_S with boundary B as the result of the following process: starting from an edgeless graph with vertex set B, for each $v_i \in B \setminus S$ add vertices u_i, u_i' and the edges $\{v_i, u_i\}, \{v_i, u_i'\}$ to I_S.

Each boundary vertex not in S thus becomes the center of a star with two leaves in I_S, and boundary vertices in S are isolated vertices in I_S. The next proposition shows that maximum independent sets of $F \oplus I_S$ reveal the value of $\mathfrak{s}_F(S)$.

▶ **Proposition 6 (★).** $\mathrm{OPT}_{\mathrm{IS}}(F \oplus I_S) = \mathfrak{s}_F(S) + 2(t - |S|)$ *for all t-boundaried graphs F.*

Using Proposition 6 we can show that the equivalence class of a boundaried graph G with respect to $\equiv_{\mathrm{IS}}$ is completely characterized by the function $\mathfrak{s}_G$.

▶ **Theorem 7 (★).** *Let G and H be two t-boundaried graphs with boundary $B = \{v_1, \ldots, v_t\}$. Then $G \equiv_{\mathrm{IS},t} H$ if and only if there exists a constant $c \in \mathbb{Z}$ such that $\mathfrak{s}_G(S) = \mathfrak{s}_H(S) + c$ for all $S \subseteq B$.*

Theorem 7 shows that two t-boundaried graphs G and H are equivalent under $\equiv_{\mathrm{IS}}$ if the functions $\mathfrak{s}_G$ and $\mathfrak{s}_H$ differ by a fixed constant for all inputs. It will be convenient to eliminate this degree of freedom by normalizing the functions.

▶ **Definition 8.** The *normalized boundary function* of a t-boundaried graph G with boundary B is the function $\mathfrak{s}_G^0 \colon 2^B \to \mathbb{N}$ given by $\mathfrak{s}_G^0(S) := \mathfrak{s}_G(S) - \mathfrak{s}_G(\emptyset)$.

Intuitively, $\mathfrak{s}_G^0(S)$ represents how much larger an independent set can be if we are allowed to use the boundary vertices from S, compared to when we are not allowed to use any boundary vertices in the independent set.

▶ **Corollary 9 (★).** *Let G and H be two t-boundaried graphs with boundary $B = \{v_1, \ldots, v_t\}$. Then $G \equiv_{\mathrm{IS}} H$ if and only if $\mathfrak{s}_G^0 = \mathfrak{s}_H^0$.*

Corollary 9 shows that equivalence classes of $\equiv_{\mathrm{IS}}$ are determined by the normalized boundary functions of the graphs in the class. To see how many different equivalence classes there can be, it is therefore useful to analyze the properties of normalized boundary functions.

▶ **Definition 10.** Let t be a positive integer and let $B := \{v_1, \ldots, v_t\}$. A function $f \colon 2^B \to \mathbb{N}$ is called a t-*representative function* if it satisfies the following three properties:
1. $f(\emptyset) = 0$.
2. *Monotonicity*: for any $S' \subseteq S \subseteq B$ we have $f(S') \leq f(S)$.
3. *Bounded increase*: For every nonempty set $S \subseteq B$ we have $f(S) \leq 1 + \min_{v \in S} f(S \setminus \{v\})$.

▶ **Lemma 11.** *Let G be a t-boundaried graph with boundary $B := \{v_1, \ldots, v_t\}$. Then $\mathfrak{s}_G^0$ is a t-representative function.*

Proof. We prove that $\mathfrak{s}_G^0$ has the three properties given in Definition 10.
1. By definition of $\mathfrak{s}_G^0$ we have $\mathfrak{s}_G^0(\emptyset) = \mathfrak{s}_G(\emptyset) - \mathfrak{s}_G(\emptyset) = 0$.
2. This follows directly from Definitions 3 and 8: the collection of independent sets over which $\mathfrak{s}_G(S')$ optimizes is a subset of the independent sets over which $\mathfrak{s}_G(S)$ optimizes.

