The threshold for SDP-refutation of random regular NAE-3SAT

Yash Deshpande[*] Andrea Montanari[†] Ryan O'Donnell[‡] Tselil Schramm[§] Subhabrata Sen[¶]

Abstract

Unlike its cousin 3SAT, the NAE-3SAT (not-all-equal-3SAT) problem has the property that spectral/SDP algorithms can efficiently refute random instances when the constraint density is a large constant (with high probability). But do these methods work immediately above the "satisfiability threshold", or is there still a range of constraint densities for which random NAE-3SAT instances are unsatisfiable but hard to refute?

We show that the latter situation prevails, at least in the context of random regular instances and SDP-based refutation. More precisely, whereas a random d-regular instance of NAE-3SAT is easily shown to be unsatisfiable (whp) once $d \geq 8$, we establish the following sharp threshold result regarding efficient refutation: If $d < 13.5$ then the basic SDP, even augmented with triangle inequalities, fails to refute satisfiability (whp); if $d > 13.5$ then even the most basic spectral algorithm refutes satisfiability (whp).

1 Introduction

A randomly chosen n-variable constraint satisfaction problem (CSP) will typically be unsatisfiable once the constraint density α (ratio of constraints to variables) is a sufficiently large constant. Taking 3SAT as an example, the conjectural satisfiability threshold [MPZ02, MMZ06] is $\alpha_c \approx 4.2667$, and the trivial first moment method already establishes unsatisfiability (whp) once $\alpha > \log_{7/8}(1/2) \approx 5.19$. Despite this, there is no known efficient algorithm that can refute random 3SAT instances (whp) for any large constant α. The best known algorithms [FGK05, GL03, CGL07, FO07, FKO06], all of which use spectral or semidefinite-programming (SDP) techniques, work only once $\alpha \gtrapprox \sqrt{n}$. Indeed, there are lower bounds [Sch10, Tul09, KMOW17] showing that any polynomial-time algorithm based on such techniques — more generally, based on the constant-degree "Sum of Squares" method — will fail to refute unless $\alpha \gtrapprox \sqrt{n}$. The most general of these results [KMOW17] applies to *any* CSP for which the constraint predicate supports a pairwise-uniform probability distribution.[1]

On the other hand, for any CSP whose predicate does *not* support a pairwise-uniform probability distribution, it has been shown [AOW15] that there *is* an efficient SDP-based algorithm for refuting random instances once the constraint density α is a sufficiently large constant.[2] For such CSPs, where "all of the action" is in the sparse regime of $O(n)$ constraints, it is more plausible to hope for an efficient refutation algorithm that works just above the satisfiability threshold — or at least to identify sharp thresholds for when efficient refutation algorithms succeed.

Perhaps the simplest and most natural NP-complete CSP of this type is NAE-3SAT. This is the variant of 3SAT in which a clause is considered "satisfied" if and only if it has at least one true literal *and* one false literal; i.e., the literals' truth values are Not All Equal. (The further variant wherein all literals appear positively is equivalent to the problem of 2-coloring a 3-uniform hypergraph.) Being a more symmetric — and in some sense, simpler — variant of 3SAT, the NAE-3SAT problem has received a great deal of attention in the study of random CSPs; see, e.g., [AS93, ACIM01, AM02, GJ03,

*Department of Mathematics, Massachusetts Institute of Technology. yash@mit.edu

†Department of Electrical Engineering and Department of Statistics, Stanford University. montanari@stanford.edu.

Supported by grants NSF CCF-1714305 and NSF IIS-1741162.

‡Computer Science Department, Carnegie Mellon University. odonnell@cs.cmu.edu.

Supported by NSF grants CCF-1618679, CCF-1717606. This material is based upon work supported by the National Science Foundation under grant numbers listed above. Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation (NSF).

§Harvard University and MIT. tselil@seas.harvard.edu.

This work was partly supported by an NSF graduate research fellowship (1106400), and took place in part while T.S. was a fellow at the Simons Institute program on Optimization.

¶Microsoft Research-NE and Department of Mathematics, Massachusetts Institute of Technology. ssen90@mit.edu

[1] That is, there is a distribution D over satisfying assignments x to the predicate, with the property that the order 1 and 2 moments of D are identical to those of the uniform distribution.

[2] In [AOW15], it is stated that $\alpha = n^{k/2-1}\text{polylog}n$ suffices when no k-wise uniform distribution is supported; however, in the particular case of $k = 2$ one can show that the polylog$n$ is unnecessary, using the (worst-case) strong refutation algorithm for 2XOR-SAT [CW04].

CNRZ03, DRZ08, DKR15, DSS16]. In particular, by 2003 Goerdt and Jurdziński [GJ03] had already proven that SDP methods could refute random NAE-3SAT instances at sufficiently high constant constraint density. NAE-3SAT is also closely related to the Max-Cut and 2XOR-SAT CSPs and has a natural basic SDP relaxation; for this reason, the problem has also been well-studied from the point of view of worst-case approximation algorithms [KLP96, AE98, Zwi98, Zwi99].

This paper is motivated by the question of whether efficient algorithms might be able to refute unsatisfiability of random NAE-3SAT instances at densities all the way down to the satisfiability threshold — or whether there is still a range of constant densities where random instances are unsatisfiable, but this is hard for efficient algorithms to certify. The latter case seems to prevail for 3SAT, and one would likely pessimistically guess the same is true for NAE-3SAT. However one may need a finer analysis for NAE-3SAT; the range of presumably-hard densities for refuting 3SAT is between a constant and $\sqrt{n}$, whereas for NAE-3SAT it is between two universal constants.

One way to give evidence for the existence of hard densities for NAE-3SAT refutation would be to study the *SDP-satisfiability threshold* for random instances; i.e., the largest density for which the basic SDP algorithm fails to refute satisfiability. The goal would be to give a lower-bound for the SDP-satisfiability threshold that exceeds the actual NAE-3SAT satisfiability threshold. In fact, the main result of this paper is a determination of the *exact* SDP-satisfiability threshold of random NAE-3SAT instances, in the setting of random regular instances. This threshold provably exceeds the actual satisfiability threshold, thus establishing a range of degrees for which random regular NAE-3SAT refutation is hard for SDP algorithms.

1.1 Our results For technical simplicity, we work in the setting of random *regular* instances of NAE-3SAT, where every variable participates in the same number, d, of 3NAE-constraints. (This is in contrast to the "Erdős–Rényi" setting with clause density α, in which the degree of each variable is like a Poisson random variable with mean 3α.) We also use the "random lift" model for d-regular instances, rather than, say, the "configuration" model. For precise details see Section 3.3, but in brief, our random d-regular instances are chosen as follows:

i Start with the bipartite graph $K_{d,3}$.

ii Choose a uniformly random n-lift $\boldsymbol{H}$, a bipartite graph with dn vertices of degree 3 in one part and $3n$ vertices of degree d in the other part.

iii Treat the degree-d vertices as CSP variables and the degree-3 vertices as 3NAE constraints on the adjacent variables

iv In each constraint, randomly replace each variable-appearance with its negation, uniformly and independently.

Notice that for *any* $(3, d)$-biregular graph H and any truth assignment to the variables, the randomness from the negations alone gives us that each constraint is independently satisfied with probability $3/4$. Thus the first moment method implies the following:

FACT 1.1. *For $d > \log_{\frac{4}{3}} 8 \approx 7.228$ (i.e., for $d \geq 8$) a random d-regular NAE-3SAT instance will be unsatisfiable with high probability (indeed, in any model with random negations).*[3]

Our main theorem is the following sharp threshold for SDP-satisfiability:

THEOREM 1.1. *Let $\boldsymbol{I}$ be a random d-regular instance of NAE-3SAT. Then with high probability (meaning probability $1 - o_{n\to\infty}(1)$):*

- *For $d < 13.5$, the natural SDP relaxation will* not *refute satisfiability of $\boldsymbol{I}$.*
- *For $d > 13.5$, the natural SDP relaxation* will *refute satisfiability of $\boldsymbol{I}$.*

Of course, since d is always an integer we could have phrased the two cases as $d \leq 13$ and $d \geq 14$. However, as will be seen below, there is a sense in which the precise non-integer 13.5 is the sharp threshold. In any case, these results show that for $d = 8, 9, 10, 11, 12, 13$ (and likely also $d = 7$), a random d-regular NAE-3SAT instance is unsatisfiable, yet this cannot be efficiently refuted using the basic SDP relaxation.

In fact, our results are somewhat stronger than what is stated in Theorem 1.1. Let us define

$$f(d) = \frac{9}{8} - \frac{3}{8} \cdot \frac{\left(\sqrt{d-1} - \sqrt{2}\right)^2}{d},$$

a quantity that decreases on $[3, \infty)$, with $f(13.5) = 1$ and $\lim_{d\to\infty} f(d) = 3/4$. We show:

[3] In fact, the unsatisfiability threshold is more likely to be lower, specifically $d \geq 7$, based on heuristics from statistical physics. The "1RSB" prediction for the unsatisfiability threshold of random NAE-3SAT — which was rigorously verified for NAE-kSAT, $k \geq k_0$, in [DSS16] — was determined to be at average degree $3 \cdot 2.105 = 6.315$ in the Erdős–Rényi case [CNRZ03], and at degree at most 7 in the regular case [DRZ08] (albeit these predictions were for the "coloring" version of NAE-3SAT without negations).

- (See Theorem 5.4 and Theorem 5.5 for details.) Even when augmented with the triangle inequalities, the SDP "thinks" that a random d-regular NAE-3SAT instance has a solution satisfying at least an $f(d) - \epsilon$ fraction of the constraints; in particular, it thinks the instance is satisfiable if $d < 13.5$. Indeed this holds for *any* d-regular NAE-3SAT instance of sufficiently large constant girth.
- (See Theorem 4.8 for details.) Even the basic "eigenvalue bound" (a special case of the SDP method) shows that a random d-regular NAE-3SAT instance has no solution satisfying at least an $f(d)+\epsilon$ fraction of the constraints; in particular, it refutes satisfiability if $d > 13.5$.

2 Strategy and related work

2.1 2XOR-SAT and semidefinite programming

One reason that semidefinite programming algorithms are particularly natural for NAE-3SAT is that the CSP is essentially a form of 2XOR-SAT. Recall that the 2XOR-SAT CSP has constraints on pairs of literals, with the constraint being satisfied if the literals are assigned unequal truth values. Now for literals ℓ_1, ℓ_2, ℓ_3:

$$\begin{aligned} &\mathrm{NAE}(\ell_1,\ell_2,\ell_3) \quad \text{satisfied} \iff \text{exactly 2 of} \\ &\mathrm{XOR}(\ell_1,\ell_2),\ \mathrm{XOR}(\ell_2,\ell_3),\ \mathrm{XOR}(\ell_3,\ell_1) \text{ satisfied;} \\ &\mathrm{NAE}(\ell_1,\ell_2,\ell_3) \text{ unsatisfied} \iff \text{exactly 0 of} \\ &\mathrm{XOR}(\ell_1,\ell_2),\ \mathrm{XOR}(\ell_2,\ell_3),\ \mathrm{XOR}(\ell_3,\ell_1) \text{ satisfied.} \end{aligned}$$

(In case all the literals are variables appearing positively, the resulting 2XOR-SAT instance is in fact a "Max-Cut" instance.) If we convert an NAE-3SAT CSP with m constraints to a 2XOR-SAT CSP with $3m$ constraints in the above way, every truth assignment satisfying a β fraction of NAE-3SAT constraints satisfies a $(2/3)\beta$ fraction of 2XOR-SAT constraints.

Indeed, the standard SDP relaxation for NAE-3SAT, first studied by Kann, Lagergren, and Panconesi [KLP96], is nothing more than 3/2 times the basic Goemans–Williamson [GW95] SDP for the associated 2XOR-SAT instance. We recall here the basic definitions:

Definition 2.1. *Let I be an instance of 2XOR-SAT with m constraints on n variables, to be assigned values in $\{\pm 1\}$. We identify the instance with its (multi)set of constraints. Each constraint is a triple (u, v, ξ) for $u, v \in [n]$ distinct and $\xi \in \{\pm 1\}$; this is thought of as the constraint $x_u x_v = -\xi$. The SDP relaxation value is defined to be*

$$\mathrm{SDP}(I) = \sup \left\{ \frac{1}{m} \sum_{(u,v,\xi)\in I} \left(\frac{1}{2} - \frac{1}{2}\xi\langle X_u, X_v\rangle \right) \right\} \in [0,1],$$

where the sup *is over all choices of vectors $(X_v)_{v\in[n]}$ satisfying $\langle X_v, X_v\rangle = 1$ for all v. Equivalently, instead of vectors, the X_v's may be jointly (centered) Gaussian random variables, with $\langle X_u, X_v\rangle$ interpreted as $\mathbf{E}[X_u X_v]$. The quantity* $\mathrm{SDP}(I)$ *always upper-bounds* $\mathrm{OPT}(I)$, *the maximum fraction of simultaneously satisfiable 2XOR-SAT constraints, since for any truth assignment $x \in \{\pm 1\}^n$ we may take the joint Gaussians $X_u = x_u Z$, where Z is a standard Gaussian. The advantage of* $\mathrm{SDP}(I)$ *is that while computing* $\mathrm{OPT}(I)$ *is* NP*-hard, one can compute* $\mathrm{SDP}(I)$ *(to additive accuracy 2^{-n}) in polynomial time.*

Definition 2.2. *A common algorithmic technique is to also enforce the* triangle inequalities, *meaning to only take the* sup *over X_v's satisfying*

$$\begin{aligned} \langle X_u, X_v\rangle + \langle X_v, X_w\rangle + \langle X_w, X_u\rangle \geq -1, \\ \langle X_u, X_v\rangle - \langle X_v, X_w\rangle - \langle X_w, X_u\rangle \geq -1. \end{aligned}$$

The resulting value, $\mathrm{SDP}_\triangle(I)$, *is a tighter relaxation:* $\mathrm{OPT}(I) \leq \mathrm{SDP}_\triangle(I) \leq \mathrm{SDP}(I)$.

