Certifying solution geometry in random CSPs: counts, clusters and balance

Jun-Ting Hsieh[*]   Sidhanth Mohanty[†]   Jeff Xu[‡]

June 25, 2021

Abstract

An active topic in the study of random constraint satisfaction problems (CSPs) is the geometry of the space of satisfying or almost satisfying assignments as the function of the density, for which a precise landscape of predictions has been made via statistical physics-based heuristics. In parallel, there has been a recent flurry of work on *refuting* random constraint satisfaction problems, via nailing refutation thresholds for spectral and semidefinite programming-based algorithms, and also on *counting* solutions to CSPs. Inspired by this, the starting point for our work is the following question:

What does the solution space for a random CSP look like to an efficient algorithm?

In pursuit of this inquiry, we focus on the following problems about random Boolean CSPs at the densities where they are unsatisfiable but no refutation algorithm is known.

1. **Counts.** For every Boolean CSP we give algorithms that with high probability certify a subexponential upper bound on the number of solutions. We also give algorithms to certify a bound on the number of large cuts in a Gaussian-weighted graph, and the number of large independent sets in a random d-regular graph.

2. **Clusters.** For Boolean 3CSPs we give algorithms that with high probability certify an upper bound on the number of *clusters* of solutions.

3. **Balance.** We also give algorithms that with high probability certify that there are no "unbalanced" solutions, i.e., solutions where the fraction of $+1$s deviates significantly from 50%.

Finally, we also provide hardness evidence suggesting that our algorithms for counting are optimal.

Contents

1 Introduction

Constraint satisfaction problems (CSPs) are fundamental in the study of algorithm design and complexity theory. They are simultaneously simple and also richly expressive in capturing a wide range of computational tasks, which has led to fruitful connections to other areas of theoretical computer science (see, for example, [Gol11, ABW10] for connections to cryptography, [DLSS14] for applications to hardness of learning, and [Fei02] for applications to average-case hardness). Hence, understanding them has received intense attention in the past few decades, leading to several comprehensive theories of their complexity. Some of the highlights include: the Dichotomy Theorem, which characterizes the worst-case complexity of satisfiability of CSPs via their algebraic properties [Sch78, BJK05, Zhu20], inapproximability results via the PCP Theorem [Hås01], and the theory of optimal inapproximability based on connections between semidefinite programming and the Unique Games conjecture [Kho02, KKMO07, Rag08].

In this work, we are interested in the algorithmic aspects of random instances of CSPs. There has been a diverse array of phenomena about random CSPs illustrated in recent work, of dramatically varying nature depending on the ratio of the number of constraints to the number of variables, known as the *density*. Of central importance is the *satisfiability threshold*, which marks a phase transition where a random CSP instance shifts from being likely satisfiable to being likely unsatisfiable. When the density is well below the satisfiability threshold, there are several algorithms for tasks such as counting and sampling assignments to a random CSP instance [Moi19, JPV21, GGGY19, BGG+19], whereas well above this threshold there are efficient algorithms for *certifying* that random CSPs are unsatisfiable [AOW15]. The densities in the interim hold mysteries that we don't yet fully understand, and this work is an effort to understand the algorithmic terrain there. To make matters concrete, for now we will specialize the discussion of the problem setup and our work to the canonical 3SAT predicate.

Consider a random 3SAT formula $\mathcal{I}$ on n variables and Δn clauses where each clause is sampled uniformly, independently, and adorned with uniformly random negations. Once the density Δ is a large enough constant, this random instance is unsatisfiable with high probability.[1] On the other hand, the widely believed Feige's random 3SAT hypothesis [Fei02] conjectures that when Δ is any constant, there is no algorithm to *certify* that a random instance is unsatisfiable. Further, the best known algorithms for efficiently certifying that it is unsatisfiable require $\Delta \gtrsim \sqrt{n}$ [GL03, COGL07, FO07, AOW15]. Moreover, when $\Delta \lesssim \sqrt{n}$ there is a lower bound against the Sum-of-Squares hierarchy [Gri01, Sch08] (known to capture many algorithmic techniques), which suggests an *information-computation gap* and earns $\sqrt{n}$ the name *refutation threshold*.

In this picture, at both densities $n^{.25}$ and $n^{.35}$, $\mathcal{I}$ is likely unsatisfiable but "looks" satisfiable to an efficient algorithm. But is there a concrete sense in which a random formula at density $n^{.25}$ is "more satisfiable" than one at density $n^{.35}$ from the lens of a polynomial-time algorithm? A natural measure of a 3SAT formula's satisfiability is its number of satisfying assignments, which motivates the following question.

What is the best efficiently certifiable upper bound on the number of assignments satisfying $\mathcal{I}$?

In the context of 3SAT, our work proves:

---

[1]In fact, it is conjectured that there is a sharp threshold for unsatisfiability once Δ crosses some constant $\alpha_{\text{SAT}} \approx 4.267$.

Theorem 1.1 (Informal). *There is an efficient algorithm to certify with high probability that a random* 3SAT *formula with density* $\Delta = n^{1/2-\delta}$ *has at most* $\exp(\widetilde{O}(n^{3/4+\delta/2}))$ *satisfying assignments.*

In addition to certifying the number of satisfying assignments, we can certify that the solutions form clusters and upper bound the number of clusters under the refutation threshold.

Clusters. Besides the satisfiability threshold, random kSAT is conjectured to go through other phase transitions too, as predicted in the work of [KMRT$^+$07]. In particular, the *clustering threshold* is the density where the solution space is predicted to change from having one giant component to roughly resembling a union of several small Hamming balls, known as *clusters*, that are pairwise far apart in Hamming distance.

Much like the refutation threshold that marks where efficient algorithms can witness unsatisfiability, it is natural to ask if there is some regime under the refutation threshold where an efficient algorithm can witness a bound on the number of clusters of solutions. The following more nuanced version of Theorem 1.1 gives an answer to this question.

Theorem 1.2 (Informal). *There is an efficient algorithm to certify with high probability that the satisfying assignments of a random* 3SAT *formula with density* $\Delta = n^{1/2-\delta}$ *are covered by at most* $\exp(\widetilde{O}(n^{1/2+\delta}))$ *diameter-*$\widetilde{O}(n^{3/4+\delta/2})$ *clusters.*

Balance in the solution space. Suppose at density Δ, a typical 3SAT formula has $\sim \exp(c_\Delta n)$ satisfying assignments, then due to the uniformly random negations in clauses, each string is satisfying with probability $\sim \exp((c_\Delta - 1)n)$. Then one can show via the first moment method that with high probability there are no satisfying assignments with Hamming weight outside $\left[\frac{1}{2} - f(c_\Delta), \frac{1}{2} + f(c_\Delta)\right]$.[2] In particular, the intersection of the solution space with the set of unbalanced strings empties out under the satisfiability threshold. This raises the question:

> *Is there an efficient algorithm to certify that a random CSP instance has no unbalanced assignments at density significantly under the refutation threshold?*

We affirmatively answer this question and in the special case of 3SAT prove:

Theorem 1.3 (Informal). *There is an efficient algorithm to certify with high probability that a random* 3SAT *formula with density* $\Delta = n^{1/2-\delta}$ *has no satisfying assignments with Hamming weight outside*

$$\left[\frac{1}{2} - \widetilde{\Theta}\left(\frac{1}{n^{1/4-\delta/2}}\right), \frac{1}{2} + \widetilde{\Theta}\left(\frac{1}{n^{1/4-\delta/2}}\right)\right].$$

We illustrate our upper bounds for counting satisfying assignments and clusters in Figure 1. We delve into the precise technical statements of our results and the techniques involved in proving them in Section 1.1. Then to put our work in context, we survey and discuss existing work on information-computation gaps, and algorithmic work on counting, sampling and estimating partition functions in Section 1.2.

---

[2]where f is chosen so that the number of strings outside that Hamming range is $\ll \exp((c_\Delta - 1)n)$.
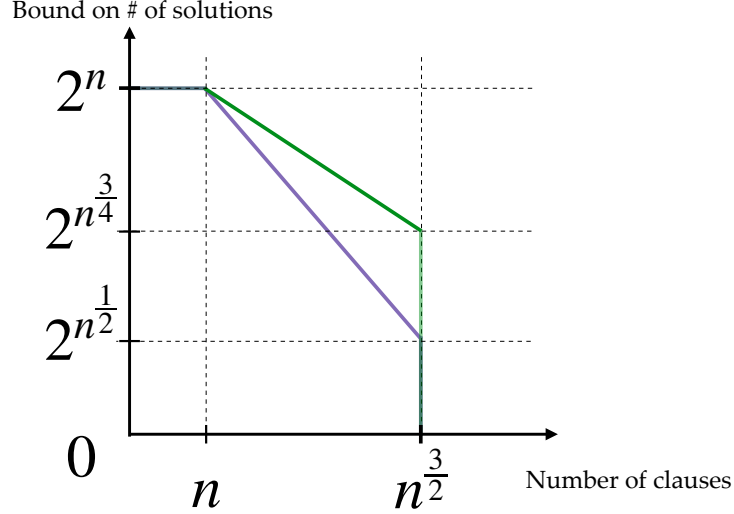
Figure 1: Our results for 3SAT. **Green**: certifiable upper bound on the number of satisfying assignments. **Purple**: upper bound on the number of clusters of satisfying assignments. In the case of kSAT, the green plot looks identical but with n replaced by $n^{(k-1)/2}$ and $n^{3/2}$ replaced by $n^{k/2}$.

1.1 Our Contributions

In this section, we give a more detailed technical description of our contributions. To set the stage for doing so, we first formally clarify the notion of *certification* and some preliminaries on constraint satisfaction problems.

Fix a sample space Ω, a probability distribution $\mathcal{D}$ over Ω, and a function $f : \Omega \to \mathbb{R}$. For example, Ω is the space of 3SAT instances, $\mathcal{D}$ is the distribution of instances given by the random 3SAT model, and f is the number of satisfying assignments.

Definition 1.4. We say that a deterministic algorithm $\mathcal{A}$ certifies that $f \leqslant C$ with probability over $1 - p$ over $\mathcal{D}$ if $\mathcal{A}$ satisfies

1. For all $\omega \in \Omega$, $f(\omega) \leqslant \mathcal{A}(\omega)$.

2. For a random sample $\omega \sim \mathcal{D}$, $\mathcal{A}(\omega) \leqslant C$ with probability over $1 - p$.

We emphasize that an algorithm that always outputs the typical value of f is *not* a certification algorithm: it will satisfy the second condition but not the first. Thus, in several average-case problems, there are gaps between the typical value and the best known certifiable upper bound.

Remark 1.5. Due to the guarantees of $\mathcal{A}$, one can think of the "transcript" of the algorithm on input ω as being a proof that $f(\omega) \leqslant \mathcal{A}(\omega)$.

Definition 1.6. A *predicate* $P : \{\pm 1\}^k \to \{0, 1\}$ is any Boolean function that is not the constant function that always evaluates to 1. An instance $\mathcal{I}$ of a *constraint satisfaction problem* on predicate P and vertex set $[n]$ is a collection of *clauses*, where a clause is a pair (c, S) with $c \in \{\pm 1\}^k$ and $S \in [n]^k$. Given $x \in \{\pm 1\}^n$, the value of $\mathcal{I}$ on x is:

$$\mathcal{I}(x) := \frac{1}{|\mathcal{I}|} \sum_{(c,S) \in \mathcal{I}} P(c_1 x_{S_1}, \ldots, c_k x_{S_k}).$$

3

We say x *satisfies* a clause (c, S) if $P(c_1 x_{S_1}, \ldots, c_k x_{S_k}) = 1$, and say x is $(1 - \eta)$*-satisfying* if $\mathcal{I}(x) \geqslant 1 - \eta$. If $\eta = 0$, we say x is *exactly satisfying*.

In this work we are concerned with random CSPs. We defer an exact description of the random model to Section 2.3 (note however that the common random models used in the literature are all qualitatively similar; cf. [AOW15, Appendix D]). Our first result is an algorithm certifying a subexponential upper bound on the number of $(1 - \eta)$-satisfying assignments for random CSPs.

Theorem 1.7. *Let $\mathcal{I}$ be a random kCSP instance on any predicate P on n variables and Δn clauses. For every $\varepsilon > 0$, there is an algorithm that certifies with high probability that the number of $(1 - \eta)$-satisfying assignments to $\mathcal{I}$ is upper bounded by:*

$$\exp\left(\widetilde{O}(\eta n)\right) \cdot \exp\left(\widetilde{O}\left(\sqrt{\frac{n^{(k+1)/2}}{\Delta}}\right)\right) \cdot \exp\left(O\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right).$$

To more easily parse the statement, let's plug in concrete parameters.

Remark 1.8. Let's fix the predicate to be kSAT for any $k \geqslant 3$, $\eta = 0$, and $\Delta = n^{k/2-1.1}$. The quantity of interest is the number of exactly satisfying solutions to a random kSAT formula at a density strictly smaller than the refutation threshold of $\widetilde{\Omega}(n^{k/2-1})$. Then, we get an algorithm that with high probability certifies that the number of exactly satisfying assignments is at most:

$$\exp(\widetilde{O}(n^{0.8})),$$

which is a subexponential bound. More generally, our algorithms certify a subexponential bound on the number of satisfying assignments for kSAT for $\Delta = n^{k/2-1.5+c}$ for any $c > 0$ and this bound improves as we increase c.

The proof of Theorem 1.7 relies on 3 ingredients of increasing complexity. The first is the simple observation that given a kCSP instance $\mathcal{I}$ on any predicate P, there is a transformation to a kSAT instance $\mathcal{I}'$ such that:

(i) For any $\eta > 0$, if x is $(1 - \eta)$-satisfying for $\mathcal{I}$, then it is also $(1 - \eta)$-satisfying for $\mathcal{I}'$.

(ii) If $\mathcal{I}$ is a random instance of a CSP on P with density Δ, then $\mathcal{I}'$ is a random instance of kSAT with density Δ.

This reduction is described in the proof of Corollary 4.8.

The second ingredient is a generalization of the "3XOR-principle" of [Fei02, FO07], which we call the "kXOR-principle". The kXOR principle, which we state below, reduces count certification/refutation for a random kSAT formula to the same task on a random kXOR formula.

Lemma 1.9. *Let $\mathcal{I}$ be a random kSAT formula on $m = \Delta n$ clauses. There is an efficient algorithm that with high probability certifies that any $(1 - \eta)$-satisfying assignment of $\mathcal{I}$ must kXOR-satisfy at least $\left(1 - O(\eta) - \widetilde{O}\left(\sqrt{\frac{n^{(k-3)/2}}{\Delta}}\right)\right) m$ clauses.*

4

We detail the proof in Section 3, which is close to the reduction from generic CSP refutation to kXOR refutation in [AOW15] based on the Fourier expansion.

For the sake of a notationally simple sketch, let's restrict ourselves to the case $\eta = 0$. We can write $k\mathsf{SAT}(x_1, \ldots, x_k) = (1 - 2^{-k}) + 2^{-k} x_1 x_2 \cdots x_k + q(x_1, \ldots, x_k)$ where q is a degree-$(k-1)$ polynomial without a constant term. Thus, given a random $k\mathsf{SAT}$ instance $\mathcal{I}$ and any satisfying assignment x:

$$1 = \mathcal{I}(x) = 1 - 2^{-k} + 2^{-k} \frac{1}{|\mathcal{I}|} \sum_{(c,S) \in \mathcal{I}} \prod_{i=1}^{k} c_i x_{S_i} + \frac{1}{|\mathcal{I}|} \sum_{c,S \in \mathcal{I}} q(c_1 x_{S_1}, \ldots, c_k x_{S_k}).$$

Once $\Delta \gtrsim n^{(k-3)/2}$ the refutation algorithm of [AOW15] can be employed to certify that the last term is insignificantly small by virtue of the last term being a degree-$(k-1)$ polynomial with no constant term. This would force $2^{-k} \frac{1}{|\mathcal{I}|} \sum_{(c,S) \in \mathcal{I}} \prod_{i=1}^{k} c_i x_{S_i}$ to be near 1, which is the same as saying x must kXOR-satisfy most clauses.

Our third ingredient for Theorem 1.7 is a count certification algorithm for kXOR, which we prove in Section 4.

Theorem 1.10. *For constant $k \geqslant 3$, consider a random kXOR instance with n variables and Δn clauses. For any constant $\varepsilon > 0$, there is a polynomial-time algorithm that certifies with high probability that the number of $(1 - \eta)$-satisfying assignments is at most*

$$\exp\left(\widetilde{O}(\eta n)\right) \cdot \exp\left(O\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right).$$

In fact, the certification algorithm only depends on the hypergraph structure of the kXOR instance and not the signings of each clause. This is crucial since our algorithm recursively looks at $(k-1)$XOR subinstances with unknown signings. The stronger statement we prove is:

Theorem 1.11. *For constant $k \geqslant 2$, consider a random k-uniform hypergraph $\boldsymbol{H}$ on n vertices and Δn hyperedges where $\Delta \gg \log n$. For $\varepsilon > 0$, there is a polynomial-time algorithm that certifies with high probability that the number of $(1 - \eta)$-satisfying assignments to any kXOR instance on $\boldsymbol{H}$ is at most*

$$\exp\left(\widetilde{O}(\eta n)\right) \cdot \begin{cases} 1 & \text{if } k = 2 \\ \exp\left(\widetilde{O}\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right) & \text{if } k \geqslant 3. \end{cases}$$

Theorem 1.11 is of interest beyond algorithmic considerations as it gives a high-probability bound on the number of approximate solutions for *any* kXOR formula on a random hypergraph.

Remark 1.12. Gaussian elimination is able to count exact solutions to an explicit kXOR instance but fails for counting $(1 - \eta)$-satisfying assignments or when the signings are unknown.

We now give a brief sketch of our proof of Theorem 1.11. Given a random k-uniform hypergraph, we would like to certify that *any* kXOR instance on this hypergraph has no more than $\exp\left(\widetilde{O}(\eta n)\right) \cdot \exp\left(\widetilde{O}\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right)$ approximate solutions. We will first present an overview in the context of 2XOR as the "base case", and then explain the algorithm for 3XOR to illustrate the "recursive step".

2XOR sketch. Let's consider a random graph G on n vertices and Δn edges where $\Delta \gg \log n$. Then, its degrees concentrate and its normalized Laplacian has a large spectral gap (more precisely, a spectral gap of $1 - O\left(\frac{1}{\sqrt{\Delta}}\right)$). As a consequence of Cheeger's inequality, any set S containing fewer than half the vertices has roughly half its edges leaving — which quantitatively would be around $\Delta|S|$. We prove that a large spectral gap and concentration of degrees is all that is necessary for any 2XOR instance to have an appropriately bounded number of satisfying assignments.

Now let $\mathcal{I}$ be any 2XOR instance on G. The key observation is that if x and x' are two $(1-\eta)$-satisfying assignments for $\mathcal{I}$, then the pointwise product $y := x \circ x'$ is $(1-2\eta)$-satisfying for $\mathcal{I}_+$, the 2XOR instance on G obtained by setting the sign on all constraints to $+1$. The constraints violated by y are the ones on the cut between S_+ and S_-, the positive and negative vertices in y respectively. There are roughly $\Delta \cdot \min\{|S_+|, |S_-|\}$, and consequently $\min\{|S_+|, |S_-|\} \leqslant 2\eta n$ since y is $(1-2\eta)$-satisfying. In particular, y either has at most $2\eta n$ positive entries or $2\eta n$ negative entries. The upshot is the number of $(1-\eta)$-satisfying assignments of $\mathcal{I}$ is at most $\exp(\widetilde{O}(\eta n))$. This sketched argument is carefully carried out in Section 4.1.

3XOR sketch. Now, let H be a random hypergraph on n vertices and Δn hyperedges. The observation here is that for any 3XOR instance $\mathcal{I}$ on H, any assignment x that $(1-\eta)$-satisfies $\mathcal{I}$ also approximately satisfies a particular *induced* 2XOR *instance* of a fixed subset of variables S (cf. Definition 2.11). The induced 2XOR instance's underlying graph G is fixed and distributed like a random graph, and only the signings on the edges vary as we vary x. That lets us run the algorithm for 2XOR on G to obtain an upper bound F on all induced instances on G, which then yields a bound of $2^{|S|} \cdot F$. This is where we use that our algorithm depends only on the underlying graph, hence avoiding an enumeration of all assignments to variables in S.

We immediately see that for a fixed subset S, the above procedure throws away most of the clauses (keeping only clauses that have 1 variable in S). Thus, it is clearly suboptimal to look at just one subset S. To resolve this, we partition the variables into subsets $S_1, \ldots, S_\ell$, run the algorithm on each of them, and aggregate the results. This is explained in detail in the proofs of Lemma 4.6 and Theorem 4.4.

Clustering. Our next result is an algorithm to upper bound the number of clusters formed by the solutions. Given $x \in \{\pm 1\}^n$, we call the Hamming ball $B(x, r)$ a *radius-r cluster* or *diameter-$2r$ cluster*. For 3CSPs we prove in Corollary 5.2:

Theorem 1.13. *Let P be any 3-ary predicate, and let $\mathcal{I}$ be a random instance of P on n variables and Δn clauses. Let $\eta \in [0, \eta_0]$ where η_0 is a universal constant, and let $\theta := 8\eta + O\left(\sqrt{\frac{\log^5 n}{\Delta}}\right)$. There is an algorithm that certifies with high probability that the $(1-\eta)$-satisfying assignments to $\mathcal{I}$ as a P-CSP instance are covered by at most*

$$\exp(O(\theta^2 \log(1/\theta))n)$$

diameter-(θn) clusters.

