Robust sublinear expander

Hong Liu

29th July 2020

Abstract

In this note, we will cover some methods on robust sublinear expander. Based on a notion of expander introduced in the 90s by Komlós and Szemerédi, this concept has been recently developed, bringing versatile building blocks that can be found in general graphs. It has proven to be a powerful tool for embedding sparse graphs, playing an essential role in the recent resolutions of several long-standing conjectures that were previously out of reach.

Contents

1 Introduction

Expanders are typically well connected sparse graphs in which vertex subsets exhibit expansions. Originally introduced for network design, expanders, apart from being a central notion in graph theory, have also close interplay with other areas of mathematics and as well as theoretical computer science. This is partially reflected by the fact that expanders have equivalent definitions from different angles. Indeed, in terms of graph expansions, an expander is a graph whose vertex subsets have large neighbourhood; while algebraically, there is a spectral gap when you look at its adjacency matrix; and probabilistically, random walks on expanders are rapidly mixing.

Expanders studied so far usually have constant expansion. In this note, we will discuss yet another variant of expander, first introduced by Komlós and Szemerédi in the 90s [3, 4]. This notion of expander, albeit having weak (sublinear) expansion, has been a key part of recent advancements on sparse embedding problems. In particular, building on this notion, several versatile building blocks have been introduced recently [1, 2, 5, 6], and methods around these structures developed settle many open problems on sparse embeddings that have been elusive for decades.

We will go over:

- the motivation of the building blocks introduced;
- how we can find such substructures in general graphs;
- how we put these "lego pieces" together to embed sparse graphs.

Below is a list of material that will be covered in this note. Our emphasis will *not* be to give an overview of the proof, but rather to talk about various useful properties of this expander, and illustrate, through taking bits here and there from [1]–[6], the use of these properties in different problems on embedding sparse graphs.

○ Introduce sublinear expander and prove that in any general graph, there is a subexpander with almost the same average degree. Then we will see its original application: the theorem of Komlós-Szemerédi stating that average degree d forces $K_{\Omega(\sqrt{d})}$-subdivision [3, 4].

♠ Introduce *unit/web* structure and illustrate how to handle expanders with medium density. For this, we will consider Komlós's conjecture (1981) which states that among all graphs with given minimum degree, complete graphs minimise the number of Hamiltonian subsets [2].

♡ Show a basic max-degree reduction. For this, we look at Mader's conjecture (1999) that C_4-free graphs with average degree d contains $K_{\Omega(d)}$-subdivision [5].

♣ Present an approach for working in almost regular sparse expanders. For this, we shall see that $(\frac{3}{2}+o(1))d$ average degree forces any d-vertex planar graph as a minor, and the constant $\frac{3}{2}$ is optimal [1]. Then introduce *nakji* structure and outline a strategy to reduced problems to almost regular graphs.

♢ Introduce *adjuster* structure and briefly explain the basic idea behind the solutions of

- ⋆ Erdős-Hajnal's problem (1966) that in graphs with infinite chromatic number, the harmonic sum of odd cycle lengths is infinite; and
- ⋆ \$1000 problems of Erdős on unavoidable cycle length in graphs with large (but constant) average degree [6].

2 Preliminaries

For $n \in \mathbb{N}$, let $[n] := \{1, \dots, n\}$. Given a set X and $k \in \mathbb{N}$, let $\binom{X}{k}$ the family of all k-sets in X. For brevity, we write v for a singleton set $\{v\}$ and xy for a set of pairs $\{x, y\}$. We write $a = b \pm c$ if $b - c \le a \le b + c$. If we claim that a result holds whenever we have $0 < a \ll b, c \ll d < 1$, it means that there exist positive functions f, g such that the result holds as long as $a < f(b, c)$ and $b < g(d)$ and $c < g(d)$. We will not compute these functions explicitly. In many cases, we treat large numbers as if they are integers, by omitting floors and ceilings if it does not affect the argument. We write log for the base-e logarithm.

A graph H is a *minor* of G, denoted by $H \prec G$, if H can be obtained from G by vertex/edge deletions and edge contractions. Here *contracting* an edge uv in G means identifying u and v to a new vertex w and setting the neighbourhood of w to be the union of u, v's neighbourhoods. Equivalently, $H \prec G$ if there are pairwise vertex disjoint sets $V_x \subseteq V(G)$, indexed by $x \in V(H)$, such that the subgraph induced $G[V_x]$ is connected for each $x \in V(H)$ and for any $xy \in E(H)$, there is a V_x, V_y-edge in G.

A special kind of minor is that of *topological minor* or *subdivision*. A subdivision of H is obtained from replacing each edge of H by pairwise internally vertex disjoint paths. This notion connects topology and graph theory. The well-known Kuratowski's theorem (1930) states that a graph is planar if and only if it does not contain K_5 or $K_{3,3}$ as subdivisions.

Given graphs H and G, in a copy of H-subdivision in G, we call the vertices that correspond to $V(H)$ the *anchor* vertices of the subdivision. For a given path $P = x_1 \dots x_t$, we write $\mathsf{Int}(P) = \{x_2, \dots, x_{t-1}\}$ to denote the set of its internal vertices. Given a graph H, a set of vertices $S \subseteq V(H)$ and a subgraph $F \subseteq H$, denote by $H - S = H[V(H) \setminus S]$ the subgraph induced on $V(H) \setminus S$ and by $H \setminus F$ the spanning subgraph obtained from H by removing edges in F.

Given a graph G, denote its average degree $2e(G)/|G|$ by $d(G)$. For two sets $X, Y \subseteq V(G)$, the (graph) *distance* between them is the length of a shortest path from X to Y. For two graphs G, H, we write $G \cup H$ to denote the graph with vertex set $V(G) \cup V(H)$ and the edge set $E(G) \cup E(H)$. A *k-star* denotes a copy of $K_{1,k}$ which is a star with k edges. Given a collection of graphs $\mathcal{F} = \{F_i : i \in I\}$, we write $V(\mathcal{F}) = \bigcup_{i \in I} V(F_i)$ and $|\mathcal{F}| = |I|$. For path P and a vertex set U, we write $P|_U$ for the induced subgraph of P on vertex set $V(P) \cap U$.

For a set of vertices $X \subseteq V(G)$ and $i \in \mathbb{N}$, denote

$$N^i(X) := \{u \in V(G) : \text{ the distance between } X \text{ and } u \text{ is exactly } i\}$$

the i-th sphere/layer around X, and write $N^0(X) = X$, $N(X) := N^1(X)$, and for $i \in \mathbb{N} \cup \{0\}$, let $B^i(X) = \bigcup_{j=0}^{i} N^j(X)$ be the ball of radius i around X. We write $\partial(X)$ for the edge-boundary of X, that is, the set of edges between X and $V(G) \setminus X$ in G. Given another set $Z \subseteq V(G)$, we write $N(X, Z) = N(X) \cap Z$ for the set of neighbours of X in Z.

3 Robust sublinear expander

To define the robust graph expansion, we need the following function. For $\varepsilon_1 > 0$ and $t > 0$, let $\rho(x)$ be the function

$$\rho(x) = \rho(x, \varepsilon_1, t) := \begin{cases} 0 & \text{if } x < t/5, \\ \varepsilon_1 / \log^2(15x/t) & \text{if } x \ge t/5, \end{cases} \tag{1}$$

where, when it is clear from context we will not write the dependency on ε_1 and t of $\rho(x)$. Note that when $x \ge t/2$, while $\rho(x)$ is decreasing, $\rho(x) \cdot x$ is increasing.

Komlós and Szemerédi [3, 4] introduced a notion of expander G in which any set X of reasonable size expands by a sublinear factor, that is, $|N_G(X)| \ge \rho(|X|)|X|$. We shall extend this notion to a robust one such that similar expansion occurs even after removing a relatively small set of edges.

Definition 3.1. (ε_1, t)**-robust-expander:** A graph G is an (ε_1, t)-*robust-expander* if

- for any subset $X \subseteq V(G)$ of size $t/2 \leq |X| \leq |G|/2$,
- and any subgraph $F \subseteq G$ with $e(F) \leq d(G) \cdot \rho(|X|)|X|$,

where $\rho(\cdot)$ is as in (1), then we have

$$|N_{G \setminus F}(X)| \geq \rho(|X|)|X|.$$

Recall that when $x \geq t/2$, while $\rho(x)$ is decreasing as x increases, $\rho(x) \cdot x$ is increasing. Therefore, though the expansion (relative to the size of the set X) becomes weaker as the size of the set X increases, the size of each layer while expanding, i.e. $N^k_G(X)$, is increasing in size.

One imperative feature of this (weak) sublinear expander is that any graph contains one such expander subgraph which, furthermore, is almost as dense as the original graph, as shown by Komlós and Szemerédi [3, 4].

Theorem 3.2 ([3]Theorem 2.2)**.** *There exists some $\varepsilon_1 > 0$ such that the following holds for every $t > 0$. Every graph G has an (ε_1, t)-expander subgraph H with $d(H) \geq d(G)/2$ and $\delta(H) \geq d(H)/2$.*

Remark 3.3. It is a good time now to point out some pros and cons.

- First of all, though almost retaining the average degree, the expander subgraph H could be much smaller than G in order. For instance, if G is a union of many vertex disjoint small cliques, then H could be just one of those cliques.

  Such drawback often makes it difficult to utilise expanders iteratively within graphs. If the expander subgraph H shrinks too much, then it might not inherit some useful properties of G. For example, as we shall see later, it is often time useful to have a 'bounded' maximum degree condition, say $\Delta(G) = (\log |G|)^{O(1)}$, which could mean nothing to H if $|H| < \log |G|$.
- Taking $\rho(x) \sim \frac{1}{\log^2 x}$ is for a clean presentation. We are allowed to take this sublinear expansion factor to be $\frac{1}{\log x \cdot (\log \log x)^{O(1)}}$. This is almost optimal, as beyond this point, it is known that there are graphs without subgraph with expansion rate $\frac{(\log \log x)^{\Omega(1)}}{\log x}$.
- Here, t is free to choose. For most of the problems that we shall see later, it will be linear in the average degree $d(G)$. The trade-off for choosing larger t is the following:
  - small sets might not expand, as $\rho(x) = 0$ for $x < t/5$;
  - large sets have (slightly) better expansion, as $\rho(x, \varepsilon_1, t)$ is increasing with t.

  One nature example to choose t other than $\Theta(d(G))$ is when G is $K_{s,t}$-free, say $2 \leq s \leq t$. We can take instead $t \sim d(G)^{\frac{s}{s-1}}$ to get better expansion for large sets without losing small sets expansion. Indeed, sets up to size $d(G)^{\frac{s}{s-1}}$ expand as $\mathrm{ex}(n, K_{s,t}) = O(n^{2-1/s})$ (see e.g. Proposition 5.2 in [5]).

We shall prove the following slightly stronger version of Theorem 3.2.