Definition 2.3. *A related quantity is the* Lovász theta function *[Lov79]; for a graph G, the Lovász theta function (of its complement), $\vartheta(\overline{G})$, is the least k such that there are centered joint Gaussians (X_u) with $\langle X_u, X_u\rangle = 1$ for all vertices u and $\langle X_u, X_v\rangle = -\frac{1}{k-1}$ for all edges (u, v). In particular, if G is thought of as a Max-Cut instance, then* $\mathrm{SDP}(G) \geq \frac{1}{2} + \frac{1}{2}\frac{1}{\vartheta(\overline{G})-1}$.

Definition 2.4. *The SDP for 2XOR-SAT is also known to have a* dual *characterization [DP93]:*

$$\mathrm{SDP}(I) = \inf_{\substack{w\in\mathbb{R}^n \\ \sum_u w_u = 0}} \left\{ \frac{n}{4m} \cdot \lambda_{max}(L_I + \mathrm{diag}(w)) \right\},$$

where L_I denotes the Laplacian *matrix for I (defined in Section 3.2), and λ_{max} denotes the largest eigenvalue. Note that by taking $w = 0$ we get an upper bound on* $\mathrm{SDP}(I)$; *we refer to this as the* eigenvalue bound,

$$\mathrm{EIG}(I) = \frac{n}{4m} \cdot \lambda_{max}(L_I) = \frac{1}{2d} \cdot \lambda_{max}(L_I),$$

the latter equality holding in case I is d-regular. The certificate $\mathrm{OPT}(I) \leq \mathrm{EIG}(I)$ *is easy to see; it is a consequence of the definitions that* $\mathrm{OPT}(I) = \frac{n}{4m} \cdot \max\{x^\top L_I x : x \in \{\pm\frac{1}{\sqrt{n}}\}^n\}$, *and $\lambda_{max}(L_I)$ allows taking the max over all unit vectors.*

2.2 Methodology To prove Theorem 1.1, we convert our random NAE3-SAT instances into random 2XOR-SAT instances, and then try to analyze whether or not the SDP-value of these instances is as large as $\frac{2}{3}$.

(Recall that this corresponds to the SDP-value of the NAE3-SAT instances being as large as 1.) There are a number of prior works on analyzing the Goemans–Williamson SDP on random graphs (see below); however, our situation is a bit different. The main difference is that the graphs underlying our random 2XOR-SAT instances are not uniformly random $2d$-regular graphs, but rather have a peculiar "triangle-structure". Recall that they are generated by first choosing a large random $(3, d)$-biregular graph (by randomly lifting $K_{d,3}$), then replacing each 3-regular vertex on the left with a triangle on the right. Thus, locally, the resulting graphs look like the graph on the right in Figure 2 (for $d = 4$). An additional small complication is that these random "triangle-graphs" effectively get random edge-signings when the random literal-negations are taken into account, converting the Max-Cut instance to a 2XOR-SAT instance. Finally, in the remainder of the paper we will focus on the generalized problem in which triangles are replaced by c-cliques, for $c \geq 3$. This generalization does not correspond to any well-known CSP, but analyzing general c turns out to be no harder than analyzing the $c = 3$ special case.

For the part of our main theorem showing that the simple eigenvalue bound succeeds as d becomes large, we need to show tight bounds on the eigenvalues of the random "triangle-graphs" (more generally, c-clique graphs) that arise in our model. If we simply had random d-regular graphs, Friedman's famous almost-Ramanujan theorem [Fri08] would have sufficed. Instead, we relate the eigenvalues of our random graphs to those of a randomly lifted (c, d)-biregular bipartite graph. We then use Bordenave's recent reproof [Bor17] of Friedman's theorem (revised to also include random edge-signings), as well as the Ihara–Bass formula, to show that with high probability the nontrivial spectrum of such random bipartite graphs is contained in $\pm[\sqrt{d-1} - \sqrt{c-1}, \sqrt{d-1} + \sqrt{c-1}]$. Inspiration for these computations comes from [FM16].

For the part of our main theorem showing that large-value SDP solutions exist, the tools we use come from a fairly recent line of work concerning "Gaussian waves" in infinite regular graphs [Elo09, CGHV15, HV15]. This work can be seen as giving a way to convert eigenfunctions on the infinite regular tree (and other vertex-transitive infinite graphs) into Goemans–Williamson SDP solutions — in fact, Lovász theta function solutions. These may be converted to such solutions on high-girth finite graphs that locally resemble the infinite graphs. Several works in this area [CGHV15, HV15, Csó16, Lyo17] used this method to show, e.g., that high-girth 3-regular graphs must contain large independent sets, using techniques resembling the randomized rounding of independent-set SDPs (cf. [KMS98]) and also local improvement techniques applicable to cubic graphs (cf. [HLZ04]). These techniques were also used to show limits on the performance of SDP for Max-Cut, Min-Bisection, and community detection problems in, e.g., [MS16, FM16]. See [BKM17] for similar approaches in the context of graph-coloring, and [JMR16] for more on phase transitions for SDPs in the context of community detection.

3 Preliminaries: graphs, lifts, and eigenvalues

3.1 Graphs, hypergraphs, and edge-labeled graphs We begin with some general notation.

H will typically denote a simple (c, d)-biregular bipartite graph with $c, d \geq 2$. The setting of most interest to us is $d \geq c = 3$. Sometimes we will refer to the vertices on the c-regular side as *constraints* and the vertices on the d-regular side as *variables*. Figure 1 shows an example, $K_{4,3}$, with the variables depicted as circles and the constraints depicted as squares.

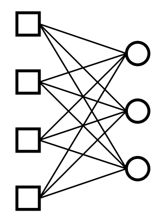

Figure 1: $H = K_{4,3}$

We may also think of H as a c-uniform d-regular hypergraph, with the variables as vertices and constraints as hyperedges. X will denote an edge-signed version of H (thought of as a bipartite graph, not a hypergraph); i.e., one in which each edge of H is labeled with ± 1. (In the unsigned case, we think of all edges as being labeled $+1$.) We say that X is a "random signing" of H if it is formed by independently labeling each edge of H with ± 1, uniformly at random.

Given H, we will write $G = G_H$ for the (loopless multi-)graph formed by first thinking of H as a hypergraph and then replacing each hyperedge by a c-clique. As a result, G is a $(c-1)d$-regular graph, called the *primal graph* for H. Given an edge-signed version X of H, we will write $I = I_X$ for the primal graph of X, an edge-signed version of G defined as follows: whenever constraint a is adjacent to variables i, j with edge-signs $\xi_{ai}, \xi_{aj} \in \{\pm 1\}$, we place the sign $\xi_{ai}\xi_{aj}$ on the resulting $\{i, j\}$ edge of G. We may think of I as a 2XOR-SAT instance, where the vertices are to be assigned values $x_i \in \{\pm 1\}$, and an edge $\{i, j\}$ with label ξ corresponds to the constraint $x_i x_j = -\xi$.

In the special case of $c = 3$, we can think of X as a NAE-3SAT instance, where the variables are to be assigned values $x_i \in \{\pm 1\}$, and a constraint a adjacent to variables i, j, k with labels $\xi_{ai}, \xi_{aj}, \xi_{ak}$ corresponds to the constraint that $\xi_{ai}x_i, \xi_{aj}x_j, \xi_{ak}x_k$ are not all equal. In this case there is a precise relationship between the NAE-3SAT instance X and the 2XOR-SAT instance I; any assignment to the vertices satisfying exactly a β fraction of the NAE-3SAT constraints will necessarily satisfy exactly a $\frac{2}{3}\beta$ fraction of the 2XOR-SAT constraints.

3.2 Associated matrices Given any of $Y \in \{H, X, G, I\}$, we will write A_Y for the adjacency matrix. More precisely, $A_Y[i, j]$ is the sum of the (positive and negative) edge-labels on all edges connecting i and j.

We will write D_Y for the diagonal degree matrix of Y, whose entry $D_Y[i, i]$ equals the degree of vertex i. (Both signed and unsigned edges count 1 toward the degree.) We write $L_Y = D_Y - A_Y$ for the Laplacian matrix of Y; we also write $L_Y(u) = (1-u^2)\mathbb{1} + u^2 D_Y - uA_Y$ for the "deformed Laplacian", parameterized by $u \in \mathbb{R}$, which reduces to the basic Laplacian when $u = 1$. (Here $\mathbb{1}$ denotes the identity operator.)

Finally, we will write B_Y for the non-backtracking matrix of Y. Recall that this matrix is formed as follows: First, each undirected edge in Y is converted to two directed edges (both having the same sign, in case Y is edge-signed). Then B_Y is the square (non-symmetric) matrix indexed by the directed edges, in which $B_Y[(i, j), (k, \ell)]$ entry is nonzero if and only if $j = k$ and $i \neq \ell$, in which case it equals the sign-label of (i, j).

3.3 Lifts Suppose now that $Y = (V, E)$ denotes any undirected (multi-)graph. For $n \in \mathbb{Z}^+$, an n-lift of Y is a graph Y_n whose vertex set is $V \times [n]$ and whose edges consist of a perfect matching between $\{u\} \times [n]$ and $\{v\} \times [n]$ for each edge $\{u, v\} \in E$. When the $|E|$ perfect matchings are chosen independently and uniformly at random, we call Y_n a random n-lift of Y. Note that if Y is a d-regular graph, then so is Y_n, and if Y is a (c, d)-biregular bipartite graph, then so is Y_n. If B (respectively, B_n) denotes the non-backtracking matrix of Y (respectively, Y_n), it is known that the multiset of B_n's eigenvalues contains the multiset of Y's eigenvalues. The remaining eigenvalues are referred to as the "new" eigenvalues of B_n.

3.4 Eigenvalues Given an N-dimensional matrix M, we write $\mathrm{spec}(M) \subset \mathbb{C}$ for its spectrum, the cardinality-N *multiset* of roots of its characteristic polynomial. We also write $\rho(M)$ for its spectral radius, $\max\{|\lambda| : \lambda \in \mathrm{spec}(M)\}$. The adjacency matrix of a (possibly edge-signed) graph is symmetric, and hence its spectrum is real; the Laplacian is furthermore positive semidefinite, and hence its spectrum is nonnegative. A non-backtracking matrix, however, will in general have complex spectrum.

We are particularly interested in bipartite graphs, so we record some facts concerning them here. Suppose X is a possibly edge-signed bipartite graph, with vertex parts of size $m \geq n$. Then it is well known that

$$\begin{aligned} \mathrm{spec}(A_X) = &\{0 : \text{with multiplicity } m - n\} \\ &\cup \{\pm\lambda : \lambda \in \mathrm{PS}(A_X)\} \end{aligned}$$

for some multiset $\mathrm{PS}(A_X) \subset \mathbb{R}^{\geq 0}$.[4] Further, if X is (c, d)-biregular, we'll have $\mathrm{PS}(A) \subset [0, \sqrt{cd}]$. The set $\pm\mathrm{PS}(A_X)$ may be called the "nontrivial" part of A_X's spectrum. A warning, though: $\pm\mathrm{PS}(A_X)$ is not the same as the "nonzero" part of A_X's spectrum, since $\mathrm{PS}(A_X)$ may contain 0 with positive multiplicity. Indeed, this happens in one of the simplest cases, as is well known:

FACT 3.1. *Let $H = K_{d,c}$, the complete bipartite graph with vertex parts of size $d \geq c$. Then $\mathrm{PS}(A_H)$ consists of $c - 1$ copies of 0 and 1 copy of $\sqrt{cd}$.*

We also record below the spectrum of the non-backtracking matrix of $K_{d,c}$, which we'll derive in Section 4.1 using the Ihara–Bass formula. But first, some notation we'll use heavily in this paper:

NOTATION 3.1. *For $c, d \geq 2$, we write*

$$s_c = \sqrt{c-1}, \quad s_d = \sqrt{d-1}, \quad \rho_1 = s_c s_d, \quad \overline{\lambda} = s_d + s_c,$$
$$\underline{\lambda} = |s_d - s_c|, \quad \kappa = (c-1)d = \rho_1^2 + s_c^2.$$

We will often assume $d \geq c$, in which case $\underline{\lambda} = s_d - s_c$.