Inspecting the proof of counting 2XOR (specifically the argument about $\mathcal{I}_+$), we see that it additionally certifies that the approximate solutions form clusters. In a similar fashion, we certify that any pair of $(1-\eta)$-satisfying assignments to a random 3SAT instance must have Hamming distance close to 0 or roughly $\frac{n}{2}$, i.e. the solutions form clusters where the clusters are roughly $\frac{n}{2}$

6

apart. The main ingredient is an efficient algorithm to certify an important structural result of random 3-uniform hypergraphs (Lemma 5.4), allowing us to reason about the constraints violated in $\mathcal{I}_+$. In fact, Lemma 5.4 will also be a crucial step in refuting CSPs under global cardinality constraints in Section 6. The upshot is that we will be able to certify that any pair of solutions is either ρ-close or $\frac{1-\rho}{2}$-far.

The second ingredient is a result in coding theory. Since the clusters are roughly $\frac{1\pm\rho}{2}n$ apart in Hamming distance, the number of clusters must be upper bounded by the cardinality of the largest ρ-balanced binary error-correcting code. The best known upper bound is $2^{O(\rho^2 \log(1/\rho))n}$ by [MRRW77] (see also [Alo09]), which yields our final result. Complete details are in Section 5.

Balance. We observe that the idea of hypergraph expansion can be applied to the problem of strongly refuting random CSPs with *global cardinality constraints*. This problem was first investigated by Kothari, O'Donnell, and Schramm [KOS18], where they proved that under the refutation threshold $n^{k/2}$, the polynomial-time regime of the Sum-of-Squares hierarchy cannot refute the instance even with the global cardinality constraint $\sum_{i=1}^n x_i = B$ for any integer $B \in [-O(\sqrt{n}), O(\sqrt{n})]$ (here we assume $x \in \{\pm 1\}^n$). On the other hand, they proved once that $|B| > n^{3/4}$, Sum-of-Squares could indeed refute a random kXOR instance up to a factor of $\sqrt{n}$ under the refutation threshold.

We say an assignment x is ρ-biased if $\frac{1}{n}\left|\sum_{i\in[n]} x_i\right| \geqslant \rho$. We give a strong refutation algorithm for random instances of all Boolean CSPs under the constraint that the solution is "unbalanced".

Theorem 1.14. *Let P be any k-variable predicate and let $\mathcal{I}$ be a random CSP instance on $m := n^{\frac{k-1}{2}+\beta}$ clauses where $\beta > 0$. For every constant $\rho > 0$, there is an efficient algorithm that certifies that $\mathcal{I}$ has no 2ρ-biased assignment which $(1 - O(\rho^k))$-satisfies $\mathcal{I}$ as a P-CSP instance.*

Remark 1.15. Compared to [KOS18], our result is a strong refutation algorithm for all CSPs, whereas their algorithm is specific for kXOR and only a weak refutation (refuting only exactly satisfying assignments). For $k = 3$ (Theorem 6.1), we match their cardinality constraint requirement (see Remark 6.2). However, for $k \geqslant 4$ (Corollary 6.4), we require a slightly stronger cardinality assumption.

The formal statements and proofs are detailed in Section 6. Akin to the case for counting solutions, we employ the reduction of every kCSP to kSAT and the kXOR principle to reduce the problem to strongly refuting kXOR under global cardinality constraints.

The first main insight is that given a graph G which is a sufficiently good spectral expander, we can efficiently certify that any 2XOR instance on G, where the number of positive constraints is roughly equal to the number of negative constraints, has no unbalanced approximately satisfying assignments. The proof of this is based on using the expander mixing lemma to show that any imbalanced assignment x must satisfy $x_u x_v = +1$ for $\gg \frac{1}{2}$ of the edges, which then lets us lower bound the number of negative constraints that are violated.

Then given a random kXOR instance $\mathcal{I}$, we pick some set of ρn vertices S and consider all clauses with exactly $k - 2$ vertices in S and 2 variables outside S. If we place an edge between the two variables outside S for every clause, we get some random graph G. Now consider any assignment y to the variables in S. For this chosen set of clauses to be (nearly) satisfied, the assignment to variables outside S must nearly satisfy the induced 2XOR instance on the graph G

7

whose signings are determined by y. The second insight is that we can efficiently certify that for any assignment y the induced 2XOR instance has a roughly equal number of positive and negative constraints. This is possible since the quantity #positive constraints(y) − #negative constraints(y) is the objective value of a particular random $(k-2)$XOR instance on assignment y, which we can certify tight bounds on using the algorithm of [AOW15].

Certified counting for subspace problems. So far, we have developed certification algorithms for CSPs mainly based on analyses of random hypergraphs. For other inherently different problems such as counting solutions to the SK model, we turn to a different technique. Our main insight is that for several problems, the approximate solutions must lie close to a small-dimensional linear subspace. Thus, we can reduce the problem to counting the number of (Boolean) vectors close to a subspace. We name this technique *dimension-based count certification* since the algorithms and their guarantees only depend on the dimension of the subspace.

Theorem 1.16. *Let V be a linear subspace of dimension αn in $\mathbb{R}^n$. For any $\varepsilon \in (0, 1/4)$, the number of Boolean vectors in $\left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ that are ε away from V is upper bounded by $2^{(H_2(4\varepsilon^2) + \alpha \log \frac{3}{\varepsilon})n}$.*

We note that the upper bound is almost tight (see Remark 7.3 for more details).

We now give a brief overview of the proof of Theorem 1.16. First, we upper bound the maximum number of (normalized) Boolean vectors that can lie within any ε'-ball. Secondly, we take an ε-net of the unit ball in the subspace V (i.e. $B_1(0) \cap V$). We simply multiply the two quantities to get the upper bound, which only depends on the dimension of V.

Next, we apply this technique to two problems: the Sherrington-Kirkpatrick model and the independent sets in random d-regular graphs.

Sherrington-Kirkpatrick (SK). Given M sampled from $\mathsf{GOE}(n)$, the SK problem is to compute

$$\mathsf{OPT}(M) = \max_{x \in \{\pm 1\}^n} x^\top M x.$$

This problem can also be interpreted as finding the largest cut in a Gaussian-weighted graph. The SK model arises from the spin-glass model studied in statistical physics [SK75]. Talagrand [Tal06] famously proved that $\mathsf{OPT}(M)$ concentrates around $2\mathsf{P}^* n^{3/2} \approx 1.526 n^{3/2}$, where P^* is the *Parisi constant*, first predicted by Parisi [Par79, Par80].

Recently, the problem of certifying an upper bound for $\mathsf{OPT}(M)$ has received wide attention. A natural algorithm is the *spectral refutation*: $\mathsf{OPT}(M) \leqslant n \cdot \lambda_{\max}(M)$. Since $\lambda_{\max}(M)$ concentrates around $2\sqrt{n}$, the algorithm certifies that $\mathsf{OPT}(M) \leqslant (2 + o(1))n^{3/2}$, which we call the *spectral bound*. Clearly, there is a gap between the spectral bound and the true value, and it is natural to ask whether there is an algorithm that beats the spectral bound. Surprisingly, building on works by [MS16, MRX20, KB20], Ghosh et. al. [GJJ+20] showed that even the powerful Sum-of-Squares hierarchy cannot certify a bound better than $(2 - o(1))n^{3/2}$ in subexponential time. We also mention an intriguing work by Montanari [Mon19] where he gave an efficient algorithm for the *search problem* — to find a solution with objective value close to $\mathsf{OPT}(M)$ with high probability (assuming a widely-believed conjecture from statistical physics). However, we emphasize that his algorithm is not a certification algorithm (recall Definition 1.4).

In the spirit of this work, a natural question is to certify an upper bound on the number of assignments $x \in \{\pm 1\}^n$ such that $x^\top M x \geqslant 2(1 - \eta)n^{3/2}$ for some $\eta > 0$.

8

Theorem 1.17. *Let $M \sim \mathsf{GOE}(n)$. Given $\eta \in (0, \eta_0)$ for some universal constant η_0, there is an algorithm certifying that at most $2^{O(\eta^{3/5} \log \frac{1}{\eta})n}$ assignments $x \in \{\pm 1\}^n$ satisfy $x^\top M x \geqslant 2(1 - \eta)n^{3/2}$.*

Our proof first looks at the eigenvalue distribution of M, which follows the *semicircle law* (Theorem 2.21). This shows that any x that achieves close to the spectral bound must be close to the top eigenspace of M (of dimension determined by the semicircle law). Then, we directly apply Theorem 1.16. See Section 7.1 for complete details.

Independent sets in d-regular graphs. The largest independent set size (the *independence number*) in a random d-regular graph has been studied extensively. It is well-known that with high probability, the independence number is $\leqslant \frac{2n \log d}{d}$ for a sufficiently large constant d (cf. [Bol81, Wor99]). The current best known *certifiable* upper bound is via the smallest eigenvalue of the adjacency matrix (often referred to as Hoffman's bound, cf. [FO05, BH11]): Let A be the adjacency matrix, and let $\lambda := -\lambda_{\min}(A)$. Then, $|S| \leqslant \frac{\lambda}{d+\lambda}n$ for all independent sets S. We give a proof for completeness in Section 7.2.

It is also well-established that $\lambda \leqslant 2\sqrt{d-1} + o(1)$ with high probability (see Theorem 2.23). Thus, we can certify that the independence number is at most $C_d n$ where $C_d := \frac{2\sqrt{d-1}}{d+2\sqrt{d-1}}$.

The natural question for us is to certify an upper bound on the number of independent sets larger than $C_d(1 - \eta)n$ for some $\eta > 0$.

Theorem 1.18. *For a random d-regular graph on n vertices, given $\eta \in (0, \eta_0)$ for some universal constant η_0, there is an algorithm certifying that there are at most $2^{O(\eta^{3/5} \log \frac{1}{\eta})n}$ independent sets of size $C_d(1 - \eta)n$.*

The proof is very similar to the SK model. We first map each independent set S to a vector $y_S \in \mathbb{R}^n$ such that if S is large, then y_S is close to the bottom eigenspace of A. Then, using a variant of Theorem 1.16, we upper bound the number of such vectors that are close to the eigenspace. We carry out the proof in full detail in Section 7.2.

Optimality for counting kCSP solutions. Finally, we give evidence suggesting that our algorithmic upper bounds are close to optimal. Our hardness results are built on the hypothesis that there is no efficient *strong* refutation algorithm for random kXOR under the refutation threshold (in the regime $n^\varepsilon \ll \Delta \ll n^{k/2-1}$). Although no NP-hardness results are known, this hypothesis is widely believed to be true. In particular, the problem was shown to be hard for the Sum-of-Squares semidefinite programming hierarchy [Gri01, Sch08, KMOW17], which is known to capture most algorithmic techniques for average-case problems. Thus, improving our results would imply a significant breakthrough.

We show that assuming this hypothesis is true, then we cannot certify an upper bound on the number of $(1 - \eta)$-satisfying assignments better than $\exp(O(\eta n))$.

Theorem 1.19. *If there is an efficient algorithm that with high probability can certify a bound of $\exp\left(\frac{\eta n}{10k}\right)$ on the number of $(1 - \eta)$-satisfying assignments to $\mathcal{I}$, then there is an efficient algorithm that with high probability can certify that $\mathcal{I}$ has no $(1 - \eta/2)$-satisfying assignments.*

This shows that the term $\exp(\widetilde{O}(\eta n))$ in Theorem 1.7 and Theorem 1.10 is tight up to log factors (see Remark 8.2). Our proof is simple: given a $(1 - \eta/2)$-satisfying assignment and a small set S, we can flip the assignments to S arbitrarily and still be $(1 - \eta)$-satisfying. Hence the number of

$(1 - \eta)$-satisfying assignments is at least $2^{|S|}$. Thus, an upper bound better than this would imply that there is no $(1 - \eta/2)$-satisfying assignments. See Section 8.1 for complete details.

Surprisingly, the optimality of Theorem 1.7 suggests that there is a phase transition for certifiable counting at the refutation threshold. For concreteness, take random kSAT for example,

Remark 1.20. At $m = \widetilde{\Omega}(n^{k/2})$, there is a strong refutation algorithm [AOW15] which certifies that no $(1 - \eta)$-satisfying assignment exists (even for constant $\eta < 1/2$). However, at $m = n^{k/2 - \varepsilon}$ and take $\eta = n^{-\frac{1}{4} + \frac{\varepsilon}{2}}$, we can at best certify that the number of $(1 - \eta)$-satisfying assignments is at most $\exp(O(n^{\frac{3}{4} + \frac{\varepsilon}{2}}))$. See also Figure 1 for illustration.

Optimality for counting independent sets. We also show barriers to improving Theorem 1.18, which can be viewed as a weak hardness evidence. Specifically, we show that improving the upper bound of Theorem 1.18 to $\exp\left(O(\eta \log(1/\eta)n)\right)$ would imply beating Hoffman's bound by a factor of $1 - \eta/2$ (for any small constant η), which would be an interesting algorithmic breakthrough.

Theorem 1.21. *Let G be a random d-regular graph. Given constant $\eta \in (0, 1/2)$, if there is an efficient algorithm that with high probability certifies a bound of $\exp\left(\frac{C_d}{4}\eta \log(1/\eta)n\right)$ on the number of independent sets of size $C_d(1 - \eta)n$, then there is an algorithm that with high probability certifies that G has no independent set of size $(1 - \eta/2)C_d n$.*

The proof is a simple observation that for any independent set S, all subsets of S are also independent sets. Thus, if S is of size $(1 - \eta/2)C_d n$, then we can lower bound the number of subsets of size $(1 - \eta)C_d n$. We give a short proof in Section 8.2. We note the interesting gap between $\eta^{3/5}$ and η in the exponent of the upper and lower bounds respectively, and we conjecture that there may be an algorithm matching the lower bound.

1.2 Context and related work

Information-computation gaps in CSPs. This work is very closely related to the line of work on information-computation gaps. In the context of certification in random CSPs, the most well-understood information-gaps are in that of refutation of random CSPs. Feige's random 3SAT hypothesis was one of the earliest conjectured gaps. As discussed earlier, while unsatisfiability for random 3SAT set in at constant density, it was conjectured by Feige that certifying this was hard at all constant densities. Further, integrality gaps for the Sum-of-Squares hierarchy of [Gri01, Sch08] seem to point to hardness up to density $\sqrt{n}$. The wide information-computation gap is a main motivation for us to understand what an efficient algorithm can certify about the landscape of solutions in the regime between the satisfiability threshold and the refutation threshold. We refer the reader to the introduction of [AOW15] for a comprehensive treatment of the literature on information-computation gaps for refuting random CSPs prior to their work, CSPs more broadly, as well as connections to other areas of theoretical computer science.

The situation for general constraint satisfaction problems beyond XOR and SAT was considered in the work of [AOW15], which gave algorithms to refute all CSPs at density $n^{t/2-1}$ where t is the smallest integer such that there is no t-wise uniform distribution supported on the predicate's satisfying assignments. Then somewhat surprisingly, the work of [RRS17] gave algorithms for refuting random CSPs between constant density and the $n^{t/2-1}$ threshold from [AOW15], whose

running time smoothly interpolated between exponential time at constant density to polynomial time at the [AOW15] threshold, with a (steadily improving) subexponential running time in the intermediate regime. The algorithms of [AOW15, RRS17] are spectral, and can be captured within the Sum-of-Squares hierarchy. Finally the work of [KMOW17] (presaged by [BCK15]) established that the algorithm of [RRS17] was tight for Sum-of-Squares in all regimes, thereby nailing a characterization for the exact gaps (up to logarithmic factors) for all random CSPs.

Solution geometry in random CSPs. One of the earlier predictions using nonrigorous physics techniques was the location of the 3SAT satisfiability threshold in the works of [MZ02, MPZ02]. In particular, they conjectured that there is a sharp threshold at a constant $\alpha_{\text{SAT}} \approx 4.267$. These works put forth the "1-step replica symmetry breaking hypothesis" (a conjectured property of the solution space in random kSAT; we refer the reader to the introduction of [DSS15] for a description), which was the starting point for several subsequent works. These techniques were used to precisely predict the kSAT satisfiability threshold for all values of k [MMZ06], proved for large k in a line of work culminating in [DSS15] and building on [AM02, AP03, COP13, COP16].

Eventually, the works of [KMRT$^+$07, MRTS08] predicted that besides the satisfiability threshold, random kSAT goes through other phase transitions too, and gave conjectures for their locations. A notable one connected to this work is the *clustering threshold*, for which there has been rigorous evidence given in the works of [MMZ05, ART06, ACO08]. Above the clustering threshold, the solution space is predicted to break into exponentially many exponential-sized clusters far away from each other in Hamming distance. More precisely, there is some function Σ for which there are $\exp(\Sigma(s, \Delta)n)$ clusters of size approximately $\exp(sn)$ each. In particular, this leads to the prediction that the number of solutions at density Δ is roughly $\max_s\{\exp((s + \Sigma(s, \Delta))n)\}$. Another phase transition of interest is the *condensation threshold*, where the number of clusters of solutions drops to a constant.

Approximate Counting for CSPs. Approximate counting of solutions in CSPs has attracted much attention in recent years. There have been numerous positive algorithmic results for approximately counting solutions in (i) sparse CSPs in the worst case, (ii) sparse random CSPs well under the satisfiability threshold. The takeaway here is that even though the problems we consider get harder as we approach the satisfiability threshold, if one goes well under the threshold the algorithmic problems once again become tractable.

One exciting line of research for worst-case CSPs is the problem of approximately counting satisfying assignments of a kSAT formula under conditions similar to those of the Lovász Local Lemma (LLL) [EL73]. A direct application of the LLL shows that if the maximum degree D of the *dependency graph* is $\leqslant 2^k/e$, then the formula is satisfiable. Building on works of [Moi19, FGYZ20, FHY20, JPV20], Jain, Pham, and Vuong [JPV21] recently showed that there is an algorithm for approximate counting well under the LLL thresholds, i.e. when $D \lesssim 2^{k/5.741}$ (hiding factors polynomial in k), using techniques similar to an algorithmic version of the LLL. Further, the algorithms of [Moi19, JPV20] are deterministic, which may suggest their techniques are amenable to obtaining certifiable counts. However, it was shown that the problem of approximately counting solutions to a kSAT formula is NP-hard when $D \gtrsim 2^{k/2}$ by [BGG$^+$19], well in the sparse regime, which suggests a hard phase between the highly sparse setting and the dense setting we are concerned with.

For random kSAT, the exact satisfiability threshold that was established by Ding, Sly, and Sun [DSS15] takes on value $\alpha_{\text{SAT}} = 2^k \ln 2 - \frac{1}{2}(1 + \ln 2) + o_k(1)$. And similarly, well below the satisfiability threshold, Galanis, Goldberg, Guo, and Yang [GGGY19] adapted Moitra's techniques [Moi19] to the random setting and developed a polynomial-time algorithm when the density $\Delta \leqslant 2^{k/301}$ and k sufficiently large.

Closely related to the counting problem is approximating the partition function of random kSAT, for which there have also been positive algorithmic results. Specifically, given a random kSAT instance $\mathcal{I}$, the partition function is defined as $Z(\mathcal{I}, \beta) := \sum_\sigma e^{-\beta H(\sigma)}$, where $H(\sigma)$ is the number of unsatisfied clauses under assignment σ. The partition function can be viewed as a weighted (or "permissive") version of the counting problem. Montanari and Shah [MS06] first showed that the Belief Propagation algorithm approximately computes the partition function at $\Delta \sim \frac{2 \log k}{k}$; their analysis is based on correlation decay (or the *Gibbs uniqueness property*). Recently, [COMR20] further showed that Belief Propagation succeeds as long as the random kSAT model satisfies a *replica symmetry* condition, conjectured to hold up to $\Delta \sim 2^k \ln k / k$. See also the works of [KMRT+07, Pan13, CO17] for further details of this matter.

Counting independent sets and related problems. Another counting problem that has been the subject of active study is that of counting independent sets, especially in the statistical physics community. For a graph G with maximum degree d, let $\mathsf{IS}(G)$ be the set of independent sets in G. The task is to estimate the *independence polynomial* $Z_G(\lambda) = \sum_{I \in \mathsf{IS}(G)} \lambda^{|I|}$, also known as the partition function of the *hard-core model* with *fugacity* λ in the physics literature. Earlier works by [DG00, Vig01] developed randomized algorithms based on *Glauber dynamics* to estimate $Z_G(\lambda)$ when $\lambda \leqslant \frac{2}{d-2}$. In a major breakthrough, Weitz [Wei06] showed a deterministic algorithm, based on correlation decay, that approximates $Z_G(\lambda)$ when $0 \leqslant \lambda < \lambda_c$, where $\lambda_c := \frac{(d-1)^{d-1}}{(d-2)^d}$. Sly and Sun [SS12] later proved that this is tight: no efficient approximate algorithm for $Z_G(\lambda)$ exists for $\lambda > \lambda_c$ unless $\mathsf{NP} = \mathsf{RP}$.

