Explicit Two-sided Unique-neighbor Expanders

Jun-Ting Hsieh[*] Theo McKenzie[†] Sidhanth Mohanty[‡] Pedro Paredes[§]

February 2, 2023

Abstract

We study the problem of constructing explicit sparse imbalanced bipartite unique-neighbor expanders. For large enough d_1 and d_2, we give a strongly explicit construction of an infinite family of (d_1, d_2)-biregular graph (assuming $d_1 \leqslant d_2$) where all sets S with fewer than $1/d_1^3$ fraction of vertices have $\Omega(d_1 \cdot |S|)$ unique-neighbors. Further, for each $\beta \in (0, 1)$, we give a construction with the additional property that the left side of each graph has roughly β fraction of the total number of vertices. Our work provides the first two-sided construction of imbalanced unique-neighbor expanders, meaning small sets contained in both the left and right side of the bipartite graph exhibit unique-neighbor expansion.

Our construction is obtained from the "line product" of a large small-set edge expander and a sufficiently good constant-sized unique-neighbor expander, a product defined in the work of Alon and Capalbo [AC02].

[*]Carnegie Mellon University. juntingh@cs.cmu.edu. Supported by NSF CAREER Award #2047933.

[†]Harvard University. tmckenzie@math.harvard.edu. Supported by NSF GRFP Grant DGE-1752814 and NSF Grant DMS-2212881.

[‡]UC Berkeley. sidhanthm@cs.berkeley.edu. Supported by a Google PhD fellowship.

[§]Princeton University. pparedes@cs.princeton.edu.

1 Introduction

We say a graph is a *unique-neighbor expander* if every sufficiently small set of vertices has many *unique-neighbors*, where a unique-neighbor of a set of vertices S is a vertex v with exactly one edge to S. In this work, we are interested in the problem of constructing bipartite unique-neighbor expanders where every small subset of vertices has many unique-neighbors. This notion of expansion and explicitly constructing such graphs are motivated by applications to error-correcting codes [DSW06, BV09], high-dimensional geometry [BGI+08, GLR10, Kar11, GMM22], routing [ALM96], and understanding hard instances of random constraint satisfaction problems [DFHT21, HL22].

Random graphs expand *losslessly*, i.e. they achieve the best unique-neighbor expansion possible. For example, all small sets S in a random d-regular graph have roughly $(d-1)|S|$ unique neighbors (see [HLW06, Theorem 4.16]). However, there is still a gap between the quality of the expansion of explicit constructions known and that of random graphs. Indeed, the construction of [CRVW02] achieves lossless expansion in bipartite graphs of arbitrary imbalance for small sets contained on the left side, whereas the only previous constructions to yield unique-neighbor expansion for all small sets [AC02, Bec16] give regular graphs of degree bounded by 6.

The contribution of this work is an explicit construction of unique-neighbor expanders that make progress towards the bounds guaranteed by random graphs.

Theorem 1.1 (See Theorem 4.4 for formal statement)**.** *For every $\beta \in (0, 1/2]$ and large enough d_1, d_2 with $1 \geqslant \frac{d_1}{d_2} \geqslant \frac{\beta}{1-\beta}$, there is an infinite family of $(2d_1, 2d_2)$-biregular graphs $(Z_n)_{n \geqslant 1}$ such that all $S \subseteq V(Z_n)$ of size at most $\frac{1}{d_1^3} \cdot |V(Z_n)|$ have at least $\Omega_\beta(d_1 \cdot |S|)$ unique-neighbors.*

Remark 1.2. Theorem 1.1 gives the first construction of two-sided unique-neighbor expanders where the left and right side have unequal sizes.[1] It also gives the only construction besides the one-sided lossless expander construction of [CRVW02] 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 1.3. Our techniques also straightforwardly generalize to constructing families of bounded degree k-partite unique-neighbor expanders for any distribution $(\beta_1, \ldots, \beta_k)$ of vertices across partitions.

A first-order question one might have is why we seemingly need distinct constructions from the rich set of spectral expander constructions available. 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.

However, the connection between the vertex 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 ε, any set S with at most εn vertices has roughly at least $|S| \cdot d/2$ distinct neighbors. Observe that once the vertex expansion of a set exceeds $d/2$, it starts to have unique-neighbors. Strikingly, the $d/2$ bound

[1] Throughout the paper we will assume the left side is larger, i.e. $d_1 \leqslant d_2$.

Table 1: Comparison of our Theorem 1.1 with prior work.

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

*Non-bipartite construction that can be made bipartite by passing to the double cover.
[†] d here refers to the degree of the left vertex set.
[‡] "Aspect ratio" refers to the ratio between the sizes of the left and right vertex sets.

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 Ramanujan graphs contain sublinear-sized sets with *zero* unique neighbors (see also [Kah95, MM21] for examples of near-Ramanujan graphs exhibiting a similar property).

1.1 Construction

Our construction is based on a special graph product, which we call 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].

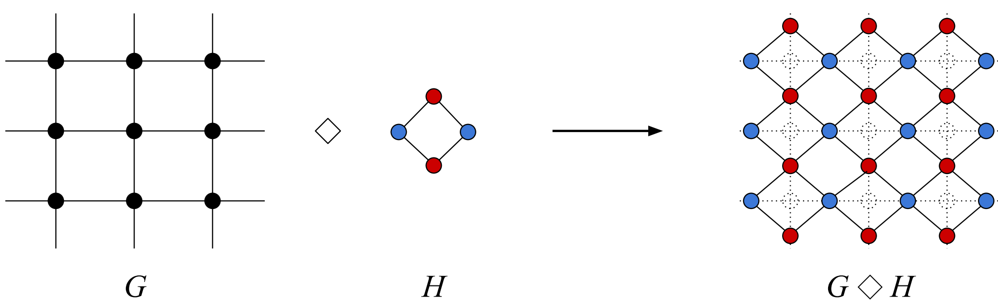

Figure 1: An example of the line product.

