Near-domination in graphs

Bruce Reed
Researcher, Projet COATI, INRIA and Laboratoire I3S, CNRS France,
and Visiting Researcher, IMPA, Brazil

Alex Scott
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

Paul Seymour\textsuperscript{1}
Princeton University, Princeton, NJ 08544, USA

May 29 2017; revised June 24, 2017

\textsuperscript{1}Supported by ONR grant N00014-14-1-0084 and NSF grant DMS-1265563.
Abstract

A vertex $u$ of a graph “$t$-dominates” a vertex $v$ if there are at most $t$ vertices different from $u, v$ that are adjacent to $v$ and not to $u$; and a graph is “$t$-dominating” if for every pair of distinct vertices, one of them $t$-dominates the other. Our main result says that if a graph is $t$-dominating, then it is close (in an appropriate sense) to being 0-dominating.

It was already known that every $t$-dominating graph on $n$ vertices has a stable set or clique of cardinality $O(n^\epsilon)$ for some $\epsilon > 0$ depending on $t$, in accordance with the Erdős-Hajnal conjecture, but from our result it follows that the same is true with $\epsilon = 1$. 
1 Introduction

In this paper, all graphs are finite and have no loops or parallel edges. We say a vertex \( u \) of a graph \( t \)-dominates a vertex \( v \) if there are at most \( t \) vertices different from \( u, v \) that are adjacent to \( v \) and not to \( u \); and a graph is \( t \)-dominating if for every pair of distinct vertices, one of them \( t \)-dominates the other.

Graphs that are 0-dominating are easily understood; they are called “threshold graphs” and have several different characterizations, which we discuss later. But for general fixed \( t \), \( t \)-dominating graphs are not so transparent, and our main result states that every \( t \)-dominating graph has bounded “edit distance” from a 0-dominating graph. Let us define this.

Let \( G, H \) be graphs on the same vertex set. We say that the edit distance between \( G, H \) is \( d \) if \( d \) is the maximum, over all vertices \( v \), of
\[ |N_G(v) \setminus N_H(v)| + |N_H(v) \setminus N_G(v)|, \]
where \( N_G(v), N_H(v) \) denote the set of neighbours of \( v \) in \( G, H \) respectively. Thus, if \( G, H \) have edit distance \( d \), then \( H \) can be obtained from \( G \) by changing the adjacency of some pairs of vertices, where the changed pairs form a graph with maximum degree \( d \). Edit distance is evidently a metric, and could be used, for instance, to describe “defective colouring”. A \( d \)-defective \( k \)-colouring of a graph is a partition of its vertex set into \( k \) subsets such that the subgraph induced on each subset has maximum degree at most \( d \), and this is the same as saying the graph has edit distance at most \( d \) from a \( k \)-colourable graph.

It is easy to see that if \( G \) is 0-dominating and \( G, H \) have edit distance at most \( d \) then \( H \) is \( 2d \)-dominating, and we prove a kind of converse:

1.1 For all integers \( t \geq 0 \), if \( G \) is \( t \)-dominating then there is a 0-dominating graph \( H \) with the same vertex set such that \( G, H \) have edit distance at most \( 646t^4 \).

The proof is in three stages: we reduce the problem to “split graphs” (graphs with vertex set partitioned into a clique and a stable set); we reduce the split graph question to a problem about matrices; and then we solve the matrix problem. They are carried out in reverse order.

This research was motivated by the Erdős-Hajnal conjecture [4], that for every graph \( H \), there exists \( \epsilon > 0 \) such that every graph \( G \) not containing \( H \) as an induced subgraph has a stable set or clique of cardinality at least \( O(|V(G)|^\epsilon) \), and we discuss an application of our theorem to that conjecture in the final section.

2 0-domination

There are several different characterizations of 0-dominating graphs, as we explain now. A graph \( G \) is 0-dominating if and only if there are no four distinct vertices \( u, u', v, v' \) such that \( uu' \) and \( vv' \) are edges and the pairs \( u, v' \) and \( v, u' \) are nonadjacent. Here \( u, v \) might or might not be adjacent, and also \( u', v' \) might or might not be adjacent; so this can restated as:

2.1 A graph is 0-dominating if and only if no induced subgraph is isomorphic to \( C_4, 2K_2, \) or \( P_4 \).
(\(C_4\) denotes the four-vertex cycle graph; \(2K_2\) denotes its complement, the graph consisting of two disjoint edges; and \(P_4\) denotes the four-vertex path.) A split graph is a graph \(G\) such that \(V(G)\) can be partitioned into a clique and a stable set, and a graph is a split graph if and only if [6] it has no induced subgraph isomorphic to \(C_4, 2K_2\) or \(C_5\). Thus every 0-dominating graph is a split graph.

A half-graph is a bipartite graph with bipartition \(X,Y\) say, such that \(X,Y\) can be ordered as \(X = \{x_1, \ldots, x_m\}\) and \(Y = \{y_1, \ldots, y_n\}\) with the following property: for all \(i,i',j,j'\) with \(1 \leq i \leq i' \leq m\) and \(1 \leq j \leq j' \leq n\), if \(x_i y_j\) is an edge then \(x_i y_{j'}\) is an edge. It is easy and well-known that a bipartite graph is a half-graph if and only if it has no induced subgraph isomorphic to \(2K_2\). Let us say a split half-graph is a graph obtained from a half-graph with bipartition \(X,Y\) by adding edges to make \(X\) a clique.

Let \(G\) be 0-dominating, and let \(V(G)\) be the disjoint union of a clique \(X\) and a stable set \(Y\). Let \(B\) be the bipartite graph with bipartition \(X,Y\) formed by the edges of \(G\) between \(X\) and \(Y\). Then \(B\) has no induced subgraph \(2K_2\), since \(G\) has no \(P_4\), and it follows that \(B\) is a half-graph, and so \(G\) is a split half-graph. Since split half-graphs are 0-dominating, we have shown that:

2.2 \(G\) is 0-dominating if and only if \(G\) is a split half-graph.

Every nonnull split half-graph has either a vertex of degree zero or a vertex adjacent to all other vertices, and deleting this vertex gives another split half-graph. A graph \(G\) is a threshold graph if it can be built starting from the null graph by the two operations of adding an isolated vertex and adding a vertex adjacent to all current vertices; and since threshold graphs are 0-dominating, we have a third characterization:

2.3 \(G\) is 0-dominating if and only if \(G\) is a threshold graph.

The characterization 2.2 is the most useful for our purposes.

3 A matrix problem

Let us denote the set of all pairs \((i,j)\) \((1 \leq i \leq m, 1 \leq j \leq n)\) by \([m] \times [n]\). A subset \(F\) of \([m] \times [n]\) is up-closed if for all \(i,i',j,j'\) with \(1 \leq i \leq i' \leq m\) and \(1 \leq j \leq j' \leq n\), if \((i,j) \in F\) then \((i',j') \in F\); and \(F\) is down-closed if \((([m] \times [n]) \setminus F)\) is up-closed. The union of two up-closed subsets is up-closed, and so every subset of \([m] \times [n]\) has a unique maximal up-closed subset.

