Sparse universal graphs

Noga Alon * Vera Asodi †

Abstract

For every n, we describe an explicit construction of a graph on n vertices with at most $O(n^{2-\epsilon})$ edges, for $\epsilon = 0.133\ldots$, that contains every graph on n vertices with maximum degree 3 as a subgraph. It is easy to see that each such graph must have at least $\Omega(n^{\frac{4}{3}})$ edges. We also show that the minimum number of edges of a graph that contains every graph with n edges as a subgraph is $\Theta(\frac{n^2}{\log^2 n})$. This improves a result of Babai, Chung, Erdös, Graham and Spencer.

1 Introduction

For a family $\mathcal{H}$ of graphs, a graph G is *$\mathcal{H}$-universal* if it contains every member of $\mathcal{H}$ as a (not necessarily induced) subgraph. The study of universal graphs for various families $\mathcal{H}$ is motivated by problems in VLSI circuit design. See, e.g., [4], [6] and their references.

In this paper we study the minimum possible number of edges in universal graphs for two families of graphs.

Let $\mathcal{H}(r, n)$ denote the family of all graphs on n vertices in which every degree is at most r. In section 2 we study the minimum possible number of edges in a graph on n vertices which is $\mathcal{H}(3, n)$-universal. The best known construction for such graphs is given in [1], where the authors describe an $\mathcal{H}(3, n)$-universal graph with n vertices and at most $O(n^{2-0.023})$ edges. Here we improve this result by giving a tighter analysis using a related technique, which is also based on some of the ideas in [5].

In subsection 2.1 we give an extension of a lemma used in [5], in subsection 2.2 we describe the construction of the graph, and in subsection 2.3 we bound its number of edges.

Let $\mathcal{E}_n$ denote the family of all graphs with at most n edges and without isolated vertices. In section 3 we study the minimum possible number of edges of an $\mathcal{E}_n$-universal graph. The best known result is given in [3], where the authors prove that the minimum possible number of edges of an $\mathcal{E}_n$-universal graph is at least $\frac{cn^2}{\log^2 n}$ and at most $(1 + o(1))\frac{n^2 \log\log n}{\log n}$. Here we determine this minimum possible number up to a constant factor and show that it is actually $\Theta(\frac{n^2}{\log^2 n})$.

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

†Department of Computer Science, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv, Israel. Email: vera@math.tau.ac.il.

2 $\mathcal{H}(3,n)$-universal graphs

2.1 The main lemma

To simplify the presentation we omit all floor and ceiling signs. It is not difficult to check that this is not crucial for the proofs.

The following lemma is proved in [5]:

Lemma 2.1 *Let $G = (V, E)$ be a graph on n vertices in which every degree is at most 3. Let $V_1, V_2, \ldots, V_m$ be any collection of pairwise disjoint subsets of V such that $|V_i| \geq \log^3 n$ for each i. Then, there exists an independent set W such that $|W \cap V_i| \geq |V_i|/5$, for all i.*

The following result extends and strengthens the above lemma:

Lemma 2.2 *The following holds for all $0 < \epsilon < 1/4$. Let $G = (V, E)$ be a graph on n vertices in which every degree is at most 3. Let $V_1, V_2, \ldots, V_m$ be any collection of pairwise disjoint subsets of V such that $|V_i| \geq \frac{10}{\epsilon^2} \log n$ for each i. Then, there exist two disjoint independent sets W_0 and W_1 such that $|W_i \cap V_j| = |V_j|(1/4 - \epsilon)$ for all i, j.*

Proof: Clearly, it suffices to prove the existence of two independent sets W_0 and W_1 such that $|W_i \cap V_j| \geq |V_j|(1/4 - \epsilon)$ for all i, j, as we can then take the desired sets as appropriate subsets of these two. We proceed with the proof that such W_0 and W_1 exist. Let $V = \{v_0, v_1, \ldots, v_{n-1}\}$ and let σ be a permutation chosen uniformly from S_n. Define W_0 and W_1 as follows: $W_0 = \{v_i \in V \mid \forall j \ s.t. \ (v_i, v_j) \in E, \ \sigma(i) < \sigma(j)\}$, and $W_1 = \{v_i \in V \mid \forall j \ s.t. \ (v_i, v_j) \in E, \ \sigma(i) > \sigma(j)\}$.

We now show that W_0 and W_1 satisfy the desired properties with positive probability. Without loss of generality, we may assume that $W_0 \cap W_1 = \emptyset$. (We can assume that there are no isolated vertices in G. If there is more than one isolated vertex, we can add edges so that there will be no isolated vertices, and if there is one isolated vertex, we can find independent sets in the rest of the graph, and add the isolated vertex to one of them.)