Recently, Barvinok initiated a line of research on estimating partition functions using the *interpolation method* (see Barvinok's recent book [Bar16]). The main idea is to estimate the low-order Taylor approximation of $\log Z_G(\lambda)$ provided that the polynomial $Z_G(\lambda)$ does not vanish in some region in $\mathbb{C}$. This approach led to deterministic algorithms that match Weitz's result and work even for negative or complex λ's [PR17, PR19]. These polynomial-based approaches were also used to obtain deterministic algorithms for counting colorings in bounded degree graphs [LSS19a], estimating the Ising model partition function [LSS19b], and algorithms for a counting version of the Unique Games problem [CDK+19].

There has also been works on worst-case upper bounds of $Z_G(\lambda)$ for d-regular graphs. Zhao proved that for any d-regular graph G and any $\lambda \geqslant 0$, $Z_G(\lambda) \leqslant (2(1 + \lambda)^d - 1)^{n/2d}$ [Zha10]. In particular, setting $\lambda = 1$, this shows that the total number of independent sets is bounded by $(2^{d+1} - 1)^{n/2d}$, settling a conjecture by Alon [Alo91] and Kahn [Kah01].

Certifying bounds on partition functions and free energy. A recent line of work [Ris16, RL16, JKR19] is focused on an approach based on a convex programming relaxation of entropy to certify upper bounds on the *free energy* of the Ising model (weighted 2XOR), both in the worst case and in the average case. While on the surface level, these approaches differ significantly from ours, an

interesting direction is to investigate if these entropy-based convex programming relaxations can achieve our algorithmic results.

1.3 Table of results

We include a table to have a succinct snapshot of our results.

	Problem	Theorem	Upper bound	Randomness
Counts	2XOR	4.1	$\exp(\widetilde{O}(\eta n))$	Hypergraph
	kXOR	4.4	$\exp(\widetilde{O}(\eta n)) \cdot \exp\left(O\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right)$	Hypergraph
	kCSP	4.8	$\exp(\widetilde{O}(\eta n)) \cdot \exp\left(O\left(\frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}\right)\right)$ $\cdot \exp\left(\widetilde{O}\left(\sqrt{\frac{n^{(k+1)/2}}{\Delta}}\right)\right)$	Hypergraph + signings
	SK model	7.5	$\exp(O(\eta^{3/5}\log\frac{1}{\eta})n)$	$M \sim \mathsf{GOE}(n)$
	Independent set	7.7	$\exp(O(\eta^{3/5}\log\frac{1}{\eta})n)$	d-regular graph
Clusters	3XOR	5.1	$\exp(O(\theta^2\log(1/\theta))n)$, for $\theta = \max(2\eta, \Delta^{-\frac{1}{2}}\log n)$	Hypergraph
	3CSP	5.2	$\exp(O(\theta^2\log(1/\theta))n)$, for $\theta = 8\eta + \widetilde{O}(\Delta^{-1/2})$	Hypergraph + signings
Balance	3CSP	6.1	bias $\leqslant \rho$ for $\rho \gg \sqrt{\frac{\log n}{\Delta}}$, $\eta = \rho/16$	Hypergraph + signings
	kCSP	6.4	bias $\leqslant \rho$ for any constant $\rho > 0$, $\eta = \rho^k/2^{k+1}$	Hypergraph + signings

Table 1: A summary of our results.
(1) **kCSP counts**: given $\mathcal{I} \sim \mathcal{H}_k^n(m)$ where $m = \Delta n$, we upper bound the number of $(1 - \eta)$-satisfying assignments.
(2) **SK counts**: given $M \sim \mathsf{GOE}(n)$, we upper bound the number of $x \in \{\pm 1\}^n$ such that $x^\top M x \geqslant 2(1 - \eta)n^{3/2}$.
(3) **Independent set counts**: given a random d-regular graph for constant $d \geqslant 3$, we upper bound the number of independent sets of size $\geqslant C_d n(1 - \eta)$.
(4) **3CSP clusters**: given $\mathcal{I} \sim \mathcal{H}_3^n(m)$ where $m = \Delta n$, we upper bound the number of diameter-(θn) clusters of $(1 - \eta)$-satisfying assignments.
(5) **Balance**: given $\mathcal{I} \sim \mathcal{H}_k^n(m)$ where $m = \Delta n$, we certify that any $(1 - \eta)$-satisfying assignment must have *bias* $\frac{1}{n}\left|\sum_{i\in[n]} x_i\right| \leqslant \rho$.

13

1.4 Open directions

In this section we suggest a couple of avenues for further investigation on the themes related to this work.

Worst-case complexity of certified counting.   In this work, we deal mostly with random CSPs. Here we present a worst-case version of the problem, specialized to 3SAT. A classic result due to [Hås01] is that it is NP-hard to distinguish between a $(7/8 + \varepsilon)$-satisfiable 3SAT formula from a fully satisfiable 3SAT formula. However, it is unclear what the complexity of a version of this question is when there is a stronger promise on the satisfiable 3SAT formula.

Question 1.22. Consider the following algorithmic task:

> Given a 3SAT formula $\mathcal{I}$ under the promise that it is either $(7/8 + \varepsilon)$-satisfiable, or has at least T fully satisfying assignments, decide which of the two categories $\mathcal{I}$ falls into.

What is the complexity of the above problem?

We remark that this problem is similar to counting-3SAT, but subtly different.

Certifying optimal bounds on number of exactly satisfying kSAT solutions.   In the context of kSAT, while our algorithms can certify subexponential bounds for both exactly satisfying assignments and approximately satisfying assignments, the matching evidence of hardness is only for the approximate version of the problem. Thus, it is still possible that there is an algorithm to certify an even tighter bound than ours for the problem of counting exactly satisfying assignments to a random kSAT formula. This motivates the following question:

Question 1.23. What is the tightest bound an efficient algorithm can certify on the number of solutions to a random kSAT instance?

We conjecture that the algorithms presented in this paper are indeed optimal. An approach to providing hardness evidence for this is to construct a hard planted distribution, and prove it is hard within the *low-degree likelihood ratio* framework of [HS17]. We outline a possible approach in Section 8.3 to construct a planted distribution for readers interested in this problem.

Properties of arbitrary CSP instances on random hypergraphs.   In the context of approximate kXOR, our certification algorithms for solution counts and cluster counts depend only on the hypergraph structure and not the random negations. Hence, they also prove nontrivial statements about the solution space of any XOR instance on a random hypergraph, which are potentially useful in the context of quiet planting or semi-random models of CSPs. However, our certification algorithms for other CSPs, such as kSAT, heavily make use of the random signings in the reduction to kXOR.

Question 1.24. Can all the results related to certifying bounds on number of solutions/clusters in this work for random kSAT instances be generalized to arbitrary kSAT instances on random hypergraphs?

14

2 Preliminaries

2.1 Graph theory

Given a graph G, we use $V(G)$ to denote its vertex set, $E(G)$ to denote its edge set, and $\deg_G(u)$ to denote the degree of a vertex u. For $S \subseteq V(G)$ and $T \subseteq V(G)$, we use $e(S,T)$ to denote the number of tuples (u,v) such that $u \in S$, $v \in T$ and $\{u,v\} \in E(G)$.

We will be concerned with its *normalized Laplacian matrix*, denoted L_G, defined as:

$$L_G := \mathbb{1} - D_G^{-1/2} A_G D_G^{-1/2},$$

where A_G is the adjacency matrix of G and D_G is the diagonal matrix of vertex degrees in G. Since L_G is a self-adjoint matrix, it has n real eigenvalues, which we sort in increasing order and denote as:

$$0 = \lambda_1(G) \leqslant \lambda_2(G) \leqslant \ldots \leqslant \lambda_n(G).$$

Of particular interest to us is $\lambda_2(G)$, which we call the *spectral gap*.

A combinatorial quantity we will be concerned with is the *conductance* of G. For a subset $S \subseteq V$, we define the *volume* of S as $\mathrm{vol}(S) := \sum_{u \in S} \deg(u)$ and let $\phi_G(S) := \frac{e(S,\bar{S})}{\mathrm{vol}(S)}$. The conductance of G is then defined as:

$$\phi_G := \min_{\substack{S \subseteq V(G) \\ \mathrm{vol}(S) \leqslant \mathrm{vol}(V)/2}} \phi_G(S).$$

The well-known Cheeger's inequality on graphs, first proved in [AM85], relates the conductance and the spectral gap. We refer the reader to [Tre17] for a good exposition of the proof.

Theorem 2.1 (Cheeger's inequality). *For any graph G,*

$$\frac{\lambda_2(G)}{2} \leqslant \phi_G \leqslant \sqrt{2\lambda_2(G)}.$$

It is well-known that dense Erdős-Rényi random graphs have large spectral gaps (cf. [CO07, HKP19]).

Theorem 2.2 ([HKP19, Theorem 1.1]). *Let G be an Erdős-Rényi random graph with $p = \omega\left(\frac{\log n}{n}\right)$, and let $d = p(n-1)$ denote the average degree. Then, there is a constant C such that*

$$\lambda_2(G) \geqslant 1 - \frac{C}{\sqrt{d}}$$

with probability at least $1 - Cn\exp(-d) - C\exp(-d^{1/4}\log n)$.

Closely related to Cheeger's inequality is the *expander mixing lemma*, which roughly states that the edges of an expander graph are *well distributed*. Here, we consider the adjacency matrix A and the "de-meaned" matrix $\bar{A} := A - \frac{d}{n}J$, where J is the all-ones matrix. We include a short proof for completeness.

Theorem 2.3 (Expander Mixing Lemma [AC88]). *Let G be a graph with n vertices and average degree d, and let $\bar{A}$ be the de-meaned adjacency matrix. Then, for any $S, T \subseteq V(G)$,*

$$\left| e(S,T) - \frac{d}{n}|S| \cdot |T| \right| \leqslant \|\bar{A}\|\sqrt{|S| \cdot |T|}.$$

Proof. Let $1_S, 1_T \in \{0,1\}^n$ be the indicator vectors for subsets S, T. Clearly, we have $1_S^\top A 1_T = e(S, T)$ and $1_S^\top J 1_T = |S| \cdot |T|$. Moreover, $\|1_S\|_2 = \sqrt{|S|}$ and $\|1_T\|_2 = \sqrt{|T|}$. Thus,

$$e(S,T) - \frac{d}{n}|S| \cdot |T| = 1_S^\top \left(A - \frac{d}{n}J \right) 1_T$$

$$\Rightarrow \left| e(S,T) - \frac{d}{n}|S| \cdot |T| \right| \leqslant \|\overline{A}\| \cdot \|1_S\|_2 \cdot \|1_T\|_2. \qquad \square$$

2.2 Fourier analysis of Boolean functions

We refer the reader to [O'D14] for an elaborate treatment of the subject. The functions $\{\prod_{i \in T} x_i\}_{T \subseteq [k]}$ form an orthogonal basis for the space of functions from $\{\pm 1\}^k$ to $\mathbb{R}$, and hence any function $f : \{\pm 1\}^k \to \mathbb{R}$ can be expressed as a multilinear polynomial:

$$f(x) = \sum_{T \subseteq [n]} \widehat{f}(T) \prod_{i \in T} x_i.$$

Further, the coefficients $\widehat{f}(T)$, which are called *Fourier coefficients*, can be obtained via the formula:

$$\widehat{f}(T) = \mathop{\mathbf{E}}_{z \sim \{\pm 1\}^k} \left[f(z) \prod_{i \in T} z_i \right].$$

A key property, called *Plancherel's theorem*, is the following:

Fact 2.4. *Let f, g be functions from $\{\pm 1\}^k$ to $\mathbb{R}$. Then:*

$$\langle f, g \rangle := \mathop{\mathbf{E}}_{z \sim \{\pm 1\}^k} f(z)g(z) = \sum_{T \subseteq [k]} \widehat{f}(T)\widehat{g}(T).$$

Given a probability distribution $\mathcal{D}$ on $\{\pm 1\}^k$, following the notation of [AOW15], we use D to denote the function equal to its probability density function multiplied by 2^k. This leads to the notational convenience that for any function $f : \{\pm 1\}^k \to \mathbb{R}$,

$$\mathop{\mathbf{E}}_{z \sim \mathcal{D}} f(z) = \mathop{\mathbf{E}}_{z \sim \{\pm 1\}^k} f(z)D(z).$$

2.3 Random hypergraphs and CSPs

In this work, we will deal with random CSPs defined as signed k-uniform hypergraphs.

Definition 2.5 (Signed k-uniform hypergraph). A *signed k-uniform hypergraph* $\mathcal{I}$ on universe $[n]$ is a collection of pairs $\{(c, U)\}$ where for each (c, U), $c \in \{\pm 1\}^k$ and U is in $[n]^k$. Each pair $(c, U) \in \mathcal{I}$ is also called a *hyperedge* in $\mathcal{I}$.

The distribution over signed k-uniform hypergraphs we work with is:

Definition 2.6. We use $\mathcal{H}_k^n(m)$ to denote the distribution on signed k-uniform hypergraphs on universe $[n]$ where a sample $\mathcal{I} \sim \mathcal{H}_k^n(m)$ is obtained by independently including each of the $2^k n^k$ potential hyperedges with probability $\frac{m}{2^k n^k}$. Moreover, we use $\mathcal{H}_{k,+}^n(m)$ to denote the distribution of (unsigned) k-uniform hypergraphs, which is the same as $\mathcal{H}_k^n(m)$ but with the signs removed.

16

Given a tuple $U \in [n]^k$ and $x \in \{\pm 1\}^n$, we use x_U to denote the tuple $(x_{U[1]}, \dots, x_{U[k]})$. And for a pair of k-tuples a, b we use $a \circ b$ to denote the entrywise product of the tuples $(a_1 \cdot b_1, \dots, a_k \cdot b_k)$.

Definition 2.7. A *constraint satisfaction problem instance* (CSP instance) on variable set $[n]$ is a signed hypergraph $\mathcal{I}$ along with a function $P : \{\pm 1\}^k \to \mathbb{R}$, called a *predicate*. Given $x \in \{\pm 1\}^n$, we use $P_{\mathcal{I}}(x)$ to denote the *objective value* on x:

$$P_{\mathcal{I}}(x) := \frac{1}{|\mathcal{I}|} \sum_{(c,U) \in \mathcal{I}} P(c \circ x_U).$$

Definition 2.8. Given a signed k-uniform hypergraph $\mathcal{I}$ on $[n]$ and any $x \in \{\pm 1\}^n$, we define $\mathcal{D}_{\mathcal{I},x}$ to be the distribution on $\{\pm 1\}^k$ with density (scaled by 2^k):

$$D_{\mathcal{I},x}(z) := \frac{2^k}{|\mathcal{I}|} \cdot |\{(c,U) \in \mathcal{I} | c \circ x_U = z\}| .$$

Note that $\widehat{D_{\mathcal{I},x}}(\varnothing) = \mathbf{E}_{z \sim \{\pm 1\}^k}[D_{\mathcal{I},x}(z)] = 1$ for all $x \in \{\pm 1\}^n$. A simple but important observation is the following.

Observation 2.9. If $\mathcal{I}$ is a signed k-uniform hypergraph and $P : \{\pm 1\}^k \to \mathbb{R}$ is some predicate, then:

$$
\begin{aligned}
P_{\mathcal{I}}(x) &= \frac{1}{|\mathcal{I}|} \sum_{(c,U) \in \mathcal{I}} P(c \circ x_U) \\
&= \mathop{\mathbf{E}}_{z \sim \{\pm 1\}^k} [P(z) D_{\mathcal{I},x}(z)] \\
&= \sum_{T \subseteq [k]} \widehat{P}(T) \widehat{D_{\mathcal{I},x}}(T).
\end{aligned}
$$

Definition 2.10 (Biased assignment). We say $x \in \{\pm 1\}^n$ is η-*biased* if $\frac{1}{n} |\sum_{i=1}^n x_i| \geqslant \eta$.

Next, we focus on kXOR. Given a pair $(c, U) \in \mathcal{I}$, $P(c \circ x_U) = 1$ if and only if $\prod_{i=1}^k x_{U[i]} = \prod_{i=1}^k c_i$, thus in the case of kXOR we will also write (b, U) as a constraint (or clause) of $\mathcal{I}$, where $b = \prod_{i=1}^k c_i \in \{\pm 1\}$ is the *signing* of the constraint. We call a constraint (b, U) *positive* if $b = +1$ and *negative* if $b = -1$. We say an instance $\mathcal{I}$ is p-*positive* if the fraction of its constraints that are positive is at most p. Moreover, we will need the notion of *induced* XOR and *truncated* XOR, which we define below.

Definition 2.11 (Induced tXOR and truncated $(k-t)$XOR). Given a kXOR instance $\mathcal{I}$, an integer $1 \leqslant t \leqslant k - 1$, a subset of variables $S \subseteq [n]$, and an assignment $\sigma \in \{\pm 1\}^S$, we define the *induced* tXOR *instance* $\mathcal{I}_{S,\sigma,t}$ on variables $\overline{S}$ as follows: for each clause $(b, U) \in \mathcal{I}$ where all variables in $U[1 : k - t]$ are in S and all variables in $U[k - t + 1 : k]$ are in $\overline{S}$, add a tXOR clause $(b', U[k - t + 1 : k])$ where $b' = b \cdot \prod_{i=1}^{k-t} \sigma_i$. Similarly, we define the *truncated* $(k-t)$XOR *instance* $\mathcal{I}|_{S,k-t}$ on variables S as follows: for each clause $(b, U) \in \mathcal{I}$ where all variables in $U[1 : k - t]$ are in S and all variables in $U[k - t + 1 : k]$ are in $\overline{S}$, add a $(k - t)$XOR clause $(b, U[1 : k - t])$.

For example, consider a 4XOR instance and a constraint $x_a x_b x_c x_d = +1$. Suppose $a, b \in S$, $c, d \notin S$, and $\sigma_a = +1, \sigma_b = -1$. Then for $t = 2$, the induced instance $\mathcal{I}_{S,\sigma,2}$ will have a constraint $x_c x_d = -1$, and the truncated instance $\mathcal{I}|_{S,2}$ will have a constraint $x_a x_b = +1$. Note that the truncated instance does not depend on the assignments to S.

Observation 2.12. Given a kXOR instance $\mathcal{I}$, subset $S \subseteq [n]$ and $\sigma \in \{\pm 1\}^S$, the following are useful relationships between the induced tXOR instance $\mathcal{I}_{S,\sigma,t}$ and the truncated $(k - t)$XOR instance $\mathcal{I}|_{S,k-t}$:

1. $|\mathcal{I}_{S,\sigma,t}| = \left| \mathcal{I}|_{S,k-t} \right|$.

2. $\displaystyle\sum_{(b,U)\in\mathcal{I}_{S,\sigma,t}} b = \sum_{(b,U)\in\mathcal{I}|_{S,k-t}} b \prod_{i \in U} \sigma_i$.

Definition 2.13 (Induced hypergraph). Given a kXOR instance $\mathcal{I}$, an integer $1 \leqslant t \leqslant k - 1$, and a subset of variables $S \subseteq [n]$, we define the *induced t-uniform hypergraph $H_{S,t}$* on $n - |S|$ variables as the underlying t-uniform hypergraph of the induced tXOR of S. Note that the hypergraph only depends on S, t and not the assignment σ to S.

For simplicity, we denote $\mathcal{I}_{S,\sigma}$ to be $\mathcal{I}_{S,\sigma,k-1}$ and H_S to be $H_{S,k-1}$.

Definition 2.14 (Primal graph). Given a hypergraph H, we define the *primal graph* of H on the same vertex set, denoted H_P, as follows: for every hyperedge $e \in H$ and every pair $(u, v) \in e$, add (u, v) to H_P. Parallel edges are allowed.

2.4 Refuting random CSPs

We will need the refutation algorithm for random CSPs of [AOW15]. A crucial notion in [AOW15] is that of approximate t-wise uniformity.

Definition 2.15 ((ε, t)-wise uniform). A distribution $\mathcal{D}$ is (ε, t)-wise uniform if for all $T \subseteq [k]$ such that $0 < |T| \leqslant t$, $|\widehat{D}(T)| \leqslant \varepsilon$.

Definition 2.16. We say a signed k-uniform hypergraph $\mathcal{I}$ on $[n]$ is (ε, t)-quasirandom if for every $x \in \{\pm 1\}^n$ the distribution $\mathcal{D}_{\mathcal{I},x}$ is (ε, t)-wise uniform.

Now we are ready to state the key statement we use from [AOW15].

Theorem 2.17. *Let $\mathcal{I} \sim \mathcal{H}_k^n(m)$ with $m \geqslant \alpha n^{t/2}$. Then there is an efficient algorithm that with probability $1 - o(1)$ certifies that $\mathcal{I}$ is $\left(\frac{2^{O(t)} \log^{5/2} n}{\sqrt{\alpha}}, t \right)$-quasirandom.*

Another statement we use from [AOW15] is their algorithm to refute random polynomials on the hypercube.

Theorem 2.18. *For $k \geqslant 2$, let $\{w_T\}_{T \in [n]^k}$ be independent centered random variables on $[-1, 1]$ such that:*

$$\mathbf{Pr}[w_T \neq 0] \leqslant p \quad \forall T \in [n]^k.$$

Then there is an efficient algorithm which certifies with high probability:

$$\sum_{T \in [n]^k} w_T x^T \leqslant 2^{O(k)} \max\{\sqrt{p}n^{3k/4}, n^{k/2}\} \log^{3/2} n$$

for all $x \in \{\pm 1\}^n$.

2.5   Random matrix theory

We will need the matrix Bernstein inequality for proving spectral norm bounds as stated in [Tro15, Theorem 6.1.1].