Now let \(A = (a_{ij} : (i,j) \in [m] \times [n])\) be a 0/1 matrix. Its support is the set of all pairs \((i,j)\) \([m] \times [n]\) with \(a_{ij} = 1\). We say that \(A\) is monotone if its support is up-closed. For \(i,i' \in \{1, \ldots, m\}\) we say that \(i,i'\) are row-comparable in \(A\) if one of the sets \(\{\bar{j} \in \{1, \ldots, n\} : a_{ij} \neq 0\}, \{\bar{j} \in \{1, \ldots, n\} : a_{ij} = 0\}\) is a subset of the other (and we define column-comparability similarly). We say that the matrix \(A\) is comparable if all \(i,i' \in \{1, \ldots, m\}\) are row-comparable in \(A\) (and therefore all \(j,j' \in \{1, \ldots, n\}\) are column-comparable in \(A\)). It follows that \(A\) is comparable if and only its rows and columns can be reordered to make a matrix that is monotone.

If \(t \geq 0\) is an integer, we say that \(A\) is \(t\)-restricted if

- for all \(i,i' \in \{1, \ldots, m\}\) with \(i < i'\), there are at most \(t\) values of \(j \in \{1, \ldots, n\}\) such that \(a_{ij} = 1\) and \(a_{ij'} = 0\); and
- for all \(j,j' \in \{1, \ldots, n\}\) with \(j < j'\), there are at most \(t\) values of \(i \in \{1, \ldots, m\}\) such that \(a_{ij} = 1\) and \(a_{ij'} = 0\).
Thus a matrix $A$ is 0-restricted if and only if it is monotone, that is, its support is up-closed. In order to prove 1.1 we need to prove something similar for $t$-restricted matrices; but before that we handle a special case.

Let $A, B$ be two 0/1 matrices both indexed by $[m] \times [n]$. The edit distance between $A, B$ is the maximum, over all rows and columns, of the number of terms in that row or column in which $A, B$ differ. One might hope that for all $t$, every $t$-restricted 0/1 matrix $A$ has bounded edit distance from some monotone matrix; but that is false, even for $t = 1$. For instance, with $m = 1$ and $n$ large, a $1 \times n$ matrix $A$ with entries $(1, \ldots, 1, 0, \ldots, 0)$ (with $n/2$ ones and $n/2$ zeroes) is 1-restricted, and yet has arbitrarily large edit distance from every monotone matrix. This is perhaps unfair in that being 1-restricted has no content for a $1 \times n$ matrix, but we could pad it by adding more rows; say $n/2$ rows of all zeroes, then the given row, and then $n/2$ rows of all ones, making an $(n + 1) \times n$ matrix, which is also a counterexample. One might try assuming in addition that the rows are in increasing order of row-sum, and the same for columns (which will be the case when we apply these results to our graph problem), but a similar counterexample still can be made, as follows. Take the $(n + 1) \times n$ matrix just described, and change the $n/2$ entries $a_{\frac{n}{2} + i, i}$ ($1 \leq i \leq n/2$) to ones and the $n/2$ entries $a_{\frac{n}{2} + i + 1, i}$ ($1 \leq i \leq n/2$) to zeroes.

Nevertheless, something like this is true; replace “monotone” by “comparable”. We will prove:

3.1 For all integers $t \geq 0$ and every $t$-restricted 0/1 matrix $A$, there is a comparable matrix $B$, such that the edit distance between $A, B$ is at most $644t^4$.

But before we prove this, we handle a special case, matrices of bounded “breadth”, and next we define this. Let $A = (a_{ij} : (i, j) \in E)$ be a 0/1 matrix, with support $E_1$, and let $E_0 = ([m] \times [n]) \setminus E_1$. Let $X$ be the maximal down-closed subset of $E_0$, and $Y$ the maximal up-closed subset of $E_1$. Thus $X \cap Y = \emptyset$. A diagonal for $A$ means a subset $F$ of $[m] \times [n]$ such that for some integer $c$, $F$ is the set of all pairs $(i, j) \in [m] \times [n]$ with $j = i + c$ and $(i, j) \notin X \cup Y$. It is easy to see that every diagonal $F$ is an interval in the sense that if $(i, i + c), (i'', i'' + c) \in F$ where $i'' > i$ then $(i', i' + c) \notin F$ for all $i'$ with $i \leq i' \leq i''$. We call the maximum cardinality of all diagonals the breadth of $A$. We first show:

3.2 For all integers $t, w \geq 0$ with $2w \geq t + 1$, and every $t$-restricted 0/1 matrix $A$ with breadth at most $w$, there is a comparable matrix $B$ such that the edit distance between $A, B$ is at most $2(t + w)^3$.

Proof. Let $A$ be a $t$-restricted 0/1 matrix indexed by $[m] \times [n]$, with breadth at most $w$, let its support be $E_1$, let $E_0 = ([m] \times [n]) \setminus E_1$, and let $X, Y$ be the maximal down-closed subset of $E_0$ and up-closed subset of $E_1$ respectively. Let $Z = ([m] \times [n]) \setminus (X \cup Y)$. We will give rules to change some of the entries $a_{ij}$ for $(i, j) \in Z$. We need to satisfy two conditions:

- after the changes, all pairs $i, i'$ will be row-comparable; and
- for each row or column, at most $(2t + w)w^3$ entries in that row or column will be changed.

Our first task is to give the rules, but that needs a number of definitions. For $1 \leq j \leq n$, the $j$th column means the set $\{(i, j) : 1 \leq i \leq m\}$ for some $j \in \{1, \ldots, n\}$, and a row is defined similarly. A post is the nonempty intersection of a column with $Z$, and a beam is the nonempty intersection of a row with $Z$. Every member of $Z$ belongs to a unique post and a unique beam, and how we change the corresponding entry of $A$ depends on the types of this post and beam. (Posts and beams again are intervals, in the natural sense.)
Let \( P, P' \) be posts, where \( P, P' \) are the intersection of the \( j \)th column with \( Z \) and the \( j' \)th column with \( Z \), respectively. They are parallel if

- for all \( i \in \{1, \ldots, m\} \), \((i, j) \in Z\) if and only if \((i, j') \in Z\); and
- for all \( i \in \{1, \ldots, m\} \) if \((i, j) \in Z\) then \( a_{ij} = a_{ij'} \).

Thus the two posts are identical in that they involve the same rows, and have the same entries in those rows. The multiplicity of a post \( P \) is the number of posts that are parallel to \( P \) (counting \( P \) itself). We define parallelness and multiplicity for beams similarly.

Let \((i, j) \in Z\). We associate four integers with \((i, j)\):

- \( p^-(i, j) \), the number of \( i' < i \) such that \((i', j) \in Z\);
- \( q^-(i, j) \), the number of \( j' < j \) such that \((i, j') \in Z\);
- \( p^+(i, j) \), the number of \( i' > i \) such that \((i', j) \in Z\);
- \( q^+(i, j) \), the number of \( j' > j \) such that \((i, j') \in Z\).

