



## Spectra of lifted Ramanujan graphs

Eyal Lubetzky<sup>a,\*</sup>, Benny Sudakov<sup>b,1</sup>, Van Vu<sup>c,2</sup>

<sup>a</sup> *Microsoft Research, One Microsoft Way, Redmond, WA 98052, USA*

<sup>b</sup> *Department of Mathematics, UCLA, Los Angeles, CA 90095, USA*

<sup>c</sup> *Department of Mathematics, Rutgers, Piscataway, NJ 08854, USA*

Received 23 July 2010; accepted 16 March 2011

Available online 1 April 2011

Communicated by Gil Kalai

---

### Abstract

A random  $n$ -lift of a base-graph  $G$  is its cover graph  $H$  on the vertices  $[n] \times V(G)$ , where for each edge  $uv$  in  $G$  there is an independent uniform bijection  $\pi$ , and  $H$  has all edges of the form  $(i, u), (\pi(i), v)$ . A main motivation for studying lifts is understanding Ramanujan graphs, and namely whether typical covers of such a graph are also Ramanujan.

Let  $G$  be a graph with largest eigenvalue  $\lambda_1$  and let  $\rho$  be the spectral radius of its universal cover. Friedman (2003) [12] proved that every “new” eigenvalue of a random lift of  $G$  is  $O(\rho^{1/2}\lambda_1^{1/2})$  with high probability, and conjectured a bound of  $\rho + o(1)$ , which would be tight by results of Lubetzky and Greenberg (1995) [15]. Linial and Puder (2010) [17] improved Friedman’s bound to  $O(\rho^{2/3}\lambda_1^{1/3})$ . For  $d$ -regular graphs, where  $\lambda_1 = d$  and  $\rho = 2\sqrt{d-1}$ , this translates to a bound of  $O(d^{2/3})$ , compared to the conjectured  $2\sqrt{d-1}$ .

Here we analyze the spectrum of a random  $n$ -lift of a  $d$ -regular graph whose nontrivial eigenvalues are all at most  $\lambda$  in absolute value. We show that with high probability the absolute value of every nontrivial eigenvalue of the lift is  $O((\lambda \vee \rho) \log \rho)$ . This result is tight up to a logarithmic factor, and for  $\lambda \leq d^{2/3-\varepsilon}$  it substantially improves the above upper bounds of Friedman and of Linial and Puder. In particular, it implies that a typical  $n$ -lift of a Ramanujan graph is nearly Ramanujan.

© 2011 Elsevier Inc. All rights reserved.

*Keywords:* Random lifts; Spectral expanders; Ramanujan graphs

---

\* Corresponding author.

*E-mail addresses:* [eyal@microsoft.com](mailto:eyal@microsoft.com) (E. Lubetzky), [bsudakov@math.ucla.edu](mailto:bsudakov@math.ucla.edu) (B. Sudakov), [vanvu@math.rutgers.edu](mailto:vanvu@math.rutgers.edu) (V. Vu).

<sup>1</sup> The author is supported by NSF CAREER award 0812005 and a USA-Israeli BSF grant.

<sup>2</sup> The author is supported by research grants DMS-0901216 and AFOSAR-FA-9550-09-1-0167.

## 1. Introduction

Over the last quarter of a century, *expander* graphs have played a vital role in a remarkable variety of areas, ranging from combinatorics to discrete geometry to theoretical computer science, while exhibiting deep connections to algebra and number theory. Notable applications of expanders, to name just a few, include the design of efficient communication networks, explicit error-correcting codes with efficient encoding and decoding schemes, derandomization of randomized algorithms, compressed sensing and the study of metric embeddings. See the expository article of Sarnak [28] on these intriguing objects, as well as the comprehensive survey of Hoory, Linial and Wigderson [16] demonstrating their many applications.

Informally, an expander is a graph where every small subset of the vertices has a relatively large edge boundary (see Section 2.1 for a formal definition). Most applications utilize  $d$ -regular sparse expanders ( $d \geq 3$  fixed), where it is well known that expansion is related to the ratio between  $d$  and  $\lambda$ , the second largest eigenvalue in absolute value of the adjacency matrix. The smaller  $\lambda$  is, the better the graph expansion becomes. As a consequence of the Alon–Boppana bound [25] (see also [11])  $\lambda \geq 2\sqrt{d-1} - o(1)$  where the  $o(1)$ -term tends to 0 as the graph size tends to  $\infty$ . Graphs for which  $\lambda \leq 2\sqrt{d-1}$  are in that respect optimal expanders and are called *Ramanujan* graphs.

A proof that  $d$ -regular expanders exist for any  $d \geq 3$  was given by Pinsker [26] in the early 70's via a simple probabilistic argument. However, constructing good expanders *explicitly* is far more challenging and particularly important in applications (see [27] and the references therein), a task that was first achieved by Margulis [21]. Thereafter Ramanujan graphs were constructed explicitly in the seminal works of Lubotzky, Phillips and Sarnak [20] and Margulis [22], relying on deep number theoretic facts. Till this date Ramanujan graphs remain mysterious: Not only are there very few constructions for such graphs, but for instance it is not even known whether they exist for any  $d \geq 3$  and arbitrarily many vertices. A striking result of Friedman [13] shows that almost every  $d$ -regular graph on  $n$  vertices is *nearly* Ramanujan — it has  $\lambda = 2\sqrt{d-1} + o(1)$  (the  $o(1)$ -term tends to 0 as  $n \rightarrow \infty$ ). What proportion of these graphs satisfy  $\lambda \leq 2\sqrt{d-1}$  remains an intriguing open problem.

The useful connection between expanders and the topological notion of *covering maps* was extensively studied by many authors over the last decade. Various properties of random covers of a given graph were thoroughly examined (see e.g. [4–6,17]), motivated in part by the problem of generating good (large) expanders from a given one.

Given two simple graphs  $G$  and  $H$ , a *covering map*  $\pi : V(H) \rightarrow V(G)$  is a homomorphism that for every  $x \in V(H)$  induces a bijection between the edges incident to  $x$  and those incident to  $\pi(x)$ . In the presence of such a covering map we say that  $H$  is a *lift* (or a *cover*) of  $G$ , or alternatively that  $G$  is a *quotient* of  $H$ . The *fiber* of  $y \in V(G)$  is the set  $\pi^{-1}(y)$ , and if  $G$  is connected then all fibers are of the same cardinality, the *covering number*.

One well-known connection between covers and expansion is the fact that the universal cover of any  $d$ -regular connected graph is the infinite  $d$ -regular tree  $\mathbb{T}_d$ , whose spectral radius is  $\rho = 2\sqrt{d-1}$ , the eigenvalue threshold in Ramanujan graphs. In fact, Greenberg and Lubotzky [15] (cf. [19, Chapter 4]) extended the Alon–Boppana bound to any family of general graphs in terms of the spectral radius of its universal cover (also see [12, Theorem 4.1]).

It is easy to see that any lift of a  $d$ -regular base-graph  $G$  is itself  $d$ -regular and inherits all the original eigenvalues of  $G$ . One hopes that the lift would also inherit the expansion properties of its base-graph, and in particular that almost every cover of a (small) Ramanujan graph will also be Ramanujan (see Fig. 1).

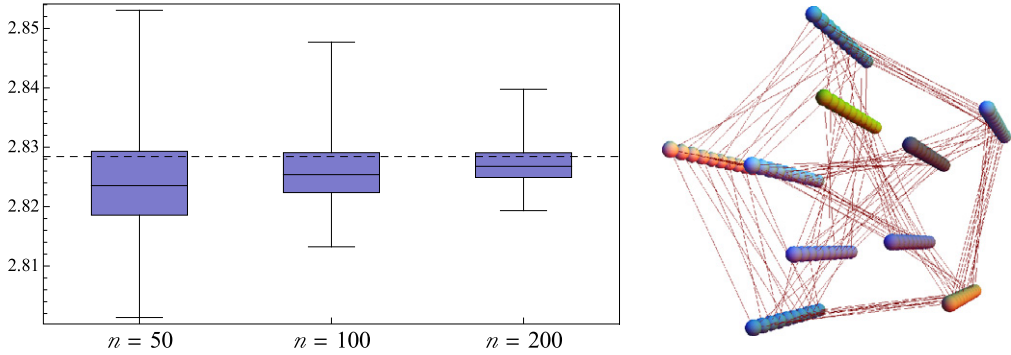


Fig. 1. Second eigenvalue (in absolute value) of a lifted Petersen graph, a 3-regular Ramanujan graph on 10 vertices, simulated for covering number  $n \in \{50, 100, 200\}$ . Dashed line marks the Ramanujan threshold  $2\sqrt{2}$ . Boxes span values from the  $\frac{1}{4}$ -quantile to the  $\frac{3}{4}$ -quantile out of 1000 lifts.

Since our focus here is on lifts of Ramanujan graphs (regular by definition) we restrict our attention to base-graphs that are  $d$ -regular for  $d \geq 3$ .

A random uniform  $n$ -lift of a base-graph  $G$  (a uniform cover of  $G$  with covering number  $n$ ) has the following convenient description: It is the graph  $H$  on the vertices  $[n] \times V(G)$ , where for each edge  $uv$  in  $G$  there is an independent uniform bijection  $\pi$ , and  $H$  has all edges of the form  $(i, u), (\pi(i), v)$ . The random lift of a complicated base-graph is thus a hybrid between the complex geometry of the quotient and the randomness due to the bijections.

In an important development in the study of the spectrum of random lifts Friedman [12] showed in 2003 that with high probability (w.h.p.) every “new” eigenvalue of an  $n$ -lift (one that is not inherited from the base-graph) is at most  $\sqrt{\rho\lambda_1} + o(1)$ , where  $\lambda_1$  is the largest eigenvalue of the base-graph and  $\rho$  is the spectral radius of its universal cover. When the base-graph  $G$  is  $d$ -regular,  $\rho = 2\sqrt{d-1}$  and Friedman’s result implies that in its random  $n$ -lift  $H$  the largest absolute value of all nontrivial eigenvalues is w.h.p.

$$\lambda(H) \leq \lambda(G) \vee O(d^{3/4}), \tag{1.1}$$

where  $(a \vee b)$  denotes  $\max\{a, b\}$ . Conversely,  $\lambda(H) \geq \lambda(G)$  and by Alon–Boppana it is also at least  $2\sqrt{d-1} - o(1)$ . This lower bound was conjectured by Friedman [12] to be tight (for general graphs he conjectured that w.h.p. all new eigenvalues are at most  $\rho + o(1)$  as in the Greenberg–Lubotzky bound).

In a recent paper [17], Linial and Puder were able to significantly improve Friedman’s bound and show that w.h.p. all the new eigenvalues of  $H$  are at most  $O(\rho^{2/3}\lambda_1^{1/3})$ . Consequently, an  $n$ -lift  $H$  of a  $d$ -regular  $G$  w.h.p. satisfies

$$\lambda(H) \leq \lambda(G) \vee O(d^{2/3}). \tag{1.2}$$

When  $G$  is a  $d$ -regular expander with nontrivial eigenvalues of  $O(\sqrt{d})$  as is the case for Ramanujan graphs, this translates to  $c\sqrt{d} \leq \lambda(H) \leq O(d^{2/3})$ .

Our main result in this work is the new near optimal upper bound of  $O((\lambda \vee \rho) \log \rho)$  when  $G$  is  $d$ -regular with all nontrivial eigenvalues at most  $\lambda$  in absolute value. For  $\lambda \leq d^{2/3-\epsilon}$  it

substantially improves the known bounds (1.1), (1.2), and when  $\lambda = O(\sqrt{d})$  as in Ramanujan graphs it is tight up to a logarithmic factor, giving  $c\sqrt{d} \leq \lambda(H) \leq O(\sqrt{d} \log d)$ .

**Theorem 1.** *Let  $G$  be a  $d$ -regular graph with all nontrivial eigenvalues at most  $\lambda$  in absolute value and let  $\rho = 2\sqrt{d-1}$  be the spectral radius of its universal cover. Let  $H$  be a random  $n$ -lift of  $G$ . For some explicit absolute constant  $C > 0$ , every nontrivial eigenvalue of  $H$  is at most  $C(\lambda \vee \rho) \log \rho$  in absolute value except with probability  $O(n^{-100})$ .*

**Corollary 2.** *Let  $G$  be a  $d$ -regular Ramanujan graph and let  $H$  be a random  $n$ -lift of  $G$ . With probability  $1 - O(n^{-100})$  every nontrivial eigenvalue of  $H$  is at most  $C\sqrt{d} \log d$  in absolute value, where  $C > 0$  is an explicit absolute constant.*

Note that the above corollary implies that typical random  $n$ -lifts of Ramanujan graphs are nearly Ramanujan. No attempt was made to optimize the explicit constant in Theorem 1. Finally, the statement of Theorem 1 holds even when the size of the base-graph  $m$  is allowed to grow with  $n$  provided that  $n$  is large enough in comparison (e.g.,  $n \geq m^{3/2}$ ).

### 1.1. Related work

The previous bounds on the spectra of random  $n$ -lifts of a fixed graph  $G$  due to Friedman [12] and Linial and Puder [17] were both obtained via Wigner's trace method. The fact that the universal cover  $\mathbb{T}$  of a connected graph  $G$  is the infinite tree of non-backtracking walks from an arbitrarily chosen vertex makes the trace method particularly useful for relating the new eigenvalues of the lift with  $\rho$ , the spectral radius of  $\mathbb{T}$ .

Even when the geometry of a graph is very well understood, bounding its nontrivial eigenvalues can be extremely challenging. For instance, a line of papers (cf. [7,14,12,11]) established various bounds for the second eigenvalue of certain random regular graphs, culminating in the optimal bound  $2\sqrt{d-1} + o(1)$  for a uniformly chosen  $d$ -regular graph on  $n$  vertices, proved by Friedman [13] using highly sophisticated arguments.

It turns out that this model is essentially the special case of an  $n$ -lift of a graph comprising a single vertex with self-loops: It is easy to see that for  $d$  even, the random  $d$ -regular graph obtained by  $d/2$  independent uniform permutations in  $S_n$  is equivalent to an  $n$ -lift of the base-graph  $G$  that has a single vertex with  $d/2$  loops (this model is in fact contiguous to the uniform random  $d$ -regular graph for  $d \geq 4$ , cf. e.g. [30]). Unfortunately, the structure of a random lift of an arbitrary  $d$ -regular graph can behave quite differently as it deeply depends on the geometry of the base-graph. Therefore, when the base-graph features a complex and rich structure (e.g. the LPS-expanders, whose expansion properties hinge on a deep theorem of Selberg) it becomes significantly harder to control the spectrum of its lifts. Indeed, there are many examples of geometric properties that have been pinpointed precisely for the random regular graph yet remain unknown for arbitrary expanders (see [18] for a recent such example). Estimating the number of closed walks in lifts of arbitrary Ramanujan graphs thus appears to be a formidable task.