Theorem 2.19 (Matrix Bernstein inequality (special case of [Tro15, Theorem 6.1.1])). *Let $S_1, \ldots, S_\ell$ be a collection of independent random symmetric matrices of dimension $d \times d$. Assume that $\mathbf{E}\, S_i = 0$ and $\|S_i\| \leqslant 1$ for all $i \in [\ell]$. Let $Z = \sum_{i \in [\ell]} S_i$. Define:*

$$v = \left\| \sum_{i \in [\ell]} \mathbf{E}[S_i^2] \right\|.$$

Then, for any $t \geqslant 0$,

$$\mathbf{Pr}[\|Z\| \geqslant t] \leqslant 2d \exp\left(\frac{-t^2}{v + t/3} \right).$$

We will also need the eigenvalue distribution of random matrices (in Section 7). We first consider Gaussian matrices. Let W be a random $n \times n$ matrix with independent standard Gaussian entries, and let $M := \frac{1}{\sqrt{2}}(W + W^\top)$. We say that M is sampled from the Gaussian Orthogonal Ensemble, denoted $\mathrm{GOE}(n)$. We recall the following results in random matrix theory,

Fact 2.20. *The spectral norm $\lambda_{\max}(M) \leqslant (2 + t)\sqrt{n}$ with probability $1 - 2\exp(-nt^2/2)$.*

Theorem 2.21 (Semicircle law [Erd11]). *The empirical distribution of eigenvalues of $M \sim \mathrm{GOE}(n)$ follows a universal pattern, the Wigner semicircle law. For any fixed real numbers $a < b$, with probability $1 - o(1)$,*

$$\frac{1}{n} \left| \left\{ i : \frac{1}{\sqrt{n}} \lambda_i(M) \in [a, b] \right\} \right| = (1 \pm o(1)) \int_a^b \rho_{sc}(x)dx, \quad \rho_{sc}(x) := \frac{1}{2\pi} \sqrt{(4 - x^2)_+},$$

where $(a)_+ = \max(a, 0)$.

Lemma 2.22. *For $M \sim \mathrm{GOE}(n)$ and any $\varepsilon \in (0, 2)$, with probability $1 - o(1)$,*

$$\frac{1}{n} \left| \left\{ i : \lambda_i(M) \geqslant (2 - \varepsilon)\sqrt{n} \right\} \right| \leqslant \frac{\varepsilon^{3/2}}{\pi} (1 \pm o(1)).$$

Proof. We apply Theorem 2.21 directly. The area under the semicircle between $2 - \varepsilon$ and 2 can be upper bounded by a rectangle of width ε and height $\sqrt{4\varepsilon - \varepsilon^2} \leqslant 2\sqrt{\varepsilon}$. Dividing by 2π completes the proof. $\qquad\square$

Next, we look at the adjacency matrix A of a random d-regular graph. It is a standard result that the largest eigenvalue of A is d, and the all-ones vector $\vec{1}$ is the top eigenvector. Thus, it is common to look at the "de-meaned" matrix $\overline{A} := A - \frac{d}{n}J$ (removing the top eigenvector component). The eigenvalue bounds for $\overline{A}$ were conjectured by Alon [Alo86] and later proved by Friedman [Fri08]. The following quantitative statement is by Bordenave [Bor15].

Theorem 2.23. *Let $d \geqslant 3$ be an integer and let A be the adjacency matrix of a random d-regular graph on n vertices. With probability $1 - \frac{1}{\mathrm{poly}(n)}$, all eigenvalues of $A - \frac{d}{n}J$ lie within $[-2\sqrt{d-1} - \varepsilon_n, 2\sqrt{d-1} + \varepsilon_n]$, where $\varepsilon_n = c \left(\frac{\log \log n}{\log n} \right)^2$ for some constant c.*

For constant d (fixed as n grows), the empirical eigenvalue distribution is given by the *Kesten–McKay law* [Kes59, McK81]. The eigenvalues of A and $A - \frac{d}{n}J$ interlace by Cauchy's interlacing theorem, and hence the limiting eigenvalue distributions are the same for both matrices.

Theorem 2.24 (Kesten–McKay law [McK81]). *Let $d \geqslant 3$ be a fixed integer. Let A be the adjacency matrix of a random d-regular graph, and let $\overline{A} = A - \frac{d}{n}J$. For any fixed real numbers $a < b$, with probability $1 - o(1)$,*

$$\frac{1}{n}\left|\{i : \lambda_i(\overline{A}) \in [a,b]\}\right| = (1 \pm o(1))\int_a^b \rho_d(x)dx,$$

$$\rho_d(x) := \frac{d\sqrt{4(d-1) - x^2}}{2\pi(d^2 - x^2)} \quad \text{for } |x| \leqslant 2\sqrt{d-1}.$$

Lemma 2.25. *Let $d \geqslant 3$ be a fixed integer. Let A be the adjacency matrix of a random d-regular graph, and let $\overline{A} = A - \frac{d}{n}J$. For any $\varepsilon \in (0,1)$, with probability $1 - o(1)$,*

$$\frac{1}{n}\left|\{i : \lambda_i(\overline{A}) \leqslant -2\sqrt{d-1}(1-\varepsilon)\}\right| \leqslant \frac{12\sqrt{2}}{\pi}\varepsilon^{3/2}(1 \pm o(1)).$$

Proof. We apply Theorem 2.24 directly. The Kesten–McKay density ρ_d can be upper bounded by $\frac{d}{2\pi(d^2 - 4(d-1))}\sqrt{4(d-1) - x^2}$, a scaled semicircle of radius $2\sqrt{d-1}$. Let $x_\varepsilon = 2\sqrt{d-1}(1-\varepsilon)$. Then, $\rho_d(x_\varepsilon) \leqslant \frac{d\sqrt{d-1}}{\pi(d-2)^2}\sqrt{2\varepsilon}$. This upper bound is increasing with ε. Thus, we can bound the area under ρ_d between $2\sqrt{d-1}(1-\varepsilon)$ and $2\sqrt{d-1}$ by $\frac{2\sqrt{2}}{\pi} \cdot \frac{d(d-1)}{(d-2)^2}\varepsilon^{3/2}$. The term $\frac{d(d-1)}{(d-2)^2}$ is decreasing for $d \geqslant 3$ and the maximum is 6. This completes the proof. $\square$

3 The kXOR principle

A crucial ingredient in our algorithms is that we can efficiently certify with high probability that any assignment that approximately satisfies a random kSAT formula $\mathcal{I}$ must also approximately satisfy the formula as kXOR. This generalizes the $k = 3$ case that appears in [Fei02, FO07] under the name "3XOR-principle".

Concretely, we show:

Lemma 3.1. *Let $\mathcal{I}$ be a random kSAT formula on $m = \alpha n^{(k-1)/2}$ clauses. There is an algorithm that with high probability certifies:*

> *Any $(1 - \eta)$-satisfying assignment of $\mathcal{I}$ must kXOR-satisfy at least $(1 - 2^{k-1}\eta)m - 2^{O(k)}\sqrt{\alpha \log^5 n} \cdot n^{(k-1)/2}$ clauses.*

To prove Lemma 3.1, we will need the Fourier expansion of kSAT. Though this is well-known and simple to derive, we present a proof for completeness.

Claim 3.2. $k\mathrm{SAT}(x_1, \ldots, x_k) = 1 - 2^{-k}\sum_{T \subseteq [k]}\prod_{i \in T} x_i$.

Proof. First note that $k\mathrm{SAT}(x_1, \ldots, x_k) = 1 - \mathbf{1}_{x=\bar{1}}$. It remains to verify that all the Fourier coefficients of $\mathbf{1}_{x=\bar{1}}$ are equal to 2^{-k}. Indeed, for any $T \subseteq [k]$:

$$\widehat{\mathbf{1}}_{x=\bar{1}}(T) = \mathop{\mathbf{E}}_{x \sim \{\pm 1\}^k} \chi_T(x)\mathbf{1}_{x=\bar{1}}(x) = 2^{-k},$$

which completes the proof. $\square$

We are now ready to prove Lemma 3.1.

Proof of Lemma 3.1. If x is a $(1 - \eta)$-satisfying assignment for $\mathcal{I}$, then on one hand:

$$\frac{1}{|\mathcal{I}|} \sum_{(c,U) \in \mathcal{I}} k\mathsf{SAT}(c \circ x_U) \geqslant 1 - \eta. \tag{1}$$

On the other hand, by Theorem 2.17, with high probability we can certify that $\mathcal{I}$ is quasirandom, specifically $\mathcal{D}_{\mathcal{I},x}$ is approximately $(k-1)$-wise uniform for all $x \in \{\pm 1\}^n$ (recall Definition 2.15 and Definition 2.16). Thus, we can certify:

$$\begin{aligned}
\frac{1}{|\mathcal{I}|} \sum_{(c,U) \in \mathcal{I}} k\mathsf{SAT}(c \circ x_U) &= \mathop{\mathbf{E}}_{z \sim \{\pm 1\}^k} [k\mathsf{SAT}(z) D_{\mathcal{I},x}(z)] && \text{(by Observation 2.9)} \\
&= (1 - 2^{-k}) \widehat{D_{\mathcal{I},x}}(\varnothing) + 2^{-k} \sum_{\substack{T \subseteq [k] \\ T \neq \varnothing}} \widehat{D_{\mathcal{I},x}}(T) && \text{(by Claim 3.2)} \\
&= 1 - 2^{-k} + 2^{-k} \widehat{D_{\mathcal{I},x}}([k]) + 2^{-k} \sum_{\substack{T \subseteq [k] \\ 1 \leqslant |T| \leqslant k-1}} \widehat{D_{\mathcal{I},x}}(T) \\
&\leqslant 1 - 2^{-k} + 2^{-k} \widehat{D_{\mathcal{I},x}}([k]) + \frac{2^{O(k)} \log^{5/2} n}{\sqrt{\alpha}}. && \text{(by Theorem 2.17)}
\end{aligned} \tag{2}$$

If the certification of the above inequality fails, then we halt the algorithm and output `failure`.

Otherwise, (1) and (2) together tell us that:

$$1 - \eta \leqslant 1 - 2^{-k} + 2^{-k} \widehat{D_{\mathcal{I},x}}([k]) + \frac{2^{O(k)} \log^{5/2} n}{\sqrt{\alpha}},$$

which can be rearranged as:

$$\widehat{D_{\mathcal{I},x}}([k]) \geqslant 1 - 2^k \eta - \frac{2^{O(k)} \log^{5/2} n}{\sqrt{\alpha}}. \tag{3}$$

By Observation 2.9 and (3) the fraction of clauses of $\mathcal{I}$ that are $k\mathsf{XOR}$-satisfied by x is:

$$\frac{1 + \widehat{D_{\mathcal{I},x}}([k])}{2} \geqslant 1 - 2^{k-1} \eta - \frac{2^{O(k)} \log^{5/2} n}{\sqrt{\alpha}},$$

which completes the proof. □

4 Count Certification for kCSPs

Thanks to the $k\mathsf{XOR}$-principle, we can first focus on $k\mathsf{XOR}$. For any $k\mathsf{XOR}$ instance, we can calculate the number of exactly satisfying assignments by Gaussian elimination. However, the problem becomes non-trivial when we turn to the number of approximate solutions. A priori, it is unclear whether we can certify a bound better than a naive $2^{O(n)}$ bound. In this section, we will show an algorithm that certifies a subexponential upper bound when the underlying graph is random and sufficiently dense.

21

4.1 Count certification for 2XOR

As a warmup, we start with 2XOR.

Theorem 4.1. *Let $G \sim \mathcal{H}^n_{2,+}(m)$ be a random graph where $m = \Delta n$ and $\Delta = n^\delta$ (for some constant $\delta > 0$). For any $\eta \in [0,1]$, there is a polynomial-time algorithm certifying that the number of $(1 - \eta)$-satisfying assignments to any 2XOR instance on G is*

- *at most 2 if $\eta \leqslant \frac{1}{3n}$,*

- *at most $2e^{3\eta n \log n}$ if $\eta > \frac{1}{3n}$,*

with probability at least $1 - \exp(-n^{\Omega(\delta)})$ over the randomness of G.

Remark 4.2. Given a random graph, the algorithm simultaneously certifies an upper bound for all instances on this graph with arbitrary signings. The simultaneous certification will be crucial in the subsequent sections.

Our first observation is that if the graph has large expansion, then given an approximate satisfying assignment, we cannot flip too many variables without violating many constraints. Specifically, if we flip $S \subseteq [n]$ ($|S| < n/2$), then the number of clauses negated is $e(S, \overline{S})$, and this is large if the graph is an expander. In fact, Theorem 4.1 holds for any instance with an expanding graph. Our second observation is the following (which holds for all $k \geqslant 2$).

Observation 4.3. Let $\mathcal{I}$ be a kXOR instance with arbitrary signings, and let $x, x' \in \{\pm 1\}^n$ be two $(1 - \eta)$-satisfying assignments. If a clause $(b, U) \in \mathcal{I}$ is satisfied by both, then $x_U \cdot x'_U = b^2 = 1$. Thus, the entry-wise product $x \circ x' \in \{\pm 1\}^n$ is a $(1 - 2\eta)$-satisfying assignment for the all-positive instance $\mathcal{I}_+$ (changing all signings to $+1$).

We now proceed to prove Theorem 4.1.

Proof of Theorem 4.1. The algorithm is as follows: given a graph with n vertices and $m = \Delta n = n^{1+\delta}$ edges, and a parameter $\eta \in [0,1]$.

(1) Check that all vertex degrees are within $2\Delta \left(1 \pm \frac{1}{\Delta^{1/3}}\right)$.

(2) Compute $\lambda_2(L)$ where L is the normalized Laplacian matrix. Check that $\lambda_2(L) \geqslant 1 - \frac{1}{\Delta^{1/4}}$.

(3) If the checks fail, output 2^n. Otherwise, if $\eta \leqslant \frac{1}{3n}$, output 2; otherwise, output $2e^{3\eta n \log n}$.

The random graph G is an Erdős-Rényi random graph sampled from $\mathcal{G}(n, p)$ with $p = \frac{2m}{n^2} = 2n^{-1+\delta}$ (removing parallel edges and self-loops allowed by the random model). It is a standard result in random graph theory that at this density, all vertex degrees concentrate around $np = 2\Delta$; specifically, the check in step (1) will succeed with probability $1 - \exp(-\Omega(\Delta^{1/3}))$ by a simple Chernoff bound (cf. [FK16]). Furthermore, the check in step (2) will succeed with probability $1 - \exp(-\Omega(\Delta^{1/4}))$ due to Theorem 2.2. Note that for any instance where the checks fail, the output 2^n is still a valid (trivial) upper bound.

Consider a maximum satisfying assignment x (assume it is $(1 - \eta)$-satisfying, otherwise our bound trivially holds), and let x' be any $(1 - \eta)$-satisfying assignment. By Observation 4.3, $y :=$

$x \circ x'$ is a $(1 - 2\eta)$-satisfying assignment to the all-positive instance $\mathcal{I}_+$ on G. Clearly, the all-ones vector $\vec{1}$ and its negation are exactly satisfying assignments to $\mathcal{I}_+$. We will show that y must be close to $\vec{1}$ or $-\vec{1}$, meaning that x' must be close to x or $-x$ in Hamming distance.

Suppose (without loss of generality) that y is closer to $\vec{1}$, and let $S = \{i : y_i = -1\}$ where $|S| \leqslant \frac{n}{2}$. Then, the constraints of $\mathcal{I}_+$ between S and $\overline{S}$ will be violated. By Cheeger's inequality (Theorem 2.1), $e(S, \overline{S}) \geqslant \frac{\lambda_2}{2} \operatorname{vol}(S)$, and by the degree concentration, $\operatorname{vol}(S) \geqslant |S| \cdot 2\Delta(1 - o(1))$. Thus, the spectral gap $\lambda_2 \geqslant 1 - o(1)$ and the degree bounds together certify that

$$e(S, \overline{S}) \geqslant \Delta |S| (1 - o(1)).$$

If $|S| \geqslant 3\eta n$, then $e(S, \overline{S}) > 2\eta \Delta n = 2\eta m$, contradicting that y is a $(1 - 2\eta)$-satisfying assignment of $\mathcal{I}_+$. Thus, any $(1 - \eta)$-satisfying assignment x' must be $\lfloor 3\eta n \rfloor$-close to x or $-x$ in Hamming distance. Note that if $\eta \leqslant \frac{1}{3n}$, then x' can only be $\pm x$.

For the $\eta > \frac{1}{3n}$ case, the number of assignments $\lfloor 3\eta n \rfloor$ away from $\pm x$ is upper bounded by

$$2 \sum_{\ell=0}^{\min(\lfloor 3\eta n \rfloor, n)} \binom{n}{\ell} \leqslant 2e^{3\eta n \log n}.$$

Note that if $\eta = O(1)$, then the upper bound trivially holds. For $\eta = o(1)$, we use the fact that $\binom{n}{3\eta n} \leqslant \frac{n^{3\eta n}}{(3\eta n)!}$. $\qquad\square$

4.2 Count certification for kXOR

For $k \geqslant 3$, we can obtain subexponential upper bounds by a recursive algorithm. Same as Theorem 4.1, our algorithm simultaneously certifies an upper bound for all kXOR instances on the given random hypergraph.

Theorem 4.4. *For constant $k \geqslant 3$, let $\boldsymbol{H} \sim \mathcal{H}^n_{k,+}(m)$ be a random k-uniform hypergraph with $m = \Delta n$ and $\Delta = n^\delta$ (for some constant $\delta \in (0, k-1)$). For any $\eta \in [0,1]$ and $\varepsilon > 0$, there is a polynomial-time algorithm certifying that the number of $(1 - \eta)$-satisfying assignments to any kXOR instance on $\boldsymbol{H}$ is at most*

$$2(2n)^{k-2} \cdot \exp\left(O(\eta n \log n)\right) \cdot \exp\left(O\left(n^{1 - \frac{\delta}{k-2} + \varepsilon}\right)\right)$$

with probability at least $1 - n^k \exp(-n^{\Omega(\varepsilon)})$ over the randomness of $\boldsymbol{H}$.

Remark 4.5. The $(2n)^{k-2}$ in the upper bound is there to handle the case when $\eta = o(1/n)$ and $\delta \geqslant k - 2$ (very dense instance). In this case, we get a $\operatorname{poly}(n)$ upper bound.

The main idea is that if we have a certification algorithm for $(k-1)$XOR, then we can obtain upper bounds for the *induced* $(k-1)$XOR defined in Definition 2.11. At a high level, the algorithm will do the following: 1) fix a set $S \subseteq V$ of a certain size, 2) look at the induced $(k-1)$-uniform hypergraph H_S (Definition 2.13), 3) run the certification algorithm for $(k-1)$XOR on H_S to obtain an upper bound, and 4) multiply by $2^{|S|}$.

The intuition is that for every assignment $\sigma_S \in \{\pm 1\}^S$, we get an induced $(k-1)$XOR instance with a random hypergraph and arbitrary signings (determined by σ_S). Here we crucially use the fact that our algorithm simultaneously certifies a bound for all signing patterns, hence we avoid enumerating every assignment σ_S. Once we have an upper bound on approximate solutions to the induced instances, we simply multiply it by $2^{|S|}$ to get the final upper bound.

We immediately see that for a fixed subset S, the above procedure throws away most of the clauses (keeping only clauses that have 1 variable in S). Thus, it is clearly suboptimal to look at just one subset S. To resolve this, we partition V into subsets $S_1, \ldots, S_\ell$, run the algorithm on each of them, and aggregate the results via the following lemma.

Lemma 4.6. *Given a kXOR instance $\mathcal{I}$, a partition of vertices $S_1, \ldots, S_\ell$, and a threshold t. Suppose for each $i \in [\ell]$ and each induced $(k-1)$XOR instance $\mathcal{I}_{S_i, \sigma_{S_i}}$, the number of assignments (on variables $[n] \setminus S_i$) that violate at most $\lfloor \frac{kt}{\ell} \rfloor$ constraints is upper bounded by u_i, then the number of assignments violating at most t constraints in $\mathcal{I}$ is upper bounded by $\sum_{i=1}^{\ell} 2^{|S_i|} u_i$.*

Proof. Let H be the underlying k-uniform hypergraph. Consider the induced $(k-1)$-uniform hypergraphs $H_{S_1}, \ldots H_{S_\ell}$. Each hyperedge in H contributes at most k hyperedges in the induced hypergraphs (i.e. the union of $H_{S_1}, \ldots, H_{S_\ell}$ will have at most k copies of the same hyperedge). Thus, for any assignment σ that violates $\leqslant t$ constraints in $\mathcal{I}$, there must be an $i \in [\ell]$ such that $\leqslant \lfloor \frac{kt}{\ell} \rfloor$ constraints are violated in the induced $(k-1)$XOR instance $\mathcal{I}_{S_i, \sigma_{S_i}}$.

