A solution to Erdős and Hajnal's odd cycle problem

Hong Liu [*]   Richard Montgomery [†]

October 29, 2020

Abstract

In 1981, Erdős and Hajnal asked whether the sum of the reciprocals of the odd cycle lengths in a graph with infinite chromatic number is necessarily infinite. Let $\mathcal{C}(G)$ be the set of cycle lengths in a graph G and let $\mathcal{C}_{\mathrm{odd}}(G)$ be the set of odd numbers in $\mathcal{C}(G)$. We prove that, if G has chromatic number k, then $\sum_{\ell \in \mathcal{C}_{\mathrm{odd}}(G)} 1/\ell \geq (1/2 - o_k(1)) \log k$. This solves Erdős and Hajnal's odd cycle problem, and, furthermore, this bound is asymptotically optimal.

In 1984, Erdős asked whether there is some d such that each graph with chromatic number at least d (or perhaps even only average degree at least d) has a cycle whose length is a power of 2. We show that an average degree condition is sufficient for this problem, solving it with methods that apply to a wide range of sequences in addition to the powers of 2.

Finally, we use our methods to show that, for every k, there is some d so that every graph with average degree at least d has a subdivision of the complete graph K_k in which each edge is subdivided the same number of times. This confirms a conjecture of Thomassen from 1984.

1   Introduction

Does the chromatic number or the average degree imply anything about the cycle lengths of a graph? For any fixed k, we can never infer the presence of a cycle with length k, as a graph may have arbitrarily high chromatic number yet no such cycle (as Erdős famously showed in 1959 [1]). Can we, however, say something about the density of cycle lengths or infer the existence of a cycle with length in some given infinite set of integers?

We will consider these two questions for both the even cycles and the odd cycles of a graph. From an average degree condition, we can only say something about the even cycles of a graph, as, of course, a graph with high average degree may have no odd cycles. The natural corresponding condition to impose for odd cycles is one on the chromatic number (see, for example, the survey by Verstraëte [24]).

In this paper, we give the first constructions for even cycles with precise lengths using only an average degree condition. We then develop our methods further to find many odd cycle lengths in graphs with a chromatic number condition. This development, while itself novel, relies on the strength of our results on even cycles and cannot be done using the previous results guaranteeing different even cycle lengths.

We now discuss further, and state our results on, even cycles in graphs with an average degree condition (in Section 1.1) and odd cycles in graphs with a chromatic number condition (in Section 1.2). We then give a further application of our work which confirms a conjecture of Thomassen [20] on subdivisions (in Section 1.3). For questions on cycle lengths in graphs under additional conditions, and the background on them, we refer the reader to the comprehensive survey by Verstraëte [24].

1.1 Average degree and even cycle lengths

In 1966, Erdős and Hajnal [8] suggested the harmonic sum of the cycle lengths in a graph as a measure of the density of its cycle lengths. In particular, letting $\mathcal{C}(G)$ be the set of cycle lengths in a graph G, Erdős and Hajnal [8] asked whether

$$\sum_{\ell \in \mathcal{C}(G)} \frac{1}{\ell} \to \infty \quad \text{as} \quad \chi(G) \to \infty, \tag{1}$$

where $\chi(G)$ is the chromatic number of G. As Erdős later wrote [4], they felt that (1) should even hold under the weaker condition $d(G) \to \infty$, where $d(G)$ is the average degree of the graph G. In 1984, Gyárfás, Komlós and Szemerédi [10] confirmed this stronger conjecture by proving that any graph G with average degree d has $\sum_{\ell \in \mathcal{C}(G)} 1/\ell = \Omega_d(\log d)$. If G is a complete balanced bipartite graph with average degree d, then $\sum_{\ell \in \mathcal{C}(G)} 1/\ell = (1/2 + o_d(1)) \log d$, so this lower bound is tight up to the implicit constant. Erdős [4] had previously stated, in 1975, that it was likely that $(1/2 + o_d(1)) \log d$ is the correct asymptotic lower bound, over all graphs G with $d(G) \geq d$, and this remained an open problem.

We say an increasing sequence $\sigma_1, \sigma_2, \sigma_3, \ldots$ of integers is *unavoidable with high average degree* if there is some d such that any graph with average degree at least d has a cycle with length in $\sigma_1, \sigma_2, \sigma_3, \ldots$. In 1977, Bollobás [2] confirmed a conjecture of Burr and Erdős by showing that, if σ_i forms an arithmetic progression containing even numbers, then σ_i is unavoidable with high average degree. Solving a problem of Erdős, in 2005 Verstraëte [23] showed that some unavoidable sequence with density 0 must exist, without giving an explicit sequence.

In 2008, Sudakov and Verstraëte [18] showed that many increasing sequences of integers are unavoidable in all but (potentially) exceptionally sparse graphs. In particular, they showed that, for any exponentially bounded increasing sequence of even integers σ_i (that is, where $\sigma_{i+1} \leq C\sigma_i$ for each $i \in \mathbb{N}$ and some fixed $C > 1$), any n-vertex graph G with $\sigma_i \notin \mathcal{C}(G)$ for each $i \in \mathbb{N}$ must have average degree at most $e^{O(\log^* n)}$, where $\log^* n$ is the iterated logarithm function. In 1984, Erdős [6] had asked whether the powers of 2 are unavoidable with high average degree, but, despite the results quoted here, there remained no explicit sequence with density 0 which was known to be unavoidable with high average degree (or even unavoidable with high chromatic number, as defined analogously in Section 1.2).

In this paper, we introduce new techniques for constructing even cycles while controlling their length. This allows us to find, in any graph G, a long interval of consecutive even numbers in $\mathcal{C}(G)$, as follows.

Theorem 1.1. *There is $d_0 > 0$ such that the following holds. If G is a graph with average degree $d \geq d_0$, then, there is some $\ell \geq d/(10 \log^{10} d)$ such that $\mathcal{C}(G)$ contains every even integer in $[\log^8 \ell, \ell]$.*

2

From the density of the even numbers in the interval $[\log^8 \ell, \ell]$ as ℓ increases, we get immediately the following improvement of the result of Gyárfás, Komlós and Szemerédi [10], which confirms the asymptotically correct lower bound on the harmonic sum of $\mathcal{C}(G)$, as conjectured by Erdős [4].

Corollary 1.2. *If a graph G has average degree d, then*

$$\sum_{\ell \in \mathcal{C}(G)} \frac{1}{\ell} \geq \left(\frac{1}{2} - o_d(1) \right) \log d.$$

By Theorem 1.1, any avoidable sequence of cycle lengths must continue to avoid some intervals $[\log^8 \ell, \ell]$ as $\ell \to \infty$. Thus, many increasing sequences of even integers are unavoidable with high average degree, as follows.

Corollary 1.3. *There is some $d_0 > 0$ such that the following holds. Given any infinite sequence σ_i, $i \in \mathbb{N}$, of increasing even integers with $\sigma_{i+1} \leq \exp(\sigma_i^{1/10})$ for each $i \in \mathbb{N}$, any graph G with average degree at least $\max\{d_0, \sigma_1^2\}$ has some $i \in \mathbb{N}$ with $\sigma_i \in \mathcal{C}(G)$.*

In particular, this answers the question of Erdős [6] mentioned above by showing that the powers of 2 are unavoidable with high average degree, as, furthermore, is any exponentially bounded sequence of increasing even numbers. This latter implication answers a question of Sudakov and Verstraëte [18]; for further questions on specific sequences answered by Corollary 1.3, we refer the readers to [7].

On the other hand, the sequence defined by $\sigma_1 = 1$ and $\sigma_{i+1} = 2(i+1)^{\sigma_i}$ for each $i \in \mathbb{N}$ is known to be avoidable (see [18]). Corollary 1.3 is thus optimal up to the fraction $1/10$ in the exponent, though it may hold with $1/10$ replaced by $1 - o_i(1)$. We have not optimised our methods to maximise the fraction $1/10$, but note that doing so could not increase it beyond $1/3$ (see Section 2.4.4).

1.2   Chromatic number and odd cycle lengths

Consider the set of odd cycle lengths in G, by letting $\mathcal{C}_{\mathrm{odd}}(G) = \{\ell \in \mathcal{C}(G) : \ell \text{ is odd}\}$. As noted above, a natural condition to impose in search of odd cycles is one on the chromatic number rather than the average degree, and, for example, Gyárfás [9] proved in 1992 that a graph with chromatic number at least $2k + 1$ must have $|\mathcal{C}_{\mathrm{odd}}(G)| \geq k$. In 1981, Erdős and Hajnal [5] asked whether

$$\sum_{\ell \in \mathcal{C}_{\mathrm{odd}}(G)} \frac{1}{\ell} \to \infty \quad \text{as} \quad \chi(G) \to \infty. \tag{2}$$

As Erdős noted [6], this gives a 'much deeper question' than the corresponding question for all cycle lengths and chromatic number that we considered in Section 1.1. Indeed, the only relevant result towards this has been by Sudakov and Verstraëte [19], who showed that $\sum_{\ell \in \mathcal{C}_{\mathrm{odd}}(G)} 1/\ell \to \infty$ if the independence ratio of G is not extremely small compared to its number of vertices.

Building on our methods for even cycles, we use the precision of these methods to construct a large interval of cycle lengths using a high chromatic number, as follows.

Theorem 1.4. *For each $\varepsilon > 0$, there is some $k_0 \in \mathbb{N}$ such that the following holds for each $k \geq k_0$. If G is a graph with chromatic number k, then, for some $\ell \in \mathbb{N}$, $\mathcal{C}(G)$ contains every odd integer in $[\ell, \ell \cdot k^{1-\varepsilon}]$.*

3

As the harmonic sum of the odd integers in $[\ell, \ell \cdot k^{1-\varepsilon}]$ diverges as $k \to \infty$ (for all values of ℓ), this solves Erdős and Hajnal's odd cycle problem by confirming (2). Furthermore, in combination with Theorem 1.1, we get the following immediate lower bound for the harmonic sum of the cycle lengths of any specific residue, which answers another question of Erdős [6].

Corollary 1.5. *Let $a, b \in \mathbb{N}$, and let $\mathcal{C}_{a,b}(G) = \{\ell \in \mathcal{C}(G) : \ell \equiv a \mod b\}$. If G has chromatic number k, then*

$$\sum_{\ell \in \mathcal{C}_{a,b}(G)} \frac{1}{\ell} \geq \left(\frac{1}{b} - o_k(1)\right) \log k.$$

We say an increasing sequence $\sigma_1, \sigma_2, \sigma_3, \ldots$ of integers is *unavoidable with high chromatic number* if there is some k such that any graph with chromatic number at least k has a cycle with length in $\sigma_1, \sigma_2, \sigma_3, \ldots$. There have previously been no non-trivial sequences of increasing odd integers which are known to be unavoidable with high chromatic number, but, similarly to in Section 1.1, Theorem 1.4 immediately implies that many such sequences are unavoidable with high chromatic number, as follows.

Corollary 1.6. *Given $C \in \mathbb{N}$, there exists $k_0 \in \mathbb{N}$ such that the following holds. Let $\sigma_1, \sigma_2, \ldots$ be an infinite increasing sequence of odd integers such that $\sigma_{i+1} \leq C\sigma_i$ for each $i \in \mathbb{N}$. Then, every graph G with chromatic number at least $\max\{k_0, \sigma_1^2\}$ has some $i \in \mathbb{N}$ with $\sigma_i \in \mathcal{C}(G)$.*

We remark that our proof in fact shows that Theorem 1.4 (and hence also Corollaries 1.5 and 1.6) holds with the weaker hypothesis that the graph G is 'not too close' to being bipartite. That is, as seen in Section 5.1.1, we use only the condition that G contains a subgraph with average degree $\Omega(k)$ despite the removal of any bipartite graph.

1.3   Balanced subdivisions

A *subdivision* of a graph G is obtained by replacing each edge of G by a path, such that the new paths are internally vertex disjoint. This notion has played a central role in topological graph theory since the seminal result of Kuratowski in 1930 that a graph is planar if and only if it does not contain a subdivision of the complete graph with five vertices or a subdivision of the complete bipartite graph with three vertices on each side [14].

In 1967, Mader [16] proved that, for each $k \in \mathbb{N}$, there is some $d = d(k)$ such that every graph with average degree at least d contains a subdivision of the complete graph K_k. After improved bounds on $d(k)$ by Mader [17], and Komlós and Szemerédi [12], Bollobás and Thomason [3] proved that, optimally, we may take $d(k) = O(k^2)$. Komlós and Szemerédi [13] later improved their own methods to give an independent proof of this, and the graph expansion methods they introduced (see Section 2.2) form the basis for many constructions in sparse graphs both here and elsewhere (see, for example, [11, 15]).

Our constructive approach to controlling the length of cycles also allows us to control the length of paths, and thus construct subdivisions in which each edge is replaced by a path of the same length. For integers $\ell, k \in \mathbb{N}$, denote by $\mathsf{TK}_k^{(\ell)}$ a subdivision of a complete graph K_k in which each edge is replaced by a path with length ℓ. We say that $\mathsf{TK}_k^{(\ell)}$ is a *balanced subdivision* of K_k.

In 1984, Thomassen [20] (see also [21], [22]) conjectured that, for each $k \in \mathbb{N}$, high average degree in a graph is sufficient to guarantee a balanced subdivision of K_k. This was open even for $k = 4$. We confirm Thomassen's conjecture, as follows.

4

Theorem 1.7. *For each $k \in \mathbb{N}$, there exists d such that every graph with average degree at least d contains a $\mathsf{TK}_k^{(\ell)}$ for some $\ell \in \mathbb{N}$.*

We note that it is conceivable that there is some $\varepsilon > 0$ such that any graph with average degree at least d in fact contains a $\mathsf{TK}_{\varepsilon\sqrt{d}}^{(\ell)}$, which would be optimal up to the value of ε. Furthermore, this may be provable using appropriate extensions of our methods (in particular along the lines of the techniques in [15]). However, though the balanced subdivision problem was the original focus of our work, we do not push these techniques further at the expense of a clear presentation of the new cycle construction techniques.

2 Structure and proof sketches

After describing our notation in Section 2.1, in Section 2.2 we recall the graph expansion concepts introduced by Komlós and Szemerédi in [12, 13]. This allows us to give our main theorem, Theorem 2.7, in Section 2.3 and derive Theorem 1.1, while noting how we use similar methods to prove Theorem 1.7. We then sketch the proof of Theorem 2.7 in Section 2.4. The proof of Theorem 1.4 is sketched in Section 2.5. Finally, in Section 2.6, we highlight one particular innovation for our constructions in expanders, which may prove useful elsewhere.

2.1 Notation

Let G be a graph, let $v \in V(G)$ be a vertex and let $W \subseteq V(G)$ be a set of vertices. We write $|G| = |V(G)|$ for the order of the graph. Let $\delta(G), d(G)$ and $\Delta(G)$ be the minimum, average and maximum degree of G respectively. Denote by $N_G(v, W)$ the set of neighbours of v in W, and denote by $d_G(v, W) = |N_G(v, W)|$ the degree of v into W in G. Denote the (external) neighbourhood of W by $N(W) = (\cup_{v \in W} N(v)) \setminus W$. Let $G[W] \subseteq G$ be the induced subgraph of G with vertex set W. Denote by $G - W$ the induced subgraph $G[V(G) \setminus W]$. We write $N_G^0(W) = W$, and, for each integer $k \geq 1$, let $B_G^k(W) = \cup_{0 \leq j \leq k} N_G^j(W)$ the ball of radius k around W in G, that is, the set of all vertices a graph distance at most k to W. We let $B(W) = B^1(W)$.

Given graphs G and H, the graph $G \cup H$ has vertex set $V(G) \cup V(H)$ and edge set $E(G) \cup E(H)$. Denote by $G \setminus H$ the graph with vertex set $V(G)$ and edge set $E(G) \setminus E(H)$. For a collection $\mathcal{P}$ of graphs, denote by $|\mathcal{P}|$ the number of graphs in $\mathcal{P}$ and write $V(\mathcal{P}) = \cup_{G \in \mathcal{P}} V(G)$.

For a path P, let its length be $\ell(P)$. Where we say P is a path from a vertex set A to a disjoint vertex set B, we mean that P has one endvertex in each of A and B, and no internal vertices in $A \cup B$. For each $\ell \in \mathbb{N}$ and $k > 0$, $\mathsf{TK}_k^{(\ell)}$ is a subdivision of a complete graph $K_{\lfloor k \rfloor}$ in which each edge is replaced by a path with length ℓ.

Many of our results that build to the theorems stated in Section 1 state that for each $\varepsilon_1, \varepsilon_2 > 0$ (and perhaps each $k \in \mathbb{N}$), there is some $d_0(\varepsilon_1, \varepsilon_2)$ (or $d_0(\varepsilon_1, \varepsilon_2, k)$) such that some property holds for $n \geq d \geq d_0$. For brevity, we do not calculate the function $d_0(\varepsilon_1, \varepsilon_2)$ (or $d_0(\varepsilon_1, \varepsilon_2, k)$) and assume implicitly in our proofs that n and d are as large as needed, depending on ε_1 and ε_2 (and perhaps k) — where it may help the reader we recall this at various points in the proofs. The results from now on in this paper are stated and proved in order, so that these functions could be chosen sequentially through the paper.

We omit the subscript G when the underlying graph G is clear. When it is not essential, we omit the floors and ceilings. All logarithms are natural.

2.2 Komlós-Szemerédi graph expansion

An expansion property in a graph G is typically one in which every set $X \subseteq V(G)$ satisfies $|N_G(X)| \geq \varepsilon|X|$ for some function ε depending on the size of X. Expansion is key to all of our constructions. Our expansion must therefore exist in some subgraph of any graph. Following Komlós and Szemerédi [12, 13], effectively we use the strongest expansion that can be found in some subgraph of any graph, based on its average degree.

Definition 2.1. For each $\varepsilon_1 > 0$ and $k > 0$, a graph G is an (ε_1, k)-*expander* if

$$|N(X)| \geq \varepsilon(|X|, \varepsilon_1, k) \cdot |X|$$

for all $X \subseteq V(G)$ with $k/2 \leq |X| \leq |G|/2$, where

$$\varepsilon(x, \varepsilon_1, k) := \begin{cases} 0 & \text{if } x < k/5, \\ \varepsilon_1/\log^2(15x/k) & \text{if } x \geq k/5. \end{cases} \tag{3}$$

Whenever the choices of ε_1, k are clear, we omit them and write $\varepsilon(x)$ for $\varepsilon(x, \varepsilon_1, k)$.

If an n-vertex graph G is an (ε_1, k)-expander and $X \subseteq V(G)$ has size at least $k/2$, then $B_G^i(X)$ increases as i increases, until the set contains at least $n/2$ vertices. The rate of expansion, $|N_G(B_G^i(X))|/|B_G^i(X)| \geq \varepsilon(|B_G^i(X)|, \varepsilon_1, k)$ guaranteed by the expansion condition decreases as i increases (see, for example, [13]). That is, $\varepsilon(x, \varepsilon_1, k)$ decreases as $x \geq k/2$ increases. However, $\varepsilon(x, \varepsilon_1, k) \cdot x$ increases as x does, so the lower bound from expansion we have for $|N_G(B_G^i(X))|$ increases as i increases.

As Komlós and Szemerédi [13] showed, every graph G contains an expander with comparable average degree to G, as follows.

Theorem 2.2 ([13]). *There exists some $\varepsilon_1 > 0$ such that the following holds for every $k > 0$. Every graph G has an (ε_1, k)-expander subgraph H with $d(H) \geq d(G)/2$ and $\delta(H) \geq d(H)/2$.*

Note that, in Theorem 2.2, the expander subgraph H can be much smaller than the original graph G in size. Indeed, G could be the disjoint union of many copies of such a graph H.

We use expansion to expand and connect vertex sets, creating paths to construct cycles of varying lengths. A typical use is the following result of Komlós and Szemerédi, though we use the comparable Lemma 3.4.

Lemma 2.3 ([13]). *Let $\varepsilon_1, k > 0$. If G is an n-vertex (ε_1, k)-expander, then any two vertex sets, each of size at least $x \geq k$, are of distance at most $\frac{2}{\varepsilon_1}\log^3(15n/k)$ apart. This remains true even after deleting $x \cdot \varepsilon(x)/4$ arbitrary vertices from G.*

It is convenient to work in a bipartite graph; to do this we use the following simple and well known result.

Proposition 2.4. *Within any graph G there is a bipartite subgraph H with $d(H) \geq d(G)/2$.*

Combining this with Theorem 2.2, we get the following immediate corollary.

Corollary 2.5. *There exists some $\varepsilon_1 > 0$ such that the following holds for every $\varepsilon_2 > 0$ and $d \in \mathbb{N}$. Every graph G with $d(G) \geq 8d$ has a bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander subgraph H with $\delta(H) \geq d$.* $\quad\square$

2.3 A stronger version of Theorem 1.1

We will prove Theorem 1.1 in the slightly stronger form of Theorem 2.7 below, which applies to an expander. Combining this with Corollary 2.5 easily gives Theorem 1.1, as shown below. We use Theorem 2.7 to prove Theorem 1.4, as outlined in Section 2.5. Theorem 1.7 is proved using very similar methods to Theorem 2.7, and we comment on this below.

In order to state Theorem 2.7, we use the follow definition, which records whether there will be even or odd length paths between two vertices u and v in a connected bipartite graph H.

Definition 2.6. For any connected bipartite graph H and $u, v \in V(H)$, let

$$\pi(u, v, H) = \begin{cases} 0 & \text{if } u = v, \\ 1 & \text{if } u \text{ and } v \text{ are in different vertex classes in the (unique) bipartition of } H, \\ 2 & \text{if } u \text{ and } v \text{ are in the same vertex class and } u \neq v. \end{cases}$$

Note that, for example by Lemma 2.3, any (ε_1, k)-expander subgraph with minimum degree at least k is connected, and this will allow us to use this definition for the bipartite expanders we use.

In common with many of our results in the rest of this paper, Theorem 2.7 applies only to $\mathsf{TK}_\ell^{(2)}$-free graphs for some ℓ (often $\ell = d/2$). A subdivision of the complete graph on ℓ vertices, with each edge subdivided into a path of length 2, has many different even cycle lengths, and many different path lengths between pairs of vertices. As our aim in applying Theorem 2.7 is to find only *some* subgraph with this property (see, for example, the proof of Corollary 5.1), we need look no further than such a subdivision. Why our constructions require the graph to be $\mathsf{TK}_\ell^{(2)}$-free is commented on in Section 3.7.

Theorem 2.7. *There exists $\varepsilon_1 > 0$, such that, for each $0 < \varepsilon_2 < 1/5$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that H is a $\mathsf{TK}_{d/2}^{(2)}$-free bipartite n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(H) \geq d$. Let $x, y \in V(H)$ be distinct, and let*

$$\ell \in [\log^7 n, n/\log^{10} n]$$

satisfy $\pi(x, y, H) = \ell \mod 2$.

Then, H contains an x, y-path with length ℓ.

Theorem 1.1 follows from Theorem 2.7 and Corollary 2.5, as follows.

Proof of Theorem 1.1. Let $\varepsilon_1 > 0$ be such that the condition in Corollary 2.5 applies, and let $\varepsilon_2 = 1/100$. Let d_0 be large (see Section 2.1), and in particular large enough that the property in Theorem 2.7 holds for ε_1 and $\varepsilon_2' = 8\varepsilon_2$, for each $d \geq d_0/8$.

Let G be a graph with average degree $d \geq d_0$ and let $\bar{d} = d/8$. By the property from Corollary 2.5, G contains a bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander subgraph H with $\delta(H) \geq \bar{d}$. As $\bar{d} = d/8$, H is an $(\varepsilon_1, 8\varepsilon_2 \bar{d})$-expander. If H contains a $\mathsf{TK}_{\bar{d}/2}^{(2)}$, then H contains every even cycle length between 6 and $\bar{d}$. As $\bar{d} = d/8$ is large, the property in the theorem holds with $\ell = \bar{d}$ as $\log^8 \bar{d} \geq 6$.

