arXiv:2108.08814v5 [math.CO] 1 Sep 2023

Rainbow subdivisions of cliques

Tao Jiang * Shoham Letzter† Abhishek Methuku‡ Liana Yepremyan§

Abstract

We show that for every integer $m \geq 2$ and large n, every properly edge-coloured graph on n vertices with at least $n(\log n)^{53}$ edges contains a rainbow subdivision of K_m. This is sharp up to a polylogarithmic factor. Our proof method exploits the connection between the mixing time of random walks and expansion in graphs.

1 Introduction

The *Turán number* of a graph H, denoted by $\mathrm{ex}(n, H)$, is the maximum possible number of edges in an n-vertex graph that does not contain a copy of H. In this paper we study a rainbow variant of Turán numbers, introduced by Keevash, Mubayi, Sudakov and Verstraëte [7]. A *proper edge-colouring* of a graph is an assignment of colours to its edges so that edges that share a vertex have distinct colours. A *rainbow* subgraph of an edge-coloured graph is a subgraph whose edges have distinct colours. The *rainbow Turán number* of a graph H, denoted by $\mathrm{ex}^*(n, H)$, is the maximum possible number of edges in a properly edge-coloured graph on n vertices with no rainbow copy of H. One can define $\mathrm{ex}(n, \mathcal{H})$ and $\mathrm{ex}^*(n, \mathcal{H})$ analogously for a family of graphs $\mathcal{H}$.

It was shown in [7] that $\mathrm{ex}^*(n, H) = (1+o(1))\,\mathrm{ex}(n, H)$ for non-bipartite H. Perhaps unsurprisingly, little is known about rainbow Turán numbers of bipartite graphs. The authors of [7] raised two problems concerning rainbow Turán numbers of even cycles, one concerning an even cycle of fixed length $2k$ and the other concerning the family $\mathcal{C}$ of all cycles. For all $k \geq 2$, they showed that $\mathrm{ex}^*(n, C_{2k}) = \Omega(n^{1+1/k})$ and conjectured that $\mathrm{ex}^*(n, C_{2k}) = \Theta(n^{1+1/k})$. The authors of [7] verified the conjecture for $k \in \{2, 3\}$. Following further progress on the conjecture by Das, Lee and Sudakov [3], Janzer [4] recently resolved the conjecture.

*Department of Mathematics, Miami University, Oxford, OH 45056, USA. Email: jiangt@miamioh.edu. Research supported by National Science Foundation grant DMS-1855542.

†Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: s.letzter@ucl.ac.uk. Research supported by the Royal Society.

‡School of Mathematics, University of Birmingham, Birmingham, B15 2TT, UK. Email: abhishekmethuku@gmail.com. Research supported by the EPSRC, grant no. EP/S00100X/1 (A. Methuku).

§Deparment of Mathematics, Emory University, Atlanta, US. Email: lyeprem@emory.edu. Research supported by Marie Sklodowska Curie Global Fellowship, H2020-MSCA-IF-2018:846304.

Regarding the rainbow Turán number of the family $\mathcal{C}$ of all cycles, Keevash, Mubayi, Sudakov and Verstraëte [7] showed that $\mathrm{ex}^*(n,\mathcal{C}) = \Omega(n\log n)$, by considering a naturally defined proper edge-colouring of the hypercube Q_k, where $k = \lfloor \log n \rfloor$ (colour the edge uv by colour i if the u and v differ in the coordinate i; it is easy to see that no cycle can be rainbow in this colouring). They also showed that $\mathrm{ex}^*(n,\mathcal{C}) = O(n^{4/3})$ and asked if $\mathrm{ex}^*(n,\mathcal{C}) = O(n^{1+o(1)})$ and furthermore, if $\mathrm{ex}^*(n,\mathcal{C}) = O(n\log n)$. Das, Lee and Sudakov [3] answered the first question affirmatively, by showing that $\mathrm{ex}^*(n,\mathcal{C}) \le ne^{(\log n)^{\frac{1}{2}+o(1)}}$. In 2022, Janzer [4] improved this bound by establishing that $\mathrm{ex}^*(n,\mathcal{C}) = O\big(n(\log n)^4\big)$, which is tight up to a polylogarithmic factor. Recently, Jiang, Methuku and Yepremyan [6] proved the following generalisation of Das, Lee and Sudakov [3] on $\mathrm{ex}^*(n,\mathcal{C})$.

Theorem 1.1 (Jiang, Methuku, Yepremyan [6])**.** *For every integer $m \ge 2$ there exists a constant $c > 0$ such that for every integer $n \ge m$ the following holds. If G is a properly edge-coloured graph on n vertices with at least $ne^{c\sqrt{\log n}}$ edges, then G contains a rainbow subdivision of K_m, where each edge is subdivided at most $1300\log^2 n$ times.*

The method used in [6] utilises robust expanders in the coloured setting together with a density increment argument, inspired in part by the method introduced by Sudakov and Tomon [16].

In this paper, we lower the $e^{O(\sqrt{\log n})}$ error term in Theorem 1.1 to a polylogarithmic term, which in conjunction with the above-mentioned $\Omega(n\log n)$ lower bound on $\mathrm{ex}^*(n,\mathcal{C})$ determines the rainbow Turán number of the family of K_m-subdivisions up to a polylogarithmic factor.

Theorem 1.2. *Fix an integer $m \ge 2$ and let n be sufficiently large. Suppose that G is a properly edge-coloured graph on n vertices with at least $n(\log n)^{53}$ edges. Then G contains a rainbow subdivision of K_m, where each edge is subdivided at most $(\log n)^6$ times.*

Theorem 1.2 provides the rainbow analogue of a fundamental (and highly influential) result of Mader [13] stating that for every integer $m \ge 2$, there exists $d = d(m)$ such that every graph with average degree at least d contains a subdivision of K_m. Research on this problem has a long history, see e.g., Mader [14], Komlós and Szemerédi [9, 10], and Bollobás and Thomason [2].

Our proof of Theorem 1.2 exploits the connection between mixing time of random walks and edge expansion. This connection is used in conjunction with counting lemmas developed by Janzer in [4] regarding homomorphisms of cycles in graphs. We also prove a strengthening of Theorem 1.2, regarding 'rooted' rainbow subdivisions of K_m in expanders (see Theorem 6.1). For this stronger version, in addition to the ingredients used for proving Theorem 1.2, we use the framework of [6] and an additional idea used by Letzter in [11] (see Lemma 3.7).

The rest of the paper is organised as follows. In Section 2, we give a short overview of our proofs, namely the proof of Theorem 1.2 and a strengthening of it (Theorem 6.1). In Section 3, we mention various preliminary results, regarding the existence of expanders which are close to being regular and properties of expanders. In Section 4, we state three lemmas due to Janzer [4] and some consequences of these lemmas. Section 5 contains the main new ideas of the paper, exploiting a

connection between the mixing time of a random walk and expansion properties in a graph. In Section 6, we prove Theorem 1.2 and a strengthening of it regarding rooted subdivisions in almost regular expanders. We complete the paper with concluding remarks in Section 7.

Throughout the paper, for convenience, we drop floor and ceiling signs for large numbers, and logarithms are in base 2.

2 Overview of the proofs

Our main idea is to use the connection between the mixing time of random walks, the notion of 'conductance' (see Definition 5.3) and our notion of expansion. It is a well-known and very useful fact that 'large' conductance implies 'small' mixing time (see, e.g., Lovász [12]). Moreover, our notion of expansion implies that our expanders have large conductance. Using these facts we show that if additionally, such expanders are almost regular, then long enough walks are close to being uniformly distributed. We also use two counting lemmas of Janzer from [4]. Below we describe these lemmas and the main ideas in more detail.

In a properly edge-coloured graph, say that a closed walk is *degenerate* if it is either not rainbow or visits a vertex more than once. The first lemma from [4] implies that in a properly edge-coloured graph which is close to being regular, the number of degenerate closed $2k$-walks is significantly smaller than the number of closed $2k$-walks, provided that k is sufficiently large.

Given two vertices x and y, a closed $2k$-walk W is said to *be hosted* by x and y if it starts at x and reaches y after k steps. We call a pair of vertices (x, y) *good* if the number of degenerate closed $2k$-walks hosted by x and y is significantly smaller than the number of closed $2k$-walks hosted by x and y. The second lemma from [4] that we use shows that if a pair (x, y) is good then there are many short pairwise colour-disjoint and internally vertex-disjoint k-paths from x to y.

In fact, we use versions of these lemmas which are applicable to edge-weighted graphs. These weighted versions can be easily deduced for the original unweighted versions.[1] We will later apply these weighted lemmas with a specific edge-weighting, namely where $w(xy) = \frac{1}{\sqrt{d(x)d(y)}}$. This weight was chosen so that the weight of a walk $W = x_0 \dots x_k$ is the probability that a random walk of length k starting at x_0 produces W, times $\sqrt{\frac{d(x_k)}{d(x_0)}}$.

Using results about random walks on graphs, which relate mixing time to expansion, we show that in an expander G on n vertices which is close to being regular, for k suitably large (at least polylogarithmic in n), the numbers of closed $2k$-walks hosted by any two pairs of vertices are within a suitable polylogarithmic factor (in n) of each other. This, combined with the fact that the number of degenerate closed $2k$-walks is small compared to the total number of closed $2k$-walks (due to the first lemma above), implies that almost all pairs of vertices are good. Thus, using Turán's theorem, we find a copy of K_m in the graph formed by good pairs. This, together with the fact that there are

[1] To deduce the weighted version of the first lemma, we actually need a multigraph version of Janzer's original one, whose proof is identical to the original version for simple graphs.

many short colour-disjoint and internally vertex-disjoint rainbow paths between any good pair of vertices (due to the second lemma above) allows us to greedily build the desired rainbow-subdivision of K_m.

We also prove a stronger version of Theorem 1.2, Theorem 6.1, asserting that in an expander G which is close to being regular and whose average degree is large enough, for any set S of m vertices, there exists a rainbow K_m-subdivision with the vertices of S being the branching vertices. The main step in this proof shows that for any two vertices x and y in G there is a short rainbow x, y-path avoiding a prescribed small set C of vertices and colours. By iterating this over all pairs of vertices in S, we can build the desired rainbow K_m-subdivision.

