Universality for graphs with bounded density

Noga Alon* Natalie Dodson† Carmen Jackson‡ Rose McCarty§
Rajko Nenadov¶ Lani Southern‖

Abstract

A graph G is *universal* for a (finite) family $\mathcal{H}$ of graphs if every $H \in \mathcal{H}$ is a subgraph of G. For a given family $\mathcal{H}$, the goal is to determine the smallest number of edges an $\mathcal{H}$-universal graph can have. With the aim of unifying a number of recent results, we consider a family of graphs with bounded density. In particular, we construct a graph with

$$O_d\left(n^{2-1/(\lceil d\rceil+1)}\right)$$

edges which contains every n-vertex graph with density at most $d \in \mathbb{Q}$ ($d \geq 1$), which is close to a lower bound $\Omega(n^{2-1/d-o(1)})$ obtained by counting lifts of a carefully chosen (small) graph. When restricting the maximum degree of such graphs to be constant, we obtain a near-optimal universality. If we further assume $d \in \mathbb{N}$, we get an asymptotically optimal construction.

1 Introduction

A graph G is *universal* for a (finite) family $\mathcal{H}$ of graphs if every $H \in \mathcal{H}$ is a (not necessarily induced) subgraph of G. The complete graph with n vertices is universal for the family of all graphs with n vertices, and this is clearly the smallest universal graph for this family. However, if we restrict our attention to a family of graphs with some additional properties, more efficient (in terms of the number of edges) universal graphs might exist. This is a natural combinatorial question, with applications in VLSI circuit design [15], data storage [23], and simulation of parallel computer architecture [14].

The problem of estimating the minimum possible number of edges in a universal graph for various families has received a considerable amount of attention. The previous work deals with families of graphs with properties which naturally bound their density, such as graphs with bounded maximum degree [5, 7, 6, 8, 9], forests [20, 21, 22, 28] and, more generally, graphs with bounded degeneracy [2, 33], as well as families of graphs with additional structural properties such as planar graphs [12, 26] and graphs with small separators [17, 18, 19], to name a few. Our focus is on the former case. Aiming to unify these results, we initiate the study of universality for a family of graphs with bounded density and no other assumptions. The density of a graph H is defined as

$$m(H) = \max_{H' \subseteq H} \frac{e(H')}{v(H')}$$

where $e(H')$ is the numbers of edges of H' and $v(H')$ is the number of its vertices. In plain words, a graph G has density $d \in \mathbb{Q}$ if not only the number of edges of G is at most $v(G)d$, but this also holds for every subgraph of G. For $d \in \mathbb{Q}$ and $n \in \mathbb{N}$, we denote with $\mathcal{H}_d(n)$ the family of all graphs with n vertices and density at most d.

*Department of Mathematics, Princeton University, USA. Email: nalon@math.princeton.edu. This author was supported by the NSF under grant DMS-2154082.

†Middlebury College, USA. Email: ndodson@middlebury.edu. This author was supported by the NSF under grant DMS-2150434.

‡Northwestern University, USA. Email: carmenjackson2024@u.northwestern.edu. This author was supported by the NSF under grant DMS-2150434.

§Department of Mathematics, Princeton University, USA. Current affiliation: School of Mathematics and School of Computer Science, Georgia Institute of Technology, USA. Email: rmccarty3@gatech.edu. This author was supported by the NSF under grants DMS-2150434 and DMS-2202961.

¶School of Computer Science, University of Auckland, New Zealand. Email: rajko.nenadov@auckland.ac.nz

‖Willamette University, USA. Email: lmsouthern@willamette.edu. This author was supported by the NSF under grant DMS-2150434.

As already hinted, the family of graphs with bounded density generalizes many interesting families. For example, graphs with maximum degree d have density at most $d/2$. Forests have density arbitrarily close to 1, and d-degenerate graphs and, more generally, graphs of arboricity d, have density arbitrarily close to d (a graph is d-degenerate if every subgraph has minimum degree at most d). Note that every graph of density at most d is also $\lfloor 2d \rfloor$-degenerate, thus bounded density implies bounded degeneracy. However, as we are aiming for an optimal dependence on the parameters, this implication does not suffice. A number of (almost-)optimal results have been obtained in some of these cases [2, 7, 21], and we conjecture that for all such families the bound on the size of a smallest universal graph is largely governed by the density. Therefore, generalizing all of these results, we believe the following is true.

Conjecture 1.1. *For every $d \in \mathbb{Q}$, $d > 1$, and $n \in \mathbb{N}$, there exists a graph with G with*

$$e(G) \leq Cn^{2-1/d}$$

edges which is $\mathcal{H}_d(n)$-universal, where $C = C(d)$.

If true, the bound in Conjecture 1.1 is the best possible up to a constant C. Indeed, a simple counting argument shows that if $e(G) = o(n^{2-1/d})$ then the number of graphs with density at most d which can possibly appear in G is less than the total number of such graphs. Moreover, we obtain such a lower bound even when restricting $\mathcal{H}_d(n)$ to graphs with bounded maximum degree. This is summarised in the following proposition.

Proposition 1.2. *For every $d \in \mathbb{Q}$ ($d \geq 1$) there exists $D \in \mathbb{N}$ and $\alpha > 0$ such that the following holds for every sufficiently large n. If G is an $\mathcal{H}_d^D(n)$-universal graph then*

$$e(G) \geq \alpha n^{2-1/d},$$

where $\mathcal{H}_d^D(n)$ is the family of all graphs $H \in \mathcal{H}_d(n)$ with maximum degree at most D.

A careful reader will notice that we require $d > 1$ in Conjecture 1.1. Indeed, if a graph H has density at most 1 then each connected component of H contains at most one cycle, thus H is almost a forest. For the family of all forests it is known that $\Theta(n \log n)$ edges are both necessary and sufficient [20, 21]. This seems to be an artifact of the fact that having only $\Theta(n)$ edges is simply too restrictive with how we can arrange them, which we believe is not the case when we have $O(n^{1+\varepsilon})$ edges for any constant $\varepsilon > 0$. While this justification is vague, one can draw analogy with the theory of random graphs, where the multiplicative $\log(n)$ factor becomes unnecessary when moving from Hamilton cycles to, say, powers of Hamilton cycles [30] (signifying the difference between unicyclic graphs and those of density $d > 1$).

A reader familiar with random graph theory [16] will notice that a random graph with n vertices and $\omega(n^{2-1/d})$ edges is likely to contain a given graph H with $m(H) \leq d$, provided $v(H)$ is significantly smaller than n. However, it is known that if $v(H)$ is large then, in some cases such as when H is a collection of many triangles, a significantly denser random graph is needed in order for H to appear. It is interesting that from the point of view of constructing universal graphs, this phenomena does not happen.

A recent result of Allen, Böttcher, and Liebenau [2] establishes the bound in Conjecture 1.1 (up to a $\log^{2/d}(n)$ factor) in the special case where we restrict attention to the graphs in $\mathcal{H}_d(n)$ which are also d-degenerate (for d an integer). Moreover, using the fact that a graph of density d is $\lfloor 2d \rfloor$-degenerate, their result also implies an upper bound of order

$$O_d\left(n^{2-1/\lfloor 2d \rfloor} \log^{1/(2d)+o(1)}(n)\right).$$

for $\mathcal{H}_d(n)$-universality. Our first main result, Theorem 1.3, significantly improves this starting from $d \geq 1.5$.