Next, we bound the number of assignments violating at most t constraints.

$$\sum_{\sigma \in \{\pm 1\}^n} \mathbf{1}(\sigma \text{ violates} \leqslant t \text{ constraints in } \mathcal{I}) \leqslant \sum_{\sigma \in \{\pm 1\}^n} \sum_{i=1}^{\ell} \mathbf{1}\left(\sigma \text{ violates} \leqslant \left\lfloor \frac{kt}{\ell} \right\rfloor \text{ constraints in } \mathcal{I}_{S_i, \sigma_{S_i}}\right)$$

$$\leqslant \sum_{i=1}^{\ell} \left| \left\{ \sigma : \sigma \text{ violates} \leqslant \left\lfloor \frac{kt}{\ell} \right\rfloor \text{ constraints in } \mathcal{I}_{S_i, \sigma_{S_i}} \right\} \right|$$

$$\leqslant \sum_{i=1}^{\ell} 2^{|S_i|} u_i.$$

The second inequality follows by switching the two summations, and the final inequality holds because an upper bound u_i for the induced $(k-1)$XOR instance implies an upper bound of $2^{|S_i|} u_i$ by enumerating all possible assignments to S_i. $\qquad \square$

In the proof of Theorem 4.4, we will partition $[n]$ into $\ell = n^{1-c}$ subsets of size n^c for some $c < 1$ chosen later. The upper bounds $u_1, \ldots, u_\ell$ will be obtained recursively and will roughly be the same with high probability. Thus, by Lemma 4.6, we get an upper bound of $\ell 2^{n^c} u$, where we can choose c to obtain the optimal result.

Proof of Theorem 4.4. First, we can assume without loss of generality that $\varepsilon < \frac{\delta}{k-2}$, otherwise the upper bound trivially holds. Our algorithm is the exact same procedure as Lemma 4.6.

Certification Algorithm for kXOR

Input: k-uniform hypergraph with n variables and $m = \Delta n = n^{1+\delta}$ edges, parameters $\eta \in [0, 1], \varepsilon \in (0, \frac{\delta}{k-2})$.

(1) If $\delta < k - 2$, choose $c = 1 - \frac{\delta}{k-2} + \varepsilon$ for $0 < \varepsilon < \frac{\delta}{k-2}$; if $\delta \geqslant k - 2$, choose $c = 0$. Partition the vertices into $\ell = n^{1-c}$ subsets $S_1, \ldots, S_\ell$ of size n^c, and extract the induced $(k-1)$-uniform hypergraphs $H_{S_1}, \ldots, H_{S_\ell}$ (removing duplicate hyperedges).

(2) Run the $(k-1)$XOR certification algorithm with $\eta' m' = \frac{k}{\ell} \eta m$ on each H_{S_i}, where m' is the number of hyperedges in H_{S_i}. Let u_i be the upper bound.

24

(3) Output $\sum_{i=1}^{\ell} 2^{|S_i|} u_i$.

We will prove the correctness of the algorithm by induction on k. The base case is $k = 2$, and our 2XOR algorithm from Theorem 4.1 achieves the same guarantees (we assume in this case $n^{1-\frac{\delta}{k-2}+\varepsilon} = 0$). Now, suppose $k \geqslant 3$ and we have a $(k-1)$XOR algorithm with performance as stated in Theorem 4.4. Then, by Lemma 4.6, the output is a valid upper bound on the number of assignments violating at most ηm constraints. Theorem 4.4 requires the number of hyperedges $m = n^{1+\delta}$ for some constant $\delta \in (0, k-1)$, and thus it suffices to prove that the induced hypergraphs $H_{S_1}, \dots, H_{S_\ell}$ have the required density, which we can control by choosing c.

For each S_i, the induced $(k-1)$XOR is a random $(k-1)$-uniform hypergraph where each $(k-1)$-tuple is included with probability $q := 1 - (1-p)^{|S_i|}$, where $p = \frac{m}{n^k} = n^{1+\delta-k}$ (recall that we treat hyperedges as tuples; removing duplicates will not affect the upper bound). Here we split into two cases,

- $\delta \leqslant k-2$: we set $c = 1 - \frac{\delta}{k-2} + \varepsilon < 1$. Then, $p|S_i| = n^{1+\delta-k} \cdot n^c < n^{-1+c} = o(1)$. Thus, $q = p|S_i|(1 \pm o(1)) = pn^c(1 \pm o(1))$.

- $\delta > k-2$: we set $c = 0$, thus $q = p = pn^c$.

In both cases, the number of hyperedges in the induced hypergraph is concentrated around

$$m' = qn^{k-1} = \frac{m}{n^k} n^c (1 \pm o(1)) \cdot n^{k-1} = n^{\delta+c}(1 \pm o(1)).$$

Thus, the density $\Delta' = \frac{m'}{n-n^c} \sim n^{\delta+c-1}$. Let $\delta' := \delta + c - 1$. Again, we split into the two cases,

- $\delta \leqslant k-2$: we have $\delta' = \delta + (1 - \frac{\delta}{k-2} + \varepsilon) - 1 = (1 - \frac{1}{k-2})\delta + \varepsilon \geqslant \varepsilon > 0$ since $\varepsilon > 0$, and $\delta' \leqslant (k-2) + c - 1 < k-2$ since $c < 1$. Further, $\frac{\delta'}{k-3} \geqslant \frac{\delta}{k-2} + \frac{\varepsilon}{k-3} > \frac{\delta}{k-2}$.

- $\delta > k-2$: we have $\delta' = \delta - 1 > k-3 \geqslant 0$ since $k \geqslant 3$, and $\delta' < k-2$ since $\delta < k-1$. Further, $\frac{\delta'}{k-3} = \frac{\delta-1}{k-3} > \frac{\delta}{k-2}$.

In both cases, the induced $(k-1)$XOR instance has the required density $\Delta' = n^{\delta'}$ with $\delta' \in (0, k-2)$, which means we can apply the $(k-1)$XOR algorithm. The parameter η' is set to $\frac{1}{m'} \cdot \frac{k}{\ell} \eta m \sim k\eta$ (capped at 1), and set $\varepsilon' = \varepsilon$.

The $(k-1)$XOR algorithm on the induced instance will certify an upper bound of

$$u_i \leqslant 2(2n)^{k-3} \exp\left(O(\eta' n \log n)\right) \cdot \exp\left(O\left(n^{1-\frac{\delta'}{k-3}+\varepsilon'}\right)\right).$$

Since $\eta' = O(\eta)$, $\varepsilon' = \varepsilon$, and $\frac{\delta'}{k-3} > \frac{\delta}{k-2}$, our final upper bound is

$$\sum_{i=1}^{\ell} 2^{|S_i|} u_i \leqslant 2(2n)^{k-2} \cdot \exp\left(O(\eta n \log n)\right) \cdot \exp\left(O\left(n^{1-\frac{\delta}{k-2}+\varepsilon}\right)\right).$$

Finally, we bound the failure probability. The $(k-1)$XOR algorithm fails with probability $< n^{k-1} \exp(-n^{\Omega(\varepsilon')}) = n^{k-1} \exp(-n^{\Omega(\varepsilon)})$. We union bound over the ℓ induced hypergraphs, we get the failure probability $< n^k \exp(-n^{\Omega(\varepsilon)})$. $\qquad \square$

4.3 Count certification for all kCSPs

First, observe that as an immediate consequence of Theorem 4.4 and Lemma 3.1, we have:

Corollary 4.7. *For constant $k \geqslant 3$, let $\mathcal{I} \sim \mathcal{H}_k^m(n)$ be a random signed hypergraph where $m = \Delta n = n^{1+\delta}$. For every constant $\varepsilon > 0$, there is an algorithm that certifies with high probability that the number of $(1 - \eta)$-satisfying assignments to $\mathcal{I}$ as an instance of kSAT is at most*

$$\exp\left(\widetilde{O}(\eta n)\right) \cdot \exp\left(\widetilde{O}\left(n^{\frac{k+1}{4} - \frac{\delta}{2}}\right)\right) \cdot \exp\left(O\left(n^{1 - \frac{\delta}{k-2} + \varepsilon}\right)\right).$$

It is simple to upgrade the statement of Corollary 4.7 from the case of kSAT to all kCSPs.

Corollary 4.8. *Let P be any predicate not equal to the constant-1 function. For constant $k \geqslant 3$, let $\mathcal{I} \sim \mathcal{H}_k^m(n)$ be a random signed hypergraph where $m = \Delta n = n^{1+\delta}$. For every constant $\varepsilon > 0$ there is an algorithm that certifies with high probability that the number of $(1 - \eta)$-satisfying assignments to $\mathcal{I}$ as an instance of P is at most*

$$\exp\left(\widetilde{O}(\eta n)\right) \cdot \exp\left(\widetilde{O}\left(n^{\frac{k+1}{4} - \frac{\delta}{2}}\right)\right) \cdot \exp\left(O\left(n^{1 - \frac{\delta}{k-2} + \varepsilon}\right)\right).$$

Proof. Let $z \in \{\pm 1\}^k$ be any string not in the support of P. Construct a new signed hypergraph $\mathcal{I}' := \{(c \circ z, U) : (c, U) \in \mathcal{I}\}$. If x is a $(1 - \eta)$-satisfying assignment to $\mathcal{I}$ as an instance of P, then x is also a $(1 - \eta)$-satisfying assignment to $\mathcal{I}'$ as an instance of kSAT. Further, it is easy to see that $\mathcal{I}'$ is also distributed as $\mathcal{H}_k^m(n)$, so we can apply the algorithm from Corollary 4.7 to certify a bound on the number of $(1 - \eta)$-satisfying assignments to $\mathcal{I}'$ as an instance of kSAT. Consequently, we get the desired statement. $\square$

5 Count certification of Solution Clusters

In this section, we bound the number of clusters of satisfying assignments. Due to the 3XOR-principle, we first focus on clusters of $(1 - \eta)$-satisfying assignments to random 3XOR instances.

Theorem 5.1. *Consider a random 3-uniform hypergraph $\boldsymbol{H} \sim \mathcal{H}_{3,+}^n(m)$ where $m = \Delta n$ and $\Delta = n^\delta$ for some constant $\delta \in (0, 2)$. Let $\eta \in [0, \eta_0]$ where η_0 is a universal constant, and let $\theta := \max(2\eta, \Delta^{-\frac{1}{2}} \log n)$. There is a polynomial-time algorithm certifying that the $(1 - \eta)$-satisfying assignments to any 3XOR instance on $\boldsymbol{H}$ are covered by at most*

$$\exp(O(\theta^2 \log(1/\theta))n)$$

diameter-(θn) clusters, with probability at least $1 - \frac{1}{\text{poly}(n)}$ over the randomness of $\boldsymbol{H}$.

As an immediate corollary of Theorem 5.1, Lemma 3.1, and the reduction in the proof of Corollary 4.8 we have:

Corollary 5.2. *Let P be any 3-ary predicate not equal to the constant-1 predicate. Let $\mathcal{I} \sim \mathcal{H}_3^m(n)$ be a random signed hypergraph where $m = \Delta n = n^{1+\delta}$ for some constant $\delta \in (0, 2)$. Let $\eta \in [0, \eta_0]$ where η_0 is a universal constant, and let $\theta := 8\eta + O(\sqrt{\frac{\log^5 n}{\Delta}})$. There is an algorithm that certifies with high probability that the $(1 - \eta)$-satisfying assignments to $\mathcal{I}$ as a P-CSP instance are covered by at most*

$$\exp(O(\theta^2 \log(1/\theta))n)$$

diameter-(θn) clusters.

Inspecting the proof of Theorem 4.1, we see that it actually proves a stronger statement: for any pair of $(1 - \eta)$-satisfying assignments x, x' to the 2XOR instance $\mathcal{I}$, x' must be $(3\eta n)$-close to x or $-x$ in Hamming distance. The proof looks at the instance $\mathcal{I}_+$ (where all signs are set to $+1$) and the expansion of the graph.

The proof of Theorem 5.1 will follow a similar path. On a high level, we will first prove that $y := x \circ x'$ must be either close to $\vec{1}$ or be roughly balanced, i.e. x and x' must have Hamming distance close to 0 or roughly $\frac{n}{2}$. The main ingredient is Lemma 5.4 which lets us certify an important structural result of random 3-uniform hypergraphs, allowing us to reason about the constraints violated by y in $\mathcal{I}_+$. Lemma 5.4 will be a crucial step in Section 6 as well.

The second ingredient is a result in coding theory. Since the clusters are roughly $\frac{n}{2}$ apart, the number of clusters must be upper bounded by the cardinality of the largest ε-balanced binary error-correcting code. The best known upper bound is $2^{O(\varepsilon^2 \log(1/\varepsilon))n}$, obtained by [MRRW77] using linear programming techniques and also by [Alo09] using an analysis of perturbed identity matrices. This gives our final result.

We begin by proving an eigenvalue bound for the *primal graph* of a random 3-uniform hypergraph (the graph obtained by adding a 3-clique for each hyperedge; see Definition 2.14). For simplicity, we will implicitly assume that all hyperedges with repeated vertices (allowed in our random model) are removed; this will not affect the results.

Lemma 5.3. *Let $H \sim \mathcal{H}_{3,+}^n(m)$ be a random 3-uniform hypergraph where $m = \Delta n$ and $\Delta = n^\delta$ (for constant $\delta \in (0, 2)$), and let A be the adjacency graph of the primal graph H_P. Then, there is a constant C such that with probability $1 - \frac{1}{\text{poly}(n)}$,*

$$\left\| A - \frac{6\Delta}{n} J \right\| \leqslant C\sqrt{\Delta \log n}.$$

Proof. The primal graph is a random graph such that for each tuple $U = (a, b, c) \in [n]^3$ (no repeated vertices), edges $(a, b), (b, c), (c, a)$ are included with probability $p := \frac{m}{n^3} = \frac{\Delta}{n^2}$. Let A_U be the adjacency matrix of the graph containing just the 3 edges. Then, the adjacency matrix $A = \sum_{U \in [n]^3} \mathbf{1}(U \in H) A_U$.

Define $S_U := (\mathbf{1}(U \in H) - p) A_U$ and

$$S := \sum_{U \in [n]^3} S_U = A - p \cdot 6(n-2)(J - \mathbb{1}).$$

Clearly, $\mathbf{E}[S] = 0$ since $\mathbf{E}[S_U] = 0$, and $\mathbf{E}[S_U^2] = p(1-p)A_U^2$. Further, $\sum_U A_U^2 = 6(n-2)(J + (n-2)\mathbb{1})$, thus

$$v := \left\| \sum_{U \in [n]^3} \mathbf{E}[S_U^2] \right\| = p(1-p) \left\| \sum_{U \in [n]^3} A_U^2 \right\| = p(1-p) \cdot 6(n-1)(2n-2) \leqslant 12\Delta.$$

Moreover, $\|S_U\| \leqslant \|A_U\| \leqslant 2$. Thus, by the matrix Bernstein inequality (Theorem 2.19) and assuming $\Delta = \omega(\log n)$, for a large enough constant C,

$$\mathbf{Pr}\left[\|S\| \geqslant C\sqrt{\Delta \log n} \right] \leqslant \frac{1}{\text{poly}(n)}.$$

27

Next, $S = A - \frac{6\Delta(n-2)}{n^2}(J - \mathbb{1}) = (A - \frac{6\Delta}{n}J) + \frac{12\Delta}{n^2}J - \frac{6\Delta(n-2)}{n^2}\mathbb{1}$. The second and third terms have norm $O(\frac{\Delta}{n})$, negligible compared to $\sqrt{\Delta \log n}$ when $\Delta \ll n^2 \log n$. Thus, by the triangle inequality,

$$\left\| A - \frac{6\Delta}{n}J \right\| \leqslant C\sqrt{\Delta \log n}. \qquad \square$$

Lemma 5.3 allows us to apply the expander mixing lemma on the primal graph and prove the following structural result for random hypergraphs.

Lemma 5.4. *Let $H \sim \mathcal{H}_{3,+}^n(m)$ be a random 3-uniform hypergraph where $m = \Delta n$ and $\Delta = n^\delta$ (for constant $\delta \in (0,2)$). There is an algorithm certifying that for all subsets $S \subseteq [n]$ such that $|S| = (\frac{1}{2} + \gamma)n$ and $\gamma \in (0, \frac{1}{2})$,*

1. *the number of hyperedges with 2 variables in S and 1 in $\overline{S}$ is at least $3m(\gamma - 2\gamma^2) - O(n\sqrt{\Delta \log n})$,*

2. *the number of hyperedges fully contained in S is at least $m(\gamma + 2\gamma^2) - O(n\sqrt{\Delta \log n})$,*

with probability $1 - \frac{1}{\text{poly}(n)}$ over the randomness of H.

Proof. We first look at the primal graph H_P. The average degree is $\frac{6m}{n} = 6\Delta$. Let A be the adjacency matrix, and let $\overline{A} = A - \frac{6\Delta}{n}J$. The certification algorithm is simply checking that $\|\overline{A}\| \leqslant C\sqrt{\Delta \log n}$ for some constant C. By Lemma 5.3, this will succeed with high probability. We proceed to prove that this is a valid certificate.

We categorize the hyperedges of H into 4 groups T_0, T_1, T_2, T_3, where T_i is the set of hyperedges with i variables in S and $3 - i$ in $\overline{S}$. We first lower bound $|T_2|$.

By the expander mixing lemma (Theorem 2.3), the number of edges (of H_P) between S and $\overline{S}$ is

$$e(S, \overline{S}) = \frac{6\Delta}{n}|S|(n - |S|) \pm C\sqrt{\Delta \log n}\sqrt{|S|(n - |S|)}.$$

Moreover, the number of edges within $\overline{S}$ (note that $e(\overline{S}, \overline{S})$ double counts the edges)

$$\frac{1}{2}e(\overline{S}, \overline{S}) = \frac{3\Delta}{n}|\overline{S}|^2 \pm C\sqrt{\Delta \log n}|\overline{S}|.$$

Observe that the edges between S and $\overline{S}$ must come from T_1 and T_2, each hyperedge contributing 2 edges: $2|T_1| + 2|T_2| = e(S, \overline{S})$. On the other hand, the the edges within $\overline{S}$ come from T_0, T_1, each hyperedge contributing 3 and 1 edges respectively: $3|T_0| + |T_1| = \frac{1}{2}e(\overline{S}, \overline{S})$. Thus, we have $|T_2| \geqslant \frac{1}{2}e(S, \overline{S}) - \frac{1}{2}e(\overline{S}, \overline{S})$. For $|S| = (\frac{1}{2} + \gamma)n$,

$$|T_2| \geqslant \frac{3\Delta}{n} \cdot \left(\left(\frac{1}{2} - \gamma\right)\left(\frac{1}{2} + \gamma\right) - \left(\frac{1}{2} - \gamma\right)^2 \right) n^2 - O(n\sqrt{\Delta \log n})$$

$$= 3m(\gamma - 2\gamma^2) - O(n\sqrt{\Delta \log n}).$$

Next, we lower bound $|T_3|$. Similar to the derivations for $|T_2|$, we observe that $3|T_3| + |T_2| = \frac{1}{2}e(S, S)$ and $|T_2| \leqslant \frac{1}{2}e(S, \overline{S})$, hence $|T_3| \geqslant \frac{1}{6}(e(S, S) - e(S, \overline{S}))$. Similar calculations show that

$$|T_3| \geqslant m(\gamma + 2\gamma^2) - O(n\sqrt{\Delta \log n}). \qquad \square$$

Note that for small γ, the error term $O(n\sqrt{\Delta \log n})$ is negligible compared to $m(\gamma \pm 2\gamma^2)$ as long as $\gamma \gg \sqrt{\frac{\log n}{\Delta}}$. Moreover, in the first claim, for γ close to $\frac{1}{2}$ (say $\gamma = \frac{1}{2} - \gamma'$, $|S| = (1 - \gamma')n$), the error term is negligible as long as $\gamma' \gg \sqrt{\frac{\log n}{\Delta}}$.

Next, we state a result by [Alo09], which was proved using an elegant argument about rank lower bounds of perturbed identity matrices. Towards doing so, we define an *ε-balanced code of length-n* as a subset L of $\{\pm 1\}^n$ such that every pair of distinct $x, y \in L$ have Hamming distance in $\frac{1 \pm \varepsilon}{2}n$.

Lemma 5.5 ([Alo09], Proposition 4.1]). *For any $\frac{1}{\sqrt{n}} \leqslant \varepsilon < \frac{1}{2}$, the cardinality of any ε-balanced code of length n is at most $2^{c\varepsilon^2 \log(1/\varepsilon)n}$ for some absolute constant c.*

Finally, we are ready to prove Theorem 5.1.

Proof of Theorem 5.1. Similar to the proof of Theorem 4.1, we consider the instance $\mathcal{I}_+$. By Observation 4.3, for any $(1 - \eta)$-satisfying assignments x, x', the product $y := x \circ x'$ is a $(1 - 2\eta)$-satisfying assignment for $\mathcal{I}_+$.

Let $S_+ = \{i : y_i = +1\}$ and $S_- = \overline{S_+}$. Assume $|S_+| = (\frac{1}{2} + \gamma)n$ for $\gamma > 0$ (x, x' agree on more than half). Since all 3XOR clauses have sign $+1$ in $\mathcal{I}_+$, the clauses that have 2 variables in S_+ and 1 in S_- must be violated. By Lemma 5.4, we can certify a lower bound of $3m(\gamma - 2\gamma^2)(1 - o(1))$ of such clauses when $\omega(\sqrt{\frac{\log n}{\Delta}}) \leqslant \gamma \leqslant \frac{1}{2} - \omega(\sqrt{\frac{\log n}{\Delta}})$. Thus, we take $\theta := \max(2\eta, \Delta^{-\frac{1}{2}} \log n)$. For a small enough η ($\eta < 1/6$ suffices), we can certify that the number of violated constraints $3m(\gamma - 2\gamma^2)(1 - o(1)) > 2\eta m$ for all $\gamma \in [\theta, \frac{1}{2} - \theta]$. This shows that $|S_+|$ must be either $\geqslant (1 - \theta)n$ or $\leqslant (\frac{1}{2} + \theta)n$.