Lemma 3.4. *Let $C > 30$, $\varepsilon_1 \leq 1/(10C)$, $t > 0$ and $\rho(x) = \rho(x, \varepsilon_1, t)$ as in (1). Then every graph G has an (ε_1, t)-robust-expander subgraph H with $d(H) \geq (1 - \delta)d(G)$, where $\delta := \frac{C\varepsilon_1}{\log 3}$, and $\delta(H) \geq d(H)/2$.*

3.1 Small diameter

Beforing proving Lemma 3.4, we would like to show first a key property of expanders which is the starting point of all constructions to come. That is, expanders have small (polylogarithmic) diameter.

The following lemma shows that in an expander, we can connect two sets X_1, X_2 using a short path while avoiding another set W as long as W is a bit smaller than X_1, X_2. The proof is straightforward, basically expand X_1, X_2 until they meet and W is too small to block any layer during expanding.

Lemma 3.5 ([4]Corollary 2.3)**.** *If G is an n-vertex (ε_1, t)-robust-expander, then for any two vertex sets X_1, X_2 each of size at least $x \geq t/2$, and a vertex set W of size at most $\rho(x)x/4$, there exists a path in $G - W$ between X_1 and X_2 of length at most*

$$\frac{2}{\varepsilon_1} \log^3\left(\frac{15n}{t}\right).$$

Proof. By the expansion property of G and that $\rho(x)x$ is increasing with x, we have, for each $i \in \mathbb{N}$, that

$$|N_G(B^{i-1}_{G-W}(X_1))| \geq \rho(|B^{i-1}_{G-W}(X_1)|)|B^{i-1}_{G-W}(X_1)| \geq \rho(x)x \geq 4|W|.$$

Thus, $|N_{G-W}(B^{i-1}_{G-W}(X_1))| \geq |N_G(B^{i-1}_{G-W}(X_1))| - |W| \geq \frac{1}{2}|N_G(B^{i-1}_{G-W}(X_1))|$.

Next, as $\rho(x)$ is decreasing with x when $x \geq t/2$, we see that

$$|N_{G-W}(B^{i-1}_{G-W}(X_1))| \geq \frac{1}{2}\rho(|B^{i-1}_{G-W}(X_1)|)|B^{i-1}_{G-W}(X_1)| \geq \frac{1}{2}\rho(n)|B^{i-1}_{G-W}(X_1)|.$$

That is, in $G - W$, X_1 expands each time at least a factor of $(1 + \frac{1}{2}\rho(n))$ until it gets larger than $n/2$:

$$|B^{i}_{G-W}(X_1)| \geq \left(1 + \frac{1}{2}\rho(n)\right)^i x.$$

Solving $(1 + \frac{1}{2}\rho(n))^r > n/2$, we see that for $r = \frac{1}{\varepsilon_1}\log^3(\frac{15n}{t})$, $B^r_{G-W}(X_i)$, $i \in [2]$, have size larger than $n/2$ each, hence they intersect, yielding the desired short X_1, X_2-path avoiding W. $\square$

Note that we do not require W to be disjoint from X_1, X_2 above.

Remark 3.6. Lemma 3.5 is a key ingredient in applying sublinear expander. One drawback, though, is that, apart from the connecting path being short ($O(\log^3 n)$), we have no control on the length of the path. For instance, not to mention a specified length, the proof above cannot even guarantee a path of length that is say $0 \mod 3$. We will come back to this in Section 8.

3.2 Proof of the robust expander lemma

In this subsection, we give the proof of Lemma 3.4. Recall that our goal is to find within G an expander subgraph.

The idea is the following. If G is an expander already, then we are home, so there is a set X that expands poorly in G. Then observe that, as $N(X)$ is small, we can pass from G to a smaller subgraph, either $G[X \cup N(X)]$ or $G - X$, without losing much average degree. To use this observation, define a graph functional (see $\phi(G)$ in (4)) that penalise passing to a smaller subgraph. Then, a graph maximising this functional is necessarily an expander.

We now make the above idea rigorous. First set up some functions as follows. For $C > 0$ and $x \geq 1$, define

$$\gamma(x) = C \int_x^\infty \frac{\rho(u)}{u}\, \mathrm{d}u.$$

Note that $\gamma(x)$ is a decreasing function. We will make use of the following two inequalities:

$$\gamma(1) = \gamma(t/5) = C\int_{t/5}^{\infty} \frac{\varepsilon_1}{u\log^2(15u/t)}\,\mathrm{d}u = \frac{C\varepsilon_1}{\log 3} = \delta < 1; \tag{2}$$

and, for $C' > 1$ and $x \geq t/2$, since $\rho(x)$ is decreasing and $\rho(x)\cdot x$ is increasing with x for $x \geq t/2$, we have

$$\gamma(x) - \gamma(C'x) = C\int_x^{C'x} \frac{\rho(u)}{u}\,\mathrm{d}u \geq C\rho(C'x)\int_x^{C'x} \frac{1}{u}\,\mathrm{d}u = C\log C'\cdot\rho(C'x) \geq \frac{C\log C'}{C'}\rho(x). \tag{3}$$

Now, for a graph G, define

$$\phi(G) = d(G)[1+\gamma(|G|)]. \tag{4}$$

We say that G is *ϕ-maximal* if

$$\phi(G) = \max_{H\subseteq G}\phi(H).$$

Claim 3.7. If G is ϕ-maximal, then $d(G) = \max_{H\subseteq G} d(H)$ and $\delta(G) \geq d(G)/2$.

Proof of claim. Since $\gamma(x)$ is decreasing, for any $H \subseteq G$, by the ϕ-maximality of G, we have that $d(G)[1+\gamma(|G|)] \geq d(H)[1+\gamma(|H|)]$ and consequently $d(G) \geq d(H)$.

Suppose there is a vertex v with $d(v) < d(G)/2$. Let $H := G - v$, then

$$d(H) = \frac{d(G)|G| - 2d(v)}{|G|-1} > d(G),$$

a contradiction. ■

Take $H \subseteq G$ such that H is ϕ-maximal. We will show that H is the desired robust expander. By Claim 3.7, $\delta(H) \geq d(H)/2$. Since H is ϕ-maximal and that $\gamma(x)$ is decreasing, we have

$$d(H) \geq \frac{d(G)(1+\gamma(|G|))}{1+\gamma(|H|)} \geq \frac{d(G)}{1+\gamma(1)} \geq (1-\gamma(1))d(G) \overset{(2)}{=} (1-\delta)d(G).$$

We are left to show that H has the claimed expansions.

Note that, since H is ϕ-maximal, for any $K \subseteq H$,

$$d(K) \leq \frac{1+\gamma(|H|)}{1+\gamma(|K|)}\cdot d(H) \leq d(H). \tag{5}$$

Fix an arbitrary $X \subseteq V(H)$ with $t/2 \leq |X| \leq |H|/2$,[1] and an arbitrary subgraph $F \subseteq H$ with $e(F) \leq d(H)\rho(|X|)|X|$. Let $Y = X \cup N_{H\setminus F}(X)$ and $\overline{X} = V(H)\setminus X$. Then,

$$\begin{aligned} d(H)(|X|+|\overline{X}|) = 2e(H) &\leq 2e(H[Y]) + 2e(H[\overline{X}]) + 2e(F) \\ &\overset{(5)}{\leq} d(H[Y])|Y| + d(H)|\overline{X}| + 2d(H)\cdot\rho(|X|)\cdot|X|. \end{aligned}$$

Consequently, we have

$$d(H)|X| \leq d(H[Y])|Y| + 2d(H)\cdot\rho(|Y|)\cdot|Y| \overset{(5)}{\leq} \left(\frac{1+\gamma(|H|)}{1+\gamma(|Y|)} + 2\rho(|Y|)\right) d(H)|Y|,$$

and so

$$\frac{|X|}{|Y|} \leq \frac{1+\gamma(|H|)}{1+\gamma(|Y|)} + 2\rho(|Y|).$$

[1] Note that if $|H| < t$, then the lemma is vacuously true.

Thus,

$$\frac{|N_{H\setminus F}(X)|}{|Y|} = 1 - \frac{|X|}{|Y|} \geq \frac{\gamma(|Y|) - \gamma(|H|)}{1+\gamma(|Y|)} - 2\rho(|Y|) \overset{(2)}{\geq} \frac{\gamma(|Y|) - \gamma(|H|)}{2} - 2\rho(|Y|).$$

Now, if $|Y| \geq 3|H|/4$, then $|N_{H\setminus F}(X)| = |Y| - |X| \geq |H|/4 \geq |X|/2 \geq \rho(|X|) \cdot |X|$, yielding the desired robust vertex expansion. If $|Y| \leq 3|H|/4$, then applying (3) with $C' = 4/3$, we have, as $\gamma(x)$ is decreasing and $C \geq 30$, that

$$\begin{aligned}\frac{|N_{H\setminus F}(X)|}{|Y|} &\geq \frac{\gamma(|Y|) - \gamma(|H|)}{2} - 2\rho(|Y|) \geq \frac{\gamma(|Y|) - \gamma(4|Y|/3)}{2} - 2\rho(|Y|)\\ &\geq \frac{C\log(4/3)}{8/3} \cdot \rho(|Y|) - 2\rho(|Y|) \geq \rho(|Y|).\end{aligned}$$

Finally, as $\rho(x)x$ is increasing with x, we have $|N_{H\setminus F}(X)| \geq \rho(|Y|)|Y| \geq \rho(|X|)|X|$.

This finishes the proof of Lemma 3.4.

3.3 High connectivity

It should not come as surprise that expanders are highly connected. We show the following version with $t \sim d(H)$.

Proposition 3.8. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$, $d > 0$ and $\rho(x) = \rho(x, \varepsilon_1, \varepsilon_2 d)$ as in (1). Let H be an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-robust-expander with $\delta(H) \geq d$, then H is νd-connected, where $\nu := \frac{\varepsilon_1}{4\log^2(\frac{15n}{2\varepsilon_2})}$.*

Proof. Suppose H has a vertex cut S of size less than νd, where $\nu = \frac{\varepsilon_1}{4\log^2(\frac{15}{2\varepsilon_2})}$. Let X be the smallest component in $H - S$. Then $x := |X| < |H|/2$. On the other hand, for any vertex $v \in X$, we have $N_H(v) \subseteq X \cup S$. Since $\delta(H) \geq d$, we have that

$$|X| \geq \delta(H) - |S| > \frac{d}{2} \geq \frac{\varepsilon_2 d}{2}.$$

Thus, by the expansion property of H, $|N_H(X)| \geq \rho(x)x$. Now, since $N_H(X) \subseteq S$ and $\rho(x)x$ is increasing, we have

$$|S| \geq |N_H(X)| \geq \rho(x)x \geq \rho\left(\frac{d}{2}\right)\frac{d}{2} = 2\nu d > |S|,$$

a contradiction. □

3.4 Applying expander lemma

We have laid out the basics of the sublinear expanders and are now ready to see some applications. Summarising what we have so far, the following version of expander lemma combining Lemma 3.4 and Proposition 3.8 is convenient to apply.

