

Approximate Cauchy-Schwarz Inequality and Improved Bounds for Sherali-Adams Refutation of Semirandom CSPs

Pravesh K. Kothari 

Department of Computer Science, Princeton University, NJ, USA

Andrew D. Lin 

Program in Applied and Computational Mathematics, Princeton University, NJ, USA

Abstract

We formulate an *approximate Cauchy-Schwarz* inequality and show that it is satisfied by solutions to the Sherali-Adams linear programming hierarchy (interpreted as “pseudo-distributions”). As a consequence, we resolve a question left open by the work of O’Donnell and Schramm [13] that they had explicitly attributed to the lack of such an inequality.

A Cauchy-Schwarz inequality is *exactly* satisfied by pseudo-distributions satisfying the constraints of the *sum-of-squares* semidefinite programming hierarchy and already has scores of applications. However, the proof there requires *global* positive semidefiniteness. Our approximate version, on the other hand, relies only on *local* positive semidefiniteness satisfied by the Sherali-Adams pseudo-distributions. Our formulation loses an additive error that scales with the L1 norm of the coefficients of the constituent polynomials, and this loss is asymptotically tight. Our proof is elementary and relies on a simple sampling argument.

As an application, we resolve a question left open in the work of O’Donnell and Schramm that gives a trade-off between constraint density and the Sherali-Adams degree for refuting random constraint satisfaction problems. Specifically, for odd arity CSPs, we show that the constraint density requirement for a given degree can be improved by a polynomial factor in n . Along the way, we observe that by a simple extension, the results in their work extend to a more general semirandom setting.

2012 ACM Subject Classification Theory of computation → Linear programming; Theory of computation → Approximation algorithms analysis

Keywords and phrases Sherali-Adams hierarchy, constraint satisfaction problems, semirandom CSPs, refutation, Cauchy-Schwarz inequality

Digital Object Identifier 10.4230/LIPIcs.APPROX/RANDOM.2026.77

Category RANDOM

1 Introduction

A degree d sum-of-squares (SoS) pseudo-distribution on $\{-1, 1\}^n$ is a signed measure $D : \{-1, 1\}^n \rightarrow \mathbb{R}$ such that the associated pseudo-expectation functional defined for any $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ by $\tilde{\mathbb{E}}_D[f] = \sum_x D(x)f(x)$ satisfies 1) $\tilde{\mathbb{E}}_D[1] = 1$ and 2) $\tilde{\mathbb{E}}_D[f^2] \geq 0$ for every polynomial f of degree $\leq d/2$. Such pseudo-distributions are equivalent to feasible points for SDPs from the SoS semidefinite programming hierarchy. Degree d SoS pseudo-distributions are equivalent to *pseudo-moment matrices* – matrices M indexed by subsets of $[n]$ of size $\leq d/2$ whose entry at (S, T) is given by $\tilde{\mathbb{E}}_D[\prod_{i \in S \Delta T} x_i]$. Pseudo-moment matrices are themselves in one-to-one correspondence with the set of all positive semidefinite matrices that satisfy some natural linear constraints¹. The SoS semidefinite programming

¹ These constraints encode that $M(S, T)$ depends only on $S \Delta T$ and the diagonals of M are 1.



hierarchy, at a high level, can be thought of as solving (in $n^{O(d)}$ time) the problem of finding a pseudo-expectation (equivalently, a pseudo-moment matrix) satisfying some additional linear constraints.

The analysis of algorithms (see the survey [7]) based on the SoS hierarchy requires reasoning about the pseudo-expectations of various polynomials arising in a given problem. Such analyses often rely on generalizing natural analytic inequalities that hold for expectations (w.r.t. probability distributions) to pseudo-expectations (w.r.t. pseudo-distributions). Indeed, the power of the pseudo-expectation view of SoS solutions comes largely from being able to “guess” true inequalities for pseudo-distributions by analogy with actual probability distributions.

The Cauchy-Schwarz Inequality

Perhaps the most fundamental of such general analytic inequalities is the Cauchy-Schwarz inequality:

$$\tilde{\mathbb{E}}_D[fg] \leq \tilde{\mathbb{E}}_D[f^2]^{1/2} \tilde{\mathbb{E}}_D[g^2]^{1/2},$$

which is satisfied whenever D is a degree d pseudo-distribution and f, g are polynomials of degree $\leq d/2$. The Cauchy-Schwarz inequality appears in scores of applications of the sum-of-squares hierarchy, especially in executing the *proofs-to-algorithms* paradigm in algorithm design. The proof (and indeed, the truth) of the inequality relies on the positive semidefiniteness constraint that is often viewed as a *global* constraint on the pseudo-moments.

Such global positive semidefiniteness constraints are not satisfied by solutions to the weaker *Sherali-Adams* linear-programming hierarchy. An R -local Sherali-Adams solution can also be viewed as a degree $2R$ pseudo-distribution $D : \{-1, 1\}^n \rightarrow \mathbb{R}$ with the key difference that the associated pseudo-expectation only satisfies a *local* version of (2). More precisely:

► **Definition 1.1** (*R -local pseudo-distributions and pseudo-expectation*). *An R -local pseudo-distribution over the n -dimensional Boolean hypercube is a function $D : \{-1, 1\}^n \rightarrow \mathbb{R}$ such that the associated pseudo-expectation defined by $\tilde{\mathbb{E}}_D[f] := \sum_{x \in \{-1, 1\}^n} D(x)f(x)$ satisfies:*

1. $\tilde{\mathbb{E}}[1] = 1$,
2. *For every pointwise non-negative $f : \{-1, 1\}^n \rightarrow \mathbb{R}$ such that f depends on only $\leq R$ out of n possible inputs, we have: $\tilde{\mathbb{E}}[f] \geq 0$.*

Equivalently, for every non-negative function f that depends only on $\leq R$ variables, $\tilde{\mathbb{E}}[f] \geq 0$. The corresponding pseudo-moment matrices now satisfy *local* positive semidefiniteness constraints – every principal submatrix that corresponds to pseudo-moments of any set of at most R variables is positive semidefinite. However, the matrix as a whole may not be PSD! This weakening allows encoding the constraints of a degree- d Sherali-Adams pseudo-expectation as an $n^{O(d)}$ -size *linear* (as opposed to a semidefinite) program. The weakening to only satisfying local constraints comes at a cost. Results like the Cauchy-Schwarz inequality no longer hold, provably (see Theorem 2.2)!

Our goal in this work is to formulate an approximate version of the Cauchy-Schwarz inequality that can hold for Sherali-Adams pseudo-distributions. Before diving into our formulation, let us discuss our main motivation:

Sherali-Adams vs Sum-of-Squares

How does the power of the Sherali-Adams linear programming hierarchy compare with that of the SoS SDP hierarchy? On the one hand, lower bounds for constraint satisfaction problems such as Max-Cut [4] show that degree 2 SoS SDPs [9] can achieve drastically better

approximation ratios even as Sherali-Adams LPs of n^ε -degree fail to appreciably beat simply returning a uniformly random assignment. On the other hand, O'Donnell and Schramm [13] and the follow-up work due to Hopkins, Schramm, and Trevisan [11] show that degree n^c (for fixed $c < 1$) Sherali-Adams LPs can qualitatively match the performance of SDPs for Max-Cut! Perhaps more strikingly, they show that $O(1/\varepsilon)$ -degree Sherali-Adams LPs (thus, in polynomial time!) can strongly refute (i.e., certify an upper bound of $1/2 + \varepsilon$ on the normalized value of) random k -XOR instances with $m \geq n^{k/2+\varepsilon}$ constraints **for even k** . This matches, up to n^ε factors for an arbitrarily small ε , the best-known refutation algorithms based on spectral methods [2, 3, 6].

The case of odd k

For random k -XOR refutation when k is odd, O'Donnell and Schramm prove a weaker bound of $m \geq n^{\lceil k/2 \rceil + \varepsilon}$ for the same guarantee that is off of the optimal $n^{k/2}$ by a fixed (i.e., that does not improve as $\varepsilon \rightarrow 0$) polynomial factor in n . This is because the argument in [13] essentially treats an odd k -XOR formula as a $(k+1)$ -XOR and then applies their result for even arity XOR.

As an aside, such a gap between the case of even and odd k s has been a fixture in this line of work. The case of odd k , for which the above bound appears suboptimal, has indeed proved to be more difficult to tackle in all previous analyses [5, 2, 1, 12, 10] even with more powerful classes of spectral methods.

In the work of O'Donnell and Schramm [13], the polynomial factor “discrepancy” can be blamed on the lack of a Cauchy-Schwarz inequality for Sherali-Adams pseudo-distributions. Indeed, the authors themselves remark (see footnote on page 4): “*This is because we do not know how to relate the Sherali-Adams value of the objective function to its square (local versions of the Cauchy-Schwarz argument result in a loss). Such a relation would allow us to apply our techniques immediately to prove that Sherali Adams matches the SOS and spectral performance for odd as well as even k .*”

Our Results

In this work, we prove an approximate Cauchy-Schwarz inequality that holds for Sherali-Adams pseudo-distributions (so a “local” version as in the above quote) but with an additive loss that becomes negligible as the locality/degree of the Sherali-Adams LP increases. As an application, we get the right conjectured bound for refuting random k -XOR formulas for odd k . Along the way, we also observe a natural generalization of the ideas of [13] to the problem of semirandom k -XOR refutation where the set of tuples participating in constraints is arbitrary (i.e., worst-case) while the “right hand sides” are chosen independently at random.

Let us begin with our approximate Cauchy-Schwarz inequality. For any functions f, g , of $x = (x_1, x_2, \dots, x_n)$ and any distribution D over x , $\mathbb{E}_{x \sim D}[f(x)g(x)] \leq \mathbb{E}[f(x)^2]^{1/2} \mathbb{E}[g(x)^2]^{1/2}$. The proof crucially relies on the positive semidefiniteness of the *moment* matrix of the pseudo-distribution. It provably does not hold for Sherali-Adams pseudo-distributions that only satisfy *local* positive semidefiniteness. Indeed, in Section 2.1, we observe the following seemingly drastic failure of the Cauchy-Schwarz inequality.

► **Theorem 1.2** (Sherali-Adams Cannot Satisfy Cauchy-Schwarz Exactly, see Theorem 2.2). *For any R and $n \gg R$, there exists an R -local Sherali-Adams pseudo-distribution on $\{-1, 1\}^n$ such that for the linear polynomial $p(x) = \sum_i x_i$, we have: $\mathbb{E}_D[p(x)^2] = 0$, and, $\mathbb{E}_D[p] \geq (1 - o_R(1)) \frac{n}{\sqrt{R-1}}$.*

Despite this failure, we show that there is a natural way to formulate a correct, approximate Cauchy-Schwarz inequality. To the best of our knowledge, such an approximate formulation has not appeared in prior works. Notice that the above counter-example rules out an approximate Cauchy-Schwarz inequality with a *multiplicative* error. Our formulation has a bound on the *additive* error instead.