On the other hand, suppose $|S_-| = (\frac{1}{2} + \gamma)n$ for $\gamma > 0$ (x, x' agree on less than half). The clauses contained in S_- must be violated. Again, Lemma 5.4 allows us to lower bound such clauses by $m(\gamma + 2\gamma^2)(1 - o(1)) > 2\eta m$ for all $\gamma \in [\theta, \frac{1}{2}]$. This shows that $|S_-|$ must be $\leqslant (\frac{1}{2} + \theta)n$.

Combining the results, we can certify that x, x' must have Hamming distance $\leqslant \theta n$ or between $[(\frac{1}{2} - \theta)n, (\frac{1}{2} + \theta)n]$. Thus, the $(1 - \eta)$-satisfying solutions form clusters of diameter θn, and the distance between any two clusters is $(\frac{1}{2} \pm \theta)n$. If we pick one assignment from each cluster, this gives a (2θ)-balanced code. Thus, by Lemma 5.5, we can upper bound the number of clusters by

$$\exp(O(\theta^2 \log(1/\theta))n).$$

This completes the proof. $\qquad\square$

6 Refuting CSPs under global cardinality constraints

In this section, we give an algorithm to strongly refute random CSPs with global cardinality constraints, i.e., constraints of the form $\sum_i x_i \geqslant B$ well under the refutation threshold for appropriate values of B.

6.1 Refuting 3CSPs under global cardinality constraints

For 3SAT, there is a strong refutation algorithm using the random hypergraph structure result of Lemma 5.4, without requiring the 3XOR-principle. Via an identical reduction as the one in the proof of Corollary 4.8 the below statement extends to all 3CSPs.

Theorem 6.1. *Given a 3SAT instance $\mathcal{I} \sim \mathcal{H}_3^n(m)$ where $m = \Delta n = n^{1+\delta}$ for constant $\delta > 0$, there is an efficient algorithm that certifies with high probability that $\mathcal{I}$ has no ρ-biased assignment which is $(1-\eta)$-satisfying where $\rho \gg \sqrt{\frac{\log n}{\Delta}}$ and $\eta = \rho/32$.*

Proof. Given a random 3SAT instance $\mathcal{I}$, we extract sub-instances $\mathcal{I}|_+$ and $\mathcal{I}|_-$ consisting of clauses with no negations and fully-negated clauses, respectively. Each sub-instance is a random 3-uniform hypergraph with density $\frac{1}{8}\Delta$. Consider an assignment $x \in \{\pm 1\}^n$, and let $S_+ = \{i : x_i = 1\}$ and $S_- = \{i : x_i = -1\}$. Our main insight is that all hyperedges of $\mathcal{I}|_-$ contained in S_+ must be violated, and similarly all hyperedges of $\mathcal{I}|_+$ contained in S_- must be violated.

First, we consider the case when $|S_+| = \frac{1+\rho}{2}n$ for $\rho \gg \sqrt{\frac{\log n}{\Delta}}$. By Lemma 5.4, we can certify with high probability that there must be more than $\frac{1}{8}m \cdot \frac{\rho}{4}$ hyperedges of $\mathcal{I}|_-$ contained in S_+. Thus, for $\eta = \rho/16$, any assignment x such that $|S_+| \geqslant \frac{1+\rho}{2}n$ cannot be $(1-\eta)$-satisfying.

Similarly, consider the case when $|S_-| = \frac{1+\rho}{2}n$ for $\rho \gg \sqrt{\frac{\log n}{\Delta}}$. Again by Lemma 5.4, we can certify with high probability that there must be more than $\frac{1}{8}m \cdot \frac{\rho}{4}$ hyperedges of $\mathcal{I}_+$ contained in S_-. Thus, any assignment x such that $|S_-| \geqslant \frac{1+\rho}{2}n$ cannot be $(1-\eta)$-satisfying.

Therefore, this certifies that any ρ-biased assignment x cannot be $(1-\eta)$-satisfying. $\qquad\square$

Remark 6.2. We compare our result to the result of [KOS18, Corollary C.2]. For random 3XOR with $m = n^{\frac{3}{2}-\varepsilon}$ (under the refutation threshold), they showed that the Sum-of-Squares algorithm can certify that there is no ρ-biased exactly satisfying assignment when $\rho = \widetilde{\Omega}(n^{-\frac{1}{4}+\frac{\varepsilon}{2}})$. Our algorithm matches this cardinality condition for ρ, and further extends to $(1-\Theta(\rho))$-satisfying assignments and to arbitrary 3CSPs.

6.2 The case of kCSPs when $k \geqslant 4$

In this section we give an algorithm to refute kCSPs under global cardinality constraint when $k \geqslant 4$. Our approach yields quantitatively weaker guarantees so we make no effort to optimize the tradeoff between refutation quality and the imbalance in the global cardinality constraint. Akin to our certified counting algorithms, we begin by first giving strong refutation algorithms for kXOR, then use the kXOR-principle to extend the algorithm to kSAT, which then implies a strong refutation algorithm for every kCSP. For kXOR we prove:

Theorem 6.3. *Let $k \geqslant 4$. Given a kXOR instance $\mathcal{I} \sim \mathcal{H}_k^n(m)$ where $m := n^{\frac{k-2}{2}+\beta}$, there is an efficient algorithm that certifies with high probability that $\mathcal{I}$ has no 2ρ-biased assignment which is $(1-\eta)$-satisfying where $\rho \gg \frac{\log^6 n}{n^{\beta/(k-2)}}$ and*

$$\eta = \rho^{k-2}\left(\frac{\rho^2}{2} - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/2}n^{(k-4)/4}n^{\beta/2}}\right) - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/4}n^{\beta/2}}\right) - \frac{2}{n}\right).$$

Via the kXOR principle (Lemma 3.1), and the arbitrary kCSP-to-kSAT reduction in the proof of Corollary 4.8, we get the following statement for refutation of kCSPs under global cardinality constraints.

Corollary 6.4. *Let $\mathcal{I} \sim \mathcal{H}_k^n(m)$ where $m := n^{\frac{k-1}{2}+\beta}$ and $\beta > 0$. For any predicate P not equal to the constant-1 function and any constant $\rho > 0$, there is an efficient algorithm that certifies that $\mathcal{I}$ has no 2ρ-biased assignment which $(1-\rho^k/2)$-satisfies $\mathcal{I}$ as a P-CSP instance.*

While a quantitatively stronger statement than the above is true, we present the simplified version for ease of exposition.

Now we turn our attention to proving Theorem 6.3, and in service of which we prove two other lemmas as ingredients.

The high level idea for our kXOR strong-refutation algorithm is to pick some set of vertices $S \subseteq [n]$ of size n^c where $0 < c < 1$ is an appropriately chosen constant. Then for any assignment y to the variables in S, there is an induced 2XOR instance $\mathcal{I}_{S,y,2}$ on $[n] \setminus S$ that must be approximately satisfied (recall the definition of induced instances in Definition 2.11). The two steps of the algorithm are then to:

1. Certify that for any y the induced 2XOR instance is $(1/2 + \varepsilon)$-positive for some small ε.

2. Simultaneously strongly-refute the family of all 2XOR instances that are $(1/2 + \varepsilon)$-positive under a global cardinality constraint.

We first describe the algorithm that certifies that the induced 2XOR instance is $(1/2 + \varepsilon)$-positive.

Lemma 6.5. *Given a kXOR instance $\mathcal{I} \sim \mathcal{H}_k^n(m)$ where $m := n^{\frac{k-2}{2} + \beta}$, constant c satisfying $c > 1 - \frac{2\beta}{k-2}$, and a fixed set of vertices S of cardinality n^c, there is an efficient algorithm to certify with high probability:*

For any $y \in \{\pm 1\}^S$ the instance $\mathcal{I}_{S,y,2}$ is $\left(\frac{1}{2} + \varepsilon_{6.5}\right)$-positive for the following choice of $\varepsilon_{6.5}$.

$$\varepsilon_{6.5} = \widetilde{O}\left(\max\left\{1, n^{\frac{\beta}{2} - \frac{k-2}{4}}\right\} \sqrt{\frac{n^{(1-c)\frac{k-2}{2}}}{n^\beta}}\right).$$

Proof. To put the statement we want to certify in a different way, we would like an algorithm that certifies:

$$\frac{1}{|\mathcal{I}_{S,y,2}|} \sum_{(b,U) \in \mathcal{I}_{S,y,2}} b \leqslant 2\varepsilon.$$

By Observation 2.12, this is equivalent to certifying the following for the truncated instance:

$$\frac{1}{\left|\mathcal{I}|_{S,k-2}\right|} \sum_{(b,U) \in \mathcal{I}|_{S,k-2}} b \prod_{i \in U} y_i \leqslant 2\varepsilon,$$

whose LHS can then be rewritten as:

$$\frac{1}{\left|\mathcal{I}|_{S,k-2}\right|} \sum_{U \in S^{k-2}} \prod_{i \in U} y_i \sum_{(b,U) \in \mathcal{I}|_{S,k-2}} b.$$

We write w_U to denote $\sum_{(b,U) \in \mathcal{I}|_{S,k-2}} b$. Each w_U is distributed as the sum of $O(n^2)$ random variables which are independent, bounded, centered and each nonzero with probability $O(n^{\beta-1-k/2})$. Denote α as the expected number of nonzero terms in the sum defining w_U; its value is $Cn^{\beta+1-k/2}$ for some constant $C > 0$. By standard binomial concentration and Hoeffding's inequality, we know that with high probability $|w_U| \leqslant B := \max\{\log n, \sqrt{\alpha} \log n\}$ for all $U \in S^{k-2}$. Further, the probability that a given w_U is nonzero is at most $p := \min\{\alpha, 1\}$. Define $\widetilde{w}_U$ as the random variable $w_U \mathbf{1}[|w_U| \leqslant B]$.

31

1. $\widetilde{w}_U$ is a centered random variable since w_U is symmetric around 0.

2. $\widetilde{w}_U$ is supported on $[-B, B]$.

3. The probability that $\widetilde{w}_U$ is nonzero is at most p.

Hence by the certification algorithm of [AOW15] (Theorem 2.18), we can certify with high probability that for any $y \in \{\pm 1\}^S$:

$$\sum_{U \in S^{k-2}} \widetilde{w}_U \prod_{i \in U} y_i \leqslant 2^{O(k)} B \max\{\sqrt{p}|S|^{3(k-2)/4}, |S|^{(k-2)/2}\} \log^{3/2} |S|.$$

Since $w_U = \widetilde{w}_U$ for all $U \in S^{k-2}$ with high probability, our algorithm can verify this and also certify an identical upper bound on:

$$\sum_{U \in S^{k-2}} w_U \prod_{i \in U} y_i \leqslant 2^{O(k)} B \max\{\sqrt{p}|S|^{3(k-2)/4}, |S|^{(k-2)/2}\} \log^{3/2} |S|. \tag{4}$$

Plugging $|S| = n^c$ and $p = \min\{Cn^{\beta+1-k/2}, 1\}$ into (4) with $c > 1 - \frac{2\beta}{k-2}$, we have $\sqrt{p}|S|^{3(k-2)/4} > |S|^{(k-2)/2}$. Moreover, since $\left|\mathcal{I}|_{S,k-2}\right|$ concentrates around $(1 \pm o(1))(\frac{|S|}{n})^{k-2}m$, we can certify with high probability that:

$$\frac{1}{\left|\mathcal{I}|_{S,k-2}\right|} \sum_{U \in S^{k-2}} w_U \prod_{i \in U} y_i \leqslant 2^{O(k)} B \sqrt{\frac{n^{(1-c)\frac{k-2}{2}}}{n^\beta}} \log^3 n$$

$$\leqslant 2^{O(k)} \max\left\{1, n^{\frac{\beta}{2} - \frac{k-2}{4}}\right\} \sqrt{\frac{n^{(1-c)\frac{k-2}{2}}}{n^\beta}} \log^{5/2} n,$$

since $B = \max\{1, \sqrt{\alpha}\} \log n = O(\max\{1, n^{\frac{\beta}{2} - \frac{k-2}{4}}\} \log n)$. $\qquad \square$

We now describe the algorithm to *simultaneously* refute the relevant family of 2XOR instances. The following definition more concretely describes the family of instances we are interested in.

Definition 6.6. Given a multi-graph G, let $\mathcal{F}(G, \varepsilon)$ be the collection of all 2XOR instances on G that are $(\frac{1}{2} + \varepsilon)$-positive.

Lemma 6.7. *Let G be an n-vertex random multi-graph with average degree $\Delta \geqslant \log^2 n$ obtained by independently adding $\mathrm{Binom}\left(r, \frac{\Delta}{nr}\right)$ edges between i and j for every pair of distinct $i, j \in [n]$. Further assume $r > \frac{2\Delta}{n}$. There is an efficient algorithm that takes in G as input and with high probability certifies that every $\mathcal{I} \in \mathcal{F}(G, \varepsilon)$ has no ρ-biased assignment that is $\left(1 - \frac{\rho^2}{2} + 8\sqrt{\frac{\log n}{\Delta}} + \varepsilon + \frac{2}{n}\right)$-satisfying.*

Proof. Let x be a ρ-biased assignment and let $Y = \{v \in [n] : x_v = +1\}$ where $|Y| = (\frac{1}{2} + c)n$ and $|c| \geqslant \frac{\rho}{2}$. To prove the statement our algorithm will certify a lower bound on the fraction of constraints that are violated in any $\mathcal{I} \in \mathcal{F}(G, \varepsilon)$. Towards doing so, we will first certify some lower bound γ on the fraction of constraints within Y or $\overline{Y}$. Now, for every constraint $\{u, v\}$ where $u, v \in Y$ or $u, v \in \overline{Y}$, $x_u x_v = +1$. Since the fraction of constraints $(b, \{u, v\})$ for which $b = +1$ is bounded by $\frac{1}{2} + \varepsilon$, the fraction of violated constraints must be at least $\gamma - (\frac{1}{2} + \varepsilon)$.

The certification of the lower bound is much like the proof of the expander mixing lemma. The number of edges that are contained within Y or $\overline{Y}$ is accounted by:

$$
\begin{aligned}
2(|E(G[Y])| + |E(G[\overline{Y}])|) &= 1_Y^\top A_G 1_Y + 1_{\overline{Y}}^\top A_G 1_{\overline{Y}} \\
&= 1_Y^\top \mathbf{E}[A_G] 1_Y + 1_{\overline{Y}}^\top \mathbf{E}[A_G] 1_{\overline{Y}} + 1_Y^\top (A_G - \mathbf{E}\, A_G) 1_Y + 1_{\overline{Y}}^\top (A_G - \mathbf{E}\, A_G) 1_{\overline{Y}} \\
&\geqslant \frac{\Delta}{n} \left(|Y|^2 + |\overline{Y}|^2 \right) - \|A_G - \mathbf{E}\, A_G\| \cdot n - \Delta.
\end{aligned}
$$

With $|Y| = \left(\frac{1}{2} + c \right) n$ and $|c| \geqslant \frac{\rho}{2}$, we can rewrite the above inequality as:

$$
2(|E(G[Y])| + |E(G[\overline{Y}])|) \geqslant \Delta n \left(\frac{1}{2} + 2c^2 \right) - \|A_G - \mathbf{E}\, A_G\| \cdot n - \Delta
$$

$$
\geqslant \frac{1}{2} \Delta n \left(1 + \rho^2 \right) - \|A_G - \mathbf{E}\, A_G\| \cdot n - \Delta
$$

With the above bound in hand, our algorithm can certify a lower bound of $\gamma - \left(\frac{1}{2} + \varepsilon \right)$ where

$$
\gamma = \frac{1}{2|E(G)|} \left(\frac{1}{2} \Delta n \left(1 + \rho^2 \right) - \|A_G - \mathbf{E}\, A_G\| \cdot n - \Delta \right).
$$

Next, to treat the factor involving $|E(G)|$ observe that:

$$
2|E(G)| = 1^\top \mathbf{E}\, A_G 1 + 1^\top (A_G - \mathbf{E}\, A_G) 1 \leqslant \Delta(n-1) + \|A_G - \mathbf{E}\, A_G\| \cdot n.
$$

To complete the proof, it suffices to give a high-probability upper bound on $\|A_G - \mathbf{E}\, A_G\|$. For the bound on $\|A_G - \mathbf{E}\, A_G\|$ we use the matrix Bernstein inequality, as stated in Theorem 2.19. Let $r \cdot K_n$ be the graph on $[n]$ with r parallel edges between every pair of vertices. Then G can be thought of as the graph obtained by sampling each $e \in E(r \cdot K_n)$ independently with probability $\frac{\Delta}{rn}$. For $e \in E(r \cdot K_n)$ define A_e as the adjacency matrix of the single edge e. Then:

$$
A_G - \mathbf{E}\, A_G = \sum_{e \in E(r \cdot K_n)} A_e \cdot \left(\mathbf{1}[e \in G] - \frac{\Delta}{rn} \right).
$$

Then the v parameter from the statement of Theorem 2.19 for the above sum of random matrices is then equal to:

$$
\left\| \sum_{e \in E(r \cdot K_n)} \mathbb{1}_e \, \mathbf{E} \left[\left(\mathbf{1}[e \in G] - \frac{\Delta}{rn} \right)^2 \right] \right\| \leqslant \left\| \Delta \cdot \left(1 - \frac{\Delta}{nr} \right) \cdot \mathbb{1} \right\| \leqslant \Delta.
$$

Thus, by Theorem 2.19:

$$
\|A_G - \mathbf{E}\, A_G\| \leqslant 4\sqrt{\Delta \log n}
$$

except with probability at most $1/n^2$. When this holds for large enough n this implies:

$$
\gamma \geqslant \frac{1}{2} + \frac{\rho^2}{2} - 8\sqrt{\frac{\log n}{\Delta}} - \frac{2}{n}.
$$

This in turn implies that the lower bound certified on the fraction of constraints violated by x is with high probability at least:

$$
\frac{\rho^2}{2} - 8\sqrt{\frac{\log n}{\Delta}} - \frac{2}{n} - \varepsilon
$$

which completes the proof. $\qquad\square$

Now we are ready to prove Theorem 6.3.

Proof of Theorem 6.3. If $n^\beta \geqslant n \log^6 n$, then we can use the algorithm of [AOW15] as stated in Theorem 2.17 to prove our statement. Thus, we assume $n^\beta \leqslant n \log^6 n$. Let S be some set, say $\{1, \ldots, \ell\}$, where $\ell := n^c$ is chosen so that $n^c = \rho n$. Hence, when $k \geqslant 4$, the value of $\varepsilon_{6.5}$ is $\widetilde{O}\left(\frac{1}{\rho^{(k-2)/4} n^{\beta/2}}\right)$. By Lemma 6.5 we can certify with high probability that simultaneously for all assignments y to variables in S, the induced 2XOR formula is $\mathcal{I}_{S,y,2}$ is $\left(\frac{1}{2} + \widetilde{O}\left(\frac{1}{\rho^{(k-2)/4} n^{\beta/2}}\right)\right)$-positive. The underlying graph in $\mathcal{I}_{S,y,2}$ remains the same as we vary y. The expected number of edges in this graph is $\rho^{k-2} n^{\frac{k-2}{2}+\beta} \gg n^{\frac{k}{2}-1} \log^6 n$ and hence the number of edges concentrates around its expectation. And thus, the average degree Δ is $\rho^{k-2} n^{\frac{k-4}{2}+\beta} \gg \log^6 n$. Further, the underlying graph is distributed exactly the same as in the hypothesis of Lemma 6.7. Thus, when $k \geqslant 4$, an application of Lemma 6.7 tells us that we can certify with high probability that any ρ-biased assignment $x \in \{\pm 1\}^{n \setminus S}$ must violate at least

$$\frac{\rho^2}{2} - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/2} n^{(k-4)/4} n^{\beta/2}}\right) - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/4} n^{\beta/2}}\right) - \frac{2}{n}$$

fraction of the constraints on the induced formula $\mathcal{I}_{S,x_S,2}$, and consequently must violate at least

$$\rho^{k-2}\left(\frac{\rho^2}{2} - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/2} n^{(k-4)/4} n^{\beta/2}}\right) - \widetilde{O}\left(\frac{1}{\rho^{(k-2)/4} n^{\beta/2}}\right) - \frac{2}{n}\right)$$

fraction of the constraints in $\mathcal{I}$. Thus, any assignment x that avoids violating at least the above fraction of constraints must satisfy $\left|\sum_{i \in [n] \setminus S} x_i\right| \leqslant \rho n$. Since $|S| = \rho n$, x must be 2ρ-biased. $\qquad\square$

7 Dimension-based count certification

We begin by upper bounding the number of Boolean vectors close to an arbitrary linear subspace.

Theorem 7.1. *Let V be a linear subspace of dimension αn in $\mathbb{R}^n$ for some $\alpha \in (0,1)$. For any $\varepsilon \in (0, 1/4)$, the number of Boolean vectors in $\left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ that are ε-close to V is upper bounded by $2^{(H_2(4\varepsilon^2) + \alpha \log \frac{3}{\varepsilon})n}$.*

Proof. Let T be the set of vectors in $\left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ that are ε-close to V, and let $B_V := B_1(0) \cap V$ be the unit ball in V. We take an ε-net $\mathcal{N}_\varepsilon$ of B_V. Every $x \in T$ is ε-close to a point in B_V (namely $\Pi_V x$), so by the triangle inequality, every $x \in T$ is 2ε-close to a point in $\mathcal{N}_\varepsilon$.

