HYPERGRAPH CYCLE ENUMERATION FOR EFFICIENT INFERENCE OF PLANTED k-XOR WITH SMOOTH TRADE-OFF

Anonymous authors
Paper under double-blind review

ABSTRACT

We introduce *hypergraph cycle enumeration* (HCE), a simple spectral framework for detecting and recovering planted signals in noisy hypergraph inference problems. In this work, we focus on planted noisy k-XOR, a canonical average-case constraint satisfaction problem (CSP) closely connected to tensor PCA, refutation of random CSPs, and computational-statistical gaps in high-dimensional inference. We show that HCE yields weak recovery with a hierarchical threshold (determined by a fixed parameter l), matching the fixed-level Kikuchi spectral threshold curve up to a constant independent of l and polylogarithmic factors for even k and, after support-based trimming, for odd k, while giving a direct cycle-counting interpretation of the mechanism. Both theoretically and empirically, HCE achieves comparable recovery behavior to Kikuchi methods while using a substantially smaller matrix representation, leading to significant runtime and memory improvements in the sparse regimes we test.

1 INTRODUCTION

Many problems in high-dimensional statistics and average-case complexity ask the same basic question: when is a weak, noisy signal hidden in random data detectable, and when can it be recovered *efficiently*? In many settings there is a striking separation between what is possible with unlimited computation and what is possible with known polynomial-time algorithms. Understanding such informational–computational gaps has become a central goal in the study of random inference problems, with examples ranging from planted clique and sparse PCA to tensor PCA, random constraint satisfaction problems, and learning parity with noise.

A particularly clean setting in which these questions arise is that of noisy parity measurements. Suppose an unknown vector $x \in \{\pm 1\}^n$ is observed only through random local tests, each revealing a noisy parity of a small number of coordinates. How many such tests are needed to identify the hidden vector, and how many are needed by an *efficient* algorithm? This captures a basic tension: although each individual observation is extremely simple, combining many noisy high-order parities appears to be computationally difficult. The resulting problem is a canonical model for the algorithmic limits of extracting a planted discrete signal from sparse, noisy, high-order data.

Here we examine the planted noisy k-XOR problem, which is closely related to tensor PCA [9; 10; 16] and to the refutation of random sparse CSPs [1; 4; 6; 15]. A k-XOR instance $\mathcal{I}$ consists of clauses (S_a, b_a), $a = 1, \dots, m$, where each support $S_a \subseteq [n]$ has size k and each label $b_a \in \{\pm 1\}$. Given an assignment $z \in \{\pm 1\}^n$, the clause is satisfied when $z_{S_a} := \prod_{v \in S_a} z_v = b_a$. In the planted noisy model there is a hidden $x \in \{\pm 1\}^n$ such that

$$b_a \;=\; \eta_a \prod_{v \in S_a} x_v, \qquad a = 1, \dots, m, \tag{1}$$

where the supports S_a are i.i.d. uniform k-subsets of $[n]$ and the noise variables $\eta_a \in \{\pm 1\}$ are independent with $\mathbb{E}[\eta_a] = \rho$. The *detection* task is to decide whether the data has a planted solution; the *recovery* task is to reconstruct x (up to a global sign when it is not identifiable). The noiseless case $\rho = 1$ is solved by Gaussian elimination whenever $m \gtrsim n$. Information-theoretically, weak recovery (non-trivial correlation with x) is possible already at $m = \Theta(n)$ clauses for any constant

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

$\rho \in (0, 1]$, while exact recovery requires at least the coupon-collector scale $m = \Omega(n \log n)$ [17] No polynomial-time algorithm is known below $m = \widetilde{\Theta}(n^{k/2})$ for even $k \geq 4$, and this threshold is conjectured optimal for efficient algorithms [1; 15; 13]. The noisy case is conjectured to exhibit a sharp statistical–computational gap [1; 15; 13; 20].

Historically this and related problems has been approached with methods such as tensor unfolding [16] and the sum-of-squares (SoS) hierarchy [8; 10; 11]. We review these algorithms in Appendix K. The dominant family of algorithms in the computational threshold regime is the *Kikuchi hierarchy* [20; 7; 17]: spectral methods indexed by a degree parameter ℓ that trade runtime $n^{O(\ell)}$ for a sample complexity that decreases smoothly with ℓ. This *smooth trade-off* is precisely the property that makes the Kikuchi and SoS hierarchies interesting, since it interpolates between the polynomial-time and the information-theoretic thresholds.

Our new contributions can be summarized as follows:

- We introduce **hypergraph cycle enumeration** (HCE), a hierarchical algorithmic approach to planted noisy k-XOR (Section 3). On level l HCE aggregates sets of l clauses each giving a weak parity signal into a low-dimensional object with detectable correlation to the planted assignment.
- We provide **detection** and **recovery guarantees** based respectively on collecting size-l even covers, and using size-l sets of clauses with symmetric difference of size 2 to generate a dense $n \times n$ matrix, the principal component of which aligns with the solution. For both detection and recovery, the parameter l traces a similar threshold curve as in the Kikuchi and SoS hierarchies: increasing l combines more clauses at higher computational cost. Specifically, HCE succeeds at same threshold clause density $\widetilde{\Omega}(\rho^{-2}(n/l)^{k/2-1})$. Thus, the HCE hierarchy exhibits a **smooth trade-off**, making it of theoretical interest as the third known hierarchy with this property after SoS and Kikuchi.
- We compare the **complexity** of HCE to the best competitor, the Kikuchi method. While the asymptotic complexity is worse for HCE, we run **experiments** for $k = 4, 5, 6$ (Section 5) showing that HCE matches Kikuchi recovery overlap at equal level while using substantially less memory and runtime, especially in the ultra-sparse regimes (e.g., at $n = 50$, $k = 4$, level 6, HCE completes in $\approx$300 s versus $\approx$5.5 h per instance for Kikuchi, a $\sim 60\times$ speed-up; Appendix J.8). Thus HCE presents a practical advantage, being able to solve some instances beyond the capability of Kikuchi, and the advantage is largest in the hardest (sparsest) regime.

Note: for recovery, we use smooth trade-off in the soft-asymptotic sense common in work on the Kikuchi hierarchy, where thresholds are compared after suppressing logarithmic factors [20]. Our detection result (Theorem 3) is stronger: it has no logarithmic loss. We conjecture that a normalized or suitably trimmed HCE recovery operator removes the logarithmic slack, uniformly over a nontrivial range of growing levels $l = l(n)$; proving this remains open.

2 Preliminaries

2.1 Kikuchi method

The Kikuchi hierarchy [20; 7; 12] is the benchmark spectral method for planted noisy k-XOR, and the main baseline for HCE. For even k and a level $\ell \geq k/2$, the symmetric level-ℓ Kikuchi matrix is indexed by ℓ-subsets of variables,

$$K_{T,U} = \sum_{a=1}^{m} b_a \, \mathbf{1}[T \triangle U = S_a], \qquad T, U \in \binom{[n]}{\ell}, \tag{2}$$

so each clause links two ℓ-sets that differ exactly by its support. Its top eigenvector correlates with the planted signal once $m \gtrsim \widetilde{\Theta}\big(n(n/\ell)^{k/2-1}\big)$, at cost $n^{O(\ell)}$; increasing ℓ gives the smooth trade-off discussed in the introduction. Odd k requires an unbalanced lifting [20].

HCE is similar in spirit: both methods multiply clause labels along chains of symmetric differences, so that planted signs cancel. The difference is where the spectral step takes place. Kikuchi works

on a $\binom{n}{\ell}$-dimensional lifted space, whereas HCE compresses all such chains onto an $n \times n$ matrix (Proposition 7, Appendix D). The two hierarchies follow the same threshold curve (Section 3.1), so we compare them at matched level in Table 1 and Section 5. In the experiments, K is stored sparsely, solved by Lanczos and decoded by voting; HCE uses the meet-in-the-middle construction of Section 4.3 and the post-processing of Appendix C (details in Appendix B).

2.2 Hypergraph Even covers and witnesses

Throughout, $[n] = \{1, \dots, n\}$ and $\mathcal{K} := \binom{[n]}{k}$ denotes the complete k-uniform hypergraph, with $M := |\mathcal{K}| = \binom{n}{k}$ the number of possible clauses. The instance hypergraph is $\mathcal{H} \subseteq \mathcal{K}$ with $|\mathcal{H}| = m$, and we parameterise the clause density by

$$m = C\, n^{k/2}. \tag{3}$$

For a set of clauses $\mathcal{S} \subseteq \mathcal{H}$ we write $\triangle(\mathcal{S}) := \triangle_{S \in \mathcal{S}} S \subseteq [n]$ for the symmetric difference (equivalently, the $\mathbb{F}_2$ sum of indicator vectors) of their supports, and $X_{\mathcal{S}} := \prod_{a \in \mathcal{S}} b_a$ for the product of their labels. A full notation table is given in Appendix E.

Because we work with hypergraphs, the word "cycle" is ambiguous; we fix the following standard notion, which is the correct analogue for $\mathbb{F}_2$ linear algebra.

Definition 1 (Even cover / hypergraph cycle)**.** A *proper even l-cover* is a set $F = \{S_1, \dots, S_l\}$ of l *distinct* hyperedges such that every vertex is contained in an even number of them, i.e. $\triangle(F) = \emptyset$. We write $\mathcal{C}_l$ for the set of proper even l-covers contained in $\mathcal{H}$ and $Z_l := |\mathcal{C}_l|$.

An estimator achieves *weak recovery* if $|\langle \hat{x}, x \rangle|/n = \Omega(1)$, *almost-exact recovery* if this overlap tends to one, and *exact recovery* if $\hat{x} = \pm x$. The recovery theorem below is a weak-recovery statement.

Definition 2 (Witness)**.** For $i \neq j$, an (i, j)*-witness of level l* in the recovery model is an index set $I \in \binom{[m]}{l}$ with $\triangle_{a \in I} S_a = \{i, j\}$. We write $\mathcal{W}_l(i, j)$ for the set of such index sets, $W_l(i, j) := |\mathcal{W}_l(i, j)|$, and $N_l := \mathbb{E}\, W_l(i, j)$ (which by symmetry does not depend on i, j).

In simple terms, a witness is an even cover "with two loose ends". HCE, as will be introduced below, is an algorithm that traces these "loose ends" in order to detect if there is a planted instances, and if so, recover it. Throughout this work, we reserve l for the *HCE level* and ℓ for the *Kikuchi level*; the two are distinct parameters and are only approximately equal by coincidence at small k (Section 3.1).

3 Hypergraph Cycle Enumeration

3.1 Cycles carry the signal

The core idea is to use the clause structure of the k-XOR instance directly. Let $F \in \mathcal{C}_l$ be an even cover. Every variable in F appears an even number of times, so all planted signs cancel and

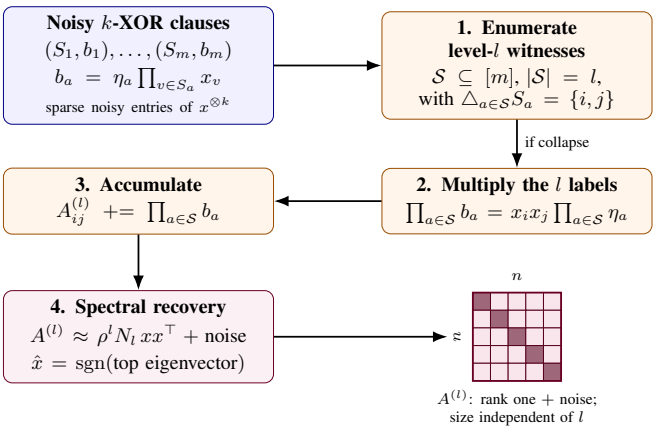
$$X_F \;=\; \prod_{S \in F} b_S \;=\; \prod_{S \in F} \eta_S, \qquad \mathbb{E}[X_F] = \rho^l, \qquad \mathrm{Var}(X_F) = 1 - \rho^{2l}. \tag{4}$$

Under the null model ($\rho = 0$) the mean vanishes, so even covers are exactly the local structures that certify planting. Collecting enough of them therefore enables detection; hence the name *hypergraph cycle enumeration*. Define the *level-l discriminator*

$$X(l) \;:=\; \begin{cases} Z_l^{-1} \sum_{F \in \mathcal{C}_l} X_F, & Z_l > 0, \\ 0, & Z_l = 0. \end{cases} \tag{5}$$

If the summands were uncorrelated — as they are for edge-disjoint covers, and asymptotically after the overlap terms are controlled through Proposition 13 — then detection with constant probability requires

$$\frac{(\mathbb{E} X(l))^2}{\mathrm{Var}\, X(l)} \;\approx\; \rho^{2l} Z_l \;\gg\; 1 \iff Z_l = \Omega(\rho^{-2l}). \tag{6}$$



HCE at level l

Input: clauses $\{(S_a, b_a)\}_{a=1}^m$ with $|S_a| = k$; level l with kl even. **Output:** $\hat{x} \in \{\pm 1\}^n$, estimating x up to a global sign.

1. Set $r \leftarrow \lfloor l/2 \rfloor$, $s \leftarrow \lceil l/2 \rceil$, and $G^{(l)} \leftarrow 0$.
2. Hash every $\mathcal{S}$ with $|\mathcal{S}| = s$ under the key $\triangle(\mathcal{S}) \in \mathbb{F}_2^n$, storing only the aggregate sum of $X_{\mathcal{S}}$ for that key.
3. For every $\mathcal{S}'$ with $|\mathcal{S}'| = r$ and every $i < j$: look up $u = \triangle(\mathcal{S}') \triangle \{i, j\}$ and add the aggregate at u, times $X_{\mathcal{S}'}$, to $G_{ij}^{(l)}$ and $G_{ji}^{(l)}$.
4. Recover $A^{(l)}$ from $G^{(l)}$ and the lower-level matrices by the overlap recurrence (138).
5. Return $\hat{x} = \mathrm{sgn}(\hat{u})$ with $\hat{u}$ the top eigenvector of $A^{(l)}$; for odd k and $l \geq 6$, use the trimmed matrix $\bar{A}^{(l)}$ of Definition 34.

Figure 1: Hypergraph cycle enumeration (HCE). *Left:* the clauses are sparse noisy entries of the rank-one order-k tensor $x^{\otimes k}$; HCE multiplies l clauses whose supports XOR to a pair $\{i, j\}$, so that every variable cancels modulo two except i and j, leaving an effective noisy 2-XOR constraint on $x_i x_j$. The case $l = 2$ is elementary: two clauses with $S_a \triangle S_b = \{i, j\}$ share a common $(k-1)$-set. *Right:* the meet-in-the-middle implementation, analysed in Section 4.3.

For random uniform hypergraphs Z_l concentrates around its expectation (Corollary 11), whose asymptotics follow from an exact Kravchuk-polynomial count of even covers in the complete hypergraph (Appendix H) combined with the sampling step (Appendix I, Theorem 41). Writing $Z_l^K = (\alpha_{k,l} + o_n(1))\, n^{kl/2}$ for the count on the complete hypergraph, sampling gives $\mu_l := \mathbb{E}[Z_l] = (\alpha_{k,l} + o_n(1))(k!\, C)^l$, and for large admissible l

$$\mathbb{E}[Z_l] = \big(L_k + o_l(1)\big)\, l^{-1/2} \left(\gamma_k^{-1} \frac{m\, l^{k/2-1}}{n^{k/2}}\right)^l, \quad \gamma_k := k^{-k/2} e^{k/2-1}, \quad L_k := \tfrac{e^{-(k-1)/2}}{\sqrt{\pi}}. \tag{7}$$

Detection is therefore possible for large n, constant k and large admissible l whenever

$$m \;\gg\; \gamma_k\, \rho^{-2}\, n \left(\frac{n}{l}\right)^{k/2-1}. \tag{8}$$

Up to logarithmic factors, the Kikuchi detection threshold [20; 7; 12] is the same condition with γ_k replaced by $\binom{k}{k/2}^{-1}$ and l replaced by the Kikuchi level ℓ. Thus HCE exhibits the same smooth trade-off, and at $k = 4$ the two constants agree to within 2%.

3.2 From cycles to a matrix: the recovery matrix $A^{(l)}$

Recovery proceeds by reducing the problem to standard PCA. Here the clauses are the i.i.d. sample $S_1, \ldots, S_m$, and witnesses are indexed by clause indices: for $I \subseteq [m]$ put $\partial I := \triangle_{a \in I} S_a$ and define the *recovery matrix*

$$A_{ij}^{(l)} := \sum_{\substack{I \in \binom{[m]}{l} \\ \partial I = \{i,j\}}} \prod_{a \in I} b_a \quad (i \neq j), \qquad A_{ii}^{(l)} := 0. \tag{9}$$

Every I in the sum has the property that all variables cancel modulo two except i and j, so under the planted model $\prod_{a \in I} b_a = x_i x_j \prod_{a \in I} \eta_a$ and, conditionally on the supports,

$$\mathbb{E}\big[A_{ij}^{(l)} \;\big|\; S_1, \ldots, S_m\big] \;=\; \rho^l\, x_i x_j\, W_l(i, j). \tag{10}$$

Hence $A^{(l)}$ is a weighted noisy estimator of the rank-one matrix $xx^\top$. It is convenient to normalise by the expected witness count and its n-scaled version,

$$N_l = \binom{m}{l} \mathbb{P}\big[\partial\{1, \ldots, l\} = \{i, j\}\big], \qquad d_l := nN_l, \qquad \theta_l := \rho^l d_l. \tag{11}$$

Writing $D_x = \mathrm{diag}(x)$ and $\widehat{A}^{(l)} = D_x A^{(l)} D_x$, the i.i.d. model gives exactly

$$\mathbb{E}\widehat{A}^{(l)} = \rho^l N_l\big(\mathbf{1}\mathbf{1}^\top - I\big), \qquad A^{(l)} = \rho^l N_l\, xx^\top + R^{(l)}, \tag{12}$$

so the planted eigenvalue is $\rho^l N_l(n-1) = \theta_l(1-1/n)$ while every orthogonal eigenvalue of the mean is $-\rho^l N_l$. The exactly centered conjugated residual $\widetilde{R}^{(l)} := \widehat{A}^{(l)} - \mathbb{E}\widehat{A}^{(l)}$ differs from $R^{(l)}$ only by the deterministic diagonal shift,

$$D_x R^{(l)} D_x = \widetilde{R}^{(l)} - \rho^l N_l I, \qquad \big|\|R^{(l)}\|_{\mathrm{op}} - \|\widetilde{R}^{(l)}\|_{\mathrm{op}}\big| \le \theta_l/n. \tag{13}$$

Recovery is, thus, a signal–noise comparison: it suffices that $\|R^{(l)}\|_{\mathrm{op}}$ be suitably smaller than θ_l. Comparing this condition across fixed levels gives the formal trade-off curve. HCE estimates x from the sign pattern of the top eigenvector of $A^{(l)}$; see Figure 1.

4 Theoretical Guarantees

We state the two main results here and defer complete proofs to Appendices F and G.

The detection and recovery analyses use different, equally natural, models for the clauses, and we keep them separate throughout. For *detection*, $\mathcal{H}$ is a uniformly random m-edge subhypergraph of $\mathcal{K}$: the supports are drawn uniformly *without replacement*, so $\mathcal{H}$ is a simple hypergraph and even covers are sets of distinct edges. For *recovery*, $S_1, \dots, S_m$ are *i.i.d.* uniform k-subsets, so repeated supports may occur and witnesses are indexed by clause indices rather than by edges; repeated supports are retained rather than discarded. The two laws do not coincide at $m = Cn^{k/2}$: the expected number of coinciding support pairs is $\binom{m}{2}/\binom{n}{k} \to C^2 k!/2$. This is immaterial because each theorem is proved entirely in its stated model; no estimate is transferred between them.

4.1 Detection

The heuristic of Section 3.1 used two unproven simplifications: that Z_l concentrates around $\mu_l = \mathbb{E}_{\mathcal{H}}[Z_l]$, and that the variance of $X(l)$ is dominated by edge-disjoint pairs of covers. Both are controlled by a single combinatorial estimate. Let n_s be the number of ordered pairs of even l-covers of $\mathcal{H}$ sharing exactly s hyperedges. Lemma 8 computes the corresponding count on the complete hypergraph *with its constant*, and after sampling this yields

$$\frac{\mathrm{Var}_{\mathcal{H}}[Z_l]}{\mathbb{E}_{\mathcal{H}}[Z_l]^2} = \sum_{s=1}^{l} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1) = \Theta_{k,C}\big(l^{-(k-2)\pi_k}\big) + o_n(1), \qquad \mu_q := \alpha_{k,q}\,(C\,k!)^q, \tag{14}$$

with π_k as in Section 2. The rate $l^{-(k-2)\pi_k}$ is the mechanism behind the trade-off: at fixed density, deeper levels make overlapping covers relatively rarer, and the discriminator behaves like an average of independent ± 1 variables of mean ρ^l. This makes the error probabilities explicit.

Theorem 3 (Smooth trade-off for detection)**.** *Fix $k > 2$, $\rho \in (0, 1]$ and a target error probability $\delta \in (0, 1)$. There is a constant $c_k > 0$, depending only on k, such that the following holds for every fixed admissible $l \ge c_k \delta^{-2/((k-2)\pi_k)}$ and every fixed $C > 0$. Set $m = Cn^{k/2}$ (assumed integral). Then, for all sufficiently large n, the test that declares the instance planted when $Z_l > 0$ and $X(l) > \rho^l/2$ has false-positive and false-negative probabilities each at most δ provided*

$$m \;\ge\; \left(\frac{32}{\delta}\right)^{1/l} \frac{\rho^{-2}}{k!\,\alpha_{k,l}^{1/l}}\, n^{k/2}. \tag{15}$$

For large l this reads $m \ge \theta_{l,k,\delta}\, \rho^{-2}\, n\, (n/l)^{k/2-1}$ with $\theta_{l,k,\delta} = (1 + o_l(1))\big(32\sqrt{l}/(\delta L_k)\big)^{1/l}\gamma_k$, i.e. exactly the threshold (8) up to an explicit factor tending to one as l grows. The same sample condition gives a test whose decision rule does not use ρ: declare the instance planted when $Z_l > 0$ and

$$X(l) > \sqrt{\frac{2}{\delta Z_l}}. \tag{16}$$

Its false-positive and false-negative probabilities are again at most δ for all sufficiently large n.

4.2 RECOVERY

By (12), $A^{(l)}$ is a rank-one spike of strength θ_l along $xx^\top$ plus the residual $R^{(l)}$. By Weyl's and the Davis–Kahan inequalities, we show that its top eigenvector aligns with x once $\|R^{(l)}\|_{\mathrm{op}}$ is small compared with θ_l. For recovery we must control $\|R^{(l)}\|_{\mathrm{op}}$ globally rather than entrywise. The exact overlap formula in Appendix G shows that the entry variance contains both the N_l term and correlation contributions from overlapping witnesses. Entrywise information alone cannot exclude their localised effects; we therefore bound high trace moments of the exactly centered residual $\widetilde{R}^{(l)}$ of (13).

