Generalized Girth Problems in Graphs and Hypergraphs

Uriel Feige* Tal Wagner†

July 29, 2016

Abstract

Given a graph G and a parameter $\ell \geq 2$, the generalized ℓ-girth of G is the size of the smallest subgraph of minimum degree ℓ. A different variation asks for the smallest subgraph of average degree at least ℓ. The case $\ell = 2$ coincides with the girth of the graph (in both variations). We provide upper bounds on the generalized girth of graphs. For example, for arbitrarily small $\epsilon > 0$, we show that every n vertex graph of average degree d has a subgraph of size at most $O(n/d^{1.8-\epsilon})$ and average degree at least 4, and improve this upper bound to $O(n/d^{2-\epsilon})$ in some special cases.

*Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot 76100, Israel. Email: `uriel.feige@weizmann.ac.il`.

†Massachussetts Institute of Technology, Cambridge, MA. Email: `talw@mit.edu`.

1 Introduction

This work deals with several extremal problems in graphs and hypergraphs, that can be described in terms of generalized notions of girth. The *girth* of a graph is the size of the smallest cycle it contains (or infinity if there are none), and it poses a fundamental graph-theoretic notion that arises in many contexts. The problems we study arise as natural generalizations, and are further motivated by connections to Complexity Theory.

All graphs under discussion are simple, unweighted and undirected unless stated otherwise. Our overall goal would be find small dense subgraphs of an input graph. We fix the following notation, to be used throughout the paper: G is the input **g**raph; n is its number of **n**odes; d is its average **d**egree; ℓ is a **l**ower bound on the average or minimum degree required of the subgraph we are looking for. We call a subgraph G' of G an *ℓ-subgraph* if G' has average degree at least ℓ. The main problem we address is the following.

Question 1.1. *What is the optimal upper bound $A_\ell(n,d)$, such that any graph of size n and average degree d must contain an ℓ-subgraph of size at most $A_\ell(n,d)$?*

Observe that a smallest 2-subgraph is a smallest cycle and vice-versa, and therefore, and hence, the $\ell = 2$ case of Question 1.1 asks how large can the girth of a graph be in terms of its size and average degree. It is well known that $A_2(n,d) = \Theta(\log_{d-1} n)$, that is, any graph of size n and average degree d contains a cycle of at most such size, and there are graphs that exclude any smaller cycles. Question 1.1 may then be viewed as a generalized girth problem.

A closely related notion that has been studied is the *ℓ-girth*, defined as the size of the smallest subgraph with minimum degree at least ℓ (or infinity if there are none). Again, the $\ell = 2$ case coincides with the usual girth. Similarly to Question 1.1, the following question arises.

Question 1.2. *What is the optimal upper bound $g_\ell(n,d)$, such that any graph of size n and average degree d has ℓ-girth at most $g_\ell(n,d)$?*

The third question we consider concerns hypergraphs. Recall that a hypergraph $H(V,E)$ is *r-uniform* (abbrev. *r-hypergraph*) if each hyperedge has cardinality r. The *degree* of a node in H is the number of hyperedges that contain it. A hypeegraph $H'(V',E')$ is a *subhypergraph* of H if $V' \subset V$ and $E' \subset E$. We call H' an *ℓ-subhypergraph* of H if it has average degree at least ℓ. Question 1.1 is then naturally posed for hypergraphs as well.

Question 1.3. *What is the optimal upper bound $A_\ell(n,d,r)$, such that any r-hypergraph of size n and average degree d must contain an ℓ-subhypergraph of size at most $A_\ell(n,d,r)$?*

1.1 Results

We prove nearly tight asymptotic bounds on $A_\ell(n,d)$. For concreteness, we focus here on the case $\ell = 4$. Standard probabilistic methods yield that $A_4(n,d) = \Omega(n/d^2)$ in the general case, and $A_4(n,d) = \omega(n/d^2)$ in the high density case $d = \Omega(\sqrt{n})$, see Section 5. We conjecture this bound is tight up to polylogarithmic factors.

Conjecture 1.4. $A_4(n,d) = \tilde{O}(n/d^2)$*, where the $\tilde{O}$-notation hides polylogarithmic terms in n. More generally,* $A_\ell(n,d) = \tilde{O}(n/d^{\ell/(\ell-2)})$.

We establish two approximate variants of this conjecture. The first is making a small compromise on the average degree of the target subgraph, from 4 to $4-\epsilon$ where $\epsilon > 0$ is an arbitrarily small constant. The following result states that $A_{4-\epsilon}(n,d) = O(n/d^2)$.

Theorem 1.5. *Let $\epsilon > 0$. Every graph of size n with average degree d contains a $(4-\epsilon)$-subgraph of size at most $O(n/d^2)$. The constant hidden in the O-notation depends on ϵ.*

The second approximate variant of Conjecture 1.4 is a compromise on the size of the target subgraph, from $O(n/d^2)$ to, ideally, $O(n/d^{2-\epsilon})$. This turns out more challenging to analyze, and constitutes the main part of our work. The following is our main result.

Theorem 1.6. *$A_4(n,d) = O(n/d^{1.8-\epsilon})$ for every $\epsilon > 0$. The constant hidden in the O-notation depends on ϵ.*

More generally, we may approach $O(n/d^{2-\epsilon})$ by requiring a tradeoff between ϵ and d.

Theorem 1.7. *Let $\epsilon > 0$, and $t \geq 2$ be an integer. Every graph of size n and average degree $d = O(n^{1/(t-\delta(t)-\epsilon)})$ contains a 4-subgraph of size $O(n/d^{2-\delta(t)-\epsilon})$, where $\delta(t) = 1/(3 \cdot 2^{t-2} + 2)$.*

Note that Theorem 1.6 is the $t = 2$ case of Theorem 1.7, which yields $\delta(t) = 0.2$. This special case applies to all graphs, since if $d = O(n^{1/(1.8-\epsilon)})$ then the hypothesis of Theorem 1.7 is satisfied, and otherwise the bound stated in Theorem 1.6 is trivial.

For general ℓ, we obtain a lower bound $A_\ell(n,d) = \Omega(n/d^{\frac{\ell}{\ell-2}})$ and (for even ℓ) an upper bound $A_\ell(n,d) = O(n/d^{\frac{\ell}{\ell-2}-O(\frac{1}{\ell})^3-\epsilon})$. A detailed proof of the upper bound appears in Appendix B.

The above results directly imply bounds on the ℓ-girth, by nothing that $A_\ell(n,d) \leq g_\ell(n,d) \leq A_{2\ell-2}(n,d)$ for every ℓ (see Lemma 2.4 for the upper bound). In Section 6 we extend our results to 3-hypergraphs as well.

1.2 Motivation and Related Work

Girth. The girth problem, which is to determine $A_2(n,d)$ or (equivalently) $g_2(n,d)$, has been intensively studied. The asymptotics $S_1(n,d) = \Theta(\log_{d-1} n)$ have well-known and short proofs: The upper bound is by a breadth-first search on the graph, and the lower bound by the probabilistic deletion method, see eg. [Bol04, page 104, Theorems 1.1 and 1.2]. A major problem is to identify the optimal leading constant c of the $\log_{d-1} n$ term, which to date remains open, despite concentrated efforts. The best upper bound is $c \leq 2$ due to Alon, Hoory and Linial [AHL02], and the best lower bound is $c \geq 4/3$, originally achieved by Margulis [Mar82, Mar88] and (independently) by Lubotzky, Phillips and Sarnak [LPS88], and subsequently reproved and extended in a sequence of works [Imr84, BB90, Mor94, LU95, Dah13]. Notablt, their proofs are by explicit constructions of extremal graphs, known as "Ramanujan graphs", that have many important applications.

ℓ-Girth. The bound $g_\ell(n,d) \leq 2\ell n/d$ is stated as a secondary result in the work of Erdős et al. [EFRS90]. Apart from this, research on $g_\ell(n,d)$ has primarily focused on borderline low densities. The precise d that guarantees $g_\ell(n,d) < \infty$ is known (see [Kez91, Lemma 23]), and graphs with that density and ℓ-girth exactly n were studied by Kézdy [Kez91, Chapter 5]. Upper bounds in the presence of one extra edge are discussed in [EFRS90], and some properties of graphs one edge short are studied in [EFGS88] and [BB89].

From a complexity-theoretic point of view, the computational problem of determining the ℓ-girth of an input graph was considered by Amini et al. [ASS12, APP+12], who have shown various hardness of approximation results for $\ell \geq 3$. (The $\ell = 2$ case is polynomial-time solvable, by performing a BFS starting at each vertex.) Computational aspects related to $A_\ell(n,d)$ are discussed in the subsequent work [FKW16].

Hypergraphs. To motivate the discussion of hypergraphs, we present yet another notion of generalized girth: An *even cover* is a subset of hyperedges that cover each vertex is covered an even (possibly zero) number of times. In graphs ($r = 2$), a smallest even cover coincides with (the edge set of) a smallest cycle, so again we are dealing with a girth-type problem.

Question 1.8. *What is the optimal upper bound $EC(n, d, r)$, such that any r-uniform hypergraph of size n and average degree d must contain an even cover of size at most $EC(n, d, r)$?*

Feige posed the conjecture $EC(n, d, r) = \tilde{O}(n/d^{2/(r-2)})$ [Fei08, Conjecture 1.2], which meets a known lower bound exhibited by random hypergraphs. The conjecture was priorly shown by Feige, Kim and Ofek [FKO06] to hold for random hypergraphs. For arbitrary hypergraphs, some progress has been made on high densities. Naor and Verstraëte [NV08] proved the highest density case for all even r, namely $EC(n, n^{(r-2)/2}, r) = O(\log n)$ (see [Fei08, Proposition 2.2]). The odd r case appears more difficult. Focusing on $r = 3$, [Fei08] and [NV08] obtain logarithmic bounds on $EC(n, d, 3)$ for various densities in the regime $\tilde{O}(\sqrt{n})$. For yet higher densities, Dellamonica et al. [DHL+12] prove[1] that $EC(n, n^{1/2+\epsilon}, 3) \le \lceil 4/\epsilon \rceil$, for any $\epsilon > 0$ and $n \ge \lceil (5/\epsilon)^{1/(2\epsilon)} \rceil$.

The connection between Question 1.3 and Question 1.8 is twofold. On one hand, an even cover in a r-hypergraph forms a 2-subhypergraph, since each hyperedge covers r vertices, and each covered vertex is covered at least twice. Question 1.3 with $\ell = 2$ is therefore suggested in [Fei08] as an intermediate step towards Question 1.8. On the other hand, a subhypergraph with more hyperedges than vertices (i.e. an ℓ-subhypergraph with $\ell > r$) must contain an even cover: Viewing each hyperedge as an indicator vector for its vertices, the hyperedges contain a linear dependency over $\mathbb{F}_2$, and that subset forms an even cover.[2] Therefore, upper bounds on Question 1.3 in the $\ell > r$ regime apply directly to Question 1.8. For example, Alon and Feige [AF09, Lemma 3.3] prove that every 3-hypergraph of size n and density d contains a subhypergraph of size $O(n \log n/d)$ with strictly more hyperedges than vertices, and hence an even cover of that size. For general densities this is currently the best upper bound on $EC(n, d, 3)$.

The study of even covers, and in particular of Question 1.8, is motivated in part by applications to Theory of Computation. In [NV08] it arises in the context of *sparse parity-check matrices*, a key notion in coding theory. The motivation in [Fei08] is the design of *refutation algorithms* for random 3CNF formulas, which is an average-case variant of the fundamental 3-satisfiability problem (3SAT). This task poses a main challenge in complexity theory, and further has known implications to hardness of approximation [Fei02]. The approach centered at even covers has so far been the most successful: Apart from achieving the best bounds for random refutation [FO07], it has also been applied to non-deterministic refutation in parameters regimes for which deterministic algorithms are unknown [FKO06], and to refutation of semi-random formulas, in which the assignment of variables to clauses is adversarial and the randomness is confined to variable polarities [Fei07]. This latter work is especially notable in our context, as it relies directly on upper bounds on $EC(n, d, 3)$.

Turán-type problems. Questions 1.1 to 1.3 and 1.8 fit into a broader theme in Extremal Combinatorics known as *Turán-type* problems, where the goal is to determine the maximum number of edges that a graph (or hypergraph) may have while avoiding certain subgraphs. Typically, however, these problems are concerned with a single forbidden subgraph (the original Turán Theorem avoids a clique) or a rather restricted forbidden family, whereas in our case the "forbidden family" is very

[1] Their result assumes that any two hyperedges may intersect on at most one vertex, but by [Fei08, Lemma 2.4] removing this assumption has only a negligible effect on the density.

[2] In the even r case this holds already for $\ell = r$, i.e. at least as many hyperedges as vertices. This is because each hyperedge e satisfies $\mathbf{1}^T e = 0$ over $\mathbb{F}_2$ (where $\mathbf{1}$ is the all-1's vector), so the hyperedges reside in a subspace with co-dimension 1.

general (say in Question 1.1, all ℓ-subgraphs up to some size). Yet, a certain well studied line of Turán-type problems on hypergraphs relates to Question 1.3 for very small sized ℓ-subgraphs. Brown, Erdős and Sós [BES73b, BES73a] initiated the study of the asymptotic growth of the maximum number of hyperedges in a r-hypergraph that excludes all subgraphs with e hyperedges and v vertices, for small constants v, e. A celebrated result of Ruzsa and Szemerédi [RS76] resolved the $r = 3, e = 6, v = 3$ case, that became known as *the* $(6, 3)$*-problem*, settling the answer at $o(n^2)$. Phrased in terms of Question 1.3, it states that for any $\epsilon > 0$ and sufficiently large n, density ϵn in 3-hypergraphs forces a $r/2$-subgraph of size 6. This was extended by Erdős, Frankl and Rödl [EFR86] to $r/(r-1)$-subgraphs of size $3r - 3$ for any uniformity r. For some more recent bounds of this flavour, see Alon and Shapira [AS06].

2 Preliminaries

This section records some simple lemmas that will be needed for our analysis. All proofs are deferred to Appendix C.

2.1 Concentration Lemmas

Lemma 2.1 (reverse Markov inequality)**.** *Let X be distributed over $0, \ldots, n$ with $\mathbb{E}[X] = \mu$. Then* $\Pr[X \geq \frac{1}{2}\mu] \geq \frac{\mu}{2n}$.

Lemma 2.2. *Let X be a non-negative random variable. There is an integer $t \geq 1$ such that* $\Pr[X \geq \frac{2}{9} \cdot 2^t \cdot \mathbb{E}X] \geq (2^t t^2)^{-1}$. *Furthermore if M is an upper bound on X, then $t \leq \log(\frac{9}{2}M/\mathbb{E}[X])$.*

Lemma 2.3 (a Chernoff bound)**.** *If X is binomially distributed, then* $\Pr[X < 2\mathbb{E}[X]] \geq 1 - (0.25e)^{\mathbb{E}[X]}$.