Theorem 1.3. *For every $n \in \mathbb{N}$ and $d \in \mathbb{Q}$, $d > 1$, there exists a graph G with*

$$e(G) \leq Cn^{2-1/(\lceil d \rceil + 1)}$$

edges which is $\mathcal{H}_d(n)$-universal, where $C = C(d)$.

As a further support towards Conjecture 1.1, we consider the family $\mathcal{H}_d^D(n)$ of graph $H \in \mathcal{H}_d(n)$ with maximum degree D. In this case, we get a nearly-optimal bound.

Theorem 1.4. *For every $D, n \in \mathbb{N}$ and $d \in \mathbb{Q}$, $d > 1$, there exists a graph G with*

$$e(G) \leq n^{2-1/d} \cdot 2^{C\sqrt{\log n}}$$

edges which is $\mathcal{H}_q^D(n)$-universal, where $C = C(D, d)$.

Finally, in the case where d is an integer, we obtain an optimal bound.

Theorem 1.5. *For every $D, n \in \mathbb{N}$ and $d \in \mathbb{N}$, there exists a graph G with*

$$e(G) \leq Cn^{2-1/d}$$

edges which is $\mathcal{H}_d^D(n)$-universal, where $C = C(D, d)$.

Let us briefly compare these results with the previous ones. Alon and Capalbo [7] constructed a graph with $O_D(n^{2-2/D})$ edges which is universal for the family of all n-vertex graphs with maximum degree at most D. Theorem 1.5 implies this result, and further generalizes it, in the case D is even. In the case D is odd, Theorem 1.4 provides a bound which is by a factor of $2^{O(\sqrt{\log n})}$ weaker than the one from [7]. However, in comparison with the proof from [7] which relies on the fact that bounded degree graphs can be decomposed into path-like pieces, we use a much more general decomposition (Lemma 3.1) which applies to all graphs with bounded density. Moreover, Theorem 1.5 improves a result of the fifth author [33] on universality of d-degenerate graphs with bounded degree, to an optimal one. That being said, improving Theorem 1.4 is a natural first step towards Conjecture 1.1.

Conjecture 1.6. *For every $D, n \in \mathbb{N}$ and $d \in \mathbb{Q}$, $d \geq 1$, there exists a graph G with*

$$e(G) \leq Cn^{2-1/d}$$

edges which is $\mathcal{H}_d^D(n)$-universal, where $C = C(D, d)$.

Unlike in the case of graphs with arbitrarily large degree, an additional multiplicative $\log(n)$ factor is not needed even in the case of forests [28]. Thus we can relax the condition to $d \geq 1$.

Finally, let us briefly note that our constructions of universal graphs are based on a *product construction* first used by Alon and Capalbo [6, 7], refining an earlier approach by Alon, Capalbo, Kohayakawa, Rödl, Ruciński, and Szemerédi [9]. The proofs here also apply some of the ideas of Beck and Fiala [13] from Discrepancy Theory, results of Feldman, Friedman and Pippenger [27] (see also [24]) from the theory of nonblocking networks, and random walks on expanders, together with the Matroid Decomposition Theorem of Edmonds [25].

The paper is organised as follows. In the next section we prove that the bound in Conjecture 1.1, if true, is optimal even if we only restrict attention to $\mathcal{H}_d^D(n) \subseteq \mathcal{H}_d(n)$, where the bound D on the maximum degree depends on d. Section 3 collects some results used in two or all three proofs. We then proceed with proofs of our main theorems. Comments on differences between proofs are given when appropriate. Throughout the paper we assume, whenever this is needed, that the parameter n is sufficiently large as a function of any other parameter. To simplify the presentation we omit all floor and ceiling signs whenever they are not crucial.

2 Lower bound

Consider a (fixed) graph F such that $m(F) = e(F)/v(F)$. Such graphs are called *balanced*. Let $n \in \mathbb{N}$ be sufficiently large and divisible by $v(F)$. We obtain a large family of n-vertex graphs H with $m(H) = m(F)$ as follows. Set $V(H) = V_1 \cup \ldots \cup V_{v(F)}$, where V_i's are disjoint sets of size $n/v(F)$, and for each $ij \in E(G)$ put a perfect matching between V_i and V_j. The resulting graph H is called a *lift* of F. It is not difficult to check that regardless of which perfect matching we choose, we have $m(H) = m(F)$. The number of such (labelled) graphs H is

$$\left(\left(\frac{n}{v(F)}\right)!\right)^{e(F)} > \left(\frac{n}{3v(F)}\right)^{ne(F)/v(F)}. \tag{1}$$

We use this bound to prove Proposition 1.2.

Proof of Proposition 1.2. The result of Rúcinski and Vince [34] implies that for every $d \in \mathbb{Q}$, $d \geq 1$, there exists a balanced graph F with $m(F) = d$. As a lift H of F has maximum degree $D = v(F)$, we have $H \in \mathcal{H}_d^D(n)$.

Suppose that a graph G contains every lift of F of order n, and let $M = e(G)$. As every lift contains exactly $ne(F)/v(F)$ edges, by (1) we necessarily have

$$\binom{M}{ne(F)/v(F)} n! > \left(\frac{n}{3v(F)}\right)^{ne(F)/v(F)}, \tag{2}$$

as otherwise there is a lift of F which does not appear in G. Note that the $n!$ term on the left hand side takes into account that every choice of $ne(F)/v(F)$ edges accounts for at most $n!$ different labeled subgraphs. We can further upper bound the left hand side of (2) as follows:

$$\binom{M}{ne(F)/v(F)} n! < \left(\frac{3M}{n}\right)^{ne(F)/v(F)} n^n = \left(\frac{3M}{n^{1-v(F)/e(F)}}\right)^{ne(F)/v(F)}.$$

Comparing this with the right hand side of (2), we conclude

$$M > \frac{1}{9v(F)} n^{2-v(F)/e(F)} = \frac{1}{9v(F)} n^{2-1/m(F)}.$$

□

3 Preliminaries

In the following lemma we identify a graph with its edge set. We say that a graph is *unicyclic* if it contains at most one cycle. The proofs of all our three main theorems are based on the decomposition given by the following lemma. Its proof applies some basic results from Matroid Theory, see, e.g, [35] for the relevant notions.

Lemma 3.1. *Let H be a simple graph satisfying $m(H) \geq 1$ and $H^{(b)}$ a multigraph obtained from H by duplicating each edge $b \in \mathbb{N}$ times. Then there exists a partition $H^{(b)} = H_1 \cup \ldots \cup H_k$, where $k = \lceil b \cdot m(H) \rceil$, such that, for every $i \in [k]$, each component of H_i is a simple unicyclic graph.*

Proof. It suffices to prove the statement in the case H is connected.