PROPOSITION 3.1. *Let B be the non-backtracking matrix of $K_{d,c}$, where $d \geq c \geq 2$, $d \neq 2$. Let i be the fourth primitive root of unity. Then*

$$\mathrm{spec}(B) = \begin{cases} \pm 1 & \text{with multiplicity } (c-1)(d-1) \text{ each;} \\ \pm i s_c & \text{with multiplicity } (d-1) \text{ each;} \\ \pm i s_d & \text{with multiplicity } (c-1) \text{ each;} \\ \pm s_c s_d & \text{with multiplicity } 1 \text{ each;} \end{cases}$$

and hence, $\rho(B) = s_c s_d = \rho_1$.

As described in Section 3.1, we will often consider forming the primal graph G of a (c, d)-biregular graph H. It is simple to work out the relationship between the eigenvalues of H and the eigenvalues of G; this is done in, e.g., [LS96, Section 4.1]. The analysis is unchanged for the edge-signed variant, and it yields:

[4] We chose "PS" to stand for Positive Spectrum, notwithstanding our warning that it may contain 0.

PROPOSITION 3.2. *Let X be an edge-signed (c,d)-biregular graph, and let $I = I_X$ be the corresponding edge-signed primal graph. Then*

$$\mathrm{spec}(A_I) = \{\lambda^2 - d : \lambda \in \mathrm{PS}(A_X)\}.$$

Since I is κ-regular, where $\kappa = cd - d$, we can also conclude that

$$\mathrm{spec}(L_I) = \{cd - \lambda^2 : \lambda \in \mathrm{PS}(A_X)\}.$$

3.5 The infinite biregular tree and distance-regular graph Since a large random (c,d)-biregular graph looks locally like a tree, we will want to study the infinite (c,d)-biregular tree, which we denoted by $\mathbb{T}_{d,c}$. More to the point, we will want to study its (infinite) primal graph, which we denote by $\mathbb{G}_{d,c}$. Fragments of these graphs, in the case $c = 3$, $d = 4$, are pictured in Figure 2.

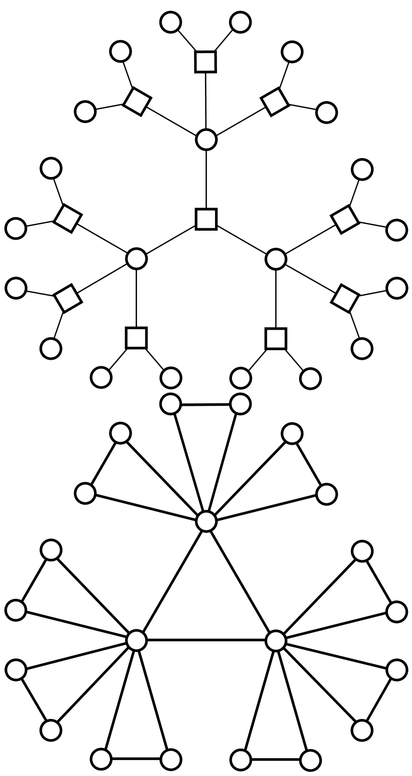

Figure 2: Fragments of the infinite biregular tree $\mathbb{T}_{4,3}$, and its primal graph $\mathbb{G}_{4,3}$

As shown by Ivanov [Iva83], the graphs $\mathbb{G}_{d,c}$ are precisely the infinite graphs G that are *distance-regular*, meaning that there exist constants $p^h_{j,k}$ such that for every pair $u, v \in V(G)$ with $\mathrm{dist}_G(u,v) = h$, the number of vertices $w \in V(G)$ having $\mathrm{dist}_G(w,u) = j$ and $\mathrm{dist}_G(w,v) = k$ is equal to $p^h_{j,k}$. It is elementary to compute these quantities for $\mathbb{G}_{d,c}$, and the results appears below. Only the cases $h = 0, 1$ are truly essential for the paper, and the reader might like to verify them while referring to Figure 2.

PROPOSITION 3.3. *In the distance-regular graph $\mathbb{G}_{d,c}$, recalling the notation*

$$s_c^2 - 1 = c - 2, \quad \rho_1^2 = (c-1)(d-1),$$
$$\rho_1^2 + s_c^2 = \kappa = (c-1)d, \quad \rho_1^2 - s_c^2 = (c-1)(d-2),$$

we have

$$p^0_{\ell,\ell} = \begin{cases} 1 & \text{if } \ell = 0 \\ (\rho_1^2 + s_c^2)\rho_1^{2(\ell-1)} & \text{if } \ell \geq 1; \end{cases}$$

and, for $h \geq 1$, $0 \leq t \leq h$,

if h and t have the same parity:

$$p^h_{\ell,\ell+t} = p^h_{\ell+t,\ell} = \begin{cases} 0 & \text{if } \ell < \frac{h-t}{2} \\ 1 & \text{if } \ell = \frac{h-t}{2} \\ \rho_1^{2\ell} & \text{if } \ell > \frac{h-t}{2} \text{ and } t = h \\ (\rho_1^2 - s_c^2)\rho_1^{2(\ell - (\frac{h-t+2}{2}))} & \text{if } \ell > \frac{h-t}{2} \text{ and } t \neq h; \end{cases}$$

if h and t have opposite parity:

$$p^h_{\ell,\ell+t} = p^h_{\ell+t,\ell} = \begin{cases} 0 & \text{if } \ell < \frac{h-t+1}{2} \\ (s_c^2 - 1)\rho_1^{2(\ell - \frac{h-t+1}{2})} & \text{if } \ell \geq \frac{h-t+1}{2}; \end{cases}$$

and finally, $p^h_{j,k} = 0$ otherwise.

The spectrum of the adjacency "matrix" (operator) of $\mathbb{G}_{d,c}$ — and indeed, the whole "spectral measure" — has been known since the early '80s. (There are appropriate definitions for these terms, generalizing the definitions in the finitary case. We will not give them here since, strictly speaking, this paper does not rely on them.) In particular,

(3.1)
$$\mathrm{spec}(A_{\mathbb{T}_{d,c}}) = \{0\} \cup \pm[\underline{\lambda}, \overline{\lambda}], \ \mathrm{spec}(A_{\mathbb{G}_{d,c}}) = [\underline{\lambda}^2 - d, \overline{\lambda}^2 - d];$$

(the latter holding under the assumption $d \geq c$; if $d < c$ then also $-d \in \mathrm{spec}(A_{\mathbb{G}_{d,c}})$). The history of these results can be found in [MW89, Section 7E] and [GM88, Section 5.2], the latter of which also shows that the spectral measures of large random (c,d)-biregular graphs converge to a measure with support $\mathrm{spec}(A_{\mathbb{T}_{d,c}})$ (and similarly for their primal graphs and $\mathrm{spec}(A_{\mathbb{G}_{d,c}})$).

4 Eigenvalues of random lifts and signings

Generalizing Friedman's celebrated characterization of the spectrum of random d-regular random graphs [Fri08], Bordenave recently proved the following theorem:

THEOREM 4.1. *([Bor17, Theorem 20].) Let Y be a connected multigraph (with more edges than vertices) having non-backtracking matrix B. Fix $\epsilon > 0$. Let $\boldsymbol{Y}_n$ be a random n-lift of Y, and let $\boldsymbol{B}_n$ be its non-backtracking matrix. Then*

$$\mathbf{Pr}[\boldsymbol{B}_n \text{ has a new eigenvalue of magnitude} \geq \sqrt{\rho(B)} + \epsilon] = o_{n\to\infty}(1).$$

We will need a variant of this theorem in which the graph is randomly lifted and then randomly signed. The statement and proof are actually a little bit simpler.

THEOREM 4.2. *Let Y be a connected graph (with more edges than vertices) having non-backtracking matrix B. Fix $\epsilon > 0$. Let $\boldsymbol{X}_n$ be a random signing of a random n-lift $\boldsymbol{Y}_n$ of Y, and let $\boldsymbol{B}_n$ denote the non-backtracking matrix of $\boldsymbol{X}_n$. Then*

$$\mathbf{Pr}[\rho(\boldsymbol{B}_n) \geq \sqrt{\rho(B)} + \epsilon] = o_{n\to\infty}(1).$$

The proof follows that of [Bor17, Theorem 20].

We will also quote some basic results about the scarcity of cycles in randomly lifted graphs:

THEOREM 4.3. *(Greenhill–Janson–Ruciński [GJR10, Lemma 5.1].) Let $\boldsymbol{Y}_n$ be as in Theorem 4.1 or Theorem 4.2, and write $\boldsymbol{Z}_k$ for the number of length-k cycles in $\boldsymbol{Y}_n$. Let $\boldsymbol{P}_2, \boldsymbol{P}_3, \dots$ be independent Poisson random variables with $\boldsymbol{P}_k$ of mean $w_k/(2k)$, where $w_k = \operatorname{tr}(B^k)$ is the number of closed non-backtracking walks in Y. Then for any $g \in \mathbb{N}^+$, the random variables $(\boldsymbol{Z}_2, \boldsymbol{Z}_3, \dots, \boldsymbol{Z}_g)$ converge jointly in distribution to $(\boldsymbol{P}_2, \boldsymbol{P}_3, \dots, \boldsymbol{P}_g)$. In particular, for a fixed g and n sufficiently large, there is a positive probability (depending only on g and Y) that $\boldsymbol{Y}_n$ has girth exceeding g.*

THEOREM 4.4. *(Easily extracted from the proof of [Bor17, Lemma 24].) Let $\boldsymbol{Y}_n$ be as in Theorem 4.1 or Theorem 4.2 and write d for the maximum degree of Y. Call a vertex of $\boldsymbol{Y}_n$ g-bad if its distance-g neighborhood contains a cycle. Then the expected number of g-bad vertices in $\boldsymbol{Y}_n$ is $O((d+1)^g)$.*

4.1 The Ihara–Bass formula

The Ihara–Bass formula relates the eigenvalues of a graph's adjacency matrix and its non-backtracking matrix. Originally proved by Ihara [Iha66] for regular graphs, it was subsequently generalized to irregular graphs [Has92, Bas92, ST96, KS00], vertex-weighted graphs [Kem16], and most generally, edge-weighted graphs [WF09, FM16]. We will need the last of these, but only in the special case that all edge-weights are ± 1. In this case, the resulting formula looks identical to the usual (irregular, unweighted) Ihara–Bass formula:

THEOREM 4.5. *([WF09, Theorem 2], specialized to all edge-weights ± 1.) Let X be a edge-signed graph, having adjacency matrix A, non-backtracking matrix B, and deformed Laplacian $L(u) = (1-u^2)\mathbb{1} + u^2 D - uA$. Then for all real $u \neq \pm 1$,*

$$\det(\mathbb{1} - uB) = \det(L(u)) \cdot (1-u^2)^{\#E(X) - \#V(X)}.$$

In the special case when X is (c,d)-biregular, one can use this formula to work out a very explicit mapping between the eigenvalues of A and the eigenvalues of B. The computations appear in [Kem16, Section 4.2]; that paper only considered unsigned edges, but the result is the same because the Ihara–Bass formula is identical. Recalling the notation from Section 3.4:

THEOREM 4.6. *(Follows from [Kem16, Theorem 6] using Theorem 4.5.) Let X be an edge-signed (c,d)-biregular graph, with m vertices on the c-regular side and n vertices on the d-regular side, so $e = cm = dn$ is the number of edges. Let A denote the adjacency matrix of X. Then B, the non-backtracking matrix of X, has the following $2e$ eigenvalues:*

- *$e - (m+n)$ copies each of ± 1.*
- *$m - n$ copies each of $\pm i s_c$.*
- *$4n$ "nontrivial" eigenvalues, all roots of $p_\lambda(u) = u^4 + (s_c^2 + s_d^2 - \lambda^2)u^2 + \rho_1^2$ for $\lambda \in \operatorname{PS}(A)$.*

We would now like to understand the location of the 4 roots of $p_\lambda(u)$ in $\mathbb{C}$ as λ varies in $[0, \sqrt{cd}]$. To do this, write

$$s_c = \frac{\overline{\lambda} - \underline{\lambda}}{2}, \quad s_d = \frac{\overline{\lambda} + \underline{\lambda}}{2}, \quad \alpha = \frac{\lambda^2 - \underline{\lambda}^2}{2}, \quad \beta = \frac{\lambda^2 - \overline{\lambda}^2}{2}, \quad U = u^2.$$