► **Theorem 1.3** (Approximate Cauchy-Schwarz Inequality). *Let $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$ be polynomials of degree $\leq d$. Let $\|f\|_1, \|g\|_1$ be the $L1$ norms of the coefficients of f, g , respectively. Then, for every R -local pseudo-distribution $\mathcal{D}$ on $\{-1, 1\}^n$ and any positive integers s, t such that $d(s + t) \leq R$, we have:*

$$\left| \mathbb{E}[fg] \right| \leq \sqrt{\left(1 - \frac{1}{s}\right) \widetilde{\mathbb{E}}[f^2] + \frac{1}{s} \|f\|_1^2} \sqrt{\left(1 - \frac{1}{t}\right) \widetilde{\mathbb{E}}[g^2] + \frac{1}{t} \|g\|_1^2}.$$

► **Remark 1.4.** Note that $\|f\|_1$ and $\|g\|_1$ are basis dependent quantities. This is to be expected as Sherali-Adams (unlike sum-of-squares) is not basis independent. Next, our bounds are written in the $\{\pm 1\}$ basis. They can, of course, be transformed into the $\{0, 1\}$ basis, but the quantities appearing will change. Finally, we present the inequality with s, t as independent parameters (as opposed to setting them both to be $R/2$, say) since we anticipate that in applications, $\|f\|_1, \|g\|_1$ could be different and the freedom to modulate s, t accordingly may lead to better bounds.

As we observe in Section 2.1, Theorem 1.2 shows that the formulation we prove in Theorem 1.3 is tight down to the exact constant as $n/R \rightarrow \infty$.

Applications

We apply Approximate Cauchy-Schwarz to prove tight bounds for CSP refutation.

► **Theorem 1.5** (Sherali-Adams Refutation of Random and Semirandom k -XOR; see Theorems 3.4 and 4.2). *Suppose we have a random or semirandom k -XOR instance on n variables with $n^{k/2+\delta}$ expected (in the random case) or total (in the semirandom case) constraints. Then with high probability, there exists some $R = O_{\varepsilon, \delta, k}(1)$ such that for all R -local Sherali-Adams pseudodistributions $\mathcal{D}$ on $\{-1, 1\}^n$, we have $\mathbb{E}_{\mathcal{D}} \psi(x) \leq 2\varepsilon |\mathcal{H}|$.*

In the case of odd k , we use our approximate Cauchy-Schwarz lemma to show that the standard spectral certificate used to refute random odd k -XOR instances in [2] implies a Sherali-Adams refutation, which allows us to refute with just $n^{k/2+\delta}$ constraints. In addition, we show that we can refute semirandom k -XOR in Sherali-Adams, matching the SOS algorithms from [1]. This was previously unknown for both even and odd k .

2 Sherali-Adams Proof of an Approximate Cauchy-Schwarz Inequality

In this section, we show that Sherali-Adams pseudo-distributions satisfy a natural approximate Cauchy-Schwarz inequality. In the next section, we will see how this inequality helps linear-programming-based refutations of random constraint satisfaction problems.

Let us formulate the approximate Cauchy-Schwarz inequality:

► **Theorem 2.1** (Approximate Cauchy-Schwarz Inequality). *Let $f, g : \{-1, 1\}^n \rightarrow \mathbb{R}$ be polynomials of degree $\leq d$. Let $\|f\|_1, \|g\|_1$ be the $L1$ norms of the coefficients of f, g respectively when written in the standard monomial basis over $\{-1, 1\}^n$. Then, for every R -local pseudo-distribution $\mathcal{D}$ on $\{-1, 1\}^n$, and any positive integers s, t such that $d(s+t) \leq R$, we have:*

$$|\tilde{\mathbb{E}}[fg]| \leq \sqrt{(1 - \frac{1}{s})\tilde{\mathbb{E}}[f^2] + \frac{1}{s}\|f\|_1^2} \sqrt{(1 - \frac{1}{t})\tilde{\mathbb{E}}[g^2] + \frac{1}{t}\|g\|_1^2}.$$

Our proof is based on a simple sampling scheme that allows us to write f and g as averages over *local* functions with some additive “sampling” error.

Proof. We present the proof for $d = 1$. The case of larger d is the same but, unlike below, where we “sample” monomials of degree 1, we sample monomials of degree $\leq d$ instead.

Let $f(x) = a_0 + \sum_{i \leq n} a_i x_i$ and $g(x) = b_0 + \sum_i b_i x_i$. We will assume that $0 < \|f\|_1 = \sum_i |a_i|$ and $0 < \|g\|_1 = \sum_i |b_i|$ (if either of these quantities is 0 then there is nothing to prove).

Let us sample an unbiased estimator for f . Define a distribution p on $\{0, 1, 2, \dots, n\}$ by setting $p_i = \frac{|a_i|}{\|f\|_1}$. For each $i \in \{0, 1, 2, \dots, n\}$, let $\chi_i(x) = x_i$ if $i > 0$ and $\chi_0(x) = 1$. Draw $i_1, i_2, \dots, i_s$ independently from p and define:

$$F = \frac{1}{s} \sum_{\ell=1}^s \frac{a_{i_\ell}}{p_{i_\ell}} \chi_{i_\ell}(x).$$

Then, clearly, F depends on at most s variables. Further, $\mathbb{E}[F] = f$, where the expectation is with respect to p . Similarly, construct G from g using t samples.

Now, let $\mathcal{D}$ be any R -local pseudo-distribution on $x_1, x_2, \dots, x_n \in \{\pm 1\}$. Since F and G are s, t -local respectively and $s+t \leq R$, we must have:

$$|\tilde{\mathbb{E}}_{\mathcal{D}}[FG]| \leq \tilde{\mathbb{E}}_{\mathcal{D}}[F^2]^{1/2} \tilde{\mathbb{E}}_{\mathcal{D}}[G^2]^{1/2}. \quad (1)$$

Now, since F and G are sampled independently, we also have:

$$\mathbb{E}_{\text{sampling}} \tilde{\mathbb{E}}[FG] = \tilde{\mathbb{E}}[fg].$$

Thus, using (1), taking expectations with respect to the sampling process and applying Jensen’s inequality for the expectation of the square root function gives

$$|\tilde{\mathbb{E}}[fg]| \leq \mathbb{E}_{\text{sampling}} \tilde{\mathbb{E}}_{\mathcal{D}}[F^2]^{1/2} \tilde{\mathbb{E}}_{\mathcal{D}}[G^2]^{1/2} \leq \sqrt{(\mathbb{E}_{\text{sampling}} \tilde{\mathbb{E}}_{\mathcal{D}}[F^2])} \sqrt{\mathbb{E}_{\text{sampling}} \tilde{\mathbb{E}}_{\mathcal{D}}[G^2]}. \quad (2)$$

We will now finish the proof by computing the RHS. Let $X_\ell = \frac{a_{i_\ell}}{p_{i_\ell}} \chi_{i_\ell}(x)$, so $F = \frac{1}{s} \sum_{\ell=1}^s X_\ell$.

For two distinct samples $\ell \neq m$, we have:

$$\mathbb{E}_{i_\ell, i_m} [\tilde{\mathbb{E}}[X_\ell X_m]] = \tilde{\mathbb{E}}[f^2].$$

On the other hand, when $\ell = m$, we have:

$$\mathbb{E}_{i_\ell} [\tilde{\mathbb{E}}[X_\ell^2]] = \sum_i \frac{a_i^2}{p_i} = (\sum_i |a_i|)^2 = \|f\|_1^2.$$

Thus,

$$\mathbb{E}_{\text{sampling}}[\tilde{\mathbb{E}}[F^2]] = \frac{s(s-1)}{s^2} \tilde{\mathbb{E}}[f^2] + \frac{s}{s^2} \|f\|_1^2 = (1 - 1/s) \tilde{\mathbb{E}}[f^2] + \frac{\|f\|_1^2}{s}.$$

A similar argument for G gives:

$$\mathbb{E}_{\text{sampling}}[\tilde{\mathbb{E}}[G^2]] = (1 - 1/t) \tilde{\mathbb{E}}[g^2] + \frac{\|g\|_1^2}{t}.$$

Plugging back in (2) finishes the proof. $\blacktriangleleft$

2.1 A Matching Lower Bound

In this section, we prove that the additive error appearing in our approximate Cauchy-Schwarz inequality is necessary and, in fact, that our bound is quantitatively optimal.

► **Theorem 2.2.** *Let R, n be such that $n/R \rightarrow \infty$. Then, there exists an R -local Sherali-Adams pseudo-distribution on $\{-1, 1\}^n$ such that for the linear polynomial $p(x) = \sum_i x_i$, we have: $\tilde{\mathbb{E}}_D[p(x)^2] = 0$, and,*

$$\tilde{\mathbb{E}}_D[p] \geq (1 - O(1/R + R/n)) \frac{n}{\sqrt{R-1}}.$$

► **Remark 2.3.** Applying Theorem 2.1 with $f = p$, $g = 1$, $s = R - 1$ and $t = 1$ gives $\tilde{\mathbb{E}}_D[p] \leq n/\sqrt{R-1}$. Thus, the additive error in Theorem 2.1 is asymptotically tight with the correct leading constant. Notice also that this example rules out an approximate Cauchy-Schwarz inequality with any finite *multiplicative* error.

Proof. We will define the pseudo-distribution D by describing a local distribution on every collection T of R variables. The local distributions will be permutation-invariant on T and thus, the marginal distribution on any $T' \subseteq T$ will depend only on $|T'|$. Thus, in particular, local consistency holds: for any $T' \subseteq T_1 \cap T_2$ for two different sets of size R each, the marginal of the local distributions on T_1 and T_2 on T' agree. This implies immediately the existence of an R -local Sherali-Adams pseudo-distribution with pseudo-moments agreeing with the local distributions we will construct.

For every $T \subseteq [n]$ of size exactly R , let $Y_T = \sum_{i \in T} x_i$. We will sample Y_T from a distribution on $[-R, R]$ and then set the constituent x_i s to be uniform conditioned on satisfying $\sum_i x_i$ being equal to the sampled value of Y_T . This will ensure permutation-invariance that we desire immediately.

What remains to be described is the specific distribution to sample the random variable Y_T on $[-R, R]$. We will choose this carefully to have prescribed first and second moments (which will govern $\tilde{\mathbb{E}}_D[p]$ and $\tilde{\mathbb{E}}_D[p^2]$ respectively).