Splitting witnesses according to whether i and j lie in the same support component (*boundary-connected*) or in different ones (*boundary-disconnected*) gives $\widetilde{R}^{(l)} = R^{(l)}_{\mathrm{bc}} + R^{(l)}_{\mathrm{bd}}$. A Hoeffding decomposition controls the first sector (Appendix G), and for even k a parity argument gives $R^{(l)}_{\mathrm{bd}} = 0$ identically.

Theorem 4 (HCE weak recovery for even k)**.** *Fix even $k \geq 4$, admissible $l \geq 2$, $\rho \in [0,1]$, $0 < C \leq n^{k/2-1}$ with $d_l \geq 1$, and set $a = (k+1)l + 3$. There are constants $B = B_{k,l}$ and $n_0 = n_0(k,l)$ such that for all $n \geq n_0$ and even $q \geq 2$,*

$$\left(\mathbb{E}\,\mathrm{tr}\,|\widetilde{R}^{(l)}|^q\right)^{1/q} \leq n^{1/q}(Bq)^a\left[d_l^{1/2} + \rho^{l-2}d_l^{1-1/l}\right]. \tag{17}$$

Hence, for every fixed $\gamma > 0$, with probability at least $1 - n^{-\gamma}$,

$$\|R^{(l)}\|_{\mathrm{op}} \leq \mathfrak{R}_{n,\gamma} := B_\gamma(\log n)^a\left[d_l^{1/2} + \rho^{l-2}d_l^{1-1/l}\right] + \frac{\theta_l}{n}. \tag{18}$$

If $\theta_l > (3+\varepsilon)\mathfrak{R}_{n,\gamma}$ for some $\varepsilon > 0$, put $\delta = \mathfrak{R}_{n,\gamma}/(\theta_l - \mathfrak{R}_{n,\gamma})$. Then the top eigenvalue of $A^{(l)}$ is isolated, and its unit eigenvector $\hat{u}$ and $\hat{x} = \mathrm{sgn}(\hat{u})$ satisfy, with $u = x/\sqrt{n}$ and up to a global sign,

$$|\langle \hat{u}, u\rangle| \geq \sqrt{1-\delta^2} > 0, \quad \textit{and} \quad \frac{|\langle \hat{x}, x\rangle|}{n} \geq 1 - 4\delta^2 > 0. \tag{19}$$

The exact signal-to-noise ratio is $\mathrm{SNR}_l = \theta_l(1-1/n)/\|\widetilde{R}^{(l)}\|_{\mathrm{op}}$, and, ignoring fixed constants, it diverges as soon as

$$\rho^{2l} d_l \gg (\log n)^{al}. \tag{20}$$

It remains to count witnesses. Appendix J shows $N_l = (1+o(1))\kappa_{l,k}\, m^l/n^{kl/2+1}$ with $\kappa_{l,k}^{1/l} \sim \frac{e}{l}\left(\frac{kl}{e}\right)^{k/2}$, so that $d_l = nN_l = (1+o(1))\kappa_{l,k}C^l$ when $m = Cn^{k/2}$. Thus the separation condition becomes a condition on the clause density alone.

Corollary 5 (Smooth trade-off for recovery)**.** *Fix even $k \geq 4$, $0 < \rho \leq 1$, $\gamma > 0$, and $\varepsilon > 0$. There is a constant $K_{k,\gamma,\varepsilon} > 0$, independent of l, such that for every fixed admissible $l \geq 2$, with $a = (k+1)l + 3$, the following holds for all sufficiently large n (where the lower bound on n may depend on l). If $m = Cn^{k/2}$, $C \leq n^{k/2-1}$, and*

$$C \;\geq\; K_{k,\gamma,\varepsilon}\,\rho^{-2}\kappa_{l,k}^{-1/l}(\log n)^{(k+1)l+3}, \tag{21}$$

then, for all sufficiently large n, the binary output of HCE satisfies, with probability at least $1-n^{-\gamma}$,

$$\frac{|\langle \hat{x}, x\rangle|}{n} \geq 1 - \frac{4}{(2+\varepsilon)^2} > 0.$$

For fixed k and large l, $\kappa_{l,k}^{-1/l} \sim \gamma_k l^{1-k/2}$, so (21) traces the Kikuchi curve up to an l-independent constant; the results are stated for fixed l only (Appendix G). For odd k, the same curve holds at every even l after trimming a sublinear, support-determined set $\mathcal{B}_l$ of coordinates, which is empty for $l \in \{2,4\}$ (Theorem 37 and Corollary 38 in Appendix G.7).

Remark 6. Since the raw HCE matrix can localise on high-degree vertices, removing the $(\log n)^a$ slack in (18) will likely require degree normalization or trimming, as in the normalized Kikuchi construction of Schmidhuber & Hastings [17].

Method	Memory	Runtime
HCE, level l	$\widetilde{O}\left(n^2 + m^{\lceil l/2 \rceil}\right)$	$\widetilde{O}\left(n^3 + m^{\lceil l/2 \rceil} + n^2 m^{\lfloor l/2 \rfloor}\right)$
HCE, even l, $m = Cn^{k/2}$	$\widetilde{O}\left(n^2 + C^{l/2} n^{kl/4}\right)$	$\widetilde{O}\left(n^3 + C^{l/2} n^{kl/4+2}\right)$
HCE, odd l, $m = Cn^{k/2}$	$\widetilde{O}\left(n^2 + C^{l/2} n^{kl/4}\right)$	$\widetilde{O}\left(n^3 + C^{l/2} n^{kl/4+1}\right)$
Kikuchi, level ℓ	$\widetilde{O}\left(n^\ell + mn^{\ell-k/2}\right)$	$\widetilde{O}\left(mn^{\ell-k/2} + n^{\ell+1}\right)$
Kikuchi, even ℓ, $m = Cn^{k/2}$	$\widetilde{O}\left((C+1)\, n^\ell\right)$	$\widetilde{O}\left(n^{\ell+1}\right)$

Table 1: Asymptotic runtime and memory for level-l HCE and level-ℓ Kikuchi spectral recovery on planted noisy k-XOR with $m = Cn^{k/2}$ clauses; k is assumed even and $\widetilde{O}(\cdot)$ suppresses logarithmic factors and constants depending on k, l and ℓ.

4.3 Complexity

A direct implementation of (9) enumerates all l-tuples of clauses and has runtime $O(m^l) = O(n^{kl/2})$ when $m = Cn^{k/2}$. In the Appendinx J.8, we outline the Meet-in-the-middle (MITM) strategy and how that sets our runtime and memory scaling, in the asymptotic limit, shown in Table 1. Naively, one can build $A^{(l)}$, by iterating over the entire k-XOR data l times, and tracking the clauses with symmetric difference of (i, j), i.e. finding (i, j) witnesses. However, this is costly in runtime. This approach would need only the $O(n^2)$ memory of $A^{(l)}$ itself, but runtime $O(m^l)$. MITM trades memory for time: it splits each l-set into halves of sizes $s = \lceil l/2 \rceil$ and $r = \lfloor l/2 \rfloor$, hashes all s-sets by their XOR boundary (memory $O(m^s)$), and, for every r-set and pair $\{i, j\}$, looks up the complementary boundary (runtime $O(n^2 m^r)$). Pairs of halves sharing clauses are removed exactly by a recurrence over lower levels, and the final $n \times n$ eigendecomposition costs $O(n^3)$. Substituting $m = Cn^{k/2}$ gives the entries of Table 1.

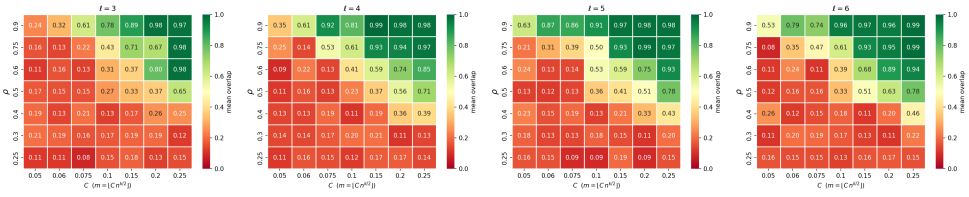
5 Experiments

Full specifications of both methods and the resource metrics are given in Appendix B. The figure of merit for recovery is the *overlap* $|\langle \hat{x}, x \rangle|/n$ between the recovered and planted bitstrings. Unless stated otherwise, results are averaged over random instances generated from identical seeds for all methods, so that HCE and Kikuchi see exactly the same instances. We compare the methods at matched level $l = \ell$. At $k = 4$ their explicit witness-count constants agree to within 2% (Section 3), making this a natural nominal statistical comparison, but it is not an equal-resource comparison.

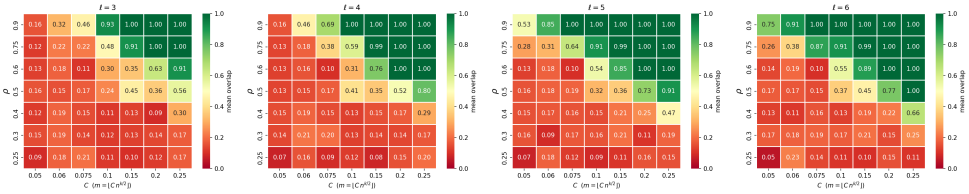
Recovery quality at $k = 4$. In Figure 2, we see that at matched level, HCE attains overlap comparable to Kikuchi, consistent with the constant-level comparison of Section 3. Since the HCE pipeline uses the post-processing described in Appendix C, whereas the recovery theorem concerns the bare top eigenvector of one layer, these heatmaps compare the practical pipelines rather than directly testing Theorem 4. We also noted that tensor power iteration did not appreciably boost the overlaps for either methods, for reasonably large number of iteration steps.

Odd and larger k. At $k = 6$, the $k = 4$ comparison repeats, with HCE matching Kikuchi recovery quality at equal level ($l = \ell$). At $k = 5$, HCE can be utilized directly for recovery, unlike symmetric Kikuchi; by the parity obstruction of Section 2, the $l = 3, 5$ matrices vanish, giving no additional recovery contributions, as seen by plateaus from $l = 2$ to $l = 3$, and $l = 4$ to $l = 5$. Every reported odd-k run thus has signal only at $l \in \{2, 4\}$, where $\mathcal{B}_l = \varnothing$ and trimmed HCE (Appendix G.7) coincides with HCE; trimming at $l \geq 6$ is left for future work. Both heatmaps are in Appendix A (Figures 5 and 6).

Scaling in the problem size. Figure 3 shows runtime and the size of the largest persistent data structure, as defined in Appendix B, for a ultra-sparse instances ($C = 0.15$). We observe that HCE is consistently faster than Kikuchi at matched level, and the memory gap grows with n: Kikuchi

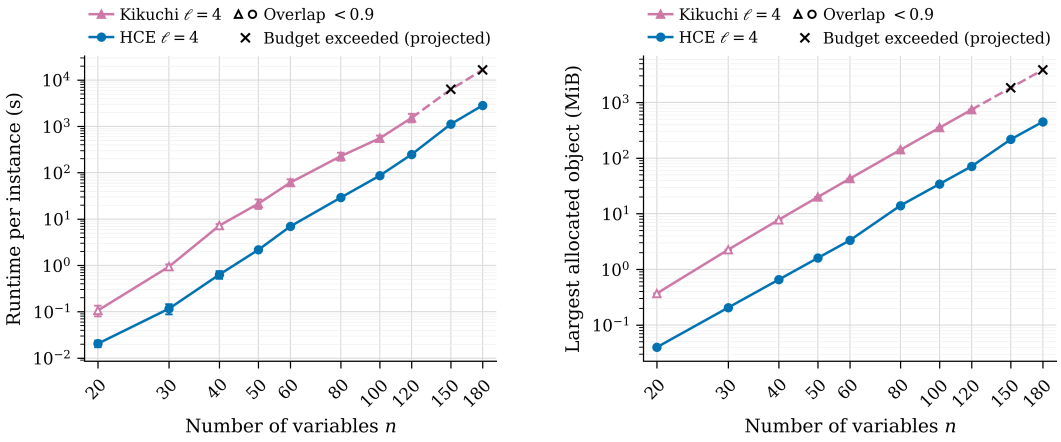


(a) Kikuchi spectral recovery.



(b) HCE recovery.

Figure 2: Recovery overlap with the planted bitstring for $n = 30$, $k = 4$, averaged over 10 random k-XOR instances (identical seeds across methods). The HCE level l and the Kikuchi level ℓ are matched, $l = \ell$. The corresponding heatmaps for $k = 5$ and $k = 6$ are reported in Appendix A, Figures 5–6.



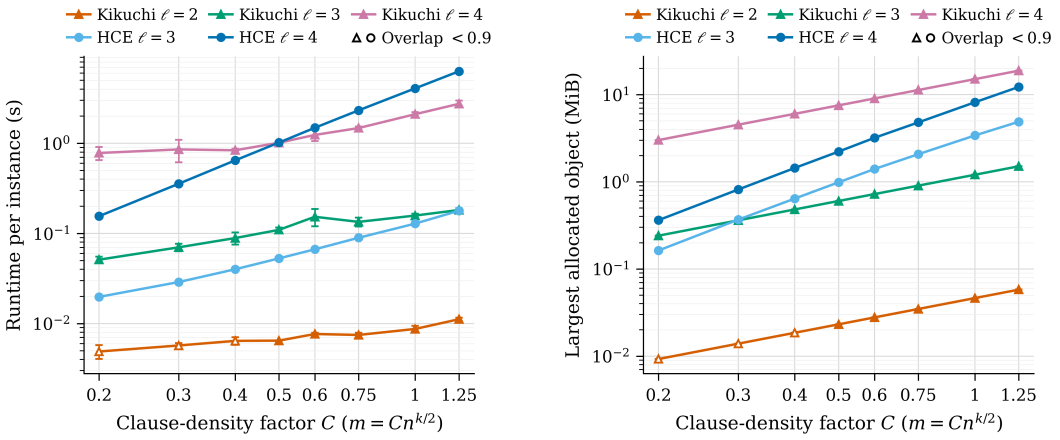
(a) Runtime versus n.

(b) Largest memory object versus n.

Figure 3: *HCE is faster than Kikuchi at the same level.* Scaling of runtime and largest-persistent-object size with the number of variables n for noisy k-XOR with $k = 4$, $C = 0.15$, $\rho = 0.75$ at fixed level $l = \ell = 4$, averaged over 3 random instances. This plot demonstrates that even though the scaling in n is algorithmically worse for HCE in the asymptotic limit, in practice, for finite sizes, it is much faster than Kikuchi, for $l = \ell$.

stores a lifted sparse operator, whereas HCE diagonalises an $n \times n$ matrix after constructing its hash table.

Scaling in the clause density. As can be noted in Figure 4, the advantage of HCE is largest in the hardest regime, i.e. when C is small: there, few clause products collapse to a pair, so the enumeration that HCE pays for is cheap, while the Kikuchi lifted operator remains $\binom{n}{\ell}$-dimensional irrespective of density. Conversely, at large C the enumeration cost of HCE grows and Kikuchi becomes preferable. This delineates three regimes: for $C \gtrsim 1$ ('sparse-but-dense') the instance is already solved at level 2; near $C \sim 1$ (sparse) the two methods are comparable at matched level; and for solvable $C \ll 1$ (ultra-sparse) HCE is empirically preferable.



(a) Runtime versus C.

(b) Memory versus C.

Figure 4: **The advantage over Kikuchi is largest in the hardest regime, where C is small.** Scaling of runtime and largest-persistent-object size with clause density C for $k = 4$, $n = 30$, $\rho = 0.75$; HCE at $l = 4$ and Kikuchi at $\ell = 2, 3, 4$, averaged over 3 random instances. Note that for $C < 0.5$ (ultra-sparse), $\ell = 2$ Kikuchi does not achieve recovery overlap > 0.9. The 'fastest' method to achieve recovery in the ultra-sparse setting is with HCE $l = 3$.

6 Discussion and Conclusion

We introduced hypergraph cycle enumeration (HCE), an elementary spectral hierarchy for planted noisy k-XOR. At level l, HCE multiplies l clauses whose supports XOR to a pair $\{i, j\}$, similar to a Gaussian elimination procedure on finite field $\mathbb{F}_2$. This reduces the problem, and hence sparse rank-one tensor recovery, to PCA on an $n \times n$ matrix at every level. In this work, we show that counting even covers gives detection without logarithmic loss. We employ a Hoeffding-decomposed trace-moment bound, to show that weak recovery is possible from eigenvector decomposition of the reduced $n \times n$ matrix. Both follow the Kikuchi threshold curve up to an l-independent constant, and recovery carries additional polylogarithmic slack.

The mechanism for HCE is transparent: every hypergraph cycle contributes an explicit parity signal, aligned with $x_i x_j$ for recovery. The threshold (8) is exactly where the noise cost ρ^{2l} balances the witness gain N_l. Unlike Kikuchi, HCE does not depend on the parity of k. For odd k, recovery only requires trimming a sublinear, support-determined set of coordinates. HCE's asymptotic runtime exponent is larger than Kikuchi's. In practice, however, it solves $k = 4$ instances with $n > 100$ at a fraction of Kikuchi's memory and runs considerably faster in the ultra-sparse regime, reflecting much smaller constant overheads. Three directions remain open. The first is removing the polylogarithmic slack in (18), which would match the sharp Kikuchi result of Schmidhuber & Hastings [17]. The second is lowering the enumeration cost below the generic meet-in-the-middle baseline by exploiting planted structure. The third is extending HCE to smoothed and semi-random CSPs, since even covers underlie both refutation and the Moore bound.

AI USE STATEMENT

In this work we used generative AI tools to assist with editing and LaTeX formatting of the manuscript, and for code-completion assistance while implementing the experimental pipeline. Additionally, we have used generative AI tools to audit our theoretical derivations, and address mathematical shortcomings. We did not use generative AI tools to generate research ideas or proofs. All mathematical statements and proofs were derived and checked by the authors, and all AI-assisted code was reviewed and tested for correctness by at least two authors. We take responsibility for the final content of this work, including all text, claims and artifacts.

ETHICS STATEMENT

This work is theoretical and algorithmic, and studies synthetic random constraint satisfaction instances. It involves no human subjects, no personal or sensitive data, and no dataset release. We note that algorithms for noisy parity problems are connected to the cryptanalysis of learning-parity-with-noise based primitives; the regimes we study are far from parameter settings of cryptographic interest, and our results are consistent with the conjectured statistical–computational gap rather than violating it.

REPRODUCIBILITY STATEMENT

The planted noisy k-XOR model is fully specified by (1) and the algorithm by the pseudocode in Figure 1; all experimental parameters (n, k, l, ℓ, C, ρ, number of instances) are stated in the captions of Figures 2, 3, 4, 5 and 6. HCE and Kikuchi are run on instances generated from identical random seeds. Complete proofs of all theoretical claims, together with the assumptions under which they hold, are given in Appendices F–J. Instance generation, solver settings, thread usage and resource-metric definitions are given in Appendix B, and the implemented decoding pipelines in Appendix C. Reference implementations of HCE, of the optimized Kikuchi baseline, and of the instance generator are provided as anonymised supplementary material.

REFERENCES

[1] Sarah R. Allen, Ryan O'Donnell, and David Witmer. How to refute a random CSP. In *Proceedings of the 56th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 689–708, 2015. doi: 10.1109/FOCS.2015.48.

[2] Amin Coja-Oghlan, Colin Cooper, and Alan Frieze. An efficient sparse regularity concept. *SIAM Journal on Discrete Mathematics*, 23(4):2000–2034, 2010. doi: 10.1137/080730160.

[3] Chandler Davis and William M. Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970. doi: 10.1137/0707001.

[4] Uriel Feige. Relations between average case complexity and approximation complexity. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 534–543, 2002. doi: 10.1145/509907.509985.

[5] Joel Friedman, Andreas Goerdt, and Michael Krivelevich. Recognizing more unsatisfiable random k-SAT instances efficiently. *SIAM Journal on Computing*, 35(2):408–430, 2005. doi: 10.1137/S009753970444096X.

[6] Andreas Goerdt and Michael Krivelevich. Efficient recognition of random unsatisfiable k-SAT instances by spectral methods. In *Proceedings of the 18th Annual Symposium on Theoretical Aspects of Computer Science (STACS)*, pp. 294–304, 2001. doi: 10.1007/3-540-44693-1_26.

[7] 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 Symposium on Theory of Computing (STOC)*, pp. 678–689, 2022. doi: 10.1145/3519935.3519955.

[8] Samuel B. Hopkins. *Statistical Inference and the Sum of Squares Method*. PhD thesis, Cornell University, 2018.

[9] Samuel B. Hopkins, Jonathan Shi, and David Steurer. Tensor principal component analysis via sum-of-square proofs. In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pp. 956–1006, 2015.

[10] Samuel B. Hopkins, Tselil Schramm, Jonathan Shi, and David Steurer. Fast spectral algorithms from sum-of-squares proofs: Tensor decomposition and planted sparse vectors. In *Proceedings of the 48th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 178–191, 2016. doi: 10.1145/2897518.2897529.

[11] Samuel B. Hopkins, Pravesh K. Kothari, Aaron Potechin, Prasad Raghavendra, Tselil Schramm, and David Steurer. The power of sum-of-squares for detecting hidden structures. In *Proceedings of the 58th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 720–731, 2017. doi: 10.1109/FOCS.2017.72.

[12] Jun-Ting Hsieh, Pravesh K. Kothari, and Sidhanth Mohanty. A simple and sharper proof of the hypergraph Moore bound. In *Proceedings of the 34th Annual ACM–SIAM Symposium on Discrete Algorithms (SODA)*, pp. 2324–2344, 2023. doi: 10.1137/1.9781611977554.ch89.

[13] 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 Symposium on Theory of Computing (STOC)*, pp. 132–145, 2017. doi: 10.1145/3055399.3055485.

[14] Dmitriy Kunisky, Alexander S. Wein, and Afonso S. Bandeira. Notes on computational hardness of hypothesis testing: Predictions using the low-degree likelihood ratio. In *Mathematical Analysis, its Applications and Computation*, volume 385 of *Springer Proceedings in Mathematics & Statistics*, pp. 1–50. Springer, 2022. doi: 10.1007/978-3-030-97127-4_1.

[15] Prasad Raghavendra, Satish Rao, and Tselil Schramm. Strongly refuting random CSPs below the spectral threshold. In *Proceedings of the 49th Annual ACM Symposium on Theory of Computing (STOC)*, pp. 121–131, 2017. doi: 10.1145/3055399.3055417.

[16] Emile Richard and Andrea Montanari. A statistical model for tensor PCA. In *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 2897–2905, 2014.

[17] Alexander Schmidhuber and Matthew B. Hastings. The Kikuchi hierarchy is sharp for kXOR. *arXiv preprint arXiv:2607.29672*, 2026. doi: 10.48550/arXiv.2607.29672.

[18] Joel A. Tropp. An introduction to matrix concentration inequalities. *Foundations and Trends in Machine Learning*, 8(1–2):1–230, 2015. doi: 10.1561/2200000048.

[19] David Wagner. A generalized birthday problem. In *Advances in Cryptology — CRYPTO 2002*, volume 2442 of *Lecture Notes in Computer Science*, pp. 288–304. Springer, 2002. doi: 10.1007/3-540-45708-9_19.

[20] Alexander S. Wein, Ahmed El Alaoui, and Cristopher Moore. The Kikuchi hierarchy and tensor PCA. In *Proceedings of the 60th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pp. 1446–1468, 2019. doi: 10.1109/FOCS.2019.000-2.