Lemma 3.9 (Sublinear expander lemma, [1]Lemma 3.2)**.** *Let $C > 30, \varepsilon_1 \leq 1/(10C), \varepsilon_2 < 1/2, d > 0$ and $\rho(x) = \rho(x, \varepsilon_1, \varepsilon_2 d)$ as in (1). Then every graph G with $d(G) = d$ has a subgraph H such that*

(i) H is an $(\varepsilon_1, \varepsilon_2 d)$-robust-expander;

(ii) $d(H) \geq (1-\delta)d$, where $\delta := \frac{C\varepsilon_1}{\log 3}$;

(iii) $\delta(H) \geq d(H)/2$;

(iv) H is νd-connected, where $\nu := \frac{\varepsilon_1}{16\log^2(\frac{15}{2\varepsilon_2})}$.

As the expander subgraph H is almost as dense as the original graph G, when a problem is about graphs with given average/min degree, by passing to an expander guaranteed in Lemma 3.9, we may assume our graph G itself is an expander. However, as mentioned in Remark 3.3, we have now no control over the density of the expander. So, in many applications, we have to work in separate cases depending on its density and employ different techniques. Writing $n = |G|$ and $d = d(G)$, often time, the following division is useful:

- *Dense case:* $d = \Omega(n)$.
- *Medium expander;* $(\log n)^{\Omega(1)} \le d = o(n)$;
- *Sparse expander:* $d = (\log n)^{O(1)}$

When graphs have positive edge density, Szemerédi's regularity lemma offers much stronger pseudorandomness than the sublinear expansion. We shall focus here the latter two cases of embedding in medium and sparse expanders.

Roughly speaking, for the applications that we shall see later, the goal is to find various paths and the strategies can be succinctly summarised as follows.

Medium expanders are still relatively dense. So as long as the sets that we connect are large enough, a greedy embedding usually suffices. The main task here then is to find vertices that has 'large vertex boundary' to serve as endpoints.

For sparse expanders, if additionally we have 'bounded' maximum degree, then we can find many vertices that are pairwise far apart to anchor. Then we can grow each of them till large enough (without running into each other) and then connect. The main issue here is that the set of high degree vertices, say L, would keep us from taking enough points far apart, and small sets in $G - L$ might not expand as $G - L$ might not be an expander anymore. This can be overcome in many cases as we shall see later.

4 Extremal problems on Subdivisions

In this section, we will see the original application of sublinear expanders by Komlós and Szemerédi [3] on embedding clique subdivisions.

The extremal problem for graph subdivisions started at Mader's work in the 1960s, who proved that constant average degree suffices to force subdivision of any fixed size. That is, for any H, we can define

$$d_{\mathsf{T}}(H) := \inf\{c: \ d(G) \ge c \Rightarrow G \ \text{ contains } \ H \ \text{ as a subdivision}\}.^2$$

After this, Mader and independently Erdős and Hajnal conjectured in the 70s that quadratic bound suffices to force a clique subdivision, i.e. $d_{\mathsf{T}}(K_t) = O(t^2)$. If true, this would be possible.

Exercise 4.1. If $K_{d,d}$ contains a K_ℓ-subdivision, then $\ell = O(\sqrt{d})$.

Mader-Erdős-Hajnal's conjecture remained open for a couple of decades. It was resolved in early 90s, by Bollobás and Thomason, and indepensently by Komlós and Szemerédi.

Applying Lemma 3.9, we may assume G is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \ge d$.

Now, if G is dense, i.e. $d = \Omega(n)$, then Alon, Krivelevich, Sudakov proved, using dependent random choice, that G contains a 1-subdivision of $K_{\Omega(\sqrt{n})}$ (i.e. the paths replacing all clique edges are of length 2).

To illustrate the basic use of sublinear expanders via Lemma 3.5, we shall prove here a weaker bound with extra polylog-factor for medium expanders as follows.

[2] Here T stands for topological minor.

Theorem 4.2. *Let $0 < \varepsilon_1 < 1$, then there exist d_0 such that the following holds for any $d \geq d_0$ and $n \in \mathbb{N}$ with $d \geq n^{0.0001}$. Let G be an n-vertex (ε_1, d)-expander with $\delta(G) \geq d$, then $\mathsf{TK}_{\sqrt{d}/\log^4 d} \subseteq G$.*

Proof. Let $t = \sqrt{d}/\log^4 d$ and $m = \frac{2}{\varepsilon_1}\log^3(\frac{15n}{d})$. To get a TK_t, take arbitrary t vertices $v_1, \ldots, v_t$ to anchor and in arbitrary order connect pairs v_i, v_j, $ij \in \binom{[t]}{2}$ with paths P_{ij} of length at most $2m$ each such that all paths P_{ij}, $ij \in \binom{[t]}{2}$, are pairwise internally vertex disjoint. This can indeed be done as, using that $d \geq n^{0.0001}$, the total number of vertices in all paths is at most $\binom{t}{2} \cdot 2m \leq t^2 m < \rho(d)d/4$. Since $\delta(G) \geq d$, for any pair v_i, v_j, we can connect $N(v_i), N(v_j)$, using Lemma 3.5 with W being the internal vertices of previous paths, with a path of length at most m. This path together with v_i, v_j forms the desired v_i, v_j-path P_{ij} of length at most $m + 2 \leq 2m$. □

5 Komlós's conjecture on Hamiltonian subsets

In this section, we give another example of working with medium expanders, proving a special case of Komlós's conjecture on Hamiltonian subsets.

A set of vertices $A \subseteq V(G)$ is *Hamiltonian* if $G[A]$ contains a Hamiltonian cycle, i.e. G contains a cycle whose vertex set is A. Denote by $c(G)$ the number of Hamiltonian subsets in G. A natural question is to ask how $c(G)$ relates to minimum degree, that is, minimising $c(G)$ given $\delta(G) \geq d$. Here, given minimum degree, as the number of vertices n of G is not fixed, it is intuitive that to minimise $c(G)$, one would like to minimise n. Note that the complete graph K_{d+1} is the unique smallest graph with $\delta(G) \geq d$.

Indeed, Komlós conjectured in 1981 that among all graphs with minimum degree at least d, K_{d+1} minimises the number of Hamiltonian subsets, i.e.

$$\delta(G) \geq d \quad \Longrightarrow \quad c(G) \geq c(K_{d+1}) = (1 + o_d(1))2^{d+1}.$$

Tuza gave a bound of $c(G) \geq 2^{d/2}$, leaving an exponential gap.

Recently, Komlós's conjecture was confirmed in [2] for large d in a strong sense. It shows a strong stability that under the weaker condition that $d(G) \geq d$, we already have about twice as many Hamiltonian subsets as K_{d+1}, i.e.

$$c(G) \geq (2 - o_d(1))c(K_{d+1}),$$

unless G is isomorphic to K_{d+1} or the graph obtained from gluing K_{d+1} and K_d at a vertex.

We will present a key step dealing with *almost regular* medium expanders as follows. It shows that almost regular medium expanders have exponentially more Hamiltonian subsets.

Theorem 5.1. *Let $0 < \varepsilon_1 < 1$ and $L \geq 1$, there exist d_0, K_0 such that the following holds for any $d \geq d_0, K \geq K_0$ and $n \in \mathbb{N}$ with $\log^{100} n \leq d \leq n/K$. Let H be an n-vertex $(\varepsilon_1, d/30)$-expander with $d/10 \leq \delta(H) \leq \Delta(H) \leq Ld$. Then $c(H) \geq 2^{50d}$.*

5.1 Proof idea

We will locate a set $Z \subseteq V(H)$ of size $200d$ (see Lemma 5.2) such that for every subset $U \subseteq Z$ of size $|Z|/2$, we can find a cycle C_U whose intersection with Z is almost the whole of U. This then implies that for a large fraction of $U \in \binom{Z}{|Z|/2}$, the corresponding $V(C_U)$ are distinct Hamiltonian subsets.

To construct such a set Z, a key structure is a 'web' (See Definition 5.4 and Figure 1), which offers a 'core' vertex with large 'exterior'. Suppose there are $200d$ almost pairwise disjoint webs, then set Z to be the set of core vertices of these webs. For each $|Z|/2$-set U in Z, to construct C_U, we will connect the webs corresponding to U via paths through their 'exteriors' in a cyclic manner. We hope to find vertex-disjoint (short) paths between the (large) exteriors of webs, avoiding previously-built paths, which together with the paths inside the webs leading to their core vertices form the desired cycle C_U (see Figure 3).

However, such short paths can still block all webs, making it impossible to integrate their core vertices into the cycle C_U. To overcome this, we will choose our paths in a more careful way, such that we avoid using too many vertices inside any particular web. Then the fact that the webs chosen are almost disjoint enables us to incorporate most of the vertices in U into C_U.

Lemma 5.2. *Let H be as in Theorem 5.1. Then $V(H)$ contains a set Z of size $200d$ such that, for every subset $U \subseteq Z$ of size $100d$, there exists a cycle C_U with $V(C_U) \cap Z \in \binom{U}{\geq 98d}$.*

Let us first see how this lemma implies what we need.

Proof. Lemma 5.2 $\implies$ Theorem 5.1. Apply Theorem 5.2 to obtain a set $Z \subseteq V(H)$ of size $200d$ such that, for every subset $U \subseteq Z$ of size $100d$, there exists a cycle C_U with $V(C_U) \cap Z \in \binom{U}{\geq 98d}$.

Fix an arbitrary cycle C in H such that $V(C) \cap Z$ has size $98d \leq |V(C) \cap Z| \leq 100d$. There are at most $\binom{|Z|-|V(C)\cap Z|}{100d-|V(C)\cap Z|}$ ways to choose a $100d$-set $U \subseteq Z$ containing $V(C) \cap Z$. In other words, for a fixed cycle C in H,

$$|\{U \subseteq Z : V(C_U) = V(C)\}| \leq \binom{|Z| - |V(C) \cap Z|}{100d - |V(C) \cap Z|} \leq \binom{102d}{2d}.$$

Therefore the number of Hamiltonian subsets in H is

$$c(H) \geq \frac{\binom{|Z|}{100d}}{\binom{102d}{2d}} \geq 2^{50d},$$

as desired. $\square$

5.2 Webs

We now introduce the key structure 'web'. The main result in this subsection is Lemma 5.5 which guarantees many webs with disjoint interiors.

Definition 5.3 ((h_1, h_2, h_3)-unit)**.** For $h_1, h_2, h_3 \in \mathbb{N}$, a graph F is an (h_1, h_2, h_3)-*unit* if it contains distinct vertices u (the *core vertex* of F) and $x_1, \ldots, x_{h_1}$, and $F = \bigcup_{i\in[h_1]}(P_i \cup S_{x_i})$, where

- $\{P_i : i \in [h_1]\}$ is a collection of pairwise internally vertex disjoint paths, each of length at most h_3, such that P_i is a u, x_i-path.
- $\{S_{x_i} : i \in [h_1]\}$ is a collection of vertex disjoint h_2-stars such that S_{x_i} has centre x_i and $\bigcup_{i\in[h_1]}(V(S_{x_i}) \setminus \{x_i\})$ is disjoint from $\bigcup_{i\in[h_1]} V(P_i)$.

Define the *exterior* $\mathsf{Ext}(F) := \bigcup_{i\in[h_1]}(V(S_{x_i}) \setminus \{x_i\})$ and *interior* $\mathsf{Int}(F) := V(F) \setminus \mathsf{Ext}(F)$. For every vertex $w \in \mathsf{Ext}(F)$, let $P(F, w)$ be the unique path from the core vertex u to w in F.