3. Consider a nonempty set $S \subseteq B$ and let X be an independent set in G of size $\mathfrak{s}_G(S)$ with $X \cap B \subseteq S$, which exists by Definition 3. For every $v \in S$ we have that $X \setminus \{v\}$ is an independent set of size $|X| - 1$ in G whose intersection with B is a subset of $S \setminus \{v\}$, implying that $\mathfrak{s}_G(S \setminus \{v\}) \geq |X| - 1 = \mathfrak{s}_G(S) - 1$. Adding $1 - \mathfrak{s}_G(\emptyset)$ on both sides we obtain $\mathfrak{s}_G^0(S) = \mathfrak{s}_G(S) - \mathfrak{s}_G(\emptyset) \leq 1 + \mathfrak{s}_G(S \setminus \{v\}) - \mathfrak{s}_G(\emptyset) = 1 + \mathfrak{s}_G^0(S \setminus \{v\})$. As this holds for all $v \in S$, it holds in particular for $v \in S$ minimizing $\mathfrak{s}_G^0(S \setminus \{v\})$.   ◀

4   Defining graphs with given boundary characteristics

Corollary 9 shows that t-boundaried graphs with the same normalized boundary function belong to the same equivalence class. Since each normalized boundary function is a t-representative function by Lemma 11, this implies that the number of equivalence classes of $\equiv_{\mathrm{IS},t}$ is *at most* the number of distinct t-representative functions. In Lemma 13 we will show that, surprisingly, the converse also holds: for each t-representative function there is a distinct equivalence class of $\equiv_{\mathrm{IS},t}$. Before proving that lemma, we first derive a useful property of t-representative functions.

▶ **Proposition 12 (★).** *Each t-representative function f satisfies $f(S') - |S' \setminus S| \leq f(S)$ for all $S, S' \subseteq B$.*

▶ **Lemma 13.** *For every t-representative function f, there exists a t-boundaried graph G with boundary $B := \{v_1, v_2, \ldots, v_t\}$, such that $\mathfrak{s}_G^0(S) = f(S)$ for every $S \subseteq B$.*

Proof. Consider an arbitrary t-representative function f, which assigns a non-negative integer to each $S \subseteq B$. We construct a t-boundaried graph G for which $\mathfrak{s}_G^0 = f$, as follows:
1. Start from an edgeless graph with vertex set B, which is the boundary of the graph.
2. For each $i \in [t]$ add a vertex u_i and the edge $\{u_i, v_i\}$.
3. For each $S \subseteq B$ with $f(S) > 0$, add a set $V_S = \{v_{S,1}, \ldots, v_{S,f(S)}\}$ consisting of $f(S)$ vertices to the graph. These vertices are false twins (all share the same open neighborhood) and are connected to the rest of the graph as follows:
   a. For each $i \in [t]$ with $v_i \in S$, all vertices of V_S are adjacent to u_i.
   b. For each $i \in [t]$ with $v_i \notin S$, all vertices of V_S are adjacent to v_i.
   c. All vertices of V_S are adjacent to all vertices $V_{S'}$ that are created for sets $S' \neq S$.

We show that $\mathfrak{s}_G(S) = t + f(S)$ for all $S \subseteq B$. This will imply that $\mathfrak{s}_G^0(S) = \mathfrak{s}_G(S) - \mathfrak{s}_G(\emptyset) = (t + f(S)) - (t + f(\emptyset)) = (t + f(S)) - (t + 0) = f(S)$ for all $S \subseteq B$, since $f(\emptyset) = 0$ by Definition 10. We therefore conclude the proof by showing that $\mathfrak{s}_G(S) = t + f(S)$ for all $S \subseteq B$, by establishing two inequalities. Consider an arbitrary $S \subseteq B$.

($\geq$) To show $\mathfrak{s}_G(S) \geq t + f(S)$ we construct an independent set X in G of size $t + f(S)$ that intersects B in a subset of S. If $f(S) = 0$ then $X = \{u_1, \ldots, u_t\}$ suffices, so assume in the remainder that $f(S) > 0$. Let X consist of the $f(S)$ vertices in V_S, together with the vertices $\{u_i \mid i \in [t], v_i \notin S\}$ and $\{v_i \mid i \in [t], v_i \in S\}$. Then $|X| = t + f(S)$, and using the construction above it is straight-forward to verify that X is an independent set. Since $X \cap B = S$, this shows that $\mathfrak{s}_G(S) \geq t + f(S)$.