[21] Yi Yu, Tengyao Wang, and Richard J. Samworth. A useful variant of the Davis–Kahan theorem for statisticians. *Biometrika*, 102(2):315–323, 2015. doi: 10.1093/biomet/asv008.

A Additional Experimental Results

This appendix collects the recovery heatmaps for $k = 5$ and $k = 6$ referred to in Section 5. All settings (instance generation, seeds, number of repetitions) are identical to those of Figure 2.

Figure 5 shows HCE overlap for $k = 5$ at levels $l = 2, 3, 4, 5$. Since kl must be even, odd l is inadmissible for odd k and the corresponding recovery matrices vanish identically; the flat rows at $l = 3$ and $l = 5$ are exactly this parity obstruction, and they are a prediction of the theory rather than an artefact of the implementation. Kikuchi methods do not apply directly to odd k, so no matched baseline is plotted. Figure 6 repeats the matched-level HCE/Kikuchi comparison of Figure 2 at $k = 6$, where the two methods again track each other closely, as predicted by Corollary 5.

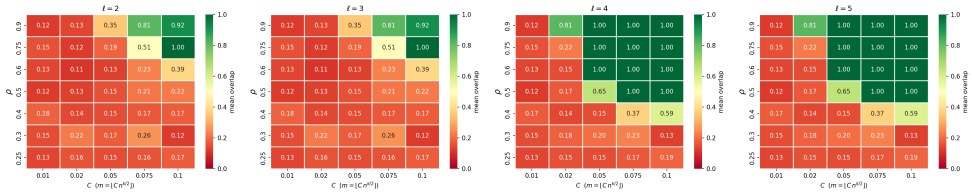
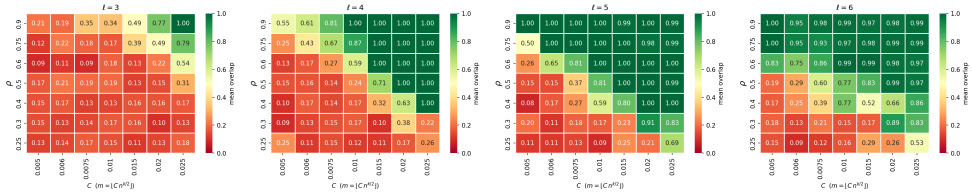
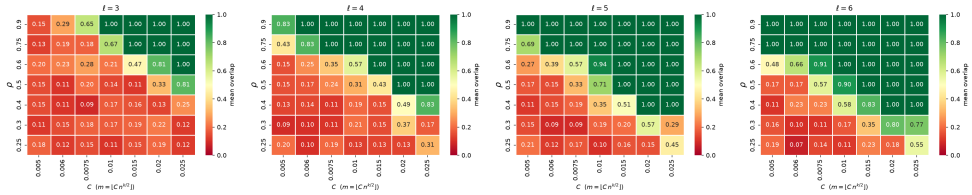


Figure 5: HCE overlap heatmap for $k = 5$, $n = 30$, levels $l = 2, 3, 4, 5$. Kikuchi does not directly apply to odd k. Odd levels carry no signal because kl must be even: this can be seen by noting that the odd levels have the same recovery as the lower even levels (e.g. heatmap of level 3 is exactly the same as that of level 2)



(a) Kikuchi spectral recovery.



(b) HCE recovery.

Figure 6: Recovery overlap for $n = 30$, $k = 6$, averaged over 10 random instances with identical seeds across methods.

B Experimental Protocol

Instances. For each (n, k, C, ρ) and seed we draw $x \in \{\pm 1\}^n$ uniformly, set $m = \lfloor Cn^{k/2} \rfloor$, draw $S_1, \ldots, S_m$ independently and uniformly from $\binom{[n]}{k}$ (the recovery model of Section 2), and set $b_a = \eta_a \prod_{v \in S_a} x_v$ with $\mathbb{P}[\eta_a = 1] = (1 + \rho)/2$. The same instances are used for every method.

Kikuchi baseline. For even k and $\ell \geq k/2$ we form $K_{T,U} = \sum_a b_a \mathbf{1}[T \triangle U = S_a]$ on $\binom{[n]}{\ell}$. No degree normalization, row pruning or diagonal shift is applied; the diagonal is zero by construction. The $m\binom{k}{k/2}\binom{n-k}{\ell-k/2}$ generated entries are stored as COO records, using 32-bit row and column indices and 8-bit values, and the matrix is applied without conversion to a dense representation. Repeated clauses may generate duplicate coordinates, so this count is a count of stored records, not

necessarily of distinct nonzero coordinates. The three largest algebraic eigenpairs are computed by ARPACK Lanczos (`scipy.sparse.linalg.eigsh`, `which=LA`, ncv $= 20$, tolerance set to machine precision). These eigenvectors are mapped to an estimate of x by the voting procedure in Appendix C. No local search is applied.

HCE. The layer matrices used by the implementation are computed exactly with hash tables of XOR boundaries. Products containing a repeated clause are removed algebraically; this is the implementation counterpart of the overlap recurrence (138). The resulting lower-level matrices are also retained for the multi-layer decoder of Appendix C. The spectral step uses dense `numpy.linalg.eigh` on an $n \times n$ matrix. No trimming is performed in the reported experiments: every odd-k run has signal only at $l \in \{2, 4\}$, where $\mathcal{B}_l = \varnothing$.

Resource metrics. *Runtime* is the wall-clock time of one solver call, from receiving the instance to returning $\hat{x}$, including matrix construction or enumeration, eigensolve and decoding, and excluding instance generation and I/O. *Memory* is the recorded payload size of the largest persistent data structure: the COO records for Kikuchi and the size-$\lceil l/2 \rceil$ hash table for HCE (all memory figures use $l = 4$). This is a proxy for the dominant memory term of Table 1, not the process peak resident set size. It excludes the interpreter, transient buffers, allocator overhead and the Lanczos workspace ($\mathrm{ncv}\binom{n}{\ell}8$ bytes for double-precision Kikuchi vectors). The proxy can therefore undercount either implementation and should be read only as a comparison of the selected persistent structures.

Hardware and software. All timings were obtained on Intel CPUs with Python 3.11, NumPy 2.4.2 and SciPy 1.17.1. The Kikuchi baseline is single-process and uses no explicit parallelism; HCE uses one probe thread except for the level-6, $n = 50$ timing quoted in Section 4.3, which used eight probe threads. These threads independently compute the matrix element calculations for each (i, j) pair.

Repetitions. Heatmap cells are means over 10 instances, and runtime curves are means over 3 instances. No uncertainty intervals are reported, so the runtime curves are descriptive. The Kikuchi memory proxy is a deterministic function of (n, k, m, ℓ), namely $9m\binom{k}{k/2}\binom{n-k}{\ell-k/2}$ bytes before coalescing duplicate COO coordinates. The HCE proxy at $l = 4$ is determined by (n, k, m) up to coincidences of XOR keys, so additional repetitions would have little effect on the plotted memory curves.

C Decoding Pipelines Used in the Experiments

Theorem 4 concerns the bare spectral estimator of Figure 1: $\hat{x} = \mathrm{sgn}(\hat{u})$, where $\hat{u}$ is the top eigenvector of the single-level matrix $A^{(l)}$. The implementation used in Section 5 wraps this estimator in data-dependent post-processing, and the Kikuchi baseline uses its own decoder. These steps are specified here to distinguish the practical pipelines from the estimators analysed in Section 4.

HCE. For a target level l, the implementation retains the layers $A^{(2)}, \ldots, A^{(l)}$ produced during enumeration. The spectral step is applied to the combined matrix

$$A_{\mathrm{tot}} = \sum_{l'=2}^{l} \hat{w}_{l'} A^{(l')}, \tag{22}$$

where $\hat{w}_{l'} \in [0, 1]$ is a data-driven estimate of $\rho^{l'}$, so that each layer is weighted by an estimate of its signal strength. No knowledge of ρ is used.

Layer weights. Alongside $A^{(l')}$ the enumeration returns the witness counts $W_{l'}(i, j)$ of (10); these are obtained by replacing every label b_a by 1 in the same hash computation. Write $\prod_{a \in I} b_a = x_i x_j \eta_I$ with $\eta_I := \prod_{a \in I} \eta_a$. Then $(A^{(l')}_{ij})^2 = \sum_{I,J} \eta_I \eta_J$, where the sum runs over pairs of level-l' witnesses of $\{i, j\}$. The $W_{l'}(i, j)$ diagonal terms $I = J$ contribute exactly 1 each. For $I \neq J$ the exact noise factor is $\mathbb{E}[\eta_I \eta_J] = \rho^{2l' - 2|I \cap J|}$. The implemented weighting rule heuristically replaces

every off-diagonal factor by its clause-disjoint value $\rho^{2l'}$, giving

$$\mathbb{E}\big[\|A^{(l')}\|_F^2 \,\big|\, S_1,\dots,S_m\big] \;\approx\; V_{l'} + \rho^{2l'}\big(Q_{l'} - V_{l'}\big),$$
$$V_{l'} := \sum_{i\neq j} W_{l'}(i,j), \qquad Q_{l'} := \sum_{i\neq j} W_{l'}(i,j)^2. \tag{23}$$

The term $V_{l'}$ is the exact conditional expectation of $\|A^{(l')}\|_F^2$ at $\rho = 0$. Solving the heuristic relation (23) for $\rho^{l'}$ gives

$$\hat{w}_{l'} \;=\; \begin{cases} \min\Big\{1,\; \big(\|A^{(l')}\|_F^2 - V_{l'}\big)^{1/2} \big/ \big(Q_{l'} - V_{l'}\big)^{1/2}\Big\}, & \text{if } \|A^{(l')}\|_F^2 > V_{l'} \text{ and } Q_{l'} > V_{l'}, \\ 0, & \text{otherwise.} \end{cases}$$

A layer whose Frobenius mass does not exceed its noise floor is therefore discarded, as is a layer with $Q_{l'} = V_{l'}$. For odd k and odd l' the layer is identically zero and receives weight zero. The cap at 1 reflects $\rho \le 1$. This weighting rule is an experimental heuristic and is not used in the recovery theorem.

Candidate bitstrings. We compute the full eigendecomposition of A_{tot} and retain the eigenvectors $\hat{u}_1, \hat{u}_2, \hat{u}_3$ of its three largest algebraic eigenvalues, together with an eigenvector $\hat{u}_{\min}$ of its smallest eigenvalue. Using $\mathrm{sgn}(0) = +1$, the candidate set consists of

1. the eight vectors $\pm\,\mathrm{sgn}(\hat{u}_1), \pm\,\mathrm{sgn}(\hat{u}_2), \pm\,\mathrm{sgn}(\hat{u}_3), \pm\,\mathrm{sgn}(\hat{u}_{\min})$;
2. for $q \in \{2,3\}$, the rotated rounding $\mathrm{sgn}(\cos\theta_t\,\hat{u}_1 + \sin\theta_t\,\hat{u}_q)$ with $\theta_t = \pi t/180$, $t = 0,\dots,179$, that violates the fewest clauses among these 180 angles.

The rotations address cases where the planted direction is split across nearly degenerate top eigenvectors.

Final bitstring. Each candidate $y \in \{\pm1\}^n$ is scored by

$$\mathrm{viol}(y) := \big|\{a \in [m] :\; b_a \neq \textstyle\prod_{v\in S_a} y_v\}\big|,$$

and the candidate with the smallest score is returned as $\hat{x}$; ties are broken by the fixed order above. Selection uses only the observed clauses, never the planted x or ρ. It is nevertheless in-sample post-processing, so the resulting curves need not equal those of the bare single-level estimator in Theorem 4.

Kikuchi. A single level ℓ is used, with no multi-level combination, candidate selection or local search.

Eigenvector combination. Let $v^{(1)}, v^{(2)}, v^{(3)} \in \mathbb{R}^{\binom{[n]}{\ell}}$ be unit eigenvectors of the three largest algebraic eigenvalues of K, with the signs returned by ARPACK. The implemented pipeline combines them with equal weights,

$$v = \frac{v^{(1)} + v^{(2)} + v^{(3)}}{\|v^{(1)} + v^{(2)} + v^{(3)}\|_2},$$

using $v := v^{(1)}$ if the sum vanishes. This implementation rule is not invariant under sign changes or basis rotations within a nearly degenerate eigenspace; the reported figures should therefore be understood as results for this specified numerical pipeline, not for a basis-invariant decoder.

Voting matrix. In the planted model the informative Kikuchi eigenspace is correlated with $(\chi_T(x))_{T\in\binom{[n]}{\ell}}$, where $\chi_T(x) := \prod_{t\in T} x_t$. For an $(\ell-1)$-set c and $u, u' \notin c$, the product $v_{c\cup\{u\}} v_{c\cup\{u'\}}$ is therefore a noisy estimate of $x_u x_{u'}$. Define $W \in \mathbb{R}^{n\times\binom{n}{\ell-1}}$ by $W_{u,c} := v_{c\cup\{u\}}$ if $u \notin c$ and $W_{u,c} := 0$ otherwise, and set

$$V := WW^\top - \mathrm{diag}(WW^\top), \qquad V_{uu'} = \sum_{\substack{c\in\binom{[n]}{\ell-1} \\ u,u'\notin c}} v_{c\cup\{u\}} v_{c\cup\{u'\}} \quad (u \neq u'), \qquad V_{uu} = 0.$$

The diagonal carries no pairwise sign information and is removed. The factorization avoids looping over pairs of Kikuchi coordinates and uses $O(n\binom{n}{\ell-1})$ memory for W.

Final bitstring. We compute the top eigenvector z of V by dense `scipy.linalg.eigh` and return $\hat{x} = \mathrm{sgn}(z)$, with $\mathrm{sgn}(0) = +1$.

D Operator Comparison with Kikuchi and Square Constructions

The relationship between HCE and earlier lifted spectral constructions can be made exact. Let $\mathcal{H}_{\mathrm{cube}} = \ell^2(2^{[n]})$ have standard basis $(e_U)_{U\subseteq[n]}$. For $S \subseteq [n]$ define the Boolean-cube translation $\mathsf{T}_S e_U = e_{U\triangle S}$, put

$$Y_a := b_a \mathsf{T}_{S_a}, \qquad \mathsf{B} := \sum_{a=1}^{m} Y_a, \tag{24}$$

and let P_r be the orthogonal projection onto the span of basis vectors indexed by r-subsets.

Proposition 7 (Full-cube polynomial and slice compressions)**.** *Define the elementary clause operator*

$$\mathsf{E}_l := \sum_{I\in\binom{[m]}{l}} \prod_{a\in I} Y_a. \tag{25}$$

Then, as an identity of formal power series,

$$\sum_{l=0}^{m} z^l \mathsf{E}_l = \prod_{a=1}^{m} (I + zY_a) = (1-z^2)^{m/2} \exp\bigl(\operatorname{artanh}(z)\mathsf{B}\bigr), \tag{26}$$

and consequently

$$\mathsf{E}_l = K_l\left(\frac{mI - \mathsf{B}}{2}, m\right), \tag{27}$$

where $K_l(\cdot, m)$ is the unnormalised Kravchuk polynomial of (104), evaluated by polynomial functional calculus. Moreover, in the simple-hypergraph detection model,

$$\sum_{F\in\mathcal{C}_l} X_F = \langle e_\varnothing, \mathsf{E}_l e_\varnothing\rangle, \tag{28}$$

$$A^{(l)} = \operatorname{offdiag}(\mathsf{P}_1 \mathsf{E}_l \mathsf{P}_1), \tag{29}$$

$$(\mathsf{P}_\ell \mathsf{B} \mathsf{P}_\ell)_{T,U} = \sum_{a=1}^{m} b_a \mathbf{1}\{T\triangle U = S_a\}, \qquad |T| = |U| = \ell. \tag{30}$$

In particular,

$$\mathsf{E}_2 = \frac{\mathsf{B}^2 - mI}{2}. \tag{31}$$

Proof. Boolean-cube translations commute, and $Y_a^2 = b_a^2 \mathsf{T}_{S_a}^2 = I$. Therefore

$$I + zY_a = (1-z^2)^{1/2} \exp\bigl(\operatorname{artanh}(z) Y_a\bigr).$$

Multiplying over a proves (26). On a simultaneous eigenvector of the commuting involutions $(Y_a)_{a\le m}$, suppose exactly r eigenvalues equal -1. Then B has eigenvalue $m - 2r$, while E_l has eigenvalue $\sum_j (-1)^j \binom{r}{j}\binom{m-r}{l-j} = K_l(r, m)$. This proves (27).

For every clause-index set I,

$$\prod_{a\in I} Y_a \, e_U = \left(\prod_{a\in I} b_a\right) e_{U\triangle\triangle_{a\in I} S_a}.$$

Taking respectively the $(\varnothing,\varnothing)$ entry, the $(\{i\},\{j\})$ entry for $i \neq j$, and the (T,U) entry of B proves (28)–(30). Finally, expanding B^2 gives $mI + 2\mathsf{E}_2$, proving (31). □

Thus HCE is a matrix-valued degree-l low-degree polynomial, but it is not in general a power of one fixed Kikuchi slice: in $\mathsf{P}_1 K_l((mI - \mathsf{B})/2, m)\mathsf{P}_1$ the intermediate Boolean-cube boundaries are unrestricted, whereas powers of $\mathsf{P}_\ell \mathsf{B} \mathsf{P}_\ell$ force every intermediate state to remain on the ℓ-slice. For odd k the balanced slice $\mathsf{P}_\ell \mathsf{B} \mathsf{P}_\ell$ vanishes because the symmetric difference of two ℓ-sets has even cardinality, while even-level HCE can be nonzero. Equation (31) recovers the usual square construction at level two.

E NOTATION

Symbol	Type	Meaning / property
n, m, k	positive integers	problem size, number of sampled clauses, clause size
l	positive integer	**HCE level**: even-cover size / witness size; admissible when kl is even
ℓ	positive integer	**Kikuchi level**: degree of the lifted Kikuchi matrix; used only in comparisons
π_k	3 (k even), 4 (k odd)	minimal admissible cover level
$[n]$	finite set $\{1, \dots, n\}$	vertex set
$\binom{[n]}{k}$	finite set of k-subsets of $[n]$	possible supports
M	$= \binom{n}{k}$	number of possible clauses
C	positive real	clause density, $m = Cn^{k/2}$
ρ	real number in $[0, 1]$	noise mean (planted advantage)
x	element of $\{\pm 1\}^n$	planted sign vector
$\mathcal{K}$	$= \binom{[n]}{k}$	complete k-uniform hypergraph
$\mathcal{H}$	subset of $\mathcal{K}$	instance clause set, size m
$\mathcal{S}$	subset of $\mathcal{K}$	any subset of $\mathcal{H}$
S_a	element of $\mathcal{K}$	clause support indexed by a
η_a	r.v. in $\{\pm 1\}$	independent noise, $\mathbb{E}[\eta_a] = \rho$
b_a	r.v. in $\{\pm 1\}$	observation $b_a = \eta_a \prod_{v \in S_a} x_v$
$\triangle$	operation	symmetric difference
$\triangle(\mathcal{S})$, ∂I	subset of $[n]$	$\triangle_{S \in \mathcal{S}} S$; $\triangle_{a \in I} S_a$ for a set I of clause indices
$X_{\mathcal{S}}$	r.v. in $\{\pm 1\}$	$\prod_{a \in \mathcal{S}} b_a$
F	subset of $\mathcal{K}$	even cover: $\lvert F \rvert = l$, $\triangle(F) = \emptyset$
$\mathcal{C}_l$, Z_l	set / integer	even l-covers in $\mathcal{H}$ and their number
$\mathcal{C}_l^{\mathcal{K}}$, $Z_l^{\mathcal{K}}$	set / integer	even l-covers in $\mathcal{K}$ and their number
$E_{n,k}(l)$	integer	$= Z_l^{\mathcal{K}}$, the exact Kravchuk count
$\alpha_{k,l}$	non-negative real	$Z_l^{\mathcal{K}} = (\alpha_{k,l} + o_n(1)) n^{kl/2}$; zero for inadmissible l; $\alpha_{k,0} = 1$
μ_q	non-negative real	$\alpha_{k,q} (C\, k!)^q$, so $\mathbb{E}_{\mathcal{H}}[Z_q] = \mu_q + o_n(1)$; $\mu_0 = 1$
n_s, $n_s^{\mathcal{K}}$	integers	ordered pairs of even l-covers of $\mathcal{H}$ (resp. $\mathcal{K}$) meeting in exactly s edges
$\mathcal{W}_l(i, j)$, $W_l(i, j)$	set / integer	(i, j)-witnesses in $\mathcal{H}$ and their number
N_l	non-negative real	$\mathbb{E}\, W_l(i, j)$
d_l, θ_l	positive reals	$d_l = nN_l$; signal scale $\theta_l = \rho^l d_l$
$\widehat{A}^{(l)}$, $\widetilde{R}^{(l)}$	$n \times n$ matrices	$D_x A^{(l)} D_x$ and its exact centering $\widehat{A}^{(l)} - \mathbb{E}\widehat{A}^{(l)}$
q, q_n	even integers	trace-moment order; $q_n = 2\lceil \log n \rceil$
$\mathcal{M}_q$, $\mathcal{N}_q$	functionals	$\mathcal{M}_q(M) = (\mathbb{E}\,\mathrm{tr}\,\lvert M \rvert^q)^{1/q}$, $\mathcal{N}_q(M) = (\mathbb{E}\lVert M \rVert_{\mathrm{op}}^q)^{1/q}$
a, a_0	positive integers	$a_0 = (k+1)l + 2$ (pattern-count exponent); $a = a_0 + 1$, the polylogarithmic exponent
$p_r(w)$	probability	$\mathbb{P}[\partial\{1, \dots, r\} = w]$, the prescribed-boundary probability
$R_{\mathrm{bc}}^{(l)}$, $R_{\mathrm{bd}}^{(l)}$	$n \times n$ matrices	boundary-connected and boundary-disconnected sectors of $\widetilde{R}^{(l)}$
$\mathcal{B}_l$, $\mathcal{G}_l$	subsets of $[n]$	singleton-boundary trimming set and its complement, (91)
Π_l, $\bar{A}^{(l)}$	$n \times n$ matrices	projection onto $\mathcal{G}_l$ and trimmed recovery matrix, (92)
$N_{l,\mathrm{bc}}$, $d_{l,\mathrm{bc}}$, $\theta_{l,\mathrm{bc}}$	non-negative reals	boundary-connected witness count, degree scale, and signal scale
$\kappa_{l,k}^{\mathrm{bc}}$	positive real	boundary-connected restriction of the full i.i.d. witness coefficient
$T_l, Q_l(i, j)$	integers	distinct-support universe count and ordered i.i.d. support-tuple count
κ_l, $\kappa_{l,k}$	positive reals	normalized ordered-tuple and sampled i.i.d. witness constants; $\kappa_{l,k} = (k!)^l \kappa_l$
γ_k, a_k, L_k	positive reals	$\gamma_k = k^{-k/2} e^{k/2-1}$; $a_k = 1/(\gamma_k k!)$; $L_k = e^{-(k-1)/2}/\sqrt{\pi}$

Two conventions are worth restating. First, the HCE level is always l and the Kikuchi level is always ℓ; the two coincide only in the matched-level experiments, where we write $l = \ell$ explicitly. Second, C denotes the clause density throughout, so the even-cover count present in the instance is written Z_l (and not C_l), the witness constants are $\kappa_l, \kappa_{l,k}$ (and not $C_l, C_{l,k}$), and the generic constants appearing in the moment bounds are written B. The trace-moment order is q, which is unrelated to the inclusion probability $p = m/\binom{n}{k}$ of Appendix J.