In this work, the bounds obtained for the spectra of lifts of arbitrary expanders rely on an approach introduced by Kahn and Szemerédi [14], which is quite different from Wigner's trace method. This approach was originally used to control the spectrum of a random regular graph, and several new ideas are required to adapt it to the more complicated geometry of the lifts considered here.

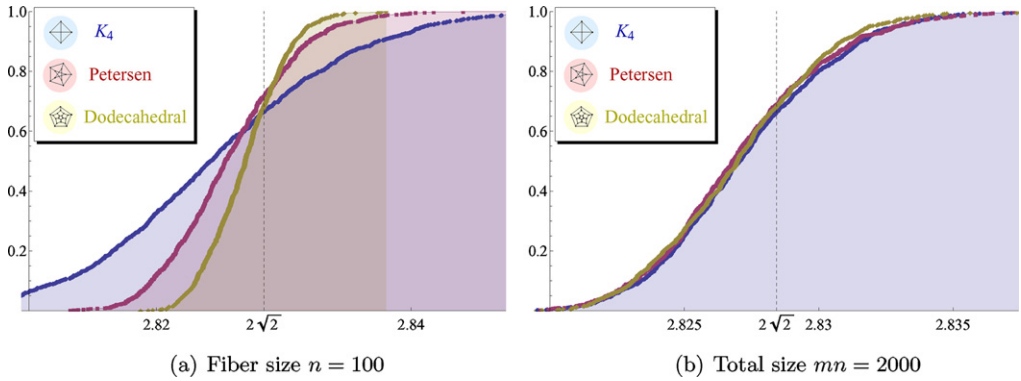


Fig. 2. Empirical estimates for the c.d.f. of the second (in absolute value) eigenvalue of  $n$ -lifts of 3-regular Ramanujan graphs on  $m$  vertices (1000 lifts were simulated per graph):  $K_4$  (complete graph on  $m = 4$  vertices), the Petersen graph ( $m = 10$ ) and the Dodecahedral graph ( $m = 20$ ). The c.d.f.'s coincide when aligning the total graph size. The probability for being strictly Ramanujan ( $\lambda \leq 2\sqrt{2}$ ) is here roughly  $2/3$ .

Another related problem in the study of spectra of lifts, yet of a rather different nature, considers the 2-lift of a base-graph (rather than  $n$ -lifts of a small fixed graph). Bilu and Linial [6] showed that for any  $d$ -regular graph  $G$  there exists a 2-lift with all new eigenvalues at most  $O(\sqrt{d \log^3 d})$ . This was shown by means of the Lovász Local Lemma, combined with the crucial observation of [6] whereby the new eigenvalues correspond precisely to the eigenvalues of a signing of the adjacency matrix of  $G$  (the matrix obtained by replacing a subset of its 1 entries by  $-1$ ). In the absence of such a characterization when the covering number  $n$  is large, different tools are needed for the problem studied here, where we seek a bound that holds for almost every  $n$ -lift with  $n$  sufficiently large.

1.2. The distribution of the second eigenvalue

As stated above, while a random  $d$ -regular graph  $G$  has second eigenvalue  $\lambda(G) \leq 2\sqrt{d-1} + o(1)$  w.h.p. (the  $o(1)$ -term tending to 0 as  $|V(G)| \rightarrow \infty$ ), the probability that  $G$  is Ramanujan is unknown. See [16,24,28] for some experimental results suggesting that this probability is bounded away from 0 and 1. As this is essentially the simplest special case of a random lift (the quotient being a single vertex with self-loops), it is natural to conjecture the following:

**Conjecture 3.** For any Ramanujan graph  $G$  there exists some  $0 < c < 1$  such that its random  $n$ -lift  $H$  satisfies  $\mathbb{P}(H \text{ is Ramanujan}) = c + o(1)$ , where the  $o(1)$ -term tends to 0 as  $n \rightarrow \infty$ .

Note that the limiting constant in the above conjecture depends on the base-graph  $G$ , as it is plausible that its structure may affect the probability of being Ramanujan. For instance, a random cover of a complete graph on  $d + 1$  vertices might behave quite differently compared to lifts of a sparse  $d$ -regular Ramanujan graph (e.g., the former deterministically contains an equipartition to  $d + 1$  color classes where each vertex has exactly one neighbor in each of the other classes). However, as we next elaborate, experimental results lead us to suspect that up to normalization this is not the case.

Fig. 2(a) shows the cumulative distribution function (c.d.f.) of  $\lambda(H)$  where  $H$  is the 100-lift of 3 different 3-regular Ramanujan base-graphs:  $K_4$  (complete graph on 4 vertices), the 10-vertex

Petersen graph and the 20-vertex Dodecahedral graph. Each curve was evaluated from 1000 random lifts. In these simulations, the probability of a random lift being Ramanujan for each of these three base-graphs was bounded between  $\frac{3}{5}$  and  $\frac{4}{5}$ .

Somewhat surprisingly, aligning the number of vertices of the graph cover  $H$  to be the same (via 100-lifts of the Dodecahedral graph, 200-lifts of the Petersen graph and 500-lifts of  $K_4$ , giving 2000-vertex covers for each graph) resulted in the curves of the individual c.d.f.’s coinciding fairly accurately. This is demonstrated in Fig. 2(b), in light of which we speculate that the following stronger version of the statement of Conjecture 3 holds.

First, it seems plausible that for any integer  $d \geq 3$  the limiting distribution of the second eigenvalue of the random cover is independent of the base-graph. Namely, there exists a distribution  $\mu_d$  on  $[0, d]$  such that for any  $d$ -regular Ramanujan graph  $G$  on  $m$  vertices, the distribution of  $\lambda(H)$  for its random  $n$ -lift  $H$  converges to  $\mu_d$  as  $n \rightarrow \infty$ . Second, the strong fit between the curves after aligning the total graph sizes suggests that even the rate of convergence to  $\mu_d$  depends on  $mn$  rather than on the geometry of the base-graph or even its relative size (in case  $m$  is allowed to depend to  $n$ ). Of-course, one clearly needs some level of “burn-in” for the covering number  $n$  compared to  $m$  since the cover  $H$  starts as Ramanujan at  $n = 1$ . For example, it may be that for any  $n \geq m$  the total-variation distance between the distribution of  $\lambda(H)$  and  $\mu_d$  decays as a function of  $mn$  alone, namely that  $\|\mathbb{P}(\lambda(H) \in \cdot) - \mu_d\|_{TV} \leq \alpha_d(mn)$  where  $\alpha_d(k)$  depends only on  $d, k$  and tends to 0 as  $k \rightarrow \infty$ .

## 2. Preliminaries and outline of the proof

### 2.1. Combinatorial vs. spectral expanders

The base-graph  $G$  from Theorem 1 corresponds to the algebraic definition of an expander known as an  $(m, d, \lambda)$ -graph. An alternative closely-related criterion is the traditional definition of an expander graph in terms of its combinatorial edge or vertex expansion. Let  $G$  be a  $d$ -regular graph on  $m$  vertices. The Cheeger constant of  $G$  (also referred to as the edge isoperimetric constant) is defined as

$$h(G) = \min_{\emptyset \neq S \subsetneq V} \frac{|\partial S|}{|S| \wedge |V \setminus S|},$$

where  $(a \wedge b)$  denotes  $\min\{a, b\}$  and  $\partial S$  is the set of edges with exactly one endpoint in  $S$ . We say that  $G$  is a  $c$ -edge-expander for some fixed  $c > 0$  if it satisfies  $h(G) > c$ . Similarly, one defines a  $c$ -vertex-expander by replacing  $\partial S$  with the vertex boundary.

For  $G$  as above the eigenvalues of the corresponding adjacency matrix are

$$d = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq -d$$

by Perron–Frobenius. We say that  $G$  is an  $(m, d, \lambda)$ -graph if  $|\lambda_i| \leq \lambda$  for all  $i \neq 1$ . This notion was introduced by Alon in the 1980’s, motivated by the fact that when  $\lambda$  is much smaller than  $d$  such graphs exhibit strong pseudo-random properties, resembling a random graph with edge density  $d/m$ . A notable example of this is captured by the *Expander Mixing Lemma*: if  $A, B$  are (not necessarily disjoint) subsets of vertices of an  $(m, d, \lambda)$ -graph then

$$\left| e(A, B) - \frac{d}{m}|A||B| \right| \leq \lambda \sqrt{|A||B|}, \tag{2.1}$$

where  $e(A, B) = \#\{(a, b) : a \in A, b \in B, ab \in E(G)\}$  [3, Chapter 9].

Relating the above two notions of expansion is the following well-known discrete analogue of Cheeger’s inequality bounding the first eigenvalue of a Riemannian manifold (Alon [1], Alon and Milman [2], Dodziuk [9], Sinclair and Jerrum [29]):

$$\frac{d - \lambda}{2} \leq h(G) \leq \sqrt{2d(d - \lambda)}.$$

See the survey [16] for further information on expanders.

### 2.2. Outline of the proof

We begin by describing the Kahn–Szemerédi [14] approach for obtaining an  $O(\sqrt{d})$  bound for all nontrivial eigenvalues of random  $d$ -regular graphs. Following Broder and Shamir [7], the actual random graph model studied by [14] is the  $2d$ -regular graph obtained from the union of  $d$  permutations, contiguous to the lift of a single vertex with  $d$  loops.

Let  $H$  be the random graph in mention and let  $A_H$  denote its adjacency matrix. By the Rayleigh quotient principle, the second (in absolute value) eigenvalue of the graph  $H$  can be written as

$$\lambda(H) = \max_{\substack{\|x\|=1 \\ \langle x, \mathbf{1} \rangle = 0}} |x^T A_H x| = \max_{\substack{\|x\|=\|y\|=1 \\ \langle x, \mathbf{1} \rangle = \langle y, \mathbf{1} \rangle = 0}} |x^T A_H y|,$$

where  $\mathbf{1} = (1, \dots, 1)$  is the trivial eigenvector of  $A_H$ . To bound  $\lambda(H)$ , the authors of [14] analyzed the maximal possible value of  $|x^T A_H y|$  separating the contribution of the pairs  $x_i, y_j$  to the bilinear form into two cases:

- (1) *Heavy pairs*: the contribution from those pairs  $x_i, y_j$  where  $|x_i y_j|$  is suitably large. Here it is shown that w.h.p. the total contribution to  $x^T A_H y$  by any pair of unit vectors  $x, y \in \mathbf{1}^\perp$  is at most  $O(\sqrt{d})$ .
- (2) *Light pairs*: the remaining pairs  $x_i, y_j$ . Here it was shown that two fixed vectors  $x, y$  are unlikely to contribute more than  $O(\sqrt{d})$  to the bilinear form, and an  $\varepsilon$ -net argument was used to extend this result to any unit vectors  $x, y \in \mathbf{1}^\perp$ .

Adapting this method to lifts of general graphs requires several additional ingredients. Even in the simpler setting of [14] (see also [8]), some of the arguments are highly nontrivial and require a rather delicate analysis. More crucially, in our case we have little knowledge of the base-graph  $G$ , hence the study of both the “heavy” and “light” parts becomes significantly more involved.

First, our only input on  $G$  is the magnitude of its second eigenvalue, which turns the analysis of the heavy part into a delicate optimization problem, requiring two levels of dyadic expansions of the potential contributions to the final bilinear form. One level of dyadic expansion reduces this optimization problem into a geometric statement on cuts between subsets of vertices in typical lifts (Proposition 3.2), whereas a second level of dyadic expansion was needed in the proof of that

result. It is this dyadic expansion that entails the extra logarithmic factor in our main theorem. It is plausible that one could obtain a tight bound on the nontrivial eigenvalues of typical random lifts by improving this specific argument.

Second, the consideration of the light part relies on a nontrivial martingale argument which may be useful in other applications: In the absence of sufficient control over the expectation (due to the unknown contribution of the heavy part) we resort to an  $L^2$  analysis of the increments in the corresponding Doob’s martingale and apply a Bernstein–Kolmogorov type large deviation inequality due to Freedman.

### 2.3. Notation

Throughout the paper we use  $G$  to denote the base-graph, a  $d$ -regular graph on  $m$  vertices, and let  $H$  denote its random  $n$ -lift. The asymptotic notation is used under the assumption that  $n \rightarrow \infty$ .

For the sake of clarity, when addressing a vertex in  $V(H) = [n] \times V(G)$  we will typically denote it either by  $ij$  or by  $i'j'$  using indices  $i, i' \in [n]$  and  $j, j' \in [m]$ . Whenever  $u, v$  are vertices in some graph whose identity is clear from the context, the abbreviation  $u \sim v$  will denote that these two vertices are adjacent. For example,  $ij \sim i'j'$  will usually stand for  $(ij, i'j') \in E(H)$ , which in turn implies that  $j \sim j'$  in  $G$  by the definition of the lift.

Unless stated otherwise, all logarithms are using base 2 and  $\|\cdot\|$  denotes the  $L^2$  norm in the appropriate Euclidean space.

### 2.4. Organization

The rest of this paper is organized as follows. Section 3 deals with the contribution of the heavy pairs to the bilinear form  $x^T A_H y$ . Section 4 deals with the contribution of the light pairs. In the final section, Section 5, we combine these results to conclude the proof of Theorem 1.

## 3. Heavy pairs and large cuts

Let  $G$  be an  $(m, d, \lambda)$ -graph (that is, a  $d$ -regular graph where all nontrivial eigenvalues are at most  $\lambda$  in absolute value) with adjacency matrix  $A_G$ , and  $H$  be a random  $n$ -lift of  $G$  with adjacency matrix  $A_H$ . As mentioned before, the largest nontrivial eigenvalue of  $H$  in absolute value is precisely

$$\max\{|x^T A_H x|: x \in \mathbb{R}^{mn}, \langle x, \mathbf{1} \rangle = 0, \|x\| = 1\},$$

where  $\mathbf{1}$  denotes the trivial eigenvector.

