

# Explicit two-sided unique-neighbor expanders

Jun-Ting Hsieh<sup>\*</sup>   Theo McKenzie<sup>†</sup>   Sidhanth Mohanty<sup>‡</sup>   Pedro Paredes<sup>§</sup>

January 12, 2024

## Abstract

We study the problem of constructing explicit sparse graphs that exhibit strong vertex expansion. Our main result is the first two-sided construction of imbalanced unique-neighbor expanders, meaning bipartite graphs where small sets contained in both the left and right bipartitions exhibit unique-neighbor expansion, along with algebraic properties relevant to constructing quantum codes.

Our constructions are obtained from instantiations of the *tripartite line product* of a large tripartite spectral expander and a sufficiently good constant-sized unique-neighbor expander, a new graph product we defined that generalizes the *line product* in the work of Alon and Capalbo [AC02] and the *routed product* in the work of Asherov and Dinur [AD23]. To analyze the vertex expansion of graphs arising from the tripartite line product, we develop a sharp characterization of subgraphs that can arise in bipartite spectral expanders, generalizing results of Kahale [Kah95], which may be of independent interest.

By picking appropriate graphs to apply our product to, we give a strongly explicit construction of an infinite family of  $(d_1, d_2)$ -biregular graphs  $(G_n)_{n \geq 1}$  (for large enough  $d_1$  and  $d_2$ ) where all sets  $S$  with fewer than a small constant fraction of vertices have  $\Omega(d_1 \cdot |S|)$  unique-neighbors (assuming  $d_1 \leq d_2$ ). Additionally, we can also guarantee that subsets of vertices of size up to  $\exp(\Omega(\sqrt{\log |V(G_n)|}))$  expand *losslessly*.

---

<sup>\*</sup>Carnegie Mellon University. [juntingh@cs.cmu.edu](mailto:juntingh@cs.cmu.edu). Supported by NSF CAREER Award #2047933.

<sup>†</sup>Stanford University. [theom@stanford.edu](mailto:theom@stanford.edu). Supported by NSF GRFP Grant DGE-1752814 and NSF Grant DMS-2212881.

<sup>‡</sup>MIT. [sidhanth@csail.mit.edu](mailto:sidhanth@csail.mit.edu). Much of this work was conducted while the author was a PhD student at UC Berkeley.

<sup>§</sup>Princeton University. [pparedes@cs.princeton.edu](mailto:pparedes@cs.princeton.edu).

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Our results . . . . .	1
1.2	Context and related work . . . . .	4
1.3	Technical overview . . . . .	6
1.4	Open questions . . . . .	9
1.5	Organization . . . . .	10
<b>2</b>	<b>Preliminaries</b>	<b>10</b>
2.1	Graph expansion . . . . .	11
2.2	Non-backtracking matrix . . . . .	12
<b>3</b>	<b>The line product of graphs</b>	<b>12</b>
<b>4</b>	<b>Algebraic unique-neighbor expanders</b>	<b>14</b>
<b>5</b>	<b>Non-backtracking matrix of subgraphs</b>	<b>16</b>
5.1	Tree extensions . . . . .	17
5.2	Regular tree extensions of subgraphs . . . . .	18
5.3	Folding regular tree extensions . . . . .	19
5.4	Proof of <a href="#">Theorem 5.1</a> . . . . .	20
<b>6</b>	<b>Expansion and density of subgraphs</b>	<b>22</b>
<b>7</b>	<b>Unique-neighbor expanders with lossless small-set expansion</b>	<b>24</b>
7.1	Lossless expansion in high-girth graphs . . . . .	24
7.2	Tripartite line product . . . . .	25
<b>8</b>	<b>One-sided lossless expanders</b>	<b>28</b>
<b>A</b>	<b>Generalization of the Moore bound</b>	<b>34</b>
<b>B</b>	<b>Expansion profile of random graphs</b>	<b>35</b>
<b>C</b>	<b>Coupling Erdős–Rényi and random regular graphs</b>	<b>37</b>
C.1	Random graph extension . . . . .	39
C.2	Proof of <a href="#">Lemma C.3</a> . . . . .	43

# 1 Introduction

A bipartite graph  $G$  is a *one-sided unique-neighbor expander* if every small subset of its left vertices has many *unique-neighbors*, where a unique-neighbor of a set  $S$  is a vertex  $v$  with exactly one edge to  $S$ . Classically, there is a wealth of applications of one-sided unique-neighbor expanders to error-correcting codes [SS96, DSW06, BV09], high-dimensional geometry [BGI<sup>+</sup>08, GLR10, Kar11, GMM22], and routing [ALM96], as well as several explicit constructions [AC02, CRVW02, AD23, Gol23, CRTS23].

A recent work of Lin & Hsieh [LH22] established a connection between quantum error-correcting codes and *two-sided unique-neighbor expanders*, which are graphs where every small subset of both the left and right vertices has many unique-neighbors. In particular, they showed that good *quantum low-density parity check (LDPC) codes* with efficient decoding algorithms can be obtained from two-sided *lossless* expanders satisfying certain algebraic properties, with the additional advantage of being simpler to analyze than earlier constructions of good quantum codes [PK22, LZ22]. Here, lossless expanders are graphs achieving the quantitatively strongest form of unique-neighbor expansion possible. A random biregular graph is a two-sided lossless expander with high probability, but no explicit constructions are known; all explicit constructions of one-sided unique-neighbor expanders are not known to satisfy two-sided expansion.

The main contribution of this work is to give explicit constructions of infinite families of two-sided unique-neighbor expanders. We now delve into our results and provide context.

## 1.1 Our results

We give a formal description of the graphs we would like to construct, motivated by constructing quantum codes, and then describe our contributions.

**Definition 1.1** (Two-sided (algebraic) unique-neighbor expander). We say a  $(d_1, d_2)$ -biregular graph  $Z$  with left and right vertex sets  $L$  and  $R$  respectively is a  $\gamma$ -two-sided unique-neighbor expander if there is a constant  $\delta$  depending on  $\gamma, d_1, d_2$  such that:

1. Every subset  $S \subseteq L$  with  $|S| < \delta|L|$  has at least  $\gamma \cdot d_1|S|$  neighbors in  $R$ .
2. Every subset  $S \subseteq R$  with  $|S| < \delta|R|$  has at least  $\gamma \cdot d_2|S|$  neighbors in  $L$ .

We say  $Z$  is a  $\gamma$ -two-sided algebraic unique-neighbor expander if additionally: there is a group  $\Gamma$  of size  $\Omega(|L| + |R|)$  that acts on  $L$  and  $R$  such that  $gv = v$  iff  $g$  is the identity element of  $\Gamma$ , and  $\{gu, gv\}$  is an edge iff  $\{u, v\}$  is an edge in  $Z$ .

The work of Lin & Hsieh [LH22] proves that for  $\gamma > 1/2$ , the existence of  $\gamma$ -two-sided algebraic unique-neighbor expanders with arbitrary aspect ratio implies the existence of linear-time decodable quantum LDPC codes.

We give an explicit construction of  $\gamma$ -two-sided algebraic unique-neighbor expanders for small constant  $\gamma$ .

**Theorem 1.2** (Two-sided algebraic unique-neighbor expanders.). *For every  $\beta \in (0, 1/2]$  there is a constant  $\gamma > 0$  such that for all large enough  $d_1, d_2$  with  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$ , there is an explicit infinite family of  $(2d_1, 2d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  where every  $Z_n$  is a  $\gamma$ -two-sided algebraic unique-neighbor expander.*

We refer the reader to [Theorem 4.4](#) for a formal statement.

**Remark 1.3.** [Theorem 1.2](#) gives the first construction of two-sided unique-neighbor expanders where the left and right side have unequal sizes.<sup>1</sup> It also gives the only construction besides the one-sided lossless expander constructions of [[CRVW02](#), [Gol23](#), [CRTS23](#)] where the number of unique-neighbors of a set  $S$  can be made arbitrarily larger than  $|S|$ . We give a detailed comparison to prior work in [Table 1](#).

**Remark 1.4.** Our techniques also straightforwardly generalize to constructing families of bounded degree  $k$ -partite unique-neighbor expanders for any distribution  $(\beta_1, \dots, \beta_k)$  of vertices across partitions.

We also give constructions of two-sided unique-neighbor expanders where we can additionally guarantee that small enough sets expand losslessly, at the expense of the algebraic property.

**Theorem 1.5** (Two-sided unique-neighbor expanders with small-set lossless expansion; see [Theorem 7.8](#)). *For every  $\beta \in (0, 1/2]$  and  $\varepsilon > 0$ , there are constants  $\gamma > 0$  and  $K$  such that for all large enough  $d_1, d_2$  with  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$  that are multiples of  $K$ , there is an explicit infinite family of  $(d_1, d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  where:*

1.  $Z_n$  is a  $\gamma$ -two-sided unique-neighbor expander,
2. every  $S \subseteq L(Z_n)$  with  $|S| \leq \exp(\Omega(\sqrt{\log |V(Z_n)|}))$  has  $(1 - \varepsilon) \cdot d_1 \cdot |S|$  unique-neighbors,
3. every  $S \subseteq R(Z_n)$  with  $|S| \leq \exp(\Omega(\sqrt{\log |V(Z_n)|}))$  has  $(1 - \varepsilon) \cdot d_2 \cdot |S|$  unique-neighbors.

At a high level, all of our constructions involve taking a certain product of a large “base graph” with a constant-sized “gadget graph”. In [Theorems 1.2](#) and [1.5](#), the unique-neighbor expansion comes from strong spectral expansion properties of the base graph; see [Section 1.3](#) for an overview. The algebraic property is also inherited from the base graph satisfying the same algebraic property. These two properties can be simultaneously achieved by choosing the base graph as Ramanujan Cayley graphs [[LPS88](#), [Mar88](#), [Mor94](#)].

**Remark 1.6** (Bicycle-free Ramanujan graph construction). In [Theorem 1.5](#), the small-set lossless expansion property comes from the base graph consisting of bipartite spectral expanders with no short bicycles ([Definition 7.3](#)): no bicycles of length- $g$  roughly translates to lossless expansion for sets of size  $\exp(g)$ ; see [Theorem 7.8](#) for a formal statement. We believe that the biregular Ramanujan graph construction of [[BFG<sup>+</sup>15](#)] should have no bicycles of length- $\Omega(\log n)$  and also endow a group action, but we do not prove it in this work. We instead use constructions from the works of [[MOP20](#), [OW20](#)], which have no bicycles of length  $\Omega(\sqrt{\log n})$  but no group action.

**Remark 1.7** (One-sided lossless expanders). As explained in more detail in the technical overview ([Section 1.3](#)), our graph product generalizes the *routed product* defined in [[AC02](#), [AD23](#), [Gol23](#)]. In particular, by instantiating the product with slightly different parameters, we are able to prove one-sided lossless expansion with essentially the same proof as [Theorem 1.5](#), recovering the result of Golowich [[Gol23](#)]. The analysis is carried out in [Section 8](#).

---

<sup>1</sup>Throughout the paper we will assume the left side is larger, i.e.  $d_1 \leq d_2$ .

**New results in spectral graph theory.** In service of proving [Theorem 1.5](#), we prove two results that we believe to be independently interesting in spectral graph theory:

- (i) we give a sharp characterization of what subgraphs can arise in bipartite spectral expanders, generalizing results of Kahale [[Kah95](#)] and Asherov & Dinur [[AD23](#)],
- (ii) we give a refinement to the well-known *irregular Moore bound* of [[AHL02](#)] on the tradeoff between girth and edge density in a graph.

In particular, we show that for any small induced subgraph of a near-Ramanujan biregular graph, the spectral radius of its *non-backtracking matrix* (see [Section 2.2](#)) must be bounded.

**Theorem 1.8** (See [Theorem 5.1](#)). *Let  $\varepsilon \in (0, 0.1)$ , and let  $3 \leq c \leq d$  be integers. Let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph and  $S \subseteq L \cup R$  such that  $|S| \leq d^{-1/\varepsilon} |L \cup R|$ . Then,*

$$\rho(B_{G[S]}) \leq \frac{1}{2} \left( \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} - \sqrt{d-1})^2} \right),$$

where  $\tilde{\lambda} = \max(\lambda_2(A_G), \sqrt{c-1} + \sqrt{d-1}) \cdot (1 + O(\varepsilon))$ .

**Remark 1.9.** For the sake of intuition, we inspect what [Theorem 1.8](#) tells us in the special case where  $G$  is a biregular near-Ramanujan graph. When we plug in  $\tilde{\lambda} = (\sqrt{c-1} + \sqrt{d-1}) \cdot (1 + \varepsilon)$ , we obtain:

$$\rho(B_{G[S]}) \leq ((c-1)(d-1))^{1/4} + \delta(\varepsilon)$$

where  $\delta(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$ .

When  $c = d$ , Kahale's result for  $d$ -regular graphs (e.g., Theorem 3 of [[Kah95](#)]) also has the form  $\frac{1}{2}(\tilde{\lambda} + \sqrt{\tilde{\lambda}^2 - 4(d-1)})$ . The above expression thus generalizes Kahale's result to biregular graphs.

**Remark 1.10** (Sharpness of [Theorem 1.8](#)). One can adapt the techniques of [[MM21](#)] to prove that for *any* graph  $H$  on  $o(n)$  vertices where

$$\rho(B_H) \leq \frac{1}{2} \left( \sqrt{\lambda^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\lambda^2 - (\sqrt{c-1} - \sqrt{d-1})^2} \right),$$

there is a graph  $G$  that contains  $H$  as a subgraph, and  $\lambda_2(A_G) \leq \lambda \cdot (1 + o(1))$ .

As a consequence, we obtain the following result which answers a question raised by [[AD23](#)].

**Theorem 1.11** (Subgraph density in (near-)Ramanujan graphs; see [Theorem 6.1](#)). *Let  $3 \leq c \leq d$  be integers, and let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph such that  $\lambda_2(G) \leq (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$ . Then, there exists  $\delta = \delta(\varepsilon, c, d) > 0$  such that for every  $S_1 \subseteq L$  and  $S_2 \subseteq R$  with  $|S_1| + |S_2| \leq \delta |L \cup R|$ , the left and right average degrees  $d_1 = \frac{|E(S_1, S_2)|}{|S_1|}$  and  $d_2 = \frac{|E(S_1, S_2)|}{|S_2|}$  in the induced subgraph  $G[S_1 \cup S_2]$  must satisfy*

$$(d_1 - 1)(d_2 - 1) \leq \sqrt{(c-1)(d-1)} \cdot (1 + O(\varepsilon + \sqrt{\gamma})).$$

Our refinement to the Moore bound involves using the spectral radius of the non-backtracking matrix of the graph instead of the degree, and yields the existence of *bicycles* — pairs of short cycles that are close in the graph.

**Theorem 1.12** (Generalized Moore bound; see [Theorem 7.4](#)). *Let  $G$  be a graph on  $n$  vertices and let  $\rho = \lambda_1(B_G)$  where  $B_G$  is the non-backtracking matrix of  $G$ . Assuming  $\rho > 1$ ,  $G$  must contain a cycle of length at most  $(2 + o_n(1)) \log_\rho n$  and  $G$  must contain a bicycle of length at most  $(3 + o_n(1)) \log_\rho n$ .*

**Remark 1.13.** For a graph  $G$  with average degree  $d$ ,  $\rho(B_G)$  is at least  $d - 1$ . Therefore, [Theorem 1.12](#) is stronger than the girth guarantee of  $2 \log_{d-1} n$  from the classical irregular Moore bound of [\[AHL02\]](#) in some cases, in particular for some graphs arising in the proof of [Theorem 1.5](#). A simple example where this yields tighter bounds is a  $(d, 2)$ -biregular graph. When  $d \gg 2$ , the average degree is  $\approx 4$  and the classical Moore bound yield a cycle of length  $2 \log_3 n$ . Nevertheless, the generalized Moore bound tells us that there is a cycle of length  $\approx 4 \log_{d-1} n$ .

## 1.2 Context and related work

**Spectral expansion vs. unique-neighbor expansion.** In contrast to unique-neighbor expanders, we have a rich set of spectral and edge expander constructions. A key conceptual difficulty in unique-neighbor expansion is the lack of an “analytic handle” for it. Several other graph properties required in applications of expander graphs, such as high conductance on cuts, low density of small subgraphs, and rapid mixing of random walks, have an excellent surrogate in the second eigenvalue of the normalized adjacency matrix, which is a highly tractable quantity.

It is natural to wonder if any form of unique-neighbor expansion can be deduced from the eigenvalues of a graph. However, the connection between the unique-neighbor expansion in a graph and its spectral properties is tenuous at best. Kahale [\[Kah95\]](#) proved that in  $d$ -regular graphs with optimal spectral expansion, small sets have vertex expansion at least  $d/2$ , i.e., for a sufficiently small constant  $\varepsilon$ , any set  $S$  with at most  $\varepsilon n$  vertices has roughly at least  $|S| \cdot d/2$  distinct neighbors. Observe that once the vertex expansion of a set exceeds  $d/2$ , it begins to be forced to have unique-neighbors. Strikingly, Kahale also showed the  $d/2$  bound is tight for spectral expanders, which makes them fall short at the cusp of the unique-neighbor expansion threshold; indeed, it was proved in [\[KK22\]](#) that certain algebraic bipartite Ramanujan graphs contain sublinear-sized sets with *zero* unique-neighbors (see also [\[Kah95, MM21\]](#) for examples of near-Ramanujan graphs exhibiting a similar property).

**Quantum codes.** Resilience to errors is essential for constructing quantum computers [\[Kit03\]](#), which makes quantum error correction fundamental for quantum computing. One approach to this problem is in the form of *quantum LDPC codes* (qLDPC codes). Recently, a flurry of work culminated in the construction of qLDPC codes with constant rate and distance [\[PK22, LZ22\]](#), which was also coupled with the construction of  *$c^3$ -locally testable codes* [\[DEL<sup>+</sup>22, PK22\]](#). At a high level, these codes are constructed by composing a structured spectral expander, a *square Cayley complex*, along with a structured inner code, a *robustly testable tensor code*. The analysis is complicated by the stringent requirements on the inner code, and poses a barrier for generalizing these constructions. Indeed, it is unclear how to generalize the square Cayley complex and the inner code to construct *quantum locally-testable codes* (qLTCs).

More recently, Lin and Hsieh [\[LH22\]](#) constructed good qLDPC codes with *linear time decoders* assuming the existence of two-sided algebraic lossless expander graphs. Their construction does not require an inner code, and as a byproduct, yields a simpler analysis and is plausibly easier to



generalize to other applications such as qLTCs.<sup>2</sup> However, two-sided algebraic lossless expander graphs are not known to exist, and obtaining them is one of the primary motivations for the goals of this paper.

The *chain complexes* arising in the qLDPC constructions have also been fruitful for other problems in theoretical computer science — constructing explicit integrality gaps for the Sum-of-Squares semidefinite programming hierarchy for the  $k$ -XOR problem [HL22] & the resolution of the quantum NLTS conjecture [ABN23].

The previously known integrality gaps for  $k$ -XOR came from *random* instances [Gri01, Sch08]. Building on the work of Dinur, Filmus, Harsha & Tulsiani [DFHT21], Hopkins & Lin [HL22] constructed *explicit* families of 3-XOR instances that are hard for the Sum-of-Squares (SoS) hierarchy of semidefinite programming relaxations (previously known lower bounds are *random* instances). Specifically, they illustrated  $k$ -XOR instances which are highly unsatisfiable but even  $\Omega(n)$  levels of SoS fail to refute them (i.e., perfect completeness).

**Previous constructions.** The first constructions of unique-neighbor expanders appeared in the work of Alon and Capalbo [AC02]. One of their constructions, which we extend in this paper, takes the line product of a large Ramanujan graph with the 8-vertex 3-regular graph obtained by the union of the octagon and edges connecting diametrically opposite vertices.

Another construction in the same work gives one-sided unique-neighbor expanders of aspect ratio  $22/21$ , and was extended in a recent work of Asherov and Dinur [AD23] to obtain one-sided unique-neighbor expanders of aspect ratio  $\alpha$  for all  $\alpha \geq 1$  where every small set on the left side has at least 1 unique-neighbor. The construction takes a graph product called the *routed product* of a large biregular Ramanujan graph with a constant-sized random graph. (See the recent work of Kopparty, Ron-Zewi & Saraf [KRZS23] for a simplified analysis with weaker ingredients.)

The work of Capalbo, Reingold, Vadhan & Wigderson [CRVW02] constructs one-sided lossless expanders of arbitrarily large degree and arbitrary aspect ratio. Their construction relies on a generalization of the *zig-zag product* of [RVW00] applied to various randomness conductors to construct lossless conductors, analyzed by tracking entropy, which then translates to lossless expanders. More recently, a simpler construction and analysis was given by Golowich [Gol23] based on the routed product (see also [CRTS23] for a similar construction, and see Remark 1.16 for a discussion on where the routed product constructions fall short of achieving two-sided expansion).

Finally, motivated by randomness extractors, the works [TSUZ07, GUV09] construct one-sided lossless expanders where the left side is polynomially larger than the right.

**Applications of unique-neighbor expanders.** Unique-neighbor expanders have several applications in theoretical computer science. In coding theory, it was shown in [DSW06, BV09] that unique-neighbor expander codes [Tan81] are “weakly smooth”, hence when tensored with a code with constant relative distance, they give *robustly testable codes*. In high-dimensional geometry, unique-neighbor expanders were used in [GLR10, Kar11] to construct  $\ell_p$ -spread subspaces as well as in [BGI<sup>+</sup>08, GMM22] to construct matrices with the  $\ell_p$ -restricted isometry property (RIP).

Unique-neighbor expanders were also used in designing *non-blocking networks* [ALM96]: given a set of input and output terminals, the network graph is connected such that no matter which input-output pairs are connected previously, there is a path between any unused input-output pair

---

<sup>2</sup> More recently qLDPCs with linear time decoders have been constructed [DHLV22, GPT22, LZ23], but they still make use of an inner code.

Table 1: Comparison of our [Theorem 1.2](#) with prior work.

Construction?	Which $d$ ? <sup>†</sup>	# unique-neighbors of $S$	2-sided?	Explicit?	Aspect ratio <sup>‡</sup>
Random graphs	any $d$	$(d - \varepsilon) \cdot  S $	✓	✗	any
[AC02]	$\{3, 4, 6\}$	$\Omega( S )$	✓*	✓	1*
[AC02]	$\subseteq [25]$	$\Omega( S )$	✗	✓	22/21
[CRVW02] [Gol23, CRTS23]	large enough $d$	$(d - o(d)) \cdot  S $	✗	✓	any
[Bec16]	6	$\Omega( S )$	✓*	✓	1*
[AD23, KRZS23]	large enough $d$	at least 1	✗	✓	any
<b>this paper</b>	large enough $d$	$\Omega(d S )$	✓	✓	any

\*Non-bipartite construction that can be made bipartite by passing to the double cover.

<sup>†</sup> $d$  here refers to the degree of the left vertex set.

<sup>‡</sup>“Aspect ratio” refers to the ratio between the sizes of the left and right vertex sets.

using unused vertices.

### 1.3 Technical overview

**Line product.** Our construction of two-sided algebraic unique-neighbor expanders, featured in [Theorem 1.2](#), is based on the *line product* between a large *base graph* and a small *gadget graph*. Let  $G$  be a  $D$ -regular graph on  $n$  vertices and  $H$  be a  $d$ -regular graph on  $D$  vertices. The line product  $G \diamond H$  is a graph on the *edges* of  $G$  where for each vertex  $v \in G$  we place a copy of  $H$  on the set of edges incident to  $v$ . See [Definition 3.1](#) for a formal definition and [Figure 1](#) for an example. This graph product was also used in the works of [AC02, Bec16].