F Proof of the Detection Theorem

Throughout this appendix $\mathcal{H}$ is a uniformly random m-edge subhypergraph of $\mathcal{K}$ with $m = Cn^{k/2}$, and $X(l)$ is the discriminator (5), with $X_F = \prod_{S_a \in F} b_a$. We keep the two sources of randomness apart by a subscript: $\mathbb{P}_{\mathcal{H}}, \mathbb{E}_{\mathcal{H}}, \mathrm{Var}_{\mathcal{H}}$ refer to the draw of the clause supports, while $\mathbb{P}_\rho, \mathbb{E}_\rho, \mathrm{Var}_\rho$ and Cov_ρ refer to the signs at planted advantage ρ, the null model being $\rho = 0$.

We only consider *admissible* levels, i.e. $l \geq 3$ with kl even, for which $\mathcal{K}$ possesses proper even l-covers; the minimal admissible level is π_k of Section 2. By Theorem 40, $Z_l^{\mathcal{K}} = \alpha_{k,l} n^{kl/2} + o_n(n^{kl/2})$ for an n-independent $\alpha_{k,l}$, with $\alpha_{k,l} = 0$ when l is inadmissible and $\alpha_{k,0} = 1$ (the empty family is the unique even 0-cover). For $k > 2$ and large admissible l,

$$\alpha_{k,l} = \big(L_k + o_l(1)\big)\, l^{-1/2} \big(a_k\, l^{k/2-1}\big)^l, \qquad a_k := \frac{k^{k/2} e^{1-k/2}}{k!} = \frac{1}{\gamma_k\, k!}, \qquad L_k := \frac{e^{-(k-1)/2}}{\sqrt{\pi}}. \tag{32}$$

Finally set

$$n_s := \big|\{(F, F') \in \mathcal{C}_l \times \mathcal{C}_l : |F \cap F'| = s\}\big|, \qquad n_s^{\mathcal{K}} := \big|\{(F, F') \in \mathcal{C}_l^{\mathcal{K}} \times \mathcal{C}_l^{\mathcal{K}} : |F \cap F'| = s\}\big|, \tag{33}$$

for the instance and for the complete hypergraph respectively.

Lemma 8 (Pairs of covers meeting in s edges)**.** *Fix k, l and $0 \leq s \leq l$, and set $r := l - s$. Then*

$$n_s^{\mathcal{K}} = \big(\alpha_{k,s}\, \alpha_{k,r}^2 + o_n(1)\big)\, n^{kl - ks/2}. \tag{34}$$

If $\mathcal{K}$ has no proper even cover of size s or of size r, the corresponding coefficient $\alpha_{k,s}$ or $\alpha_{k,r}$ vanishes and the statement is to be read as $n_s^{\mathcal{K}} = o_n(n^{kl-ks/2})$.

Proof. Define $J := F \cap F'$, $A := F \setminus J$ and $B := F' \setminus J$. Every triple (J, A, B) corresponds to a unique pair (F, F'), so we count triples with $|J| = s$ and $|A| = |B| = r$. Since F and F' are even covers, $\triangle(J) = \triangle(A) = \triangle(B)$; call this common vertex set U and put $u := |U|$.

A family of q hyperedges with symmetric difference U uses at most $u + (kq - u)/2$ vertices: the vertices of U consume at least one incidence each, while every other used vertex has positive even degree and hence consumes at least two. Counting the vertices of U only once across J, A, B, the whole triple uses at most

$$u + \frac{ks - u}{2} + 2 \cdot \frac{kr - u}{2} = kl - \frac{ks}{2} - \frac{u}{2} \tag{35}$$

distinct vertices. Therefore $U \neq \emptyset$ contributes at a strictly lower order in n. If $U = \emptyset$, then J, A, B are themselves even covers, of sizes s, r, r, and the leading contribution comes from triples with pairwise disjoint vertex sets, since sharing a vertex again reduces the number of free labels. The number of such triples is

$$Z_s^{\mathcal{K}} \big(Z_r^{\mathcal{K}}\big)^2 = \big(\alpha_{k,s} \alpha_{k,r}^2 + o_n(1)\big) n^{ks/2 + kr}, \tag{36}$$

and removing the vertex-intersecting triples does not affect the leading term. Since $ks/2 + kr = kl - ks/2$, the claim follows. □

Proposition 9. *Under the detection sampling model, $\mathbb{E}_{\mathcal{H}}[Z_l] = Z_l^{\mathcal{K}} \dfrac{(m)_l}{(M)_l}$, where $(y)_j$ denotes the falling factorial and $M = \binom{n}{k}$. Consequently, for $m = Cn^{k/2}$ and fixed k, l, $\mathbb{E}_{\mathcal{H}}[Z_l] = \mu_l + o_n(1)$ with $\mu_l = \alpha_{k,l}(C\, k!)^l$.*

Proof. The first identity is linearity of expectation: a fixed l-set of edges lies in $\mathcal{H}$ with probability $(m)_l/(M)_l$. The second follows from $Z_l^{\mathcal{K}} = (\alpha_{k,l} + o_n(1))n^{kl/2}$, $(m)_l/(M)_l = (1 + o_n(1))(m/M)^l$ and $M = (1 + O(k^2/n))n^k/k!$. □

Lemma 10 (Normalised second moment)**.** *Fix k, l with $\alpha_{k,l} > 0$ and let $m = Cn^{k/2}$ for a constant $C > 0$. With $\mu_q := \alpha_{k,q}(C\,k!)^q$ and $\mu_0 = 1$,*

$$\frac{\mathrm{Var}_{\mathcal{H}}[Z_l]}{\mathbb{E}_{\mathcal{H}}[Z_l]^2} = \sum_{s=1}^{l} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1). \tag{37}$$

For large admissible l the right-hand side is $\Theta_{k,C}\big(l^{-(k-2)\pi_k}\big) + o_n(1)$.

Proof. Writing $Z_l = \sum_{F \in \mathcal{C}_l^{\mathcal{K}}} \mathbf{1}(F \subseteq \mathcal{H})$ gives $Z_l^2 = \sum_{F,F'} \mathbf{1}(F \cup F' \subseteq \mathcal{H})$. For two even covers with $|F \cap F'| = s$ we have $|F \cup F'| = 2l - s$, hence $\mathbb{E}_{\mathcal{H}}[\mathbf{1}(F \cup F' \subseteq \mathcal{H})] = (m)_{2l-s}/(M)_{2l-s}$ and

$$\frac{\mathbb{E}_{\mathcal{H}}[Z_l^2]}{\mathbb{E}_{\mathcal{H}}[Z_l]^2} = \sum_{s=0}^{l} \frac{n_s^{\mathcal{K}}}{\mathbb{E}_{\mathcal{H}}[Z_l]^2} \frac{(m)_{2l-s}}{(M)_{2l-s}}. \tag{38}$$

By Lemma 8, for every fixed s,

$$\frac{n_s^{\mathcal{K}}}{\big(Z_l^{\mathcal{K}}\big)^2} = \left(\frac{\alpha_{k,s}\alpha_{k,l-s}^2}{\alpha_{k,l}^2} + o_n(1)\right) n^{-ks/2}, \tag{39}$$

while, for fixed l and s,

$$\frac{(m)_{2l-s}}{(M)_{2l-s}} \left(\frac{(M)_l}{(m)_l}\right)^2 = \big(1 + o_n(1)\big) \frac{n^{ks/2}}{(C\,k!)^s}. \tag{40}$$

Multiplying the two displays and using the definition of μ_q gives $\mu_s\mu_{l-s}^2/\mu_l^2 + o_n(1)$ for the s-th summand; the $s = 0$ term equals $1 + o_n(1)$. As the sum has only $l + 1$ terms, subtracting one from the normalised second moment proves (37).

For the large-l estimate the summation must be uniform in s. Put

$$g_{k,t} := \frac{(kt-1)!!}{t!\,(k!)^t} \qquad (kt \text{ even}).$$

The pairing argument in the proof of Theorem 40 gives constants $0 < c_k < C_k < \infty$ such that $c_k g_{k,t} \le \alpha_{k,t} \le C_k g_{k,t}$ for every sufficiently large admissible t; enlarging the constants covers the finitely many smaller admissible levels. We use the two-sided factorial bounds

$$\sqrt{2\pi}\, t^{t+1/2} e^{-t} e^{1/(12t+1)} \le t! \le \sqrt{2\pi}\, t^{t+1/2} e^{-t} e^{1/(12t)}, \qquad t \ge 1. \tag{41}$$

Applied to the factorials in $g_{k,t}$, these bounds give constants $\eta_k, c_{k,C}, B_{k,C} > 0$ such that, uniformly over admissible indices,

$$\frac{\mu_s\mu_{l-s}^2}{\mu_l^2} \le \left[B_{k,C}\left(\frac{s}{l}\right)^{k-2}\right]^s, \qquad 1 \le s \le \eta_k l,$$

and the same bound with s replaced by $r = l - s$ holds for $1 \le r \le \eta_k l$. On the interior $\eta_k l \le s \le (1 - \eta_k)l$, the same substitution gives $e^{-c_{k,C}\, l \log l}$. For example, the power-carrying part of the logarithm at the lower endpoint is

$$\frac{k-2}{2}\big[s\log s + 2(l-s)\log(l-s) - 2l\log l\big] \le (k-2)s\log(s/l) + O_k(s),$$

uniformly after choosing η_k small; the square-root factors in (41) are absorbed into $B_{k,C}^s$. The first nonzero lower-endpoint index is $s = \pi_k$, and its summand is $\Theta_{k,C}(l^{-(k-2)\pi_k})$. The opposite endpoint and the interior are smaller, while $s = l$ equals $1/\mu_l$ and is superpolynomially small at fixed C. Summing the three ranges proves the stated $\Theta_{k,C}(l^{-(k-2)\pi_k})$ estimate. □

Corollary 11 (Concentration of the cover count)**.** *Fix an admissible l and $t > 0$. Then*

$$\mathbb{P}_{\mathcal{H}}\left(\left|\frac{Z_l}{\mathbb{E}_{\mathcal{H}}[Z_l]} - 1\right| > t\right) \le \frac{1}{t^2}\left(\sum_{s=1}^{l} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1)\right) = O_{k,C}\big(t^{-2} l^{-(k-2)\pi_k}\big) + o_n(1) \quad (42)$$

for large admissible l.

Proof. Immediate from Lemma 10 and Chebyshev's inequality. □

Lemma 12 (Overlap sums at the detection scale)**.** *Fix $k > 2$, $\rho \in (0, 1]$ and $\delta \in (0, 1)$. If*

$$C \ge \left(\frac{32}{\delta}\right)^{1/l} \frac{\rho^{-2}}{k!\, \alpha_{k,l}^{1/l}},$$

then, for large admissible l,

$$\sum_{s=1}^{l} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} = \frac{1}{\mu_l} + O_k\Big(\rho^{2\pi_k} l^{-(k-2)\pi_k/2}\Big), \quad (43)$$

$$\sum_{s=1}^{l-1} \rho^{-2s} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} = O_k\Big(l^{-(k-2)\pi_k/2}\Big). \quad (44)$$

Proof. Put $p = (k-2)/2$. By (32), after enlarging constants to cover the finitely many small admissible indices,

$$c_k\, t^{-1/2}(a_k t^p)^t \le \alpha_{k,t} \le C_k\, t^{-1/2}(a_k t^p)^t$$

for every admissible t. The sample lower bound implies $\rho^{-2}(Ck!)^{-1} \le B_k \alpha_{k,l}^{1/l}$. Therefore the sth summand of (44), denoted $T_{l,s}$, satisfies

$$T_{l,s} \le B_k^s \alpha_{k,s} \alpha_{k,l-s}^2 \alpha_{k,l}^{-2+s/l}.$$

Applying (41) to this display gives constants $\eta_k, c_k, B_k > 0$ such that

$$\begin{aligned} T_{l,s} &\le [B_k (s/l)^p]^s, && 1 \le s \le \eta_k l, \\ T_{l,l-r} &\le [B_k (r/l)^{2p}]^r, && 1 \le r \le \eta_k l, \\ T_{l,s} &\le e^{-c_k l}, && \eta_k l \le s \le (1-\eta_k) l. \end{aligned}$$

Indeed, after writing $x = s/l$, the exponential part in the interior is

$$pl\,[x \log x + 2(1-x)\log(1-x)],$$

which is strictly negative on every compact subinterval of $(0, 1)$. At the lower endpoint, the inequality $\log(1 - s/l) \le -s/l$ yields the first bound uniformly for $s \le \eta_k l$; at the upper endpoint, setting $r = l - s$ gives the second. The square-root remainders in (41) are absorbed into B_k^s or B_k^r.

Only admissible indices contribute, and the first positive one is π_k. Summing the lower endpoint gives $O_k(l^{-p\pi_k})$, the upper endpoint gives $O_k(l^{-2p\pi_k})$, and the interior is exponentially small. This proves (44). The corresponding unweighted summand is $\rho^{2s} T_{l,s}$; hence its lower endpoint acquires $\rho^{2\pi_k}$, while the remaining ranges are no larger. Finally the omitted term $s = l$ is exactly $1/\mu_l$, proving (43). □

We now address the second heuristic assumption.

Proposition 13. *Let $F, F' \in \mathcal{C}_l$ with $|F| = |F'| = l$ and $|F \cap F'| = s$. Then, conditionally on $F, F' \subseteq \mathcal{H}$, $\mathrm{Cov}_\rho(X_F, X_{F'}) = \rho^{2l-2s} - \rho^{2l}$.*

Proof. $X_F X_{F'} = X_{F \triangle F'}$, and since even covers form a closed cycle space, $F \triangle F'$ is an even cover of size $2l - 2s$. Hence $\mathbb{E}_\rho[X_F X_{F'}] = \rho^{2l-2s}$, while $\mathbb{E}_\rho[X_F]\mathbb{E}_\rho[X_{F'}] = \rho^{2l}$. □

Proof of Theorem 3. Let $m = Cn^{k/2}$ and declare the model planted when $Z_l > 0$ and $X(l) > \rho^l/2$. We first control the hypergraph randomness. Let G be the event $\{Z_l \geq \mathbb{E}_{\mathcal{H}}[Z_l]/2\}$. Corollary 11 with $t = 1/2$ bounds

$$\mathbb{P}_{\mathcal{H}}(G^c) \leq 4 \sum_{s=1}^{l} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1); \tag{45}$$

we bound this after imposing the sample condition below, l being fixed from this point onwards.

For $\mathcal{H} \in G$ under the null model ($\rho = 0$), $X(l)$ has mean zero and distinct covers have zero covariance, so $\mathrm{Var}_0(X(l)) = 1/Z_l$. Chebyshev's inequality, together with the contribution of G^c and $\mathbb{E}_{\mathcal{H}}[Z_l] = \mu_l + o_n(1)$, gives

$$\mathbb{P}_0(\text{false positive}) \leq \mathbb{P}_{\mathcal{H}}(G^c) + \frac{8}{\rho^{2l}\mu_l} + o_n(1). \tag{46}$$

Under the planted model the conditional mean is ρ^l, and Proposition 13 gives

$$\frac{\mathrm{Var}_\rho(X(l))}{\rho^{2l}} = \frac{\rho^{-2l} - 1}{Z_l} + \frac{1}{Z_l^2} \sum_{s=1}^{l-1} \big(\rho^{-2s} - 1\big) n_s, \qquad \frac{\mathbb{E}_{\mathcal{H}}[n_s]}{\mu_l^2} = \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1), \tag{47}$$

the second identity being Lemma 8 together with the sampling step of Proposition 9. Chebyshev's inequality and averaging over $\mathcal{H}$ then yield

$$\mathbb{P}_\rho(\text{false negative}) \leq \mathbb{P}_{\mathcal{H}}(G^c) + \frac{8}{\rho^{2l}\mu_l} + 16 \sum_{s=1}^{l-1} \rho^{-2s} \frac{\mu_s \mu_{l-s}^2}{\mu_l^2} + o_n(1). \tag{48}$$

It remains to choose C and l. Requiring $8/(\rho^{2l}\mu_l) \leq \delta/4$ is exactly

$$C \geq \left(\frac{32}{\delta}\right)^{1/l} \frac{\rho^{-2}}{k!\, \alpha_{k,l}^{1/l}}, \tag{49}$$

which is (15) after substituting $C = m/n^{k/2}$, and which for large l forces $C = \Omega(\rho^{-2} l^{1-k/2})$. Lemma 12 gives the required uniform bounds on the two sums above. The sample condition gives $1/\mu_l \leq \delta\rho^{2l}/32$, hence $4/\mu_l \leq \delta/8$; since $\rho \leq 1$, the second remainder also controls the first. Choosing the implicit constant in $l = \Omega\big(\delta^{-2/((k-2)\pi_k)}\big)$ large enough makes four times the first remainder at most $\delta/8$ and sixteen times the second sum at most $\delta/4$. Then $\mathbb{P}_{\mathcal{H}}(G^c) \leq \delta/4 + o_n(1)$ by (45), and both error probabilities are below δ for all sufficiently large n.

Finally, inserting the large-l form (32) of $\alpha_{k,l}$ into (15) gives

$$m \geq \theta_{l,k,\delta}\, \rho^{-2} n \left(\frac{n}{l}\right)^{k/2-1}, \qquad \theta_{l,k,\delta} = \big(1 + o_l(1)\big) \left(\frac{32\sqrt{l}}{\delta L_k}\right)^{1/l} \gamma_k, \tag{50}$$

which is (8) up to a factor tending to one as l grows.

It remains to verify the adaptive rule (16). Under the null model, conditionally on every hypergraph with $Z_l > 0$, $\mathbb{E}_0 X(l) = 0$ and $\mathrm{Var}_0(X(l)) = 1/Z_l$. Hence Chebyshev's inequality gives

$$\mathbb{P}_0\left[X(l) > \sqrt{\frac{2}{\delta Z_l}} \,\middle|\, \mathcal{H}\right] \leq \frac{\delta}{2}.$$

Under the planted model, on the event $G = \{Z_l \geq \mathbb{E}_{\mathcal{H}} Z_l/2\}$, the sample condition implies

$$\rho^{2l} Z_l \geq \frac{1}{2}\rho^{2l}\mu_l + o_n(1) \geq \frac{16}{\delta} + o_n(1).$$

In particular, for all sufficiently large n the left-hand side exceeds $8/\delta$, so the adaptive threshold is at most $\rho^l/2$. Its false-negative event on G is contained in the one already bounded above, while G^c has already been included in that bound. Thus both adaptive errors are at most δ. □

G Proof of the Recovery Theorem

This appendix proves Theorem 4 and Corollary 5, together with the odd-k trimming result Theorem 37 and Corollary 38. Ground-variable labels (elements of $[n]$) and clause-index labels (elements of $[m]$) are kept strictly distinct throughout. The even-k argument consists of the boundary, component, Hoeffding, diagram, exposure and connected-moment subsections G.1–G.6; the odd-k contraction is isolated in Subsection G.7. Throughout, B denotes a positive constant depending only on k and l whose value may change from line to line, $a_0 = (k+1)l+2$, $a = a_0+1 = (k+1)l+3$, and for a random matrix M and an even integer $q \geq 2$ we write

$$\mathcal{M}_q(M) := \big(\mathbb{E}\,\mathrm{tr}\,|M|^q\big)^{1/q}, \qquad \mathcal{N}_q(M) := \big(\mathbb{E}\|M\|_{\mathrm{op}}^q\big)^{1/q}, \qquad \mathcal{N}_q(M) \leq \mathcal{M}_q(M) \leq n^{1/q}\mathcal{N}_q(M). \tag{51}$$

G.1 Setup, boundary probabilities and the exact entry variance

We fix the planted vector $x \in \{\pm 1\}^n$ and condition on it throughout. Recovery uses the i.i.d. sampling convention of Section 2:

1. the supports $S_1, \dots, S_m$ are drawn i.i.d. uniformly from $\binom{[n]}{k}$ — repeated supports are retained and carry independent labels, so that clauses are indexed by $[m]$ rather than by a hypergraph;
2. the noise $\eta_1, \dots, \eta_m \in \{\pm 1\}$ is independent of the supports and of itself, with $\mathbb{E}[\eta_a] = \rho \in [0,1]$ and $\eta_a^2 = 1$.

Each observation is $b_a = \eta_a \prod_{v \in S_a} x_v$, and $A^{(l)}$ is as in (9). The two sources of randomness are tracked separately: the witness *structure* (the event $\partial I = \{i,j\}$) depends only on the supports, whereas the witness *value* $\prod_{a\in I} b_a$ involves both supports and noise. This separation drives the entire argument.

For $I \subseteq [m]$ with $\partial I = \{i,j\}$ we have $\prod_{a\in I} b_a = x_i x_j \prod_{a\in I}\eta_a$, which gives the conditional mean (10); averaging over the supports and using permutation symmetry gives the rank-one decomposition (12) with the normalisation (11). It is convenient to write, for $w \in \mathbb{F}_2^n$ and $0 \leq r \leq l$,

$$p_r(w) := \mathbb{P}\big[\,\partial\{1,\dots,r\} = w\,\big], \qquad w_{ij} := \mathbf{1}_{\{i\}} + \mathbf{1}_{\{j\}}, \tag{52}$$

so that $N_l = \binom{m}{l} p_l(w_{ij})$, $d_l = nN_l$ and $\theta_l = \rho^l d_l$. Fixing $i \neq j$ and conditioning on the supports, the same computation as in Proposition 13 gives

$$\mathrm{Var}\big(A_{ij}^{(l)} \mid S_1, \dots, S_m\big) = (1-\rho^{2l})\, W_l(i,j) + \sum_{\substack{I,I' \,:\, \partial I = \partial I' = \{i,j\} \\ I \neq I'}} \big(\rho^{|I \triangle I'|} - \rho^{2l}\big), \tag{53}$$

and averaging over the supports yields the exact overlap expansion in Proposition 16. For $l \geq 4$, its endpoint scales include both $N_l = d_l/n$ and $n^{-1}\rho^{2l-4} d_l^{2-2/l}$, as summarised by Theorem 18. Such entrywise information is *not* sufficient for an operator-norm bound because overlapping witnesses can localise on a few rows. The remainder of this appendix therefore bounds high trace moments of the exactly centered residual $\widetilde{R}^{(l)}$ of (13). We begin with the exact scale of the objects appearing in (53).