Let $A = \frac{R(n-R)}{n-1}$. This will be our prescribed 2nd moment for Y_T . Notice that $A = R(1 - O(R/n))$ and $\sqrt{A} = \sqrt{R}(1 - O(R/n))$. Choose a to be the largest positive integer in $\{-R, -R+2, \dots, R-2, R\}$ such that $a \leq \sqrt{A}$. Set $b = a + 2$. Choose θ so that $(1 - \theta)a^2 + \theta b^2 = A$, that is, set $\theta = \frac{A-a^2}{b^2-a^2}$. Observe that since $a^2 \leq A \leq b^2$, $\theta \in [0, 1]$.

Now, consider the following distribution for Y_T . With probability $1 - \theta$, output a , with probability θ , output b . Choose $x_T \in \{-1, 1\}^T$ to be uniform over all assignments satisfying $\sum_i x_i = Y_T$.

Let's compute $\tilde{\mathbb{E}}_D[x_i x_j]$ for $i \neq j$. Observe that $Y_T^2 = R + 2 \sum_{i < j, i, j \in T} x_i x_j$. Since $\mathbb{E}[Y_T^2] = A$, we must have $\tilde{\mathbb{E}}_D[x_i x_j] = \frac{(A-R)}{R(R-1)} = -\frac{1}{n-1}$. Thus, $\tilde{\mathbb{E}}_D[p^2] = 0$.

Next, a simple computation gives $\tilde{\mathbb{E}}_D[x_i] = 1/R \tilde{\mathbb{E}}_D[Y_T] = 1/R[(1-\theta)a + \theta b] = \frac{1}{\sqrt{R}}(1 - O(1/R + R/n))$.

Thus, $\tilde{\mathbb{E}}_D[p] = (1 - O(1/R + R/n)) \frac{n}{\sqrt{R}} = (1 - O(1/R + R/n)) \frac{n}{\sqrt{R-1}}$. $\blacktriangleleft$

3 Refuting Random k -XOR

In this section, we will show that Sherali-Adams can refute random k -XOR instances, which we define below, as an application of approximate Cauchy-Schwarz.

► **Definition 3.1.** Let $k \geq 2$ be an integer. We say that ψ is a random k -XOR instance with pn^k expected clauses if for all $U \in [n]^k$, we add U to $\mathcal{H}$ with probability p and sample an independent Rademacher (± 1) sign b_U . Then we write $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$, where $x^U := \prod_{i \in U} x_i$.

► **Remark 3.2.** Notice that if an $\frac{1}{2} + \gamma$ fraction of constraints are satisfied, i.e. we have $b_U \cdot x^U = 1$ for $\frac{1}{2} + \gamma$ fraction of $U \in \mathcal{H}$, then $\psi(x) = 2\gamma|\mathcal{H}|$. Therefore, certifying that we cannot satisfy a large fraction of constraints is equivalent to certifying an upper bound on $\psi(x)$.

We also define the following notation that we will use:

► **Definition 3.3.** Let $\psi = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with clause set $\mathcal{H}$. For all $U \notin \mathcal{H}$, we let $b_U = 0$. Define $\mathcal{H}_i := \{C \in \mathcal{H} : C_k = i\}$ be the subset of clauses in $\mathcal{H}$ whose last coordinate is i .

We will prove the following theorem:

► **Theorem 3.4.** Let k be an odd integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a random k -XOR instance with $n^{k/2+\delta}$ expected constraints. Suppose we have $2^{O(k)} \log^3 n \leq n^{\delta/2}$. Then for all $\Omega(1) = \gamma \leq 1/2$, there exists some $R = O_{\gamma, \delta, k}(1)$ such that every R -local pseudodistribution $\mathcal{D}$ satisfies $\tilde{\mathbb{E}}_{\mathcal{D}}\psi(x) \leq 2\gamma|\mathcal{H}|$.

► **Remark 3.5.** Using a standard Fourier analytic reduction, we can generalize this result to general k -CSPs.

The state-of-the-art refutation algorithms for random (and semirandom) k -XOR instances are based on spectral certificates. Informally, the idea is to show that the “value” polynomial of any k -XOR instance (the polynomial that computes the advantage over $1/2$ of an assignment) can be upper bounded at every x by a quadratic form of a certain matrix built from the instance. One then establishes that when the input instance is random (or semirandom) with appropriately many constraints, then, the spectral norm of this matrix is small. We will call a matrix satisfying sufficient conditions for obtaining a Sherali-Adams refutations a spectral certificate.

► **Definition 3.6 (Spectral Certificates).** Let k be an odd integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance. Define $A(\psi)$ to be a $n^{k-1} \times n^{k-1}$ matrix defined as follows: For all $(S_1, T_1) \neq (S_2, T_2) \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, we have

$$A(\psi)_{(S_1, T_1), (S_2, T_2)} := \sum_{i \in [n]} b_{S_1 T_1 i} b_{S_2 T_2 i} + b_{T_1 S_1 i} b_{T_2 S_2 i},$$

77:8 Approximate Cauchy-Schwarz for Sherali-Adams

and $A(\psi)_{(S_1, S_1), (S_2, S_2)} = 0$. We call $A(\psi)$ a δ -spectral certificate for $\mathcal{H}$ if

$$\|A(\psi)\| \leq n^{-\delta} \left(\min_{(S_1, S_2)} \|A(\psi)_{(S_1, S_2)}\|_1 \right).$$

We let $D(\psi)_{S_1, S_2} := \|A_{S_1, S_2}\|_1$ and write $D(\psi) := \text{diag}(\{D(\psi)_{S_1, S_2}\})$. Observe that

$$x^{\odot(k-1)} A(\psi) x^{\odot(k-1)} = 2 \sum_i \sum_{U \neq U' \in \mathcal{H}_i} b_U b_{U'} x^U x^{U'}.$$

With high probability over the choice of ψ , the matrix $A(\psi)$ is a spectral certificate. The proof of this is simple and follows from the work of [2], and is deferred to the appendix.

► **Lemma 3.7.** *Suppose ψ is a random k -XOR instance with $n^{k/2+\delta}$ constraints such that $2^{O(k)} \log^3 n \leq n^{\delta/2}$. Then $A(\psi)$ is a $\frac{\delta}{2}$ -spectral certificate for ψ with high probability.*

One can view the main idea of O'Donnell and Schramm's work as showing (under appropriate conditions) that spectral bounds on certifying matrices such as above can be proven using the Sherali-Adams linear program. We will abstract their analysis into the following lemma:

► **Lemma 3.8** (Abstracting out [13]). *Let A be a symmetric matrix with integer entries with rows and columns indexed by t -tuples of $[n]$, and let D be its degree matrix. Let $\pi_* = \frac{1}{\sum_{S,T} |A_{S,T}|} \min_S \|A_S\|_1$, $\rho(\cdot)$ denote the spectral radius of a matrix, and $\eta > \min(\pi_*^{-1/2}, \rho(D^{-1}A))$. Take $\ell := \left\lceil \frac{1}{4} \frac{\log(\eta^2 \pi_*)}{\log(\rho(D^{-1}A)/\eta)} \right\rceil$ and $p := \lceil (\frac{2}{5\eta})^{2\ell} \rceil$. Then for $R(A) := tp\ell + t$, any R -local pseudodistribution $\mathcal{D}$ on $x_1, \dots, x_n$ satisfies*

$$\left| \tilde{\mathbb{E}}_{\mathcal{D}}(x^{\odot t})^\top A x^{\odot t} \right| \leq \eta \sum_{S,T} |A_{S,T}|,$$

where $x^{\odot t}$ is the n^t -dimensional vector whose entries are indexed by t -tuples of $[n]$ and are equal to the product of the corresponding entries.

► **Lemma 3.9.** *Let k be an odd integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance such that $|\mathcal{H}| \geq \frac{36n}{\gamma^2}$ and $|\mathcal{H}_i| \leq \frac{|\mathcal{H}|}{n}(1 + \gamma)$ for all i . Suppose that $A(\psi)$ is a δ -spectral certificate for ψ . Then for all $\Omega(1) = \gamma \leq \frac{1}{2}$, $R = O_{\gamma, \delta, k}(1)$ -rounds of Sherali-Adams can certify that $\psi(x) \leq 2\gamma|\mathcal{H}|$; in other words, for all R -local pseudodistributions $\mathcal{D}$, $|\tilde{\mathbb{E}}_{\mathcal{D}}\psi(x)| \leq 2\gamma|\mathcal{H}|$.*

Proof. Let $\psi_i = \sum_{U \in \mathcal{H}_i} b_U x^U$, $R := \max\{R(A(\psi)), \frac{10k(1+\gamma)^2}{\gamma^2}\} = O_{\gamma, \delta, k}(1)$, where $R(A(\psi))$ is the quantity from Lemma 3.8 when $\eta = \gamma^2$, and $\tilde{\mathbb{E}}$ be any R -local pseudodistribution. A δ -spectral certificate $A(\psi)$ satisfies $\rho(D^{-1}(\psi)A(\psi)) \leq n^{-\delta}$ (recall that $\rho(\cdot)$ is the spectral radius of a matrix) and therefore by Lemma 3.8,

$$\begin{aligned} \tilde{\mathbb{E}} \sum_i \psi_i^2(x) &= \tilde{\mathbb{E}} \sum_{U \in \mathcal{H}} b_U^2 + \frac{1}{2} \tilde{\mathbb{E}} x^{\odot(k-1)} A(\psi) x^{\odot(k-1)} \\ &\leq |\mathcal{H}| + \frac{\gamma^2}{2} \sum_i |\mathcal{H}_i| (|\mathcal{H}_i| - 1) \leq |\mathcal{H}| + \frac{\gamma^2}{2} \frac{|\mathcal{H}|^2}{n} (1 + \gamma)^2, \end{aligned}$$

so by Theorem 2.1, we have for $s \geq \frac{9(1+\gamma)^2}{\gamma^2}$,

$$\begin{aligned} \left| \widetilde{\mathbb{E}}\psi(x) \right| &= \left| \widetilde{\mathbb{E}} \sum_i \psi_i(x) \right| \leq \sum_i \sqrt{\left(1 - \frac{1}{s}\right) \widetilde{\mathbb{E}}\psi_i^2 + \frac{1}{s} \|\psi_i\|_1^2} \\ &\leq \sqrt{n \left(\left(1 - \frac{1}{s}\right) \sum_i \widetilde{\mathbb{E}}\psi_i^2 + \frac{1}{s} \sum_i |\mathcal{H}_i|^2 \right)} \\ &\leq \sqrt{n|\mathcal{H}|} + \gamma(1+\gamma)|\mathcal{H}| + \frac{|\mathcal{H}|}{\sqrt{s}}(1+\gamma) \leq 2\gamma|\mathcal{H}|. \end{aligned} \quad \blacktriangleleft$$

Proof of Theorem 3.4. For random k -XOR, Chernoff bounds imply that $|\mathcal{H}_i| \leq \frac{|\mathcal{H}|}{n}(1+\gamma)$ for all i with high probability. Then the result follows directly from Lemma 3.7 and Lemma 3.9. $\blacktriangleleft$

4 Refuting Semirandom k -XOR

In this section, we show that Sherali-Adams can refute semirandom instances of k -XOR.