Then

$$p_\lambda(u) = U^2 - (\alpha + \beta)U + \left(\frac{\alpha - \beta}{2}\right)^2,$$

which has roots

$$U = \frac{1}{2}\left(\sqrt{\alpha} \pm \sqrt{\beta}\right)^2.$$

If $\underline{\lambda}^2 \leq \lambda^2 \leq \overline{\lambda}^2$ then $\beta \leq 0 \leq \alpha$ and

$$|U| = \frac{1}{2}\left(\sqrt{\alpha}^2 + \sqrt{-\beta}^2\right) = \frac{\alpha - \beta}{2} = \frac{\overline{\lambda}^2 - \underline{\lambda}^2}{4} = s_c s_d = \rho_1.$$

On the other hand, if $\lambda^2 \notin \left[\underline{\lambda}^2, \overline{\lambda}^2\right]$, then α and β have the same sign and

$$|U| = \frac{1}{2}(|\alpha| + |\beta| \pm 2|\alpha| \cdot |\beta|),$$

$$\text{the larger of which exceeds } \frac{\overline{\lambda}^2 - \underline{\lambda}^2}{4} = \rho_1.$$

We conclude:

PROPOSITION 4.1. *For real λ, the roots of $p_\lambda(u)$ simultaneously have magnitude at most $\sqrt{\rho_1}$ if and only if $\lambda^2 \in \left[\underline{\lambda}^2, \overline{\lambda}^2\right]$ (i.e., $\lambda \in \pm[\underline{\lambda}, \overline{\lambda}]$).*

Also, when $\lambda = 0$ we have $p_\lambda(u) = u^4 + (s_c^2 + s_d^2)u^2 + s_c^2 s_d^2$, and when $\lambda = \sqrt{cd}$ we have $p_\lambda(u) = u^4 - (\rho_1^2+1)u^2 + \rho_1^2$. Thus we can directly verify:

PROPOSITION 4.2. *For $\lambda = 0$, the 4 roots of $p_\lambda(u)$ are $\pm i s_c$, $\pm i s_d$. And, for $\lambda = \sqrt{cd}$, the 4 roots of $p_\lambda(u)$ are $\pm\rho_1$, ± 1.*

At this point, we can combine Theorem 4.6, Fact 3.1, and Proposition 4.2 to obtain Proposition 3.1 as stated in Section 3.4. We may furthermore put together all the results in this section:

THEOREM 4.7. *Let $d \geq c \geq 2$, $d \neq 2$. Fix $\epsilon > 0$. Let $\boldsymbol{X}_n$ be a random signing of a random n-lift of the complete bipartite graph $K_{d,c}$, and let $\boldsymbol{A}_n$ denote its adjacency matrix. Then*

$$\mathbf{Pr}\big[\mathrm{PS}(\boldsymbol{A}_n) \not\subseteq [\underline{\lambda}-\epsilon, \overline{\lambda}+\epsilon]\big] = o_{n\to\infty}(1).$$

Proof. We apply Theorem 4.2 with $Y = K_{d,c}$ and some sufficiently small $\epsilon' = \epsilon'(\epsilon, c, d) > 0$. The non-backtracking matrix B of Y has spectral radius ρ_1, by Proposition 3.1. Thus if $\boldsymbol{B}_n$ is the non-backtracking matrix of the randomly signed random lift $\boldsymbol{X}_n$ of Y, we get

$$\mathbf{Pr}\big[\rho(\boldsymbol{B}_n) \geq \sqrt{\rho_1} + \epsilon'\big] = o_{n\to\infty}(1),$$

Thus with probability $1-o(1)$ we have $\rho(\boldsymbol{B}_n) < \sqrt{\rho_1}+\epsilon'$. In this case, taking ϵ' sufficiently small and using the fact that the roots of a polynomial are continuous in its coefficients, Proposition 4.1 and Theorem 4.6 imply that $\mathrm{PS}(\boldsymbol{A}_n) \subset [\underline{\lambda}-\epsilon, \overline{\lambda}+\epsilon]$. The proof is complete.

REMARK 4.1. *This theorem is "to be expected" in light of the Godsil–Mohar work on spectral convergence mentioned at the end of Section 3.5. But of course one needs the hard work of Bordenave's Theorem to show that random (c,d)-biregular graphs typically do not any eigenvalues outside the spectral bulk. In fact, to emphasize that care is needed, we remark that the random signing in Theorem 4.7 is essential; without it, it's not hard to show that $\mathrm{PS}(\boldsymbol{A}_n)$ will contain 0 with probability 1.*

COROLLARY 4.1. *Let $d \geq c \geq 2$, $d \neq 2$. Fix $\epsilon > 0$. Let $\boldsymbol{X}_n$ be a random signing of a random n-lift of the complete bipartite graph $K_{d,c}$, let $\boldsymbol{I}_n$ be the associated 2XOR-SAT instance (as in Section 3.1), and let $\boldsymbol{L}_n$ be its Laplacian matrix. Then*

$$\mathbf{Pr}\big[\boldsymbol{L}_n \text{ has an eigenvalue outside } [(1-\rho_1)^2-\epsilon, (1+\rho_1)^2+\epsilon]\big] = o_{n\to\infty}(1).$$

Proof. This follows from Proposition 3.2, $cd - \overline{\lambda}^2 = (1-\rho_1)^2$, and $cd - \underline{\lambda}^2 = (1+\rho_1)^2$.

Corollary 4.1 now directly implies the following:

THEOREM 4.8. *Let $d \geq c \geq 2$, $d \neq 2$. Fix $\epsilon > 0$. Let $\boldsymbol{I}_n$ be a random 2XOR-SAT instance as in Corollary 4.1, so $\boldsymbol{I}_n$ is κ-regular ($\kappa = (c-1)d$) with cn variables and $\binom{c}{2}dn$ constraints. Then*

$$\mathbf{Pr}\left[\mathrm{EIG}(\boldsymbol{I}_n) \geq \frac{(1+\rho_1)^2}{2\kappa} + \epsilon\right] = o_{n\to\infty}(1),$$

where $\rho_1 = \sqrt{c-1}\sqrt{d-1}$.

In case $c = 3$, if we view $\boldsymbol{I}_n$ as a random d-regular NAE-3SAT instance on $3n$ variables (chosen according to the random lift/sign model), we have

$$\mathbf{Pr}\left[\mathrm{EIG}(\boldsymbol{I}_n) \geq \frac{9}{8} - \frac{3}{8}\cdot\frac{\left(\sqrt{d-1}-\sqrt{2}\right)^2}{d} + \epsilon\right] = o_{n\to\infty}(1),$$

As mentioned in Section 1.1, the quantity $\frac{9}{8} - \frac{3}{8}\cdot\frac{(\sqrt{d-1}-\sqrt{2})^2}{d}$ decreases from $\frac{9}{8}$ to $\frac{3}{4}$ on $[3,\infty)$ and takes value 1 at $d = 13.5$. Thus the above theorem shows that the basic eigenvalue bound refutes a random d-regular instance of NAE-3SAT (whp) provided $d > 13.5$.

5 SDP solutions for random instances

As a guide for our construction, let us imagine SDP solutions for the Max-Cut problem on the infinite graph $\mathbb{G}_{d,c}$. (As these imaginings are only for intuition's sake, we will not be completely formal.) To lower bound $\mathrm{SDP}(\mathbb{G}_{d,c})$, it is necessary and sufficient to construct jointly standard Gaussian random variables $(\boldsymbol{X}_v)_{v\in V(\mathbb{G}_{d,c})}$ for which the correlation $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v]$ — "on average", over all edges $\{u,v\} \in E(\mathbb{G}_{d,c})$ — is very negative. It's simpler, and stronger, to look for such a Gaussian process in which $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \varrho$ for *every* edge $\{u,v\}$, with ϱ as negative as possible. Such solutions would give an upper bound for the Lovász theta value, $\vartheta(\overline{\mathbb{G}_{d,c}}) \leq 1 - 1/\varrho$, while still giving an SDP lower bound of $\mathrm{SDP}(\mathbb{G}_{d,c}) \geq \frac{1}{2} - \frac{1}{2}\varrho$. In turn, we would have such a Gaussian process provided it satisfied

$$\frac{1}{\kappa}\sum_{u\sim v}\boldsymbol{X}_u = \varrho\boldsymbol{X}_v \quad \text{for all } v \in V(\mathbb{G}_{d,c}), \tag{5.2}$$

where, as before, $\kappa = (c-1)d$ is the degree of each v. This is the "eigenvalue equation" for $A_{\mathbb{G}_{d,c}}$ for $\lambda = \kappa\varrho$. Thus one may suspect that Equation (5.2) is possible whenever $\lambda = \kappa\varrho \in \mathrm{spec}(A_{\mathbb{G}_{d,c}})$. Given $\mathrm{spec}(A_{\mathbb{G}_{d,c}})$ as in Equation (3.1), we may therefore hope to obtain the desired Gaussian process for any

$$\varrho \in \left[\frac{\underline{\lambda}^2-d}{\kappa}, \frac{\overline{\lambda}^2-d}{\kappa}\right] = \left[1-\frac{(1+\rho_1)^2}{\kappa}, 1-\frac{(1-\rho_1)^2}{\kappa}\right]; \tag{5.3}$$

in particular, for the most negative such value,

$$\varrho^* = 1 - \frac{(1+\rho_1)^2}{\kappa}. \tag{5.4}$$

This would lead to the lower bound

$$\mathrm{SDP}(\mathbb{G}_{d,c}) \geq \frac{1}{2} - \frac{1}{2}\varrho^* = \frac{(1+\rho_1)^2}{2\kappa}.$$

In fact, since $\mathbb{G}_{d,c}$ is a vertex-transitive graph, it follows from a theorem of Harangi and Virág that such Gaussian processes do exist, and they can be constructed in a simple fashion as "linear block factors of IIDs":

THEOREM 5.1. *([HV15, Theorem 4].) Let G be an infinite vertex-transitive graph with adjacency operator A_G. Then for each $\lambda \in \mathrm{spec}(A_G)$, there is an $\mathrm{Aut}(G)$-invariant standard Gaussian process $(\boldsymbol{X}_v)_{v\in V(G)}$ for which $\sum_{u\sim v} \boldsymbol{X}_u = \lambda \boldsymbol{X}_v$ holds for all $v \in V(G)$. Furthermore, the process can be approximated (in distribution) by a "linear block factor of IID process", meaning one that is constructed as follows: $(\boldsymbol{Z}_v)_{v\in V(G)}$ are chosen as IID standard Gaussians, and then $\boldsymbol{X}_v$ is set to be a fixed linear function f of those $\boldsymbol{Z}_u$'s which have $\mathrm{dist}_G(u,v) \leq L$, where L is a finite "radius".*

As mentioned in Section 2.2, results of this nature date back at least to the work of Elon [Elo09], who constructed such "Gaussian waves" on the infinite d-regular tree $\mathbb{T}_d$. An important aspect of Theorem 5.1 is the "block" aspect, meaning that each $\boldsymbol{X}_v$ is defined just from a "local", finite number of $\boldsymbol{Z}_u$'s. Thus we can hope to use the construction for (primal graphs of) large but finite (c,d)-biregular graphs with large girth, which locally look tree-like.

That said, we cannot quite use the Theorem 5.1 as a black box for our purposes, for a few reasons. One reason is that we want to apply it to large random biregular graphs, which will not strictly speaking have low girth, but will merely have "few", "far apart" short cycles. Second, we will be constructing SDP solutions for *edge-signed* graphs, a slight generalization of Theorem 5.1's framework. Finally, it will be nice for us to reason about $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v]$ not just for adjacent u, v.

On the other hand, the construction of the linear block factor of IID process for $\mathbb{G}_{d,c}$ is a fairly straightforward generalization of earlier concrete constructions for $\mathbb{T}_d$ such as the one in [CGHV15]. We present it in the next section.

5.1 Linear factors of IIDs Here we essentially prove Theorem 5.1 in the special case of $\mathbb{G}_{d,c}$. The proof closely follows [CGHV15, Section 3].