We observe that the post containing \((i, j)\) has cardinality \( p^-(i, j) + p^+(i, j) + 1 \), and a similar statement holds for the beam.

1. For each \((i, j) \in Z\), at least one of \( p^-(i, j), q^+(i, j) < w \), and at least one of \( q^-(i, j), p^+(i, j) < w \).
2. For every post \( P \), if \((i, j) \in P\) then \( q^-(i, j) + q^+(i, j) + 1 \) is at least the multiplicity of \( P \).

The beam containing \((i, j)\) has cardinality \( q^-(i, j) + q^+(i, j) + 1 \); but it intersects all posts parallel to \( P \), and so its cardinality is at least the multiplicity of \( P \). This proves (2).

Now we can give the rules. For each \((i, j) \in [m] \times [n]\) define \( b_{ij} \) as follows. If \((i, j) \notin Z\) then \( b_{ij} = a_{ij} \), so we may assume that \((i, j) \in Z\); let \((i, j)\) belong to a post \( P \) and a beam \( Q \).

- If both \( p^-(i, j), q^-(i, j) \geq w \) then \( b_{ij} = 1 \);
- if both \( p^+(i, j), q^+(i, j) \geq w \) then \( b_{ij} = 0 \);
- if both \( p^-(i, j), p^+(i, j) < w \) and \( P \) has multiplicity at least \( 2w \) then \( b_{ij} = a_{ij} \);
- if both \( q^- (i, j), q^+(i, j) < w \) and \( Q \) has multiplicity at least \( 2w \) then \( b_{ij} = a_{ij} \);
- if both \( p^- (i, j), p^+(i, j) < w \) and \( P \) has multiplicity less than \( 2w \) then \( b_{ij} = 0 \);
- if both \( q^- (i, j), q^+(i, j) < w \) and \( Q \) has multiplicity less than \( 2w \) then \( b_{ij} = 0 \).
We claim the rules are consistent; for let \((i, j) \in Z\). By (1), only one of the first two rules applies to \((i, j)\), and if one of the first two applies then none of the other rules apply. If say the third rule applies to \((i, j)\), then \(P\) has multiplicity at least \(2w\), and so by (2) \(q^- (i, j) + q^+(i, j) + 1 \geq 2w\), so at least one of \(q^- (i, j), q^+(i, j) \geq w\) and consequently the fourth, fifth and sixth rules do not apply; and similarly if the fourth rule applies then the fifth and sixth do not. Finally, both the fifth and sixth may apply simultaneously, but they assign the same value to \(b_{ij}\). Thus the rules are consistent. Furthermore, we observe that every \((i, j) \in Z\) falls under one of the rules, by (1), and so the matrix \(B = (b_{ij})\) is well-defined.

Now we must check the two bullets given at the start of this proof. We say \(B:\) is increasing in row \(i\) if \(b_{ij} \leq b_{ij'}\) for all \(j, j'\) with \(1 \leq j < j' \leq n\), and we define increasing in column \(j\) similarly.

(3) Let \(P\) be a post, a subset of the \(j\)th column. If \(|P| < 2w\) then \(B\) is increasing in column \(j\).

Suppose that there exist \(i, i'\) with \(1 \leq i < i' \leq m\), such that \(b_{ij} = 1\) and \(b_{i'j} = 0\). It follows that \((i, j) \notin X\), and so \((i', j) \notin X\), since \(X\) is down-closed; and similarly \((i, j), (i', j) \notin Y\), and consequently \((i, j), (i', j) \in Z\). Let \(Q, Q'\) be the beams containing \((i, j), (i', j)\) respectively. Since \(|P| < 2w\), both \(Q, Q'\) have multiplicity less than \(2w\). Since \(b_{ij} = 1\), not both \(q^-(i, j), q^+(i, j) < w\). If \(p^+(i, j) \geq w\), then by (1) \(q^-(i, j) < w\), so \(q^+(i, j) \geq w\) and \(b_{ij} = 0\), a contradiction. Thus \(p^+(i, j) < w\). Suppose that also \(p^-(i, j) < w\). Since \(b_{ij} = 1\) it follows that \(P\) has multiplicity at least \(2w\), and \(a_{ij} = b_{ij} = 1\). If \(a_{ij} = 0\), then since \(P\) has multiplicity at least \(2w \geq t + 1\), there are \(t + 1\) values of \(j' \in \{1, \ldots, n\}\) such that \(a_{ij'} = 1\) and \(a_{ij'} = 0\), contradicting that \(A\) is \(t\)-restricted. Thus \(a_{ij} = 1\), and in particular \(a_{ij} \neq b_{ij}\). Since \(P\) has multiplicity at least \(2w\) it follows that one of \(p^-(i', j), p^+(i', j) \geq w\), and hence \(p^-(i', j) \geq w\), since \(p^+(i', j) < p^+(i, j) < w\). Moreover, since \(P\) has multiplicity at least \(2w\), (2) implies that \(q^-(i', j) + q^+(i', j) + 1 \geq 2w\), and so one of \(q^-(i', j), q^+(i', j) \geq w\). Since \(b_{ij} = 0\), this contradicts the rules. This proves that \(p^-(i, j) \geq w\). We have seen that not both \(q^-(i, j), q^+(i, j) < w\), and so from (1), \(q^-(i, j) \geq w\). But \(q^-(i', j) \geq q^-(i, j)\) since \(X\) is down-closed, and so \(q^-(i', j) \geq w\); and also \(p^-(i', j) > p^-(i, j) \geq w\), and yet \(b_{ij} = 0\), contrary to the rules. This proves (3).

(4) Let \(1 \leq i < i' \leq m; \) then \(i, i'\) are row-comparable in \(B\).

Suppose not, then in particular, there exists \(j \in \{1, \ldots, n\}\) such that \(b_{ij} = 1\) and \(b_{ij} = 0\). Now \((i, j) \notin X\) since \(b_{ij} = 1\), and since \(X\) is down-closed it follows that \((i', j) \notin X\). Also \((i, j) \notin Y\) since \(b_{ij} = 0\), and so \((i, j) \notin Y\) since \(Y\) is up-closed. Consequently \((i, j), (i', j) \in Z\). Let \(P\) be the post containing them both, and let \(Q, Q'\) be the beams containing \((i, j), (i', j)\) respectively. Since \(B\) is not increasing in the \(j\)th column, \(|P| \geq 2w\) by (3). Also, since \(i, i'\) are not row-comparable in \(B\), \(B\) is not increasing in one of the \(i\)th row or the \(i'\)th row; and so one of \(|Q|, |Q'| \geq 2w\) by (3).

Suppose that \(|Q| \geq 2w\). Then one of \(q^-(i, j), q^+(i, j) \geq w\), and one of \(p^-(i, j), p^+(i, j) \geq w\), and since \(b_{ij} = 1\), it follows from (1) and the rules that \(p^-(i, j), q^+(i, j) \geq w\). But \(q^-(i', j) \geq q^-(i, j)\) since \(X\) is down-closed, and so \(q^-(i', j) \geq w\); and also \(p^-(i', j) > p^-(i, j) \geq w\), and yet \(b_{ij} = 0\), contrary to the first rule. Thus \(|Q| < 2w\), and so \(|Q'| \geq 2w\).