Let $\mathcal{B}$ be the family of all spanning subgraphs $B \subseteq H^{(b)}$ which are simple, connected, and unicyclic. We will shortly prove that $\mathcal{M} = (E(H^{(b)}), \mathcal{B})$ is a matroid, with $\mathcal{B}$ being the family of bases of $\mathcal{M}$. Assume for now that this is true. Consider some $H' \subseteq H^{(b)}$, and let H'' be the graph obtained from H' by removing duplicate edges. Then the rank $r(H')$ of H' in $\mathcal{M}$ is $v(H'')$ if H'' contains a cycle, and $v(H'') - 1$ otherwise (and note that $v(H') = v(H'')$). In the former case we have

$$|H'|/r(H') = e(H')/r(H') \leq b \cdot e(H'')/v(H'') \leq b \cdot m(H).$$

In the latter case we necessarily have $e(H'') < v(H'')$, thus

$$e(H')/r(H') \leq b \cdot (v(H'') - 1)/(v(H'') - 1) = b \leq b \cdot m(H).$$

By a result of Edmonds [25], one can cover H with $\lceil b \cdot m(H) \rceil$ disjoint independent sets from $\mathcal{M}$, which proves the lemma.

Let us verify that $\mathcal{M}$ is indeed a matroid. Consider some distinct bases $X, Y \in \mathcal{B}$, and let $e \in X \setminus Y$. We aim to find $e' \in Y \setminus X$ such that $(X \setminus \{e\}) \cup \{e'\} \in \mathcal{B}$. If e lies on the cycle of X, then $X \setminus \{e\}$ is a tree thus adding any edge $e' \in Y \setminus X$ produces a spanning unicyclic graph. Otherwise, $X \setminus \{e\}$ breaks X into two components, a tree X_1 and a unicyclic X_2. As $e \notin Y$, there has to exist $e' \in Y \setminus X$ with one endpoint in $V(X_1)$ and the other in $V(X_2)$. Therefore, $(X \setminus \{e\}) \cup \{e'\}$ is again a spanning unicyclic subgraph. □

In the proofs of Theorem 1.3 and Theorem 1.5, we make use of the following simple known lemma (see, e.g., [3, Lemma 2.2]).

Lemma 3.2 ([3]). *For every forest F with n vertices and every $r \in \mathbb{N}$, there exists a subset $R \subseteq V(F)$ of size $|R| \leq r$ such that each connected component in $F \setminus R$ is of size at most n/r.*

In the proofs of Theorems 1.3 and 1.4, we use the notion of a graph *blowup*, thus we define it here for the reference.

Definition 3.3. Given a graph G and $b \in \mathbb{N}$, we define a b-*blowup* of G to be a graph Γ on the vertex set $V(\Gamma) = \bigcup_{u \in V(G)} V_u$, where each V_u is of size b and all the sets are disjoint, and there is an edge between $a \in V_u$ and $b \in V_w$ iff $uv \in E(G)$ or $u = v$. In particular, Γ has $v(G)b$ vertices and $v(G)\binom{b}{2} + e(G)b^2$ edges.

In the proofs of Theorem 1.4 and Theorem 1.5 we make use of the so-called (n, t, λ)-*graphs*. These are t-regular graphs with n vertices and the second largest absolute eigenvalue at most λ. For small λ such graphs are known to be good expanders. We use the following result of the first author [4] which provides an explicit construction of such graphs for an almost optimal value of λ (note that a construction of Lubotzky, Phillips, and Sarnak [31] can be used as well, even though it does not provide a construction for every t).

Theorem 3.4. *For every $t \in \mathbb{N}$ and $n \geq n_0(t)$ such that nt is even, there exist an explicit construction of an (n, t, λ)-graph with $\lambda \leq 3\sqrt{t}$.*

It is worth noting that the proof in [4] and the known results about the Linnik problem imply that $n_0(t) \leq t^{O(1)}$. In particular, this is relevant for the proof of Theorem 1.4.

4 Graphs with bounded density

We need the following lemma from Discrepancy Theory. The proof applies the Beck-Fiala method [13]. We only use it with $q = 1/2$, however, as this does not make the proof any easier, we state it in greater generality.

Lemma 4.1. *The following holds for any positive integers t, d and real $q \in [0, 1]$ Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$ be a sequence of vectors in $\mathbb{R}^d$, each having ℓ_1-norm at most 1. Then there is a subset $I \subseteq \{1, 2, \ldots, t\}$ so that*

$$\left\| \sum_{i \in I} \mathbf{v}_i - q \sum_{i=1}^{t} \mathbf{v}_i \right\|_\infty < 1.$$

Proof. Associate each vector $\mathbf{v}_i = (v_{i1}, v_{i2}, \ldots, v_{id})$ with a real variable $x_i \in [0, 1]$. Starting with $x_i = q$ for all i, we describe an algorithm for rounding each x_i to either 0 or 1 without changing the sum $\sum x_i v_i$ by much. During this algorithm, call a variable x_i *floating* if x_i lies in the open interval $(0, 1)$, otherwise (that is, if $x_i \in \{0, 1\}$) call it *fixed*. Once a variable becomes fixed it will stay fixed until the end. In each phase, the algorithm proceeds as follows. If all variables are fixed, terminate. Otherwise, let $F \subseteq \{1, 2, \ldots, t\}$ denote the set of all indices of floating variables and consider the following linear system of equations in these variables.

For every coordinate $j \in \{1, 2, \ldots, d\}$ for which

$$\sum_{i \in F} |v_{ij}| > 1, \tag{3}$$

include the equation

$$\sum_{i=1}^{t} x_i v_{ij} = q \sum_{i=1}^{t} v_{ij}. \tag{4}$$

Note that only the floating variables $\{x_i \colon i \in F\}$ are considered as variables at this point; the fixed variables are already fixed and are treated as constants. During the algorithm, we maintain the property that for every coordinate j for which (3) holds, the equality (4) holds as well. This is certainly true at the beginning, when $x_i = q$ for all i.

By the assumption about the ℓ_1-norm of the vectors v_i, we have

$$\sum_{i \in F} \sum_{j=1}^{d} |v_{ij}| \leq |F|$$

and therefore the number of indices j for which (3) holds is strictly smaller than $|F|$. Thus there are more variables than equations and hence there is a line of solutions. One can move along this line starting with the existing point on it (that corresponds to the current value of the floating variables x_i) until the first point in which at least one of the floating variables x_i becomes 0 or 1. Fix this variable (as well as any other floating variables that become 0 or 1, if any), and continue with the next phase. Note that the desired property is maintained since no new coordinate j can satisfy (3) as F gets smaller.

Since each phase fixes at least one floating variable, the algorithm must terminate when all the variables x_i are fixed. Now, for a coordinate j, consider the first phase when condition (3) is not satisfied. The equality (4) holds at this phase since it is the first time (3) is not satisfied. Moreover, for the rest of the algorithm, the value of the sum $\sum_{i=1}^{t} x_i v_{ij}$ can only change by strictly less than $\sum_{i\in F} |v_{ij}| \leq 1$, since only the floating variables change after that. This shows that upon termination, for every index $j \in \{1, 2, \ldots, d\}$,

$$\left|\sum_{i=1}^{t} x_i v_{ij} - q\sum_{i=1}^{t} v_{ij}\right| < 1.$$

Therefore the set $I = \{i : x_i = 1\}$ of all indices in which the final value of x_i is 1 satisfies the conclusion of the lemma. □

By repeated application of the previous lemma, we get the following.