Let us define for each vertex u, the indicator random variables $\psi_i(u)$ for $i = 0, 1$ as follows: $\psi_i(u) = 1$ iff $u \in W_i$. We make the following observations.

Observation 1: For every $u \in V$, $E[\psi_i(u)] \geq 1/4$.

Observation 2: Let X be any independent set in G^2, where G^2 is the graph whose vertices are the vertices of G, and there is an edge (u, v) in G^2 iff there is a path of length 1 or 2 between u and v in G. For any distinct $u, v \in X$, u and v are not neighbours in G, and they do not share neighbours in G. Thus, for every fixed $i \in \{0, 1\}$, the $\psi_i(u)$'s, with $u \in X$ are mutually independent. This is because the value of $\psi_i(u)$ is determined by the induced permutation on the indices of u and its neighbours.

To prove the lemma it suffices to show that

$$Pr[\forall i, j \sum_{u \in V_j} \psi_i(u) \geq |V_j|(1/4 - \epsilon)] > 0.$$

Since $m < n/2$, it suffices to show that for each admissible i, j

$$Pr[\sum_{u \in V_j} \psi_i(u) < |V_j|(1/4 - \epsilon)] \leq \frac{1}{n}.$$

Fix i, j. Since the maximum degree in G is at most 3, it follows that the maximum degree in G^2 is at most 9, so by the Hajnal-Szemerédi theorem [7], there is a way to partition the induced subgraph of G^2 on V_j into 10 independent sets $X_1, X_2, \ldots, X_{10}$ of nearly equal sizes, that is $\lfloor |V_j|/10 \rfloor \leq |X_l| \leq \lceil |V_j|/10 \rceil$ for all l.

By observation 1, $\forall u \in V_j,\ E[\psi_i(u)] \geq 1/4$, and by observation 2, the $\psi_i(u)$'s for $u \in X_l$ are mutually independent. Therefore, by Chernoff's inequality (c.f., e.g., [2], Appendix A),

$$Pr[\sum_{u \in X_l} \psi_i(u) < |X_l|(1/4 - \epsilon)] \leq e^{-2\epsilon^2 |X_l|}.$$

If $\sum_{u \in X_l} \psi_i(u) > |X_l|(1/4 - \epsilon)]\ \forall 1 \leq l \leq 10$, then, clearly, $\sum_{u \in V_j} \psi_i(u) > |V_j|(1/4 - \epsilon)]$. Therefore, for a sufficiently large n, using the fact that $|V_j| \geq \frac{10}{\epsilon^2} \log n$, we conclude that

$$Pr[\sum_{u \in V_j} \psi_i(u) < |V_j|(1/4 - \epsilon)] \leq 10 e^{-2\epsilon^2 \frac{10}{10\epsilon^2} \log n} = \frac{10}{n^2} < \frac{1}{n}.$$

□

2.2 The construction of an $\mathcal{H}(3, n)$-universal graph

Define $g = \frac{2}{\epsilon^2} \log n$. We construct an $\mathcal{H}(3, n)$-universal graph $G = (V, E)$ as follows: First we construct a graph $G' = (V', E')$, where $|V'| = \Theta(n/g)$, and then we construct G from G'.

For a vector $v \in \{0, 1, 2\}^*$, denote by $N_j(v)$ the number of coordinates in v whose value is j, for $j = 0, 1, 2$. The set of vertices V' consists of all vectors v (of different lengths), such that

$$n(1/4 - \epsilon)^{N_0(v) + N_1(v)} (1/2 + 2\epsilon)^{N_2(v)} < 5g, \tag{1}$$

but this inequality fails for every prefix v' of v. Note that in this case the left hand side of the inequality is always at least g. There is an edge (u, v) in G' if and only if there is no coordinate in which both u and v have the value 0, and there is no coordinate in which they both have the value 1.

We construct G from G' by replacing each vertex v of G' by a set S_v of roughly g vertices, where the precise numbers will be given later. For each edge (u, v) in G', we add edges between each vertex of S_u and each vertex of S_v. We also add edges between all the pairs of vertices in the group that replaces the vertex in G' whose vector contains only 2's. Clearly, $|E| \leq |E'| g^2 + g^2 = O(|E'| \log^2 n)$.

Proposition 2.3 *The graph $G = (V, E)$ with the sizes of sets S_v as described below is $\mathcal{H}(3, n)$-universal.*

Proof: Suppose $H \in \mathcal{H}(3, n)$. We assign vectors of the type described above to the vertices of H, map each vertex to one in a group labeled by its vector, and show that all the required edges exist in G. Here are the details.