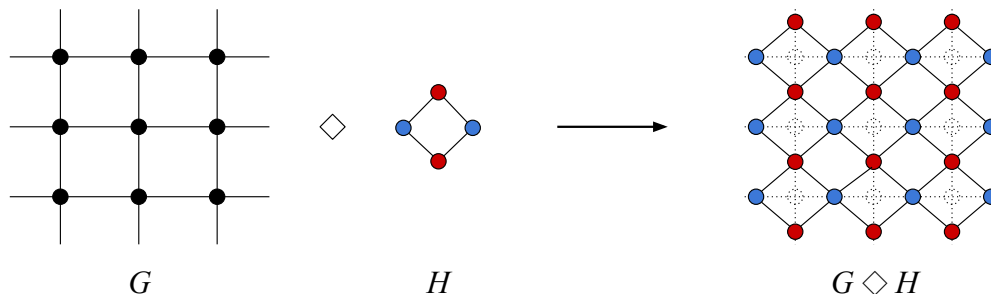


Figure 1: An example of the line product.

Observe that  $G \diamond H$  has  $nD/2$  vertices and is  $2d$ -regular. Note also that the *line graph* of  $G$  (where two edges are connected if they share a vertex) is exactly the line product between  $G$  and the  $D$ -clique, hence the name.

The key lemma ([Lemma 3.3](#)) is that if  $G$  is a small-set (edge) expander and  $H$  is a good unique-neighbor expander, then  $G \diamond H$  is a unique-neighbor expander as well. For the base graph  $G$ , we simply use the explicit Ramanujan graph construction [LPS88, Mor94]. For the gadget  $H$ , we



show that a *random* biregular graph is a good unique-neighbor expander with high probability (Lemma 4.3). Then, since  $D$  is a constant, we can find such a graph by brute force.

**Remark 1.14** (On importance of Ramanujan base graphs). We require an  $O(\sqrt{D})$  bound on the average degree of small subgraphs in a  $D$ -regular expander, which is proved using the fact that the second eigenvalue of a  $D$ -regular Ramanujan graph is  $O(\sqrt{D})$ . Typically, applications of expanders only need a second eigenvalue bound of  $o(D)$ , so we find it noteworthy that the analysis of our construction seems to require being within a constant factor of the Ramanujan bound.

**Tripartite line product.** Our constructions with stronger vertex expansion guarantees for small sets, featured in Theorem 1.5, are based on a suitably generalized version of the line product, which we call the *tripartite line product*. The first ingredient is a large tripartite *base graph*  $G$  on vertex set  $L \cup M \cup R$ , where  $L, R, M$  denote the left, middle, right partitions respectively, and there are bipartite graphs between  $L, M$  and between  $M, R$ . The second ingredient is a small constant-sized bipartite *gadget graph*  $H$ , which is chosen to be an excellent unique-neighbor expander — as before, we can brute-force search to find  $H$  that has expansion as good as a random graph.

The tripartite line product  $G \diamond H$  is a bipartite graph on  $L \cup R$  whose edges are obtained by placing a copy of  $H$  between the left neighbors of  $v$  and the right neighbors of  $v$  for each  $v \in M$ . See Definition 7.6 for a formal definition and Figure 2 for an example.

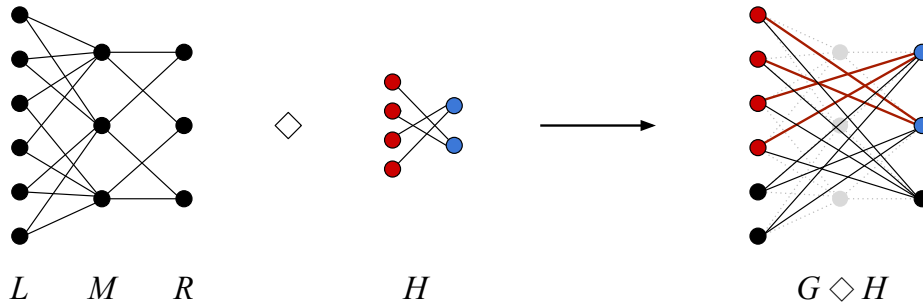


Figure 2: An example of the tripartite line product. The gadget placed on the first vertex in  $M$  is highlighted in  $G \diamond H$ .

To prove Theorem 1.5, we construct the base graph by choosing a  $(K_1, D_1)$ -biregular spectral expander between  $L$  and  $M$ , and a  $(D_2, K_2)$ -biregular spectral expander between  $M$  and  $R$  for large enough and suitably chosen parameters  $K_1, D_1, K_2, D_2$ . The gadget  $H$  is chosen to be a  $(\tilde{d}_1, \tilde{d}_2)$ -biregular graph on vertex set  $[D_1] \times [D_2]$ , where  $d_1 = K_1 \tilde{d}_1$  and  $d_2 = K_2 \tilde{d}_2$ .

**Remark 1.15** (Generalizing the line product and routed product). The *line product* and *routed product*, which feature in [AC02, AD23, Gol23], arise from instantiating the tripartite line product with appropriate base graphs. The line product can be obtained by choosing a  $(2, D_1)$ -biregular graph between  $L$  and  $M$ , and a  $(D_2, 2)$ -biregular graph between  $M$  and  $R$  in the base graph. The routed product arises by choosing a  $(K, D_1)$ -biregular graph between  $L$  and  $M$ , and a  $(D_2, 1)$ -biregular graph between  $M$  and  $R$  in the base graph.

**Remark 1.16** (One-sided vs. two-sided expanders). A key difference between our work and previous constructions that only achieve one-sided expansion is in the choice of the graph between

$M$  and  $R$ . [AD23, Gol23] choose the graph between  $M$  and  $R$  to be a  $(D_2, 1)$ -biregular graphs, equivalently a disjoint collection of stars centered at the vertices in  $M$ . This results in very small sets on the right with no unique-neighbors: for example, for any  $v \in M$ , consider the set of all its right neighbors.

**Overview of the analysis of the tripartite line product.** Let  $Z = G \diamond H$ , and let  $G^{(1)}$  and  $G^{(2)}$  be the bipartite graphs between  $L, M$  and  $M, R$  in  $G$  respectively. We choose the gadget  $H$  to be a  $(\tilde{d}_1, \tilde{d}_2)$ -biregular graph on  $D = D_1 + D_2$  vertices such that  $\tilde{d}_1 \gg \frac{1}{\varepsilon} \sqrt{K_2 D_2}$ . For a subset  $S \subseteq L$ , let  $U = N_G(S) \subseteq M$  (neighbors of  $S$  in the base graph). Our analysis for the expansion of  $S$  roughly follows two steps:

- (1) We partition  $U$  according to the number of edges going to  $S — U_\ell$  (“low  $S$ -degree”) and  $U_h$  (“high  $S$ -degree”). We then show that if we partition  $U$  according to a suitable threshold, then most edges leaving  $S$  go to  $U_\ell$ .
- (2) Since each vertex in  $U_\ell$  has small  $S$ -degree, in the local gadget graph it has “large” unique-neighbor expansion (here we rely on the expansion profile of the gadget; see Lemma 4.3). Then, we show that most of the unique-neighbors in the gadgets are also unique-neighbors of  $S$  in  $Z$ .

Both steps rely on Theorem 1.11. For step (1), we apply Theorem 1.11 on the induced subgraph  $G^{(1)}[S \cup U_h]$  (since  $G^{(1)}$  is near-Ramanujan). If we choose the threshold to be  $\approx \sqrt{D}$ , then since the right average degree of  $G^{(1)}[S \cup U_h]$  is  $\gtrsim \sqrt{D}$ , the left average degree is  $\lesssim 1 + \frac{\sqrt{K_1 D_1}}{\sqrt{D}} \ll K_1$ , thus most edges from  $S$  go to  $U_\ell$  instead of  $U_h$ .

For step (2), let  $\tilde{T} \subseteq R$  be the union of the unique-neighbors within each gadget. By the expansion of the gadgets on vertices in  $U_\ell$ , we have a lower bound on  $|\tilde{T}|$ . However, a unique-neighbor in one gadget may also have edges from other gadgets, in which case it is *not* a unique-neighbor of  $S$  in  $Z$ . To resolve this issue, we analyze the induced subgraph  $G^{(2)}[U \cup \tilde{T}]$  and show that a large fraction of  $\tilde{T}$  are unique-neighbors of  $U$  in  $G^{(2)}$ , thus must also be unique-neighbors of  $S$  in  $Z$ . This is done by observing that the left average degree of  $G^{(2)}[U \cup \tilde{T}]$  must be  $\gtrsim \tilde{d}_1$ . Thus, the right average degree is  $\lesssim 1 + \frac{\sqrt{K_2 D_2}}{\tilde{d}_1} \leq 1 + \varepsilon$  since we choose  $\tilde{d}_1 \gg \frac{1}{\varepsilon} \sqrt{K_2 D_2}$ .

One might attempt to tweak the parameters of the construction to obtain *two-sided* lossless expansion. However, this fails because in step (1) we need the threshold to be large enough such that  $S$  has lossless expansion in  $G^{(1)}$ , but then it is not possible to set the parameters of the gadget such that (i) each vertex in  $U_\ell$  expands losslessly, and (ii)  $\tilde{d}_1 \gg \frac{1}{\varepsilon} \sqrt{K_2 D_2}$  (and  $\tilde{d}_2 \gg \frac{1}{\varepsilon} \sqrt{K_1 D_1}$ ) for the analysis in step (2). See Section 7.2 (the proof of Theorem 7.8) for details.

**Lossless expansion of small sets.** For small subsets  $S \subseteq L$ , we directly show that  $S$  expands losslessly into  $U$  under the assumption that  $G^{(1)}$  has no short bicycles. Specifically, in Lemma 7.5 we prove that if a graph has no bicycle of length  $g$ , then all sufficiently small subsets (in particular, of size at most  $\exp(O(g))$ ) expand losslessly. To prove this, we first show that if a degree- $K$  set  $S$  has expansion less than  $(1 - \varepsilon)K$ , then we can lower bound the spectral radius of the non-backtracking matrix of the graph induced on  $S \cup N(S)$  by  $C(\varepsilon) := \sqrt{\varepsilon(K - 1)}$ . Now, by the generalized Moore bound (Theorem 1.12), there is a bicycle of size  $O(\log_{C(\varepsilon)} |S|)$  in  $G$ . Since  $G$  has no bicycle of length- $g$ , it lower bounds  $|S|$  via the inequality  $g \leq O(\log_{C(\varepsilon)} |S|)$ . This tells us that small sets in  $G^{(1)}$  exhibit lossless expansion. To establish that small sets in  $Z$  are losslessly expanding, we follow

the same strategy as before for step (2): since most vertices in  $U$  has only 1 edge to  $S$ , they expand by a factor of  $\tilde{d}_1$  (from the gadget), and we use [Theorem 1.11](#) to show that  $T = N_Z(S)$  are mostly unique-neighbors, proving the small set expansion result in [Theorem 1.5](#).

**One-sided lossless expanders.** As mentioned in [Remark 1.7](#), we are able to use the tripartite line product to construct *one-sided* lossless expanders, recovering the result of Louis [[Gol23](#)]. Our proof is almost the same as [Theorem 1.5](#), but with  $K_1 \gg K_2 = 1$  (hence  $G^{(2)}$  is just a collection of stars). In step (1), we choose a smaller  $S$ -degree threshold  $\approx \varepsilon\sqrt{D}$  and large enough  $K_1$  to ensure that  $1 - \varepsilon$  fraction of edges from  $S$  go to  $U_\ell$  (via [Theorem 1.11](#)). Then, with the smaller threshold, each vertex in  $U_\ell$  expands losslessly, i.e., by a factor of  $(1 - \varepsilon)\tilde{d}_1$ . Since all gadgets have disjoint right vertices, there is no collision between gadgets, which finishes the proof.

**Subgraphs in near-Ramanujan bipartite graphs.** We now give an overview of the proofs of [Theorems 1.8](#) and [1.11](#). The main ingredient is the *Bethe Hessian* of a graph  $G$  defined as  $H_G(t) := (D_G - \mathbb{1})t^2 - A_G t + \mathbb{1}$ . Specifically, for an induced subgraph  $G[S]$ , we identify  $\alpha > 0$  such that  $H_{G[S]}(t) \succ 0$  for  $t \in [0, \alpha]$  ([Theorem 5.1](#)), which then implies [Theorem 1.8](#) via the well-known Ihara–Bass formula ([Fact 2.3](#)).

To prove  $H_{G[S]}(t) \succ 0$  for  $t \in [0, \alpha]$ , we show that  $f^\top H_{G[S]}(t)f > 0$  for any  $f : G[S] \rightarrow \mathbb{R}$ . To establish  $f^\top H_{G[S]}(t)f > 0$  for the claimed values of  $t$ , we need to relate  $f^\top H_{G[S]}(t)f$  to the spectrum of the entire graph  $G$ . To this end, we consider the *regular tree extension*  $T$  of  $G[S]$  ([Definition 5.6](#)), which is obtained by attaching trees (of depth  $\ell$ ) to each vertex in  $S$  such that the resulting graph is  $(c, d)$ -biregular except for the leaves. Intuitively, the tree extension serves as a proxy of the  $\ell$ -step neighborhood of  $S$  in  $G$ . We then define the appropriate function extension  $f_t$  of  $f$  ([Definition 5.4](#)) on the tree extension with “sufficient decay” down the tree (where the decay results from an upper bound on  $t$ ). For appropriately bounded  $t$ , we can show  $f^\top H_{G[S]}(t)f$  is approximately equal to  $f_t^\top H_G(t)f_t$ , which can be controlled via the spectrum of  $G$ .

**Erdős–Rényi vs. random regular graphs.** One technical subtlety is that the edges of a random regular graph are (slightly) *correlated*, which makes it difficult to directly analyze its unique-neighbor expansion. On the other hand, the analysis is straightforward for Erdős–Rényi graphs since the edges are drawn independently ([Lemma B.1](#)).

Thus, we give a slight generalization of an embedding theorem given in [[FK16](#)] between Erdős–Rényi graphs and random regular graphs ([Lemma B.2](#)) which allows us to extend the analysis to random regular graphs.

## 1.4 Open questions

**Quantum codes from unique-neighbor expanders.** Lin & Hsieh [[LH22](#)] construct quantum codes assuming the existence of two-sided algebraic lossless expanders, and their current proof requires the unique-neighbor expansion of sets of vertices with degree- $d$  to exceed  $d/2$ .

In contrast, in the setting of classical LDPC codes, if all subsets of vertices of size at most  $\Delta$  have even a single unique-neighbor, the resulting code is guaranteed to have distance at least  $\Delta$ , albeit without a clear decoding algorithm.

**Question 1.17.** Does the construction of [[LH22](#)] yield a good quantum code when instantiated with a  $\gamma$ -two-sided algebraic unique-neighbor expander for small  $\gamma > 0$ ?

**Algebraic two-sided lossless expanders from random graphs.** Two-sided lossless expanders with relevant algebraic properties are not known to exist, even using randomness. Random bipartite graphs exhibit two-sided lossless expansion (see, for example, [HLW06, Theorem 4.16]), but do not admit any nontrivial group actions.

A natural approach is to use an algebraic random graph, such as a random Cayley graph, and study its vertex expansion properties. It is also possible to achieve lossless expansion using the tripartite line product if the small-set edge expansion of the base graph is far better than that guaranteed by spectral expansion. A concrete question in this direction that could result in a lot of progress is the following.

**Question 1.18** (Beyond spectral certificates). Let  $\Gamma$  be a group and let  $S$  be a set of  $D$  generators chosen independently and uniformly from  $\Gamma$ . Let  $G := \text{Cay}(\Gamma, S)$  be the Cayley graph given by the generator set  $S$ . Is it true that with high probability over the choice of  $S$ : for all subsets of vertices  $T \subseteq \Gamma$  where  $|T| \leq |\Gamma|/\text{poly}(D)$ , the number of edges inside  $T$  is at most  $0.1\sqrt{D}|T|$ ?

**Remark 1.19.** Spectral expansion can at best guarantee that the number of edges inside a set  $T$  is at most  $\sqrt{D}|T|$ . A resolution to the above question would necessarily use other properties of the random graph and group, beyond merely the magnitude of its eigenvalues.

## 1.5 Organization

In [Section 2](#), we give some technical preliminaries. In [Section 3](#), we define the line product and prove its unique-neighbor expansion assuming good expansion properties of the base and gadget graphs. In [Section 4](#), we prove [Theorem 1.2](#) by instantiating the line product with a suitably chosen base graph and gadget graph.

In [Section 5](#), we prove [Theorem 1.8](#). In [Section 6](#), we use [Theorem 1.8](#) to prove [Theorem 1.11](#). In [Section 7](#), we define the tripartite line product, and instantiate it with an appropriately chosen base graph and gadget graph to prove [Theorem 1.5](#) via [Theorem 1.11](#) and [Theorem 1.12](#). In [Section 8](#), we show how we can recover Golowich’s [Gol23] construction and analysis of one-sided lossless expanders as an instantiation of the tripartite line product and an application of [Theorem 1.11](#).

In [Appendix A](#), we prove the sharpened Moore bound [Theorem 1.12](#). In [Appendices B](#) and [C](#), we analyze the expansion profile of the gadget graph.

## 2 Preliminaries

**Notation.** Given a graph  $G$ , we use  $V(G)$  to denote its set of vertices,  $E(G)$  to denote its set of edges. If  $G$  is bipartite, we use  $L(G)$  and  $R(G)$  to denote its left and right vertex sets respectively. We write  $\deg_G(v)$  to denote the degree of vertex  $v$  in  $G$  (we will omit the subscript  $G$  if clear from context). We say that a bipartite graph  $G$  is  $(d_1, d_2)$ -biregular if  $\deg_G(v) = d_1$  for all  $v \in L(G)$  and  $\deg_G(v) = d_2$  for all  $v \in R(G)$ .

For any subset of edges  $F$ , we use  $F(v) \subseteq F$  to denote the set of edges in  $F$  incident to  $v$ , hence  $|F(v)| = \deg_F(v)$ . For a set of vertices  $S$ , we use  $e_F(S)$  to denote the number of edges in  $F$  with both endpoints in  $S$ , and  $e_F(S, T)$  to denote the number of  $(u, v) \in S \times T$  such that  $\{u, v\} \in F$  (for disjoint sets  $S$  and  $T$ ). We will omit the subscript if  $F = E(G)$ . We denote the eigenvalues of the

normalized adjacency matrix of  $G$  by  $1 = \lambda_1(G) \geq \dots \geq \lambda_n(G) \geq -1$ . We say  $G$  is a  $\lambda$ -spectral expander if  $\lambda_2(G) \leq \lambda$ .<sup>3</sup>

**Random bipartite graph models.** Throughout the paper, we will write random variables in **boldface**. Fix  $n_1, n_2, m, d_1, d_2 \in \mathbb{N}$  such that  $n_1 d_1 = n_2 d_2$ . We denote  $K_{n_1, n_2}$  as the complete bipartite graph with left/right vertex sets  $L = [n_1]$  and  $R = [n_2]$ . We use  $\mathbf{G} \sim \mathbf{G}_{n_1, n_2, m}$  to denote a random graph sampled from the uniform distribution over (simple) bipartite graphs on  $L = [n_1], R = [n_2]$  with exactly  $m$  edges. With slight abuse of notation, we use  $\mathbf{H} \sim \mathbf{G}_{n_1, n_2, p}$  for  $p \in (0, 1)$  to denote a random graph such that each potential edge is included with probability  $p$ . Similarly, we use  $\mathbf{R} \sim \mathbf{R}_{n_1, n_2, d_1, d_2}$  to denote a graph from the uniform distribution over  $(d_1, d_2)$ -biregular bipartite graphs on  $L = [n_1], R = [n_2]$ .

## 2.1 Graph expansion

It is a standard fact that small subgraphs in spectral expanders have bounded number of edges. The following is a special case of the Expander Mixing Lemma [AC88], and we include a short proof for completeness.

**Lemma 2.1.** *Let  $G$  be a  $D$ -regular  $\lambda$ -spectral expander. Then for any set  $S \subseteq V(G)$ , where  $|S| = \varepsilon|V(G)|$ :*

$$e(S) \leq D|S| \cdot \frac{\lambda + \varepsilon}{2}.$$

*Proof.* Let  $n = |V(G)|$ ,  $A$  be the (unnormalized) adjacency matrix of  $G$ , and  $\vec{1}_S \in \{0, 1\}^n$  be the indicator vector of  $S$ . We can decompose  $\vec{1}_S$  as  $\vec{1}_S = \frac{|S|}{n}\vec{1} + u$  where  $u \perp \vec{1}$  and  $\|u\|_2^2 = |S|(1 - \frac{|S|}{n}) \leq |S|$ . Then,  $2e(S) = \vec{1}_S^\top A \vec{1}_S \leq \frac{D|S|^2}{n} + \lambda_2(G) \cdot D\|u\|_2^2 \leq D|S|(\varepsilon + \lambda)$ .  $\square$

Within graphs of low ‘‘hereditary’’ average degree, a significant fraction of edges are incident to low-degree vertices.

**Lemma 2.2.** *For any  $\gamma > 0$ , let  $F$  be a graph such that for all  $S \subseteq V(F)$ ,  $2e(S) \leq \gamma|S|$ . Write  $V(F) = F_\ell \sqcup F_h$  where  $F_\ell$  comprises all vertices  $v$  such that  $\deg(v) \leq 2\gamma$  and  $F_h$  to denote the remaining vertices. Then:*

$$2e(F) \leq 3 \sum_{v \in F_\ell} \deg(v).$$

*Proof.* On one hand, by assumption we have:

$$2e(F_h) \leq \gamma|F_h|.$$

On the other hand,

$$2e(F_h) + e(F_h, F_\ell) = \sum_{v \in F_h} \deg(v) \geq 2\gamma|F_h| \geq 4e(F_h).$$

Consequently  $e(F_h, F_\ell) \geq 2e(F_h)$ . Since  $F_h$  and  $F_\ell$  are disjoint:

$$2e(F) = 2e(F_h) + e(F_h, F_\ell) + e(F_\ell, F_h) + 2e(F_\ell).$$

This gives us:  $2e(F) \leq 3e(F_h, F_\ell) + 2e(F_\ell) \leq 3 \sum_{v \in F_\ell} \deg(v)$ .  $\square$

<sup>3</sup> This is in contrast to most scenarios where one requires both  $\lambda_2(G)$  as well as  $-\lambda_n(G)$  to be bounded by  $\lambda$ .

## 2.2 Non-backtracking matrix

**Notation.** Given an undirected graph  $G = (V, E)$  with  $|V| = n$  vertices and  $|E| = m$  edges, we denote  $A_G \in \{0, 1\}^{n \times n}$  to be its adjacency matrix,  $D_G \in \mathbb{R}^{n \times n}$  to be its diagonal degree matrix, and finally  $B_G \in \{0, 1\}^{2m \times 2m}$  to be its non-backtracking matrix (introduced by [Has89]) defined as follows: for *directed* edges  $(u_1, v_1), (u_2, v_2)$  in the graph,

$$B_G[(u_1, v_1), (u_2, v_2)] = \mathbf{1}(v_1 = u_2) \cdot \mathbf{1}(u_1 \neq v_2).$$

Note that the non-backtracking matrix is not symmetric. Let  $\lambda_1, \dots, \lambda_{2m} \in \mathbb{C}$  be the eigenvalues of  $B_G$  ordered such that  $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{2m}|$ . The Perron-Frobenius theorem implies that  $\lambda_1$  is real and non-negative. We denote  $\rho(B_G) = \lambda_1$  to be the spectral radius of  $B_G$ .

A crucial identity we will need is the Ihara-Bass formula [Iha66, Bas92] which gives a translation between the eigenvalues of the adjacency matrix and the eigenvalues of the non-backtracking matrix:

**Fact 2.3** (Ihara-Bass formula). *For any graph  $G$  with  $n$  vertices and  $m$  edges, the following identity on univariate polynomials is true:*

$$\det(\mathbb{1} - B_G t) = \det(H_G(t)) \cdot (1 - t^2)^{m-n}$$

where  $H_G(t) := (D_G - \mathbb{1})t^2 - A_G t + \mathbb{1}$  is the Bethe Hessian of  $G$ .

