Near-Optimal Induced Universal Graphs for Bounded Degree Graphs

Abrahamsen, Mikkel; Alstrup, Stephen; Holm, Jacob; Knudsen, Mathias Bæk Tejs; Stöckel, Morten

Published in:
44th International Colloquium on Automata, Languages, and Programming (ICALP 201)

DOI:
10.4230/LIPIcs.ICALP.2017.128

Publication date:
2017

Document version
Publisher’s PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Near-Optimal Induced Universal Graphs for Bounded Degree Graphs

Mikkel Abrahamsen\(^1\), Stephen Alstrup\(^2\), Jacob Holm\(^3\), Mathias Bæk Tejs Knudsen\(^4\), and Morten Stöckel\(^5\)

1 University of Copenhagen, Copenhagen, Denmark
miab@di.ku.dk
2 University of Copenhagen, Copenhagen, Denmark
s.alstrup@di.ku.dk
3 University of Copenhagen, Copenhagen, Denmark
jaho@di.ku.dk
4 University of Copenhagen, Copenhagen, Denmark
mathias@tejs.dk
5 University of Copenhagen, Copenhagen, Denmark
morten.stockel@gmail.com

Abstract

A graph \(U\) is an induced universal graph for a family \(\mathcal{F}\) of graphs if every graph in \(\mathcal{F}\) is a vertex-induced subgraph of \(U\).

We give upper and lower bounds for the size of induced universal graphs for the family of graphs with \(n\) vertices of maximum degree \(D\). Our new bounds improve several previous results except for the special cases where \(D\) is either near-constant or almost \(n/2\). For constant even \(D\) Butler [Graphs and Combinatorics 2009] has shown \(O\left(\frac{n^D}{2}\right)\) and recently Alon and Nenadov [SODA 2017] showed the same bound for constant odd \(D\). For constant \(D\) Butler also gave a matching lower bound. For generals graphs, which corresponds to \(D = n\), Alon [Geometric and Functional Analysis, to appear] proved the existence of an induced universal graph with \((1 + o(1)) \cdot 2^{(n-1)/2}\) vertices, leading to a smaller constant than in the previously best known bound of \(16 \cdot 2^{n/2}\) by Alstrup, Kaplan, Thorup, and Zwick [STOC 2015].

In this paper we give the following lower and upper bound of

\[
\left(\frac{n^D}{2}\right) \cdot n^{-O(1)} \quad \text{and} \quad \left(\frac{n^D}{2}\right) \cdot 2^{\Omega(\sqrt{D \log D \log(n/D)\log n})},
\]

respectively, where the upper bound is the main contribution. The proof that it is an induced universal graph relies on a randomized argument. We also give a deterministic upper bound of \(O\left(\frac{n^k}{\log n}\right)\). These upper bounds are the best known when \(D \leq n/2 - \Omega(n^{3/4})\) and either \(D\) is even and \(D = o(1)\) or \(D\) is odd and \(D = \omega(1)\). In this range we improve asymptotically on the previous best known results by Butler [Graphs and Combinatorics 2009], Esperet, Arnaud and Ochem [IPL 2008], Adjiashvili and Rotbart [ICALP 2014], Alon and Nenadov [SODA 2017], and Alon [Geometric and Functional Analysis, to appear].

1998 ACM Subject Classification E.1 Data Structures, 2.2 Graph Theory

\(^*\) Research partly supported by Advanced Grant DFF-0602-02499B from the Danish Council for Independent Research under the Sapere Aude research career programme.
\(^†\) Research partly supported by the FNU project AlgoDisc, Villum Fonden, and the innovationsfonden project DABAI.
\(^‡\) Research partly supported by the FNU project AlgoDisc.
\(^§\) Research partly supported by Villum Fonden and the innovationsfonden project DABAI.
1 Introduction

A graph \( G = (V, E) \) is said to be an induced universal graph for a family \( \mathcal{F} \) of graphs if it contains each graph in \( \mathcal{F} \) as a vertex-induced subgraph. A graph \( H = (V', E') \) is contained in \( G \) as a vertex-induced subgraph if \( V' \subseteq V \) and \( E' = \{vw \mid v, w \in V' \land vw \in E\} \). Induced universal graphs have been studied since the 1960s [45, 49], and bounds on the sizes of induced universal graphs have been given for many families of graphs, including general, bipartite [11], and bounded arboricity graphs [10]. In Table 2 in Section 2.3 we give an overview of the state of the art for various graph families along with the results in this paper.

1.1 Adjacency labeling schemes and induced universal graphs

An adjacency labeling scheme for a given family \( \mathcal{F} \) of graphs assigns labels to the vertices of each graph in \( \mathcal{F} \) such that a decoder given the labels of two vertices from a graph, and no other information, can determine whether or not the vertices are adjacent in the graph. The labels are assumed to be bit strings, and the goal is to minimize the maximum label size. A \( b \)-bit labeling scheme uses at most \( b \) bits per label. Information theoretical studies of adjacency labeling schemes go back to the 1960s [16, 17], and efficient labeling schemes were introduced in [35, 47]. The first labeling schemes for bounded degree graphs were given in [17]. Let \( g_v(\mathcal{F}) \) be the smallest number of vertices in any induced universal graph for a family of graphs \( \mathcal{F} \). In the families of graphs we study in this paper, a graph always has \( n \) vertices, unless explicitly stated otherwise.