($\leq$) Now we argue that $\mathfrak{s}_G(S) \leq t + f(S)$. Consider a maximum independent set X in G that intersects B in a subset of S, which has size $\mathfrak{s}_G(S)$ by Definition 3. If X contains no vertices of $V_{S'}$ for any $S' \subseteq B$, then X has at most t vertices: an independent set contains at most one vertex of each edge $\{v_i, u_i\}$ for each $i \in [t]$. Hence $|X| \leq t$ in this case, which is at most $t + f(S)$ since $f(S) \geq 0$ by Properties 1 and 2. In the remainder, assume X contains a vertex of $V_{S'}$ for some $S' \subseteq B$. This implies that X contains no vertices from $V_{S''}$

for any $S'' \neq S'$, since all vertices of $V_{S'}$ are adjacent to all vertices of $V_{S''}$ by construction of G. Hence besides the vertices from $V_{S'}$, the set X only contains vertices of edges $\{v_i, u_i\}$ for $i \in [t]$. The independent set X contains at most one vertex from each such edge. For each $v_i \in S' \setminus S$, observe that X does not contain v_i (since $X \cap B \subseteq S$), and X does not contain u_i either (since u_i is adjacent to all members of $V_{S'}$). So X has at most $f(S')$ vertices from $V_{S'}$, no vertices of $\{v_i, u_i\}$ for each $v_i \in S' \setminus S$, and at most one vertex from each of the remaining $t - |S' \setminus S|$ edges. It follows that $|X| \leq f(S') + (t - |S' \setminus S|)$. By Proposition 12 we have $f(S') - |S' \setminus S| \leq f(S)$, which shows that $|X| \leq t + f(S)$ and concludes the proof. ◀

5 Counting t-representative functions

We say that two t-representative functions are distinct if their function values differ on some input. Lemma 13 shows that for each t-representative function f, there exists a t-boundaried graph whose normalized boundary function equals f. Together with Corollary 9, which says that boundaried graphs with the same normalized boundary function are equivalent under $\equiv_{\text{IS},t}$, this establishes a bijection between the equivalence classes of $\equiv_{\text{IS},t}$ and the t-representative functions. To bound the number of equivalence classes of $\equiv_{\text{IS},t}$ it therefore suffices to bound the number of t-representative functions. Recall that $M(t)$ denotes the t-th *Dedekind number*, the number of distinct monotone Boolean functions of t variables.

▶ **Lemma 14 (★).** *There are at least $M(t) - 1$ distinct t-representative functions.*

It is known that $M(t) \geq 2^{\binom{t}{\lfloor t/2 \rfloor}}$. To see this, consider the subsets $\mathcal{S}_t = \binom{[t]}{\lfloor t/2 \rfloor}$ of $[t]$ of size $\lfloor t/2 \rfloor$. For each subset $\mathcal{S}'_t \subseteq \mathcal{S}_t$ we obtain a different monotone set-function by saying that $f(S) = 1$ if and only if S contains one of the subsets in $\mathcal{S}'_t$. By Stirling's approximation we have $\binom{t}{\lfloor t/2 \rfloor} \geq 2^t/\sqrt{4t}$, which implies that $M(t) \geq 2^{2^t/\sqrt{4t}}$. The following lemma gives an upper bound on the number of t-representative functions.

▶ **Lemma 15 (★).** *The number of distinct t-representative functions is at most $2^{2^t - 1}$.*

Lemmata 14 and 15 give the following corollary for each positive integer t.

▶ **Corollary 16.** *The number of equivalence classes of $\equiv_{\text{IS},t}$ lies between $2^{2^t/\sqrt{4t}}$ and $2^{2^t - 1}$.*

6 Defining planar graphs with given boundary characteristics

In Lemma 13 we constructed nonequivalent t-boundaried graphs based on distinct t-representative functions. The graphs constructed in that lemma have large treewidth and are far from being planar; they contain cliques of size roughly 2^t. To derive lower bounds that are meaningful even when protrusion replacement is applied for planar graphs of bounded treewidth, we present an alternative construction to lower bound the number of equivalence classes that contain a planar graph of small pathwidth (and therefore have small treewidth). The following gadget, of which several variations were used in earlier work (cf. [13, Theorem 5.3] and [9, 16]), will be useful in our construction.

▶ **Definition 17.** Let k be a positive integer. The *clause gadget of size k* is the graph $\mathcal{C}_k$ constructed as follows (see Figure 1a). For each $i \in [k]$ create a triangle on vertices $\{u_i, v_i, w_i\}$. Connect these into a path by adding all edges $\{w_i, u_{i+1}\}$ for $i \in [k-1]$. Finally, add vertices $v_{\text{start}}, v_{\text{end}}$ and the edges $\{v_{\text{start}}, u_1\}$ and $\{w_k, v_{\text{end}}\}$. The vertices $(v_1, \ldots, v_k)$ are the *terminals* of the clause gadget.