To show that there is a short rainbow x, y-path in G, we first apply tools due to Jiang, Methuku and Yepremyan [6] and Letzter [11] to show that there is a set of vertices U of size $\Omega(n)$ such that for each $v \in U$ there is such a short rainbow x, v-path $P(v)$ and a short rainbow y, v-path $Q(v)$, both of which avoid C, such that no colour is used on too many of these paths $P(v)$ and $Q(v)$. It easily follows that for almost all pairs (u, v) with $u, v \in U$, the paths $P(u)$ and $Q(v)$ are colour-disjoint. This, combined with the fact that most pairs in U are good (in the sense mentioned earlier), implies that there exists at least one good pair (u, v) for which $P(u)$ and $Q(v)$ are colour-disjoint. This allows us to find a suitable short rainbow u, v-path L such that $P(u) \cup L \cup Q(v)$ is a rainbow x, y-walk which contains the desired rainbow x, y-path.

3 Preliminaries

Let G be a graph. We denote by $d(G)$ the average degree of G. For a subset $S \subseteq V(G)$, let $e(S) = e(G[S])$, and for subsets $S, T \subseteq V(G)$, let $e(S, T) = e(G[S, T])$. We will use the notions of d-minimality and expanders, defined below, following [6].

Definition 3.1. A graph G is said to be *d-minimal* if $d(G) \geq d$ but $d(H) < d$ for every proper subgraph $H \subseteq G$.

It is easy to see that every graph G contains a $d(G)$-minimal subgraph. The following observation was used in [6]. For completeness, we include its short proof.

Observation 3.2. *If G is d-minimal, then every subset $S \subseteq V(G)$ satisfies $e(S) + e(S, S^c) \geq \frac{d|S|}{2}$. In particular, $\delta(G) \geq \frac{d}{2}$.*

Proof. Suppose otherwise. Then $e(S^c) \geq \frac{d|V|}{2} - \frac{d|S|}{2} \geq \frac{d(|V|-|S|)}{2}$, contradicting d-minimality. □

Definition 3.3. Given $d \geq 1$, $\eta \in (0, 1)$ and $\varepsilon \in (0, \frac{1}{2}]$, an n-vertex graph G is called a (d, η, ε)-*expander* if G is d-minimal, and for every subset $S \subseteq V(G)$ of size at most $(1 - \varepsilon)n$, we have $d(S) \leq (1 - \eta)d$.

Note that, by definition, for $0 < \varepsilon' \leq \varepsilon \leq \frac{1}{2}$ and $0 < \eta \leq \eta' < 1$, every (d, η', ε')-expander is also a (d, η, ε)-expander. Also, if G is a (d, η, ε)-expander then it is a $(d(G), \eta, \varepsilon)$-expander. It will be useful to note the following 'edge-expansion' property of (d, η, ε)-expanders.

Observation 3.4. *Let $n, d \geq 1$, let $\eta \in (0,1)$ and let $\varepsilon \in (0, \frac{1}{2}]$. Suppose that G is a (d, η, ε)-expander on n vertices. Then every $S \subseteq V(G)$ with $|S| \leq (1-\varepsilon)n$ satisfies $e(S, S^c) \geq \frac{\eta d}{2}|S|$.*

Proof. Let $S \subseteq V(G)$ satisfy $|S| \leq (1-\varepsilon)n$. Since G is d-minimal and by Observation 3.2, we have $e(S) + e(S, S^c) \geq \frac{d|S|}{2}$. Since G is a (d, η, ε)-expander, by definition, we also have $e(S) = \frac{d(S)}{2}|S| \leq \frac{(1-\eta)d}{2}|S|$. It follows that $e(S, S^c) \geq \frac{\eta d}{2}|S|$, as claimed. □

Lemma 3.5 (Lemma 2.5 from [6])**.** *Let $n, d \geq 1$, let $\varepsilon \in (0, \frac{1}{2}]$ and let $\eta = \frac{\varepsilon}{2 \log n}$. Suppose that G is a graph on n vertices with average degree d. Then G contains a (d', η, ε)-expander, with $d' \geq \frac{d}{2}$.*

The following lemma from [6] asserts that in a properly edge-coloured expander, one can reach almost every vertex by a short rainbow path starting at a specified vertex.

Lemma 3.6 (Lemma 2.7 from [6])**.** *Let $n, \ell, d, M \geq 1$, let $\eta \in (0,1)$, and let $\varepsilon \in (0, \frac{1}{2}]$. Suppose that $\ell = \frac{4 \log n}{\eta}$ and $d \geq \frac{4\ell + 8M}{\eta}$. Let G be a properly edge-coloured (d, η, ε)-expander on n vertices, let $x \in V(G)$ and let F be a set of vertices and colours of size at most M. Then at least $(1-\varepsilon)n$ vertices can be reached from x by a rainbow path of length at most $\ell + 1$ that avoids the vertices and colours in F.*

We will need a stronger version of the previous lemma, where we require that no colour is used too many times in the short rainbow paths. A similar idea was used in [11] (see Lemma 5) in the context of tight paths.

Lemma 3.7. *Let $n, \ell, d, q, M \geq 1$, let $\eta \in (0,1)$ and let $\varepsilon \in (0, \frac{1}{2}]$. Suppose that $\ell = \frac{4 \log n}{\eta}$ and $d \geq \frac{20q\ell + 8M}{\eta}$. Let G be a properly edge-coloured (d, η, ε)-expander on n vertices, let $x \in V(G)$, and let F be a set of colours and vertices of size at most M. Then there is a set $U \subseteq V(G)$ of size at least $(1-\varepsilon)n$, and a collection $\mathcal{P} = \{P(u) : u \in U\}$ where for each $u \in U$, $P(u)$ is a rainbow path from x to u of length at most $\ell + 1$ that avoids the vertices and colours in F and no colour appears in more than $\frac{n}{q}$ of the paths in $\mathcal{P}$.*

Proof. Let U be a largest set satisfying that for every $u \in U$ there is a rainbow path $P(u)$ from x to u of length at most $\ell + 1$ that avoids F, such that no colour appears in more than $\frac{n}{q}$ of the paths $P(u)$. Say that a colour is *bad* if it appears on exactly $\frac{n}{q}$ of the paths $P(u)$ with $u \in U$, and let C_{bad} be the set of bad colours. Since each path $P(u)$ has length at most $\ell + 1$, we have

$$|C_{\text{bad}}| \leq \frac{n(\ell+1)}{n/q} \leq 2q\ell.$$

Since $d \geq \frac{20q\ell + 8M}{\eta} \geq \frac{4\ell + 8(M + |C_{\text{bad}}|)}{\eta}$, by Lemma 3.6 with $F \cup C_{\text{bad}}$ playing the role of F, there is a set U' with $|U'| \geq (1-\varepsilon)n$ such that for every $v \in U'$, there is a rainbow path $P'(v)$ from x to v of

length at most $\ell+1$ that avoids the colours and vertices in $F \cup C_{\text{bad}}$. If $|U| < (1-\varepsilon)n$, then there is a vertex $v \in U' \setminus U$. The set $U \cup \{v\}$ (along with the paths $P(u)$ for $u \in U$ and $P'(v)$) contradicts the maximality of U. It follows that $|U| \geq (1-\varepsilon)n$, as required. $\square$

We would like to work with expanders that are close to regular. For this, we use the following lemma, which is a slight adaptation of Lemma 3.2 of [15].

Lemma 3.8. *Let $n \geq 2$ and $d \geq 36 \log n$. Let G be a bipartite graph on n vertices with minimum degree at least d. Then there exists a subgraph H of G with average degree at least $\frac{d}{12 \log n}$ and maximum degree at most d.*

Proof. Let $\{A, B\}$ denote a bipartition of G with $|A| \geq |B|$. Let G' be obtained from G by keeping exactly d edges incident to each vertex in A. Then for each $v \in A$ we have $d_{G'}(v) = d$, and hence $e(G') = d|A|$.

Let $m = \lceil \log n \rceil$. For each $i \in [m]$, let $B_i = \{v \in B : 2^{i-1} \leq d_{G'}(v) < 2^i\}$. Denoting the set of isolated vertices in B by B_0, we have $B \setminus B_0 = \cup_{i \in [m]} B_i$. By the pigeonhole principle, there exists an $i \in [m]$ for which $e(G'[A, B_i]) \geq \frac{e(G')}{m} \geq \frac{d|A|}{2 \log n}$. Fix such i and let $t = 2^{i-1}$. Then, by definition, each $v \in B_i$ has degree between t and $2t$ in $G'[A, B_i]$. If $2t \leq d$, then $G'[A, B_i]$ has maximum degree at most d and average degree at least $\frac{2e(G'[A,B_i])}{|A|+|B_i|} \geq \frac{d}{2 \log n}$. So the lemma holds. Hence, we may assume that $2t > d$.

Set $p = \frac{d}{4t}$; then $0 < p < \frac{1}{2}$. Now let $A' \subseteq A$ be chosen by including each vertex in A independently with probability p. For convenience, write $G_i = G'[A, B_i]$ and $G'_i = G'[A', B_i]$. Then

$$\mathbb{E}[e(G'_i)] = p \cdot e(G_i) \geq \frac{pd|A|}{2 \log n}. \tag{1}$$

Now, let $B' = \{v \in B_i : d_{G'_i}(v) \leq d\}$.

For each $v \in B_i$, the degree $d_{G'_i}(v)$ is binomially distributed with expectation $p \cdot d_{G_i}(v)$. Since $t \leq d_{G_i}(v) \leq 2t$, we have $\frac{d}{4} = pt \leq \mathbb{E}[d_{G'_i}(v)] \leq 2pt = \frac{d}{2}$. Therefore, using Chernoff's bound (see, e.g., Appendix A of [1]), we have

$$\begin{aligned}\mathbb{E}[|B_i \setminus B'|] &= \sum_{v \in B_i} \mathbb{P}(v \in B_i \setminus B') = \sum_{v \in B_i} \mathbb{P}\Big[d_{G'_i}(v) \geq d\Big] \leq \sum_{v \in B_i} \mathbb{P}\Big[d_{G'_i}(v) \geq 2\mathbb{E}[d_{G'_i}(v)]\Big] \\ &\leq \sum_{v \in B_i} 2 \cdot \exp\left(-\frac{\mathbb{E}[d_{G'_i}(v)]}{3}\right) \leq \sum_{v \in B_i} 2 \cdot \exp\left(-\frac{d}{12}\right) \leq n \cdot 2e^{-3 \log n} < \frac{1}{n}.\end{aligned}$$

This together with the fact that for any A' and the corresponding B', $e(G'[A', B_i \setminus B']) \leq n|B_i \setminus B'|$, implies

$$\mathbb{E}[e(G'[A', B_i \setminus B'])] \leq n \cdot \mathbb{E}[|B_i \setminus B'|] \leq 1. \tag{2}$$

By (1) and (2),

$$\mathbb{E}[e(G'[A', B']) \geq \frac{pd|A|}{2 \log n} - 1.$$

Note that $\mathbb{E}[|A'|] = p|A|$. Also, as $t|B_i| \le e(G'[A, B_i]) \le d|A|$, we have $|B'| \le |B_i| \le \frac{d}{t}|A| = 4p|A|$. Hence $|A'| + |B'| \le 5p|A|$. Let $d_0 = \frac{d}{12 \log n}$. We have

$$\mathbb{E}\big[e(G'[A', B']) - d_0(|A'| + |B'|)\big] \ge \frac{pd|A|}{2 \log n} - 1 - \frac{5dp|A|}{12 \log n} \ge \frac{pd|A|}{12 \log n} - 1 \ge 0,$$

where the last inequality used $pd|A| = \frac{d^2|A|}{4t} \ge 18^2(\log n)^2 \ge 12 \log n$ which holds since $t \le |A|$ and $d \ge 36 \log n$. Thus, there is a choice of A' for which $e(G'[A', B']) - d_0(|A'| + |B'|) \ge 0$. Taking $H = G'[A', B']$, we have $d(H) \ge d_0 = \frac{d}{12 \log n}$ and $\Delta(H) \le d$, as desired. □

Our final preliminary result combines Lemmas 3.5 and 3.8 to show that every relatively dense graph contains an expander which is close to regular.