Corollary 4.2. *The following holds for any three positive integers $t, d, m = 2^k$. Let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_t$ be a sequence of vectors in $\mathbb{R}^d$, each having ℓ_1-norm at most 1. Then there is a partition of the vectors into m pairwise disjoint sets $\{v_i : i \in I_p\}$, $1 \leq p \leq m$, where $[t] = I_1 \cup \ldots \cup I_m$, the sets I_j are pairwise disjoint, and for every $p \in [m]$ we have*

$$\left\|\sum_{i\in I_p} \mathbf{v}_i - \frac{1}{m}\sum_{i=1}^{t} \mathbf{v}_i\right\|_\infty \leq \sum_{i=0}^{k-1} 2^{-i} < 2.$$

Proof. We prove the statement by induction on k. For $k = 1$ it is equivalent to Lemma 4.1 with $q = 1/2$. Suppose that it hold for $m = 2^{k-1}$, for $k \geq 2$. We show that then it also holds for $m = 2^k$.

Apply Lemma 4.1 with $q = 1/2$ to split the vectors into two collections, $[t] = C_1 \cup C_2$, such that for $i \in \{1, 2\}$ we have

$$\left\|\sum_{i\in C_i} \mathbf{v}_i - \frac{1}{2}\sum_{i=1}^{t} \mathbf{v}_i\right\|_\infty < 1. \tag{5}$$

By the induction hypothesis, there is a partition $C_1 = I_1 \cup \ldots \cup I_{m/2}$ such that for each $p \in [m/2]$ we have

$$\left\|\sum_{i\in I_p} \mathbf{v}_i - \frac{2}{m}\sum_{i\in C_1} \mathbf{v}_i\right\|_\infty \leq \sum_{i=0}^{k-2} 2^{-i}. \tag{6}$$

By the triangle inequality, from (5) and (6) we conclude

$$\left\|\sum_{i\in I_p} \mathbf{v}_i - \frac{1}{m}\sum_{i=1}^{t} \mathbf{v}_i\right\|_\infty = \left\|\left(\sum_{i\in I_p} \mathbf{v}_i - \frac{2}{m}\sum_{i\in C_1} \mathbf{v}_i\right) + \frac{2}{m}\left(\sum_{i\in C_1} \mathbf{v}_i - \frac{1}{2}\sum_{i=1}^{t} \mathbf{v}_i\right)\right\|_\infty \leq \sum_{i=0}^{k-2} 2^{-i} + 2/m.$$

The same argument applies to C_2, which gives a desired partition $[t] = I_1 \cup \ldots \cup I_m$. □

Proof of Theorem 1.3. Note that it suffices to prove the theorem for $d \in \mathbb{N}$.

Let m be the smallest power of 2 that is at least $n^{1/(d+1)}$, and suppose $m \geq 4$. First, form a graph Γ on the vertex set $[m]^d$ where two vertices $\mathbf{u} = (u_1, u_2, \ldots, u_d)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_d)$ are connected if $u_i = v_i$ for some $i \in [d]$. The graph Γ^+ is the $(3m + 3)$-blowup of Γ (see Definition 3.3) together with another set V^+ of $2dn/m$ vertices and all edges incident with at least one of them. Note that Γ^+ has $O(dn^{1-1/(d+1)})$ edges. We proceed to show that it is $\mathcal{H}_d(n)$-universal.

Consider some $H \in \mathcal{H}_d(n)$. Let $H = H_1 \cup \ldots \cup H_d$ be a decomposition given by Lemma 3.1 (with $b = 1$), and recall that each component, of each H_i, is unicyclic. First form $R' \subset V(H)$ as follows: For every $i \in [d]$ and every component of H_i of size at least m, take one vertex from a cycle in that component (if such exist). This adds up to at most dn/m vertices. Next, by applying Lemma 3.2 with $F = H_i \setminus R'$ (which is now a forest) and $r = n/m$ for each $i \in [d]$, we obtain a set $R \subseteq V(H)$ of size $|R| \leq dn/m$ such that each connected component of $H_i \setminus (R \cup R')$ is of size at most m. All the vertices of $R \cup R'$ will be mapped into V^+, thus we can set $H = H \setminus (R \cup R')$ and $H_i = H_i \setminus (R \cup R')$.

Let $\mathcal{C}_i$ denote the family of connected components in H_i (we identify a connected component by its vertex set). For each $h \in V(H)$, let $c_i(h)$ denote the component $K \in \mathcal{C}_i$ that contains h. We show, by induction on i, that there exist functions $\phi_i \colon \mathcal{C}_i \mapsto [m]$ such that for each $i \in [d]$ and every $\mathbf{v} = (v_1, v_2, \ldots, v_i) \in [m]^i$ we have

$$|S_{\mathbf{v}}| \leq \frac{n}{m^i} + 2m + 3 \tag{7}$$

where

$$S_{\mathbf{v}} = \{h \in V(H) : \phi_1(c_1(h)) = v_1, \phi_2(c_2(h)) = v_2, \ldots, \phi_i(c_i(h)) = v_i\}.$$

Once we have this for $i = d$, by injectively mapping each $S_{\mathbf{v}}$ into the blowup of $\mathbf{v}$, for $\mathbf{v} \in [m]^d$, we obtain an embedding of H in Γ^+.

Inequality (7) trivially holds for $i = 0$. Suppose that (7) holds for some $i - 1$, for $i \in [d]$. We show we can find ϕ_i so that it holds for i. For each connected component $K \in \mathcal{C}_i$ define a vector $\mathbf{v}_K$ of length m^{i-1} indexed by the vectors $\mathbf{u} = (u_1, u_2, \ldots, u_{i-1}) \in [m]^{i-1}$ as follows: The coordinate of $\mathbf{v}_K$ indexed by $\mathbf{u}$ is the number of vertices $h \in K$ such that

$$\phi_1(c_1(h)) = u_1,\ \phi_2(c_2(h)) = u_2, \ldots,\ \phi_{i-1}(c_{i-1}(h)) = u_{i-1}.$$

Note that the ℓ_1-norm of each $\mathbf{v}_K$ is the number of vertices of K, which is at most m. In addition, the sum of all the vectors $\mathbf{v}_K$ in each coordinate $\mathbf{u}$ is exactly $|S_{\mathbf{u}}|$, which, by the induction hypothesis, is at most $n/m^{i-1} + 2m + 3$.

By Corollary 4.2 these vectors can be partitioned into m pairwise disjoint collections so that the sum of the vectors in each collection, and with respect to each coordinate, is at most

$$\frac{n/m^{i-1} + 2m + 3}{m} + 2m \leq n/m^i + 2m + 3.$$

The value of $\phi_i(K)$ is now set to be the index of the collection containing K, implying the required inequality for i and completing the proof. □

5 Graphs with bounded density and degree

The basic idea behind the proofs of Theorem 1.4 and Theorem 1.5 is similar to that in the proof of Theorem 1.3. In Theorem 1.3 we obtain $n^{-1/(d+1)}$ instead of desired $n^{-1/d}$ because, in each of the d steps, we assign the same coordinate to all the vertices of a connected component in H_i. Intuitively, if same vertices of H belong to the same connected component across each H_i, this is not sufficient to disambiguate them and we are forced to take a small blowup at the end.

When H has bounded maximum degree, we avoid this by using the following idea, at least in the case $d \in \mathbb{N}$: Our basic graph Γ is again defined on the vertex set $[m]^d$ (now with $m \approx n^{1/d}$), however this time we connect $\mathbf{v}, \mathbf{u} \in [m]^d$ by an edge if some v_i and u_i are connected by an edge in a bounded-degree expander G on the vertex set $[m]$, which we fix upfront. Instead of mapping all the vertices of one component of H_i into a single coordinate, we disperse them across $[m]$ by using edges of the expander G. In Theorem 1.5 we can make this approach disambiguate all the vertices of H, thus avoiding the use of a final blowup all together. In Theorem 1.4 the number of vertices which are pairwise ambiguous ends up being of order $2^{O(\sqrt{\log n})}$, thus we take a very small blowup at the end – significantly smaller than in the proof of Theorem 1.3 – to deal with this.