THEOREM 5.2. *Let $c, d \geq 2$ and let $\lambda \in \mathrm{spec}(A_{\mathbb{G}_{d,c}})^\circ = (\underline{\lambda}^2 - d, \overline{\lambda}^2 - d)$. Then there exist $L \in \mathbb{N}$ and reals $a_0, a_1, \dots, a_L$ such that the following holds: When $(\boldsymbol{Z}_v)_{v\in V(\mathbb{G}_{d,c})}$ are IID standard Gaussians, and the random variables $(\boldsymbol{X}_v)_{v\in V(\mathbb{G}_{d,c})}$ are formed via*

$$\boldsymbol{X}_v = \sum_{\ell=0}^{L} \sum_{\substack{w\in V(\mathbb{G}_{d,c})\\ \mathrm{dist}(w,v)=\ell}} a_\ell \boldsymbol{Z}_w, \tag{5.5}$$

then we have $\mathbf{E}[\boldsymbol{X}_v^2] = 1$ for all v (so that the $\boldsymbol{X}_v$'s are jointly standard Gaussians), and $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \frac{\lambda}{\kappa}$ for all $\{u,v\} \in E(\mathbb{G}_{d,c})$. In other words (cf. Equation (5.3)):

$$\text{for any} \quad 1 - \frac{(1+\rho_1)^2}{\kappa} < \varrho < 1 - \frac{(1-\rho_1)^2}{\kappa} \tag{5.6}$$
$$\text{we can achieve } \mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \varrho \quad \forall \{u,v\} \in E(\mathbb{G}_{d,c}).$$

Proof. Let us temporarily relax the requirement that L be finite. To that end, we will consider defining

$$\boldsymbol{X}_v = \gamma \cdot \sum_{\ell=0}^{\infty} \sum_{\substack{w\in V(\mathbb{G}_{d,c})\\ \mathrm{dist}(w,v)=\ell}} r^\ell \boldsymbol{Z}_w, \tag{5.7}$$

for constants $\gamma \in \mathbb{R}^+$, $r \in \mathbb{R}$. It follows that for two vertices $u, v \in V(\mathbb{G}_{d,c})$ with $\mathrm{dist}(u,v) = h$, we have

$$\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \gamma^2 \cdot \sum_{j,k=0}^{\infty} p^h_{j,k} r^{j+k}. \tag{5.8}$$

In this proof we focus only on $h = 0, 1$, saving $h > 1$ for Theorem 5.3. By Proposition 3.3 we have

$$\#\{w : \mathrm{dist}(w,v) = \ell\} = p^0_{\ell,\ell} = \begin{cases} 1 & \text{if } \ell = 0, \\ (\rho_1^2 + s_c^2)\cdot \rho_1^{2(\ell-1)} & \text{if } \ell > 0, \end{cases}$$

where recall $\rho_1^2 + s_c^2 = (c-1)d$ and $\rho_1^2 = (c-1)(d-1)$. Thus

(5.9)

$$\mathbf{E}[\boldsymbol{X}_v^2] = \mathbf{Var}[\boldsymbol{X}_v] =$$

(5.10)

$$\gamma^2 \cdot \left(1 + \sum_{\ell=1}^{\infty} (\rho_1^2 + s_c^2)\cdot \rho_1^{2(\ell-1)} \cdot r^{2\ell}\right) = \gamma^2 \cdot \frac{1+(s_c r)^2}{1-(\rho_1 r)^2},$$

(5.11)

provided $|r| < \rho_1^{-1}$.

By choosing γ such that

$$\gamma^2 = \frac{1-(\rho_1 r)^2}{1+(s_c r)^2}$$

we get $\mathbf{Var}[\boldsymbol{X}_v] = 1$. On the other hand, for fixed u, v with $\mathrm{dist}(u, v) = 1$ we have

$$\#\{w : \mathrm{dist}(u,w) = \ell_1, \mathrm{dist}(v,w) = \ell_2\} = p^1_{\ell_1,\ell_2}$$
$$= \begin{cases} (s_c^2 - 1)\cdot \rho_1^{2(\ell-1)} & \text{if } \ell_1 = \ell_2 > 0, \\ \rho_1^{2\ell_1} & \text{if } \ell_2 = \ell_1 + 1, \\ \rho_1^{2\ell_2} & \text{if } \ell_1 = \ell_2 + 1, \\ 0 & \text{else,} \end{cases}$$

where recall $s_c^2 - 1 = c - 2$. Thus

(5.12)
$$\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v] =$$
(5.13)
$$\gamma^2 \cdot \left(\sum_{\ell=1}^{\infty} (s_c^2 - 1)\cdot \rho_1^{2(\ell-1)} \cdot r^{2\ell} + \sum_{\ell=0}^{\infty} 2 \cdot \rho_1^{2\ell} \cdot r^{2\ell+1} \right)$$
(5.14)
$$= \gamma^2 \cdot \frac{1 + (s_c r)^2 - (1-r)^2}{1 - (\rho_1 r)^2},$$

and so by our choice of γ we conclude

$$\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v] = 1 - \frac{(1-r)^2}{1 + (s_c r)^2}.$$

Calculus shows that the expression on the right is increasing for r in the range $[-s_c^{-2}, 1]$, which is a superset of the range that Equation (5.9) allows us for r, namely $(-\rho_1^{-1}, \rho_1^{-1})$. This establishes Equation (5.6); the only catch is that we haven't used a finite L. But this can be achieved by truncating the sum in Equation (5.7) to $\ell \le L$ for L sufficiently large. This truncation only changes Equations (5.9) and (5.12) by a quantity that decays like $(\rho_1 r)^L$. Thus the change in $\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v]$ from truncation can be made arbitrarily small, and this is acceptable for the conclusion Equation (5.6) because the desired interval of ϱ's is open.

Corollary 5.1. *Theorem 5.2 also holds for the primal graph $\mathbb{I}$ of any edge-signed version $\mathbb{X}$ of $\mathbb{T}_{d,c}$ (as defined in Section 3.1), in the sense of having $\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v] = \xi_{uv}\varrho$ for all $\{u, v\} \in E(\mathbb{I})$, where ξ_{uv} denotes the sign of edge $\{u, v\}$.*

Proof. Assume we have signs $\xi_{av} \in \{\pm 1\}$ for each constaint/variable edge $\{a, v\}$ in $\mathbb{X}$, and therefore signs $\xi_{uv} = \xi_{au}\xi_{av}$ for each edge $\{u, v\}$ in $\mathbb{I}$. It's clear that for any closed walk in the tree $\mathbb{X}$, the product of the edge-signs along the walk is 1; by construction, it follows that the same is true in $\mathbb{I}$. Thus for any $u, v \in V(\mathbb{I})$ (not necessarily adjacent) we can unambiguously define $\xi[u \leftrightarrow v]$ as the product of edge-signs along any uv-path in $\mathbb{I}$. We now alter the construction in Equation (5.7) as follows:

$$\boldsymbol{X}_v = \gamma \cdot \sum_{\ell=0}^{\infty} \sum_{\substack{w \in V(\mathbb{G}_{d,c}) \\ \mathrm{dist}(w,v)=\ell}} \xi[w \leftrightarrow v] r^\ell \boldsymbol{Z}_w,$$

Clearly $\mathbf{Var}[\boldsymbol{X}_v]$ is unchanged. As for $\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v]$, the contribution from each $\boldsymbol{Z}_w$ now yields an additional factor of $\xi[w \leftrightarrow u]\xi[w \leftrightarrow v] = \xi[u \leftrightarrow v] = \xi_{uv}$. Thus each $\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v]$ changes by a factor of ξ_{uv}, as desired. The rest of the proof is the same.

Theorem 5.3. *In the $L = \infty$ setting of Theorem 5.2, we in fact obtain, for all $r \in (-\rho_1^{-1}, \rho_1^{-1})$ and all $u, v \in V(\mathbb{G}_{d,c})$,*

$$\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v] = r^h \left(1 + \frac{h(1-r)(1+s_c^2 r)}{1 + (s_c r)^2}\right),$$
$$\textit{where } h = \mathrm{dist}(u, v).$$

(The $r = 0$ case is of course trivial, with $\boldsymbol{X}_v = \boldsymbol{Z}_v$.)

Proof. Allowing L to be infinite and returning to Equation (5.8): for $u, v \in V(\mathbb{G}_{d,c})$ with $\mathrm{dist}(u, v) = h$, one can use Proposition 3.3 to show (calculations omitted) that

$$\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v] = \gamma^2 \cdot \frac{r^h(1 + (s_c r)^2 + h(1-r)(1 + s_c^2 r))}{1 - (\rho_1 r)^2}$$

provided $|r| < \rho_1^{-1}$. The result follows.

Remark 5.1. *One can show that the expression in Theorem 5.3 has the property that its absolute value is a strictly decreasing function of h for every $r \neq 0$. (Indeed, it decreases exponentially.) This is the key takeaway of the theorem, implying that in the setting of Corollary 5.1, $|\mathbf{E}[\boldsymbol{X}_u \boldsymbol{X}_v]| \le |\varrho|$ for* all *distinct pairs $u, v \in \mathbb{I}$ (with equality when $\{u, v\} \in E(\mathbb{G}_{d,c})$).*

5.2 SDP solutions for randomly lifted/signed graphs In this section, let us fix $d \ge c \ge 2$, a small $\epsilon > 0$,

$$\varrho = 1 - \frac{(1+\rho_1)^2}{\kappa} + \epsilon,$$

and an $L = L(\epsilon, c, d)$ such that Theorem 5.2 and Corollary 5.1 hold. Since each $\boldsymbol{X}_v$ constructed therein depends only on the $\boldsymbol{Z}_v$'s at distance at most L in $\mathbb{G}_{d,c}$ (and hence distance at most $2L$ in $\mathbb{T}_{d,c}$), we see that the exact same construction works equally well on any finite primal graph constructed from a (c, d)-biregular graph of girth exceeding $4L$. Thus (using also Remark 5.1) we immediately obtain:

THEOREM 5.4. *Let H be any edge-signed (c,d)-biregular graph of girth exceeding $4L$ and let I be its associated primal graph, with edge signs ξ_{uv}, $\{u,v\} \in E(I)$. Then one can assign joint standard Gaussians $\boldsymbol{X}_v$ to the vertices $v \in V(I)$ such that $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \xi_{uv}\varrho$ for each edge $\{u,v\} \in E(I)$. Furthermore, $|\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v]| \le |\varrho|$ for all distinct $u,v \in V(I)$. As consequences:*

(i) If H is unsigned, $\vartheta(\overline{I}) \le 1 - 1/\varrho$.

(ii) If we view I as a 2XOR-SAT instance, we have $\mathrm{SDP}_\triangle(I) \ge \frac{1}{2} - \frac{1}{2}\varrho = \frac{(1+\rho_1)^2}{2\kappa} - \epsilon$.

(iii) If $c = 3$ and we view I as a d-regular NAE-3SAT instance, we have $\mathrm{SDP}_\triangle(I) \ge \frac{1}{3} - \frac{1}{3}\varrho = \frac{9}{8} - \frac{3}{8}\cdot\frac{\left(\sqrt{d-1}-\sqrt{2}\right)^2}{d} - \epsilon$.

We have the following corollary:

THEOREM 5.5. *Let Y be a (c,d)-biregular bipartite graph and let $\boldsymbol{Y}_n$ be a random n-lift of Y. Let $\boldsymbol{H}_n$ denote an* arbitrary *edge-signing of $\boldsymbol{Y}_n$, and $\boldsymbol{I}_n$ its associated primal graph. Then:*

1. *With positive probability (depending only on d and ϵ), Items (i) to (iii) of Theorem 5.4 all hold.*
2. *With high probability, Items (ii) and (iii) of Theorem 5.4 hold with an additive loss of $O(1/n)$.*

Proof. The first statement is an immediate consequence of Theorem 4.3. As for the second statement, Theorem 4.4 and Markov's inequality imply that, with high probability, only an $O((d+1)^{2L+2})/n = O(1/n)$ fraction of vertices in $\boldsymbol{Y}_n$ are "$(2L+2)$-bad" (i.e., have a cycle within their distance-$(2L+2)$ neighborhood). Assuming this holds, we use the linear block factors of IID solution from Theorem 5.2 and Corollary 5.1 but with a small twist: For each vertex v that is $2L$-bad in $\boldsymbol{Y}_n$, rather than using Equation (5.5) we simply set $\boldsymbol{X}_v = \boldsymbol{Z}'_v$, where the random variables $\boldsymbol{Z}'_v$ are new standard Gaussians independent of all other random variables. Now for the $1 - O(1/n)$ fraction of "$(2L+2)$-good" vertices, all their neighbors are still $2L$-good and thus are using the linear block factors of IID solution. We therefore still have $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = \xi_{uv}\varrho$ for each edge $\{u,v\} \in E(I)$ where u or v is $(2L+2)$-good. Furthermore, we still have $|\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v]| \le |\varrho|$ for all distinct $u,v \in V(I)$, since $\mathbf{E}[\boldsymbol{X}_u\boldsymbol{X}_v] = 0$ when one of u or v is $2L$-bad. The second statement in the theorem therefore follows.

6 Conclusions

In this work we have shown a sharp threshold for the SDP-satisfiability of random d-regular NAE-3SAT instances in the model of random lifts. Some open questions that remain are the following:

- Can we show similar sharp threshold results in the configuration model? The main challenge is proving Friedman-style bounds on the spectra of random (c,d)-biregular bipartite graphs in this model. An advantage to doing this would be the potential to show similar sharp thresholds for 2-coloring random d-regular 3-uniform hypergraphs (i.e., random d-regular NAE-3SAT *without* negations).
- Can we show similar sharp threshold results in the Erdős–Rényi random model?
- Can our analysis of the 2XOR-SAT SDP / Lovász theta function for the infinite biregular tree $\mathbb{T}_{d,c}$, and its primal graph $\mathbb{G}_{d,c}$ be extended to other interesting classes of infinite graphs (say, vertex-transitive)? Are there application to other finite CSPs?
- A difficult but important open question: can we analyze the performance higher-degree "Sum of Squares" relaxations for refuting random sparse CSPs (that do not support pairwise-uniform distributions)? Even analyzing the degree-4 Sum of Squares relaxation for NAE-3SAT or graph 3-colorability seems very challenging.