### 3.1. Heavy pairs

We first analyze the typical contribution to  $x^T A_H y$  from pairs  $x_{ij}, y_{i'j'}$  with fairly large products. More precisely, we say that a pair  $x_{ij}, y_{i'j'}$  is *heavy* if  $|x_{ij}y_{i'j'}| \geq \lambda/mn$ , and otherwise it is *light*. For  $x, y \in \mathbb{R}^{mn}$ , define  $R_h(x, y)$  to be the random variable

$$R_h(x, y) = \sum_{ij \sim i'j'} x_{ij}y_{i'j'} \mathbb{1}_{\{|x_{ij}y_{i'j'}| \geq \lambda/mn\}}.$$



The next theorem estimates the contribution of the heavy pairs along the edges of  $H$ . The exponent of  $m$  in the requirement  $n \geq m^{3/2}$  was selected to simplify the exposition and can be replaced by  $n \geq m^{1+\delta}$  for any  $\delta > 0$ .

**Theorem 3.1.** *Let  $G$  be an  $(m, d, \lambda)$ -graph with  $\lambda \geq \sqrt{d}$  and let  $H$  be a random  $n$ -lift of  $G$  for  $n \geq m^{3/2}$ . Then with probability at least  $1 - O(n^{-100})$  every  $x, y \in \mathbb{R}^{mn}$  with  $\|x\| \leq 1$  and  $\|y\| \leq 1$  satisfy  $|R_h(x, y)| \leq 3500\lambda \log d$  and moreover  $|R_h(x, y) - \mathbb{E}[R_h(x, y)]| \leq 7000\lambda \log d$ .*

The main ingredient in the proof of Theorem 3.1 is the following lemma, which provides an upper bound on the number of edges in a cut between subsets of vertices in  $H$ .

**Proposition 3.2.** *Let  $H$  be a random  $n$ -lift of an  $(m, d, \lambda)$ -graph  $G$  with  $\lambda \geq \sqrt{d}$  and  $n \geq m^{3/2}$ . Then except with probability  $O(n^{-100})$ , every two subsets of vertices  $A, B \subset V(H)$  with  $|A||B| \leq (2mn/\lambda)^2$  satisfy*

$$e(A, B) \leq 802\lambda\sqrt{|A||B|} + 75(|A| + |B|)\log^2 d. \tag{3.1}$$

We will next show how to derive Theorem 3.1 from this lemma, whose proof is postponed to Subsection 3.2.

**Proof of Theorem 3.1.** Consider the following dyadic expansion of  $x, y$ :

$$\begin{aligned} \mathcal{D}_\ell &= \left\{ (i, j) : 2^\ell \leq |x_{ij}| \sqrt{\frac{mn}{\lambda}} < 2^{\ell+1} \right\} \quad (\ell \in \mathbb{Z}), \\ \mathcal{D}'_{\ell'} &= \left\{ (i', j') : 2^{\ell'} \leq |y_{i'j'}| \sqrt{\frac{mn}{\lambda}} < 2^{\ell'+1} \right\} \quad (\ell' \in \mathbb{Z}), \end{aligned}$$

and identify any element  $(i, j)$  in  $\mathcal{D}_\ell$  or  $\mathcal{D}'_{\ell'}$  with the vertex  $ij$  in  $H$ . These definitions, together with the assumption on  $\|x\|$  and  $\|y\|$ , imply that

$$\sum_\ell 4^\ell |\mathcal{D}_\ell| \frac{\lambda}{mn} \leq \sum_{i,j} x_{ij}^2 \leq 1, \quad \sum_{\ell'} 4^{\ell'} |\mathcal{D}'_{\ell'}| \frac{\lambda}{mn} \leq \sum_{i',j'} y_{i'j'}^2 \leq 1. \tag{3.2}$$

Furthermore, if  $ij \in \mathcal{D}_\ell$  and  $i'j' \in \mathcal{D}'_{\ell'}$  then a necessary condition for  $|x_{ij}y_{i'j'}| \geq \lambda/mn$  is that  $\ell + \ell' + 2 > 0$ , and so

$$\sum_{ij \sim i'j'} |x_{ij}y_{i'j'}| \mathbb{1}_{\{|x_{ij}y_{i'j'}| \geq \lambda/mn\}} \leq 4 \sum_{\ell+\ell'>-2} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn}, \tag{3.3}$$

where  $e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'})$  is the number of edges between the two subsets of vertices in  $H$  corresponding to  $\mathcal{D}_\ell$  and  $\mathcal{D}'_{\ell'}$ .

To prove that  $|R_h(x, y)| = O(\lambda \log d)$ , we set

$$D = \log \left( \frac{d}{\lambda \log d} \right) - 3,$$

and analyze the sum in (3.3) according to whether or not  $|\ell - \ell'| > D$ .

First, consider  $D \leq 0$ . As  $H$  is  $d$ -regular, trivially  $e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \leq d|\mathcal{D}_\ell|$ , yielding

$$\sum_{\substack{\ell+\ell'>-2 \\ \ell \geq \ell'}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq \sum_{\ell} 4^\ell \cdot d|\mathcal{D}_\ell| \cdot \frac{\lambda}{mn} \sum_{\ell \geq \ell'} 2^{-(\ell-\ell')} \leq 2d,$$

where the second inequality used (3.2). Similarly,

$$\sum_{\substack{\ell+\ell'>-2 \\ \ell' \geq \ell}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 2d,$$

and since  $D \leq 0$  occurs if and only if  $d \leq 8\lambda \log d$ , altogether in this case

$$\sum_{\ell+\ell'>-2} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 4d \leq 32\lambda \log d. \tag{3.4}$$

We now focus on  $D > 0$ . Consider the case where  $\ell - \ell' > D > 0$ . Repeating the above argument, we now get

$$\begin{aligned} \sum_{\substack{\ell+\ell'>-2 \\ \ell-\ell'>D}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} &\leq \sum_{\ell} 4^\ell \cdot d|\mathcal{D}_\ell| \cdot \frac{\lambda}{mn} \sum_{\ell-\ell'>D} 2^{-(\ell-\ell')} \\ &\leq d2^{-D} = 8\lambda \log d. \end{aligned}$$

By symmetry, the same argument holds for the case  $\ell' - \ell > D$ , and we infer that

$$\sum_{\substack{\ell+\ell'>-2 \\ |\ell'-\ell|>D}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 16\lambda \log d. \tag{3.5}$$

It remains to treat  $|\ell - \ell'| \leq D$  with  $D > 0$ . This will be achieved with the help of Proposition 3.2, which estimates the size of the cut between two subsets  $A, B$  in case  $|A||B| \leq (2mn/\lambda)^2$ . Indeed, for  $\ell + \ell' \geq -1$  we have

$$\begin{aligned} \frac{1}{4} |\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}| \frac{\lambda^2}{(mn)^2} &\leq 4^{\ell+\ell'} |\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}| \frac{\lambda^2}{(mn)^2} \\ &\leq \sum_{ij \in \mathcal{D}_\ell} \sum_{i'j' \in \mathcal{D}'_{\ell'}} x_{ij}^2 y_{i'j'}^2 \\ &\leq \sum_{ij} x_{ij}^2 \sum_{i'j'} y_{i'j'}^2 \leq 1, \end{aligned}$$

and therefore  $|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}| \leq (2mn/\lambda)^2$ . Thus, Proposition 3.2 gives that with probability  $1 - O(n^{-100})$ ,

$$\sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 802 \sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} \frac{\lambda}{mn} \cdot \lambda \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|} + 75 \sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} \frac{\lambda}{mn} \cdot (|\mathcal{D}_\ell| + |\mathcal{D}'_{\ell'}|) \log^2 d. \tag{3.6}$$

For the first expression in the right-hand side of (3.6), note that there are at most  $2D + 1 \leq 2 \log(d/\lambda) \leq \log d$  integers  $k$  such that  $|k| \leq D$  (here we used the fact that  $\lambda \geq \sqrt{d}$ ). For each such value, we can combine (3.2) with Cauchy–Schwarz to get that

$$\sum_{\substack{\ell+\ell'>-2 \\ \ell-\ell'=k}} 2^{\ell+\ell'} \frac{\lambda}{mn} \cdot \lambda \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|} \leq \lambda \sqrt{\sum_{\ell} 4^{\ell} |\mathcal{D}_\ell| \frac{\lambda}{mn}} \sqrt{\sum_{\ell'} 4^{\ell'} |\mathcal{D}'_{\ell'}| \frac{\lambda}{mn}} \leq \lambda,$$

and summing over  $k$  it follows that

$$\sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} \frac{\lambda}{mn} \cdot \lambda \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|} \leq \lambda \log d.$$

For the second expression in (3.6), again recall that  $\lambda \geq \sqrt{d}$ , and so

$$\begin{aligned} \sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} \frac{\lambda}{mn} |\mathcal{D}_\ell| \log^2 d &\leq \sum_{\ell} 4^{\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell| \log^2 d \sum_{|k|\leq D} 2^{-k} < 2^{D+1} \log^2 d \\ &= \frac{d}{4\lambda \log d} \log^2 d \leq \frac{1}{4} \lambda \log d, \end{aligned}$$

and the same applies to the analogous quantity for  $|\mathcal{D}'_{\ell'}|$ .

Altogether, these two estimates for (3.6) sum up to

$$\sum_{\substack{\ell+\ell'>-2 \\ |\ell-\ell'|\leq D}} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq \left(802 + 2 \cdot \frac{75}{4}\right) \lambda \log d < 840 \lambda \log d,$$

and combining this with (3.4) and (3.5) gives that

$$\sum_{\ell+\ell'>-2} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 856 \lambda \log d.$$

Recalling (3.3), we deduce that

$$|R_h(x, y)| \leq 4 \sum_{\ell+\ell'>-2} 2^{\ell+\ell'} e(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \frac{\lambda}{mn} \leq 3424 \lambda \log d.$$

To obtain the statement on  $|R_h(x, y) - \mathbb{E}[R_h(x, y)]|$  first note that since  $H$  is  $d$ -regular  $|R_h(x, y)| \leq \lambda_1(H) = d$  with probability 1 for any two unit vectors  $x, y$ . We have already established that, except with probability  $O(n^{-100})$ , every pair of vectors  $x, y$  with norm at most 1 satisfies  $|R_h(x, y)| \leq 3424\lambda \log d$ . Hence,

$$\mathbb{E}|R_h(x, y)| \leq 3424\lambda \log d + dO(n^{-100}) < 3425\lambda \log d,$$

where the last inequality holds for any sufficiently large  $n$ . Reapplying the result on  $R_h(x, y)$  (along with the triangle inequality) now completes the proof of the theorem (with room to spare).  $\square$

### 3.2. Proof of Proposition 3.2

Write  $|A| = \alpha n$  and  $|B| = \beta n$  where  $0 < \alpha, \beta \leq m$ . Our assumption on  $|A||B|$  then translates into

$$\alpha\beta \leq (2m/\lambda)^2. \tag{3.7}$$

We aim to show that, except with probability  $1 - O(n^{-100})$ , for any such  $A, B$  we have  $e(A, B) = O(\lambda\sqrt{|A||B|} + (|A| + |B|)\log^2 d)$ , or in terms of  $\alpha, \beta$ , that

$$e(A, B)/n = O(\lambda\sqrt{\alpha\beta} + (\alpha + \beta)\log^2 d).$$

Define the following partition of the fibers according to a dyadic expansion of their proportion that is included in  $A$ ,

$$S_i = \left\{ v \in V(G) : 2^{-i-1} < \frac{|A \cap ([n] \times \{v\})|}{n} \leq 2^{-i} \right\} \quad (i = 0, 1, \dots, \log n),$$

$$A_i = A \cap \bigcup_{v \in S_i} ([n] \times \{v\}), \quad s_i = |S_i|, \quad \alpha_i = s_i 2^{-i}.$$

Notice that by these definitions,  $s_i$  is the number of fibers with about  $2^{-i}n$  vertices from  $A$ , and so  $|A_i| \approx s_i 2^{-i}n = \alpha_i n$ . In other words, there are about  $\alpha_i n$  vertices of  $A$  in fibers of type  $S_i$ , and more precisely,

$$\frac{1}{2}\alpha_i n < |A_i| \leq \alpha_i n, \quad \frac{1}{2} \sum_i \alpha_i < \alpha \leq \sum_i \alpha_i. \tag{3.8}$$

Similarly, we perform an analogous dyadic expansion for  $B$ :

$$T_j = \left\{ v \in V(G) : 2^{-j-1} < \frac{|B \cap ([n] \times \{v\})|}{n} \leq 2^{-j} \right\} \quad (j = 0, 1, \dots, \log n),$$

$$B_j = B \cap \bigcup_{v \in T_j} ([n] \times \{v\}), \quad t_j = |T_j|, \quad \beta_j = t_j 2^{-j},$$

and again have that

$$\frac{1}{2}\beta_j n < |B_j| \leq \beta_j n, \quad \frac{1}{2} \sum_j \beta_j < \beta \leq \sum_j \beta_j. \tag{3.9}$$

Clearly,

$$e(A, B) = \sum_{i,j} e(A_i, B_j),$$

and our bound on  $e(A, B)$  will follow from an analysis of the number of edges between the various types of  $A_i$ 's and  $B_j$ 's.

First, consider the case  $i = j = 0$ . Here we have  $\frac{1}{2}s_0 n < |A_0| \leq |A|$  and  $\frac{1}{2}t_0 n < |B_0| \leq |B|$ . Since there are  $n$  edges in  $H$  between any pair of fibers that correspond to adjacent vertices in  $G$ , the Expander Mixing Lemma (see (2.1)) applied to the base-graph  $G$  gives that

$$\begin{aligned} e(A_0, B_0) &\leq ne(S_0, T_0) \leq \frac{dn}{m}s_0 t_0 + n\lambda\sqrt{s_0 t_0} < \frac{4d}{mn}|A||B| + 2\lambda\sqrt{|A||B|} \\ &= \frac{4d}{m}\alpha\beta n + 2\lambda\sqrt{\alpha\beta}n. \end{aligned}$$

Recalling that  $\sqrt{\alpha\beta} \leq 2m/\lambda$  (see (3.7)) it follows that

$$e(A_0, B_0) \leq 2\sqrt{\alpha\beta}n \left( \frac{2d}{m} \cdot \frac{2m}{\lambda} + \lambda \right) = 2\sqrt{\alpha\beta}n \left( \frac{4d}{\lambda} + \lambda \right) \leq 10\lambda\sqrt{\alpha\beta}n, \tag{3.10}$$

where the last inequality used the fact that  $\lambda \geq \sqrt{d}$  and so  $\lambda \geq d/\lambda$ .