Assume then that H is $\mathsf{TK}_{\bar{d}/2}^{(2)}$-free. As $\delta(H) \geq \bar{d} > 0$, we can pick distinct vertices $x, y \in V(H)$ such that $xy \in E(H)$. Note that $\pi(x, y, H) = 1$. Let $n = |H|$ and $\ell = n/\log^{10} n \geq \bar{d}/\log^{10} \bar{d} \geq d/(10\log^{10} d)$, as $\bar{d} = d/8$ is large.

For every even $\ell' \in [\log^8 \ell, \ell]$, $(\ell' - 1)$ is an odd number in $[\log^7 n, n/\log^{10} n]$. Then, by the property from Theorem 2.7 applied with x and y, there is an x, y-path with length $\ell' - 1$ in H, and therefore a cycle with length ℓ' in H. □

7

For Theorem 1.7, essentially, we find a copy of $\mathsf{TK}_k^{(\ell)}$, for some $\ell \in \mathbb{N}$, by first taking an expander subgraph and k distinct vertices within it to be the *core vertices*. Core vertices in a K_k-subdivision are the vertices which are not interior vertices of a path which replaces an edge. Using the same construction as for Theorem 2.7 multiple times, we then find internally vertex disjoint paths with the same length between each pair of core vertices. This is done in Section 4.4.

2.4 Proof sketch for Theorem 2.7

To discuss the proof of Theorem 2.7, let H be a $\mathsf{TK}_{d/2}^{(2)}$-free bipartite n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander, with $0 < \varepsilon_1, \varepsilon_2 < 1$, such that $\delta(H) \geq d$, and let $x, y \in V(H)$ be distinct.

Our aim is to find a sequence of x, y-paths in H whose lengths increase by 2 each time. In Section 2.4.1 we describe how one cycle can be used to find two paths with the same endvertices and lengths differing by 2, and how a connected sequence of cycles can create a longer sequence of paths with lengths increasing by 2 each time. In Section 2.4.2, we discuss how a sequence of cycles can be connected, and introduce the concept of an *adjuster*. In Section 2.4.3, we describe how a polylogarithmic number of adjusters can be used to find long paths with the lengths required by Theorem 2.7. In Section 2.4.4, we discuss the natural limitations of these methods.

2.4.1 Creating a little adjustment with small cycles

Our proof of Theorem 2.7 is based on the following simple idea. Suppose we can find in H a short cycle C, with length 2ℓ say, which does not contain x or y. Take two vertices v_1 and v_2 a distance $\ell - 1$ apart on C. Then, C contains one v_1, v_2-path with length $\ell - 1$ and another with length $\ell + 1$. If we can find, using new internal vertices, disjoint paths from x to v_1 and from y to v_2, then we have two x, y-paths whose lengths differ by 2 (see Figure 1(a)).

If we can chain together many cycles between x and y in this fashion, then, by choosing the length of the path we take around each cycle, we can vary the length of the path between x and y in increments of 2 (see Figure 1(c)).
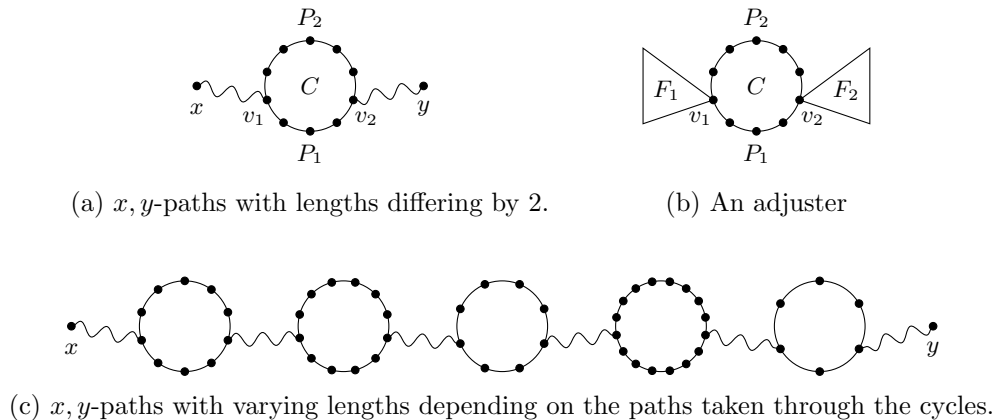

(a) x, y-paths with lengths differing by 2.   (b) An adjuster

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

Figure 1: Creating x, y-paths of different lengths using cycles.

2.4.2 Connecting the cycles

In the situation above, if C is a shortest cycle in H, then it is not too difficult to connect x to v_1 and y to v_2 (relabelling v_1 and v_2 if necessary). Indeed, for each $i \geq 0$ and $v \in V(H)$, any neighbours of the ball $B^i_{G-V(C)+v}(v)$ in $V(C)$ must be within distance $2i+2$ of each other on the cycle C, for otherwise H contains a shorter cycle than C. Given our expansion conditions, this will mean that $B^i_{G-V(C)+v_1}(v_1)$ and $B^i_{G-V(C)}(x)$ will both expand as i increases (see Section 3.1). Expanding these sets until they intersect allows us to find a short path from v_1 to x while avoiding $V(C) \setminus \{v_1\}$. Roughly speaking, setting aside a shortest path P from $\{v_1\}$ to $\{x, y\}$ which avoids $V(C) \setminus \{v_1\}$, we then expand around v_2 and the vertex left in $\{x, y\} \setminus V(P)$ while avoiding $(V(P) \cup V(C)) \setminus \{v_2\}$ to find the second path described above.

Connecting multiple cycles is more difficult, simply as there are more vertices to avoid. Due to this, instead of cycles, we will find structures we call *simple adjusters*. These correspond to the cycle C, vertices v_1 and v_2, as well as vertex disjoint subgraphs $F_1, F_2 \subseteq H$ such that $V(F_1) \cap V(C) = \{v_1\}$ and $V(F_2) \cap V(C) = \{v_2\}$ (see Figure 1(b) and Definition 4.1). The graph F_1 (and analogously F_2) has the property that every vertex is a distance $O_{\varepsilon_1}(\log^3 n)$ away from v_1 in F_1, and therefore a path leading into F_1 can be extended with a few additional vertices from $V(F_1)$ to one ending with v_1. Furthermore, $|F_1|$ will be comfortably larger than C so that F_1 can be expanded and connected while avoiding $V(C) \setminus \{v_1\}$. The distance $O_{\varepsilon_1}(\log^3 n)$ here comes from an application of Lemma 3.4.

The heart of our paper is the robust construction of these simple adjusters – that is, having some set of vertices U which is not too large, we find an adjuster in $H - U$. More discussion of this construction we defer till Section 4, but we will highlight the key innovation that allows this in Section 2.6.

2.4.3 Increasing the size of the interval of path lengths

In an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander H, any two vertices x, y are a distance at most $m = O_{\varepsilon_1}(\log^3 n)$ apart (see Lemma 3.4). Using fairly straightforward methods (see those in Section 3.6), we can use the expansion conditions to find, for any $20m \leq \ell \leq n/\log^{10} n$, a path with length between $\ell - 20m$ and ℓ between x and y (see Corollary 3.15). If we can find such a path which contains $10m$ simple adjusters, then, as long as $\ell = \pi(x, y, H) \mod 2$, we can use these adjusters to increase the path by length 2 until it has length exactly ℓ (see Section 4.3).

2.4.4 The limitations of our methods

Our methods are limited by the length of the paths we find between vertex pairs using the Komlós-Szemerédi graph expansion. Here, we can only guarantee a path with length at most $m = \Theta_{\varepsilon_1}(\log^3 n)$. Therefore the lower bound of the interval of path/cycle lengths in Theorem 2.7 and Theorem 1.1 cannot be reduced below $\Omega(\log^3 n)$ while using connection methods with this graph expansion. Correspondingly, optimising our methods cannot increase the fraction $1/10$ in Corollary 1.3 beyond $1/3$.

For simplicity, we connect together around $m = \Theta_{\varepsilon_1}(\log^3 n)$ cycles which can each be used to adjust the length of the paths we find by 2. We note that, by using different cycles to adjust the length by different amounts, we could use perhaps $O(\log m) = O_{\varepsilon_1}(\log \log n)$ cycles instead of around m cycles. However, this saving does not reach a plausible optimal bound as we cannot guarantee our connecting paths are any shorter.

2.5 Outline of the proof of Theorem 1.4

Note that in a graph H which is a copy of $\mathsf{TK}_d^{(2)}$, if $d \geq 3$, then every edge between two vertices u, v can be replaced with a path with any odd length in $[5, 2d-1]$. If d is large, then this includes paths with any odd length in $[\log^8 d, d]$. Using Corollary 2.5 and Theorem 2.7, then, in any graph with average degree at least $16d$, we can find a bipartite subgraph H satisfying the following property.

P There is some $\ell_H \geq d/\log^{10} d$ such that each edge uv in H can be replaced with a path with any odd length in $[\log^8 \ell_H, \ell_H]$.

Now, suppose G has chromatic number at least $300d$ and find in G a maximal collection $\mathcal{H}$ of edge disjoint bipartite graphs H satisfying **P** with some integer ℓ_H. We can show that the chromatic number of the union of these graphs $\cup_{H \in \mathcal{H}} H$ is at least 3, for otherwise $G \setminus (\cup_{H \in \mathcal{H}} H)$ has high enough degree that some other bipartite subgraph satisfying **P** must exist, contradicting the maximality of $\mathcal{H}$. Therefore, $\cup_{H \in \mathcal{H}} H$ contains some odd cycle, C say. Our aim is to use **P** to take many edges e of C, and, where $H_e \in \mathcal{H}$ is such that $e \in E(H_e)$, replace e with odd paths in H_e with many different lengths. This would allow us to transform C into odd cycles of different lengths. The challenge is to do this so that all of the paths replacing these edges are vertex disjoint. This is possible by taking a certain minimal odd cycle C and replacing edges e corresponding to vertex disjoint graphs $H_e \in \mathcal{H}$. This is discussed in more detail at the start of Section 5, before Theorem 1.4 is then proved.

2.6 A new construction method: Robust construction of gadgets.

Here, we will highlight an innovation for constructions using Komlós-Szemerédi expansion. Roughly speaking, this technique overcomes the sublinear property of the expansion we work with. For more details, see Section 3.3. Recall from (3) that in an expander H, any vertex set $A \subseteq V(G)$ of suitable size has neighbourhood with size $\Omega(|A|/\log^2 |A|)$, expanding with at least a factor of $\varepsilon(|A|) = \Theta(1/\log^2 |A|)$.

Our aim is to find some special subgraph – let us call it here a *gadget* – in an expander H while avoiding a vertex set W. Roughly speaking, it is natural to pick a vertex v in $H - W$ and try to construct a gadget locally in H around v using the expansion property. If this fails there will then be some vertex set A_v containing v which does not expand in $H - W$, not even by a reduced factor of $\varepsilon(|A_v|)/10$. If we fail repeatedly, then we find many disjoint sets A_v, say for the vertices $v \in V$, which do not expand in $H - W$ by a factor of $\varepsilon(|A_v|)/10$. Eventually, $\cup_{v \in V} A_v$ will be much larger than W, so it also expands in $H - W$ by a factor of a least $\varepsilon(|\cup_{v \in V} A_v|)/2$. However, as the expansion function in Section 2.2 is sublinear in x, it is not a contradiction that $\cup_{v \in V} A_v$ expands, yet no set A_v does even if the expansion factor required is reduced by a factor of 5.

Instead, we reach a contradiction by exploiting our particular circumstance. We will have that W is polylogarithmic size (in n), that $|V| \geq n^{1/2}$, and the $\mathsf{TK}_{d/2}^{(2)}$-free condition will allow us to have that $|N_H(A_v, W)| \leq |A_v|^2$ for each $v \in V$. Note that, as, for each $v \in V$, A_v does not expand in $H - W$ by a factor of $\varepsilon(|A_v|)/10$, the set A_v must have size at most $|W|\log^3 n$.

Now, firstly, if many of the sets A_v, $v \in V$, have size at least $(\log n)^{1/3}$, then we find a set $V' \subseteq V$ indexing some of these vertices so that $|W|\log^3 n \leq |\cup_{v \in V'} A_v| \leq 2|W|\log^3 n$. Then, $A := \cup_{v \in V'} A_v$ is large enough to expand past W, and hence expand in $H - W$ with a factor of least $\varepsilon(|A|)/2$. As, for each $v \in V'$, $\log |A_v| \approx \log \log n \approx \log |A|$, the expansion function ε for these sets

is essentially the same – near enough that we can get a contradiction from each set A_v, $v \in V'$, not expanding in $H - W$ by a factor of at least $\varepsilon(|A_v|)/10$ yet $A = \cup_{v \in V'} A_v$ expanding in $H - W$ by a factor of at least $\varepsilon(|A|)/2$.

Thus, we can assume that most of the sets A_v, $v \in V$, have size at most $(\log n)^{1/3}$. This gives us, for such $v \in V$, that $|N_H(A_v, W)| \leq |A_v|^2 \leq (\log n)^{2/3}$. As W is polylogarithmic in size, we have at most $|W|^{(\log n)^{2/3}} = n^{o(1)}$ subsets of W with size at most $(\log n)^{2/3}$. Thus, for some $r \leq (\log n)^{1/3}$, there must be at least r^2 vertices $v \in V$ – say those in V'' – such that $|A_v| = r$ for each $v \in V''$ and, over $v \in V''$, $N_H(A_v, W)$ is the same set, $W' \subseteq W$ say. Let $A = \cup_{v \in V''} A_v$, so that $W' = \cup_{v \in V''} N_H(A_v, W) = N_H(A, W)$, and, for each $v \in V''$, $|A_v| = r$, and $|W'| \leq |A_v|^2 = r^2$. As $|A| = r|V''| = r^3$, and $|N_H(A, W)| = |W'| \leq r^2$, A does not expand well into W in H, and hence must expand in $H - W$ by a factor of at least $\varepsilon(|A|)/2$. Again, we have that A_v does not expand in $H - W$ by a factor of at least $\varepsilon(|A_v|)/10$ and $\log|A_v| \approx \log|A|$, for each $v \in V''$. As before, this gives a contradiction.

3 Preliminary expansion results

In this section we cover the preliminary results on expansion that we use in Section 4. The main results of this section are Lemmas 3.2, 3.4, 3.5, 3.11, 3.12 and 3.14 which we prove in Sections 3.1, 3.2, 3.3, 3.4, 3.5 and 3.6 respectively.

3.1 Expanding while avoiding sets

In an expander H, we often want to expand a vertex set A, while avoiding another set X. We can do this if the set satisfies one of three conditions (matching the conditions **A1**–**A3** in Lemma 3.2).

1. Firstly, if X is much smaller than A.

2. Secondly, if X is far enough (in graph distance) from A in H that A expands to become much larger than X before it encounters X.

3. Finally, if X does not intersect too much with each sphere around A during the expansion. That is, if $N_H(B^i_{H-X}(A))$ does not intersect too much with X for each $i \geq 0$.

For the last condition above, if $B^i_{H-X}(A)$ grows with i, then we can permit the intersection of $N_H(B^i_{H-X}(A))$ with X to increase as i increases. This condition is formally captured in the following notion, before we give our main lemma expanding a set vertex A while avoiding some other vertex sets.

Definition 3.1. A vertex set A has *k-limited contact* with a vertex set X in a graph H if, for each $i \in \mathbb{N}$,
$$|N_H(B^{i-1}_{H-X}(A)) \cap X| \leq ki.$$

Lemma 3.2. *Let $0 < \varepsilon_1, \varepsilon_2 < 1$ and $k \in \mathbb{N}$. There is some $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ for which the following holds for each $n \geq d \geq d_0$.*

Suppose H is an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander. Let $m = \frac{16}{\varepsilon_1} \log^3 n$ and $\ell_0 = (\log \log n)^5$. Let $A \subseteq V(H)$ with $|A| \geq \varepsilon_2 d/2$ and let $X, Y, Z \subseteq V(H) \setminus A$ be such that the following hold.

A1 $|X| \leq |A|\varepsilon(|A|)/4$.

A2 $B_{H-X-Z}^{\ell_0}(A) \cap Y = \varnothing$ *and* $|Y| \leq m^{300k}$.

A3 *A has k-limited contact with Z in H.*

Then,

(i) $|B_{H-X-Y-Z}^{\ell_0}(A)| > m^{400k}$, *and*

(ii) $|B_{H-X-Y-Z}^{m}(A)| > n/2$.

Proof. We first prove (i). Note that, by **A2**, we have $B_{H-X-Y-Z}^{\ell_0}(A) = B_{H-X-Z}^{\ell_0}(A)$. Let $F = H - X - Z$, and suppose that $|B_F^{\ell_0}(A)| \leq m^{400k}$, for otherwise (i) holds. We will show the following claim.

Claim 3.3. *For each $0 \leq r \leq \ell_0 - 1$,*

$$|N_F(B_F^r(A))| \geq \frac{1}{4}|B_F^r(A)| \cdot \varepsilon(|B_F^r(A)|). \tag{4}$$

Given Claim 3.3, we reach a contradiction as follows, and thus (i) must hold. For each $0 \leq r < \ell_0$, we have $|B_F^r(A)| \leq |B_F^{\ell_0}(A)| \leq m^{400k}$, and hence, as $n \geq d \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large,

$$\varepsilon(|B_F^r(A)|) \geq \varepsilon(m^{400k}) = \frac{\varepsilon_1}{\log^2(15m^{400k}/\varepsilon_2 d)} \geq \frac{4}{(\log\log n)^3}.$$

Therefore, for each $0 \leq r < \ell_0$, by Claim 3.3,

$$|B_F^{r+1}(A)| = |B_F^r(A)| + |N_F(B_F^r(A))| \geq \left(1 + \frac{\varepsilon(|B_F^r(A)|)}{4}\right)|B_F^r(A)| \geq \left(1 + \frac{1}{(\log\log n)^3}\right)|B_F^r(A)|.$$

Therefore,

$$|B_F^{\ell_0}(A)| \geq \left(1 + \frac{1}{(\log\log n)^3}\right)^{\ell_0}|A| \geq \exp\left(\frac{\ell_0}{2(\log\log n)^3}\right) = \exp\left(\frac{(\log\log n)^2}{2}\right) = \omega(m^{400k}), \tag{5}$$

which contradicts our assumption that (i) does not hold as $n \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large. For (i), it is left then to prove Claim 3.3.

Proof of claim. We prove this by induction on r. For the base case $r = 0$, as $x \cdot \varepsilon(x)$ increases in $x \geq \varepsilon_2 d/2$, we have

$$|A|\varepsilon(|A|) \geq \frac{\varepsilon_2 d}{2} \cdot \varepsilon(\varepsilon_2 d/2) = \frac{\varepsilon_1\varepsilon_2 d}{2\log^2(15/2)} \geq 4k,$$

where we have used that $d \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large. In combination with **A3**, we have $|N_H(A) \cap Z| \leq k \leq |A|\varepsilon(|A|)/4$. Thus, from the expansion of H and **A1**, we have

$$|N_F(A)| \geq |N_H(A)| - |N_H(A) \cap X| - |N_H(A) \cap Z| \geq \left(1 - \frac{1}{4} - \frac{1}{4}\right)|A|\varepsilon(|A|) \geq \frac{1}{4}|A|\varepsilon(|A|).$$

Suppose then that $r \geq 1$ and that (4) holds with r replaced by r' for each $0 \leq r' < r$. Then, similarly to (5),

$$|B_F^r(A)| \geq |A|\left(1 + \frac{\varepsilon(|B_F^r(A)|)}{4}\right)^r. \tag{6}$$

Now, let α be defined by $|B_F^r(A)| = \alpha\varepsilon_2 d/15$. As $|B_F^r(A)| \geq |A| \geq \varepsilon_2 d/2$, $\alpha \geq 15/2$, and thus $\varepsilon(|B_F^r(A)|) = \varepsilon_1/\log^2\alpha \leq 1/2$. Therefore, from (6), we have

$$\alpha \geq \frac{|B_F^r(A)|}{|A|} \geq \left(1 + \frac{\varepsilon_1}{4\log^2\alpha}\right)^r \geq \exp\left(\frac{\varepsilon_1 r}{8\log^2\alpha}\right),$$

which gives $r \leq 8\log^3\alpha/\varepsilon_1$. As $\alpha \geq 15/2$, we have $\log^5\alpha/\alpha \leq 100$, so that

$$r + 1 \leq 2r \leq \frac{1600\alpha}{\varepsilon_1\log^2\alpha} = \frac{1600\alpha \cdot \varepsilon(|B_F^r(A)|)}{\varepsilon_1^2} = \frac{1600 \cdot 15}{\varepsilon_2 d \cdot \varepsilon_1^2} \cdot |B_F^r(A)|\varepsilon(|B_F^r(A)|). \tag{7}$$

As $d \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large, by **A3**, we have

$$|N_H(B_{H-Z}^r(A)) \cap Z| \leq k(r+1) \overset{(7)}{\leq} k \cdot \frac{1600 \cdot 15}{\varepsilon_2 d \cdot \varepsilon_1^2} \cdot |B_F^r(A)|\varepsilon(|B_F^r(A)|) \leq \frac{1}{4}|B_F^r(A)|\varepsilon(|B_F^r(A)|). \tag{8}$$

As $x \cdot \varepsilon(x)$ increases with $x \geq \varepsilon_2 d/2$, by **A1** we have $|X| \leq |A|\varepsilon(|A|)/4 \leq |B_F^r(A)|\varepsilon(|B_F^r(A)|)/4$. Therefore, using the expansion of H, we have

$$\begin{aligned}
|N_F(B_F^r(A))| &\geq |N_H(B_F^r(A))| - |N_H(B_F^r(A)) \cap X| - |N_H(B_{H-Z}^r(A)) \cap Z| \\
&\geq |N_H(B_F^r(A))| - |X| - |N_H(B_{H-Z}^r(A)) \cap Z| \overset{(8)}{\geq} \frac{1}{4}|B_F^r(A)| \cdot \varepsilon(|B_F^r(A)|),
\end{aligned}$$

and hence (4) holds for r. ∎

We now prove (ii). Suppose, for contradiction, that $|N_{H-X-Y-Z}^m(A)| \leq n/2$. Let $F' = H - X - Y - Z$, and for each $\ell_0 \leq r \leq m-1$, let $A_r = B_{F'}^r(A)$, so that, by (i), we have $|A_r| \geq m^{400k}$. Now, as $n \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large,

$$\varepsilon(|A_r|) \geq \varepsilon(n) \geq \varepsilon_1/\log^2 n \geq 1/m, \tag{9}$$

and thus $|A_r|\varepsilon(|A_r|)/4 \geq m^{400k-1}/4 \geq km^{300k}$. Therefore, by **A2**, $|A_r|\varepsilon(|A_r|)/4 \geq |Y|$ and, by **A3**,

$$|N_H(A_r) \cap Z| \leq |N_H(B_{H-Z}^r(A)) \cap Z| \leq k(r+1) \leq km \leq |A_r|\varepsilon(|A_r|)/4.$$

As $|A_r| \geq |A|$ and $x \cdot \varepsilon(x)$ increases with $x \geq \varepsilon_2 d/2$, we have, by **A1**, that $|A_r|\varepsilon(|A_r|)/4 \geq |X|$. In total, then,

$$|X| + |Y| + |N_H(A_r) \cap Z| \leq \frac{3}{4}|A_r|\varepsilon(|A_r|). \tag{10}$$

Thus, using the expansion of H, we have

$$|N_{F'}(A_r)| \geq |N_H(A_r)| - |X \cup Y| - |N_H(A_r) \cap Z| \overset{(10)}{\geq} \frac{1}{4}|A_r| \cdot \varepsilon(|A_r|) \overset{(9)}{\geq} \frac{\varepsilon_1|A_r|}{4\log^2 n}.$$

As this holds for all r with $\ell_0 \leq r \leq m-1$, we have

$$\begin{aligned}
|B_{F'}^m(A)| &\geq \left(1 + \frac{\varepsilon_1}{4\log^2 n}\right)^{m-\ell_0}|A_{\ell_0}| \geq \left(1 + \frac{\varepsilon_1}{4\log^2 n}\right)^{m/2} \\
&\geq \exp\left(\frac{\varepsilon_1 m}{16\log^2 n}\right) = \exp(\log n) > n/2,
\end{aligned}$$

a contradiction. □

3.2 Connecting sets with paths

Lemma 3.2 allows us to find paths between sets A and B in an expander, as follows.