Lemma 3.9. *Let $n, d \ge 1$ and let $\varepsilon \in (0, \frac{1}{2}]$, and suppose that $d \ge 10^7(\log n)^3$. Suppose that G is a bipartite graph on n vertices with average degree at least d. Then G has a subgraph H with the following properties.*

1. *H is a (d', η, ε)-expander on n' vertices, where $d' \ge \frac{d}{2500(\log n)^2}$ and $\eta \ge \frac{\varepsilon}{100(\log n')^2}$,*
2. *H has maximum degree at most $2500(\log n')^2 d'$.*

Proof. Let $G_0 = G$. We run the following process, generating graphs G_i for $i \ge 0$. For each graph G_i, we write $d_i = d(G_i)$ and $n_i = |V(G_i)|$.

(a) Let H_i be a subgraph of G_i with average degree at least $\frac{d_i}{24 \log n_i}$ and maximum degree at most d_i. Such a subgraph H_i exists by Lemma 3.8, using the fact that every graph with average degree d contains a subgraph with minimum degree at least $\frac{d}{2}$. To apply the lemma we need to verify that $d_i \ge 72 \log n_i$, which we shall do below.

(b) Write $n_i' = |V(H_i)|$. Let G_{i+1} be a subgraph of H_i which is a $(d_{i+1}, \eta_{i+1}, \varepsilon)$-expander, where $d(G_{i+1}) = d_{i+1} \ge \frac{d(H_i)}{2} \ge \frac{d_i}{48 \log n_i}$ and $\eta_{i+1} = \frac{\varepsilon}{2 \log n_i'}$. Such a subgraph exists by Lemma 3.5, using the observation that if G is a (d, η, ε)-expander then it is a $(d(G), \eta, \varepsilon)$-expander.

Claim 3.10. *If $48 \log n_i < \sqrt{48 \log n_{i-1}}$ for $i \in [t]$ then the graphs $G_1, \ldots, G_t$ can be defined as above and are non-empty.*

Proof. Notice that to prove the statement, it suffices to show that for t as described, $d_i \ge 72 \log n_i$ for $i \in [t-1]$.

We prove this by induction. It is easy to check that the statement is true for $t = 0$ (note that the condition on t holds vacuously here). Indeed, we just need to check that $d_0 \ge 72 \log n_0$, which is the case as $d_0 = d$, $n_0 = n$ and $d \ge 10^7(\log n)^3$.

Now suppose that $48 \log n_i < \sqrt{48 \log n_{i-1}}$ for $i \in [t]$ and that the inductive statement holds for $t' \le t$. This means that the process runs as described for all $i \in [t]$, and it remains to show that the

$(t+1)$-st step can be performed, namely that $d_t \geq 72 \log n_t$. By the assumption on t, the following holds for every $i \in [t]$.

$$48 \log n_i \leq (48 \log n_{i-1})^{1/2} \leq \ldots \leq (48 \log n_0)^{2^{-i}} = (48 \log n)^{2^{-i}}.$$

Since $d_i \geq \frac{d_{i-1}}{48 \log n_{i-1}}$ for $i \in [t]$ (see (b)), it follows that

$$d_t \geq \frac{d_{t-1}}{48 \log n_{t-1}} \geq \ldots \geq \frac{d_0}{48 \log n_{t-1} \cdot \ldots \cdot 48 \log n_0} \geq \frac{d}{(48 \log n)^{2^0 + \ldots + 2^{-(t-1)}}} \geq \frac{d}{(48 \log n)^2}. \quad (3)$$

Using $d \geq 10^7 \log n$ and $n \geq n_t$, we find that $d_t \geq 10^3 \log n > 72 \log n_t$, as required. □

Let ℓ be minimum such that $48 \log n_\ell \geq \sqrt{48 \log n_{\ell-1}}$. We claim that such ℓ exists. If not, then by the previous claim the process can be run forever and $n_i \geq 1$ for all $i \geq 0$, implying that $(48 \log n_i)_{i \geq 0}$ is decreasing, and thus $(n_i)_{i \geq 0}$ is an infinite decreasing sequence of positive integers, a contradiction.

We will show that G_ℓ satisfies the requirements of Lemma 3.9. Indeed, G_ℓ is a $(d_\ell, \eta_\ell, \varepsilon)$-expander. Using the proof of the above claim, inequality (3) holds for $t = \ell$, showing $d_\ell \geq \frac{d}{2500(\log n)^2}$. By choice of G_ℓ, we have $\eta_\ell = \frac{\varepsilon}{2 \log n_{\ell-1}} \geq \frac{\varepsilon}{96(\log n_\ell)^2} \geq \frac{\varepsilon}{100(\log n_\ell)^2}$ (using $48 \log n_\ell \geq \sqrt{48 \log n_{\ell-1}}$). It follows that property 1 of the lemma holds. To see property 2, note that G_ℓ has maximum degree at most $d_{\ell-1}$ and $d_\ell \geq \frac{d_{\ell-1}}{48 \log n_{\ell-1}} \geq \frac{d_{\ell-1}}{(48 \log n_\ell)^2} \geq \frac{d_{\ell-1}}{2500(\log n_\ell)^2}$ (again using $48 \log n_\ell \geq \sqrt{48 \log n_{\ell-1}}$). □

4 Counting rainbow cycles in weighted graphs

In this section we state several lemmas regarding (weighted) counts of homomorphic copies of paths and cycles in weighted graphs G. Three of these are weighted versions of lemmas from Janzer [4], and we will show how to deduce these from the original, unweighted versions.

First, let us introduce some notation. Let G be a graph with a weighting $\omega : E(G) \to \mathbb{R}^{\geq 0}$. We denote by $\omega_{\max}, \omega_{\min}$ the maximum and minimum edge weights; that is, $\omega_{\max} = \max_{e \in E(G)} \omega(e)$ and $\omega_{\min} = \min_{e \in E(G)} \omega(e)$. The *weighted maximum degree of* G, denoted $\Delta_\omega(G)$, is

$$\Delta_\omega(G) := \max_{u \in V(G)} \sum_{v \colon uv \in E(G)} \omega(uv).$$

For any $t \geq 1$, the *weight* of a walk $P := u_0 u_1 \ldots u_t$ is defined as $\omega(P) := \prod_{i=0}^{t-1} \omega(u_i u_{i+1})$, and the *weight* of a closed walk $C := u_0 u_1 \ldots u_t u_0$ is defined as $\prod_{i=0}^{t} \omega(u_i u_{i+1})$, where $u_{t+1} = u_0$.

In a graph G, denote by $\mathrm{Hom}(P_k^{xy})$ the family of walks of length k from x to y. Similarly, let $\mathrm{Hom}(C_{2k}^{xy})$ be the family of closed walks of length $2k$ that start at x and reach y after k steps. Write $\hom(P_k^{xy}) = |\mathrm{Hom}(P_k^{xy})|$ and $\hom(C_{2k}^{xy}) = |\mathrm{Hom}(C_{2k}^{xy})|$. Given a weighting $\omega : E(G) \to \mathbb{R}^{\geq 0}$, we

define the weighted homomorphism counts as follows:

$$\hom_\omega(P_k^{xy}) = \sum_{P \in \mathrm{Hom}(P_k^{xy})} \omega(P)$$
$$\hom_\omega(C_{2k}^{xy}) = \sum_{C \in \mathrm{Hom}(P_k^{xy})} \omega(C)$$

The following relation between $\hom_\omega(P_k^{xy})$ and $\hom_\omega(C_{2k}^{xy})$ is very useful.

$$\hom_\omega(C_{2k}^{xy}) = (\hom_w(P_k^{xy}))^2. \tag{4}$$

We also define $\hom_\omega(C_{2k})$ to be the total weight of the homomorphic copies of C_{2k}, namely

$$\hom_\omega(C_{2k}) = \sum_{x,y \in V(G)} \hom_\omega(C_{2k}^{xy}).$$

Similarly, we define $\hom_\omega(P_k) = \sum_{x,y \in V(G)} \hom_\omega(P_k^{xy})$. In a properly edge-coloured graph G, let $\mathrm{Hom}^*(C_{2k}^{xy})$ be the family of all the closed walks in $\mathrm{Hom}(C_{2k}^{xy})$ that do not form a rainbow cycle of length $2k$. Define $\hom^*(C_{2k}^{xy}) = |\mathrm{Hom}^*(C_{2k}^{xy})|$. If the graph is weighted with a weighting $\omega : E(G) \to \mathbb{R}^{\geq 0}$, let $\hom_\omega^*(C_{2k}^{xy}) = \sum_{C \in \mathrm{Hom}^*(C_{2k}^{xy})} \omega(C)$. Let

$$\mathrm{Hom}^*(C_{2k}) = \bigcup_{x,y \in V(G)} \mathrm{Hom}^*(C_{2k}^{xy}),$$

and write $\hom^*(C_{2k}) = |\mathrm{Hom}^*(C_{2k})|$ and $\hom_\omega^*(C_{2k}) = \sum_{x,y \in V(G)} \hom_\omega^*(C_{2k}^{xy})$.

We will make use of the following two lemmas from a recent paper of Janzer [4].