(a) Clause gadget $\mathcal{C}_3$

(b) Crossover gadget $G_\times$

(c) Values of OPT_{IS} in $G_\times$
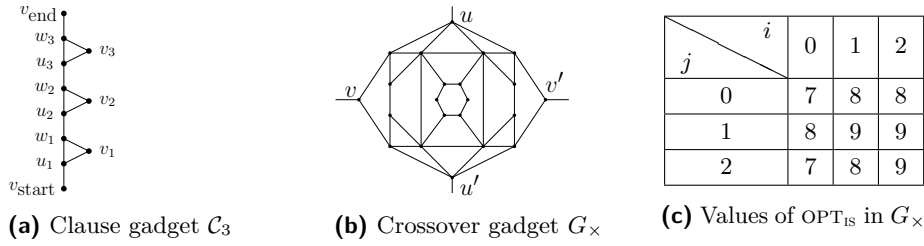
Figure 1 Gadgets for INDEPENDENT SET. The crossover gadget is due to Garey et al. [11, Fig. 11 and Table 1]. The table on the right shows for all relevant combinations of i and j what the maximum size is of an independent set X satisfying $|\{v, v'\} \cap X| = i$ and $|\{u, u'\} \cap X| = j$.

▶ **Observation 18** (Cf. [9, Obs. 6–8]). *For each positive $k \in \mathbb{N}$, the clause gadget $\mathcal{C}_k$ has the following properties:*
1. $\text{OPT}_{\text{IS}}(\mathcal{C}_k) = k + 2$.
2. *Every maximum independent set in $\mathcal{C}_k$ contains a terminal vertex v_i for some $i \in [k]$.*
3. $\forall i \in [k]$ *there is a maximum independent set in $\mathcal{C}_k$ containing v_i but no other terminals.*
4. $\mathcal{C}_k$ *is planar and $\boldsymbol{pw}(\mathcal{C}_k) = 2$.*

To ensure our construction yields a planar graph, we use a *crossover gadget* for INDE-PENDENT SET due to Garey et al. [11]. It was originally designed for VERTEX COVER, but since the complement of a maximum independent set is a minimum vertex cover, we can rephrase the properties of the gadget in terms of independent sets. The crossover gadget $G_\times$ is the 22-vertex graph illustrated in Figure 1b, which has four terminals (u, u', v, v'). When we have a drawing of a graph G in which exactly two edges $\{a, b\}, \{c, d\}$ cross in a common point, we can *planarize the crossing* by removing edges $\{a, b\}$ and $\{c, d\}$, introducing a new copy of $G_\times$ at the position of the crossing, and adding the edges $\{a, v\}, \{v', b\}, \{c, u\}, \{u', d\}$. Garey et al. [11] analyzed the size of a maximum independent set in $G_\times$ when restricting which terminal vertices may occur in the set, as shown in Figure 1c. As $G_\times$ is symmetric in both the horizontal and vertical axis, and the table shows that a maximum size independent set size of nine can already be obtained using $i = 1$ of the terminals $\{v, v'\}$ and $j = 1$ of the terminals $\{u, u'\}$, we observe the following.

▶ **Observation 19.** *For any choice of terminals $v^* \in \{v, v'\}$ and $u^* \in \{u, u'\}$ there is a maximum independent set of size nine in $G_\times$ that does not contain v^* or u^*.*

The following proposition summarizes the essential features of a planarization operation.

▶ **Proposition 20** (★). *Let G be a graph drawn in the plane such that no edge contains a vertex in its interior and no more than two edges cross in any single point. Let G' be the result of planarizing an edge crossing by a crossover gadget. The following holds.*
1. *For every independent set X in G there is an independent set X' in G' of size $|X| + 9$ such that $X' \cap V(G) = X$.*
2. *For every independent set X' in G' there is an independent set X'' in G' with $|X'| = |X''|$ containing exactly nine vertices from $G_\times$ with $X'' \cap V(G) \subseteq X' \cap V(G)$.*
3. *For every independent set X' in G' there is an independent set X in G of size $|X'| - 9$ such that $X \subseteq X' \cap V(G)$.*
4. $\text{OPT}_{\text{IS}}(G') = \text{OPT}_{\text{IS}}(G) + 9$.