Definition 5.4 ((h_0, h_1, h_2, h_3)-web)**.** For $h_0, h_1, h_2, h_3 \in \mathbb{N}$, a graph W is an (h_0, h_1, h_2, h_3)-*web* if it contains distinct vertices v (the *core vertex* of W) and $u_1, \ldots, u_{h_0}$, and $W = \bigcup_{i\in[h_0]}(Q_i \cup F_{u_i})$, where

- $\{Q_i : i \in [h_0]\}$ is a collection of pairwise internally vertex disjoint paths such that Q_i is a v, u_i-path of length at most h_3.
- $\{F_{u_i} : i \in [h_0]\}$ is a collection of vertex disjoint (h_1, h_2, h_3)-units such that F_{u_i} has core vertex u_i and $\bigcup_{i\in[h_0]}(V(F_{u_i}) \setminus \{u_i\})$ is disjoint from $\bigcup_{i\in[h_0]} V(Q_i)$.

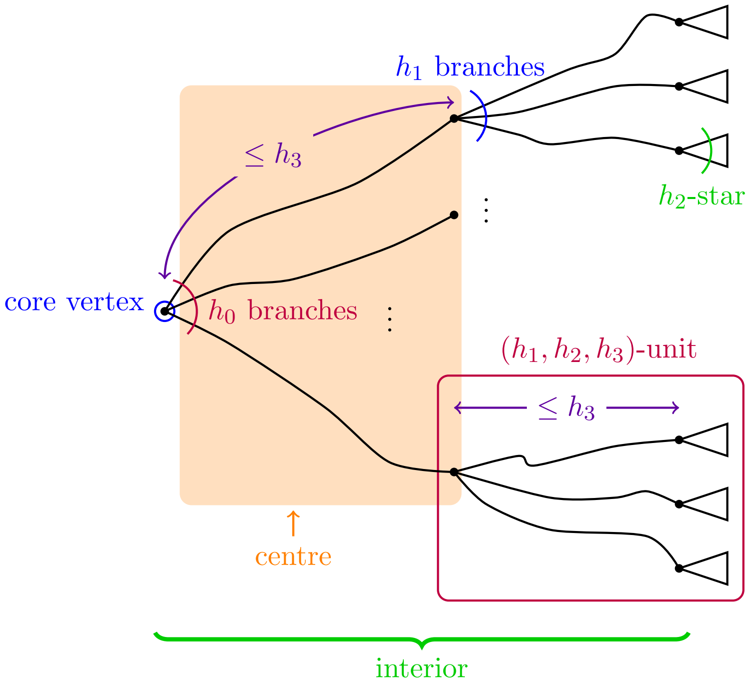


Figure 1: An (h_0, h_1, h_2, h_3)-web.

Define the *exterior* $\mathsf{Ext}(W) := \bigcup_{i\in[h_0]} \mathsf{Ext}(F_{u_i})$, *interior* $\mathsf{Int}(W) := V(W) \setminus \mathsf{Ext}(W)$ and *centre* $\mathsf{Ctr}(W) := \bigcup_{i\in[h_0]} V(Q_i)$. For every vertex $w \in \mathsf{Ext}(W)$, let $P(W, w)$ be the unique path from the core vertex v to w in W.

Lemma 5.5. *Let $0 < \varepsilon_1 < 1$ and $L \geq 1$, there exist d_0, K_0 such that the following holds for any $d \geq d_0, K \geq K_0$ and $n \in \mathbb{N}$ with $\log^{100} n \leq d \leq n/K$. Let H be an n-vertex $(\varepsilon_1, d/30)$-expander with $d/10 \leq \delta(H) \leq \Delta(H) \leq Ld$. Then H contains $(m^3, m^3, d/100, 4m)$-webs $W_1, \dots, W_{200d}$ with pairwise disjoint interiors, where $m = \frac{2}{\varepsilon_2}\log^3(\frac{450n}{d})$.*

The idea of contructing webs with disjoint webs is depicted in Figure 2. First we find disjoint units. Set previously found units aside, say X', our task is to find one more unit in $H - X'$. For this, we find many disjoint stars, connect them in a bipartite fashion (between U and V) and then argue that one of the star is linked to many on the other side, yielding one more unit. The construction of webs are similar, except that we now connect units instead of stars.

Using the 'bounded' maximum degree, we can iteratively pull out many disjoint starts as follows.

Exercise 5.6. Let H be as in Lemma 5.5 and $X' \subseteq V(H)$ be a set of size at most dm^{90}. Then $H - X'$ contains m^{80} vertex disjoint $d/25$-stars.

As outlined above, the following lemma constructs many disjoint units.

Lemma 5.7. *Let H be as in Lemma 5.5 and $X \subseteq V(H)$ be a set of size at most dm^{30}. Then $H - X$ contains vertex disjoint $(m^{10}, d/50, m + 2)$-units $F_1, \dots, F_{m^{20}}$, where $m = \frac{2}{\varepsilon_2}\log^3(\frac{450n}{d})$.*

Proof. Indeed, if we have $F_1, \dots, F_\ell$ for some $0 \leq \ell < m^{20}$, then the set $X' := \bigcup_{i\in[\ell]} V(F_i)$ has size at most $m^{20}(2 \cdot d/25 \cdot m^{10}) \leq dm^{30}$ vertices and $|X \cup X'| \leq 2dm^{30}$. Hence Exercise 5.6 implies that there are vertex disjoint $d/25$-stars $S_{v_1}, \dots, S_{v_{m^{40}}}, S_{u_1}, \dots, S_{u_{m^{60}}}$, with centres $v_1, \dots, v_{m^{40}}, u_1, \dots, u_{m^{60}}$ respectively.

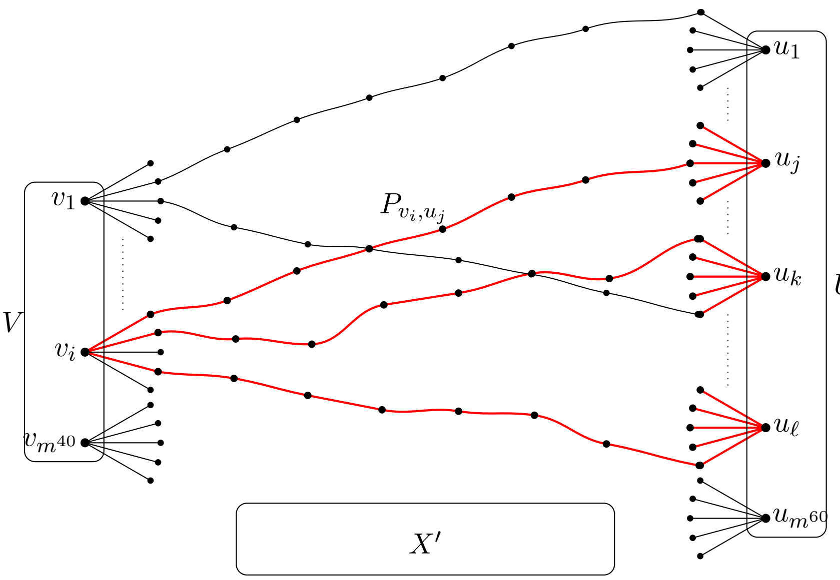


Figure 2: The proof of Lemma 5.5: a unit with core vertex $v_i \in V$ which avoids X'.

Let $P_1, \ldots, P_s$ be pairwise internally disjoint paths of length at most $m+2$ with $s < m^{50}$ where P_i is a v_{c_i}, u_{d_i}-path for each $i \in [s]$ and $d_1, \ldots, d_s$ are all distinct and each P_i does not contain any of $\{v_1, \ldots, v_{m^{40}}, u_1, \ldots, u_{m^{60}}\}$ as an internal vertex. Let

$$V' := \bigcup_{i \in [m^{40}]} (V(S_{v_i}) \setminus \{v_i\}) \text{ and } U' := \bigcup_{i \in [m^{60}] - \{d_1, \ldots, d_s\}} (V(S_{u_i}) \setminus \{v_i\}),$$

then we have $|V'| = dm^{40}$ and $|U'| \geq (m^{60} - m^{50}) \cdot d/25 > dm^{40}$. Further set

$$P' := \bigcup_{i \in [s]} (V(P_i) \setminus \{v_{c_i}\}), \text{ and } U = \{v_1, \ldots, v_{m^{40}}, u_1, \ldots, u_{m^{60}}\}.$$

Then, as $d \geq \log^{100} n$, we have $|X \cup X'| + |P'| + |U| \leq 2dm^{30} + m^{50}(m+2) + m^{40} + m^{60} \leq 3dm^{30} \leq \rho(dm^{40}) \cdot dm^{40}/4$. Hence, applying Lemma 3.5 with $V', U', X \cup X' \cup P' \cup U$ playing the roles of X_1, X_2, W respectively, we can find a path of length at most m between a vertex in $V(S_{v_{d_{s+1}}}) \setminus \{v_{s+1}\} \subseteq V'$ and a vertex in $V(S_{u_{d_{s+1}}}) \setminus \{u_{d_{s+1}}\} \subseteq U'$ avoiding vertices in $X \cup X' \cup P' \cup U$. This yields a $v_{d_{s+1}}, u_{d_{s+1}}$-path P_{s+1}, which is internally disjoint from $X \cup X' \cup P' \cup U$ and $d_{s+1} \notin \{d_1, \ldots, d_s\}$. Hence, this is internally disjoint from $P_1, \ldots, P_s$ and U.

Repeating this for $s = 0, 1, \ldots, m^{50}$, we obtain $P_1, \ldots, P_{m^{50}}$. By pigeonhole principle, at least m^{10} of $v_{c_1}, \ldots, v_{c_{m^{50}}}$ coincide, so there exists a vertex v_j and m^{10} internally disjoint paths from v_j to $V(S_{u'_1}), \ldots, V(S_{u'_{m^{10}}})$ for pairwise disjoint stars $S_{u'_1}, \ldots, S_{u'_{m^{10}}}$ where each star has $d/25$ leaves. These paths and stars together yields a $(m^{10}, d/50, m+2)$-unit $F_{\ell+1}$ disjoint from $X \cup X'$. Repeating this for $\ell = 0, 1 \ldots, m^{20}$, yields that the graph $H - X$ contains vertex disjoint $(m^{10}, d/50, m+2)$-units $F_1, \ldots, F_{m^{20}}$ as desired. $\square$

Now, given many disjoint units, we can connect them in a similar way as in Lemma 5.7 to obtain desired webs.

Exercise 5.8. Prove Lemma 5.5.

5.3 Proof of Lemma 5.2

We now carry on the proof of Lemma 5.2 outlined in Section 5.1.