Lemma 4.1 (multigraph version of Lemma 2.2 from [4])**.** *Let $k \geq 2$ be an integer and let $G = (V, E)$ be a multigraph on n vertices. Let $\sim$ be a symmetric binary relation on V such that for every $u, v \in V$, there are at most t edges vw (counted with multiplicity) for which $u \sim w$. Then the number of homomorphic $2k$-cycles $(x_1, \ldots, x_{2k})$ in G such that $x_i \sim x_j$ for some $i \neq j$ is at most*

$$32k^{3/2}t^{1/2}\Delta(G)^{1/2}n^{\frac{1}{2k}} \hom(C_{2k})^{1-\frac{1}{2k}}.$$

Note that Lemma 2.2 in [4] is phrased for simple graphs, where the condition was that for every $u, v \in V$ there are at most t neighbours w for which $u \sim w$. However, the same proof works for multigraphs with a modified condition as stated above. Similarly, the next lemma is a multigraph version of Lemma 2.1 from [4], where again the same proof works.

Lemma 4.2 (multigraph version of Lemma 2.1 from [4])**.** *Let $k \geq 2$ be an integer and let $G = (V, E)$ be a multigraph on n vertices. Suppose that $\sim$ is a symmetric binary relation on E such that for every $uv \in E$ and $w \in V$, there are at most t edges zw (counted with multiplicity) for which $uv \sim zw$. Then the number of homomorphic $2k$-cycles $(x_1, \ldots, x_{2k})$ in G such that $x_i x_{i+1} \sim x_j x_{j+1}$*

for some $i \neq j$ is at most

$$32k^{3/2}t^{1/2}\Delta(G)^{1/2}n^{\frac{1}{2k}}\hom(C_{2k})^{1-\frac{1}{2k}}.$$

The following lemma is a variant of Lemma 4.2 which is applicable for edge-weighted graphs.

Lemma 4.3. *Let $k \geq 2$ be an integer and let $G = (V, E)$ be a graph on n vertices with a given weighting $\omega : E(G) \to \mathbb{R}^{>0}$. Suppose that $\sim$ is a symmetric binary relation on E such that for every $uv \in E$ and $w \in V$, there are at most t edges wz for which $uv \sim wz$. Let $\hom^*_\omega(C_{2k})$ be the sum of weights of homomorphic $2k$-cycles $(x_1, \ldots, x_{2k})$ in G such that $x_ix_{i+1} \sim x_jx_{j+1}$ for some $i \neq j$. Then*

$$\hom^*_\omega(C_{2k}) \leq 32k^{3/2}t^{1/2} \cdot \Delta_\omega(G)^{1/2} \cdot \omega_{\max}^{1/2} \cdot n^{\frac{1}{2k}} \cdot \hom_\omega(C_{2k})^{1-\frac{1}{2k}}.$$

We prove this lemma using Lemma 4.2.[2]

Proof of Lemma 4.3. We start by proving the lemma in the case where all the weights given by ω are integers. Let G' be the multigraph on $V(G)$ where, for each edge e in G, we add $\omega(e)$ copies of e to G'. Notice that $\Delta(G') = \Delta_\omega(G)$. Let $t' := t \cdot \omega_{\max}$ and notice that for every $uv \in E(G')$ and $w \in V(G')$, there are at most t' edges zw such that $uv \sim zw$. Let $h' = \hom(C_{2k})$ in G'. Observe that $\hom_\omega(C_{2k})$ in G equals h', and that $\hom^*_\omega(C_{2k})$ in G is the number of homomorphic $2k$-cycles $(x_1, \ldots, x_{2k})$ in G' such that $x_ix_{i+1} \sim x_jx_{j+1}$ for some $i \neq j$. Thus, by Lemma 4.2,

$$\begin{aligned}\hom^*_\omega(C_{2k}) &\leq 32k^{3/2}(t')^{1/2}\Delta(G')^{1/2} \cdot n^{\frac{1}{2k}} \cdot (h')^{1-\frac{1}{2k}} \\ &\leq 32k^{3/2}t^{1/2}\Delta_\omega(G)^{1/2} \cdot \omega_{\max}^{1/2} \cdot n^{\frac{1}{2k}} \cdot \hom_\omega(C_{2k})^{1-\frac{1}{2k}}.\end{aligned}$$

Now suppose that all weights given by w are rational. Then there exists a positive integer L such that the weighting $L\omega : E(G) \to \mathbb{R}^{\geq 0}$ (defined as $L\omega(e) = L \cdot \omega(e)$ for every $e \in E(G)$) is integer-valued. By the previous paragraph,

$$\hom^*_{L\omega}(C_{2k}) \leq 32k^{3/2}t^{1/2}\Delta_{L\omega}(G)^{1/2} \cdot ((L\omega)_{\max})^{1/2} \cdot n^{\frac{1}{2k}} \cdot \hom_{L\omega}(C_{2k})^{1-\frac{1}{2k}}.$$

Noting that $\hom^*_{L\omega}(C_{2k}) = L^{2k} \cdot \hom^*_\omega(C_{2k})$, $\hom_{L\omega}(C_{2k}) = L^{2k} \cdot \hom_\omega(C_{2k})$, $\Delta_{L\omega}(G) = L \cdot \Delta_\omega(G)$ and $(L\omega)_{\max} = L \cdot \omega_{\max}$, we get

$$\hom^*_\omega(C_{2k}) \leq 32k^{3/2}t^{1/2}\Delta_\omega(G)^{1/2} \cdot \omega_{\max}^{1/2} \cdot n^{\frac{1}{2k}} \cdot \hom_\omega(C_{2k})^{1-\frac{1}{2k}}.$$

Finally, for a general positive weighting ω, let (ω_m) be a sequence of positive and rational weightings such that $\lim_{m\to\infty} \omega_m(e) = \omega(e)$ for every edge $e \in E$. Then,

$$\hom^*_{\omega_m}(C_{2k}) \leq 32k^{3/2}t^{1/2}\Delta_{\omega_m}(G)^{1/2} \cdot (\omega_m)_{\max}^{1/2} \cdot n^{\frac{1}{2k}} \cdot \hom_{\omega_m}(C_{2k})^{1-\frac{1}{2k}},$$

[2] Alternatively, one can prove this by following the proof of Lemma 2.1 from [4] and adapting it straightforwardly to the weighted version.

for every m. Taking the limits as m tends to infinity, we get

$$\hom^*_\omega(C_{2k}) \le 32k^{3/2}t^{1/2}\Delta_\omega(G)^{1/2}\cdot\omega_{\max}^{1/2}\cdot n^{\frac{1}{2k}}\cdot\hom_\omega(C_{2k})^{1-\frac{1}{2k}}. \qquad \square$$

Similarly, using Lemma 4.1 instead of Lemma 4.2, one can prove the following

Lemma 4.4. *Let $k \ge 2$ be an integer and let $G = (V, E)$ be a graph on n vertices with a given weighting $\omega : E(G) \to \mathbb{R}^{>0}$. Suppose that $\sim$ is a symmetric binary relation on V such that for every $u, v \in V$ there are at most t edges vw for which $u \sim w$. Let $\hom^*_\omega(C_{2k})$ be the sum of weights of homomorphic $2k$-cycles $(x_1, \dots, x_{2k})$ in G such that $x_i \sim x_j$ for some $i \ne j$. Then*

$$\hom^*_\omega(C_{2k}) \le 32k^{3/2}t^{1/2}\cdot\Delta_\omega(G)^{1/2}\cdot\omega_{\max}^{1/2}\cdot n^{\frac{1}{2k}}\cdot\hom_\omega(C_{2k})^{1-\frac{1}{2k}}.$$

Next we show that in an appropriately weighted almost-regular properly edge-coloured graphs the number of degenerate homomorphic copies of $2k$-cycles is a 'small' proportion of all possible copies. Here by degenerate we mean the copies which are either non-rainbow or are not isomorphic copies of C_{2k}.

Lemma 4.5. *Let $n, d, \mu, k, S \ge 1$ and let $\eta \in (0, 1)$. Suppose that $d \ge 2^{16}\mu^3k^3S^2n^{1/k}$. Let G be a properly edge-coloured graph with minimum degree at least $\frac{d}{2}$ and maximum degree at most μd, and define $\omega : E(G) \to \mathbb{R}^{\ge 0}$ by setting $\omega(xy) = 1/\sqrt{d(x)d(y)}$ for every edge xy. Then $\hom^*_\omega(C_{2k}) \le \frac{1}{S}\hom_\omega(C_{2k})$.*

Proof. We first give a lower bound on $\hom_\omega(C_{2k})$. For this, note that

$$\hom_\omega(P_k) = \sum_{W:\,W=x_0\dots x_k}\omega(W) \ge n\cdot\left(\frac{d}{2}\right)^k(\omega_{\min})^k \ge n\cdot\left(\frac{1}{2\mu}\right)^k,$$

where in the last inequality we used that $\omega_{\min} \ge \frac{1}{\Delta(G)} \ge \frac{1}{\mu d}$. Hence,

$$\begin{aligned}\hom_\omega(C_{2k}) &= \sum_{x,y\in V(G)}\left(\hom_\omega(P_k^{xy})\right)^2\\ &\ge \frac{1}{n^2}\left(\sum_{x,y\in V(G)}\hom_\omega(P_k^{xy})\right)^2 = \left(\frac{\hom_\omega(P_k)}{n}\right)^2 \ge \left(\frac{1}{2\mu}\right)^{2k},\end{aligned} \tag{5}$$

where the first inequality follows by convexity.