In most applications of crossover gadgets, the only important property is that they have a fixed effect on the optimum (Property 4). In our case we also have to ensure that the
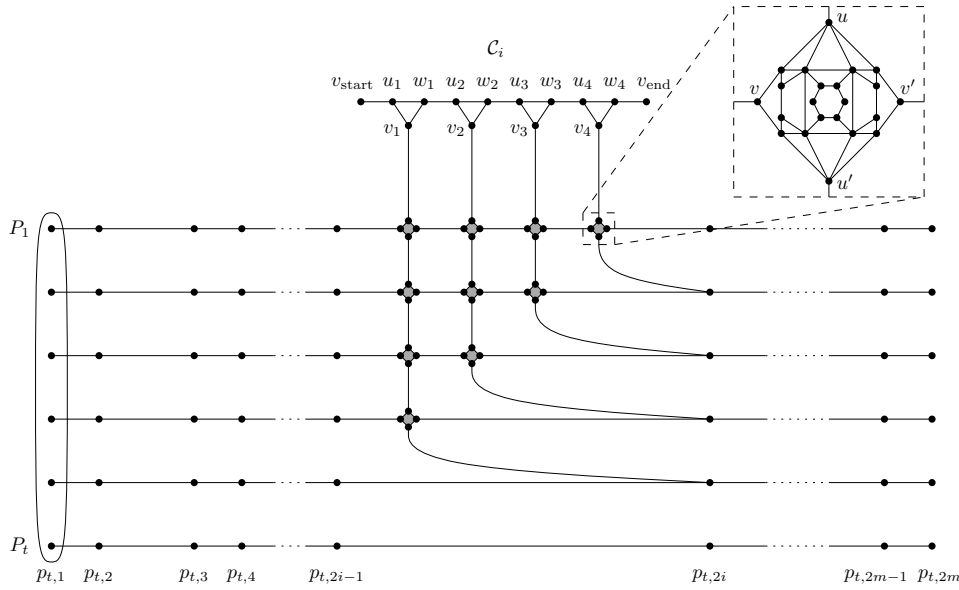
Figure 2 Planarizing the graph G_ϕ to obtain G'_ϕ in the proof of Lemma 21. Only the clause gadget for the clause $C_i = (x_5 \vee x_4 \vee x_3 \vee x_2)$ is shown. Shaded diamonds represent crossover gadgets. The boundary B of the graph is circled, containing the first vertex from each path.

crossover gadgets do not disturb how the solutions intersect the boundary of the graph. Properties 1–3 will be used for this purpose. Using these gadgets we present the construction.

▶ **Lemma 21.** *Let t be a positive integer and $B := \{p_{1,1}, p_{2,1}, \ldots, p_{t-1,1}, p_{t,1}\}$. For every non-constant monotone set-function $f \colon 2^B \to \{0, 1\}$ there is a planar graph G with boundary B such that $\boldsymbol{pw}(G) \leq t + \mathcal{O}(1)$ and for every $S \subseteq B$: $f(S) = 1$ if and only if $\mathfrak{s}_G(S) = \mathrm{OPT}_{\mathrm{IS}}(G)$.*

Proof. Consider a monotone set-function f and let ϕ be a monotone CNF formula that represents f in the sense of Proposition 1. Let the clauses of ϕ be $C_1, \ldots, C_m$ such that each clause C_i is a subset of $[t]$ giving the indices of the variables appearing in the clause. Since ϕ is monotone, all variables appear positively. The number of literals in C_i is denoted $|C_i|$.

We first construct a nonplanar graph G_ϕ of small pathwidth such that for all $S \subseteq B$ we have $f(S) = 1$ if and only if $\mathfrak{s}_{G_\phi}(S) = \mathrm{OPT}_{\mathrm{IS}}(G_\phi)$. Then we will use crossover gadgets to turn G_ϕ into a planar graph G'_ϕ while preserving these properties. The construction is inspired by a reduction of Lokshtanov et al. [16, Thm. 3.1], and proceeds as follows.

1. We start by creating t paths $P_1, \ldots, P_t$, where every path P_i for $i \in [t]$ consists of $2m$ vertices $p_{i,1}, \ldots, p_{i,2m}$. The boundary $B = \{p_{1,1}, \ldots, p_{t,1}\}$ of graph G_ϕ contains the first vertex from each path.
2. For each clause $i \in [m]$, add a copy of the clause gadget $\mathcal{C}_{|C_i|}$ to the graph and denote its terminals by $(v_1, \ldots, v_{|C_i|})$. Let $\ell(j)$ denote the j-th variable in the clause for each $j \in [|C_i|]$ and sort these such that $\ell(1) > \ell(2) > \ldots > \ell(|C_i|)$; this will be useful later on when planarizing the graph. For each $j \in [|C_i|]$ make terminal v_j in the clause gadget adjacent to vertex $p_{\ell(j),2i}$ on path $P_{\ell(j)}$. Observe that clause gadgets only connect to *even-numbered* vertices on the paths.