By Lemma 2.2, H contains two disjoint independent sets W_0 and W_1, of size $(1/4-\epsilon)n$ each. The first coordinate of the vectors of the vertices in W_i will be i, and the first coordinate of the vectors of the vertices in $V \setminus (W_0 \cup W_1)$ will be 2. Now, let us assume that we have already completed l steps and assigned to each vertex either a complete vector, if the length of this vector is at most l, or a prefix of length l. Let us classify the vertices in sets V_i according to the vectors assigned to them so far, and let $V_1, V_2, \ldots, V_m$ be all sets of vertices in this classification such that $|V_i| \geq 5g$. (Note that for each i, all the vertices in V_i have the same prefix.) Since all the sets V_j are of size at least $5g$, we can apply Lemma 2.2 and conclude that there exist two disjoint independent sets W_0 and W_1 such that $|W_i \cap V_j| = |V_j|(1/4-\epsilon)$. The $(l+1)^{th}$ coordinate of the vectors of the vertices in W_i will be i, and the $(l+1)^{th}$ coordinate of the vectors of the vertices in $V \setminus (W_0 \cup W_1)$ will be 2. If there is an edge between two vertices of H then they never belong to the same independent set during this process, which means that there is no coordinate in which the vectors assigned to them both have 0 and there is no coordinate in which they both have 1. To embed H in G simply map all the vertices whose vectors are v to S_v bijectively. Note that this process also determines the sizes of the sets S_v, as a function of n only. This is because the construction guarantees that the number of vertices to which we assign any fixed vector during the process is uniquely determined by the vector and n (and is independent of the structure of H). Indeed, in each step the size of each quantity of the form $|W_i \cap V_j| = |V_j|(1/4-\epsilon)$ is independent of the structure of H. Finally, if a fixed vector v is assigned to x vertices during the process, and $x \geq 5g$, then the concatanations $v0$ and $v1$ are assigned to $(1/4-\epsilon)x$ vertices each, and $v2$ is assigned to $(1/2+2\epsilon)x$ vertices. This implies that the process terminates with sets S_v corresponding to vectors v satisfying (1). □

2.3 Bounding the number of edges

Throughout the proof we denote by ϵ_1, ϵ_2 etc. positive constants, where $\epsilon_i \leq c_i\epsilon$ for some easily computable absolute constant c_i.

For every $v \in V$

$$\frac{n}{5g} \leq \left(\frac{1}{1/4-\epsilon}\right)^{N_0(v)+N_1(v)} \left(\frac{1}{1/2+2\epsilon}\right)^{N_2(v)} \leq \frac{n}{g}.$$

Let $k = \log_{\frac{1}{1/2+2\epsilon}} n$, and let $\delta = \log_{\frac{1}{1/2+2\epsilon}} (5g)$. Thus, for every $v \in V$

$$k - \delta \leq (N_0(v)+N_1(v)) \log_{\frac{1}{1/2+2\epsilon}} \left(\frac{1}{1/4-\epsilon}\right) + N_2(v) = k - \log_{\frac{1}{1/2+2\epsilon}} g \leq k.$$

Therefore,

$$\begin{aligned} N_2(v) &\leq k - (N_0(v)+N_1(v)) \log_{\frac{1}{1/2+2\epsilon}} \left(\frac{1}{1/4-\epsilon}\right) \\ &= k - (2+\epsilon_1)((N_0(v)+N_1(v)) \\ &< k - 2((N_0(v)+N_1(v)) \end{aligned}$$

and also

$$N_2(v) \geq k - \delta - (N_0(v)+N_1(v)) \log_{\frac{1}{1/2+2\epsilon}} \left(\frac{1}{1/4-\epsilon}\right)$$

$$= k - \delta - (2+\epsilon_1)((N_0(v) + N_1(v)).$$

For $v \in V'$, let $r = N_0(v) + N_1(v)$. Then $N_2(v) < k - 2r$, and thus the length of v is less than $k - r$. We bound the number of edges from v to vertices u such that the length of the vector u is greater than or equal to the length of the vector v as follows:

If u is a vertex as described above, then there can be $0 \leq j \leq r$ coordinates in which v has 0 or 1 and u has 2. In the other coordinates in which v has 0 or 1, if v has 0 then u has 1 and vice versa. There are $\binom{r}{j}$ possibilities for choosing the coordinates in which v has 0 or 1 and u has 2, for each $0 \leq j \leq r$.

Note also that u can also have 0 or 1 in p coordinates in which v has 2. Moreover, $0 \leq p < k - 2r$, since $N_2(v) < k - 2r$.

The number of different ways to choose these p coordinates is $\binom{N_2(v)}{p} \leq \binom{k-2r}{p}$, and the number of possibilities for the values of u in these coordinates (0 or 1) is 2^p.