Let $\sim_e$ be the binary relation on $E(G)$ where $e \sim_e f$ if and only if e and f have the same colour. Because G is properly edge-coloured, for every edge uv and vertex w, there is at most one edge wz for which $uv \sim_e wz$. Let $\sim_v$ be the binary relation defined on $V(G)$ where $u \sim_v w$ if and only if $u = w$. Then, trivially, for every $u, v \in V$ there is at most one edge vw (namely uv) for which $u \sim_v w$. Apply Lemmas 4.3 and 4.4 with $\sim_e$, $\sim_v$ in place of $\sim$, respectively (so t is taken to be 1

in both lemmas), to obtain the desired upper bound on $\hom^*_\omega(C_{2k})$, as follows.

$$\begin{aligned}\hom^*_\omega(C_{2k}) &\le 64k^{3/2}\cdot(\Delta_\omega(G))^{1/2}\cdot\omega_{\max}^{1/2}\cdot n^{\frac{1}{2k}}\cdot\hom_\omega(C_{2k})^{1-\frac{1}{2k}}\\ &\le 64k^{3/2}\cdot(2\mu)^{1/2}\cdot\left(\frac{2}{d}\right)^{1/2}\cdot n^{\frac{1}{2k}}\cdot(2\mu)\cdot\hom_\omega(C_{2k})\\ &= \frac{2^8k^{3/2}\mu^{3/2}n^{\frac{1}{2k}}}{d^{\frac{1}{2}}}\cdot\hom_\omega(C_{2k})\\ &\le \frac{1}{S}\cdot\hom_\omega(C_{2k}).\end{aligned}$$

In the inequalities above we used that $\omega_{\max}\le\frac{1}{\delta(G)}\le\frac{2}{d}$ and $\Delta_\omega(G)\le\Delta(G)\cdot\omega_{\max}\le 2\mu$ and finally, the inequality $\hom_\omega(C_{2k})\ge\left(\frac{1}{2\mu}\right)^{2k}$ proved in (5). □

It would be useful to be able to find many pairwise colour-disjoint and vertex-disjoint paths between a given pair (x,y) of vertices. To that end, we use Lemma 4.7 which will immediately follow from Lemma 4.6 stated below, whose proof uses the arguments of Theorem 3.7 in [4]. For a graph G, and vertices $x,y\in V(G)$, and walks $P,Q\in\mathrm{Hom}(P_k)$, let PQ denote the closed walk in $\mathrm{Hom}(C^{xy}_{2k})$ obtained by concatenating P and the reverse of Q.

Lemma 4.6. *Let $k,s\ge 1$ be integers. Let G be a graph with a given weighting $\omega: E(G)\to\mathbb{R}^{>0}$ and let $x,y\in V(G)$. Let $\mathcal{B}$ be a subfamily of $\mathrm{Hom}(C^{xy}_{2k})$ satisfying $\sum_{F\in\mathcal{B}}\omega(F)<\frac{1}{s^2}\hom_\omega(C^{xy}_{2k})$. Then there exist walks $P_1,\dots,P_s\in\mathrm{Hom}(P^{xy}_k)$ such that $P_iP_j\notin\mathcal{B}$ for every distinct $i,j\in[s]$.*

Proof. Randomly and independently choose s members $P_1,\dots,P_s$ of $\mathrm{Hom}(P^{xy}_k)$, where for each $i\in[s]$ and $P\in\mathrm{Hom}(P^{xy}_k)$, we have $\mathbb{P}(P_i=P)=\frac{\omega(P)}{\hom_w(P^{xy}_k)}$. Note that this is indeed a probability distribution, and $\mathrm{Hom}(C^{xy}_{2k})=\{PQ: P,Q\in\mathrm{Hom}(P^{xy}_k)\}$. For any distinct $i,j\in[s]$ and any $P,Q\in\mathrm{Hom}(P^{xy}_k)$,

$$\mathbb{P}(P_i=P,\,P_j=Q)=\frac{\omega(P)}{\hom_\omega(P^{xy}_k)}\cdot\frac{\omega(Q)}{\hom_\omega(P^{xy}_k)}=\frac{\omega(PQ)}{\hom_w(C^{xy}_{2k})}.$$

Thus, the probability that $P_iP_j\in\mathcal{B}$ is $\frac{\sum_{F\in\mathcal{B}}\omega(F)}{\hom_\omega(C^{xy}_{2k})}<\frac{1}{s^2}$. Hence, by the union bound, with positive probability, $P_iP_j\notin\mathcal{B}$ for every distinct $i,j\in[s]$. □

Lemma 4.7. *Let $k,s\ge 1$ be integers. Let G be a properly edge-coloured graph with a given weighting $\omega: E(G)\to\mathbb{R}^{>0}$ and let $x,y\in V(G)$. Suppose that $\hom^*_\omega(C^{xy}_{2k})<\frac{1}{s^2}\hom_\omega(C^{xy}_{2k})$. Then there are s pairwise colour-disjoint and internally vertex-disjoint rainbow paths of length k from x to y.*

Proof. By Lemma 4.6 applied to $\mathcal{B}=\mathrm{Hom}^*(C^{xy}_{2k})$ there exist walks $P_1,\dots,P_s\in\mathrm{Hom}(P^{xy}_k)$ satisfying $P_iP_j\notin\mathrm{Hom}^*(C^{xy}_k)$ for every distinct $i,j\in[s]$. In other words, P_iP_j is a rainbow copy of C_{2k} for every distinct $i,j\in[s]$. This means that $P_1,\dots,P_s$ are pairwise colour-disjoint and internally vertex-disjoint paths of length k from x to y, as desired. □

5 Counting walks in expanders

In this section we exploit the connection between the mixing time of a random walk on a graph G and expansion properties of G. A lot of the notation and results that we use can be found in [12].

Suppose $G = (V, E)$ is a connected graph where $V = [n]$. Consider a random walk on $V(G)$, where we start at some vertex v_0 and at the i-th step we move from v_i to one of its neighbours, denoted by v_{i+1}, where each neighbour of v_i is chosen as v_{i+1} with probability $\frac{1}{d(v_i)}$. The sequence of vertices $(v_i)_{i\geq 0}$ defines a Markov chain. Let M be the $n \times n$ matrix of transition probabilities of the Markov chain, namely $M_{v,u}$ is the probability of stepping from v to u; so $M_{v,u} = \frac{1}{d(v)}$ if $vu \in E(G)$, and $M_{v,u} = 0$ otherwise. So the probability that a random walk starting at vertex v reaches u in t steps is $(M^t)_{v,u}$.

Definition 5.1. Let G be a graph on the vertex set $[n]$. Let $D = D(G)$ denote the diagonal $n \times n$ matrix where $D_{v,v} = \frac{1}{d(v)}$ for each $v \in [n]$. Let $A = A(G)$ be the adjacency matrix of G, let $M(G) = DA$ and let $N(G) = D^{1/2}AD^{1/2}$. Note that the matrix $N(G)$ is symmetric, so it has n real eigenvalues. Let $\lambda_1(N) \geq \lambda_2(N) \geq \cdots \geq \lambda_n(N)$ denote the eigenvalues of $N := N(G)$.

Lemma 5.2. *Let G be a bipartite graph, with a bipartition $\{X, Y\}$, on the vertex set $[n]$ with m edges and no isolated vertices. Let $M = D(G)A(G)$ and $N = N(G)$. Then for every $v, u \in V(G)$ and integer $k \geq 1$, we have*

$$\left|(M^k)_{v,u} - \frac{d(u)}{2m}\left(1 + (-1)^{k+\mathbb{1}(v\in X)+\mathbb{1}(u\in X)}\right)\right| \leq \sqrt{\frac{d(u)}{d(v)}} \cdot \big(\lambda_2(N)\big)^k.$$

Note that Lemma 5.2 says that when k is even and both v, u are in the same part or when k is odd and v, u are in different parts then $\left|(M^k)_{v,u} - \frac{d(u)}{m}\right| \leq \sqrt{\frac{d(u)}{d(v)}} \cdot \big(\lambda_2(N)\big)^k$. Note that when k is even and v and u are in different parts or when k is odd and v and u are in the same part then $(M^k)_{v,u} = 0$.

Proof. For any vector $w = (w_1, \ldots, w_n)^T$, let $\overline{w}$ be the vector $(w'_1, \ldots, w'_n)^T$ where $w'_i = w_i$ when $i \in X$ and $w'_i = -w_i$ when $i \in Y$. It is easy to check that $N\overline{w} = -\overline{Nw}$. Hence, if w is an eigenvector of N with eigenvalue λ, then $\overline{w}$ is an eigenvector of N with eigenvalue $-\lambda$. It follows that $\lambda_i = -\lambda_{n+1-i}$ for $i \in [n]$. In particular, $|\lambda_i| \leq \lambda_2$ for every $i \in \{2, \ldots, n-1\}$.

One can check that w_1, defined as follows, is a unit eigenvector of N with eigenvalue 1.

$$w_1 = \frac{1}{\sqrt{2m}}\left(\sqrt{d(1)}, \sqrt{d(2)}, \ldots, \sqrt{d(n)}\right)^T.$$

By the Frobenius–Perron theorem, since the entries of N are non-negative and the entries of w_1 are positive, we have $\lambda_1 = 1$. As explained above, it follows that $\overline{w_1}$ is a unit eigenvector of N with eigenvalue $\lambda_n = -1$. Write $w_n = \overline{w_1}$, and let w_i be a unit eigenvector of N with eigenvalue λ_i for $i \in \{2, \ldots, n-1\}$, such that $w_1, \ldots, w_n$ are orthogonal to each other. Note that w_i is an eigenvector

of N^k with eigenvalue $(\lambda_i)^k$, for each $i \in [n]$. Since N^k is a symmetric matrix and $\{w_1, \ldots, w_n\}$ is an orthonormal eigenbasis for N^k, we may write N^k in the spectral form as follows (using $\lambda_1 = 1$ and $\lambda_n = -1$).

$$N^k = \sum_{i=1}^{n} (\lambda_i)^k w_i (w_i)^T = w_1 (w_1)^T + (-1)^k \, \overline{w_1} (\overline{w_1})^T + \sum_{i=2}^{n-1} (\lambda_i)^k w_i (w_i)^T.$$

We also have $D^{1/2} N D^{-1/2} = DA = M$. Therefore,

$$\begin{aligned} M^k &= D^{1/2} N^k D^{-1/2} \\ &= D^{1/2} w_1 (w_1)^T D^{-1/2} + (-1)^k D^{1/2} \, \overline{w_1} (\overline{w_1})^T D^{-1/2} + \sum_{i=2}^{n-1} (\lambda_i)^k D^{1/2} w_i (w_i)^T D^{-1/2} \end{aligned}$$

Let $Q = D^{1/2} w_1 (w_1)^T D^{-1/2} + (-1)^k D^{1/2} \, \overline{w_1} (\overline{w_1})^T D^{-1/2}$. Then

$$M^k = Q + \sum_{i=2}^{n-1} (\lambda_i)^k D^{1/2} w_i (w_i)^T D^{-1/2}.$$

Hence

$$(M^k)_{v,u} = Q_{v,u} + \sum_{i=2}^{n-1} (\lambda_i)^k w_{i,v} w_{i,u} \sqrt{\frac{d(u)}{d(v)}}. \tag{6}$$