2.2 Minimum Degree Guarantees

Lemma 2.4. *A graph with average degree d has a subgraph with minimum degree $\geq \lfloor \frac{1}{2}d + 1 \rfloor$.*

Lemma 2.5 (half-matching)**.** *Let $G(V, U; E)$ be a bipartite graph such that each $v \in V$ has degree d, and each $u \in U$ has degree at most 2. There is a subset of edges E' such that in $G'(V, U; E')$, each $v \in V$ has degree at least $\lfloor \frac{1}{2}d \rfloor$, and each $u \in U$ has degree at most 1.*

Lemma 2.6 (edge orientation)**.** *Let G be a graph with minimum degree d. Its edges can be oriented such that each vertex has at least $\lfloor \frac{1}{2}d \rfloor$ edges oriented towards it.*

Lemma 2.7. *Let $G(V, U; E)$ be a bipartite graph such that side V has average degree d and maximum degree D. There is a subset $V' \subset V$ with size $|V'| \geq \frac{d}{2D}|V|$ such that each $v \in V'$ has degree $\geq \frac{1}{2}d$.*

3 Degree Compromise

We start by considering a relaxation that allows a small compromise on the target average degree, from ℓ to $\ell - \epsilon$.

Theorem 3.1 (degree compromise)**.** *For every integer $\ell \geq 1$ and $\epsilon > 0$, $A_{\ell-\epsilon} = O(n \cdot d^{-(\ell-2)/\ell})$. The constant hidden in the O-notation depends on ℓ and ϵ.*

The proof is fairly straight-forward, and we use it to lay the foundations for the more involved proofs of Section 4. To keep the presentation simple, we focus on the following special case that captures the main ideas. The full proof of Theorem 3.1 is deferred to Appendix A. In what follows, we use $o_x(1)$ to denote a term that tends to 0 as a variable x tends to infinity, and $o_x(f) = o_x(1) \cdot f$.

Theorem 3.2. *Let $\epsilon > 0$ be arbitrary. There is a constant $C = C(\epsilon)$ such that every graph $G(V,E)$ with n vertices and average degree d satisfying $4 - \epsilon \leq d = o_n(\sqrt{n/\log n})$, contains a subgraph of size at most $C \cdot n/d^2$ with average degree $\geq 4 - \epsilon$.*

Proof. By Lemma 2.4, and a loss in the constant C, we may assume G has minimum degree d. Moreover it is enough to prove the statement for all sufficiently large values of d, and smaller values can then be handled by a proper choice of C.

Let α be a large constant that will be determined later. Sample a random subset $A \subset V$ by including each vertex in A with independent probability $p = \alpha/d^2$. We refer to vertices in A as *marked*. Note that $|A|$ is binomially distributed with parameters n, p, and $\mathbb{E}|A| = \alpha n/d^2$.

For each vertex v we fix an arbitrary subset $N(v)$ of exactly d of its neighbours. Define B to be the random subset of vertices v that are not marked, and have exactly two marked neighbours in $N(v)$. We then have,

$$\Pr[v \in B] = (1-p) \cdot \binom{d}{2} p^2 (1-p)^{d-2} = \frac{1}{2}(1 - o_d(1)) p^2 d^2 = \frac{1 - o_d(1)}{2} \alpha^2 / d^2,$$

where the middle equality holds since $(1-p)^{d-1} = (1 - \frac{1}{d^2})^{d-1} = 1 - o_d(1)$. Hence by linearity of expectation, $\mathbb{E}|B| = \frac{1-o_d(1)}{2}\alpha^2 n/d^2$.

By the Chernoff bound (Lemma 2.3) applied to $|A|$, we get:

$$\Pr\left[|A| < 2\mathbb{E}|A|\right] \geq 1 - (0.25e)^{\alpha n/d^2} > 1 - o_n(d^{-2}),$$

where the final inequality is since $d = o_n(\sqrt{n/\log n})$ and by choosing α sufficiently large. On the other hand, by Lemma 2.1, $|B|$ attains half its expected value with probability at least $\frac{\mathbb{E}|B|}{2n} = \frac{1-o_d(1)}{4}\alpha^2/d^2 = \Omega(d^{-2})$. Summing the bounds yields:

$$\Pr\left[|A| \leq 2\mathbb{E}|A|\right] + \Pr\left[|B| \geq \frac{1}{2}\mathbb{E}|B|\right] \geq 1 - o_n(d^{-2}) + \Omega(d^{-2}).$$

The bound on the right-hand side is strictly more than 1 for sufficiently large n, and hence there is positive probability that both of the events $|A| \leq 2\mathbb{E}[|A|]$ and $|B| \geq \frac{1}{2}\mathbb{E}[|B|]$ occur. We fix this event, and arbitrarily remove vertices from B until $|B| = \frac{1}{2}\mathbb{E}[|B|]$. The following bounds now hold:

$$|A| + |B| \leq 2\mathbb{E}|A| + \frac{1}{2}\mathbb{E}|B| = \left(2\alpha + \frac{1 - o_d(1)}{4}\alpha^2\right) \cdot \frac{n}{d^2}, \tag{3.1}$$

$$\frac{|B|}{|A|} \geq \frac{\frac{1}{2}\mathbb{E}|B|}{2\mathbb{E}|A|} = \frac{1 - o_d(1)}{8} \cdot \alpha. \tag{3.2}$$

We take our target subgraph H to be that induced by $A \cup B$. By eq. (3.1), its size is bounded by $C \cdot n/d^2$ for, say, $C = \frac{1}{2}\alpha^2$. To bound its average degree, note that each vertex in B is incident to two edges connecting it to A, and since A and B are disjoint (recall that vertices in B are not marked), each such edge has a unique end in B. Hence we count at least $2|B|$ different edges in H, and find that its average degree is:

$$\text{avgdeg}(H) \geq 2 \cdot \frac{2|B|}{|A| + |B|} = 4 - \frac{4}{1 + \frac{|B|}{|A|}} \geq 4 - \frac{4}{1 + \frac{1 - o_d(1)}{8} \cdot \alpha}$$

using eq. (3.2) for the final inequality. The bound on the right-hand side is guaranteed to be at least $4 - \epsilon$ as long as we pick $\alpha > 33\epsilon^{-1}$, and the proof is complete. $\square$

4 Size Compromise

In this section we prove our main results, Theorems 1.6 and 1.7. We begin with an overview of our approach.

Our starting point is the proof of Theorem 3.2. The outline was to mark each vertex with independent probability p, and then use structures induced by the marking as building-blocks in our target subgraph. Our building-block was a vertex with two marked neighbours, as such a vertex contributes two edges to the target subgraph, thus "paying for itself" towards the end of attaining average degree 4. The marked vertices were not paid for, so the overall average degree was only $4-\epsilon$.

To avoid this loss, we require some vertices to contribute *three* edges to the target subgraph, and use the extra edges to pay for the marked vertices. As a first step we may repeat the proof of Theorem 3.2, but this time take a building-block to be a vertex with *three* marked neighbours instead of two. Again we denote the set of marked vertices as A and the set of building-blocks as B. Each vertex is included in B with probability (roughly) d^3p^3, for choosing 3 out of d neighbours and marking each of them. We get $\mathbb{E}|A| = np$ and $\mathbb{E}|B| = nd^3p^3$, and requiring the sizes to be similar (up to a constant), we obtain $p \sim 1/d^{1.5}$. Hence we get a subgraph of size $O(n/d^{1.5})$ and average degree $6-\epsilon$, which is greater than 4. This proves $A_4(n,d) = O(n/d^{1.5})$.

Now let us try to extend this outline to prove $A_4(n,d) = O(n/d^{1.8-\epsilon})$, as guaranteed by Theorem 1.6. This time our building-block would be a vertex with three neighbours, each of them having two marked neighbours. See Figure 1 for illustration. We expect each vertex to be in B (the root of such a structure) with probability $\Theta(d^9p^6)$, for choosing 9 edges (each out of d possibilities) and marking 6 vertices. Therefore $\mathbb{E}|A| = np$ and $\mathbb{E}|B| = nd^9p^6$, and equating the sizes yields $p \sim 1/d^{1.8}$. We would thus obtain a subgraph of size $O(n/d^{1.8})$ and average degree $4\frac{2}{9} - \epsilon$, which is greater than 4. Albeit, this proof outline contains a gap: The argument explaining why each vertex is included in B with probability d^9p^6 inherently assumes that the radius-2 neighbourhood of each vertex is a tree. For example, if the radius-2 neighbourhood of a vertex contains $k = o(d^{3/2})$ vertices, the probability that it has six marked vertices is upper bounded by $k^6p^6 = o(d^9p^6)$. A valid argument should take into account the specific topology of the host graph and the probabilistic dependencies it may induce. In the proof of Theorem 3.2 this difficulty did not arise, since the radius-1 neighbourhood of any vertex trivially forms a tree.

Our main lemma will show that if we set $o = 1/d^{1.8-\epsilon}$, then each vertex is the root of the building-block depicted in Figure 1 with probability at least p. This falls short of d^9p^6 due to introducing ϵ, but is nonetheless sufficient for the above proof outline to carry through, since we get $\mathbb{E}|A| = \mathbb{E}|B| = np$ (up to constants) and thus find a 4-subgraph of size $O(n/d^{1.8-\epsilon})$.

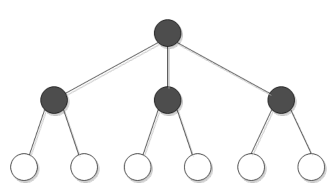

Figure 1: A "building block" with 6 leaves and 9 edges. White vertices are marked and shaded vertices are non-marked.

4.1 Definitions

Definition 4.1 (graph with random subset model)**.** *Let $G(V,E)$ be a graph. For $p \in [0,1]$, we define $G(V,E,p)$ to be a random model in which each vertex in V is marked with independent* *probability* p.

Definition 4.2 (root)**.** *In* $G(V, E, p)$*, a vertex is a* 1-root *if it is marked. For an integer* $k > 1$*, a vertex is a* k-root *if, inductively, it has two distinct neighbours which are a* k_1*-root and a* k_2 *root, and* $k_1 + k_2 = k$*.*

Definition 4.3 (excited vertex)**.** *In* $G(V, E, p)$*, for an integer* $k \geq 1$*, a vertex is* k-excited *if it has a* k*-root neighbour.*

Remark. A vertex may be a k-root concurrently for several values of k, or for the same k due to several combinations of neighbours. The same goes for being k-excited. Moreover, a vertex may be a k-root and k-excited at the same time.

Definition 4.4 (tree-like graph)**.** *For an integer* $k \geq 1$*,* $p \in [0, 1]$ *and* $\gamma > 0$*, a graph* $G(V, E)$ *with minimum degree* d *is* (k, p, γ)-tree-like *if in* $G(V, E, p)$*, each vertex is* k*-excited with probability at least* $\min\{\gamma \cdot p^k d^{2k-1}, 0.99\}$*.*

The meaningful case that should be in mind is when d is arbitrarily large, $p \ll d^{-1}$, and k and γ are constants. In this case, the following proposition is important to note. It motivates the terminology of Definition 4.4.

Claim 4.5. *Let* $t \geq 2$ *be a constant integer and* $k = 2^{t-2}$*. Let* $T(V, E)$ *be a* d*-ary tree with* d *arbitrarily large, rooted by* $v \in V$ *and with all leaves in level* t*. (The root is considered to be in level* 1*.) Then in* $T(V, E, p)$ *with* $p = o(d^{-1})$*, the probability that* v *is* k*-excited is* $\min\{\Omega(p^k d^{2k-1}), 0.99\}$*.*

Proof. We only sketch the proof as it is very clear. Let $u_1, \ldots, u_d$ be the children of v, and let A_i denote the event that u_i is a k-root. We show by induction on t that $\Pr[A_i] = \Omega(p^k d^{2k-2})$: In the base case $t = 2$ we get $k = 1$, so we need to show that u_i is a 1-root w.p. p, which holds by definition. For $t > 2$, by induction, each child of u_i is a $(\frac{1}{2}k)$-root w.p. $q = \Omega(p^{k/2} d^{k-2})$, and these events are independent since T is a tree. Hence the number of $(\frac{1}{2}k)$-roots among the children of u_i is binomially distributed with parameters (d, q), and the probability that two of them are $(\frac{1}{2}k)$-roots is $\Omega(d^2 q^2) = \Omega(p^k d^{2k-2})$. This renders u_i a k-root, so the proof by induction is complete.

The events $A_1, \ldots, A_d$ are independent since T is a tree, so the number of k-roots among $u_1, \ldots, u_d$ is binomially distributed with parameters (d, p') for $p' = \Omega(p^k d^{2k-2})$. If this number is at least one then v is k-excited, and this occurs w.p. $\Omega(dp') = \Omega(p^k d^{2k-1})$. □

We see that in the setting described above, a graph is tree-like if for the purpose of counting excited vertices in $G(V, E, p)$, it behaves roughly as if the k-neighbourhood of each vertex was a tree. We now get to our main definition.

Definition 4.6 (good graph)**.** *For integer* $k \geq 1$*,* $p \in [0, 1]$ *and* $\gamma > 0$*, a graph* G *is* (k, p, γ)-good *if for each vertex* v *in* G*, the graph* $G - \{v\}$ *is* (k, p, γ)*-tree-like.*

Put simply, a good graph is a "robust" tree-like graph. We will reduce proving upper bounds on $A_4(n, d)$ to proving that graphs are good. As we will be interested in very general classes of graphs, our "goodness" proofs will only use global properties (such as minimum degree, girth, etc.) and not rely on any specific topologies. Hence we will actually be proving that the graphs under discussion are tree-like, and the robustness will follow without additional effort.

4.2 Reduction

Following is our main technical theorem.

Theorem 4.7 (reduction theorem)**.** *Let $k \geq 1$ be an integer and $\epsilon > 1/(3k+2)$. There is a constant $C = C(k, \epsilon)$ such that for sufficiently large d the following holds: If G is a graph with $n \geq \Omega(d^2)$ vertices and minimum degree d, and G is (k, p, γ)-good with $p = 1/d^{2-\epsilon}$ and $\gamma = \Omega(1/\text{polylog}\,(d))$, then G contains a subgraph of average degree 4 and size at most $C \cdot n/d^{2-\epsilon}$. (The power of the* $\text{polylog}\,(d)$ *term in the bound on γ may depend on k.)*

The proof is in Section 4.3. In the remainder of this section we demonstrate the use of this result. In particular, we can already prove two notable special cases of Theorem 1.7 (which in its full form requires more effort). One is Theorem 1.6.

Proof of Theorem 1.6. As in the proof of Theorem 3.2, it is sufficient to prove the statement for d sufficiently large, and we may assume w.l.o.g that G has minimum degree d. Let $\epsilon > 0.2$ and set $p = d^{-(2-\epsilon)}$. In $G(V, E, p)$, the probability for each vertex to be 1-excited (that is, to have a marked neighbour) is $\geq (1 - o(1))dp$, hence G is $(1, p, 1 - o(1))$-tree-like. This holds even if we remove any single vertex from G, hence it is $(1, p, 1 - o(1))$-good. Noting that $\epsilon > 1/(3k + 2)$ for $k = 1$, the claim follows from Theorem 4.7. □

The other special case is an improved bound for graphs with large girth.