Hence one of \(q^-(i', j), q^+(i', j) \geq w\), and one of \(p^-(i', j), p^+(i', j) \geq w\), and since \(b_{ij} = 0\), it follows from (1) and the rules that \(p^+(i', j), q^+(i', j) \geq w\). But \(q^+(i', j) \geq q^+(i', j)\) since \(X\) is down-closed, and so \(q^+(i, j) \geq w\); and also \(p^+(i, j) > p^+(i', j) \geq w\), and yet \(b_{ij} = 1\), contrary to the rules. This
proves (4).

(5) The edit distance between $A, B$ is at most $2(t + w)w^3$.

Let $1 \leq j \leq n$; by the symmetry between axes, it suffices to show that there are at most $(2t + w)w^3$ values of $i$ such that $a_{ij} \neq b_{ij}$. For every such value of $i$, it follows that $(i, j) \in \mathbb{Z}$; so we may assume that there is a post $P$ included in the $j$th column with $|P| > (2t + w)w^3$, and choose $i_1, i_2$ with $1 \leq i_1 \leq i_2 \leq m$ such that

- for $1 \leq i < i_1$, $(i, j) \in X$;
- for $i_1 \leq i \leq i_2$, $(i, j) \in P$; and
- for $i_2 < i \leq m$, $(i, j) \in Y$.

Let $i_1 \leq i \leq i_2$, and suppose that $a_{ij} = 0$ and $b_{ij} = 1$. From the rules, it follows that $p^-(i, j), q^-(i, j) \geq w$; and so from (1), $p^+(i, j) < w$. Consequently $i_2 - w + 1 \leq i \leq i_2$, and so there are at most $w$ such values of $i$. Now let $i_1 \leq i \leq i_2$, and suppose that $a_{ij} = 1$ and $b_{ij} = 0$. From the rules, it follows that either

- $p^+(i, j), q^+(i, j) \geq w$, or
- both $p^-(i, j), p^-(i, j) < w$ and $P$ has multiplicity less than $2w$, or
- both $q^-(i, j), q^+(i, j) < w$ and $Q_i$ has multiplicity less than $2w$, where $Q_i$ is the beam containing $(i, j)$.

At most $w$ values of $i$ satisfy the first bullet, as before. If $i$ satisfies the second bullet then $|P| \leq 2w - 1$, a contradiction since $|P| > (2t + w)w^3$; so we may assume that $i$ satisfies the third bullet. There are therefore only $w^2$ possibilities for the pair $q^-(i, j), q^+(i, j)$. Let $h, k \geq 0$ with $h, k < w$, and let $I(h, k)$ be the set of all $i$ that satisfy $i_1 \leq i \leq i_2$, and $a_{ij} = 1$, and $b_{ij} = 0$, and $q^-(i, j) = h$ and $q^+(i, j) = k$, and $Q_i$ has multiplicity less than $2w$. We need to bound $|I(h, k)|$. We can evidently get an exponential bound, since there are only $2^{2w+1}$ possibilities for the entries of $Q_i$, and only $2w - 1$ values of $i \in I(h, k)$ in which all the entries are the same; but we can do better.

If $i \in I(h, k)$, $B$ may or may not be increasing in the $i$th row. Let $I_1$ be the set of $i \in I(h, k)$ such that $B$ is increasing in the $i$th row, and $I_2 = I(h, k) \setminus I_1$. For each $i \in I_1$, there are only $h + 1$ possibilities for the entries of $Q_i$, and at most $2w - 1$ distinct $i \in I_1$ with the same set of entries, and so $|I_1| \leq (h + 1)(2w - 1)$. If $i \in I_2$ then there exists $j'$ with $j - h \leq j' \leq j + k - 1$ such that $a_{ij'} = 1$ and $a_{i,j'+1} = 0$. There are only $h + k$ choices for $j'$, and for each such $j'$ there are at most $t$ values of $i \in I_2$ such that $a_{ij'} = 1$ and $a_{i,j'+1} = 0$, since $A$ is $t$-restricted; and so $|I_2| \leq t(h + k)$. Consequently $|I(h, k)| \leq (h + 1)(2w - 1) + t(h + k)$. Since $h, k < w$, it follows that $|I(h, k)| \leq w(2w - 1) + 2(w - 1)t$.

In summary, at most $w$ values of $i$ satisfy the first bullet above; none satisfy the second bullet; and (since there are at most $w^2$ choices for $h, k$), at most $(w(2w - 1) + 2(w - 1)t)w^3$ satisfy the third. Consequently there are at most $(w(2w - 1) + 2(w - 1)t)w^2 + w$ values of $i$ such that $a_{ij} = 1$ and $b_{ij} = 0$. As we saw, there are at most $w$ values of $i$ such that $a_{ij} = 0$ and $b_{ij} = 1$; so there are at most $(w(2w - 1) + 2(w - 1)t)w^2 + 2w$ values of $i$ such that $a_{ij} \neq b_{ij}$. This proves (5).

From (4) and (5), this proves 3.2.
3.3 For all integers \( t \geq 0 \) and every \( t \)-restricted \( 0/1 \) matrix \( A \), there is a \( t \)-restricted \( 0/1 \) matrix \( B \) with breadth at most \( 4t \), such that the edit distance between \( A, B \) is at most \( 4t \).

**Proof.** Let \( A \) be a \( 0/1 \) matrix indexed by \([m] \times [n]\), let its support be \( E_1 \), let \( E_0 = ([m] \times [n]) \setminus E_1 \), and let \( X, Y \) be the maximal down-closed subset of \( E_0 \) and up-closed subset of \( E_1 \) respectively. Let \( Z = ([m] \times [n]) \setminus (X \cup Y) \). Let \( P_i \) be the set of all \((i, j) \in [m] \times [n]\) such that there exist at least \( 2t \) values of \( i' < i \) with \((i', j) \in E_1 \), and let \( Q_j \) be the set of all \((i, j) \) such that there are at least \( 2t \) values of \( j' < j \) with \((i, j') \in E_1 \). Let \( P_0 \) be the set of all \((i, j) \) such that there exist at least \( 2t \) values of \( i' > i \) with \((i', j) \in E_0 \), and let \( Q_0 \) be the set of all \((i, j) \) such that there are at least \( 2t \) values of \( j' > j \) with \((i, j') \in E_0 \).

1. There do not exist \((i_1, j_1) \in P_1 \) and \((i_0, j_0) \in Q_0 \) such that \( i_0 \geq i_1 \) and \( j_0 \geq j_1 \). Also, there do not exist \((i_1, j_1) \in Q_1 \) and \((i_0, j_0) \in P_0 \) such that \( i_0 \geq i_1 \) and \( j_0 \geq j_1 \).