Let W be the matrix whose rows are $w_1, \ldots, w_n$. Then $WW^T = I$, implying $W^T W = I$. For each $v \in [n]$, since $(W^T W)_{v,v} = 1$, we have $\sum_{i=1}^{n} |w_{i,v}|^2 = 1$, so $\sum_{i=2}^{n-1} |w_{i,v}|^2 \leq 1$. By the Cauchy-Schwarz inequality, $\sum_{i=2}^{n-1} |w_{i,v} w_{i,u}| \leq \sqrt{\sum_{i=2}^{n-1} |w_{i,v}|^2} \sqrt{\sum_{i=2}^{n-1} |w_{i,u}|^2} \leq 1$. Since $|\lambda_i| \leq \lambda_2$ for every $i \in \{2, \ldots, n-1\}$, the inequality in (6) implies

$$\left| (M^k)_{v,u} - Q_{v,u} \right| \leq \sum_{i=2}^{n-1} |\lambda_i|^k |w_{i,u} w_{i,v}| \cdot \sqrt{\frac{d(u)}{d(v)}} \leq (\lambda_2)^k \sqrt{\frac{d(u)}{d(v)}}.$$

Finally, a straightforward calculation shows that $Q_{v,u} = \frac{d(u)}{2m} \left(1 + (-1)^{k + \mathbb{1}(v \in X) + \mathbb{1}(u \in X)}\right)$ for all $v, u \in [n]$, as desired. □

Definition 5.3. For a graph G with m edges, let $\pi(v) = \frac{d(v)}{2m}$, and for any $S \subseteq V(G)$, let $\pi(S) := \sum_{s \in S} \pi(s)$; observe that $\pi(S) \leq 1$ for every $S \subseteq V(G)$. Define the *conductance* of a set S, denoted by $\Phi(S)$, as

$$\Phi(S) := \frac{e(S, S^c)}{2m \cdot \pi(S) \pi(S^c)},$$

and let the *conductance* of a graph G, denoted by Φ_G, be defined as

$$\Phi_G := \min_{S \subseteq V(G)} \Phi(S).$$

Theorem 5.4 (Theorem 5.3 in [12]). *Let G be a graph and let $\lambda_2 = \lambda_2(N(G))$. Then $\lambda_2 \leq 1 - \frac{\Phi_G^2}{8}$.*

In light of Theorem 5.4 and Lemma 5.6, it will be useful to have a lower bound on Φ_G for a (d, η, ε)-expander G. This is easy to achieve, as can be seen in the following lemma.

Lemma 5.5. *Let $d \geq 1$, $\eta \in (0,1)$, $\varepsilon \in (0, \frac{1}{2}]$. Let G be a (d, η, ε)-expander on n vertices. Then $\Phi_G \geq \frac{\eta}{3}$.*

Proof. Let $S \subseteq V(G)$. Since $\Phi(S) = \Phi(S^c)$, we may assume that $|S| \leq \frac{n}{2} \leq (1-\varepsilon)n$. Observation 3.4 thus implies $e(S, S^c) \geq \frac{1}{2}\eta d|S|$. Let γ be such that $e(S, S^c) = \gamma d|S|$, so $\gamma \geq \frac{\eta}{2}$. By definition of a (d, η, ε)-expander, G is also d-minimal. Hence $e(G[S]) \leq \frac{1}{2}d|S|$. Hence, $\sum_{v \in S} d(v) = 2e(G[S]) + e(S, S^c) \leq d|S| + \gamma d|S|$. Also, observe that $\pi(S^c) \leq 1$. Hence

$$\Phi(S) = \frac{e(S, S^c)}{2e(G)\pi(S)\pi(S^c)} \geq \frac{e(S, S^c)}{\sum_{v \in S} d(v)} \geq \frac{\gamma d|S|}{d|S| + \gamma d|S|} \geq \frac{\gamma}{1+\gamma} \geq \frac{\eta}{\eta + 2} \geq \frac{\eta}{3}.$$

The above inequality thus implies $\Phi_G \geq \frac{\eta}{3}$. □

Recall that given a graph G and two vertices x, y, the quantity $\hom_\omega(P_k^{xy})$ denotes the sum of weights of walks of length k from x to y. The following lemma and its immediate corollary will allow us to compare the values of $\hom_\omega(P_k^{xy})$, where $w(xy) = \frac{1}{\sqrt{d(x)d(y)}}$ for every edge xy, for different pairs of vertices (x, y) in G.

Lemma 5.6. *Let $d \geq 1$, $\eta \in (0,1)$, $\varepsilon \in (0, \frac{1}{2}]$. Let G be a bipartite (d, η, ε)-expander on n vertices with bipartition $\{X, Y\}$. Define a weighting $\omega : E(G) \to \mathbb{R}^{\geq 0}$, such that for each edge xy in G, $\omega(xy) = \frac{1}{\sqrt{d(x)d(y)}}$. The following holds for any two vertices $x, y \in V(G)$ and every integer $k \geq 1$.*

$$\left| \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\sum_{z \in V(G)} \hom_\omega(P_k^{xz})\sqrt{d(z)}} - \frac{d(y)}{2e(G)}\left(1 + (-1)^{k + \mathbb{1}(x \in X) + \mathbb{1}(y \in X)}\right) \right| \leq \sqrt{\frac{d(y)}{d(x)}}\left(1 - \frac{\eta^2}{72}\right)^k.$$

Proof. Let $M = M(G)$, $N = N(G)$ and let λ_2 be the second largest eigenvalue of N. Let x be any vertex in G. Let $\mathbb{W}_k^x$ be a random walk of length k starting at x. For any walk $P = x_0 \ldots x_k$ in G, where $x_0 = x$,

$$\mathbb{P}[\mathbb{W}_k^x = P] = \frac{1}{d(x_0)} \cdot \ldots \cdot \frac{1}{d(x_{k-1})}.$$

Note that

$$\begin{aligned} \omega(P) &= \frac{1}{\sqrt{d(x_0)d(x_1)}} \cdot \ldots \cdot \frac{1}{\sqrt{d(x_{k-1})d(x_k)}} \\ &= \frac{\sqrt{d(x_0)}}{\sqrt{d(x_k)}} \cdot \frac{1}{d(x_0) \cdot \ldots \cdot d(x_{k-1})} = \frac{\sqrt{d(x_0)}}{\sqrt{d(x_k)}} \cdot \mathbb{P}[\mathbb{W}_k^x = P]. \end{aligned}$$

For any vertex y in G, by definition,

$$\mathbb{P}[\mathbb{W}_k^x \text{ ends at } y] = \sum_{P:=xx_1\ldots x_{k-1}y} \mathbb{P}(\mathbb{W}_k^x = P) = \frac{\sqrt{d(y)}}{\sqrt{d(x)}} \sum_{P:P=xx_1\ldots x_{k-1}y} \omega(P) = \hom_\omega(P_k^{xy}) \cdot \sqrt{\frac{d(y)}{d(x)}}.$$

On the other hand, notice that

$$1 = \sum_{z\in V(G)} \mathbb{P}[\mathbb{W}_k^x \textit{ ends at } z] = \sum_{z\in V(G)} \hom_\omega(P_k^{xz}) \cdot \sqrt{\frac{d(z)}{d(x)}}.$$

Recall from our discussion before Definition 5.1, that the probability that a random walk of length k starting at x ends at y is exactly $(M^k)_{x,y}$, thus

$$(M^k)_{x,y} = \mathbb{P}[\mathbb{W}_k^x \textit{ ends at } y] = \frac{\hom_\omega(P_k^{xy})\sqrt{\frac{d(y)}{d(x)}}}{\sum_{z\in V(G)} \hom_\omega(P_k^{xz})\sqrt{\frac{d(z)}{d(x)}}} = \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\sum_{z\in V(G)} \hom_\omega(P_k^{xz})\sqrt{d(z)}}.$$

By Lemma 5.2,

$$\left| \frac{\sqrt{d(y)}\hom_\omega(P_k^{xy})}{\sum_{z\in V(G)} \sqrt{d(z)}\hom_\omega(P_k^{xz})} - \frac{d(y)}{2e(G)}\left(1 + (-1)^{k+\mathbb{1}(x\in X)+\mathbb{1}(y\in X)}\right)\right| \le \sqrt{\frac{d(y)}{d(x)}} \cdot (\lambda_2)^k. \tag{7}$$

Theorem 5.4 gives that $\lambda_2 \le 1 - \frac{\Phi_G^2}{8}$, and by Lemma 5.5 we have $\Phi_G \ge \frac{\eta}{3}$. It follows that $\lambda_2 \le 1 - \frac{\eta^2}{72}$. Combining this inequality with (7), the lemma follows. □

Corollary 5.7. *Let $d \ge 1$, $\eta \in (0,1)$, $\varepsilon \in (0, \frac{1}{2}]$. Let G be a bipartite (d,η,ε)-expander on n vertices with a bipartition $\{X, Y\}$ and with a given weighting $\omega : E(G) \to \mathbb{R}^{>0}$ such that for each edge xy in G, $\omega(xy) = \frac{1}{\sqrt{d(x)d(y)}}$. The following holds for any two vertices $x, y \in X$ and even integer $k \ge 2$.*

$$\left| \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\sum_{z\in X} \hom_\omega(P_k^{xz})\sqrt{d(z)}} - \frac{d(y)}{e(G)} \right| \le \sqrt{n}\left(1 - \frac{\eta^2}{72}\right)^k.$$

Proof. Let $x, y \in X$ be given. Since k is even, for each $z \in Y$, $\mathrm{Hom}(P_k^{xz}) = \emptyset$. The claim follows by applying Lemma 5.6 and the fact that $\sqrt{\frac{d(y)}{d(x)}} \le \sqrt{n}$. □

The next lemma contains the main takeaway from our discussion about random walks in almost regular bipartite expanders. It tells us that for relatively large k, and for the weighting defined by $w(xy) = \frac{1}{\sqrt{d(x)d(y)}}$, the values of $\hom_\omega(C_{2k}^{xy})$ do not differ by much over the range of pairs (x, y) where x and y are in the largest part of the bipartition.

Lemma 5.8. *Let $n, d, \mu, k \ge 1$ and $\eta \in (0,1)$, $\varepsilon \in (0, \frac{1}{2}]$. Suppose that k is an even integer satisfying $k \ge \frac{2^9 \log n}{\eta^2}$. Let G be a bipartite (d,η,ε)-expander with maximum degree at most μd and with a*

given weighting $\omega : E(G) \to \mathbb{R}^{>0}$ *such that for each edge* xy *in* G, $\omega(xy) = \frac{1}{\sqrt{d(x)d(y)}}$. *Let* $\{X, Y\}$ *be the bipartition of* G *and suppose that* $|X| \geq \frac{n}{2}$. *Let*

$$\rho_{\min} = \min\{\hom_\omega(C_{2k}^{xy}) \,:\, x, y \in X\}$$
$$\rho_{\max} = \max\{\hom_\omega(C_{2k}^{xy}) \,:\, x, y \in X\}.$$

Then $\rho_{\max} \leq 2^{10}\mu^2 \cdot \rho_{\min}$.