A labelling scheme for \( \mathcal{F} \) is said to have unique labels if no two vertices in the same graph from \( \mathcal{F} \) are given the same label. We have the following connection between induced universal graph sizes and label sizes.

\[ \text{Theorem 1 (}[35]\). A family \( \mathcal{F} \) of graphs has a \( b \)-bit adjacency labeling scheme with unique labels iff } g_v(\mathcal{F}) \leq 2^b. \]

1.2 Our results

The contribution of this paper are stronger bounds on the size of induced universal graphs for bounded degree graphs. Our new bounds are the best known for a significant part of the parameter space, specifically we improve on previous results unless \( D \) is either near-constant or almost \( n/2 \). The best known results for the entire parameter range of induced universal graphs for bounded degree \( D \) graphs are shown in Table 1. When the bounded degree \( D \) is constant then Butler [18] along with Alon and Nenadov [8] gave optimal bounds. When \( D \) is even and of size \( \omega(1) \) and when \( D \) is odd and of size \( \omega(\log n/\log \log n) \) our first new upper bound is the best known as long as \( D = O((\log n) \log \log n) \). Let \( \mathcal{G}_D \) be the family of graphs with \( n \) vertices and maximum degree \( D \). We show the following.

\[ \text{Theorem 2. For the family } \mathcal{G}_D \text{ of graphs with bounded degree } D \text{ on } n \text{ nodes } \]

\[ g_v(\mathcal{G}_D) \leq 8 \cdot \frac{n^k}{(k-1)!}, \text{ where } k = \lceil D/2 \rceil. \]
Table 1: The state-of-the-art landscape for induced universal graph sizes. The first column denotes in which range the corresponding bound is the best known.

<table>
<thead>
<tr>
<th>Range of D</th>
<th>Bound</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D$ even and $D = O(1)$</td>
<td>$O(n^{D/2})$</td>
<td>[18]</td>
</tr>
<tr>
<td>$D$ odd and $D \in [3, D \log \log n]$</td>
<td>$O(D)^n n^{D/2}$</td>
<td>[8]</td>
</tr>
<tr>
<td>$D$ even and $D \in [\omega(1), O((\log n)^2 \log \log n)]$</td>
<td>$O\left(\frac{D}{D/2-1}\right)$</td>
<td>Theorem 2</td>
</tr>
<tr>
<td>$D$ odd and $D \in [\omega\left(\frac{\log n}{\log \log n}\right), O((\log n)^2 \log \log n)]$</td>
<td>$O\left(\frac{n(D+1)/2}{(D+1)/2}\right)$</td>
<td>Theorem 2</td>
</tr>
<tr>
<td>$D \leq \frac{n}{2} - O\left(\frac{n^{1/4} \log^3 n}{2}\right)$</td>
<td>(2^O\left(\sqrt{D \log D \log(n/D)}\right))</td>
<td>Theorem 3</td>
</tr>
<tr>
<td>$D \geq \frac{n}{2} - O\left(\frac{n^{1/4} \log^3 n}{2}\right)$</td>
<td>((1+o(1))2^{(n-1)/2})</td>
<td>[5]</td>
</tr>
</tbody>
</table>

Our second new upper bound is the smallest induced universal graph for the interval starting where Theorem 2 ends and as long as $D \leq \frac{n}{2} - O\left(\frac{n^{3/4} \log^3 n}{2}\right)$. The previous best upper bound for such large $D$ was $\left(\frac{n}{D/2}\right)^n n^{O(1)}$ due to Adjiashvili and Rotbart [3]. The bound presented in Theorem 3 is a randomized construction, which works for any $D$, and which improves asymptotically on Adjiashvili and Rotbart [3] for $D = \omega(1)$. We show the following.

**Theorem 3.** For the family $\mathcal{G}_D$ of graphs with bounded degree $D$ on $n \geq 2D$ nodes

$$g_v(\mathcal{G}_D) \leq \left(\frac{n/2}{D/2}\right)^{\frac{1}{2}} 2^O\left(\sqrt{D \log D \log(n/D)}\right).$$

We note that our bound together with the lower bound from Corollary 8 shows that for $D = \omega(1)$, $g_v(\mathcal{G}_D) = \left(\frac{n/2}{D/2}\right)^{\frac{1}{2}} 2^O(1)$. In contrast when $D \leq n/2(1-\Omega(1))$ and $D = \Omega(n)$ the bound $\left(\frac{n}{D/2}\right)^n n^{O(1)}$ due to Adjiashvili and Rotbart [3] is $\left(\frac{n/2}{D/2}\right)^{1+\Omega(1)}$, so we give the first near-optimal induced universal graph when $D$ is superconstant.

From a labeling scheme perspective, the combination of Theorems 2 and 3 shows the existence of an adjacency labeling scheme for $\mathcal{G}_D$ of size

$$\log \left(\frac{n/2}{D/2}\right) + O\left(\min\left\{D + \log n, \sqrt{D \log D \log(n/D)}\right\}\right).$$

This new labeling scheme improves upon previous in the same ranges as the improvements for the induced universal graphs as shown in Table 1.

In Corollary 8 we show that the any adjacency labeling scheme for $\mathcal{G}_D$ must have labels of size at least $\log \left(\frac{n/2}{D/2}\right) - O(\log n)$. Our new lower bounds differ from our upper bounds by $O(\min\{D + \log n, \sqrt{D \log D \log(n/D)}\})$, which is at most $O(\sqrt{n \log n})$.