► **Definition 4.1.** Let k be an integer and $\mathcal{H} \subseteq [n]^k$. We say that ψ is a semirandom k -XOR instance with clause set $\mathcal{H}$ if for $U \in \mathcal{H}$, we sample an independent Rademacher (± 1) sign b_U and write $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$.

► **Theorem 4.2.** Let k be an integer. $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with clause set $\mathcal{H}$ such that $|\mathcal{H}| = n^{k/2+\delta}$ total constraints. Suppose we have $\gamma \geq n^{-\delta/4} 4\sqrt{15}$ and $16384k^4 \log^3 n \leq n^{\delta/18}$. Then for all $\Omega(1) = \gamma \leq 1/2$, with high probability, there exists some $R = O_{\gamma,\delta,k}(1)$ such that every R -local pseudodistribution $\mathcal{D}$ satisfies $\widetilde{\mathbb{E}}_{\mathcal{D}}\psi(x) \leq 2\gamma|\mathcal{H}|$.

For semirandom odd k -XOR, the $\mathcal{H}_i$ can be much larger than $\frac{|\mathcal{H}|}{n}$, in which case we need to partition large $\mathcal{H}_i$ into smaller sets to ensure the matrix $A(\psi)$ does not have too many nonzero entries. In addition, we can no longer guarantee that $A(\psi)$ is a spectral certificate with high probability. We thus define a notion of a regular partition and a refined notion of a spectral certificate that allows us to delete a negligible fraction of rows and columns of the matrix.

► **Definition 4.3.** For all $\mathcal{H}_1 \sqcup \dots \sqcup \mathcal{H}_n$, Let $\mathcal{H} \subseteq [n]^k$. We call a partition $P := \bigsqcup_i P_i := \bigsqcup_i \bigsqcup_{j \in [p_i]} \mathcal{H}_i^{(j)}$ of the $\mathcal{H}_i$ s regular if it has a total of at most $2n$ parts and $|\mathcal{H}_i^{(j)}| \leq \frac{|\mathcal{H}|}{n}$ for all i, j .

► **Definition 4.4** $((\gamma, \delta)$ -Spectral Certificates). Let k be an odd integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with clause set $\mathcal{H}$. Suppose we have a partition $P_i := \bigsqcup_{j \in [p_i]} \mathcal{H}_i^{(j)}$ of each $\mathcal{H}_i$, and let $P = (P_1, \dots, P_n)$. Define $A'_P(\psi)$ to be a $n^{k-1} \times n^{k-1}$ matrix defined as follows: For all $(S_1, T_1) \neq (S_2, T_2) \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, we have

$$\begin{aligned} A'_P(\psi)_{(S_1, S_2), (T_1, T_2)} &:= \sum_{i \in [n]} \sum_{j \in [p_i]} b_{S_1 T_1 i} b_{S_2 T_2 i} \cdot \mathbf{1}\{S_1 T_1 i, S_2 T_2 i \in \mathcal{H}_i^{(j)}\} \\ &\quad + b_{T_1 S_1 i} b_{T_2 S_2 i} \cdot \mathbf{1}\{T_1 S_1 i, T_2 S_2 i \in \mathcal{H}_i^{(j)}\}, \end{aligned}$$

and $A'_P(\psi)_{(S_1, S_1), (S_2, S_2)} = 0$. Let $A(\psi)$ be a principal submatrix of $A'_P(\psi)$, and define $D(\psi)_{S_1, S_2} := \|A_{S_1, S_2}\|_1$ and write $D(\psi) := \text{diag}(\{D(\psi)_{S_1, S_2}\})$. We call $A(\psi)$ a (γ, δ) -spectral certificate of ψ with respect to P if the following hold:

77:10 Approximate Cauchy-Schwarz for Sherali-Adams

- $A(\psi)$ can be formed by removing a collection of rows of $A'_P(\psi)$ and their corresponding columns whose ℓ_1 norms sum to at most $\frac{\gamma^2}{2} \sum_{i \in [n]} \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| (|\mathcal{H}_i^{(j)}| - 1)$,
- $\rho(D^{-1}(\psi)A(\psi)) \leq n^{-\delta/2}$, where $\rho(\cdot)$ denotes the spectral radius of a matrix.

Observe that (abusing notation and treating $A(\psi)$ as a $n^{k-1} \times n^{k-1}$ matrix where the removed rows are replaced by zeros), we have

$$x^{\odot(k-1)} A'_P(\psi) x^{\odot(k-1)} = 2 \sum_i \sum_{j \in [p_i]} \sum_{U \neq U' \in \mathcal{H}_i^{(j)}} b_U b_{U'} x^U x^{U'},$$

and

$$x^{\odot(k-1)} (A'_P(\psi) - A(\psi)) x^{\odot(k-1)} \leq \frac{\gamma^2}{2} \sum_{S_1, S_2} |A'_P(\psi)_{S_1, S_2}|.$$

► **Definition 4.5.** Let k be an integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with clause set $\mathcal{H}$. Then we define $\psi^1(x) := \sum_{U \in \mathcal{H}} x^U$ to be the k -XOR instance with the same constraints but all ones on the RHS.

► **Remark 4.6.** Note that any entry $A'_P(\psi)_{(S_1, S_2), (T_1, T_2)}$ is the sum of $A'_P(\psi^1)_{(S_1, S_2), (T_1, T_2)}$ independent Rademacher random variables. In particular, if $A'_P(\psi^1)_{(S_1, S_2), (T_1, T_2)} \neq 0$, then $A'_P(\psi)_{(S_1, S_2), (T_1, T_2)} = 0$ with probability at most $\frac{1}{2}$.

We will show that the existence of a (γ, δ) -spectral certificate of ψ implies that Sherali-Adams admits a refutation of ψ .

► **Lemma 4.7.** Let k be an odd integer and $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance. Suppose that there exists a (γ, δ) -spectral certificate for ψ with respect to a regular partition P of all $\mathcal{H}_i$. Then if $\gamma \leq \frac{1}{2}$ and $|\mathcal{H}| \geq \frac{4n}{\gamma^2}$, then $R = O_{\gamma, \delta, k}(1)$ -rounds of Sherali-Adams can certify that $\psi(x) \leq 2\gamma|\mathcal{H}|$; in other words, for all R -local pseudodistributions $\mathcal{D}$, $|\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x)| \leq 2\gamma|\mathcal{H}|$.

Proof of Lemma 4.7. Let $\psi_{ij} = \sum_{U \in \mathcal{H}_i^{(j)}} b_U x^U$ and $A(\psi)$ be a (γ, δ) -spectral certificate of ψ . Let $R := \max\{R(A(\psi)), \frac{5k}{\gamma^2}\} = O_{\gamma, \delta, k}(1)$, where $R(A(\psi))$ is the quantity from Lemma 3.8 when $\eta = \frac{\gamma^2}{2}$, and $\tilde{\mathbb{E}}$ be an R -local pseudodistribution. Then $\rho(D^{-1}(\psi)A(\psi)) \leq n^{-\delta/2}$ and thus by Lemma 3.8,

$$\begin{aligned} \tilde{\mathbb{E}} \sum_i \sum_{j \in [p_i]} \psi_{ij}^2(x) &= \tilde{\mathbb{E}} \sum_{U \in \mathcal{H}} b_U^2 + \frac{1}{2} \tilde{\mathbb{E}} x^{\odot(k-1)} A'(\psi) x^{\odot(k-1)} \\ &\leq |\mathcal{H}| + \frac{1}{2} \tilde{\mathbb{E}} x^{\odot(k-1)} A(\psi) x^{\odot(k-1)} + \frac{1}{2} \tilde{\mathbb{E}} x^{\odot(k-1)} (A'(\psi) - A(\psi)) x^{\odot(k-1)} \\ &\leq |\mathcal{H}| + \left(\frac{\gamma^2}{2} + \frac{\gamma^2}{4} \right) \sum_i \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| (|\mathcal{H}_i^{(j)}| - 1) \\ &\leq |\mathcal{H}| + \frac{3\gamma^2}{4} \frac{|\mathcal{H}|^2}{n}, \end{aligned}$$

where the last line uses the fact that our partition is regular and therefore each $\mathcal{H}_i^{(j)}$ has size at most $\frac{|\mathcal{H}|}{n}$. Thus by Theorem 2.1, since P has at most $2n$ parts, we have, for $s \geq \frac{4}{\gamma^2}$,

$$\begin{aligned}
|\tilde{\mathbb{E}}\psi| &= \left| \tilde{\mathbb{E}} \sum_i \sum_{j \in [p_i]} \psi_{ij}(x) \right| \leq \sum_i \sum_{j \in [p_i]} \sqrt{\left(1 - \frac{1}{s}\right) \tilde{\mathbb{E}}\psi_{ij}^2 + \frac{1}{s} \|\psi_{ij}\|_1^2} \\
&\leq \sqrt{2n \left(\left(1 - \frac{1}{s}\right) \sum_i \sum_{j \in [p_i]} \tilde{\mathbb{E}}\psi_{ij}^2 + \frac{|\mathcal{H}_i|^2}{sn} \right)} \\
&\leq \sqrt{2n|\mathcal{H}| + 3\gamma^2|\mathcal{H}|^2 + \frac{2|\mathcal{H}|^2}{s}} \leq 2\gamma|\mathcal{H}|. \quad \blacktriangleleft
\end{aligned}$$

Finally, we note that given sufficient regularity conditions on $\mathcal{H}$, there exists a spectral certificate of ψ with high probability.

► **Lemma 4.8.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ such that $|\mathcal{H}| = n^{k/2+\delta}$ and let P be a regular partition of the $\mathcal{H}_i$ s. Suppose $\mathcal{H}$ has the following properties:*

- *For all $S \in [n]^{(k-1)/2}$, $i \in [n]$, there are at most $n^{\delta/2}$ total clauses of the form STi or TSi in $\mathcal{H}$.*
- *For all $S_1T_1 \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, there are at most $n^{\delta/2}$ total clauses of the form $S_1T_1i \in \mathcal{H}$.*

Suppose that $\gamma \geq n^{-\delta/4}\sqrt{60}$ and $16384k^4 \log^3 n \leq n^{\delta/18}$. Then with high probability, there exists a $(\gamma, \delta/9)$ -spectral certificate of ψ with respect to P .

We defer the proof of this lemma to Section 4.3.