Acknowledgments

This work began at the American Institute of Mathematics workshop "Phase transitions in randomized computational problems"; the authors would like to thank AIM, as well as the organizers Amir Dembo, Jian Ding, and Nike Sun, for the invitation. R. O. would like to thank Charles Bordenave, Sidhanth Mohanty, Doron Puder, Nike Sun, and David Witmer for helpful comments.

References

[ACIM01] Dimitris Achlioptas, Arthur Chtcherba, Gabriel Istrate, and Cristopher Moore. The phase transition in 1-in-k SAT and NAE 3-SAT. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 721–722, 2001. 1

[AE98] Gunnar Andersson and Lars Engebretsen. Better approximation algorithms for Set Splitting and Not-All-Equal Sat. *Information Processing Letters*, 65(6):305–311, 1998. 1

[AM02] Dimitris Achlioptas and Cristopher Moore. The asymptotic order of the random k-sat threshold. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 779–788, 2002. 1

[AOW15] Sarah Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science*, pages 689–708, 2015. 1, 2

[AS93] Noga Alon and Joel Spencer. A note on coloring random k-sets. Unpublished, 1993. 1

[Bas92] Hyman Bass. The Ihara–Selberg zeta function of a tree lattice. *International Journal of Mathematics*, 3(6):717–797, 1992. 4.1

[BKM17] Jess Banks, Robert Kleinberg, and Cristopher Moore. The Lovász theta function for random regular graphs and community detection in the hard regime. In *Proceedings of the 21st Annual International Conference on Randomization and Computation*, pages 28:1–28:22, 2017. 2.2

[Bor17] Charles Bordenave. A new proof of Friedman's second eigenvalue theorem and its extension to random lifts. Technical Report 1502.04482, arXiv, 2017. 2.2, 4.1, 4, 4.4, A, A.1, A.2, A.1, A.4, A.2, A.3

[CGHV15] Endre Csóka, Balázs Gerencsér, Viktor Harangi, and Bálint Virág. Invariant Gaussian processes and independent sets on regular graphs of large girth. *Random Structures & Algorithms*, 47(2):284–303, 2015. 2.2, 5, 5.1

[CGL07] Amin Coja-Oghlan, Andreas Goerdt, and André Lanka. Strong refutation heuristics for random k-SAT. *Combinatorics, Probability and Computing*, 16(1):5–28, 2007. 1

[CNRZ03] Tommaso Castellani, Vincenzo Napolano, Federico Ricci-Tersenghi, and Riccardo Zecchina. Bicolouring random hypergraphs. *Journal of Physics A: Mathematical and General*, 36(43):11037, 2003. 1, 3

[Csó16] Endre Csóka. Independent sets and cuts in large-girth regular graphs. Technical Report 1602.02747, arXiv, 2016. 2.2

[CW04] Moses Charikar and Anthony Wirth. Maximizing quadratic programs: Extending Grothendieck's inequality. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 54–60, 2004. 2

[DKR15] Adam Douglass, Andrew King, and Jack Raymond. Constructing SAT filters with a quantum annealer. In *Proceedings of the 18th Annual International Conference on Theory and Applications of Satisfiability Testing*, pages 104–120, 2015. 1

[DP93] Charles Delorme and Svatopluk Poljak. Laplacian eigenvalues and the maximum cut problem. *Mathematical Programming*, 62(3, Ser. A):557–574, 1993. 2.4

[DRZ08] Luca Dall'Asta, Abolfazl Ramezanpour, and Riccardo Zecchina. Entropy landscape and non-Gibbs solutions in constraint satisfaction problems. *Physical Review E*, 77(3):031118, 2008. 1, 3

[DSS16] Jian Ding, Allan Sly, and Nike Sun. Satisfiability threshold for random regular NAE-SAT. *Communications in Mathematical Physics*, 341(2):435–489, 2016. 1, 3

[Elo09] Yehonatan Elon. Gaussian waves on the regular tree. Technical Report 0907.5065, arXiv, 2009. 2.2, 5

[FGK05] Joel Friedman, Andreas Goerdt, and Michael Krivelevich. Recognizing more unsatisfiable random k-SAT instances efficiently. *SIAM Journal on Computing*, 35(2):408–430, 2005. 1

[FKO06] Uriel Feige, Jeong Han Kim, and Eran Ofek. Witnesses for non-satisfiability of dense random 3CNF formulas. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 497–508, 2006. 1

[FM16] Zhou Fan and Andrea Montanari. How well do local algorithms solve semidefinite programs? Technical Report 1610.05350, arXiv, 2016. 2.2, 4.1

[FO07] Uriel Feige and Eran Ofek. Easily refutable subformulas of large random 3CNF formulas. *Theory of Computing. An Open Access Journal*, 3:25–43, 2007. 1

[Fri08] Joel Friedman. A proof of Alon's second eigenvalue conjecture and related problems. *Memoirs of the American Mathematical Society*, 195(910), 2008. 2.2, 4

[GJ03] Andreas Goerdt and Tomasz Jurdziński. Some results on random unsatisfiable k-Sat instances and approximation algorithms applied to random structures. *Combinatorics, Probability and Computing*, 12(3):245–267, 2003. 1

[GJR10] Catherine Greenhill, Svante Janson, and Andrzej Ruciński. On the number of perfect matchings in random lifts. *Combinatorics, Probability and Computing*, 19(5-6):791–817, 2010. 4.3

[GL03] Andreas Goerdt and André Lanka. Recognizing more random unsatisfiable 3-SAT instances efficiently. *Electronic Notes in Discrete Mathematics*, 16:21–46, 2003. 1

[GM88] Chris Godsil and Bojan Mohar. Walk generating functions and spectral measures of infinite graphs. *Linear Algebra and its Applications*, 107:191–206, 1988. 3.5

[GW95] Michel Goemans and David Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the Association for Computing Machinery*, 42(6):1115–1145, 1995. 2.1

[Has92] Ki-ichiro Hashimoto. Artin type L-functions and the density theorem for prime cycles on finite graphs. *International Journal of Mathematics*, 3(6):809–826, 1992. 4.1

[HLZ04] Eran Halperin, Dror Livnat, and Uri Zwick. MAX CUT in cubic graphs. *Journal of Algorithms. Cognition, Informatics and Logic*, 53(2):169–185, 2004. 2.2

[HV15] Viktor Harangi and Bálint Virág. Independence ratio and random eigenvectors in transitive graphs. *The Annals of Probability*, 43(5):2810–2840, 2015. 2.2, 5.1

[Iha66] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over $\mathfrak{p}$-adic fields. *Journal of the Mathematical Society of Japan*, 18:219–235, 1966. 4.1

[Iva83] Alexander Ivanov. Bounding the diameter of a distance-regular graph. *Doklady Akademii Nauk SSSR*, 271(4):789–792, 1983. 3.5

[JMR16] Adel Javanmard, Andrea Montanari, and Federico Ricci-Tersenghi. Phase transitions in semidefinite relaxations. *Proceedings of the National Academy of Sciences of the United States of America*, 113(16), 2016. 2.2

[Kem16] Mark Kempton. Non-backtracking random walks and a weighted Ihara's theorem. Technical Report 1603.05553, arXiv, 2016. 4.1, 4.1, 4.6

[KLP96] Viggo Kann, Jens Lagergren, and Alessandro Panconesi. Approximability of maximum splitting of k-sets and some other APX-complete problems. *Information Processing Letters*, 58(3):105–110, 1996. 1, 2.1

[KMOW17] Pravesh Kothari, Ryuhei Mori, Ryan O'Donnell, and David Witmer. Sum of squares lower bounds for refuting any CSP. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing*, pages 132–145, 2017. 1

[KMS98] David Karger, Rajeev Motwani, and Madhu Sudan. Approximate graph coloring by semidefinite programming. *Journal of the ACM*, 45(2):246–265, 1998. 2.2

[KS00] Motoko Kotani and Toshikazu Sunada. Zeta functions of finite graphs. *The University of Tokyo. Journal of Mathematical Sciences*, 7(1):7–25, 2000. 4.1

[Lov79] László Lovász. On the Shannon capacity of a graph. *Institute of Electrical and Electronics Engineers. Transactions on Information Theory*, 25(1):1–7, 1979. 2.3

[LS96] Wen-Ch'ing Winnie Li and Patrick Solé. Spectra of regular graphs and hypergraphs and orthogonal polynomials. *European Journal of Combinatorics*, 17(5):461–477, 1996. 3.4

[Lyo17] Russell Lyons. Factors of IID on trees. *Combinatorics, Probability and Computing*, 26(2):285–300, 2017. 2.2

[MMZ06] Stephan Mertens, Marc Mézard, and Riccardo Zecchina. Threshold values of random k-sat from the cavity method. *Random Structures & Algorithms*, 28(3):340–373, 2006. 1

[MPZ02] Marc Mézard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297:812–815, 2002. 1

[MS16] Andrea Montanari and Subhabrata Sen. Semidefinite programs on sparse random graphs and their application to community detection. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing*, pages 814–827, 2016. 2.2

[MW89] Bojan Mohar and Wolfgang Woess. A survey on spectra of infinite graphs. *The Bulletin of the London Mathematical Society*, 21(3):209–234, 1989. 3.5

[Sch10] Grant Schoenebeck. *Limitations of Linear and Semidefinite Programs.* PhD thesis, University of California, Berkeley, 2010. 1

[ST96] Harold Stark and Audrey Terras. Zeta functions of finite graphs and coverings. *Advances in Mathematics*, 121(1):124–165, 1996. 4.1

[Tul09] Madhur Tulsiani. CSP gaps and reductions in the Lasserre hierarchy. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing*, pages 303–312, 2009. 1

[WF09] Yusuke Watanabe and Kenji Fukumizu. Graph zeta function in the Bethe free energy and loopy belief propagation. In *Proceedings of the 23rd annual Annual Conference and Workshop on Neural Information Processing Systems*, pages 2017–2025, 2009. 4.1, 4.5

[Zwi98] Uri Zwick. Approximation algorithms for constraint satisfaction problems involving at most three variables per constraint. In *Proceedings of the 9th Annual ACM-SIAM Symposium on Discrete Algorithms*, volume 98, pages 201–210, 1998. 1

[Zwi99] Uri Zwick. Outward rotations: a tool for rounding solutions of semidefinite programming relaxations, with applications to MAX CUT and other problems. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 679–687, 1999. 1

A Bordenave's Theorem for random signed lifts

In this appendix we will prove the following theorem.

THEOREM A.1. (RESTATEMENT OF THEOREM 4.2) *Let Y be a connected graph (with more edges than vertices) having non-backtracking matrix B. Fix $\epsilon > 0$. Let $\boldsymbol{X}_n$ be a random signing of a random n-lift $\boldsymbol{Y}_n$ of Y, and let $\boldsymbol{B}_n$ denote the non-backtracking matrix of $\boldsymbol{X}_n$. Then*

$$\mathbf{Pr}[\rho(\boldsymbol{B}_n) \geq \sqrt{\rho(B)} + \epsilon] = o_{n\to\infty}(1).$$

Our theorem requires minor modifications to the trace-method proof of [Bor17, Theorem 20], and we follow it closely. The differences occur because [Bor17, Theorem 20] pertains to the spectrum of unsigned lifts, and for that reason the arguments therein must take into account the uninteresting top eigenspace of the non-backtracking matrix; this introduces some technical complications. Since we are working with randomly signed edges, we need not worry about these eigenspaces, and our arguments will be somewhat pared down (though to our knowledge they cannot be extracted from [Bor17] in a black-box fashion).

A.1 Setup and notation We set the stage for the proof by introducing some notation and definitions. Let $Y = (V, E)$ be an undirected graph, and let $\vec{E}$ be the set of directed edges associated with E, so that

$$\vec{E} = \{(u, v) : \{u, v\} \in E\},$$

and $|\vec{E}| = 2|E|$. To limit confusion, we will use plain, bold letters $\boldsymbol{e}$ to denote edges in E and decorated bold letters $\vec{\boldsymbol{e}}$ to denote arcs in $\vec{E}$. For an arc $\vec{\boldsymbol{e}} = (u, v)$, we let $(\vec{\boldsymbol{e}})^{-1} = (v, u)$.