## 2 Related results

### 2.1 Maximum degree $D$

Let $k = \lceil D/2 \rceil$. To give an upper bound for any value of $D$ Butler [18] showed the following corollary, which follows from the classic decomposition theorem by Petersen (see [41]):

**Corollary 4 ([18]).** Let $G \in \mathcal{G}_D$ be a graph on $n$ vertices with maximum degree $D$. Then $G$ can be decomposed into $k$ edge disjoint subgraphs where the maximum degree of each subgraph is at most 2.

To achieve an upper bound for $g_v(\mathcal{G}_D)$ this can be combined with:
Near-Optimal Induced Universal Graphs for Bounded Degree Graphs

**Theorem 5 ([29]).** Let \( F \) and \( Q \) be two families of graphs and let \( G \) be an induced universal graph for \( F \). Suppose that every graph in the family \( Q \) can be edge-partitioned into \( \ell \) parts, each of which forms a graph in \( F \). Then \( g_v(Q) \leq |V(G)|^{\ell} \).

Using Theorem 5, Butler [18] concluded that \( g_v(G_D) \leq (6.5)n^k \). Similarly Esperet et al. [29] achieved \( g_v(G_D) \leq (2.5n + O(1))^k \), and most recently it was shown by Abrahamsen et al. [2] that \( g_v(G_D) \leq (2n - 1)^k < 2^{c_k}n^k \) due to an induced universal graph for \( G_2 \) of size \( 2n - 1 \).

For constant maximum degree \( D \), Butler [18] also showed \( g_v(G_D) = \Omega(n^{D/2}) \), hence when \( D \) is even and constant, \( g_v(G_D) = \Theta(n^{D/2}) \) is the right answer up to constant factors due to the above bounds.

### 2.2 Constant odd degree

A universal graph for a family of graphs \( F \) is a graph that contains each graph from \( F \) as a subgraph (not necessarily vertex induced). It is a natural question how to construct universal graphs with as few edges as possible.

A graph has arboricity \( k \) if the edges of the graph can be partitioned into at most \( k \) forests. Graphs with maximum degree \( D \) have arboricity bounded by \( \lfloor D/2 \rfloor + 1 \) [19, 40].

When \( D \) is odd and constant, some improvement has been achieved on the above bounds on \( g_v(G_D) \) by arguments involving universal graphs and graphs with bounded arboricity [7, 29]. Let \( A_k \) denote a family of graphs with arboricity at most \( k \).

**Theorem 6 ([29]).** Let \( G \) be a universal graph for \( A_k \) and \( d_i \) the degree of vertex \( i \) in \( G \). Then \( g_v(A_k) \leq \sum (d_i + 1)^k \).

Alon and Capalbo [6] described a universal graph with \( n \) vertices of maximum degree \( c(D)n^{1−2/D} \log^{4/D}n \) for the family \( G_D \), where \( D \geq 3 \) and \( c(D) \) is a constant. Using this bound in Theorem 6, Esperet et al. [29] noted that for odd \( D \) (and hence arboricity \( k = \lfloor D/2 \rfloor \)), we get \( g_v(G_D) \leq c_1(D)n^{k−2+\frac{1}{D}} \), for a constant \( c_1(D) \).\(^1\) Using the slightly better universal graphs from [7] the maximum degree is reduced to \( c(D)n^{1−2/D} \) [4], giving \( g_v(G_D) \leq c_2(D)n^{k−2} \), for a constant \( c_2(D) \). Note that applying Theorem 6 along with universal graph [7] as above, then for even values of \( D \) this would give \( g_v(G_D) \leq c_3(D)n^{k−1} \), for a constant \( c_3(D) \). Recently, Alon and Nenadov [8] showed an upper bound \( g_v(G_D) = O(n^{D/2}) \), coinciding with Butler’s lower bound up to constant factors for any constant \( D \).

### 2.3 Other graph families

For the family of general, undirected graphs on \( n \) vertices, Alstrup et al. [11] gave an induced universal graph with \( O(2^{n/2}) \) vertices, which matches a lower bound by Moon [45]. More recently Alon [5] showed a construction that is tight up to an additive lower order term. We note that whereas the construction of [11] is presented as a labeling scheme, with efficient encoding and constant decoding time. However, it is not obvious if the the induced universal graph from [5] can be transformed into a labeling scheme without requiring that either the encoder or decoder use exponential space or time.

\(^1\) In [29] a typo states that the maximum degree for the universal graph in [6] is \( c(D)n^{1−2/D} \log^{4/D}n \). The theorem in [6] only states the total number of edges being \( c(D)n^{1−2/D} \log^{4/D}n \), however the maximum degree is \( c(D)n^{1−2/D} \log^{4/D}n \) [4].
Table 2: Induced-universal graphs for various families of graphs. “P” is results in this paper. For the max degree results $k = \lceil D/2 \rceil$. In the result for families of graphs with an excluded minor, the $O(1)$ term in the exponent depends on the fixed minor excluded.

The upper bounds from [11] are labeling schemes with efficient encoding and constant decoding time, but the upper bounds are larger by a constant factor. It is not obvious if the induced universal graph from [5] can be transformed into a labeling scheme without requiring that either the encoder or decoder use exponential space or time.