The Ihara-Bass formula gives a direct relationship between the spectral radius of  $B_G$  and the positive definiteness of  $H_G(t)$ . The following is classic (e.g. [FM17, Proof of Theorem 5.1]), though we include a proof for completeness.

**Lemma 2.4.** *Let  $G$  be a graph and  $0 < \alpha < 1$ . Then, the spectral radius  $\rho(B_G) \leq \frac{1}{\alpha}$  if and only if  $H_G(t) \succ 0$  for all  $t \in [0, \alpha)$ . As a result, if  $H_G(\frac{1}{\rho})$  has a non-positive eigenvalue for some  $\rho > 0$ , then  $\rho(B_G) \geq \rho$ .*

*Proof.* First observe that  $H_G(0) = \mathbb{1} \succ 0$ . Since  $H_G(t)$  is symmetric, the eigenvalues of  $H_G(t)$  are real and move continuously on the real line as  $t$  increases from 0. Note also that by the Perron-Frobenius theorem,  $\rho(B_G) = \lambda_1(B_G) \geq 0$ .

Suppose for contradiction that  $\rho(B_G) \leq \frac{1}{\alpha}$  but  $H_G(t)$  has a non-positive eigenvalue for some  $t \in [0, \alpha)$ . Due to  $H_G(0) = \mathbb{1}$  and continuity of the eigenvalues, there must be a  $t^* \in (0, t]$  such that  $H_G(t^*)$  has a zero eigenvalue, meaning  $\det(H_G(t^*)) = 0$ . By Fact 2.3, this means that  $\det(\mathbb{1} - B_G t^*) = 0$ , i.e.,  $B_G$  has an eigenvalue  $\frac{1}{t^*} \geq \frac{1}{t} > \frac{1}{\alpha}$ . This contradicts that  $\rho(B_G) \leq \frac{1}{\alpha}$ .

On the other hand, if  $H_G(t) \succ 0$  for all  $t \in [0, \alpha)$ , then by Fact 2.3  $\det(\mathbb{1} - B_G t) > 0$  for all  $t \in [0, \alpha)$ . Since  $\det(\mathbb{1} - B_G / \lambda_1) = 0$ , it follows that  $\frac{1}{\lambda_1} \geq \alpha$ , i.e.,  $\lambda_1 \leq \frac{1}{\alpha}$ .

Finally, suppose  $H_G(\frac{1}{\rho}) \not\succeq 0$  for some  $\rho > 0$ , then setting  $\alpha = \frac{1}{\rho - \varepsilon} > \frac{1}{\rho}$  for any  $\varepsilon \rightarrow 0^+$ , we have that  $\rho(B_G) > \rho - \varepsilon$ . This implies that  $\rho(B_G) \geq \rho$ .  $\square$

## 3 The line product of graphs

Our construction is based on taking the *line product* of a suitably chosen spectral expander and unique-neighbor expander. See Figure 1 for an example.



**Definition 3.1** (Line product). Let  $G$  be a  $D$ -regular graph on vertex set  $[n]$ , and let  $H$  be a graph on vertex set  $[D]$ . For each  $v \in V(G)$  and  $i \in [D]$ , let  $e_v^i$  denote the  $i$ -th incident edge to  $v$ . The *line product*  $G \diamond H$  is the graph on vertex set  $E(G)$  and edges obtained by placing a copy of  $H$  on  $E(G)$  for each  $v \in V(G)$ , such that  $\{e_v^i, e_v^j\}$  is an edge in  $H(v)$  if and only if  $\{i, j\}$  is an edge in  $H$ .

For convenience, we denote  $H(v)$  to be the subgraph of  $G \diamond H$  given by the copy of  $H$  associated with  $v$ .

**Definition 3.2.** Given a graph  $H$ , we denote  $\text{UN}_H(S)$  to be the set of unique-neighbors of  $S$ . The *unique-neighbor expansion profile* of a graph  $H$ , denoted  $P_H$ , is defined:

$$P_H(t) := \min_{S \subseteq V(H): |S| \leq t} \frac{|\text{UN}_H(S)|}{|S|} \quad \text{for } t \geq 1.$$

When  $H$  is a bipartite graph, we use  $P_H^{(L)}(t)$  and  $P_H^{(R)}(t)$  to denote the analogous quantity where the minimum is taken only over subsets of the left and right respectively.

**Lemma 3.3** (Expansion profile of the line product). *Let  $\gamma > 0$  and  $\varepsilon \in (0, 1)$ . Suppose*

1.  $G$  is a  $D$ -regular graph such that  $2e(S) \leq \gamma|S|$  for all  $S \subset V(G)$  with  $|S| \leq \varepsilon|V(G)|$ , and
2.  $H$  is a graph on  $[D]$  such that  $P_H(t) \geq 12\gamma/t$  for  $1 \leq t \leq 2\gamma$ .

Then for  $Z := G \diamond H$ ,  $P_Z(\frac{\varepsilon}{D} \cdot |V(Z)|) \geq \frac{P_H(2\gamma)}{3}$ .

*Proof.* Let  $T \subseteq V(Z)$ , viewed as a collection of edges in  $G$ , be such that  $|T| \leq \frac{\varepsilon}{D} \cdot |V(Z)| = \frac{\varepsilon}{2} \cdot |V(G)|$  (since  $|V(Z)| = |E(G)| = \frac{D}{2}|V(G)|$ ). Let  $S \subseteq V(G)$  be the set of vertices of  $G$  touched by  $T$ . Note that  $|S| \leq 2|T| \leq \varepsilon|V(G)|$ . Recall that we denote  $T(v) \subseteq T$  to be the set of edges in  $T$  incident to  $v$ , and we have  $\deg_T(v) = |T(v)| \leq \deg_{G[S]}(v)$ . Then,

$$|\text{UN}_Z(T)| = \sum_{v \in S} \left| \text{UN}_{H(v)}(T(v)) \right| - \sum_{\substack{v \in S \\ v' \in S \\ v \neq v'}} \left| \text{UN}_{H(v)}(T(v)) \cap \text{UN}_{H(v')}(T(v')) \right|.$$

Each summand in the first term of the right-hand side can be lower bounded using  $P_H$  (see below). A vertex  $\{v_1, v_2\} \in V(Z)$  (an edge in  $G$ ) is contained in exactly two subgraphs  $H(v_1)$  and  $H(v_2)$ , so it can only be counted twice in the second term, which means we can bound the whole sum by  $2e(S)$ .

Since  $|S| \leq \varepsilon|V(G)|$ , by the assumption on  $G$ , both the induced subgraph  $G[S]$  and  $T$  (viewed as a subgraph of  $G[S]$ ) satisfy the assumption of [Lemma 2.2](#), i.e., all  $S' \subseteq S$  satisfies  $2e_T(S') \leq 2e(S') \leq \gamma|S'|$ . Thus, define  $S_1 := \{v \in S : \deg_T(v) \leq 2\gamma\}$  and  $S_2 := \{v \in S : \deg_{G[S]}(v) \leq 2\gamma\} \subseteq S_1$ . Applying [Lemma 2.2](#) to both  $T$  and  $G[S]$ , we get that  $2|T| = 2e_T(S) \leq 3 \sum_{v \in S_1} \deg_T(v)$  and  $2e(S) \leq 3 \sum_{v \in S_2} \deg_{G[S]}(v) \leq 6\gamma|S_2| \leq 6\gamma|S_1|$ , hence

$$\begin{aligned} |\text{UN}_Z(T)| &\geq \sum_{v \in S} |T(v)| \cdot P_H(|T(v)|) - 2e(S) \\ &\geq \sum_{v \in S} |T(v)| \cdot P_H(|T(v)|) - 6\gamma|S_1| \\ &= \sum_{v \in S_1} |T(v)| \cdot \left( P_H(|T(v)|) - \frac{6\gamma}{|T(v)|} \right) + \sum_{v \in S \setminus S_1} |T(v)| \cdot P_H(|T(v)|) \end{aligned}$$

$$\geq \frac{1}{2} \sum_{v \in S_1} |T(v)| \cdot P_H(|T(v)|),$$

where the last inequality is by  $|T(v)| = \deg_T(v) \leq 2\gamma$  for all  $v \in S_1$  and the assumption that  $P_H(t) \geq 12\gamma/t$  for all  $t \leq 2\gamma$ . Then since  $P_H(t)$  is monotonically decreasing with  $t$  and  $\sum_{v \in S_1} |T(v)| \geq \frac{2}{3}|T|$ ,

$$\frac{|\text{UN}_Z(T)|}{|T|} \geq \frac{1}{2|T|} \sum_{v \in S_1} |T(v)| \cdot P_H(2\gamma) \geq \frac{P_H(2\gamma)}{3}. \quad \square$$

## 4 Algebraic unique-neighbor expanders

We use a Ramanujan graph equipped with symmetries bestowed by its Cayley graph structure as our base spectral expander.

**Fact 4.1** (Ramanujan graph construction [LPS88, Mor94]). *For every  $D = p^r + 1$  where  $p$  is prime and  $r \in \mathbb{N}$ , there is an infinite family of groups  $(\Gamma_n)_{n \in \mathbb{N}}$  and a collection of generators  $A \subseteq \Gamma_n$  closed under inversion where  $|A| = D$  such that the Cayley graph  $G := \text{Cay}(\Gamma_n, A)$  is a  $D$ -regular Ramanujan graph, i.e., it is a  $\frac{2\sqrt{D-1}}{D}$ -spectral expander.*

For arbitrary  $D$ , by deleting a few edges, we can get  $D$ -regular Cayley graphs that satisfy the expanding condition in [Lemma 3.3](#) with similar parameters as Ramanujan graphs.

**Lemma 4.2** (Expanding Cayley graphs of every degree). *For every  $D \in \mathbb{N}$ ,  $D \geq 3$ , there is an infinite family of groups  $(\Gamma_n)_{n \in \mathbb{N}}$  and a collection of generators  $A \subseteq \Gamma_n$  closed under inversion where  $|A| = D$  such that the Cayley graph  $G := \text{Cay}(\Gamma_n, A)$  is a  $D$ -regular graph such that*

$$2e(S) \leq 3D|S| \left( \frac{2\sqrt{D-1}}{D} + \varepsilon \right), \quad \forall S \subseteq V(G), |S| = \varepsilon|V(G)|.$$

*Proof.* For odd  $D$ , there exists an  $r \in \mathbb{N}$  such that  $D - 1 \leq 2^r \leq 2(D - 1)$ , thus we set  $D' = 2^r + 1 \leq 2D$ . For even  $D$ , there exists an  $r \in \mathbb{N}$  such that  $D - 1 \leq 3^r \leq 3(D - 1)$ , thus we set  $D' = 3^r + 1 \leq 3D$ . We construct the  $D'$ -regular graph  $G' = \text{Cay}(\Gamma_n, A')$  via [Fact 4.1](#) such that  $\lambda_2(G') \leq \frac{2\sqrt{D'-1}}{D'} \leq \frac{2\sqrt{D-1}}{D}$  (since  $\frac{2\sqrt{x-1}}{x}$  is decreasing with  $x$  for  $x \geq 2$ ).

Since  $D$  and  $D'$  have the same parity, we can remove pairs of generators  $g \neq g^{-1}$  from  $A'$  until there are  $D$  elements left (if at some point only self-inverse elements remain, then we start removing them one at a time). Let  $A$  be the remaining generators with  $|A| = D$ . By construction,  $G := \text{Cay}(\Gamma_n, A)$  is  $D$ -regular.

Now, we upper bound  $e(S)$ . Deleting edges can only decrease  $e(S)$ , so by [Lemma 2.1](#),

$$2e(S) \leq D'|S|(\lambda_2(G') + \varepsilon) \leq 3D|S| \left( \frac{2\sqrt{D-1}}{D} + \varepsilon \right).$$

This completes the proof. □

For the gadget, we will need a unique-neighbor expander with strong quantitative guarantees.

**Lemma 4.3.** Let  $\beta \in (0, 1/2]$ ,  $\theta > 0$ , and  $\tau > 0$  be constants. For integers  $d_1, d_2, D_1, D_2$  and  $D := D_1 + D_2$  satisfying

1.  $\frac{d_1}{d_2} = \frac{D_2}{D_1}$ ,
2.  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$ ,
3.  $\theta\sqrt{D}/2 \leq d_1 + d_2 \leq \theta\sqrt{D}$ ,

there is a  $(d_1, d_2)$ -biregular graph  $H$  with  $D_1$  vertices on the left and  $D_2$  vertices on the right such that:

$$P_H(t) \geq (1 - o_D(1)) \cdot d_1 \cdot \exp(-\theta t / \sqrt{D})$$

for  $1 \leq t \leq \tau\sqrt{D}$  where  $o_D(1)$  hides constant factors depending only on  $\beta, \theta$  and  $\tau$ .

We defer the proof of [Lemma 4.3](#) to [Appendix B](#), and prove our main theorem below.

**Theorem 4.4** (Formal version of [Theorem 1.2](#)). For every  $\beta \in (0, 1/2]$ , there are  $d_0 = d_0(\beta) > 0$  and  $\delta = \delta(\beta) > 0$  such that for all even  $d_1, d_2 \geq d_0$  with  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$ , there is an infinite family of  $(2d_1, 2d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  with

$$|\text{UN}_{Z_n}(S)| \geq \delta \cdot d_1 \cdot |S|, \quad \forall S \subseteq V(Z_n) \text{ with } |S| \leq \frac{1}{d_1^3} \cdot |V(Z_n)|. \quad (1)$$

*Proof.* The construction of  $Z_n$  is based on taking the line product of  $G_n$  from [Fact 4.1](#) and a bipartite gadget graph  $H$  from [Lemma 4.3](#) for suitably chosen parameters.

Fix parameters  $\tau = 18$ ,  $\theta = \frac{40\tau}{\beta}$ , and choose  $d_0$  to be large enough such that  $d_0 \geq \frac{8}{\beta}e^{\tau\theta}$  and such that for  $D \geq 4d_0^2/\theta^2$ , the  $o_D(1)$  terms in [Lemma 4.3](#) and all subsequent occurrences in this proof are smaller than 0.1. For  $d_1, d_2 \geq d_0$ , let  $D_1 := \left\lceil \frac{d_1+d_2}{\theta^2} \right\rceil \cdot d_2$  and  $D_2 := \left\lceil \frac{d_1+d_2}{\theta^2} \right\rceil \cdot d_1$ , and define  $D := D_1 + D_2$ . This choice of parameters satisfies the requirements of [Lemma 4.3](#), and hence there is a  $(d_1, d_2)$ -biregular graph  $H$  with  $D_1$  left vertices,  $D_2$  right vertices, and

$$P_H(t) \geq 0.9 \cdot d_1 \cdot \exp(-\theta t / \sqrt{D}) \quad (2)$$

for  $t \leq \tau\sqrt{D}$ .

We choose  $G_n$  as an  $n$ -vertex  $D$ -regular expander graph from [Lemma 4.2](#). It remains to prove that  $Z_n = G_n \diamond H$  has the claimed unique-neighbor expansion guarantee, and is indeed a  $(2d_1, 2d_2)$ -biregular graph (which requires an appropriate ordering of  $H$ ).

Note that by [Lemma 4.2](#),  $G_n$  satisfies  $2e(S) \leq 9\sqrt{D} \cdot |S|$  for all  $S \subseteq V(G_n)$  with  $|S| \leq \frac{1}{\sqrt{D}}n$ , which satisfies the first condition in [Lemma 3.3](#) with  $\varepsilon = \frac{1}{\sqrt{D}}$  and  $\gamma = 9\sqrt{D} = \frac{\tau}{2}\sqrt{D}$ . Thus, it suffices to show that  $P_H(t)$  satisfies the second condition in [Lemma 3.3](#) (a weaker lower bound):

$$(2) \geq \frac{12\gamma}{t} = \frac{6\tau\sqrt{D}}{t}, \quad \text{for } 1 \leq t \leq 2\gamma = \tau\sqrt{D}. \quad (3)$$

This allows us to apply [Lemma 3.3](#) and the (stronger) lower bound of  $P_H(t)$  in [Eq. \(2\)](#) to get

$$P_{Z_n} \left( \frac{1}{D^{3/2}} \cdot |V(Z_n)| \right) \geq \frac{1}{3} P_H(\tau\sqrt{D}) \geq 0.3e^{-\tau\theta} \cdot d_1 := \delta \cdot d_1$$

for constant  $\delta = 0.3e^{-\tau\theta}$ . Since  $D \leq \frac{4}{\theta^2}(d_1 + d_2)^2$ ,  $\beta(d_1 + d_2) \leq d_1$ , and  $\beta\theta = 40\tau$ , we have  $\frac{1}{D^{3/2}} \geq \frac{1}{d_1^3}$ . As  $P_{Z_n}(t)$  is a decreasing function with  $t$ , this establishes the desired unique-neighbor expansion as articulated in Eq. (1), finishing the proof of the theorem.

Now, to establish Eq. (3), observe that the function  $xe^{-x}$  is monotone increasing for  $x \leq 1$  and monotone decreasing for  $x \geq 1$ , hence for  $x \in [a, b]$ ,  $xe^{-x} \geq \min\{ae^{-a}, be^{-b}\}$ . Thus, for  $1 \leq t \leq \tau\sqrt{D}$ ,

$$\frac{\theta t}{\sqrt{D}} \cdot e^{-\theta t/\sqrt{D}} \geq \min \left\{ \frac{\theta}{\sqrt{D}} \cdot e^{-\theta/\sqrt{D}}, \tau\theta \cdot e^{-\tau\theta} \right\}.$$

By using  $d_1 \geq \beta(d_1 + d_2)$ ,  $d_1 + d_2 \geq \theta\sqrt{D}/2$  and the above, from Eq. (2) we get  $P_H(t) \geq \frac{\sqrt{D}}{t} \cdot 0.45\beta\theta \cdot \min\{e^{-\theta/\sqrt{D}}, \tau\sqrt{D} \cdot e^{-\tau\theta}\}$ .

With our choice of parameters,  $\beta\theta \geq 40\tau$  and  $\sqrt{D} \geq 2d_0/\theta \geq \theta$  imply that  $0.45\beta\theta \cdot e^{-\theta/\sqrt{D}} \geq 6\tau$ . Furthermore,  $\theta\sqrt{D} \geq d_1 + d_2 \geq 2d_0 \geq \frac{16}{\beta}e^{\tau\theta}$  implies that  $0.45\beta\theta \cdot \tau\sqrt{D}e^{-\tau\theta} \geq 6\tau$ . Therefore, we have established Eq. (3).

Finally, we show that  $Z_n$  is a  $(2d_1, 2d_2)$ -biregular graph. Since  $G_n = \text{Cay}(\Gamma, A)$  is a Cayley graph with generators  $A$ , each edge  $\{u, v\}$  is labeled by group elements  $a$  and  $a^{-1}$  in  $A$ , i.e.,  $\{u, v\} = e_u^a = e_v^{a^{-1}}$ . To construct the line product  $G_n \diamond H$ , we need a bijective map  $\varphi$  between  $A$  and  $V(H) = L(H) \cup R(H)$  such that each pair  $a, a^{-1} \in A$  gets assigned to the same side of  $H$ . This can be done as long as  $d_1$  and  $d_2$  are even.

Let  $L(Z_n) = \{e_v^a : v \in V(G_n), \varphi(a) \in L(H)\}$  and  $R(Z_n) = \{e_v^a : v \in V(G_n), \varphi(a) \in R(H)\}$ . First,  $L(Z_n)$  and  $R(Z_n)$  is a disjoint partition of  $E(G_n) = V(Z_n)$  since  $a, a^{-1}$  are assigned to the same side of  $H$ . Moreover, all edges of  $Z_n$  are between  $L(Z_n)$  and  $R(Z_n)$ , establishing bipartiteness of  $Z_n$ . Finally, observe that the degree of  $e \in V(Z_n)$ , an edge between  $u$  and  $v$ , is  $d_1$  in both  $H(u)$  and  $H(v)$  if  $e \in L(Z_n)$ , and  $d_2$  in both if  $e \in R(Z_n)$ , which implies  $(2d_1, 2d_2)$ -biregularity.  $\square$

## 5 Non-backtracking matrix of subgraphs

In this section, we bound the spectral radius of the nonbacktracking matrix of subgraphs of bipartite spectral expanders. This gives us tight control over the degree profile of subgraphs, improving on bounds provided by classic tools like the expander mixing lemma [AC88].

The following is a generalization of an analogous result of Kahale for regular graphs [Kah95, Theorem 1].

**Theorem 5.1** (Formal version of Theorem 1.8). *Let  $\varepsilon \in (0, 0.1)$ , and let  $3 \leq c \leq d$  be integers. Let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph and  $S \subseteq L \cup R$  such that  $|S| \leq d^{-1/\varepsilon}|L \cup R|$ . Then, for any  $t \geq 0$  such that*

$$\frac{1}{t} \geq \frac{1}{2} \left( \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} - \sqrt{d-1})^2} \right), \quad (4)$$

where  $\tilde{\lambda} = \max(\lambda_2(A_G), \sqrt{c-1} + \sqrt{d-1}) \cdot (1 + O(\varepsilon))$ , we have

$$H_{G[S]}(t) \succ 0.$$

To understand Eq. (4), consider  $\tilde{\lambda} \approx \sqrt{c-1} + \sqrt{d-1}$ . Then, Eq. (4) simplifies to  $\frac{1}{t} \gtrsim ((c-1)(d-1))^{1/4}$ . More specifically, suppose  $\tilde{\lambda} = (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$ , then denoting  $\eta :=$

$(\sqrt{c-1} + \sqrt{d-1})^2(\frac{2\gamma}{d} + \frac{\gamma^2}{d^2})$ , Eq. (4) becomes  $\frac{1}{t} \geq \sqrt{\sqrt{(c-1)(d-1)} + \eta/4} + \frac{1}{2}\sqrt{\eta}$ . In particular, we have the following corollary.

**Corollary 5.2.** *Let  $3 \leq c \leq d$ , and suppose  $\tilde{\lambda} = (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$  for  $\gamma \in [0, 2]$ , then*

$$\frac{1}{t^2} \geq \sqrt{(c-1)(d-1)} \cdot (1 + 3\sqrt{\gamma})$$

*implies Eq. (4).*

**Overview of the proof.** We prove  $H_{G[S]}(t) \succ 0$  by showing that  $\langle f, H_{G[S]}(t)f \rangle > 0$  for all  $f : S \rightarrow \mathbb{R}$ . The way we prove  $\langle f, H_{G[S]}(t)f \rangle > 0$  is by relating it to a quadratic form against the matrix  $H_G(t)$ , which we can control via the spectrum of  $G$ . In particular, we consider the depth- $\ell$  regular tree extension  $T$  of  $G[S]$ , and for  $f$  we define an appropriate function extension  $f_t$  on the tree depending on  $t$  (Definition 5.4) such that  $\langle f, H_{G[S]}(t)f \rangle = \langle f_t, H_T(t)f_t \rangle$ . The function  $f_t$  additionally has the property that its  $\ell_2$  mass on vertices  $r$ -far from  $G[S]$  decays exponentially in  $r$ . At a high level, we use the tree extension as a proxy for the  $\ell$ -step neighborhood of  $S$  in  $G$ , and this is made precise in Section 5.3 as we define a natural folded function  $\tilde{f}_t$  of  $f_t$  into  $G$  (Definition 5.9).

This allows us to lower bound  $\langle f, H_{G[S]}(t)f \rangle$  by  $\langle \tilde{f}_t, H_G(t)\tilde{f}_t \rangle$  with some errors. The errors can be bounded using the decay of  $f_t$  from the definition, though this requires  $t < ((c-1)(d-1))^{-1/4}$  (see Lemma 5.7). Ignoring errors, it comes down to showing that

$$\frac{1}{t^4} - \frac{1}{t^2} (\lambda^2 - (c-1) - (d-1)) + (c-1)(d-1) > 0,$$

and we solve the quadratic formula in Lemma 5.12 and show that the above gives rise to Eq. (4). The full proof is presented in Section 5.4.