Theorem 4.8. *Let $\epsilon > 0$, and $t \geq 2$ be an integer. There is a constant $C(t, \epsilon)$ such that every graph $G(V, E)$ with n vertices, average degree d (satisfying $4 \leq d \leq O(\sqrt{n})$), and girth $\geq 2t - 1$, contains a subgraph of average degree 4 and size at most $C(\epsilon) \cdot n/d^{2-\delta(t)-\epsilon}$, where $\delta(t) = 1/(3 \cdot 2^{t-2} + 2)$.*

Proof. Again we prove for sufficiently large d and assume d is the minimum degree. Let $k = 2^{t-2}$. Since G has no cycles of length $2t - 1$ or less, the radius-$(t - 1)$ neighbourhood of each vertex is a tree. In other words, each vertex v is the root of a d-ary tree with all leaves in level t. Hence by Claim 4.5, v is k-excited with probability $\Omega(p^k d^{2k-1})$ for $p = d^{-(2-\epsilon)}$. This means G is $(k, p, \Omega(1))$-tree-like, and the argument remains intact (up to a small change of constants) even if we remove any single vertex from G, so it is $(k, p, \Omega(1))$-good.

The theorem now follows by applying from Theorem 4.7 with $\epsilon' = \delta(t) + \epsilon$. By the choice of $\delta(t)$, this ensures $\epsilon' > 1/(3k + 2)$, so the condition of Theorem 4.7 is satisfied. □

Theorem 4.8 also follows from Theorem 1.7, since by [AHL02], a large girth implies an upper bound on the average degree.

4.3 Proof of the Reduction

In this section we prove Theorem 4.7. Let us first present an overall description of the proof. As explained earlier, we use roots in $G(V, E, p)$ to identify small structures with two edges per non-marked vertex, and our plan is to use them as building-blocks to construct a 4-subgraph. These structures are referred to as *arrangements* and are discussed in Section 4.3.1.

The key to the proof is to combine arrangements in a way that earns an additional edge, that would eventually be used to pay for the marked vertices. This is done in the main lemma, proven in Section 4.3.2. It states that with sufficiently high probability, we can identify structures in which one vertex contributes *three* edges, and not just two. That extra edge makes all the difference in achieving average degree 4.

In Section 4.3.3 we prove Theorem 4.7, in a way quite similar to the proof of Theorem 3.2. However, one additional point requires attention: The structures we use to build our target subgraph are rather large and may overlap in edges, causing us to pay with the same edge for two different vertices (its two endpoints). In Section 3 this issue did not arise, as each edge in a structure had only one non-marked endpoint and was counted to pay for it. To solve this now, we equip each arrangement with an orientation of its edges and decide that each edge pays for its destination vertex.

4.3.1 Arrangements in $G(V, E, p)$

We now formalize the notion explained above, of structures that have two edges per each non-marked vertex.

Definition 4.9 (arrangement; arrangeable subgraph). *Given a fixed marking of vertices sampled from $G(V, E, p)$, an* arrangement *is a pair (G', O) of a subgraph $G'(V', E')$ of G, and an orientation O of the edges in E' such that each non-marked vertex in V' has at least two edges oriented towards it.*

G' is an arrangeable subgraph *if there is an orientation O of its edges such that (G', O) is an arrangement.*

Lemma 4.10 (union of arrangements). *Given a fixed marking of vertices sampled from $G(V, E, p)$, let H_1 and H_2 be arrangeable subgraphs in G. The union subgraph $H = H_1 \cup H_2$ is arrangeable.*

Proof. Let O_1, O_2 be orientations of the edges in H_1, H_2 respectively, such that (H_1, O_1) and (H_2, O_2) are arrangements. We orient the edges in H as follows: For an edge e, if it is present in H_1 then we orient it by O_1. Otherwise e is present in H_2, and then we orient it by O_2. Call the resulting orientation O.

To see that (H, O) is an arrangement, let $v \in H$ be any vertex. If $v \in H_1$ then some two edges $e, e' \in H_1$ are oriented towards v in O_1, and by the above they are oriented towards v also in O. Otherwise we have $v \in H_2$, and then some two edges $e, e' \in H_2$ are oriented towards v in O_2. Neither of them can be in H_1 (as $v \notin H_1$), so they are oriented in O as they are in O_2, towards v. Thus, every vertex in H has two edges oriented towards it by O. ☐

The next lemma states that k-roots are part of arrangements with small (constant) size.

Proposition 4.11. *Let $k \geq 1$ be an integer. Given a fixed marking of vertices sampled from $G(V, E, p)$, if v is a k-root, then v is part of an arrangeable subgraph with at most $k-1$ non-marked vertices.*

Proof. By induction on k: The base case $k = 1$ is trivial, since a 1-root is just a marked vertex, which by itself constitutes an arrangement with 0 non-marked vertices.

For $k > 1$, by definition v has two distinct neighbours u_1 and u_2 which are a k_1-root and a k_2-roots respectively, such that $k_1 + k_2 = k$. By induction, u_1 is part of an arrangeable subgraph H_1 with at most $k_1 - 1$ non-marked vertices, and u_2 is part of an arrangeable subgraph H_2 with at most $k_2 - 1$ non-marked vertices.

Let H be the union of H_1 and H_2; it is an arrangeable subgraph by Lemma 4.10, and it has at most $k-2$ non-marked vertices. Let O be an orientation such that (H, O) is an arrangement. If $v \in H$, then we are done. Otherwise, we add v to H together with the two edges connecting v to u_1 and u_2, orienting both towards v. (Note that $v \notin H$ implies that the edges vu_1 and vu_2 are not already orientated by O, so we can orient them as we wish.) We have added a single non-marked vertex to H, so it now has at most $k-1$ non-marked vertices, as required. ☐

4.3.2 Main Lemma: That Extra Edge

Lemma 4.12 (main)**.** *Let $k \geq 1$ be an integer, $M > 0$ an arbitrary constant, and $\epsilon > 1/(3k+2)$. For sufficiently large d, the following holds: If $G(V, E)$ is a graph with minimum degree d that is (k, p, γ)-good for $p = d^{-(2-\epsilon)}$ and $\gamma = \Omega(1/\text{polylog}(d))$, then in $G(V, E, 2p)$, each vertex $v \in V$ has probability $\geq M \cdot p$ to be part of an arrangement H_v with the following properties:*

- *v has three edges oriented towards it in H_v.*
- *H_v has at most $3k+1$ non-marked vertices.*

(The power of the polylog (d) *term in the bound on γ may depend on k.)*

Proof. We assume d is sufficiently large wherever needed, possibly without explicitly stating so. Moreover in order to simplify the presentation, we do not attempt to optimize the constants involved in the proof.

Fix a vertex $v \in V$ for the remainder of the proof. We begin by removing v from G, along with all of its incident edges. Since G is (k, p, γ)-good, the resulting graph $G - \{v\}$ is (k, p, γ)-tree-like. To ease notation, we will refer to $G - \{v\}$ as $G(V, E)$, keeping in mind that it no longer contains v.

To sample from $G(V, E, 2p)$, we mark the vertices of G in two independent phases, each with probability p. That is, a vertex that was not marked in the first phase has probability p to be marked in the second phase, hence total probability of $2p - p^2$ to be marked at all. Since $p \ll 1$, this is equivalent to $G(V, E, 2p)$ up to a small variation of constants. Importantly, note that the first phase is a sample from $G(V, E, p)$.

Let $N(v)$ denote the set of neighbours of v. We restrict our attention to exactly d neighbours of v (arbitrarily chosen), so $|N(v)| = d$. Our current (and main) goal is to show that with probability $M \cdot p$, three vertices in $N(v)$ are $(k+1)$-roots after the second phase.

Preliminaries. We begin by orienting the edges incident to vertices in $N(v)$. In principle we would like all edges incident to $u \in N(v)$ to point towards u, but there might be edges with two ends in $N(v)$. Bypassing this issue is a mere technicality: Since each vertex in $N(v)$ has at least d incident edges,[3] we can use Lemma 2.6 to find an orientation by which each vertex has $\frac{1}{2}d$ edges oriented towards it. Fix this orientation henceforth.

Consider an arbitrary vertex $u \in N(v)$, and let $e_i = w_i u$ be the edges oriented towards it, for $i = 1, \dots, \frac{1}{2}d$. Suppose u is k-excited after the first phase, which means it has a k-root neighbour w. In this case, we call each edge $e_i = uw_i$ with $w_i \neq w$ an *excited* edge. So if u is k-excited after the first phase, it is incident to $\frac{1}{2}d$ excited edges[4] that are oriented towards it. (There may be additional excited edges incident to u but oriented away from it, towards other vertices in $N(v)$.)

The source vertex w_i of an excited edge e_i will be called a *touched* vertex. Observe that if w_i is marked in the second phase, then u becomes a $(k+1)$-root. This is illustrated in fig. 2. Lastly, we say that a touched vertex is *light* if it is the source of one or two excited edges, and *heavy* if it is the source of at least three excited edges.

Counting excited edges. Let X denote the subset of k-excited vertices in $N(v)$ after the first phase. Since G is (k, p, γ)-tree-like, each vertex has probability $q = \gamma p^k d^{2k-1}$ to be k-excited, and since $|N(v)| = d$, we get $\mathbb{E}|X| = dq$.

[3] In fact $d - 1$ edges, having removed v from G, but we suppress the -1 for simplicity.

[4] In fact u is incident to either $\frac{1}{2}d$ (if w is one of the w_i's) or $\frac{1}{2}d - 1$ (otherwise) excited edges, but again we suppress the -1.

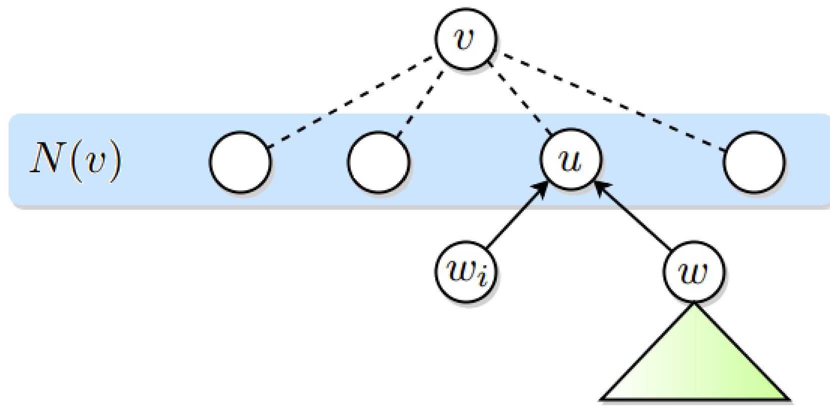

Figure 2: After the first phase, w is a k-root, rendering u a k-excited vertex in $N(v)$. The edge $w_i u$ is thus *excited*, and w_i is a *touched* vertex (which is either heavy or light, depending on how many additional excited edge are oriented away from it). If w_i is marked in the second phase, u would become a $(k+1)$-root.

Let Y be the set of excited edges. As each vertex in X renders $\frac{1}{2}d$ edges excited, we have

$$|Y| = \frac{1}{2}d|X| \tag{4.1}$$

and therefore,

$$\mathbb{E}|Y| = \frac{1}{2}d\mathbb{E}|X| = \frac{1}{2}qd^2 \tag{4.2}$$

Applying Lemma 2.2 to $|Y|$, we get:

$$\Pr[|Y| \geq \frac{1}{9} \cdot 2^t qd^2] \geq \frac{1}{2^t t^2} \tag{4.3}$$

for some integer $t \geq 1$.

Claim 4.13. *$t < 2\log_2 d$.*

Proof. We take the bound M in the statement Lemma 2.2 of to be $\frac{1}{2}d^2$, as this is the total number of edges that can be excited ($\frac{1}{2}d$ edges per vertex in $N(v)$). Hence:

$$t < \log(\frac{9}{2}M/\mathbb{E}|Y|) \leq \log(9/2q)$$

(plugging eq. (4.2) for $\mathbb{E}|Y|$.) And by recalling that $p = d^{-(2-\epsilon)}$ and $\gamma = \Omega(1/\text{polylog}\,(d))$:

$$q = \gamma p^k d^{2k-1} = \Omega(1/\text{polylog}\,(d))\frac{d^{k\epsilon}}{d} > \frac{1}{d}$$

(for sufficiently large d.) Combining the latter two inequalities, we obtain: $t < \log(\frac{9}{2}) + \log d < 2\log d$. $\square$

Recall that an excited edge has a touched vertex as its source. Each touched vertex is either light or heavy, so for each fixed marking of the vertices in phase 1, either half the excited edges are sourced in light vertices, or half are sourced in heavy vertices. By applying an averaging argument on eq. (4.3), we get that with probability $\frac{1}{2}(2^t t^2)^{-1}$, one of these two cases holds concurrently with the event $|Y| \geq \frac{1}{9} \cdot 2^t qd^2$. We now handle each case separately.

Case 1 - Light vertices: With probability $\frac{1}{2}(2^t t^2)^{-1}$, we have at least $\frac{1}{9} \cdot 2^t q d^2$ excited edges, half of which are sourced in light vertices.

In this case, we want the second phase to turn three k-excited vertices in $N(v)$ into $(k+1)$-roots by marking a light vertex for each. See fig. 3 for illustration. Let L be the set of light vertices. Since each light vertex is incident to at most two excited edges, we have $|L| \geq \frac{1}{2}|Y| = \frac{1}{4}d|X|$, where the last equality is by eq. (4.1).

Our intention is now to uniquely assign light vertices to k-excited vertices in $N(v)$ (i.e. those in X). Towards this end, we consider the bipartite graph with sides X and L and with the excited edges as the edge set. In fact, X and L may intersect; in such case we make two copies of each vertex in the intersection, putting one copy on the X-side and the other on the L-side. Note that all the edges are oriented from L to X.

Side X has average degree $|X|/|L| \geq \frac{1}{4}d$ and maximum degree $\frac{1}{2}d$ (as each vertex in X is the end of only $\frac{1}{2}d$ excited edges), so by Lemma 2.7, there is a subset $X' \subset X$ with size $|X'| \geq \frac{1}{4}|X|$, such that each vertex in X' is adjacent to at least $\frac{1}{8}d$ vertices in L. Now consider the bipartite graph with sides X' and L: Side X' has minimum degree $\frac{1}{8}d$, and side L has degree at most 2 (since a light vertex is adjacent to at most two vertices in X, and hence in X'). So by applying Lemma 2.5, we can assign the vertices in L to the vertices in X' in a way that each L-vertex is assigned to at most one X'-vertex, and each X'-vertex has at least $\frac{1}{16}d$ L-vertices assigned to it. This concludes the assignment of light vertices to (a large subset of the) k-excited vertices.