Lemma 3.4. *For each $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$ and $x \geq 1$. Let G be an n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d - 1$.*

Let $A, B \subseteq V(G)$ with $|A|, |B| \geq x$, and let $W \subseteq V(G) \setminus (A \cup B)$ satisfy $|W| \log^3 n \leq 10x$. Then, there is a path from A to B in $G - W$ with length at most $\frac{40}{\varepsilon_1} \log^3 n$.

Proof. If $x \geq \varepsilon_2 d/2$, then, as $n \geq d \geq d_0(\varepsilon_1, \varepsilon_2)$ is large, we have

$$x \cdot \frac{\varepsilon(x)}{4} = \frac{\varepsilon_1 x}{4 \log^2(15x/\varepsilon_2 d)} \geq \frac{\varepsilon_1 x}{4 \log^2(15n)} \geq \frac{\varepsilon_1 x}{8 \log^2 n} \geq \frac{10x}{\log^3 n} \geq |W|. \tag{11}$$

Therefore, letting $m = \frac{16}{\varepsilon_1} \log^3 n$, by Lemma 3.2 applied with $X = W$ and $Y = Z = \emptyset$, we have $|B_{G-W}^m(A)|, |B_{G-W}^m(B)| > n/2$. Thus, there is a path from A to B in $G - W$ with length at most $2m \leq \frac{40}{\varepsilon_1} \log^3 n$.

Suppose then that $x < \varepsilon_2 d/2 \leq d/2$. Let $x' = \min\{|A \cup N_{G-W}(A)|, |B \cup N_{G-W}(B)|\}$. As $x < d/2$ and $A, B \neq \varnothing$, we have $x' \geq \delta(G) - |W| \geq d - 1 - 10d/\log^3 n \geq d/2 \geq \varepsilon_2 d/2$, as $n \geq d \geq d_0(\varepsilon_1, \varepsilon_2)$ is large. Therefore, as above, there is a path from $A \cup N_{G-W}(A)$ to $B \cup N_{G-W}(B)$ in $G - W$ with length at most $2m$, and hence a path with length at most $2m + 2 \leq \frac{40}{\varepsilon_1} \log^3 n$ between A and B in $G - W$, as required. $\qquad\square$

3.3 Expansion of sets of lower degree vertices

The following lemma is the key new technicality that allows our constructions, as discussed in Section 2.6. We then develop it for convenience of use to get Lemma 3.7.

Lemma 3.5. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

Let $U \subseteq V(G)$ satisfy $|U| \leq \exp((\log \log n)^2)$, and let $K = G - U$. Let I be any set and $V_i \subseteq V(K)$, $i \in I$, be pairwise disjoint sets such that, for each $i \in I$,

B1 $\varepsilon_2 d \leq |V_i| \leq \exp((\log \log n)^2)$,

B2 $|N_K(V_i)| \leq \frac{5|V_i|}{\log^{10} |V_i|}$, *and*

B3 $d_G(v, U) \leq d/2$ *for each $v \in V_i$.*

Then, $|\cup_{i \in I} V_i| < n^{1/8}$.

Proof. Suppose to the contrary that $|\cup_{i \in I} V_i| \geq n^{1/8}$ and let $D = \exp((\log \log n)^2)$. Let $I_1 = \{i \in I : |V_i| \geq (\log n)^{1/10}\}$ and $I_2 = I \setminus I_1$.

First, suppose that $|\cup_{i \in I_1} V_i| \geq n^{1/8}/2$. Then, as $|V_i| \leq D$ for each $i \in I$, $|I_1| \geq n^{1/9} \geq D^2$. Let $I_0 \subseteq I_1$ be a subset of I_1 of size D^2, and let $W = \cup_{i \in I_0} V_i$. Now, for each $i \in I_0$, $\log |V_i| \geq (\log \log n)/10$, while $|W| \leq |I_0| D \leq D^3$, so that $\log |W| \leq 3(\log \log n)^2$. Thus, we have

$$|N_G(W)| \leq |U| + |N_K(\cup_{i \in I_0} V_i)| \overset{\textbf{B2}}{\leq} D + \sum_{i \in I_0} |N_K(V_i)| \leq D + \sum_{i \in I_0} \frac{5|V_i|}{\log^{10} |V_i|}$$

$$\leq D + \sum_{i \in I_0} \frac{5(10)^{10}|V_i|}{(\log \log n)^{10}} = D + \frac{5(10)^{10}|W|}{(\log \log n)^{10}} < \frac{\varepsilon_1 |W|}{\log^3 |W|} \leq \varepsilon(|W|)|W|,$$

14

as $|W| \geq D^2$, $\log |W| \leq 3(\log \log n)^2$ and $n \geq d_0(\varepsilon_1, \varepsilon_2)$ is large. As $|W| \leq D^3$, this contradicts the fact that G is an $(\varepsilon_1, \varepsilon_2 d)$-expander. Thus, we have $|\cup_{i \in I_1} V_i| \leq n^{1/8}/2$.

Therefore, we have $|\cup_{i \in I_2} V_i| \geq n^{1/8}/2$. Thus, $I_2 \neq \varnothing$, and hence by definition, taking any $i \in I_2$, we have $\varepsilon_2 d \leq |V_i| \leq (\log n)^{1/10}$ so that $d \leq (\log n)^{1/10}/\varepsilon_2$. Furthermore, by the pigeonhole principle there must be some $r \in \mathbb{N}$ with $\varepsilon_2 d \leq r \leq (\log n)^{1/10}$ for which there are at least $|I_2|/(\log n)^{1/10} \geq |\cup_{i \in I_2} V_i|/((\log n)^{1/10})^2 \geq n^{1/9}$ indices $i \in I_2$ with $|V_i| = r$. Taking such an r, let $I_3 = \{i \in I_2 : |V_i| = r\}$, so that $|I_3| \geq n^{1/9}$.

Now, for each $i \in I_3$, as $d_G(v, U) \leq d \leq r/\varepsilon_2$ for each $v \in V_i$, we have $|N_G(V_i) \cap U| \leq r^2/\varepsilon_2 \leq (\log n)^{1/4}$, as $n \geq d_0(\varepsilon_1, \varepsilon_2)$ is large. As $|U| \leq D$, the number of sets of size at most $(\log n)^{1/4}$ in U is at most

$$\sum_{i=0}^{(\log n)^{1/4}} \binom{D}{i} \leq (\log n)^{1/4} D^{(\log n)^{1/4}} \leq \exp((\log n)^{1/3}).$$

Therefore, there must be at least $n^{1/9}/\exp((\log n)^{1/3}) \geq n^{1/10}$ indices $i \in I_3$ for which $N_G(V_i) \cap U$ is the same set, Z say. Taking any $i \in I_3$, note that, by **B3**, $|Z| = |N_G(V_i) \cap U| \leq dr \leq r^2/\varepsilon_2$. Let I_4 be a set of $r^2 \leq (\log n)^{1/5}$ indices $i \in I_3$ for which $N_G(V_i) \cap U = Z$.

Let $Y = \cup_{i \in I_4} V_i$, so that $N_G(Y) \cap U = Z$ and $|Y| = r|I_4| = r^3$. Then, as $d \geq d_0(\varepsilon_1, \varepsilon_2)$ is large and $r \geq \varepsilon_2 d$, we have

$$|N_G(Y)| \leq |Z| + \sum_{i \in I_4} |N_K(V_i)| \stackrel{\text{B2}}{\leq} \frac{r^2}{\varepsilon_2} + \frac{5r^3}{\log^{10} r} \leq \frac{\varepsilon_1 r^3}{\log^3(r^3)} = \frac{\varepsilon_1 |Y|}{\log^3 |Y|} \leq \varepsilon(|Y|)|Y|,$$

contradicting that G is an $(\varepsilon_1, \varepsilon_2 d)$-expander. $\qquad\square$

Lemma 3.5 is used three times, each in a similar situation, so for its application, we prove Lemma 3.7 below. In this lemma, with $r = n^{1/8}$, we have sets A_i, $i \in [r]$, and wish to find some set A_j which expands while avoiding some set $B_j \cup C_j$ which depends on j, as well as avoiding some large common set U. To find such an A_j, we assume for contradiction that no such set A_j exists, before recording as V_i the first ball around A_i which does not expand nicely. Applying Lemma 3.5 to the sets V_i, $i \in [r]$, will then reach a contradiction.

To prove the lemma, we will also use the following very simple proposition.

Proposition 3.6. *For each $i, \ell \geq 1$, we have $\ell^{2^{-i}} - (\ell - 1)^{2^{-i}} \leq \ell^{-1+2^{-i}}$.*

Proof. Not that this is true for any ℓ with equality if $i = 0$. Assume then for induction that $i > 0$ and it is true for ℓ with i replaced with $i - 1$. We have

$$\ell^{-1+2^{-(i-1)}} \geq \ell^{2^{-(i-1)}} - (\ell-1)^{2^{-(i-1)}} = (\ell^{2^{-i}} + (\ell-1)^{2^{-i}})(\ell^{2^{-i}} - (\ell-1)^{2^{-i}}) \geq \ell^{2^{-i}}(\ell^{2^{-i}} - (\ell-1)^{2^{-i}}),$$

and therefore $\ell^{2^{-i}} - (\ell - 1)^{2^{-i}} \leq \ell^{-1+2^{-(i-1)}-2^{-i}} = \ell^{-1+2^{-i}}$, as required. $\qquad\square$

Lemma 3.7. *For each $0 < \varepsilon_1 < 1$, $0 < \varepsilon_2 < 1/5$ and $k \in \mathbb{N}$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$. Let $U \subseteq V(G)$ satisfy $|U| \leq \exp((\log \log n)^2)$. Let $r = n^{1/8}$ and $\ell_0 = (\log \log n)^{20}$. Suppose (A_i, B_i, C_i), $i \in [r]$, are such that the following hold for each $i \in [r]$.*

C1 $|A_i| \geq d_0$.

C2 $B_i \cup C_i$ and A_i are disjoint sets in $V(G) \setminus U$, with $|B_i| \le |A_i|/\log^{10}|A_i|$.

C3 A_i has 4-limited contact with C_i in $G - U - B_i$.

C4 Each vertex in $B^{\ell_0}_{G-U-B_i-C_i}(A_i)$ has at most $d/2$ neighbours in U.

C5 For each $j \in [r] \setminus \{i\}$, A_i and A_j are at least a distance $2\ell_0$ apart in $G - U - B_i - C_i - B_j - C_j$.

   Then, for some $i \in [r]$, $|B^{\ell_0}_{G-U-B_i-C_i}(A_i)| \ge \log^k n$.

Proof. Suppose, for contradiction, that $|B^{\ell_0}_{G-U-B_i-C_i}(A_i)| < \log^k n$ for each $i \in [r]$. Let $\alpha = 1/16$, and note that $\exp(\ell_0^\alpha) = \exp((\log\log n)^{1.25}) > \log^k n$, as $n \ge d_0(\varepsilon_1, \varepsilon_2, k)$ is large, and hence $|B^{\ell_0}_{G-U-B_i-C_i}(A_i)| < \exp(\ell_0^\alpha)$ for each $i \in [r]$. Therefore, for each $i \in [r]$, we can let ℓ_i be the smallest $\ell \in [\ell_0]$ such that

$$|B^{\ell}_{G-U-B_i-C_i}(A_i)| \le \exp(\ell^\alpha). \tag{12}$$

For each $i \in [r]$, let $V_i = B^{\ell_i-1}_{G-U-B_i-C_i}(A_i)$. By the definition of ℓ_i, we have that

$$|V_i| \ge \exp((\ell_i - 1)^\alpha) \quad \text{and} \quad |B_{G-U-B_i-C_i}(V_i)| \le \exp(\ell_i^\alpha). \tag{13}$$

Claim 3.8. *For each $i \in [r]$, we have $|N_{G-U}(V_i)| \le \frac{5|V_i|}{\log^{10}|V_i|}$.*

Proof of claim. Fix $i \in [r]$. By Proposition 3.6, we have $\ell_i^\alpha - (\ell_i - 1)^\alpha \le \ell_i^{-1+\alpha} \le (\ell_i^\alpha)^{-10}$. Note that, as $d \ge d_0(\varepsilon_1, \varepsilon_2, k)$ is large, we have that $\ell_i \ge \log d_0$ is large by **C1** and (12). As $\exp(1/x) - 1 \le 2/x$ for large $x > 0$, we thus have

$$\exp(\ell_i^\alpha - (\ell_i - 1)^\alpha) - 1 \le \exp((\ell_i^\alpha)^{-10}) - 1 \le 2/(\ell_i^\alpha)^{10}. \tag{14}$$

Then,

$$|N_{G-U-B_i-C_i}(V_i)| \le |B_{G-U-B_i-C_i}(V_i)| - |V_i| = \left(\frac{|B_{G-U-B_i-C_i}(V_i)|}{|V_i|} - 1\right)|V_i|$$
$$\overset{(13)}{\le} \left(\frac{\exp(\ell_i^\alpha)}{\exp((\ell_i-1)^\alpha)} - 1\right)|V_i| \overset{(14)}{\le} \frac{2|V_i|}{(\ell_i^\alpha)^{10}} \overset{(13)}{\le} \frac{2|V_i|}{\log^{10}|V_i|}. \tag{15}$$

   Now, due to **C3**, we have

$$|N_{G-U-B_i}(V_i) \cap C_i| \le 4\ell_i \overset{(13)}{\le} 4(\log^{16}|V_i| + 1) \le \frac{|V_i|}{\log^{10}|V_i|}, \tag{16}$$

as, by **C1**, $|V_i| \ge |A_i| \ge d_0(\varepsilon_1, \varepsilon_2, k)$ is large. Therefore, as

$$|N_{G-U}(V_i)| \le |N_{G-U-B_i-C_i}(V_i)| + |B_i| + |N_{G-U-B_i}(V_i) \cap C_i|,$$

the claim follows from (15), (16), $|A_i| \le |V_i|$ and **C2**. $\blacksquare$

   We now check the conditions to appy Lemma 3.5 to the sets V_i, $i \in [r]$. By **C5**, the sets V_i, $i \in [r]$, are pairwise disjoint. Note that $|V_i| \ge |B_{G-U-B_i-C_i}(A_i)| \ge \varepsilon_2 d$. Indeed, if $|A_i| \ge \varepsilon_2 d$, this holds clearly; if $|A_i| < \varepsilon_2 d$, then by **C2**, **C3** and **C4**, $|B_{G-U-B_i-C_i}(A_i)| \ge \delta(G) - |B_i| - 4 - d/2 \ge \varepsilon_2 d$. Thus, for each $i \in [r]$, $\varepsilon_2 d \le |V_i| < \log^k n \le \exp((\log\log n)^2)$. By Claim 3.8, if $K = G - U$, then $|N_K(V_i)| \le 5|V_i|/\log^{10}|V_i|$ for each $i \in [r]$. By **C4**, for each $i \in [r]$ and $v \in V_i$, we have $d_G(v, U) \le d$. Therefore, by Lemma 3.5, we have $r \le |\cup_{i \in [r]} V_i| < n^{1/8}$, a contradiction. $\square$

3.4 Disjoint vertex expansions

In order to connect structures together in an expander, we typically find the structures we want with an extra subgraph attached to the vertex, v say, we wish to make connections from. This extra subgraph, F say, should have enough vertices that either Lemma 3.2 or Lemma 3.4 can be used to connect $V(F)$ to another vertex set while avoiding some structures that we have already found. The graph F should also have a short path from v to any other vertex in F, so that a path to $V(F)$ can be extended to one to v without many additional vertices. This motivates the following definition.

Definition 3.9. 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 of F.

Before we find vertex expansions, we first prove the following simple proposition which finds a smaller expansion within any expansion.

Proposition 3.10. *Let* $D, m \in \mathbb{N}$ *and* $1 \leq D' \leq D$. *Then, any graph F which is a (D, m)-expansion of v contains a subgraph which is a (D', m)-expansion of v.*

Proof. We prove this for each $1 \leq D' \leq D$ by induction on D' for $D' = D, D-1, \ldots, 1$. Note that F demonstrates this is true for $D' = D$. Suppose then it is true for $D' \geq 2$, and let F' with $|F'| = D'$ be a (D', m)-expansion of v. Let $w \in V(F')$ maximise the graph distance from v to w in F'. As $D' \geq 2$, $v \neq w$. Noting that $F' - w$ is a $(D'-1, m)$-expansion of v completes the proof of the inductive step, and hence the proposition. $\square$

We now give our lemma which finds vertex expansions. Its proof is different according to whether there are many vertices in the graph of high degree (Case I) or not (Case II). We will construct structures using short cycles later, so for the application of this lemma we need to find vertex expansions while avoiding a short cycle as much as possible.

Lemma 3.11. *For each $k \in \mathbb{N}$ and any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following holds for each $n \geq d \geq d_0$.*

Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d-1$. Let $m = \frac{40}{\varepsilon_1} \log^3 n$. Let C be a shortest cycle in G, and let $x_1, \ldots, x_k$ be distinct vertices in G. For each $i, j \in [k]$, let $D_{i,j} \in [1, \log^{5k} n]$.

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 which contains no vertices other than x_i in $V(C) \cup \{x_1, \ldots, x_k\}$.*

- *The sets $V(F_{i,j}) \setminus \{x_i\}$, $i, j \in [k]$, are pairwise disjoint.*

Proof. Note first that, as $\delta(G) \geq d-1$, $|C| \leq 2 \log n / \log(d-1)$. Let $D = \log^{5k} n$. By Proposition 3.10, we can assume that $D_{i,j} = D$ for all $i, j \in [k]$. Let $r = k^2$ and let L be the set of vertices in G with degree at least $\Delta = D^2$. We will split into two cases depending on whether $|L \setminus V(C)| \geq 2r$ or not.

Case I: First suppose that there are $2r$ vertices $v_1, \ldots, v_{2r} \in L \setminus V(C)$. Assume, by relabelling if necessary, that $V := \{v_1, \ldots, v_r\}$ is disjoint from $X := \{x_1, \ldots, x_k\}$.

Let $\mathcal{P}$ be a maximal collection of paths in $G - V(C)$ from X to V such that

17

- each path in $\mathcal{P}$ has length at most $3m$ and internal vertices in $V(G) \setminus (V(C) \cup X \cup V)$,

- all the paths in $\mathcal{P}$ are disjoint outside of X, and

- there are at most k paths containing each vertex in X.

Subject to $|\mathcal{P}|$ being maximal, suppose that $\sum_{P \in \mathcal{P}} \ell(P)$ is minimised. Note that, as there are at most k paths containing each vertex in X, $|\mathcal{P}| \leq k^2$.

Now, suppose for contradiction that there is some $x \in X$ which is in fewer than k paths in $\mathcal{P}$, and so $|\mathcal{P}| < k^2$. Let $U = (V \cup X \cup V(\mathcal{P}) \cup V(C)) \setminus \{x\}$. For each path $P \in \mathcal{P}$ and $\ell \in \mathbb{N}$, at most $\ell + 1$ vertices in $N_G(B_{G-U}^{\ell-1}(x))$ can lie on P, otherwise we can find a shorter path than P from x to $V(P) \cap V$ in $G - (U \setminus V(P))$. Swapping P for this shorter path contradicts the minimality of $\sum_{P' \in \mathcal{P}} \ell(P')$. Therefore, for each $\ell \in \mathbb{N}$,

$$|N_G(B_{G-U}^{\ell-1}(x)) \cap V(\mathcal{P})| \leq \sum_{P \in \mathcal{P}} |N_G(B_{G-U}^{\ell-1}(x)) \cap V(P)| \leq (\ell+1)|\mathcal{P}| \leq (\ell+1)k^2. \qquad (17)$$

Furthermore, for any vertex v and any integer $\ell \in \mathbb{N}$, at most $2\ell + 1$ vertices in $B_G^\ell(v)$ can lie on C, as C is a shortest cycle in G. Therefore,

$$|N_G(B_{G-U}^{\ell-1}(x)) \cap U| \leq |V \cup X| + |N_G(B_{G-U}^{\ell-1}(x)) \cap V(\mathcal{P})| + |N_G(B_{G-V(C)+x}^{\ell-1}(x)) \cap V(C)|$$

$$\overset{(17)}{\leq} 2k^2 + (\ell+1)k^2 + |B_G^\ell(x) \cap V(C)| \leq (\ell+3)k^2 + (2\ell+1) \leq 10\ell k^2.$$

Thus, $\{x\}$ has $10k^2$-limited contact with U in G, so that $B_{G-U}(x)$ has $20k^2$-limited contact with U in G. We also have $|B_{G-U}(x)| \geq \delta(G) - 10k^2 \geq d/2 \geq \varepsilon_2 d/2$. Therefore, applying Lemma 3.2 with $(A, X, Y, Z, k)_{3.2} = (B_{G-U}(x), \varnothing, \varnothing, U, 20k^2)$, we get that $|B_{G-U}^{m+1}(x)| = |B_{G-U}^m(B_{G-U}(x))| > n/2$.

Note that, by the choice of $\mathcal{P}$, each $P \in \mathcal{P}$ has a distinct vertex in V. As $|\mathcal{P}| < k^2$, we can choose a vertex $v \in V \setminus V(\mathcal{P})$. Note that $|U| \leq 2k^2 + k^2 \cdot (3m+1) + 2\log n / \log(d-1) \leq \log^4 n$. As $d_G(v) \geq \Delta = \log^{10k} n \geq |U| \log^3 n / 10$ and $|B_{G-U}^{m+1}(x)| > n/2 \geq |U| \log^3 n / 10$, using Lemma 3.4, we can connect $B_{G-U}^{m+1}(x)$ and $N_G(v)$ with a path of length at most m in $G - U$, which extends in $B_{G-U}^{m+1}(x) \cup \{v\}$ to an x,v-path in $G - U$ with length at most $3m$, contradicting the maximality of $\mathcal{P}$.

Therefore, each vertex in X is in exactly k paths in $\mathcal{P}$. Label the paths in $\mathcal{P}$ as $P_{i,j}$, $i, j \in [k]$, so that x_i is an endvertex of $P_{i,j}$, and let $v_{i,j}$ be the endvertex of $P_{i,j}$ in V. Greedily, using that $|N_G(v_{i,j}) \setminus (V \cup X \cup V(C) \cup V(\mathcal{P}))| \geq \Delta - \log^4 n \geq k^2 D$ for each $i, j \in [k]$, pick disjoint sets $A_{i,j} \subseteq N_G(v_{i,j}) \setminus (V \cup X \cup V(C) \cup V(\mathcal{P}))$, $i, j \in [k]$, with size $D - |P_{i,j}|$. Then $F_{i,j} = G[A_{i,j}] \cup P_{i,j}$, $i, j \in [k]$, are easily seen to be the $(D, 5m)$-expansions we require.