Next, we bound the number of vectors in any ε-ball.

Claim 7.2. *For any $\varepsilon \in (0, 1/\sqrt{2})$ and vector $u \in \mathbb{R}^n$, there can be at most $2^{H_2(\varepsilon^2)n}$ Boolean vectors in $\left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ contained in the ε-ball $B_\varepsilon(u)$.*

Proof. If there are no Boolean vectors in $B_\varepsilon(u)$ we are done. So assume there is a Boolean vector $x \in B_\varepsilon(u)$. For any Boolean vector x' at Hamming distance at least $\varepsilon^2 n$:

$$\|x - x'\|_2 \geqslant \sqrt{\varepsilon^2 n \cdot \frac{4}{n}} = 2\varepsilon,$$

34

which means $x' \notin B_\varepsilon(u)$ and any Boolean vector in $B_\varepsilon(u)$ must be Hamming distance at most $\varepsilon^2 n$ from x, of which there are:

$$\sum_{i=0}^{\varepsilon^2 n} \binom{n}{i} \leqslant 2^{H_2(\varepsilon^2)n},$$

where the inequality follows from the assumption $\varepsilon < 1/\sqrt{2}$. $\qquad\square$

A standard volume argument shows that there exists an ε-net with cardinality $|\mathcal{N}_\varepsilon| \leqslant (\frac{3}{\varepsilon})^{\alpha n}$ (see, for example, [Ver18, Corollary 4.2.13]). Finally, we bound the cardinality of T. Since $T \subseteq \bigcup_{u \in \mathcal{N}_\varepsilon} B_{2\varepsilon}(u)$,

$$|T| \leqslant |\mathcal{N}_\varepsilon| \cdot 2^{H_2(4\varepsilon^2)n} \leqslant 2^{(H_2(4\varepsilon^2) + \alpha \log \frac{3}{\varepsilon})n}. \qquad\square$$

Remark 7.3. The upper bound of Theorem 7.1 is almost tight. For some small constants $\alpha, \varepsilon > 0$, consider the subcube $T = \left\{ \pm\frac{1}{\sqrt{n}}(\vec{1}, y) : y \in \{\pm 1\}^{\alpha n} \right\} \subset \left\{ \pm\frac{1}{\sqrt{n}} \right\}^n$, and let $V = \text{span}(T)$. Clearly, $|T| = 2^{\alpha n + 1}$ and $\dim(V) = \alpha n + 1$. For any $x \in T$, there are

$$\sum_{i=1}^{\frac{\varepsilon^2 n}{4}} \binom{(1-\alpha)n}{i} \geqslant 2^{H_2(\frac{\varepsilon^2}{4(1-\alpha)})(1-\alpha)n - O(\log n)}$$

number of Boolean vectors ε-close to x and differ from x in the first $(1-\alpha)n$ coordinates. Multiplied by $|T|$, the number of Boolean vectors ε-close to V is at least $2^{(H_2(\frac{\varepsilon^2}{4(1-\alpha)}) + \Omega(\alpha))n - O(\log n)}$. This shows that the exponent in the upper bound of Theorem 7.1 is tight up to a $\log(1/\varepsilon)$ factor.

The idea of bounding the number of structured vectors close to a subspace will be the main theme in the following sections. Specifically, for a matrix $M \in \mathbb{R}^{n \times n}$, if $x^\top M x \approx \lambda_{\max}(M)\|x\|_2^2$, then x must be close to the top eigenspace of M. This allows us to apply Theorem 7.1 (or a variant of it for independent sets). For a linear subspace V, we denote $\Pi_{V^\perp}$ as the projection matrix to the orthogonal subspace $V^\perp$, i.e. $\|\Pi_{V^\perp} x\|_2$ is the distance from x to V. We will first prove the following useful lemma.

Lemma 7.4. *Let M be a symmetric $n \times n$ matrix with eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$ and orthonormal eigenvectors $v_1, \ldots, v_n$ such that $\lambda_1 > 0$. Further, let $V := \text{span}\{v_i : \lambda_i \geqslant \lambda_1(1-\delta)\}$ for some constant $\delta \in (0, 1)$. Suppose $x \in \mathbb{R}^n$ satisfies $x^\top M x \geqslant \lambda_1(1-\eta)\|x\|_2^2$ for some $\eta > 0$, then $\|\Pi_{V^\perp} x\|_2 \leqslant \sqrt{\frac{\eta}{\delta}} \|x\|_2$.*

Proof. Let $\alpha := \frac{1}{n} \dim(V)$. Let $x = \sum_{i=1}^n \widehat{x}_i v_i$ written in the eigenvector basis. Clearly, we have $\sum_{i=1}^n \widehat{x}_i^2 = \|x\|_2^2$ and $\sum_{i=\alpha n+1}^n \widehat{x}_i^2 = \|\Pi_{V^\perp} x\|_2^2$.

$$x^\top M x = \sum_{i=1}^n \lambda_i \widehat{x}_i^2 \leqslant \sum_{i=1}^{\alpha n} \lambda_1 \widehat{x}_i^2 + \sum_{i=\alpha n+1}^n \lambda_1(1-\delta)\widehat{x}_i^2 \leqslant \lambda_1 \|x\|_2^2 - \delta \lambda_1 \|\Pi_{V^\perp} x\|_2^2.$$

Along with $x^\top M x \geqslant \lambda_1(1-\eta)\|x\|_2^2$, we conclude $\|\Pi_{V^\perp} x\|_2^2 \leqslant \frac{\eta}{\delta}\|x\|_2^2$. $\qquad\square$

35

7.1 Sherrington-Kirkpatrick

Recall the Sherrington-Kirkpatrick (SK) problem: given M sampled from $\mathsf{GOE}(n)$, compute

$$\mathsf{OPT}(M) = \max_{x \in \{\pm 1\}^n} x^\top M x.$$

A simple spectral refutation algorithm gives a *spectral bound* of $\mathsf{OPT}(M) \leqslant (2 + o(1))n^{3/2}$. We will certify an upper bound on the number of assignments achieving value close to the spectral bound.

Theorem 7.5. *Let $M \sim \mathsf{GOE}(n)$. Given $\eta \in (0, \eta_0)$ for some universal constant η_0, there is an algorithm certifying with high probability that at most $2^{O(\eta^{3/5} \log \frac{1}{\eta})n}$ assignments $x \in \{\pm 1\}^n$ satisfy $x^\top M x \geqslant 2(1 - \eta)n^{3/2}$.*

Proof. The algorithm is as follows.

(1) Choose $\delta = \eta^{2/5}$, and let $\varepsilon = \sqrt{\frac{\eta}{\delta}} = \eta^{3/10}$.

(2) Compute the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ of M, and compute $\alpha = \frac{1}{n}|\{i : \lambda_i \geqslant 2(1 - \delta)\sqrt{n}\}|$. Check that $|\frac{\lambda_1}{\sqrt{n}} - 2| < n^{-1/4}$; output 2^n if this fails.

(3) Output $2^{(H_2(16\varepsilon^2) + \alpha \log \frac{3}{\varepsilon})n}$.

Let $v_1, \ldots, v_n$ be the corresponding eigenvectors of M, and let $V_\delta := \mathrm{span}\{v_i : \lambda_i \geqslant 2(1 - \delta)\sqrt{n}\}$ be the top eigenspace of dimension αn. By the semicircle law (Lemma 2.22), with high probability $\alpha \leqslant O(\delta^{3/2})$. Moreover, the check in step (2) will succeed with high probability due to Fact 2.20, and thus $\lambda_1 \leqslant (2 + o(1))\sqrt{n}$.

Next, consider a normalized Boolean vector $y \in \left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ such that $y^\top M y \geqslant 2(1 - \eta)\sqrt{n}$. By Lemma 7.4, we have $\|\Pi_{V_\delta^\perp} y\|_2 \leqslant \sqrt{\frac{\eta}{\delta}} + o(1) = \varepsilon + o(1)$, i.e. y is 2ε-close to V_δ.

By Theorem 7.1, the number of $y \in \left\{\pm \frac{1}{\sqrt{n}}\right\}^n$ that are 2ε-close to a αn-dimensional subspace is

$$2^{(H_2(16\varepsilon^2) + \alpha \log \frac{3}{\varepsilon})n}.$$

Finally, we use the fact that $H_2(p) \leqslant 2p \log_2 \frac{1}{p}$ for $p \leqslant \frac{1}{2}$. Thus, for small enough $\eta < \eta_0$, $H_2(16\varepsilon^2) \leqslant O(\varepsilon^2 \log \frac{1}{\varepsilon})$. Since $\alpha \leqslant O(\delta^{3/2})$, our choice $\delta = \eta^{2/5}$ gives us an upper bound

$$2^{O(\eta^{3/5} \log \frac{1}{\eta})n}.$$

This completes the proof. $\qquad\square$

7.2 Independent sets

Recall that the best known *certifiable* upper bound of the largest independent set size (the independence number) in a random d-regular graph is by the smallest eigenvalue of the adjacency matrix (known as Hoffman's bound). We first present a proof of the certification, which will give us some insights for the counting problem.

For a set $S \subseteq V$, let $1_S \in \{0,1\}^n$ be the indicator vector of S. We will heavily use the "centered" vector $y_S \in \mathbb{R}^n$ defined as follows,

$$y_S = 1_S - \frac{\langle 1_S, \vec{1} \rangle}{n} \vec{1}, \quad y_S(i) = \begin{cases} 1 - \frac{|S|}{n} & i \in S, \\ -\frac{|S|}{n} & i \notin S. \end{cases}$$

In words, y_S is the projection of 1_S onto the subspace orthogonal to the all-ones vector. Crucially, we have $\langle y_S, \vec{1} \rangle = 0$ and $\|y_S\|_2^2 = |S|(1 - \frac{|S|}{n})$.

For an adjacency matrix A, let $\overline{A} := A - \frac{d}{n}J$ be the "de-meaned" adjacency matrix, i.e., the matrix obtained by projecting away the Perron eigenvector. We will mainly use $\overline{A}$ because its eigenvalues are well-distributed, whereas A has an outlier eigenvalue d. However, note that they have the same minimum eigenvalue: $\lambda_{\min}(A) = \lambda_{\min}(\overline{A}) > 0$. The following lemma, widely known as Hoffman's bound (see also [FO05]), relates the independence number to $-\lambda_{\min}(A)$.

Lemma 7.6 (Certifiable upper bound on independence number). *Let G be a d-regular graph with n vertices, let A be the adjacency matrix, and let $\lambda := -\lambda_{\min}(A)$. Suppose $S \subseteq [n]$ is an independent set, then*

$$|S| \leqslant \frac{\lambda}{d + \lambda} n.$$

Proof. Since S is an independent set, $1_S^\top A 1_S = 0$. Further, we have $1_S^\top J 1_S = |S|^2$. Thus,

$$1_S^\top \left(\frac{d}{n}J - A \right) 1_S = \frac{d}{n}|S|^2.$$

Denote $\overline{A} = A - \frac{d}{n}J$. Since $\vec{1}$ is in the kernel of $\overline{A}$, by the definition of y_S, we have $1_S^\top(-\overline{A})1_S = y_S^\top(-\overline{A})y_S$. Thus,

$$\frac{d}{n}|S|^2 = y_S^\top(-\overline{A})y_S \leqslant \lambda \|y_S\|_2^2 = \lambda \cdot \frac{|S|(n - |S|)}{n}, \tag{5}$$

where $\lambda = \lambda_{\max}(-\overline{A}) = -\lambda_{\min}(A)$ since A and $\overline{A}$ have the same minimum eigenvalue. This gives us the upper bound. $\qquad\square$

For random d-regular graphs, $\lambda \leqslant 2\sqrt{d-1} + o(1)$ due to Friedman's Theorem (Theorem 2.23). Denote $r_d := \frac{2\sqrt{d-1}}{d}$ and $C_d := \frac{r_d}{1+r_d}$. Note that for all $d \geqslant 2$, $r_d \leqslant 1$, thus $C_d \leqslant \frac{1}{2}$. Lemma 7.6 allows us to certify with high probability that all independent sets in a random d-regular graph have size $\leqslant C_d n(1 + o(1))$.

We then turn to the problem of counting large independent sets.

Theorem 7.7. *Let $d \geqslant 3$ be a constant. For a random d-regular graph G on n vertices, given $\eta \in (0, \eta_0)$ for some universal constant η_0, there is an algorithm certifying with high probability that there are at most $2^{O(\eta^{3/5} \log \frac{1}{\eta})n}$ independent sets of size $C_d(1 - \eta)n$.*

Remark 7.8. A trivial upper bound is $\binom{n}{C_d(1-\eta)n} \approx 2^{H_2(C_d(1-\eta))n} = 2^{\Omega_d(n)}$ for small η. Thus, for constant d and small η, our upper bound is significantly better than this trivial bound.

Let $\mathcal{S}_\eta(G)$ be the set of independent sets of size at least $C_d(1 - \eta)n$ in a d-regular graph G, and let $\mathcal{Y}_\eta(G) = \{y_S : S \in \mathcal{S}_\eta(G)\}$. Clearly, we have $|\mathcal{S}_\eta(G)| = |\mathcal{Y}_\eta(G)|$. To bound $|\mathcal{Y}_\eta(G)|$ for a

random d-regular graph G, we follow the same idea of bounding the number of vectors close to a subspace. We first show that any $y_S \in \mathcal{Y}_\eta(G)$ is close to the top eigenspace of $-\overline{A}$.

Let $v_1, \ldots, v_n$ be the eigenvectors of $-\overline{A}$, and let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues. For a constant $\delta \in (0,1)$, let $V_\delta = \text{span}\{v_i : \lambda_i \geqslant dr_d(1-\delta)\}$.

Lemma 7.9. *Suppose $\lambda_{\max}(-\overline{A}) \leqslant dr_d(1+o(1))$. Let $\eta \in (0,1)$. Then, for any independent set S of size $C_d(1-\eta)n$, the vector y_S satisfies $\|\Pi_{V_\delta^\perp} y_S\|_2 \leqslant \sqrt{\frac{2\eta}{\delta}}\|y_S\|_2$.*

Proof. Recall that $\|y_S\|_2^2 = |S| \left(1 - \frac{|S|}{n}\right)$ and that $y_S^\top(-\overline{A})y_S = \frac{d}{n}|S|^2$. Hence:

$$\frac{y_S^\top(-\overline{A})y_S}{\|y_S\|_2^2} = \frac{d|S|}{n\left(1 - \frac{|S|}{n}\right)} = \frac{dC_d(1-\eta)}{1 - C_d(1-\eta)} = \frac{dr_d(1-\eta)}{1+\eta r_d} \geqslant dr_d(1-2\eta)$$

where the last inequality uses $\frac{1}{1+t} \geqslant 1-t$ for $t \in (0,1)$ and $r_d \leqslant 1$.

Since $\lambda_{\max}(-\overline{A}) \leqslant dr_d(1+o(1))$, the statement follows from Lemma 7.4. $\qquad\square$

Next, similar to Lemma 7.2, we upper bound the number of $y_S \in \mathcal{Y}_\eta(G)$ that can be in the same $\varepsilon\sqrt{n}$-ball. Here, we have a factor of $\sqrt{n}$ in the radius because each $y_S \in \mathcal{Y}_\eta(G)$ has norm $\Theta(\sqrt{n})$.

Lemma 7.10. *Let $\varepsilon > 0$ such that $\varepsilon < \frac{1}{4\sqrt{2}}$, and let G be a d-regular graph whose maximum independent set is bounded by $\frac{n}{2}$. There can be at most $2^{(32\varepsilon^2 \log_2 \frac{1}{\varepsilon})n}$ vectors in $\mathcal{Y}_\eta(G)$ contained in any $(\varepsilon\sqrt{n})$-ball.*

Proof. For any sets $S, T \subseteq [n]$ and $|S|, |T| \leqslant \frac{n}{2}$, we have $\|y_S - y_T\|_2^2 \geqslant |S\Delta T|/4$ where $S\Delta T$ is the symmetric difference. Thus, if $|S\Delta T| > 16\varepsilon^2 n$, then they cannot be in the same $\varepsilon\sqrt{n}$-ball.

Pick any set S for which $y_S \in \mathcal{Y}_\eta(G)$ (if no such set exists we are trivially done). Then the number of sets T such that $|S\Delta T| \leqslant 16\varepsilon^2 n$ is at most $\binom{n}{16\varepsilon^2 n}$, which is at most $2^{H_2(16\varepsilon^2)n}$, which is at most $2^{(32\varepsilon^2 \log_2 \frac{1}{\varepsilon})n}$ since $H_2(p) \leqslant 2p \log_2 \frac{1}{p}$ when $p \leqslant \frac{1}{2}$. This completes the proof. $\qquad\square$

Now, we are ready to prove Theorem 7.7.

Proof of Theorem 7.7. Recall that we defined $r_d := \frac{2\sqrt{d-1}}{d}$ and $C_d := \frac{r_d}{1+r_d}$, and the de-meaned adjacency matrix $\overline{A} = A - \frac{d}{n}J$. Further, define $c_d := \frac{\sqrt{r_d}}{1+r_d}$. Note that $c_d < \sqrt{C_d} < 1$. The algorithm is as follows.

(1) Choose $\delta = \eta^{2/5}$, $\varepsilon = \sqrt{\frac{2\eta}{\delta}}$, and $\varepsilon_d' = 2\varepsilon c_d$.

(2) Compute the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ of $-\overline{A}$, and compute $\alpha = \frac{1}{n}|\{i : \lambda_i \geqslant 2\sqrt{d-1}(1-\delta)\}|$. Check that $|\lambda_1 - 2\sqrt{d-1}| < \frac{\log\log n}{\log n}$; output 2^n if this fails.

(3) Output $2^{(32\varepsilon_d'^2 \log(1/\varepsilon_d') + \alpha \log(3/\varepsilon))n}$.

Let $v_1, \ldots, v_n$ be the eigenvectors of $-\overline{A}$, and let $V_\delta := \text{span}\{v_i : \lambda_i \geqslant 2\sqrt{d-1}(1-\delta)\}$ be the space spanned by the top αn eigenvectors. First, the check in step (2) will succeed with high probability due to Friedman's Theorem (Theorem 2.23). Thus, $\lambda_{\max}(-\overline{A}) = 2\sqrt{d-1}(1+o(1)) = dr_d(1+o(1))$. Moreover, by the Kesten–McKay law (Lemma 2.25), $\alpha \leqslant O(\delta^{3/2})$ with high probability.

38

$\lambda_{\max}(-\overline{A})$ certifies that the maximum independent set has size at most $(1+o(1))C_d n$. Then, we have that every $y_S \in \mathcal{Y}_\eta(G)$ has norm $\|y_S\|_2^2 \leqslant (1+o(1))nC_d(1-C_d) = (1+o(1))n\frac{r_d}{(1+r_d)^2} = (1+o(1))c_d^2 n$. Since $d \geqslant 3$, $(1+o(1))C_d \leqslant \frac{1}{2}$ and by Lemma 7.9, every $y_S \in \mathcal{Y}_\eta(G)$ satisfies

$$\|\Pi_{V_\delta^\perp} y\|_2 \leqslant \sqrt{\frac{2\eta}{\delta}}\|y_S\|_2 \leqslant \varepsilon(1+o(1))c_d\sqrt{n}.$$

Thus, $\mathcal{Y}_\eta(G)$ is contained in the centered ball of radius $2c_d\sqrt{n}$ in $\mathbb{R}^n$, and is within distance $2\varepsilon c_d\sqrt{n}$ from the subspace V_δ.

We take an ε-net $\mathcal{N}_\varepsilon$ of $B_1(0) \cap V_\delta$, the unit ball within V_δ, and scale it by $2c_d\sqrt{n}$. Then $\mathcal{Y}_\eta(G) \subseteq \bigcup_{u \in \mathcal{N}_\varepsilon} B_{\varepsilon_d'\sqrt{n}}(u)$. By Lemma 7.10, each $(\varepsilon_d'\sqrt{n})$-ball contains at most

$$2^{(32\varepsilon_d'^2 \log \frac{1}{\varepsilon_d'})n}$$

vectors in $\mathcal{Y}_\eta(G)$, provided that $\varepsilon_d' < \frac{1}{4\sqrt{2}}$, which holds as long as $\eta < \eta_0$ for some universal constant η_0.

The cardinality of $\mathcal{N}_\varepsilon$ is bounded by $\left(\frac{3}{\varepsilon}\right)^{\alpha n} \leqslant 2^{(\alpha \log \frac{3}{\varepsilon})n}$. Since $\alpha \leqslant O(\delta^{3/2})$ and $c_d < 1$, our choice $\delta = \eta^{2/5}$ gives us an overall upper bound of

$$2^{O(\eta^{3/5} \log \frac{1}{\eta})n},$$

which completes the proof. $\qquad\square$

8   Hardness evidence

In this section, we provide hardness evidence for some of the algorithmic problems we consider. In these cases, improving slightly or significantly on our bounds would also result in improved algorithms for the refutation versions of these problems. While we don't make confident claims of improvements on the refutation problems being computationally intractable, algorithms for them would certainly bypass several known barriers.

8.1   Refutation-to-certified-counting reduction for kXOR

We first address the problem of counting solutions to a random kXOR instance. Let $\mathcal{I} \sim \mathcal{H}_k^n(m)$ be a random kXOR instance on $m = \Delta n$ clauses.