<table>
<thead>
<tr>
<th>Graph family</th>
<th>Lower bound</th>
<th>Upper bound</th>
<th>Lower/Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>General *</td>
<td>$2^{\frac{n}{2}+1}$</td>
<td>$(1 + o(1)) \cdot 2^{\frac{n}{2}+1}$</td>
<td>[45]/[5, 11]</td>
</tr>
<tr>
<td>Tournaments *</td>
<td>$2^{\frac{n}{2}+1}$</td>
<td>$(1 + o(1)) \cdot 2^{\frac{n}{2}+1}$</td>
<td>[46]/[5, 11]</td>
</tr>
<tr>
<td>Bipartite *</td>
<td>$(1 - o(1)) \cdot 2^\frac{n}{2}$</td>
<td>$(1 + o(1)) \cdot 2^\frac{n}{2}$</td>
<td>[42]/[5, 11]</td>
</tr>
<tr>
<td>A: Max degree $D$</td>
<td>$\binom{n/2}{D/2} \cdot n^{-O(1)}$</td>
<td>$O\left(\frac{n^2}{(n-1)}\right)$</td>
<td>$P([39, 43, 44])/P$</td>
</tr>
<tr>
<td>B: Max degree $D$</td>
<td>$\binom{n/2}{D/2} \cdot n^{-O(1)}$</td>
<td>$\binom{n/2}{D/2} \cdot 2^{O(\sqrt{D} \log D \log(n/D))}$</td>
<td>$P([39, 43, 44])/P$</td>
</tr>
<tr>
<td>C: Constant max degree $D$</td>
<td>$\Omega(n^{D/2})$</td>
<td>$O(n^{D/2})$</td>
<td>[18]/[8]</td>
</tr>
<tr>
<td>Max degree 2</td>
<td>$11 \cdot</td>
<td>n/6</td>
<td>$</td>
</tr>
<tr>
<td>Acyclic, max degree 2</td>
<td>$[3/2n]$</td>
<td>$[3/2n]$</td>
<td>[2]/[2]</td>
</tr>
<tr>
<td>Excluding a fixed minor</td>
<td>$\Omega(n)$</td>
<td>$n^2(\log n)^{O(1)}$</td>
<td>[33]</td>
</tr>
<tr>
<td>Planar</td>
<td>$\Omega(n)$</td>
<td>$n^2(\log n)^{O(1)}$</td>
<td>[33]</td>
</tr>
<tr>
<td>Planar, constant degree</td>
<td>$\Omega(n)$</td>
<td>$O(n^2)$</td>
<td>[20]</td>
</tr>
<tr>
<td>Outerplanar</td>
<td>$\Omega(n)$</td>
<td>$n(\log n)^{O(1)}$</td>
<td>[33]</td>
</tr>
<tr>
<td>Outerplanar, constant degree</td>
<td>$\Omega(n)$</td>
<td>$O(n)$</td>
<td>[20]</td>
</tr>
<tr>
<td>Treewidth $l$</td>
<td>$n^{2^{O(l)}}$</td>
<td>$n(\log \frac{n}{2})^{O(l)}$</td>
<td>[33]</td>
</tr>
<tr>
<td>Constant arboricity $l$</td>
<td>$\Omega(n^l)$</td>
<td>$O(n^l)$</td>
<td>[12]/[10]</td>
</tr>
</tbody>
</table>

It follows from [10, 12] that $g_c(\mathcal{A}_k) = \Theta(n^k)$ for the family $\mathcal{A}_k$ of graphs with constant arboricity $k$ and $n$ vertices. Using universal graphs constructed by Babai et al. [13], Bhatt et al. [14], and Chung et al. [21, 22, 23, 24], Chung [20] obtained the best currently known bounds for e.g. induced universal graphs for planar and outerplanar bounded degree graphs.

Labeling schemes are being widely used and well-studied in the theory community: Chung [20] gave labels of size $\log n + O(\log \log n)$ for adjacency labeling in trees, which was improved to $\log n + O(\log^2 n)$ [12] and in [15, 20, 31, 32, 36] to $\log n + \Theta(1)$ for various special cases of trees. Finally it was improved to $\log n + \Theta(1)$ for general trees [10].

Using labeling schemes, it is possible to avoid costly access to large global tables and instead only perform local and distributed computations. Such properties are used in applications such as XML search engines [1], network routing and distributed algorithms [26, 27, 30, 51], dynamic and parallel settings [25, 38], and various other applications [37, 48, 50].

A survey on induced universal graphs and adjacency labeling can be found in [11]. See [34] for a survey on labeling schemes for various queries. We give an overview in Table 2 of dominating existing and new results. In the table, “P” refer to a result in this paper and we define $k = \lceil D/2 \rceil$. The “A” and “B” case below represent two different constructions. The upper bound in “B” is a randomized construction.
3 Preliminaries

Let \([n] = \{0, \ldots, n - 1\}\), \(\mathbb{N}_0 = \{0, 1, 2, \ldots\}\), \(\mathbb{N} = \{1, 2, \ldots\}\), and let \(\log n\) refer to \(\log_2 n\). For a graph \(G\), let \(V[G]\) be the set of vertices and \(E[G]\) be the set of edges of \(G\), and let \(|G| = |V[G]|\) be the number of vertices. We denote the maximum degree of graph \(G\) as \(\Delta(G)\).

For \(i \in \mathbb{N}\), let \(P_i\) denote a path with \(i\) vertices, and for \(i > 2\), let \(C_i\) denote a simple cycle with \(i\) vertices. We let \(G^2\) denote the square of the unweighted graph \(G\), i.e., there is an edge between two nodes in \(G^2\) if they have at most distance two in \(G\). For a boolean statement \(B\) we will denote by \([B]\) the value 1 if \(B\) is true and 0 otherwise.