## 5.1 Tree extensions

We start with defining tree extensions of a graph.

**Definition 5.3** (Tree extension). For a graph  $G = (V, E)$ , we say that  $T = (V(T), E(T))$  is a tree extension of  $G$  if  $T$  is obtained by attaching a tree  $T_r$  to each vertex  $r \in V$ , with  $r$  being the root. Each vertex  $x \in T$  belongs to a unique tree  $T_r$  rooted at  $r$ . For any  $x \in T$ , we write  $\text{depth}(x)$  to be the distance between  $x$  and the root of the tree containing  $x$ .

Fix a tree extension  $T$  of  $G$ , for functions  $f, g : V(T) \rightarrow \mathbb{R}$ , define  $\langle f, g \rangle = \sum_{x \in T} f(x)g(x)$  and  $\|f\|_2^2 = \sum_{x \in T} f(x)^2$ .

**Definition 5.4** (Function extension). Given a function  $f : V(G) \rightarrow \mathbb{R}$ , a tree extension  $T$  of  $G$ , and parameter  $t \in \mathbb{R}$ , we define  $f_t : V(T) \rightarrow \mathbb{R}$  to be the *extension* of  $f$  to  $T$  such that for  $x \in T$ ,

$$f_t(x) = f(r) \cdot t^{\text{depth}(x)}, \quad \text{if } x \in T_r. \quad (5)$$

The following simple but crucial lemma establishes a relationship between  $H_G$  and  $H_T$ , which also motivates the definition of  $f_t$ .

**Lemma 5.5.** *Let  $G$  be a graph and  $T$  be any tree extension of  $G$ . Then, for any  $t \in \mathbb{R}$  and  $f : V(G) \rightarrow \mathbb{R}$ , the extension  $f_t : V(T) \rightarrow \mathbb{R}$  defined in Eq. (5) satisfies*

$$(H_T(t) \cdot f_t)(x) = \begin{cases} (H_G(t) \cdot f)(x) & x \in V(G), \\ 0 & x \notin V(G). \end{cases}$$

*Proof.* Recall that  $H_G(t) = (D_G - \mathbb{1})t^2 - A_G t + \mathbb{1}$ . For  $x \notin V(G)$ , let  $d(x)$  be its degree and let  $r \in V(G)$  be the root of the tree containing  $x$ . Observe that  $x$  has 1 parent (with value  $f(r)t^{\text{depth}(x)-1}$ ) and  $d(x) - 1$  children (with value  $f(r)t^{\text{depth}(x)+1}$ ) in the tree  $T_r$ . Thus,

$$\begin{aligned} (H_T(t) \cdot f_t)(x) &= ((d(x) - 1)t^2 + 1) \cdot f(r)t^{\text{depth}(x)} - t \cdot f(r) \left( t^{\text{depth}(x)-1} + (d(x) - 1)t^{\text{depth}(x)+1} \right) \\ &= 0. \end{aligned}$$

For  $x \in V(G)$ , let  $d_G(x)$  be its degree in  $G$  and  $d_T(x)$  be its degree in  $T$ . Then,  $x$  has  $d_T(x) - d_G(x)$  children (with value  $t \cdot f(x)$ ) in the tree  $T_x$ .

$$\begin{aligned} (H_T(t) \cdot f_t)(x) &= ((d_T(x) - 1)t^2 + 1) \cdot f(x) - t \left( (A_G f)(x) + (d_T(x) - d_G(x)) \cdot t f(x) \right) \\ &= ((d_G(x) - 1)t^2 + 1) \cdot f(x) - t(A_G f)(x) \\ &= (H_G(t) \cdot f)(x). \end{aligned}$$

This completes the proof. □

## 5.2 Regular tree extensions of subgraphs

For a subgraph  $G[S]$  in a regular (or biregular) graph, we consider its regular tree extension.

**Definition 5.6** (Regular tree extension). Let  $G = (V, E)$  be a  $d$ -regular graph,  $S \subseteq V$ ,  $\ell \in \mathbb{N}$ , and consider the induced subgraph  $G[S]$ . We define the *depth- $\ell$  regular tree extension* of  $G[S]$  to be the tree extension  $T$  of  $G[S]$  where depth- $\ell$  trees are attached to vertices in  $S$  such that the resulting graph is  $d$ -regular except for the leaves. Let  $\text{Leaves}(T)$  denote the set of leaves.

Similarly, let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph,  $S \subseteq L \cup R$ , and  $\ell \in \mathbb{N}$ . The *depth- $\ell$  regular tree extension* of  $G[S]$  is the tree extension such that the resulting graph is  $(c, d)$ -biregular except for the leaves.

We show that given a graph  $G = (V, E)$  and  $S \subseteq V$ , for any function  $f : S \rightarrow \mathbb{R}$  and its extension  $f_t$  to the depth- $\ell$  regular tree extension of  $G[S]$ , the contribution from the leaves decays exponentially with  $\ell$  when  $t < ((c - 1)(d - 1))^{-1/4}$ .

**Lemma 5.7** (Decay of  $f_t$ ). Let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph with  $c \leq d$ , let  $S \subseteq L \cup R$ , let  $\ell \in \mathbb{N}$  be even, and let  $T$  be the depth- $\ell$  regular tree extension of  $G[S]$ . Moreover, let  $t \in \mathbb{R}$  such that  $t^2 \sqrt{(c - 1)(d - 1)} = 1 - \delta$  for some  $\delta \in (0, 1)$ . Given any function  $f : S \rightarrow \mathbb{R}$ , let  $f_t : V(T) \rightarrow \mathbb{R}$  be the function extension (as defined in Eq. (5)), and let  $f_t^{\text{leaf}}$  be  $f_t$  restricted to the leaves of  $T$ . Then,

$$\|f_t^{\text{leaf}}\|_2^2 \leq \frac{2\delta}{e^{\delta\ell} - 1} \cdot \|f_t\|_2^2.$$

Since the function  $\frac{2x}{e^{x\ell} - 1}$  is monotone decreasing, we have for any  $\delta' > \delta$ ,

$$\|f_t^{\text{leaf}}\|_2^2 \leq \frac{2\delta'}{e^{\delta'\ell} - 1} \cdot \|f_t\|_2^2.$$

*Proof.* We will lower bound  $\|f_t\|_2^2$  and upper bound the contribution from the leaves at depth  $\ell$ . Fix a vertex  $r \in R$  (with  $\deg_G(r) = d$ ) and consider the tree  $T_r$  rooted at  $r$ . Let  $\deg_{T_r}(r)$  be the degree of



$r$  in  $T_r$ . The number of children of vertices in the tree alternates between  $c - 1$  and  $d - 1$  as we go down the tree. Thus, for an even integer  $k \leq \ell$ , the number of vertices in the  $k$ -th level is

$$\deg_{T_r}(r)(c-1)((c-1)(d-1))^{\frac{k}{2}-1} = \frac{\deg_{T_r}(r)}{\deg_G(r)-1}((c-1)(d-1))^{\frac{k}{2}}. \quad (6)$$

The same argument shows that the above also holds for  $r \in L$  (with  $\deg_G(r) = c$ ). Thus, the contribution of the tree  $T_r$  to  $\|f_t\|_2^2$  can be lower bounded by the product of the following two terms:

$$\sum_{\substack{0 \leq k \leq \ell \\ k \text{ even}}} t^{2k}((c-1)(d-1))^{\frac{k}{2}} = \sum_{i=0}^{\ell/2} (1-\delta)^{2i} = \frac{1-(1-\delta)^{\ell+2}}{1-(1-\delta)^2} \geq \frac{1-(1-\delta)^\ell}{2\delta}.$$

Next, the contribution from the leaves of  $T_r$  to  $\|f_t^{-\ell}\|_2^2$  is also given by [Eq. \(6\)](#). Thus, we have

$$\frac{\|f_t^{-\ell}\|_2^2}{\|f_t\|_2^2} \leq (1-\delta)^\ell \frac{2\delta}{1-(1-\delta)^\ell} \leq \frac{2\delta}{e^{\delta\ell}-1},$$

using  $(1-\delta)^\ell \leq e^{-\delta\ell}$ , finishing the proof.  $\square$

### 5.3 Folding regular tree extensions

Given a regular tree extension  $T$  of an induced subgraph  $G[S]$ , there is a natural folding into  $G$  via breadth-first search from  $S$ .

**Definition 5.8** (Folding into  $G$ ). Let  $G = (V, E)$  be a  $d$ -regular or  $(c, d)$ -biregular graph, let  $S \subseteq V$ , and let  $T$  be the depth- $\ell$  regular tree extension of  $G[S]$ . There is a natural homomorphism  $\sigma : T \rightarrow G$  such that

- $\sigma(x) = x$  for all  $x \in S$ ;
- $\deg_T(x) = \deg_G(\sigma(x))$  for all  $x \in V(T) \setminus \text{Leaves}(T)$ ;
- Two edges  $\{x, y\}$  and  $\{y, z\}$  in  $T$  sharing a vertex are not mapped to the same edge in  $E$ , i.e., all edges in  $T$  that map to the same edge in  $E$  are vertex-disjoint.

**Definition 5.9** (Folded function). Fix a map  $\sigma : T \rightarrow G$ . Given any  $f : V(T) \rightarrow \mathbb{R}$ , we associate each vertex  $v \in G$  with a function  $f^v : V(T) \rightarrow \mathbb{R}$  such that for  $x \in T$ ,

$$f^v(x) = \begin{cases} f(x) & \text{if } \sigma(x) = v, \\ 0 & \text{otherwise.} \end{cases}$$

We define the *folded* function  $\tilde{f} : V(G) \rightarrow \mathbb{R}$  to be

$$\tilde{f}(v) = \|f^v\|_2.$$

**Observation 5.10.** The  $f^v$ 's have disjoint support, thus  $\|\tilde{f}\|_2^2 = \sum_{v \in G} \|f^v\|_2^2 = \|f\|_2^2$ . More generally, let  $\Gamma, \tilde{\Gamma}$  be diagonal operators such that  $(\Gamma f)(x) = \gamma(\deg_G(\sigma(x)))f(x)$  for  $x \in T$  and  $(\tilde{\Gamma}g)(v) = \gamma(\deg_G(v))g(v)$  for  $v \in G$ . Then,  $\langle \tilde{f}, \tilde{\Gamma}\tilde{f} \rangle = \langle f, \Gamma f \rangle$ .

We next prove the following useful lemma that relates the quadratic forms of  $f$  and  $\tilde{f}$  with  $A_G$ .

**Lemma 5.11.** *Let  $G = (V, E)$  be a  $d$ -regular or  $(c, d)$ -biregular graph, let  $S \subseteq V$ , and let  $T$  be a regular tree extension of  $G[S]$ . For any  $f : V(T) \rightarrow \mathbb{R}$  and its folded function  $\tilde{f} : V(G) \rightarrow \mathbb{R}$ , we have*

$$\langle f, A_T f \rangle \leq \langle \tilde{f}, A_G \tilde{f} \rangle.$$

*Proof.* Recall from [Definition 5.8](#) that the map  $\sigma : T \rightarrow G$  satisfies that if  $\{x, y\}$  is an edge in  $T$ , then  $\{\sigma(x), \sigma(y)\} \in E$ . Then,

$$\langle f, A_T f \rangle = 2 \sum_{\{x, y\} \in E(T)} f(x)f(y) = 2 \sum_{\{u, v\} \in E(G)} \sum_{\{x, y\} \in E(T)} \mathbf{1}(\sigma(\{x, y\}) = \{u, v\}) \cdot f^{\sigma(x)}(x)f^{\sigma(y)}(y).$$

Moreover, all edges in  $T$  that map to the same edge are vertex-disjoint. Thus, for any  $\{u, v\} \in E$ ,  $\sum_{\{x, y\} \in E(T)} \mathbf{1}(\sigma(\{x, y\}) = \{u, v\}) \cdot f^{\sigma(x)}(x)f^{\sigma(y)}(y)$  can be expressed as an inner product between some permutations of  $f^u$  and  $f^v$ , which is upper bounded by  $\|f^u\|_2 \cdot \|f^v\|_2 = \tilde{f}(u)\tilde{f}(v)$  by Cauchy-Schwarz. Thus, we have

$$\langle f, A_T f \rangle \leq 2 \sum_{\{u, v\} \in E(G)} \tilde{f}(u)\tilde{f}(v) = \langle \tilde{f}, A_G \tilde{f} \rangle. \quad \square$$

## 5.4 Proof of [Theorem 5.1](#)

Before we prove [Theorem 5.1](#), we first prove the following lemma for convenience.

**Lemma 5.12.** *Let  $3 \leq c \leq d \in \mathbb{N}$  and  $\varepsilon \in (0, 1)$ . Let  $\lambda \geq \sqrt{c-1} + \sqrt{d-1}$  and  $\tilde{\lambda} = \lambda(1 + \varepsilon)$ . Then, for all  $x$  such that*

$$x \geq \frac{1}{2} \left( \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} + \sqrt{d-1})^2} + \sqrt{\tilde{\lambda}^2 - (\sqrt{c-1} - \sqrt{d-1})^2} \right),$$

*we have*

$$x^4 - x^2(\lambda^2(1 + \varepsilon) - (c + d - 2)) + (c - 1)(d - 1) > 0.$$

*Proof.* Denote  $a := c - 1$  and  $b := d - 1$  for convenience. Then, to show that  $x^4 - x^2(\lambda^2(1 + \varepsilon) - a - b) + ab \geq 0$ , it suffices to verify that

$$x^2 > \frac{1}{2} (\lambda^2(1 + \varepsilon) - a - b) + \frac{1}{2} \sqrt{(\lambda^2(1 + \varepsilon) - a - b)^2 - 4ab}.$$

Squaring both sides of  $x \geq \frac{1}{2} \left( \sqrt{\tilde{\lambda}^2 - (\sqrt{a} + \sqrt{b})^2} + \sqrt{\tilde{\lambda}^2 - (\sqrt{a} - \sqrt{b})^2} \right)$ , we get

$$\begin{aligned} x^2 &\geq \frac{1}{2} (\tilde{\lambda}^2 - a - b) + \frac{1}{2} \sqrt{(\tilde{\lambda}^2 - (\sqrt{a} + \sqrt{b})^2)(\tilde{\lambda}^2 - (\sqrt{a} - \sqrt{b})^2)} \\ &= \frac{1}{2} (\tilde{\lambda}^2 - a - b) + \frac{1}{2} \sqrt{\tilde{\lambda}^4 - 2(a + b)\tilde{\lambda}^2 + (a - b)^2} \\ &= \frac{1}{2} (\tilde{\lambda}^2 - a - b) + \frac{1}{2} \sqrt{(\tilde{\lambda}^2 - a - b)^2 - 4ab}, \end{aligned}$$

which completes the proof with  $\tilde{\lambda} = \lambda(1 + \varepsilon)$ . □

*Proof of Theorem 5.1.* We first verify that the assumption on  $t$  (Eq. (4)) implies that

$$t^2 \leq \frac{1 - \varepsilon}{\sqrt{(c-1)(d-1)}}. \quad (7)$$

Indeed, as  $\tilde{\lambda} = \lambda(1 + O(\varepsilon)) \geq (\sqrt{c-1} + \sqrt{d-1})(1 + \varepsilon)$ , Eq. (4) implies that

$$\begin{aligned} \frac{1}{t^2} &\geq \frac{1}{4} \left( (\sqrt{c-1} + \sqrt{d-1})^2 (1 + \varepsilon)^2 - (\sqrt{c-1} - \sqrt{d-1})^2 \right) \\ &\geq \sqrt{(c-1)(d-1)} + \frac{1}{2} (\sqrt{c-1} + \sqrt{d-1})^2 \varepsilon \end{aligned}$$

which implies Eq. (7).

We would like to show that  $\langle f, H_{G[S]}(t)f \rangle > 0$  for any function  $f : S \rightarrow \mathbb{R}$ . Let  $\ell = \lceil \frac{1}{2\varepsilon} \rceil$  be an even integer and let  $T$  be the depth- $\ell$  regular tree extension of  $G[S]$  (Definition 5.6). Let  $f_t : V(T) \rightarrow \mathbb{R}$  be the function extension of  $f$  to  $T$  with parameter  $t$ . By Lemma 5.5, we have

$$\langle f, H_{G[S]}(t)f \rangle = \langle f_t, H_T(t)f_t \rangle = \langle f_t, ((D_T - \mathbb{1})t^2 - tA_T + \mathbb{1})f_t \rangle.$$

Note that all internal vertices  $x \in T \setminus \text{Leaves}(T)$  have degree  $c$  or  $d$  while the leaves have degree 1. Let  $D'_T$  be the diagonal matrix such that the leaves have the ‘‘correct’’ degree, i.e., for  $x \in \text{Leaves}(T)$  in the tree  $T_r$  rooted at  $r \in S$ ,  $D'_T[x, x] = \deg_G(r)$  (since  $\ell$  is even). Then, by Eq. (7), Lemma 5.7 states that  $f_t^{\ell}$  decays with a factor  $\frac{2\varepsilon}{e^{\ell} - 1} \leq 4\varepsilon$ , thus

$$\langle f_t, (D_T - \mathbb{1})f_t \rangle = \langle f_t, (D'_T - \mathbb{1})f_t \rangle - \langle f_t^{\ell}, (D'_T - \mathbb{1})f_t^{\ell} \rangle \geq \langle f_t, (D'_T - \mathbb{1})f_t \rangle (1 - 4\varepsilon).$$

Consider the folded function  $\tilde{f}_t : V(G) \rightarrow \mathbb{R}$  as defined in Definition 5.9. By Observation 5.10, we have  $\langle f_t, D'_T f_t \rangle = \langle \tilde{f}_t, D_G \tilde{f}_t \rangle$  and  $\|f_t\|_2^2 = \|\tilde{f}_t\|_2^2$ . Moreover, by Lemma 5.11,  $\langle f_t, A_T f_t \rangle \leq \langle \tilde{f}_t, A_G \tilde{f}_t \rangle$ . Thus,

$$\begin{aligned} \langle f, H_{G[S]}(t)f \rangle &\geq t^2 \langle \tilde{f}_t, (D_G - \mathbb{1})\tilde{f}_t \rangle (1 - 4\varepsilon) - t \langle \tilde{f}_t, A_G \tilde{f}_t \rangle + \|\tilde{f}_t\|_2^2 \\ &\geq (1 - 4\varepsilon) \langle \tilde{f}_t, (t^2(D_G - \mathbb{1}) + \mathbb{1})\tilde{f}_t \rangle - t \langle \tilde{f}_t, A_G \tilde{f}_t \rangle. \end{aligned} \quad (8)$$

We would like to show that the above is non-negative. Denote  $\Gamma_G := t^2(D_G - \mathbb{1}) + \mathbb{1}$ , and  $\gamma_1 := t^2(c-1) + 1$  and  $\gamma_2 := t^2(d-1) + 1$ . Note that  $\gamma_2 \geq \gamma_1 > 0$  as we assume that  $c \leq d$ . Since  $G$  is a  $(c, d)$ -biregular graph,  $\Gamma_G$  and  $A_G$  have the following block structure,

$$\Gamma_G = \begin{pmatrix} \gamma_1 \mathbb{1} & 0 \\ 0 & \gamma_2 \mathbb{1} \end{pmatrix}, \quad A_G = \begin{pmatrix} 0 & A_{L,R} \\ A_{L,R}^\top & 0 \end{pmatrix}.$$

In particular,

$$(1 - 4\varepsilon)\Gamma_G - tA_G = \Gamma_G^{1/2} \left( (1 - 4\varepsilon)\mathbb{1} - t \cdot \Gamma_G^{-1/2} A_G \Gamma_G^{-1/2} \right) \Gamma_G^{1/2} = \Gamma_G^{1/2} \left( (1 - 4\varepsilon)\mathbb{1} - \frac{t}{\sqrt{\gamma_1 \gamma_2}} A_G \right) \Gamma_G^{1/2}.$$

Then, denoting  $g := \Gamma_G^{1/2} \tilde{f}_t$ , we can write Eq. (8) as

$$\langle f, H_{G[S]}(t)f \rangle \geq \langle \tilde{f}_t, ((1 - 4\varepsilon)\Gamma_G - tA_G)\tilde{f}_t \rangle = (1 - 4\varepsilon)\|g\|_2^2 - \frac{t}{\sqrt{\gamma_1 \gamma_2}} \langle g, A_G g \rangle. \quad (9)$$

Next, we upper bound  $\langle g, A_G g \rangle$ . For any  $(c, d)$ -biregular graph, the (normalized) top eigenvector of  $A_G$  is  $\frac{1}{\sqrt{2|E|}} D_G^{1/2} \vec{1}$  with eigenvalue  $\sqrt{cd}$ . Thus,

$$\langle g, A_G g \rangle \leq \frac{\sqrt{cd}}{2|E|} \langle g, D_G^{1/2} \vec{1} \rangle^2 + \lambda \|g\|_2^2,$$

where  $\lambda = \max(\lambda_2(A_G), \sqrt{c-1} + \sqrt{d-1})$  is the second eigenvalue.

Since  $T$  has depth  $\ell$ , the support of  $\tilde{f}_t$  (and  $g$ ) must be contained in  $B := \{v \in V(G) : \text{dist}(v, S) \leq \ell\}$ . We have  $|B| \leq |S| \sum_{i=0}^{\ell} d^i \leq |S| d^{\ell+1}$ . Thus, by Cauchy-Schwarz,

$$\frac{\sqrt{cd}}{2|E|} \langle g, D_G^{1/2} \vec{1} \rangle^2 \leq \frac{\sqrt{cd}}{2|E|} \cdot d|B| \cdot \|g\|_2^2 \leq d^{-1/4\epsilon} \|g\|_2^2 \leq \epsilon \|g\|_2^2,$$

since  $|S| \leq d^{-1/\epsilon} |L \cup R|$ ,  $\ell = \lceil \frac{1}{2\epsilon} \rceil$ ,  $|E| = c|L| = d|R|$ , and  $\epsilon \leq 0.1$  (note that  $d^{-1/4\epsilon} \leq \epsilon$  for all  $d \geq 3$  and  $\epsilon \leq 0.1$ ).

Thus,  $\langle g, A_G g \rangle \leq (\lambda + \epsilon) \|g\|_2^2 \leq \lambda(1 + \epsilon) \|g\|_2^2$ , and from Eq. (9),

$$\langle f, H_{G[S]}(t)f \rangle \geq \frac{1}{\sqrt{\gamma_1 \gamma_2}} ((1 - 4\epsilon)\sqrt{\gamma_1 \gamma_2} - t\lambda(1 + \epsilon)).$$

As  $\frac{1+\epsilon}{1-4\epsilon} \leq 1 + 5\epsilon$ , to prove that the above is positive, it suffices to prove that  $t^2 \lambda^2 (1 + 5\epsilon) < \gamma_1 \gamma_2 = (t^2(c-1) + 1)(t^2(d-1) + 1)$ , or equivalently,

$$\frac{1}{t^4} - \frac{1}{t^2} (\lambda^2(1 + 5\epsilon) - (c-1) - (d-1)) + (c-1)(d-1) > 0.$$

With  $\tilde{\lambda} = \lambda(1 + 5\epsilon)$  and the assumption on  $t$  (Eq. (4)), the above holds via Lemma 5.12.  $\square$

## 6 Expansion and density of subgraphs

The spectral radius of the nonbacktracking matrix of a subgraph of a bipartite graph imposes constraints on the left and right degrees, articulated by the following.