Theorem 8.1. *If there is an efficient algorithm that with high probability certifies a bound of $\exp\left(\frac{\eta n}{10k}\right)$ on the number of $(1-\eta)$-satisfying assignments to $\mathcal{I}$, then there is an efficient algorithm that with high probability can certify that $\mathcal{I}$ has no $(1-\eta/2)$-satisfying assignments.*

Proof. Assume we ran the algorithm from the hypothesis of the theorem statement on $\mathcal{I}$, and obtained a bound on the count of $\exp\left(\frac{\eta n}{10k}\right) \leqslant 2^{\frac{\eta n}{4k}}$. Take any set $S \subseteq [n]$ of size $\ell := \frac{\eta n}{3k}$, say $\{1, \ldots, \ell\}$. If the number of clauses with at least one variable in S is at most $\frac{\eta m}{2}$ (which holds with high probability), we output "$\mathcal{I}$ has no $(1-\eta/2)$-satisfying assignments". Now, if there exists a $(1-\eta/2)$-satisfying assignment x to $\mathcal{I}$, then any x' which differs from x on a subset of indices in S must be $(1-\eta)$-satisfying. However, since there are $2^{\frac{\eta n}{3k}}$ choices for x' but a contradictory bound of $2^{\frac{\eta n}{4k}}$ on the number of such strings, there cannot be a $(1-\eta/2)$ satisfying assignment. $\qquad\square$

Remark 8.2. If $n^\varepsilon \ll \Delta \ll n^{k/2-1}$, and $\eta n = \frac{n^{1+\varepsilon}}{\Delta^{1/(k-2)}}$ for some $\varepsilon > 0$, we are in a regime where:

1. there are no known algorithms to certify that there are no $(1 - \eta/2)$-satisfying assignments, and

2. our algorithm from Theorem 4.4 certifies a bound of $\exp(O(\eta n \log n))$.

Thus, beating our algorithm in the above regime of η by more than a logarithmic factor in the exponent would break a long-standing algorithmic barrier.

8.2 Refutation-to-certified-counting reduction for Independent Set

In this section, we show evidence that our upper bound for independence number in a random d-regular graph (Theorem 7.7) cannot be improved significantly.

First, recall that we defined $r_d := \frac{2\sqrt{d-1}}{d}$ and $C_d := \frac{r_d}{1+r_d}$. Moreover, the best known *certifiable* upper bound (Hoffman's bound) for the independence number of a random d-regular graph is $C_d n$. It is widely believed that beating this bound (getting an upper bound of $(1 - \varepsilon)C_d n$ for some constant ε) requires some new algorithmic ideas.

Theorem 8.3. *Let G be a random d-regular graph. Given constant $\eta \in (0, 1/2)$, if there is an efficient algorithm that with high probability certifies a bound of $\exp\left(\frac{C_d}{4}\eta \log(1/\eta)n\right)$ on the number of independent sets of size $C_d(1 - \eta)n$, then there is an algorithm that with high probability certifies that G has no independent set of size $(1 - \eta/2)C_d n$.*

Proof. Assume we have an algorithm that obtains a bound of $\exp\left(\frac{C_d}{4}\eta \log(1/\eta)n\right)$. Further, assume that there is an independent set S of size $(1 - \eta/2)C_d n$. Then, by the fact that any subset of S is also an independent set, the number of independent sets of size $(1 - \eta)C_d n$ must be at least

$$\binom{(1 - \frac{\eta}{2})C_d n}{(1 - \eta)C_d n} = \binom{(1 - \eta/2)C_d n}{\frac{\eta}{2}C_d n} \geqslant 2^{H_2(\frac{\eta/2}{1-\eta/2})(1-\eta/2)C_d n - O(\log n)} > \exp\left(\frac{C_d}{4}\eta \log(1/\eta)n\right),$$

using the fact that $H_2(p) \geqslant p \log(1/p)$ and $\eta < 1/2$. This contradicts the upper bound. □

Remark 8.4. For constant $d \geqslant 3$ and a small constant $\eta > 0$, Theorem 7.7 certifies a bound of $\exp\left(O(\eta^{3/5}\log(1/\eta)n)\right)$. However, beating a bound of $\exp\left(O(\eta \log(1/\eta)n)\right)$ would also beat the Hoffman's bound by a factor of $(1 - \eta/2)$. We conjecture that there is an efficient algorithm to certify an upper bound matching the lower bound, but we leave that as an open direction.

8.3 Approach for proving low-degree hardness for certified counts in kSAT

We note that the counting hardness evidence for kXOR we provide is only evidence for optimality of our algorithms for counting *approximately satisfying assignments* to CSPs. It is desirable to give evidence suggesting that our algorithm's guarantees for certifying bounds on, say, the number of exactly satisfying assignments to a random kSAT formula are tight. One approach for doing so is to construct a *planted distribution* with an appropriately large number of kSAT assignments, and prove that it is impossible for low-degree polynomials to distinguish this planted distribution from random kSAT instances. Here, we provide a blueprint for constructing such a planted distribution:

1. Sample a random k-uniform hypergraph with a planted independent set S (where S is an independent set if no hyperedge contains $\geqslant 2$ vertices in S).

2. Sample a random assignment x on variables outside S.

3. Place random negations c so the clauses completely outside S are satisfied by x as 3XOR, and for hyperedges $U = (u, v, w)$ with $u \in U$, the clause $(c_{U,v}v, c_{U,w}w)$ is satisfied as 2XOR.

One of the challenges is in planting an independent set that doesn't "stand out" to spectral or low-degree polynomial distinguishers. To this end we suspect that the techniques from [BBK+20] might be useful.

Acknowledgments

We would like to thank Pravesh Kothari, Prasad Raghavendra, and Tselil Schramm for their encouragement and thorough feedback on an earlier draft. We would also like to thank Tselil Schramm for enlightening discussions on refuting random CSPs.

References

[ABW10]   Benny Applebaum, Boaz Barak, and Avi Wigderson. Public-key cryptography from different assumptions. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 171–180, 2010. 1

[AC88]   Noga Alon and Fan RK Chung. Explicit construction of linear sized tolerant networks. *Discrete Mathematics*, 72(1-3):15–19, 1988. 15

[ACO08]   Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 793–802. IEEE, 2008. 11

[Alo86]   Noga Alon. Eigenvalues and expanders. *Combinatorica*, 6(2):83–96, 1986. 19

[Alo91]   Noga Alon. Independent sets in regular graphs and sum-free subsets of finite groups. *Israel journal of mathematics*, 73(2):247–256, 1991. 12

[Alo09]   Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability & Computing*, 18(1-2):3, 2009. 7, 27, 29

[AM85]   Noga Alon and Vitali D Milman. λ_1, isoperimetric inequalities for graphs, and superconcentrators. *Journal of Combinatorial Theory, Series B*, 38(1):73–88, 1985. 15

[AM02]   Dimitris Achlioptas and Cristopher Moore. The asymptotic order of the random k-SAT threshold. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 779–788. IEEE, 2002. 11

[AOW15]   Sarah R Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 689–708. IEEE, 2015. 1, 4, 5, 8, 10, 11, 16, 18, 32, 34

[AP03]   Dimitris Achlioptas and Yuval Peres. The threshold for random k-sat is $2^k(\ln 2 - o(k))$. In *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 223–231, 2003. 11

[ART06]   Dimitris Achlioptas and Federico Ricci-Tersenghi. On the solution-space geometry of random constraint satisfaction problems. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 130–139, 2006. 11

[Bar16]   Alexander Barvinok. *Combinatorics and complexity of partition functions*, volume 9. Springer, 2016. 12

[BBK⁺20]   Afonso S Bandeira, Jess Banks, Dmitriy Kunisky, Cristopher Moore, and Alexander S Wein. Spectral planting and the hardness of refuting cuts, colorability, and communities in random graphs. *arXiv preprint arXiv:2008.12237*, 2020. 41

[BCK15]   Boaz Barak, Siu On Chan, and Pravesh K Kothari. Sum of squares lower bounds from pairwise independence. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 97–106, 2015. 11

[BGG⁺19]   Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Daniel Stefankovic. Approximation via correlation decay when strong spatial mixing fails. *SIAM Journal on Computing*, 48(2):279–349, 2019. 1, 11

[BH11]   Andries E Brouwer and Willem H Haemers. *Spectra of graphs*. Springer Science & Business Media, 2011. 9

[BJK05]   Andrei Bulatov, Peter Jeavons, and Andrei Krokhin. Classifying the complexity of constraints using finite algebras. *SIAM journal on computing*, 34(3):720–742, 2005. 1

[Bol81]   Béla Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, pages 433–436, 1981. 9

[Bor15]   Charles Bordenave. A new proof of Friedman's second eigenvalue Theorem and its extension to random lifts. *arXiv preprint arXiv:1502.04482*, 2015. 19

[CDK⁺19]   Matthew Coulson, Ewan Davies, Alexandra Kolla, Viresh Patel, and Guus Regts. Statistical physics approaches to Unique Games. *arXiv preprint arXiv:1911.01504*, 2019. 12

[CO07]   Amin Coja-Oghlan. On the Laplacian eigenvalues of $G_{n,p}$. *Combinatorics, Probability & Computing*, 16(6):923, 2007. 15

[CO17]   Amin Coja-Oghlan. Belief propagation guided decimation fails on random formulas. *Journal of the ACM (JACM)*, 63(6):1–55, 2017. 12

[COGL07]   Amin Coja-Oghlan, Andreas Goerdt, and André Lanka. Strong refutation heuristics for random *k*-SAT. *Combinatorics, Probability & Computing*, 16(1):5, 2007. 1

[COMR20]   Amin Coja-Oghlan, Noëla Müller, and Jean B Ravelomanana. Belief Propagation on the random *k*-SAT model. *arXiv preprint arXiv:2011.02303*, 2020. 12

[COP13]   Amin Coja-Oghlan and Konstantinos Panagiotou. Going after the *k*-SAT threshold. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 705–714, 2013. 11

[COP16]   Amin Coja-Oghlan and Konstantinos Panagiotou. The asymptotic *k*-SAT threshold. *Advances in Mathematics*, 288:985–1068, 2016. 11

[DG00]   Martin Dyer and Catherine Greenhill. On Markov chains for independent sets. *Journal of Algorithms*, 35(1):17–49, 2000. 12

[DLSS14]   Amit Daniely, Nati Linial, and Shai Shalev-Shwartz. From average case complexity to improper learning complexity. In *Proceedings of the forty-sixth annual ACM symposium on Theory of computing*, pages 441–448, 2014. 1

[DSS15]   Jian Ding, Allan Sly, and Nike Sun. Proof of the satisfiability conjecture for large *k*. In *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, pages 59–68, 2015. 11, 12

[EL73]   Paul Erdős and László Lovász. Problems and results on 3-chromatic hypergraphs and some related questions. In *Colloquia Mathematica Societatis Janos Bolyai 10. Infinite and Finite Sets, Keszthely (Hungary)*. Citeseer, 1973. 11

[Erd11]   László Erdős. Universality of Wigner random matrices: a survey of recent results. *Russian Mathematical Surveys*, 66(3):507, 2011. 19

[Fei02]   Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 534–543, 2002. 1, 4, 20

[FGYZ20]   Weiming Feng, Heng Guo, Yitong Yin, and Chihao Zhang. Fast sampling and counting *k*-SAT solutions in the local lemma regime. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 854–867, 2020. 11

[FHY20]   Weiming Feng, Kun He, and Yitong Yin. Sampling Constraint Satisfaction Solutions in the Local Lemma Regime. *arXiv preprint arXiv:2011.03915*, 2020. 11

[FK16]   Alan Frieze and Michał Karoński. *Introduction to random graphs*. Cambridge University Press, 2016. 22

[FO05]   Uriel Feige and Eran Ofek. Spectral techniques applied to sparse random graphs. *Random Structures & Algorithms*, 27(2):251–275, 2005. 9, 37

[FO07]   Uriel Feige and Eran Ofek. Easily refutable subformulas of large random 3CNF formulas. *Theory of Computing*, 3(1):25–43, 2007. 1, 4, 20

[Fri08]   Joel Friedman. *A proof of Alon's second eigenvalue conjecture and related problems*. American Mathematical Soc., 2008. 19

[GGGY19]   Andreas Galanis, Leslie Ann Goldberg, Heng Guo, and Kuan Yang. Counting solutions to random CNF formulas. *arXiv preprint arXiv:1911.07020*, 2019. 1, 12

[GJJ+20] Mrinalkanti Ghosh, Fernando Granha Jeronimo, Chris Jones, Aaron Potechin, and Goutham Rajendran. Sum-of-squares lower bounds for Sherrington-Kirkpatrick via planted affine planes. *arXiv preprint arXiv:2009.01874*, 2020. 8

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

[Gol11] Oded Goldreich. Candidate one-way functions based on expander graphs. In *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 76–87. Springer, 2011. 1

[Gri01] Dima Grigoriev. Complexity of Positivstellensatz proofs for the knapsack. *computational complexity*, 10(2):139–154, 2001. 1, 9, 10

[Hås01] Johan Håstad. Some optimal inapproximability results. *Journal of the ACM (JACM)*, 48(4):798–859, 2001. 1, 14

[HKP19] Christopher Hoffman, Matthew Kahle, and Elliot Paquette. Spectral gaps of random graphs and applications. *International Mathematics Research Notices*, 2019. 15

[HS17] Samuel B Hopkins and David Steurer. Efficient Bayesian estimation from few samples: community detection and related problems. In *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 379–390. IEEE, 2017. 14

[JKR19] Vishesh Jain, Frederic Koehler, and Andrej Risteski. Mean-field approximation, convex hierarchies, and the optimality of correlation rounding: a unified perspective. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, pages 1226–1236, 2019. 12

[JPV20] Vishesh Jain, Huy Tuan Pham, and Thuy Duong Vuong. Towards the sampling Lovász Local Lemma. *arXiv preprint arXiv:2011.12196*, 2020. 11

[JPV21] Vishesh Jain, Huy Tuan Pham, and Thuy-Duong Vuong. On the sampling Lovász Local Lemma for atomic constraint satisfaction problems. *arXiv preprint arXiv:2102.08342*, 2021. 1, 11

[Kah01] Jeff Kahn. An entropy approach to the hard-core model on bipartite graphs. *Combinatorics, Probability & Computing*, 10(3):219, 2001. 12

[KB20] Dmitriy Kunisky and Afonso S Bandeira. A tight degree 4 sum-of-squares lower bound for the Sherrington–Kirkpatrick Hamiltonian. *Mathematical Programming*, pages 1–39, 2020. 8

[Kes59] Harry Kesten. Symmetric random walks on groups. *Transactions of the American Mathematical Society*, 92(2):336–354, 1959. 20

[Kho02] Subhash Khot. On the power of unique 2-prover 1-round games. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 767–775, 2002. 1

[KKMO07]   Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inap-
           proximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on
           Computing*, 37(1):319–357, 2007. 1

[KMOW17]   Pravesh K 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 SIGACT
           Symposium on Theory of Computing*, pages 132–145, 2017. 9, 11

[KMRT⁺07] Florent Krzakała, Andrea Montanari, Federico Ricci-Tersenghi, Guilhem Semerjian,
           and Lenka Zdeborová. Gibbs states and the set of solutions of random constraint
           satisfaction problems. *Proceedings of the National Academy of Sciences*, 104(25):10318–
           10323, 2007. 2, 11, 12

[KOS18]   Pravesh Kothari, Ryan O'Donnell, and Tselil Schramm. SOS lower bounds with hard
           constraints: think global, act local. *arXiv preprint arXiv:1809.01207*, 2018. 7, 30

[LSS19a]   Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. A deterministic algorithm for
           counting colorings with 2Δ colors. In *2019 IEEE 60th Annual Symposium on Founda-
           tions of Computer Science (FOCS)*, pages 1380–1404. IEEE, 2019. 12

[LSS19b]   Jingcheng Liu, Alistair Sinclair, and Piyush Srivastava. The Ising partition function:
           Zeros and deterministic approximation. *Journal of Statistical Physics*, 174(2):287–315,
           2019. 12

[McK81]   Brendan D McKay. The expected eigenvalue distribution of a large regular graph.
           Linear Algebra and its Applications, 40:203–216, 1981. 20

[MMZ05]   Marc Mézard, Thierry Mora, and Riccardo Zecchina. Clustering of solutions in the
           random satisfiability problem. *Physical Review Letters*, 94(19):197205, 2005. 11

[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.
           11

[Moi19]   Ankur Moitra. Approximate counting, the Lovász local lemma, and inference in
           graphical models. *Journal of the ACM (JACM)*, 66(2):1–25, 2019. 1, 11, 12

[Mon19]   Andrea Montanari. Optimization of the sherrington-kirkpatrick hamiltonian. In *Pro-
           ceedings of the 60th IEEE Symposium on Foundations of Computer Science*, pages 1417–
           1433, 2019. 8

[MPZ02]   Marc Mézard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solu-
           tion of random satisfiability problems. *Science*, 297(5582):812–815, 2002. 11

[MRRW77]   Robert McEliece, Eugene Rodemich, Howard Rumsey, and Lloyd Welch. New upper
           bounds on the rate of a code via the delsarte-macwilliams inequalities. *IEEE Transac-
           tions on Information Theory*, 23(2):157–166, 1977. 7, 27

[MRTS08] Andrea Montanari, Federico Ricci-Tersenghi, and Guilhem Semerjian. Clusters of solutions and replica symmetry breaking in random k-satisfiability. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(04):P04004, 2008. 11

[MRX20] Sidhanth Mohanty, Prasad Raghavendra, and Jeff Xu. Lifting sum-of-squares lower bounds: degree-2 to degree-4. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 840–853, 2020. 8

[MS06] Andrea Montanari and Devavrat Shah. Counting good truth assignments of random k-SAT formulae. *arXiv preprint cs/0607073*, 2006. 12

[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. 8

[MZ02] Marc Mézard and Riccardo Zecchina. Random k-satisfiability problem: From an analytic solution to an efficient algorithm. *Physical Review E*, 66(5):056126, 2002. 11

[O'D14] Ryan O'Donnell. *Analysis of boolean functions*. Cambridge University Press, 2014. 16

[Pan13] Dmitry Panchenko. Spin glass models from the point of view of spin distributions. *Annals of Probability*, 41(3A):1315–1361, 2013. 12

[Par79] Giorgio Parisi. Infinite number of order parameters for spin-glasses. *Physical Review Letters*, 43(23):1754, 1979. 8

[Par80] Giorgio Parisi. A sequence of approximated solutions to the SK model for spin glasses. *Journal of Physics A: Mathematical and General*, 13(4):L115, 1980. 8

[PR17] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for partition functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017. 12

[PR19] Han Peters and Guus Regts. On a conjecture of Sokal concerning roots of the independence polynomial. *The Michigan Mathematical Journal*, 68(1):33–55, 2019. 12

[Rag08] Prasad Raghavendra. Optimal algorithms and inapproximability results for every CSP? In *Proceedings of the fortieth annual ACM symposium on Theory of computing*, pages 245–254, 2008. 1

[Ris16] Andrej Risteski. How to calculate partition functions using convex programming hierarchies: provable bounds for variational methods. In *Conference on Learning Theory*, pages 1402–1416. PMLR, 2016. 12

[RL16] Andrej Risteski and Yuanzhi Li. Approximate maximum entropy principles via goemans-williamson with applications to provable variational methods. *Advances in Neural Information Processing Systems*, 29:4628–4636, 2016. 12

[RRS17] Prasad Raghavendra, Satish Rao, and Tselil Schramm. Strongly refuting random CSPs below the spectral threshold. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 121–131, 2017. 10, 11

[Sch78]   Thomas J Schaefer. The complexity of satisfiability problems. In *Proceedings of the tenth annual ACM symposium on Theory of computing*, pages 216–226, 1978. 1

[Sch08]   Grant Schoenebeck. Linear level Lasserre lower bounds for certain k-CSPs. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 593–602. IEEE, 2008. 1, 9, 10

[SK75]   David Sherrington and Scott Kirkpatrick. Solvable model of a spin-glass. *Physical review letters*, 35(26):1792, 1975. 8

[SS12]   Allan Sly and Nike Sun. The computational hardness of counting in two-spin models on *d*-regular graphs. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pages 361–369. IEEE, 2012. 12

[Tal06]   Michel Talagrand. The Parisi formula. *Annals of mathematics*, pages 221–263, 2006. 8

[Tre17]   Luca Trevisan. Lecture notes on graph partitioning, expanders, and spectral methods. *University of California, Berkeley, https: // lucatrevisan. github. io/ books/ expanders-2016. pdf* , 2017. 15

[Tro15]   Joel A Tropp. An introduction to matrix concentration inequalities. *arXiv preprint arXiv:1501.01571*, 2015. 19

[Ver18]   Roman Vershynin. *High-dimensional probability: An introduction with applications in data science*, volume 47. Cambridge university press, 2018. 35

[Vig01]   Eric Vigoda. A note on the Glauber dynamics for sampling independent sets. *the electronic journal of combinatorics*, 8(1):R8, 2001. 12

[Wei06]   Dror Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 140–149, 2006. 12

[Wor99]   Nicholas C Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999. 9

[Zha10]   Yufei Zhao. The number of independent sets in a regular graph. *Combinatorics, Probability and Computing*, 19(2):315–320, 2010. 12

[Zhu20]   Dmitriy Zhuk. A Proof of the CSP Dichotomy Conjecture. *Journal of the ACM (JACM)*, 67(5):1–78, 2020. 1