Let $n \in \mathbb{Z}^+$, let $Y_n = (V_n, E_n)$ be an n-lift of Y as defined in Section 3.3, and let $X_n = (V_n, E_n, \xi_n)$ be random signing of Y_n with signs $\xi_n : E_n \to \mathbb{R}$.[5] In the n-lift, each edge $e \in E_n$ (arc $\vec{e} \in \vec{E_n}$) is associated with an edge $\{u, v\} \in E$ (arc $(u, v) \in \vec{E}$), and with a pair of labels $i, j \in [n]$, so that $e = \{(u, i), (v, j)\}$ ($\vec{e} = ((u, i), (v, j))$. Again to limit confusion, we will use non-bold, plain letters to denote edges $e \in E_n$ and decorated, non-bold letters to denote arcs $\vec{e} \in \vec{E_n}$. We let S_n^E be the set of tuples of $|E|$ permutations on $[n]$. Each n-lift is associated with some $\sigma = \{\sigma_{\boldsymbol{e}}\}_{\boldsymbol{e}\in E} \in S_n^E$, so that $E_n = \{\{(u, i), (v, \sigma_{u,v}(i))\}\}$ (where we take u to proceed v lexicographically, in order to ensure that the bijection between σ and lifts is unique).[6] We sometimes refer to the lift specified by $\sigma \in S_n^E$ as $Y_n(\sigma)$.

We also define B_n to be the weighted non-backtracking matrix of X_n as in Section 3.2, so that for directed edges $(u, v), (x, y) \in \vec{E_n}$,

$$B_n[(u, v), (x, y)] = \xi_n(\{u, v\}) \cdot \mathbb{1}[v = x] \cdot \mathbb{1}[u \neq y].$$

We will apply the trace method to B_n; that is, we will relate $\rho(B_n)$ to the expected trace of a power of B_n.

Fact A.1. *If $A \in \mathbb{C}^{n\times n}$ is a random complex matrix, $m, \ell \in \mathbb{Z}^+$, $\epsilon, c \in \mathbb{R}^+$, and $\mathbf{E}[\mathrm{tr}((A^\ell(A^\ell)^*)^k)] \leq R^{2m\ell}$, then for $\ell \cdot m \geq \frac{c}{\epsilon} R \log n$ and $\epsilon < R/2$,*

$$\mathbf{Pr}[\rho(A) \geq R + \epsilon] \leq n^{-c}$$

Proof. This follows by noticing that $\rho(A)^\ell \leq \sup_{x\in\mathbb{R}^n} \frac{\|A^\ell x\|_2}{\|x\|_2} = \|A^\ell(A^\ell)^*\|^{1/2}$, and then applying Markov's inequality:

$$\mathbf{Pr}[\|A^\ell(A^\ell)^*\|^{1/2\ell} \geq t] \leq \frac{\mathbf{E}[\mathrm{tr}((A^\ell(A^\ell)^*)^m)]}{t^{2m\ell}} \leq \left(\frac{R}{t}\right)^{2m\ell},$$

and choosing $t = R + \epsilon$ with $2\epsilon < R$,

$$\left(\frac{1}{1 + \epsilon/R}\right)^{2m\ell} \leq \left(1 - \frac{\epsilon}{2R}\right)^{2m\ell} \leq \exp\left(-\frac{\epsilon m\ell}{R}\right)$$

for $\ell \cdot m \geq \frac{c}{\epsilon} R \log n$ the conclusion follows.

In our computations, we will bound the contribution of sequences of *half-edges* (so as to be consistent with [Bor17]).

Definition A.1. (half-edge) *A* half-edge γ *is given by an arc $(u, v) \in \vec{E}$, and an index $i \in [n]$ corresponding to the index of u. We think of $\gamma = ((u, v), i)$ as an arc leaving the ith copy of u in the lift, and going to vertex v at some unspecified index; colloquially, $\gamma = ((u, i), (v, ?))$.*

We call the set of all possible half-edges Π. In the interest of promoting clarity, we point out that Π does not depend on the specific choice of lift, σ.

Definition A.2. (valid sequence of half-edges) *We will say that a sequence of half-edges $(\gamma_1, \ldots, \gamma_{2k})$ is* valid *if it satisfies the following constraints:*

1. Admissibility of pairs*: consecutive pairs of half-edges correspond to the same edge in Y. Formally, for each $t \in [k]$ with $\gamma_{2t-1} = (\vec{\boldsymbol{e}}_{2t-1}, i_{2t-1})$ and $\gamma_{2t} = (\vec{\boldsymbol{e}}_{2t}, i_{2t})$, we have that $\vec{\boldsymbol{e}}_{2t-1} = (\vec{\boldsymbol{e}}_{2t})^{-1}$.*

2. Consistency*: if two half-edges are paired once, they remain paired for the remainder of the sequence. Formally, if there exists t^* such that the half-edge $g = \gamma_{2t^*-1}$ is succeeded by the half-edge $h = \gamma_{2t^*}$, then for all t such that $\gamma_{2t-1} = g$, we must also have $\gamma_{2t} = h$. Similarly, for all t with $\gamma_{2t} = g$, we must also have $\gamma_{2t-1} = h$.*

3. Consecutiveness*: the sequence of half-edges, when glued together, must correspond to a valid walk. Formally, for every t, if we have $\gamma_{2t} = ((u_{2t}, v_{2t}), i_{2t})$ and $\gamma_{2t+1} = ((v_{2t+1}, u_{2t+1}), i_{2t+1})$, then we must have $v_{2t+1} = v_{2t}$ and $i_{2t+1} = i_{2t}$.*

Colloquially, if two half-edges $\gamma = (\boldsymbol{e}, i), \gamma' = ((\boldsymbol{e})^{-1}, j)$ appear consecutively in a sequence with γ in an odd position and γ' in an even position, we will say that they are *glued together* to give the edge $\{(\boldsymbol{e}_1, i), (\boldsymbol{e}_2, j)\}$ (where $\boldsymbol{e}_1, \boldsymbol{e}_2$ are the first and second endpoints of $\boldsymbol{e}$, respectively).

Definition A.3. (non-backtracking sequence) *A sequence of half-edges $(\gamma_1, \ldots, \gamma_k)$ is called* non-backtracking *if it does not define a walk that backtracks; that is, for each $t \in [k]$, if $\gamma_{2t} = (\boldsymbol{e}_{2t}, i_{2t})$ and $\gamma_{2t+1} = (\boldsymbol{e}_{2t+1}, i_{2t+1})$, we require that $\boldsymbol{e}_{2t} \neq \boldsymbol{e}_{2t+1}$.*

We define Γ^{2k} to be the set of all valid, non-backtracking sequences of $2k$ half-edges.

A.2 Walk decomposition

For $\boldsymbol{e} = \{u, v\} \in E$, define $M_{\boldsymbol{e}}$ to be the $n \times n$ signed permutation matrix which encodes σ_e, so that $(M_{\boldsymbol{e}})_{ij} = \xi(\{(u, i), (v, j)\})$ if and only if $\sigma_{\boldsymbol{e}}(i) = j$. Further, for two half edges $\gamma = (\vec{\boldsymbol{e}}, i), \gamma' = (\vec{\mathbf{f}}, j)$, we let $M_{\gamma,\gamma'} = \mathbb{1}[\vec{\boldsymbol{e}} = (\vec{\mathbf{f}})^{-1}] \cdot \mathbb{1}[\sigma_{\boldsymbol{e}}(i) = j] \cdot \xi((\boldsymbol{e}_1, i), (\boldsymbol{e}_2, j))$ (where $\boldsymbol{e}$ is the undirected version of $\vec{\boldsymbol{e}}$

For two arcs $\vec{e}, \vec{f} \in \vec{E}_n$, let $\Gamma^{2k}_{\vec{e},\vec{f}}$ be the set of all valid, non-backtracking sequences of $2k$ half-edges

[5] In our setting, we will choose $\xi_n(e) \in \{\pm 1\}$ independently and uniformly for each $e \in E_n$.

[6] Again, in our setting we will choose each $\sigma_{\boldsymbol{e}}$ uniformly at random in S_n.

$(\gamma_1, \ldots, \gamma_{2k})$, such that γ_1, γ_2 form e when glued together, with the direction of $\vec{e}$ specified by γ_1, and such that $\gamma_{2k-1}, \gamma_{2k}$ form $\vec{f}$ when glued together, with the direction of $\vec{f}$ specified by γ_{2k-1}. We have by definition that

$$(B_n^k)_{ef} = \sum_{\gamma \in \Gamma^{2k+2}_{\vec{e},\vec{f}}} \prod_{s=1}^{k} M_{\gamma_{2s-1}\gamma_{2s}}, \tag{A.1}$$

since if a sequence γ is not valid or non-backtracking, it will have value 0.

We now define *tangles*, which are undesirable, low-probability walk structures (we will be able to discard their contribution to Equation (A.1)).

Definition A.4. (tangle-free) *For a positive integer ℓ, a graph G is* ℓ-tangle free *if it contains at most one cycle in every neighborhood of radius at most ℓ. A valid sequence $\gamma \in \Gamma^{2k}$ is ℓ-tangle free if the graph given by the edges and vertices visited by γ does not contain more than one cycle in any neighborhood of radius at most ℓ.*

The following lemma from [Bor17] proves that with high probability, Y_n is ℓ-tangle free.

Lemma A.1. ([Bor17, Lemma 24]) *If $\ell \leq \kappa \log_{d-1} n$ with $\kappa \in [0, 1/4]$ and d the maximum degree of a vertex in Y, then with high probability Y_n is tangle-free.*

Finally, we will require the following definition.

Definition A.5. *A valid sequence γ is* even *if the walk it induces contains every undirected edge with even multiplicity.*

A.3 Bounding the expectation of a single walk

Now, we bound the expectation of the product of entries along a walk.

For a sequence $\gamma = (\gamma_1, \ldots, \gamma_{2\ell})$ of length 2ℓ, with $\gamma_t = ((u_t, v_t), i_t)$, let E_γ be the set of lifted edges in γ,

$$E_\gamma = \{\{(u_{2t-1}, i_{2t-1}), (v_{2t-1}, i_{2t})\} \mid t \in [k]\}.$$

Proposition A.1. *Suppose that γ is a valid sequence of length $2k \ll \sqrt{n}$. Let $\ell < \frac{1}{4}\log_{d-1} n$. Then we have*

$$\mathop{\mathbf{E}}_{\sigma,\xi}\left[\prod_{s=1}^{k} M_{\gamma_{2s-1}\gamma_{2s}}\right] \leq \mathbb{1}[\gamma \text{ even}] \cdot (1 + o_n(1)) \cdot \left(\frac{1}{n}\right)^{|E_\gamma|}.$$

Proof. Consider some valid sequence of half-edges $\gamma = (\gamma_1, \ldots, \gamma_{2k})$, and let $\gamma_t = ((u_t, v_t), i_t)$ and $\vec{\boldsymbol{e}}_t = (u_t, v_t)$, $\boldsymbol{e}_t = \{u_t, v_t\}$ for convenience. We have that

$$\mathop{\mathbf{E}}_{\sigma,\xi}\left[\prod_{s=1}^{k} M_{\gamma_{2s-1}\gamma_{2s}}\right] = \mathop{\mathbf{E}}_{\sigma,\xi}\left[\prod_{\boldsymbol{e}\in\gamma} \prod_{\substack{t\in[k]\\ \boldsymbol{e}_{2t-1}=\boldsymbol{e}}} M_{\gamma_{2t-1}\gamma_{2t}}\right] \tag{A.2}$$

$$= \prod_{\boldsymbol{e}\in\gamma} \mathop{\mathbf{E}}_{\sigma,\xi}\left[\prod_{\substack{t\in[k]\\ \boldsymbol{e}_{2t-1}=\boldsymbol{e}}} M_{\gamma_{2t-1}\gamma_{2t}}\right], \tag{A.3}$$

since for $\boldsymbol{e} \neq \boldsymbol{e}'$, $\sigma_{\boldsymbol{e}}$ and $\sigma_{\boldsymbol{e}'}$ are independent, and by the independence of ξ_n. Expanding the entries of M according to M's definition,

$$\text{eq. (A.3)} = \prod_{\boldsymbol{e}\in\gamma} \mathop{\mathbf{E}}_{\sigma,\xi}\Big[\prod_{\substack{t\in[k]\\ \boldsymbol{e}_{2t-1}=\boldsymbol{e}}} \mathbb{1}[\sigma_{e_{2t-1}}(i_{2t-1}) = i_{2t}] \times$$

$$\xi((u_{2t-1}, i_{2t-1}), (v_{2t-1}, i_{2t}))\Big]. \tag{A.4}$$

By the independence of the signing ξ, we have that the expectation of any sequence in which any (undirected) edge is visited an odd number of times is 0. Assimilating this fact,

$$\text{eq. (A.4)} = \mathbb{1}[\gamma \text{ even}] \cdot \prod_{e\in\gamma} \mathop{\mathbf{E}}_{\sigma}\left[\prod_{\substack{t\in[k]\\ \boldsymbol{e}_{2t-1}=\boldsymbol{e}}} \mathbb{1}[\sigma_{\boldsymbol{e}_{2t-1}}(i_{2t-1}) = i_{2t}]\right]. \tag{A.5}$$