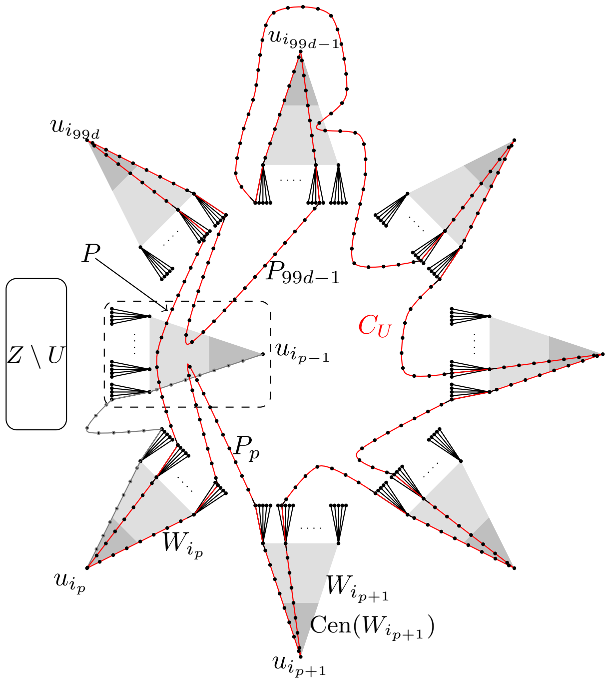


Figure 3: Constructing a cycle C_U. Here $p \in [99d]$ is the smallest index such that W_{i_p} is a good web (so the web $W_{i_{p-1}}$ enclosed in a dashed box is bad, i.e. its interior is over-used by the paths P_j). Paths P_j can intersect the interior of a web but not its centre.

Let $m := \frac{2}{\epsilon_1}\log^3(\frac{450n}{d})$. Recall that $\rho(x)$ is decreasing when $d/30 \le x \le n$ and by choosing $n/d \ge K \ge K_0$ large enough, we have

$$\rho(x) \ge \rho(n) > \frac{1}{m}; \quad \text{and also} \quad n \ge Ldm^{100} \quad \text{and} \quad d \ge m^{30}. \tag{6}$$

By Lemma 5.5, we can find in H a collection $W_1, \dots, W_{200d}$ of $(m^3, m^3, \frac{d}{100}, 4m)$-webs whose interiors $\mathsf{Int}(W_1), \dots, \mathsf{Int}(W_{200d})$ are pairwise disjoint. Let $Z := \{u_1, \dots, u_{200d}\}$ where u_i is the core vertex of W_i for all $i \in [200d]$. Fix an arbitrary $100d$-set U in Z. Without loss of generality, assume that $U = \{u_1, \dots, u_{100d}\}$. First, we show that there exists an index set $I = \{i_1, \dots, i_{99d}\} \subseteq [100d]$ and a collection $\mathcal{Q} = \{P_\ell : \ell \in [99d-1]\}$ of paths satisfying the following. For each $\ell \in [99d-1]$,

A1 P_ℓ is a $u_{i_\ell}, u_{i_{\ell+1}}$-path of length at most $18m$;

A2 $\mathsf{Int}(P_\ell)$ is disjoint from $\bigcup_{k \in [200d] \setminus \{i_\ell, i_{\ell+1}\}} \mathsf{Ctr}(W_k) \cup Z$;

A3 $\mathsf{Int}(P_\ell)$ and $\mathsf{Int}(P_k)$ are disjoint for all $k \in [99d-1] \setminus \{\ell\}$;

A4 $|\mathsf{Int}(W_{i_{\ell+1}}) \cap \bigcup_{k \in [\ell]} V(P_k)| < 2m^2$.

To find such an $(I, \mathcal{Q})$, we will build a path between pairs in U avoiding vertices used in previously-built paths and the centres of all other webs. During the process, we will skip a web if its interior is 'over-used'.

Assume we have built $P_1, \dots, P_s$ and determined $i_1, \dots, i_{s+1}$ with $s < 99d-1$ satisfying **A1**–**A4**. Since index set $\{1\}$ with the empty collection of paths satisfies **A1**–**A4**, such a collection $\{P_1, \dots, P_s\}$ exists. Let $P' := \bigcup_{k \in [s]} \mathsf{Int}(P_k)$. For $i \in [200d]$, we say a web W_i is *bad* if $|\mathsf{Int}(W_i) \cap P'| \ge 2m^2$, and *good* otherwise. Note that **A4** implies that $W_{i_{s+1}}$ is good. By **A1**,

$$|P'| \le 18m \cdot s \le 1800dm. \tag{7}$$

As good webs by definition are mostly intact, we can use Lemma 3.5 as follows.

Exercise 5.9. Let W_{j_1} and W_{j_2} be two good webs. Then there exists a path P of length at most $18m$ in H from u_{j_1} to u_{j_2} such that $\mathsf{Int}(P)$ is disjoint from $P' \cup Z \cup \bigcup_{k\in[200d]\setminus\{j_1,j_2\}} \mathsf{Ctr}(W_k)$.

Since the interiors of $W_1, \ldots, W_{200d}$ are pairwise disjoint, (7) implies that the number of webs whose interiors contain at least m^2 vertices of P' is at most

$$\frac{1800dm}{m^2} < \frac{d}{2}. \tag{8}$$

Since $s < 99d - 1$, we can choose $i_{s+2} \in U \setminus \{i_1, \ldots, i_{s+1}\}$ such that

$$|\mathsf{Int}(W_{i_{s+2}}) \cap P'| \leq m^2. \tag{9}$$

Recall that $W_{i_{s+1}}$ is good. Thus by Exercise 5.9, there is a $u_{i_{s+1}}, u_{i_{s+2}}$-path P_{s+1} of length at most $18m$. Then it is easy to see that $\{i_1, \ldots, i_{s+2}\}$ together with $P_1, \ldots, P_{s+1}$ satisfy **A1**–**A3** since $\mathsf{Int}(P_{s+1})$ is disjoint from $P' \cup \bigcup_{k\in[200d]\setminus\{i_\ell,i_{\ell+1}\}} \mathsf{Ctr}(W_k) \cup Z$. Moreover,

$$\left|\mathsf{Int}(W_{i_{s+2}}) \cap \bigcup_{k\in[s+1]} V(P_k)\right| = \left|\mathsf{Int}(W_{i_{s+2}}) \cap P'\right| + |\mathsf{Int}(W_{i_{s+2}}) \cap P_{s+1}| \overset{(9)}{\leq} m^2 + 18m < 2m^2,$$

so **A4** also holds. Therefore, we can repeat this process until $s = 99d - 1$, upon which we obtain the desired $(I, \mathcal{Q})$ satisfying **A1**–**A4**.

Observe that, as before, (8) implies that less than $d/2$ indices $k \in [100d] \setminus I$ are such that W_k is bad. Let $p \in [99d]$ be the minimum index such that W_{i_p} is a good web (see Figure 3). Note that $W_{i_{99d}}$ is good by **A4**. Then $p \leq d/2$ and so $|\{i_p, i_{p+1}, \ldots, i_{99d}\}| > 98d$. By **A1**–**A3**, the concatenation of $P_p, P_{p+1}, \ldots, P_{99d-1}$ is a $u_{i_p}, u_{i_{99d}}$-path avoiding $Z \setminus U$. By Exercise 5.9, there exists a $u_{i_p}, u_{i_{99d}}$-path P of length at most $18m$ such that $\mathsf{Int}(P)$ is disjoint from $\bigcup_{k\in[99d-1]} \mathsf{Int}(P_k) \cup Z \cup \bigcup_{k\in[200d]\setminus\{i_p,i_{99d}\}} \mathsf{Ctr}(W_k)$. Thus, the concatenation of $P_p, P_{p+1}, \ldots, P_{99d-1}, P$ form a cycle C_U, as in Figure 3. Finally, by **A1**, **A2** and Exercise 5.9,

$$V(C_U) \cap Z = \{u_{i_p}, u_{i_{p+1}}, \ldots, u_{i_{99d}}\} \in \binom{U}{\geq 98d},$$

completing the proof of Lemma 5.2.

6 Mader's conjecture on subdivisions in C_4-free graphs

We have seen that 'bounded' maximum degree is often useful, for instance when we tried to build webs in Lemma 5.5, we need it for extracting disjoint stars in Exercise 5.6. However, as mentioned in Remark 3.3, when passing from G to an expander subgraph H, we do not know anything about the order of H, so neither do we have any bound its maximum degree. In this section, we will show an example in which reduction to expanders with 'bounded' maximum degree follows straightforwardly from a basic property of expanders. That is, (see Proposition 6.3), an expander still expands (though with slightly weaker expansion) if a very small set of vertices is removed from it.

The example we shall look at is a continuation of the problems discussed in Section 4. Recall that Komlós and Szemerédi proved that every graph with average degree d contains a $\mathsf{TK}_{\Omega(\sqrt{d})}$, which is best possible as shown by $K_{d,d}$ in Exercise 4.1. Note that in this example, the complete bipartite graph $K_{d,d}$ is locally dense in the sense that it contains all sorts of small bipartite graph. It is natural to suspect that the square root bound can be improved if we step away from $K_{d,d}$. Indeed, this was conjectured by Mader as follows: if G is C_4-free and $d(G) \geq d$, then it contains a $\mathsf{TK}_{\Omega(d)}$. Note that, if true, linear in d is obviously the optimal order of magnitude, as the graph G could be simply d-regular.

Mader's conjectured was confirmed in a strong sense in [5] as follows. Let $2 \le s \le t$, and G be a $K_{s,t}$-free graph with average degree d. Then G contains a $\mathsf{TK}_{\Omega\left(d^{\frac{1}{2}\frac{s}{s-1}}\right)}$. We leave it as an exercise to show that, conditioning on $\mathrm{ex}(n, K_{s,t}) = \Theta(n^{2-1/s})$, the exponent $\frac{1}{2}\frac{s}{s-1}$ is best possible.

Let us consider the baby case $s = t = 2$, i.e. when G is $K_{2,2} = C_4$-free. We will show the starting step which reduces the problem to expanders with 'bounded' maximum degree.

The idea is that, if there linearly in d many vertices with high degree, then we can greedily embed a desired clique subdivision, see Lemma 6.1. If there is only sublinearly many high degree vertices, then removing this small set of vertices leaves us still an expander with a slightly weaker expansion, see Proposition 6.3.

Lemma 6.1. *Let $0 < \varepsilon_1, \varepsilon_2, c < 1$, then there exists $c' > 0$ such that the following holds. Let G be an n-vertex C_4-free $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \ge d$. If there are at least cd vertices with degree at least $d \log^{10} n$, then $\mathsf{TK}_{c'd} \subseteq G$.*

In the above, if we have $\Omega(d)$ vertices of degree at least $d^2 \log^{10} n$, then it is an easy exercise, using Lemma 3.5, to embed $\mathsf{TK}_{\Omega(d)}$. We can use the C_4-freeness to reduce the degree requirement.

Exercise 6.2. Prove Lemma 6.1.

Proposition 6.3. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$, then there exists $c > 0$ such that the following holds. Let G be an $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \ge d$, and $A \subseteq V(G)$ be a set of size at most cd. Then $G - A$ is an $(\varepsilon_1/2, \varepsilon_2 d)$-expander with minimum degree at least $d/2$.*

This basic but useful property follows from the definition of expander. We leave it as an exercise.