Suppose that such \((i_1, j_1), (i_0, j_0) \) exist. From the symmetry between axes, we may assume that \((i_1, j_1) \in P_1 \), and hence \((i_0, j_0) \in Q_0 \). Consequently \((i_0, j_1) \in P_1 \cap Q_0 \). Choose \( I_1 \subseteq \{1, \ldots, i_0 - 1\} \) with cardinality \( 2t \) such that \((i, j_1) \in E_1 \) for each \( i \in I_1 \), and choose \( J_0 \subseteq \{j_1 + 1, \ldots, n\} \) with cardinality \( 2t \) such that \((i_0, j) \in E_0 \) for each \( j \in J_0 \). From the symmetry between zeroes and ones, we may assume without loss of generality that \((i_0, j_1) \in E_0 \). If \( E_1 \) contains at least half of the pairs \((i, j) \) with \( i \in I_1 \) and \( j \in J_0 \), then there exists \( i \in I_1 \) such that \((i, j) \in E_1 \) for at least \( t \) values of \( j \in J_0 \), and hence for at least \( t + 1 \) values of \( j \in J_0 \cup \{j_0\} \); and since \((i_0, j_0) \in E_0 \) for each \( j \), this contradicts that \( A \) is \( t \)-restricted. On the other hand, if \( E_0 \) contains more than half of the pairs \((i, j) \) with \( i \in I_1 \) and \( j \in J_0 \), then there exists \( j \in J_0 \) such that \((i, j) \in E_0 \) for at least \( t + 1 \) values of \( i \in I_1 \); and since \((i_1, j) \in E_1 \) for each \( i \), this also contradicts that \( A \) is \( t \)-restricted. This proves (1).

2. There do not exist \((i_1, j_1) \in P_1 \cup Q_1 \) and \((i_0, j_0) \in P_0 \cup Q_0 \) such that \( i_0 \geq i_1 + 4t \) and \( j_0 \geq j_1 + 4t \).

Suppose that such \((i_1, j_1), (i_0, j_0) \) exist. From the symmetry between axes, we may assume that \((i_1, j_1) \in P_1 \). By (1) it follows that \((i_0, j_0) \notin Q_0 \), and so \((i_0, j_0) \in P_0 \). Choose \( I_1 \subseteq \{1, \ldots, i_1 - 1\} \) with cardinality \( 2t \) such that \((i, j_1) \in E_1 \) for each \( i \in I_1 \), and choose \( I_0 \subseteq \{i_0 + 1, \ldots, m\} \) with cardinality \( 2t \) such that \((i_0, j) \in E_0 \) for each \( i \in I_0 \). If \( E_0 \) contains more than half of the pairs \((i, j) \) with \( i \in I_1 \) and \( j_1 < j < j_0 \), then there exists \( j \) with \( j_1 < j < j_0 \) such that \((i, j) \in E_0 \) for at least \( t + 1 \) values of \( i \in I_1 \), contradicting that \( A \) is \( t \)-restricted. If \( E_1 \) contains at least half of the pairs \((i, j) \) with \( i \in I_1 \) and \( j_1 < j < j_0 \), then there exists \( i \in I_1 \) such that \((i, j) \in E_1 \) for at least half of the values of \( j \) with \( j_1 < j < j_0 \), and since \((i, j_1) \in E_1 \), it follows that \((i, j_0) \in Q_1 \), contrary to (1) since \((i_0, j_0) \in P_0 \). This proves (2).

Construct a matrix \( B = (b_{ij}) \) as follows. Let \((i, j) \in [m] \times [n] \).

- If \((i, j) \in Z \) and there is no \((i_0, j_0) \in P_0 \cup Q_0 \) such that \( i_0 \geq i \) and \( j_0 \geq j \) then \( b_{ij} = 1 \);
- if \((i, j) \in Z \) and there exists \((i_0, j_0) \in P_0 \cup Q_0 \) such that \( i_0 \geq i \) and \( j_0 \geq j \), and there is no \((i_1, j_1) \in P_1 \cup Q_1 \) such that \( i_1 \leq i \) and \( j_1 \leq j \), then \( b_{ij} = 0 \);
- if neither of these applies then \( b_{ij} = a_{ij} \).
We need to show that $B$ is $t$-restricted, $B$ has breadth at most $4t$, and the edit distance between $A, B$ is at most $4t$. Let the support of $B$ be $F_1$, let $F_0 = ([m] \times [n]) \setminus F_1$, and let $X', Y'$ be the maximal down-closed subset of $F_0$ and up-closed subset of $F_1$ respectively.

(3) For all $(i,j) \in [m] \times [n]$, if there is no $(i_0, j_0) \in F_0 \cup Q_0$ such that $i_0 \geq i$ and $j_0 \geq j$ then $(i,j) \in Y'$. Also, if there is no $(i_1, j_1) \in P_1 \cup Q_1$ such that $i_1 \leq i$ and $j_1 \leq j$, then $(i,j) \in X'$.

Suppose there is no $(i_0, j_0) \in F_0 \cup Q_0$ such that $i_0 \geq i$ and $j_0 \geq j$. Let $F$ be the set of all $(i', j') \in [m] \times [n]$ such that $i \leq i'$ and $j \leq j'$; then for every such pair, $(i', j') \in Z \cup Y$, and there is no $(i_0, j_0) \in P_0 \cup Q_0$ such that $i_0 \geq i'$ and $j_0 \geq j'$, and so $(i', j') \in F_1$. Hence $F$ is an up-closed subset of $F_1$, and since $Y'$ is the unique maximal up-closed subset of $F_1$, it follows that $F \subseteq Y'$, and in particular $(i,j) \in Y'$ as required. The second statement follows similarly. This proves (3).

(4) $B$ is $t$-restricted.

Suppose that there exist $i, i'$ with $1 \leq i < i' \leq m$ such that $b_{ij} = 1$ and $b_{i'j} = 0$ for $t + 1$ values of $j \in \{1, \ldots, n\}$. For each such $j$, $(i,j) \notin Y'$ since $(i', j) \in F_0$; $(i,j) \notin X'$ since $(i,j) \notin F_0$; and similarly $(i', j) \notin X', Y'$. By (3), $b_{ij} = a_{ij}$ and $b_{i'j} = a_{i'j}$, contradicting that $A$ is $t$-restricted. From the symmetry between axes, this proves (4).

(5) $B$ has breadth at most $4t$.

Let $Z' = ([m] \times [n]) \setminus (X' \cup Y')$. Since $X \subseteq X'$ and $Y \subseteq Y'$, it follows that $Z' \subseteq Z$. Suppose that for some $i, j$ all of the pairs $(i + h, j + h)$ ($0 \leq h \leq 4t$) belong to $Z'$. Thus $(i,j) \in Z$, and since $(i,j) \notin X'$, it follows from (3) that there exists $(i_1, j_1) \in P_1 \cup Q_1$ such that $i_1 \leq i$ and $j_1 \leq j$. Similarly, since $(i+4t, j+4t) \notin Y'$, there exists $(i_0, j_0) \in P_0 \cup Q_0$ such that $i_0 \geq i+4t$ and $j_0 \geq j+4t$. But then $i_0 \geq i+4t$ and $j_0 \geq j+4t$, contrary to (2). This proves (5).