The event we are interested in, denoted henceforth as $\mathcal{A}$, is that the second phase turns three vertices in X into $(k+1)$-roots. Given the assignment we have just worked out, it is sufficient to pick three vertices in X' and for each of them to mark in the second phase an assigned vertex in L. Since each vertex in X' has $\frac{1}{16}d$ uniquely assigned vertices in L, the probability to mark one of them is $\frac{1}{16}dp(1-p)^{d/16}$, which is lower-bounded by $\frac{1}{32}dp$ for sufficiently large d. Moreover these events are independent for distinct vertices in X' (by the unique assignment), hence in total we get:

$$\Pr[\mathcal{A}] \geq \frac{1}{2 \cdot 2^t t^2} \cdot \binom{|X'|}{3} \cdot \left(\frac{1}{32} d \cdot p\right)^3 \geq \frac{1}{2 \cdot 2^t t^2} \cdot \frac{|X'|^3}{3^3} \cdot \left(\frac{1}{32} d \cdot p\right)^3 \tag{4.4}$$

(Recall that $\binom{a}{b} \geq (\frac{a}{b})^b$ for any integers a, b.) To lower-bound eq. (4.4), we recall that $|X'| \geq \frac{1}{4}|X|$, that $|X| = \frac{2}{d}|Y|$ (by eq. (4.1)) and that $|Y| \geq \frac{1}{9} \cdot 2^t q d^2$ (by the assumption of the current case). Putting these together, we get $|X'| \geq \frac{1}{18} \cdot 2^t q d$, and plugging into eq. (4.4):

$$\Pr[\mathcal{A}] \geq \frac{1}{2 \cdot (3 \cdot 32 \cdot 18)^3} \cdot \frac{2^{2t}}{t^2} \cdot (q d^2 \cdot p)^3$$

Next we recall that $1 \leq t < 2 \log d$ and $q = \gamma p^k d^{2k-1} = \Omega(1/\mathrm{polylog}\,(d)) \cdot p^k d^{2k-1}$. Plugging these into the above, we get:

$$\Pr[\mathcal{A}] \geq \frac{(p^{k+1} d^{2k+1})^3}{\Omega(\mathrm{polylog}\,(d))}$$

We need this bound to be at least $M \cdot p$. Suppressing the constant M into the Ω notation, we need the following to hold:

$$p^{3k+2} d^{6k+3} \geq \Omega(\mathrm{polylog}\,(d))$$

Recalling that $p = d^{-(2-\epsilon)}$, the latter is rewritten as:

$$d^{(3k+2)\epsilon - 1} \geq \Omega(\mathrm{polylog}\,(d))$$

For this to hold for sufficiently large d, it is enough to require:

$$(3k+2)\epsilon - 1 > 0$$

and this holds by the hypothesis of the lemma. Hence we have obtained: $\Pr[\mathcal{A}] \geq M \cdot p$.

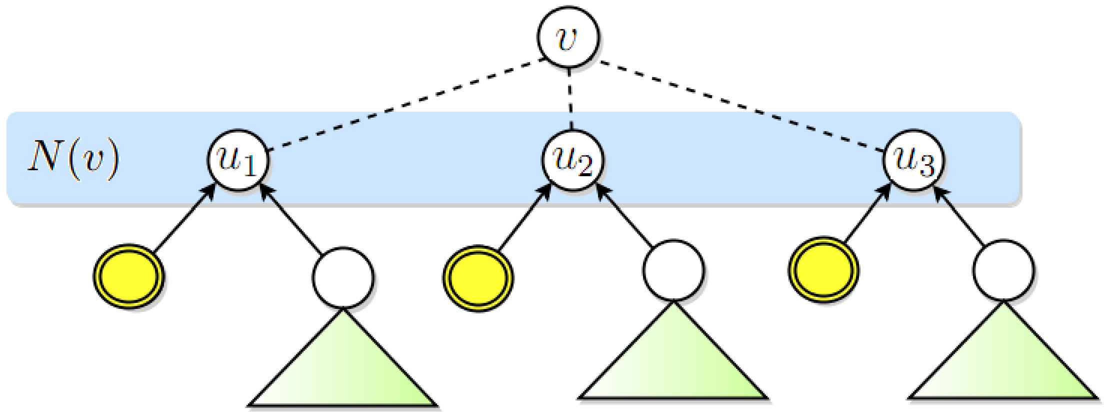

Figure 3: u_1, u_2, u_3 were k-excited after the first phase. In the second phase, marking a light vertex (in double circle) for each turns them into a $(k+1)$-roots.

Case 2 - Heavy vertices: With probability $\frac{1}{2}(2^t t^2)^{-1}$, we have at least $\frac{1}{9} \cdot 2^t q d^2$ excited edges, half of which are incident to heavy vertices.

Let L be the subset of heavy vertices. Each such vertex is incident to at most d excited edges (as there are only d vertices in $N(v)$), so we have $|L| \geq \frac{1}{18} \cdot 2^t q d$.

Again we let $\mathcal{A}$ denote the event that the second phase turns three vertices in X into $(k+1)$-roots. A heavy vertex is the source of at least three excited edges, so it is adjacent to three k-excited vertices in $N(v)$. Hence, marking any heavy vertex in the second phase is enough for $\mathcal{A}$ to occur. See fig. 4 for illustration. The probability to mark a heavy vertex in the second phase is $(1 - o(1))|L|p$, and hence:

$$\Pr[\mathcal{A}] \geq \frac{1}{2 \cdot 2^t t^2} \cdot (1 - o(1))|L|p \geq \frac{(1 - o(1))qdp}{36 \cdot t^2}$$

Plugging $t < 2 \log d$ and $q = \gamma p^k d^{2k-1} = \Omega(1/\text{polylog}\,(d)) \cdot p^k d^{2k-1}$, we get

$$\Pr[\mathcal{A}] \geq \frac{p^{k+1} d^{2k}}{\Omega(\text{polylog}\,(d))}$$

We need this bound to be at least $M \cdot p$. Suppressing the constant M into the Ω notation, we need the following to hold:

$$p^k d^{2k} \geq \Omega(\text{polylog}\,(d))$$

Recalling that $p = d^{-(2-\epsilon)}$, the latter is equivalent to:

$$d^{\epsilon k} \geq \Omega(\text{polylog}\,(d))$$

and this clearly holds for sufficiently large d. Hence we have obtained: $\Pr[\mathcal{A}] \geq M \cdot p$.

Conclusion. In both of the above cases, we have shown that with probability at least $M \cdot p$ over the two phases, there are three $(k+1)$-roots in $N(v)$. By Proposition 4.11, each such root is part of an arrangement with at most k non-marked vertices, and by Lemma 4.10 we can unite the three arrangements. In summation, we've shown that with probability $M \cdot p$, the graph $G - \{v\}$ has an

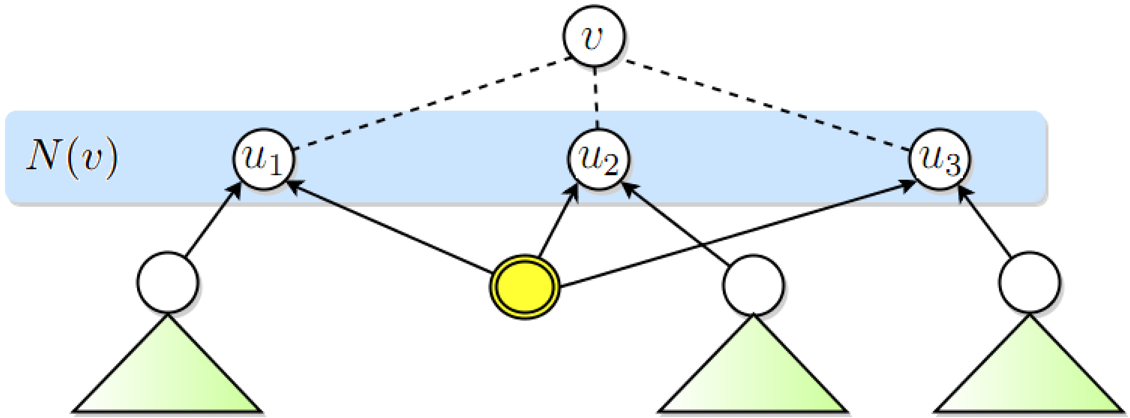

Figure 4: u_1, u_2, u_3 were k-excited after the first phase. In the second phase, marking a single heavy vertex (in double circle) turns all of them into $(k+1)$-roots.

arrangement H'_v that contains three neighbours of v, say u_1, u_2, u_3, and has at most $3k$ non-marked vertices.

The claim now follows easily: All the above was shown to hold for $G - \{v\}$, and hence holds also for G with the guarantee that $v \notin H'_v$. We therefore can construct an arrangement H_v as required in the claim, by adding v to H'_v, and orienting towards it the three edges connecting it to u_1, u_2, u_3. The proof of the claim is complete. $\square$

4.3.3 Proof of Theorem 4.7

Consider a sample from $G(V, E, 2p)$. Let A denote the subset of marked vertices, so $\mathbb{E}[|A|] = 2np$. Let $\mathcal{O}$ be a random orientation of the edges in G, sampled as follows: Each edge is oriented with independent probability $\frac{1}{2}$ towards each of its two ends. We define B to be the random subset that contains each vertex v if,

- v satisfies the conclusion of Lemma 4.12. That is, v is part of an arrangement H_v with at most $3k+1$ non-marked vertices, and has three edges oriented towards it in H_v.
- The orientation of H_v coincides with the orientation $\mathcal{O}$.

By Lemma 4.12, for sufficiently large d, each vertex v has probability $\geq M \cdot p$ to satisfy the first item, for a constant M that we will set right away. As for the second item, H_v has at most $3k+1$ non-marked vertices that each of whom has two edges oriented towards it (by definition of arrangement), plus one additional edge oriented towards v (by the assertion of Lemma 4.12). Hence H_v has at most $6k+3$ edges, and therefore, $\mathcal{O}$ has probability $\geq (\frac{1}{2})^{6k+3}$ to coincide with the orientation of H_v. In conclusion, v has probability $\geq (\frac{1}{2})^{6k+3} Mp$ to be in B, over the choices of both $\mathcal{O}$ and the sample of $G(V, E, 2p)$. We set $M = 16 \cdot 2^{6k+3}$ and obtain that $\mathbb{E}[|B|] \geq 16np$.

By applying the Chernoff bound Lemma 2.3 to $|A|$ and plugging $n \geq \Omega(d^2)$, we get:

$$\Pr[|A| < 2\mathbb{E}|A|] \geq 1 - (0.25e)^{2np} = 1 - (0.25e)^{2nd^{\epsilon-2}} \geq 1 - (0.25e)^{\Omega(d^\epsilon)}$$

On the other hand, by Lemma 2.1, $|B|$ attains half its expected value with probability at least $\frac{\mathbb{E}|B|}{2n} = 8p = 8d^{-(2-\epsilon)}$. Summing the bounds yields:

$$\Pr[|A| \leq 2\mathbb{E}|A|] + \Pr\left[|B| \geq \frac{1}{2}\mathbb{E}|B|\right] \geq 1 - (0.25e)^{\Omega(d^\epsilon)} + 8d^{-(2-\epsilon)}$$

Since $(0.25e)^{\Omega(d^\epsilon)} \ll 8d^{-(2-\epsilon)}$ for sufficiently large d, the above RHS is strictly larger than 1. Therefore, there is a positive probability that both the events $|A| \leq 2\mathbb{E}|A|$ and $|B| \geq \frac{1}{2}\mathbb{E}|B|$ occur. These imply $|A| \leq 4np$ and $|B| \geq 8np$, respectively. We fix this event from now on (note that this also fixes an orientation $\mathcal{O}$), and arbitrarily remove vertices from B until $|B| = 8np$. We take our target subgraph H to be the one induced by the union of A and the arrangements H_v for all $v \in B$.

Size of H. For each $v \in B$, H_v has at most $3k+1$ non-marked vertices. Hence:

$$|H| \leq |A| + (3k+1)|B| \leq (2 + 4(3k+1))np = (12k+6)n/d^{2-\epsilon}$$

which is as required, if we set $C(k) = 12k + 6$.

Average degree of H. To show that H has average degree at least 4, we need to count two edges per vertex. We use the orientation $\mathcal{O}$ to assign edges to vertices, to ensure that each edge is counted in favour of only one vertex.

Consider a non-marked vertex u in H. It is part of an arrangement H_v for some $v \in B$, and being an arrangement, u has two edges oriented towards it in H_v. Since the orientation of H_v coincides with $\mathcal{O}$, we see that u has two edges oriented towards it in $\mathcal{O}$.

Additionally, each $v \in B$ has a third edge oriented towards it in H_v, and hence in $\mathcal{O}$. Together, we have $|B|$ edges that we have not yet used, and we now count them to cover for the marked vertices. Since the number of marked vertices is $|A| \leq 4np$, and the number of remaining edges is $|B| = 8np$, we can indeed count two edges per marked vertex. In conclusion, we see that H has average degree ≥ 4, and the proof is complete. □

4.4 Tighter Bounds for Lower Densities

We now prove Theorem 1.7. The idea is that either some vertex neighbourhood is dense enough to constitute our target subgraph, or all neighbourhoods are sparse enough to be considered as trees, which be handled as in Claim 4.5.

Lemma 4.14. *Let $t \geq 1$ be a fixed integer, $0 < \alpha < 1$, and $r > 0$ sufficiently large. Let T be an r-ary tree with all leaves in level t.*[5] *Suppose we remove all but at least an α-fraction of the leaves from T. The remaining tree contains an r'-ary tree with $r' = (\frac{1}{2})^{t-1}\alpha r$, with the same root as T and with all leaves in level t.*

Proof. By induction on t. In the base case $t = 1$ there is nothing to show, as the root itself is the only leaf and $\alpha > 0$, so we cannot remove any leaves. Suppose now $t > 1$. Let L be the subset of remaining leaves after the removal, and let M be the subset of vertices in the $(t-1)^{th}$ level of T. In the bipartite graph with sides L and M, side M has maximum degree r and average degree $\frac{|L|}{|M|} = \frac{\alpha \cdot r^{t-1}}{r^{t-2}} = \alpha r$. Hence by Lemma 2.7, there is a subset $M' \subset M$ with size $|M'| \geq \frac{\alpha}{2}|M|$ such that each vertex in M' has $\frac{1}{2}\alpha r$ neighbours in L.

Consider the r-ary tree T' given by the top $t-1$ levels of T. The subset of its leaves is M, and suppose we remove all leaves but those in M'. This removes all but at least a $\frac{1}{2}\alpha$-fraction of the leaves, so by the inductive hypothesis, T' contains an r'-ary tree T" with $r' = (\frac{1}{2})^{t-2} \cdot \frac{1}{2}\alpha \cdot r = (\frac{1}{2})^{t-1}\alpha r$, which has the same root as T and all leaves in M'. We extend T" by one more level, by picking for each leaf in M' an arbitrary subset of r' neighbours in L, which it is guaranteed to have by the choice of M' (as explained above). T" constitutes the required subtree of T. □

[5] Recall that as in Claim 4.5, we consider the root to be in level 1.

Lemma 4.15 (main towards proving Theorem 1.7)**.** *Let $t \geq 1$ be an integer. There is a constant $\gamma_t > 0$ such that for every graph G with minimum degree d sufficiently large and an arbitrary vertex v in G, at least one of the following holds:*

- *G contains a subgraph of size at most $\gamma_t d^{t-2}$ and average degree ≥ 4.*
- *G contains a $(\gamma_t d)$-ary tree rooted by v, with all leaves in level t.*

Proof. For each vertex u in G we restrict our attention to an arbitrary subset of exactly d of its neighbours, and refer only to them as its neighbours. This approach has been taken in our proofs before; as usual, it may happen that for an adjacent pair or vertices u, u' we consider u' to be a neighbour of u but not vice-versa, and this would not interfere with our reasoning. Moreover, keep in mind that we will assume d is sufficiently large wherever necessary.