Case II: Suppose then that $|L \setminus V(C)| < 2r$. Relabelling if necessary, let $k' \leq k$ be such that $\{x_1, \ldots, x_k\} \setminus L = \{x_1, \ldots, x_{k'}\}$. Let $G' = G - L$, $X = \{x_1, \ldots, x_{k'}\}$, $r' = k'k$ and $\ell_0 = 2(\log \log n)^5$.

Let $s \leq r'$ be the largest integer for which there are vertices $w_1, \ldots, w_s \in V(G')$ such that

- the sets $B_{G'}^{5\ell_0}(w_i)$, $i \in [s]$, X and $V(C) \setminus L$ are all pairwise disjoint.

Suppose $s < r'$. Then, we must have $V(G') = B_{G'}^{10\ell_0}((\{w_1, \ldots, w_s\} \cup X \cup V(C)) \setminus L)$. However, as $\Delta(G') \leq \Delta = D^2$ and $|C| \leq 2\log n / \log(d-1) \leq 2\log n$,

$$|G'| = |B_{G'}^{10\ell_0}((\{w_1, \ldots, w_s\} \cup X \cup V(C)) \setminus L)| \leq 2 \cdot (r' + k' + 2\log n) \cdot \Delta^{10\ell_0} \leq \exp((\log \log n)^7) < n/2,$$

contradicting $|G'| \geq n - |L \setminus V(C)| - |C| \geq n - 2r - 2\log n \geq n/2$. Therefore, $s = r'$.

Now, fixing an arbitrary $i \in [r']$, similarly to before, for each $\ell \in \mathbb{N}$ at most $2\ell + 1$ vertices in $B_G^\ell(w_i)$ can lie on C, otherwise there is a shorter cycle in G than C. Thus,

$$|B_{G'-V(C)}(w_i)| \geq \delta(G) - |L \setminus V(C)| - 3 \geq \delta(G) - 2r - 3 \geq d/2, \qquad (18)$$

and, for each $\ell \in \mathbb{N}$,

$$|N_G(B_{G-V(C)}^{\ell-1}(B_{G'-V(C)}(w_i))) \cap V(C)| \leq |B_G^{\ell+1}(w_i) \cap V(C)| \leq 2\ell + 3 \leq 5\ell.$$

Therefore, $B_{G'-V(C)}(w_i)$ has 5-limited contact with $V(C)$ in G. Let $z = |B_{G'-V(C)}(w_i)|$, and note that, as $|L \setminus V(C)| \leq 2r = 2k^2$ and, by (18), $z \geq d/2 \geq d_0(\varepsilon_1, \varepsilon_2, k)/2$ is large, we have that $|L \setminus V(C)| \leq \varepsilon(z)z/4$. Thus, by Lemma 3.2 with $(A, X, Y, Z, k)_{3.2} = (B_{G'-V(C)}(w_i), L \setminus V(C), \varnothing, V(C), k+5)$ we get that $|B_{G'-V(C)}^{\ell_0+1}(w_i)| \geq D^2 = \Delta$. Hence, by Proposition 3.10 we can pick a subgraph $F_i \subseteq G'$ induced on a subset of $B_{G'-V(C)}^{\ell_0}(w_i)$ which is a $(\Delta, 2\ell_0)$-expansion of w_i.

Now, let $\mathcal{P}$ be a maximal collection of paths in G' from X to $V := \cup_{i \in [r']} V(F_i)$ such that

- each path in $\mathcal{P}$ has length at most $3m$ and internal vertices in $V(G') \setminus (V(C) \cup X \cup V)$,

- all the paths in $\mathcal{P}$ are disjoint outside of X,

- at most 1 path in $\mathcal{P}$ has a vertex in $V(F_i)$, for each $i \in [r']$, and

- there are at most k paths containing each vertex in X.

Subject to $|\mathcal{P}|$ being maximal, suppose that $\sum_{P \in \mathcal{P}} \ell(P)$ is minimised.

Suppose again there is some $x \in X$ in fewer than k paths in $\mathcal{P}$ and let $U = (L \cup X \cup V(\mathcal{P}) \cup V(C)) \setminus \{x\}$. As in Case I, by the minimality of $\sum_{P' \in \mathcal{P}} \ell(P')$, for each path $P \in \mathcal{P}$ and $\ell \in \mathbb{N}$, at most $\ell + 1$ vertices in $N_G(B_{G-U}^{\ell-1}(x))$ can lie on P. Therefore, for each $\ell \in \mathbb{N}$,

$$|N_G(B_{G-U}^{\ell-1}(x)) \cap V(\mathcal{P})| \leq \sum_{P \in \mathcal{P}} |N_G(B_{G-U}^{\ell-1}(x)) \cap V(P)| \leq (\ell+1)|\mathcal{P}| \leq (\ell+1)k^2.$$

Again, for any integer $\ell \in \mathbb{N}$, at most $2\ell + 1$ vertices in $B_G^\ell(x)$ can lie on C, as C is a shortest cycle in G. Thus,

$$|N_G(B_{G-U}^{\ell-1}(x)) \cap U| \leq |X \cup (L \setminus V(C))| + |N_G(B_{G-U}^{\ell-1}(x)) \cap V(\mathcal{P})| + |B_G^\ell(x) \cap V(C)|$$
$$\leq k + 2r + (\ell+1)k^2 + (2\ell+1) \leq 10\ell k^2.$$

Therefore, $\{x\}$ has $10k^2$-limited contact with U in G, and so $B_{G-U}(x)$ has $20k^2$-limited contact with U in G. Furthermore, by the choice of the w_i, $B_{G-U}^{\ell_0}(x) \cap V = \varnothing$, thus

$$|B_{G-U-V}(x)| = |B_{G-U}(x)| \geq d - 10k^2 \geq d/2.$$

We also have that $|V| \leq k^2\Delta = k^2 \log^{10k} n$. Therefore, by Lemma 3.2 with $(A, X, Y, Z, k)_{3.2} = (B_{G-U}(x), \varnothing, V, U, 20k^2)$, $|B_{G-U-V}^{m+1}(x)| > n/2$.

By the choice of $\mathcal{P}$, each $P \in \mathcal{P}$ has exactly one vertex in $V(F_i)$ for some $i \in [r']$. As $|\mathcal{P}| < k'k = r'$, there is some $j \in [r']$ such that $V(\mathcal{P})$ has no vertex in $V(F_j)$. Now, the vertices w_i, $i \in [r']$, are pairwise at least $10\ell_0$-far in G', and F_i is a (Δ, ℓ_0)-expansion around w_i. Therefore,

19

the subgraphs F_i, $i \in [r']$, are pairwise at least $8\ell_0$-far from each other in G', so that F_j is at least $8\ell_0$-far from $V \setminus V(F_j)$.

Now, as $|U| \leq |X| + |L \setminus V(C)| + |C| + |V(\mathcal{P})| \leq k + 2r + 2\log n + (3m+1)r \leq \log^4 n$, we have $|F_j| = \Delta \geq m|U|$. As $L \neq V(G)$, we have $\Delta > \delta(G) \geq \varepsilon_2 d$, so that $|F_j| \geq \varepsilon_2 d$. Therefore, by Lemma 3.2 with $(A, X, Y, Z, k)_{3.2} = (V(F_j), U, V \setminus V(F_j), \varnothing, k)$, we have

$$|B^m_{G-U-(V \setminus V(F_j))}(V(F_j))| > n/2.$$

Thus, as $|B^{m+1}_{G-U-V}(x)| > n/2$, there is a path from $\{x\}$ to $V(F_j)$ in G' with length at most $3m$ which is internally disjoint from $V(C) \cup X \cup V$. This contradicts the maximality of $\mathcal{P}$.

Therefore, each vertex in X is in exactly k paths in $\mathcal{P}$. Label the paths in $\mathcal{P}$ with $P_{i,j}$, $i \in [k']$ and $j \in [k]$, and graphs $F_{i'}$, $i' \in [r']$, as $F'_{i,j}$, $i \in [k']$ and $j \in [k]$, so that, for each $i \in [k']$ and $j \in [k]$, $P_{i,j}$ is a path from x_i to $F'_{i,j}$. Note that $P_{i,j} \cup F'_{i,j}$ is a $(|P_{i,j} \cup F'_{i,j}|, \ell(P_{i,j}) + 2\ell_0)$-expansion of x_i, for each $i \in [k']$ and $j \in [k]$. For each $i \in [k']$ and $j \in [k]$, apply Proposition 3.10 to obtain a $(D, 5m)$-expansion $F_{i,j} \subseteq P_{i,j} \cup F'_{i,j}$ around x_i.

Lastly, for each $k'+1 \leq i \leq k$ and $j \in [k]$, as $x_i \in L$ we can greedily pick pairwise disjoint $(D, 1)$-expansions $F_{i,j}$ induced on a subset of $N_G(x_i) \setminus (V(C) \cup (\cup_{i' \in [k'], j' \in [k]} V(F_{i',j'})))$. □

3.5 Enlarging vertex expansions

In this section, we take up to 4 disjoint vertex expansions, and expand them disjointly to get larger vertex expansions around the same vertices (see Lemma 3.13). This enlargement allows us to later connect vertex expansions with very long paths, as our path lengths need to be smaller than the expansions (see Section 3.6).

We first show for Lemma 3.12 that we can always find a linear size vertex set with polylogarithmic diameter in G while avoiding an arbitrary set of up to $\Theta(n/\log^2 n)$ vertices. This is used for Lemma 3.13 and Lemma 3.14, as well as later in Section 4.2

Lemma 3.12. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

For any set $W \subseteq V(G)$ with $|W| \leq \varepsilon_1 n/100 \log^2 n$, there is a set $B \subseteq G - W$ with size at least $n/25$ and diameter at most $\frac{100}{\varepsilon_1} \log^3 n$.

Proof. Let $\ell_0 = \frac{50}{\varepsilon_1} \log^2 n$. Suppose that G and $W \subseteq V(G)$ satisfy the conditions in the lemma and let $G' = G - W$. Take the largest integer $r \leq \log n$ such that there is a set of at most $1 + n/(10 \cdot 4^r)$ vertices $V \subseteq V(G')$ with $|B^{\ell_0 r}_{G'}(V)| \geq n/25$. Note that such a set of vertices exists for $r = 0$, as $|G'| \geq n - \varepsilon_1 n/100 \log^2 n \geq n/25$ and $n \geq d_0(\varepsilon_1, \varepsilon_2)$ is large. Suppose, for contradiction, that $|V| > 1$, and hence, that $r \leq \log n - 1$.

Let $A = B^{\ell_0 r}_{G'}(V)$. As $|W| \leq \varepsilon_1 n/100 \log^2 n \leq \varepsilon_1 |A|/4 \log^2 n$, for each ℓ with $|B^\ell_{G'}(A)| < n/2$, we have, by the expansion property of G, and as $\varepsilon(|B^\ell_{G'}(A)|) \geq \varepsilon(n) \geq \varepsilon_1/\log^2 n$,

$$|N_{G'}(B^\ell_{G'}(A))| \geq |N_G(B^\ell_{G'}(A))| - |W| \geq \frac{\varepsilon_1}{\log^2 n} \cdot |B^\ell_{G'}(A)| - \frac{\varepsilon_1 |A|}{4 \log^2 n} \geq \frac{\varepsilon_1}{2 \log^2 n} \cdot |B^\ell_{G'}(A)|,$$

so that $|B^{\ell+1}_{G'}(A)| \geq (1 + \varepsilon_1/2 \log^2 n)|B^\ell_{G'}(A)|$. If $|B^{\ell_0}_{G'}(A)| < n/2$, then

$$|B^{\ell_0}_{G'}(A)| \geq \left(1 + \frac{\varepsilon_1}{2 \log^2 n}\right)^{\ell_0} |A| \geq \frac{\varepsilon_1 \ell_0}{2 \log^2 n} \cdot \frac{n}{25} \geq n/2,$$

a contradiction. Therefore, we have

$$|B_{G'}^{\ell_0(r+1)}(V)| = |B_{G'}^{\ell_0}(A)| \geq n/2.$$

Consequently, by averaging, there exists a set of at most $\lceil |V|/12 \rceil$ vertices $V' \subseteq V$ such that $|B_{G'}^{\ell_0(r+1)}(V')| \geq |B_{G'}^{\ell_0(r+1)}(V)|/12 \geq n/25$. Noting that, as $|V| \geq 2$,

$$\lceil |V|/12 \rceil \leq 1 + (|V| - 1)/12 \leq 1 + n/(10 \cdot 4^{r+1}),$$

this contradicts the maximality of r.

Therefore, we have $|V| = 1$. That is, there is some vertex $v \in V(G')$ with $|B_{G'}^{\ell_0 r}(v)| \geq n/25$. Letting $B = B_{G'}^{\ell_0 r}(v)$, we have that $|B| \geq n/25$ and B has diameter at most $2\ell_0 r \leq \frac{100}{\varepsilon_1} \log^3 n$, as required. $\qquad\square$

The large vertex sets with small diameter found in Lemma 3.12 are large vertex expansions around some vertex. To find large vertex expansions around particular vertices, which themselves sit in smaller vertex expansions, we take disjointly many large vertex expansions, then expand and connect the smaller vertex expansions to the larger ones, for the following lemma.

Lemma 3.13. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*
Let $\log^{10} n \leq D \leq n/\log^{10} n$ and $m = \frac{100}{\varepsilon_1} \log^3 n$. Let $A \subseteq V(G)$ satisfy $|A| \leq D/\log^3 n$. Let $F_1, \ldots, F_4 \subseteq G - A$ be vertex disjoint subgraphs and $v_1, \ldots, v_4$ be vertices such that, for each $i \in [4]$, F_i is a (D, m)-expansion of v_i. Then, $G - A$ contains vertex disjoint subgraphs $F'_1, \ldots, F'_4$ such that, for each $i \in [4]$, F'_i is an $(n/m^2, 3m)$-expansion of v_i.

Proof. Let $W = V(F_1) \cup \ldots \cup V(F_4) \cup A$ so that $|W| \leq 5D \leq 5n/\log^{10} n$. Applying Lemma 3.12 iteratively $32m$ times, we can find disjoint sets $B_1, \ldots, B_{32m}$ in $G - W$ such that each B_i has size n/m^2 and diameter at most m. Note that this is possible, as after we have found B_i, where $i \in [32m]$, we have $W \cup (\cup_{j \in [i]} B_j) \leq 5n/\log^{10} n + 32m \cdot n/m^2 \leq \varepsilon_1 n/100 \log^2 n$, as $n \geq d_0(\varepsilon_1, \varepsilon_2)$ is large.

Take a maximal set $I \subseteq [4]$ such that there are paths $P_{i,j}$, with $i \in I$ and $j \in [8m]$, and distinct $k_{i,j} \in [32m]$, satisfying the following.

D1 $P_{i,j}$ is a path from v_i to $B_{k_{i,j}}$ of length at most $2m$.

D2 The sets $V(P_{i,j}) \setminus V(F_i)$ are vertex disjoint across $i \in I$ and $j \in [8m]$.

D3 There is an ordering σ on I such that $V(P_{i,j})$ is disjoint from $\cup_{i' \in [4] \setminus I'} V(F_{i'})$, where $I' = \{i' \in I : \sigma(i') \leq \sigma(i)\}$.

Suppose, for contradiction, that $J := [4] \setminus I \neq \varnothing$. Let $W' = A \cup (\cup_{i \in I, j \in [8m]} V(P_{i,j}))$. Take a maximal set $K \subseteq [32m] \setminus \{k_{i,j} : i \in I, j \in [8m]\}$ for which there are paths P_k, $k \in K$, with length at most m from $\cup_{i \in J} V(F_i)$ to B_k, which avoid W' and which are vertex disjoint. Suppose, to contradict the maximality of K, that there is some $j' \in [32m] \setminus \{k_{i,j} : i \in I, j \in [8m]\}$. Note that $|\cup_{i \in J} V(F_i)| \geq D$ and $|B_{j'}| = n/m^2 \geq D$. Noting that $|W' \cup (\cup_{k \in K} V(P_k))| \leq D/\log^3 n + 32m(2m+1) \leq 2D/\log^3 n$, by Lemma 3.4, there is a path between $\cup_{i \in J} V(F_i)$ and $B_{j'}$ with length at most m which avoids $W' \cup (\cup_{k \in K} V(P_k))$, contradicting the maximality of K.

Therefore, we have $K = [32m] \setminus \{k_{i,j} : i \in I, j \in [8m]\}$, and hence $|K| = 32m - 8m|I| = 8m|J|$. Consequently, for some $i' \in J$, there are at least $8m$ values of $k \in K$ for which P_k has a vertex in $V(F_{i'})$. Taking $k_{i',1}, \ldots, k_{i',8m}$ to be distinct such values of k, for each $j \in [8m]$, as $F_{i'}$ is a (D, m)-expansion of $v_{i'}$, we can find a path $P_{i',j} \subseteq P_{k_{i',j}} \cup F_{i'}$ from $v_{i'}$ to $B_{k_{i',j}}$ with length at most $2m$. The paths $P_{i,j}$, $i \in I \cup \{i'\}$ and $j \in [8m]$, satisfy **D1**–**D3**, contradicting the maximality of I.

Thus, we have $I = [4]$. By relabelling if necessary, assume from **D3** that $V(P_{i,j})$ is disjoint from $V(F_{i'})$ for each $i' > i$ and $j \in [8m]$. Now, for each $1 \leq i \leq 4$, greedily select $r_i \in [8m]$ in turn such that $V(P_{i,r_i}) \cup B_{k_{i,r_i}}$ has no vertices in $\cup_{i' < i} P_{i',r_{i'}}$. Note that this is possible as, for each $i \in [4]$, $\cup_{i' < i} V(P_{i',r_{i'}})$ contains at most $3(2m+1)$ vertices, none of which are in $V(F_i)$, and the sets $V(P_{i,r_i}) \cup B_{k_{i,j}}$, $j \in [8m]$ are disjoint outside of $V(F_i)$.

For each $i \in [4]$, let $F_i'' = P_{i,r_i} \cup G[B_{k_{i,r_i}}]$. Note that the subgraphs F_i'', $i \in [4]$, are vertex disjoint, and, as each set B_i, $i \in [32m]$, has diameter at most m, for each $i \in [4]$, F_i'' is a $(D_i, 3m)$-expansion of v_i for some $D_i \geq n/m^2$. For each $i \in [4]$, using Proposition 3.10, let $F_i' \subseteq F_i''$ be an $(n/m^2, 3m)$-expansion of v_i, completing the proof. $\square$

3.6 Long paths between vertex expansions

Our goal now is, given some desired path length, to connect two vertices with an initial path with close to this desired length. Section 4 then makes the fine adjustment to this path so that it has exactly the desired length. Given vertex expansions around the vertices to be connected, we find such an initial path, as follows.

Lemma 3.14. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

Let $\log^3 n \leq D \leq n/\log^4 n$ and $\frac{300}{\varepsilon_1} \log^3 n \leq m \leq 3 \log^4 n$. Suppose F_1, F_2 are vertex disjoint (D, m)-expansions of vertices $v_1, v_2 \in V(G)$ respectively. Suppose $W \subseteq V(G) \setminus (V(F_1) \cup V(F_2))$ satisfies $|W| \leq D/\log^3 n$. Then, for any $\ell \leq D/\log^3 n$, there is a v_1, v_2-path in $G - W$ with length between ℓ and $\ell + 5m$.

Proof. Let $(P_1, v_3, F_3, P_2, v_4, F_4)$ be such that $\ell(P_1) + \ell(P_2)$ is maximised subject to the following properties.

E1 For each $i \in [2]$, P_i is a v_i, v_{i+2}-path in $G - W$.

E2 $\ell(P_1) + \ell(P_2) \leq \ell + 2m$.

E3 For each $i \in \{3, 4\}$, F_i is a (D, m)-expansion of v_i in $G - W$ with $V(F_i) \cap V(P_{i-2}) = \{v_i\}$.

E4 $V(P_1 \cup F_3)$ and $V(P_2 \cup F_4)$ are vertex disjoint.

Note that $P_1 = G[\{v_3\}]$, $v_3 = v_1$, $F_3 = F_1$, $P_2 = G[\{v_4\}]$, $v_4 = v_2$, $F_4 = F_2$ satisfy **E1**–**E4**, and therefore such a sextuple $(P_1, v_3, F_3, P_2, v_4, F_4)$ exists.

We claim that $\ell(P_1) + \ell(P_2) \geq \ell$. Suppose for contradiction that $\ell(P_1) + \ell(P_2) < \ell$. Note that $|W \cup V(F_3 \cup F_4 \cup P_1 \cup P_2)| \leq 3D + \ell \leq n/\log^3 n$. By Lemma 3.12, then, $G - W - V(F_3 \cup F_4 \cup P_1 \cup P_2)$ contains a set, B say, with size at least D and diameter at most m. Note furthermore that $|W \cup V(P_1) \cup V(P_2)| \leq D/\log^3 n + \ell + 2 \leq 10D/\log^3 n$. By Lemma 3.4, there is a path, Q' say, from B to $V(F_3) \cup V(F_4)$ which avoids $(W \cup V(P_1) \cup V(P_2)) \setminus \{v_3, v_4\}$ and has length at most m. Say without loss of generality that Q' has endvertices $v_3'' \in V(F_3)$ and $v_3' \in B$. By **E3** and **E4**, we

22

can extend Q' in F_3 to a v_3, v_3'-path Q with length at most $2m$ which is vertex disjoint from P_2 and $P_1 - v_3$.

Using Proposition 3.10, let $F_3' \subseteq G[B]$ be a (D, m)-expansion around v_3'. Let $P_1' = P_1 \cup Q$ and note that, as $v_3' \in B$ and $\ell(Q) \leq 2m$, this is a v_1, v_3'-path with length at least $\ell(P_1) + 1$ and at most $\ell(P_1) + 2m$. Then, $P_1', v_3', F_3', P_2, v_4, F_4$ satisfy **E1**–**E4** with P_1', v_3', F_3' in place of P_1, v_3, F_3, and $\ell(P_1') + \ell(P_2) > \ell(P_1) + \ell(P_2)$, a contradiction. Therefore, $\ell(P_1) + \ell(P_2) \geq \ell$.

Now, as $|W \cup V(P_1) \cup V(P_2)| \leq 10D/\log^3 n$, by Lemma 3.4 there is a path, R say, from some $r_1 \in V(F_3)$ to some $r_2 \in V(F_4)$ avoiding $(W \cup V(P_1) \cup V(P_2)) \setminus \{v_3, v_4\}$ with length at most m. For each $i \in [2]$, let Q_i be a path from v_{i+2} to r_i in F_{i+2} with length at most m. Then, $P_1 \cup Q_1 \cup R \cup Q_2 \cup P_2$ is a v_1, v_2-path in $G - W$ with length at least $\ell(P_1) + \ell(P_2) \geq \ell$ and at most, by **E2**, $\ell + 2m + 3m \leq \ell + 5m$. $\qquad\square$

Combining Lemma 3.14 with our result on extending vertex expansions, we now convert Lemma 3.14 into the precise form we apply later. The following corollary finds not one but two paths, whose combined length is close to some desired length. We later apply it so that these two paths connect two vertices with a string of simple adjusters in the middle.

Corollary 3.15. *For any $0 < \varepsilon_1, \varepsilon_2 < 1$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

Let $\log^{10} n \leq D \leq n/\log^{10} n$, $\frac{100}{\varepsilon_1} \log^3 n \leq m \leq \log^4 n$ and $\ell \leq n/\log^{10} n$. Let $A \subseteq V(G)$ satisfy $|A| \leq D/\log^3 n$. Let $F_1, \ldots, F_4 \subseteq G - A$ be vertex disjoint subgraphs and $v_1, \ldots, v_4$ be vertices such that, for each $i \in [4]$, F_i is a (D, m)-expansion of v_i.

Then, $G - A$ contains vertex disjoint paths P and Q with $\ell \leq \ell(P) + \ell(Q) \leq \ell + 20m$ such that both P and Q connect $\{v_1, v_2\}$ to $\{v_3, v_4\}$.