**Theorem 6.1** (Subgraph density in (near-)Ramanujan graphs; restatement of Theorem 1.11). *Let  $3 \leq c \leq d$  be integers,  $\gamma \in [0, 1]$ , and  $\epsilon \in (0, 0.1)$ . Let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph such that  $\lambda_2(G) \leq (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$ . Then, there exists  $\delta = \delta(\epsilon, c, d) > 0$  such that for every  $S_1 \subseteq L$  and  $S_2 \subseteq R$  with  $|S_1| + |S_2| \leq \delta |L \cup R|$ , the left and right average degrees  $d_1 = \frac{|E(S_1, S_2)|}{|S_1|}$  and  $d_2 = \frac{|E(S_1, S_2)|}{|S_2|}$  in the induced subgraph  $G[S_1 \cup S_2]$  must satisfy*

$$(d_1 - 1)(d_2 - 1) \leq \sqrt{(c-1)(d-1)} \cdot (1 + O(\epsilon + \sqrt{\gamma})).$$

Theorem 6.1 is a direct consequence of Theorem 5.1 and the following lemma:

**Lemma 6.2.** *Let  $G = (L \cup R, E)$  be a bipartite graph, and let the left and right average degrees be  $d_1 = \frac{|E|}{|L|}$  and  $d_2 = \frac{|E|}{|R|}$ , respectively. Then, for any  $t \in (-1, 1) \setminus \{0\}$  such that  $H_G(t) \succeq 0$ , we have*

$$(d_1 - 1)(d_2 - 1) \leq \frac{1}{t^2}.$$

*Proof.* We can assume that  $d_1, d_2 > 1$ , otherwise the statement holds trivially with  $|t| < 1$ . Let  $x$  be the vector such that for  $u \in L \cup R$ ,

$$x_u = \begin{cases} 1 & u \in L, \\ \alpha & u \in R, \end{cases}$$

where  $\alpha \in \mathbb{R}$  will be determined later. Recall that  $H_G(t) = (D_G - \mathbb{1})t^2 - tA_G + \mathbb{1}$ . Since  $|E| = d_1|L| = d_2|R|$ , we have  $x^\top D_G x = d_1|L| + \alpha^2 d_2|R| = (1 + \alpha^2)d_1|L|$  and  $x^\top A_G x = 2\alpha|E| = 2\alpha d_1|L|$ , and substituting  $|R| = \frac{d_1}{d_2}|L|$  we get

$$\begin{aligned} x^\top H_G(t)x &= x^\top \left( (D_G - \mathbb{1})t^2 - tA_G + \mathbb{1} \right) x \\ &= t^2 \left( (d_1 - 1)|L| + \alpha^2(d_2 - 1)|R| \right) - t \cdot 2\alpha d_1|L| + (|L| + \alpha^2|R|) \\ &= |L| \left( (d_1 - 1)t^2 - 2t\alpha d_1 + 1 \right) + |R| \left( \alpha^2(d_2 - 1)t^2 + \alpha^2 \right) \\ &= |L| \left( (d_1 - 1)t^2 + \alpha^2 d_1 t^2 - 2t\alpha d_1 + 1 + \frac{d_1}{d_2} \cdot \alpha^2(1 - t^2) \right). \end{aligned}$$

Then,  $H_G(t) \succeq 0$  and  $t \in (-1, 1)$  imply that

$$\frac{1}{d_2} \geq \frac{1}{1 - t^2} \left( -\frac{(d_1 - 1)t^2 + 1}{d_1 \alpha^2} + \frac{2t}{\alpha} - t^2 \right).$$

To maximize the right-hand side, we choose  $\frac{1}{\alpha} = \frac{d_1 t}{(d_1 - 1)t^2 + 1}$ , which gives

$$\begin{aligned} \frac{1}{d_2} &\geq \frac{1}{1 - t^2} \left( \frac{d_1 t^2}{(d_1 - 1)t^2 + 1} - t^2 \right) = \frac{1}{1 - t^2} \cdot \frac{t^2(d_1 - 1)(1 - t^2)}{(d_1 - 1)t^2 + 1} = \frac{1}{1 + \frac{1}{(d_1 - 1)t^2}} \\ &\implies d_2 \leq 1 + \frac{1}{(d_1 - 1)t^2}. \end{aligned}$$

Rearranging the above gives  $(d_1 - 1)(d_2 - 1) \leq 1/t^2$ .  $\square$

We can now prove [Theorem 6.1](#).

*Proof of Theorem 6.1.* We consider the induced subgraph  $G[S_1 \cup S_2]$ . By [Theorem 5.1](#), we can choose  $t$  such that  $\frac{1}{t^2} \approx \sqrt{(c-1)(d-1)}$  and  $H_{G[S_1 \cup S_2]}(t) \succeq 0$  — specifically, set  $\delta = d^{-d/\varepsilon^2}$  (depending only on  $\varepsilon, d$ ) and  $\tilde{\lambda} = (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d) \cdot (1 + O(\varepsilon^2/d))$ , then by [Corollary 5.2](#),  $H_{G[S_1 \cup S_2]} \succeq 0$  as long as

$$\frac{1}{t^2} \geq \sqrt{(c-1)(d-1)} \cdot \left( 1 + O\left(\sqrt{\gamma + \varepsilon^2}\right) \right) = \sqrt{(c-1)(d-1)} (1 + O(\varepsilon + \sqrt{\gamma})).$$

Plugging the above into [Lemma 6.2](#) completes the proof.  $\square$

As a corollary of [Theorem 6.1](#), we recover the following result of Asherov and Dinur [[AD23](#)] proved for Ramanujan graphs, which further extends to near-Ramanujan graphs.

**Corollary 6.3.** *Let  $3 \leq c \leq d$  be integers,  $\gamma \in [0, 1]$ , and  $\varepsilon \in (0, 0.1)$ . Let  $G = (L \cup R, E)$  be a  $(c, d)$ -biregular graph such that  $\lambda_2(G) \leq (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma/d)$ . Then, there exists  $\delta = \delta(\varepsilon, c, d) > 0$  such that for every  $S \subseteq L$  of size  $|S| \leq \delta|L|$ ,*

$$\frac{c|S|}{|N(S)|} \leq 1 + (1 + O(\varepsilon + \sqrt{\gamma})) \sqrt{\frac{d-1}{c-1}}.$$

## 7 Unique-neighbor expanders with lossless small-set expansion

In this section, we prove [Theorem 1.5](#).

**Theorem** (Restatement of [Theorem 1.5](#)). *For every  $\beta \in (0, 1/2]$  and  $\varepsilon > 0$ , there are constants  $\gamma > 0$  and  $K$  such that for all large enough  $d_1, d_2$  with  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$  that are multiples of  $K$ , there is an explicit infinite family of  $(d_1, d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  where:*

1.  $Z_n$  is a  $\gamma$ -two-sided unique-neighbor expander,
2. every  $S \subseteq L(Z_n)$  with  $|S| \leq \exp(\Omega(\sqrt{\log |V(Z_n)|}))$  has  $(1 - \varepsilon) \cdot d_1 \cdot |S|$  unique-neighbors,
3. every  $S \subseteq R(Z_n)$  with  $|S| \leq \exp(\Omega(\sqrt{\log |V(Z_n)|}))$  has  $(1 - \varepsilon) \cdot d_2 \cdot |S|$  unique-neighbors.

The graphs from [Theorem 1.5](#) are constructed by taking the tripartite line product of a tripartite graph on vertex set  $L \cup M \cup R$  where  $L \cup M$  and  $M \cup R$  are (different) near-Ramanujan biregular graphs with no short bicycles, with suitably chosen degrees. In particular, from the works of [[MOP20](#), [OW20](#)], we have the following.

**Theorem 7.1** (Special case of main theorem in [[OW20](#)]). *For every  $c, d \geq 3$  and  $\gamma > 0$ , there is an explicit construction of an infinite family of  $(c, d)$ -biregular graphs  $(G_n)_{n \geq 1}$  where  $\lambda_2(G_n) \leq (\sqrt{c-1} + \sqrt{d-1}) \cdot (1 + \gamma)$ , and  $G_n$  has no bicycle on  $o(\sqrt{\log |V(G_n)|})$  vertices.*

[Theorem 1.5](#) is then a consequence of choosing the graphs between  $L$  &  $M$ , and  $M$  &  $R$  using the above theorem in conjunction with the forthcoming [Theorem 7.8](#).

**Remark 7.2.** Replacing the above choice of base graph with one equipped with a group action and improved bicycle-free radius would result in a group action for the construction in [Theorem 1.5](#) as well as improved parameters. It is plausible that the construction of [[BFG<sup>+</sup>15](#)] satisfies these properties, but we do not prove it in this work.

### 7.1 Lossless expansion in high-girth graphs

We first define the notion of *bicycles* from [[MOP20](#)].

**Definition 7.3** (Excess). Given a graph  $G = (V, E)$ , its *excess* is  $\text{exc}(G) = |E| - |V|$ . In particular, a graph with excess 0 and 1 is called cyclic and bicyclic respectively.

We will also need the following refinement of the well-known irregular Moore bound for graphs.

**Theorem 7.4** (Generalized Moore bound; formal version of [Theorem 1.12](#)). *Suppose  $G$  is a graph on  $n$  vertices, and let  $\rho = \lambda_1(B_G)$  be the spectral radius of its non-backtracking matrix  $B_G$ . Suppose  $\rho > 1$ , then  $G$  contains a cycle of size at most  $2(\lfloor \log_\rho n \rfloor + 1)$  and a bicycle of size at most  $3(\lfloor \log_\rho 2n \rfloor + 1)$ .*

We defer the proof of [Theorem 7.4](#) to [Appendix A](#). As mentioned in [Remark 1.13](#), for a graph  $G$  with average degree  $d$ ,  $\rho(B_G)$  is at least  $d - 1$ . This follows from the fact that  $\vec{1}^\top H_G(\frac{1}{d-1})\vec{1} = 0$  and [Lemma 2.4](#). Therefore, [Theorem 7.4](#) (the cycle case) is potentially stronger than the girth guarantee of  $2 \log_{d-1} n$  from the classical Moore bound.

Finally, we will need the following statement about the expansion of small sets in graphs with no short cycles or bicycles, which generalizes [[Kah95](#), Theorem 10].



**Lemma 7.5** (Expansion of small sets). *Let  $G = (L \cup R, E)$  be a  $d$ -left-regular bipartite graph, and let  $\varepsilon \in (0, 1)$  such that  $\varepsilon(d-1) > 1$ . Suppose  $G$  has no cycle of length at most  $g$ , then for all  $S \subseteq L$  with  $|S| \leq \frac{1}{d+1}(\varepsilon(d-1))^{\frac{1}{4}g-\frac{1}{2}}$  we have  $|N_G(S)| \geq (1-\varepsilon)d|S|$ .*

*Similarly, suppose  $G$  has no bicycle of length at most  $g$ , then for all  $S \subseteq L$  with  $|S| \leq \frac{1}{2(d+1)}(\varepsilon(d-1))^{\frac{1}{6}g-\frac{1}{2}}$  we have  $|N_G(S)| \geq (1-\varepsilon)d|S|$ .*

*Proof.* Let  $T := N_G(S) \subseteq R$ . Suppose  $S$  does not expand losslessly, i.e.,  $|T| < (1-\varepsilon)d|S|$ . Then, the subgraph  $G[S \cup T]$  must have right average degree at least  $\frac{d|S|}{(1-\varepsilon)d|S|} \geq \frac{1}{1-\varepsilon} \geq 1 + \varepsilon$ . Let  $\rho > 0$  be the spectral radius of the non-backtracking matrix  $B_{G[S \cup T]}$  so that  $H_{G[S \cup T]}(1/\rho) \geq 0$ . Then, applying [Lemma 6.2](#), we have

$$1 + \varepsilon \leq 1 + \frac{\rho^2}{(d-1)} \implies \rho \geq \sqrt{\varepsilon(d-1)}.$$

Next, by [Theorem 7.4](#),  $G[S \cup T]$  must contain a cycle of size at most

$$2 \log_\rho(|S| + |T|) + 2 \leq \frac{2 \log((d+1)|S|)}{\log \sqrt{\varepsilon(d-1)}} + 2.$$

Suppose  $|S| \leq \frac{1}{d+1}(\varepsilon(d-1))^{\frac{1}{4}(g-2)}$ , then there exists a cycle of length at most  $g$ , which is a contradiction.

Similarly, by [Theorem 7.4](#),  $G[S \cup T]$  must contain a bicycle of size at most

$$3 \log_\rho(2(|S| + |T|)) + 3 \leq \frac{3 \log(2(d+1)|S|)}{\log \sqrt{\varepsilon(d-1)}} + 3.$$

Suppose  $|S| \leq \frac{1}{2(d+1)}(\varepsilon(d-1))^{\frac{1}{6}(g-3)}$ , then there exists a bicycle of length at most  $g$ , which is a contradiction.  $\square$

## 7.2 Tripartite line product

We now define a generalization of the line product.

**Definition 7.6** (Tripartite line product). Let  $G = (L \cup M \cup R, E_1 \cup E_2)$  be a tripartite graph consisting of a  $(K_1, D_1)$ -biregular graph  $G^{(1)} = (L \cup M, E_1)$  and a  $(D_2, K_2)$ -biregular graph  $G^{(2)} = (M \cup R, E_2)$ . Let  $H$  be a bipartite graph on vertex set  $[D_1] \cup [D_2]$ . The tripartite line product  $G \diamond H$  is the bipartite graph on vertex set  $L \cup R$  and edges obtained by placing a copy of  $H$  on the neighbors of  $v$  for each  $v \in M$ .

**Remark 7.7.** Note that [Definition 7.6](#) is indeed a generalization of the line product in [Definition 3.1](#) (where  $H$  is bipartite). To see this, consider a  $D$ -regular graph  $G$  and define a tripartite graph  $G'$  as follows: set  $M = V(G)$ , set  $L, R$  to be a partition of  $E(G)$ , and for  $v \in M$  and  $e = L \cup R$ , connect  $\{v, e\}$  if and only if  $v \in e$ . Note that in this case  $K_1 = K_2 = 2$ . If  $L, R$  satisfy that each  $v \in M$  has  $D_1$  neighbors in  $L$  and  $D_2$  neighbors in  $R$  (with  $D = D_1 + D_2$ ), then  $G' \diamond H$  is exactly the same as  $G \diamond H$  in [Definition 3.1](#).

We now prove [Theorem 1.5](#); we use the tripartite line product to construct two-sided unique-neighbor expanders where we can additionally guarantee that small enough sets expand losslessly.

**Theorem 7.8** (Formal version of [Theorem 1.5](#)). *Suppose for some  $c, d \geq 3$  and all  $\gamma > 0$ , there exists an explicit infinite family of  $(c, d)$ -biregular near-Ramanujan graphs  $(F_n)_{n \geq 1}$  with  $\lambda_2(F_n) \leq (\sqrt{c-1} + \sqrt{d-1})(1 + \gamma)$  and has no bicycle of size at most  $g_n = \omega_n(1)$ .*

*For every  $\beta \in (0, 1/2]$  and  $\varepsilon \in (0, 1)$ , there are  $K = K(\varepsilon)$ ,  $d_0 = d_0(\beta, \varepsilon)$  and  $\delta = \delta(\beta, \varepsilon) > 0$  such that for all  $d_1, d_2 \geq d_0$  which are multiples of  $K$  and satisfy  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$ , the following holds: there exists  $\mu = \mu(\beta, \varepsilon, d_1, d_2) > 0$  such that there is an infinite family of  $(d_1, d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  where*

$$|\text{UN}_{Z_n}(S)| \geq \delta \cdot d_1 \cdot |S|, \quad \forall S \subseteq V(Z_n) \text{ with } |S| \leq \mu \cdot |V(Z_n)|.$$

Moreover,

$$|N_{Z_n}(S)| \geq (1 - \varepsilon) \cdot d_1 \cdot |S|, \quad \forall S \subseteq L(Z_n) \text{ with } |S| \leq \exp(g_n),$$

$$|N_{Z_n}(S)| \geq (1 - \varepsilon) \cdot d_2 \cdot |S|, \quad \forall S \subseteq R(Z_n) \text{ with } |S| \leq \exp(g_n).$$

*Proof.* The construction of  $Z_n$  is based on taking the tripartite line product of some  $G_n$  and a bipartite gadget graph  $H$  from [Lemma 4.3](#) with suitably chosen parameters.

- Let  $K = K_1 = K_2 = \frac{1000}{\varepsilon^2}$ .
- Let  $d_0 \geq C_0 K$  where  $C_0 = C_0(\varepsilon, \beta)$  is some large enough constant chosen later.
- Let  $\tilde{d}_1 := d_1/K$  and  $\tilde{d}_2 := d_2/K$ , both at least  $C_0$ .
- Let  $\theta := \frac{C}{\varepsilon} \sqrt{K/\beta}$  (depending only on  $\varepsilon, \beta$ ) where  $C$  is a universal constant.
- Let  $D_1 := \left\lceil \frac{\tilde{d}_1 + \tilde{d}_2}{\theta^2} \right\rceil \cdot \tilde{d}_2$  and  $D_2 := \left\lceil \frac{\tilde{d}_1 + \tilde{d}_2}{\theta^2} \right\rceil \cdot \tilde{d}_1$ , and define  $D := D_1 + D_2$ .

Note that  $1 \geq \tilde{d}_1/\tilde{d}_2 = d_1/d_2 \geq \frac{\beta}{1-\beta}$ . One can verify that  $KD_1 \leq 2\varepsilon^2 \tilde{d}_2^2 / C^2$  and  $KD_2 \leq 2\varepsilon^2 \tilde{d}_1^2 / C^2$ .

The above choice of parameters satisfy the requirements of [Lemma 4.3](#). Thus, applying [Lemma 4.3](#) with parameters  $\theta$  and  $\tau = 1$ , there is a  $(\tilde{d}_1, \tilde{d}_2)$ -biregular graph  $H$  with  $D_1$  left vertices and  $D_2$  right vertices such that

$$\min_{S \subseteq V(H): 1 \leq |S| \leq t} \frac{|\text{UN}_H(S)|}{|S|} \geq (1 - o_D(1)) \cdot \tilde{d}_1 \cdot \exp(-\theta t / \sqrt{D}) \quad (10)$$

for  $1 \leq t \leq \sqrt{D}$ . We can set  $C_0$  large enough (depending only on  $\beta, \theta$  which only depend on  $\varepsilon, \beta$ ) such that the  $o_D(1)$  term is at most 0.1.

For the tripartite base graph  $G_n = (L \cup M \cup R, E_1 \cup E_2)$ , we construct  $G_n^{(1)} = (L \cup M, E_1)$  and  $G_n^{(2)} = (M \cup R, E_2)$  to be  $(K_1, D_1)$  and  $(D_2, K_2)$ -biregular near-Ramanujan graphs respectively, i.e.,  $\lambda_2(G_n^{(1)}) \leq (\sqrt{K_1-1} + \sqrt{D_1-1})(1 + 0.01/D_1)$  and  $\lambda_2(G_n^{(2)}) \leq (\sqrt{K_2-1} + \sqrt{D_2-1})(1 + 0.01/D_2)$ , along with the small-bicycle-free assumption.

**Unique-neighbor expansion.** Next, we analyze the vertex expansion of a subset  $S \subseteq L(Z_n)$  in the product graph  $Z_n$ . Recall that  $L(Z_n) = L$ . Let  $U := N_{G_n}(S) \subseteq M$  be the neighbors of  $S$  in  $G_n^{(1)}$ , and we partition  $U$  into  $U_\ell := \{v \in U : |E_1(v, S)| \leq \sqrt{D}\}$  (the ‘‘low  $S$ -degree’’ vertices) and  $U_h := U \setminus U_\ell$  (the ‘‘high  $S$ -degree’’ vertices). Consider the bipartite subgraph induced by  $S \cup U_h$ . By definition, the average right-degree in  $G_n^{(1)}[S \cup U_h]$  is at least  $\sqrt{D}$ . By the upper bound on  $\lambda_2(G_n^{(1)})$ , we can apply [Theorem 6.1](#) and bound the average left-degree by

$$d_{\text{left}}(S, U_h) \leq 1 + \frac{\sqrt{(K_1-1)(D_1-1)}}{\sqrt{D}-1} \cdot 1.1 \leq 1 + 1.2\sqrt{K},$$

as long as  $|S| \leq \mu|L|$  for some  $\mu = \mu(K, D_1) > 0$  (depending only on  $\varepsilon, \beta, d_1, d_2$ ). For any  $K \geq 100$ , the above is at most  $0.2K$ . Thus, we know that  $|E_1(S, U_\ell)| \geq 0.8K|S|$ , i.e., a constant fraction of edges incident to  $S$  go to  $U_\ell$ . This also implies that  $|U_h| \leq 0.2|U|$ .

For each  $v \in U$ , let  $S_v \subseteq S$  be the vertices in  $S$  incident to  $v$ . Consider the gadget  $H$  placed on  $v$ , and let  $T_v \subseteq R$  be the set of unique-neighbors of  $S_v$  in the gadget. Further, let  $\tilde{T} := \bigcup_{v \in U} T_v$ . Note that each vertex in  $\tilde{T}$  is a unique-neighbor *within some gadget*, but there may be edges coming from other gadgets, so not all of  $\tilde{T}$  are unique-neighbors of  $S$  in the final product graph. Our goal is to show that a large fraction of  $\tilde{T}$  are unique-neighbors of  $S$ .

We will analyze the induced subgraph  $G_n^{(2)}[U \cup \tilde{T}]$ , and we claim that a large fraction of  $\tilde{T}$  are unique-neighbors of  $U$  in  $G_n^{(2)}$ , thus are also unique-neighbors of  $S$  in  $Z_n$ . We first lower bound the left average degree of  $G_n^{(2)}[U \cup \tilde{T}]$ . For each  $v \in U_\ell$ , we have  $1 \leq |S_v| \leq \sqrt{D}$ , and by the expansion profile of the gadget (Eq. (10)),  $v$  has degree at least

$$|S_v| \cdot 0.9 \cdot \tilde{d}_1 \cdot \exp(-\theta|S_v|/\sqrt{D}) \geq 0.9 \cdot \tilde{d}_1 \cdot \min \left\{ e^{-\theta/\sqrt{D}}, \sqrt{D}e^{-\theta} \right\}$$

in  $G_n^{(2)}[U \cup \tilde{T}]$ . Since  $\theta$  depends only on  $\varepsilon, \beta$ , we choose  $C_0 = C_0(\varepsilon, \beta)$  to be large enough (thus also  $D$ ) such that the above is at least  $0.8 \cdot \tilde{d}_1$ .

Next, for  $v \in U_h$ , we have no control over its degree in  $G_n^{(2)}[U \cup \tilde{T}]$ . However, since  $|U_h| \leq 0.2|U| \leq \frac{1}{4}|U_\ell|$ , we have

$$d_{\text{left}}(U, \tilde{T}) \geq \frac{0.8\tilde{d}_1 \cdot |U_\ell|}{|U_\ell| + |U_h|} \geq 0.64 \cdot \tilde{d}_1.$$

Then, for  $|S| \leq \mu|L|$  where  $\mu$  is small enough, we have  $|U| \leq \mu'|M|$  where  $\mu'$  (depending on  $\varepsilon, \beta, K, D_2$ ) is small enough to apply [Theorem 6.1](#) and conclude that the right average degree

$$d_{\text{right}}(U, \tilde{T}) \leq 1 + \frac{\sqrt{(K_2 - 1)(D_2 - 1)}}{0.64 \cdot \tilde{d}_1 - 1} \cdot O(1) \leq 1.1,$$

since  $\tilde{d}_1 + \tilde{d}_2 \leq \frac{1}{\beta}\tilde{d}_1$  and  $K_2 D_2 \leq \tilde{d}_1^2 / C$  with some large  $C$  by our choice of  $\theta$  and  $D_2$ . This implies that 0.9 fraction of  $\tilde{T}$  are unique-neighbors of  $S$ .