The proof of Theorem 1.4, presented next, borrows ideas of using random walks in expanders from [6]. One significant difficulty in the proof of Theorem 1.4 is that we are not able to split H_i into small connected components and we have to deal with the whole H_i at once, which further emphasizes dispersion via expanders. The proof of Theorem 1.5 generalizes the approach from [7] from embedding paths in expanders, in a specific way, to embedding bounded-degree trees. This is done using some of the ideas in [27] and [24].

5.1 Density bounded by a rational

We use the following well known property of random walks on expanders, see, e.g., [29].

Lemma 5.1. *Let G be an (n, t, λ)-graph, and consider a random walk starting in a given vertex $v \in V(G)$. The probability that after exactly ℓ steps we finish in a vertex $w \in V(G)$ is at most*

$$1/n + (\lambda/t)^\ell.$$

Randomized tree homomorphism. Given a tree T with the designated root r and a graph G, we use the following randomized procedure for constructing a homomorphism $\phi\colon T \hookrightarrow G$:

(i) Consider any ordering $h_1, \dots, h_n$ of $V(T)$ such that $h_1 = r$ and, for each $i \geq 2$, h_i has exactly one neighbour within $\{h_1, \dots, h_{i-1}\}$.

(ii) Take $s_1 \in V(G)$ to be some upfront chosen vertex in $V(G)$.

(iii) For $i = \{2, \dots, n\}$, sequentially, take s_i to be a neighbour of s_j in G chosen uniformly at random, where $j < i$ is a unique index such that $h_j h_i \in T$.

The homomorphism is then given by $\phi(h_i) := s_i$. Note that the ordering of the vertices $h_2, \dots, h_n$ plays no role in the distribution of ϕ, as long as each vertex other than h_1 has exactly one predecessor.

Lemma 5.2. *Let G be an $(m, t, 3\sqrt{t})$-graph where $t = 2^{\sqrt{\log n}}$ and $n \geq m$. Suppose T is a tree with the root r, and $U \subseteq V(T) \setminus \{r\}$ a subset such that every two $t, t' \in U \cup \{r\}$ are at distance at least $16\sqrt{\log n}$ in T. Let ϕ be a random homomorphism $\phi\colon T \hookrightarrow G$ obtained by the described procedure. Then, for any $v \in V(G)$, the size of the set*

$$U_v = \{u \in U \colon \phi(u) = v\}$$

is stochastically dominated by a binomial random variable $B(|U|, 1/m + 1/n^3)$.

Proof. Let $u_1, \dots, u_k$ be an ordering of the vertices in U such that if u_j is closer to r than u_i, then $j < i$. Let P_1 be the path from r to u_1 and set $x_1 = r$. For each $2 \leq i \leq k$, define the path P_i as follows:

- Let $x_i \in V(T)$ be the first vertex on a path from u_i to r which belongs to $\bigcup_{j<i} V(P_j)$;
- Set P_i to be the path from x_i to u_i.

Importantly, for every $i \in [k]$ we have $|P_i| \geq 8\sqrt{\log n}$. Let us quickly prove this. As $x_i \in \bigcup_{j<i} V(P_j)$ we have $x_i \in V(P_j)$, for some $j < i$. That implies u_j is not further from r than u_i, thus the path from x_i to u_j is not larger than $|P_i|$. Therefore u_i and u_j are at distance at most $2|P_i|$, which gives the desired lower bound.

We now describe an equivalent way of generating ϕ:

(i) Set $\phi(r)$ to be the upfront chosen vertex s_1 in $V(G)$.

(ii) For each $i \in [k]$, sequentially, extend the partial mapping ϕ to $V(P_i) \setminus \{x_i\}$ by taking a random walk of length $|P_i|$ which starts in $\phi(x_i)$.

(iii) Let $f_1, \dots, f_{k'}$ be an ordering of the vertices in

$$V_P = V(T) \setminus \bigcup_{i \in [k]} V(P_i)$$

such that each f_i has exactly one neighbour $f_i' \in V_P \cup \{f_1, \dots, f_{i-1}\}$. Sequentially, for $i \in [k']$, extend ϕ to f_i by taking a random neighbour of $\phi(f_i')$.

By Lemma 5.1 we have

$$\Pr[\phi(u_i) = v \mid \phi(u_1), \dots, \phi(u_{i-1})] \leq 1/m + (4/\sqrt{t})^{|P_i|} < 1/m + 1/n^3,$$

thus the conclusion of the lemma follows. ☐

Finally, we make use of the following lemma which allows us to treat unicyclic graphs as trees.

Lemma 5.3. *Let H be a connected unicyclic graph. Then there exists a tree T on the same vertex set such that $\Delta(T) \leq \Delta(H)$ and $H \subseteq T^2$.*

Proof. Let $v_1, \dots, v_k \in V(H)$ be the vertices along the cycle in H. Form a tree T by removing the edges on the cycle in H, and adding the edges on the path $v_1 v_k v_2 v_{k-1} v_3 v_{k-2} \dots v_{k'}$, where $k' = \lceil (k+1)/2 \rceil$. ☐

We are ready to prove Theorem 1.4.

Proof of Theorem 1.4. Suppose $d = a/b$, for some $a, b \in \mathbb{N}$ with $a \geq b$. Let $m = n^{1/a}$ and $t = 2^{\sqrt{\log n}}$, and let G be an $(m, t, 3\sqrt{t})$-graph on the vertex set $[m]^a$ (see Theorem 3.4). We form the graph Γ as follows: $V(\Gamma) = [m]^a$, and two vertices $\mathbf{v} = (v_1, \ldots, v_a)$ and $\mathbf{w} = (w_1, \ldots, w_a)$ are connected by an edge iff there exist at least b distinct indices $i_1, \ldots, i_b \in [a]$ such that $v_j w_j \in G^2$, for each $j \in \{i_1, \ldots, i_b\}$. Finally, take Γ^+ to be a $(2^{C\sqrt{\log n}})$-blowup of Γ, for C being a sufficiently large constant. The graph Γ^+ has

$$O\left(n^{2-b/a} \cdot 2^{2C\sqrt{\log n}}\right)$$

edges. It remains to show that Γ^+ is $\mathcal{H}_d^D(n)$-universal.

Consider some $H \in \mathcal{H}_d^D(n)$. Applying Lemma 3.1 with b, we obtain subgraphs $H_1, \ldots, H_a \subseteq H$ such that each connected component, of each H_i, is unicyclic, and each edge $e \in H$ belongs to exactly b of these subgraphs. By Lemma 5.3, for each component H_i there exists a forest T_i such that $H_i \subseteq T_i^2$ and $\Delta(T_i) \leq D$. Therefore, any homomorphism of T_i into G is also a homomorphism of H_i into G^2. By adding edges across leaves of some components in T_i, we can assume that T_i is a spanning tree on the vertex set $V(H)$. Let $r \in V(H)$ be an arbitrary vertex which will serve as the root of every tree T_i.