Next, consider  $e(A_i, B_j)$  in case  $|i - j| > 2 \log d$ . There is a total of  $s_i$  fibers in  $A_i$ , thus by definition of the  $n$ -lift of a  $d$ -regular graph there are at most  $ds_i$  fibers, where  $B_j$  may have vertices that contribute to  $e(A_i, B_j)$ . Since  $B_j$  has at most  $2^{-j}n$  vertices in each fiber, and each vertex has  $d$  neighbors in  $H$ , we deduce that

$$\begin{aligned} \sum_{j-i > 2 \log d} e(A_i, B_j) &\leq \sum_i \sum_{j > i + 2 \log d} ds_i \cdot 2^{-j} n \cdot d \\ &= \sum_i s_i 2^{-i} n \cdot d^2 \sum_{j-i > 2 \log d} 2^{-(j-i)} \\ &\leq \sum_i s_i 2^{-i} n = \sum_i \alpha_i n \leq 2\alpha n, \end{aligned}$$

where the last inequality followed from (3.8). Similarly, we have

$$\sum_{i-j > 2 \log d} e(A_i, B_j) \leq \sum_j \beta_j n \leq 2\beta n,$$

and conclude that

$$\sum_{|i-j|>2\log d} e(A_i, B_j) \leq 2(\alpha + \beta)n. \tag{3.11}$$

It remains to treat the case  $|i - j| \leq 2 \log d$  for all  $(i, j) \neq (0, 0)$ , where the required bound will only hold w.h.p.

Consider a prescribed set of  $k$  pairs of vertices  $(i_l j_l, i'_l j'_l)$  ( $l \in [k]$ ) in  $H$ . We wish to bound the probability that  $\{i_l j_l \sim i'_l j'_l$  for all  $l\}$  by  $(3/n)^k$ . By the independence of the different pairs of fibers in the lift, it clearly suffices to show this when all the  $i_l j_l$ 's are on one fiber and all the  $i'_l j'_l$ 's are on another, i.e., for some  $j \neq j'$  and all  $l$  we have  $j_l = j$  and  $j'_l = j'$ . When  $k \leq \frac{2}{3}n$  then it is straightforward that this probability is indeed at most  $(3/n)^k$ . To see this, expose the pairings of  $i_1 j, i_2 j, \dots, i_k j$  one by one, and note that for  $l \leq k \leq \frac{2}{3}n$ , the probability to match  $i_l j$  to  $i'_l j'$ , given that so far we succeeded in matching all the  $l - 1$  previous pairs, is  $1/(n - l + 1) \leq 3/n$ .

Further note that, when considering potential edges between  $A_i$  and  $B_j$ , the case  $k \geq \frac{2}{3}n$  can only arise when  $i = j = 0$ , otherwise no two fibers have more than  $n/2$  points of  $A_i$  and  $B_j$  respectively. Since we excluded the case  $i = j = 0$ , the above estimate holds for any of our sets  $A_i, B_j$ .

Write

$$W_{ij} = e(A_i, B_j)/n,$$

and recall that  $A_i$  and  $B_j$  are contained in the fibers corresponding to  $S_i$  and  $T_j$  respectively, and have at most  $2^{-i}n$  and  $2^{-j}n$  vertices on each of these respective fibers. Suppose first that the identity of the fibers  $S_i$  and  $T_j$  are given (we will account for these later). In this case, the number of configurations of the vertices of  $A_i$  on the fibers  $S_i$  can be bounded from above by  $2^{s_0 n}$  if  $i = 0$  and by  $[\sum_{r \leq 2^{-i}n} \binom{n}{r}]^{s_i} \leq 2^{(i+2)s_i 2^{-i}n}$  if  $i \neq 0$ , here using the well-known fact that  $\sum_{i \leq \gamma m} \binom{m}{i} \leq 2^{H(\gamma)m}$  where  $H(x) = -x \log_2 x - (1 - x) \log_2 (1 - x)$  is the binary entropy function. Similarly, the number of configuration of  $B_j$  on the fibers  $T_j$  is at most  $2^{(j+2)t_j 2^{-j}n}$ . For each such configuration of the vertices of  $A_i, B_j$  there are at most  $2^{-i-j}n^2 e_G(S_i, T_j)$  pairs which may potentially be connected in  $H$ . From the above estimate on the probability of  $k$  pairs being adjacent in  $H$ , it now follows that for any  $w_{ij} > 0$  and choice of  $S_i$  and  $T_j$ ,

$$\begin{aligned} \mathbb{P}(W_{ij} = w_{ij}) &\leq 2^{(i+2)s_i 2^{-i}n} 2^{(j+2)t_j 2^{-j}n} \binom{2^{-i-j}n^2 e(S_i, T_j)}{w_{ij}n} \left(\frac{3}{n}\right)^{w_{ij}n} \\ &\leq 2^{((i+2)\alpha_i + (j+2)\beta_j)n} (9 \cdot 2^{-i-j} e(S_i, T_j)/w_{ij})^{w_{ij}n}. \end{aligned}$$

(Here we used the well-known fact that  $\binom{a}{b} \leq (ea/b)^b$ .) Defining

$$z_{ij} = \frac{2^{i+j} w_{ij}}{9e(S_i, T_j)},$$

we then get

$$\begin{aligned} \mathbb{P}(W_{ij} = w_{ij}) &\leq (2^{(i+2)\alpha_i + (j+2)\beta_j} 2^{-w_{ij} \log z_{ij}})^n \\ &= 2^{[(i+2)\alpha_i + (j+2)\beta_j - 9e(S_i, T_j) 2^{-i-j} z_{ij} \log z_{ij}]n}. \end{aligned} \tag{3.12}$$

We will next establish a threshold for  $z_{ij}$  such that the above probability would be at most  $\exp(-n^{3/4-o(1)})$  and then translate this bound to the cut  $e(A_i, B_j)$  via the corresponding  $w_{ij}$ 's.

Consider the equation  $x \log x = b$  and note that for  $b > 0$  it has a unique solution  $x > 1$  monotone increasing in  $b$ . Let  $z_{ij}^*$  be the solution to

$$z_{ij}^* \log z_{ij}^* = \frac{2^{i+j}}{9e(S_i, T_j)} [(i+2)\alpha_i + (j+2)\beta_j + n^{-1/4}], \tag{3.13}$$

and define its counterpart (similar to the relation between  $w_{ij}$  and  $z_{ij}$ )

$$w_{ij}^* = \frac{9e(S_i, T_j)}{2^{i+j}} (z_{ij}^* \vee 2). \tag{3.14}$$

Combining these definitions with the probability bound (3.12), while noting that this bound in that equation is monotone decreasing in  $z_{ij}$  (and hence in  $w_{ij}$ ) in the range  $z_{ij} \geq 1$ , we deduce that for any  $k \geq w_{ij}^*$ ,

$$\mathbb{P}(W_{ij} = k) \leq 2^{-n^{3/4}}.$$

Since  $W_{ij} = e(A_i, B_j)/n$  with  $e(A_i, B_j) \leq e(H) = dmn/2$  we can sum  $k$  over all possible values that  $W_{ij}$  can accept and infer that

$$\mathbb{P}(W_{ij} \geq w_{ij}^*) \leq dmn 2^{-n^{3/4}}.$$

Next recall that the above estimate was for  $A_i, B_j$  with a given choice of the fibers  $S_i, T_j$ . Summing the above probability over all possible choices for such fibers (using a trivial bound of  $2^m$  options for each of the sets) and then further summing over at most  $\log^2 n$  pairs of  $i, j$  we deduce that

$$\mathbb{P}\left(\bigcup_{i,j} \{e(A_i, B_j) \geq w_{ij}^* n\}\right) \leq (dmn \log^2 n) 2^{2m-n^{3/4}} < n^{-100},$$

with the last inequality valid for any sufficiently large  $n$  since  $m \leq n^{2/3}$ . Collecting (3.10) and (3.11) this yields that, except with probability  $n^{-100}$ , any two sets  $A, B$  with  $\alpha\beta \leq (2m/\lambda)^2$  (as per (3.7)) satisfy

$$e(A, B)/n \leq 10\lambda\sqrt{\alpha\beta} + 2(\alpha + \beta) + \sum_{\substack{i+j>0 \\ |i-j|\leq 2\log d}} w_{ij}^*. \tag{3.15}$$

It thus suffices to bound  $\sum_{i,j} w_{ij}^*$  in order to complete the proof.

**Lemma 3.3.** Let  $z_{ij}^*$  be the solution to (3.13) and let  $w_{ij}^*$  be its counterpart as given in (3.14). Then for any  $d \geq 320$  and large enough  $n$  we have

$$\sum_{\substack{i+j>0 \\ |i-j| \leq 2 \log d}} w_{ij}^* \leq 792\lambda\sqrt{\alpha\beta} + 74(\alpha + \beta) \log^2 d + n^{-1/3}.$$

**Proof.** We first consider pairs  $i, j$  such that  $z_{ij}^* \geq 2$ .

By the definition of  $z_{ij}^*$  as the solution of (3.13), the right-hand side of that equation, which we denote by  $b_{ij}$ , necessarily then satisfies

$$b_{ij} = \frac{2^{i+j}}{9e(S_i, T_j)} [(i + 2)\alpha_i + (j + 2)\beta_j + n^{-1/4}] \geq 2. \tag{3.16}$$

Furthermore, since for any  $b > 1$  the solution of  $z \log z = b$  satisfies  $z < 2\frac{b}{\log b}$  (this is easy to verify using the monotonicity of  $z \log z$ ) we can infer an upper bound on  $z_{ij}^*$  in the form of

$$z_{ij}^* \leq 2 \frac{2^{i+j}}{9e(S_i, T_j) \log b_{ij}} [(i + 2)\alpha_i + (j + 2)\beta_j + n^{-1/4}],$$

and as a consequence

$$w_{ij}^* \mathbb{1}_{\{z_{ij}^* \geq 2\}} = \frac{9e(S_i, T_j)}{2^{i+j}} z_{ij}^* \leq \frac{2}{\log b_{ij}} [(i + 2)\alpha_i + (j + 2)\beta_j + n^{-1/4}].$$

Immediately by (3.16) the last denominator is at least 1 and so

$$\sum_{|i-j| \leq 2 \log d} \frac{n^{-1/4}}{\log b_{ij}} \leq \sum_{i,j} n^{-1/4} \leq n^{-1/4} \log^2 n = o(n^{-1/3}),$$

where  $0 \leq i, j \leq \log n$  by definition. Next, note that for the same reason

$$\begin{aligned} \sum_{|i-j| \leq 2 \log d} \frac{\alpha_i + \beta_j}{\log b_{ij}} &\leq (1 + 4 \log d) \left( \sum_i \alpha_i + \sum_j \beta_j \right) \\ &\leq (2 + 8 \log d)(\alpha + \beta), \end{aligned}$$

using the fact that  $\sum_i \alpha_i \leq 2\alpha$  and  $\sum_j \beta_j \leq 2\beta$  as given in (3.8), (3.9). The combination of the last three equations implies that for large enough  $n$ ,

$$\begin{aligned} \sum_{\substack{i+j>0 \\ |i-j| \leq 2 \log d}} w_{ij}^* \mathbb{1}_{\{z_{ij}^* \geq 2\}} &\leq 2 \sum_{|i-j| \leq 2 \log d} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} \\ &\quad + (32 \log d + 8)(\alpha + \beta) + n^{-1/3}. \end{aligned} \tag{3.17}$$



Before we further analyze the expressions  $(i\alpha_i + j\beta_j)/\log b_{ij}$  we wish to narrow down the range of pairs  $(i, j)$ . First, we can quickly move to  $i, j \geq 20$ . Indeed, if for instance  $i < 20$  then using (3.16)

$$\begin{aligned} \sum_{\substack{i < 20 \\ |i-j| \leq 2\log d}} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} &\leq \sum_{\substack{i \leq 19 \\ j \leq 19+2\log d}} i\alpha_i + j\beta_j \\ &\leq (40\log d + 380) \left( \sum_i \alpha_i + \sum_j \beta_j \right) \\ &\leq (80\log d + 760)(\alpha + \beta). \end{aligned}$$

An analogous calculation holds for  $j < 20$ , yielding that

$$\sum_{\substack{i < 20 \\ |i-j| \leq 2\log d}} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} + \sum_{\substack{j < 20 \\ |i-j| \leq 2\log d}} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} \leq (160\log d + 1520)(\alpha + \beta).$$

The case  $20 \leq i, j \leq 4\log d$  is treated similarly:

$$\begin{aligned} \sum_{20 \leq i, j \leq 4\log d} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} &\leq 4\log d(4\log d - 19) \left( \sum_i \alpha_i + \sum_j \beta_j \right) \\ &\leq (32\log^2 d - 152\log d)(\alpha + \beta). \end{aligned}$$

Plugging the last two equations in (3.17) and defining

$$\Gamma = \left\{ (i, j) : \begin{array}{l} i, j \geq 20 \\ |i - j| \leq 2\log d \\ i \geq 4\log d \text{ or } j \geq 4\log d \end{array} \right\} \tag{3.18}$$

it follows that

$$\begin{aligned} \sum_{\substack{i+j>0 \\ |i-j|\leq 2\log d}} w_{ij}^* \mathbb{1}_{\{z_{ij}^* \geq 2\}} &\leq 2 \sum_{(i,j) \in \Gamma} \frac{i\alpha_i + j\beta_j}{\log b_{ij}} \\ &\quad + (32\log^2 d + 40\log d + 1528)(\alpha + \beta) + n^{-1/3}. \end{aligned} \tag{3.19}$$

Recalling the definition (3.16) of  $b_{ij}$ , it now remains to bound  $\sum_{(i,j) \in \Gamma} \xi_{ij}$  with  $\xi_{ij}$  given by

$$\xi_{ij} = \frac{i\alpha_i + j\beta_j}{\log \left[ \frac{2^{i+j} [(i+2)\alpha_i + (j+2)\beta_j + n^{-1/4}]}{9e(S_i, T_j)} \right]} \leq \frac{i\alpha_i + j\beta_j}{\log \left[ \frac{2^{i+j} (i\alpha_i + j\beta_j)}{9e(S_i, T_j)} \right]}.$$

Note that the last inequality (where we reduced the argument of the  $\log(\cdot)$ ) is only legitimate provided that

$$2^{i+j} \frac{i\alpha_i + j\beta_j}{9e(S_i, T_j)} > 1. \tag{3.20}$$

In what follows we will show that this is indeed the case and then proceed to bound  $\sum \xi_{ij}$ . This will be achieved by splitting the analysis into two cases, according to the structure of  $e(S_i, T_j)$  in the base-graph  $G$ . Recall that we have  $e(S_i, T_j) \leq (d/m)s_i t_j + \lambda\sqrt{s_i t_j}$  as  $G$  is an  $(m, d, \lambda)$ -graph.