Finally, we lower bound  $|E_2(U, \tilde{T})|$ . Again by [Eq. \(10\)](#),

$$\begin{aligned} |E_2(U, \tilde{T})| &\geq \sum_{v \in U_\ell} |S_v| \cdot 0.9 \cdot \tilde{d}_1 \cdot \exp(-\theta|S_v|/\sqrt{D}) \geq 0.9 \cdot \tilde{d}_1 \cdot \exp(-\theta) \sum_{v \in U_\ell} |S_v| \\ &= 2\delta \cdot \tilde{d}_1 \cdot |E_1(S, U_\ell)| \geq 1.6\delta d_1 |S|, \end{aligned}$$

where  $\delta = \delta(\varepsilon, \beta) > 0$ . The last inequality uses the fact that  $|E_1(S, U_\ell)| \geq 0.8K|S|$  and  $d_1 = K\tilde{d}_1$ . With  $d_{\text{right}}(U, \tilde{T}) \leq 1.1$ , it follows that

$$|\text{UN}_{Z_n}(S)| \geq \delta d_1 |S|.$$

For  $S \subseteq R(Z_n)$ , the analysis is completely symmetric. Indeed, we have  $K_1 = K_2$ , and we can verify that  $K_1 D_1 \leq \tilde{d}_2^2 / C$ . Since  $d_1 \leq d_2$ , the unique-neighbor lower bound holds for all  $S \subseteq V(Z_n)$  with  $|S| \leq \mu|V(Z_n)|$ .

**Small set lossless expansion.** We now turn to the expansion of small subsets  $S \subseteq L(Z_n)$ . Let  $U := N_{G_n^{(1)}}(S) \subseteq M$  and  $T := N_{Z_n}(S) \subseteq R$ . By assumption,  $G_n^{(1)}$  has no bicycle of size at most  $g_n$ , thus [Lemma 7.5](#) states that  $|U| \geq (1 - \varepsilon/2)K|S|$  (i.e.,  $S$  expands losslessly in  $G_n^{(1)}$ ) assuming that

$$|S| \leq \frac{1}{2(K+1)} \left( \frac{\varepsilon}{2}(K-1) \right)^{\frac{1}{6}g_n - \frac{1}{2}}.$$

With our choice of  $K$  and  $g_n = \omega_n(1)$ , it suffices that  $|S| \leq \exp(g_n)$ .

Next, as each gadget on  $v \in U$  expands with a factor of at least  $\tilde{d}_1$ , we can lower bound  $|E_2(U, T)|$  by  $\tilde{d}_1 \cdot |U|$ . Moreover, the left average degree of the induced subgraph  $G_n^{(2)}[U \cup T]$  is at least  $\tilde{d}_1$ . Then, by [Theorem 6.1](#), the right average degree of  $G_n^{(2)}[U \cup T]$  is

$$d_{\text{right}}(U, T) \leq 1 + \frac{\sqrt{(K_2 - 1)(D_2 - 1)}}{\tilde{d}_1 - 1} \cdot O(1) \leq 1 + \frac{\varepsilon}{2},$$

given that  $K_2 D_2 \leq \varepsilon^2 \tilde{d}_1^2 / C$  for a large enough constant  $C$ . Thus, since  $|E_2(U, T)| \geq \tilde{d}_1 \cdot |U| \geq (1 - \varepsilon/2) \cdot K \tilde{d}_1 \cdot |S| = (1 - \varepsilon/2) d_1 |S|$ ,

$$|T| \geq \frac{1}{1 + \varepsilon/2} |E_2(U, T)| \geq \frac{1 - \varepsilon/2}{1 + \varepsilon/2} \cdot d_1 |S| \geq (1 - \varepsilon) d_1 |S|.$$

For  $S \subseteq R(Z_n)$ , the analysis is symmetric with  $d_1, \tilde{d}_1$  replaced by  $d_2, \tilde{d}_2$ . □

## 8 One-sided lossless expanders

In this section, we illustrate how the construction of Golowich [[Gol23](#)] can be instantiated using the tripartite line product, and give a succinct proof of lossless expansion using our results on subgraphs of spectral expanders.

**Theorem 8.1 (One-sided lossless expanders).** *For every  $\beta \in (0, 1/2]$  and  $\varepsilon \in (0, 1)$ , there are  $K = K(\varepsilon)$ ,  $d_0 = d_0(\beta, \varepsilon)$  such that for all  $d_1, d_2 \geq d_0$  which are multiples of  $K$  and satisfy  $1 \geq \frac{d_1}{d_2} \geq \frac{\beta}{1-\beta}$ , the following holds: there exists  $\mu = \mu(\beta, \varepsilon, d_1, d_2) > 0$  such that there is an infinite family of  $(d_1, d_2)$ -biregular graphs  $(Z_n)_{n \geq 1}$  where*

$$|N_{Z_n}(S)| \geq (1 - \varepsilon) \cdot d_1 \cdot |S|, \quad \forall S \subseteq L(Z_n), |S| \leq \mu |L(Z_n)|.$$

*Proof.* The construction of  $Z_n$  is the same as [Theorem 7.8](#), except this time we take  $K_1 \gg K_2 = 1$  and  $\theta = 1$ .

- Let  $K_1 = \left\lceil \frac{2^{10}}{\varepsilon^4} \right\rceil$  and  $K_2 = 1$ .
- Let  $d_0 \geq C_0 K_1$  where  $C_0 = C_0(\varepsilon, \beta)$  is some large enough constant chosen later.
- Let  $\tilde{d}_1 := d_1 / K_1$  and  $\tilde{d}_2 := d_2$ , both at least  $C_0$ .
- Let  $\theta := 1$ .
- Let  $D_1 := (\tilde{d}_1 + \tilde{d}_2) \cdot \tilde{d}_2$  and  $D_2 := (\tilde{d}_1 + \tilde{d}_2) \cdot \tilde{d}_1$ , and define  $D := D_1 + D_2$ .

Since  $d_1 \leq d_2$ , we have  $\tilde{d}_1 \leq \tilde{d}_2$ . This choice of parameters satisfy the requirements of [Lemma 4.3](#). Thus, applying [Lemma 4.3](#) with parameters  $\theta = \tau = 1$ , there is a  $(\tilde{d}_1, \tilde{d}_2)$ -biregular graph  $H$  with  $D_1$  left vertices and  $D_2$  right vertices such that

$$\min_{S \subseteq V(H); |S| \leq t} \frac{|N_H(S)|}{|S|} \geq (1 - o_D(1)) \cdot \tilde{d}_1 \cdot \exp(-t/\sqrt{D}) \quad (11)$$

for  $1 \leq t \leq \sqrt{D}$ . We can set  $C_0$  large enough (depending only on  $\varepsilon, \beta$ ) such that the  $o_D(1)$  term is at most  $\varepsilon/8$ .

For the tripartite base graph  $G_n = (L \cup M \cup R, E_1 \cup E_2)$ , we construct  $G_n^{(1)} = (L \cup M, E_1)$  to be a  $(K_1, D_1)$ -biregular near-Ramanujan graph, and  $G_n^{(2)} = (M \cup R, E_2)$  is simply a  $(D_2, 1)$ -biregular graph.

Next, we analyze the vertex expansion of a subset  $S \subseteq L(Z_n)$  in the product graph  $Z_n$ . Similar to the proof of [Theorem 7.8](#), let  $U := N_{G_n}(S) \subseteq M$  be the neighbors of  $S$  in  $G_n^{(1)}$ , and we partition  $U$  into  $U_\ell := \{v \in U : |E_1(v, S)| \leq \frac{\varepsilon}{4}\sqrt{D}\}$  (the “low  $S$ -degree” vertices) and  $U_h := U \setminus U_\ell$  (the “high  $S$ -degree” vertices). Consider the bipartite subgraph induced by  $S \cup U_h$ . By definition, the right average degree in  $G_n^{(1)}[S \cup U_h]$  is at least  $\frac{\varepsilon}{4}\sqrt{D}$ . By [Theorem 6.1](#), we can bound the left average degree by

$$d_{\text{left}}(S, U_h) \leq 1 + \frac{\sqrt{(K_1 - 1)(D_1 - 1)}}{\frac{\varepsilon}{4}\sqrt{D} - 1} \cdot 1.1 \leq \frac{8}{\varepsilon} \sqrt{K_1}$$

as long as  $|S| \leq \mu|L|$  for some  $\mu = \mu(K_1, D_1) > 0$  (depending only on  $\varepsilon, \beta, d_1, d_2$ ). Since  $K_1 \geq \frac{2^{10}}{\varepsilon^4}$ , the above is at most  $\frac{\varepsilon}{4}K_1$ . Thus, we know that  $|E_1(S, U_\ell)| \geq (1 - \varepsilon/4)K_1|S|$ , i.e., most edges incident to  $S$  go to  $U_\ell$ .

Moreover, for each  $v \in U_\ell$ , the gadget placed on  $H$  has at most  $\frac{\varepsilon}{4}\sqrt{D}$  vertices on the left, thus by [Eq. \(11\)](#) each gadget expands losslessly. Specifically, denoting  $T := N_{Z_n}(S) \subseteq R$ , we have that

$$|E_2(v, T)| \geq |E_1(v, S)| \cdot (1 - o_D(1)) \cdot \tilde{d}_1 \cdot \exp(-\varepsilon/4) \geq (1 - \varepsilon/2) \cdot \tilde{d}_1 \cdot |E_1(v, S)|$$

for large enough  $C_0$  (hence large enough  $D$ ).

Finally, since  $G_n^{(2)}$  is a  $(D_2, 1)$ -biregular graph,

$$|T| = \sum_{v \in U} |E_2(v, T)| \geq (1 - \varepsilon/2) \cdot \tilde{d}_1 \cdot |E_1(S, U_\ell)| \geq (1 - \varepsilon) \cdot d_1 \cdot |S|.$$

This completes the proof. □

## References

- [ABN23] Anurag Anshu, Nikolas P Breuckmann, and Chinmay Nirkhe. NLTS Hamiltonians from good quantum codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 1090–1096, 2023. [5](#)
- [AC88] Noga Alon and Fan RK Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988. [11](#), [16](#)
- [AC02] Noga Alon and Michael Capalbo. Explicit unique-neighbor expanders. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 73–79. IEEE, 2002. [0](#), [1](#), [2](#), [5](#), [6](#), [7](#)
- [AD23] Ron Asherov and Irit Dinur. Bipartite unique-neighbour expanders via Ramanujan graphs. *arXiv preprint arXiv:2301.03072*, 2023. [0](#), [1](#), [2](#), [3](#), [5](#), [6](#), [7](#), [8](#), [23](#)
- [AHL02] Noga Alon, Shlomo Hoory, and Nathan Linial. The Moore bound for irregular graphs. *Graphs and Combinatorics*, 18(1):53–57, 2002. [3](#), [4](#), [34](#)
- [ALM96] Sanjeev Arora, FT Leighton, and Bruce M Maggs. On-Line Algorithms for Path Selection in a Nonblocking Network. *SIAM Journal on Computing*, 25(3):600–625, 1996. [1](#), [5](#)
- [Bas92] Hyman Bass. The Ihara-Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(06):717–797, 1992. [12](#)
- [Bec16] Oren Becker. Symmetric unique neighbor expanders and good LDPC codes. *Discrete Applied Mathematics*, 211:211–216, 2016. [6](#)
- [BFG<sup>+</sup>15] Cristina Ballantine, Brooke Feigon, Radhika Ganapathy, Janne Kool, Kathrin Maurischat, and Amy Wooding. Explicit construction of Ramanujan bigraphs. In *Women in Numbers Europe: Research Directions in Number Theory*, pages 1–16. Springer, 2015. [2](#), [24](#)
- [BGI<sup>+</sup>08] Radu Berinde, Anna C Gilbert, Piotr Indyk, Howard Karloff, and Martin J Strauss. Combining geometry and combinatorics: A unified approach to sparse signal recovery. In *2008 46th Annual Allerton Conference on Communication, Control, and Computing*, pages 798–805. IEEE, 2008. [1](#), [5](#)
- [BV09] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5(1):239–255, 2009. [1](#), [5](#)
- [CRTS23] Itay Cohen, Roy Roth, and Amnon Ta-Shma. HDX condensers. In *2023 IEEE 64th Annual Symposium on Foundations of Computer Science (FOCS)*. IEEE, 2023. [1](#), [2](#), [5](#), [6](#)
- [CRVW02] Michael Capalbo, Omer Reingold, Salil Vadhan, and Avi Wigderson. Randomness conductors and constant-degree lossless expanders. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing*, pages 659–668, 2002. [1](#), [2](#), [5](#), [6](#)
- [DEL<sup>+</sup>22] Irit Dinur, Shai Evra, Ron Livne, Alexander Lubotzky, and Shahar Mozes. Locally Testable Codes with constant rate, distance, and locality. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 357–374, 2022. [4](#)



- [DFHT21] Irit Dinur, Yuval Filmus, Prahladh Harsha, and Madhur Tulsiani. Explicit sos lower bounds from high-dimensional expanders. In *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, 2021. [5](#)
- [DFRŠ17] Andrzej Dudek, Alan Frieze, Andrzej Ruciński, and Matas Šileikis. Embedding the Erdős–Rényi hypergraph into the random regular hypergraph and Hamiltonicity. *Journal of Combinatorial Theory, Series B*, 122:719–740, 2017. [37](#)
- [DHLV22] Irit Dinur, Min-Hsiu Hsieh, Ting-Chun Lin, and Thomas Vidick. Good Quantum LDPC Codes with Linear Time Decoders. *arXiv preprint arXiv:2206.07750*, 2022. [5](#)
- [DSW06] Irit Dinur, Madhu Sudan, and Avi Wigderson. Robust local testability of tensor products of LDPC codes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques.*, pages 304–315. Springer, 2006. [1](#), [5](#)
- [FK16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016. [9](#), [37](#), [40](#)
- [FM17] Zhou Fan and Andrea Montanari. How well do local algorithms solve semidefinite programs? In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 604–614, 2017. [12](#)
- [GLR10] Venkatesan Guruswami, James R Lee, and Alexander Razborov. Almost Euclidean subspaces of  $\ell_1^N$  via expander codes. *Combinatorica*, 30(1):47–68, 2010. [1](#), [5](#)
- [GMM22] Venkatesan Guruswami, Peter Manohar, and Jonathan Mosheiff.  $\ell_p$ -Spread and Restricted Isometry Properties of Sparse Random Matrices. In *37th Computational Complexity Conference (CCC 2022)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2022. [1](#), [5](#)
- [Gol23] Louis Golowich. New Explicit Constant-Degree Lossless Expanders. *arXiv preprint arXiv:2306.07551*, 2023. [1](#), [2](#), [5](#), [6](#), [7](#), [8](#), [9](#), [10](#), [28](#)
- [GPT22] Shouzhen Gu, Christopher A Pattison, and Eugene Tang. An efficient decoder for a linear distance quantum LDPC code. *arXiv preprint arXiv:2206.06557*, 2022. [5](#)
- [Gri01] Dima Grigoriev. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1-2):613–622, 2001. [5](#)
- [GUV09] Venkatesan Guruswami, Christopher Umans, and Salil Vadhan. Unbalanced expanders and randomness extractors from Parvaresh–Vardy codes. *Journal of the ACM (JACM)*, 56(4):1–34, 2009. [5](#)
- [Has89] Ki-ichiro Hashimoto. Zeta functions of finite graphs and representations of p-adic groups. In *Automorphic forms and geometry of arithmetic varieties*, pages 211–280. Elsevier, 1989. [12](#)
- [HKM23] Jun-Ting Hsieh, Pravesh K Kothari, and Sidhanth Mohanty. A simple and sharper proof of the hypergraph Moore bound. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2324–2344. SIAM, 2023. [34](#)

- [HL22] Max Hopkins and Ting-Chun Lin. Explicit lower bounds against  $\omega$  (n)-rounds of sum-of-squares. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 662–673. IEEE, 2022. [5](#)
- [HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. [10](#)
- [Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963. [43](#)
- [Iha66] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over p-adic fields. *Journal of the Mathematical Society of Japan*, 18(3):219–235, 1966. [12](#)
- [Kah95] Nabil Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM (JACM)*, 42(5):1091–1106, 1995. [0](#), [3](#), [4](#), [16](#), [24](#)
- [Kar11] Zohar S Karnin. Deterministic construction of a high dimensional  $\ell_p$  section in  $\ell_1^n$  for any  $p < 2$ . In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 645–654, 2011. [1](#), [5](#)
- [Kit03] A Yu Kitaev. Fault-tolerant quantum computation by anyons. *Annals of physics*, 303(1):2–30, 2003. [4](#)
- [KK22] Amitay Kamber and Tali Kaufman. Combinatorics via closed orbits: number theoretic Ramanujan graphs are not unique neighbor expanders. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 426–435, 2022. [4](#)
- [KRZS23] Swastik Kopparty, Noga Ron-Zewi, and Shubhangi Saraf. Simple constructions of unique neighbor expanders from error-correcting codes. *arXiv preprint arXiv:2310.19149*, 2023. [5](#), [6](#)
- [LH22] Ting-Chun Lin and Min-Hsiu Hsieh. Good quantum LDPC codes with linear time decoder from lossless expanders. *arXiv preprint arXiv:2203.03581*, 2022. [1](#), [4](#), [9](#)
- [LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. [2](#), [6](#), [14](#)
- [LZ22] Anthony Leverrier and Gilles Zémor. Quantum tanner codes. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 872–883. IEEE, 2022. [1](#), [4](#)
- [LZ23] Anthony Leverrier and Gilles Zémor. Efficient decoding up to a constant fraction of the code length for asymptotically good quantum codes. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1216–1244. SIAM, 2023. [5](#)
- [Mar88] Grigorii Aleksandrovich Margulis. Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators. *Problemy peredachi informatsii*, 24(1):51–60, 1988. [2](#)

- [MM21] Theo McKenzie and Sidhanth Mohanty. High-Girth Near-Ramanujan Graphs with Lossy Vertex Expansion. In *48th International Colloquium on Automata, Languages, and Programming (ICALP 2021)*. Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2021. [3](#), [4](#)
- [MOP20] Sidhanth Mohanty, Ryan O’Donnell, and Pedro Paredes. Explicit near-Ramanujan graphs of every degree. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 510–523, 2020. [2](#), [24](#)
- [Mor94] Moshe Morgenstern. Existence and explicit constructions of  $q+1$  regular Ramanujan graphs for every prime power  $q$ . *Journal of Combinatorial Theory, Series B*, 62(1):44–62, 1994. [2](#), [6](#), [14](#)
- [OW20] Ryan O’Donnell and Xinyu Wu. Explicit near-fully X-Ramanujan graphs. In *2020 IEEE 61st Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1045–1056. IEEE, 2020. [2](#), [24](#)
- [PK22] Pavel Panteleev and Gleb Kalachev. Asymptotically good quantum and locally testable classical LDPC codes. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing*, pages 375–388, 2022. [1](#), [4](#)
- [RVW00] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders and extractors. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 3–13. IEEE, 2000. [5](#)
- [Sch08] Grant Schoenebeck. Linear level lasserre lower bounds for certain  $k$ -csp. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602. IEEE, 2008. [5](#)
- [SS96] Michael Sipser and Daniel Spielman. Expander codes. *IEEE Trans. Inform. Theory*, 42(6, part 1):1710–1722, 1996. [1](#)
- [Tan81] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on information theory*, 27(5):533–547, 1981. [5](#)
- [TSUZ07] Amnon Ta-Shma, Christopher Umans, and David Zuckerman. Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27:213–240, 2007. [5](#)

## A Generalization of the Moore bound

In this section, we strengthen the classical Moore bound of Alon, Hoory and Linial [AHL02] and generalize the result to *bicycles* (recall Definition 7.3). Our proof closely follows Section A of [HKM23], which is an alternative proof of the Moore bound.

**Theorem** (Restatement of Theorem 7.4). *Suppose  $G$  is a graph on  $n$  vertices, and let  $\rho = \lambda_1(B_G)$  be the spectral radius of its non-backtracking matrix  $B_G$ . Suppose  $\rho > 1$ , then  $G$  contains a cycle of size at most  $2(\lfloor \log_\rho n \rfloor + 1)$  and a bicycle of size at most  $3(\lfloor \log_\rho 2n \rfloor + 1)$ .*

The proof of Theorem 7.4 is based on *non-backtracking walks*, which are walks such that no edge is the inverse of its preceding edge. For a graph  $G$  on  $n$  vertices with adjacency matrix  $A$ , we define  $A^{(s)}$  to be the  $n \times n$  matrix whose  $(u, v)$  entry counts the number of length- $s$  non-backtracking walks between vertices  $u$  and  $v$  in  $G$ . The following is a standard fact.

**Fact A.1** (Recurrence and generating function of  $A^{(s)}$ ). *The non-backtracking matrices  $A^{(s)}$  satisfy the following recurrence:*

$$\begin{aligned} A^{(0)} &= \mathbb{1}, \\ A^{(1)} &= A, \\ A^{(2)} &= A^2 - D, \\ A^{(s)} &= A^{(s-1)}A - A^{(s-2)}(D - \mathbb{1}), \quad s > 2. \end{aligned}$$

The recurrences imply that these matrices have a generating function:

$$J(t) := \sum_{s=0}^{\infty} A^{(s)} t^s = (1 - t^2) \cdot H(t)^{-1}$$

for  $t \in [0, 1)$  whenever the series converges, where we recall that  $H(t) = (D - \mathbb{1})t^2 - At + \mathbb{1}$ .

We first state the following simple lemma from [HKM23],

**Lemma A.2.** *Let  $s, k \in \mathbb{N}$ ,  $s \geq k$ , and let  $q, r$  be the quotient and remainder of  $s$  divided by  $k$ , i.e.  $s = qk + r$ . Then,*

$$\mathrm{tr}(A^{(s)}) \leq \sqrt{n} \cdot \|A^{(k)}\|_2^q \cdot \|A^{(r)}\|_F.$$

With Fact A.1 and Lemma A.2, we now prove Theorem 7.4 by analyzing the convergence of  $J(t)$  as  $t$  increases from 0.

*Proof of Theorem 7.4.* Let  $A$  be the adjacency matrix of  $G$  with average degree  $d > 2$ , let  $D$  be the diagonal degree matrix  $G$ , and let  $\rho = \lambda_1(B_G)$ . We will analyze the convergence of  $\mathrm{tr}(J(t)) = \sum_{s=0}^{\infty} \mathrm{tr}(A^{(s)})t^s$  as  $t$  increase from 0 to  $1/\rho$ . In particular, by Lemma 2.4 we have that  $H_G(t) \succ 0$  (thus  $\mathrm{tr}(J(t)) < \infty$ ) for all  $t \in [0, 1/\rho)$ , and  $\mathrm{tr}(J(1/\rho))$  diverges.

Fix  $k \in \mathbb{N}$ . For each  $s \in \mathbb{N}$  we can write  $s = qk + r$ , and

$$J(t) = \sum_{s=0}^{\infty} A^{(s)} t^s = \sum_{r=0}^{k-1} \sum_{q=0}^{\infty} A^{(qk+r)} t^{qk+r}.$$

By [Lemma A.2](#), we have

$$\mathrm{tr}(J(t)) \leq \sum_{r=0}^{k-1} t^r \sqrt{n} \|A^{(r)}\|_F \sum_{q=0}^{\infty} \|A^{(k)}\|_2^q \cdot t^{qk} = \sum_{r=0}^{k-1} t^r \sqrt{n} \|A^{(r)}\|_F \sum_{q=0}^{\infty} \left( \|A^{(k)}\|_2 \cdot t^k \right)^q. \quad (12)$$

Now, let  $k := \lfloor \log_{\rho} n \rfloor + 1$  and suppose for contradiction that  $G$  contains no cycle of size  $\leq \ell = 2k$ . Observe that every entry of  $A^{(k)}$  must be either 0 or 1, otherwise if  $A^{(k)}[i, j] > 1$  then there are two distinct length- $k$  paths from  $i$  to  $j$ , meaning there is a cycle of length at most  $2k = \ell$ , a contradiction. Therefore, the  $L_1$  norm of each row of  $A^{(k)}$  is at most  $n$ , hence  $\|A^{(k)}\|_2 \leq n$ . Then, setting  $t = 1/\rho$ , we have  $\|A^{(k)}\|_2 \cdot (1/\rho)^k < 1$  since  $k > \log_{\rho} n$ , and [Eq. \(12\)](#) shows that  $\mathrm{tr}(J(1/\rho)) < \infty$ . This contradicts that  $\mathrm{tr}(J(1/\rho))$  must diverge.