After choosing the above we have a prefix of length as that of v, of vectors which are adjacent to v. Let us denote this prefix by x. We claim that the number of vertices in V' with prefix x is at most

$$\frac{n}{g}(\frac{1}{2} + 2\epsilon)^{N_2(x)}(\frac{1}{4} - \epsilon)^{N_0(x)+N_1(x)}.$$

To prove this claim note that V' can be constructed by starting with a single vertex of weight n indexed by the empty vector, and by repeatedly splitting each vertex of weight $w \geq 5g$ indexed by u into three vertices, indexed by $u0$, $u1$ and $u2$, of weights $w(\frac{1}{4} - \epsilon)$, $w(\frac{1}{4} - \epsilon)$ and $w(\frac{1}{2} + 2\epsilon)$ respectively. Since the weight of the vertex indexed by x is $n(\frac{1}{2} + 2\epsilon)^{N_2(x)}(\frac{1}{4} - \epsilon)^{N_0(x)+N_1(x)}$, and by the end of the splitting the weight of each vertex is at least g, the desired claim follows. Since $(\frac{1}{2} + 2\epsilon)^{k-\delta} = \frac{5g}{n}$ the number of vectors in V' with prefix x is at most

$$\frac{n}{g}(\frac{1}{2} + 2\epsilon)^{N_2(x)}(\frac{1}{4} - \epsilon)^{N_0(x)+N_1(x)}$$
$$\leq \frac{n}{g}(\frac{1}{2} + 2\epsilon)^{k-\delta-(2+\epsilon_1)r+j-p}(\frac{1}{4} - \epsilon)^{r-j+p}$$
$$= 5(\frac{1}{1/2 + 2\epsilon})^{(2+\epsilon_1)r-j+p}(\frac{1}{4} - \epsilon)^{r-j+p}$$
$$< 5 \cdot 2^{(2+\epsilon_1)r-j+p}(\frac{1}{4})^{r-j+p}$$
$$= 5 \cdot 2^{j-p}2^{\epsilon_1 r}$$
$$\leq 5n^{\epsilon_2}2^{j-p},$$

where the last inequality follows from the fact that $r \leq k/2$ and thus $2^r \leq 2^{k/2} < n$. Therefore the number of edges from a vertex v with $N_0(v) + N_1(v) = r$ to vertices with vectors of at least the same length is bounded by:

$$\sum_{j=0}^{r}\binom{r}{j}\sum_{p=0}^{k-2r}\binom{k-2r}{p}2^p 5n^{\epsilon_2}2^{j-p} = 5n^{\epsilon_2}\sum_{j=0}^{r}\binom{r}{j}2^j\sum_{p=0}^{k-2r}\binom{k-2r}{p}.$$

As $r \le k/2$ and the number of vertices of V' with $N_0(v) + N_1(v) = r$ is at most $\binom{k-r}{r}2^r$, the total number of edges is bounded by

$$5n^{\epsilon_2} \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} 2^r \sum_{j=0}^{r} \binom{r}{j} 2^j \sum_{p=0}^{k-2r} \binom{k-2r}{p}$$

$$\le 5n^{\epsilon_2} \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} 2^r \sum_{j=0}^{r} \binom{r}{j} 2^j \sum_{p=0}^{k-2r} \binom{k-2r}{p}$$

$$= 5n^{\epsilon_2} \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} 2^r \sum_{j=0}^{r} \binom{r}{j} 2^j 2^{k-2r}$$

$$= 5n^{\epsilon_2} 2^k \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} 2^{-r} \sum_{j=0}^{r} \binom{r}{j} 2^j$$

$$= 5n^{\epsilon_2 + \log_{\frac{1}{1/2+2\epsilon}} 2} \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} 2^{-r} 3^r$$

$$= 5n^{1+\epsilon_3} \sum_{r=0}^{\lfloor k/2 \rfloor} \binom{k-r}{r} (\frac{3}{2})^r.$$

The last formula is a sum of $\lfloor k/2 \rfloor + 1$ terms. Although it is not difficult to compute the sum precisely, we prefer to bound it as follows. Its value is clearly at most $5n^{1+\epsilon_3}(\lfloor k/2 \rfloor + 1)max(f(r))$, where $f(r) = \binom{k-r}{r}(\frac{3}{2})^r$, and $0 \le r \le \lfloor k/2 \rfloor$. Observe that

$$\frac{f(r+1)}{f(r)} = \frac{\binom{k-(r+1)}{r+1}(\frac{3}{2})^{r+1}}{\binom{k-r}{r}(\frac{3}{2})^r} = \frac{3}{2}\frac{(k-2r)(k-2r-1)}{(k-r)(r+1)}.$$