Lemma 14 (Prescribed boundary probability)**.** *For fixed r, k and prescribed w of Hamming weight h,*

$$p_r(w) = O_{k,r,h}\big(n^{-(kr+h)/2}\big). \tag{54}$$

For feasible $h \in \{0,2\}$ this is a matching Θ estimate. The feasibility tests are

$$p_r(\varnothing) > 0 \iff r = 0 \text{ or } (r \geq 2 \text{ and } kr \text{ is even}), \tag{55}$$

$$p_r(w) > 0 \quad (|w| = 2) \iff kr \text{ is even and } \big(r \geq 2 \text{ or } (r = 1 \text{ and } k = 2)\big). \tag{56}$$

Proof. An ordered r-tuple of supports has kr clause–ground-variable incidence slots. Each prescribed odd-degree variable consumes at least one slot and every other used variable consumes at least two; hence at most $(kr-h)/2$ unprescribed ground-variable labels occur. Counting labels and dividing by $\binom{n}{k}^r$ gives (54). For a feasible zero- or two-point boundary, reserve the prescribed

slots. The remaining row counts have even total and no row count exceeds the sum of the others, so the slots can be paired so that every pair joins two distinct rows (for instance by repeatedly pairing slots from the two rows of largest positive count). Assigning a distinct fresh ground variable to each pair produces valid k-sets and the matching lower bound. Parity and the single-row exception give (55)–(56). □

Definition 15 (Overlap probability with prescribed final boundary)**.** For $0 \le u \le l$ and a final boundary $w \in \mathbb{F}_2^n$, set

$$\Gamma_u(w) := \sum_{v \in \mathbb{F}_2^n} p_u(v)\, p_{l-u}(w+v)^2. \tag{57}$$

Equivalently, $\Gamma_u(w)$ is the probability-weighted contribution of two l-clause witnesses that share exactly u clauses and whose common final boundary is w.

Proposition 16 (Exact overlap formula)**.** *For $i \neq j$,*

$$\mathbb{E}\big(\widetilde{R}_{ij}^{(l)}\big)^2 = \sum_{u=0}^{l} \Xi_u \Big(\rho^{2(l-u)}\Gamma_u(w_{ij}) - \rho^{2l} p_l(w_{ij})^2\Big), \qquad \Xi_u = \binom{m}{l}\binom{l}{u}\binom{m-l}{l-u}, \tag{58}$$

and the $u = 0$ summand vanishes.

Proof. Expand the square and group ordered pairs of witnesses by $u = |I \cap I'|$. Conditionally on the boundary v of the u common clauses, both private parts must have boundary $v + w_{ij}$. Shared signs are squared to one and private signs have mean ρ, which produces $\rho^{2(l-u)}\Gamma_u(w_{ij})$; the centering subtracts $\rho^{2l}p_l(w_{ij})^2$. At $u = 0$ the two witnesses are independent and the two terms cancel exactly. □

Definition 17 (Active overlaps and variance proxy)**.** An overlap size $u \in \{1, \dots, l\}$ is *active* if there is a $w \in \mathbb{F}_2^n$ with $|w| = 2$ such that $p_u(w) > 0$ and, in addition, $p_{l-u}(\varnothing) > 0$. The associated variance proxy is

$$\mathcal{V}_l(d, \rho) := \sum_{\substack{u \in \{1,\dots,l\} \\ u \text{ active}}} \rho^{2(l-u)}\, d^{\,2-u/l}. \tag{59}$$

Theorem 18 (Sharp second moment)**.** *Suppose $0 \le d_l = nN_l \le (\log n)^A$ for some constant A. Then*

$$\mathbb{E}\,\mathrm{tr}\big[(\widetilde{R}^{(l)})^2\big] = \Theta_{k,l,\rho}\big(n\,\mathcal{V}_l(d_l, \rho)\big), \tag{60}$$

and the smallest active overlap is

$$u^\star = \begin{cases} l, & l \in \{2, 3\},\ (k,l) \neq (2,3), \\ 1, & k = 2,\ l \ge 3, \\ 2, & k \ge 3,\ l \ge 4. \end{cases} \tag{61}$$

Proof. For active u the leading term of $\Gamma_u(w_{ij})$ is $v = w_{ij}$ and equals $p_u(w_{ij})p_{l-u}(\varnothing)^2$. Lemma 14 together with $p_l(w_{ij}) \asymp n^{-kl/2-1}$ gives the one-entry scale $n^{-1}\rho^{2(l-u)}d_l^{\,2-u/l}$. Writing $h = |v|$ and $\alpha = |w_{ij} \cap v|$, summing over all v with this profile produces, apart from common factors, the exponent $(h-\alpha) - h/2 - (h+2-2\alpha) = -h/2 + \alpha - 2 \le -1$, with equality only at $v = w_{ij}$; every other profile loses at least a factor $n^{-1/2}$. If u is inactive, the unique maximising profile vanishes, so the whole class carries this fixed power loss and is negligible for polylogarithmic d_l. For $u \ge 1$ the centering contribution obeys $\Xi_u p_l(w_{ij})^2 = O(n^{-ku/2-2}d_l^{\,2-u/l})$, smaller than the active scale by a fixed power of n, and the $u = 0$ term cancels exactly by Proposition 16. Summing the entry variances over the $n^2 - n$ off-diagonal entries gives (60), and the classification follows from (55)–(56). □

For $k \ge 3$ and $l \ge 4$ the active terms form a subset of a finite geometric progression whose endpoints $u = 2$ and $u = l$ are both present; since k, l are fixed, the sum is comparable to its larger endpoint, so that uniformly in $\rho \le 1$ and $d_l \ge 1$

$$\sqrt{\mathcal{V}_l(d_l, \rho)} \;\asymp\; \underbrace{d_l^{1/2}}_{\text{Wigner-like}} \;+\; \underbrace{\rho^{\,l-2} d_l^{\,1-1/l}}_{\text{correlations}}. \tag{62}$$

The two terms of (17) are exactly the trace-moment counterparts of these two endpoints.

G.2 Support components and boundary sectors

For a witness I let $\mathsf{G}(I)$ be the graph on its clauses that joins two clauses whose supports intersect; its connected components are called *support components*.

Lemma 19 (Boundary-component dichotomy)**.** *If $\partial I = \{i, j\}$, then either one support component has boundary $\{i, j\}$, or two support components have boundaries $\{i\}$ and $\{j\}$. Every other component has empty boundary, and the two alternatives are disjoint.*

Proof. Distinct support components use disjoint ground-variable sets, so their boundaries are disjoint. The statement lists the only partitions of the two-set $\{i, j\}$ into non-empty parts. □

Let $\widehat{A}^{(l)}_{\mathrm{bc}}$ and $\widehat{A}^{(l)}_{\mathrm{bd}}$ be the corresponding sub-sums of the gauged matrix $\widehat{A}^{(l)} = D_x A^{(l)} D_x$, and set $R^{(l)}_\sigma = \widehat{A}^{(l)}_\sigma - \mathbb{E}\widehat{A}^{(l)}_\sigma$ for $\sigma \in \{\mathrm{bc}, \mathrm{bd}\}$. Then

$$\widetilde{R}^{(l)} = R^{(l)}_{\mathrm{bc}} + R^{(l)}_{\mathrm{bd}}. \tag{63}$$

Boundary-connected does *not* mean that the whole witness is connected: additional zero-boundary components are allowed.

Proposition 20 (Disconnected-sector vanishing)**.** *If k is even, or if $l < 6$, then $R^{(l)}_{\mathrm{bd}} = 0$ identically.*

Proof. Let J be a support component consisting of r clauses. Its total number of clause–ground-variable incidences is kr, and if $d_J(v)$ denotes the number of clauses of J containing v, then

$$kr = \sum_{v \in [n]} d_J(v) \equiv \#\{v : d_J(v) \text{ odd}\} = |\partial J| \pmod 2.$$

Hence a component with singleton boundary has kr odd, which forces both k and r odd; in particular no such component exists when k is even. If $r = 1$ then $\partial J = S_a$ has size $k \geq 2$, so $r \neq 1$, and since r is odd every singleton-boundary component contains at least three clauses. By Lemma 19 a boundary-disconnected witness contains two distinct singleton-boundary components, hence at least six clauses in total. Therefore $\widehat{A}^{(l)}_{\mathrm{bd}} = 0$ whenever k is even or $l < 6$, and so is its centered version. □

For intuition, let $U_r(i)$ count connected r-clause sets of boundary $\{i\}$. At $(k, l) = (3, 6)$ the matrix $\widehat{A}^{(6)}_{\mathrm{bd}}$ is, off the diagonal, $U_3 U_3^\top$ minus intersecting component pairs: a low-rank contraction that is invisible to the entry-variance analysis of Theorem 18. This term is removed exactly by the trimming in Subsection G.7.

G.3 Hoeffding decomposition and trace diagrams

Definition 21 (One-witness and sector kernels)**.** Write a clause variable as $z_t = (S_t, \eta_t)$, where $S_t \in \binom{[n]}{k}$ is its support and $\eta_t \in \{\pm 1\}$ its noise sign. The symmetric *one-witness kernel* is the matrix

$$H(z_1, \dots, z_l)_{ij} = \mathbf{1}\{S_1 \triangle \cdots \triangle S_l = \{i, j\}\} \prod_{t=1}^{l} \eta_t \quad (i \neq j), \qquad H_{ii} = 0, \tag{64}$$

i.e. the contribution of a single unordered set of l clause samples before summing over $I \in \binom{[m]}{l}$. Form the support-intersection graph on the argument positions $[l]$, joining α and β when $S_\alpha \cap S_\beta \neq \varnothing$. Define H_{bc} by retaining the entry in (64) when i and j lie in the same support component, and H_{bd} by retaining it when they lie in two different components; all other entries are set to zero, so that $H = H_{\mathrm{bc}} + H_{\mathrm{bd}}$ pointwise.

For $1 \leq s \leq l$ and $\Lambda \subseteq [s]$ — a set of *argument positions*, equivalently clause slots, not of ground variables — put $w_t = z_t$ for $t \in \Lambda$ and $w_t = (S'_t, \varepsilon_t)$ for $t \in [l] \setminus \Lambda$, and define the projection

$$\begin{aligned}(H_\sigma)_\Lambda(z_\Lambda) &:= \mathbb{E}_{(Z'_t)_{t \in [l] \setminus \Lambda}} H_\sigma(w_1, \dots, w_l) \\ &= \frac{1}{\binom{n}{k}^{l - |\Lambda|}} \sum_{(S'_t)_{t \notin \Lambda}} \sum_{(\varepsilon_t)_{t \notin \Lambda}} \Big(\prod_{t \notin \Lambda} \tfrac{1 + \rho \varepsilon_t}{2} \Big) H_\sigma(w_1, \dots, w_l).\end{aligned} \tag{65}$$

Thus the arguments in Λ are held fixed while every argument position outside Λ is independently averaged over both its uniform k-support and its noise sign.

With $Z_a = (S_a, \eta_a)$, define the order-s canonical kernel and the multiplicity coefficient

$$h_{\sigma,s}(z_1, \dots, z_s) = \sum_{\Lambda \subseteq [s]} (-1)^{s-|\Lambda|} (H_\sigma)_\Lambda(z_\Lambda), \qquad \beta_s = \binom{m-s}{l-s}. \tag{66}$$

Lemma 22 (Sectorwise Hoeffding decomposition)**.** *Every $h_{\sigma,s}$ is canonical: for each argument position $r \in [s]$,*

$$\mathbb{E}_{Z_r}\big[h_{\sigma,s}(z_1, \dots, z_{r-1}, Z_r, z_{r+1}, \dots, z_s)\big] = 0, \tag{67}$$

where Z_r is a fresh clause variable and the remaining arguments are fixed. Moreover

$$R_\sigma^{(l)} = \sum_{s=1}^{l} Q_{\sigma,s}, \qquad Q_{\sigma,s} = \beta_s \sum_{J \in \binom{[m]}{s}} h_{\sigma,s}(Z_J). \tag{68}$$

In a trace expansion, canonicality annihilates every clause-index equality block that occurs in only one matrix factor.

Proof. Fix $r \in [s]$ and average (66) over its rth argument. Group the subsets $\Lambda \subseteq [s]$ into pairs Λ' and $\Lambda' \cup \{r\}$ with $\Lambda' \subseteq [s] \setminus \{r\}$. The projection $(H_\sigma)_{\Lambda'}$ does not depend on the rth argument, while averaging the fixed argument Z_r in $(H_\sigma)_{\Lambda' \cup \{r\}}$ turns it back into a fresh clause variable, so $\mathbb{E}_{Z_r}(H_\sigma)_{\Lambda' \cup \{r\}}(z_{\Lambda'}, Z_r) = (H_\sigma)_{\Lambda'}(z_{\Lambda'})$. The two members of each pair carry opposite signs in (66) and cancel, proving (67).

Next fix one l-clause witness with clause-index set I. Möbius inversion on the subset lattice gives $H_\sigma(Z_I) - \mathbb{E}H_\sigma = \sum_{\varnothing \neq J \subseteq I} h_{\sigma,|J|}(Z_J)$. Summing over all $I \in \binom{[m]}{l}$, a fixed s-set J occurs once for each l-set containing it, i.e. $\binom{m-s}{l-s} = \beta_s$ times; grouping by $s = |J|$ proves (68). Finally, if a clause index occurs in only one matrix factor of a trace product, condition on all other clause variables and integrate that one: the factor vanishes by (67), hence so does the whole term. □

Fix an even $q \geq 2$ and Hoeffding orders $\boldsymbol{s} = (s_1, \dots, s_q) \in [l]^q$, and let $\Omega_{\boldsymbol{s}} = \{(t, \alpha) : 1 \leq t \leq q,\ 1 \leq \alpha \leq s_t\}$ be the resulting slot set. A partition π of $\Omega_{\boldsymbol{s}}$ is *admissible* if no block contains two slots with the same first coordinate and every block has size at least two; write $\mathfrak{P}(\boldsymbol{s})$ for the set of admissible partitions and $[t, \alpha]_\pi$ for the block containing (t, α). An injection $\phi : \pi \hookrightarrow [m]$ assigns distinct clause indices to distinct blocks, hence realises exactly the equality pattern encoded by π. For factor t set

$$J_t(\phi) := \big(\phi([t,1]_\pi), \dots, \phi([t,s_t]_\pi)\big), \tag{69}$$

whose entries are distinct by admissibility and injectivity, and define the diagram value

$$D_\sigma(\boldsymbol{s}, \pi) = \frac{\prod_{t=1}^q \beta_{s_t}}{\prod_{t=1}^q s_t!} \sum_{\phi : \pi \hookrightarrow [m]} \mathbb{E}\,\mathrm{tr} \prod_{t=1}^{q} h_{\sigma,s_t}\big(Z_{J_t(\phi)}\big). \tag{70}$$

Proposition 23 (Exact sectorwise trace-diagram identity)**.** *For $\sigma \in \{\mathrm{bc}, \mathrm{bd}\}$ and every even $q \geq 2$,*

$$\mathbb{E}\,\mathrm{tr}\big[(R_\sigma^{(l)})^q\big] = \sum_{\boldsymbol{s} \in [l]^q} \sum_{\pi \in \mathfrak{P}(\boldsymbol{s})} D_\sigma(\boldsymbol{s}, \pi). \tag{71}$$

Proof. Insert (68) into each of the q trace factors and choose $\boldsymbol{s}$; since h_{σ,s_t} is symmetric, replace each unordered J_t by an ordered tuple of distinct clause indices and divide by $s_t!$. Every term then assigns one clause index to each slot of $\Omega_{\boldsymbol{s}}$, and grouping slots with equal indices produces a unique partition π. Indices within one tuple are distinct, so no block contains two slots from the same factor; and a singleton block would make the term vanish by Lemma 22. Hence every surviving π is admissible. Conversely each pair (π, ϕ) reconstructs one term through (69). Collecting terms with the same π gives (70), and summing over $\boldsymbol{s}$ and π proves (71). □

Thus the exact trace expansion retains arbitrary multiway blocks: different trace factors may reuse clauses exposed by earlier factors. To count the ground-variable labels of such a pattern we expose factors sequentially and charge only the clauses appearing for the first time.

G.4 SEQUENTIAL EXPOSURE AND PATTERN COUNTING

Lemma 24 (Sequential exposure)**.** *Suppose some clauses have already been exposed and let O contain every ground variable incident to them. Expose one further witness; let f be the number of its clauses whose indices have not appeared before (equivalently, the number of fresh independent supports), and let v_{new} be the number of its ground variables outside O.*

(a) *If its boundary is $\{i, j\}$ with $i \in O$ and i, j in one support component, then $v_{\mathrm{new}} \le kf/2$.*
(b) *A singleton-boundary witness satisfies $v_{\mathrm{new}} \le (kf+1)/2$; if it is connected, its root is old and $f > 0$, then $v_{\mathrm{new}} \le (kf-1)/2$.*
(c) *A zero-boundary witness satisfies $v_{\mathrm{new}} \le kf/2$.*

Proof. The f fresh clauses carry kf incidences in total. Let b be the number of these whose ground-variable endpoint lies in O; since previously exposed clauses cannot contain a variable outside O, the fresh clauses contribute exactly $kf - b$ incidences to variables outside O. Let r_∂ be the number of boundary variables of the new witness lying outside O. Every outside non-boundary variable has positive even degree, hence degree at least two, and every outside boundary variable has degree at least one, so

$$kf - b \;\ge\; 2(v_{\mathrm{new}} - r_\partial) + r_\partial = 2v_{\mathrm{new}} - r_\partial, \qquad \text{i.e.} \qquad v_{\mathrm{new}} \le \frac{kf - b + r_\partial}{2}. \tag{72}$$

For (a), $i \in O$ leaves at most j outside. If $j \in O$ then $r_\partial = 0$ and (72) gives $v_{\mathrm{new}} \le kf/2$. If $j \notin O$ then $r_\partial = 1$, and because i, j lie in one support component a path from i to j must cross from O to its complement; at the first crossing a fresh clause meets a variable of O, so $b \ge 1$ and the terms $-b$ and $+r_\partial$ cancel. For (b), $r_\partial \le 1$ and $b \ge 0$ give $v_{\mathrm{new}} \le (kf+1)/2$; in the rooted connected case $r_\partial = 0$, and if $f > 0$ connectivity forces some fresh clause to meet an old variable, so $b \ge 1$. For (c), $r_\partial = 0$ and $b \ge 0$. □

Corollary 25 (Completed connected trace patterns)**.** *A completed product of q boundary-connected entries along a closed trace walk, using N independent supports and v ground variables, satisfies*

$$v \le 1 + kN/2. \tag{73}$$

Proof. Expose the first trace variable, then the witnesses in cyclic order. Each left endpoint is then old, so Lemma 24(a) applies, and each support is new exactly once. □

Lemma 26 (Pattern count)**.** *There is a constant $C_k > 0$, depending only on k, such that an expansion with q factors has at most $(C_k lq)^{a_0 q}$ abstract clause- and ground-variable equality patterns, where $a_0 = (k+1)l + 2$. A pattern with c sampled clause labels, N actually used independent supports and v ground-variable labels contributes at most*

$$C_k^{lq} m^c n^v n^{-kN} \tag{74}$$

before deterministic coefficients.

Proof. An abstract pattern records which of the at most ql clause slots are equal and, separately, which of the at most $(kl+2)q$ incidence and trace positions represent the same ground variable, i.e. two set partitions of at most $a_0 q$ positions; since the number of partitions of r objects is at most r^r, there are at most $(C_k lq)^{a_0 q}$ of them. Once a pattern is fixed, its c clause classes and v ground-variable classes can be labelled in at most m^c and n^v ways, and each of the N distinct independent supports carries probability $\binom{n}{k}^{-1}$; as $N \le ql$, $\binom{n}{k}^{-N} \le C_k^{lq} n^{-kN}$ for all sufficiently large n. □

Lemma 27 (Tensorized completion of a canonical diagram)**.** *Fix $1 \le s \le l$, an even $q \ge 2$, and a surviving clause-index partition of the qs core slots in the expansion of $\mathbb{E}\,\mathrm{tr}\,Q_s^q$. Let c be its number of blocks. For factor t, choose $\Lambda_t \subseteq [s]$ in the expansion of the canonical kernel, and put*

$$\Lambda := \sum_{t=1}^{q} |\Lambda_t|, \qquad D := qs - \Lambda.$$

Let c' be the number of clause-index blocks containing at least one retained slot and let $F := c - c'$ be the number of completely discarded blocks. Then the corresponding projection term has the following properties.

(i) *On an enlarged product probability space it is the expectation of a product of q genuine boundary-connected l-clause kernels. Retained slots in the same block use the same sampled clause variable, and every omitted slot is filled by a fresh auxiliary clause variable; all auxiliary variables are mutually independent and independent of the sampled variables.*

(ii) *The number of independent support draws and the exponent used in the sampled clause-label upper bound are respectively*

$$N = c' + ql - \Lambda, \qquad E_0 = c + q(l - s). \tag{75}$$

Every block of a surviving partition has size at least two, so $D \geq 2F$ and

$$N - E_0 = D - F \geq F \geq 0. \tag{76}$$

(iii) *Distinct independent draws remain distinct probability factors even if their realised k-supports coincide. After fixing a ground-variable equality pattern with v labels, its absolute contribution, including the sum over sampled clause indices and the factors β_s^q, is at most*

$$C_k^{lq} m^{E_0} n^v n^{-kN}. \tag{77}$$

The pattern may contain additional zero-boundary support components; they are included in the same N draws and v labels.

The projection signs have absolute value one, there are at most 2^{sq} choices of $(\Lambda_1, \ldots, \Lambda_q)$, and the abstract pattern count is bounded by Lemma 26.

Proof. For a fixed factor, Definition 21 writes the selected projection as an expectation over its omitted clause arguments. Conditional on all retained sampled clauses, the product of these q expectations equals one expectation over q mutually independent collections of auxiliary arguments. This is the product-measure tensorization identity. It proves (i) and gives one independent draw for each of the c' retained blocks and one for each of the $ql-\Lambda$ auxiliary arguments, proving the formula for N.

The injection assigning sampled clause indices to the c blocks has at most m^c choices, while $\beta_s^q = \binom{m-s}{l-s}^q \leq m^{q(l-s)}$. This proves the formula for E_0. A completely discarded block contributes all of its slots to D; by canonicality every surviving block has size at least two. Hence $D \geq 2F$, and direct substitution gives $N - E_0 = D - F$, proving (ii).

The N source variables just counted are independent draws from $\binom{[n]}{k}$. Equality of two realised support values is therefore an event in the ground-variable equality pattern, not an identification of their source variables; it still carries two factors $\binom{n}{k}^{-1}$. Once a pattern is fixed, label its v ground-variable classes in at most n^v ways and bound $\binom{n}{k}^{-N} \leq C_k^{lq} n^{-kN}$. This proves (77). Each completed kernel is boundary-connected by construction, but its other support components may have zero boundary; no step of this count discards them. The remaining coefficient and pattern-count statements follow from the Möbius expansion and Lemma 26. □

G.5 The boundary-connected moment bound

Write $Q_s = Q_{\mathrm{bc},s}$. Since $m = Cn^{k/2}$, the density C plays the role of the natural expansion parameter, and $d_l \asymp C^l$ at fixed k, l.

Theorem 28 (All connected Hoeffding orders)**.** *There is a constant $D_k > 0$, depending only on k, such that, for $1 \leq s \leq l$, $C > 0$, $\rho \in [0, 1]$, even $q \geq 2$, and all sufficiently large n (depending on k, l),*

$$\mathcal{M}_q(Q_s) \leq n^{1/q} (D_k lq)^a \rho^{l-s} \begin{cases} C^{l-s+s/q}, & 0 < C \leq 1, \\ C^{l-s/2}, & C \geq 1, \end{cases} \qquad a = (k+1)l + 3. \tag{78}$$

If $\rho = 0$ and $s < l$ then $Q_s = 0$.