Form an auxiliary graph A by taking an edge between $h, h' \in V(H)$ iff they are at distance at most $16\sqrt{\log n}$ in some T_i. Then

$$\Delta(A) \leq aD^{16\sqrt{\log n}}.$$

Take $U_0 \subseteq V(H)$ to be the set of all vertices in $V(H)$ which are neighbours of r in A, together with r itself. Using Hajnal-Szemerédi theorem, partition $V(H) \setminus U_0$ into independent sets $U_1, \ldots, U_{\Delta(A)+1}$ in A.

Our goal is to find homomorphisms $\phi_i \colon T_i \hookrightarrow G$ such that, for each $i \in [a]$, $\mathbf{v} = (v_1, \ldots, v_i) \in [m]^i$ and $j \in \{1, \ldots, \Delta(A) + 1\}$, we have

$$|S_{\mathbf{v}}^j| \leq \max\{2n^{(a-i)/a}, 4\log n\}, \tag{8}$$

where

$$S_{\mathbf{v}}^j = \{h \in U_j \colon \phi_1(h) = v_1, \ldots, \phi_i(h) = v_i\}.$$

This implies that, for every $\mathbf{v} = (v_1, \ldots, v_a) \in [m]^a$, the set

$$S_{\mathbf{v}} = \{h \in V(H) \colon \phi_1(h) = v_1, \ldots, \phi_a(h) = v_a\}$$

is of size

$$|S_{\mathbf{v}}| \leq |U_0| + (\Delta(A) + 1) \cdot 4\log n < 2^{C\sqrt{\log n}},$$

thus by injectivelly mapping $S_{\mathbf{v}}$ into the blowup of $\mathbf{v}$, we obtain a copy of H in Γ^+.

Suppose that we have found $\phi_1, \ldots, \phi_{i-1}$ such that (8) holds. Let $\phi_i \colon T_i \hookrightarrow G$ be a random homomorphism generated as described at the beginning of this section. By Lemma 5.2 and standard estimates of the binomial distribution, this holds for one particular choice of $\mathbf{v} = (v_1, \ldots, v_i)$ and $j \in \{1, \ldots, \Delta(A)+1\}$ with probability at least $1 - 1/n^2$. Therefore, by the union-bound, it holds for all choices with positive probability, thus a desired homomorphism exists. ☐

5.2 Density bounded by an integer

The following lemma replaces the use of randomness in the proof of Theorem 1.4 and is the core of the proof of Theorem 1.5.

Lemma 5.4. *For every $D \in \mathbb{N}$ there exist $t \in \mathbb{N}$ and $\varepsilon > 0$ such that the following holds. Let G be an $(n, t, 3\sqrt{t})$-graph, and let T be a tree with $v(T) \leq n/(3t)$ vertices and maximum degree $\Delta(T) \leq D$. Then for any family of subsets $\{S_v \subseteq V(G)\}_{v \in V(T)}$ with $|S_v| \geq (1 - \varepsilon)n$ for each $v \in V(T)$, there exists an embedding $\phi \colon T \to G$ such that $\phi(v) \in S_v$ for every $v \in V(T)$.*

The main machinery underlying the proof of Lemma 5.4 is a result from the theory of *nonblocking networks*, due to Feldman, Friedman, and Pippenger [27, Proposition 1]. An efficient algorithmic version of this result was obtained by Aggarwal et al. [1].

Definition 5.5. Given $t, s \in \mathbb{N}$, we say that a bipartite graph $B = (V_1 \cup V_2, E)$ is (t, s)-*nonblocking* if there exists a family $\mathcal{S}$ of subsets of E, called the *safe* states, such that the following holds:

(P1) $\emptyset \in \mathcal{S}$,

(P2) if $E'' \subseteq E'$ and $E' \in \mathcal{S}$ then $E'' \in \mathcal{S}$, and

(P3) given $E' \in \mathcal{S}$ of size $|E'| < s$ and a vertex $v \in V_1$ with $\deg_{E'}(v) < t$ (that is, v is incident to less than t edges in E'), there exists an edge $e = (v, w) \in E \setminus E'$ such that $E' \cup \{e\} \in \mathcal{S}$ and w is not incident to any edge in E'.

Lemma 5.6 ([27]). *Let $B = (V_1 \cup V_2, E)$ be a bipartite graph and $a, t \in \mathbb{N}$. If*

$$|N_B(X)| \geq 2t|X|$$

for every $X \subseteq V_1$ of size $1 \leq |X| \leq 2a$, then B is (t, ta)-nonblocking.

We are now ready to prove Lemma 5.4.

Proof of Lemma 5.4. Let $h_1, \ldots, h_r$ be an ordering of the vertices of T such that for each $2 \leq i \leq r = v(T)$, h_i has exactly one neighbour within $\{h_1, \ldots, h_{i-1}\}$. For $i \in \{r, \ldots, 1\}$, iteratively, define the set $A_i \subseteq S_{h_i}$ as follows: If h_i does not have a neighbour within $\{h_{i+1}, \ldots, h_r\}$, set $A_i = S_{h_i}$; otherwise, let $R_i = \{j > i \colon h_i h_j \in T\}$ and set

$$A_i = \{v \in S_{h_i} \colon |N_G(v, A_j)| \geq (1 - \beta)t \text{ for every } j \in R_i\},$$

where $\beta > 0$ is a sufficiently small constant we will specify shortly. We show that each A_i is of size $|A_i| \geq (1 - 2\varepsilon)n$. This clearly holds for $i = r$. Suppose that it holds for $A_{i+1}, \ldots, A_r$, for some $1 \leq i \leq r - 1$. We show that it then holds for A_i as well. We can assume $R_i \neq \emptyset$, as otherwise we are immediately done. Suppose, towards a contradiction, that $|A_i| < (1 - 2\varepsilon)n$. As $|R_i| \leq D$, there exists $j \in R_i$ such that the set

$$X = \{v \in S_{h_i} \colon |N_G(v, A_j)| < (1 - \beta)t\}$$

is of size $|X| \geq \varepsilon n / D$. By the definition of X and [11, Theorem 9.2.4], we have

$$|X|(\beta - 2\varepsilon)^2 t^2 \leq \sum_{v \in X} (|N_G(v, A_j)| - (1 - 2\varepsilon)t)^2 \leq 9t \cdot 2\varepsilon n.$$

For $t > 18D/(\beta - 2\varepsilon)^2$ this gives $|X| < \varepsilon n / D$, thus a contradiction.

Before we move to the embedding of T, we need another bit of preparation. Let B be the bipartite graph on the vertex set $V(G) \times \{1, 2\}$ where (v, i) and (w, j) are connected by an edge iff $vw \in G$ and $i \neq j$. By, e.g. [10, Lemma 2.4], for sufficiently small $\beta > 0$ we have

$$|N_B(X \times \{1\})| \geq 4\beta t|X|$$

for each $X \subseteq V(G)$ of size $|X| \leq 2n/t$. Therefore, by Lemma 5.6, B is $(2\beta t, n/t)$-nonblocking.

We find distinct vertices $s_1, \ldots, s_r \in V(G)$ such that mapping h_i to h_i gives a copy of T in G. Throughout the procedure we maintain a *safe* subset $E \subseteq B$ (see Definition 5.5), which is initially empty. First choose arbitrary $s_1 \in A_1$. Then, for each $2 \leq i \leq r$, sequentially, do the following:

(i) Let $j < i$ be the unique index such that $h_j h_i \in T$.