(6) The edit distance between $A, B$ is at most $4t$.

Let $1 \leq i \leq m$, and let $J$ be the set of all $j$ with $1 \leq j \leq n$ such that $b_{ij} \neq a_{ij}$. We need to bound $|J|$. Let $J_1$ be the set of all $j \in J$ such that $b_{ij} = 1$ and $a_{ij} = 0$, and let $J_0$ be the set of all $j \in J$ such that $b_{ij} = 0$ and $a_{ij} = 1$. If $j \in J_1$, then there is no $(i_0, j_0) \in P_0 \cup Q_0$ such that $i_0 \leq i$ and $j_0 \geq j$, and in particular $(i,j) \notin Q_0$; and so $(i,j) \in F_0$ for at most $2t - 1$ values of $j' < j$. Since $J_1 \subseteq E_0$, it follows (by choosing $j \in J_1$ maximal, if possible) that $|J_1| \leq 2t$.

If $j \in J_0$, then there is no $(i_1, j_1) \in P_1 \cup Q_1$ such that $i_1 \leq i$ and $j_1 \leq j$, and in particular $(i,j) \notin Q_1$, and so $(i,j) \in E_1$ for at most $2t - 1$ values of $j' > j$; and so $|J_0| \leq 2t$. Summing, it follows that $|J| \leq 4t$. From the symmetry between axes, this proves (6).

From (4), (5), (6), this completes the proof of 3.3.

Combining the previous two theorems, we deduce 3.1, which we restate:

**3.4** For all integers $t \geq 0$ and every $t$-restricted 0/1 matrix $A$, there is a comparable matrix $B$, such that the edit distance between $A, B$ is at most $644t^4$. 

Proof. By 3.3 there is a $t$-restricted 0/1 matrix $A'$ with breadth at most $4t$, such that the edit distance between $A,A'$ is at most $4t$. By 3.2 applied to $A'$, with $w = 4t$, there is a comparable matrix $B$ such that the edit distance between $A',B$ is at most $640t^4$. But then the edit distance between $A,B$ is at most $640t^4 + 4t \leq 644t^4$. This proves 3.4.

4 Near-dominating graphs

Now we use the results of the previous section to study $t$-dominating graphs. First we show:

4.1 Let $G$ be a $t$-dominating split graph. Then there is a split half-graph $H$ with $V(H) = V(G)$ such that the edit distance between $G,H$ is at most $644t^4$.

Proof. Let $G$ be a $t$-dominating split graph, and let $V(G)$ be the disjoint union of a clique $M$ and a stable set $N$. Let $M = \{u_1, \ldots, u_m\}$, where for $1 \leq i < j \leq m$ the degree of $u_i$ is at most that of $u_j$, and let $N = \{v_1, \ldots, v_n\}$, where for $1 \leq i < j \leq n$ the degree of $v_i$ is at most that of $v_j$. Let $A = (a_{ij})$ be the 0/1 matrix where $a_{ij} = 1$ if $u_i,v_j$ are adjacent and $a_{ij} = 0$ otherwise. We claim that $A$ is $t$-restricted. For let $1 \leq i < i' \leq m$, and suppose that there are $t$ values of $j \in \{1, \ldots, n\}$ such that $a_{ij} = 1$ and $a_{ij'} = 0$. Thus for each such $j$, $v_j$ is adjacent to $u_i$ and not to $u_{i'}$. Since the degree of $u_{i'}$ is at least that of $u_i$, it follows that there are at least $t$ vertices adjacent to $u_{i'}$ and not to $u_i$, contradicting that $G$ is $t$-dominating. A similar argument shows that there do not exist $j,j' \in \{1, \ldots, n\}$ with $j < j'$ such that for $t$ values of $i$, $a_{ij} = 1$ and $a_{ij'} = 0$. Hence $A$ is $t$-restricted. By 3.4, there is a comparable matrix $B$ such that the edit distance between $A,B$ is at most $644t^4$. Let $H$ be the graph with vertex set $M \cup N$ in which $u_i,v_j$ are adjacent if $b_{ij} = 1$, and $M$ is a clique and $N$ is a stable set. It follows that $H$ is a split half-graph, and its edit distance from $G$ is at most $644t^4$. This proves 4.1.

4.2 Let $G$ be a $t$-dominating graph. Then there is a $t$-dominating split graph $H$ with $V(H) = V(G)$ and with edit distance at most $2t$ from $G$.

Proof. Let $V(G) = \{v_1, \ldots, v_n\}$, where for $1 \leq i < j \leq n$, the degree of $v_j$ is at least that of $v_i$. Choose $i$ minimum such that some vertex in $\{v_1, \ldots, v_i\}$ has at least $2t+1$ neighbours in $\{v_1, \ldots, v_i\}$. (If this is not possible then every vertex has degree at most $2t$, and so $G$ has edit distance at most $2t$ from the graph $H$ obtained by deleting all the edges, which is a $t$-dominating split graph.) Let $N = \{v_1, \ldots, v_{i-1}\}$ and $M = \{v_i, \ldots, v_n\}$. Thus every vertex in $N$ has at most $2t$ neighbours in $N$. Choose $x \in N \cup \{v_i\}$ and $X \subseteq N \setminus \{x\}$ with $|X| \geq 2t$ (and $X \neq \emptyset$ if $t = 0$), such that $X$ is the set of neighbours of $x$ in $N$. For all $j \geq i$, there are at most $t$ vertices adjacent to $x$ and not to $v_j$, since the degree of $v_j$ is at least that of $x$ and $G$ is $t$-dominating. In particular, $v_j$ is adjacent to at least half the vertices in $X$. Let $j \geq i$, and let $Y$ be the set of vertices in $M$ that are different from and nonadjacent to $v_j$. Since every vertex in $Y$ is adjacent to at least half the vertices in $X$, and $X \neq \emptyset$, some vertex $v_h \in X$ is adjacent to at least half the vertices in $Y$. But there are at most $t$ vertices adjacent to $v_h$ and not to $v_j$, since the degree of $v_h$ is at most that of $v_j$ and $G$ is $t$-dominating; and consequently $|Y|/2 \leq t$. Thus every vertex $v \in M$ is nonadjacent to at most $2t$ vertices in $M \setminus \{v\}$.

Let $H$ be the split graph obtained from $G$ by deleting all edges with both ends in $N$ and making adjacent all pairs of vertices in $M$; then the edit distance between $G,H$ is at most $2t$. 9
We claim that $H$ is $t$-dominating. For let $u, v \in V(H) = V(G)$. If $u, v \in M$ then since one of $u, v$ $t$-dominates the other in $G$, the same is true in $H$ (since in $H$, $u, v$ have the same neighbours in $M \setminus \{u, v\}$). Similarly if $u, v \in N$ then one of $u, v$ $t$-dominates the other. If $u \in M$ and $v \in N$ then $v$ 0-dominates and hence $t$-dominates $u$. This proves that $H$ is $t$-dominating, and so proves 4.2.