Let \(G\) and \(U\) be two graphs and let \(\lambda: V[G] \to V[U]\) be an injective function. If \(\lambda\) has the property that \(uv \in E[G]\) if and only if \(\lambda(u)\lambda(v) \in E[U]\), we say that \(\lambda\) is an embedding function of \(G\) into \(U\). \(G\) is an induced subgraph of \(U\) if there exists an embedding function of \(G\) into \(U\), and in that case, we say that \(G\) is embedded in \(U\) and that \(U\) embeds \(G\). Let \(\mathcal{F}\) be a family of graphs. \(U\) is an induced universal graph for \(\mathcal{F}\) if \(G\) is an induced subgraph of \(U\) for each \(G \in \mathcal{F}\).

4 General \(D\)

In this section we present two upper bounds on \(g_v(\mathcal{G}_D)\), the number of nodes in the smallest induced universal graph for graphs on \(n\) nodes with bounded degree \(D\). In Theorem 2 we give a deterministic construction of an induced universal graph for \(\mathcal{G}_D\) that relies on the fact that \(P^*_n\) is a sparse universal graph for \(\mathcal{G}_2\).

In Theorem 3 we give a randomized labeling scheme for \(\mathcal{G}_D\). For every graph in the family \(\mathcal{G}_D\) we give a randomized assignment of labels to the nodes of the graph and show that the labels are short with non-zero probability, thereby showing that there exist short labels for every graph in \(\mathcal{G}_D\). This in turn implies an upper bound on \(g_v(\mathcal{G}_D)\). Combining the two results shows the existence of an adjacency labeling scheme for \(\mathcal{G}_D\) of size \(\log \left(\lfloor n/2 \rfloor \right) + O(\min \{ D + \log n, \sqrt{D \log D \log(n/D)} \})\).

In Section 4.2 we show to use the results by Liebenau and Wormald [39] to give lower bounds on \(g_v(\mathcal{G}_D)\). These lower bounds imply that any adjacency labeling scheme for \(\mathcal{G}_D\) must have labels of size at least \(\log \left(\lfloor n/2 \rfloor \right) = O(\log n)\), which means that the upper bounds are tight up to an additive term of size \(O(\min \{ D + \log n, \sqrt{D \log D \log(n/D)} \})\), which is at most \(O(\sqrt{n \log n})\).

4.1 Upper bounds on \(g_v(\mathcal{G}_D)\)

We present the proof of our first upper bound stated in Theorem 2.

**Proof.** For a set \(S\) we let \(S^{\leq \ell}\) denote the set of all subsets of \(S\) of size \(\leq \ell\). We note that \(|S^{\leq \ell}| \leq 2 \frac{|S|^\ell}{\ell!}\) whenever \(S\) is finite.

Fix \(n, D\), let \(k = \lfloor D/2 \rfloor\) and let \(H_n = P^*_n\) be the square of the path of length \(n\). Identify the vertices of \(H_n\) with \([n]\) in the obvious way. Then two nodes \(i, j\) in \(H_n\) are adjacent if and only if they are different and \(|i - j| \leq 2\). We define the graph \(G\) to have vertex set \(V[G] = [n] \times [2]^2 \times [n]^{\leq k-1}\). For a node \(u = (i, x, y, S)\) in \(G\), we define \(id(u) = i\). We also
We now orient the edges of each of the graphs \( G \) so assume that there is an edge between vertex set function id by sorting the values \( i, j \) be a bijection that assigns an identifier to each node of a sufficiently large constant. Let \( G \) is an induced universal graph for \( G_D \). We will show that \( H \) is an induced subgraph of \( G \). By Corollary 4 we know that we can decompose the edges of \( H \) into \( k \) edge disjoint subgraphs, \( H_0, H_1, \ldots, H_{k-1} \), such that each \( H_i \) has vertex set \( V[H_i] = V[H] \) and maximum degree at most 2. First we order the nodes of \( H \) as \( u_0, u_1, \ldots, u_{n-1} \) such that all edges \((u_i, u_j)\) in \( H_0 \) satisfy \(|i - j| \leq 2\). This is possible since \( H_0 \) has maximum degree at most 2, and therefore consists of only paths and cycles. We let \( x_i \) (resp. \( y_i \)) be 1 if there is an edge from \( u_i \) to \( u_{i+1} \) (resp. \( u_{i+2} \)) in \( H_0 \). That is:

\[
\begin{align*}
x_i & = [(u_i, u_{i+1}) \in E[H_0]}, \quad y_i = [(u_i, u_{i+2}) \in E[H_0]].
\end{align*}
\]

We now orient the edges of each of the graphs \( H_1, \ldots, H_{k-1} \) such that the out degree of each node is at most 1. This is possible since each of \( H_i \) has maximum degree at most 2. We let \( S_i \) be the set of all \( u_i \)'s outgoing neighbours in the graphs \( H_1, \ldots, H_{k-1} \), and note that \( S_i \) contains at most \( k - 1 \) elements. We let \( \lambda : H \to G \) be defined by \( \lambda(i) = (i, x_i, y_i, S_i) \). It is now straightforward to check that \( \lambda \) is an embedding function and therefore that \( H \) is an induced subgraph of \( G \). Since \( H \) was arbitrarily chosen this shows that \( G \) is an induced universal graph of \( G_D \). The number of nodes in \( G \) is exactly \( 4n \cdot [n]^{\leq k-1} \) which yields the desired result.