(ii) Obtain $E \subset E' \in \mathcal{S}$ by repeatedly applying (P3) with $v = (s_j, 1)$, such that at the end we have $\deg_{E'}((s_j, 1)) = 2\beta t$.

(iii) Choose an edge $((s_j, 1), (w, 2)) \in E' \setminus E$ such that $w \in A_i \setminus \{s_1\}$. Set $s_i := w$ and $E = E \cup \{(s_i, 2)\}$.

Note that E remains a safe subset throughout the procedure. It is also evident from the description of the procedure that $\deg_E(v) \leq \Delta(T)$ for every $v \in V(G) \times \{1\}$, and $|E| < v(T)$. As $v(T) + 2\beta h < n/(2h)$, step (ii) is well-defined. It remains to show that a desired edge in (iii) always exists. Consider some step i. As $s_j \in A_j$ and $\deg_{E'}(s_j) = 2\beta t$, by the definition of A_j we have

$$|N_{E'}(s_j, A_i)| \geq \beta t,$$

thus $\deg_E(s_j) \leq D < \beta t - 1$ (by choosing t to be sufficiently large) implies the desired edge indeed exists. Note that we need to explicitly exclude s_1 in (iii) as $(s_1, 2)$ is not an endpoint of any edge in E. ☐

Note that in the previous proof, one could as well use [24, Theorem 2.8]. For our purposes, we find Lemma 5.6 to provide cleaner framework.

Proof of Theorem 1.5. Let $t \in \mathbb{N}$ and $\varepsilon > 0$ be as given by Lemma 5.4. Furthermore, let $m = Cn^{1/d}$, for sufficiently large C, and let G be an $(m, t, 3\sqrt{t})$-graph on the vertex set $[m]$ (see Theorem 3.4). We first construct Γ as follows: $V(\Gamma_n) = [m]^d$ and two vertices $\mathbf{v} = (v_1, \ldots, v_k)$ and $\mathbf{w} = (w_1, \ldots, w_k)$ are connected by an edge iff there exists $i \in [d]$ such that $v_i w_i \in G$. Note that Γ has $O(n)$ vertices and $O(n^{2-1/d})$ edges. Finally, we construct Γ^+ by adding a new set V^+ of $2dn/m$ vertices and adding all the edges incident to at least one vertex in V^+. The number of edges of Γ^+ remains $O(n^{2-1/d})$. We show that Γ^+ is $\mathcal{H}_d^D(n)$-universal.

Consider some $H \in \mathcal{H}_d^D(n)$, and let $H = H_1 \cup \ldots \cup H_d$ be a decomposition of H given by Lemma 3.1 (with $b = 1$). We clean-up H_i's as in the proof of Theorem 1.3. First form $R' \subset V(H)$ as follows: For every $i \in [d]$ and every component of H_i of size at least m, take one vertex from a cycle in that component (if such exist). This adds up to at most dn/m vertices. Next, by applying Lemma 3.2 with $F = H_i \setminus R'$ (which is now a forest) and $r = n/m$ for each $i \in [d]$, we obtain a set $R \subseteq V(H)$ of size $|R| \leq dn/m$ such that each connected component of $H_i \setminus (R \cup R')$ is of size at most m. All the vertices of $R \cup R'$ will be mapped into V^+, thus we can set $H = H \setminus (R \cup R')$ and $H_i = H_i \setminus (R \cup R')$.

We iteratively find homomorphisms $\phi_i \colon H_i \hookrightarrow G$ such that, for each $\mathbf{v} = (v_1, \ldots, v_i) \in [m]^i$, we have

$$|S_{\mathbf{v}}| \leq n^{(d-i)/d}, \tag{9}$$

where

$$S_{\mathbf{v}} = \{h \in V(H) \colon \phi_1(h) = v_1, \ldots, \phi_i(h) = v_i\}.$$

Once we have this, a homomorphism $\phi \colon H \hookrightarrow \Gamma$ given by $\phi(h) = (\phi_1(h), \ldots, \phi_k(h))$ is an injection, thus H is indeed a subgraph of Γ.

Suppose we have found homomorphisms $\phi_1, \ldots, \phi_{i-1}$, for some $i \in [d]$, such that (9) holds. To find a desired homomorphism $\phi_i \colon H_i \hookrightarrow G$, we do the following. Take trees in H_i one at a time, in an arbitrary order, and extend ϕ_i to the current tree T by taking an embedding of T into G with each $w \in T$ mapped to $S_w = V(G) \setminus R_w$, where

$$R_w = \{v \in V(G) \colon \exists S_{\mathbf{v}} = (v_1, \ldots, v_{i-1}, v) \text{ such that } |S_{\mathbf{v}}| = n^{(d-i)/d} \text{ and } w \in S_{\mathbf{v}}^-\},$$

where $S_{\mathbf{v}}$ is defined with respect to the current partial homomorphism ϕ_i and $S_{\mathbf{v}}^- = (v_1, \ldots, v_{i-1})$. By taking C to be sufficiently large we have $|R_w| < \varepsilon m$, thus we can apply Lemma 5.4 to find a desired embedding. □

Finally, we are in position to say something about the difference between the proofs of Theorem 1.4 and Theorem 1.5. In Theorem 1.5 we are able to 'cut' forests in such a way that each tree is of size $o(v(G))$, where G is an expander. This greatly helps us with planning how to embed the vertices such that the homomorphisms are as dispersed as possible: Embed one tree, revise forbidden subsets for images of some vertices, embed the next tree, revise, and so on. The fact that for each next tree we can freely choose where the root is embedded makes it possible to implement this strategy. In contrast, in the proof of Theorem 1.4 we cannot 'cut' forests in this way: We would need to remove $O(n^{1-1/a})$ vertices, resulting in $O(n^{2-1/a})$ edges in Γ^+ which is way too much. Instead we need to find a homomorphism of the whole H_i at once, and consequently we cannot do the planning one tree at a time. We resort to randomness, and drift away from the optimal bound in order to beat certain union bound. It would be interesting to improve this and resolve Conjecture 1.6.

6 Concluding remarks and open problems

- It is possible to decrease the number of vertices in the constructions in all three main theorems to $(1+\varepsilon)n$, increasing the number of edges by a factor of $c(\varepsilon)$. This can be done following the construction in Theorem 5 of [9] which is based on an appropriate concentrator (unbalanced expander).
- The proofs of all theorems provide efficient (deterministic or randomized) algorithms for embedding a given input graph H of the corresponding family in the appropriate universal graph.
- The proof of Theorem 1.3 can be easily extended to provide economical universal graphs for any family of graphs on n vertices in which the edges of each graph in the family can be partitioned into a given number d of subgraphs from a family with strongly sublinear separators. Indeed it need only be possible to break each of these subgraphs into small connected components by removing a relatively small number of vertices. The number of edges will depend on the size of the separators.

- There are several natural classes of sparse graphs that are subsets of the family of graphs with appropriate bounded density. Notable examples are graphs with a bounded acyclic chromatic number, graphs with a bounded arboricity, degenerate graphs and graphs with a bounded maximum degree. Here are some brief details.

  The acyclic chromatic number of a graph H is the minimum integer k so that there is a proper vertex coloring of H by k colors and the vertices of each cycle of H receive at least 3 distinct colors. Equivalently this means that there is a proper vertex coloring of H by k colors so that the induced subgraph on the union of any two color classes is ayclic, that is, a forest. A graph H is k-degenerate if every subgraph of it contains a vertex of degree at most k. A graph has arboricity k if its edge-set can be partitioned into k forests.