7 Reed and Wood's question on planar minors

In Sections 4 and 5, we have seen two examples for medium expanders. In particular, we have seen that the web structure, due to its large exterior, is particularly useful when carrying out connections robustly using Lemma 3.5. However, recalling the proof of Lemma 5.7, the way we construct webs requires that the expander is not too sparse (average degree at least polylog$n$). We can no longer guarantee webs when expander is sparse, at least not with the same approach.

The purpose of this section is twofold.

1. We study sparse expanders and present a different approach that works even when the expander has constant average degree. Essentially, provided again 'bounded' maximum degree, we want to find vertices that are pairwise far apart so that there is room to grow them to have large boundaries, which are easy to connect, see Section 7.2.

   In doing so, we will introduce in Section 7.1 yet another important expansion lemma, Lemma 7.4, showing that sets expand almost exponentially in an expander.

2. The above new approach still requires 'bounded' maximum degree condition. In the example we consider here, Proposition 6.3 is not strong enough to do the max-degree reduction. Instead, we will rely on a new structure nakji. This will be introduced in Section 7.3 along with an explaination of why it is useful. Then in Section 7.4, we sketch a proof for a more involved reduction method.

The problem we want to look at here is about forcing sparse graphs, rather than the cliques, as minors. Intuitively, given the order of H, it is easier to find an H-minor, when H is sparse.

Indeed, the seminal result of Kostochka, and independently Thomason from the 80s, states that the average degree needed to force K_t -minor is $O(t\sqrt{\log t})$. On the other hand, Reed and Wood showed that for any t-vertex graph H, the average degree needed to force H-minor is $O(t\sqrt{\log d(H)})$. Along with this result, Reed and Wood raised several interesting questions about forcing sparse minors. Among others, they asked for the minimum constant $c > 0$ such that average degree ct forces every t-vertex planar graph as minor. This was recently answered in [1] asymptotically as follows.

Theorem 7.1. *Let G be a graph. If $d(G) \geq (3/2 + o_t(1))t$, then it contains every t-vertex planar graph as a minor. Furthermore, the constant $3/2$ is best possible.*

In fact, somewhat surprisingly, $(3/2+o(1))t$ is the tight bound even for graphs with bounded genus. A key step in the proof is a general result embedding subdivisions of sparse bipartite graphs. The example we will consider here is the following baby case dealing with sparse expanders with additional 'bounded' maximum degree.

Lemma 7.2. *Let $0 \ll 1/d \ll \varepsilon_1, \varepsilon_2 \ll \varepsilon, 1/\Delta < 1$ and let H be a graph with at most d vertices and $\Delta(H) \leq \Delta$. Suppose F is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-robust-expander with $\delta(F) \geq \varepsilon^2 d$. If $\Delta(F) \leq \log^{30000} n$, then F contains an H-subdivision.*

To prove Lemma 7.2, we need a useful expansion lemma, Lemma 7.4, which shows that small sets grow almost exponentially in a sublinear expander.

7.1 Exponential growth

In an (ε_1, t)-robust-expander graph, for a set X with size at least $t/2$, the ball $B^i(X)$ grows with the radius i. We need a robust version, quantifying how resilient this growth is to deletion of a thin set around X. The particular type of thin set we use is the following.

Definition 7.3. For a set $X \subseteq W$ of vertices, the paths $P_1, \ldots, P_q$ are *consecutive shortest paths from X within W* if the following holds. For each $i \in [q]$, $P_i|_W$ is a shortest path from X to some vertex $v_i \in W \setminus X$ in the graph restricted to $W - \bigcup_{j \in [i-1]} V(P_j)$.

Basically, if $P_1, \ldots, P_q$ are consecutive shortest paths from X, then collectively they should not take up too many vertices per layer around X, as otherwise shorter paths could be found. As such, we can expand X past these paths. This is made precise as follows.

Lemma 7.4. *Let $0 < 1/d \ll \varepsilon_1, \varepsilon_2 \ll 1$. Suppose G is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-robust-expander and X, Y are sets of vertices with $|X| = x \geq \varepsilon_2 d$ and $|Y| \leq \frac{1}{4} \cdot \rho(x) \cdot x$. Let $P_1, \ldots, P_q$ be consecutive shortest paths in $G - Y$ from X within $B^r_{G-Y}(X)$, where $1 \leq r \leq \log n$ and $q < x/\log^8 x$, and let $P = \bigcup_{i \in [q]} V(P_i)$. Then for each $i \in [r]$, we have*

$$|B^i_{G-P-Y}(X)| \geq \exp(i^{1/4}).$$

Proof. For each $i \geq 0$, let $Z_i = B^i_{G-P-Y}(X)$. As $P_1, \ldots, P_q$ are consecutive shortest paths from X, for each $i \geq 0$, each path P_j, $j \in [q]$, can intersect with the set $N_{G-Y}(Z_i)$ on at most $i+2$ vertices. Indeed, otherwise we can replace the initial segment of P_j with a path in $Z_i \cup N_{G-Y}(Z_i)$ of length $i+1$ to get a shorter path in $G - Y - \bigcup_{k \in [j-1]} V(P_k)$, contradicting the choice of P_j. Thus, $|N_{G-Y}(Z_i) \cap P| \leq (i+2)q$. Consequently, the expansion of G implies for each $i \geq 0$ that

$$\begin{aligned} |Z_{i+1}| = |Z_i| + |N_G(Z_i) \setminus (Y \cup P)| &\geq |Z_i| + \rho(|Z_i|)|Z_i| - |Y| - |N_{G-Y}(Z_i) \cap P| \\ &\geq |Z_i| + \frac{3}{4}\rho(|Z_i|)|Z_i| - (i+2)q. \end{aligned}$$

Let

$$f(z) = \exp(z^{1/4}) \quad \text{and} \quad g(z) := x + \frac{1}{2}\rho(x)xz.$$

We first use induction on i to show that for each $0 \le i \le \log^4 x$, $|Z_i| \ge g(i)$. Indeed, $|Z_0| = |X| = x = g(0)$. Then, as $\rho(z)z$ is increasing when $z \ge x$ and $\frac{1}{4}\rho(x)x \ge \frac{(i+2)x}{\log^8 x} > (i+2)q$ due to $i \le \log^4 x$, we see that

$$|Z_{i+1}| \ge |Z_i| + \frac{3}{4}\rho(|Z_i|)|Z_i| - (i+2)q \ge |Z_i| + \frac{1}{2}\rho(x)x = |Z_i| + g(i+1) - g(i) \ge g(i+1).$$

We may then assume $i > \log^4 x$, as $f(i) \le g(i) \le |Z_i|$ when $i \le \log^4 x$. Now, as $i > \log^4 x$, $\frac{f(i)}{i^{7/4}} \ge \frac{f(\log^4 x)}{(\log^4 x)^{7/4}} = \frac{x}{\log^7 x}$ and so

$$(i+2)q < i \cdot \frac{2x}{\log^8 x} \le i \cdot \frac{f(i)}{i^{7/4}} = \frac{f(i)}{i^{3/4}}.$$

Also note that $f(i+1) - f(i) \le \frac{f(i)}{i^{3/4}}$ and $\rho(|Z_i|)|Z_i| \ge \rho(f(i))f(i) \ge \frac{\varepsilon_1 f(i)}{i^{1/2}}$. Thus, we have

$$\begin{aligned}|Z_{i+1}| &\ge |Z_i| + \frac{3}{4}\rho(|Z_i|)|Z_i| - (i+2)q \ge |Z_i| + \frac{3\varepsilon_1 f(i)}{4i^{1/2}} - \frac{f(i)}{i^{3/4}} \\ &\ge |Z_i| + \frac{f(i)}{i^{3/4}} \ge |Z_i| + f(i+1) - f(i) \ge f(i+1),\end{aligned}$$

as desired. □

7.2 'Bounded' degree sparse expander

In this subsection, we prove Lemma 7.2.

We start with the following simple proposition which finds many vertices pairwise far apart.

Proposition 7.5. *Suppose that F is an n-vertex graph with $\Delta(F) \le \log^{30000} n$, and n sufficiently large. Then there is a set of at least $n^{1/5}$ vertices pairwise having distance at least $\frac{\log n}{20000 \log\log n}$.*

Exercise 7.6. Prove Proposition 7.5.

Proof of Lemma 7.2. Let

$$r := (\log\log n)^5 \quad \text{and} \quad r' = \sqrt{\log n}.$$

As $d \le \delta(F)/\varepsilon^2 \le \Delta(F)/\varepsilon^2 \le \log^{30001} n$, Proposition 7.5 implies that we can find vertices $v_1, \dots, v_h$, where $h = |H| \le d$, such that the distance between any two of them is at least $2r + 2r'$. Let $x_1, \dots, x_h$ be the vertices of H and $e_1 = x_{a_1}x_{b_1}, \dots, e_{h'} = x_{a_{h'}}x_{b_{h'}}$ be the edges of H where $h' = e(H) \le \Delta d/2$.

Suppose that we have $Q_1, \dots, Q_\ell$ for some $0 \le \ell < h'$ such that:

- **B1** for each $i \in [\ell]$, Q_i is a v_{a_i}, v_{b_i}-path with length at most $2\log^4 n$;
- **B2** for distinct $i, j \in [\ell]$, the paths Q_i and Q_j are internally vertex disjoint;
- **B3** for each $i \in [h]$, for the edges $\{e_{k_1}, \dots, e_{k_s}\} = \{e_{k_j} : k_j \in [\ell], x_i \in e_{k_j}\}$ with $k_1 < \cdots < k_s$, the paths $Q_{k_1}, \dots, Q_{k_s}$ form consecutive shortest paths from v_i in $B^r(v_i)$; and
- **B4** for any $i \in [h]$ and $j \in [\ell]$ with $x_i \notin e_j$, $B^r(v_i)$ and $V(Q_j)$ are disjoint.