We go by induction on t. The base case $t = 1$ is trivial, since v alone is a tree with one level (of any arity). Now fix $t > 1$. By induction, G either contains a subgraph of size d^{t-3} and average degree ≥ 4, or a $(\gamma_{t-1}d)$-ary tree T rooted by v with all leaves in level $t-1$. The former immediately implies the proposition, so we now focus on the latter. Let L denote the set of leaves in T. Note that $|L| = (\gamma_{t-1}d)^{t-2}$ and,

$$|T| = \frac{(\gamma_{t-1}d)^{t-1} - 1}{\gamma_{t-1}d - 1} \leq 2(\gamma_{t-1}d)^{t-2}.$$

Let E_L be the subset of edges incident to vertices in L. Since G has minimum degree d, we have $|E_L| \geq d|L|$. We write E_L as a disjoint union $E_L = E_L^{in} \cup E_L^{out}$, where E_L^{in} are the edges going back into T (i.e. have both endpoints in T), and E_L^{out} are all the other edges. If $|E_L^{in}| \geq 2|T|$ then T is a subgraph with average degree 4 and size $\leq 2(\gamma_{t-1}d)^{t-2}$, which meets the requirement of the proposition if we set $\gamma_t = 2\gamma_{t-1}^{t-2}$. Otherwise,

$$|E_L^{out}| > |E_L| - 2|T| \geq \gamma_{t-1}^{t-2}d^{t-1} - 4(\gamma_{t-1}d)^{t-2} \geq \tfrac{1}{2}\gamma_{t-1}^{t-2} \cdot d^{t-1} \geq \frac{1}{2}d|L| \tag{4.5}$$

for sufficiently large d.

Let N_L be the set of all endpoints of edges in E_L^{out} which are not in T. (Note that each edge in E_L^{out} has exactly one endpoint not in T.) We say a vertex in N_L is *light* if it is a neighbour of at most two vertices in L, and *heavy* otherwise. Now, either half the edges in E_L^{out} are incident to light vertices, or half are incident to heavy vertices. We handle the two cases separately.

- *Case I - Light vertices:* Half the edges in E_L^{out} are incident to light vertices in N_L.

  Let $N_l \subset N_L$ denote the set of light vertices. Since each light vertex is incident to at most 2 edges in E_L^{out}, we have $|N_l| \geq \frac{1}{4}|E_L^{out}|$. Hence in the bipartite graph with sides L and N_l, side L has maximum degree d and average degree (using eq. (4.5)):

  $$\frac{|N_l|}{|L|} \geq \frac{\frac{1}{4}|E_L^{out}|}{|L|} \geq \frac{\frac{1}{8}d|L|}{|L|} = \tfrac{1}{8}d$$

  Now by Lemma 2.7, there is a subset $L' \subset L$ of size at least $\frac{1}{16}|L|$ such that each vertex in L' has $\frac{1}{16}d$ neighbours in N_l. Remove from T all leaves except those in L'. By Lemma 4.14 (with $\alpha = \frac{1}{16}$ and $r = \gamma_{t-1}d$), T contains a $((\frac{1}{2})^{t+2}\gamma_{t-1}d)$-ary subtree T' with all leaves in level $t-1$, that is in L'. We extend T by one more level using Lemma 2.5: Recall that in the bipartite graph with sides L' and N_l, side L has minimum degree $\frac{1}{16}d$ and side N_l has maximum degree 2. Hence, Lemma 2.5 picks a subset of edges such that each vertex in N_l is adjacent to at most one vertex in L', and each vertex in L' has at least $\frac{1}{32}d$ neighbours in N_l, thus adding a t^{th} level to T' while keeping it r-ary with $r = \min\{(\frac{1}{2})^{t+2}\gamma_{t-1}d, \frac{1}{32}d\}$. This meets the requirement of the proposition.

- *Case II - Heavy vertices:* Half the edges in E_L^{out} are incident to heavy vertices in N_L.

  Let $N_h \subset N_L$ be the subset of heavy vertices.

  - If $|N_h| < 2|L|$, consider the subgraph H induced by $L \cup N_h$. It has at most $3|L|$ vertices, and it contains all the edges in E_L^{out} which are incident to heavy vertices. By the assumption of the current case there are at least $\frac{1}{2}|E_L^{out}|$ such edges, and by eq. (4.5), this is at least $\frac{1}{2}d|L|$. Hence for $d \geq 12$, H has average degree ≥ 4.
  - If $|N_h| \geq 2|L|$, pick an arbitrary subset $N_h' \subset N_h$ of size exactly $2|L|$, and consider the subgraph H induced by $L \cup N_h'$. It has $3|L|$ vertices and at least $3|N_h'| = 6|L|$ edges, as each vertex in N_h' is incident to at least 3 edges in E_L^{out}. Hence H has average degree ≥ 4.

  In both cases the size of H is bounded by $3|L| = 3(\gamma_{t-1}d)^{t-2}$, and hence satisfies the requirement of the proposition if we set $\gamma_t = 3\gamma_{t-1}^{t-2}$.

Conclusion. Considering all the arising cases, the claim is proven with

$$\gamma_t = \min\{\frac{1}{16}, (\frac{1}{2})^{t+1}\gamma_{t-1}, 3\gamma_{t-1}^{t-2}\}.$$

□

Proof of Theorem 1.7. As usual we assume that d is the minimum degree and that it is sufficiently large. Set $k = 2^{t-2}$ (and note this adheres to Claim 4.5). The motivation for the choice t in the statement is that in order to apply Theorem 4.7, we need to pick t such that $\epsilon > 1/(3k+2)$. As already remarked in the proof of Theorem 4.8, one may verify by rearranging that the value set for t is the smallest integer that meets this requirement.

We apply Lemma 4.15 to each vertex in G, and handle two cases:

- Case I: For some vertex v, the first option of Lemma 4.15 is met. This means G contains a subgraph of size at most $\gamma_t d^{t-2}$ and average degree 4.

- Case II: For each vertex v, the second option of Lemma 4.15 is met. This means v is the root of a $(\gamma_t d)$-ary tree with all leaves in level t, so by Claim 4.5, v has probability $\Omega(1) \cdot d^{2k-1}p^k$ to be k-excited with $p = 1/d^{2-\epsilon}$. This holds for each vertex, hence G is $(k, p, \Omega(1))$-tree-like, and this holds even if we remove any single vertex from the graph, hence G is $(k, p, \Omega(1))$-good. We can now apply Theorem 4.7 with $\epsilon' = \delta(t) + \epsilon$. Recalling that $\delta(t) = 1/(3k + 2)$, we have $\epsilon' > 1/(3k + 2)$ so the condition of Theorem 4.7 is satisfied. Thus, there is a constant $C = C(k, \epsilon)$ such that G contains a subgraph of size $C \cdot n/d^{2-\epsilon}$ and average degree 4.

Combining the two cases, we obtain that G contains a subgraph of average degree 4 and size at most $\max\{\gamma_t d^{t-2}, Cn/d^{2-\delta(t)-\epsilon}\}$. It remains to show

$$\gamma_t d^{t-2} \leq Cn/d^{2-\delta(t)-\epsilon},$$

and this holds by the hypothesis $d = O(n^{1/(t-\delta(t)-\epsilon)})$, and by letting C be sufficiently large. □

5 Lower Bounds

In this section we provide lower bounds on $A_\ell(n,d)$, by establishing the existence of arbitrarily large graphs that exclude all ℓ-subgraphs up to a certain size.

We note that a graph may have an ℓ-subgraph of a certain size but not of any larger size, so in a sense the property is non-monotone. (To illustrate this, consider an $(\ell+1)$-clique joined with arbitrarily many isolated vertices.) Therefore in order to prove $A_\ell(n,d) \geq s$, we need explicitly to rule out ℓ-subgraph of all sizes up to s, and not just s.

We now define the random graph model that we will work with.

Definition 5.1. *The distribution $\mathcal{G}_{min}(n,d)$ over simple graphs on the vertex set $[n] = \{1,\ldots,n\}$ is defined by the following sampling process:*

- *In the first stage, each vertex chooses uniformly at random a subset of size d of the remaining $n-1$ vertices, and connects to them with an (undirected) edge. Parallel edges are allowed.*
- *In the second stage, parallel edges are unified into a single edge. The resulting (simple) graph is the output sample.*

Proposition 5.2. *Let G be sampled from $\mathcal{G}_{min}(n,d)$ with $d \geq 2$. Then,*

1. *G has minimum degree at least d.*
2. *G has average degree between d and $2d$.*
3. *Any subset F of possible edges on the vertex set $[n]$ occurs in G w.p. $\leq (\frac{2d}{n})^{|F|}$.*

Proof. (*1*) In the first stage of the sampling process, each vertex chooses d neighbours, and remains connected to all of them after the second stage.

(*2*) The first stage places exactly dn edges in G. Then, as each pair of vertices is connected with at most two parallel edges, the second stage removes at most half the edges.

(*3*) Each edge occurs in G with probability $p = \frac{d}{n-1} + \left(1 - \frac{d}{n-1}\right)\frac{d}{n-1} \leq \frac{2d}{n}$ (for $d \geq 2$), and concurrent appearance of edges is either independent (if they are vertex-disjoint) or negatively correlated (otherwise). $\square$

We turn to establish lower bounds. The next theorem states that $A_\ell(n,d) = \Omega(n/d^{\ell/(\ell-2)})$, for all densities d.

Theorem 5.3. *Let $\ell > 2$ be even. There is a constant $c_\ell > 0$ such that for all sufficiently large n and $d = O(n^{(\ell-2)/\ell})$, there is a graph on n vertices with minimum degree d, without any ℓ-subgraphs of size $\leq c_\ell \cdot n/d^{\ell/(\ell-2)}$.*

Proof. Denote $h = \frac{1}{2}\ell$. Let G be a sample of $\mathcal{G}_{min}(n,d)$. For each subset U of $[n]$, let A_U denote the event that G contains an ℓ-subgraph on the vertex set U. We recall this means that there is a subset F of edges with size $\lceil h|U| \rceil$ and with all endpoints in U. We assume for simplicity that $h|U|$ is an integer (even though this is not necessary), so $|F| = h|U|$. By Proposition 5.2, each such F occurs in G w.p. $\leq \left(\frac{2d}{n}\right)^{h|U|}$, so a union bound over the possible choices of F out of the edges that may be induced by U, implies:

$$\Pr[A_U] \leq \binom{|U|^2}{h|U|} \cdot \left(\frac{2d}{n}\right)^{h|U|} \leq \left(\frac{2ed|U|}{hn}\right)^{h|U|} \tag{5.1}$$

where we have applied the known bound $\binom{m}{k} \le \left(\frac{em}{k}\right)^k$. For $s > 0$, we apply another union bound to bound the probability p_s that G contains an ℓ-subgraph of size at most s:

$$p_s \le \sum_{U \subset [n], |U| \le s} \Pr[A_U] = \sum_{t=1}^{s} \sum_{U \subset [n], |U| = t} \Pr[A_U] \le \sum_{t=1}^{s} \binom{n}{t} \left(\frac{2edt}{hn}\right)^{ht} \le \sum_{t=1}^{s} \left(\frac{en}{t}\right)^t \left(\frac{2edt}{hn}\right)^{ht}$$

$$= \sum_{t=1}^{s} \left(e \left(\frac{2e}{h}\right)^h \cdot \frac{d^h}{n^{h-1}} \cdot t^{h-1} \right)^t \le \sum_{t=1}^{s} \left(e \left(\frac{2e}{h}\right)^h \cdot \frac{d^h}{n^{h-1}} \cdot s^{h-1} \right)^t \tag{5.2}$$

Plugging $s = c_\ell \cdot n / d^{h/(h-1)}$ with $c_\ell = \left(\frac{1}{3e} (\frac{h}{2e})^h\right)^{1/(h-1)}$, we find that $p_s \le \sum_{t=1}^{s} (\frac{1}{3})^t < \sum_{t=1}^{\infty} (\frac{1}{3})^t = \frac{1}{2} < 1$. Hence there is a sample G without ℓ-subgraphs of size at most s. $\square$

The next theorem states that $A_\ell(n, d) = \omega(n/d^{\ell/(\ell-2)})$ for the highest density case, $d = \Theta(n^{(\ell-2)/\ell})$. This amounts to finding arbitrarily large graphs of such density without any constant-sized ℓ-subgraphs.

Theorem 5.4. *Let $\ell > 2$ be even and $c, s > 0$ be arbitrary constants. For all sufficiently large n, there is a graph on n vertices with minimum degree $d = c \cdot n^{(\ell-2)/\ell}$, without any ℓ-subgraphs of size $\le s$.*

Proof. Our proof parallels that of Hoory [Hoo02, Theorem A.4], which addresses the closely related problem of showing there are graphs with large girth. Denote $h = \frac{1}{2}\ell$. Let G be a sample of $\mathcal{G}_{min}(n, d)$. For each subset U of $[n]$, let A_U denote the event that G contains an ℓ-subgraph on the vertex set U. For U with size $|U| = t \le s$, we have:

$$\Pr[A_U] \le \left(\frac{2edt}{hn}\right)^{ht} \le \frac{\alpha}{n^t} \tag{5.3}$$

for a sufficiently large constant α, where the first inequality is by eq. (5.1). We need to show that with positive probability, none of the events $\{A_U : |U| \le s\}$ occurs. To this end we invoke the Local Lemma, stated next. For a proof see [AS11, Lemma 5.1.1].

Lemma 5.5 (Local Lemma)**.** *Let $\mathcal{A}$ be a finite set of events in an arbitrary probability space. For $A \in \mathcal{A}$, let $\Gamma(A) \subset \mathcal{A}$ be such that A is independent of the collection of events $\mathcal{A} \setminus (A \cup \Gamma(A))$. If there is an assignment of reals $x : \mathcal{A} \to (0, 1)$ such that for all $A \in \mathcal{A}$,*

$$\Pr[A] \le x(A) \cdot \prod_{B \in \Gamma(A)} (1 - x(B)) \tag{5.4}$$

then with positive probability, none of the events in $\mathcal{A}$ occurs.

Observe that in $\mathcal{G}_{min}(n, d)$, the event A_U is determined solely by the choices of the vertices in U, and hence is independent of $A_{U'}$ for all $U' \subset [n]$ such that $U \cap U' = \emptyset$. In other words, A_U may be dependent only of events $A_{U'}$ for which U, U' have mutual vertices. Therefore, applying Lemma 5.5 with $x(A_U) = 2\Pr[A_U]$, the condition eq. (5.4) becomes:

$$\Pr[A_U] \le 2\Pr[A_U] \cdot \prod_{r=1}^{s} \prod_{\substack{|U'| = r \\ U \cap U' \neq \emptyset}} (1 - 2\Pr[A_{U'}]) \tag{5.5}$$

The number of subsets U' of size r that share any vertices with U is:

$$\binom{n}{r} - \binom{n - |U|}{r} \leq \frac{n^r}{r!} - \frac{(n - r - |U| + 1)^r}{r!} \leq \beta n^{r-1}$$

for a sufficiently large constant β. Using this and eq. (5.3) we get,

$$\prod_{r=1}^{s} \prod_{\substack{|U'|=r \\ U \cap U' \neq \emptyset}} (1 - 2\Pr[A_{U'}]) \geq \prod_{r=1}^{s} \left(1 - \frac{2\alpha}{n^r}\right)^{\beta n^{r-1}} \geq \exp\left(-\sum_{r=1}^{s} \frac{4\alpha\beta}{n}\right) \geq \frac{1}{2}$$

for sufficiently large n. (The middle inequality is because $1 - z \geq \exp(-2z)$ holds for all, say, $0 < z < \frac{1}{2}$.) Consequently eq. (5.5) is satisfied, so by Lemma 5.5 we get positive probability that none of the events occurs, and the theorem follows. $\square$

6 Hypergraphs

In this section we provde bounds for Question 1.3. The starting point of this work was in fact the following conjecture of Feige [Fei08].