Proof. By Lemma 3.13, $G - A$ contains disjoint subgraphs $F_1', \ldots, F_4'$ such that, for each $i \in [4]$, F_i' is an $(n/m^2, 3m)$-expansion of v_i. By Lemma 3.4, there is a path $P' \subseteq G - A$ from $V(F_1') \cup V(F_2')$ to $V(F_3') \cup V(F_4')$ with length at most m. Note that we can assume, without loss of generality, that P' goes from $V(F_1')$ to $V(F_3')$. Using that F_1' and F_3' are $(n/m^2, 3m)$-vertex expansions of v_1 and v_3, respectively, let P be a v_1, v_3-path with length at most $7m$ in $F_1' \cup P' \cup F_3'$.

Let $W = A \cup V(P)$, noting that $|W| \leq D/\log^3 n + 7m + 1 \leq n/m^2 \log^3 n$. Note that $0 \leq \ell - \ell(P) + 7m \leq 2n/\log^{10} n \leq n/m^2 \log^3 n$. Therefore, by Lemma 3.14 with $(F_1, F_2, D, m, W, \ell)_{3.14} = (F_2', F_4', n/m^2, 3m, W, \ell - \ell(P) + 7m)$, there is a path Q in $G - W$ from v_2 to v_4 with length between $\ell - \ell(P) + 7m$ and $\ell - \ell(P) + 20m$. As $\ell \leq \ell(P) + \ell(Q) \leq \ell + 20m$, the paths P and Q satisfy the property in the corollary. $\qquad\square$

3.7 Subdivisions in skewed bipartite graphs

As commented on before Theorem 2.7, we often work in graphs without a subdivision of a certain size clique with each edge divided once, that is, in a $\mathsf{TK}_\ell^{(2)}$-free graph for some ℓ. This is because we often construct structures in a graph G while avoiding a vertex set, say W, for which we need many edges in $G - W$. For the sizes of W and values of ℓ that we use, if the graph G is $\mathsf{TK}_\ell^{(2)}$-free, then the following simple proposition shows that there cannot be too many edges between W and $V(G) \setminus W$. In our implementation this will imply that $G - W$ contains many edges.

Proposition 3.16. *Let $d \in \mathbb{N}$ and let G be a graph containing disjoint vertex sets U and W such that $|U| \geq |W|^2$ and every vertex in U has at least d neighbours in W. Then, G contains a $\mathsf{TK}_d^{(2)}$.*

Proof. Take a maximal set $I \subseteq W^{(2)}$ for which there is a set of distinct vertices $v_{\{x,y\}}$, $\{x,y\} \in I$, in U such that $x, y \in N(v_{\{x,y\}})$ for each $\{x,y\} \in I$. Now, as $|U| \geq |W|^2 > |W^{(2)}|$, there is some $u \in U \setminus \{v_{\{x,y\}} : \{x,y\} \in I\}$. Let $A = N(u, W)$, so that $|A| \geq d$. In the choice of I, the vertex u is a good candidate for $v_{\{x,y\}}$ for each $\{x,y\} \in A^{(2)}$. Thus, by the maximality of I, we have $A^{(2)} \subseteq I$.

Taking the $\mathsf{TK}_d^{(2)}$ with vertex set $A \cup \{v_{\{x,y\}} : \{x,y\} \in A^{(2)}\}$ and edge set $\{xv_{\{x,y\}}, yv_{\{x,y\}} : \{x,y\} \in A^{(2)}\}$, then gives a $\mathsf{TK}_d^{(2)}$ in G, as required. $\square$

4 Proof of Theorem 2.7

In this section, we prove Theorem 2.7. As discussed in Section 2.4, the basic mechanism we use to adjust the length of a path is an adjuster, which we formally define as follows (see Figure 1(b) for an illustration).

Definition 4.1. 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 $\ell \in \mathbb{N}$.

  F1 A, $V(F_1)$ and $V(F_2)$ are pairwise disjoint.

  F2 For each $i \in [2]$, F_i is a (D, m)-expansion around v_i.

  F3 $|A| \leq 10mk$.

  F4 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 $\ell + 2i$.

We call the smallest such ℓ for which these properties hold the *length of the adjuster* and denote it $\ell(\mathcal{A})$. Note that it immediately follows that $\ell(\mathcal{A}) \leq |A| + 1 \leq 10mk + 1$. We call a $(D, m, 1)$-adjuster a *simple adjuster*. We refer to the subgraphs F_1 and F_2 of an adjuster $\mathcal{A} = (v_1, F_1, v_2, F_2, A)$ as the *ends* of the adjuster, and let $V(\mathcal{A}) = V(F_1) \cup V(F_2) \cup A$.

In this section, we start by finding one simple adjuster in an expander for Lemma 4.2 in Section 4.1. We then find such an adjuster despite the removal of any medium-sized vertex set from the expander, giving Lemma 4.3 in Section 4.2. In Section 4.3, we chain simple adjusters together for Lemma 4.7, before using this to join vertex expansions by paths with precise lengths for Lemma 4.8. Finally, we prove Theorem 1.7 in Section 4.4 and Theorem 2.7 in Section 4.5.

4.1 Finding one simple adjuster

Here, we find one simple adjuster, proving Lemma 4.2. The adjuster (v_1, F_1, v_2, F_2, A) is found for prespecified vertices v_1 and v_2, as required by one of the applications of Lemma 4.2.

Lemma 4.2. *For any $0 < \varepsilon_1 < 1$, $0 < \varepsilon_2 < 1/5$ and $k \in \mathbb{N}$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following is true for each $n \geq d \geq d_0$. Suppose that G is an n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d - 1$.*

Let C be a shortest cycle in G and let x_1, x_2 be distinct vertices in $V(G) \setminus V(C)$. Let $m = \frac{200}{\varepsilon_1} \log^3 n$ and $D \leq \log^{5k} n$.

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, as G is bipartite, C has even length, let ℓ_0 be such that $2\ell_0$ is the length of C. Since $\delta(G) \geq d - 1$, we must have $\ell_0 \leq \log n / \log(d-1) \leq m$, as $n \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large. Pick vertices $x_3, x_4 \in V(C)$ which are distance $\ell_0 - 1$ apart on C and let the paths separating them in C be R_1 and R_2, where R_1 is the shorter path.

Let $D_{1,1} = D_{2,1} = D$, $D_{1,2} = D_{3,1} = m^3 D$, $D_{2,2} = D_{4,1} = m^2 D$, and note that $m^3 D \leq \log^{15k} n$. Using any arbitrary vertices $x_5, \ldots, x_{15k}$ and $D_{i,j} = D$ for any $i, j \in [15k]$ not already chosen, apply Lemma 3.11 to $x_1, \ldots, x_{15k}$ and C with $(k, m)_{3.11} = (15k, m/5)$ to get graphs $F_{i,j}$, $i, j \in [2]$ and $F_{3,1}, F_{4,1}$, for which the following hold.

- For each $i, j \in [2]$ or $i \in \{3,4\}$ and $j = 1$, $F_{i,j}$ is a $(D_{i,j}, m)$-expansion around x_i in G which contains no vertices other than x_i in $\{x_1, \ldots, x_4\} \cup V(C)$.

- The sets $V(F_{i,j}) \setminus \{x_i\}$, $i, j \in [2]$ or $i \in \{3,4\}$ and $j = 1$, are pairwise disjoint.

Now, we have $|V(C) \cup V(F_{1,1} \cup F_{2,1} \cup F_{2,2} \cup F_{4,1})| \leq m + 2D + 2m^2 D \leq 10m^3 D / \log^3 n$. Therefore, as $|F_{1,2}| = |F_{3,1}| = m^3 D$, by Lemma 3.4, we can find a path P' with length at most m from $V(F_{1,2})$ to $V(F_{3,1})$ with no vertices in $(V(C) \cup V(F_{1,1} \cup F_{2,1} \cup F_{2,2} \cup F_{4,1})) \setminus \{x_1, x_3\}$. As, $F_{1,2}$ is a $(D_{1,2}, m)$-expansion of x_1, and $F_{3,1}$ is a $(D_{3,1}, m)$-expansion of x_3, we can extend P' using vertices from $V(F_{1,2} \cup F_{1,3})$ to get an x_1, x_3-path, say P, with length at most $3m$, which has no vertices in $(V(C) \cup V(F_{1,1} \cup F_{2,1} \cup F_{2,2} \cup F_{4,1})) \setminus \{x_1, x_3\}$.

Next, observe that $|V(C) \cup V(P) \cup V(F_{1,1}) \cup V(F_{2,1})| \leq m + 3m + 1 + 2D \leq 10m^2 D / \log^3 n$. Therefore, as $|F_{2,2}| = |F_{4,1}| = m^2 D$, by Lemma 3.4, we can find a path Q' with length at most m from $V(F_{2,2})$ and $V(F_{4,1})$ which has no vertices in $(V(C) \cup V(P) \cup V(F_{1,1}) \cup V(F_{2,1})) \setminus \{x_2, x_4\}$. As, $F_{2,2}$ is a $(D_{2,2}, m)$-expansion of x_2, and $F_{4,1}$ is a $(D_{4,1}, m)$-expansion of x_4, we can extend Q' using vertices from $V(F_{2,2} \cup F_{4,1})$ to get an x_2, x_4-path, say Q, with length at most $3m$ and no vertices in $(V(C) \cup V(P) \cup V(F_{1,1}) \cup V(F_{2,1})) \setminus \{x_2, x_4\}$.

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 $(D, m, 1)$-adjuster, as desired. $\square$

4.2   Finding simple adjusters robustly

In this section, we prove Lemma 4.3, a key component of our proof. This finds a simple adjuster robustly in an expander G – that is, given any subset $U \subseteq V(G)$ with moderate size, we construct an adjuster in $G - U$.

Lemma 4.3. *There exists some $\varepsilon_1 > 0$ such that, for every $0 < \varepsilon_2 < 1$ and $k \in \mathbb{N}$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following is true for each $n \geq d \geq d_0$. Suppose that G is a $\mathsf{TK}^{(2)}_{d/2}$-free n-vertex bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$. Let $m = \frac{200}{\varepsilon_1} \log^3 n$ and $D \leq \log^k n$. Let $U \subseteq V(G)$ satisfy $|U| \leq 10D$.*
   Then, $G - U$ contains a $(D, 2m, 1)$-adjuster.

The following proof sketch is illustrated in Figure 2. Essentially, we find an expander subgraph in $H \subseteq G - U$ and apply Lemma 4.2 to find a simple adjuster in H. However, H may be much smaller than G, so this simple adjuster may be far too small to satisfy Lemma 4.3. We thus find many of these simple adjusters and use Lemma 3.7 to expand the ends of one of them to make them large enough to satisfy Lemma 4.3.

More precisely, to prove Lemma 4.3, we assume no such adjuster exists, before collecting the high degree vertices in a set L. We take a maximal set of adjusters $\mathbf{A}_0$ in $G - U$ so that their ends (the sets $V(F_1)$ and $V(F_2)$ in an adjuster (v_1, F_1, v_2, F_2, A)) are in $G - L$, and furthermore the ends of different adjusters in $\mathbf{A}_0$ are far apart in $G - L$. The adjusters in $\mathbf{A}_0$ will each not have large enough ends to be a $(D, m, 1)$-adjuster, so we wish to expand the ends of some adjuster to make them larger. The challenge is to do this while avoiding U.

We first show that $\mathbf{A}_0$ contains many adjusters (see Claim 4.4). If this is not the case, then, collecting together U with the adjusters in $\mathbf{A}_0$ and any vertices near their ends in $G - L$, we remove them and show that there must be an expander subgraph H in what remains. Applying Lemma 4.2 to (essentially) H, we get an adjuster that either satisfies the lemma (a contradiction) or should have belonged in $\mathbf{A}_0$ (another contradiction).

If many adjusters in $\mathbf{A}_0$ have an end with a short path to $L \setminus U$, then applying Lemma 3.7 shows that in one of these adjusters the other end must expand while avoiding the short path to $L \setminus U$. Larger ends can then be chosen for this adjuster respectively from the expansion and from the short path to $L \setminus U$ and the neighbourhood of its endvertex in $L \setminus U$. This gives an adjuster satisfying the lemma (a contradiction). Thus, many adjusters in $\mathbf{A}_0$ have no short path to $L \setminus U$ (see Claim 4.5) – we collect such adjusters in $\mathbf{A}_1 \subseteq \mathbf{A}_0$.

We then find a large set Z in $G - L$ with small diameter using Lemma 3.12 – destined to provide a large expansion for the end of an adjuster. If many adjusters in $\mathbf{A}_1$ have an end with a short path to Z, then applying Lemma 3.7 shows that, for one of these adjusters, the other end must expand while avoiding the short path to Z. Larger ends can then be chosen for this adjuster respectively from the expansion and from the short path to Z and Z itself. This gives an adjuster satisfying the lemma (a contradiction). Thus, many adjusters in $\mathbf{A}_1$ have no short path to Z (see Claim 4.6) – we collect such adjusters in $\mathbf{A}_2 \subseteq \mathbf{A}_1$.

However, by Lemma 3.7, the ends of the adjusters in $\mathbf{A}_2$ must expand while avoiding U. Therefore, by Lemma 3.4, one of them must connect to Z, giving the final contradiction which completes the proof.
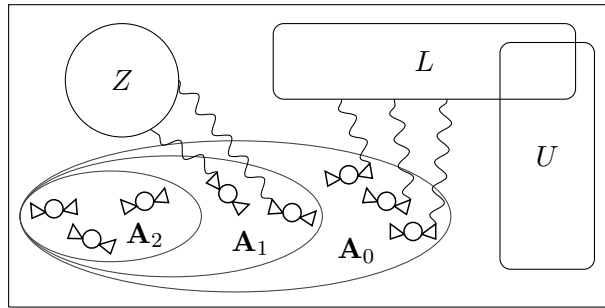

Figure 2: An illustration of the proof of Lemma 4.3. We find $\mathbf{A}_0$, a large set of adjusters in $G - U$, before discarding those with a short path to $L \setminus U$, and then those with a short path to Z. Showing that many adjusters still remain, in the set $\mathbf{A}_2$, leads to a contradiction.

Proof of Lemma 4.3. Let $0 < \varepsilon_1 < 1$ be small enough that the property in Corollary 2.5 holds. Suppose, for contradiction, that $G - U$ contains no $(D, 2m, 1)$-adjuster. Let $\Delta = 200mD$, $L = \{v \in V(G) : d_G(v) \geq \Delta\}$ and $G' = G - L$, so that $\Delta(G') \leq \Delta$.

Set $\ell_0 = (\log\log n)^{20}$. Let $U_0 = \{v \in V(G) \setminus U : d_G(v, U) \geq d/2\}$. Note that if $|U_0| \geq 100D^2 \geq |U|^2$, then by Proposition 3.16 with $(U, W)_{3.16} = (U_0, U)$, G contains a $\mathsf{TK}^{(2)}_{d/2}$, a contradiction. Therefore, we can assume that $|U_0| \leq 100D^2$, and hence, as $\delta(G) \geq d$ and $n \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large, $G - U$ contains at least $(n - |U| - |U_0|) \cdot (d/2)/2 \geq nd/8$ edges. Let $U_1 = U \cup U_0$, so that $|U_1| \leq 200D^2 \leq 200\log^{2k} n$.

Take a maximal collection $\mathbf{A}_0$ of adjusters in $G - U$, such that the following hold.

G1 The sets $V(F_1 \cup F_2)$, $(v_1, F_1, v_2, F_2, A) \in \mathbf{A}_0$, are subsets of $V(G')$ and are all at least a distance $10\ell_0$ apart from each other and from $U_1 \setminus L$ in G'.

G2 For each $\mathcal{A} \in \mathbf{A}_0$, for some $m_{\mathcal{A}}$ with $\log^3 d_0 \leq m_{\mathcal{A}} \leq m$, $\mathcal{A}$ is an $(m_{\mathcal{A}}^2, m_{\mathcal{A}}, 1)$-adjuster.

Claim 4.4. $|\mathbf{A}_0| \geq n^{1/4}$.

Proof of claim. Suppose, for contradiction, that $|\mathbf{A}_0| < n^{1/4}$. Let $W = (U_1 \cup (\cup_{\mathcal{A} \in \mathbf{A}_0} V(\mathcal{A}))) \setminus L$. For each $\mathcal{A} = (v_1, F_1, v_2, F_2, A) \in \mathbf{A}_0$, $|V(\mathcal{A})| = |F_1| + |F_2| + |A| \leq 2m_{\mathcal{A}}^2 + 10m_{\mathcal{A}} \leq 3m^2$, and therefore $|W| \leq n^{1/4} \cdot 3m^3 + 200\log^{2k} n \leq n^{1/3}$. Let $W' = B_{G'}^{10\ell_0}(W)$, so, as $\Delta(G') \leq \Delta$, we have that $|W'| \leq 2|W| \cdot \Delta^{10\ell_0} \leq n^{1/2}$.

Now, there are at most $|W'|\Delta \leq \Delta n^{1/2} \leq nd/16$ edges in G with some vertex in W'. Let $\bar{d} = d/64$. As $G - U$ contains at least $nd/8$ edges, $G - U - W'$ contains at least $nd/16$ edges, so that $d(G - U - W') \geq d/8 = 8\bar{d}$. Then, by Corollary 2.5, $G - U - W'$ contains an $(\varepsilon_1, \varepsilon_2\bar{d})$-expander H with $\delta(H) \geq \bar{d}$. Let C be a shortest cycle in H. We will consider two cases, depending on how many vertices of L there are in $V(H) \setminus V(C)$.

Case I: $|(V(H) \setminus V(C)) \cap L| \leq 1$. Let $H' = H - (V(H) \setminus V(C)) \cap L$, so that $\delta(H') \geq \bar{d} - 1$. Note that, for each $X \subseteq V(H')$ with $\varepsilon_2\bar{d}/2 \leq |X| \leq |H'|/2 < |H|/2$, we have

$$|N_{H'}(X)| \geq |N_H(X)| - 1 \geq |X| \cdot \varepsilon(|X|, \varepsilon_1, \varepsilon_2\bar{d}) - 1$$

$$\geq \frac{1}{2}|X| \cdot \varepsilon(|X|, \varepsilon_1, \varepsilon_2\bar{d}) + \frac{\varepsilon_2\bar{d}}{4} \cdot \varepsilon(\varepsilon_2\bar{d}/2, \varepsilon_1, \varepsilon_2\bar{d}) - 1$$

$$\geq |X| \cdot \varepsilon(|X|, \varepsilon_1/2, \varepsilon_2\bar{d}) + \frac{\varepsilon_2\bar{d}}{4} \cdot \frac{\varepsilon_1}{\log^2(15/2)} - 1 \geq |X| \cdot \varepsilon(|X|, \varepsilon_1/2, \varepsilon_2\bar{d}),$$

where the last inequality follows as $\bar{d} \geq d_0(\varepsilon_1, \varepsilon_2, k)/64$ is large. Therefore, H' is a $(\varepsilon_1/2, \varepsilon_2\bar{d})$-expander with $\delta(H') \geq \bar{d} - 1$. Note that C is a shortest cycle in H'.

Let $m_{H'} = 200\log^3 |H'|/\varepsilon_1 \leq m$, and note that, as $|H'| \geq \delta(H') + 1 \geq \bar{d} \geq d_0/64$, and $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ is large, $m_{H'} \geq \log^3 d_0$. Picking arbitrary vertices $x_1, x_2 \in V(H') \setminus V(C)$ and noting that $\bar{d} \geq d_0(\varepsilon_1, \varepsilon_2, k)/64$ is large, by Lemma 4.2 with $(k, D)_{4.2} = (10, m_{H'}^2)$, H' contains an $(m_{H'}^2, m_{H'}, 1)$-adjuster (v_1, F_1, v_2, F_2, A) with $V(C) \subseteq A$. As A is disjoint from $V(F_1 \cup F_2)$, $V(C) \subset A$ and $(V(H') \setminus V(C)) \cap L = \varnothing$, we have that $V(F_1 \cup F_2)$ is disjoint from L, and hence lies in $V(G')$. Together with $V(F_1 \cup F_2) \subseteq V(H')$ being disjoint from W' and so $10\ell_0$-far in G' from the ends of the adjusters in $\mathbf{A}_0$ and from $U_1 \setminus L$, this violates the maximality of $\mathbf{A}_0$, a contradiction.

Case II: $|(V(H) \setminus V(C)) \cap L| \geq 2$. Let $x_1, x_2 \in (V(H) \setminus V(C)) \cap L$ be distinct and let $m_{H'} = 200\log^3 |H'|/\varepsilon_1 \leq m$. By Lemma 4.2 with $(k, D)_{4.2} = (1, 1)$, H contains a $(1, m_{H'}, 1)$-adjuster (v_1, F_1, v_2, F_2, A) with $v_1 = x_1$ and $v_2 = x_2$. Using that $|A| \leq 10m_{H'} \leq 10m$, $|U| \leq 10D$, and $d_G(x_1), d_G(x_2) \geq \Delta = 200mD$, pick disjointly sets $X_1 \subseteq N_G(x_1) \setminus (U \cup A \cup \{x_2\})$ and $X_2 \subseteq N_G(x_2) \setminus (U \cup A \cup \{x_1\})$ with $|X_1| = |X_2| = D - 1$. Letting $F_i' = G[\{x_i\} \cup X_i]$ for each $i \in [2]$, and noting $|A| \leq 20m$, we have that $(x_1, F_1', x_2, F_2', A)$ is a $(D, 2m, 1)$-adjuster in $G - U$, a contradiction. $\blacksquare$

Now, let $\mathbf{A}_1 \subseteq \mathbf{A}_0$ be the set of adjusters $(v_1, F_1, v_2, F_2, A) \in \mathbf{A}_0$ for which there is no path with length at most ℓ_0 from $V(F_1) \cup V(F_2)$ to $L \setminus U$ in $G - U - A$.

Claim 4.5. $|\mathbf{A}_1| \geq n^{1/4}/2$.

Proof of claim. Let $r = n^{1/8}$. Suppose, for contradiction, that we can label distinct $\mathcal{A}_1, \ldots, \mathcal{A}_r \in \mathbf{A}_0 \setminus \mathbf{A}_1$. Say, for each $i \in [r]$, that $\mathcal{A}_i = (v_{i,1}, F_{i,1}, v_{i,2}, F_{i,2}, \bar{A}_i)$ and let P'_i be a shortest path with length at most ℓ_0 from $V(F_{i,1}) \cup V(F_{i,2})$ to $L \setminus U$ in $G - U - \bar{A}_i$. Relabelling, if necessary, for each $i \in [r]$ suppose the endvertex of P'_i in $V(F_{i,1} \cup F_{i,2})$ is in $V(F_{i,1})$, and let Q_i be a path from this endvertex of P'_i to $v_{i,1}$ in $F_{i,1}$ with length at most $m_{\mathcal{A}_i}$.

For each $i \in [r]$, let x_i be the endpoint of P'_i in $L \setminus U$, and let $P_i = P'_i - x_i$. We shall apply Lemma 3.7 by setting, for each $i \in [r]$, $A_i = V(F_{i,2})$, $B_i = \bar{A}_i \cup V(Q_i) \cup \{x_i\}$ and $C_i = V(P_i)$. Firstly, as $|A_i| = m^2_{\mathcal{A}_i} \geq \log^6 d_0$ by **G2** and $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ is large, we have that $|A_i| \geq d_0^{\text{L}3.7}$, where $d_0^{\text{L}3.7}$ is the function in Lemma 3.7, so that **C1** holds.