Let

$$W_1 = \bigcup_{i \in [\ell]} \mathsf{Int}(Q_i), \quad W_2 := \bigcup_{i \in [h]:\ x_i \notin e_{\ell+1}} B^r(v_i) \quad \text{and} \quad W = W_1 \cup W_2.$$

Note that $v_1, \dots, v_h$ are pairwise a distance at least $2r + 2r'$ apart, so by **B4** we have

$$|B^r_{F-W}(v_{k_{\ell+1}})| = |B^r_{F-W_1}(v_{k_{\ell+1}})| = |B^{r-1}_{F-W_1}(B^1_{F-W_1}(v_{k_{\ell+1}}))|$$

for each $k \in \{a,b\}$. Note that $|B^1_{F-W_1}(v_{k_{\ell+1}})| \geq \varepsilon^2 d - \Delta \geq \varepsilon^2 d/2$ by **B3** and **B4**. By **B3**, we can apply Lemma 7.4 with $B^1_{F-W_1}(v_{k_{\ell+1}}), W_1, \varnothing, \Delta$ playing the roles of X, P, Y, q, and then for each $k \in \{a,b\}$ we have

$$|B^r_{F-W}(v_{k_{\ell+1}})| = |B^r_{F-W_1}(v_{k_{\ell+1}})| \geq \exp((r-1)^{1/4}) \geq d \log^8 n,$$

where the last inequality follows from $d \leq \log^{30001} n$. This implies that

$$|W_1| \leq h' \cdot 2\log^4 n \leq \Delta d \log^4 n < \frac{1}{4}\rho(|B^r_{F-W}(v_{k_{\ell+1}})|) \cdot |B^r_{F-W}(v_{k_{\ell+1}})|.$$

Hence, by applying Lemma 7.4, now with $B^r_{F-W}(v_{k_{\ell+1}}), \varnothing, W_1$ playing the roles of X, P, Y for each $k \in \{a,b\}$, we similarly have

$$|B^{r+r'}_{F-W}(v_{k_{\ell+1}})| = |B^{r+r'}_{F-W_1}(v_{k_{\ell+1}})| \geq \exp((r')^{1/4}) \geq \exp(\sqrt[9]{\log n}).$$

As $\Delta(F) \leq \log^{30000} n$ and $d \leq \log^{30001} n$, we then have

$$|W| \leq |W_1| + |W_2| \leq \Delta d \log^4 n + d \cdot 2(\log^{30000} n)^r < \frac{1}{4}\rho\left(\exp(\sqrt[9]{\log n})\right)\exp(\sqrt[9]{\log n}).$$

Therefore, by Lemma 3.5, there is a path in $F-W$ of length at most $\log^4 n$ between $B^{r+r'}_{F-W}(v_{a_{\ell+1}})$ and $B^{r+r'}_{F-W}(v_{b_{\ell+1}})$. So we can let $Q_{\ell+1}$ be a shortest path between $v_{a_{\ell+1}}$ and $v_{b_{\ell+1}}$ in $F-W$, which has length at most $\log^4 n + 2r + 2r' \leq 2\log^4 n$. Hence, $Q_1, \ldots, Q_{\ell+1}$ satisfy **B1**–**B4**. Repeating this for $\ell = 0, 1, \ldots, h'-1$, then the union of all paths $\bigcup_{i\in[h']} Q_i$ is an H-subdivision. $\square$

7.3 Nakji

As we see in the baby case Lemma 7.2, it is relatively simple to handle expanders with an additional 'bounded' maximum degree condition. As mentioned in our discussion in Section 3.4, the bulk of the work, when dealing with sparse expanders, is often time dealing with the set of high degree vertices, say L. Indeed, without a bound on maximum degree, our approach in Section 7.2 breaks down as the starting point Proposition 7.5 is simply not true anymore.

Nonetheless, we will outline in Section 7.4 a way to get around this difficulty. In this subsection, we introduce a key structure involved: nakji, which in sparse expanders is a good substitute for webs.

Definition 7.7 ((t,s,r,τ)-nakji)**.** Given $t, s, r, \tau \in \mathbb{N}$, a graph N is a (t,s,r,τ)-*nakji* in G if it contains vertex disjoint sets M, D_i, $i \in [t]$, each having size at most s, and paths P_i, $i \in [t]$ such that for each $i \in [t]$

- P_i is an M, D_i-path with length at most $10r$, and all paths P_i, $i \in [t]$, are pairwise internally disjoint;
- D_i has diameter at most r, and all D_i, $i \in [t]$, are a distance at least τ in G from each other and from M, and they are disjoint from internal vertices of $\bigcup_{i\in[t]} P_i$.

We call M the *head* of the nakji and each D_i, $i \in [t]$, a *leg*. See Figure 4.

To obtain subdivision from nakjis is straightforward. If there are $d = |H|$ nakjis that are pairwise far apart, then we can link nakjis together to get an H-subdivision by first expanding each leg without bumping into any other part to an enormous size, so that each connection made leaves irrelevant structures untouched.

We shall discuss in the next subsection a way to handle high degree vertices L and carry out our previous approach. At a high level, the idea has the following three new ingredients:

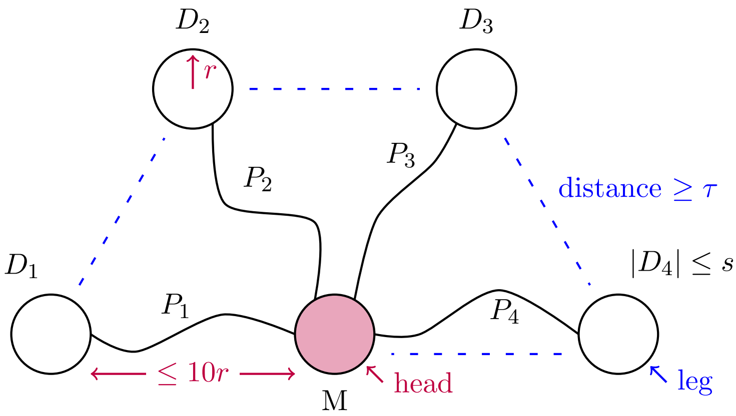


Figure 4: A $(4, s, r, \tau)$-nakji.

- **C1** use nakji to reduce the task from finding vertices with large boundary in $G - L$ to the much easier one of finding sets with large boundary;
- **C2** iterative use of expanders that allows us to construct nakjis;
- **C3** an averaging argument that finds (plenty of) small sets that expand past the set L of high degree vertices.

7.4 Proof sketch for generic sparse expanders

The statement we shall see is the following lemma which deals with generic sparse expanders.

Lemma 7.8. *Let $0 < 1/d \ll \delta \ll \varepsilon_1, \varepsilon_2 \ll \varepsilon, 1/\Delta < 1$ and H be a bipartite graph on at most $(1-\varepsilon)d$ vertices and $\Delta(H) \leq \Delta$. Suppose G is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-robust-expander with $d(G) \geq (1-\delta)d$ and $\delta(G) \geq d(G)/2$. If $d < \log^{300} n$, then G contains an H-subdivision.*

To motivate the definition of nakji, let us first look at the easier problem of finding H-minors, where $|H| = d$ and $e(H) = O(d)$. For minors, we just need to find d large (connected) balls (and contract them afterwards) and find $O(d)$ internally disjoint paths between them.

Notice first that, writing L for the set of high degree vertices, we may assume that $|L| < d$, as if there are at least $d = |H|$ high degree vertices (degree at least $d \log^{50} n$ say), we can easily build even an H-subdivision anchored at these high degree vertices, as in Lemma 6.1. Write again $m = O(\log^3 n)$ for the diameter of the expander G. Since G is now sparse, from above we see then that $|L| < d \leq m^{100}$ is quite small.

We then argue that $G - L$ still has average degree $\Omega(d)$. To see this, if there are too many edge going between L and its complement, then we do an asymmetric bipartite Turán problem to embed H as a subgraph. Having $G - L$ still dense allows us to find subexpanders within $G - L$ with average degree linear in d.

Suppose now within G, we can find d vertices, $v_1, \ldots, v_d$, pairwise a distance $\sqrt{\log n}$ apart, such that for each v_i, the ball B_i of radius say $(\log\log n)^{20}$ around it has size at least m^{200}. So each B_i is large enough to enjoy exponential growth (Lemma 7.4) avoiding all paths previously built. Now to get, say, a v_i, v_j-path, we first expand B_i, B_j to larger balls with radius say $(\log n)^{1/10}$. These larger balls are so gigantic that we can connect them avoiding all the smaller balls $\bigcup_{i \in [d]} B_i$. It is left to find such v_i and B_i. We can find them one by one, by collectively growing a set U of pairwise far apart vertices past L and using an averaging argument to locate the next v_i that expands well in $G - L$.

Coming back to embedding H-subdivisions, we shall follow the general strategy as that of finding minors. However, an immediate obstacle we encounter is the following. To get a subdivision instead of a minor, we need to be able to lead up to $\Delta(H) = O(1)$ many paths arriving at B_i disjointly to v_i. In other words, each anchor vertex v_i has to expand even after removing $O(1)$ disjoint paths starting from itself. Here comes the problem: in the minor case, we just need to expand U ignoring a smaller set L; whereas now U is asked to expand past a larger set of $\Theta(|U|)$ vertices that are used in the previous connections. Our sublinear expansion property is simply too weak for this.

This is where the nakji structure comes into play. It is designed precisely to circumvent this problem by doing everything in reverse order. Basically, instead of looking for anchor vertices that expand robustly, we rather anchor on nakjis and link them via their legs first and then extend the paths from the legs in each nakji's head using connectivity. This illustrates **C1**.

The remaining task is then to find many nakjis in $G-L$. This is done essentially by expanding and linking small subexpanders within $G-L$. For this, we need to find many small expanders in $G-L$. Recall that $G-L$ is still dense and hence by the expander lemma, Lemma 3.9, it contains at least one subexpander F. Now if F is large in order, then if additionally F has medium density, we can embed H-subdivision in F using web structure. So if F is large subexpander, it must be sparse. But then large $F \subseteq G - L$ naturally inherits the 'bounded' max-degree of $G - L$ and so we can invoke Lemma 7.2 on F to go home. Therefore, all subexpanders in $G - L$ is small and consequently we can greedily pull out many small expanders as claimed. This illustrate **C2**.

Finally we arrive at the last problem that we could encounter. We only have $|L| < d$, while each of the subexpanders, though having size $\Omega(d)$, could be smaller than L. This keeps us from expanding and linking each subexpander in $G - L$. An averagin argument again comes to the rescue. Here, a subtlety worth pointing out is that , given that L is not large, for the averaging, one would naturally like to take a huge set of subexpanders, whose union is so large and thus grows easily past L. This, however, would not work as the expanding function $\rho(\cdot)$ is sublinear, if there are too many subexpanders to begin with, say e.g. $\log n$ expanders each of size x, then after averaging, the expansion rate of subexpander in $G - L$ that we can guarantee would be $\rho(x \log n)$, which could be much smaller than our target rate $\rho(x)$. To overcome this difficulty, instead, we shall average over a set of subexpanders of appropriate size that is just big enough to ignore L and on the other hand just small enough that $\rho(\cdot)$ does not decay too much. This illustrates **C3**.

8 Erdős and Hajnal's problem on cycles

In this section, we will address the drawback of Lemma 3.5 in Remark 3.6. In previous problems, we were happy finding short connecting paths. When attacking extremal problems on cycles, it would be useful to be able to connect sets with paths of specified length.

There is a long line of research on what cycles could appear in general graphs. Define the *cycle space* $\mathcal{C}(G)$ of a graph G to be the set of distinct cycle lengths in G. One could ask many natural questions about $\mathcal{C}(G)$ given its chromatic number $\chi(G)$ or its average degree $d(G)$. For example, $d(G) \geq 2$ implies that $\mathcal{C}(G) \neq \varnothing$ and Dirac's theorem states that if $\delta(G) \geq |G|/2$, then $|G| \in \mathcal{C}(G)$. What about finer structure of $\mathcal{C}(G)$?

A positive increasing sequence of integers $\sigma_1, \sigma_2, \dots$ is *unavoidable with high average degree* (*respectively, with high chromatic number*) if there exists some d such that every graph G with average degree at least d (respectively, chromatic number at least d) has some $i \in \mathbb{N}$ with $\sigma_i \in \mathcal{C}(G)$. Note that there are graphs, e.g. $K_{d,d}$, with no odd cycles yet arbitrarily high average degree. Thus, the two overarching questions are as follows.