4.1 Refuting induced 2-XOR instances

Unlike the random case, a semirandom XOR instance ψ does not necessarily have a spectral certificate: when large groups of variables occur together in a large number of clauses, the spectral norm of $A'(\psi)$ can be large and making it sufficiently small requires pruning a large (possibly almost all) of the rows. Instead, we can partition our set of clauses $\mathcal{H}$ into two subparts $\mathcal{H}'$ and $\mathcal{H}''$, where for $\mathcal{H}'$ is the “approximately regular” part where we can construct the aforementioned spectral certificate, and for the second part $\mathcal{H}''$, we instead construct a 2-XOR instance that we can refute using Sherali-Adams. At a high level, this resembles the approach of [1], however, their method for refuting the 2-XOR instance for $\mathcal{H}''$ relies on bounding the value of the instance by an infinity-to-1 norm and using the fact that SOS can recognize Grothendieck’s inequality. This is not known for Sherali-Adams, so we must handle this case differently.

We define the notion of an induced 2-XOR instance below. This is a 2-XOR instance formed by splitting each clause U of a k -XOR instance into two parts $S \sqcup T$, defining a variable y_S and y_T for each of the two parts, and replacing the k -XOR constraint $b_U x^U$ with the 2-XOR constraint $b_U y_S y_T$. Formally, we have a bijective map between our clause $U \in \mathcal{H}$ and the partition (S, T) , and using the transformation $y_S = \prod_{i \in S} x_i$, for all x , there exists y such that $\phi(y) = \psi(x)$.

► **Definition 4.9.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$. We call a 2-XOR instance ϕ an f -induced 2-XOR instance of ψ if each of its variables y_S is indexed by some $S \subseteq [n]^{\leq k}$ and there exists a set $\mathcal{B} = \mathcal{B}(\phi)$ of pairs of variables (S, T) and a bijection $f : \mathcal{B} \rightarrow \mathcal{H}$ such that ST is a permutation of $f(S, T)$ and ϕ can be written as*

$$\phi = \sum_{(S, T) \in \mathcal{B}} b_{f(S, T)} y_S y_T,$$

77:12 Approximate Cauchy-Schwarz for Sherali-Adams

Let $d^\phi(S) = |\{U \in \mathcal{H} : f^{-1}(U) \ni S\}|$ denote the number of times the variable y_S appears in the above expression for ϕ , which we call the degree of S in ϕ . Then we define the weighted adjacency matrix $A'(\phi)$ of ϕ with rows and columns indexed by variables of ϕ :

$$(A'(\phi))_{S,T} = b_{f(S,T)} \mathbf{1}\{(S,T) \in \mathcal{B}\} + b_{f(T,S)} \mathbf{1}\{(T,S) \in \mathcal{B}\}.$$

Let $A(\phi)$ be a principal submatrix of $A'(\phi)$ and let $D(\phi)$ be the degree matrix of $A(\phi)$. We call $A(\phi)$ an induced (γ, δ) -spectral certificate for ψ if the following hold:

- $A(\phi)$ can be formed by deleting rows of $A'(\phi)$ and their corresponding columns whose ℓ_1 norms sum to at most $\gamma \sum_{S,T} |A'(\phi)_{S,T}|$,
- $\rho(D^{-1}(\phi)A(\phi)) \leq n^{-\delta}$.

A Sherali-Adams refutation of an induced 2-XOR instance of ψ implies a refutation of ψ itself, which we prove in the appendix.

► **Lemma 4.10.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ and ϕ be an f -induced 2-XOR instance of ψ . Suppose that for all R -local pseudodistributions $\mathcal{D}'$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}'} \phi(y)| \leq 2\gamma|\mathcal{H}|$. Then for all kR -local pseudodistributions $\mathcal{D}$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x)| \leq 2\gamma|\mathcal{H}|$.*

Using this, we prove that Sherali-Adams can refute XOR instances with an induced spectral certificate.

► **Lemma 4.11.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance and suppose there exists an induced (γ, δ) -spectral certificate for ψ . Then $R = O_{\gamma, \delta, k}(1)$ -rounds of Sherali-Adams can certify that $\psi(x) \leq 2\gamma|\mathcal{H}|$; in other words, for all R -local pseudodistributions $\mathcal{D}$, $|\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x)| \leq 2\gamma|\mathcal{H}|$.*

Proof. Let ϕ be an f -induced 2-XOR instance of ψ and $A(\phi)$ be a corresponding induced (γ, δ) -spectral certificate. Let $R := k \cdot R(A(\phi))$, where $R(A(\phi))$ is the quantity from Lemma 3.8 where $\eta = \gamma$, and let $\mathcal{D}$ be any $R = O_{\gamma, \delta, k}(1)$ -local pseudodistribution. Since $\rho(D^{-1}(\phi)A(\phi)) \leq n^{-\delta}$, by Lemma 3.8, for any $R(A(\phi))$ -local pseudodistribution $\mathcal{D}'$, we have

$$\begin{aligned} |\tilde{\mathbb{E}}_{\mathcal{D}'} \phi(y)| &= \frac{1}{2} \left| \tilde{\mathbb{E}} y^\top A'(\phi) y \right| \\ &= \frac{1}{2} \left| \tilde{\mathbb{E}} y^\top A(\phi) y \right| + \frac{1}{2} \left| \tilde{\mathbb{E}} y^\top (A'(\phi) - A(\phi)) y \right| \\ &\leq \gamma|\mathcal{H}| + \gamma|\mathcal{H}| = 2\gamma|\mathcal{H}|, \end{aligned}$$

so $\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x) \leq 2\gamma|\mathcal{H}|$ by Lemma 4.10. ◀

Furthermore, induced 2-XOR instances without too many variables admit an induced spectral certificate for its parent instance, which we show in Section 4.3.

► **Lemma 4.12.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ and ϕ be an f -induced 2-XOR instance of ψ with variable set $\mathcal{S}$ such that $|\mathcal{S}| \leq n^{-\delta}|\mathcal{H}|$. If $\gamma \geq n^{-\delta/2}$ and $4\sqrt{k \log n} \leq n^{\delta/20}$, then with high probability, there exists an induced $(\gamma, \frac{\delta}{5})$ -spectral certificate for ψ with respect to $(\mathcal{B}, f)$.*

4.2 Proof of Theorem 4.2

Here we prove the main result of this section.

Proof of Theorem 4.2. First, suppose that k is even. For all $U = ST \in \mathcal{H}$, where $S, T \in [n]^{k/2}$ define $f(S, T) := U$. Construct the f -induced 2-XOR instance ϕ of ψ . Then we have at most $n^{k/2} = n^{-\delta}|\mathcal{H}|$ total variables, so by Lemma 4.12, there exists an induced $(\gamma, \frac{\delta}{5})$ -spectral certificate for ψ with respect to f . Thus by Lemma 4.11, there exists some $R = O_{\gamma, \delta, k}(1)$ such that for all R -local pseudodistributions $\mathcal{D}$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}}\psi(x)| \leq 2\gamma|\mathcal{H}|$.

Now suppose that k is odd. We construct a partition $\mathcal{H} = \mathcal{H}' \sqcup \mathcal{H}'' \sqcup \mathcal{H}'''$ as follows:

1. While there exists some $S \in [n]^{(k-1)/2}$ and $i \in [n]$ such that there are more than $n^{\delta/2}$ total clauses of the form STi or TSi in $\mathcal{H}$, move all STi and TSi to a new hypergraph $\mathcal{H}'$.
2. Afterwards, while there exist some $(S, T) \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$ such that there are more than $n^{\delta/2}$ total clauses of the form $STi \in \mathcal{H}$, move all such STi from $\mathcal{H}$ to a new hypergraph $\mathcal{H}''$.
3. Let $\mathcal{H}'''$ consist of the remaining clauses.

Now, consider $\mathcal{H}'$. If $|\mathcal{H}'| \leq \frac{\gamma}{2}|\mathcal{H}|$, then $\psi(\mathcal{H}') \leq \frac{\gamma}{2}|\mathcal{H}|$ trivially. Otherwise, each edge $STi \in \mathcal{H}'$ consists of one part $(S, i) \in [n]^{(k-1)/2} \times [n]$ and one part $T \in [n]^{(k-1)/2}$. By construction, we have at most $n^{k/2+\delta/2}$ different (S, i) that occur in $\mathcal{H}'$, since each must occur in at least $n^{\delta/2}$ tuples for it to be added to $\mathcal{H}'$, and there are at most $n^{(k-1)/2}$ total T . Thus, when we use these clauses to construct an induced 2-XOR instance (letting $f((S, i), T) = STi$ if (S, i) was the heavy part and $f(S, (T, i)) = STi$ if (T, i) was the heavy part), we will have $\leq n^{k/2+\delta/2} + n^{(k-1)/2} \leq n^{-\delta/3}|\mathcal{H}'|$ total variables. Thus by Lemma 4.12, there exists an induced $(\frac{\gamma}{2}, \frac{\delta}{15})$ -spectral certificate for $\psi(\mathcal{H}')$, so by Lemma 4.11 there exists some $R' = O_{\gamma/2, \delta/15, k}(1) = O_{\gamma, \delta, k}(1)$ such that for all R' -local pseudodistributions $\mathcal{D}'$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}'}\psi(\mathcal{H}')| \leq \gamma|\mathcal{H}'|$.

Similarly for $\mathcal{H}''$ either $|\mathcal{H}''| \leq \frac{\gamma}{2}|\mathcal{H}|$ and thus $\psi(\mathcal{H}'') \leq \frac{\gamma}{2}|\mathcal{H}|$, or we can construct an induced 2-XOR instance whose variables are $ST \in [n]^{k-1}$ which occur together in at least $n^{\delta/2}$ clauses and $i \in [n]$. Therefore we will have $\leq n^{k/2+\delta/2} + n \leq n^{-\delta/3}|\mathcal{H}''|$ total variables, so like before there exists an induced $(\frac{\gamma}{2}, \frac{\delta}{15})$ -spectral certificate for $\psi(\mathcal{H}'')$ so for all R' -local pseudodistributions $\mathcal{D}'$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}'}\psi(\mathcal{H}'')| \leq \gamma|\mathcal{H}''|$.

Finally, for $\mathcal{H}'''$, if $|\mathcal{H}'''| \leq \frac{\gamma}{2}|\mathcal{H}|$, then $\psi(\mathcal{H}''') \leq \frac{\gamma}{2}|\mathcal{H}|$. Otherwise, construct a regular partition P of $\mathcal{H}'''$ by taking each $\mathcal{H}_i'''$ of size larger than $\frac{|\mathcal{H}'''}{n}$ and repeatedly peeling off $\frac{|\mathcal{H}'''}{n}$ at a time. Then $\mathcal{H}'''$ satisfies the conditions of Lemma 4.8 so there exists a $(\frac{\gamma}{2}, \frac{\delta}{9})$ -spectral certificate for $\psi(\mathcal{H}''')$. Thus by Lemma 4.7, there exists some $R''' = O_{\gamma/2, \delta/9, k}(1) = O_{\gamma, \delta, k}(1)$ such that for all R''' -local pseudodistributions $\mathcal{D}'$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}'}\psi(\mathcal{H}''')| \leq \gamma|\mathcal{H}'''|$.