- **Case (i):**  $e(S_i, T_j) \leq 2(d/m)s_i t_j$ .

Since  $\alpha_i = s_i 2^{-i}$  and  $\beta_j = t_j 2^{-j}$ , in this case we have

$$\frac{2^{i+j}}{e(S_i, T_j)} \geq \frac{2^{i+j}}{2(d/m)s_i t_j} = \frac{m}{2d\alpha_i \beta_j}.$$

With the regime of  $(i, j)$  as in (3.18) in mind, suppose first that  $i \geq 4 \log d$ . It follows that

$$\sum_{\substack{i \geq 4 \log d \\ j \geq 20 \\ |i-j| \leq 2 \log d}} \xi_{ij} \leq \sum_{\substack{i \geq 4 \log d \\ j \geq 20 \\ |i-j| \leq 2 \log d}} \frac{i\alpha_i + j\beta_j}{\log \left[ \frac{m(i\alpha_i + j\beta_j)^2}{18d\alpha_i \beta_j} \right]} \leq \sum_{\substack{i \geq 4 \log d \\ j \geq 20 \\ |i-j| \leq 2 \log d}} \frac{i\alpha_i + j\beta_j}{\log \frac{mj}{18d\alpha_i}}.$$

As  $\alpha_i 2^i = s_i \leq m$  we have  $\alpha_i \leq m 2^{-i}$  and plugging in the fact that  $j \geq 20$ ,

$$\begin{aligned} \sum_{\substack{i \geq 4 \log d \\ j \geq 20 \\ |i-j| \leq 2 \log d}} \frac{i\alpha_i + j\beta_j}{\log \frac{mj}{18d\alpha_i}} &\leq \sum_{\substack{i \geq 4 \log d \\ |i-j| \leq 2 \log d}} \frac{i\alpha_i + j\beta_j}{\log(2^i/d)} \\ &\leq \sum_{\substack{i \geq 4 \log d \\ |i-j| \leq 2 \log d}} \frac{i\alpha_i + j\beta_j}{\frac{3}{4}i} \\ &\leq \frac{4}{3} \sum_{|i-j| \leq 2 \log d} \alpha_i + 2 \sum_{|i-j| \leq 2 \log d} \beta_j \\ &\leq (1 + 4 \log d)(3\alpha + 4\beta), \end{aligned}$$

where we used the fact that  $j \leq i + 2 \log d \leq \frac{3}{2}i$  for the above  $i, j$ . Note that we have just verified Eq. (3.20) by showing that its left-hand side is at least  $2^i/d \geq d^3$ . Similarly, if  $j \geq 4 \log d$  then

$$\sum_{\substack{j \geq 4 \log d \\ i \geq 20 \\ |i-j| \leq 2 \log d}} \xi_{ij} \leq (1 + 4 \log d)(4\alpha + 3\beta).$$

Altogether, in this case we have

$$\sum_{(i,j) \in \Gamma} \xi_{ij} \leq (28 \log d + 7)(\alpha + \beta). \tag{3.21}$$

- **Case (ii):**  $e(S_i, T_j) \leq 2\lambda\sqrt{s_i t_j}$ .

Rewriting the assumption on  $e(S_i, T_j)$  in terms of  $\alpha_i$  and  $\beta_j$ , we have

$$\frac{2^{i+j}}{e(S_i, T_j)} \geq \frac{2^{i+j}}{2\lambda\sqrt{s_i t_j}} = \frac{2^{\frac{i+j}{2}}}{2\lambda\sqrt{\alpha_i \beta_j}},$$

which gives that

$$\xi_{ij} \leq \frac{i\alpha_i + j\beta_j}{\log\left(\frac{i\alpha_i + j\beta_j}{18\lambda\sqrt{\alpha_i \beta_j}} 2^{\frac{i+j}{2}}\right)} \leq \frac{i\alpha_i + j\beta_j}{\log\left(\frac{\sqrt{ij}}{18\lambda} 2^{\frac{i+j}{2}}\right)} \leq \frac{i\alpha_i + j\beta_j}{\log(2^{\frac{i+j}{2}}/\lambda)},$$

where the second inequality was derived from the fact that  $x + y \geq 2\sqrt{xy}$  for any  $x, y \geq 0$ , and the last one by the fact that  $i, j \geq 20$ .

Notice that if  $2^{\frac{i+j}{4}} \leq \lambda$  then  $i, j \leq 4 \log d$  and thus  $(i, j) \notin \Gamma$ . We therefore have  $2^{\frac{i+j}{4}} > \lambda$  and so

$$\log(2^{\frac{i+j}{2}}/\lambda) > \log(2^{\frac{i+j}{4}}) = (i + j)/4.$$

This verifies (3.20) and further implies that  $\xi_{ij} \leq 4(\alpha_i + \beta_j)$ . Altogether,

$$\begin{aligned} \sum_{(i,j) \in \Gamma} \xi_{ij} &\leq 4 \sum_{|i-j| \leq 2 \log d} (\alpha_i + \beta_j) \leq 4(1 + 4 \log d) \left( \sum_i \alpha_i + \sum_j \beta_j \right) \\ &\leq (32 \log d + 8)(\alpha + \beta). \end{aligned} \tag{3.22}$$

Combining (3.19) with the two cases (3.21), (3.22) for the  $\xi_{ij}$ 's proves that

$$\begin{aligned} \sum_{\substack{i+j>0 \\ |i-j| \leq 2 \log d}} w_{ij}^* \mathbb{1}_{\{z_{ij}^* \geq 2\}} &\leq (32 \log^2 d + 160 \log d + 1558)(\alpha + \beta) + n^{-1/3} \\ &\leq 74(\alpha + \beta) \log^2 d + n^{-1/3}, \end{aligned} \tag{3.23}$$

where in the last inequality we plugged in the fact that  $d \geq 320$ .

It remains to consider the case  $z_{ij}^* < 2$  where by definition

$$w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} = 18 \frac{e(S_i, T_j)}{2^{i+j}}.$$

Since the  $(m, d, \lambda)$ -graph  $G$  satisfies  $e(S_i, T_j) \leq (d/m)s_i t_j + \lambda\sqrt{s_i t_j}$ , we have the following two cases:

- **Case (i):**  $e(S_i, T_j) \leq 2(d/m)s_i t_j$ .

The above bound on  $w_{ij}^*$  then translates into

$$w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} \leq 36 \frac{d}{m} \frac{s_i t_j}{2^{i+j}} = 36 \frac{d}{m} \alpha_i \beta_j,$$

and summing over all such  $i, j$  while recalling that  $\sqrt{\alpha\beta} \leq 2m/\lambda$  we get

$$\begin{aligned} \sum_{i,j} w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} &\leq 36 \frac{d}{m} \sum_i \alpha_i \sum_j \beta_j \leq 144 \frac{d}{m} \alpha\beta \\ &\leq 144 \frac{d}{m} \sqrt{\alpha\beta} \frac{2m}{\lambda} \leq 288\lambda \sqrt{\alpha\beta}, \end{aligned} \tag{3.24}$$

where we used the inequalities  $\sum_i \alpha_i \leq 2\alpha$ ,  $\sum_j \beta_j \leq 2\beta$  and  $\lambda \geq \sqrt{d}$ .

- **Case (ii):**  $e(S_i, T_j) \leq 2\lambda\sqrt{s_i t_j}$ .

Here we have

$$w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} \leq 36\lambda \frac{\sqrt{s_i t_j}}{2^{i+j}} = \frac{36\lambda}{2^{(i+j)/2}} \sqrt{\alpha_i \beta_j},$$

and so

$$\begin{aligned} \sum_{i,j} w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} &\leq 36\lambda \sum_{i,j} \sqrt{\frac{\alpha_i}{2^i}} \sqrt{\frac{\beta_j}{2^j}} \\ &= 36\lambda \sum_{k \geq 0} \sum_i \sqrt{\frac{\alpha_i}{2^i}} \sqrt{\frac{\beta_{i+k}}{2^{i+k}}} + 36\lambda \sum_{k > 0} \sum_j \sqrt{\frac{\beta_j}{2^j}} \sqrt{\frac{\alpha_{j+k}}{2^{j+k}}}. \end{aligned}$$

By Cauchy–Schwarz,

$$\begin{aligned} \sum_{k \geq 0} \sum_i \sqrt{\frac{\alpha_i}{2^i}} \sqrt{\frac{\beta_{i+k}}{2^{i+k}}} &\leq \sum_{k \geq 0} \sqrt{\left(\sum_i \frac{\alpha_i}{2^i}\right) \left(\sum_i \frac{\beta_{i+k}}{2^{i+k}}\right)} \leq \sum_{k \geq 0} \sqrt{2^{-k} \sum_i \alpha_i \sum_j \beta_j} \\ &\leq 2\sqrt{\alpha\beta} \sum_{k \geq 0} 2^{-k/2} \leq 7\sqrt{\alpha\beta}, \end{aligned}$$

and similarly,

$$\sum_{k > 0} \sum_j \sqrt{\frac{\beta_j}{2^j}} \sqrt{\frac{\alpha_{j+k}}{2^{j+k}}} \leq 7\sqrt{\alpha\beta}.$$

We deduce that in this case

$$\sum_{i,j} w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} \leq 504\lambda \sqrt{\alpha\beta}. \tag{3.25}$$

Adding the bounds obtained for the two cases (3.24), (3.25) gives

$$\sum_{i,j} w_{ij}^* \mathbb{1}_{\{z_{ij}^* < 2\}} \leq 792\lambda\sqrt{\alpha\beta}.$$

The proof of Lemma 3.3 is now concluded by combining the above bound with (3.23).  $\square$

Notice that for proving Proposition 3.2 we can assume that  $d \geq 320$ , since otherwise  $d < 5 \log^2 d$  and the statement immediately follows from the trivial bound

$$e(A, B) \leq d(|A| \wedge |B|) \leq \frac{d}{2}(|A| + |B|) < \frac{5}{2}(|A| + |B|) \log^2 d.$$

For  $d \geq 320$  we can apply Lemma 3.3 (recalling the discussion preceding this lemma) combined with (3.15) and obtain that, except with probability  $n^{-100}$ , every two subsets  $A, B$  with  $\alpha\beta \leq (2m/\lambda)^2$  satisfy

$$e(A, B)/n \leq 802\lambda\sqrt{\alpha\beta} + 74(\alpha + \beta) \log^2 d + 2(\alpha + \beta) + n^{-1/3}.$$

When the subsets  $A, B$  satisfy in addition

$$|A| + |B| \geq n^{2/3}$$

then  $\alpha + \beta \geq n^{-1/3}$  and the above bound (for  $d \geq 320$ ) translates to

$$e(A, B)/n \leq 802\lambda\sqrt{\alpha\beta} + 75(\alpha + \beta) \log^2 d. \tag{3.26}$$

Altogether, we have established the statement of Proposition 3.2 under the additional assumption  $|A| + |B| \geq n^{2/3}$  for the subsets  $A, B$  in mention.

The separate case of  $|A| + |B| < n^{2/3}$  is much simpler to handle, and is treated by the next claim using a standard first moment argument.

**Claim 3.4.** *Let  $G$  be an arbitrary graph on  $m$  vertices and let  $H$  be a random  $n$ -lift of  $G$  with  $n \geq m$ . Then with probability  $1 - O(n^{-100})$ , every two subsets  $A, B \subset V(H)$  of size  $|A| + |B| \leq n^{2/3}$  have  $e(A, B) \leq 50(|A| + |B|)$ .*

**Proof.** Suppose that  $A, B$  are two subsets that satisfy  $|A| + |B| \leq n^{2/3}$  and  $e(A, B) \geq 50(|A| + |B|)$ , and consider their union  $R = A \cup B$ . Clearly,

$$|R| \leq |A| + |B| \leq n^{2/3}$$

whereas the number of edges in the induced subgraph on  $R$  satisfies

$$e(R) \geq e(A, B)/2 \geq 25|R|.$$

As argued below Eq. (3.11), if  $(i_1 j_1, i'_1 j'_1), \dots, (i_k j_k, i'_k j'_k)$  are  $k$  arbitrary distinct pairs of vertices in  $H$  of which no  $\frac{2}{3}n$  vertices share the same fiber,

$$\mathbb{P}(i_l j_l \sim i'_l j'_l \text{ for all } l \in [k]) \leq (3/n)^k$$

(that argument applies to any base-graph  $G$  by definition of the  $n$ -lift). Consider  $k$  distinct pairs of vertices in  $R$  that may potentially be adjacent in  $H$ . Clearly, for large enough  $n$  these do not contain any  $\frac{2}{3}n$  points on the same fiber since  $|R| = o(n)$ , hence the probability that they are all adjacent is at most  $(3/n)^k$ .

It now follows that the probability there exists a subset  $R \subset V(H)$  of size  $|R| = r \leq n^{2/3}$  with  $e(R) \geq 25r$  is at most

$$\binom{mn}{r} \binom{\binom{r}{2}}{25r} \left(\frac{3}{n}\right)^{25r} \leq \left(\frac{emn}{r} \left(\frac{3er}{50n}\right)^{25}\right)^r \leq (cm(r/n)^{24})^r$$

where  $c > 0$  is an absolute constant. As  $r \leq n^{2/3}$  and  $m \leq n$ , the base of the exponent in the last expression is at most  $O(n^{-7})$  whereas  $r \geq 50$  necessarily to allow  $e(R) \geq 25r$ . Summing this error probability over the  $n^{2/3}$  possible values of  $r$  completes the proof.  $\square$

As the bound given in the above claim is clearly smaller than the bound (3.26) for  $d \geq 320$ , together they imply that (3.26) holds with probability  $1 - O(n^{-100})$  for any two sets  $A, B$  with  $|A||B| \leq (2mn/\lambda)^2$ . This concludes the proof of Proposition 3.2.