Similarly, let  $k' := \lfloor \log_{\rho} 2n \rfloor + 1$  and suppose for contradiction that  $G$  has no bicycle of size  $\leq \ell' = 3k'$ . We claim that three distinct non-backtracking walks of a given length- $k'$  between any two vertices must form a bicycle, hence every entry of  $A^{(k')}$  must be at most 2. Suppose the union of the three distinct nonbacktracking walks between vertices  $u$  and  $v$ , called  $H_{uv}$ , did not give rise to a bicycle, its excess must be at most 0. Since  $H_{uv}$  is connected, it must have at most one cycle. If there are no cycles, then there is exactly one nonbacktracking walk from  $u$  to  $v$ , so we assume there is exactly one cycle. Any nonbacktracking walk in  $H_{uv}$  can enter and exit the cycle at most once. Further, there is a unique way to start from  $u$  and enter the cycle, and a unique way to exit the cycle and arrive at  $v$ . Between entering and exiting the cycle, there are only two choices: walking in the cycle clockwise or counterclockwise. There are at most two ways to walk between  $u$  and  $v$  in  $k'$  steps — either the shortest path between them is of length exactly  $k'$  and does not touch the cycle, or a length- $k$  nonbacktracking walk must enter the cycle, which we established gives at most 2 distinct walks.

Thus,  $\|A^{(k')}\|_2 \leq 2n$  and  $\|A^{(k')}\|_2 \cdot (1/\rho)^{k'} < 1$  since  $k' > \log_{\rho} 2n$ . Again, [Eq. \(12\)](#) shows that  $\mathrm{tr}(J(1/\rho)) < \infty$ , a contradiction. This completes the proof.  $\square$

## B Expansion profile of random graphs

In this section we prove [Lemma 4.3](#) (existence of biregular graphs with good expansion profile). We first prove the desired statement for Erdős–Rényi graphs given in [Lemma B.1](#), and then transfer the result to random regular graphs via a coupling articulated in [Lemma B.2](#). See [Section 2](#) for the notations of various random bipartite graph models.

**Lemma B.1.** *Let  $H \sim \mathbb{G}_{n_1, n_2, p}$  where  $n_1 \geq n_2$ . Then with probability  $1 - O\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$ , for all  $t$ :*

$$P_H(t) \geq p(1-p)^{t-1}n_2 - \sqrt{4p(1-p)^{t-1}n_1 \log n_1}.$$

**Lemma B.2** (Embedding Erdős–Rényi graphs into random regular graphs). *Fix  $n_1, n_2, d_1, d_2 \in \mathbb{N}$  such that  $m = n_1 d_1 = n_2 d_2$ . Then for  $p = \left(1 - C\left(\frac{d_1 d_2}{m} + \frac{\log m}{\min\{d_1, d_2\}}\right)^{1/3}\right) \frac{m}{n_1 n_2}$ , there is a joint distribution of  $H \sim \mathbb{G}_{n_1, n_2, p}$  and  $R \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  such that*

$$\Pr[H \subset R] = 1 - o(1).$$

We first give a proof of [Lemma 4.3](#) assuming the [Lemmas B.1](#) and [B.2](#). [Lemma B.1](#) is proved later in this section, and [Lemma B.2](#) is proved in [Appendix C](#).

*Proof of Lemma 4.3.* Recall that we would like to show that for  $1 \geq \frac{d_1}{d_2} = \frac{D_2}{D_1} \geq \frac{\beta}{1-\beta}$  and  $\theta\sqrt{D}/2 \leq d_1 + d_2 \leq \theta\sqrt{D}$ , there exists a  $(d_1, d_2)$ -biregular graph  $R$  with  $D_1$  and  $D_2$  vertices on the left and right respectively such that  $P_R(t)$  is large.

By Lemma B.2 there is a coupling between  $\mathbf{R} \sim \mathbb{R}_{D_1, D_2, d_1, d_2}$  and  $\mathbf{H} \sim \mathbb{G}_{D_1, D_2, p}$  such that  $\mathbf{H} \subset \mathbf{R}$  with probability  $1 - o_D(1)$  where  $p = \left(1 - \frac{C_\beta \log^{1/3} D}{D^{1/6}}\right) \frac{d_1}{D_2}$  where  $C_\beta$  is a constant depending on  $\beta$  and  $\theta$ . Note that  $p \leq \frac{d_1}{D_2} = \frac{d_1 + d_2}{D_1 + D_2} \leq \frac{\theta}{\sqrt{D}}$ .

By concentration of the binomial random variable and the union bound, with probability  $1 - o_D(1)$  all vertices have degree  $o(d_2) = o(\sqrt{D})$  in  $\mathbf{R} \setminus \mathbf{H}$ . Consequently:  $P_R(t) \geq P_H(t) - o(\sqrt{D})$ . Thus, it suffices to lower bound  $P_H(t)$  to obtain a lower bound on  $P_R(t)$ .

By Lemma B.1, for  $t \geq 1$ ,

$$\begin{aligned} P_H(t) &\geq p(1-p)^{t-1}D_2 - \sqrt{4p(1-p)^{t-1}D_1 \log D_1} \\ &\geq \left(d_1(1-p)^{t-1} - \sqrt{4d_2(1-p)^{t-1} \log D}\right) \cdot (1 - o_D(1)). \end{aligned}$$

For  $t \leq L\sqrt{D}$ ,  $(1-p)^{t-1}$  is  $\Omega(1)$  because  $p \leq \frac{\theta}{\sqrt{D}}$ , and since  $d_2 \leq \theta\sqrt{D}$  and  $d_1 \geq \beta(d_1 + d_2) \geq \beta \cdot \theta\sqrt{D}/2$ , we can conclude that the above is at least  $(1 - o_D(1)) \cdot d_1 \cdot (1-p)^{t-1}$ . Finally, since  $p \leq \frac{\theta}{\sqrt{D}}$ , for  $t \leq L\sqrt{D}$ ,

$$(1-p)^{t-1} \geq (1 - o_D(1)) \cdot \exp(-pt) \geq (1 - o_D(1)) \cdot \exp(-\theta t / \sqrt{D}),$$

which completes the proof.  $\square$

We now prove Lemma B.1: we show a lower bound on the expansion profile of  $\mathbb{G}_{n_1, n_2, p}$  using standard concentration inequalities and union bound.

*Proof of Lemma B.1.* Write  $S \subseteq V(\mathbf{H})$ , write  $S := S_L \cup S_R$  where  $S_L := S \cap L(\mathbf{H})$  and  $S_R := S \cap R(\mathbf{H})$ . Observe that  $|\text{UN}_H(S)| = |\text{UN}_H(S_L)| + |\text{UN}_H(S_R)|$ . Therefore, without loss of generality we can study  $S$  completely in  $L(\mathbf{H})$  or  $R(\mathbf{H})$ .

For  $S \subseteq R(\mathbf{H})$  with  $|S| = t$ , we have:

$$|\text{UN}_H(S)| = \sum_{v \in L(\mathbf{H})} \mathbf{1}[v \in \text{UN}_H(S)].$$

For each  $v \in L(\mathbf{H})$ , the number of edges between  $v$  and  $S$  is distributed as  $\text{Bin}(t, p)$ , so each  $\mathbf{1}[v \in \text{UN}_H(S)]$  is an independent Bernoulli with bias  $q_t := tp(1-p)^{t-1}$ . By the Chernoff bound:

$$\Pr[|\text{UN}_H(S)| \leq q_t n_1 - s\sqrt{q_t n_1}] \leq \exp(-s^2/2),$$

which in particular implies that  $|\text{UN}_H(S)| \geq q_t n_1 - \sqrt{4q_t n_1 t \log n_2}$  except with probability at most  $n_2^{-2t}$ . By a union bound over all  $S \subseteq R(\mathbf{H})$  of size  $t$ ,

$$\forall S \subseteq R(\mathbf{H}) \text{ s.t. } |S| = t : |\text{UN}_H(S)| \geq q_t n_1 - \sqrt{4q_t n_1 t \log n_2}$$

with probability at least  $1 - n_2^{-t}$ . By an identical argument,

$$\forall S \subseteq L(\mathbf{H}) \text{ s.t. } |S| = t : |\text{UN}_H(S)| \geq q_t n_2 - \sqrt{4q_t n_2 t \log n_1}$$

with probability at least  $1 - n_1^{-t}$ . In both cases, since  $n_1 \geq n_2$ , we have

$$\frac{|\text{UN}_H(S)|}{|S|} \geq p(1-p)^{t-1}n_2 - \sqrt{4p(1-p)^{t-1}n_1 \log n_1}.$$

Finally, taking a union bound over all  $t \geq 1$  completes the proof.  $\square$

## C Coupling Erdős–Rényi and random regular graphs

We will closely follow [FK16, Section 11.5] (which is a special case of [DFRŠ17]), adapted to the case of bipartite graphs. As before, we will write random variables in **boldface**, and see Section 2 for a reminder of the notation for various random bipartite graph models.

**Theorem C.1** (Embedding theorem). *Fix  $n_1, n_2, d_1, d_2 \in \mathbb{N}$  such that  $m = n_1 d_1 = n_2 d_2$ . There is a universal constant  $C$  such that if  $\gamma \in (0, 1)$  satisfies  $\gamma \geq C \left( \frac{d_1 d_2}{m} + \frac{\log m}{\min\{d_1, d_2\}} \right)^{1/3}$ , then for  $\tilde{m} \leq \lfloor (1 - \gamma)m \rfloor$ , there is a joint distribution of  $\mathbf{G} \sim \mathbb{G}_{n_1, n_2, \tilde{m}}$  and  $\mathbf{R} \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  such that*

$$\Pr[\mathbf{G} \subset \mathbf{R}] = 1 - o(1).$$

Furthermore, let  $p = \frac{(1-2\gamma)m}{n_1 n_2}$ . There is a joint distribution of  $\mathbf{H} \sim \mathbb{G}_{n_1, n_2, p}$  and  $\mathbf{R} \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  such that

$$\Pr[\mathbf{H} \subset \mathbf{R}] = 1 - o(1).$$

To prove Theorem C.1, we need to introduce some more notation. With slight abuse of notation, we write

$$\mathbf{G} = (e_1, \dots, e_m), \quad \mathbf{R} = (f_1, \dots, f_m)$$

to be random orderings of the edges. Moreover, for  $t = 1, 2, \dots, m$ , we define random variables  $\mathbf{G}_t = (e_1, \dots, e_t)$  and  $\mathbf{R}_t = (f_1, \dots, f_t)$ .

Note that for any bipartite graph  $G$  of size  $t$  and any edge  $e \in K_{n_1, n_2} \setminus G$ , the conditional probability

$$\Pr[e_{t+1} = e | \mathbf{G}_t = G] = \frac{1}{n_1 n_2 - t}.$$

This motivates the following definition.

**Definition C.2.** Fix  $\varepsilon \in (0, 1)$ . We define  $A_{\varepsilon, t}$  to be the event that for all  $e \in K_{n_1, n_2} \setminus \mathbf{R}_t$ ,

$$\Pr[f_{t+1} = e | \mathbf{R}_t] \geq \frac{1 - \varepsilon}{n_1 n_2 - t}. \quad (13)$$

Further, we define the stopping time

$$T_\varepsilon = \max\{u : A_{\varepsilon, t} \text{ occurs for all } t \leq u\}.$$

Intuitively, suppose we sample the edges of  $\mathbf{R} \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  one by one. At each time step  $t \leq T_\varepsilon$ , the conditional distribution of the next edge is close to uniform, which is the case for  $\mathbb{G}_{n_1, n_2, m}$ . Thus, the main ingredient in the proof of Theorem C.1 is to show that  $T_\varepsilon$  is large with high probability, i.e., for most steps, the sampling process behaves roughly like  $\mathbb{G}_{n_1, n_2, m}$ . The following lemma is analogous to Lemma 11.18 of [FK16].



**Lemma C.3** (Large stopping time). *There is a universal constant  $C$  such that if  $\varepsilon \in (0, 1)$  satisfies  $\varepsilon \geq C \left( \frac{d_1 d_2}{m} + \frac{\log m}{\min\{d_1, d_2\}} \right)^{1/3}$ , then  $T_\varepsilon \geq (1 - \varepsilon)m$  with probability  $1 - o(1)$ .*

We will defer the proof to [Appendix C.2](#). This lemma suffices to prove [Theorem C.1](#).

*Proof of [Theorem C.1](#) by [Lemma C.3](#).* Recall that  $m = n_1 d_1 = n_2 d_2$ . We will define a graph process  $\mathbf{R}' = (f'_1, \dots, f'_m)$  coupled with  $\mathbf{G} = (e_1, \dots, e_m) \sim \mathbb{G}_{n_1, n_2, m}$  and show that (1)  $f'_t$  and  $f_t$  have the same conditional distribution, and (2)  $\mathbf{R}' \cap \mathbf{G}$  is large and contains a random subgraph  $\mathbf{G}' \subset \mathbf{G}$  of size  $\tilde{m} < m$  with high probability. Since a subgraph  $\mathbf{G}'$  is also distributed as  $\mathbb{G}_{n_1, n_2, \tilde{m}}$ , this gives a coupling between  $\mathbf{G}' \sim \mathbb{G}_{n_1, n_2, \tilde{m}}$  and  $\mathbf{R}' \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  such that  $\mathbf{G}' \subset \mathbf{R}'$  with high probability.

Recall that  $\Pr[e_{t+1} = e | \mathbf{G}_t] = \frac{1}{n_1 n_2 - t}$  for all  $t$  and  $e \in K_{n_1, n_2} \setminus \mathbf{G}_t$ , and we write

$$p_{t+1}(e | \mathbf{R}_t) := \Pr[f_{t+1} = e | \mathbf{R}_t]$$

which is at least  $\frac{1-\varepsilon}{n_1 n_2 - t}$  for  $t \leq T_\varepsilon$  by [Definition C.2](#).

The graph process  $\mathbf{R}'$  is sampled as follows: at time step  $t \leq T_\varepsilon$ ,

1. Sample a Bernoulli random variable  $\zeta_{t+1} \in \{0, 1\}$  with bias  $1 - \varepsilon$ .
2. Sample a random edge  $\mathbf{g}_{t+1} \in K_{n_1, n_2} \setminus \mathbf{R}'_t$  according to the conditional distribution

$$\Pr[\mathbf{g}_{t+1} = e | \mathbf{R}'_t, \mathbf{G}_t] := \frac{1}{\varepsilon} \left( p_{t+1}(e | \mathbf{R}'_t) - \frac{1 - \varepsilon}{n_1 n_2 - t} \right) \geq 0.$$

Note that this is a valid probability distribution over  $K_{n_1, n_2} \setminus \mathbf{R}'_t$  since the above is non-negative due to [Eq. \(13\)](#) and the sum is 1 because  $|K_{n_1, n_2} \setminus \mathbf{R}'_t| = n_1 n_2 - t$ .

3. Fix any bijection map  $h : \mathbf{R}'_t \setminus \mathbf{G}_t \rightarrow \mathbf{G}_t \setminus \mathbf{R}'_t$ . Set

$$f'_{t+1} = \begin{cases} e_{t+1}, & \text{if } \zeta_{t+1} = 1, e_{t+1} \notin \mathbf{R}'_t, \\ h(e_{t+1}), & \text{if } \zeta_{t+1} = 1, e_{t+1} \in \mathbf{R}'_t, \\ \mathbf{g}_{t+1}, & \text{if } \zeta_{t+1} = 0. \end{cases}$$

Note that if  $\zeta_{t+1} = 1$ , then  $f'_{t+1} \in \mathbf{G}_{t+1}$ .

For  $t > T_\varepsilon$ , we sample  $f'_{t+1}$  according to the probabilities  $p_{t+1}(e | \mathbf{R}'_t)$  without coupling, and we keep sampling  $\zeta_{t+1}$  for notational convenience.

We first show that the conditional distribution of  $f'_{t+1}$  is the same as  $f_{t+1}$ . For  $e \in K_{n_1, n_2} \setminus \mathbf{R}'_t$ ,

$$\begin{aligned} \Pr[f'_{t+1} = e | \mathbf{R}'_t, \mathbf{G}_t] &= (1 - \varepsilon) \cdot \Pr[f'_{t+1} = e | \mathbf{R}'_t, \mathbf{G}_t, \zeta_{t+1} = 1] + \varepsilon \cdot \Pr[\mathbf{g}_{t+1} = e | \mathbf{R}'_t, \mathbf{G}_t] \\ &= \frac{1 - \varepsilon}{n_1 n_2 - t} + \left( p_{t+1}(e | \mathbf{R}'_t) - \frac{1 - \varepsilon}{n_1 n_2 - t} \right) \\ &= p_{t+1}(e | \mathbf{R}'_t). \end{aligned}$$

This is because if  $\zeta_{t+1} = 1$  then  $f'_{t+1}$  is an edge in  $\mathbf{G}_{t+1}$ , and if  $\zeta_{t+1} = 0$  then  $f'_{t+1} = \mathbf{g}_{t+1}$ . This shows that  $f'_{t+1}$  and  $f_{t+1}$  have the same conditional distribution.

Next, we claim that in the end,  $\mathbf{R}'$  and  $\mathbf{G}$  share many edges. Let

$$\mathbf{S} := \{f'_t : \zeta_t = 1, 0 \leq t \leq (1 - \varepsilon)m\} \subset \mathbf{R}'.$$

By [Lemma C.3](#),  $T_\varepsilon \geq (1 - \varepsilon)m$  with probability  $1 - o(1)$ . Conditioned on this, we know that all edges in  $S$  lie in  $G$ . Moreover,  $|S|$  is distributed as  $\text{Bin}((1 - \varepsilon)m, 1 - \varepsilon)$ , so  $\mathbb{E}[|S|] = (1 - \varepsilon)^2 m$ , and by Chernoff bound,

$$\Pr[|S| \leq (1 - \varepsilon)^3 m] \leq \exp\left(-\frac{1}{2}\varepsilon^2(1 - \varepsilon)^2 m\right) = o(1).$$

Let  $\gamma = 3\varepsilon$  and fix  $\tilde{m} \leq \lfloor (1 - \gamma)m \rfloor \leq (1 - \varepsilon)^3 m$ . We now take the first  $\tilde{m}$  edges from  $S \subset R'$ , and the resulting graph  $G'$  is distributed as  $\mathbb{G}_{n_1, n_2, \tilde{m}}$ . Thus, we have obtained a joint distribution between  $G' \sim \mathbb{G}_{n_1, n_2, \tilde{m}}$  and  $R' \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  such that  $G' \subset R'$  with probability  $1 - o(1)$ .

The second statement of the theorem is a simple modification. We sample  $\mathbb{G}_{n_1, n_2, p}$  as follows: (1) sample  $m' \sim \text{Bin}(n_1 n_2, p)$ , and (2) sample  $H \sim \mathbb{G}_{n_1, n_2, m'}$  coupled with  $R' \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  as described before (if  $m' > m$  then sample the extra edges randomly). By the Chernoff bound,

$$\Pr[m' \geq (1 + \gamma)n_1 n_2 p] \leq \exp(-\gamma^2 n_1 n_2 p / 3) = o(1).$$

Since  $p = \frac{(1 - 2\gamma)m}{n_1 n_2}$ ,  $\lfloor (1 + \gamma)n_1 n_2 p \rfloor \leq \lfloor (1 - \gamma)m \rfloor$ , and conditioned on  $m' \leq \lfloor (1 - \gamma)m \rfloor$ , the exact same analysis goes through. Thus, we get  $\Pr[H \subset R'] = 1 - o(1)$ , completing the proof.  $\square$

## C.1 Random graph extension

To prove [Lemma C.3](#), we first need a few definitions and lemmas about extensions of graphs. As before, fix  $n_1, n_2, d_1, d_2 \in \mathbb{N}$  such that  $m := n_1 d_1 = n_2 d_2$ . We first introduce the following definitions.

**Definition C.4** (Graph extension). Given an ordered bipartite graph  $G = (e_1, \dots, e_t)$ , we say that an ordered simple  $(d_1, d_2)$ -biregular graph  $H = (f_1, \dots, f_m)$  with  $m$  edges is an *extension* of  $G$  if  $e_i = f_i$  for  $i \leq t$ . We write  $\mathcal{S}_G := \mathcal{S}_G(n_1, n_2, d_1, d_2)$  to denote the set of extensions of  $G$ , and write  $\mathbb{S}_G$  as a random graph sampled uniformly from  $\mathcal{S}_G$  (we will drop the dependence on  $n_1, n_2, d_1, d_2$  when clear from context).

Given  $G$  and an extension  $H$ , for vertices  $u, v \in L$  or  $u, v \in R$ ,

$$\deg_{H|G}(u, v) = |\{w : (u, w) \in H \setminus G \text{ and } (v, w) \in H\}|.$$

Note that  $\deg_{H|G}(u, v)$  is not symmetric in  $u$  and  $v$ .

Although [Definition C.4](#) is a natural definition, it is difficult to analyze since we require the extension of  $G$  to be *simple*. On the other hand, if we allow *multigraphs* (parallel edges allowed), then there is a very simple process to sample a multigraph extension from  $G$ , namely the ‘‘configuration model’’. Furthermore, it is easy to see that conditioned on the sampled multigraph being simple, the process gives the uniform distribution over  $\mathcal{S}_G$ .

**Definition C.5** (Random multigraph extension). Given a graph  $G = (e_1, \dots, e_t)$  of size  $t$ , we denote  $M_G$  to be an ordered random *multigraph* extension of  $G$  sampled as follows:

1. Set  $U$  to be a random permutation of  $(1, \dots, 1, \dots, n_1, \dots, n_1)$  where each  $u \in [n_1]$  has multiplicity  $d_1 - \deg_G(u)$ .  $U$  has length  $n_1 d_1 - |G| = m - t$ .

2. Set  $V$  to be a random permutation of  $(1, \dots, 1, \dots, n_2, \dots, n_2)$  where each  $v \in [n_2]$  has multiplicity  $d_2 - \deg_G(v)$ .  $V$  also has length  $m - t$ .
3. Set the  $i$ -th edge of  $M_G$  to be  $e_i$  for  $i \leq t$ , and set the  $(t + j)$ -th edge of  $M_G$  to be  $(U[j], V[j])$  for each  $1 \leq j \leq m - t$ .

**Fact C.6.** There are  $\frac{(m-t)!}{\prod_{u \in [n_1]} (d_1 - \deg_G(u))!}$  distinct permutations of  $U$  and  $\frac{(m-t)!}{\prod_{v \in [n_2]} (d_2 - \deg_G(v))!}$  distinct permutations of  $V$ . Suppose  $H$  is a simple (ordered) extension of  $G$ , then

$$\Pr[M_G = H] = \frac{\prod_{u \in [n_1]} (d_1 - \deg_G(u))! \cdot \prod_{v \in [n_2]} (d_2 - \deg_G(v))!}{((m-t)!)^2}.$$

In particular, conditioned on  $M_G$  being simple, it has the same distribution as  $S_G$ .

The main ingredient of the proof of [Lemma C.3](#) is the following lemma, which states that the probability of a random multigraph extension of  $G \cup e$  being simple is roughly the same for all  $e \notin G$ , assuming that  $G$  is not too “saturating”.

**Lemma C.7.** Let  $\varepsilon \geq C \left(\frac{d_1 d_2}{m}\right)^{1/3} + C \sqrt{\frac{\log m}{\min\{d_1, d_2\}}}$  for a large enough constant  $C$ . Let  $G$  be a bipartite graph with  $t \leq (1 - \varepsilon)m$  edges such that  $S_G = \mathcal{S}_G(n_1, n_2, d_1, d_2)$  is non-empty. If  $\deg_G(u) \leq (1 - \varepsilon/2)d_1$  and  $\deg_G(v) \leq (1 - \varepsilon/2)d_2$  for all  $u \in [n_1], v \in [n_2]$ , then for every  $e, e' \notin G$ , we have

$$\frac{\Pr[M_{G \cup e'} \in \mathcal{S}_{G \cup e'}]}{\Pr[M_{G \cup e} \in \mathcal{S}_{G \cup e}]} \geq 1 - \frac{\varepsilon}{2}.$$

**Typical random extension.** The following is analogous to Lemma 11.20 in [\[FK16\]](#) and will be used to prove [Lemma C.7](#). It roughly states that a random extension  $S_G$  of a graph  $G$  behaves nicely.