Combining the above and letting $R := \max(R', R''')$, for all R -local pseudodistributions $\mathcal{D}$, we have

$$\begin{aligned} |\tilde{\mathbb{E}}_{\mathcal{D}}\psi(x)| &\leq |\tilde{\mathbb{E}}_{\mathcal{D}}\psi(\mathcal{H}')(x)| + |\tilde{\mathbb{E}}_{\mathcal{D}}\psi(\mathcal{H}'')(x)| + |\tilde{\mathbb{E}}_{\mathcal{D}}\psi(\mathcal{H}''')(x)| \\ &\leq \frac{\gamma}{2}|\mathcal{H}| + \frac{\gamma}{2}|\mathcal{H}| + \gamma(|\mathcal{H}'| + |\mathcal{H}''| + |\mathcal{H}'''|) \\ &\leq 2\gamma|\mathcal{H}|. \end{aligned}$$

4.3 Existence of spectral certificates

Here we prove Lemma 4.8 and Lemma 4.12. We use the following result, which we defer to the appendix, in the proof of Lemma 4.8.

► **Lemma 4.13.** Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ such that $|\mathcal{H}| = n^{k/2+\delta}$ and let P be a regular partition of the $\mathcal{H}_i$ s. Suppose $\mathcal{H}$ has the following properties:

- For all $S \in [n]^{(k-1)/2}$, $i \in [n]$, there are at most $n^{\delta/2}$ total clauses of the form STi or TSi in $\mathcal{H}$.
- For all $S_1T_1 \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, there are at most $n^{\delta/2}$ total clauses of the form $S_1T_1i \in \mathcal{H}$.

Then with high probability, there exists a principal submatrix $A(\psi)$ of $A'(\psi)$ such that all of the following hold:

- The deleted rows and columns of $A'(\psi)$ have ℓ_1 norms summing to at most

$$30n^{-\delta/2} \sum_{i \in [n]} \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| (|\mathcal{H}_i^{(j)}| - 1).$$

- Each row of $A(\psi)$ has at least n^δ nonzero entries.
- Each row of $A(\psi)$ has at least $\frac{1}{3}$ as many nonzero entries as the corresponding row of $A(\psi^1)$, where we recall that $A'(\psi^1)$ is the matrix defined with the same $\mathcal{H}$ as ψ but with all b_U s equal to 1.
- Each entry of $A(\psi)$ has absolute value at most $4kn^{\delta/4}\sqrt{\log n}$.

► **Lemma 4.8.** Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ such that $|\mathcal{H}| = n^{k/2+\delta}$ and let P be a regular partition of the $\mathcal{H}_i$ s. Suppose $\mathcal{H}$ has the following properties:

- For all $S \in [n]^{(k-1)/2}$, $i \in [n]$, there are at most $n^{\delta/2}$ total clauses of the form STi or TSi in $\mathcal{H}$.
- For all $S_1T_1 \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, there are at most $n^{\delta/2}$ total clauses of the form $S_1T_1i \in \mathcal{H}$.

Suppose that $\gamma \geq n^{-\delta/4}\sqrt{60}$ and $16384k^4 \log^3 n \leq n^{\delta/18}$. Then with high probability, there exists a $(\gamma, \delta/9)$ -spectral certificate of ψ with respect to P .

Proof. Let $A(\psi)$ be the principal submatrix of $A'(\psi)$ given by Lemma 4.13. Since $\gamma \geq \sqrt{60}n^{-\delta/4}$, it satisfies the first condition of a (γ, δ) -spectral certificate with high probability. Now we prove the second condition. Let $D(\psi)$ be the degree matrix of $A(\psi)$, $D(\psi^1)$ be the degree matrix of $A(\psi^1)$, and $M(\psi)$ be the degree matrix of the support matrix of $A(\psi^1)$, i.e. a diagonal entry of $M(\psi)$ is the number of nonzero entries in the corresponding row of $A(\psi^1)$.

Note that we have $M(\psi) \leq 3D(\psi)$ entrywise. Define $\Gamma := M(\psi)^{-1/2}A(\psi)M(\psi)^{-1/2}$, so we have

$$\rho(D(\psi)^{-1}A(\psi)) \leq 3\rho(M(\psi)^{-1}A(\psi)) = 3\rho(\Gamma) \leq 3\|\Gamma\|.$$

We start by upper bounding $\mathbb{E}\text{tr}(\Gamma^\ell) = \mathbb{E}\text{tr}((M^{-1}(\psi)A(\psi))^\ell)$ for even ℓ by analyzing closed length- ℓ walks $V_1 \mapsto V_2 \mapsto \dots \mapsto V_\ell \mapsto V_1$ in the weighted multigraph whose vertices are variables of ψ corresponding to the weighted signed adjacency matrix $M^{-1}(\psi)A(\psi)$. Note that for any edge between two vertices $V_j = (S_1, S_2)$ and $V_{j+1} = (T_1, T_2)$, there exists some i such that either both $S_1T_1i, S_2T_2i \in \mathcal{H}$ or both $T_1S_1i, T_2S_2i \in \mathcal{H}$. Since each $U \in \mathcal{H}$ comes with an independent Rademacher sign ξ_U , if some clause $U \in \mathcal{H}$ appears only once as one of the two clauses inducing an edge in our walk, then the expected value of the walk is 0. Therefore, in a walk with a nonzero contribution to the expected trace, each clause which induces an edge must occur at least twice.

Then we can define an encoding of any closed walk $V_1 \mapsto V_2 \mapsto \dots \mapsto V_\ell \mapsto V_1$ by first choosing a variable $V_j = (S_1, S_2)$, of which there are at most n^{k-1} choices, then choosing $b \in \{0, 1\}^\ell$ such that b_j is the indicator that both of the two tuples which induce the edge from V_j to V_{j+1} have not induced a previous step in the walk. Since each edge must occur at least twice, we note that $b_j = 0$ for at least $\frac{\ell}{2}$ values of j .

At each step j , we choose the next step of the walk based on the value of b_j : At $V_j = (S_1, S_2)$, if $b_j = 1$, we simply pick an arbitrary neighbor of V_j , while if $b_j = 0$, there are at most $2j \leq 2\ell$ possible $U \in \mathcal{H}$ to choose from which have induced a previous step in the walk. Therefore, when $b_j = 1$, there are at most $D(\psi)(V_j)$ possible choices, so $\sum_{V_{j+1} \in N_1(V_j)} \frac{1}{M(\psi)(V_j)} \leq \frac{D(\psi)(V_j)}{M(\psi)(V_j)} \leq 4kn^{\delta/4} \sqrt{\log n}$. When $b_j = 0$, without loss of generality, suppose the tuple we pick is of the form $S_1 T_1 i \in \mathcal{H}$ for some $i \in [n]$. Then the second tuple inducing this edge must be of the form $(S_2, T_2, i) \in \mathcal{H}_i^{(j)}$ for some $i \in [n], j \in [p_i]$, and there are at most $n^{\delta/2}$ total choices for that, so $\sum_{V_{j+1} \in N_0(V_j)} \frac{1}{M(\psi)(V_j)} \leq \frac{2\ell n^{\delta/2}}{n^\delta} \leq 2\ell n^{-\delta/2}$, since $M(\psi)(V_j) \geq n^\delta$. Thus we get

$$\begin{aligned} \mathbb{E} \text{tr}(\Gamma^\ell) &\leq \sum_{b \in \{0,1\}^\ell} \sum_{V_1 \in [n]^{k-1}} \frac{1}{M(\psi)(V_1)} \sum_{S_2 \in N_{b_1}(S_1)} \frac{1}{M(\psi)(S_2)} \\ &\quad \dots \sum_{S_\ell \in N_{b_{\ell-1}}(S_{\ell-1})} \frac{\mathbf{1}\{S_1 \in N_{b_\ell}(S_\ell)\}}{M(\psi)(S_\ell)} \\ &\leq 2^\ell n^{k-1} (4kn^{\delta/4} \sqrt{\log n} \cdot 2\ell n^{-\delta/2})^{\ell/2} \\ &= 2^\ell n^{k-1} (2\sqrt{2k\ell} (\log n)^{1/4} n^{-\delta/8})^\ell. \end{aligned}$$

Therefore, we set $\ell = 2 \left\lceil \frac{k \log n}{2} \right\rceil$ so by Markov's inequality, we have

$$\Pr \left[\|\Gamma\| \geq n^{-\delta/9} \geq 8\sqrt{2}kn^{-\delta/8} \log^{3/4} n \right] \leq \frac{\mathbb{E} \text{tr}(\Gamma^\ell)}{(8\sqrt{2}kn^{-\delta/8} \log^{3/4} n)^\ell} = \frac{1}{\text{poly}(n^k)}. \quad \blacktriangleleft$$

► **Lemma 4.12.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ and ϕ be an f -induced 2-XOR instance of ψ with variable set $\mathcal{S}$ such that $|\mathcal{S}| \leq n^{-\delta} |\mathcal{H}|$. If $\gamma \geq n^{-\delta/2}$ and $4\sqrt{k} \log n \leq n^{\delta/20}$, then with high probability, there exists an induced $(\gamma, \frac{\delta}{5})$ -spectral certificate for ψ with respect to $(\mathcal{B}, f)$.*

Proof. While $A'(\phi)$ contains a row with at most $n^{\delta/2}$ entries, delete that row and its corresponding column, and let $A(\phi)$ be the result. Then note that the deleted rows contain at most $|\mathcal{S}| n^{\delta/2} \leq n^{-\delta/2} |\mathcal{H}|$ entries. Now let $D(\phi)$ be the degree matrix of $A(\phi)$, and note that all diagonal entries of $D(\phi)$ are at least $n^{\delta/2}$. Define $\Gamma = D(\phi)^{-1/2} A(\phi) D(\phi)^{-1/2}$, and note that $\rho(D(\phi)^{-1} A(\phi)) = \rho(\Gamma) \leq \|\Gamma\|$. We start by upper bounding $\mathbb{E} \text{tr}(\Gamma^\ell) = \mathbb{E} \text{tr}((D^{-1}(\phi) A(\phi))^\ell)$ by analyzing closed even length- ℓ walks $S_1 \mapsto S_2 \mapsto \dots \mapsto S_\ell \mapsto S_1$ in the weighted graph whose vertices are variables of ϕ which corresponds to the weighted signed adjacency matrix $(D^{-1}(\phi) A(\phi))$. Note that there exists an edge between two vertices S_i, S_{i+1} if either $f(S_i, S_{i+1}) \in \mathcal{H}$ or $f(S_{i+1}, S_i) \in \mathcal{H}$. Since each $U \in \mathcal{H}$ comes with an independent Rademacher sign ξ_U , if some unordered tuple (S, T) appears only once as an edge of the walk, then the expected value of the walk is 0. Therefore, each edge in a walk with a nonzero contribution to the expected trace must occur at least twice. Then we can define an encoding of any closed walk $S_1 \mapsto S_2 \mapsto \dots \mapsto S_\ell \mapsto S_1$ by first choosing a variable S_1 , of which there are $|[n]^{k-1}| < n^k$ choices, then choosing $b \in \{0, 1\}^\ell$ such that b_i is the indicator that the i th edge of the walk has not occurred in a previous step. Since each edge must occur at least twice, we note that $b_i = 0$ for at least $\frac{\ell}{2}$ values of i .