Proof. Since q is even and Q_s symmetric, $\operatorname{tr}|Q_s|^q = \operatorname{tr} Q_s^q$, and by (68)

$$\mathbb{E} \operatorname{tr} Q_s^q = \beta_s^q \sum_{J_1, \ldots, J_q \in \binom{[m]}{s}} \mathbb{E} \operatorname{tr}\big[h_{\mathrm{bc},s}(Z_{J_1}) \cdots h_{\mathrm{bc},s}(Z_{J_q})\big]. \tag{79}$$

Regard the s arguments of each of the q kernels as qs *core slots*. Grouping slots with equal clause index gives an equality partition with, say, c blocks; singleton blocks vanish by canonicality (67), *before* absolute values are taken, and the s slots of one factor lie in distinct blocks. Hence

$$s \le c \le qs/2. \tag{80}$$

Fix a surviving partition and expand each canonical kernel via (66): a term chooses $\Lambda_t \subseteq [s]$, keeping those arguments and averaging the others out. Apply Lemma 27. Its notation gives $N = c' + ql - \Lambda$, $E_0 = q(l-s) + c$, and $N - E_0 \ge 0$. Moreover,

$$q(l-s) + s \le E_0 \le q(l - s/2), \tag{81}$$

by (80).

Corollary 25 allows at most $1 + kN/2$ ground-variable labels; this includes every additional zero-boundary component. By Lemma 27, a fixed completed pattern is bounded by

$$C_k^{lq} m^{E_0} n^{1-kN/2} = C_k^{lq} n^{1-k(N-E_0)/2} C^{E_0} \le C_k^{lq} n C^{E_0},$$

using $m = Cn^{k/2}$ and $N \ge E_0$. The two bounds on c give $C^{E_0} \le C^{q(l-s/2)}$ when $C \ge 1$ and $C^{E_0} \le C^{q(l-s)+s}$ when $0 < C \le 1$. In factor t, averaging the $l - |\Lambda_t| \ge l - s$ omitted noise signs contributes at most ρ^{l-s}, so the product over factors is at most $\rho^{q(l-s)}$. Finally Lemma 26, together with the choices of the sets Λ_t, gives at most $(C_k lq)^{a_0 q} 2^{lq}$ terms. Together with the factor C_k^{lq} from a fixed pattern, the factor 2^{lq} is absorbed into the base by enlarging D_k, since $a_0 \ge (k+1)l$. Thus the total is at most $(D_k lq)^{a_0 q}$ times the displayed density and noise factors. Taking qth roots and using $a = a_0 + 1$ yields (78). If $\rho = 0$ and $s < l$, every projection averages at least one omitted noise sign, so $h_{\mathrm{bc},s} = 0$. □

Canonical cancellation must precede completion: it removes singleton core blocks without taking absolute values.

We use the following standard consequence of the self-adjoint matrix Bernstein inequality [18]. If $X_1, \dots, X_m$ are independent centered self-adjoint $n \times n$ matrices with $\|X_a\|_{\mathrm{op}} \le \mathsf{L}$ almost surely and $\|\sum_a \mathbb{E} X_a^2\|_{\mathrm{op}} \le \mathsf{v}$, then, for every $q \ge 2$,

$$\big(\mathbb{E}\,\mathrm{tr}\,|\textstyle\sum_a X_a|^q\big)^{1/q} \le n^{1/q} B\big(\sqrt{q\mathsf{v}} + q\mathsf{L}\big). \tag{82}$$

It follows by integrating the usual matrix Bernstein tail bound; the numerical constant B is universal.

Lemma 29 (Improved first projection for $k \ge 3$)**.** *For even $q \ge 2$ and $k \ge 3$, with the constant D_k of Theorem 28,*

$$\mathcal{M}_q(Q_1) \le n^{1/q} (D_k lq)^a \rho^{l-1} \big[\sqrt{q}\, n^{1/2-k/4} C^{l-1/2} + q\, n^{1-k/2} C^{l-1}\big]. \tag{83}$$

Proof. For a fixed support S let $P_S(i,j)$ be the probability that S, completed by $l-1$ fresh supports, forms a boundary-connected witness of boundary $\{i,j\}$, and put $G_S = \binom{m-1}{l-1} P_S$. If $u = |S \cap \{i,j\}|$ then Lemma 14 gives $G_S(i,j) \le (D_k l)^a n^{-k/2-1+u} C^{l-1}$. Indeed, its slot count uses at most $k(l-1)+2$ positions and hence at most $(C_k l)^{k(l-1)+2}$ equality patterns, which is bounded by $(D_k l)^a$ because $a = (k+1)l + 3$. Block row sums therefore yield

$$\|G_S\|_{\mathrm{op}} \le (D_k l)^a n^{1-k/2} C^{l-1}, \qquad \|G_S \mathbf{1}\|_2^2 \le (D_k l)^{2a} n^{2-k} C^{2l-2}. \tag{84}$$

The law of G_S is invariant under permutations of $[n]$, so the non-negative matrix $\mathbb{E} G_S^2$ has constant row sums and therefore

$$\|\mathbb{E} G_S^2\|_{\mathrm{op}} = \tfrac{1}{n} \mathbb{E}\, \|G_S \mathbf{1}\|_2^2 \le (D_k l)^{2a} n^{1-k} C^{2l-2}. \tag{85}$$

Now $Q_1 = \sum_a X_a$ with independent centered summands $X_a = \rho^{l-1} \eta_a G_{S_a} - \mathbb{E}[\rho^{l-1} \eta_a G_{S_a}]$, and the matrix Bernstein moment inequality (82), with the variance and norm proxies of (84)–(85) gives (83). □

Theorem 30 (Connected-sector bound)**.** *For $k \ge 3$, $\rho \in [0,1]$, even $q \ge 2$, $C > 0$, and all sufficiently large n (depending on k, l),*

$$\mathcal{M}_q\big(R_{\mathrm{bc}}^{(l)}\big) \le n^{1/q} (D_k lq)^a \begin{cases} C^{l/q} + \rho^{l-2} C^{l-2+2/q}, & 0 < C \le 1, \\ C^{l/2} + \rho^{l-2} C^{l-1}, & 1 < C \le n^{k/2-1}, \\ C^{l/2} + \rho^{l-1} C^{l-1/2}, & C > n^{k/2-1}. \end{cases} \tag{86}$$

Proof. By (68) and the triangle inequality $\mathcal{M}_q(R_{\mathrm{bc}}^{(l)}) \le \sum_{s=1}^{l} \mathcal{M}_q(Q_s)$. Sum first the orders $2 \le s \le l$. If $\rho = 0$, the last assertion of Theorem 28 gives $Q_s = 0$ for every $s < l$, so only the $s = l$ term remains and all three displayed bounds follow immediately. Assume $\rho > 0$ below. For $0 < C \le 1$, Theorem 28 gives, suppressing the common prefactor $n^{1/q}(D_k lq)^a$, the geometric sequence $u_s = \rho^{l-s} C^{l-s+s/q}$ with ratio $u_{s+1}/u_s = \rho^{-1} C^{-1+1/q}$, so that regardless of whether the ratio exceeds one

$$\sum_{s=2}^{l} u_s \le (l-1)(u_2 + u_l) = (l-1)\big(\rho^{l-2} C^{l-2+2/q} + C^{l/q}\big), \tag{87}$$

the fixed factor $l - 1$ being absorbed into $(D_k lq)^a$. Likewise for $C \ge 1$ with $w_s = \rho^{l-s} C^{l-s/2}$,

$$\sum_{s=2}^{l} w_s \le (l-1)\big(\rho^{l-2} C^{l-1} + C^{l/2}\big). \tag{88}$$

It remains to absorb Q_1. If $C \le n^{k/2-1}$, dividing the first term of (83) by the relevant $s = 2$ endpoint gives, up to a factor already present in $(D_k lq)^a$, the quantity $n^{1/2-k/4} C^{3/2-2/q} \le 1$ for $0 < C \le 1$ (using $q \ge 2$, $k \ge 3$) and $n^{1/2-k/4} C^{1/2} \le 1$ for $1 < C \le n^{k/2-1}$; for the second term the corresponding ratios are $n^{1-k/2} C^{1-2/q} \le 1$ and $n^{1-k/2} \le 1$. The factors $\rho \le 1$, q and $\sqrt{q}$ are absorbed into $(D_k lq)^a$, so $\mathcal{M}_q(Q_1)$ is dominated by the $s = 2$ term of (87) or (88). Finally, if $C > n^{k/2-1}$, apply the dense branch of Theorem 28 to all $1 \le s \le l$ and repeat the endpoint estimate with endpoints $s = 1$ and $s = l$. □

Corollary 31 (Level-explicit dense deviation)**.** *Fix $k \ge 3$, $l \ge 2$, and $\gamma > 0$, and let $a = (k+1)l+3$. There is a constant $D_{k,\gamma} > 0$, independent of l, such that, whenever $1 < C \le n^{k/2-1}$, for all sufficiently large n,*

$$\mathbb{P}\Big[\|R_{\mathrm{bc}}^{(l)}\|_{\mathrm{op}} > D_{k,\gamma}^{a} (\log n)^a \big(C^{l/2} + \rho^{l-2} C^{l-1}\big)\Big] \le n^{-\gamma}. \tag{89}$$

Proof. For an even q, Theorem 30 and Markov's inequality give

$$\mathbb{P}[\|R_{\mathrm{bc}}^{(l)}\|_{\mathrm{op}} > t] \le n \left[\frac{(D_k lq)^a (C^{l/2} + \rho^{l-2} C^{l-1})}{t}\right]^q.$$

Choose the least even integer $q \ge (\gamma+1)\log n/a$. For fixed l and sufficiently large n, this choice satisfies $q \ge 2$, $n^{(\gamma+1)/q} \le e^a$, and $lq \le C_{k,\gamma} \log n$. Taking

$$t = n^{(\gamma+1)/q} (D_k lq)^a (C^{l/2} + \rho^{l-2} C^{l-1})$$

proves (89) after enlarging $D_{k,\gamma}$. □

G.6 Even-k residual, signal-to-noise and the threshold

Proposition 32 (Even-k residual)**.** *Fix even $k \ge 4$, an admissible $l \ge 2$, $0 < C \le n^{k/2-1}$, $d_l \ge 1$ and even $q \ge 2$. For all sufficiently large n depending on k, l, $R_{\mathrm{bd}}^{(l)} = 0$ and*

$$\mathcal{M}_q\big(\widetilde{R}^{(l)}\big) \le n^{1/q} (Bq)^a \Big[d_l^{1/2} + \rho^{l-2} d_l^{1-1/l}\Big]. \tag{90}$$

Proof. Proposition 20 gives $\widetilde{R}^{(l)} = R_{\mathrm{bc}}^{(l)}$ for even k. For $1 < C \le n^{k/2-1}$, insert the dense branch of Theorem 30 and use the finite-n calibration of Lemma 42. Its large-m condition is automatic here for sufficiently large n: Lemma 14 gives $d_l \le B_{k,l} C^l$, so $d_l \ge 1$ bounds C away from zero and hence $m = Cn^{k/2} \to \infty$. For $0 < C \le 1$, use the sparse branch and the facts that $d_l \ge 1$ and $q \ge 2$ imply $d_l^{1/q} \le d_l^{1/2}$ and $d_l^{1-2/l+2/(ql)} \le d_l^{1-1/l}$. Theorem 18 and $\mathcal{M}_q \ge \mathcal{N}_q \ge \mathcal{N}_2 \ge (n^{-1}\mathbb{E}\,\mathrm{tr}(\cdot)^2)^{1/2}$ show, when $d_l \le (\log n)^A$, that the same dependence on d_l is sharp up to fixed constants (for $l \in \{2, 3\}$ only the $d_l^{1/2}$ endpoint is active). At the corollary's threshold one may take any fixed $A \ge al$. □

Proof of Theorem 4. Estimate (17) is exactly (90). For the deviation bound, let $q_n = 2\lceil \log n \rceil$ and abbreviate $\Delta_n := d_l^{1/2} + \rho^{l-2} d_l^{1-1/l}$. Using $\|\widetilde{R}^{(l)}\|_{\mathrm{op}}^{q_n} \le \operatorname{tr} |\widetilde{R}^{(l)}|^{q_n}$ and Markov's inequality, for any $t > 0$

$$\mathbb{P}\big[\|\widetilde{R}^{(l)}\|_{\mathrm{op}} \ge t\big] \le t^{-q_n}\, \mathbb{E} \operatorname{tr} |\widetilde{R}^{(l)}|^{q_n} \le n \left(\frac{(Bq_n)^a \Delta_n}{t} \right)^{q_n},$$

where the factor $n^{1/q_n} \le e^{1/2}$ from (90) has been absorbed into B. Choosing $t = B_\gamma (\log n)^a \Delta_n$ with B_γ large enough that the bracket is at most $e^{-(1+\gamma)}$ (possible because $q_n \le 2 \log n + 2$, so $(Bq_n)^a \le B'(\log n)^a$) gives $\mathbb{P}[\|\widetilde{R}^{(l)}\|_{\mathrm{op}} \ge t] \le n\, e^{-q_n(1+\gamma)} \le n^{-\gamma-1}$. Adding the deterministic diagonal shift of (13), which costs at most θ_l/n, turns this into the bound (18) for $\|R^{(l)}\|_{\mathrm{op}}$, with probability at least $1 - n^{-\gamma}$.

For the correlation bound, work in the gauged frame: $D_x x = \mathbf{1}$, so with $u = \mathbf{1}/\sqrt{n}$ the decomposition (12) reads

$$\widehat{A}^{(l)} = \theta_l\, uu^\top - \frac{\theta_l}{n} I + \widetilde{R}^{(l)}.$$

The multiple of the identity shifts the whole spectrum and does not affect eigenvectors. On the event $\|\widetilde{R}^{(l)}\|_{\mathrm{op}} \le \mathfrak{R}_{n,\gamma}$, and under the assumption $\theta_l > (3+\varepsilon)\mathfrak{R}_{n,\gamma}$, Weyl's inequality separates the top eigenvalue of $\widehat{A}^{(l)}$ from the rest of its spectrum by a gap of at least $\theta_l - 2\mathfrak{R}_{n,\gamma} > 0$, so that eigenvalue is isolated. The Davis–Kahan $\sin\theta$ theorem [3; 21] then gives, for a unit top eigenvector $\hat{u}$,

$$\sin \angle(\hat{u}, u) \le \frac{\mathfrak{R}_{n,\gamma}}{\theta_l - \mathfrak{R}_{n,\gamma}},$$

and $|\langle \hat{u}, u \rangle| = \cos \angle(\hat{u}, u) = \sqrt{1 - \sin^2 \angle(\hat{u}, u)}$ gives (19). Conjugating back by D_x maps u to $x/\sqrt{n}$ and leaves every norm involved unchanged. Choose the global sign of $\hat{u}$ so that $\langle \hat{u}, u \rangle \ge 0$. With $\delta = \mathfrak{R}_{n,\gamma}/(\theta_l - \mathfrak{R}_{n,\gamma})$, the correlation bound gives $\|\hat{u} - u\|_2^2 \le 2\delta^2$. Each disagreement between $\operatorname{sgn}(\hat{u})$ and x contributes at least $1/n$ to this squared distance, proving (19). Finally, $\theta_l > (3+\varepsilon)\mathfrak{R}_{n,\gamma}$ implies $\delta < 1/2$, so the displayed binary overlap is strictly positive. □

The exact signal-to-noise ratio is $\mathrm{SNR}_l = \theta_l(1 - 1/n)/\|\widetilde{R}^{(l)}\|_{\mathrm{op}}$. Comparing $\theta_l = \rho^l d_l$ with the two endpoints in (18) gives respectively $\rho^{2l} d_l \gg (\log n)^{2a}$ and $\rho^{2l} d_l \gg (\log n)^{al}$. Since $l \ge 2$, the latter is binding, which is (20).

Proof of Corollary 5. Let $D = D_{k,\gamma}$ be the constant in Corollary 31, let $a_l = (k+1)l + 3$, and put $\lambda = 3 + \varepsilon$. The witness asymptotic (135) implies

$$\kappa_{l,k}^{1/l} = c_k l^{k/2-1}(1 + o_l(1)) \quad \text{for some } c_k > 0.$$

Consequently the two suprema over admissible $l \ge 2$,

$$M_1 := \sup_l D^{a_l} \kappa_{l,k}^{-(1-1/l)},$$
$$M_2 := \sup_l D^{2a_l/l} \kappa_{l,k}^{-1/l},$$

are finite: their summands tend respectively to zero like $[D^{k+1}/(c_k l^{k/2-1})]^{l-1}$ and $D^{2(k+1)}/(c_k l^{k/2-1})$. Choose $K_{k,\gamma,\varepsilon}$ larger than a fixed multiple of $\lambda \max\{1, M_1, M_2\}$.

On the event of (89), write

$$t_l := D^{a_l} (\log n)^{a_l} (C^{l/2} + \rho^{l-2} C^{l-1}).$$

Since $\theta_l = (1 + o_n(1))\rho^l \kappa_{l,k} C^l$, condition (21) and the definitions of M_1, M_2 give, for all sufficiently large n,

$$\theta_l > 2\lambda t_l.$$

Indeed, the correlated endpoint uses $K \ge \lambda D^{a_l} \kappa_{l,k}^{-(1-1/l)}$, while the Wigner endpoint uses $K \ge \lambda^{2/l} D^{2a_l/l} \kappa_{l,k}^{-1/l}$; the extra powers of $\log n$ are nonnegative because $l \ge 2$.

For even k, $R^{(l)}_{\mathrm{bd}} = 0$, so $\widetilde{R}^{(l)} = R^{(l)}_{\mathrm{bc}}$. Adding the deterministic diagonal shift costs θ_l/n. For $n \geq 2\lambda$, the preceding display yields

$$\theta_l > \lambda(t_l + \theta_l/n).$$

Repeating the Weyl, Davis–Kahan, and sign-rounding argument of Theorem 4 with radius $t_l + \theta_l/n$ gives the stated overlap. Corollary 31 gives failure probability at most $n^{-\gamma}$. □

The large-l asymptotic (135), together with the l-independence of $K_{k,\gamma,\varepsilon}$, identifies the algebraic large-l recovery curve without any hidden level-dependent prefactor. The lower bound on n in the fixed-level theorem is still not uniform in growing l, and the exponent $a = (k+1)l + 3$ grows with l; thus no theorem for $l = l(n)$ is claimed.

G.7 Odd k: trimming the boundary-disconnected sector

Throughout this subsection $k \geq 3$ is odd and l is even. For $l \geq 6$ the sector $R^{(l)}_{\mathrm{bd}}$ need not vanish. Instead of bounding its localized low-rank contribution, we remove it exactly by deleting a small, support-determined set of coordinates. For $l \in \{2, 4\}$ this set is empty and no trimming is needed; for $l \geq 6$, the resulting threshold contains the additional fixed-level factor $\kappa_{l,k}/\kappa^{\mathrm{bc}}_{l,k}$ (Corollary 38).

Remark 33 (Why the untrimmed matrix is not used)**.** This heuristic is not used below. At $l = 6$, let $U_3(i)$ be the gauged signed count of connected three-clause sets with boundary $\{i\}$. Apart from sparse compatibility corrections, $\widehat{A}^{(6)}_{\mathrm{bd}}$ is $U_3U_3^\top$ off the diagonal. In the sparse-certificate regime $C^3 = o(\sqrt{n})$, which includes fixed and polylogarithmic C, Lemma 14 gives $O(C^3\sqrt{n})$ expected three-clause singleton-boundary certificates. Neglecting lower-order endpoint collisions then predicts $\mathbb{E}|\operatorname{supp} U_3| = \Theta(C^3\sqrt{n})$ and a norm of the same order, whereas $\theta_6 \asymp \rho^6 C^6$. Thus the untrimmed top eigenvector can localize at polylogarithmic C.

Definition 34 (Singleton-boundary set and trimmed recovery matrix)**.** For even l define

$$\mathcal{B}_l := \big\{v \in [n] : \partial P = \{v\} \text{ for some } r \in \{3, \dots, l-3\} \text{ and } P \in \tbinom{[m]}{r}\big\}, \qquad \mathcal{G}_l := [n]\backslash\mathcal{B}_l. \tag{91}$$

Let $\Pi_l := \operatorname{diag}(\mathbf{1}_{\mathcal{G}_l})$ and

$$\bar{A}^{(l)} := \Pi_l A^{(l)} \Pi_l. \tag{92}$$

We regard $\bar{A}^{(l)}$ as acting on $\operatorname{Ran}(\Pi_l)$, equivalently as the principal submatrix indexed by $\mathcal{G}_l$. The trimmed estimator takes the signs of its top eigenvector on $\mathcal{G}_l$ and assigns arbitrary signs on $\mathcal{B}_l$.

Only odd values of r can occur in (91); moreover $r \neq 1$ because one k-support has boundary size $k \geq 3$. Hence $\mathcal{B}_l = \varnothing$ for $l \in \{2, 4\}$.

Lemma 35 (Exact removal of the disconnected sector)**.** *Every nonzero entry (i, j) of $\widehat{A}^{(l)}_{\mathrm{bd}}$ satisfies $i, j \in \mathcal{B}_l$. Consequently*

$$\Pi_l \widehat{A}^{(l)}_{\mathrm{bd}} = \widehat{A}^{(l)}_{\mathrm{bd}} \Pi_l = 0, \qquad \Pi_l \widehat{A}^{(l)} \Pi_l = \Pi_l \widehat{A}^{(l)}_{\mathrm{bc}} \Pi_l. \tag{93}$$

Proof. If I is a boundary-disconnected (i, j)-witness, Lemma 19 gives disjoint support components $K_i, K_j \subseteq I$ with boundaries $\{i\}$ and $\{j\}$. Each has odd size at least three, so $|K_i|, |K_j| \leq l - 3$. Thus $i, j \in \mathcal{B}_l$, proving the claim. □

Lemma 36 (Size of the singleton-boundary set)**.** *For even $l \geq 6$ and $C > 0$,*

$$\mathbb{E}|\mathcal{B}_l| \leq B\sqrt{n} \sum_{\substack{r \text{ odd} \\ 3 \leq r \leq l-3}} C^r \leq B\sqrt{n} \max\{C^3, C^{l-3}\}. \tag{94}$$

Consequently, for every $\tau > 0$,

$$\mathbb{P}[|\mathcal{B}_l| > \tau n] \leq \frac{B \max\{C^3, C^{l-3}\}}{\tau\sqrt{n}}.$$

Proof. A union bound over the certifying index sets and Lemma 14 with $h = 1$ give

$$\mathbb{E}|\mathcal{B}_l| \le \sum_r n\binom{m}{r} B n^{-(kr+1)/2} \le B\sqrt{n}\sum_r C^r.$$

The second bound absorbs the fixed number of terms into B, and Markov's inequality gives the tail estimate. □

Let $W_{l,\mathrm{bc}}(i,j)$ count boundary-connected witnesses and set

$$N_{l,\mathrm{bc}} := \mathbb{E}W_{l,\mathrm{bc}}(i,j), \quad d_{l,\mathrm{bc}} := nN_{l,\mathrm{bc}}, \quad \theta_{l,\mathrm{bc}} := \rho^l d_{l,\mathrm{bc}}.$$

Permutation symmetry gives

$$\mathbb{E}\widehat{A}^{(l)}_{\mathrm{bc}} = \rho^l N_{l,\mathrm{bc}}(\mathbf{1}\mathbf{1}^\top - I). \tag{95}$$

Proposition 45 shows that $N_{l,\mathrm{bc}} = (\kappa^{\mathrm{bc}}_{l,k} + o_n(1))m^l/n^{kl/2+1}$ with $\kappa^{\mathrm{bc}}_{l,k} > 0$.