Conjecture 6.1 (Conjecture 1.7 from [Fei08])**.** *Let c be sufficiently large. Every 3-uniform hypergraph on n vertices and $m = c \cdot dn$ hyperedges (with $1 < d \leq O(\sqrt{n})$) has a set of $n' \leq \tilde{O}(n/d^2)$ vertices that induce at least $2n'/3$ hyperedges. Put otherwise, $A_2(n, d, 3) = \tilde{O}(n/d^2)$.*

Note that a 2-subhypergraph in a 3-hypergraphs contains at least 2/3 edges per vertex, so Conjecture 6.1 concerns the $\ell = 2$ case. A matching lower bound of $A_2(n, d, 3) = \Omega(n/d^2)$ can be obtaiedn similarly to Theorem 5.3, by considering random hypergraph model in which each hyperedge is present with independent probability $p = d/n^2$. Further details are omitted.

We will show that our upper bounds extend to hypergraphs. This brings us short of proving Conjecture 6.1, but we obtain the following weaker variants, which are analogues of Theorems 1.5, 1.6 and 1.7 respectively.

Theorem 6.2. *For every $\epsilon > 0$,*

1. $A_{2-\epsilon}(n, d, 3) = O(n/d^2)$.
2. $A_2(n, d, 3) = O(n/d^{1.8-\epsilon})$.
3. *Let $t \geq 2$ be an integer. Every 3-hypergraph of size n and average degree $d = O(n^{1/(t-\delta(t)-\epsilon)})$ contains a 2-subgraph of size $O(n/d^{2-\delta(t)-\epsilon})$, where $\delta(t) = 1/(3 \cdot 2^{t-2} + 2)$.*

The constants hidden in the O-notations depend on ϵ.

The proof is by simply replacing each hyperedge with a clique of pairwise edges, and considering the resulting graph. It may have parallel edges, which we have to take into account in order to avoid a fatal loss in the average degree.

Observation 6.3. *Theorems 1.5, 1.6 and 1.7 hold as stated for graphs with parallel edges.*

Proof. If any two vertices in the graph are connected with 4 of more parallel edges, they form a 4-subgraph of size 2. Otherwise, eliminating all parallel edges decreases the average degree from d to no less than $\frac{1}{3}d$, and Theorems 1.5, 1.6 and 1.7 hold with only a change of constants. □

We now formalize the reduction from 3-uniform hypergraphs to graphs.

Definition 6.4 (skeleton of a hypergraph)**.** *Let $H(V, E)$ be a hypergraph. The* skeleton multigraph *G_H of H has vertex set V, and for each $u, v \in V$, the number of parallel edges connecting u, v in G_H is the number of hyperedges in E covering both u and v. Equivalently, each hyperedge in H induces an edge in G_H between each pair it covers.*

Lemma 6.5 (main for Section 6)**.** *Let $H(V, E)$ be a 3-uniform hypergraph with skeleton multigraph G_H. Suppose G_H has a $(4 - \epsilon)$-subgraph G' of size k, for $0 \leq \epsilon < 4$. Then H has a $(2 - \frac{1}{2}\epsilon)$-subhypergraph of size at most $3k$.*

Proof. We may assume w.l.o.g. that G' is an induced subgraph. Let E' be the subset of hyperedges in H that induce any edges in G'. We can partition E' to disjoint subsets A, B as follows:

- A - hyperedges that induce exactly *one* edge in G'. Let $a = |A|$.

- B - hyperedges that induce exactly *three* edge in G'. Let $b = |B|$.

(Note that since we have assumed G' is an induced subgraph, there are no hyperedges that induce two edges.) With this notation G' has exactly $a + 3b$ edges, and since its average degree is at least $4 - \epsilon$, we have

$$a + 3b \geq (2 - \frac{1}{2}\epsilon)k. \tag{6.1}$$

We now arbitrarily remove edges from A (and thus from E') either until A is empty, or until $a + 3b \leq 2k$ (whichever happens first). Note eq. (6.1) remains satisfied.

Let V' be the subset of vertices in H that are covered by any edges in E', and consider the subgraph $H'(V', E')$ of H. Let $U \subset V'$ be the subset of vertices in H' that are not in G'. Observe that each hyperedge in A contributes one vertex to U, and a hyperedge in B contributes no vertices to U. Hence $|U| \leq a$ (note that a vertex in U may be contributed by more than one hyperedge in A), and we conclude:

$$|V'| = k + |U| \leq k + a. \tag{6.2}$$

We bound the size of H'. Recall that by the above, one of two cases must hold: Either $A = \emptyset$, in which case $U = \emptyset$ and hence $|V'| = k$, or $a + 3b \leq 2k$, in which case $a \leq 2k$ and hence $|V'| = k + |U| \leq k + a \leq 3k$. In both cases we have $|V'| \leq 3k$.

We bound the average degree of H'. Using eqs. (6.1) and (6.2), we get:

$$|V'| \leq k + a \leq \frac{a + 3b}{2 - \epsilon} + a = \frac{3a + 3b - \frac{1}{2}\epsilon a}{2 - \epsilon} = \frac{3|E'| - \epsilon a}{2 - \frac{1}{2}\epsilon}$$

(recalling that $|E'| = a + b$), and hence:

$$\text{avgdeg}(H') = \frac{3|E'|}{|V'|} \geq \frac{3(2 - \frac{1}{2}\epsilon)|E'|}{3|E'| - \epsilon a} = \frac{2 - \frac{1}{2}\epsilon}{1 - \frac{\epsilon a}{3|E'|}} \geq 2 - \tfrac{1}{2}\epsilon,$$

as required. □

Proof of Theorem 6.2. Apply either of Theorems 1.5, 1.6 and 1.7 to the skeleton multigraph G_H of H, relying on Observation 6.3. Then apply Lemma 6.5 to the resulting subgraph G'. □

Acknowledgements. This work was supported in part by the Israel Science Foundation (grant No. 621/12) and by the I-CORE Program of the Planning and Budgeting Committee and the Israel Science Foundation (grant No. 4/11).

References

[AF09] Noga Alon and Uriel Feige, *On the power of two, three and four probes*, proceedings of the 20th annual ACM-SIAM symposium on discrete algorithms (SODA), 2009, pp. 346–354.

[AHI02] Yuichi Asahiro, Refael Hassin, and Kazuo Iwama, *Complexity of finding dense subgraphs*, Discrete Applied Mathematics **121** (2002), no. 1-3, 15–26.

[AHL02] Noga Alon, Shlomo Hoory, and Nathan Linial, *The moore bound for irregular graphs*, Graphs and Combinatorics **18** (2002), no. 1, 53–57.

[APP+12] Omid Amini, David Peleg, Stéphane Pérennes, Ignasi Sau, and Saket Saurabh, *On the approximability of some degree-constrained subgraph problems*, Discrete Applied Mathematics **160** (2012), no. 12, 1661–1679.

[AS06] Noga Alon and Asaf Shapira, *On an extremal hypergraph problem of Brown, Erdős and Sós*, Combinatorica **26** (2006), no. 6, 627–645.

[AS11] Noga Alon and Joel H. Spencer, *The probabilistic method*, Wiley Series in Discrete Mathematics and Optimization, Wiley, 2011.

[ASS12] Omid Amini, Ignasi Sau, and Saket Saurabh, *Parameterized complexity of finding small degree-constrained subgraphs*, J. Discrete Algorithms **10** (2012), 70–83.

[BB89] Béla Bollobás and Graham Brightwell, *Long cycles in graphs with no subgraphs of minimal degree 3*, Discrete Mathematics **75** (1989), no. 1-3, 47–53.

[BB90] Norman L. Biggs and Alan G. Boshier, *Note on the girth of ramanujan graphs*, J. Comb. Theory, Ser. B **49** (1990), no. 2, 190–194.

[BES73a] William G. Brown, Paul Erdős, and Vera T. Sós, *On the existence of triangulated spheres in 3-graphs, and related problems*, Periodica Mathematica Hungarica **3** (1973), no. 3-4.

[BES73b] ———, *Some extremal problems on r-graphs*, New Directions in the Theory of Graphs, Proceedings of the 3rd Ann Arbor Conference on Graph Theory (1973).

[Bol04] Béla Bollobás, *Extremal graph theory*, Dover Books on Mathematics Series, Dover Publications, 2004.

[Dah13] Xavier Dahan, *Regular graphs of large girth and arbitrary degree*, in arXiv.org, CoRR abs/1110.5259v4. To appear in Combinatorica (2013).

[DHL+12] Domingos Jr. Dellamonica, Penny E. Haxell, Tomasz Luczak, Dhruv Mubayi, Brendan Nagle, Yury Person, Vojtech Rödl, Mathias Schacht, and Jacques Verstraëte, *On even-degree subgraphs of linear hypergraphs*, Combinatorics, Probability & Computing **21** (2012), no. 1-2, 113–127.

[EFGS88] Paul Erdős, Ralph J. Faudree, András Gyárfás, and Richard H. Schelp, *Cycles in graphs without proper subgraphs of minimum degree 3*, Ars Combinatorica **25(B)** (1988), 195201.

[EFR86] Paul Erdős, Peter Frankl, and Vojtěch Rödl, *The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent*, Graphs and Combinatorics **2** (1986), no. 1, 113–121 (English).

[EFRS90] Paul Erdős, Ralph J. Faudree, Cecil C. Rousseau, and Richard H. Schelp, *Subgraphs of minimal degree k*, Discrete Mathematics **85** (1990), no. 1, 53–58.

[Fei02] Uriel Feige, *Relations between average case complexity and approximation complexity*, proceedings of the 34th annual ACM symposium on theory of computing (STOC), 2002, pp. 534–543.

[Fei07] ———, *Refuting smoothed 3CNF formulas*, proceedings of the 48th annual symposium on foundations of computer science (FOCS), 2007, pp. 407–417.

[Fei08] ———, *Small linear dependencies for binary vectors of low weight*, Building Bridges Between Mathematics and Computer Science, Bolyai Society Mathematical Studies **19** (2008), 283–307.

[FKO06] Uriel Feige, Jeong Han Kim, and Eran Ofek, *Witnesses for non-satisfiability of dense random 3CNF formulas*, proceedings of the 47th annual symposium on foundations of computer science (FOCS), 2006, pp. 497–508.

[FKW16] Uriel Feige, Shimon Kogan, and Tal Wagner, *Hardness of the average degree constrained subgraph problem*, in preparation (2016).

[FO07] Uriel Feige and Eran Ofek, *Easily refutable subformulas of large random 3CNF formulas*, Theory of Computing **3** (2007), no. 1, 25–43.

[FS97] Uriel Feige and Michael Seltser, *On the densest k-subgraph problem*, Technical Report CS97-16, Faculty of Mathematics and Computer Science, Weizmann Institute of Science (1997).

[Hoo02] Shlomo Hoory, *On graphs of high girth*, Hebrew University of Jerusalem, 2002.

[Imr84] Wilfried Imrich, *Explicit construction of regular graphs without small cycles*, Combinatorica **4** (1984), no. 1, 53–59 (English).

[Kez91] André E. Kezdy, *Studies in connectivity*, University of Illinois at Urbana-Champaign, 1991.

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak, *Ramanujan graphs*, Combinatorica **8** (1988), no. 3, 261–277.

[LU95] Felix Lazebnik and Vasiliy A Ustimenko, *Explicit construction of graphs with an arbitrary large girth and of large size*, Discrete Applied Mathematics **60** (1995), no. 1, 275–284.

[Mar82] Gregori A. Margulis, *Explicit constructions of graphs without short cycles and low density codes*, Combinatorica **2** (1982), no. 1, 71–78.

[Mar88] ———, *Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators*, Probl. Peredachi Inf. **24** (1988), no. 1, 39–46.

[Mor94] Moshe Morgenstern, *Existence and explicit constructions of q+1 regular Ramanujan graphs for every prime power q*, Journal of Combinatorial Theory, Series B **62** (1994), no. 1, 44–62.

[NV08] Assaf Naor and Jacques Verstraëte, *Parity check matrices and product representations of squares*, Combinatorica **28** (2008), no. 2, 163–185.

[RS76] Imre Z. Ruzsa and Endre Szemerédi, *Triple systems with no six points carrying three triangles*, Combinatorics, Proc. Colloq. Math. Soc. János Bolyai **2** (1976), no. 18, 939–945.

[WS11] David P. Williamson and David B. Shmoys, *The design of approximation algorithms*, 1st ed., Cambridge University Press, New York, NY, USA, 2011.

A Appendix: Proof of Theorem 3.1

Theorem A.1 (restatement of Theorem 3.1)**.** *Let $\ell > 1$ be an integer and $c, \epsilon > 0$. There is a constant $C = C(\ell, \epsilon, c)$, such that every graph $G(V, E)$ on n vertices with average degree d satisfying $\ell - \epsilon \leq d \leq c \cdot n^{(\ell-2)/\ell}$, contains a subgraph of size at most $C \cdot n/d^{\ell/(\ell-2)}$ with average degree $\geq \ell - \epsilon$.*

Proof. By Lemma 2.4 we assume, up to a slight variation of constants, that G has minimum degree d. Moreover it is enough to prove the theorem for all sufficiently large values of d, and the lower values can then be handled by a proper choice of constant C. We use $o(1)$ to denote a term that tends to 0 as d grows.

For each vertex, we fix an arbitrary subset of exactly d of its neighbours, and refer only to them as its neighbours.

We first assume that ℓ is even, $\ell = 2h$ for an integer $h \geq 2$. We handle two separate cases, according to the range of the density d.

Case I - Low density: Suppose $d \leq n^{(h-1)/1.99h}$. The proof in this case is very similar to that of Theorem 3.2. Let α be a large constant that will be determined later. Sample a random subset $A \subset V$ by including each vertex in A with independent probability $p = \alpha/d^{h/(h-1)}$. We refer to vertices in A as *marked*. Note that $|A|$ is binomially distributed with parameters n, p, and that $\mathbb{E}|A| = \alpha n/d^{h/(h-1)}$.

For each vertex v we fix an arbitrary subset $N(v)$ of exactly d of its neighbours. Define B to be the random subset of vertices v that are not marked, and have exactly ℓ marked neighbours in $N(v)$. We then have,

$$\Pr[v \in B] = (1-p) \cdot \binom{d}{h} p^h (1-p)^{d-h} \geq \tfrac{(1-o(1))}{h^h}(dp)^h = \tfrac{(1-o(1))}{h^h}\alpha^h d^{-h/(h-1)}$$

where the inequality is by the known bound $\binom{a}{b} \geq (\frac{a}{b})^b$ and by observing that $1 - p = 1 - o(1)$. By linearity of expectation we get $\mathbb{E}|B| = \frac{(1-o(1))}{h^h}\alpha^h n/d^{h/(h-1)}$.