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 crucial lemma (Lemma 3.3) we will prove 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.

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 5.1).

Thus, we give a slight generalization of an embedding theorem given in [FK16] between Erdős–Rényi graphs and random regular graphs (Lemma 5.2) which allows us to extend the analysis to random regular graphs.

1.2 Related work

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, can be interpreted as taking the line product of a large Ramanujan graph with the 8-vertex 3-regular 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 independent work of Asherov and Dinur [AD23] to obtain one-sided unique-neighbor expanders of aspect ratio α for all $\alpha \geqslant 1$ 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.

Notably, both of these works require the precise eigenvalue bounds that come from having a Ramanujan graph, as well as a "better-than-naive" way to translate an eigenvalue bound to a bound on the average degree of small subgraphs [Kah95, Theorem 4.2]. We similarly require an $O(\sqrt{D})$ bound on the average degree of small subgraphs in a D-regular expander, but the precise constant in front of $\sqrt{D}$ does not make or break unique-neighbor expansion, and only affects the number of unique-neighbors of a set.

The remarkable work of Capalbo, Reingold, Vadhan, and 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.

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 ℓ_p-*spread* subspaces as well as in [BGI$^+$08, GMM22] to construct matrices with the ℓ_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 using unused vertices.

Recently, [DFHT21, 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 showed instances which are highly unsatisfiable but even $\Omega(n)$ levels of SoS fail to refute them (i.e., perfect completeness). While their constructions are based on high-dimensional expanders, it is known that perfect completeness holds as long as the factor graph of the instance is a unique-neighbor expander [Gri01, Sch08].[2] It would be interesting to see if our result yields simpler constructions of hard instances for SoS (without resorting to high-dimensional expanders).

We also mention the importance of *two-sided* lossless expanders in quantum LDPC codes [LH22b]. In the wake of recent breakthroughs in quantum LDPC codes (qLDPC) and c^3-LTCs [DEL$^+$22, PK22, LH22a, LZ22], Lin and Hsieh [LH22b] constructed good qLDPC codes with *linear time decoders* assuming the existence of 2-sided lossless expander graphs (though more recently qLDPCs with linear time decoders have been constructed [DHLV22, GPT22, LZ23]).

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 a vertex v and any subset of edges F, we use $F(v)$ to denote the set of edges in F incident to v. For a set of vertices S, we use $e(S)$ to denote the number of edges with both endpoints in S, and $e(S,T)$ to denote the number of $(u,v) \in S \times T$ such that $\{u,v\}$ is an edge (for disjoint sets S and T). We denote the eigenvalues of the normalized adjacency matrix of G by $1 = \lambda_1(G) \geqslant \dots \geqslant \lambda_n(G) \geqslant -1$. We say G is a λ-spectral expander if $\lambda_2(G) \leqslant \lambda$.[3]

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 $\boldsymbol{G} \sim \mathbb{G}_{n_1,n_2,m}$ 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 $\boldsymbol{H} \sim \mathbb{G}_{n_1,n_2,p}$ for $p \in (0,1)$ to denote a random graph such that each potential edge is included with probability p. Similarly, we use $\boldsymbol{R} \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ 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 sets in spectral expanders have a bounded number of edges.

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

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

[2] In fact, perfect completeness requires that the number of unique-neighbors scales linearly in the size of the set.

[3] This is in contrast to most scenarios where one requires both $\lambda_2(G)$ as well as $-\lambda_n(G)$ to be bounded by λ.

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 $\|u\|_2^2 = |S|(1 - \frac{|S|}{n}) \leqslant |S|$. Then, $2e(S) = \vec{1}_S^\top A \vec{1}_S \leqslant \frac{D|S|^2}{n} + \lambda_2(G) \cdot D\|u\|_2^2 \leqslant D|S|(\varepsilon + \lambda)$. $\square$

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 $S \subseteq V(F)$, $2e(S) \leqslant \gamma|S|$. Write $V(F) = F_\ell \sqcup F_h$ where F_ℓ comprises all vertices v such that $\deg(v) \leqslant 2\gamma$ and F_h to denote the remaining vertices. Then:*

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

Proof. On one hand, by assumption we have:

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

On the other hand,

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

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

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

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

2.2 Concentration inequalities

Fact 2.3 ([Hoe63])**.** *Fix $1 \leqslant n \leqslant N$. Let $\Omega = (x_1, \dots, x_N)$ be a finite set of points in $\mathbb{R}$. Let $\boldsymbol{X}_1, \dots, \boldsymbol{X}_n$ be a random sample without replacement from Ω, let $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$ be a random sample with replacement from Ω, and let $\boldsymbol{X} = \sum_{i=1}^n \boldsymbol{X}_i$ and $\boldsymbol{Y} = \sum_{i=1}^n \boldsymbol{Y}_i$. Suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous and convex, then*

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

Fact 2.3 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 2.4 (Concentration for sampling without replacement)**.** *Fix $1 \leqslant k, n \leqslant N$. Let $S \subseteq [N]$, let $\boldsymbol{T}$ be a random sample of n elements from $[N]$ without replacement, and let $p = \frac{n}{N}$. Then, for all $\delta \in (0,1)$,*

$$\mathbf{Pr}\left[\big||S \cap \boldsymbol{T}| - p|S|\big| \geqslant \delta p|S|\right] \leqslant 2\exp\left(-\frac{1}{3}\delta^2 p|S|\right),$$
$$\mathbf{Pr}\left[\big||S \cap \boldsymbol{T}| - p|S|\big| \geqslant \delta(1-p)|S|\right] \leqslant 2\exp\left(-\frac{1}{3}\delta^2 (1-p)|S|\right).$$