The function $f(r)$ is ascending when $\frac{f(r+1)}{f(r)} > 1$. This happens when $r < (1+o(1))\frac{7-\sqrt{7}}{14}k$ and when $r > (1+o(1))\frac{7+\sqrt{7}}{14}k$. But $r \le \lfloor k/2 \rfloor$, therefore $f(r)$ reaches its maximum for $r = (1+o(1))\frac{7-\sqrt{7}}{14}k$. Thus,

$$max(f(r)) = f((1+o(1))\frac{7-\sqrt{7}}{14}k)$$

$$\le 2^{\epsilon_4 k} \binom{\frac{7+\sqrt{7}}{14}k}{\frac{7-\sqrt{7}}{14}k} (\frac{3}{2})^{\frac{7-\sqrt{7}}{14}k}$$

$$\le 2^{\epsilon_4 k} (\frac{3}{2})^{\frac{7-\sqrt{7}}{14}k} \frac{(\frac{7+\sqrt{7}}{14}k)^{\frac{7+\sqrt{7}}{14}k}}{(\frac{7-\sqrt{7}}{14}k)^{\frac{7-\sqrt{7}}{14}k}(\frac{1}{\sqrt{7}}k)^{\frac{1}{\sqrt{7}}k}}$$

$$= n^{\alpha+\epsilon_5},$$

where $\alpha = \frac{7-\sqrt{7}}{14}\log\frac{3}{2} + \frac{7+\sqrt{7}}{14}\log\frac{7+\sqrt{7}}{14} - \frac{7-\sqrt{7}}{14}\log\frac{7-\sqrt{7}}{14} - \frac{1}{\sqrt{7}}\log\frac{1}{\sqrt{7}} = 0.866\ldots$

Thus, the number of edges of G' is less than

$$5\lfloor k/2 \rfloor n^{1+\epsilon_3+\alpha+\epsilon_5} = O(n^{1.866\ldots+\epsilon_6}),$$

and since the number of edges of G is $|E| = O(|E'|\log^2 n)$ we get the desired result.

We have thus proved the following:

Theorem 2.4 *There exists an explicit $\mathcal{H}(3,n)$-universal graph with $O(n^{1.867})$ edges.*

Remark: From the analysis it follows that the number of edges of G above is at least

$$n^{-\epsilon_7}\sum_{r=0}^{\lfloor k/2\rfloor}\binom{k-r}{r}2^r\sum_{j=0}^{r}\binom{r}{j}2^j\sum_{p=0}^{k-2r}\binom{k-2r}{p}\geq$$

$$\geq n^{-\epsilon_7}\cdot max(\binom{k-r}{r}2^r\binom{r}{j}2^j\binom{k-2r}{p})$$

where the maximum is taken over all admissible r, j, p.

We now find the values of p, j, and r that provide the maximum. The value of p in this term is the value for which $\binom{k-2r}{p}$ reaches its maximum, which is $k/2-r$. For $p=k/2-r$, $\binom{k-2r}{p}=\Theta(\frac{2^{k-2r}}{\sqrt{k-2r}})$, and therefore the term behaves like $\binom{k-r}{r}2^r\binom{r}{j}2^j2^{k-2r}$. The value of j in the maximal term is the j for which $\binom{r}{j}2^j$ reaches its maximum, which is $\frac{2}{3}r$, and then $\binom{r}{j}2^j=\Theta(\frac{3^r}{\sqrt{r}})$. Then the maximal term is roughly $\binom{k-r}{r}2^r3^r2^{k-2r}=2^k\binom{k-r}{r}(\frac{3}{2})^r$. We now have the same expression that we had when we calculated the sum. We have seen that it reaches its maximum for $r\approx\frac{7-\sqrt{7}}{14}k$. Then we have $p=k/2-r\approx\frac{\sqrt{7}}{14}k$, and $j=\frac{2}{3}r\approx\frac{7-\sqrt{7}}{21}k$. Thus $p\leq j$, and hence the bound we have found is essentially the correct number of edges of G.

It will be interseting to close the gap between the $O(n^{1.866\ldots})$ upper bound for the minimum possible number of edges of an $\mathcal{H}(3,n)$-universal graph on n vertices proved here, and the simple lower bound of $\Omega(n^{\frac{4}{3}})$ mentioned in [1].

3 $\mathcal{E}_n$-universal graphs

Let $f(\mathcal{H})$ denote the minimum possible number of edges in an $\mathcal{H}$-universal graph. In this section we study the minimum possible number of edges in an $\mathcal{E}_n$-universal graph. The best known result is given in [3], where the authors prove that

$$\frac{cn^2}{\log^2 n}<f(\mathcal{E}_n)<(1+o(1))\frac{n^2\log\log n}{\log n},$$

for some absolute constant $c>0$.