#### 4. Light pairs and epsilon-nets

We now move on to estimating the expected total of all light pairs along the edges of  $H$ . Recall that a pair  $x_{ij}, y_{i'j'}$  is *light* if  $|x_{ij}y_{i'j'}| < \lambda/mn$ . To bound the bilinear form  $x^T A_H y$  with respect to the light pairs, we will approximate each such vector using an  $\varepsilon$ -net, where

$$\varepsilon = \frac{1}{d\sqrt{mn}}.$$

More precisely, we consider the  $mn$ -dimensional lattice  $\mathcal{L} = (\varepsilon\mathbb{Z})^{mn}$ , and show that the required statement on the bilinear form holds for any two vectors  $x, y$  with norm at most 1 in this lattice.

**Theorem 4.1.** *Let  $G$  be an  $(m, d, \lambda)$ -graph for  $d \geq 3$  and  $\lambda \geq \sqrt{d}$  and let  $H$  be a random  $n$ -lift of  $G$ . For  $x, y \in \mathbb{R}^{mn}$ , let  $R_l(x, y)$  be the random variable*

$$R_l(x, y) = \sum_{ij \sim i'j'} x_{ij} y_{i'j'} \mathbb{1}_{\{|x_{ij}y_{i'j'}| < \lambda/mn\}}.$$

Let  $\mathcal{L}$  denote the  $mn$ -dimensional lattice  $(\frac{1}{d\sqrt{mn}}\mathbb{Z})^{mn}$ . Then except with probability  $O(\exp(-mn))$ , every  $x, y \in \mathcal{L}$  with  $\|x\| \leq 1$  and  $\|y\| \leq 1$  satisfy  $|R_l(x, y) - \mathbb{E}[R_l(x, y)]| \leq 250\lambda \log d$ .

**Proof.** In order to establish the above concentration result, we must first estimate the variance of  $R_l(x, y)$ .

**Lemma 4.2.** *Let  $G$  be an  $(m, d, \lambda)$ -graph with  $\lambda \geq \sqrt{d}$ , and let  $x, y \in \mathbb{R}^{mn}$  be two fixed vectors satisfying  $\|x\| \leq 1$  and  $\|y\| \leq 1$ . Then*

$$\sum_{i,i'} \sum_{j \sim j'} x_{ij}^2 y_{i'j'}^2 \mathbb{1}_{\{|x_{ij}y_{i'j'}| < \lambda/mn\}} \leq 50 \frac{\lambda^2 \log d}{m}.$$

**Proof.** As in the treatment of the heavy pairs, we consider the following dyadic expansion of  $x$  and  $y$ :

$$\begin{aligned} \mathcal{D}_\ell &= \left\{ (i, j) : 2^{-\ell} \leq |x_{ij}| \sqrt{\frac{mn}{\lambda}} < 2^{-\ell+1} \right\} \quad (\ell \in \mathbb{Z}), \\ \mathcal{D}'_{\ell'} &= \left\{ (i', j') : 2^{-\ell'} \leq |y_{i'j'}| \sqrt{\frac{mn}{\lambda}} < 2^{-\ell'+1} \right\} \quad (\ell' \in \mathbb{Z}), \end{aligned}$$

and the assumptions  $\|x\| \leq 1$  and  $\|y\| \leq 1$  translate into

$$\sum_{\ell} 4^{-\ell} |\mathcal{D}_\ell| \frac{\lambda}{mn} \leq 1, \quad \sum_{\ell'} 4^{-\ell'} |\mathcal{D}'_{\ell'}| \frac{\lambda}{mn} \leq 1. \tag{4.1}$$

Further note that, if  $ij \in \mathcal{D}_\ell$  and  $i'j' \in \mathcal{D}'_{\ell'}$ , then a necessary condition for  $|x_{ij}y_{i'j'}| < \lambda/mn$  is that  $\ell + \ell' > 0$ .

Consider the graph  $G'$  where every two fibers that are connected in  $G$  have a complete bipartite graph between them in  $G'$ . That is,  $ij \sim i'j'$  in  $G'$  if  $jj' \in E(G)$ . It follows that

$$\sum_{j \sim j'} \sum_{i,i'} x_{ij}^2 y_{i'j'}^2 \mathbb{1}_{\{|x_{ij}y_{i'j'}| < \lambda/mn\}} \leq 16 \sum_{\ell+\ell'>0} 4^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2, \tag{4.2}$$

and we aim to bound the sum in the right-hand side by at most  $3 \frac{\lambda^2 \log d}{m}$ .

The adjacency matrix of  $G'$  is therefore precisely  $A_G \otimes J_n$ , where  $J_n$  is the all-ones matrix of order  $n$  and  $\otimes$  denotes tensor product, and so by the definition of  $G$  (and the properties of tensor products) it follows that  $G'$  is an  $(mn, dn, \lambda n)$ -graph. As such, for any two subsets  $A, B$  of its vertices,

$$e_{G'}(A, B) \leq \frac{d}{m} |A||B| + \lambda n \sqrt{|A||B|}. \tag{4.3}$$

We now separate the sum in (4.2) into two cases, comparing  $|\ell - \ell'|$  to

$$D = \log\left(\frac{d}{\lambda \log d}\right) + 2 \leq \frac{1}{2} \log d - 1.$$

To justify the last inequality, note first that we may assume that  $d \geq 256$  otherwise the statement of the lemma holds trivially. Indeed, since there are  $dm/2$  edges in  $G$ , summing over  $\frac{1}{2}dmn^2$  pairs, each of which contributes at most  $(\lambda/mn)^2$ , gives at most  $\frac{\lambda^2 d}{2m}$ . For  $d \leq 256$  we have  $d \leq 32 \log d$  and so this is clearly at most  $16 \frac{\lambda^2 \log d}{m}$  and we are done. Assume therefore that

$d \geq 256$ , in which case  $D + 1 = \log(\frac{8d}{\lambda \log d}) \leq \log(d/\lambda)$  and the above inequality follows from the fact that  $\lambda \geq \sqrt{d}$ .

In case  $D \leq 0$  we have  $d \leq \frac{1}{4}\lambda \log d$ . Here, applying the trivial bound  $e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \leq dn|\mathcal{D}_\ell|$  gives

$$\begin{aligned} \sum_{\substack{\ell+\ell'>0 \\ \ell' \geq \ell}} 2^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 &\leq \frac{\lambda}{mn} \sum_{\ell} 4^{-\ell} \cdot dn|\mathcal{D}_\ell| \cdot \frac{\lambda}{mn} \sum_{\ell' \geq \ell} 2^{-(\ell'-\ell)} \\ &\leq \frac{2\lambda d}{m} \leq \frac{\lambda^2 \log d}{2m}, \end{aligned}$$

where the inequality between the two lines used (4.1). Performing the same calculation for the sum over  $\ell \geq \ell'$  gives the same bound. Altogether these two bounds sum up to  $\frac{\lambda^2 \log d}{m}$  and we get that for  $D \leq 0$ ,

$$\begin{aligned} \sum_{\ell+\ell'>0} 4^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 \\ \leq \frac{1}{2} \sum_{\ell+\ell'>0} 2^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 \leq \frac{\lambda^2 \log d}{2m} \end{aligned}$$

(the factor of  $\frac{1}{2}$  in the first inequality due to the fact that  $\ell + \ell' > 0$ ). Thus, Eq. (4.2) translates this bound to  $8\frac{\lambda^2 \log d}{m}$  and confirms the statement of the lemma for the case  $D \leq 0$ . It remains to handle  $D > 0$ .

• **Case (i):**  $|\ell - \ell'| \geq D$ .

In this case, we use the trivial bound  $e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \leq dn|\mathcal{D}_\ell|$ , giving

$$\begin{aligned} \sum_{k \geq D} \sum_{\substack{\ell+\ell'>0 \\ \ell'-\ell=k}} 4^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 \\ \leq \sum_{k \geq D} \sum_{\substack{\ell+\ell'>0 \\ \ell'-\ell=k}} 2^{-4\ell-2k} dn|\mathcal{D}_\ell| \left(\frac{\lambda}{mn}\right)^2 \\ = \sum_{\ell} 4^{-\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell| \sum_{\substack{k \geq D \\ \ell+(k+\ell)>0}} 2^{-(2\ell+k)} (2^{-k}d) \frac{\lambda}{m} \\ \leq \sum_{\ell} 4^{-\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell| \frac{\lambda^2 \log d}{4m}, \end{aligned}$$

where we used the facts that  $2^k \geq 4d/(\lambda \log d)$  for  $k \geq D$ , and that  $\sum 2^{-(2\ell+k)} \leq 1$ . By (4.1), it now follows that



$$\sum_{k \geq D} \sum_{\substack{\ell + \ell' > 0 \\ \ell' - \ell = k}} 4^{-(\ell + \ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left( \frac{\lambda}{mn} \right)^2 \leq \frac{\lambda^2 \log d}{4m},$$

and adding the symmetric case where we sum over  $\ell - \ell' = k$ , we get

$$\sum_{\substack{\ell + \ell' > 0 \\ |\ell' - \ell| \geq D}} 4^{-(\ell + \ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left( \frac{\lambda}{mn} \right)^2 \leq \frac{\lambda^2 \log d}{2m}.$$

• **Case (ii):**  $|\ell - \ell'| < D$ .

Here we have two bounds according to the two expressions in the upper bound (4.3). That is, we break  $e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'})$  into the sum of the two expressions corresponding to  $\frac{d}{m} |\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|$  and to  $\lambda n \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|}$  and bound each of them separately.

First, by (4.1) the sum corresponding to  $\frac{d}{m} |\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|$  contributes

$$\begin{aligned} & \sum_{0 \leq k < D} \sum_{\substack{\ell + \ell' > 0 \\ \ell' - \ell = k}} 4^{-(\ell + \ell')} \frac{d}{m} |\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}| \left( \frac{\lambda}{mn} \right)^2 \\ & \leq \lceil D \rceil \frac{d}{m} \sum_{\ell} 4^{-\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell| \sum_{\ell'} 4^{-\ell'} \frac{\lambda}{mn} |\mathcal{D}'_{\ell'}| \leq \lceil D \rceil \frac{d}{m} \leq \frac{\lambda^2 \log d}{2m}, \end{aligned}$$

where the last inequality is due to the facts  $\lambda^2 \geq d$  and  $\lceil D \rceil \leq \frac{1}{2} \log d$ . Second, the sum corresponding to  $\lambda n \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|}$  contributes

$$\begin{aligned} & \sum_{0 \leq k < D} \sum_{\substack{\ell + \ell' > 0 \\ \ell' - \ell = k}} 4^{-(\ell + \ell')} \lambda n \sqrt{|\mathcal{D}_\ell| |\mathcal{D}'_{\ell'}|} \left( \frac{\lambda}{mn} \right)^2 \\ & = \sum_{0 \leq k < D} \frac{\lambda^2}{m} \sum_{\substack{\ell + \ell' > 0 \\ \ell' - \ell = k}} 2^{-(\ell + \ell')} \sqrt{2^{-2\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell|} \sqrt{2^{-2\ell'} \frac{\lambda}{mn} |\mathcal{D}'_{\ell'}|} \\ & \leq \sum_{0 \leq k < D} \frac{\lambda^2}{2m} \sqrt{\sum_{\ell > -k/2} 2^{-2\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell|} \sqrt{\sum_{\ell' > k/2} 2^{-2\ell'} \frac{\lambda}{mn} |\mathcal{D}'_{\ell'}|} \\ & \leq \frac{\lceil D \rceil \lambda^2}{2m} \sqrt{\sum_{\ell} 2^{-2\ell} \frac{\lambda}{mn} |\mathcal{D}_\ell|} \sqrt{\sum_{\ell'} 2^{-2\ell'} \frac{\lambda}{mn} |\mathcal{D}'_{\ell'}|} \\ & \leq \frac{\lceil D \rceil \lambda^2}{2m} \leq \frac{\lambda^2 \log d}{4m}, \end{aligned}$$

where the inequality in the third line followed from Cauchy–Schwarz and in the last line we used (4.1) together with the fact  $\lceil D \rceil \leq \frac{1}{2} \log d$ . The last two inequalities now give a

combined bound of  $\frac{3\lambda^2 \log d}{4m}$ . As the same holds for the sum over  $\ell - \ell' = k$ , altogether we have

$$\sum_{\substack{\ell+\ell'>0 \\ |\ell'-\ell|<D}} 4^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 \leq \frac{3\lambda^2 \log d}{2m}.$$

Adding together Cases (i), (ii) implies that when  $D \geq 0$ ,

$$\sum_{\ell+\ell'>0} 4^{-(\ell+\ell')} e_{G'}(\mathcal{D}_\ell, \mathcal{D}'_{\ell'}) \left(\frac{\lambda}{mn}\right)^2 \leq \left(\frac{1}{2} + \frac{3}{2}\right) \frac{\lambda^2 \log d}{m} = \frac{2\lambda^2 \log d}{m},$$

and (4.2) now translates this bound to  $32\frac{\lambda^2 \log d}{m}$ , confirming the statement of the lemma (with room to spare) for the case  $D > 0$  as required.  $\square$

Next, we need to address the support of  $x$ . A vector  $x \in \mathbb{R}^{mn}$  is called *sparse* if it has at most  $n/2$  non-zero entries on each fiber, that is, if

$$|\{i \in [n]: x_{ij} \neq 0\}| \leq n/2 \quad \text{for all } j \in V(G).$$

The next lemma establishes concentration for  $R_l(x, y)$  provided that  $x$  is sparse.