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

By Fact 2.3, we can apply the Chernoff bound as if the $\boldsymbol{X}_i$'s are sampled with replacement. Let $\boldsymbol{Y}_1, \dots, \boldsymbol{Y}_n$ be samples from Ω with replacement, then $\boldsymbol{Y}_i$ are i.i.d. Bernoulli random variables with $\mathbf{E}[\boldsymbol{Y}_i] = \frac{|S|}{N}$, and $\mathbf{E}\sum_{i=1}^n \boldsymbol{Y}_i = \frac{n|S|}{N} = p|S|$. The first inequality then follows from the Chernoff bound.

For the second inequality, we look at $|S \setminus \boldsymbol{T}|$. Due to symmetry, sampling n elements from $[N]$ without replacement is equivalent to sampling $N-n$ elements and taking the complement. Let $\boldsymbol{X}'_1, \dots, \boldsymbol{X}'_{N-n}$ be a random sample without replacement from Ω. Then, $|S \setminus \boldsymbol{T}|$ is distributed as $\boldsymbol{X}' = \sum_{i=1}^{N-n} \boldsymbol{X}'_i$, and $\mathbf{E}[\boldsymbol{X}'] = (N-n)\frac{|S|}{N} = (1-p)|S|$. By Fact 2.3 (transferring to sampling with replacement) and the Chernoff bound,

$$\mathbf{Pr}\left[\big||S \setminus \boldsymbol{T}| - (1-p)|S|\big| \geqslant \delta(1-p)|S|\right] \leqslant 2\exp\left(-\frac{1}{3}\delta^2(1-p)|S|\right).$$

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

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 d-regular 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 $E(v)$ 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 $\mathrm{UN}_H(S)$ to be the set of unique-neighbors of S. The *unique-neighbor expansion profile* of a graph H, denoted P_H, is defined:

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

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

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

Proof. Let $T \subseteq V(Z)$ be such that $|T| \leqslant \frac{\varepsilon}{D} \cdot |V(Z)| = \frac{\varepsilon}{2} \cdot |V(G)|$, and let $S \subseteq V(G)$ be the set of vertices of G touched by T.[4] Note that $|S| \leqslant 2|T| \leqslant \varepsilon|V(G)|$. Then:

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

[4] Recall that T is also a collection of edges in G.

Each summand in the first term of the right-hand side can be bounded from below using P_H. A vertex $\{v_1, v_2\} \in Z$ (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 $|S| \leqslant \varepsilon|V(G)|$, by the assumption on G, the induced subgraph $G[S]$ satisfies the assumption of Lemma 2.2. Thus, defining S_ℓ as the set of all vertices with $\deg_{G[S]}(v) \leqslant 2\gamma$, and S_h as the remaining vertices in S, we have that $2e(S) \leqslant 3\sum_{v\in S_\ell} \deg_{G[S]}(v) \leqslant 6\gamma|S_\ell|$. Then,

$$\begin{aligned}
|\mathrm{UN}_Z(T)| &\geqslant \sum_{v\in S} |T(v)| \cdot P_H(|T(v)|) - 2e(S) \\
&\geqslant \sum_{v\in S} |T(v)| \cdot P_H(|T(v)|) - 6\gamma|S_\ell| \\
&= \sum_{v\in S_\ell} |T(v)| \cdot \left(P_H(|T(v)|) - \frac{6\gamma}{|T(v)|}\right) + \sum_{v\in S_h} |T(v)| \cdot P_H(|T(v)|) \\
&\geqslant \frac{1}{2} \sum_{v\in S_\ell} |T(v)| \cdot P_H(|T(v)|),
\end{aligned}$$

where the last inequality is by $|T(v)| = \deg_{G[S]}(v) \leqslant 2\gamma$ for all $v \in S_\ell$ and the assumption that $P_H(t) \geqslant 12\gamma/t$ for $t \leqslant 2\gamma$. Then by the fact that $P_H(t)$ is monotone non-increasing with t and Lemma 2.2 (namely $\sum_{v\in S_\ell} |T(v)| \geqslant \frac{2}{3}e(S) \geqslant \frac{2}{3}|T|$):

$$\frac{|\mathrm{UN}_Z(T)|}{|T|} \geqslant \frac{1}{2|T|} \sum_{v\in S_\ell} |T(v)| \cdot P_H(2\gamma) \geqslant \frac{P_H(2\gamma)}{3}. \qquad \square$$

4 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 $G := \mathrm{Cay}(\Gamma_n, A)$ 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 $D \in \mathbb{N}$, $D \geqslant 3$, 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 $G := \mathrm{Cay}(\Gamma_n, A)$ is a D-regular graph such that*

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

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

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, $G := \mathrm{Cay}(\Gamma_n, A)$ is D-regular.

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

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

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 $L > 0$ 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. $1 \geqslant \frac{d_1}{d_2} \geqslant \frac{\beta}{1-\beta}$,
3. $\theta\sqrt{D}/2 \leqslant d_1 + d_2 \leqslant \theta\sqrt{D}$,

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:

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

for $1 \leqslant t \leqslant L\sqrt{D}$ where $o_D(1)$ hides constant factors depending only on β, θ and L.

We defer the proof of Lemma 4.3 to Section 5, and prove our main theorem below.