Theorem 37 (HCE weak recovery for odd k after trimming)**.** *Fix odd $k \ge 3$, even $l \ge 2$, and $\rho \in [0,1]$, and suppose $0 < C \le n^{k/2-1}$ and $d_l \ge 1$. Put $a = (k+1)l + 3$ and $q_n = 2\lceil \log n \rceil$. The conclusions below hold for all sufficiently large n depending on k, l.*

(i) *For every even $q \ge 2$,*

$$\mathcal{M}_q(R^{(l)}_{\mathrm{bc}}) \le n^{1/q}(Bq)^a[d_l^{1/2} + \rho^{l-2}d_l^{1-1/l}]. \tag{96}$$

For every fixed $\gamma > 0$, with probability at least $1 - n^{-\gamma}$,

$$\|R^{(l)}_{\mathrm{bc}}\|_{\mathrm{op}} + \frac{\theta_{l,\mathrm{bc}}}{n} \le \mathfrak{R}^{\mathrm{bc}}_{n,\gamma} := B_\gamma(\log n)^a[d_l^{1/2} + \rho^{l-2}d_l^{1-1/l}] + \frac{\theta_{l,\mathrm{bc}}}{n}. \tag{97}$$

(ii) *If $l \ge 6$, let $\tau \in (0,1)$; for $l \le 4$ set $\tau = 0$. On the event in (97) and $|\mathcal{B}_l| \le \tau n$, suppose*

$$(1-\tau)\theta_{l,\mathrm{bc}} > (3+\varepsilon)\mathfrak{R}^{\mathrm{bc}}_{n,\gamma}. \tag{98}$$

Then the top eigenvalue of $\bar{A}^{(l)}$ is isolated. If $\hat{u}$ is a unit top eigenvector, then

$$\frac{|\langle \hat{u}, x\rangle|}{\sqrt{n}} \ge (1-\tau)^{1/2}(1-\delta^2)^{1/2}, \qquad \delta := \frac{\mathfrak{R}^{\mathrm{bc}}_{n,\gamma}}{(1-\tau)\theta_{l,\mathrm{bc}} - \mathfrak{R}^{\mathrm{bc}}_{n,\gamma}}. \tag{99}$$

The rounded trimmed estimator satisfies

$$\frac{|\langle \hat{x}, x\rangle|}{n} \ge 1 - 4\delta^2 - 2\tau. \tag{100}$$

In particular it achieves weak recovery whenever the right-hand side is positive.

The probability of the joint event is at least

$$1 - n^{-\gamma} - \begin{cases} 0, & l \le 4, \\ \dfrac{B\max\{C^3, C^{l-3}\}}{\tau\sqrt{n}}, & l \ge 6, \end{cases}$$

with $\tau \in (0,1)$ in the second case.

Proof. Theorem 30 is valid for every $k \ge 3$. Repeating the proof of Proposition 32 for $R^{(l)}_{\mathrm{bc}}$ gives (96), and the same Markov argument gives (97).

Work in the gauged frame and write $g := |\mathcal{G}_l|$ and $u_{\mathcal{G}} := \mathbf{1}_{\mathcal{G}_l}/\sqrt{g}$. By Lemma 35 and (95),

$$\Pi_l \widehat{A}^{(l)}\Pi_l = \lambda_{\mathcal{G}} u_{\mathcal{G}} u_{\mathcal{G}}^\top + E_{\mathcal{G}}, \quad \lambda_{\mathcal{G}} := \frac{g}{n}\theta_{l,\mathrm{bc}}, \quad E_{\mathcal{G}} := \Pi_l R^{(l)}_{\mathrm{bc}}\Pi_l - \frac{\theta_{l,\mathrm{bc}}}{n}\Pi_l. \tag{101}$$

Since $\|\Pi_l\|_{\mathrm{op}} \le 1$, $\|E_{\mathcal{G}}\|_{\mathrm{op}} \le \mathfrak{R}^{\mathrm{bc}}_{n,\gamma}$ on the deviation event, while $g/n \ge 1 - \tau$. Weyl and Davis–Kahan therefore give the eigengap under (98) and $\sin\angle(\hat{u}^g, u_{\mathcal{G}}) \le \delta$ for the gauged top eigenvector $\hat{u}^g = D_x\hat{u}$, viewed on $\mathrm{Ran}(\Pi_l)$ and extended by zero on $\mathcal{B}_l$. Hence

$$\frac{|\langle \hat{u}, x\rangle|}{\sqrt{n}} = \sqrt{\frac{g}{n}}\,|\langle \hat{u}^g, u_{\mathcal{G}}\rangle|,$$

which proves (99).

For rounding, orient the gauged top eigenvector toward $u_{\mathcal{G}}$. Its squared distance from $u_{\mathcal{G}}$ is at most $2\delta^2$, so at most $2g\delta^2$ retained coordinates have the wrong sign. The discarded coordinates contribute at worst $-|\mathcal{B}_l|$ to the absolute overlap, which yields (100). The probability statement follows from Lemma 36 and a union bound. □

Corollary 38 (Smooth trade-off for recovery, odd k)**.** *Fix odd $k \geq 3$, even $l \geq 2$, $0 < \rho \leq 1$, $\gamma > 0$, and $\varepsilon > 0$. There is a constant $K^{\mathrm{odd}}_{k,\gamma,\varepsilon} > 0$, independent of l, such that for every fixed even $l \geq 2$, with $a = (k+1)l + 3$, the following holds for all sufficiently large n (where the lower bound on n may depend on l). Let $C = C(n) \leq n^{k/2-1}$ satisfy, if $l \geq 6$, $\max\{C^3, C^{l-3}\} = o(\sqrt{n})$, and suppose*

$$C \geq K^{\mathrm{odd}}_{k,\gamma,\varepsilon} \rho^{-2} \frac{\kappa_{l,k}}{\kappa^{\mathrm{bc}}_{l,k}} \kappa_{l,k}^{-1/l} (\log n)^{(k+1)l+3}. \tag{102}$$

Then, for all sufficiently large n, the trimmed estimator has overlap at least $1 - 4/(2+\varepsilon)^2 - o(1) > 0$ with probability $1 - o(1)$.

Proof. For $l \geq 6$ take $\tau_n = (\max\{C^3, C^{l-3}\}/\sqrt{n})^{1/2} \to 0$; for $l \leq 4$ take $\tau_n = 0$. Proposition 45 gives $\theta_{l,\mathrm{bc}} = (\kappa^{\mathrm{bc}}_{l,k}/\kappa_{l,k} + o_n(1))\theta_l$. Apply Corollary 31 to $R^{(l)}_{\mathrm{bc}}$. The two endpoint comparisons are the same as in the proof of Corollary 5: in the correlated endpoint the factor $\kappa^{\mathrm{bc}}_{l,k}$ cancels against the displayed ratio $\kappa_{l,k}/\kappa^{\mathrm{bc}}_{l,k}$, while the Wigner endpoint is no larger because $\kappa^{\mathrm{bc}}_{l,k} \leq \kappa_{l,k}$. Thus the same finite suprema M_1, M_2 show that one may choose $K^{\mathrm{odd}}_{k,\gamma,\varepsilon}$ independently of l so that $(1 - \tau_n)\theta_{l,\mathrm{bc}}$ exceeds the level-explicit radius by a factor $3 + \varepsilon$ for all sufficiently large n. Lemma 36 gives $\mathbb{P}[|\mathcal{B}_l| > \tau_n n] = o(1)$, and the deviation event fails with probability at most $n^{-\gamma}$. Finally $\delta < 1/(2+\varepsilon)$, so the proof of Theorem 37, with this level-explicit radius, gives the stated overlap. □

For fixed k followed by large even l, Proposition 45 and (135) give

$$\frac{\kappa_{l,k}}{\kappa^{\mathrm{bc}}_{l,k}} \kappa_{l,k}^{-1/l} \sim \gamma_k l^{1-k/2}.$$

Since $K^{\mathrm{odd}}_{k,\gamma,\varepsilon}$ is independent of l, this is the algebraic large-l recovery curve without any hidden level-dependent prefactor. The lower bound on n is still not uniform in growing l.

The side condition holds in the fixed-ρ threshold regime, where C is polylogarithmic. No claim is made here for arbitrary n-dependent ρ.

Computing $\mathcal{B}_l$. The set $\mathcal{B}_l$ can be computed with the same meet-in-the-middle machinery used for HCE:

1. Fix an odd $r \in \{3, \ldots, l-3\}$. Hash every $P_1 \subseteq [m]$ with $|P_1| = (r+1)/2$ under the key ∂P_1.
2. For every $P_2 \subseteq [m]$ with $|P_2| = (r-1)/2$ and every $v \in [n]$, look up the key $\partial P_2 \triangle \{v\}$.
3. Declare $v \in \mathcal{B}_l$ as soon as a lookup succeeds.

No disjointness check between P_1 and P_2 is needed. If $\partial P_1 \triangle \partial P_2 = \{v\}$, then $Q := P_1 \triangle P_2$ has $\partial Q = \{v\}$ and $3 \leq |Q| \leq r \leq l - 3$; the lower bound follows because a singleton-boundary set has odd size and cannot contain one clause when $k \geq 3$. Hence $v \in \mathcal{B}_l$. Conversely, every certificate P of odd size r has a disjoint split of the displayed sizes, so the procedure computes $\mathcal{B}_l$ exactly. Its expected cost is

$$\widetilde{O}_{k,l}(m^{(l-2)/2} + n\, m^{(l-4)/2})$$

time, below the level-l MITM cost. Trimming then only removes the corresponding rows and columns before the eigensolve; equivalently, witnesses with an endpoint in $\mathcal{B}_l$ need not be accumulated.

G.8 Pure-noise refinement

Proposition 39 (Improved pure-noise bound)**.** *If $\rho = 0$ and $C \geq 1$ then, for every even $q \geq 2$,*

$$\mathcal{M}_q\big(R^{(l)}_{\mathrm{bc}}\big) \leq n^{1/q} (Bq)^a C^{l/2}, \qquad \mathcal{M}_q\big(\widetilde{R}^{(l)}\big) \leq n^{1/q} (Bq)^a \sqrt{n}\, C^{l/2}. \tag{103}$$

Consequently, at zero bias the full residual has high-probability upper scale $B(\log n)^a \sqrt{n\, d_l}$.

Proof. All proper noise projections vanish, so the first assertion is the $s = l$ case of Theorem 28. In the full trace expansion a clause occurring an odd number of times has zero noise expectation, so the number c of distinct clauses is at most $ql/2$. Exposing a general boundary-two witness along the trace with its left endpoint old, at most one outside ground variable can have odd degree, so a step with f new clauses introduces at most $(kf+1)/2$ new ground variables; starting from one trace root gives $v \le 1+kc/2+q/2$. By Lemma 26 a pattern weighs at most $B^q n^{1+q/2} C^c \le B^q n^{1+q/2} C^{\,ql/2}$; count patterns and take qth roots. □

H Counting Cycles in Complete Uniform Hypergraphs

We count proper even l-covers (Definition 1) on the complete k-uniform hypergraph $\mathcal{K} = ([n], \binom{[n]}{k})$, assuming kl even. Such identities are classical in the analysis of random CSP refutation and appear e.g. in Coja-Oghlan et al. [2]; Allen et al. [1]; Raghavendra et al. [15]; Wein et al. [20]; Hsieh et al. [12]. We use the unnormalised Kravchuk polynomials

$$K_j(x, M) := \sum_{r=0}^{j} (-1)^r \binom{x}{r} \binom{M-x}{j-r}. \tag{104}$$

Theorem 40. *Fix $k > 2$ and $l \ge 1$. The number of proper even l-covers (up to permutation) in $\mathcal{K}$ is*

$$Z_l^{\mathcal{K}} = E_{n,k}(l) = \frac{1}{2^n} \sum_{w=0}^{n} \binom{n}{w} K_l\!\left(\frac{M - K_k(w,n)}{2},\, M\right), \qquad M = \binom{n}{k}. \tag{105}$$

For fixed k, l, as $n \to \infty$,

$$Z_l^{\mathcal{K}} = \alpha_{k,l}\, n^{kl/2} + o_n\!\left(n^{kl/2}\right), \qquad \alpha_{k,l} = \frac{1}{l!}\left(\frac{1}{k!}\right)^{l/2} \mathbb{E}\left[H_l\!\left(\frac{H_k(Z)}{\sqrt{k!}}\right)\right], \quad Z \sim N(0,1), \tag{106}$$

with $\alpha_{k,l}$ an n-independent constant, equal to zero when l is inadmissible and equal to 1 when $l = 0$. For fixed $k > 2$ and large admissible l, $\alpha_{k,l}$ obeys (32). In the iterated limit $n \to \infty$ followed by admissible $l \to \infty$, this is equivalently

$$Z_l^{\mathcal{K}} \;=\; \big(1 + o_l(1)\big) \frac{e^{-(k-1)/2}}{\sqrt{\pi}} \binom{M}{l} \left(\frac{2}{n}\right)^{kl/2} \Gamma\!\left(\frac{kl+1}{2}\right) \big(1 + o_n(1)\big). \tag{107}$$

Proof. Let $v_S \in \mathbb{F}_2^n$ be the characteristic vector of hyperedge S. Then

$$E_{n,k}(l) = \sum_{\substack{T \subseteq \binom{[n]}{k} \\ |T| = l}} \mathbf{1}\left[\bigoplus_{S \in T} v_S = 0\right] = \frac{1}{2^n} \sum_{y \in \mathbb{F}_2^n} \sum_{\substack{T \\ |T| = l}} (-1)^{y \cdot \sum_{S \in T} v_S}, \tag{108}$$

using Boolean Fourier analysis. The sum over T may be seen as a sum over $T \in \binom{[M]}{l}$; defining $X(y) \in \{0,1\}^M$ by $X_S(y) = y \cdot v_S$, the Kravchuk lemma gives $\sum_T (-1)^{X \cdot v_T} = K_l(|X(y)|, M)$. By symmetry $|X(y)|$ depends only on $|y|$:

$$|X(y)| = \sum_S \frac{1 - (-1)^{X_S(y)}}{2} = \frac{M}{2} - \frac{1}{2} \sum_S (-1)^{y \cdot v_S} = \frac{M - K_k(|y|, n)}{2}. \tag{109}$$

Grouping by Hamming weight yields (105).

For the fixed-l asymptotics, let $W \sim \mathrm{Bin}(n, 1/2)$ and $Z_n := (n - 2W)/\sqrt{n}$. For every fixed j the central limit of the Kravchuk polynomial reads $M^{-j/2} K_j\big(\frac{M - \sqrt{M} z}{2}, M\big) \to H_j(z)/j!$, where H_j is the jth (probabilists') Hermite polynomial. The scaled expressions are fixed-degree polynomials whose coefficients converge, and Z_n converges in moments to $Z \sim N(0,1)$, so applying this first to the inner and then to the outer polynomial gives

$$Z_l^{\mathcal{K}} = \frac{M^{l/2}}{l!} \mathbb{E}\left[H_l\!\left(\frac{H_k(Z)}{\sqrt{k!}}\right)\right] + o_n(n^{kl/2}) = \alpha_{k,l} n^{kl/2} + o_n(n^{kl/2}),$$

using $M = \binom{n}{k} \sim n^k/k!$, which is (106). Since $\alpha_{k,l}$ is a fixed multiple of $\mathbb{E}[H_l(H_k(Z)/\sqrt{k!})]$, it vanishes exactly when no proper even l-cover exists, i.e. when l is inadmissible.

It remains to estimate $\alpha_{k,l}$ for large l. View the kl factors in $\mathbb{E}[Z^{kl}]$ as l boxes of size k. By Wick's formula $(kl-1)!!$ counts all pairings of these factors. Put $A_{k,l} := \mathbb{E}[H_k(Z)^l]$; this counts those with no pair inside a box. The number of within-box pairs in a uniform pairing converges to a Poisson variable of mean $(k-1)/2$ — its rth factorial moment tends to $((k-1)/2)^r$ — whence $A_{k,l} = \big(e^{-(k-1)/2} + o_l(1)\big)(kl-1)!!$.

Let $P^\star_{k,l}$ count pairings with neither an internal box pair nor two boxes paired completely to one another. Intersections of complete-pairing events correspond exactly to matchings of disjoint pairs of boxes, so inclusion–exclusion gives

$$P^\star_{k,l} = \sum_{r=0}^{\lfloor l/2 \rfloor} (-1)^r \frac{l!}{2^r r!(l-2r)!} (k!)^r A_{k,l-2r}. \tag{110}$$

The probabilists' identity $H_l(y) = l! \sum_r (-1)^r y^{l-2r}/(2^r r!(l-2r)!)$ therefore yields exactly

$$\alpha_{k,l} = \frac{P^\star_{k,l}}{l!(k!)^l}.$$

Moreover, a union bound over the $\binom{l}{2}$ possible completely paired box pairs gives

$$0 \le A_{k,l} - P^\star_{k,l} \le \binom{l}{2} k!\, A_{k,l-2}, \qquad \frac{A_{k,l} - P^\star_{k,l}}{A_{k,l}} = O_k(l^{2-k}) = o_l(1).$$

Thus the outer Hermite polynomial changes the pairing count only by this relative error. Inserting $(kl-1)!! = (kl)!/(2^{kl/2}(kl/2)!)$ and Stirling's formula into (106) gives

$$\alpha_{k,l} = \big(L_k + o_l(1)\big) l^{-1/2} \left(\frac{k^{k/2} e^{1-k/2}}{k!}\, l^{k/2-1} \right)^l, \qquad L_k = \frac{e^{-(k-1)/2}}{\sqrt{\pi}},$$

which is (32). Combining this formula with Stirling's identity for $\Gamma((kl+1)/2)$ yields (107). The factor $e^{-(k-1)/2}$ is essential: omitting the Hermite exclusion terms would, for example, give the wrong fixed-level constant already at $(k,l) = (4,3)$. □

I Expected Cycles in Random Uniform Hypergraphs

Now sample m hyperedges uniformly from $\binom{[n]}{k}$ and let Z_l be the number of proper even l-covers in the resulting random hypergraph. Its mean follows from Proposition 9 together with Theorem 40.

Theorem 41. *For fixed $k > 2$ and admissible l, with $m = Cn^{k/2}$,*

$$\mathbb{E}_{\mathcal{H}}[Z_l] = \binom{m}{l} \frac{Z_l^{\mathcal{K}}}{\binom{M}{l}} \;=\; \big(\alpha_{k,l} + o_n(1)\big)\,(C\,k!)^l \;=\; \mu_l + o_n(1), \tag{111}$$

so that, for large admissible l,

$$\mu_l = \big(L_k + o_l(1)\big)\, l^{-1/2} \left(\frac{m\, l^{k/2-1}}{\gamma_k\, n^{k/2}} \right)^l, \qquad \gamma_k = k^{-k/2} e^{k/2-1}. \tag{112}$$

Proof. The first equality is Proposition 9. For the second, l is fixed and $m \ll M$, so $\binom{m}{l}/\binom{M}{l} = (1+o_n(1))m^l/M^l$; inserting $M = \binom{n}{k} = (1+o_n(1))n^k/k!$, $m = Cn^{k/2}$ and $Z_l^{\mathcal{K}} = (\alpha_{k,l} + o_n(1))n^{kl/2}$ gives $\mathbb{E}_{\mathcal{H}}[Z_l] = (\alpha_{k,l} + o_n(1))(Ck!)^l = \mu_l + o_n(1)$, which is the definition of μ_l used in (14). Equation (112) then follows from (32) and $a_k = 1/(\gamma_k k!)$. □

In particular $\mu_l \rho^{2l} \gg 1$ is equivalent, to leading exponential order, to $m \gg \gamma_k \rho^{-2} n (n/l)^{k/2-1}$, which is (8); the exact form of the resulting condition, with its constant, is (15).

J COUNTING THE WITNESSES N_l

It is useful to separate the purely combinatorial count from the random sampling of clauses. Fix a pair $\{i, j\}$ and let

$$T_l(i,j) := \#\left\{\{D_1,\dots,D_l\} \subseteq \binom{[n]}{k} : D_r \text{ distinct},\ \triangle_{r=1}^{l} D_r = \{i,j\}\right\}. \tag{113}$$

By symmetry this does not depend on $\{i, j\}$, so we write T_l. In the Poissonised inclusion model, where every k-set is included independently with probability $p = m/\binom{n}{k}$, the expected number of witnesses for a fixed pair is exactly

$$N_l = T_l\, p^l = T_l \left(\frac{m}{\binom{n}{k}}\right)^l. \tag{114}$$

For the i.i.d. clause model used in the main text, clause indices are distinct but their supports need not be. The distinct-support contribution is $(m)_l T_l/\binom{n}{k}^l$, while repeated-support tuples must be counted separately. They can contribute at the same power $n^{lk/2-1}$; their relative weight is small only in the subsequent large-l comparison, not in the $n \to \infty$ limit at fixed l.

J.1 THE $\mathbb{F}_2$ PARITY CONSTRAINT

Identify every support D_r with its indicator vector in $\mathbb{F}_2^n$. The witness condition is $\sum_{r=1}^{l} \mathbf{1}_{D_r} = e_i + e_j \pmod 2$: every coordinate outside $\{i, j\}$ has even multiplicity, while i and j have odd multiplicity. Taking cardinalities modulo 2 gives the necessary condition $2 = |\{i,j\}| \equiv \sum_r |D_r| = lk \pmod 2$. Hence $T_l = 0$ unless lk is even: for odd k only even l contributes, while for even k every $l \geq 2$ is admissible.