As $V(F_{i,2}) \subseteq V(G') = V(G) \setminus L$ by **G1**, and $V(F_{i,2})$ is disjoint from $V(F_{i,1})$ and $\bar{A}_i$ by **F1**, we have that A_i and $B_i \cup C_i$ are disjoint. Furthermore, $|B_i| \leq |\bar{A}_i| + |Q_i| + 1 \leq 20 m_{\mathcal{A}_i} \leq m^2_{\mathcal{A}_i}/\log^{10}(m^2_{\mathcal{A}_i})$ as $m_{\mathcal{A}_i} \geq \log^3(d_0(\varepsilon_1, \varepsilon_2, k))$ is large, and thus **C2** holds.

Now, as P'_i is a shortest path from $V(F_{i,1}) \cup V(F_{i,2})$ to $L \setminus U$ in $G - U - \bar{A}_i$, which has an endvertex in $V(F_{i,1})$, and $A_i = V(F_{i,2})$, we have, for each $\ell \in \mathbb{N}$, that $B^\ell_{G - U - \bar{A}_i}(A_i)$ has at most $\ell + 1$ vertices in P'_i, and hence P_i. Therefore, A_i has 4-limited contact with C_i in $G - U - \bar{A}_i$, and hence in $G - U - B_i$, and thus **C3** holds.

Suppose there is a path, R_i say, with length at most $10\ell_0$ from A_i to $L \setminus (U \cup \{x_i\})$ in $G - U - B_i - C_i$. Then, there is a path $R'_i \subseteq R_i \cup F_{i,2}$ from $v_{i,2}$ to some vertex $y_i \in L \setminus (U \cup \{x_i\})$ with length at most $10\ell_0 + m_{\mathcal{A}_i} \leq 2m - 1$, and the path $Q_i \cup P'_i$ is a path from $v_{i,1}$ to x_i with length at most $m_{\mathcal{A}_i} + \ell_0 \leq 2m - 1$ in $G - U - \bar{A}_i$ with vertices in $B_i \cup C_i$. Then, as $|U \cup A_i \cup V(R'_i) \cup V(Q_i \cup P'_i)| \leq 10D + 10m_{\mathcal{A}_i} + 4m \leq 10D + 15m$, as $x_i, y_i \in L$ both have degree at least $\Delta = 200mD$, we can comfortably choose $X_i \subseteq N_G(x_i)$ and $Y_i \subseteq N_G(y_i)$ which are disjoint from each other and from $U \cup A_i \cup V(R'_i) \cup V(Q_i \cup P'_i)$ and have size $D - |P'_i \cup Q_i|$ and $D - |R'_i|$ respectively. Then, $(v_{i,1}, G[X_i \cup V(P'_i) \cup V(Q_i)], v_{i,2}, G[Y_i \cup V(R'_i)], \bar{A}_i)$ is a $(D, 2m, 1)$-adjuster in $G - U$, a contradiction. Therefore, there is no such path R_i, so that

$$B^{\ell_0}_{G - U - B_i - C_i}(A_i) = B^{\ell_0}_{G' - U - B_i - C_i}(A_i),$$

which, by **G1**, is disjoint from U_1. By the choice of $U_0 \subseteq U_1$, we have that **C4** holds.

Now, similarly, for any $j \in [r] \setminus \{i\}$, we have that $B^{\ell_0}_{G - U - B_j - C_j}(A_j) = B^{\ell_0}_{G' - U - B_j - C_j}(A_j)$, so that, by **G1**, $B^{\ell_0}_{G - U - B_j - C_j}(A_j)$ and $B^{\ell_0}_{G - U - B_i - C_i}(A_i)$ are disjoint. In particular, A_i and A_j are a distance at least $2\ell_0$ apart in $G - U - B_i - C_i - B_j - C_j$, and therefore **C5** holds.

Thus, by Lemma 3.7, there is some $j \in [r]$ for which $|B^{\ell_0}_{G - U - B_j - C_j}(A_j)| \geq \log^k n \geq D$. As $F_{j,2}$ is an $(m^2_{\mathcal{A}_j}, m_{\mathcal{A}_j})$-expansion of $v_{j,2}$ in $G' - U - B_j - C_j$, $m_{\mathcal{A}_j} \leq m$ and $A_j = V(F_{j,2})$, we have that $|B^{2m}_{G - U - B_j - C_j}(v_{j,2})| \geq D$. Therefore, by Proposition 3.10, we can pick a $(D, 2m)$-expansion, $F'_{j,2}$ say, of $v_{j,2}$ in $G - U - B_i - C_j$.

As $x_j \in L$, we can then pick a set U' of neighbours of x_j disjoint from $U \cup V(F'_{j,2}) \cup \bar{A}_j \cup V(Q_j) \cup V(P'_j)$ with $|U'| = D - V(P'_j \cup Q_j)$. Let $F'_{j,1} = G[U' \cup V(P'_j) \cup V(Q_j)]$. Note that $F'_{j,1}$ is then a $(D, 2m)$-expansion of $v_{j,1}$ as $Q_j \cup P'_j$ is a $v_{j,1}, x_j$-path with length at most $m_{\mathcal{A}_j} + \ell_0 \leq 2m - 1$. Finally, note that $(v_{j,1}, F'_{j,1}, v_{j,2}, F'_{j,2}, \bar{A}_j)$ is a $(D, 2m, 1)$-adjuster in $G - U$, a contradiction. Therefore, by Claim 4.4, we have $|\mathbf{A}_1| > n^{1/4} - r \geq n^{1/4}/2$. ∎

Let $\mathbf{A}_1' \subseteq \mathbf{A}_1$ satisfy $|\mathbf{A}_1'| = n^{1/4}/2$. Then, $|\cup_{\mathcal{A} \in \mathbf{A}_1'} V(\mathcal{A})| \leq n^{1/4} \cdot 3m^2 \leq n^{1/3}$ by **G2**. Therefore,

$$|U \cup B_{G'}^{\ell_0}(\cup_{\mathcal{A} \in \mathbf{A}_1'}(V(\mathcal{A}) \setminus L))| \leq 10D + n^{1/3} \cdot 2\Delta^{\ell_0} \leq n^{1/2}.$$

Thus, by Lemma 3.12, there is a set $Z \subseteq V(G) \setminus U$ which has diameter at most $m/2$ and size $10m^2D$, and is a distance at least ℓ_0 in G' from $V(\mathcal{A}) \setminus L$ for each $\mathcal{A} \in \mathbf{A}_1'$.

Let $\mathbf{A}_2 \subseteq \mathbf{A}_1'$ be the set of adjusters $(v_1, F_1, v_2, F_2, A) \in \mathbf{A}_1'$ for which there is no path with length at most $m/2$ from $V(F_1) \cup V(F_2)$ to Z in $G - U - A$.

Claim 4.6. $|\mathbf{A}_2| \geq n^{1/4}/4$.

Proof of claim. Let $r = n^{1/8}$. Suppose, for contradiction, we can label distinct $\mathcal{A}_1, \ldots, \mathcal{A}_r \in \mathbf{A}_1' \setminus \mathbf{A}_2$. Say, for each $i \in [r]$, that $\mathcal{A}_i = (v_{i,1}, F_{i,1}, v_{i,2}, F_{i,2}, \bar{A}_i)$ and let P_i be a shortest path with length at most $m/2$ from $V(F_{i,1}) \cup V(F_{i,2})$ to Z in $G - U - \bar{A}_i$. Relabelling, if necessary, for each $i \in [r]$ suppose the endvertex of P_i in $V(F_{i,1} \cup F_{i,2})$ is in $V(F_{i,1})$, and let Q_i be a path from this endvertex of $V(P_i)$ to $v_{i,1}$ in $F_{i,1}$ with length at most $m_{\mathcal{A}_i}$.

We will apply Lemma 3.7 to $A_i = V(F_{i,2})$, $B_i = \bar{A}_i \cup V(Q_i)$ and $C_i = V(P_i)$, for each $i \in [r]$. For each $i \in [r]$, similarly to the proof of Claim 4.5, we have that **C1**–**C3** hold. By the choice of $\mathbf{A}_1$, for each $i \in [r]$, there is no path of length at most ℓ_0 from A_i to $L \setminus U$ in $G - U - B_i - C_i$. Therefore, by **G1**, $B_{G-U-B_i-C_i}^{\ell_0}(A_i) = B_{G'-U-B_i-C_i}^{\ell_0}(A_i)$ is disjoint from U_1. Thus, **C4** holds by the definition of U_1. It similarly follows that $B_{G-U-B_i-C_i}^{\ell_0}(A_i)$ and $B_{G-U-B_j-C_j}^{\ell_0}(A_j)$ are vertex disjoint for each $j \in [r] \setminus \{i\}$, and thus **C5** holds.

Thus, by Lemma 3.7, there is some $j \in [r]$ for which $|B_{G'-U-B_j-C_j}^{\ell_0}(A_j)| = |B_{G-U-B_j-C_j}^{\ell_0}(A_j)| \geq D$. Thus, as $F_{j,2}$ is an $(m_{\mathcal{A}_j}^2, m_{\mathcal{A}_j})$-expansion of $v_{j,2}$ in $G' - U - B_j - C_j$ by **G1** and **G2**, and $A_j = V(F_{j,2})$, by Proposition 3.10, there is a $(D, 2m)$-expansion, $F_{j,2}'$ say, of $v_{j,2}$ in $B_{G'-U-B_j-C_j}^{\ell_0}(V(F_{j,2}))$. As Z was chosen to have a distance at least ℓ_0 in G' from $V(\mathcal{A}_j) \setminus L$, we have that $V(F_{j,2}')$ is disjoint from Z.

Now, as Z has diameter at most $m/2$ in G, $Q_j \cup P_j \cup G[Z]$ is an expansion of $v_{j,1}$ with radius at most $\ell(Q_j) + \ell(P_j) + m/2 \leq 2m$ and size at least D. Therefore, by Proposition 3.10, we can find within $Q_j \cup P_j \cup G[Z]$ a $(D, 2m)$-expansion, $F_{j,1}'$ say, of $v_{j,1}$, which then must be vertex-disjoint from $\bar{A}_j$ and from $V(F_{j,2}') \subset B_{G'-U-B_j-C_j}^{\ell_0}(V(F_{j,2}))$. Thus, we have that $(v_{j,1}, F_{j,1}', v_{j,2}, F_{j,2}', \bar{A}_j)$ is a $(D, 2m, 1)$-adjuster in $G - U$, a contradiction. Thus, $|\mathbf{A}_2| \geq |\mathbf{A}_1| - r \geq n^{1/4}/4$, by Claim 4.5. ∎

Let $r = n^{1/8}$. Using Claim 4.6, label distinct $\mathcal{A}_1, \ldots, \mathcal{A}_r \in \mathbf{A}_2$, and say, for each $i \in [r]$, that $\mathcal{A}_i = (v_{i,1}, F_{i,1}, v_{i,2}, F_{i,2}, \bar{A}_i)$. We shall apply Lemma 3.7 to $A_i = V(F_{i,1} \cup F_{i,2})$, $B_i = \bar{A}_i$ and $C_i = \varnothing$. Similarly as in the proof of Claim 4.6, the only difference being that **C3** holds trivially as $C_i = \varnothing$ and A_i is slightly larger, we have that **C1**–**C5** hold.

Thus, by Lemma 3.7, there is some $j \in [r]$ with $|B_{G-U-B_j-C_j}^{\ell_0}(A_j)| \geq 10m^2D \geq 10\log^3 n|U \cup B_j|$. Therefore, by Lemma 3.4, there is a path in $G - U - B_j$ from $B_{G-U-B_j}^{\ell_0}(A_j)$ to Z with length at most $m/4$. Then, as $A_j = V(F_{j,1} \cup F_{j,2})$ and $B_j = \bar{A}_j$, there is a path in $G - U - \bar{A}_j$ from $V(F_{j,1} \cup F_{j,2})$ to Z with length at most $m/2$, contradicting $\mathcal{A}_j \in \mathbf{A}_2$, and completing the proof. ☐

4.3 Connecting simple adjusters for paths with specific lengths

Using Lemma 4.3, we can find many vertex disjoint simple adjusters. We now connect them together into a larger adjuster, for Lemma 4.7, before using these to construct paths with specific lengths

for Lemma 4.8.

Lemma 4.7. *There exists some $\varepsilon_1 > 0$ such that, for any $0 < \varepsilon_2 < 1/5$ and $k \geq 10$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex $\mathsf{TK}_{d/2}^{(2)}$-free bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

Let $m = \frac{800}{\varepsilon_1} \log^3 n$. Suppose $\log^{10} n \leq D \leq \log^k n$, $1 \leq r \leq 20m$ and $U \subseteq V(G)$ with $|U| \leq D$. Then, there is a (D, m, r)-adjuster in $G - U$.

Proof. Let $\varepsilon_1 > 0$ be such that the property in Lemma 4.3 holds. By this property, as $d \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large, for every set $V \subseteq V(G)$ with $|V| \leq \log^{2k} n$, $G - V$ contains a $(D, m/2, 1)$-adjuster. By Lemma 3.4, as $d \geq d_0(\varepsilon_1, \varepsilon_2, k)$ is large, for any sets X and Y with size at least $2D$, and any set $V \subseteq V(G) \setminus (X \cup Y)$ with size at most $20D/\log^3 n$, there is a path from X to Y in $G - V$ with length at most m.

We now prove the property in the lemma by induction on r. Note that, we already have this property for $r = 1$ as $|U| \leq D \leq \log^{2k} n$, and a $(D, m/2, 1)$-adjuster is also a $(D, m, 1)$-adjuster. Suppose then, for some r with $1 \leq r < 20m$, $G - U$ contains a (D, m, r)-adjuster, $(v_1, F_1, v_2, F_2, A_1)$ say. Let $U' = U \cup A_1 \cup V(F_1) \cup V(F_2)$, so that $|U'| \leq 4D \leq \log^{2k} n$. Therefore, $G - U'$ contains a $(D, m/2, 1)$-adjuster, $(v_3, F_3, v_4, F_4, A_2)$ say. As $|F_1 \cup F_2| = |F_3 \cup F_4| = 2D$, and $|A_1 \cup A_2| \leq 20rm \leq 400m^2 \leq D/\log^3 n$, there is a path, P say, with length at most m, from $V(F_1) \cup V(F_2)$ to $V(F_3) \cup V(F_4)$ avoiding $A_1 \cup A_2$.

Note that, without loss of generality, we can assume that P is a path from $V(F_1)$ to $V(F_3)$. Using that F_1 and F_3 is a (D, m)-expansion of v_1 and v_3 respectively, take a v_1, v_3-path $Q \subseteq F_1 \cup P \cup F_3$ with length at most $5m$. Then, we claim $(v_2, F_2, v_4, F_4, A_1 \cup A_2 \cup V(Q))$ is a $(D, m, r+1)$-adjuster. Indeed, we easily have that **F1** and **F2** hold, and $|A_1 \cup A_2 \cup V(Q)| \leq 5m + 10 \cdot (m/2) + 10mr = 10(r+1)m$, so that **F3** holds.

Finally, let $\ell_1 = \ell((v_1, F_1, v_2, F_2, A_1))$, $\ell_2 = \ell((v_3, F_3, v_4, F_4, A_2))$ and $\ell = \ell_1 + \ell_2 + \ell(Q)$. If $i \in \{0, 1, \ldots, r+1\}$, then there is some $i_1 \in \{0, 1, \ldots, r\}$ and $i_2 \in \{0, 1\}$ with $i = i_1 + i_2$. Let P_1 be a v_2, v_1-path in $G[A_1 \cup \{v_1, v_2\}]$ with length $\ell_1 + 2i_1$ and let P_2 be a v_3, v_4-path with length $\ell_2 + 2i_2$ in $G[A_2 \cup \{v_3, v_4\}]$. Then, $P_1 \cup Q \cup P_2$ is a v_2, v_4-path in $G[A_1 \cup A_2 \cup V(Q)]$ with length $\ell + 2i$, and thus ℓ satisfies **F4**. $\square$

Combining Lemma 4.7 with Corollary 3.15, we can finally find paths with exactly some desired length, as follows.

Lemma 4.8. *There exists some $\varepsilon_1 > 0$ such that, for any $0 < \varepsilon_2 < 1/5$ and $k \geq 10$, there exists $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ such that the following holds for each $n \geq d \geq d_0$. Suppose that G is an n-vertex $\mathsf{TK}_{d/2}^{(2)}$-free bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(G) \geq d$.*

Suppose $\log^{10} n \leq D \leq \log^k n$, and $U \subseteq V(G)$ with $|U| \leq D/2\log^3 n$, and let $m = \frac{800}{\varepsilon_1} \log^3 n$. Suppose $F_1, F_2 \subseteq G - U$ are vertex disjoint such that F_i is a (D, m)-expansion of v_i, for each $i \in [2]$. Let $\log^7 n \leq \ell \leq n/\log^{10} n$ be such that $\ell = \pi(v_1, v_2, G) \mod 2$.

Then, there is a v_1, v_2-path with length ℓ in $G - U$.

Proof. By Lemma 4.7, there is a $(D, m, 20m)$-adjuster, $\mathcal{A} = (v_3, F_3, v_4, F_4, A)$ say, in $G - U$ with length $\ell(\mathcal{A}) \leq |A| + 1 \leq 400m^2$. Let $\bar{\ell} = \ell - 20m - \ell(\mathcal{A})$, so that $0 \leq \bar{\ell} \leq n/\log^{10} n$. As $|A \cup U| \leq 400m^2 + D/2\log^3 n \leq D/\log^3 n$, by Corollary 3.15, there are paths P and Q in $G - U - A$ which are vertex disjoint, both connect $\{v_1, v_2\}$ to $\{v_3, v_4\}$ and so that $\bar{\ell} \leq \ell(P) + \ell(Q) \leq \bar{\ell} + 20m$. Note that we can assume, without loss of generality, that P is a v_1, v_3-path and Q is a v_2, v_4-path.

Now, $0 \leq \ell - \ell(P) - \ell(Q) - \ell(\mathcal{A}) \leq 20m$. As $\mathcal{A}$ is a $(D, m, 20m)$-adjuster there is a v_3, v_4-path in $G[A \cup \{v_3, v_4\}]$ with length $\ell(\mathcal{A})$, and therefore $\ell(\mathcal{A}) = \pi(v_3, v_4, G) \mod 2$. Then, as $\ell(P) = \pi(v_1, v_3, G) \mod 2$, $\ell(Q) = \pi(v_2, v_4, G) \mod 2$, $\ell = \pi(v_1, v_2, G) \mod 2$ and $\pi(v_1, v_2, G) = \pi(v_1, v_3, G) + \pi(v_3, v_4, G) + \pi(v_4, v_2, G) \mod 2$, we have $\ell - \ell(P) - \ell(Q) - \ell(\mathcal{A}) = 0 \mod 2$. That is, there is some $i \in \mathbb{N}$ with $2i = \ell - \ell(P) - \ell(Q) - \ell(\mathcal{A})$, where $i \leq 10m$.

Therefore, by the property of the adjuster, there is a v_3, v_4-path, R say, with length $\ell(\mathcal{A}) + 2i = \ell - \ell(P) - \ell(Q)$ in $G[A \cup \{v_3, v_4\}]$. Then, $P \cup R \cup Q$ is a v_1, v_2-path with length ℓ in $G - U$. $\qquad\square$

4.4   Proof of Theorem 1.7

We can now prove Theorem 1.7. We take some core vertices $v_1, \ldots, v_k$ in an expander, find expansions around them using Lemma 3.11 and then connect each pair of core vertices using Lemma 4.8.

Proof of Theorem 1.7. Let $\varepsilon_1 > 0$ be such that the properties in Corollary 2.5 and Lemma 4.8 hold. Let $\varepsilon_2 = 1/10$. Let $d_0 = d_0(\varepsilon_1, \varepsilon_2, k)$ be large, and let $d = 8d_0$. Let G be a graph with $d(G) \geq d$.

By Corollary 2.5, we can find a bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander $H \subseteq G$ with $\delta(H) \geq d_0$. Let $K = \binom{k}{2}$, $n = |H| \geq d_0$, $m = \frac{800}{\varepsilon_1} \log^3 n$ and $\ell = \log^7 n$. Take k distinct vertices in the same partition in H, say $v_1, \ldots, v_k$. As $d_0(\varepsilon_1, \varepsilon_2, k)$ is large and $m^{10K} \leq \log^{30k^2} n$, by Lemma 3.11 with $(k)_{3.11} = 30k^2$, we can find, for each $i, j \in [k]$ an (m^{10K}, m)-expansion $F_{i,j} \subseteq H$ of v_i so that the sets $V(F_{i,j}) \setminus \{v_i\}$ are pairwise disjoint over $i, j \in [k]$.

Let $f : [K] \to [k]^{(2)}$ be a bijection and let $g, h : [K] \to [k]$ be such that $f(i) = \{g(i), h(i)\}$ for each $i \in [K]$. For each $i \in [K]$, using Proposition 3.10, let $H_{i,1} \subseteq F_{g(i),h(i)}$ be such that $H_{i,1}$ is an $(m^{10(K+1-i)}, m)$-expansion of $v_{g(i)}$ and let $H_{i,2} \subseteq F_{h(i),g(i)}$ be such that $H_{i,2}$ is an $(m^{10(K+1-i)}, m)$-expansion of $v_{h(i)}$.

We shall connect pairs of core vertices, in the order given by f. For each $i \in [K]$, the expansions $H_{i,1}$ and $H_{i,2}$ will be used to connect $v_{g(i)}$ and $v_{h(i)}$. We will make sure that the expansions that are not yet used are protected. More precisely, we will find paths $P_1, \ldots, P_K$, each with length ℓ, so that the following hold.

H1 For each $i \in [K]$, P_i is a $v_{g(i)}, v_{h(i)}$-path with length ℓ.

H2 For each $i \in [K]$, $V(P_i)$ is disjoint from

$$U_i := (\{v_j : j \in [k]\} \cup (\cup_{j>i}(V(H_{j,1}) \cup V(H_{j,2}))) \cup (\cup_{j<i} V(P_j))) \setminus \{v_{g(i)}, v_{h(i)}\}.$$

This is sufficient to prove the theorem. Indeed, by **H2**, for each $1 \leq i < j \leq K$, as $V(P_i) \setminus (\{v_{g(i)}, v_{h(i)}\} \cap \{v_{g(j)}, v_{h(j)}\}) \subseteq U_j$, we have that P_i and P_j are internally disjoint. Therefore, $\cup_{i \in [K]} P_i$ is a copy of $\mathsf{TK}_k^{(\ell)}$.

Suppose then that $1 \leq i \leq K$, and we have found paths $P_1, \ldots, P_{i-1}$ satisfying **H1** and **H2**. Note that

$$|U_i| \leq k + \sum_{j>i} 2m^{10(K+1-j)} + \sum_{j<i} \ell \leq k + 4m^{10(K-i)} + k^2 m^3 \leq \frac{m^{10(K+1-i)}}{m^3} = \frac{|H_{i,1}|}{m^3} = \frac{|H_{i,2}|}{m^3}.$$

Therefore, by Lemma 4.8 with $(v_1, F_1, v_2, F_2, U)_{4.8} = (v_{g(i)}, H_{i,1}, v_{h(i)}, H_{i,2}, U_i)$, there is a path P_i with length ℓ between $v_{g(i)}$ and $v_{h(i)}$ which does not intersect U_i. This completes the proof. $\qquad\square$

4.5 Proof of Theorem 2.7

Finally, we combine Lemmas 3.11 and 4.8 to prove Theorem 2.7.

Proof of Theorem 2.7. Let $\varepsilon_1 > 0$ be such that the property in Lemma 4.8 holds. Let $k = 10$, let $d_0 = d_0(\varepsilon_1, \varepsilon_2)$ be large and let $n \geq d \geq d_0$. Suppose then that H is a $\mathsf{TK}_{d/2}^{(2)}$-free bipartite n-vertex $(\varepsilon_1, \varepsilon_2 d)$-expander with $\delta(H) \geq d$ and let $x, y \in V(H)$ be distinct. Let $\ell \in [\log^7 n, n/\log^{10} n]$ satisfy $\ell = \pi(x, y, H) \mod 2$. We will show that H contains an x, y path with length ℓ.