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

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

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 $L = 18$, $\theta = \frac{40L}{\beta}$, and choose d_0 to be large enough such that $d_0 \geqslant \frac{8}{\beta} e^{L\theta}$ and such that for $D \geqslant 4d_0^2/\theta^2$, the $o_D(1)$ terms in Lemma 4.3 and all subsequent occurrences in this proof are smaller than 0.1. For $d_1, d_2 \geqslant d_0$, 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 graph (d_1, d_2)-biregular graph H with D_1 left vertices, D_2 right vertices, and

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

for $t \leqslant L\sqrt{D}$.

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

$$(2) \geqslant \frac{12\gamma}{t} = \frac{6L\sqrt{D}}{t}, \quad \text{for } 1 \leqslant t \leqslant 2\gamma = L\sqrt{D}. \tag{3}$$

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

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

for constant $\delta = 0.3e^{-L\theta}$. Since $D \leqslant \frac{4}{\theta^2}(d_1 + d_2)^2$, $\beta(d_1 + d_2) \leqslant d_1$, and $\beta\theta = 40L$, we have $\frac{1}{D^{3/2}} \geqslant \frac{1}{d_1^3}$. 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 $x \leqslant 1$ and monotone decreasing for $x \geqslant 1$, hence for $x \in [a, b]$, $xe^{-x} \geqslant \min\{ae^{-a}, be^{-b}\}$. Thus, for $1 \leqslant t \leqslant L\sqrt{D}$,

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

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

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

Finally, we show that Z_n is a $(2d_1, 2d_2)$-biregular graph. Since $G_n = \mathrm{Cay}(\Gamma, A)$ 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 φ 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. □

5 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 5.1, and then transfer the result to random regular graphs via a coupling articulated in Lemma 5.2. See Section 2 for the notations of various random bipartite graph models.

Lemma 5.1. *Let $\boldsymbol{H} \sim \mathbb{G}_{n_1,n_2,p}$ where $n_1 \geqslant n_2$. Then with probability $1 - O\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$, for all t:*

$$P_{\boldsymbol{H}}(t) \geqslant p(1-p)^{t-1} n_2 - \sqrt{4p(1-p)^{t-1} n_1 \log n_1}.$$

Lemma 5.2 (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 $\boldsymbol{H} \sim \mathbb{G}_{n_1,n_2,p}$ and $\boldsymbol{R} \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ such that*

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

We first give a proof of Lemma 4.3 assuming the Lemmas 5.1 and 5.2. Lemma 5.1 is proved later in this section, and Lemma 5.2 is proved in Section 6.

Proof of Lemma 4.3. Recall that we would like to show that for $1 \geqslant \frac{d_1}{d_2} = \frac{D_2}{D_1} \geqslant \frac{\beta}{1-\beta}$ and $\theta\sqrt{D}/2 \leqslant d_1 + d_2 \leqslant \theta\sqrt{D}$, 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 5.2 there is a coupling between $\boldsymbol{R} \sim \mathbb{R}_{D_1,D_2,d_1,d_2}$ and $\boldsymbol{H} \sim \mathbb{G}_{D_1,D_2,p}$ such that $\boldsymbol{H} \subset \boldsymbol{R}$ 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_β is a constant depending on β and θ. Note that $p \leqslant \frac{d_1}{D_2} = \frac{d_1+d_2}{D_1+D_2} \leqslant \frac{\theta}{\sqrt{D}}$.

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 $\boldsymbol{R} \setminus \boldsymbol{H}$. Consequently: $P_{\boldsymbol{R}}(t) \geqslant P_{\boldsymbol{H}}(t) - o(\sqrt{D})$. Thus, it suffices to lower bound $P_{\boldsymbol{H}}(t)$ to obtain a lower bound on $P_{\boldsymbol{R}}(t)$.

By Lemma 5.1, for $t \geqslant 1$,

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

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

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

which completes the proof. $\square$

We now prove Lemma 5.1: 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 5.1. Write $S \subseteq V(\boldsymbol{H})$, write $S := S_L \cup S_R$ where $S_L := S \cap L(\boldsymbol{H})$ and $S_R := S \cap R(\boldsymbol{H})$. Observe that $|\mathrm{UN}_{\boldsymbol{H}}(S)| = |\mathrm{UN}_{\boldsymbol{H}}(S_L)| + |\mathrm{UN}_{\boldsymbol{H}}(S_R)|$. Therefore, without loss of generality we can study S completely in $L(\boldsymbol{H})$ or $R(\boldsymbol{H})$.

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

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

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

$$\mathbf{Pr}[|\mathrm{UN}_{\boldsymbol{H}}(S)| \leqslant q_t n_1 - s\sqrt{q_t n_1}] \leqslant \exp(-s^2/2),$$

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

$$\forall S \subseteq R(\boldsymbol{H}) \text{ s.t. } |S| = t : |\mathrm{UN}_{\boldsymbol{H}}(S)| \geqslant q_t n_1 - \sqrt{4 q_t n_1 t \log n_2}$$

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

$$\forall S \subseteq L(\boldsymbol{H}) \text{ s.t. } |S| = t : |\mathrm{UN}_{\boldsymbol{H}}(S)| \geqslant q_t n_2 - \sqrt{4 q_t n_2 t \log n_1}$$

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

$$\frac{|\mathrm{UN}_{\boldsymbol{H}}(S)|}{|S|} \geqslant p(1-p)^{t-1} n_2 - \sqrt{4p(1-p)^{t-1} n_1 \log n_1}.$$