▶ **Claim 22 (★).** *The graph G_ϕ with boundary $B := \{p_{1,1}, \ldots, p_{t,1}\}$ satisfies:*
1. $\mathfrak{s}_{G_\phi}(B) = \mathrm{OPT}_{\mathrm{IS}}(G_\phi) \leq mt + \sum_{1 \leq i \leq m}(|C_i| + 2)$.
2. $\mathfrak{s}_{G_\phi}(B) = \mathrm{OPT}_{\mathrm{IS}}(G_\phi) = mt + \sum_{1 \leq i \leq m}(|C_i| + 2)$.
3. *For each $S \subseteq B$ we have $f(S) = 1$ if and only if $\mathfrak{s}_{G_\phi}(S) = \mathrm{OPT}_{\mathrm{IS}}(G_\phi)$.*

Claim 22 shows that the boundary function of G_ϕ expresses the monotone Boolean function f. The same argumentation as used by Lokshtanov et al. [16, Lemma 3.3] shows that G_ϕ has pathwidth $t + \mathcal{O}(1)$. However, we will not prove this here for the non-planar graph G_ϕ; we will prove a pathwidth bound after planarizing the graph. The planarization starts from a drawing of G_ϕ in the plane in which the crossings have a fixed structure. This drawing is defined as follows (see Figure 2):

- Draw each path $P_1, \ldots, P_t$ horizontally. Place the paths above each other so that P_1 is the highest and P_t is the lowest.
- For each clause $i \in [m]$ of ϕ, draw the clause gadget in a planar fashion above the paths, so that its terminals stick out at the bottom, the lowest-indexed terminal on the left and the highest-numbered terminal on the right. Draw the gadget for clause i between the vertical lines containing the $2i - 1$-th and the $2i$-th vertices on each path. Consider the set of edges E_{C_i} connecting the gadget for clause C_i to the vertices of the paths. By construction of G_ϕ, the gadget only connects to vertices with index $2i$ on the paths. Draw the edges from E_{C_i} in such a way that $e \in E_{C_i}$ only crosses the edges between the vertices $p_{j,2i-1}$ and $p_{j,2i}$ of the paths P_j for $j \in [t]$, and do not cross any other edge $e' \in E_{C_i}$. Since the left-to-right order of the variables in a clause matches the order in which the paths are laid out from top to bottom, this is possible.

Based on this drawing we planarize the graph G_ϕ by repeatedly replacing crossings by crossover gadgets, resulting in a planar graph G'_ϕ as shown in Figure 2. Let N denote the number of crossover gadgets which were introduced during the planarization process. By Proposition 20 we know that $\text{OPT}_{\text{IS}}(G'_\phi) = \text{OPT}_{\text{IS}}(G_\phi) + 9N = mt + 9N + \sum_{1 \le i \le m}(|C_i| + 2)$, where we use Property 2 of Claim 22 for the second equality. To conclude the proof, it remains to show that $\mathbf{pw}(G'_\phi) \le t + \mathcal{O}(1)$ (Claim 24) and that for all subsets $S \subseteq B$ we have $f(S) = 1$ if and only if $\mathfrak{s}_{G'_\phi}(S) = \text{OPT}_{\text{IS}}(G'_\phi)$ (Claim 23).

▶ **Claim 23 (★).** *For every $S \subseteq B$ we have $f(S) = 1$ if and only if $\mathfrak{s}_{G'_\phi}(S) = \text{OPT}_{\text{IS}}(G'_\phi)$.*

▶ **Claim 24 (★).** *The graph G'_ϕ has pathwidth $t + \mathcal{O}(1)$.*

This concludes the proof of Lemma 21.   ◀

7   Lower bound for protrusion replacement

To leverage the construction of Lemma 21 into a lower bound on the critical size of a set of representatives, we need the following lemma. Observe that its second condition shows that no pair of graphs from the constructed set $\mathcal{G}$ is equivalent under $\equiv_{\text{IS},t}$, and this is witnessed already by gluing planar graphs of pathwidth one onto them. This implies that in any protrusion reduction scheme applied to planar graphs that aims to replace occurrences of bounded-pathwidth protrusions by representatives, there should be a distinct representative for each graph in $\mathcal{G}$.