Let $m = \frac{800}{\varepsilon_1} \log^3 n$ and $D = \log^{10} n$. Then, by Lemma 3.11 (applied with C taken to be an arbitrary shortest cycle in H), there are vertex disjoint graphs $F_x, F_y \subseteq H$ so that F_x is a (D, m)-expansion of x and F_y is a (D, m)-expansion of y. Then, by Lemma 4.8 with $U = \emptyset$, there is a x, y-path with length ℓ in H, as required. □

5 Proof of Theorem 1.4

We will now prove Theorem 1.4 using Theorem 2.7. For convenience, before we discuss the proof further, we will prove the following corollary of Theorem 2.7.

Corollary 5.1. *For each $\varepsilon > 0$, there is some d_0 such that the following holds for each $d \geq d_0$. If a graph G has $d(G) \geq 10d$, then it contains a connected bipartite subgraph H for which there is some positive integer ℓ such that the following holds.*

> ***I*** *For any $u, v \in V(H)$ with $u \neq v$ and $t \in [\ell, \ell \cdot d^{1-\varepsilon}]$ with $t = \pi(u, v, H) \mod 2$, there is a u, v-path in H with length t.*

Proof. Let $\varepsilon_1 > 0$ be such that the properties in Theorem 2.7 and Corollary 2.5 hold, and let $\varepsilon_2 = 1/10$. Let $d_0 = d_0(\varepsilon, \varepsilon_1, \varepsilon_2)$ be large. Suppose then that the graph G has $d(G) \geq 8d$.

If possible, let $H \subseteq G$ be a copy of $\mathsf{TK}_{d/2}^{(2)}$, and let $\ell = 6$. Then, as $d \geq d_0(\varepsilon, \varepsilon_1, \varepsilon_2)$ is large, for any two distinct vertices u, v in H and any integer $t \in [\ell, \ell \cdot d^{1-\varepsilon}] \subseteq [6, d-8]$ with $t = \pi(u, v, H) \mod 2$, it is easy to see that there is a u, v-path in H with length t.

Assume then that G is $\mathsf{TK}_{d/2}^{(2)}$-free. By Corollary 2.5, G contains a bipartite $(\varepsilon_1, \varepsilon_2 d)$-expander H with $\delta(H) \geq d(G)/8 \geq d$. Let $\ell = \log^7 |H|$. Note that $|H| \geq \delta(H) + 1 \geq d + 1$, and $d \geq d_0 = d_0(\varepsilon, \varepsilon_1, \varepsilon_2)$ is large, so that $|H|/\log^{10} |H| \geq \ell \cdot d^{1-\varepsilon}$. As $n \geq d \geq d_0(\varepsilon, \varepsilon_1, \varepsilon_2)$ is large, by Theorem 2.7, for any two distinct vertices u, v in H and any integer $t \in [\ell, \ell \cdot d^{1-\varepsilon}] \subseteq [\log^7 |H|, |H|/\log^{10} |H|]$ with $t = \pi(u, v, H) \mod 2$, there is a u, v-path in H with length t. □

The following sketch is illustrated in Figure 3. For Theorem 1.4, we have a graph G with $\chi(G) = k$ and wish to find a long interval in the set of odd cycle lengths in G. Letting $d \approx k/30$, we first find a maximal collection H_i, $i \in [s]$, of edge disjoint bipartite graphs and corresponding integers ℓ_i, $i \in [s]$, which satisfy **I**. As $\chi(G) = k \approx 30d$ is large, it will follow from the maximality of this collection and Corollary 5.1 that $\cup_{i \in [s]} H_i$ has high enough chromatic number that it must contain some odd cycle.

Now, say each bipartite H_i has vertex classes A_i and B_i, and consider the auxilliary graph K formed from $\cup_{i \in [s]} H_i$ by including any missing edges between A_i and B_i for each $i \in [s]$. As $\cup_{i \in [s]} H_i$ has an odd cycle, so does K. Consider a shortest odd cycle C in K. Each edge in C, say the edge e between $A_{i(e)}$ and $B_{i(e)}$, can be replaced with a path with any odd length in $[\ell_{i(e)}, \ell_{i(e)} \cdot k^{1-o(1)}]$
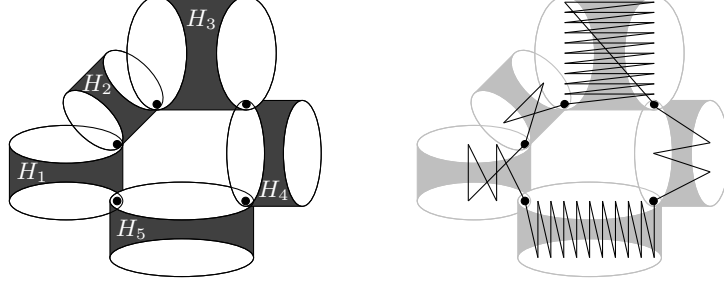
Figure 3: An illustration of the proof of Theorem 1.4. As seen on the left, we find a collection of edge disjoint bipartite expander graphs, here $H_1, \ldots, H_5$, so that H_i intersects with H_{i-1} and H_{i+1} on at least one vertex each (working mod 5 in the indices), and any cycle around the 'cycle of subgraphs' is odd.
We then form different length odd cycles by choosing short paths between the intersecting vertices in some of the expanders H_i, while varying the length of the paths between a vertex disjoint collection of the expanders (here, H_3 and H_5).

by $\mathbf{I}$. Roughly speaking, doing this for each edge in C creates cycles with all possible odd lengths in $[\ell, \ell \cdot k^{1-o(1)}]$, with $\ell = \sum_{e \in E(C)} \ell_{i(e)}$.

The main complication omitted in this sketch is that the paths replacing the edges in C need to be internally vertex disjoint from each other and $V(C)$. To ensure this, we only replace some of the edges with paths of varying length, only doing so for a suitable large collection of the edges on C. If two edges on C are far apart, yet their containing bipartite graphs H_i intersect, then we can find a shorter odd cycle in K than C, a contradiction. Therefore, we choose a subset $E \subseteq E(C)$ of edges which are pairwise nicely separated on C (and which maximises $\sum_{e \in E} \ell_{i(e)}$ subject to this). We replace each edge in E by an odd length path with length from $[\ell_{i(e)}, \ell_{i(e)} k^{1-o(1)}]$, while we (essentially) replace each edge in $E(C) \setminus E$ with some minimal path between the same vertices in the corresponding graph H_i, in order to connect the paths corresponding to E into a cycle.

Recalling Definition 2.6, we use the following simple proposition.

Proposition 5.2. *Given any connected bipartite graph H containing three distinct vertices a_1, a_2 and a_3, we have*

$$\pi(a_1, a_3, H) + \pi(a_2, a_3, H) - \pi(a_1, a_2, H) \in \{0, 2\}. \tag{19}$$

Furthermore, for any (not necessarily distinct) vertices a_1, a_2 and a_3, we have

$$\pi(a_1, a_3, H) + \pi(a_2, a_3, H) = \pi(a_1, a_2, H) \mod 2 \tag{20}$$

Proof. First, suppose that a_1, a_2 and a_3 are distinct. Let the vertex classes of H be A and B, and assume without loss of generality that $a_3 \in A$. If $a_1, a_2 \in A$, then $\pi(a_1, a_3, H) + \pi(a_2, a_3, H) - \pi(a_1, a_2, H) = 2+2-2 = 2$. If $a_1, a_2 \in B$, then $\pi(a_1, a_3, H) + \pi(a_2, a_3, H) - \pi(a_1, a_2, H) = 1+1-2 = 0$. We may then assume that $a_1 \in A$ and $a_2 \in B$, and so $\pi(a_1, a_3, H) + \pi(a_2, a_3, H) - \pi(a_1, a_2, H) = 2 + 1 - 1 = 2$. Therefore, whenever a_1, a_2 and a_3 are distinct, (19) holds, and furthermore (20) holds in this case as well.

Now, note that if $a_1 = a_2 = a$, then (20) holds as $2\pi(a, a_3, H) = 0 = \pi(a, a, H) \mod 2$, while if $a_1 = a_3 = a$, then $0 + \pi(a_2, a, H) = \pi(a, a_2, H) \mod 2$, and, similarly, (20) holds if $a_2 = a_3$. This completes the remaining cases when a_1, a_2 and a_3 are not necessarily distinct. □

33

We are now ready to prove Theorem 1.4, which we do throughout Section 5.1.

5.1 Proof of Theorem 1.4

Let $\varepsilon > 0$. To prove Theorem 1.4, we will show that there is some $k_0 \in \mathbb{N}$ such that the following holds for each $k \geq k_0$. If G is a graph with chromatic number k, then, for some $\ell \in \mathbb{N}$, $\mathcal{C}(G)$ contains every odd integer in $[\ell, \ell \cdot k^{1-\varepsilon}]$.

Let then d_0 be large enough that $d^{1-\varepsilon/4} \geq (30(d+1))^{1-\varepsilon/2}$ holds for each $d \geq d_0$ and the property in Corollary 5.1 holds for d_0 with $\varepsilon_{5.1} = \varepsilon/4$. Let $k_0 = 30d_0$. Suppose $k \geq k_0$ and that the graph G has $\chi(G) = k$. Let $d = \lfloor k/30 \rfloor$.

As outlined at the beginning of this section, using the high chromatic number of G, we first find a minimal 'cycle of subgraphs' that could potentially offer many distinct odd cycle lengths (in Section 5.1.1). We then prove that non-consecutive subgraphs in this cycle are vertex disjoint (in Section 5.1.2). We then choose some of these non-consecutive subgraphs in which to vary the path lengths (in Section 5.1.3). Finally, we take different path lengths in the chosen subgraphs and connect them up into a cycle, getting many odd cycles (in Section 5.1.4).

5.1.1 A minimal 'cycle of subgraphs'

Let $H_1, \ldots, H_s$ be a maximal collection of edge disjoint connected bipartite subgraphs of G such that, for each $i \in [s]$, there is a positive integer ℓ_i for which the following holds.

> **J** For any two distinct vertices u, v in H_i and any integer $t \in [\ell_i, \ell_i \cdot k^{1-\varepsilon/2}]$ with $t = \pi(u, v, H_i)$ mod 2, there is a u, v-path in H_i with length t.

Let $H = \cup_{i \in [s]} H_i$. Let $G' = G \setminus H$. As $d^{1-\varepsilon/4} \geq (30(d+1))^{1-\varepsilon/2} \geq k^{1-\varepsilon/2}$, by the maximality of the collection $H_1, \ldots, H_s$ and Corollary 5.1, G' has no subgraph with average degree at least $8d$. Therefore, every subgraph of G' has a vertex with degree in that subgraph less than $8d$, and hence, as is well known, chromatic number less than $8d$.

We will now show that $\chi(H) \geq 3$. Indeed, suppose to the contrary that $\chi(H) < 3$. Let $c_1 : V(G') \to [8d]$ be a proper colouring of G' and let $c_2 : V(H) \to [2]$ be a proper colouring of H. Then, $c : V(G) \to [8d] \times [2]$ defined by $c(v) = (c_1(v), c_2(v))$ is easily seen to be a proper colouring of G. Therefore, $\chi(G) \leq 16d < k$, a contradiction. Thus, we have that $\chi(H) \geq 3$.

Therefore, we can choose an odd cycle C in H. Say that C has length r' and vertices $a_1 \ldots a_{r'+1}$ where $a_1 = a_{r'+1}$. We now take a certain minimal sequence, where C will demonstrate that such a sequence exists. That is, we take a sequence

$$\mathcal{S} = b_1 F_1 b_2 F_2 b_3 \ldots b_r F_r b_{r+1}$$

such that, setting $r(\mathcal{S}) = r$, we have

K1 $b_1, \ldots, b_r$ are distinct vertices and $b_1 = b_{r+1}$,

K2 for each $i \in [r]$, $F_i \in \{H_1, \ldots, H_s\}$ and $b_i, b_{i+1} \in V(F_i)$,

K3 $\pi(\mathcal{S}) := \sum_{i=1}^{r} \pi(b_i, b_{i+1}, F_i)$ is odd, and

K4 subject to **K1–K3**, $\pi(\mathcal{S}) + r(\mathcal{S})$ is minimised.

Indeed, such a sequence exists as the sequence $\mathcal{S}' = a_1 G_1 a_2 G_2 a_3 \ldots a_{r'} G_{r'} a_{r'+1}$ satisfies **K1**–**K3**, where for each $i \in [r']$, $G_i = H_j$ for the $j \in [s]$ with $a_i a_{i+1} \in E(H_j)$, and we have that $\pi(\mathcal{S}') = \sum_{i=1}^{r'} \pi(a_i, a_{i+1}, G_i) = r'$ is odd.

Note that we must have $r \geq 2$. Indeed, if $r = 1$, then **K1** implies that $\pi(b_1, b_2, F_1) = \pi(b_1, b_1, F_1) = 0$, violating **K3**.

5.1.2 Non-consecutive subgraphs are vertex disjoint

We will now use the minimality of $\mathcal{S}$ (that is, **K4**) to infer two key properties. These are (roughly) that non-consecutive graphs in $F_1, \ldots, F_r$ are vertex disjoint (Claim 5.3) and that each graph is a different graph in $\{H_1, \ldots, H_s\}$ (Claim 5.4).

Claim 5.3. *For each $i, j \in [r]$ with $i \neq j$ and $i \neq j \pm 1 \mod r$, F_i and F_j are vertex disjoint.*

Proof of claim. Suppose, for contradiction, we have some distinct $i, j \in [r]$ with $i \neq j \pm 1 \mod r$ (and thus $r \geq 4$) and that F_i and F_j are not vertex disjoint. We consider separately the case when F_i and F_j do not share some vertex in $\{b_1, \ldots, b_r\}$ (Case I) and when they share some vertex in $\{b_1, \ldots, b_r\}$ (Case II).

Case I. Suppose then that F_i and F_j share some vertex $a \notin \{b_1, \ldots, b_r\}$. Assume, without loss of generality, that $i < j$. Consider the two sequences

$$\mathcal{S}_1 = b_1 F_1 \ldots b_i F_i a F_j b_{j+1} \ldots F_r b_{r+1} \quad \text{and} \quad \mathcal{S}_2 = b_{i+1} F_{i+1} b_{i+2} F_{i+2} \ldots b_j F_j a F_i b_{i+1}.$$

We will show that one of these sequences satisfies **K1**–**K3** in place of $\mathcal{S}$ and contradicts the minimality of $\mathcal{S}$ in **K4**. That each sequence $\mathcal{S}_1$ and $\mathcal{S}_2$ satisfies the corresponding version of **K1** and **K2** follows immediately from **K1** and **K2** and as $a \notin \{b_1, \ldots, b_r\}$ is in both $V(F_i)$ and $V(F_j)$.

Define $\pi(\mathcal{S}_1)$ and $\pi(\mathcal{S}_2)$ similarly to $\pi(\mathcal{S})$ in **K3**, that is, let

$$\pi(\mathcal{S}_1) = \sum_{1 \leq i' < i} \pi(b_{i'}, b_{i'+1}, F_{i'}) + \pi(b_i, a, F_i) + \pi(a, b_{j+1}, F_j) + \sum_{j < i' \leq r} \pi(b_{i'}, b_{i'+1}, F_{i'}),$$

and

$$\pi(\mathcal{S}_2) = \sum_{i < i' < j} \pi(b_{i'}, b_{i'+1}, F_{i'}) + \pi(b_j, a, F_j) + \pi(a, b_{i+1}, F_i).$$

Then,

$$\pi(\mathcal{S}_1) + \pi(\mathcal{S}_2) - \pi(\mathcal{S}) \tag{21}$$
$$= \pi(b_i, a, F_i) + \pi(a, b_{j+1}, F_j) + \pi(b_j, a, F_j) + \pi(a, b_{i+1}, F_i) - \pi(b_i, b_{i+1}, F_i) - \pi(b_i, b_{j+1}, F_j).$$

By **K1**, and as $a \notin \{b_1, \ldots, b_r\}$, a, b_i and b_{i+1} are distinct. Thus, by Proposition 5.2, $\pi(b_i, a, F_i) + \pi(b_{i+1}, a, F_i) - \pi(b_i, b_{i+1}, F_i) \in \{0, 2\}$. Similarly, $\pi(b_j, a, F_j) + \pi(b_{j+1}, a, F_j) - \pi(b_j, b_{j+1}, F_j) \in \{0, 2\}$. Therefore, (21) implies that $\pi(\mathcal{S}_1) + \pi(\mathcal{S}_2) - \pi(\mathcal{S}) \in \{0, 2, 4\}$.

Now, as $i \neq j \pm 1 \mod r$, both $\pi(\mathcal{S}_1)$ and $\pi(\mathcal{S}_2)$ are sums of at least three numbers from $\{1, 2\}$. Therefore, if $\pi(\mathcal{S}_1)$ or $\pi(\mathcal{S}_2)$ is even, they must be at least 4. As $\pi(\mathcal{S})$ is odd, if $\pi(\mathcal{S}_1)$ is even, then $\pi(\mathcal{S}_2) \leq 4 + \pi(\mathcal{S}) - \pi(\mathcal{S}_1) \leq \pi(\mathcal{S})$ is odd. Similarly, if $\pi(\mathcal{S}_2)$ is even, then $\pi(\mathcal{S}_1) \leq \pi(\mathcal{S})$ is odd.

Including the repetition of $b_1 = b_{r+1}$ and b_{i+1} respectively, as $i \neq j \pm 1 \mod r$, the sequence $\mathcal{S}_1$ contains $i + 1 + (r + 1) - j < r + 1$ vertices and the sequence $\mathcal{S}_2$ contains $j - i + 2 < r + 1$

35

vertices. Thus, $r(\mathcal{S}_1), r(\mathcal{S}_2) < r = r(\mathcal{S})$. Therefore, taking the sequence $\mathcal{S}_{i'}$ with $i' \in [2]$ such that $\pi(\mathcal{S}_{i'}) \leq \pi(\mathcal{S})$ is odd (so that the corresponding versions of **K1**–**K3** hold for $\mathcal{S}_{i'}$) contradicts **K4**.

Case II. Suppose then that F_i and F_j share some vertex, say $b_{i'}$, in $\{b_1, \ldots, b_r\}$. We first show that we can assume that $b_j \in V(F_i)$. If $i' = j$, then this already holds. If $i' = i$ then take instead $i = j$ and $j = i = i'$. If $i' \neq i - 1, i, i + 1 \mod r$, then take instead $i = i$ and $j = i'$. If $i' \neq j - 1, j, j + 1 \mod r$, then take instead $j = i$ and $i = i'$. In each case we then get $i \neq j \pm 1 \mod r$, $i \neq j$ and $b_j \in V(F_i)$.

This leaves only the case that $i' \in \{i - 1, i + 1\} \mod r$ and $i' \in \{j - 1, j + 1\} \mod r$. Therefore, switching i and j if necessary, we have that $i + 1 = i' = j - 1 \mod r$. We have that $b_{i'} = b_{i+1}$ is in $V(F_j)$, and now reverse the sequence $\mathcal{S}$ so that, roughly speaking, b_{i+1} and F_i are assigned the same index. That is, consider the sequence $\mathcal{S}'$ created by reversing $\mathcal{S}$, taking that

$$\mathcal{S}' = b_1' F_1' b_2' F_2' b_3' \ldots b_r' F_r' b_{r+1}' = b_{r+1} F_r b_r F_{r-1} b_{r-1} \ldots b_2 F_1 b_1,$$

where $b_{j'}' = b_{r+2-j'}$ for each $j' \in [r + 1]$ and $F_{j'}' = F_{r+1-j'}$ for each $j' \in [r]$. Then, $b_{r+1-i}' = b_{i+1} \in V(F_j) = V(F_{r+1-j}')$. Thus, relabelling $i = r + 1 - j$ and $j = r + 1 - i$ we have $i \neq j \pm 1 \mod r$, $i \neq j$ and $b_j' \in V(F_i')$. Using the corresponding definition for $\mathcal{S}'$ in **K3**, we have $\pi(\mathcal{S}') = \pi(\mathcal{S})$ and $r(\mathcal{S}') = r = r(\mathcal{S})$, so that $\mathcal{S}'$ satisfies its corresponding versions of **K1**–**K4**. Thus, taking $\mathcal{S}'$ instead of $\mathcal{S}$ if necessary, we can always assume that $i \neq j \pm 1 \mod r$, $i \neq j$ and $b_j \in V(F_i)$.

Now we show that we can assume that $j > i$. If $j < i$, then consider the rotated sequence of $\mathcal{S}$ in which the index i is shifted to 1, that is,

$$\mathcal{S}'' = b_1'' F_1'' b_2'' F_2'' b_3'' \ldots b_r'' F_r'' b_{r+1}'' = b_i F_i b_{i+1} \ldots b_r F_r b_{r+1} F_1 b_2 \ldots b_{i-1} F_{i-1} b_i,$$

where $b_{j'}'' = b_{j'+i-1}$ and $F_{j'}'' = b_{j'+i-1}$ for each $j' \in [r]$ with addition modulo r in the indices and $b_{r+1}'' = b_i$. Let $i'' = 1$ and $j'' = r + j - i + 1$, and note that $i'' < j''$ and $i'' \neq j'' \pm 1 \mod r$. Furthermore, $b_{j''}'' = b_j \in V(F_i) = V(F_{i''}'')$. Similarly to $\mathcal{S}'$, $\mathcal{S}''$ satisfies its corresponding versions of **K1**–**K4**, and therefore, taking $\mathcal{S}''$ instead of $\mathcal{S}$ if necessary, we can assume that $i < j$.

Thus, we have that $i < j$, $i \neq j \pm 1 \mod r$ and $b_j \in V(F_i)$. Consider now the sequences

$$\mathcal{S}_3 = b_1 F_1 b_2 F_2 \ldots b_i F_i b_j F_j b_{j+1} \ldots F_r b_{r+1} \quad \text{and} \quad \mathcal{S}_4 = b_j F_i b_{i+1} F_{i+1} b_{i+2} \ldots F_{j-1} b_j.$$

We will show that one of these sequences satisfies **K1**–**K3** in place of $\mathcal{S}$ and contradicts the minimality of $\mathcal{S}$ in **K4**. That each sequence $\mathcal{S}_3$ and $\mathcal{S}_4$ satisfies the corresponding versions of **K1** and **K2** follows immediately from **K1** and **K2** as $b_j \in V(F_j)$.

Writing

$$\pi(\mathcal{S}_3) = \sum_{1 \leq i' < i} \pi(b_{i'}, b_{i'+1}, F_{i'}) + \pi(b_i, b_j, F_i) + \sum_{j \leq i' \leq r} \pi(b_{i'}, b_{i'+1}, F_{i'}),$$

and

$$\pi(\mathcal{S}_4) = \pi(b_j, b_{i+1}, F_i) + \sum_{i < i' < j} \pi(b_{i'}, b_{i'+1}, F_{i'}),$$

we have

$$\pi(\mathcal{S}_3) + \pi(\mathcal{S}_4) - \pi(\mathcal{S}) = \pi(b_i, b_j, F_i) + \pi(b_j, b_{i+1}, F_i) - \pi(b_i, b_{i+1}, F_i). \tag{22}$$

By **K1** and Proposition 5.2, we have $\pi(b_i, b_j, F_i) + \pi(b_j, b_{i+1}, F_i) - \pi(b_i, b_{i+1}, F_i) \in \{0, 2\}$. Thus, by (22), we have $\pi(\mathcal{S}_3) + \pi(\mathcal{S}_4) - \pi(\mathcal{S}) \in \{0, 2\}$.