The intuition behind the randomized bound below is the following. Consider placing all \( n \) vertices on a circle in a randomly chosen order and rename the vertices with indices \([n]\) following the order on the circle. Now, a vertex \( v \in [n] \) remembers its neighbours in the next half of the circle, i.e., \( v \) stores all the adjacent vertices among \( \{v + 1, \ldots, v + [n/2]\} \) (where indices are taken modulo \( n \)). If two vertices \( u, v \) are adjacent, then clearly either \( u \) stores the index of \( v \) or conversely, hence an adjacency query can be answered. See Figure 1. A standard application of Chernoff bounds implies that vertex \( v \) with high probability stores at most \( D/2 + O(\sqrt{D \log n}) \) indices. However, this can be tightened by a Lovász Local Lemma argument, since each random variable that denote which indices should be stored depend on at most \( D^2 \) other such random variables. This allows us to tighten the number of stored indices to \( D/2 + O(\sqrt{D \log D}) \), and it follows that there exists an ordering of the points on the circle where every vertex stores that many neighbours and the theorem follows.

We are ready to show Theorem 3.

**Proof.** Fix \( n, D \) and wlog assume that \( n \) is odd. For \( D \leq \log n \) the result follows from Theorem 2 so assume that \( D \geq \log n \). We assume that \( n \) and \( D \) are bounded from below by a sufficiently large constant. Let \( G \) be a graph in \( G_D \), and wlog assume that \( V[G] = [n] \). Let \( t_0, t_1, \ldots, t_{n-1} \in [n] \) be chosen independently and uniformly at random, and let \( \text{id} : [n] \to [n] \) be a bijection that assigns an identifier to each node of \( G \) such that

\[
(t_i, i) < (t_j, j) \Rightarrow \text{id}(i) < \text{id}(j),
\]

for all values of \( i, j \in [n] \), where \( < \) is the lexicographical ordering. We construct the function \( \text{id} \) by sorting the values \( t_i \), and then choosing \( \text{id} \) to be a bijection such that
Figure 1 The intuition of the randomized upper bound. Pictured are adjacency relations stored by $v_1$ and $v_2$ – an edge denotes adjacency and a directed edge $(v, u)$ denotes $v$ stores its relation to $u$. Here $v_1$ stores $v_2$ but not $u_2$ as $u_2$ is on the wrong side of $v_1$’s bisection, and $v_2$ stores $u_1$ and $u_2$, but not $v_1$ as it is on the wrong side of $v_2$’s bisection.

$(t_{\text{id}(0)}, \text{id}(0)), \ldots, (t_{\text{id}(n-1)}, \text{id}(n-1))$ is a non-decreasing sequence. We note that the values $t_i$ determine $\text{id}$ uniquely.

For each $i \in [n]$ we let $S_i \subseteq [n]$ be the set that contains all neighbours $j$ of $i$ for which it holds that:

$$(t_j - t_i) \mod n \in \left\{ 0, 1, 2, \ldots, \frac{n-1}{2} \right\}$$

That is, we define $S_i$ by:

$$S_i = \left\{ j \in [n] \mid \{i, j\} \in E[G], \ (t_j - t_i) \mod n \in \left\{ 0, 1, 2, \ldots, \frac{n-1}{2} \right\} \right\}$$ (1)

We say that the values $t_i$ are good if the following properties hold for all $i \in [n]$:

$$|S_i| \leq \frac{D}{2} + C \sqrt{D \log D}$$ (2)

$$\max_{j \in S_i} \{|\text{id}(j) - \text{id}(i)| \mod n\} \leq \frac{n}{2} + \max \left\{ \frac{n}{D}, C \sqrt{n \log n} \right\}$$ (3)

where $C > 0$ is a (sufficiently large) constant to be chosen later. Firstly, we will show that, when choosing the $t_i$’s randomly, they are good with non-zero probability. For each $i \in [n]$ let $A_i$ be the event that (2) does not hold for $S_i$. Let $\mathcal{A} = \{ A_i \mid i \in [n] \}$ and for each $A_i \in \mathcal{A}$ let $\Gamma(A_i)$ denote the set of all events $A_j$ where $j \neq i$ has distance at most two to $i$ in $G$. We note that since $G$ has maximum degree at most $D$ we have that $|\Gamma(A_i)| \leq D^2$. For each $i \in [n]$ the event $A_i$ is independent of all events $\mathcal{A} \setminus (\{ A_i \} \cup \Gamma(A_i))$ for the following reason. The event $A_i$ is determined exclusively by the values $t_j$ where $j = i$ or $j$ is a neighbour of $i$ in $G$. For each $A_j$ such that $A_j \in \mathcal{A} \setminus (\{ A_i \} \cup \Gamma(A_i))$ we have that $j$ has distance at least three to $i$, and $A_j$ is determined by the values $t_{j'}$ where $j' = j$ or $j'$ is a neighbour of $j$. No such $j'$ can also be a neighbour of $i$ since $j$ has distance three to $i$ and we conclude that $A_i$ is independent of the events $\mathcal{A} \setminus (\{ A_i \} \cup \Gamma(A_i))$ for each $i \in [n]$. By a Chernoff bound we have that $A_i$ happens, i.e. (2) is false for $S_i$, with probability at most $e^{-\Theta(C^2 \log D)}$. Choosing $C$
We note that (4) is positive if and only if
\begin{align*}
x(A_i) \prod_{A_j \in \Gamma(A_i)} (1 - x(A_j)) \geq D^{-10} (1 - D^{-10})^{D^2} = D^{-10} e^{-\Theta(D^{-10})} D^2 \\
> \frac{1}{2} D^{-10} \geq P(A_i)
\end{align*}