Now, suppose that $k_{\boldsymbol{e}}$ distinct lifted copies of the edge $\boldsymbol{e} \in E$ appear in γ. Since γ is consistent, and because we may assume every edge appears with even multiplicity, the term within the expectation just corresponds to fixing $k_{\boldsymbol{e}}$ edges of a permutation on n elements. Thus we simplify,

$$\text{eq. (A.5)} = \mathbb{1}[\gamma \text{ even}] \cdot \prod_{\boldsymbol{e}\in\gamma} \frac{(n-k_{\boldsymbol{e}})!}{n!} \tag{A.6}$$

$$\leq \mathbb{1}[\gamma \text{ even}] \cdot \prod_{\boldsymbol{e}\in\gamma} \left(\frac{1}{n}\left(1 + \frac{2k_{\boldsymbol{e}}}{n}\right)\right)^{k_{\boldsymbol{e}}}, \tag{A.7}$$

where to obtain the last inequality we have used that for $i \leq k_{\boldsymbol{e}} \ll \sqrt{n}$,

$$\frac{1}{n-i} \leq \frac{1}{n}\left(1 + \frac{2i}{n}\right) \leq \frac{1}{n}\left(1 + \frac{2k_{\boldsymbol{e}}}{n}\right).$$

And now since $\sum_{\boldsymbol{e}\in\gamma} k_{\boldsymbol{e}} = |E_\gamma|$ is the number of distinct lifted edges in γ, and the number of base edges is at most the number of lifted edges,

$$\leq \mathbb{1}[\gamma \text{ even}] \cdot \left(\frac{1}{n}\right)^{|E_\gamma|} \left(1 + \frac{2k}{n}\right)^{2k}. \tag{A.8}$$

Using that $2k \ll \sqrt{n}$ we obtain our conclusion.

A.4 Counting walks To apply Fact A.1, we will need to bound the trace of a power of $B_n^\ell(B_n^\ell)^*$. Since the trace corresponds to a sum over walks, and because in Section A.3 we have a bound on the expectation of each walk as a function of the number of distinct edges and the evenness of the walk, we have reduced our problem to counting the number of walks of various types. We will follow the definitions of Bordenave rather closely, so we may recycle his bounds.

We have that

$$\operatorname{tr}\left((B^\ell(B^\ell)^*)^m\right) = \sum_{e_1,\ldots,e_{2m-1}\in E_n^{2m-1}} \prod_{s=1}^{2m-1} (B_n^\ell)_{e_s,e_{s+1}}, \tag{A.9}$$

where we have taken $s+1$ modulo $2m-1$. To characterize the summation, it is useful for us to define the following set of sequences of half-edges, which have the property that large sub-sequences are tangle-free.

Definition A.6. *Let $W_{\ell,m}$ be the set of sequences of half-edges γ of length $2\ell \times 2m$ with the properties that, if we write γ as a sequence of sub-sequences $\gamma = (\gamma^{(1)},\ldots,\gamma^{(2m)})$*

1. *For each $s \in [2m]$, the sub-sequence $\gamma^{(s)}$ is valid, non-backtracking, and* **tangle-free***.*
2. *For each $s \in [m]$, the final edge in $\gamma^{(s)}$ is equal to the first edge in $\gamma^{(s+1)}$ (where we take addition mod $2m$). Formally, if $\gamma^{(t)} = (((u_1^{(t)}, v_1^{(t)}), i_1^{(t)}),\ldots,((u_{2\ell}^{(t)}, v_{2\ell}^{(t)}), i_{2\ell}^{(t)}))$, then we require $u_{2\ell-1}^{(s)} = u_1^{(s+1)}$, $v_{2\ell-1}^{(s)} = v_1^{(s+1)}$, $i_{2\ell-1}^{(s)} = i_1^{(s+1)}$ and $i_{2\ell}^{(s)} = i_2^{(s+1)}$.*

Recall we have defined Π to be the set of all half-edges (not necessarily present in Y_n).

Definition A.7. *We define an equivalence relation on Π^m: $\gamma,\gamma' \in \Pi^m$, with $\gamma_t = ((u_t, v_t), i_t)$ and with $\gamma'_t = ((u'_t, v'_t), i'_t)$ for $t \in [m]$. We'll say that for $\gamma \sim \gamma'$ if for all $t \in [m]$ we have $(u_t, v_t) = (u'_t, v'_t)$, and if in addition there exists a tuple of permutations in S_n, one for each vertex $u \in V$ from the base graph, $(\sigma_u)_{u\in V}$, so that $i'_t = \sigma_{u_t}(i_t)$.*

We observe that if γ is *even*, then any $\gamma' \sim \gamma$ is even as well. Similarly, if $\gamma \sim \gamma'$, then $|E_\gamma| = |E_{\gamma'}|$. We choose a canonical representative for each equivalence class:

Definition A.8. (Canonical sequence) *Let $V_\gamma(u) \subseteq \{u\} \times [n]$ be the set of all vertices of Y_n visited by γ which include u. We'll call $\gamma \in \Pi^m$* canonical *if for all $u \in V$, $V_\gamma(u) = \{(u,1),\ldots,(u,|V_\gamma(u)|)\}$, and if the vertices of $V_\gamma(u)$ appear in lexicographical order in γ.*

The following lemmas are given in [Bor17].

Lemma A.2. ([Bor17, Lemma 27]) *Let $\gamma \in \Pi^m$, and let $V_\gamma \subseteq V \times [n]$ be the set of vertices of Y_n which appear in γ. Suppose that $|V_\gamma| = s$. Then γ is isomorphic to at most n^s elements in Π^m.*

Lemma A.3. ([Bor17, Lemma 28]) *Let $\mathcal{W}_{\ell,m}(s,a)$ be the subset of canonical paths in $W_{\ell,m}$ with $|V_\gamma| = s$ and $|E_\gamma| = a$. Then for any fixed $\overline{\rho} > \rho(B)$, there exists a constant κ (depending also on Y) such that we have*

$$|\mathcal{W}_{\ell,m}(s,a)| \le \overline{\rho}^s(\kappa\ell m)^{8m(a-s+1)+10m}.$$

We are now ready to bound the contribution of the sums of tangle-free sections.

Proposition A.2. *For $m = \lfloor \frac{\log n}{17\log\log n} \rfloor$, $n \ge 3$, and $\ell \le \frac{1}{4}\log_{d-1} n$, and $\overline{\rho} > 1$ as above, there is a constant c independent of n such that*

$$\mathbf{E}\left[\sum_{\gamma\in W_{\ell,m}} \prod_{i=1}^{2m}\prod_{t=1}^{\ell} M_{\gamma_{2t-1}^{(i)},\gamma_{2t}^{(i)}}\right] \le n(c\ell m)^{10m}\overline{\rho}^{(\ell+2)m}.$$

Proof. We split the left-hand side according to the $\mathcal{W}$ equivalence classes,

$$\mathbf{E}\left[\sum_{\gamma\in W_{\ell,m}} \prod_{i=1}^{2m}\prod_{t=1}^{\ell} M_{\gamma_{2t-1}^{(i)},\gamma_{2t}^{(i)}}\right] \le \sum_{s=1}^{\infty}\sum_{a=s-1}^{\infty} n^s \times \sum_{\gamma\in\mathcal{W}_{\ell,m}(s,a)} \mathbf{E}\left[\prod_{i=1}^{2m}\prod_{t=1}^{\ell} M_{\gamma_{2t-1}^{(i)},\gamma_{2t}^{(i)}}\right], \tag{A.10}$$

where we have used that $|V_\gamma| - 1 \le |E_\gamma|$, since G_γ is connected. Now applying Proposition A.1 (using that $\ell m \ll \sqrt{n}$), we have that for $\gamma \in \mathcal{W}_{\ell,m}(s,a)$,

$$\mathbf{E}\left[\prod_{i=1}^{2m}\prod_{t=1}^{\ell} M_{\gamma_{2t-1}^{(i)},\gamma_{2t}^{(i)}}\right] \le \mathbb{1}[\gamma \text{ even}]\cdot(1+o_n(1))\cdot\left(\frac{1}{n}\right)^a.$$

Plugging this in above, along with the bound on $|\mathcal{W}_{\ell,m}(s,a)|$ from Lemma A.3, we have

$$\text{eq. (A.10)} \le \sum_{s=1}^{(\ell+2)m+1} n^s \sum_{a=s-1}^{(\ell+2)m} \overline{\rho}^s(\kappa\ell m)^{8m(a-s+1)+10m} \times (1+o_n(1))\cdot n^{-a},$$

where we use the fact that γ must be even to obtain that $|E_\gamma| = s \le (\ell+2)m$, (as there are only $2(\ell+2)m$ edges in the sequence γ, and each must appear twice), and adjusted the upper limits of the summation accordingly.

We re-index the above summation, setting $a' = a-s+1$ and beginning to sum from $a' = 0$ (and summing till $a' = \infty$, as this yields a valid upper bound),

$$\text{eq. (A.10)} \leq (1+o_n(1)) \cdot (\kappa\ell m)^{10m} \cdot \sum_{s=1}^{(\ell+2)m+1} n^s \overline{\rho}^s \times \left(\frac{1}{n}\right)^{s-1} \sum_{a'=0}^{\infty} (\kappa\ell m)^{8ma'} \cdot \left(\frac{1}{n}\right)^{a'}$$

$$= (1+o_n(1)) \cdot n(\kappa\ell m)^{10m} \times \sum_{s=1}^{(\ell+2)m+1} \overline{\rho}^s \sum_{a'=0}^{\infty} (\kappa\ell m)^{8ma'} \cdot \left(\frac{1}{n}\right)^{a'}. \tag{A.11}$$

For our chosen m, when n is large enough, $\frac{(\kappa\ell m)^{8m}}{n} \leq \frac{(\log n)^{16m}}{n} \leq n^{-1/17}$. Combining this observation with the fact that the rightmost sum is a geometric sum, there is a constant c such that

$$\text{eq. (A.11)} \leq cn(\kappa\ell m)^{10m} \cdot \sum_{s=1}^{(\ell+2)m+1} \overline{\rho}^s.$$

Finally, we are left again with a geometric sum; since we have $\overline{\rho} > 1$, there is a constant c' so that

$$\leq c' n(\kappa\ell m)^{10m} \cdot \overline{\rho}^{(\ell+2)m+1}.$$

Using that $\overline{\rho}$ is independent of n to push $\overline{\rho}$ into the constant, we have our conclusion.

A.5 Putting things together We now finally have the ingredients to prove Theorem 4.2.

Proof. [Proof of Theorem 4.2] Fix $\epsilon > 0$, choose $\overline{\rho} > \rho(B)$ such that $\sqrt{\overline{\rho}} < \sqrt{\rho(B)} + \epsilon/2$, let $\ell = \kappa \log_{d-1} n$ for a constant $\kappa \in (0, 1/4)$, and let $m = \lfloor \frac{\log n}{17 \log\log n} \rfloor$. By Lemma A.1, if $\mathcal{E}$ is the event that Y_n is ℓ-tangle-free,

$$\begin{aligned}
\mathbf{Pr}[\rho(B_n) &\geq \sqrt{\rho(B)} + \epsilon] \\
&\leq \mathbf{Pr}[\rho(B_n) \geq \sqrt{\overline{\rho}} + \epsilon/2, \mathcal{E}] + o(1) \\
&\leq \mathbf{Pr}\, \|B_n^\ell (B_n^\ell)^*\|^{1/2\ell} \geq \sqrt{\overline{\rho}} + \epsilon/2, \mathcal{E}] + o(1).
\end{aligned}$$

If Y_n is ℓ-tangle-free, then only sequences $\gamma \in W_{\ell,m}$ contribute to Equation (A.9), as any (consecutive) subsequence $\gamma^{(i)} \subset \gamma$ of length 2ℓ defines a length-ℓ walk in Y_n. So using Fact A.1 in conjunction with Equation (A.9) and Proposition A.2, we have that

$$\mathbf{E}[\mathrm{tr}((B_n^\ell (B_n^\ell)^*)^m) \cdot \mathbb{1}[\mathcal{E}]] \leq n(c\ell m)^{10m} \overline{\rho}^{(\ell+2)m}.$$

Taking the $2\ell m$th root on the right, by our choice of $\ell = \Theta(\log n)$ and $m = \Theta(\log n / \log\log n)$, $(c\ell m)^{5/\ell} = o(\log^2 n)^{1/\log n} = 1 + o(1)$, $n^{1/2\ell m} \leq 2^{\Theta(\log\log n/\log n)} = 1 + o(1)$, and since $\overline{\rho}$ is independent of n, $\overline{\rho}^{1/\ell} = 1 + o_n(1)$, and we have the desired conclusion.