By the Chernoff bound Lemma 2.3 applied to $|A|$, we get:

$$\Pr\left[|A| < 2\mathbb{E}|A|\right] \geq 1 - (0.25e)^{\mathbb{E}|A|} = 1 - (0.25e)^{\alpha n/d^{h/(h-1)}} \geq 1 - (0.25e)^{\alpha d^{0.99h/(h-1)}}$$

where the final inequality is by the assumption of the current case, $d \leq n^{(h-1)/1.99h}$. On the other hand, by Lemma 2.1, $|B|$ attains half its expected value with probability at least $\frac{\mathbb{E}|B|}{2n} = \frac{(1-o(1))}{h^h}\alpha^h d^{-h/(h-1)} = \Omega(d^{-h/(h-1)})$. Summing the bounds yields:

$$\Pr\left[|A| \leq 2\mathbb{E}|A|\right] + \Pr\left[|B| \geq \tfrac{1}{2}\mathbb{E}|B|\right] \geq 1 - (0.25e)^{\alpha d^{0.99h/(h-1)}} + \Omega(d^{-h/(h-1)})$$

The second term in the above right-hand side shrinks exponentially in d, whereas the third term shrinks polynomially. Hence for sufficiently large d the above right-hand side is strictly more than 1, and hence there is a positive probability that both of the events $|A| \leq 2\mathbb{E}[|A|]$ and $|B| \geq \frac{1}{2}\mathbb{E}[|B|]$ occur. We fix this event from now on, and arbitrarily remove vertices from B until $|B| = \frac{1}{2}\mathbb{E}[|B|]$. The following bounds now hold:

$$|A| + |B| \leq 2\mathbb{E}|A| + \frac{1}{2}\mathbb{E}|B| = \left(2\alpha + \tfrac{(1-o(1))}{h^h}\alpha^h\right) \cdot \frac{n}{d^{h/(h-1)}} \tag{A.1}$$

$$\frac{|B|}{|A|} \geq \frac{\frac{1}{2}\mathbb{E}|B|}{2\mathbb{E}|A|} = \frac{1-o(1)}{4h^h} \cdot \alpha^{h-1} \tag{A.2}$$

We take our target subgraph H to be that induced by $A \cup B$. By eq. (A.1), its size is bounded by $C \cdot n/d^{h/(h-1)}$, which equals $C \cdot n/d^{\ell/(\ell-2)}$, for $C = 2\alpha + \frac{(1-o(1))}{h^h}\alpha^h$. To bound its average degree, note that each vertex in B is incident to h edges connecting it to A, and since A and B are disjoint (recall that vertices in B are not marked), each such edge has a unique end in B. Hence we count at least $h|B|$ different edges in H, and find that its average degree is:

$$\mathrm{avgdeg}(H) \geq 2 \cdot \frac{h|B|}{|A| + |B|} = 2h - \frac{2h}{1 + \frac{|B|}{|A|}} \geq 2h - \frac{2h}{1 + \frac{1-o(1)}{4h^h} \cdot \alpha^{h-1}}$$

using eq. (A.2) for the final inequality. The bound on the right-hand side is guaranteed to be at least $2h - \epsilon$ (which equals $\ell - \epsilon$) as long as we pick α such that $\alpha^{h-1} > 16h^{h+1}/\epsilon$, and the proof for this case is complete.

Case II - High density: Suppose $d > n^{(h-1)/1.99h}$. In this case, we choose the subset A uniformly at random over all subsets of V with size exactly $a = \alpha n/d^{h/(h-1)}$. (Again, $\alpha > 0$ is a large constant that will be set later.) Again we refer to vertices in A as *marked*, and define B similarly to the previous case, as the subset of non-marked vertices with exactly h marked neighbours. We now lower-bound the probability of a vertex $v \in V$ to be in B:

$$\Pr[v \in B] = \frac{\binom{d}{h}\binom{n-d-1}{a-h}}{\binom{n}{a}} \geq \frac{\left(\frac{d}{h}\right)^h \cdot \frac{(n-d-1)!}{(a-h)!(n-d-1-a+h)!}}{\frac{n!}{a!(n-a)!}} = \frac{\left(\frac{d}{h}\right)^h \cdot \frac{a!}{(a-h)!} \cdot \frac{(n-d-1)!}{(n-d-1-a+h)!}}{\frac{n!}{(n-a)!}}$$

Next we apply the inqeualities $(m-k+1)^k \leq \frac{m!}{(m-k)!} \leq m^k$ that holds for all positive integers m, k, and obtain:

$$\Pr[v \in B] \geq \frac{\left(\frac{d}{h}\right)^h \cdot (a-h+1)^h \cdot (n-d-a+h)^{a-h}}{n^a} \geq \frac{\left(\frac{d}{h}\right)^h \cdot \left(\frac{1}{2}a\right)^h \cdot (n-d-a+h)^{a-h}}{n^a}$$

the second inequality is because a tends to infinity with d, so for sufficiently large d we have $a - h + 1 \geq \frac{1}{2}a$. (We remark that this is where the constant c from the statement of the theorem comes into play: the growth rate of a depends on c.) We can now rearrange and write:

$$\Pr[v \in B] \geq \left(\frac{ad}{2hn}\right)^h \cdot \left(\frac{n-d-a+h}{n}\right)^{a-h} = \left(\frac{ad}{2hn}\right)^h \cdot \left(1 - \frac{d+a-h}{n}\right)^{a-h}$$

We claim that the term $\left(1 - \frac{d+a-h}{n}\right)^{a-h}$ is $1 - o(1)$. For this it is sufficient to show that $\frac{d+a-h}{n} \cdot (a-h) = o(1)$, and since h is constant, this is equivalent to showing that $\frac{d+a}{n} \cdot a = o(1)$. Indeed, recalling that $a = \alpha n/d^{h/(h-1)}$,

$$\frac{d+a}{n} \cdot a = \frac{da}{n} + \frac{a^2}{n} = \frac{\alpha}{d^{1/(h-1)}} + \frac{\alpha^2 n}{d^{2h/(h-1)}} = o(1) + o(1) = o(1)$$

where we have used the assumption of the current case, that $n < d^{1.99h/(h-1)}$. We conclude,

$$\Pr[v \in B] \geq (1 - o(1))\left(\frac{ad}{2hn}\right)^h = \frac{1-o(1)}{(2h)^h} \cdot \alpha^h \cdot \frac{1}{d^{h/(h-1)}}$$

and from this point the proof proceeds as in the low density case. This concludes the proof for the case ℓ is even.

Odd values of ℓ. The proof in this case is a close variant of the above proof for even values of ℓ, so we only sketch the differences. Suppose $\ell = 2h+1$. Again we mark each vertex with independent probability $p = \alpha/d^{\ell/(\ell-2)}$, so the subset A of marked vertices has the "correct" expected size, $\mathbb{E}|A| = \alpha n/d^{\ell/(\ell-2)}$.

The difference is that we define B to be the subset of *edges* that each of their two endpoints is non-marked, and has exactly h marked neighbours. As in the even-ℓ case, the probability for a vertex to be non-marked and to have h marked neighbours is roughly $q = (dp)^h$. We claim that in the current setting, each edge has probability roughly q^2 to be in B.

This is a subtle but technical point. Fix an edge $e = uv$, let X_u denote the event that u is non-marked and has h marked neighbours, and similarly define X_v for v. As stated above, each of X_u and X_v occurs with probability q, and we wish to show that both occur (which means e is included in B) with probability $\Omega(q^2)$. In the low density case random model, where the vertices are marked independently, the events X_u, X_v are either positively correlated (if u, v have any mutual neighbours) or independent (otherwise), so the probability for both to occur is indeed at least q^2. In the high density case random model, where we pick a random subset of marked vertices with fixed size, the events X_u, X_v may in fact be negatively correlated, but it is a technicality to calculate the probability for both X_u, X_v to occur and to see that it remains approximately q^2.

Having established that each edge in included in B with probability about q^2, we recall that there are $\frac{1}{2}nd$ edges in G, and hence the expected size of B is roughly (recall that $p = \alpha/d^{\ell/(\ell-2)} = \alpha/d^{(2h+1)/(2h-1)}$):

$$\mathbb{E}|B| = nd \cdot q^2 = nd(dp)^{2h} = \alpha^{2h} n/d^{(2h+1)/(2h-1)} = \alpha^{2h} n/d^{\ell/(\ell-2)}$$

which is also the "correct" expected size. The target subgraph H is taken to be the one induced by A and all the endpoints of edges in B. The proof then proceeds as in the even-ℓ case. Full details are omitted. $\square$

B Appendix: Upper Bound on $A_\ell(n,d)$

In this appendix we prove the extension of Theorem 1.6 for general even ℓ, in full detail. It may be beneficial as a simplified, self-contained proof of Theorem 1.6, outside the general setting that allowed the extension to Theorem 1.7.

Theorem B.1. *Let $\ell \geq 4$ be an even integer, $\ell = 2h$, and $\epsilon > 0$. There is a constant $C = C(\ell,\epsilon)$ such that every graph on n vertices with average degree d (satisfying $\ell \leq d \leq O(n^{(\ell-2)/\ell})$) contains a subgraph of size at most $C \cdot n/d^{\frac{\ell}{\ell-2}-\frac{8}{\ell^3-4\ell+8}-\epsilon}$ with average degree ℓ.*

Proof. Let $G(V,E)$ be a graph as in the statement. We prove the theorem for sufficiently large d, and the lower values can then be handled by a proper choice of constant C. We will use $o(1)$ to denote a term that tends to 0 as d grows.

By Lemma 2.4 we may assume that G has minimal degree $\geq \frac{1}{2}d$. Now by Lemma 2.6, we can orient the edges in G such that each vertex has at least $\frac{1}{4}d$ edges oriented towards it. For each vertex v, we then fix an arbitrary subset of exactly $\frac{1}{4}d$ edges oriented towards v, and call their other endpoints the *in-neighbours* of v.

Set $p = 1/d^{h/(h-1)-\epsilon}$. We mark the vertices of G in two independent phases: In the first phase, each vertex is marked with independent probability p, and in the second phase, each non-marked vertex is again marked with independent probability p. Let A be the subset of marked vertices; we have $\mathbb{E}[|A|] = (2p-p^2)n$.

Definition B.2. *Let $k \geq 1$ be an integer, and consider a fixed marking of the vertices in G. A vertex $v \in V$ is a k-*root *if it is non-marked and has exactly k marked in-neighbours.*

Proposition B.3. *After the first phase of marking vertices in G, each $u \in V$ has probability $\geq \gamma_h \cdot (dp)^{h-1}$ to be an $(h-1)$-root, where γ_h is a constant that depends only on h.*

Proof. The number of marked in-neighbours of u is binomially distributed with parameters $\frac{1}{4}d$ and p, so its probability to be an $(h-1)$-root is at least:

$$(1-p)\cdot\binom{\frac{1}{4}d}{h-1}p^{h-1}(1-p)^{\frac{1}{4}d-(h-1)} \geq \gamma_h \cdot (dp)^{h-1}$$

for $\gamma_h \geq \frac{1}{2}(4(h-1))^{-(h-1)}$. Note that the leading $(1-p)$ in the above is the probability for u to be non-marked. For the lower bound, we have used the known bound $\binom{m}{k} \geq (\frac{m}{k})^k$, and $(1-p)^{\frac{1}{4}d-(h-1)+1} \geq \frac{1}{2}$, which holds for sufficiently large d as we recall $p = 1/d^{h/(h-1)-\epsilon}$. $\square$

Lemma B.4 (main)**.** *Let $c > 0$ be any constant. If $\epsilon > 1/(h^3-2h+1)$, then for d sufficiently large, each $v \in V$ has probability $\geq cp$ to be non-marked and to have at least $h+1$ in-neighbours which are h-roots.*

Proof. Fix $v \in V$ and let $N(v)$ be the set of in-neighbours of v. (Recall that $|N(v)| = \frac{1}{4}d$.) Consider G after the first phase of marking vertices. Suppose $u \in N(v)$ is an $(h-1)$-root. We call an edge e oriented towards u an *excited* edge, if its source vertex is a non-marked in-neighbour of u. Recall that if u is an $(h-1)$-root then it has exactly $h-1$ marked in-neighbours, so at least $\frac{1}{4}d-(h-1) \geq \frac{1}{8}d$ non-marked in-neighbours. This means u renders at least $\frac{1}{8}d$ edges excited, and for simplicity, we assume henceforth that it renders *exactly* $\frac{1}{8}d$ edges excited (arbitrarily chosen).

Let X be the set of $(h-1)$-roots in $N(v)$ (after only the first phase of marking vertices), and let Y be the set of excited edges. By Proposition B.3, each $u \in N(v)$ has probability $\geq \gamma_h (pd)^{h-1}$ to be an $(h-1)$-root and hence,

$$\mathbb{E}|X| \geq \tfrac{1}{4} d \cdot \gamma_h (pd)^{h-1} = \tfrac{1}{4} \gamma_h d^h p^{h-1}$$

and since each vertex in X renders $\frac{1}{8}d$ edges excited,

$$\mathbb{E}|Y| \geq \tfrac{1}{8} d \cdot \mathbb{E}|X| \geq \tfrac{1}{32} \gamma_h d^{h+1} p^{h-1} \tag{B.1}$$

Applying Lemma 2.2 we get that for some $t \geq 1$,

$$\Pr[|Y| \geq \tfrac{2^t}{9 \cdot 16} \gamma_h d^{h+1} p^{h-1}] \geq (2^t t^2)^{-1} \tag{B.2}$$

Proposition B.5. $t < 2 \log d$.

Proof. Recall that $p = d^{h/(h-1)-\epsilon}$ and hence,

$$(dp)^{h-1} = d^{1-(h-1)\epsilon} > d^{-1} \tag{B.3}$$

Now observe that there can be at most $\frac{1}{4}d$ excited edges per vertex in $N(v)$, so a total of at most $\frac{1}{16}d^2$ excited edges. Hence, when applying Lemma 2.2 to derive eq. (B.2), we can set $M = \frac{1}{16}d^2$ in the statement of Lemma 2.2 and obtain:

$$t \leq \log\left(\frac{9M}{2\mathbb{E}|Y|}\right) \leq \log\left(\frac{9}{\gamma_h (dp)^{h-1}}\right) < \log\left(\frac{9}{\gamma_h} \cdot d\right) < 2 \log d$$

where the first inequality is by the guarantee of Lemma 2.2; the second one is by plugging eq. (B.1) for $\mathbb{E}|Y|$ and $M = \frac{1}{16}d^2$; the next one is by eq. (B.3); and the final inequality holds for sufficiently large d, since γ_h is constant. $\square$

We proceed with the proof of the Main Lemma B.4. By eq. (B.2), after the first phase, there is probability $(2^t t^2)^{-1}$ to have $\frac{2^t}{9 \cdot 16} d^{h+1} p^{h-1}$ excited edges. Let W be the set of source vertices of all the excited edges. We recall that by definition (of an excited edge), each vertex in W is a non-marked in-neighbour of a vertex in X.

For each $w \in W$, we say w is *light* if it is the source of at most h excited edges, and *heavy* if it is the source of at least $h+1$ excited edges. For a fixed marking of the vertices in G after the first phase, either half the excited edges are sourced at light vertices, or half are sourced at heavy vertices. By an averaging argument applied to eq. (B.2), we see that one of these cases must hold with probability $\frac{1}{2}(2^t t^2)^{-1}$. We handle the two cases separately.

Case I - Light vertices: With probability $\frac{1}{2}(2^t t^2)^{-1}$ we have $\frac{2^t \gamma_h}{9 \cdot 16} d^{h+1} p^{h-1}$ excited edges, and half of them are sourced at light vertices of W.