  It is not difficult to check that if the acylic chromatic number of a graph H is k, then every nonempty subset U of its vertices spans at most $(k-1)(|U|-1)$ edges. Therefore, by Edmonds' Matroid Decomposition Theorem [25] (which for the graphic matroid that is the one relevant here has been proved earlier by Nash-Williams [32]), the arboricity of H is at most $k-1$. If the aroboricity is at most $k-1$ then the density $m(H)$ is also clearly at most $k-1$. Another simple observation is that if H is d-degenerate, then its arboricity (and hence also its density) is at most d. Finally it is obvious that if the maximum degree of H is t then its density is at most $t/2$. It follows that the main theorems in this paper also provide economical constructions of universal graphs for the families of n-vertex graphs in each of these classes.

- The main open problem remaining is the assertion of Conjecture 1.1 for all rationals $d > 1$. The results proved here as well as the special cases established in [2, 7, 21] indicate that it is likely to hold in full generality.

Acknowledgements

This project began at the 2023 REU "Combinatorics and Coding Theory in the Tropics" which was run by Anant Godbole and Fernando Piñero. The authors are grateful to them for their impact, as well as to other participants of the REU, especially Mackenzie Bookamer, Sarah Capute, and Liza Ter-Saakov. We also thank Kalina Petrova for pointing out a helpful reference.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, and M. Sudan. Efficient routing in optical networks. *J. ACM*, 43(6):973–1001, 1996.

[2] P. Allen, J. Böttcher, and A. Liebenau. Universality for graphs of bounded degeneracy. *arXiv preprint arXiv:2309.05468*, 2023.

[3] N. Alon. Covering a hypergraph of subgraphs. *Discrete Math.*, 257(2-3):249–254, 2002.

[4] N. Alon. Explicit expanders of every degree and size. *Combinatorica*, 41(4):447–463, 2021.

[5] N. Alon and V. Asodi. Sparse universal graphs. *J. Comput. Appl. Math.*, 142(1):1–11, 2002.

[6] N. Alon and M. Capalbo. Sparse universal graphs for bounded-degree graphs. *Random Struct. Algorithms*, 31(2):123–133, 2007.

[7] N. Alon and M. Capalbo. Optimal universal graphs with deterministic embedding. In *Proceedings of the nineteenth annual ACM-SIAM symposium on discrete algorithms, SODA 2008, San Francisco, CA, January 20–22, 2008*, pages 373–378. New York, NY: Association for Computing Machinery (ACM); Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2008.

[8] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rodl, A. Rucinski, and E. Szemerédi. Universality and tolerance. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 14–21. IEEE, 2000.

[9] N. Alon, M. Capalbo, Y. Kohayakawa, V. Rödl, A. Ruciński, and E. Szemerédi. Near-optimum universal graphs for graphs with bounded degrees. In *International Workshop on Randomization and Approximation Techniques in Computer Science*, pages 170–180. Springer, 2001.

[10] N. Alon, E. Mossel, and R. Pemantle. Distributed corruption detection in networks. *Theory Comput.*, 16:23, 2020. Id/No 1.

[11] N. Alon and J. H. Spencer. *The probabilistic method.* Wiley-Intersci. Ser. Discrete Math. Optim. Hoboken, NJ: John Wiley & Sons, 4th edition edition, 2016.

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

[13] J. Beck and T. Fiala. "Integer-making" theorems. *Discrete Appl. Math.*, 3:1–8, 1981.

[14] S. Bhatt, F. Chung, T. Leighton, and A. Rosenberg. Optimal simulations of tree machines. In *27th Annual Symposium on Foundations of Computer Science*, pages 274–282. IEEE, 1986.

[15] S. Bhatt and T. Leighton. A framework for solving VLSI graph layout problems. *J. Comput. Syst. Sci.*, 28:300–343, 1984.

[16] B. Bollobás. Threshold functions for small subgraphs. *Math. Proc. Camb. Philos. Soc.*, 90:197–206, 1981.

[17] M. Capalbo. Small universal graphs for bounded-degree planar graphs. *Combinatorica*, 22(3):345–359, 2002.

[18] M. R. Capalbo and S. R. Kosaraju. Small universal graphs. In *Proceedings of the 31st annual ACM symposium on theory of computing, STOC 1999. Atlanta, GA, USA, May 1–4, 1999*, pages 741–749. New York, NY: ACM, Association for Computing Machinery, 1999.

[19] F. R. K. Chung. Separator theorems and their applications. Paths, flows, and VLSI-layout, Proc. Meet., Bonn/Ger. 1988, Algorithms Comb. 9, 17-34., 1990.

[20] F. R. K. Chung and R. L. Graham. On graphs which contain all small trees. *J. Comb. Theory, Ser. B*, 24:14–23, 1978.

[21] F. R. K. Chung and R. L. Graham. On universal graphs for spanning trees. *J. Lond. Math. Soc., II. Ser.*, 27:203–211, 1983.

[22] F. R. K. Chung, R. L. Graham, and N. Pippenger. On graphs which contain all small trees. II. Combinatorics, Keszthely 1976, Colloq. Math. Soc. Janos Bolyai 18, 213-223 (1978)., 1978.

[23] F. R. K. Chung, A. L. Rosenberg, and L. Snyder. Perfect storage representations for families of data structures. *SIAM J. Algebraic Discrete Methods*, 4:548–565, 1983.

[24] N. Draganić, M. Krivelevich, and R. Nenadov. Rolling backwards can move you forward: on embedding problems in sparse expanders. *Trans. Am. Math. Soc.*, 375(7):5195–5216, 2022.

[25] J. Edmonds. Minimum partition of a matroid into independent subsets. *J. Res. Natl. Bur. Stand., Sect. B*, 69:67–72, 1965.

[26] L. Esperet, G. Joret, and P. Morin. Sparse universal graphs for planarity. *J. Lond. Math. Soc., II. Ser.*, 108(4):1333–1357, 2023.

[27] P. Feldman, J. Friedman, and N. Pippenger. Wide-sense nonblocking networks. *SIAM J. Discrete Math.*, 1(2):158–173, 1988.

[28] J. Friedman and N. Pippenger. Expanding graphs contain all small trees. *Combinatorica*, 7:71–76, 1987.

[29] S. Hoory, N. Linial, and A. Wigderson. Expander graphs and their applications. *Bull. Amer. Math. Soc.*, 43(04):439???562.

[30] J. Kahn, B. Narayanan, and J. Park. The threshold for the square of a Hamilton cycle. *Proc. Am. Math. Soc.*, 149(8):3201–3208, 2021.

[31] A. Lubotzky, R. Phillips, and P. Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[32] C. St. J. A. Nash-Williams. Decomposition of finite graphs into forests. *J. London Math. Soc.*, 39 (1964), 12.

[33] R. Nenadov. *Ramsey and universality properties of random graphs.* PhD thesis, ETH Zurich, 2016.

[34] A. Ruciński and A. Vince. Strongly balanced graphs and random graphs. *J. Graph Theory*, 10(2):251–264, 1986.

[35] D. J. A. Welsh. *Matroid Theory.* L. M. S. Monographs, No. 8. Academic Press, London - New York, 1976, xi+433 pp.