Combining 4.1 and 4.2, we obtain our main result, which we restate:

**4.3** Let $G$ be a $t$-dominating graph. Then there is a split half-graph $H$ with $V(H) = V(G)$ and with edit distance at most $646t^4$ from $G$.

**Proof.** Let $G$ be $t$-dominating. By 4.2 there is a $t$-dominating split graph $G'$ with $V(G') = V(G)$ and with edit distance at most $2t$ from $G$. By 4.1, there is a split half-graph $H$ with $V(H) = V(G)$ such that the edit distance between $G', H$ is at most $644t^4$. Thus the edit distance between $G, H$ is at most $644t^4 + 2t \leq 646t^4$. This proves 4.3.

**5 A counterexample**

Here is a concept similar to $t$-domination for graphs and $t$-restriction for matrices. Let $G$ be a bipartite graph with bipartition $(A, B)$; we say $G$ is $t$-nested on $(A, B)$ if for all distinct $u, v \in B$, one of them has at most $t$ neighbours nonadjacent to the other. Thus, 0-nested graphs are half-graphs.

Let $G, H$ both be bipartite graphs with bipartition $(A, B)$; we say the bipartite edit distance between $G, H$ is the maximum over $v \in B$ of $|N_G(v) \setminus N_H(v)| + |N_H(v) \setminus N_G(v)|$. One might hope that for all $t$ there exists $f(t)$ such that if $G$ is $t$-nested on $(A, B)$ then there is a 0-nested bipartite graph $H$ with the same bipartition $(A, B)$, with bipartite edit distance at most $f(t)$ from $G$. (Indeed, this looked like an easier question than our main result, and we tried it first as a warm-up.) But this is false, even if $t = 1$. Here is a counterexample.

Let $S$ be the set of all 0/1 sequences with at most $k$ terms, and let $T$ be the binary tree defined by $S$; thus $V(T) = S$, and $s, s'$ are adjacent if their lengths differ by one and one of them is an initial subsequence of the other. Let $I$ be the set of all $s \in S$ of length less than $k$, and $B$ those of length exactly $k$. (Thus $B$ is the set of leaves of $T$, and $I$ the set of internal vertices.) Each $s \in B$ is a sequence of $k$ terms $(s_1, \ldots, s_k)$ say, and we define $n(s) = \sum_{1 \leq i \leq k} 2^{k-i}$. Thus all the numbers $n(s)$ ($s \in B$) are different and range from 0 to $2^k - 1$. For $s, s' \in B$, we write $s < s'$ if $n(s) < n(s')$.

For each $s \in B$ take a set $W_s$ of $2^{k+1} + 2k$ new vertices, so that all the sets $V(T)$ and $W_s$ ($s \in B$) are pairwise disjoint. Let $W = \bigcup_{s \in B} W_s$, and let $A = I \cup W$. We construct a bipartite graph $G$ with bipartition $(A, B)$ as follows. Let $s \in B$. For $s' \in B$, $s$ has no neighbour in $W_{s'}$ if $s' > s$, and $s$ is adjacent to every vertex in $W_{s'}$ if $s' \leq s$. For $s' \in I$, we decide the adjacency of $s, s'$ in $G$ by the following rule. Choose $i$ minimum such that either $i$ is greater than the number of terms of $s'$, or the $i$th terms of $s$ and of $s'$ are different, and let $x$ be the $i$th term of $s$; we make $s, s'$ adjacent if and only if $x = 0$.

This graph is 1-nested; and indeed, for $s, s' \in B$ with $s < s'$, there is at most one $v \in I$ adjacent to $s$ and not to $s'$, and no such $v \in W$. Suppose that there is a 0-nested graph $H$ with bipartition $(A, B)$ that has bipartite edit distance less than $k$ from $G$. For each $s \in S$, its degree in $G$ is at least $(2^{k+1} + 2k)n(s)$ and at most $(2^{k+1} + 2k)n(s) + 2^{k+1} - 1$; and so its degree in $H$ is at least $(2^{k+1} + 2k)n(s) - k$ and at most $(2^{k+1} + 2k)n(s) + 2^{k+1} - 1 + k < (2^{k+1} + 2k)(n(s) + 1) - k$. 

10
Consequently, for distinct $s, s' \in B$, if $s < s'$ then the degree of $s$ in $H$ is less than that of $s$. Since $H$ is 0-nested, it follows that every vertex of $A$ adjacent to $s$ in $H$ is also adjacent to $s'$.

Let $t_0$ be the null sequence, and let $L_0, R_0$ be the sets of members of $B$ with first term 0 and 1 respectively. It follows that $t_0$ is adjacent in $G$ to every vertex in $L_0$, and nonadjacent to every vertex in $R_0$; and yet $s < s'$ for every $s \in L_0$ and $s' \in R_0$. Consequently, in $H$ either $t_0$ is nonadjacent to every member of $L_0$, or adjacent to every member of $R_0$. If the first let $t_1$ be the one-term sequence $(0)$, and otherwise $t_1 = (1)$. Let $L_1$ be the set of members of $B$ such that $t_1$ is an initial segment, and the second term is 0, and let $R_1$ be those such that $t_1$ is an initial segment with second term 1. Again, $t_1$ is adjacent in $G$ to every member of $L_1$, and nonadjacent to every member of $R_1$, and in $H$ either $t_1$ is nonadjacent to all vertices in $L_1$ or adjacent to all in $R_1$; let $t_2$ be the corresponding two-term sequence. By continuing this process we obtain a sequence $t_0, t_1, t_2, \ldots, t_{k-1} \in I$, and a vertex $t_k \in B$, where for $0 \leq i \leq k$ $t_i$ has $i$ terms and $t_{i-1}$ is an initial segment of $t_i$, and for $0 \leq i \leq k - 1$, $t_k$ is adjacent to $t_i$ in exactly one of $G, H$. Since $t_k \in B$, this contradicts that the bipartite edit distance between $G, H$ is at most $k - 1$.

This could be viewed another way: for each vertex in $B$, take its set of neighbours in $A$. Then we obtain a collection of subsets $\mathcal{C}$ of $A$ such that for every two of them, say $X, Y$, one of $|X \setminus Y|$, $|Y \setminus X| \leq 1$. But if we want to change this last 1 to a 0 by adding and subtract elements of $A$ from the sets of $\mathcal{C}$, then some set has to have an arbitrarily large number of elements added or subtracted.

6 The Erdős-Hajnal conjecture

Let us say an ideal of graphs is a class $\mathcal{C}$ of graphs, such that if $G \in \mathcal{C}$ and $H$ is isomorphic to an induced subgraph of $G$ then $H \in \mathcal{C}$; and an ideal is proper if some graph is not in it. The Erdős-Hajnal conjecture [4] asserts:

6.1 Conjecture: For every proper ideal $\mathcal{C}$, there exist $c, \epsilon > 0$ such that for every graph $G \in \mathcal{C}$, $G$ has a clique or stable set of cardinality at least $c|V(G)|^\epsilon$.