In this section we prove that $f(\mathcal{E}_n)=\Theta(\frac{n^2}{\log^2 n})$.

Theorem 3.1 *There exist two positive constants c_1 and c_2 such that for all n*

$$c_1\frac{n^2}{\log^2 n}\leq f(\mathcal{E}_n)\leq c_2\frac{n^2}{\log^2 n}.$$

The fact that $f(\mathcal{E}_n)\geq\Omega(\frac{n^2}{\log^2 n})$ is proved in [3] by a simple counting argument. It also follows from the fact that for $r=\lfloor\log n\rfloor$ and $M=2\lfloor\frac{n}{\log n}\rfloor$, any $\mathcal{E}_n$-universal graph must contain all the r-regular graphs on M vertices. Therefore, by a result proved in [1],

$$f(\mathcal{E}_n)=\Omega(M^{2-2/r})=\Omega((\frac{n}{\log n})^{2-2/\log n})=\Omega(\frac{n^2}{\log^2 n}).$$

In the rest of this section we prove the upper bound. We make no attempt to optimize the absolute constants, and we omit all floor and ceiling signs whenever these are not crucial. All logarithms are in base 2. Throughout the proof we assume, whenever this is needed, that n is sufficiently large.

We construct an $\mathcal{E}_n$-universal graph $G = (V, E)$. Let $V = V_0 \cup V_1 \cup \ldots \cup V_k$, where $k = \lceil \log \log n \rceil$, V_0 is set of $2x_0 = \frac{4n}{\log^2 n}$ vertices, and for all $1 \le i \le k$, V_i is a set of $2x_i = \frac{4n2^i}{\log n}$ vertices. Each vertex in V_0 is connected to all the other vertices of G, the graph on V_1 is a complete graph, and for all $2 \le i \le k$, for every $u \in V_i$ and $v \in V_1 \cup V_2 \cup \ldots \cup V_i$, $u \ne v$, we let (u, v) be an edge, randomly and independently, with probability $\min(1, \frac{c}{8^i})$, for some constant c to be specified later.

The number of edges between the vertices of V_0 and all the vertices of G is

$$\binom{|V_0|}{2} + |V_0| \cdot |V \setminus V_0|$$

$$< \frac{(\frac{4n}{\log^2 n})^2}{2} + \frac{4n}{\log^2 n}(\sum_{i=1}^{k} \frac{4n2^i}{\log n})$$

$$= \frac{8n^2}{\log^4 n} + \frac{16n^2}{\log^3 n}(2^{k+1} - 2)$$

$$< \frac{8n^2}{\log^4 n} + \frac{16n^2}{\log^3 n} 2^{\lceil \log \log n \rceil + 1}$$

$$< \frac{8n^2}{\log^4 n} + \frac{64n^2}{\log^2 n}$$

$$\le \frac{72n^2}{\log^2 n}.$$

The number of edges in V_1 is $\binom{|V_1|}{2} < \frac{(\frac{8n}{\log n})^2}{2} = \frac{32n^2}{\log^2 n}$.

For each $2 \le i \le k$, the expected number of edges between the vertices of V_i and the vertices of $V_1 \cup V_2 \cup \ldots \cup V_i$ is at most

$$\frac{c}{8^i} \cdot |V_i| \cdot |V_1 \cup V_2 \cup \ldots \cup V_i|$$

$$= \frac{c}{8^i} \cdot \frac{4n2^i}{\log n}(\sum_{j=1}^{i} \frac{4n2^j}{\log n})$$

$$= \frac{16cn^2}{4^i \log^2 n}(2^{i+1} - 2)$$

$$< \frac{32cn^2}{2^i \log^2 n}.$$

Thus, the expected number of edges in G is less than

$$\frac{72n^2}{\log^2 n} + \frac{32n^2}{\log^2 n} + \sum_{i=2}^{k} \frac{32cn^2}{2^i \log^2 n} <$$

$$< \frac{104n^2}{\log^2 n} + \frac{16cn^2}{\log^2 n}$$

$$= \frac{(104+16c)n^2}{\log^2 n}.$$

Therefore, by Markov's inequality, with probability at least $\frac{1}{2}$, G contains at most $\frac{(208+32c)n^2}{\log^2 n}$ edges.