**Lemma C.8.** Let  $\varepsilon m \geq 4d_1 d_2$ . Let  $G$  be a graph with  $t \leq (1 - \varepsilon)m$  edges such that  $S_G$  is non-empty, and let  $S_G$  be a uniform sample from  $S_G$ . For every  $e \notin G$ , we have

$$\Pr[e \in S_G] \leq \frac{2d_1 d_2}{\varepsilon m}.$$

Moreover, for every  $u_1, u_2 \in [n_1]$  and  $\ell \geq \ell_1 := \lceil \frac{4d_1^2 d_2}{\varepsilon m} \rceil$ .

$$\Pr[\deg_{S_G|G}(u_1, u_2) > \ell] \leq 2^{-(\ell - \ell_1)},$$

and for every  $v_1, v_2 \in [n_2]$  and  $\ell \geq \ell_2 := \lceil \frac{4d_1 d_2^2}{\varepsilon m} \rceil$ .

$$\Pr[\deg_{S_G|G}(v_1, v_2) > \ell] \leq 2^{-(\ell - \ell_2)}.$$

*Proof.* Fix  $e = (u, v) \notin G$ . We first define

$$\mathcal{G}_{\in e} = \{H \in S_G : e \in H\}, \quad \mathcal{G}_{\notin e} = \{H \in S_G : e \notin H\}.$$

Then,  $\Pr[e \in S_G] = \frac{|\mathcal{G}_{\in e}|}{|\mathcal{G}_{\in e}| + |\mathcal{G}_{\notin e}|} \leq \frac{|\mathcal{G}_{\in e}|}{|\mathcal{G}_{\notin e}|}$ . We will proceed to upper bound this ratio.

Define a bipartite graph  $B$  between  $\mathcal{G}_{\in e}$  and  $\mathcal{G}_{\notin e}$  as follows. We connect  $H \in \mathcal{G}_{\in e}$  and  $H' \in \mathcal{G}_{\notin e}$  if we can obtain  $H'$  from  $H$  with the following switching operation: choose an edge  $(w, x) \in H \setminus G$  disjoint from  $(u, v)$  such that  $(u, x)$  and  $(w, v)$  are not edges in  $H$ , and replace  $(u, v), (w, x)$  by  $(u, x), (w, v)$ . We write  $\deg_B(H)$  to denote the degree of  $H$  in  $B$ .

For  $H \in \mathcal{G}_{\in e}$ , we can choose edge  $(w, x) \in H \setminus G$  as long as  $w$  is not a neighbor of  $v$  and  $x$  is not a neighbor of  $u$  in  $H$ . Thus, there are at least  $|H| - |G| - \deg_H(u) \cdot \deg_H(v) = m - t - d_1 d_2$  choices, meaning  $\deg_B(H) \geq m - t - d_1 d_2$ .

On the other hand, for  $H' \in \mathcal{G}_{\notin e}$ , we must select  $x$  to be a neighbor of  $u$  and  $w$  a neighbor of  $v$  in  $H'$ . Thus,  $\deg_B(H') \leq d_1 d_2$ .

Since  $B$  is a bipartite graph, we must have

$$|\mathcal{G}_{\in e}| \cdot \min_{H \in \mathcal{G}_{\in e}} \deg_B(H) \leq |\mathcal{G}_{\notin e}| \cdot \max_{H' \in \mathcal{G}_{\notin e}} \deg_B(H') \implies \frac{|\mathcal{G}_{\in e}|}{|\mathcal{G}_{\notin e}|} \leq \frac{d_1 d_2}{m - t - d_1 d_2} \leq \frac{2d_1 d_2}{\varepsilon m},$$

since  $t \leq (1 - \varepsilon)m$  and  $\varepsilon m \geq 4d_1 d_2$ .

To prove the second statement, for  $u_1, u_2 \in [n_1]$ , we define

$$\mathcal{G}_\ell = \left\{ H \in \mathcal{S}_G : \deg_{H|G}(u_1, u_2) = \ell \right\}$$

for  $\ell = 0, 1, \dots$ , and we adopt a similar strategy of constructing an auxiliary bipartite graph  $B_\ell$  between  $\mathcal{G}_\ell$  and  $\mathcal{G}_{\ell-1}$ . We connect  $H \in \mathcal{G}_\ell$  to  $H' \in \mathcal{G}_{\ell-1}$  if we can obtain  $H'$  from  $H$  by the following: (1) select a vertex  $w$  contributing to  $\deg_{H|G}(u_1, u_2)$ , i.e.,  $(u_1, w) \in H \setminus G$  and  $(u_2, w) \in H$ , (2) select a disjoint edge  $(u', w') \in H \setminus G$  such that there is no edge between  $\{u', w'\}$  and  $\{u_1, u_2, w\}$  in  $H$ , and (3) replace edges  $(u_1, w)$  and  $(u', w')$  with  $(u', w)$  and  $(u, w')$ .

By a similar analysis, fix  $H \in \mathcal{G}_\ell$ , there are  $\ell$  choices for  $w$  and at least  $m - t - 2d_1 d_2$  choices for  $(u', w')$ , thus  $\deg_{B_\ell}(H) \geq \ell(m - t - 2d_1 d_2) \geq \frac{1}{2}\ell\varepsilon m$ . On the other hand, fix  $H' \in \mathcal{G}_{\ell-1}$ , there are at most  $d_1$  choices for  $w$ ,  $d_1$  choices for  $u'$ , and  $d_2$  choices for  $w'$ , thus  $\deg_{B_\ell}(H') \leq d_1^2 d_2$ . Therefore,

$$\frac{|\mathcal{G}_\ell|}{|\mathcal{G}_{\ell-1}|} \leq \frac{2d_1^2 d_2}{\ell\varepsilon m} \leq \frac{1}{2}$$

for all  $\ell \geq \ell_1 := \lceil \frac{4d_1^2 d_2}{\varepsilon m} \rceil$ . Therefore,

$$\Pr \left[ \deg_{\mathcal{S}_G|G}(u_1, u_2) > \ell \right] = \frac{\sum_{k>\ell} |\mathcal{G}_k|}{\sum_{k \geq 0} |\mathcal{G}_k|} \leq \frac{\sum_{k>\ell} |\mathcal{G}_k|}{|\mathcal{G}_{\ell_1}|} \leq \sum_{k>\ell} 2^{-(k-\ell_1)} \leq 2^{-(\ell-\ell_1)}.$$

For  $v_1, v_2 \in [n_2]$ , the same analysis shows that  $\Pr[\deg_{\mathcal{S}_G|G}(v_1, v_2) > \ell] \leq 2^{-(\ell-\ell_2)}$  for all  $\ell \geq \ell_2 := \lceil \frac{4d_1 d_2^2}{\varepsilon m} \rceil$ . This completes the proof.  $\square$

We can now complete the proof of [Lemma C.7](#).

*Proof of Lemma C.7.* We will write  $M = M_{G_{Ue}}$  and  $M' = M_{G_{Ue'}}$  for simplicity. We will construct a coupling of  $M$  and  $M'$  such that they differ in at most 3 positions. Given  $M$  and  $e = (u, v)$ ,  $e' = (u', v')$ , we perform a switching operation:

1. Delete  $e$  and add  $e'$  to  $M$ .
2. Recall  $U$  and  $V$  of length  $m - (t + 1)$  defined in [Definition C.5](#).

- If  $e$  and  $e'$  are disjoint, then randomly select a copy of  $u'$  in  $U$  and change to  $u$ , and similarly randomly select a copy of  $v'$  in  $V$  and change to  $v$ .
- If  $e$  and  $e'$  are not disjoint (w.l.o.g. assume  $v = v'$ ), then just change a random copy of  $u'$  in  $U$  to  $u$ .

Then, connect edges according to  $U$  and  $V$  as in [Definition C.5](#).

Note that step 2 is equivalent to sampling a random edge  $(u', w_R)$  incident to  $u'$  in  $M \setminus (G \cup e)$  and replacing  $(u', w_R)$  with  $(u, w_R)$ , and similarly replacing a random  $(w_L, v')$  with  $(w_L, v)$ .

We denote the resulting graph as  $M^*$ . It is clear that the resulting vector  $V$  after step 2 is distributed as a random permutation of  $(1, \dots, 1, \dots, n_2, \dots, n_2)$  with multiplicity  $d_2 - \deg_{G \cup e'}(v)$  for each  $v$ , thus  $M^*$  has the same distribution as  $M'$ .

We now analyze the probability of  $M^*$  being simple conditioned on  $M$  being simple. We first identify some nice properties of  $M$ , which we will show to occur with high probability. We define

$$\mathcal{G}_{\text{nice}} = \left\{ H \in \mathcal{S}_{G \cup e} : e' \notin H, \deg_{H|G \cup e}(u', u) \leq \ell_1 + \log m \text{ and } \deg_{H|G \cup e}(v', v) \leq \ell_2 + \log m \right\},$$

where  $\ell_1 = \lceil \frac{4d_1^2 d_2}{\varepsilon m} \rceil$  and  $\ell_2 = \lceil \frac{4d_1 d_2^2}{\varepsilon m} \rceil$  as defined in [Lemma C.8](#).

First, by [Fact C.6](#) and [Lemma C.8](#),  $\Pr[e' \in M | M \in \mathcal{S}_{G \cup e}] = \Pr[e' \in \mathcal{S}_{G \cup e}] \leq \frac{2d_1 d_2}{\varepsilon m} \leq \varepsilon/8$  due to our lower bound on  $\varepsilon$ . Moreover, we have  $\deg_{M|G \cup e}(u', u) > \ell_1 + \log m$  and  $\deg_{M|G \cup e}(v', v) > \ell_2 + \log m$  with probability  $\leq \frac{1}{m}$ . Thus,

$$\Pr[M \in \mathcal{G}_{\text{nice}} | M \in \mathcal{S}_{G \cup e}] \geq 1 - \frac{\varepsilon}{4}.$$

Now suppose  $M \in \mathcal{G}_{\text{nice}}$ . A parallel edge can occur in three ways: (1) if we replace  $(u', w_R) \in M \setminus (G \cup e)$  with  $(u, w_R)$  but  $(u, w_R) \in M$  already, (2) similarly for  $v'$ , (3) if we select  $(u', v)$  and  $(u, v')$  (i.e.,  $w_R = v$  and  $w_L = u$ ) resulting in  $(u, v)$  being parallel. By the union bound over these 3 cases,

$$\Pr[M^* \notin \mathcal{S}_{G \cup e'} | M \in \mathcal{S}_{G \cup e}] \leq \frac{\deg_{M|G \cup e}(u', u)}{\deg_{M \setminus (G \cup e)}(u')} + \frac{\deg_{M|G \cup e}(v', v)}{\deg_{M \setminus (G \cup e)}(v')} + \frac{1}{\deg_{M \setminus (G \cup e)}(u') \cdot \deg_{M \setminus (G \cup e)}(v')}.$$

By the assumption on the degrees of  $G$ , we know that  $\deg_{M \setminus (G \cup e)}(u') \geq \varepsilon d_1/2$  and  $\deg_{M \setminus (G \cup e)}(v') \geq \varepsilon d_2/2$ .

$$\begin{aligned} \Pr[M^* \notin \mathcal{S}_{G \cup e'} | M \in \mathcal{S}_{G \cup e}] &\leq \frac{\ell_1 + \log m}{\varepsilon d_1/2} + \frac{\ell_2 + \log m}{\varepsilon d_2/2} + \frac{1}{\varepsilon^2 d_1 d_2/4} \\ &\leq \frac{8d_1 d_2}{\varepsilon^2 m} + \frac{2 \log m}{\varepsilon} \left( \frac{1}{d_1} + \frac{1}{d_2} \right) + \frac{4}{\varepsilon^2 d_1 d_2}. \end{aligned}$$

For  $\varepsilon \geq C \left( \frac{d_1 d_2}{m} \right)^{1/3} + C \sqrt{\frac{\log m}{\min\{d_1, d_2\}}}$  for some large enough constant  $C$ , the above can be bounded by  $\varepsilon/4$ .

Finally, we can finish the proof. As  $M^*$  is distributed as  $M'$ ,

$$\begin{aligned} \frac{\Pr[M' \in \mathcal{S}_{G \cup e'}]}{\Pr[M \in \mathcal{S}_{G \cup e}]} &\geq \frac{\Pr[M \in \mathcal{G}_{\text{nice}}, M \in \mathcal{S}_{G \cup e}]}{\Pr[M \in \mathcal{S}_{G \cup e}]} \cdot \frac{\Pr[M^* \in \mathcal{S}_{G \cup e'}]}{\Pr[M \in \mathcal{G}_{\text{nice}}]} \\ &\geq \left( 1 - \frac{\varepsilon}{4} \right)^2 \geq 1 - \frac{\varepsilon}{2}, \end{aligned}$$

completing the proof. □

## C.2 Proof of Lemma C.3

**Degree bounds.** We first prove a degree concentration bound. To do so, we will need concentration results for random processes without replacement.

**Fact C.9** ([Hoe63]). Fix  $1 \leq n \leq N$ . Let  $\Omega = (x_1, \dots, x_N)$  be a finite set of points in  $\mathbb{R}$ . Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be a random sample without replacement from  $\Omega$ , let  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  be a random sample with replacement from  $\Omega$ , and let  $\mathbf{X} = \sum_{i=1}^n \mathbf{X}_i$  and  $\mathbf{Y} = \sum_{i=1}^n \mathbf{Y}_i$ . Suppose  $f : \mathbb{R} \rightarrow \mathbb{R}$  is continuous and convex, then

$$\mathbf{E}[f(\mathbf{X})] \leq \mathbf{E}[f(\mathbf{Y})].$$

Fact C.9 implies that many concentration results known for sampling *with* replacement, such as the Chernoff bound, can be transferred to the case of sampling *without* replacement. In particular, the following is a standard result for which we include the proof for completeness.

**Lemma C.10** (Concentration for sampling without replacement). Fix  $1 \leq k, n \leq N$ . Let  $S \subseteq [N]$ , let  $T$  be a random sample of  $n$  elements from  $[N]$  without replacement, and let  $p = \frac{n}{N}$ . Then, for all  $\delta \in (0, 1)$ ,

$$\begin{aligned} \Pr [||S \cap T| - p|S|| \geq \delta p|S|] &\leq 2 \exp\left(-\frac{1}{3}\delta^2 p|S|\right), \\ \Pr [||S \cap T| - p|S|| \geq \delta(1-p)|S|] &\leq 2 \exp\left(-\frac{1}{3}\delta^2(1-p)|S|\right). \end{aligned}$$

*Proof.* Let  $\Omega = (x_1, \dots, x_N)$  where  $x_i = 1$  if  $i \in S$  and 0 otherwise. Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be a random sample without replacement from  $\Omega$ . Then,  $\mathbf{X} = \sum_{i=1}^n \mathbf{X}_i$  has the same distribution as  $|S \cap T|$ .

By Fact C.9, we can apply the Chernoff bound as if the  $\mathbf{X}_i$ 's are sampled with replacement. Let  $\mathbf{Y}_1, \dots, \mathbf{Y}_n$  be samples from  $\Omega$  with replacement, then  $\mathbf{Y}_i$  are i.i.d. Bernoulli random variables with  $\mathbf{E}[\mathbf{Y}_i] = \frac{|S|}{N}$ , and  $\mathbf{E}\sum_{i=1}^n \mathbf{Y}_i = \frac{n|S|}{N} = p|S|$ . The first inequality then follows from the Chernoff bound.

For the second inequality, we look at  $|S \setminus T|$ . Due to symmetry, sampling  $n$  elements from  $[N]$  without replacement is equivalent to sampling  $N - n$  elements and taking the complement. Let  $\mathbf{X}'_1, \dots, \mathbf{X}'_{N-n}$  be a random sample without replacement from  $\Omega$ . Then,  $|S \setminus T|$  is distributed as  $\mathbf{X}' = \sum_{i=1}^{N-n} \mathbf{X}'_i$ , and  $\mathbf{E}[\mathbf{X}'] = (N - n)\frac{|S|}{N} = (1 - p)|S|$ . By Fact C.9 (transferring to sampling with replacement) and the Chernoff bound,

$$\Pr [||S \setminus T| - (1 - p)|S|| \geq \delta(1 - p)|S|] \leq 2 \exp\left(-\frac{1}{3}\delta^2(1 - p)|S|\right).$$

But  $|S \cap T| + |S \setminus T| = |S|$ , so  $||S \setminus T| - (1 - p)|S|| = ||S \cap T| - p|S||$ . This completes the proof.  $\square$

Recall that  $\mathbf{R} = (f_1, \dots, f_m) \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$  is a (ordered) random  $(d_1, d_2)$ -biregular graph on  $L = [n_1]$ ,  $R = [n_2]$ , and we write  $\mathbf{R}_t = (f_1, \dots, f_t)$ .

**Lemma C.11** (Degree concentration). Consider the random graph process  $\mathbf{R}_t = (f_1, \dots, f_t)$  for  $t = 0, 1, \dots, m$ . Let  $\varepsilon \in (0, 1)$ . With probability  $1 - O(m^{-1})$ , for all  $t \leq (1 - \varepsilon)m$ , letting  $p = \frac{t}{m}$ , we have

$$\begin{aligned} |\deg_{\mathbf{R}_t}(u) - pd_1| &\leq 3\sqrt{(1-p)d_1 \log m}, \quad \forall u \in [n_1], \\ |\deg_{\mathbf{R}_t}(v) - pd_2| &\leq 3\sqrt{(1-p)d_2 \log m}, \quad \forall v \in [n_2]. \end{aligned}$$

*Proof.* Fix a time step  $1 \leq t \leq (1 - \varepsilon)m$  and a vertex  $u \in [n_1]$ . Since  $\mathbf{R}_t$  can be viewed as sampling  $t$  edges from  $m$  total edges without replacement, we can apply the second inequality in [Lemma C.10](#) with  $|S| = d_1$  and  $\delta = 3\sqrt{\frac{\log m}{(1-p)d_1}}$  (note that  $1 - p \geq \varepsilon > 0$ ):

$$\Pr \left[ \left| \deg_{\mathbf{R}_t}(u) - pd_1 \right| \geq 3\sqrt{(1-p)d_1 \log m} \right] \leq 2m^{-3}.$$

By the same analysis, the inequality is true for  $v \in [n_2]$  if we replace  $d_1$  with  $d_2$ .

The lemma now follows from taking the union bound over all  $t \leq (1 - \varepsilon)m$  and  $u \in [n_1]$ ,  $v \in [n_2]$ .  $\square$

We are now ready to prove [Lemma C.3](#).

*Proof of Lemma C.3.* We would like to prove that for all  $e \in K_{n_1, n_2} \setminus \mathbf{R}_t$ ,  $\Pr[f_{t+1} = e | \mathbf{R}_t] \geq \frac{1-\varepsilon}{n_1 n_2 - t}$  for every  $t \leq (1 - \varepsilon)m$ . It suffices to prove that for every  $e, e' \in K_{n_1, n_2} \setminus \mathbf{R}_t$ ,

$$\frac{\Pr[f_{t+1} = e' | \mathbf{R}_t]}{\Pr[f_{t+1} = e | \mathbf{R}_t]} \geq 1 - \varepsilon,$$

since the average  $\Pr[f_{t+1} = e | \mathbf{R}_t]$  over all  $e$  must be  $\frac{1}{n_1 n_2 - t}$ , hence  $\max_e \Pr[f_{t+1} = e | \mathbf{R}_t] \geq \frac{1}{n_1 n_2 - t}$ .

Recalling the definition of extensions in [Definition C.4](#), we have

$$\frac{\Pr[f_{t+1} = e' | \mathbf{R}_t]}{\Pr[f_{t+1} = e | \mathbf{R}_t]} = \frac{|\mathcal{S}_{\mathbf{R}_t \cup e'}|}{|\mathcal{S}_{\mathbf{R}_t \cup e}|}. \quad (14)$$

We now consider multigraph extensions of  $\mathbf{R}_t \cup e$  and  $\mathbf{R}_t \cup e'$ . By [Fact C.6](#),

$$\begin{aligned} \Pr[M_{\mathbf{R}_t \cup e} \in \mathcal{S}_{\mathbf{R}_t \cup e}] &= |\mathcal{S}_{\mathbf{R}_t \cup e}| \cdot \frac{\prod_{u \in [n_1]} (d_1 - \deg_{\mathbf{R}_t \cup e}(u))! \cdot \prod_{v \in [n_2]} (d_2 - \deg_{\mathbf{R}_t \cup e}(v))!}{((m-t)!)^2}, \\ \Pr[M_{\mathbf{R}_t \cup e'} \in \mathcal{S}_{\mathbf{R}_t \cup e'}] &= |\mathcal{S}_{\mathbf{R}_t \cup e'}| \cdot \frac{\prod_{u \in [n_1]} (d_1 - \deg_{\mathbf{R}_t \cup e'}(u))! \cdot \prod_{v \in [n_2]} (d_2 - \deg_{\mathbf{R}_t \cup e'}(v))!}{((m-t)!)^2}. \end{aligned}$$

Thus, let  $e = (u, v)$  and  $e' = (u', v')$ ,

$$\frac{\Pr[M_{\mathbf{R}_t \cup e'} \in \mathcal{S}_{\mathbf{R}_t \cup e'}]}{\Pr[M_{\mathbf{R}_t \cup e} \in \mathcal{S}_{\mathbf{R}_t \cup e}]} = \frac{|\mathcal{S}_{\mathbf{R}_t \cup e'}|}{|\mathcal{S}_{\mathbf{R}_t \cup e}|} \cdot \frac{(d_1 - \deg_{\mathbf{R}_t}(u))(d_2 - \deg_{\mathbf{R}_t}(v))}{(d_1 - \deg_{\mathbf{R}_t}(u'))(d_2 - \deg_{\mathbf{R}_t}(v'))}. \quad (15)$$

Let  $p = \frac{t}{m}$ . By the concentration of degrees ([Lemma C.11](#)), with probability  $1 - O(m^{-1})$ , for  $t \leq (1 - \varepsilon)m$  (hence  $1 - p \geq \varepsilon$ ),

$$\frac{d_1 - \deg_{\mathbf{R}_t}(u)}{(1-p)d_1} \in 1 \pm 3\sqrt{\frac{\log m}{(1-p)d_1}} \in 1 \pm 3\sqrt{\frac{\log m}{\varepsilon d_1}}.$$

Our assumption on  $\varepsilon$  implies that  $\varepsilon \geq C \left( \frac{\log m}{\min\{d_1, d_2\}} \right)^{1/3}$  for a large enough  $C$ , so the above is bounded by  $1 \pm \varepsilon/8$ . The same also holds for  $d_2 - \deg_{\mathbf{R}_t}(v)$  and  $d_2 - \deg_{\mathbf{R}_t}(v')$ . Thus, we have

$$(15) \leq \frac{|\mathcal{S}_{\mathbf{R}_t \cup e'}|}{|\mathcal{S}_{\mathbf{R}_t \cup e}|} (1 + \varepsilon/2).$$



On the other hand, the degree bounds also allow us to apply [Lemma C.7](#):

$$(15) \geq 1 - \frac{\varepsilon}{2}.$$

Therefore, by [Eq. \(14\)](#) we have  $\frac{\Pr[f_{t+1}=e'|R_t]}{\Pr[f_{t+1}=e|R_t]} \geq 1 - \varepsilon$  with probability  $1 - O(m^{-1})$ , completing the proof.  $\square$