Now, as $i \neq j \pm 1 \mod r$, both $\pi(\mathcal{S}_3)$ and $\pi(\mathcal{S}_4)$ are sums of at least two numbers from $\{1, 2\}$, and are therefore at least 2. As $\pi(\mathcal{S})$ is odd, if $\pi(\mathcal{S}_3)$ is even then $\pi(\mathcal{S}_4) \leq 2 + \pi(\mathcal{S}) - \pi(\mathcal{S}_3) \leq \pi(\mathcal{S})$ is odd. Similarly, if $\pi(\mathcal{S}_4)$ is even, then $\pi(\mathcal{S}_3) \leq \pi(\mathcal{S})$ is odd.

Including the repetition of $b_1 = b_{r+1}$ and b_j respectively, the sequence $\mathcal{S}_3$ contains $i + (r + 1) - j + 1 < r + 1$ vertices and the sequence $\mathcal{S}_4$ contains $j - i + 1 < r + 1$ vertices. Thus, $r(\mathcal{S}_3), r(\mathcal{S}_4) < r = r(\mathcal{S})$. Therefore, taking the sequence $\mathcal{S}_{i'}$ with $i' \in \{3, 4\}$ such that $\pi(\mathcal{S}_{i'})$ is odd (so that the corresponding versions of **K1**–**K3** hold for $\mathcal{S}_{i'}$) contradicts **K4**, completing the proof of the claim. ∎

Using Claim 5.3, we can show that each graph F_i in the minimal sequence $\mathcal{S}$ is distinct.

Claim 5.4. $F_1, \ldots, F_r$ *are distinct graphs in* $\{H_1, \ldots, H_s\}$*, and thus are pairwise edge disjoint.*

Proof of claim. If $r = 2$ and $F_1 = F_2$, then **K1** implies that $\sum_{i=1}^{r} \pi(b_i, b_{i+1}, F_i) = \pi(b_1, b_2, F_1) + \pi(b_2, b_1, F_2) = 2\pi(b_1, b_2, F_1) = 0 \mod 2$, violating **K3**. Thus, the claim holds for $r = 2$ by **K2**.

Now suppose $r \geq 3$ and that, for distinct $i, j \in [r]$, $F_i = F_j$. Then, by Claim 5.3, it must be that $i = j \pm 1 \mod r$. Say then, without loss of generality, that $i = j - 1 \mod r$. Furthermore, by relabelling the indices in the sequence $\mathcal{S}$ cyclically (as for $\mathcal{S}''$ above), we can assume that $i < r$, and hence $j = i + 1$, so that $F_i = F_{i+1}$. Consider the sequence $\mathcal{S}'$ formed from $\mathcal{S}$ by replacing $F_i b_{i+1} F_{i+1}$ by F_i in $\mathcal{S}$, so that

$$\mathcal{S}' = b_1 F_1 b_2 F_2 b_3 \ldots b_{i-1} F_{i-1} b_i F_i b_{i+2} F_{i+2} \ldots b_r F_r b_{r+1},$$

and note that $\mathcal{S}'$ satisfies its corresponding versions of **K1** and **K2**. By Proposition 5.2 and **K1**, as $F_i = F_{i+1}$, we have that $\pi(b_i, b_{i+1}, F_i) + \pi(b_{i+1}, b_{i+2}, F_{i+1}) - \pi(b_i, b_{i+2}, F_i) \in \{0, 2\}$. Then, defining $\pi(\mathcal{S}')$ as in **K3**,

$$\pi(\mathcal{S}') = \pi(\mathcal{S}) - \pi(b_i, b_{i+1}, F_i) - \pi(b_{i+1}, b_{i+2}, F_{i+1}) + \pi(b_i, b_{i+2}, F_i) \in \{\pi(\mathcal{S}), \pi(S) - 2\}.$$

As $\pi(\mathcal{S})$ is odd, $\pi(\mathcal{S}') \leq \pi(\mathcal{S})$ is odd, and therefore, as $r(\mathcal{S}') = r - 1$, this contradicts **K4**. ∎

5.1.3 The right subgraphs in which to vary path lengths

Using the previous claims in this section, we can now choose a large subcollection of graphs F_i in $\mathcal{S}$, varying the paths in which leads to many odd cycles of distinct lengths. Firstly, by Claim 5.4 and relabelling, we can assume that $F_i = H_i$ for each $i \in [r]$. Recall then that we have positive integers ℓ_i, $i \in [r]$, such that **J** holds with $H_i = F_i$.

Now, partition $[r]$ as $I_1 \cup I_2 \cup I_3$ and find distinct vertices $u_1 = u_{r+1}, u_2, \ldots, u_r$ with $u_i, u_{i+1} \in V(F_i)$ for each $i \in [r]$, and paths P_i, $i \in I_2 \cup I_3$, such that the following hold.

L1 $3 \sum_{i \in I_1} (\ell_i + 1) \geq \sum_{i \in [r]} (\ell_i + 1)$.

L2 Any collection of paths containing exactly one u_i, u_{i+1}-path in F_i, for each $i \in I_1$, and $\{P_i\}_{i \in I_2 \cup I_3}$ form a cycle.

L3 The paths P_i, $i \in I_2 \cup I_3$, have total length $\ell_0 \leq \sum_{i \in I_2 \cup I_3} (\ell_i + 1)$, where ℓ_0 is such that $\ell_0 + \sum_{i \in I_1} \pi(u_i, u_{i+1}, F_i)$ is odd.

We find the partition, vertices and paths differently depending on the length of the sequence $\mathcal{S}$. If $r \geq 4$, then we are in **Case 1**, if $r = 2$, then we are in **Case 2**, and if $r = 3$ we are in **Case 3**.

Case 1. Suppose first that $r \geq 4$. Partition $[r] = I_1 \cup I_2 \cup I_3$ so that, for each $j \in [3]$, there is no solution to $x = y + 1 \mod r$ with $x, y \in I_j$. Thus, by Claim 5.3, for every $j \in [3]$, graphs in $\{F_i : i \in I_j\}$ are pairwise vertex disjoint. By averaging, there exists some $j \in [3]$ such that $\sum_{i \in I_j} (\ell_i + 1) \geq \sum_{i \in [r]} (\ell_i + 1)/3$. By relabelling, we may assume that $j = 1$, and hence **L1** holds.

Now, for each $i \in I_2$, find a shortest path P_i between $V(F_{i-1})$ and $V(F_{i+1})$ in F_i, and label vertices so that this is a u_i, u_{i+1}-path with $u_i \in V(F_{i-1})$ and $u_{i+1} \in V(F_{i+1})$. Note that, by minimality of P_i, all internal vertices of P_i lie in $V(F_i) \setminus (V(F_{i-1}) \cup V(F_{i+1}))$. Furthermore, as $r \geq 4$, by Claim 5.3, we have that $u_i \neq u_{i+1}$.

For each $i \in I_3$, if $i-1, i+1 \in I_2$, let P_i be a shortest u_i, u_{i+1}-path in F_i. If $i-1 \in I_2$ and $i+1 \notin I_2$, let P_i be a shortest path from u_i to $V(F_{i+1})$ in F_i and label its endpoint in $V(F_{i+1})$ by u_{i+1}. If $i-1 \notin I_2$ and $i+1 \in I_2$, let P_i be a shortest path from $V(F_{i-1})$ to u_{i+1} in F_i and label its endpoint in $V(F_{i-1})$ by u_i. If $i-1, i+1 \notin I_2$, let P_i be a shortest path between $V(F_{i-1})$ and $V(F_{i+1})$ in F_i, and label vertices so that this is a u_i, u_{i+1}-path with $u_i \in V(F_{i-1})$ and $u_{i+1} \in V(F_{i+1})$.

Note that we have chosen a vertex u_i for each $i \in [r]$ with i or $i-1$ in $I_2 \cup I_3$, and, for each $i \in I_2 \cup I_3$, a path P_i. As there is no solution to $x = y + 1 \mod r$ in I_1, $\{i, i+1 : i \in I_2 \cup I_3\} = [r]$. Thus, we have in fact chosen vertices u_i for all $i \in [r]$. To see that vertices u_i, $i \in [r]$, are distinct, note that, for each $i \in [r]$, we chose $u_i \in V(F_{i-1}) \cap V(F_i)$. Therefore, if there is some $i \neq j$ with $u_i = u_j$, then as $u_j \in V(F_{i-1}) \cap V(F_i)$ and $r \geq 4$, by Claim 5.3, we must have $j = i - 1 \mod r$. However, similarly, as $u_i \in V(F_{j-1}) \cap V(F_j)$, we must have $i = j - 1 \mod r$, a contradiction.

Let $u_{r+1} = u_1$. For each $i \in I_2 \cup I_3$, P_i was a shortest path between two vertices in F_i, and thus by **J**, has length at most $\ell_i + 1$. Observe crucially that, by our careful selection of shortest paths, the paths in $\{P_i : i \in I_2 \cup I_3\}$ are pairwise internally disjoint, and furthermore internally disjoint from any u_i, u_{i+1}-path in F_i for any $i \in I_1$. In particular, this, together with the graphs F_i, $i \in I_1$, being pairwise vertex disjoint, implies that **L2** holds. It is left to show that **L3** holds. Let ℓ_0 be the total length of the paths P_i, $i \in I_2 \cup I_3$. As, for each $i \in I_2 \cup I_3$, P_i has length at most $\ell_i + 1$, we have $\ell_0 \leq \sum_{i \in I_2 \cup I_3} (\ell_i + 1)$, as required in **L3**.

For the rest of **L3**, we need to show that $\ell_0 + \sum_{i \in I_1} \pi(u_i, u_{i+1}, F_i)$ is odd. Now, as $r \geq 4$, we have $\pi(\mathcal{S}) \geq 5$ by **K3**, and hence $\pi(\mathcal{S}) + r(\mathcal{S}) \geq 9$. Now, suppose for some $i, j \in [r]$ with $j = i - 1 \mod r$, $\pi(u_i, b_i, F_i) + \pi(u_i, b_i, F_j)$ is odd. Then, $u_i \neq b_i$, and therefore $\pi(u_i, b_i, F_i) + \pi(u_i, b_i, F_j) = 3$. Thus, the sequence $\mathcal{S}' = u_i F_i b_i F_j u_i$ has $\pi(\mathcal{S}') + r(\mathcal{S}') = 5$ and satisfies its corresponding version of **K1**–**K3**, contradicting **K4**. Therefore, working mod r in the indices, for each $i \in [r]$, we have

$$\pi(u_i, b_i, F_i) + \pi(u_i, b_i, F_{i-1}) = 0 \mod 2. \tag{23}$$

Futhermore, for each $i \in [r]$, we have by two applications of (the second part of) Proposition 5.2, that

$$\pi(u_i, u_{i+1}, F_i) = \pi(u_i, b_i, F_i) + \pi(b_i, u_{i+1}, F_i) = \pi(u_i, b_i, F_i) + \pi(b_i, b_{i+1}, F_i) + \pi(b_{i+1}, u_{i+1}, F_i) \mod 2.$$

Thus, working mod r in the indices, we have

$$\sum_{i\in[r]}\pi(u_i,u_{i+1},F_i) = \sum_{i\in[r]}\pi(u_i,b_i,F_i) + \sum_{i\in[r]}\pi(b_i,b_{i+1},F_i) + \sum_{i\in[r]}\pi(b_{i+1},u_{i+1},F_i) \mod 2$$

$$= \sum_{i\in[r]}\pi(b_i,b_{i+1},F_i) + \sum_{i\in[r]}(\pi(u_i,b_i,F_i) + \pi(u_i,b_i,F_{i-1})) \mod 2$$

$$\overset{(23)}{=} \sum_{i\in[r]}\pi(b_i,b_{i+1},F_i) \mod 2.$$

For each $i \in I_2 \cup I_3$, P_i is a u_i,u_{i+1}-path in F_i, and therefore has length equal to $\pi(u_i,u_{i+1},F_i)$ mod 2. Thus, ℓ_0 is equal to $\sum_{i\in I_2\cup I_3}\pi(u_i,u_{i+1},F_i) \mod 2$. Therefore,

$$\ell_0 + \sum_{i\in I_1}\pi(u_i,u_{i+1},F_i) = \sum_{i\in[r]}\pi(u_i,u_{i+1},F_i) = \sum_{i\in[r]}\pi(b_i,b_{i+1},F_i) \mod 2.$$

Combined with **K3**, we have that $\ell_0 + \sum_{i\in I_1}\pi(u_i,u_{i+1},F_i)$ is odd, completing the proof of **L3**.

Case 2. Suppose then that $r = 2$. Assume, by relabelling that $\ell_1 \geq \ell_2$ and let $I_1 = \{1\}$, $I_2 = \{2\}$ and $I_3 = \varnothing$. Note that **L1** holds. By **K3**, we have that $\pi(b_1,b_2,F_1) + \pi(b_1,b_2,F_2)$ is odd. Find distinct vertices $u_1,u_2 \in V(F_1) \cap V(F_2)$ and a u_1,u_2-path P_2 in F_2 so that $\ell(P_2) + \pi(u_1,u_2,F_1)$ is odd, and subject to this P_2 has the shortest possible length. Note that this is possible as taking $u_1 = b_1$, $u_2 = b_2$ and letting P_2 be any u_1,u_2-path in F_2 (which exists due to **J**) satisfies these conditions, and, furthermore, by **J**, P_2 has length at most $\ell_2 + 1$. Therefore, **L3** holds for P_2.

Now, suppose P_2 has some internal vertex u in F_1. Then, note $u \in V(P_2) \subset V(F_2)$, and split P_2 as a u_1,u-path Q_1 and a u,u_2-path Q_2. By Proposition 5.2, we have, working mod 2, that

$$1 = \ell(P_2) + \pi(u_1,u_2,F_1) = \ell(Q_1) + \pi(u_1,u,F_1) + \ell(Q_2) + \pi(u,u_2,F_1) \mod 2.$$

Therefore, one of $\ell(Q_1) + \pi(u_1,u,F_1)$ or $\ell(Q_2) + \pi(u,u_2,F_1)$ must be odd, contradicting the minimality of P_2. Therefore, P_2 has no internal vertices in F_1, and hence **L2** holds.

Case 3. Suppose finally that $r = 3$. Assume, by relabelling, that $\ell_1 = \max_{i\in[r]}\ell_i$ and let $I_1 = \{1\}$, $I_2 = \{2\}$ and $I_3 = \{3\}$. Note that **L1** holds.

Suppose, for contradiction, that $F_2 \cup F_3$ is not bipartite, and let A, B and A', B' be the bipartitions of F_2 and F_3, labelled so that $b_3 \in A \cap A'$. As $A \cup A', B \cup B'$ is not a bipartition of $F_2 \cup F_3$, there must be some vertex, u say, in $A \cap B'$ or $A' \cap B$. Then, $\pi(b_3,u,F_2) + \pi(u,b_3,F_3) = 3$. Thus, $\mathcal{S}' = b_3 F_2 u F_3 b_3$ satisfies the corresponding version of **K1**–**K3** with $r(\mathcal{S}') + \pi(\mathcal{S}') = 5$. By **K1** and **K3**, $\pi(S) \geq 3$, and therefore $\pi(\mathcal{S}) + r(\mathcal{S}) \geq 6$, contradicting **K4**. Therefore, $F_2 \cup F_3$ must be bipartite. Similarly, $F_1 \cup F_2$ and $F_1 \cup F_3$ are bipartite.

As $F_2 \cup F_3$ is bipartite, we have, by Proposition 5.2, that

$$\pi(b_2,b_4,F_2\cup F_3) = \pi(b_2,b_3,F_2\cup F_3) + \pi(b_3,b_4,F_2\cup F_3) = \pi(b_2,b_3,F_2) + \pi(b_3,b_4,F_3) \mod 2.$$

Let Q be a shortest b_2,b_4-path in $F_2 \cup F_3$, so that $\ell(Q) = \pi(b_2,b_4,F_2\cup F_3) \mod 2$. As F_2 and F_3 have diameter at most $\ell_2 + 1$ and $\ell_3 + 1$ respectively by **J**, we have $\ell(Q) \leq \ell_2 + \ell_3 + 2$. As $b_4 = b_1$, Q is a b_1,b_2-path. Using Q and **K3**, find distinct vertices u_1,u_2 with $u_1 \in V(F_1) \cap V(F_3)$ and $u_2 \in V(F_1) \cap V(F_2)$, and a u_1,u_2-path P_2 in $F_2 \cup F_3$ so that $\ell(P_2) + \pi(u_1,u_2,F_1)$ is odd, and

subject to this $\ell(P_2)$ has the shortest possible length. Let P_3 be the path with only the vertex u_2. Therefore, **L3** holds for P_2 and P_3.

Similarly to Case 2, the minimality of P_2 implies that P_2 has no internal vertices in F_1, and hence **L2** holds. Now, combining P_2 with any u_1, u_2-path in F_1 with length equivalent to $\pi(u_1, u_2, F_1)$ mod 2 gives an odd cycle. Thus, as both $F_1 \cup F_2$ and $F_1 \cup F_3$ are bipartite, P_2 must have have some edge in both F_2 and F_3, and hence we can pick u_3 as an arbitrary internal vertex of P_2 in $V(F_2) \cap V(F_3)$. This completes the partition $[3] = I_1 \cup I_2 \cup I_3$, distinct vertices u_1, u_2 and u_3, with $u_1, u_2 \in V(F_1)$, $u_2, u_3 \in V(F_2)$ and $u_3, u_1 \in V(F_3)$, and paths P_2 and P_3 such that **L1**–**L3** hold.

5.1.4 Varying paths in the chosen subgraphs

Thus, in each of the three cases, we have a partition $[r] = I_1 \cup I_2 \cup I_3$, distinct vertices $u_1 = u_{r+1}, u_2, \ldots, u_r$ with $u_i, u_{i+1} \in V(F_i)$ for each $i \in [r]$, and paths P_i, $i \in I_2 \cup I_3$, for which **L1**–**L3** hold. Let

$$\ell = \ell_0 + \sum_{i \in I_1} (\ell_i + 1) \overset{\textbf{L3}}{\leq} \sum_{i \in [r]} (\ell_i + 1) \overset{\textbf{L1}}{\leq} 3 \sum_{i \in I_1} (\ell_i + 1). \tag{24}$$

Now, if ℓ_i', $i \in I_1$, is a collection of integers satisfying $\ell_i' \in [\ell_i, \ell_i \cdot k^{1-\varepsilon/2}]$ and $\ell_i' = \pi(u_i, u_{i+1}, F_i)$ mod 2, for each $i \in I_1$, then, by taking a u_i, u_{i+1}-path in F_i with length ℓ_i' by **J**, for each $i \in I_1$, and combining these paths with the paths P_i, $i \in I_2 \cup I_3$, by **L2** we get a cycle with length $\ell_0 + \sum_{i \in I_1} \ell_i'$.

As, by **L3**, $\ell_0 + \sum_{i \in I_1} \pi(u_i, u_{i+1}, F_i)$ is odd, there are sets of such numbers ℓ_i', $i \in I_1$, with $\ell_0 + \sum_{i \in I_1} \ell_i' = t$ for any odd number t such that

$$\ell_0 + \sum_{i \in I_1} (\ell_i + 1) \leq t \leq \ell_0 + \sum_{i \in I_1} (\lfloor \ell_i \cdot k^{1-\varepsilon/2} \rfloor - 1).$$

Now, we have $\ell = \ell_0 + \sum_{i \in I_1} (\ell_i + 1)$, so, as $k^{\varepsilon/2} \geq d_0^{\varepsilon/2} \geq 8$, we have

$$\ell_0 + \sum_{i \in I_1} (\lfloor \ell_i \cdot k^{1-\varepsilon/2} \rfloor - 1) \geq \sum_{i \in I_1} (\ell_i \cdot k^{1-\varepsilon/2} - 2) \geq \sum_{i \in I_1} (8\ell_i \cdot k^{1-\varepsilon} - 2) \geq \sum_{i \in I_1} (3\ell_i + 3) \cdot k^{1-\varepsilon} \overset{(24)}{\geq} \ell \cdot k^{1-\varepsilon}.$$

Thus, for each odd integer $t \in [\ell, \ell \cdot k^{1-\varepsilon}]$, G contains a cycle with length t. This completes the proof of Theorem 1.4.

Acknowledgements

References

[1] N. Alon and J. H. Spencer. **The Probabilistic Method**. John Wiley & Sons, 2004.

[2] B. Bollobás. Cycles modulo k. *Bulletin of the London Mathematical Society*, 9(1):97–98, 1977.

[3] B. Bollobás and A. Thomason. Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs. *European Journal of Combinatorics*, 19(8):883–887, 1998.

[4] P. Erdős. Some recent progress on extremal problems in graph theory. *Congr. Numer*, 14:3–14, 1975.

[5] P. Erdős. Problems and results in graph theory. *The theory and applications of graphs (Kalamazoo, MI, 1980)*, pages 331–341, 1981.

[6] P. Erdős. Some new and old problems on chromatic graphs. *Combinatorics and applications*, pages 118–126, 1984.

[7] P. Erdős. Some old and new problems in various branches of combinatorics. *Discrete Mathematics*, 165:227–231, 1997.

[8] P. Erdős and A. Hajnal. On chromatic number of graphs and set-systems. *Acta Mathematica Hungarica*, 17(1-2):61–99, 1966.

[9] A. Gyárfás. Graphs with k odd cycle lengths. *Discrete Mathematics*, 103(1):41–48, 1992.

[10] A. Gyárfás, J. Komlós, and E. Szemerédi. On the distribution of cycle lengths in graphs. *Journal of Graph Theory*, 8(4):441–462, 1984.

[11] J. Kim, H. Liu, M. Sharifzadeh, and K. Staden. Proof of Komlós's conjecture on Hamiltonian subsets. *Proceedings of the London Mathematical Society*, 115(5):974–1013, 2017.

[12] J. Komlós and E. Szemerédi. Topological cliques in graphs. *Combinatorics, Probability and Computing*, 3(2):247–256, 1994.

[13] J. Komlós and E. Szemerédi. Topological cliques in graphs II. *Combinatorics, Probability and Computing*, 5(1):79–90, 1996.

[14] K. Kuratowski. Sur le probleme des courbes gauches en topologie. *Fund. Math.*, 16:271–283, 1930.

[15] H. Liu and R. Montgomery. A proof of Mader's conjecture on large clique subdivisions in C_4-free graphs. *Journal of the London Mathematical Society*, 95(1):203–222, 2017.

[16] W. Mader. Homomorphieeigenschaften und mittlere Kantendichte von Graphen. *Mathematische Annalen*, 174(4):265–268, 1967.

[17] W. Mader. Hinreichende Bedingungen für die Existenz von Teilgraphen, die zu einem vollständigen Graphen homöomorph sind. *Mathematische Nachrichten*, 53(1-6):145–150, 1972.

[18] B. Sudakov and J. Verstraëte. Cycle lengths in sparse graphs. *Combinatorica*, 28:357–372, 2008.

[19] B. Sudakov and J. Verstraëte. Cycles in graphs with large independence ratio. *Journal of Combinatorics*, 2(1):83–102, 2011.

[20] C. Thomassen. Subdivisions of graphs with large minimum degree. *Journal of Graph Theory*, 8(1):23–28, 1984.

[21] C. Thomassen. Problems 20 and 21. In *Graphs, Hypergraphs and Applications*. H. Sachs, Ed.: 217. Teubner. Leipzig., 1985.

[22] C. Thomassen. Configurations in graphs of large minimum degree, connectivity, or chromatic number. In *Proceedings of the third international conference on Combinatorial mathematics*, pages 402–412, 1989.

[23] J. Verstraëte. Unavoidable cycle lengths in graphs. *J. Graph Theory*, 49:151–167, 2005.

[24] J. Verstraëte. Extremal problems for cycles in graphs. In *Recent trends in combinatorics*, pages 83–116. Springer, 2016.