We are interested in the way the (optimal) coefficient $\epsilon$ depends on $\mathcal{C}$. In particular, when does taking $\epsilon = 1$ work? If $\mathcal{C}$ is the set of all graphs not containing one particular graph $H$ as an induced subgraph, then there are almost no choices of $H$ for which $\epsilon = 1$ works — only those graphs $H$ with at most two vertices, as is easily seen. But there are ideals defined by excluding more than one graph.

For instance, let $S_t$ be the star with centre of degree $t$, and $\mathcal{C}$ the class of all graphs that do not contain $S_t$ as an induced subgraph; then $\epsilon = O(1/t)$, since the Ramsey number $R(k, t)$ is at least $(k/\log k)^{(t+1)/2}$ for fixed $t$ and large $k$ [2] (and $\epsilon$ is known to exist because of a result of Alon, Pach and Solymosi [1]). On the other hand, if we take $\mathcal{C}$ to be the class of graphs that contain neither $S_t$ nor its complement as an induced subgraph, then [3] for every graph $G \in \mathcal{C}$, either $G$ or its complement has maximum degree bounded by a function of $t$, and so there exist $c, \epsilon$ with $\epsilon = 1$.

Another ideal of interest is the class of all $t$-dominating graphs, for fixed $t$. Every split graph has a clique or stable set containing at least half its vertices, and so by 4.2, in every $t$-dominating graph $G$, there is a subset $X \subseteq V(G)$ with $|X| \geq |V(G)|/2$ such that either the subgraph induced on $X$ has maximum degree at most $2t$ or its complement graph does. Consequently $G$ has a clique or stable set with cardinality at least $|V(G)|/(4t + 2)$; and so we may take $\epsilon = 1$ in 6.1 for this class.
Take a “substar” (a graph obtained from a star by deleting some edges) and the complement of a substar, and let \( \mathcal{C} \) be the class containing neither of these graphs; in this case \( \epsilon = 1 \) does not work in general (consider a disjoint union of \( n^{1/2} \) cliques each with \( n^{1/2} \) vertices), but \( \epsilon = 1/2 \) works. This is a special case of the following. (If \( G \) is a graph, \( \omega(G) \) and \( \alpha(G) \) denote the cardinalities of its largest clique and largest stable set respectively, and we denote \( \max(\omega(G), \alpha(G)) \) by \( \rho(G) \).)

6.2 Let \( H_1, H_2 \) be threshold graphs, with \( |V(H_1)| + |V(H_2)| = m \) and \( \omega(H_1) + \alpha(H_2) = k \). For every graph \( G \), if \( G \) has no induced subgraph isomorphic to \( H_1 \) or to \( H_2 \), then \( |V(G)| \leq (2^m - 1)\rho(G)^{k-2} \).

**Proof.** We proceed by induction on \( m \). If one of \( H_1, H_2 \) has at most one vertex the claim is trivial, so we assume they both have at least two vertices. A vertex is isolated if it has degree zero, and universal if it is adjacent to all other vertices. For \( i = 1, 2 \), let \( H'_i \) be obtained from \( H_i \) by deleting an isolated vertex if there is one, and if not let \( H'_i = H_i \); and let \( H''_i \) be obtained from \( H_i \) by deleting a universal vertex if there is one, and if not let \( H''_i = H_i \). Let \( G \) contain neither of \( H_1, H_2 \). For \( X \subseteq V(G) \), \( G[X] \) denotes the subgraph induced on \( X \). Let \( v \in V(G) \), and let \( N \) be the set of neighbours of \( v \), and \( M = V(G) \setminus (N \cup \{v\}) \).

Suppose first that one of \( H_1, H_2 \) has an isolated vertex, and one has a universal vertex. Since \( G[N] \) contains neither of \( H'_1, H'_2 \), it follows that \( |M| \leq (2^{m-1} - 1)\rho(G)^{k-2} \); and similarly \( |N| \leq (2^{m-1} - 1)\rho(G)^{k-2} \). Consequently

\[
|V(G)| = |N| + |M| + 1 \leq 2(2^{m-1} - 1)\rho(G)^{k-2} + 1 \leq (2^m - 1)\rho(G)^{k-2}
\]

as required.

By taking complements if necessary, we may therefore assume that neither of \( H_1, H_2 \) has an isolated vertex. It follows that \( H_1, H_2 \) both have universal vertices, since they are threshold graphs; and in particular, \( \omega(H''_1) + \alpha(H''_2) = k - 1 \).

Choose \( v \in V(G) \) with maximum degree, and let \( N \) be its set of neighbours. Since \( G[N] \) contains neither of \( H''_1, H''_2 \), it follows from the inductive hypothesis that

\[
|N| \leq (2^{m-2} - 1)\rho(G)^{k-3} \leq 2^{m-2}\rho(G)^{k-3} - 1.
\]

Since \( v \) has maximum degree, it follows that \( G \) is \((2m^2 - 1)\rho(G)^{k-3}\)-colourable, and so some stable set has cardinality at least \( |V(G)|/(2^{m-2}\rho(G)^{k-3}) \). In particular \( \rho(G) \) is at least this quantity, and so

\[
2^{m-2}\rho(G)^{k-2} \geq |V(G)|,
\]

and again the result follows. This proves 6.2. \( \square \)

One can also ask, what is the size of the largest clique or stable set in almost all graphs in an ideal? This may be much larger than we can guarantee for every graph in the ideal. Thus while for every \( \epsilon > 0 \) there exists \( t \) such that there are \( K_t \)-free \( n \)-vertex graphs with no stable set of size at least \( n^{\epsilon} \), it is known [5] that for every \( t \) there exists \( C_t > 0 \) such that almost every \( K_t \)-free \( n \)-vertex graph has a stable set of size at least \( C_t n \).

It remains an open problem to determine whether for some \( \epsilon > 0 \) and every graph \( H \), there exists \( C_H > 0 \) such that almost every graph \( G \) that does not contain \( H \) as an induced subgraph has a stable set or clique of size at least \( C_H |V(G)|^{\epsilon} \). This would be true (see Lemma 3 of [7] and the discussion around it) if for some \( \epsilon > 0 \) and every integer \( t \geq 0 \), there exists \( C_t > 0 \) such that, for every graph \( G \), either
• there are two vertices \( u, v \) and a stable set \( S \) with
\[ |S \cap (N(u) - N(v))|, |S \cap (N(v) - N(u))|, |S \cap N(u) \cap N(v)| \]
all of cardinality \( t \), or

• there are two vertices \( u, v \) and a clique \( T \) with
\[ |T \cap (N(u) - N(v))|, |T \cap (N(v) - N(u))|, |T \setminus (N(u) \cup N(v))| \]
all of cardinality \( t \), or

• \( G \) contains a clique or stable set of size at least \( C_t|V(G)|^{t^c} \).

Our results are a first step in proving this statement.

References