Proof. By Corollary 5.7, for every $x, y \in X$ we have

$$\left| \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\sum_{z\in X} \hom_\omega(P_k^{xz})\sqrt{d(z)}} - \frac{d(y)}{e(G)} \right| \leq \sqrt{n}\left(1 - \frac{\eta^2}{72}\right)^k$$
$$\leq \sqrt{n} \cdot \exp\left(-\frac{k\eta^2}{72}\right) \leq \sqrt{n} \cdot \exp(-4\log n) \leq \frac{1}{n^3} \leq \frac{d(y)}{2e(G)}.$$

It follows that

$$\frac{d(y)}{2e(G)} \leq \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\sum_{z\in X} \hom_\omega(P_k^{xz})\sqrt{d(z)}} \leq \frac{2d(y)}{e(G)}.$$

Writing the same for $x, w \in X$ we obtain

$$\frac{d(w)}{2e(G)} \leq \frac{\hom_\omega(P_k^{xw})\sqrt{d(w)}}{\sum_{z\in X} \hom_\omega(P_k^{xz})\sqrt{d(z)}} \leq \frac{2d(w)}{e(G)}.$$

Hence, every $x, y, w \in X$ satisfy

$$\frac{\hom_\omega(P_k^{xy})}{\hom_\omega(P_k^{xw})} = \frac{\hom_\omega(P_k^{xy})\sqrt{d(y)}}{\hom_\omega(P_k^{xw})\sqrt{d(w)}} \cdot \sqrt{\frac{d(w)}{d(y)}} \leq \frac{4d(y)}{d(w)} \cdot \sqrt{\frac{d(w)}{d(y)}} = 4\sqrt{\frac{d(y)}{d(w)}} \leq 4\sqrt{2\mu},$$

where the second inequality follows from the fact that the minimum degree of G is at least $\frac{d}{2}$ (by Observation 3.2) and the maximum degree is at most μd.

Observe that $\hom_\omega(P_k^{xy}) = \hom_\omega(P_k^{yx})$ for $x, y \in X$. It follows that, for every $x, y, z, w \in X$,

$$\frac{\hom_\omega(P_k^{xy})}{\hom_\omega(P_k^{zw})} = \frac{\hom_\omega(P_k^{xy})}{\hom_\omega(P_k^{xz})} \cdot \frac{\hom_\omega(P_k^{xz})}{\hom_\omega(P_k^{zw})} \leq 32\mu.$$

Let $x, y, z, w \in X$ satisfy $\rho_{\max} = \hom_\omega(C_{2k}^{xy})$ and $\rho_{\min} = \hom_\omega(C_{2k}^{zw})$. Then

$$\frac{\rho_{\max}}{\rho_{\min}} = \frac{\hom_\omega(C_{2k}^{xy})}{\hom_\omega(C_{2k}^{zw})} = \left(\frac{\hom_\omega(P_k^{xy})}{\hom_\omega(P_k^{zw})}\right)^2 \leq 2^{10}\mu^2,$$

proving the lemma. □

Recall that, by Lemma 4.7, pairs of vertices (x, y) for which $\hom_\omega^*(C_{2k}^{xy})$ is considerably smaller than

$\hom_\omega(C_{2k}^{xy})$ can be used to build many colour-disjoint and internally vertex-disjoint rainbow paths. The following lemma shows that for large enough k and d, almost all pairs of vertices in one of the parts of an almost regular bipartite expander satisfy this property.

Lemma 5.9. *Let $n, d, \mu, k, s, p \geq 1$ and $\eta \in (0,1)$, $\varepsilon \in (0, \frac{1}{2}]$. Suppose that k is even and satisfies $k \geq \frac{2^9 \log n}{\eta^2}$ and that $d \geq 2^{38} k^3 \mu^7 s^4 p^2 n^{1/k}$. Let G be a bipartite (d, η, ε)-expander on n vertices with maximum degree at most μd and with a given weighting $\omega : E(G) \to \mathbb{R}^{>0}$ such that for each edge xy in G, $\omega(xy) = \frac{1}{\sqrt{d(x)d(y)}}$. Let $\{X, Y\}$ be the bipartition of G and suppose that $|X| \geq \frac{n}{2}$. Then for all but at most $\frac{n^2}{p}$ pairs (x, y) with $x, y \in X$ the following holds.*

$$\hom^*_\omega(C_{2k}^{xy}) \leq \frac{1}{s^2} \hom_\omega(C_{2k}^{xy}).$$

Proof. Let $S = 2^{11}\mu^2 s^2 p$, and let $\rho_{\min}$ and $\rho_{\max}$ be defined as in the statement of Lemma 5.8. Then, by the same lemma we have $\rho_{\max} \leq 2^{10}\mu^2 \rho_{\min}$. Let A be the collection of (ordered) pairs (x, y) with $x, y \in X$ that satisfy $\hom^*_\omega(C_{2k}^{xy}) \geq \frac{1}{s^2} \hom_\omega(C_{2k}^{xy})$. Then

$$\hom^*_\omega(C_{2k}) \geq \sum_{(x,y)\in A} \hom^*_\omega(C_{2k}^{xy}) \geq \frac{1}{s^2} \sum_{(x,y)\in A} \hom_\omega(C_{2k}^{xy}) \geq \frac{|A| \cdot \rho_{\min}}{s^2}.$$

Note that $\sum_{x,y\in X} \hom_\omega(C_{2k}^{xy}) = \sum_{x,y\in Y} \hom_\omega(C_{2k}^{xy})$, and so $\hom_\omega(C_{2k}) = 2\sum_{x,y\in X} \hom(C_{2k}^{xy})$. Hence, by Lemma 4.5 (which is applicable since $d \geq 2^{16}\mu^3 k^3 S^2 n^{1/k}$),

$$\hom^*_\omega(C_{2k}) \leq \frac{1}{S} \cdot \hom_\omega(C_{2k}) = \frac{2}{S} \sum_{x,y\in X} \hom_\omega(C_{2k}^{xy}) \leq \frac{2n^2 \cdot \rho_{\max}}{S}.$$

Combining the lower and upper bounds on $\hom^*_\omega(C_{2k})$, we obtain the required inequality, as follows.

$$|A| \leq \frac{2n^2 \cdot s^2 \cdot \rho_{\max}}{S \cdot \rho_{\min}} \leq \frac{n^2 \cdot 2^{11}\mu^2 s^2}{S} = \frac{n^2}{p},$$ □

as desired.

6 Rainbow paths and subdivisions in expanders

We now prove our first main result about rainbow subdivisions.

Proof of Theorem 1.2. Let G' be a bipartite subgraph of G with at least $\frac{e(G)}{2}$ edges. Then $d(G') \geq (\log n)^{53}$. We apply Lemma 3.9 to obtain a subgraph H of G' with the following properties.

1. H is a (d, η, ε)-expander on n' vertices, where $d \geq \frac{d(G')}{2500(\log n)^2} \geq (\log n)^{50}$, $\varepsilon = \frac{1}{2}$ and $\eta = \frac{\varepsilon}{100(\log n')^2} = \frac{1}{200(\log n')^2}$,

2. H has maximum degree at most μd, where $\mu = 2500(\log n')^2$.

Denote the bipartition of H by $\{X, Y\}$, and suppose that $|X| \geq \frac{n'}{2}$. Next, we aim to apply Lemma 5.9 to H. With that in mind, we let k be the smallest even integer which is at least $\frac{2^9 \log n'}{\eta^2}$, so $k = \Theta\big((\log n')^5\big)$. Let $s = 2\binom{m}{2}k$ and $p = 8m$. Let A be the set of pairs (x, y) with $x, y \in X$ that satisfy $\hom^*_\omega(C^{xy}_{2k}) > \frac{1}{s^2}\hom_\omega(C^{xy}_{2k})$. Note that $2^{38}k^3\mu^7 s^4 p^2 (n')^{1/k} = \Theta\big((\log n')^{49}\big) < d$ and so, by Lemma 5.9, we have $|A| \leq \frac{(n')^2}{p} = \frac{(n')^2}{8m} < \frac{1}{m-1}\binom{|X|}{2}$, noting that n' is sufficiently large.

Hence by Turán's theorem, there is a subset $Z \subseteq X$ of size m, such that $(x, y) \notin A$ for every distinct $x, y \in Z$. By Lemma 4.7, for any two vertices x and y in Z, there exist s many pairwise colour-disjoint and internally vertex-disjoint rainbow paths of length k from x to y. By the choice of s, one can find greedily $\binom{m}{2}$ many paths of length k, which are pairwise colour-disjoint and internally vertex-disjoint, and are internally vertex-disjoint from Z, each of which connecting a different pair of vertices in Z. This gives us the desired rainbow subdivision of K_m. □

Given a set of m vertices in a graph G, a *K_m-subdivision rooted at Z* is a subgraph consisting of $\binom{m}{2}$ paths, each joining a different pair of distinct vertices in Z, whose interiors are pairwise vertex-disjoint and disjoint from Z. By slightly adapting the proof of Theorem 1.2, one can show that any bipartite (d, η, ε)-expander G (with suitable parameters) contains a rainbow K_m-subdivision rooted at Z for almost all the m-sets Z in $V(G)$ (not just in X). By using some additional tools, we next show that in a bipartite (d, η, ε)-expander with suitable parameters, in fact, one can find a rainbow K_m-subdivision rooted at Z for every m-set Z in $V(G)$.

Theorem 6.1. *Let $n, L, d, \mu, m \geq 2$, $\eta \in (0, 1)$ and $\varepsilon \in (0, \frac{1}{8}]$, and suppose that $L = \frac{2^{10}\log n}{\eta^2}$ and $d \geq \frac{2^{122} m^8 \mu^7 (\log n)^7}{\eta^{14}}$. Let G be a bipartite (d, η, ε)-expander with maximum degree at most μd, and let Z be a set of m vertices in G. Then there is a rainbow K_m-subdivision, rooted at Z, where every edge is subdivided at most L times.*

Proof. Let $\ell = \frac{4\log n}{\eta}$ and let k be the smallest even integer satisfying $k \geq \frac{2^9 \log n}{\eta^2}$. One can check that $L \geq 2(\ell + 1) + k$.