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

6 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 6.1 (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 $\gamma \geqslant C\left(\frac{d_1 d_2}{m} + \frac{\log m}{\min\{d_1, d_2\}}\right)^{1/3}$, then for $\widetilde{m} \leqslant \lfloor (1-\gamma) m \rfloor$, there is a joint distribution of $\boldsymbol{G} \sim \mathbb{G}_{n_1, n_2, \widetilde{m}}$ and $\boldsymbol{R} \sim \mathbb{R}_{n_1, n_2, d_1, d_2}$ such that*

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

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

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

To prove Theorem 6.1, we need to introduce some more notation. With slight abuse of notation, we write

$$\boldsymbol{G} = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m), \quad \boldsymbol{R} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m)$$

to be random orderings of the edges. Moreover, for $t = 1, 2, \ldots, m$, we define random variables $\boldsymbol{G}_t = (\boldsymbol{e}_1, \ldots, \boldsymbol{e}_t)$ and $\boldsymbol{R}_t = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_t)$.

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

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

This motivates the following definition.

Definition 6.2. Fix $\varepsilon \in (0,1)$. We define $\boldsymbol{A}_{\varepsilon,t}$ to be the event that for all $e \in K_{n_1,n_2} \setminus \boldsymbol{R}_t$,

$$\mathbf{Pr}[\boldsymbol{f}_{t+1} = e | \boldsymbol{R}_t] \geqslant \frac{1-\varepsilon}{n_1 n_2 - t}. \tag{4}$$

Further, we define the stopping time

$$\boldsymbol{T}_\varepsilon = \max\{u : \boldsymbol{A}_{\varepsilon,t} \text{ occurs for all } t \leqslant u\}.$$

Intuitively, suppose we sample the edges of $\boldsymbol{R} \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ one by one. At each time step $t \leqslant \boldsymbol{T}_\varepsilon$, 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 6.1 is to show that $\boldsymbol{T}_\varepsilon$ 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 6.3 (Large stopping time). *There is a universal constant C such that if $\varepsilon \in (0,1)$ satisfies $\varepsilon \geqslant C\left(\frac{d_1 d_2}{m} + \frac{\log m}{\min\{d_1,d_2\}}\right)^{1/3}$, then $\boldsymbol{T}_\varepsilon \geqslant (1-\varepsilon)m$ with probability $1-o(1)$.*

We will defer the proof to Section 6.2. This lemma suffices to prove Theorem 6.1.

Proof of Theorem 6.1 by Lemma 6.3. Recall that $m = n_1 d_1 = n_2 d_2$. We will define a graph process $\boldsymbol{R}' = (\boldsymbol{f}'_1, \dots, \boldsymbol{f}'_m)$ coupled with $\boldsymbol{G} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_m) \sim \mathbb{G}_{n_1,n_2,m}$ and show that (1) $\boldsymbol{f}'_t$ and $\boldsymbol{f}_t$ have the same conditional distribution, and (2) $\boldsymbol{R}' \cap \boldsymbol{G}$ is large and contains a random subgraph $\boldsymbol{G}' \subset \boldsymbol{G}$ of size $\widetilde{m} < m$ with high probability. Since a subgraph $\boldsymbol{G}'$ is also distributed as $\mathbb{G}_{n_1,n_2,\widetilde{m}}$, this gives a coupling between $\boldsymbol{G}' \sim \mathbb{G}_{n_1,n_2,\widetilde{m}}$ and $\boldsymbol{R}' \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ such that $\boldsymbol{G}' \subset \boldsymbol{R}'$ with high probability.

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

$$p_{t+1}(e|\boldsymbol{R}_t) := \mathbf{Pr}\left[\boldsymbol{f}_{t+1} = e | \boldsymbol{R}_t\right]$$

which is at least $\frac{1-\varepsilon}{n_1 n_2 - t}$ for $t \leqslant \boldsymbol{T}_\varepsilon$ by Definition 6.2.

The graph process $\boldsymbol{R}'$ is sampled as follows: at time step $t \leqslant \boldsymbol{T}_\varepsilon$,

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

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

   Note that this is a valid probability distribution over $K_{n_1,n_2} \setminus \boldsymbol{R}'_t$ since the above is non-negative due to Eq. (4) and the sum is 1 because $|K_{n_1,n_2} \setminus \boldsymbol{R}'_t| = n_1 n_2 - t$.
3. Fix any bijection map $\boldsymbol{h} : \boldsymbol{R}'_t \setminus \boldsymbol{G}_t \to \boldsymbol{G}_t \setminus \boldsymbol{R}'_t$. Set

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

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

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

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

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

$$= p_{t+1}(e|\boldsymbol{R}'_t).$$

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

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

$$\boldsymbol{S} := \left\{ \boldsymbol{f}'_t : \boldsymbol{\xi}_t = 1, 0 \leqslant t \leqslant (1-\varepsilon)m \right\} \subset \boldsymbol{R}'.$$

By Lemma 6.3, $\boldsymbol{T}_\varepsilon \geqslant (1-\varepsilon)m$ with probability $1 - o(1)$. Conditioned on this, we know that all edges in $\boldsymbol{S}$ lie in $\boldsymbol{G}$. Moreover, $|\boldsymbol{S}|$ is distributed as $\mathrm{Bin}((1-\varepsilon)m, 1-\varepsilon)$, so $\mathbf{E}[|\boldsymbol{S}|] = (1-\varepsilon)^2 m$, and by Chernoff bound,