Lemma 3.2 *Let $G(l,p)$ be a random graph on l vertices, where $l = \sum_{j=1}^{i} 2x_j = \sum_{j=1}^{i} \frac{4n2^j}{\log n} = \frac{4n(2^{i+1}-2)}{\log n} < \frac{8n2^i}{\log n}$, $p = \min(1, \frac{c}{8^i})$ for some constant $c \geq 8^3$ and $i \leq \lceil \log\log n \rceil$. Let W be a subset of the vertices of G, such that $|W| = \frac{2x_i}{\log^2 n} = \frac{4n2^i}{\log^3 n}$. Then the following holds with probability at least $1-e^{-n^{0.3}}$. For every $r \leq \frac{|W|}{2}$ and every collection $\{S_1, S_2, \ldots, S_r\}$ of pairwise disjoint sets outside W, such that for all $1 \leq j \leq r$, $|S_j| \leq \frac{2\log n}{2^i}$, and every subset X of W satisfying $|X| = |W| - r + 1$, there exists a vertex $u \in X$ and $1 \leq j \leq r$ such that u is connected to all the vertices in S_j.*

Proof: Fix $u \in X$ and $1 \leq j \leq r$. The probability that u is connected to all the vertices in S_j is at least $\min(1, (\frac{c}{8^i})^{2\log n/2^i})$. As $c \geq 8^3$, it follows that for $i \leq 3$, this probability is 1, and for $i \geq 4$, if $c \geq 8^i$ then the probability is 1, and otherwise it is at least $(\frac{c}{8^i})^{2\log n/2^i} \geq (\frac{1}{8^{i-3}})^{2\log n/2^i} = (\frac{1}{2^{3i-9}})^{2\log n/2^i} = \frac{1}{n^{2(3i-9)/2^i}} > \frac{1}{\sqrt{n}}$.

Therefore, the probability that there is no vertex $u \in X$ and $1 \leq j \leq r$ such that u is connected to all the vertices in S_j is at most

$$(1 - \frac{1}{\sqrt{n}})^{(|W|-r+1)r}$$
$$< (1 - \frac{1}{\sqrt{n}})^{\frac{|W|}{2}r}$$
$$\leq e^{-\frac{2n2^i}{\log^3 n}\frac{1}{\sqrt{n}}r}$$
$$\leq e^{-n^{0.4}r},$$

for all sufficiently large n.

The number of possibilities to choose $r, S_1, S_2, \ldots, S_r$ and X is at most

$$\frac{|W|}{2}\binom{l}{\frac{2\log n}{2^i}}^r \binom{|W|}{|W|-r} < n(8n)^{r\log n}n^r < n^{4r\log n}n^{2r} \leq e^{5r\log^2 n}.$$

Thus, with probability at least $1 - e^{-n^{0.3}}$ the assertion of the lemma holds. □

Corollary 3.3 *There exists a graph G on the vertices $V = V_0 \cup V_1 \cup \ldots \cup V_k$, where $k = \lceil \log\log n \rceil$, with the following properties.*

1. *$|V_0| = 2x_0 = \frac{4n}{\log^2 n}$, and for all $1 \leq i \leq k$, $|V_i| = 2x_i = \frac{4n2^i}{\log n}$, and $V_i = V_{i1} \cup V_{i2} \cup \ldots \cup V_{i\log^2 n}$, where for all $1 \leq j \leq \log^2 n$, $|V_{ij}| = \frac{|V_i|}{\log^2 n}$, and all sets V_{ij} are pairwise disjoint.*

2. *The number of edges of G is at most $\frac{10000n^2}{\log^2 n}$.*

3. *The vertices of V_0 are connected to all the vertices of G.*

4. *The induced subgraph on V_1 is a complete graph.*

5. *For all* $2 \le i \le k$, $1 \le j \le \log^2 n$ *and* $1 \le r \le \frac{|V_{ij}|}{2}$, *and for every collection* $\{S_1, S_2, \ldots, S_r\}$ *of pairwise disjoint subsets of* $\cup_{s=1}^{i} V_s \setminus V_{ij}$, *such that for all* $1 \le t \le r$, $|S_t| \le \frac{2 \log n}{2^i}$, *and every subset* X *of* V_{ij} *satisfying* $|X| = |V_{ij}| - r + 1$, *there exists a vertex* $u \in X$ *and* $1 \le t \le r$ *such that* u *is connected to all the vertices in* S_t.

Proof: This follows directly from Lemma 3.2, by taking, say, $c = 8^3$. □

We next show that every graph G satisfying the assertion of Corollary 3.3 is $\mathcal{E}_n$-universal.