**Lemma 4.3.** *Let  $x, y \in \mathbb{R}^{mn}$  be two fixed vectors such that  $x$  is sparse,  $\|x\| \leq 1$  and  $\|y\| \leq 1$ . Let  $G$  be an  $(m, d, \lambda)$ -graph for  $d \geq 3$  and  $\lambda \geq \sqrt{d}$  and let  $H$  be a random  $n$ -lift of  $G$ . For any  $a \geq 125$ ,*

$$\mathbb{P}(|R_l(x, y) - \mathbb{E}[R_l(x, y)]| > a\lambda \log d) \leq 2d^{-amn/12}.$$

**Proof.** For any  $jj' \in E(G)$  and  $i \in [n]$ , define

$$X_{jj'}(i) = x_{ij} \sum_{i'=1}^n \mathbb{1}_{\{i \sim i'j'\}} y_{i'j'} \mathbb{1}_{\{|x_{ij} y_{i'j'}| < \frac{\lambda}{mn}\}}.$$

By this definition,

$$R_l(x, y) = \sum_{j \sim j'} \sum_i X_{jj'}(i),$$

and we can now expose the values of  $X_{jj'}$  sequentially by going over the pairs of fibers  $jj' \in E(G)$  one-by-one (in an arbitrary order), and for each such pair revealing the relevant part of the bijection between the fibers. More precisely, when processing a given pair of fibers  $j \sim j'$ , we proceed as follows:

- (1) Without loss of generality, suppose  $|x_{1j}| \geq |x_{2j}| \geq \dots \geq |x_{nj}|$ , and let  $q \in [n]$  be the largest index such that  $x_{qj} \neq 0$ .

(2) Sequentially go over  $i = 1, \dots, q$  and expose the neighbor of  $ij$  in the bijection between the two fibers, i.e.,  $J_i \in [n]$  such that  $ij \sim J_i j'$ , thereby determining  $X_{jj'}(i)$ .

Crucially, since the vector  $x$  is sparse, it contains at most  $n/2$  non-zero entries in any given fiber, and so in the above defined process  $q \leq n/2$ .

Denote by  $(\mathcal{F}_t)$  the filter corresponding to this process (that is,  $\mathcal{F}_t$  is the  $\sigma$ -algebra generated by the first  $t$  exposed edges), and let  $(S_t)$  be Doob’s martingale corresponding to the function  $R_t(x, y)$  with respect to  $(\mathcal{F}_t)$ :

$$S_t = \mathbb{E} \left[ \sum_{j \sim j'} \sum_i X_{jj'}(i) \mid \mathcal{F}_t \right].$$

As usual,  $S_0 = \mathbb{E}[R_t(x, y)]$  whereas at the end of the process the martingale equals  $R_t(x, y)$ .

We wish to analyze the increment  $S_t - S_{t-1}$ . Suppose that in step  $t$  we are now exposing an edge between the fibers  $j \sim j'$ . Clearly, if  $(k, k') \neq (j, j')$  and our process already exposed the edges between the fibers  $k \sim k'$ , then their contribution is canceled in  $S_t - S_{t-1}$ . Furthermore, if the edges between  $k \sim k'$  are to be exposed in the future, then

$$\mathbb{E} \left[ \sum_i X_{kk'}(i) \mid \mathcal{F}_t \right] = \mathbb{E} \left[ \sum_i X_{kk'}(i) \mid \mathcal{F}_{t-1} \right],$$

since the bijections between distinct pairs of connected fibers are independent. That is to say,  $S_t - S_{t-1}$  contains only terms that arise from the effect of the edge exposed at time  $t$  on  $\mathbb{E}[\sum_i X_{jj'}(i) \mid \mathcal{F}_t]$ .

In light of this, it suffices to treat the case where  $j \sim j'$  is the first pair of fibers processed, and the analysis of  $S_t - S_{t-1}$  will hold analogous for any other pair. In what follows, since we are now concentrating solely on the two fibers  $j \sim j'$ , we omit the subscripts  $j, j'$  from  $x$  and  $y$  to simplify the notation. Similarly, we use the abbreviation

$$L_{i,i'} = \mathbb{1}_{\{|x_i y_{i'}| < \frac{\lambda}{m_i}\}}. \tag{4.4}$$

At step  $t = 1, 2, \dots, q$  we are therefore exposing the match of  $x_t$ . For simplicity we will analyze  $S_1 - S_0$  and by merely changing the indices the same argument would carry to all other values of  $t$ . Recall that

$$S_0 = \frac{1}{n!} \sum_{\pi} \sum_i x_i y_{\pi(i)} L_{i,\pi(i)} = \frac{1}{n} \sum_i \sum_{i'} x_i y_{i'} L_{i,i'} \tag{4.5}$$

and that given the event that  $x_1$  is matched to some  $I \in [n]$  we have

$$\begin{aligned} S_1 &= x_1 y_I L_{1,I} + \frac{1}{(n-1)!} \sum_{\pi: \pi(1)=I} \sum_{i \geq 2} x_i y_{\pi(i)} L_{i,\pi(i)} \\ &= x_1 y_I L_{1,I} + \frac{1}{n-1} \sum_{i \geq 2} \sum_{i' \neq I} x_i y_{i'} L_{i,i'}. \end{aligned} \tag{4.6}$$

Let  $S_1$  and  $\tilde{S}_1$  be two possible values after revealing the match of  $x_1$ , denoting its index by  $\pi(1)$  and  $\tilde{\pi}(1)$  respectively. We can now couple the distributions over the remaining entries of  $\pi$  and  $\tilde{\pi}$  via switching  $\pi(1), \tilde{\pi}(1)$ . That is, if we let  $z = \tilde{\pi}^{-1}(\pi(1))$  then  $\pi, \tilde{\pi}$  agree everywhere except possibly on  $\{1, z\}$  and there we have

$$\pi(z) = \tilde{\pi}(1), \quad \tilde{\pi}(z) = \pi(1).$$

Clearly, in any pair of coupled  $\pi, \tilde{\pi}$  all summands of the form  $x_i y_{i'} L_{i,i'}$  in Eq. (4.6) cancel from  $S_1 - \tilde{S}_1$  except when  $i \in \{1, z\}$ , hence

$$\begin{aligned} S_1 - \tilde{S}_1 &= x_1 y_{\pi(1)} L_{1,\pi(1)} + \frac{1}{n-1} \sum_{z \neq 1} x_z y_{\tilde{\pi}(1)} L_{z,\tilde{\pi}(1)} \\ &\quad - x_1 y_{\tilde{\pi}(1)} L_{1,\tilde{\pi}(1)} - \frac{1}{n-1} \sum_{z \neq 1} x_z y_{\pi(1)} L_{z,\pi(1)}. \end{aligned}$$

By definition (4.4) each of the above terms  $x_i y_{i'} L_{i,i'}$  is at most  $\lambda/mn$  in absolute value, thus repeating this argument for any step  $t$  gives that with probability 1,

$$|S_t - S_{t-1}| \leq \frac{4\lambda}{mn} \quad \text{for } t = 1, \dots, q. \tag{4.7}$$

Obtaining an  $L^2$  bound on the increments  $S_t - S_{t-1}$  is slightly more delicate than the above  $L^\infty$  bound. To this end, we will write  $S_1 - S_0$  explicitly: Recall from (4.5), (4.6) that  $S_0$  averages over permutations  $\pi$  on  $[n]$  whereas  $S_1$  averages over all such permutations  $\tilde{\pi}$  that have  $\tilde{\pi}(1) = I$  for some  $I \in [n]$ , which is exposed in  $\mathcal{F}_1$  and identifies the match of  $x_1$ .

In other words,  $S_0$  is the mean of sums analogous to  $S_1$  with all possible values  $\pi(1) \in \{1, \dots, n\}$  replacing  $I$ . Each such value has equal probability and the case  $\pi(1) = I$  does not contribute to  $S_1 - S_0$ . In the remaining cases we can go over the possible values of  $z = \pi^{-1}(I) \in \{2, \dots, n\}$  (each with equal probability) and couple  $\pi, \tilde{\pi}$  using the switching that was used to establish the  $L^\infty$  bound, letting them agree everywhere except on  $\{1, z\}$ . Altogether we obtain that

$$\begin{aligned} S_1 - S_0 &= x_1 y_I L_{1,I} - \frac{1}{n-1} \sum_{z \neq 1} x_z y_I L_{z,I} - \frac{1}{n} \sum_{i'} x_1 y_{i'} L_{1,i'} \\ &\quad + \frac{1}{n(n-1)} \sum_{i'} \sum_{z \neq 1} x_z y_{i'} L_{z,i'}. \end{aligned}$$

Since the expressions in the last line do not depend on  $I$ , we conclude that  $\text{Var}(S_1 - S_0) = \text{Var}(Z)$  where

$$Z = x_1 y_I L_{1,I} - \frac{1}{n-1} \sum_{z \neq 1} x_z y_I L_{z,I} \quad \text{with } I \text{ uniform on } [n].$$

Estimating  $\text{Var}(Z)$  requires extra care due to the indicators  $L_{i,i'}$ . There are two possible cases:

- (i) If  $L_{1,I} = 1$  then by our ordering of the  $x_i$ 's according to decreasing absolute values we have  $L_{z,I} = 1$  for all  $z \geq 1$  and so

$$Z = x_1 y_I L_{1,I} - \frac{1}{n-1} \sum_{z \neq 1} x_z y_I L_{z,I} = \left( x_1 - \frac{1}{n-1} \sum_{z \neq 1} x_z \right) y_I.$$

Moreover,  $|x_1|$  is at least the average of the  $|x_z|$ 's for  $z > 1$ , and so in this case

$$Z^2 \leq 4x_1^2 y_I^2 L_{1,I}.$$

- (ii) Otherwise,  $L_{1,I} = 0$  and there exists some  $T = T(I) > 1$  such that  $L_{z,I} = 1$  for all  $z \geq T$ . In this case

$$Z = -\frac{1}{n-1} \sum_{z \geq T} x_z y_I.$$

By Cauchy–Schwarz, in this case we thus have

$$Z^2 \leq \frac{1}{n-1} \sum_{z > 1} x_z^2 y_I^2 L_{z,I}.$$

Combining the cases, since  $I$  is uniform on  $[n]$  it now follows that

$$\text{Var}(Z) \leq \mathbb{E}Z^2 \leq \frac{4}{n} \sum_{i'} x_1^2 y_{i'}^2 L_{1,i'} + \frac{1}{n(n-1)} \sum_{z > 1} \sum_{i'} x_z^2 y_{i'}^2 L_{z,i'}.$$

Applying the same analysis to a general  $t \in [q]$ , while assuming without loss of generality that the remaining unmatched  $y_{i'}$ 's are  $\{y_t, \dots, y_n\}$ , yields

$$\begin{aligned} \text{Var}(S_t - S_{t-1} \mid \mathcal{F}_{t-1}) &\leq \frac{4}{n-t+1} \sum_{i' \geq t} x_t^2 y_{i'}^2 L_{t,i'} \\ &\quad + \frac{1}{(n-t+1)(n-t)} \sum_{z > t} \sum_{i' \geq t} x_z^2 y_{i'}^2 L_{z,i'}. \end{aligned} \tag{4.8}$$

At this point our assumption that  $q \leq n/2$  due to the fact that  $x$  is sparse plays its important role. For some  $z, i'$  consider the total coefficient of  $x_z^2 y_{i'}^2 L_{z,i'}$  after summing (4.8) over  $t = 1, \dots, q$ . The first expression in (4.8) contributes at most  $4/(n-q+1) \leq 8/n$  whereas the second one adds up to

$$\sum_{t \leq q} \frac{1}{(n-t+1)(n-t)} \leq \frac{2}{n}.$$

Altogether we conclude that

$$\sum_{t=1}^q \text{Var}(S_t - S_{t-1} | \mathcal{F}_{t-1}) \leq \frac{10}{n} \sum_{i,i'} x_i^2 y_{i'}^2 L_{i,i'}$$

and by extending this analysis to all  $md/2$  pairs of connected fibers we get

$$\begin{aligned} \sum_t \text{Var}(S_t - S_{t-1} | \mathcal{F}_{t-1}) &\leq \frac{10}{n} \sum_{j \sim j'} \sum_{i,i'} x_{ij}^2 y_{i'j'}^2 \mathbb{1}_{\{|x_{ij} y_{i'j'}| < \lambda/mn\}} \\ &\leq 500 \frac{\lambda^2 \log d}{mn}, \end{aligned} \tag{4.9}$$

where the last step was by Lemma 4.2. We can now apply the following large deviation inequality for martingales, which is a special case of a result of Freedman [10] (see also [23] for a variant of this inequality).

**Theorem 4.4.** *Let  $(S_i)_{i=0}^n$  be a martingale with respect to a filter  $(\mathcal{F}_i)$  and let  $\Delta_i = S_i - S_{i-1}$  denote its increments. Suppose that  $|\Delta_i| \leq M$  for all  $i$  and that  $\sum_{i=1}^n \text{Var}(\Delta_i | \mathcal{F}_{i-1}) \leq \sigma^2$ . Then for any  $s > 0$  we have*

$$\mathbb{P}(|S_t - S_0| \geq s \text{ for some } t \in [n]) \leq 2 \exp\left[-\frac{s^2}{2(\sigma^2 + Ms)}\right].$$

In our case by (4.7) and (4.9) we have  $M = 4 \frac{\lambda}{mn}$  and  $\sigma^2 = 500 \frac{\lambda^2 \log d}{mn}$ . Since the final value of  $(S_t)$  is  $R_l(x, y)$  whereas  $S_0 = \mathbb{E}[R_l(x, y)]$ , we now get

$$\mathbb{P}(|R_l(x, y) - \mathbb{E}[R_l(x, y)]| \geq s) \leq 2 \exp\left(-\frac{s^2 mn}{8\lambda(125\lambda \log d + s)}\right).$$

In particular, for  $s = a\lambda \log d$  with  $a \geq 125$  we have  $8\lambda(125\lambda \log d + s) \leq 16\lambda s$  and so

$$\mathbb{P}(|R_l(x, y) - \mathbb{E}[R_l(x, y)]| \geq s) \leq 2 \exp\left(-\frac{smn}{16\lambda}\right) = 2d^{-amn/(16 \ln 2)} < 2d^{-amn/12},$$

as required.  $\square$

As a corollary, we can now infer the concentration result of Lemma 4.3 without requiring that  $x, y$  should be sparse.