- Which sequences of even numbers are unavoidable with large average degree?
- Which sequences of odd numbers are unavoidable with large chromatic number?

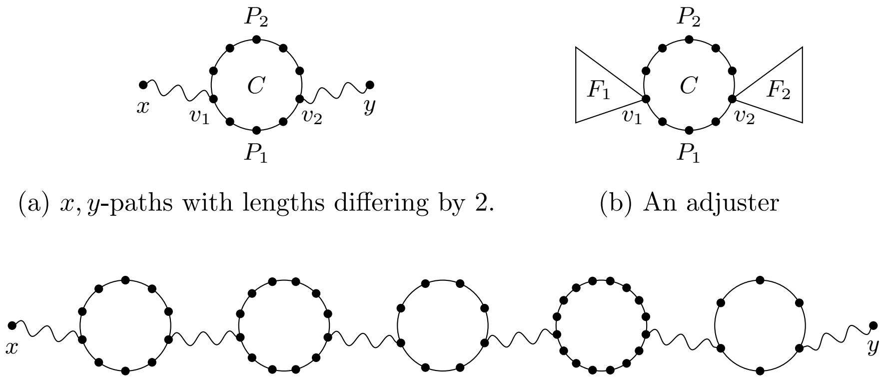
On the first type, Erdős raised several questions and offered \$1000 for solutions. Among others, he asked whether the powers of 2, i.e. $\sigma_i = 2^i$, is avoidable or not with average degree. In fact, nothing is known even for the much slower growing sequence such as $\sigma_i = i^2$ or $\sigma_i = p_i + 1$, where p_i is the ith prime. For the second type, the major open problem was the odd cycle problem of Erdős and Hajnal from 1966, which askes whether $\sum_{\ell \in \mathcal{C}_o(G)} \frac{1}{\ell} \to \infty$ as $\chi(G) \to \infty$.

These problems had been elusive, with not even an explicit unavoidable sequence with zero density known. Very recently, we [6] are able to resolve all of these questions using methods building on sublinear expanders. A key ingredient is the 'adjuster' structure that enables us to vary the length of connection path in expanders.

In the rest of this section, we will introduce this adjuster structure and show, conditioning on some more expansion lemmas, the basic step of how to find one single such structure in sublinear expanders.

8.1 adjusters

The basic idea is that by taking almost antipodal points v_1, v_2 on a cycle C, we can view C as a 'twin path' between v_1, v_2, see Figure 5(a), which gives us adjustment by 2. To get larger adjustments, we will chain many 'twin paths' up, see Figure 5(c). For this purpose, we shall define an adjuster to be a cycle C together with two graphs F_1, F_2 attached to the almost antipodal points v_1, v_2, see Figure 5(b). The point is that large F_1, F_2 help with the chaining process.



(a) x, y-paths with lengths differing by 2.

(b) An adjuster

(c) x, y-paths with varying lengths depending on the path taken through the cycles.

Figure 5: Creating x, y-paths of different lengths using cycles.

Here is the formal definition of an adjuster.

Definition 8.1. A (D, m, k)-*adjuster* $\mathcal{A} = (v_1, F_1, v_2, F_2, A)$ in a graph G consists of vertices $v_1, v_2 \in V(G)$, graphs $F_1, F_2 \subseteq G$ and a vertex set $A \subseteq V(G)$ such that the following hold for some $\ell \leq mk$.

- A, $V(F_1)$ and $V(F_2)$ are pairwise disjoint.
- For each $i \in [2]$, F_i is a (D, m)-expansion around v_i.
- $|A| \leq 10mk$.
- For each $i \in \{0, 1, \ldots, k\}$, there is a v_1, v_2-path in $G[A \cup \{v_1, v_2\}]$ with length $\ell + 2i$.

When $k = 1$, we call $(D, m, 1)$-adjuster a *simple adjuster*.

A crucial step in [6] is to construct *robustly* many adjusters in an expander. We will illustrate here the first step of finding a single simple adjuster, see Lemma 8.6. For this, we first collect some more expansion lemmas. We will omit the proofs of these expansion lemmas.

8.2 Robust expansion lemma

In this subsection, we present a useful robust expansion lemma strengthening Lemma 7.4.

Basically, we shall see below that we can expand a set A in an expander past

- any set X that is much smaller than A, and
- any set Y that is far enough from A in H that A can expand to become much larger than Y before they run into each other, and
- any set Z that is 'thin' around A, that is, it has little intersection with each sphere around A.

The following definition makes it precise what we mean by a 'thin' set.

Definition 8.2. A set V has λ-*limited contact* with a set S in a graph G if, for each $i \in \mathbb{N}$,

$$|N_G(B^{i-1}_{G-S}(V)) \cap S| \leq \lambda i.$$

Lemma 8.3. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$ and $\lambda \in \mathbb{N}$. There is some $d_0 = d_0(\varepsilon_1, \varepsilon_2, \lambda)$ for which the following holds for any $n \geq d \geq d_0$. Suppose H is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(H) \geq d$. Let $m = 32\log^3 n/\varepsilon_1$ and $\ell_0 = (\log\log n)^5$. Let $A \subseteq V(H)$ with $|A| \geq \varepsilon_2 d$ and let $X, Y, Z \subseteq V(H) \setminus A$ be such that the following hold.*

- $|X| \leq |A|\varepsilon(|A|)/4$.
- $B^{\ell_0}_{H-X-Z}(A) \cap Y = \varnothing$ *and* $|Y| \leq m^{300\lambda}$.
- *A has λ-limited contact with Z in H.*

Then,

$$|B^{\ell_0}_{H-X-Y-Z}(A)| > m^{400\lambda}.$$

8.3 Vertex expansions

In order to connect structures together in an expander, we attach a graph say F to the connecting vertex say v say. Provided that F is large, we can then use Lemma 3.5 or Lemma 8.3 to connect $V(F)$ to other structures. Furthermore, F should be 'close' to v so that paths to $V(F)$ can be extended within F to v. Formally, we define

Definition 8.4. Given a vertex v in a graph F, F is a (D, m)-*expansion of* v if $|F| = D$ and $v \in V(F)$ is a distance at most m in F from any other vertex.

The following lemma finds disjoint vertex expansions. It can be proved using Lemma 8.3.

Lemma 8.5. *For each $k \in \mathbb{N}$ and any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists d_0 such that the following holds for each $d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$. Let $m = \frac{32}{\varepsilon_1}\log^3 n$. Let C be a shortest cycle in G, and $x_1, \ldots, x_k$ be distinct vertices in G. For each $i, j \in [k]$, let $1 \leq D_{i,j} \leq \log^k n$. Then, there are graphs $F_{i,j} \subseteq G$, $i, j \in [k]$, such that the following hold.*

- *For each $i, j \in [k]$, $F_{i,j}$ is a $(D_{i,j}, 5m)$-expansion around x_i in G which contains no vertices in $V(C) \setminus \{x_i\}$.*
- *For each $i, j \in [k]$, $V(F_{i,j}) \setminus \{x_i\}$ are pairwise disjoint.*

8.4 Constructing a single simple adjuster

Lemma 8.6. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$ and $k \in \mathbb{N}$, there exists d_0 such that the following is true for each $d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$. Let C be a shortest cycle in G and x_1, x_2 be distinct vertices in $V(G) \setminus V(C)$. Let $m = \frac{200}{\varepsilon_1} \log^3 n$ and $D \leq \log^k n$. Then, G contains a $(D, m, 1)$-adjuster (v_1, F_1, v_2, F_2, A) with $v_1 = x_1$, $v_2 = x_2$ and $V(C) \subseteq A$.*

Proof. Noting that C must have even length, let ℓ_0 be such that $2\ell_0$ is the length of C. Since $\delta(G) \geq d$, we must have $\ell_0 \leq \log n / \log d \leq m$. Pick vertices $x_3, x_4 \in V(C)$ which are distance $\ell_0 - 1$ apart on C and let the paths separating them be R_1 (the shorter one, say) and R_2.

Let $D_{1,1} = D_{2,1} = D$, $D_{1,2} = D_{3,1} = m^3 D$, $D_{2,2} = D_{4,1} = m^2 D$. Apply Lemma 8.5 to x_1, x_2, x_3, x_4 and C with $(m, k)_{8.5} = (m/5, k + 10)$ to get graphs $F_{i,j}$, $i, j \in [2]$ and $F_{3,1}, F_{4,1}$, for which the following hold.

- Each $F_{i,j}$ is a $(D_{i,j}, m)$-expansion around x_i in G which contains no vertices in $V(C) \setminus \{x_i\}$.
- For $i \in [4]$, $j \in [2]$, $V(F_{i,j}) \setminus \{x_i\}$ are pairwise disjoint.

Then, by Corollary 3.5, we can connect $V(F_{1,2})$ and $V(F_{3,1})$, while avoiding $V(C) \cup V(F_{1,1}) \cup V(F_{2,1}) \cup V(F_{2,2}) \cup V(F_{4,1})$, say using the path P' with length at most m. Next, connect $V(F_{2,2})$ and $V(F_{4,1})$ while avoiding $V(C) \cup V(P) \cup V(F_{1,1}) \cup V(F_{2,1})$, say using the path Q' with length at most m. As each $F_{i,j}$ is a $(D_{i,j}, m)$-expansion around x_i, we can extend P' and Q' in them to get a x_1, x_3-path, say P, and a x_2, x_4-path, say Q, each of length at most $3m$.

Set now $v_1 = x_1$, $v_2 = x_2$, $F_1 = F_{1,1}$, $F_2 = F_{2,1}$ and $A = V(P \cup Q \cup R_1 \cup R_2) \setminus \{v_1, v_2\}$. Then, $|A| \leq 2(3m + 1) + 2\ell_0 \leq 10m$ and note that A is disjoint from $V(F_1) \cup V(F_2)$. Letting $\ell = \ell(P \cup R_1 \cup Q)$, note that $P \cup R_1 \cup Q$ and $P \cup R_2 \cup Q$ are v_1, v_2-paths in $G[A \cup \{v_1, v_2\}]$ with length ℓ and $\ell + 2$ respectively. Thus, (v_1, F_1, v_2, F_2, A) is a desired $(D, m, 1)$-adjuster. □

References

[1] J. Halsegrave, J. Kim, H. Liu, Extremal functions for sparse graphs, *preprint*, 2020.

[2] J. Kim, H. Liu, M. Sharifzadeh and K. Staden, Proof of Komlós's conjecture on Hamiltonian subsets. *Proc. Lond. Math. Soc.* **115**(5) (2017), 974–1013.

[3] J. Komlós and E. Szemerédi, Topological cliques in graphs. *Combin. Probab. Comput.* **3** (1994), 247–256.

[4] J. Komlós and E. Szemerédi, Topological cliques in graphs II. *Combin. Probab. Comput.* **5** (1996), 79–90.

[5] H. Liu, R.H. Montgomery, A proof of Mader's conjecture on large clique subdivisions in C_4-free graphs. *J. London Math. Soc.*, **95**(1), (2017), 203–222.

[6] H. Liu, R.H. Montgomery, A solution to Erdős and Hajnal's odd cycle problem, *preprint*, (2020).