By the Lovász Local Lemma [28, 9] we conclude that the probability that none of \( A_i, i \in [n] \) happen is bounded below by
\begin{align*}
\prod_{i \in [n]} (1 - x(A_i)) = (1 - D^{-10})^n = e^{-\Theta(nD^{-10})}.
\end{align*}

That is, (2) holds for all \( S_i \) with probability at least \( e^{-\Theta(nD^{-10})} \). For any value of \( i \in [n] \) the probability that (3) is false is at most \( \min \left\{ e^{-\Theta(nD^{-2})}, n^{-\Omega(C^2)} \right\} \) by a standard Chernoff bound. And by a union bound over all choices of \( i \in [n] \) (3) holds for all values of \( i \) with probability a least \( 1 - n \min \left\{ e^{-\Theta(nD^{-2})}, n^{-\Omega(C^2)} \right\} \). Therefore, the probability that the chosen \( t_s \)s are good is at least
\begin{align*}
e^{-\Theta(nD^{-10})} + \left( 1 - n \min \left\{ e^{-\Theta(nD^{-2})}, n^{-\Omega(C^2)} \right\} \right) - 1.
\end{align*}

We note that (4) is positive if and only if
\begin{align*}
e^{-\Theta(nD^{-10})} > n \min \left\{ e^{-\Theta(nD^{-2})}, n^{-\Omega(C^2)} \right\},
\end{align*}

and this can be verified, e.g. by considering the cases \( D \leq n^{1/3} \) and \( D > n^{1/3} \). That is, the values \( t_s \) are good with non-zero probability.

Now fix a good choice of \( t_s \) and the corresponding identifier function, \( \text{id} \), and the sets \( S_i \). We can now encode the values \( \text{id}(i) \) and the set \( S_i \) using at most \( O(\log n) + \log \left( \frac{n'}{D'} \right) \) bits where \( n' \) and \( D' \) are defined by:
\begin{align*}
n' = \left\lceil \frac{n}{2} + \max \left\{ \frac{n}{D}, C\sqrt{n \log n} \right\} \right\rceil, \quad D' = \left\lceil \frac{D}{2} + C\sqrt{D \log D} \right\rceil.
\end{align*}

Let \( D'' = \min \left\{ \left\lceil \frac{n'}{2} \right\rceil, D' \right\} \). Then for any node \( i \) we can encode \( \text{id}(i) \) and the set \( \{ \text{id}(j) - \text{id}(i) \mod n | j \in S_i \} \) using at most \( O(\log n) + \log \left( \frac{n'}{D''} \right) \) bits for the following reason: Firstly, \( \text{id}(i) \) can clearly be stored using \( O(\log n) \) bits. Secondly, the set of differences \( \{ \text{id}(j) - \text{id}(i) \mod n | j \in S_i \} \) contains at most \( D' \) elements which are all contained in \( \{1, 2, \ldots, n'\} \), and hence it can be stored using at most \( \log \left( \frac{n'}{D''} \right) \) bits. Given the labels of two nodes \( i, j \) we can compute their ids, \( \text{id}(i) \) and \( \text{id}(j) \), and infer whether \( \text{id}(i) \in S_j \) or \( \text{id}(j) \in S_i \), i.e. whether \( i \) and \( j \) are adjacent. Hence we have described a labeling scheme for \( \mathcal{G}_D \) using at most \( O(\log n) + \log \left( \frac{n'}{D''} \right) \) bits, and therefore
\begin{align*}
g_\varepsilon(\mathcal{G}_D) \leq \left( \frac{n'}{D''} \right)^{n^{O(1)}}. \quad (6)
\end{align*}

We first note that:
\begin{align*}
\left( \frac{n'}{D''} \right) \leq \left( \frac{n'}{|D/2|} \right)^{D'' - \left\lfloor D/2 \right\rfloor} \leq \left( \frac{n'}{|D/2|} \right)^{2O(\sqrt{D \log D \log(n/D)})}. \quad (7)
\end{align*}
Furthermore we also have:

\[
\binom{n'}{\lfloor D/2 \rfloor} \leq \binom{n/2}{\lfloor D/2 \rfloor} \cdot \binom{n' - \lfloor D/2 \rfloor}{\lfloor n/2 \rfloor - \lfloor D/2 \rfloor}^{\lfloor D/2 \rfloor}
\]

\[
= \left( \frac{n/2}{\lfloor D/2 \rfloor} \right) \cdot \left( 1 + \frac{n' - \lfloor n/2 \rfloor}{\lfloor n/2 \rfloor - \lfloor D/2 \rfloor} \right)^{\lfloor D/2 \rfloor}. \tag{8}
\]

We have that \( \lfloor n/2 \rfloor - \lfloor D/2 \rfloor = \Omega(n) \) since \( D \leq \frac{n}{2} \) and therefore we get:

\[
\left( 1 + \frac{n' - \lfloor n/2 \rfloor}{\lfloor n/2 \rfloor - \lfloor D/2 \rfloor} \right)^{\lfloor D/2 \rfloor} = \left( 1 + O\left( \frac{n' - \lfloor n/2 \rfloor}{n} \right) \right)^{\lfloor D/2 \rfloor} \leq e^{O\left( \frac{n' - \lfloor n/2 \rfloor}{n} \right) \cdot \lfloor D/2 \rfloor}. \tag{9}
\]