**Corollary 4.5.** *Let  $x, y \in \mathbb{R}^{mn}$  be two fixed vectors with  $\|x\| \leq 1$  and  $\|y\| \leq 1$ . Let  $G$  be an  $(m, d, \lambda)$ -graph for  $d \geq 3$  and  $\lambda \geq \sqrt{d}$  and let  $H$  be a random  $n$ -lift of  $G$ . Then for any  $a \geq 250$ ,*

$$\mathbb{P}(|R_l(x, y) - \mathbb{E}[R_l(x, y)]| > a\lambda \log d) \leq 4d^{-amn/24}.$$

**Proof.** Let  $x, y \in \mathbb{R}^{mn}$  satisfy  $\|x\| \leq 1$  and  $\|y\| \leq 1$ . Define the sparse vectors  $x', x'' \in \mathbb{R}^{mn}$  by

$$x'_{ij} = \begin{cases} x_{ij}, & i \leq n/2, \\ 0, & i > n/2, \end{cases} \quad \text{and} \quad x''_{ij} = \begin{cases} 0, & i > n/2, \\ x_{ij}, & i \leq n/2 \end{cases} \quad \text{for all } i, j.$$

Since for every  $i, j$  we have that  $x_{ij}$  is precisely one of  $\{x'_{ij}, x''_{ij}\}$  while the other is 0, we deduce that

$$R_l(x, y) = R_l(x', y) + R_l(x'', y). \tag{4.10}$$

Clearly, by the triangle inequality if  $|R_l(x, y) - \mathbb{E}R_l(x, y)| \geq a\lambda \log d$  then at least one of the variables  $R_l(x', y), R_l(x'', y)$  must deviate from its mean by at least  $\frac{1}{2}a\lambda \log d$ . For each of the pairs  $x', y$  and  $x'', y$  we may apply Lemma 4.3 for a choice of  $a/2 \geq 125$  and obtain that

$$\mathbb{P}(|R_l(x', y) - \mathbb{E}[R_l(x', y)]| > (a/2)\lambda \log d) \leq 2d^{-amn/24},$$

and the same applies to  $x''$ . The required result immediately follows.  $\square$

To carry the result from the above corollary to every pair of vectors in the lattice  $\mathcal{L} = (\frac{1}{d\sqrt{mn}}\mathbb{Z})^{mn}$  we need the following simple claim:

**Claim 4.6.** *There are at most  $(4\sqrt{2}d)^{mn}$  vectors  $x \in \mathcal{L}$  such that  $\|x\| \leq 1$ .*

**Proof.** Let  $T = \{x \in \mathcal{L}: \|x\| \leq 1\}$ , set  $r = (d + 1)/d$  and consider  $B_0(r)$ , the  $mn$ -dimensional ball centered at 0 with radius  $r$ . For each  $x \in T$ , define the set

$$Z_x = \{z: x_{ij} < z_{ij} < x_{ij} + 1/(d\sqrt{mn})\}.$$

Clearly, each  $z \in Z_x$  satisfies  $\|z\| \leq \|x\| + \sqrt{mn}/(d\sqrt{mn})^2 = (d + 1)/d$ , and so  $Z_x \subset B_0(r)$ . Furthermore, for any  $x \neq y \in \mathcal{L}$  we have  $Z_x \cap Z_y = \emptyset$ , and altogether, if we let  $\text{vol}(\cdot)$  denote the Lebesgue measure on  $\mathbb{R}^{mn}$  then

$$\sum_{x \in T} \text{vol}(Z_x) \leq \text{vol}(B_0(r)) \leq \frac{(\pi r^2)^{mn/2}}{[mn/2]!} \leq \left( \sqrt{2\pi e} \frac{d + 1}{d\sqrt{mn}} \right)^{mn},$$

with the last inequality following from the fact that  $k! \geq (k/e)^k$  for all  $k$ . Since for every  $x$  we have  $\text{vol}(Z_x) = (d\sqrt{mn})^{-mn}$ , we now deduce that

$$|T| \leq (\sqrt{2\pi e}(d + 1))^{mn}. \quad \square$$

Applying Corollary 4.5 to all  $x, y \in T$  with a choice of  $a = 250$  and then taking a union bound over all  $|T|^2$  possible pairs now completes the proof of the Theorem 4.1.

### 5. The second eigenvalue of a random lift

**Proof of Theorem 1.** First, note that since  $\lambda$  is an upper bound on all nontrivial eigenvalues in absolute value, we can always increase it and take  $\lambda \geq \sqrt{d}$ , and this would not effect the result (recall that the bound we target for is  $(\lambda \vee \rho) \log \rho$  where  $\rho = 2\sqrt{d-1}$ ). In this case, as  $\rho = 2\sqrt{d-1}$ , it suffices to show that every nontrivial eigenvalue of  $G$  is  $O(\lambda \log d)$  except with probability  $O(n^{-100})$ .

The following lemma establishes the expected value of  $x^T A_H y$  for any two unit vectors orthogonal to the all-ones vector.

**Lemma 5.1.** *Let  $G$  be an  $(m, d, \lambda)$ -graph and  $H$  be a random  $n$ -lift of  $G$ . Let  $x, y \in \mathbb{R}^{mn}$  satisfy  $\langle x, \mathbf{1} \rangle = \langle y, \mathbf{1} \rangle = 0$  and  $\|x\| = \|y\| = 1$ . Then  $\mathbb{E}[x^T A_H y] \leq \lambda$ .*

**Proof.** Clearly, since the  $n$ -lift is comprised of a uniform perfect matching between any two fibers that are adjacent in  $G$ , we have

$$\begin{aligned} \mathbb{E}x^T A_H y &= \mathbb{E} \sum_{ij \sim i'j'} x_{ij} y_{i'j'} = \frac{1}{n} \sum_{j \sim j'} \sum_{i, i'} x_{ij} y_{i'j'} \\ &= \frac{1}{n} \sum_{j \sim j'} \sum_i x_{ij} \sum_{i'} y_{i'j'} = \frac{1}{n} w^T A_G z, \end{aligned}$$

where  $w, z \in \mathbb{R}^m$  are defined by  $w_j = \sum_i x_{ij}$  and  $z_{j'} = \sum_{i'} y_{i'j'}$  for  $j \in [m]$ . The assumptions on  $x, y$  give that

$$\sum_j w_j = \langle x, \mathbf{1} \rangle = 0, \quad \sum_{j'} z_{j'} = \langle y, \mathbf{1} \rangle = 0,$$

and furthermore, by Cauchy–Schwarz,

$$\|w\|^2 = \sum_j \left( \sum_i x_{ij} \right)^2 \leq \sum_j n \sum_i x_{ij}^2 = n \|x\|^2 = n,$$

and similarly  $\|z\|^2 \leq n$ . Altogether, as  $w, z$  are orthogonal to the trivial eigenvector, and since every nontrivial eigenvalue of  $A_G$  is at most  $\lambda$  in absolute value,

$$\mathbb{E}x^T A_H y = \frac{1}{n} w^T A_G z \leq \frac{1}{n} \lambda \|w\| \|z\| \leq \lambda,$$

as required.  $\square$

To prove Theorem 1, assume the events described in Theorem 3.1 and Theorem 4.1 occur. Let  $x, y \in \mathbb{R}^{mn}$  satisfy  $\langle x, \mathbf{1} \rangle = \langle y, \mathbf{1} \rangle = 0$  and  $\|x\| \leq 1, \|y\| \leq 1$ . Consider  $\tilde{x}, \tilde{y}$ , the closet vectors to  $x, y$  respectively among all vectors in  $\{z \in \mathcal{L} : \|z\| \leq 1\}$ , where  $\mathcal{L} = (\frac{1}{d\sqrt{mn}} \mathbb{Z})^{mn}$ . Theorem 3.1 and Theorem 4.1 now give that



$$|R_h(\tilde{x}, \tilde{y}) - \mathbb{E}[R_h(\tilde{x}, \tilde{y})]| \leq 7000\lambda \log d,$$

$$|R_l(\tilde{x}, \tilde{y}) - \mathbb{E}[R_l(\tilde{x}, \tilde{y})]| \leq 250\lambda \log d,$$

and since  $\tilde{x}^T A_H \tilde{y} = R_l(\tilde{x}, \tilde{y}) + R_h(\tilde{x}, \tilde{y})$  by definition, we get

$$\begin{aligned} |\tilde{x}^T A_H \tilde{y}| &\leq |\tilde{x}^T A_H \tilde{y} - \mathbb{E}[\tilde{x}^T A_H \tilde{y}]| + |\mathbb{E}[\tilde{x}^T A_H \tilde{y}]| \\ &\leq |R_l(\tilde{x}, \tilde{y}) - \mathbb{E}[R_l(\tilde{x}, \tilde{y})]| + |R_h(\tilde{x}, \tilde{y}) - \mathbb{E}[R_h(\tilde{x}, \tilde{y})]| + |\mathbb{E}[\tilde{x}^T A_H \tilde{y}]| \\ &\leq \lambda + 7250\lambda \log d. \end{aligned}$$

Finally, by the definition of the lattice  $\mathcal{L}$ , both  $x' = x - \tilde{x}$  and  $y' = y - \tilde{y}$  satisfy

$$\|x'\|_\infty \leq \frac{1}{d\sqrt{mn}}, \quad \|y'\|_\infty \leq \frac{1}{d\sqrt{mn}},$$

and so  $\|x'\| \leq 1/d$  and  $\|y'\| \leq 1/d$ . Therefore, for instance,

$$|x'^T A_H \tilde{y}| \leq \|x'\| \|A_H \tilde{y}\| \leq \frac{1}{d} \cdot d = 1,$$

and similarly  $|\tilde{x}^T A_H y'| \leq 1$  and  $|x'^T A_H y'| \leq 1/d$ . Combining these inequalities, it now follows that

$$\begin{aligned} |x^T A_H y| &\leq |\tilde{x}^T A_H \tilde{y}| + |x'^T A_H \tilde{y}| + |\tilde{x}^T A_H y'| + |x'^T A_H y'| \\ &\leq |\tilde{x}^T A_H \tilde{y}| + 3 < 7500\lambda \log d, \end{aligned}$$

completing the proof.

## References

- [1] N. Alon, Eigenvalues and expanders, *Combinatorica* 6 (2) (1986) 83–96.
- [2] N. Alon, V.D. Milman,  $\lambda_1$ , isoperimetric inequalities for graphs, and superconcentrators, *J. Combin. Theory Ser. B* 38 (1) (1985) 73–88.
- [3] N. Alon, J.H. Spencer, *The Probabilistic Method*, 3rd edition, John Wiley & Sons Inc., Hoboken, NJ, 2008.
- [4] A. Amit, N. Linial, Random graph coverings. I. General theory and graph connectivity, *Combinatorica* 22 (1) (2002) 1–18.
- [5] A. Amit, N. Linial, Random lifts of graphs: edge expansion, *Combin. Probab. Comput.* 15 (3) (2006) 317–332.
- [6] Y. Bilu, N. Linial, Lifts, discrepancy and nearly optimal spectral gap, *Combinatorica* 26 (5) (2006) 495–519.
- [7] A. Broder, E. Shamir, On the second eigenvalue of random regular graphs, in: *Proc. of the 28th Annual Symposium on the Foundations of Computer Science (FOCS 1987)*, pp. 286–294.
- [8] A. Broder, A.M. Frieze, S. Suen, E. Upfal, Optimal construction of edge-disjoint paths in random graphs, *SIAM J. Comput.* 28 (2) (1999) 541–573 (electronic).
- [9] J. Dodziuk, Difference equations, isoperimetric inequality and transience of certain random walks, *Trans. Amer. Math. Soc.* 284 (2) (1984) 787–794.
- [10] D.A. Freedman, On tail probabilities for martingales, *Ann. Probab.* 3 (1975) 100–118.
- [11] J. Friedman, Some geometric aspects of graphs and their eigenfunctions, *Duke Math. J.* 69 (3) (1993) 487–525.
- [12] J. Friedman, Relative expanders or weakly relatively Ramanujan graphs, *Duke Math. J.* 118 (1) (2003) 19–35.
- [13] J. Friedman, A proof of Alon’s second eigenvalue conjecture and related problem, *Mem. Amer. Math. Soc.* 195 (910) (2008).

- [14] J. Friedman, J. Kahn, E. Szemerédi, On the second eigenvalue in random regular graphs, in: Proc. of the 21st Annual ACM Symposium on Theory of Computing (STOC 1989), pp. 587–598.
- [15] Y. Greenberg, On the spectrum of graphs and their universal coverings, PhD thesis, Hebrew University of Jerusalem, 1995 (in Hebrew).
- [16] S. Hoory, N. Linial, A. Wigderson, Expander graphs and their applications, *Bull. Amer. Math. Soc.* 43 (4) (2006) 439–561.
- [17] N. Linial, D. Puder, Word maps and spectra of random graph lifts, *Random Structures Algorithms* 37 (1) (2010) 100–135.
- [18] E. Lubetzky, A. Sly, Cutoff phenomena for random walks on random regular graphs, *Duke Math. J.* 153 (3) (2010) 475–510.
- [19] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, *Progr. Math.*, vol. 125, Birkhäuser Verlag, Basel, 1994, with an appendix by Jonathan D. Rogawski.
- [20] A. Lubotzky, R. Phillips, P. Sarnak, Ramanujan graphs, *Combinatorica* 8 (3) (1988) 261–277.
- [21] G.A. Margulis, Explicit constructions of expanders, *Problemy Peredachi Informatsii* 9 (4) (1973) 71–80 (in Russian).
- [22] G.A. Margulis, Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators, *Probl. Inf. Transm.* 24 (1) (1988) 39–46.
- [23] C. McDiarmid, Concentration, in: *Probabilistic Methods for Algorithmic Discrete Mathematics*, in: *Algorithms Combin.*, vol. 16, Springer, Berlin, 1998, pp. 195–248.
- [24] S.J. Miller, T. Novikoff, A. Sabelli, The distribution of the largest nontrivial eigenvalues in families of random regular graphs, *Experiment. Math.* 17 (2) (2008) 231–244.
- [25] A. Nilli, On the second eigenvalue of a graph, *Discrete Math.* 91 (2) (1991) 207–210.
- [26] M. Pinsker, On the complexity of a concentrator, in: Proc. of the 7th International Teletraffic Conference, 1973, pp. 318/1–318/4.
- [27] O. Reingold, S. Vadhan, A. Wigderson, Entropy waves, the zig-zag graph product, and new constant-degree expanders, *Ann. of Math.* 155 (1) (2002) 157–187.
- [28] P. Sarnak, What is... an expander?, *Notices Amer. Math. Soc.* 51 (7) (2004) 762–763.
- [29] A. Sinclair, M. Jerrum, Approximate counting, uniform generation and rapidly mixing Markov chains, *Inform. and Comput.* 82 (1) (1989) 93–133.
- [30] N.C. Wormald, Models of random regular graphs, in: *Surveys in Combinatorics*, Canterbury, 1999, in: *London Math. Soc. Lecture Note Ser.*, vol. 267, Cambridge Univ. Press, Cambridge, 1999, pp. 239–298.