Let L be the subset of light vertices. We have $\frac{2^t \gamma_h}{9 \cdot 32} d^{h+1} p^{h-1}$ excited edges incident to light vertices, and each light vertex is the source of at most h excited edges. Hence, $|L| \geq \frac{2^t \gamma_h}{9 \cdot 32 \cdot h} d^{h+1} p^{h-1}$. Arbitrarily remove vertices from L until equality holds. Moreover, the number of excited edges is $\frac{1}{8}d \cdot |X|$, and hence we get $|X| \geq \frac{2^t \gamma_h}{36} d^h p^{h-1}$.

Our intention is now to uniquely assign L-vertices to X-vertices. To this end we consider the bipartite graph with sides X and L and with the excited edges as the edge set. In fact, X and L may intersect; in such case we make two copies of each vertex in the intersection, putting one copy on the X-side and the other on the L-side. Note that all the edges are oriented from L to X.

Side X has maximum degree $\frac{1}{8}d$ and average degree:

$$\frac{|L|}{|X|} \geq \frac{\frac{2^t\gamma_h}{9\cdot 32\cdot h}d^{h+1}p^{h-1}}{\frac{2^t\gamma_h}{36}d^hp^{h-1}} = \frac{2^t}{36\cdot h}d \geq \frac{1}{36\cdot h}d$$

Hence by Lemma 2.7, there is a subset $X' \subset X$ with size $|X'| \geq \frac{1}{4.5h}|X|$ such that each vertex in X' is adjacent to at least $\frac{1}{72h}d$ vertices in L. Now consider the bipartite graph with sides X' and L: Side X' has degree (at least) $\frac{1}{72h}d$, and side L has degree at most h (since a light vertex is adjacent to at most h vertices in X, and hence in X'). We use following lemma:

Lemma B.6. *Let $G(V,U;E)$ be a bipartite graph such that each $v \in V$ has degree d, and each $u \in U$ has degree at most h. There is a subset of edges E' such that in $G'(V,U;E')$, each $v \in V$ has degree at least $\lfloor\frac{1}{h}d\rfloor$, and each $u \in U$ has degree at most 1.*

Proof. Similar to the proof of Lemma 2.5, with every "2" replaced by "h". □

By Lemma B.6, there is an assignment of L-vertices to X'-vertices such that each L-vertex is uniquely assigned, and each X'-vertex has at least $\frac{1}{72h^2}d$ vertices assigned to it. Fixed this assignment henceforth.

Let $u \in X'$. Recall that u is in X, which means it is an $(h-1)$-roots after the first phase. Moreover by the above it has $\frac{1}{16h^2}d$ neighbours in L assigned to it. If one of them is marked in the second phase, then u turns into an h-root. Denote this event by A_u. Its probability is $(1-o(1))\cdot\frac{1}{72h^2}d\cdot p$, for picking an assigned neighbour and marking it.

Let $\mathcal{A}$ denote the event that $h+1$ of the events $\{A_u : u \in X'\}$ occur concurrently after the second phase. If $\mathcal{A}$ occurs, then v has $h+1$ neighbours which are h-roots, and the conclusion of Lemma B.4 (which we are now proving) is satisfied. To lower bound its probability, we observe that the events $\{A_u : u \in X'\}$ are pairwise independent, by the uniqueness of the assignment of L-vertices to X'-vertices. Hence (recalling that the probability for the current case is $\frac{1}{2}(2^tt^2)^{-1}$):

$$\Pr[\mathcal{A}] \geq \frac{(1-o(1))}{2\cdot 2^tt^2}\binom{|X'|}{h+1}\left(\frac{1}{72h^2}dp\right)^{h+1} \geq \frac{(1-o(1))}{2\cdot 2^tt^2}\cdot\left(\frac{|X'|}{h+1}\cdot\frac{1}{72h^2}dp\right)^{h+1} \tag{B.4}$$

where we have used the known bound $\binom{a}{b} \geq (\frac{a}{b})^b$. We recall:

$$|X'| \geq \frac{1}{4.5h}|X| \geq \frac{2^t\gamma_h}{162h}d^hp^{h-1}$$

Plugging this back into eq. (B.4), we get:

$$\Pr[\mathcal{A}] \geq \frac{(1-o(1))}{2\cdot 2^tt^2}\cdot\left(\frac{\frac{2^t\gamma_h}{162h}}{h+1}\cdot\frac{1}{72h^2}\cdot d^{h+1}p^h\right)^{h+1}$$

Now applying Proposition B.5 and suppressing all the constants into a constant C_h, we get:

$$\Pr[\mathcal{A}] \geq \frac{C_h}{\log^2 d}\cdot d^{h^2+2h+1}p^{h^2+h}$$

We need this bound to be at least cp. Rearranging and suppressing c into C_h, we need:

$$\frac{C_h}{\log^2 d}\cdot d^{h^2+2h+1}p^{h^2+h-1} \geq 1$$

We recall that $p = d^{-h/(h-1)+\epsilon}$ and plug this into the above, which then becomes:

$$\frac{C_h}{\log^2 d} \cdot d^{\epsilon(h^3-2h+1)-1} \geq 1$$

and this is satisfied as long as $\epsilon > 1/(h^3 - 2h + 1)$, which holds by hypothesis (of the current Lemma B.4). The proof for this case is complete.

Case II - Heavy vertices: With probability $\frac{1}{2}(2^t t^2)^{-1}$ we have $\frac{2^t \gamma_h}{9 \cdot 16} d^{h+1} p^{h-1}$ excited edges, and half of them are sourced at heavy vertices of W.

A heavy vertex is the source of at least $h+1$ excited edges. Each such edge is oriented towards a vertex in X, which we recall is an $(h-1)$-root. Hence marking a single heavy vertex in the second phase turns $h+1$ neighbours of v into h-roots, as the lemma requires.

We have $\frac{2^t \gamma_h}{9 \cdot 32} d^{h+1} p^{h-1}$ excited edges incident to heavy vertices. Each heavy vertex is the source of at most $\frac{1}{4}d$ excited edges (since an excited edge is oriented towards a vertex in $X \subset N(v)$, and $|N(v)| = \frac{1}{4}d$). Hence, the number of heavy vertices is at least:

$$\frac{\frac{2^t \gamma_h}{9 \cdot 32} d^{h+1} p^{h-1}}{\frac{1}{4}d} = \frac{2^t \gamma_h}{72} d^h p^{h-1}$$

so the probability to mark one in the second phase is at least:

$$\frac{1}{2}(2^t t^2)^{-1} \cdot (1 - o(1))p \cdot \frac{2^t \gamma_h}{72} d^h p^{h-1} = \frac{1 - o(1)}{2t^2} \cdot \frac{\gamma_h}{72}(dp)^h \geq \frac{1 - o(1)}{16 \log^2 d} \cdot \frac{\gamma_h}{72}(dp)^h$$

using $t < 2 \log d$ by Proposition B.5 for the last inequality. To prove the lemma, we need this lower bound on the probability to be at least cp. Rearranging, we need:

$$p \geq \frac{16 \cdot 72 \cdot c \cdot \log^2 d}{(1 - o(1))\gamma_h} \cdot d^{-h/(h-1)}$$

By recalling that $p = d^{-h/(h-1)+\epsilon}$, we see that the latter inequality indeed holds for sufficiently large d, which proves the current case.

Concluding the proof of Lemma B.4. Having handled both the light and heavy cases, we have proven that v has probability cp to have $h+1$ in-neighbours which are h-roots. We further need it to be non-marked, which happens with probability $(1-p)^2$ (over the two phases), independently of the marking of any other vertices. Hence the probability that v satisfies the conclusion of the lemma is $(1-p)^2 cp = (1 - o(1))cp$. The lemma is proven by slightly rescaling c. □

Back to the proof of Theorem B.1. Suppose $\epsilon > 1/(h^3 - 2h + 1)$. Consider the graph after both phases of marking vertices. Recall that A is the subset of marked vertices. Denote by B the subset of vertices that are non-marked, and have $h+1$ in-neighbours which are h-roots. Note that by Lemma B.4, each vertex is included in B with probability $\geq cp$, for a constant $c > 0$ of our choice (that will be set later).

We recall that $\mathbb{E}|A| = (2p - p^2)n \geq pn$, that $p = d^{-h/(h-1)+\epsilon}$, and that $n \geq \Omega(d^{h/(h-1)})$ (by hypothesis of the theorem). Together we get $\mathbb{E}|A| \geq \Omega(d^\epsilon)$. Hence by a Chernoff bound (Lemma 2.3) applied to $|A|$,

$$\Pr\left[|A| < 2\mathbb{E}|A|\right] \geq 1 - (0.25e)^{\mathbb{E}|A|} \geq 1 - (0.25e)^{\Omega(d^\epsilon)}$$

On the other hand we have $\mathbb{E}|B| = cpn$, and by Lemma 2.1:

$$\Pr\left[|B| \geq \tfrac{1}{2}\mathbb{E}|B|\right] \geq \tfrac{\mathbb{E}|B|}{2n} = \tfrac{1}{2}cp = \tfrac{1}{2}cd^{-h/(h-1)+\epsilon}$$

Summing the bounds yields:

$$\Pr\left[|A| \leq 2\mathbb{E}|A|\right] + \Pr\left[|B| \geq \tfrac{1}{2}\mathbb{E}|B|\right] \geq 1 - (0.25e)^{\Omega(d^{\epsilon})} + \tfrac{1}{2}cd^{-h/(h-1)+\epsilon}$$

The second term shrinks exponentially in d whereas the third term shrinks polynomially, and hence for sufficiently large d, the above RHS is strictly larger than 1. Therefore, there is a positive probability that both the events $|A| \leq 2\mathbb{E}|A|$ and $|B| \geq \frac{1}{2}\mathbb{E}|B|$ occur. These imply $|A| \leq 4np$ (as we recall $\mathbb{E}|A| \leq 2np$) and $|B| \geq \frac{1}{2}cnp$, respectively. We fix this event from now on, and arbitrarily remove vertices from B until $|B| = \frac{1}{2}cnp$.

Recall that each vertex in B has $h+1$ in-neighbours which are h-roots. Let Z be the subset of all the h-roots in-neighbours of vertices in B. Note that B and Z may intersect, and that $|Z| \leq (h+1)|B|$. We take our target subgraph H to be one induced by $A \cup B \cup Z$. Its size is bounded by,

$$|H| \leq |A| + |B| + |Z| \leq |A| + (h+2)|B| \leq (4 + (h+2)\tfrac{1}{2}c)np = (4 + (h+2)\tfrac{1}{2}c) \cdot n/d^{h/(h-1)-\epsilon}$$

which is as required if we set $C = 4 + (h+2)\frac{1}{2}c$. We move on to establish that H has h edges per vertex. Each $v \in Z$ is an h-root, i.e. has h marked in-neighbours. Those in-neighbours are in A and hence in H, so the edges oriented from them to v are present in H, and we count them in its favour. Each $v \in B$ has $h+1$ in-neighbours in Z, so the edges going from those neighbours into v are present in H. We count h of them in favour of v. Note that so far, each edge was counted in favour of its destination endpoint, and hence was counted only once. We are left to handle vertices in A. Note that for each vertex in B, there is one edge oriented towards it (from a vertex in Z) that we have not yet counted. Together we have $|B|$ edges yet unused, and we now count them in favour of the vertices in A. We now only need $|B| \geq h|A|$ to hold; by recalling that $|A| \leq 4np$ and $|B| = \frac{1}{2}cnp$, we achieve this by choosing $c = 8h$.

The proof is complete with graph size at most $C \cdot n/d^{h/(h-1)-\epsilon}$, under the assumption $\epsilon > 1/(h^3 - 2h + 1)$. This is equivalent to the statement of the theorem, by resetting $\epsilon' = \epsilon - 1/(h^3 - 2h + 1) = 8/(\ell^3 - 4\ell + 8)$. $\square$

C Appendix: Omitted Proofs from Section 2

Proof of Lemma 2.1. For each $i = 0, \ldots, n$ denote $p_i = \Pr[X = i]$, and let $p = \sum_{i \geq \frac{1}{2}\mu} p_i$. We have,

$$\mu = \sum_{i=0}^{n} p_i i = \sum_{i<\frac{1}{2}\mu} p_i i + \sum_{i\geq\frac{1}{2}\mu} p_i i \leq \sum_{i<\frac{1}{2}\mu} p_i \frac{1}{2}\mu + \sum_{i\geq\frac{1}{2}\mu} p_i n = \frac{1}{2}\mu(1-p) + np \leq \frac{1}{2}\mu + np$$

Rearranging gives $p \geq \frac{\mu}{2n}$. $\square$

Proof of Lemma 2.2. Suppose no t satisfies the statement, then

$$\mathbb{E}[X] \leq \frac{2\mathbb{E}[X]}{9} + \sum_{t\geq 1} \frac{2\mathbb{E}[X]}{9} 2^{t+1} \frac{2^{-t}}{t^2} = \frac{2\mathbb{E}[X]}{9}(1 + 2\sum_{t\geq 1}\frac{1}{t^2}) = \frac{2\mathbb{E}[X]}{9}(1 + 2\frac{\pi^2}{6}) < \mathbb{E}[X],$$

a contradiction. This proves the first assertion of the lemma. The second assertion follows, because a value of $\frac{2}{9} \cdot 2^t \cdot \mathbb{E}[X]$ or larger is now known to be in the support of X, and hence it cannot exceed the bound M. Rearranging $\frac{2}{9} \cdot 2^t \cdot \mathbb{E}[X] \leq M$ gives $t \leq \log(\frac{9}{2}M/\mathbb{E}[X])$. $\square$

Proof of Lemma 2.3. See for example [WS11, Theorem 5.24]. $\square$

Proof of Lemma 2.4. Let G be a graph on n vertices and average degree d, so $\frac{1}{2}dn$ edges. Iteratively, as long as there are vertices with degree $\leq \frac{1}{2}d$, pick an arbitrary one and remove it from the graph. The resulting subgraph H has minimum degree greater than $\frac{1}{2}d$ as long as it is non-empty.

Supposing by contradiction that H is empty, we have performed n iterations; in each one we have removed at most $\frac{1}{2}d$ edges, and in the last one no edges were removed (as the graph then contained only a single vertex). Altogether we have removed at most $\frac{1}{2}d(n-1)$ edges, less than the total number of edges in G, and hence there are surviving edges H, which contradicts its being empty. $\square$

Proof of Lemma 2.5. Construct an auxiliary bipartite graph $H(V_H, U; E_H)$ from G, by replacing each $v \in V$ with $\lfloor\frac{1}{2}d\rfloor$ copies, each connected to the neighbours of v in U. For $W \subset V_H$, each $w \in W$ has d outgoing edges, and each $u \in U$ has (in H) at most $2\lfloor\frac{1}{2}d\rfloor$ incoming edges, so W has neighbourhood of size at least $N(W) \geq d|W|/2\lfloor\frac{1}{2}d\rfloor \geq |W|$. Hence H satisfies the condition of Hall's Theorem, and thus has a perfect matching $E'_H \subset E_H$. Re-unify all copies of each vertex $v \in V$ into a single vertex. The resulting subgraph of G is $G'(V, U; E')$. $\square$

Proof of Lemma 2.6. Consider the bipartite *incidence graph* B_G of G which has sides V and E, and $v \in V$, $e \in E$ are adjacent in B_G iff e is incident to v in G. Apply Lemma 2.5 on B_G to get an assignment of each edge in G to at most one of its end vertices. Orient the edges according to that assignment. $\square$

Proof of Lemma 2.7. By Lemma 2.1 with $\mu = d$, $n = D$, and X the degree of a random vertex in V. $\square$