At each step i , we choose the next step of the walk based on the value of b_i : At S_{i-1} , if $b_i = 1$, we simply pick an arbitrary neighbor of S_{i-1} , while if $b_i = 0$, we pick some S_i such that the edge (S_{i-1}, S_i) has occurred previously. Then, when $b_i = 1$, there are at most $D(\phi)(S_{i-1})$ possible choices, so $\sum_{S_i \in N_1(S_{i-1})} 1/D(\phi)(S_{i-1}) \leq 1$. When $b_i = 0$, there are at most $i - 1 \leq \ell$ previous edges to choose from, so $\sum_{S_i \in N_0(S_{i-1})} 1/D(\phi)(S_{i-1}) \leq \ell n^{-\delta/2}$. Combining the above, we get

$$\begin{aligned} \mathbb{E} \text{tr}(\Gamma^\ell) &\leq \sum_{b \in \{0,1\}^\ell} \sum_{S_1 \in \mathcal{S}} \frac{1}{D(\phi)(S_1)} \sum_{S_2 \in N_{b_1}(S_1)} \frac{1}{D(\phi)(S_2)} \cdots \sum_{S_\ell \in N_{b_{\ell-1}}(S_{\ell-1})} \frac{\mathbf{1}\{S_1 \in N_{b_\ell}(S_\ell)\}}{D(\phi)(S_\ell)} \\ &\leq 2^\ell n^k (\ell n^{-\delta/2})^{\ell/2}. \end{aligned}$$

Therefore, we set $\ell = 2 \left\lceil \frac{3k \log_2 n}{4} \right\rceil$ so by Markov's inequality, we have

$$\Pr \left[\|\Gamma\| \geq n^{-\delta/5} \geq 4n^{-\delta/4} \sqrt{k \log_2 n} \right] \leq \frac{\mathbb{E} \text{tr}(\Gamma^\ell)}{(4n^{-\delta/4} \sqrt{k \log_2 n})^\ell} = \frac{1}{\text{poly}(n^k)}. \quad \blacktriangleleft$$

References

- 1 Jackson Abascal, Venkatesan Guruswami, and Pravesh K. Kothari. Strongly refuting all semi-random boolean CSPs. In *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms (SODA 2021)*, pages 454–472. SIAM, 2021. doi:10.1137/1.9781611976465.28.
- 2 Sarah R. Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *56th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2015)*, pages 689–708. IEEE Computer Society, 2015. doi:10.1109/FOCS.2015.48.
- 3 Boaz Barak and Ankur Moitra. Noisy tensor completion via the sum-of-squares hierarchy. In *Proceedings of the 29th Conference on Learning Theory (COLT 2016)*, volume 49 of *JMLR Workshop and Conference Proceedings*, pages 417–445. JMLR.org, 2016. URL: <https://proceedings.mlr.press/v49/barak16.html>.
- 4 Moses Charikar, Konstantin Makarychev, and Yury Makarychev. Integrality gaps for Sherali-Adams relaxations. In *Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC 2009)*, pages 283–292. ACM, 2009. doi:10.1145/1536414.1536455.
- 5 Amin Coja-Oghlan, Andreas Goerdt, and André Lanka. Strong refutation heuristics for random k -SAT. *Combinatorics, Probability and Computing*, 16(1):5–28, 2007. doi:10.1017/S096354830600784X.
- 6 Tommaso d'Orsi and Luca Trevisan. An Ihara-Bass formula for non-boolean matrices and strong refutations of random CSPs. In *38th Computational Complexity Conference (CCC 2023)*, volume 264 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:16. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023. doi:10.4230/LIPIcs.CCC.2023.27.
- 7 Noah Fleming, Pravesh Kothari, and Toniann Pitassi. Semialgebraic proofs and efficient algorithm design. *Foundations and Trends in Theoretical Computer Science*, 14(1–2):1–221, 2019. doi:10.1561/04000000086.
- 8 Dmitry Gavinsky, Shachar Lovett, Michael Saks, and Srikanth Srinivasan. A tail bound for read- k families of functions. *Random Structures and Algorithms*, 47(1):99–108, 2015. doi:10.1002/rsa.20532.
- 9 Michel X. Goemans and David P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42(6):1115–1145, 1995. doi:10.1145/227683.227684.
- 10 Venkatesan Guruswami, Pravesh K. Kothari, and Peter Manohar. Algorithms and certificates for boolean CSP refutation: Smoothed is no harder than random. In *Proceedings of the 54th Annual ACM SIGACT Symposium on Theory of Computing (STOC 2022)*, pages 678–689. ACM, 2022. doi:10.1145/3519935.3519955.

- 11 Samuel B. Hopkins, Tselil Schramm, and Luca Trevisan. Subexponential LPs approximate max-cut. In *61st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2020)*, pages 943–953. IEEE, 2020. doi:10.1109/FOCS46700.2020.00092.
- 12 Jun-Ting Hsieh, Pravesh K. Kothari, and Sidhanth Mohanty. A simple and sharper proof of the hypergraph moore bound. In *Proceedings of the 2023 ACM-SIAM Symposium on Discrete Algorithms (SODA 2023)*, pages 2324–2344. SIAM, 2023. doi:10.1137/1.9781611977554.CH89.
- 13 Ryan O'Donnell and Tselil Schramm. Sherali-Adams strikes back. In *34th Computational Complexity Conference (CCC 2019)*, volume 137 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 8:1–8:30. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CCC.2019.8.

A Proofs of auxiliary statements

In this section, we fill in the proofs of Lemma 3.7 and Lemma 3.8. The proofs here are all reformulations of ideas in prior works [2, 13] and are included here for completeness.

A.1 Auxiliary statements for random k -XOR

We show that with high probability, the matrix $A(\psi)$ defined in Definition 3.6 is a spectral certificate.

► **Lemma 3.7.** *Suppose ψ is a random k -XOR instance with $n^{k/2+\delta}$ constraints such that $2^{O(k)} \log^3 n \leq n^{\delta/2}$. Then $A(\psi)$ is a $\frac{\delta}{2}$ -spectral certificate for ψ with high probability.*

This follows directly from the following two statements.

► **Fact A.1** ([2, Lemma A.5]). *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with $n^{k/2+\delta}$ expected constraints. Then with high probability, $\|A(\psi)\| \leq 2^{O(k)} n^{\delta} \log^3 n$.*

► **Proposition A.2.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with $n^{k/2+\delta}$ expected constraints. Then with probability $1 - 1/\text{poly}(n)$, we have $\min_{(S_1, S_2)} \|A(\psi)_{(S_1, S_2)}\|_1 \geq \min\left(\frac{n^{2\delta}}{4}, \frac{n}{6}\right)$.*

Proof. First, suppose that $\delta < \frac{1}{2}$. Note that each $U \in [n]^k$ is included in $\mathcal{H}$ independently with probability $p = n^{-k/2+\delta}$. For all $S \in [n]^{(k-1)/2}$ and $i \in [n]$, the number of $(STi) \in \mathcal{H}$ is $\text{Binomial}(n^{(k-1)/2}, n^{-k/2+\delta})$. Then this is nonzero with probability $1 - (1 - n^{-k/2+\delta})^{n^{(k-1)/2}} \geq 1 - e^{-n^{\delta-1/2}} \geq n^{\delta-1/2}(1 - o(1))$. Recall the notation ψ^1 from Definition 4.5, and note that for any $(S_1, S_2) \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, $A(\psi^1)_{(S_1, S_2)}$ has a total of $\sum_{i \in [n]} \mathbf{1}\{\exists S_1 Ti \in \mathcal{H}\} \mathbf{1}\{\exists S_2 Ti \in \mathcal{H}\}$ nonzero entries, which is Binomial with mean $n^{2\delta}(1 - o(1))$ when $S_1 \neq S_2$ or mean $n^{\frac{1}{2}+\delta}(1 - o(1)) \geq n^{2\delta}$ when $S_1 = S_2$, and thus it has at least $n^{2\delta}/2$ nonzero entries by Chernoff with probability $1 - \exp(-\Omega(n^{2\delta}))$. Recall that if an entry of $A(\psi^1)$ is nonzero, then the corresponding entry of $A(\psi)$ is nonzero with probability at least $\frac{1}{2}$ (see Remark 4.6), so $A(\psi)_{(S_1, S_2)}$ has at least $n^{2\delta}/4$ nonzero entries with probability $1 - \exp(-\Omega(n^{2\delta}))$. The result thus holds simultaneously for all $(S_1, S_2) \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$ by a union bound.

Now suppose that $\delta \geq \frac{1}{2}$. Then following the $\delta < \frac{1}{2}$ case, the number of $(STi) \in \mathcal{H}$ is nonzero with probability at least $1 - (1 - n^{-k/2+\frac{1}{2}})^{n^{(k-1)/2}} \geq 1 - \frac{1}{e} - o(1)$. Then $\sum_{i \in [n]} \mathbf{1}\{\exists S_1 Ti \in \mathcal{H}\} \mathbf{1}\{\exists S_2 Ti \in \mathcal{H}\}$ is Binomial with mean at least $(1 - \frac{1}{e} - o(1))^2 n$, and thus $A(\psi^1)_{(S_1, S_2)}$ has at least $n/3$ nonzero entries with probability $1 - \exp(-\Omega(n))$ by Chernoff, and thus $A(\psi)_{(S_1, S_2)}$ has at least $n/6$ nonzero entries with probability $1 - \exp(-\Omega(n))$, and as before this holds simultaneously for all (S_1, S_2) by a union bound. ◀

► **Lemma 3.8** (Abstracting out [13]). *Let A be a symmetric matrix with integer entries with rows and columns indexed by t -tuples of $[n]$, and let D be its degree matrix. Let $\pi_* = \frac{1}{\sum_{S,T} |A_{S,T}|} \min_S \|A_S\|_1$, $\rho(\cdot)$ denote the spectral radius of a matrix, and $\eta > \min(\pi_*^{-1/2}, \rho(D^{-1}A))$. Take $\ell := \left\lceil \frac{1}{4} \frac{\log(\eta^2 \pi_*)}{\log(\rho(D^{-1}A)/\eta)} \right\rceil$ and $p := \lceil (\frac{2}{5\eta})^{2\ell} \rceil$. Then for $R(A) := tp\ell + t$, any R -local pseudoexpectation $\mathcal{D}$ on $x_1, \dots, x_n$ satisfies*

$$\left| \mathbb{E}_{\mathcal{D}}(x^{\odot t})^\top A x^{\odot t} \right| \leq \eta \sum_{S,T} |A_{S,T}|,$$

where $x^{\odot t}$ is the n^t -dimensional vector whose entries are indexed by t -tuples of $[n]$ and are equal to the product of the corresponding entries.