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

Let $\gamma = 3\varepsilon$ and fix $\widetilde{m} \leqslant \lfloor (1-\gamma)m \rfloor \leqslant (1-\varepsilon)^3 m$. We now take the first $\widetilde{m}$ edges from $\boldsymbol{S} \subset \boldsymbol{R}'$, and the resulting graph $\boldsymbol{G}'$ is distributed as $\mathbb{G}_{n_1,n_2,\widetilde{m}}$. Thus, we have obtained a joint distribution between $\boldsymbol{G}' \sim \mathbb{G}_{n_1,n_2,\widetilde{m}}$ and $\boldsymbol{R}' \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ such that $\boldsymbol{G}' \subset \boldsymbol{R}'$ 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 $\boldsymbol{m}' \sim \mathrm{Bin}(n_1 n_2, p)$, and (2) sample $\boldsymbol{H} \sim \mathbb{G}_{n_1,n_2,\boldsymbol{m}'}$ coupled with $\boldsymbol{R}' \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ as described before (if $\boldsymbol{m}' > m$ then sample the extra edges randomly). By the Chernoff bound,

$$\mathbf{Pr}[\boldsymbol{m}' \geqslant (1+\gamma)n_1 n_2 p] \leqslant \exp\left(-\gamma^2 n_1 n_2 p/3\right) = o(1).$$

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

6.1 Random graph extension

To prove Lemma 6.3, 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 6.4 (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 $i \leqslant t$. 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 $\boldsymbol{S}_G$ 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 6.4 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 6.5 (Random multigraph extension)**.** Given a graph $G = (e_1, \dots, e_t)$ of size t, we denote $\boldsymbol{M}_G$ 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 $\boldsymbol{M}_G$ to be e_i for $i \leqslant t$, and set the $(t + j)$-th edge of $\boldsymbol{M}_G$ to be $(U[j], V[j])$ for each $1 \leqslant j \leqslant m - t$.

Fact 6.6. *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*

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

In particular, conditioned on $\boldsymbol{M}_G$ *being simple, it has the same distribution as* $\boldsymbol{S}_G$*.*

The main ingredient of the proof of Lemma 6.3 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 6.7. *Let* $\varepsilon \geqslant C\left(\frac{d_1 d_2}{m}\right)^{1/3} + C\sqrt{\frac{\log m}{\min\{d_1, d_2\}}}$ *for a large enough constant* C*. Let* G *be a bipartite graph with* $t \leqslant (1-\varepsilon)m$ *edges such that* $\mathcal{S}_G = \mathcal{S}_G(n_1, n_2, d_1, d_2)$ *is non-empty. If* $\deg_G(u) \leqslant (1 - \varepsilon/2)d_1$ *and* $\deg_G(v) \leqslant (1 - \varepsilon/2)d_2$ *for all* $u \in [n_1]$*,* $v \in [n_2]$*, then for every* $e, e' \notin G$*, we have*

$$\frac{\mathbf{Pr}[\boldsymbol{M}_{G \cup e'} \in \mathcal{S}_{G \cup e'}]}{\mathbf{Pr}[\boldsymbol{M}_{G \cup e} \in \mathcal{S}_{G \cup e}]} \geqslant 1 - \frac{\varepsilon}{2}.$$

Typical random extension. The following is analogous to Lemma 11.20 in [FK16] and will be used to prove Lemma 6.7. It roughly states that a random extension $\boldsymbol{S}_G$ of a graph G behaves nicely.

Lemma 6.8. *Let* $\varepsilon m \geqslant 4 d_1 d_2$*. Let* G *be a graph with* $t \leqslant (1-\varepsilon)m$ *edges such that* $\mathcal{S}_G$ *is non-empty, and let* $\boldsymbol{S}_G$ *be a uniform sample from* $\mathcal{S}_G$*. For every* $e \notin G$*, we have*

$$\mathbf{Pr}[e \in \boldsymbol{S}_G] \leqslant \frac{2 d_1 d_2}{\varepsilon m}.$$

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

$$\mathbf{Pr}\left[\deg_{\boldsymbol{S}_G | G}(u_1, u_2) > \ell\right] \leqslant 2^{-(\ell - \ell_1)},$$

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

$$\mathbf{Pr}\left[\deg_{\boldsymbol{S}_G | G}(v_1, v_2) > \ell\right] \leqslant 2^{-(\ell - \ell_2)}.$$

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

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

Then, $\mathbf{Pr}[e \in \boldsymbol{S}_G] = \frac{|\mathcal{G}_{\in e}|}{|\mathcal{G}_{\in e}|+|\mathcal{G}_{\notin e}|} \leqslant \frac{|\mathcal{G}_{\in e}|}{|\mathcal{G}_{\notin e}|}$. 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 $\deg_B(H) \geqslant m - t - d_1 d_2$.

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, $\deg_B(H') \leqslant d_1 d_2$.

Since B is a bipartite graph, we must have

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

since $t \leqslant (1 - \varepsilon) m$ and $\varepsilon m \geqslant 4 d_1 d_2$.

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 $\ell = 0, 1, \ldots$, and we adopt a similar strategy of constructing an auxiliary bipartite graph B_ℓ 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 ℓ choices for w and at least $m - t - 2 d_1 d_2$ choices for (u', w'), thus $\deg_{B_\ell}(H) \geqslant \ell(m - t - 2 d_1 d_2) \geqslant \frac{1}{2} \ell \varepsilon m$. 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 $\deg_{B_\ell}(H') \leqslant d_1^2 d_2$. Therefore,

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

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

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

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

We can now complete the proof of Lemma 6.7.

Proof of Lemma 6.7. We will write $\boldsymbol{M} = \boldsymbol{M}_{G \cup e}$ and $\boldsymbol{M}' = \boldsymbol{M}_{G \cup e'}$ for simplicity. We will construct a coupling of $\boldsymbol{M}$ and $\boldsymbol{M}'$ such that they differ in at most 3 positions. Given $\boldsymbol{M}$ and $e = (u, v)$, $e' = (u', v')$, we perform a switching operation:

1. Delete e and add e' to $\boldsymbol{M}$.

2. Recall U and V of length $m-(t+1)$ defined in Definition 6.5.

   - 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 6.5.