Claim 6.2. *Let M be any set of colours and vertices such that $|M| \leq 2\binom{m}{2}(L + 1)$. Let x, y be any two vertices in G. There exists a rainbow x, y-path of length at most L in G that avoids M.*

Proof of Claim 6.2. Let $q = 256\ell$. By Lemma 3.7 (using $d \geq \frac{20q\ell + 8[2\binom{m}{2}(L+1)]}{\eta}$), there exists a subset $U_x \subseteq V(G)$ of size at least $(1 - \varepsilon)n$ and a collection of paths $\mathcal{P} = \{P(u) : u \in U_x\}$, where for each $u \in U_x$ the path $P(u)$ is a rainbow path from x to u of length at most $\ell + 1$ that avoids M, and no colour appears in more than $\frac{n}{q}$ of the paths in $\mathcal{P}$. Similarly, there exists a subset $U_y \subseteq V(G)$ of size at least $(1 - \varepsilon)n$ and a collection of paths $\mathcal{Q} = \{Q(u) : u \in U_y\}$ where for each $u \in U_y$ the path $Q(u)$ is a rainbow path from y to u of length at most $\ell + 1$ that avoids M, and no colour appears in more than $\frac{n}{q}$ of the paths in $\mathcal{Q}$. Write $U = U_x \cap U_y$; then $|U| \geq (1 - 2\varepsilon)n \geq \frac{3n}{4}$.

We call an ordered pair (u, v) with $u, v \in U$ *colour-bad* if there is a colour that appears on both paths $P(u)$ and $Q(v)$. We next show that the number of colour-bad pairs in U is small compared to the number of all pairs. Indeed, let H be the auxiliary graph on the vertex set U where uv is an edge whenever at least one of (u, v) and (v, u) is colour-bad. Note that $d_H(u) \leq \frac{2(\ell+1)n}{q}$, for every $u \in U$, since $P(u)$ has length at most $\ell+1$ and any colour on $P(u)$ appears at most $\frac{n}{q}$ times in the collection $\mathcal{Q}$ (and similarly with the roles of P and Q reversed). Thus, $e(H) \leq \frac{(\ell+1)n}{q}|U| \leq \frac{2\ell n}{q}|U| \leq \frac{n^2}{128}$.

Let $s = |M| + 2(\ell + 1) + 1$; then $s \leq 2\binom{m}{2}(L + 1) + 2\ell + 3 \leq 2m^2L$. Denote the bipartition of G by $\{X, Y\}$ and suppose that $|X| \geq \frac{n}{2}$. Call a pair (u, v), with $u, v \in X$, *s-bad* if $\hom^*_\omega(C^{uv}_{2k}) > \frac{1}{s^2}\hom(C^{uv}_{2k})$. Applying Lemma 5.9 with $p = 64$ and verifying the condition on d (namely, that $d \geq 2^{38}k^3\mu^7s^4p^2(n)^{1/k}$), at most $\frac{n^2}{64}$ ordered pairs (u, v), with $u, v \in X$, are s-bad.

We claim that there is a pair (u, v), with $u, v \in U \cap X$, which is neither colour-bad nor s-bad. Indeed, the total number of ordered pairs (u, v) with $u, v \in U \cap X$ where (u, v) is either colour-bad or s-bad is at most $\frac{n^2}{128} + \frac{n^2}{64} \leq \frac{n^2}{32}$. Since $|U| \geq \frac{3n}{4}$ and $|X| \geq \frac{n}{2}$, we have $|X \cap U| \geq \frac{n}{4}$, so the number of ordered pairs (u, v) with $u, v \in U \cap X$ is certainly more than $\frac{n^2}{32}$. Hence, there is a pair (u, v) which is neither colour-bad nor s-bad, as claimed. Since (u, v) is not s-bad, by Lemma 4.7 there are s many pairwise colour-disjoint and internally vertex-disjoint rainbow paths of length k from u to v. By the choice of s, there is at least one such path $T(uv)$ which shares no colours with the paths $P(u)$ and $Q(v)$ and also avoids M. Since (u, v) is not colour-bad, $P(u)$ and $Q(v)$ are colour-disjoint. It follows that $P(u)T(uv)Q(v)$ is a rainbow x, y-walk avoiding M which contains a rainbow x, y-path, as desired. □

Let $(x_1, y_1), \ldots, (x_{\binom{m}{2}}, y_{\binom{m}{2}})$ be an arbitrary ordering of the unordered pairs (x, y) where $x, y \in Z$ and $x \neq y$. We iteratively build paths P_i for $i \in [\binom{m}{2}]$, as follows. Let P_1 be any rainbow x_1, y_1-path of length at most L, which exists by Claim 6.2. In general, suppose $P_1, \ldots, P_i$ have been defined, where $i < \binom{m}{2}$. We let M_i denote the set of vertices and colours used in $\cup_{j=1}^{i} P_j$ and let P_{i+1} be a rainbow x_{i+1}, y_{i+1}-path of length at most L that avoids $M_i \setminus \{x_{i+1}, y_{i+1}\}$. Since $|M_i| \leq 2\binom{m}{2}(L+1)$, by Claim 6.2 such a path P_{i+1} exists. Hence, we are able to find $P_1, \ldots, P_{\binom{m}{2}}$ as described above. Now, $\bigcup_{i=1}^{\binom{m}{2}} P_i$ forms a rainbow K_m-subdivision rooted at Z in which each edge is subdivided at most L times. □

7 Conclusion

In this paper we showed that there is a constant $c \leq 53$ such that for fixed m and sufficiently large n any n-vertex properly edge-coloured graph G with at least $n(\log n)^c$ edges contains a rainbow subdivision of K_m. On the other hand, an immediate lower bound is given by the best known lower bound from [7] on $\mathrm{ex}^*(n, \mathcal{C})$, which is $\Omega(n \log n)$. This shows that our bound is tight up to a polylogarithmic factor. We pose the following question.

Question 7.1. *Fix $m \geq 2$. What is the smallest c such that for all sufficiently large n the following holds: if G is a properly edge-coloured graph on n vertices with at least $\Omega(n(\log n)^c)$ edges, then it contains a rainbow subdivision of K_m? In particular, is $c = 1$?*

For the clarity of presentation, we did not optimise our arguments to obtain the best possible value of c. However, to answer Question 7.1, new ideas will be needed. Note that even the correct order of magnitude of $\mathrm{ex}^*(n, \mathcal{C})$ is still unknown. On another note, as mentioned previously, Janzer [4] proved $\mathrm{ex}^*(n, \mathcal{C}) = O(n(\log n)^4)$, using the counting lemmas on closed walks. It is worth to mention that from our methods, an alternative proof of $\mathrm{ex}^*(n, \mathcal{C}) = O(n(\log n)^5)$ can be obtained by using the basic expansion property of an expander (Lemma 3.7) and a digraph idea used by Letzter [11] regarding the Turán number of the family of tight cycles.

Acknowledgements

We would like to thank the anonymous referee for useful comments, in particular for pointing out an error in the previous version of our proof which we have now rectified. Since the submission of the first version of our paper to arXiv, several interesting developments regarding rainbow Turán numbers of cycles and of clique subdivisions occurred. First, Tomon [17] showed that $\mathrm{ex}^*(n, \mathcal{C}) \leq n(\log n)^{2+o(1)}$, and that every n-vertex graph with at least $n(\log n)^{6+o(1)}$ edges contains a rainbow subdivision of any constant sized clique, using a novel argument about sampling in expanders. The latter bound was subsequently improved to $n(\log n)^{2+o(1)}$ by Wang [18]. The bound on the rainbow Turán number of cycles was later improved to $\mathrm{ex}^*(n, \mathcal{C}) \leq O(n(\log n)^2)$ by Janzer and Sudakov [5] and independently by Kim, Lee, Liu, and Tran [8].

References

[1] N. Alon and J. H. Spencer, *The Probabilistic Method*, John Wiley & Sons, 2016. 6

[2] B. Bollobás and A. Thomason, *Proof of a conjecture of Mader, Erdős and Hajnal on topological complete subgraphs*, Europ. J. Combin. **19** (1998), no. 8, 883–887. 2

[3] S. Das, C. Lee, and B. Sudakov, *Rainbow Turán problem for even cycles*, Europ. J. Combin. **34** (2013), 905–915. 1, 2

[4] O. Janzer, *Rainbow Turán number of even cycles, repeated patterns and blow-ups of cycles*, Israel J. Math. (2022). 1, 2, 3, 8, 9, 10, 12, 21

[5] O. Janzer and B. Sudakov, *On the Turán number of the hypercube*, arXiv:2211.02015 (2022). 21

[6] T. Jiang, A. Methuku, and L. Yepremyan, *Rainbow Turán number of clique subdivisions*, Europ. J. Combin. **110** (2023), no. C. 2, 4, 5

[7] P. Keevash, D. Mubayi, B. Sudakov, and J. Verstraëte, *Rainbow Turán problems*, Combin. Probab. Comput. **16** (2007), 109–126. 1, 2, 20

[8] J. Kim, J. Lee, H. Liu, and T. Tran, *Rainbow cycles in properly edge-colored graphs*, (2022), arXiv:2211.03291. 21

[9] J. Komlós and E. Szemerédi, *Topological cliques in graphs*, Combin. Probab. Comput. **3** (1994), 247–256. 2

[10] ———, *Topological cliques in graphs II*, Combin. Probab. Comput. **5** (1996), 79–90. 2

[11] S. Letzter, *Hypergraphs with no tight cycles*, Proc. Amer. Math. Soc. **151** (2023), 455–462. 2, 4, 5, 21

[12] L. Lovász, *Random walks on graphs: A survey*, Combinatorics, Paul Erdős is eighty **2** (1993), 1–46. 3, 13, 15

[13] W. Mader, *Homomorphieeigenschaften und mittlere Kantendichte von Graphen*, Math. Ann. **174** (1967), 265–268. 2

[14] ———, *Hinreichende Bedingungen für die Existenz von Teilgraphen, die zu einem vollständigen Graphen homöomorph sind*, Math. Nachr. **53** (1972), 145–150. 2

[15] R. Montgomery, A. Pokrovskiy, and B. Sudakov, *C_4-free subgraphs of large average degree*, Israel J. Math. **246** (2021), 55–71. 6

[16] B. Sudakov and I. Tomon, *The extremal number of tight cycles*, Int. Math. Res. Not. **2022** (2022), 9663–9684. 2

[17] I. Tomon, *Robust (rainbow) subdivisions and simplicial cycles*, arXiv:2201.12309 (2022). 21

[18] Y. Wang, *Rainbow clique subdivisions*, (2022), arXiv:2204.08804. 21