By the definition of \( n' \) we have that

\[
\frac{n' - \lfloor n/2 \rfloor}{\lfloor D/2 \rfloor} = \max \left\{ \frac{1}{D}, C \sqrt{\frac{\log n}{n}} \right\} \cdot \lfloor D/2 \rfloor = O \left( \sqrt{D \log D} \right). \tag{10}
\]

Combining (8), (9) and (10) we get that

\[
\binom{n'}{\lfloor D/2 \rfloor} \leq 2^{O(\sqrt{D \log D})} \binom{n/2}{\lfloor D/2 \rfloor}. \tag{11}
\]

Combining (6) with (7) and (11) gives us the desired upper bound on \( g_v(G_D) \)

\[
g_v(G_D) \leq \binom{n/2}{\lfloor D/2 \rfloor} \cdot 2^{O(\sqrt{D \log D \log(n/D)})}.
\]

\[\blacktriangleright\]

### 4.2 Lower bounds on \( g_v(G_D) \)

In this section we show how to apply the bounds from [39] on the number of graphs of a given degree sequence. For a graph \( G \) with nodes \((u_1, u_2, \ldots, u_n)\) the degree sequence of \( G \) is \((d_1, d_2, \ldots, d_n)\) where \( d_i \) is the degree of \( u_i \). Applying [39, Conjecture 1.1] on a degree sequence \((d, d, \ldots, d)\) we obtain

\[\blacktriangleright\text{Corollary 7 [(39)]}. \text{ Let } n, d \text{ be integers such that } nd \text{ is even and } 1 \leq d \leq n - 1. \text{ Let } \mu = \frac{d}{n-1}. \text{ The number of } d\text{-regular graphs on } n \text{ nodes is}

\[
(1 + o(1))\sqrt{2e}^{1/4} \left( \mu^\mu (1 - \mu)^{1-\mu} \right)^{n(n-1)/2} \left( \frac{n-1}{d} \right)^n.
\]

We now show that the bound from Corollary 7 implies a lower bound on the size of the induced universal graph for bounded degree graphs:

\[\blacktriangleright\text{Corollary 8}. \text{ For the family } G_D \text{ of graphs with bounded degree } D \text{ on } n \text{ nodes}

\[
g_v(G_D) = \Omega \left( \frac{1}{\sqrt{D}} \binom{n}{D/2} \right). \tag{12}
\]

We remark that together with Stirling’s approximation, Corollary 8 implies that \( g_v(G_D) \geq \binom{n/2}{\lfloor D/2 \rfloor} n^{-O(1)} \).
Proof. It is clearly enough to prove (12) when $D \leq \left\lfloor \frac{n}{2} \right\rfloor$, since the right hand side is non-increasing for $D \geq \left\lfloor \frac{n}{2} \right\rfloor$. Let $N = 2 \left\lfloor \frac{n}{2} \right\rfloor$ be the largest even integer not greater than $n$. So we assume that $2D \leq N$.

Let $X$ be the number of $D$-regular graphs on $N$ nodes. By Corollary 7 we have that

$$X = \Theta \left( \mu^n (1 - \mu)^{N-1} \frac{N-1}{D} \right)^N,$$

where $\mu = \frac{D}{N-1}$. By Stirling’s approximation we have that:

$$\binom{N-1}{D} = \Theta \left( \frac{\sqrt{N-1} \left( \frac{N-1}{e} \right)^{N-1}}{\sqrt{D} \left( \frac{N-1}{e} \right)^D} \right)$$

$$= \Theta \left( \frac{N-1}{D(N-1-D)} \left( \mu^n (1 - \mu)^{1 - \mu} \right)^{-(N-1)} \right)$$

$$= \Theta \left( \frac{1}{D} \left( \mu^n (1 - \mu)^{1 - \mu} \right)^{-(N-1)} \right).$$

Rearranging gives that:

$$(\mu^n (1 - \mu)^{1 - \mu})^{(N-1)/2} = \Theta \left( \sqrt{\frac{1}{D} \left( \frac{N-1}{D} \right)^{1 - \mu}} \right)^{N/2}$$

If we insert this into (13) we get that:

$$X = \Theta \left( \sqrt{\frac{1}{D} \left( \frac{N-1}{D} \right)^{1 - \mu}} \right)^{N/2} \left( \frac{N-1}{D} \right)^N = \Theta \left( \sqrt{\frac{1}{D} \left( \frac{N-1}{D} \right)^{1 - \mu}} \right)^N.$$ 

Since $2D \leq N$ we have that $\frac{N-1}{D} = \Theta \left( \left( \frac{n}{D} \right) \right)$. Clearly $X$ is smaller than $G_D$, and therefore:

$$|G_D|^{1/n} = \Omega \left( \sqrt{\frac{1}{D} \left( \frac{n}{D} \right)} \right)$$

(14)

Let $G$ be the induced universal graph for the family $G_D$. Let $V = [n]$. Any graph $H$ from $G_D$ on the vertex set $V$ is uniquely defined by the embedding function $f$ of $H$ in $G$. Since there are no more than $|V[G]|^n$ ways to choose $f$ we get that $|V[G]|^n \geq |G_D|$. Inserting (14) this shows (12) the following way,

$$g_v(G_D) = |V[G]| \geq |G_D|^{1/n} = \Omega \left( \sqrt{\frac{1}{D} \left( \frac{n}{D} \right)} \right).$$

References


Near-Optimal Induced Universal Graphs for Bounded Degree Graphs


Near-Optimal Induced Universal Graphs for Bounded Degree Graphs