Let $H \in \mathcal{E}_n$. Then H has n edges and $m \le 2n$ vertices. Let $v_1, v_2, \ldots, v_m$ be the vertices of H such that $d(v_1) \ge d(v_2) \ge \ldots \ge d(v_m)$, where $d(v_i)$ is the degree of v_i in H. Partition $v_1, v_2, \ldots, v_m$ into blocks $B_0 = v_1, v_2, \ldots, v_{x_0}$, $B_1 = v_{x_0+1}, v_{x_0+2}, \ldots, v_{x_0+x_1}$ and so on. To complete the proof we show that there is an embedding of H in G such that for all i, B_i is mapped injectively into V_i.

We choose an arbitrary injective mapping from B_0 to V_0 and from B_1 to V_1. Let us assume that we have already found a mapping from B_j to V_j for all $j < i$, $i \ge 2$, such that all the needed edges in the induced subgraph on the images of the vertices of $\cup_{j<i} B_j$ exist. For all $v \in B_i$, $2n \ge \sum_{u \in V(H)} d(u) \ge \sum_{u \in B_{i-1}} d(u) \ge \frac{d(v) 2n 2^{i-1}}{\log n}$. Thus, $d(v) \le \frac{2 \log n}{2^i}$.

Let F be the graph whose vertices are all the vertices in B_i, where two vertices are connected iff they are either connected in H or have a common neighbour in $H \setminus B_0$. Each vertex v of B_i is adjacent in H to at most $\frac{2 \log n}{2^i}$ other vertices, and as the degree of each vertex of $H \setminus B_0$ is at most $\log n$, there are at most $\frac{2 \log n}{2^i} \log n$ paths of length 2 in $H \setminus B_0$ starting at v. Therefore, as $i \ge 2$, the maximum degree in F is at most $\frac{2 \log n}{2^i} + \frac{2 \log n}{2^i} \log n < \log^2 n - 1$. By the Hajnal-Szemerédi theorem [7], there is a partition of B_i into $\log^2 n$ independent sets B_{ij} of equal sizes, such that no two vertices in the same set have a common neighbour outside B_0.

We now embed the sets B_{ij} into V_{ij} one by one.

Let $B_{ij} = \{v_1, v_2, \ldots, v_{\frac{2n2^i}{\log^3 n}}\}$. For each $1 \le t \le \frac{2n2^i}{\log^3 n}$, let S_t be the set of the vertices in $V_1 \cup V_2 \cup \ldots \cup V_i$ to which the neighbours of v_t that have already been mapped were mapped. Since no two vertices in B_{ij} have a common neighbour outside V_0, the sets $S_1, S_2, \ldots, S_{\frac{2n2^i}{\log^3 n}}$ are pairwise disjoint. Let $G' = (B_{ij}, V_{ij}, E')$ be a bipartite graph, where for each $1 \le t \le \frac{2n2^i}{\log^3 n}$ and $u \in V_{ij}$, $(v_t, u) \in E'$ iff u is connected to all the vertices of S_t. By Corollary 3.3, for all $r \le \frac{2n2^i}{\log^3 n}$ and for all $U \subseteq B_{ij}$ such that $|U| = r$, the set of neighbours of U in V_{ij} is of size at least r. Thus, by Hall's theorem, there is a matching in G' saturating all members of B_{ij}. The mapping of B_{ij} into V_{ij} is obtained by this matching. This completes the proof of Theorem 3.1. □

Note that our constraction is probabilistic. It may be interesting to describe an explicit construction of an $\mathcal{E}_n$-universal graph with $\Theta(\frac{n^2}{\log^2 n})$ edges.

4 Acknowledgement

We would like to thank the referees for many helpful comments.

References

[1] N. Alon, M. R. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński and E. Szemerédi, Universality and tolerance, 41^{st} IEEE FOCS, IEEE (2000), 14-21.

[2] N. Alon and J. H. Spencer, The Probabilistic Method, Wiley (1992).

[3] L. Babai, F. R. K. Chung, P. Erdös, R. L. Graham and J. H. Spencer, On graphs which contain all sparse graphs, Ann. Discrete Math. 12 (1982), 21-26.

[4] S. N. Bhatt, F. Chung, F. T. Leighton and A. Rosenberg, Universal graphs for bounded-degree trees and planar graphs, SIAM J. Disc. Math. 2 (1989), 145-155.

[5] M. R. Capalbo, A smaller universal graph for general max-degree 3 graphs, preprint (1999).

[6] M. R. Capalbo and S. R. Kosaraju, Small universal graphs, Proc. STOC (1999), 741-749.

[7] A. Hajnal and E. Szemerédi, Proof of a conjecture of Erdös, in: *Combinatorial Theory and its Applications*, Vol. II (P. Erdös, A. Renyi and V. T. Sós eds.), Colloq. Math. Soc. J. Bolyai 4, North Holland, Amsterdam 1970, 601-623.