We utilize the following reparameterization of results from [13]:

► **Fact A.3** ([13, Theorem 3.6, Corollary 3.7]). *Let $G = (V, E)$ be a multigraph with signs b_{uv} for all $uv \in E$ and signed transition operator $\bar{K}$. Suppose we have $\pi_* = \min_{v \in V} \frac{\deg(v)}{2|E|}$. Given $\gamma > \min(\pi_*^{-1/2}, \rho(\bar{K}))$, take $\ell = \left\lceil \frac{1}{4} \frac{\log(\gamma^2 \pi_*)}{\log(\rho(\bar{K})/\gamma)} \right\rceil$, and take $k = \lceil (\frac{2}{5\gamma})^{2\ell} \rceil$. Then for $R = k\ell + 1$, it holds that R -local degree-2 SOS can deduce that for all $x \in \{\pm 1\}^V$,*

$$-\gamma \leq \frac{1}{|E|} \sum_{uv \in E} b_{uv} x_u x_v \leq \gamma.$$

Proof of Lemma 3.8. For all $S \subseteq [n]^t$, define a variable y_S . Define an R -local pseudoexpectation $\mathcal{D}'$ over the y_S by setting, for $r \leq R$,

$$\tilde{\mathbb{E}}_{\mathcal{D}'}[y_{S_1} y_{S_2} \cdots y_{S_r}] := \tilde{\mathbb{E}}_{\mathcal{D}} \prod_{i \in S_1} x_i \prod_{i \in S_2} x_i \cdots \prod_{i \in S_r} x_i.$$

Applying Fact A.3 with $\eta = \gamma$, we see that $\tilde{\mathbb{E}}_{\mathcal{D}'} \sum_{S,T} y_S y_T A_{S,T} \leq \eta \sum_{S,T} |A_{S,T}|$, and therefore

$$\begin{aligned} \tilde{\mathbb{E}}_{\mathcal{D}}(x^{\odot t})^\top A x^{\odot t} &= \sum_{S,T} A_{S,T} \tilde{\mathbb{E}}_{\mathcal{D}} \prod_{i \in S} x_i \prod_{i \in T} x_i \\ &= \tilde{\mathbb{E}}_{\mathcal{D}'} \sum_{S,T} A_{S,T} y_S y_T \leq \eta \sum_{S,T} |A_{S,T}|. \end{aligned} \quad \blacktriangleleft$$

A.2 Auxiliary statements for semirandom k -XOR

► **Lemma 4.10.** *Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ and ϕ be an f -induced 2-XOR instance of ψ . Suppose that for all R -local pseudodistributions $\mathcal{D}'$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}'} \phi(y)| \leq 2\gamma|\mathcal{H}|$. Then for all kR -local pseudodistributions $\mathcal{D}$, we have $|\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x)| \leq 2\gamma|\mathcal{H}|$.*

Proof. Let $\mathcal{D}$ be a kR -local pseudodistribution over variables $x_1, \dots, x_n$. Letting $\mathcal{S}$ be the variable set of ϕ , define an R -local pseudodistribution $\mathcal{D}'$ over variables y_S of ϕ by setting, for $r \leq R$,

$$\tilde{\mathbb{E}}_{\mathcal{D}'}[y_{S_1} y_{S_2} \cdots y_{S_r}] := \tilde{\mathbb{E}}_{\mathcal{D}} \prod_{i \in S_1} x_i \prod_{i \in S_2} x_i \cdots \prod_{i \in S_r} x_i.$$

Then we have

$$\tilde{\mathbb{E}}_{\mathcal{D}} \psi(x) = \tilde{\mathbb{E}}_{\mathcal{D}} \sum_{(S,T) \in \mathcal{B}} b_{f(S,T)} \prod_{i \in S} x_i \prod_{i \in T} x_i = \tilde{\mathbb{E}}_{\mathcal{D}'} \sum_{(S,T) \in \mathcal{B}} b_{f(S,T)} y_S y_T = \tilde{\mathbb{E}}_{\mathcal{D}'} \phi(y) \leq 2\gamma|\mathcal{H}|. \quad \blacktriangleleft$$

► **Lemma 4.13.** Let $\psi(x) = \sum_{U \in \mathcal{H}} b_U \cdot x^U$ be a semirandom k -XOR instance with constraint set $\mathcal{H} \subseteq [n]^k$ such that $|\mathcal{H}| = n^{k/2+\delta}$ and let P be a regular partition of the $\mathcal{H}_i$ s. Suppose $\mathcal{H}$ has the following properties:

- For all $S \in [n]^{(k-1)/2}$, $i \in [n]$, there are at most $n^{\delta/2}$ total clauses of the form ST_i or TS_i in $\mathcal{H}$.
- For all $S_1 T_1 \in [n]^{(k-1)/2} \times [n]^{(k-1)/2}$, there are at most $n^{\delta/2}$ total clauses of the form $S_1 T_1 i \in \mathcal{H}$.

Then with high probability, there exists a principal submatrix $A(\psi)$ of $A'(\psi)$ such that all of the following hold:

- The deleted rows and columns of $A'(\psi)$ have ℓ_1 norms summing to at most

$$30n^{-\delta/2} \sum_{i \in [n]} \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| (|\mathcal{H}_i^{(j)}| - 1).$$

- Each row of $A(\psi)$ has at least n^δ nonzero entries.
- Each row of $A(\psi)$ has at least $\frac{1}{3}$ as many nonzero entries as the corresponding row of $A(\psi^1)$, where we recall that $A'(\psi^1)$ is the matrix defined with the same $\mathcal{H}$ as ψ but with all b_U s equal to 1.
- Each entry of $A(\psi)$ has absolute value at most $4kn^{\delta/4}\sqrt{\log n}$.

We use the following concentration result for dependent indicators.

► **Fact A.4** (Chernoff Bound for Weakly Dependent Random Variables). [8] Suppose we have a set S of independent random variables and indicator random variables $Y_1, \dots, Y_r$ with mean μ , such that each Y_i is a function of some $T_i \subseteq S$ and no $s \in S$ is contained in more than k total T_i 's. Then for any $\varepsilon > 0$, we have

$$\Pr[Y_1 + \dots + Y_r \notin (\mu \pm \varepsilon)r] \leq 2e^{-2\varepsilon^2 r/k}.$$

Proof of Lemma 4.13. Note that since P has at most $2n$ parts, the total ℓ_1 norm of $A'(\psi^1)$ is $\sum_i \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| \geq 2n \cdot \frac{1}{2} \left(\frac{|\mathcal{H}|}{2n} \right)^2 (1 - o(1)) \geq \frac{1}{5} n^{k-1+2\delta}$. Also, since for all $(S_1, S_2), (T_1, T_2)$, there are at most $n^{\delta/2}$ total i such that $S_1 T_1 i \in \mathcal{H}$ or $S_2 T_2 i \in \mathcal{H}$, there are at most $n^{\delta/2}$ pairs of clauses that contribute to each entry of $A'(\psi)$; in other words, every entry of $A'(\psi^1)$ is at most $n^{\delta/2}$. Consider an arbitrary entry $(S_1, S_2), (T_1, T_2)$. We note that $A(\psi)_{(S_1, S_2), (T_1, T_2)}$ is the sum of $A(\psi^1)_{(S_1, S_2), (T_1, T_2)}$ products of independent Rademacher random variables, all of which are distinct, and thus it is also the sum of independent Rademachers, so $A(\psi)_{(S_1, S_2), (T_1, T_2)} \leq 4k \sqrt{A(\psi^1)_{(S_1, S_2), (T_1, T_2)} \log n} \leq 4kn^{\delta/4} \sqrt{\log n}$ with probability $1 - \frac{1}{n^{2k}}$ and thus it holds for all n^{2k-2} entries with probability $\geq 1 - 1/n^2$.

Now, while there exists some row of $A'(\psi^1)$ indexed by (S_1, S_2) whose total sum is less than $3n^{3\delta/2}$, delete that row and its corresponding column. Then the sum of the elements of $A'(\psi^1)$ we delete is at most $2 \cdot 3n^{k-1+3\delta/2} \leq 30n^{-\delta/2} \sum_{i \in [n]} \sum_{j \in [p_i]} |\mathcal{H}_i^{(j)}| (|\mathcal{H}_i^{(j)}| - 1)$, and first condition then follows from the fact that the ℓ_1 norm of rows in $A'(\psi)$ is upper bounded by the sum of the corresponding row in $A'(\psi^1)$. Let $A(\psi)$ and $A(\psi^1)$ be the matrices after the deletion step.

Now consider a surviving row indexed by some (S_1, S_2) , whose ℓ_1 norm is at least $3n^{3\delta/2}$. Since every entry of $A(\psi^1)$ is at most $n^{\delta/2}$, there are at least $3n^\delta$ nonzero entries in the row indexed by (S_1, S_2) , each of which are indexed by some (T_1, T_2) . Recall that if an entry $A(\psi^1)_{(S_1, S_2), (T_1, T_2)}$ is nonzero, then $A(\psi)_{(S_1, S_2), (T_1, T_2)}$ is nonzero with probability at least $\frac{1}{2}$ (Remark 4.6). Furthermore, each entry in row (S_1, S_2) of $A(\psi)$ depends on variables of the form $b_{S_1 T_i}$ and $b_{S_2 T_i}$. Note that anytime some $b_{S_1 T_i}$ appears in row (S_1, S_2) , it must be multiplied with some $b_{S_2 T' i}$ (in which case it corresponds to the (T, T') entry of that

row). However, we are given that there are at most $n^{\delta/2}$ total $S_2 T' i \in \mathcal{H}$, and thus $b_{S_1 T i}$ appears in at most $n^{\delta/2}$ entries in row (S_1, S_2) . Therefore, the family of indicators of entries in row (S_1, S_2) being nonzero is read- $n^{\delta/2}$ with respect to the b_U s, and thus after coupling to a family of indicators which occur with probability $\frac{1}{2}$, by Fact A.4, with probability $1 - \exp(-\Omega(n^{\delta/2}))$, the number of nonzero entries in row (S_1, S_2) of $A(\psi)$ is at least $\frac{1}{3}$ the number of nonzero entries in row (S_1, S_2) of $A(\psi^1)$, which is at least n^δ . Union bounding over all $\leq n^{k-1}$ rows of $A'(\psi)$ gives the desired result. $\blacktriangleleft$