▶ **Lemma 25 (★).** *For each positive integer t there is a set $\mathcal{G}$ of $M(t) - 2$ distinct t-boundaried planar graphs of pathwidth $t + \mathcal{O}(1)$, such that for each pair of distinct graphs $G_f, G_{f'} \in \mathcal{G}$ there are two indicator graphs I_S and I_B as in Definition 5 such that:*
1. *The graphs $G_f \oplus I_S, G_f \oplus I_B, G_{f'} \oplus I_S, G_{f'} \oplus I_B$ are planar and have pathwidth $t + \mathcal{O}(1)$.*
2. $\text{OPT}_{\text{IS}}(G_f \oplus I_S) - \text{OPT}_{\text{IS}}(G_{f'} \oplus I_S) \neq \text{OPT}_{\text{IS}}(G_f \oplus I_B) - \text{OPT}_{\text{IS}}(G_{f'} \oplus I_B)$.

Finally, we can combine our lower bound on the number of distinct equivalence classes of Lemma 25 with an upper bound on the number of small graphs to obtain our main result.

▶ **Theorem 26 (★).** *Let $t \geq t_0$ be a sufficiently large positive integer. Let $\mathcal{R}_t$ be a set of t-boundaried planar graphs such that every equivalence class of $\equiv_{\mathrm{IS},t}$ that contains a planar graph of pathwidth $t + \mathcal{O}(1)$ is represented by some graph in $\mathcal{R}_t$. Then $\mathcal{R}_t$ contains a graph with $\Omega(\log M(t)) \geq \Omega(2^t/\sqrt{4t})$ vertices.*

8   Conclusion

We presented lower and upper bounds on the number of equivalence classes of the canonical equivalence relation $\equiv_{\mathrm{IS},t}$ for INDEPENDENT SET on t-boundaried graphs. We combined these lower bounds with upper bounds on the number of small graphs to give lower bounds for the critical sizes of sets of representatives. For a set of *planar* representatives that represent all equivalence classes containing a bounded-pathwidth planar graph, we gave a lower bound of $\Omega(\log M(t)) \geq \Omega(2^t/\sqrt{4t})$ on the critical size. The same argumentation can also be used to obtain lower bounds on the critical size of sets of potentially *nonplanar* representatives. The number of distinct t-boundaried (unrestrained) graphs is at most $2^{\binom{n}{2}} \cdot \binom{n}{t} \leq 2^{n^2/2}$. Using this bound in the proof of Theorem 26 yields a lower bound of $\Omega(\sqrt{\log M(t)}) \geq \Omega(2^{t/2}/\sqrt[4]{4t})$ on the critical size of a set of representatives that contains at least $M(t) - 2$ distinct graphs.

In their work, Garnero et al. [12] (roughly) show that each equivalence class of $\equiv_{\mathrm{IS},t}$ containing a planar graph of treewidth at most t can be represented by a planar graph with $2^{(t+1)^{2^t}}$ vertices and treewidth at most t. Our lower bound shows that to represent all equivalence classes containing a planar graph of *pathwidth $t + \mathcal{O}(1)$* (a subset of the graphs of treewidth $t + \mathcal{O}(1)$), requires a graph with $\Omega(2^t/\sqrt{4t})$ vertices. Our single-exponential lower bound is very far from the triple-exponential upper bound. However, we believe that the correct bound is single-exponential. Since Corollary 9 shows that each equivalence class is completely characterized by its normalized boundary function, and the construction of Lemma 13 produces a boundaried graph with $2^{\mathcal{O}(t)}$ vertices for any given boundary function, it follows that every equivalence class of $\equiv_{\mathrm{IS},t}$ has a representative with $2^{\mathcal{O}(t)}$ vertices. Note, however, that the representatives constructed in this way are nonplanar and have pathwidth and treewidth $2^{\Theta(t)}$.

The main conceptual contribution of this work is the fact that nontrivial lower bounds can be obtained by counting equivalence classes. The fact that a significant portion of the equivalence classes (at least $M(t) \geq 2^{2^t/\sqrt{4t}}$ out of the total of at most 2^{2^t}) can be generated from monotone Boolean functions was useful in the construction of nonequivalent planar graphs of bounded pathwidth. We showed that the lower bound construction of Lokshtanov et al. [16] can be planarized while increasing the pathwidth by an additive constant. The planarization argument employed here can also be used to strengthen the SETH-based runtime lower bound of $\Omega((2 - \varepsilon)^w \cdot n^{\mathcal{O}(1)})$ for solving INDEPENDENT SET on graphs of treewidth w, to *planar* graphs of treewidth w. Not all bounded-pathwidth graphs can be planarized with a bounded increase in pathwidth. In particular, when planarizing $K_{3,n}$ for sufficiently large n the pathwidth grows arbitrarily large [6].