Note that step 2 is equivalent to sampling a random edge (u', w_R) incident to u' in $\boldsymbol{M} \setminus (G \cup e)$ 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 $\boldsymbol{M}^*$. 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 $\boldsymbol{M}^*$ has the same distribution as $\boldsymbol{M}'$.

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

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

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 6.8.

First, by Fact 6.6 and Lemma 6.8, $\mathbf{Pr}[e' \in \boldsymbol{M} | \boldsymbol{M} \in \mathcal{S}_{G\cup e}] = \mathbf{Pr}[e' \in \boldsymbol{S}_{G\cup e}] \leqslant \frac{2d_1 d_2}{\varepsilon m} \leqslant \varepsilon/8$ due to our lower bound on ε. Moreover, we have $\deg_{\boldsymbol{M}|G\cup e}(u', u) > \ell_1 + \log m$ and $\deg_{\boldsymbol{M}|G\cup e}(v', v) > \ell_2 + \log m$ with probability $\leqslant \frac{1}{m}$. Thus,

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

Now suppose $\boldsymbol{M} \in \mathcal{G}_{\text{nice}}$. A parallel edge can occur in three ways: (1) if we replace $(u', w_R) \in \boldsymbol{M} \setminus (G \cup e)$ with (u, w_R) but $(u, w_R) \in \boldsymbol{M}$ 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,

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

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

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

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

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

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

completing the proof. □

6.2 Proof of Lemma 6.3

Degree bounds. We first prove a degree concentration bound. Recall that $\boldsymbol{R} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m) \sim \mathbb{R}_{n_1,n_2,d_1,d_2}$ is a (ordered) random (d_1, d_2)-biregular graph on $L = [n_1]$, $R = [n_2]$, and we write $\boldsymbol{R}_t = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_t)$.

Lemma 6.9 (Degree concentration)**.** *Consider the random graph process $\boldsymbol{R}_t = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_t)$ for $t = 0, 1, \ldots, m$. Let $\varepsilon \in (0,1)$. With probability $1 - O(m^{-1})$, for all $t \leqslant (1-\varepsilon)m$, letting $p = \frac{t}{m}$, we have*

$$\left|\deg_{\boldsymbol{R}_t}(u) - pd_1\right| \leqslant 3\sqrt{(1-p)d_1 \log m}, \quad \forall u \in [n_1],$$
$$\left|\deg_{\boldsymbol{R}_t}(v) - pd_2\right| \leqslant 3\sqrt{(1-p)d_2 \log m}, \quad \forall v \in [n_2].$$

Proof. Fix a time step $1 \leqslant t \leqslant (1-\varepsilon)m$ and a vertex $u \in [n_1]$. Since $\boldsymbol{R}_t$ can be viewed as sampling t edges from m total edges without replacement, we can apply the second inequality in Lemma 2.4 with $|S| = d_1$ and $\delta = 3\sqrt{\frac{\log m}{(1-p)d_1}}$ (note that $1 - p \geqslant \varepsilon > 0$):

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

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 $t \leqslant (1-\varepsilon)m$ and $u \in [n_1]$, $v \in [n_2]$. □

We are now ready to prove Lemma 6.3.

Proof of Lemma 6.3. We would like to prove that for all $e \in K_{n_1,n_2} \setminus \boldsymbol{R}_t$, $\mathbf{Pr}[\boldsymbol{f}_{t+1} = e | \boldsymbol{R}_t] \geqslant \frac{1-\varepsilon}{n_1 n_2 - t}$ for every $t \leqslant (1-\varepsilon)m$. It suffices to prove that for every $e, e' \in K_{n_1,n_2} \setminus \boldsymbol{R}_t$,

$$\frac{\mathbf{Pr}[\boldsymbol{f}_{t+1} = e' | \boldsymbol{R}_t]}{\mathbf{Pr}[\boldsymbol{f}_{t+1} = e | \boldsymbol{R}_t]} \geqslant 1 - \varepsilon,$$

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

Recalling the definition of extensions in Definition 6.4, we have

$$\frac{\mathbf{Pr}[\boldsymbol{f}_{t+1} = e' | \boldsymbol{R}_t]}{\mathbf{Pr}[\boldsymbol{f}_{t+1} = e | \boldsymbol{R}_t]} = \frac{|\mathcal{S}_{\boldsymbol{R}_t \cup e'}|}{|\mathcal{S}_{\boldsymbol{R}_t \cup e}|}. \tag{5}$$

We now consider multigraph extensions of $\boldsymbol{R}_t \cup e$ and $\boldsymbol{R}_t \cup e'$. By Fact 6.6,

$$\mathbf{Pr}\left[\boldsymbol{M}_{\boldsymbol{R}_t \cup e} \in \mathcal{S}_{\boldsymbol{R}_t \cup e}\right] = |\mathcal{S}_{\boldsymbol{R}_t \cup e}| \cdot \frac{\prod_{u \in [n_1]}(d_1 - \deg_{\boldsymbol{R}_t \cup e}(u))! \cdot \prod_{v \in [n_2]}(d_2 - \deg_{\boldsymbol{R}_t \cup e}(v))!}{((m-t)!)^2},$$

$$\mathbf{Pr}\left[\boldsymbol{M}_{\boldsymbol{R}_t\cup e'} \in \mathcal{S}_{\boldsymbol{R}_t\cup e'}\right] = |\mathcal{S}_{\boldsymbol{R}_t\cup e'}| \cdot \frac{\prod_{u\in[n_1]}(d_1 - \deg_{\boldsymbol{R}_t\cup e'}(u))! \cdot \prod_{v\in[n_2]}(d_2 - \deg_{\boldsymbol{R}_t\cup e'}(v))!}{((m-t)!)^2}.$$