The lower bounds for INDEPENDENT SET given in Theorem 26 carry over to the DOMINATING SET problem, for which protrusion replacement is used frequently. In the full version [14] we describe this extension, which is based on the folklore planarity-preserving NP-completeness reduction from VERTEX COVER to DOMINATING SET.

───── **References** ─────

1   Hans L. Bodlaender, Fedor V. Fomin, Daniel Lokshtanov, Eelko Penninkx, Saket Saurabh, and Dimitrios M. Thilikos. (Meta) Kernelization. In *Proc. 50th FOCS*, pages 629–638. IEEE Computer Society, 2009. `doi:10.1109/FOCS.2009.46`.

2   Hans L. Bodlaender, Fedor V. Fomin, Daniel Lokshtanov, Eelko Penninkx, Saket Saurabh, and Dimitrios M. Thilikos. (Meta) Kernelization. *CoRR*, 2013. `arXiv:0904.0727`.

3   Hans L. Bodlaender and Babette van Antwerpen-de Fluiter. Reduction algorithms for graphs of small treewidth. *Inf. Comput.*, 167(2):86–119, 2001. `doi:10.1006/inco.2000.2958`.

4   Marek Cygan, Fedor V. Fomin, Lukasz Kowalik, Daniel Lokshtanov, Dániel Marx, Marcin Pilipczuk, Michal Pilipczuk, and Saket Saurabh. *Parameterized Algorithms*. Springer, 2015. `doi:10.1007/978-3-319-21275-3`.

5   Babette de Fluiter. *Algorithms for Graphs of Small Treewidth*. PhD thesis, Utrecht University, 1997.

6   David Eppstein. Pathwidth of planarized drawing of $K_{3,n}$. TheoryCS StackExchange question, 2016. URL: `http://cstheory.stackexchange.com/questions/35974/`.

7   Fedor V. Fomin, Daniel Lokshtanov, Neeldhara Misra, Geevarghese Philip, and Saket Saurabh. Hitting forbidden minors: Approximation and kernelization. *SIAM J. Discrete Math.*, 30(1):383–410, 2016. `doi:10.1137/140997889`.

8   Fedor V. Fomin, Daniel Lokshtanov, Neeldhara Misra, and Saket Saurabh. Planar $\mathcal{F}$-deletion: Approximation, kernelization and optimal FPT algorithms. In *Proc. 53rd FOCS*, pages 470–479. IEEE Computer Society, 2012. `doi:10.1109/FOCS.2012.62`.

9   Fedor V. Fomin and Torstein J. F. Strømme. Vertex cover structural parameterization revisited. *CoRR*, 2016. `arXiv:1603.00770`.

10   Jakub Gajarský, Petr Hliněný, Jan Obdržálek, Sebastian Ordyniak, Felix Reidl, Peter Rossmanith, Fernando Sanchez Villaamil, and Somnath Sikdar. Kernelization using structural parameters on sparse graph classes. In *Proc. 21st ESA*, pages 529–540. Springer, 2013. `doi:10.1007/978-3-642-40450-4_45`.

11   M. R. Garey, David S. Johnson, and Larry J. Stockmeyer. Some simplified NP-complete graph problems. *Theor. Comput. Sci.*, 1(3):237–267, 1976. `doi:10.1016/0304-3975(76)90059-1`.

12   Valentin Garnero, Christophe Paul, Ignasi Sau, and Dimitrios M. Thilikos. Explicit linear kernels via dynamic programming. *SIAM J. Discrete Math.*, 29(4):1864–1894, 2015. `doi:10.1137/140968975`.

13   Bart M. P. Jansen. *The Power of Data Reduction: Kernels for Fundamental Graph Problems*. PhD thesis, Utrecht University, The Netherlands, 2013. URL: `http://igitur-archive.library.uu.nl/dissertations/2013-0612-200803/UUindex.html`.

14   Bart M. P. Jansen and Jules J. H. M. Wulms. Lower bounds for protrusion replacement by counting equivalence classes. *CoRR*, 2016. `arXiv:1609.09304`.

15   Eun Jung Kim, Alexander Langer, Christophe Paul, Felix Reidl, Peter Rossmanith, Ignasi Sau, and Somnath Sikdar. Linear kernels and single-exponential algorithms via protrusion decompositions. *ACM Trans. Algorithms*, 12(2):21, 2016. `doi:10.1145/2797140`.

16   Daniel Lokshtanov, Dániel Marx, and Saket Saurabh. Known algorithms on graphs on bounded treewidth are probably optimal. In *Proc. 22nd SODA*, pages 777–789. SIAM, 2011. `doi:10.1137/1.9781611973082.61`.