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

$$\frac{\mathbf{Pr}[\boldsymbol{M}_{\boldsymbol{R}_t\cup e'} \in \mathcal{S}_{\boldsymbol{R}_t\cup e'}]}{\mathbf{Pr}[\boldsymbol{M}_{\boldsymbol{R}_t\cup e} \in \mathcal{S}_{\boldsymbol{R}_t\cup e}]} = \frac{|\mathcal{S}_{\boldsymbol{R}_t\cup e'}|}{|\mathcal{S}_{\boldsymbol{R}_t\cup e}|} \cdot \frac{(d_1 - \deg_{\boldsymbol{R}_t}(u))(d_2 - \deg_{\boldsymbol{R}_t}(v))}{(d_1 - \deg_{\boldsymbol{R}_t}(u'))(d_2 - \deg_{\boldsymbol{R}_t}(v'))}. \tag{6}$$

Let $p = \frac{t}{m}$. By the concentration of degrees (Lemma 6.9), with probability $1 - O(m^{-1})$, for $t \leqslant (1-\varepsilon)m$ (hence $1 - p \geqslant \varepsilon$),

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

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

$$(6) \leqslant \frac{|\mathcal{S}_{\boldsymbol{R}_t\cup e'}|}{|\mathcal{S}_{\boldsymbol{R}_t\cup e}|}(1 + \varepsilon/2).$$

On the other hand, the degree bounds also allow us to apply Lemma 6.7:

$$(6) \geqslant 1 - \frac{\varepsilon}{2}.$$

Therefore, by Eq. (5) we have $\frac{\mathbf{Pr}[f_{t+1}=e' | \boldsymbol{R}_t]}{\mathbf{Pr}[f_{t+1}=e | \boldsymbol{R}_t]} \geqslant 1 - \varepsilon$ with probability $1 - O(m^{-1})$, completing the proof. □

References

[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, 3

[AD23] Ron Asherov and Irit Dinur. Bipartite unique-neighbour expanders via Ramanujan graphs. *arXiv preprint arXiv:2301.03072*, 2023. 2, 3

[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, 3

[Bec16] Oren Becker. Symmetric unique neighbor expanders and good LDPC codes. *Discrete Applied Mathematics*, 211:211–216, 2016. 1, 2

[BGI+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, 3

[BV09] Eli Ben-Sasson and Michael Viderman. Tensor products of weakly smooth codes are robust. *Theory of Computing*, 5(1):239–255, 2009. 1, 3

[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, 3

[DEL+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. 1, 4

[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. 11

[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. 4

[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, 3

[FK16] Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016. 3, 11, 12, 14

[GLR10] Venkatesan Guruswami, James R Lee, and Alexander Razborov. Almost Euclidean subspaces of ℓ_1^N via expander codes. *Combinatorica*, 30(1):47–68, 2010. 1, 3

[GMM22] Venkatesan Guruswami, Peter Manohar, and Jonathan Mosheiff. ℓ_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, 3

[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. 4

[Gri01] Dima Grigoriev. Linear lower bound on degrees of Positivstellensatz calculus proofs for the parity. *Theoretical Computer Science*, 259(1-2):613–622, 2001. 4

[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. 3

[HL22] Max Hopkins and Ting-Chun Lin. Explicit lower bounds against ω (n)-rounds of sum-of-squares. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 662–673. IEEE, 2022. 1, 4

[HLW06] Shlomo Hoory, Nathan Linial, and Avi Wigderson. Expander graphs and their applications. *Bulletin of the American Mathematical Society*, 43(4):439–561, 2006. 1

[Hoe63] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association*, 58:13–30, 1963. 5

[Kah95] Nabil Kahale. Eigenvalues and expansion of regular graphs. *Journal of the ACM (JACM)*, 42(5):1091–1106, 1995. 1, 2, 3

[Kar11] Zohar S Karnin. Deterministic construction of a high dimensional ℓ_p section in ℓ_1^n for any $p < 2$. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 645–654, 2011. 1, 3

[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. 2

[LH22a] Ting-Chun Lin and Min-Hsiu Hsieh. c^3-Locally Testable Codes from Lossless Expanders. In *2022 IEEE International Symposium on Information Theory (ISIT)*, pages 1175–1180. IEEE, 2022. 4

[LH22b] Ting-Chun Lin and Min-Hsiu Hsieh. Good quantum LDPC codes with linear time decoder from lossless expanders. *arXiv preprint arXiv:2203.03581*, 2022. 4

[LPS88] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8:261–277, 1988. 3, 7

[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. 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. 4

[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)*, volume 198, page 96. Schloss Dagstuhl–Leibniz-Zentrum f {\" u} r Informatik, 2021. 2

[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. 3, 7

[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. 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. 3

[Sch08] Grant Schoenebeck. Linear level lasserre lower bounds for certain k-csps. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602. IEEE, 2008. 4

[Tan81] R Tanner. A recursive approach to low complexity codes. *IEEE Transactions on information theory*, 27(5):533–547, 1981. 3

[TSUZ07] Amnon Ta-Shma*, Christopher Umans, and David Zuckerman. Lossless condensers, unbalanced expanders, and extractors. *Combinatorica*, 27:213–240, 2007. 3