J.2 THE EXACT COUNT FOR $l = 2$

For two supports, $D_1 \triangle D_2 = \{i, j\}$ if and only if the k-sets agree outside $\{i, j\}$ and differ on both distinguished coordinates. Thus there is a unique $(k-1)$-set $R \subseteq [n] \setminus \{i, j\}$ with $\{D_1, D_2\} = \{R \cup \{i\}, R \cup \{j\}\}$, and for $n \geq k + 1$,

$$T_2 = \binom{n-2}{k-1}. \tag{115}$$

Consequently, in the i.i.d. clause model,

$$N_2 = \mathbb{E}W_2(i,j) = \binom{m}{2}\frac{2T_2}{\binom{n}{k}^2} = \frac{(m)_2}{\binom{n}{k}^2}\binom{n-2}{k-1}, \tag{116}$$

and to leading order

$$N_2 = \left(1 + O\left(\frac{k^2}{n}\right) + O\left(\frac{1}{m}\right)\right)\frac{(k!)^2}{(k-1)!}\frac{m^2}{n^{k+1}} = \left(1 + O\left(\frac{k^2}{n}\right) + O\left(\frac{1}{m}\right)\right) k\,k!\,\frac{m^2}{n^{k+1}}. \tag{117}$$

J.3 THE EXACT COUNT FOR $l = 3$

The parity constraint forces k even. Assume $k \geq 4$ even and $n \geq 3k/2 + 1$. Given two supports, the third is determined by

$$D_3 = D_1 \triangle D_2 \triangle \{i, j\}, \tag{118}$$

so an ordered triple is specified by (D_1, D_2) subject to $|D_3| = k$; for $k \geq 4$ the three supports are then automatically distinct. Let $s := |D_1 \setminus D_2| = |D_2 \setminus D_1|$ and $t := |\{i, j\} \cap (D_1 \triangle D_2)| \in \{0, 1, 2\}$. Since $|D_1 \triangle D_2| = 2s$, (118) gives $|D_3| = 2s + 2 - 2t$, so $|D_3| = k$ is equivalent to $s = s_t := \frac{k}{2} - 1 + t$. Classify coordinates by membership,

$$M_\cap = D_1 \cap D_2, \quad A = D_1 \setminus D_2, \quad B = D_2 \setminus D_1, \quad O = [n] \setminus (D_1 \cup D_2), \tag{119}$$

with $|M_\cap| = k - s_t$, $|A| = |B| = s_t$ and $|O| = n - k - s_t$. If $\vec{m} = (m_M, m_A, m_B, m_O)$ records where i, j fall, then $m_M + m_A + m_B + m_O = 2$ and $m_A + m_B = t$. Filling the remaining slots with the other $n - 2$ coordinates gives the exact count

$$T_3 = \frac{1}{6}\sum_{t=0}^{2} \sum_{\substack{m_M,m_A,m_B,m_O\geq 0 \\ m_M+m_A+m_B+m_O=2,\ m_A+m_B=t}} \frac{2!}{m_M!\, m_A!\, m_B!\, m_O!} \times \frac{(n-2)!}{(k-s_t-m_M)!\,(s_t-m_A)!\,(s_t-m_B)!\,(n-k-s_t-m_O)!}, \tag{120}$$

terms with negative factorial arguments being omitted. The first factor in the summand counts the ways of choosing $\vec{m}$ summing to 2; the second counts the ways of filling the remaining slots with the $n - 2$ remaining variables. The factor $1/6$ converts ordered triples to unordered families. Hence $N_3 = (m)_3 T_3 / \binom{n}{k}^3$ for even $k \geq 4$, and $N_3 = 0$ for odd k. The leading part of (120) is

$$T_3 = \left(\kappa_3 + O\left(\frac{k^2}{n}\right)\right) n^{3k/2-1}, \qquad \kappa_3 = \frac{1}{2(\frac{k}{2}+1)!(\frac{k}{2}-1)!^2} + \frac{1}{(\frac{k}{2})!^2(\frac{k}{2}-1)!}, \tag{121}$$

so that $N_3 = \big(\kappa_3 + O(k^2/n) + O(1/m)\big)(k!)^3 m^3 / n^{3k/2+1}$.

J.4 General fixed l: multigraph counting

We now count the leading ordered i.i.d. support tuples entering N_l, for any fixed l with lk even. Support values may repeat, although the l sampled clause indices are distinct. In a collapsing tuple, a generic coordinate $v \notin \{i, j\}$ must appear in an even number of positions, while i and j each appear an odd number. The largest power of n comes from using as many distinct generic coordinates as possible: every generic coordinate appears exactly twice and each of i, j exactly once. Any coordinate with multiplicity ≥ 4, or a distinguished coordinate with multiplicity ≥ 3, reduces the number of free generic coordinates and is lower order.

Label the tuple positions $1, \ldots, l$. A generic coordinate appearing in supports s and t is an edge $\{s, t\}$ of a loopless multigraph on $[l]$; let $d_{st} \geq 0$ be its multiplicity. Let α be the support containing i and β the support containing j, allowing $\alpha = \beta$. Since every support has size k,

$$\sum_{t \neq v} d_{vt} + \mathbf{1}[v = \alpha] + \mathbf{1}[v = \beta] = k, \qquad v = 1, \ldots, l, \tag{122}$$

and the number of generic doubled coordinates is forced to be

$$D := \sum_{1 \leq s < t \leq l} d_{st} = \frac{lk - 2}{2} = \frac{lk}{2} - 1. \tag{123}$$

For fixed (α, β) and feasible (d_{st}), choosing and assigning the D generic coordinates contributes $(n-2)_D / \prod_{s<t} d_{st}! = n^D \prod_{s<t}(d_{st}!)^{-1}(1 + O(D^2/n))$. Let $Q_l(i, j)$ be the number of ordered support tuples with boundary $\{i, j\}$. Summing over labelled configurations gives

$$Q_l(i,j) = l!\,\kappa_l\, n^{lk/2-1}\left(1 + O_{k,l}\left(\frac{1}{n}\right)\right), \qquad \kappa_l = \frac{1}{l!}\sum_{\alpha,\beta\in[l]} \sum_{\substack{(d_{st})_{s<t}\in\mathbb{Z}_{\geq 0}^{\binom{l}{2}} \\ \text{satisfying (122)}}} \prod_{1\leq s<t\leq l} \frac{1}{d_{st}!}. \tag{124}$$

This is valid for fixed l with $lk = o(\sqrt{n})$, the relative error being controlled by the birthday-type collision term $O((lk)^2/n)$. The sum does not require the incidence rows of two tuple positions to differ, so it already includes repeated support values. For example, at $l = 4$ the choice $d_{12} = k$, $d_{34} = k - 1$, $\alpha = 3$, $\beta = 4$ is precisely the leading configuration with the first two supports equal.

The checks for $l = 2, 3$ recover the formulas above. For $l = 2$ the single edge has multiplicity $d_{12} = k - 1$ and the two choices $(\alpha, \beta) = (1, 2), (2, 1)$ give $\kappa_2 = 1/(k-1)!$. For $l = 3$, writing $g_v = \mathbf{1}[v = \alpha] + \mathbf{1}[v = \beta]$, the unique solution is $d_{12} = \frac{k}{2} - 1 + g_3$, $d_{13} = \frac{k}{2} - 1 + g_2$, $d_{23} = \frac{k}{2} - 1 + g_1$, which yields κ_3 in (121).

J.5 THE SAMPLED-WITNESS CONSTANT

For a fixed clause-index set, its ordered supports are independent uniform k-sets, so the unrestricted tuple count above gives

$$N_l = \mathbb{E}W_l(i,j) = \binom{m}{l}\frac{Q_l(i,j)}{\binom{n}{k}^l}. \tag{125}$$

Substituting (124), $\binom{m}{l} = m^l(1+O_l(1/m))/l!$, and $\binom{n}{k} = n^k(1+O_k(1/n))/k!$ yields

$$N_l = \kappa_{l,k}\frac{m^l}{n^{lk/2+1}}\left(1+O_{k,l}\left(\frac{1}{n}+\frac{1}{m}\right)\right), \qquad \kappa_{l,k} := (k!)^l\kappa_l = \frac{(k!)^l}{l!}\sum_{\alpha,\beta\in[l]}\sum_{\substack{(d_{st})_{s<t}\in\mathbb{Z}_{\geq 0}^{\binom{l}{2}}\\ \text{satisfying (122)}}}\prod_{1\leq s<t\leq l}\frac{1}{d_{st}!}. \tag{126}$$

Thus repeated supports enter the full i.i.d. coefficient exactly once, through the same unrestricted multigraph sum. The factor $(k!)^l$ comes from the sampling normalisation $\binom{n}{k}^{-l} \sim (k!)^l n^{-kl}$.

Lemma 42 (Finite-n witness calibration)**.** *Fix admissible k, l. Uniformly over integers $m \geq l$, if $m = Cn^{k/2}$, then*

$$d_l = \kappa_{l,k}C^l\left(1+O_{k,l}\left(\frac{1}{n}+\frac{1}{m}\right)\right). \tag{127}$$

Consequently there are $n_0(k,l)$ and $m_0(k,l)$ such that, whenever $n \geq n_0(k,l)$ and $m \geq m_0(k,l)$,

$$\frac{1}{2}\kappa_{l,k}C^l \leq d_l \leq \frac{3}{2}\kappa_{l,k}C^l.$$

Proof. The displayed estimate for $Q_l(i,j)$ counts all leading incidence patterns. Every remaining pattern has at most $lk/2-2$ free generic ground-variable labels; there are only finitely many such patterns at fixed k, l, so their total is $O_{k,l}(n^{lk/2-2})$. Now use

$$\binom{m}{l} = \frac{m^l}{l!}\left(1+O_l(m^{-1})\right), \qquad \binom{n}{k}^{-l} = (k!)^l n^{-kl}\left(1+O_{k,l}(n^{-1})\right)$$

in $N_l = \binom{m}{l}Q_l(i,j)/\binom{n}{k}^l$, and multiply by n. The two-sided bound follows after increasing $n_0(k,l)$ so that the relative error is at most $1/2$. □

J.6 ASYMPTOTICS OF $\kappa_{l,k}$

We evaluate (126) for large l and fixed k. The multigraph sum has a generating-function form: let

$$\Phi(x_1,\dots,x_l) = \exp\left(\sum_{1\leq s<t\leq l}x_sx_t\right), \qquad S(a) := \left[\prod_{v=1}^{l}x_v^{a_v}\right]\Phi(x_1,\dots,x_l). \tag{128}$$

This coefficient is exactly $S(a) = \sum_{(d_{st}):\sum_{t\neq v}d_{vt}=a_v}\prod_{s<t}(d_{st}!)^{-1}$, because $\exp(x_sx_t)$ records how many generic variables are shared by supports s and t. For a fixed ordered pair (α,β) the degree sequence is $a_v = k - \mathbf{1}[v=\alpha] - \mathbf{1}[v=\beta]$. There are two symmetry classes: if $\alpha = \beta$ there are l choices and the degree sequence is $(k-2,k,\dots,k)$ up to permutation; if $\alpha \neq \beta$ there are $l(l-1)$ choices and it is $(k-1,k-1,k,\dots,k)$. Therefore

$$\kappa_l = \frac{S_2}{(l-1)!} + \frac{S_{11}}{(l-2)!}, \qquad S_2 = S(k-2,k,\dots,k), \quad S_{11} = S(k-1,k-1,k,\dots,k). \tag{129}$$

To extract the exponential-in-l scale, compare $S(a)$ with the configuration model in which loops are temporarily allowed; replacing (128) by $\widetilde{\Phi} = \exp\left(\frac{1}{2}(\sum_v x_v)^2\right)$ and writing $\sum_v a_v = 2D$,

$$\widetilde{S}(a) = \left[\prod_v x_v^{a_v}\right]\widetilde{\Phi} = \frac{(2D)!}{2^D D!\prod_v a_v!} = \frac{(2D-1)!!}{\prod_v a_v!}. \tag{130}$$

Lemma 43 (Loopless configuration coefficient)**.** *Fix $k \geq 3$. For either degree sequence $a = (k-2, k, \dots, k)$ or $a = (k-1, k-1, k, \dots, k)$,*

$$S(a) = \Big(e^{-(k-1)/2} + o_l(1)\Big)\, \widetilde{S}(a). \tag{131}$$

In particular, $c_k \widetilde{S}(a) \leq S(a) \leq \widetilde{S}(a)$ for all sufficiently large l, for some $c_k > 0$.

Proof. Attach a_v labelled half-edges to each vertex v and choose a uniform perfect matching of the $2D = kl - 2$ half-edges. Dividing the number of all matchings by $\prod_v a_v!$ gives $\widetilde{S}(a)$, while restricting to matchings with no pair internal to one vertex gives $S(a)$. Thus $S(a)/\widetilde{S}(a) = \mathbb{P}[L = 0]$, where L is the number of internal pairs.

For every fixed r, the contribution to $\mathbb{E}(L)_r$ from choices of r vertex-disjoint internal half-edge pairs is

$$\frac{r!\, e_r\big(\big(\binom{a_v}{2}\big)_{v\in[l]}\big)}{(2D-1)(2D-3)\cdots(2D-2r+1)},$$

where e_r is the rth elementary symmetric polynomial. Choices sharing a half-edge contribute $O_k(l^{r-1})/\Theta_k(l^r) = o_l(1)$. Since all but at most two degrees equal k, the display tends to $((k-1)/2)^r$. Hence the factorial moments of L converge to those of $\mathrm{Poisson}((k-1)/2)$, so $\mathbb{P}[L = 0] \to e^{-(k-1)/2}$. □

For both degree sequences in (129), $2D = lk - 2$ and

$$\widetilde{S}_2 = \frac{(2D-1)!!}{(k-2)!\,(k!)^{l-1}}, \qquad \widetilde{S}_{11} = \frac{(2D-1)!!}{(k-1)!^2 (k!)^{l-2}}. \tag{132}$$

Their ratio is $(k-1)/k = \Theta_k(1)$, while the prefactors in (129) differ by $(l-2)!/(l-1)! = 1/(l-1)$. Thus the diagonal contribution is smaller by $\Theta_k(1/l)$ and

$$\kappa_l^{1/l} \sim \left(\frac{(2D-1)!!}{(l-2)!\,(k-1)!^2 (k!)^{l-2}}\right)^{1/l}. \tag{133}$$

Using Stirling's formula, $((2D-1)!!)^{1/l} \sim (lk/e)^{k/2}$ and $((l-2)!)^{1/l} \sim l/e$, while the fixed factors $(k-1)!^2$ vanish under the l-th root. Hence

$$\kappa_l^{1/l} \sim \frac{(lk/e)^{k/2}}{(l/e)\,k!} = \frac{k^{k/2} e^{1-k/2}}{k!}\, l^{k/2-1}, \tag{134}$$

and multiplying by the sampling factor gives the main-text constant

$$\kappa_{l,k}^{1/l} = k!\,\kappa_l^{1/l} \sim \frac{(lk/e)^{k/2}}{l/e} = \frac{e}{l}\left(\frac{kl}{e}\right)^{k/2}. \tag{135}$$

J.7 Boundary-connected witnesses for odd k

For a feasible configuration $(\alpha, \beta, (d_{st}))$ in (126), let G_d be the multigraph on $[l]$ with d_{st} edges between s and t. In a leading tuple, two support positions intersect exactly when they are adjacent in G_d, so the tuple is boundary-connected exactly when α and β lie in the same connected component of G_d (with $\alpha = \beta$ allowed). Define

$$\kappa_{l,k}^{\mathrm{bc}} := \frac{(k!)^l}{l!} \sum_{\alpha,\beta\in[l]} \sum_{\substack{(d_{st}) \text{ satisfying (122)} \\ \alpha,\beta \text{ connected in } G_d}} \prod_{s<t} \frac{1}{d_{st}!}. \tag{136}$$

This is a restriction of the same unrestricted ordered-tuple sum that defines $\kappa_{l,k}$; repeated support values are therefore included exactly once in both constants.

Lemma 44 (Separated boundary vertices are rare)**.** *Fix odd $k \geq 3$ and even l. In the weighted multigraph sum defining $\kappa_{l,k}$, the fraction of configurations with $\alpha \neq \beta$ in different connected components is $O_k(l^{-(3k-5)/2})$.*

Proof. Fix $\alpha \neq \beta$. The degree sequence is $a_\alpha = a_\beta = k-1$ and $a_v = k$ otherwise, with total degree $2D = kl - 2$. Attach a_v labelled half-edges to vertex v and choose a uniform perfect matching. Conditional on the no-loop event L, the resulting multiplicity vector has weight proportional to $\prod_{s<t}(d_{st}!)^{-1}$. Lemma 43 gives $\mathbb{P}(\mathsf{L}) \geq c_k > 0$ uniformly in l.

If α and β are disconnected, the component vertex set V of α is closed under the pairing, contains α but not β, and has odd size r with $3 \leq r \leq l-3$. For fixed such V, put $2a_r = kr - 1$. Then

$$\mathbb{P}(V \text{ closed}) = \frac{(2a_r-1)!!(2D-2a_r-1)!!}{(2D-1)!!} = \frac{\binom{D}{a_r}}{\binom{2D}{2a_r}} \leq \binom{D}{a_r}^{-1}.$$

There are $\binom{l-2}{r-1}$ choices for V. Pairing V with its complement and applying a union bound gives

$$\mathbb{P}(\alpha \not\leftrightarrow \beta, \mathsf{L}) \leq 2 \sum_{\substack{3 \leq r \leq l/2 \\ r \text{ odd}}} \binom{l-2}{r-1} \binom{D}{a_r}^{-1}. \tag{137}$$

Vandermonde gives

$$\binom{D}{a_r} \geq \binom{l-2}{r-1}\binom{D'}{a'_r}, \quad D' = \frac{(k-2)l+2}{2}, \quad a'_r = \frac{(k-2)r+1}{2}.$$

The $r = 3$ term is $O_k(l^{-(3k-5)/2})$. For $r \geq 5$, convexity of $y \log(y/D')$ on $[a'_5, D'/2]$ bounds the remaining sum by $O_k(l^{1-(5k-9)/2}) + l2^{-l/4}$, which is of the same or smaller order. Dividing (137) by $\mathbb{P}(\mathsf{L}) \geq c_k$ proves the claim. □

Proposition 45 (Boundary-connected witness constant)**.** *Fix odd $k \geq 3$ and even $l \geq 2$.*

(a)
$$N_{l,\mathrm{bc}} = (\kappa^{\mathrm{bc}}_{l,k} + o_n(1)) \frac{m^l}{n^{kl/2+1}}, \qquad 0 < \kappa^{\mathrm{bc}}_{l,k} \leq \kappa_{l,k}.$$
For $l \in \{2, 4\}$ equality holds between the two constants.

(b) *As $l \to \infty$ through even values,*
$$\frac{\kappa^{\mathrm{bc}}_{l,k}}{\kappa_{l,k}} = 1 - O_k\left(l^{-(3k-5)/2}\right).$$
Consequently $(\kappa^{\mathrm{bc}}_{l,k})^{1/l} \sim \kappa^{1/l}_{l,k}$.

Proof. Part (a) follows by repeating the leading-pattern count of (124) with the connectivity restriction in (136). At $l = 2$, positivity follows from the exact count (115). For $l \geq 4$, it follows from the feasible $\alpha = \beta$ class, in which both boundary variables lie in one support position. Equality for $l \in \{2, 4\}$ follows from Proposition 20.

For part (b), configurations with $\alpha = \beta$ are always boundary-connected. Lemma 44, applied to every $\alpha \neq \beta$ class, shows that the complementary weighted sum is at most $O_k(l^{-(3k-5)/2})\kappa_{l,k}$. This proves the ratio estimate, and the lth-root conclusion follows from (135). □

J.8 MITM ANALYSIS

A direct implementation of (9) enumerates all l-tuples of clauses and has runtime $O(m^l) = O(n^{kl/2})$ when $m = Cn^{k/2}$. This naive cost is reduced by the standard meet-in-the-middle (MITM) strategy shown in Figure 1. Put $r = \lfloor l/2 \rfloor$ and $s = \lceil l/2 \rceil$, and aggregate rather than list the hash buckets:

$$F_s(u) := \sum_{\substack{\mathcal{S} \in \binom{[m]}{s} \\ \triangle(\mathcal{S}) = u}} X_{\mathcal{S}}, \qquad u \in \mathbb{F}_2^n.$$

The raw MITM matrix is then

$$G^{(l)}_{ij} := \sum_{\mathcal{S}' \in \binom{[m]}{r}} X_{\mathcal{S}'} F_s(\triangle(\mathcal{S}') \triangle \{i, j\}) \quad (i \neq j),$$

which takes $\widetilde{O}(m^s + n^2 m^r)$ time to construct regardless of hash collisions. See detailed discussion of this object in Appendix J.8. Terms with $\mathcal{S} \cap \mathcal{S}' \neq \varnothing$ are removed algebraically. If $t = |\mathcal{S} \cap \mathcal{S}'|$, the common labels square to one and the remaining $l - 2t$ labels form a lower-level witness. Counting the choices of the common labels and the split of the remaining labels gives the exact recurrence

$$G^{(l)} = \binom{l}{s} A^{(l)} + \sum_{t=1}^{r} \binom{m-l+2t}{t} \frac{(l-2t)!}{(s-t)!(r-t)!} A^{(l-2t)}, \qquad A^{(0)} := 0. \tag{138}$$

Here $A^{(1)}$ is defined by (9): it vanishes for $k \geq 3$, while for $k = 2$ it records the observed clauses with support $\{i, j\}$. Computing the levels $l - 2, l - 4, \ldots$ recursively therefore recovers the exact matrix (9) in deterministic time $\widetilde{O}_{k,l}(n^3 + m^s + n^2 m^r)$ and memory $\widetilde{O}_{k,l}(n^2 + m^s)$, without enumerating the possibly $\Theta(m^l)$ individual witnesses. The resulting complexities are collected in Table 1.

The enumeration task underpinning HCE — given a list of m vectors in $\mathbb{F}_2^n$, find size-l subsets with a prescribed total XOR — is an instance of the l-XOR problem, a relative of generalized birthday and subset-sum problems [19]. The MITM procedure above is a standard exact generic baseline for this enumeration problem. We do not claim that its exponent is optimal for all-witness aggregation on random planted instances, where additional structure may be exploitable.

The structural difference is that HCE always performs spectral recovery on an $n \times n$ matrix and pays for clause-product enumeration, whereas Kikuchi avoids the enumeration but uses $\widetilde{O}((1 + C)n^\ell)$ memory for a lifted operator on a $\binom{n}{\ell}$-dimensional space. For $k = 4$ and matched even levels $l = \ell$, the leading HCE runtime exponent is one factor of n larger (n^{l+2} versus $n^{\ell+1}$), so we do not claim an asymptotic runtime improvement. The representations nevertheless differ substantially, and this matters at the tested sizes. In our experiments HCE has substantially lower memory usage and runtime; for example at $n = 50$, $k = 4$, level-6 Kikuchi becomes computationally prohibitive (roughly 5.5 hours per instance), whereas a level-6 HCE run with eight probe threads completes the densest tested instance ($C = 0.25$) in roughly 300 seconds. The two timings use different thread budgets (Appendix B) and are indicative rather than resource-matched.

K Related Work

Sum of squares. The sum-of-squares (SoS) hierarchy is a common benchmark for planted inference and random-CSP refutation. In tensor PCA, low-degree SoS relaxations and spectral algorithms derived from their moment matrices attain the conjectured polynomial-time recovery scale [9; 10; 11]. For random k-XOR, the spectral/SoS method of Raghavendra et al. [15] gives strong-refutation certificates, while Kothari et al. [13] prove complementary SoS degree lower bounds. Kikuchi was introduced as a simpler spectral hierarchy attaining the same degree/runtime–signal trade-off as SoS [20], and Kikuchi matrices also underlie SoS certificates for random CSPs [7]. The normalized analysis of Schmidhuber & Hastings [17] makes this trade-off log-free for planted noisy k-XOR detection and weak recovery while retaining degree-$O(\ell)$ SoS refutation certificates. HCE is a low-degree spectral statistic in this lineage, but is not itself an SoS relaxation.

Tensor unfolding and spectral baselines. A simpler and historically important approach is *tensor unfolding* (matricisation): represent the observations as a sparse order-k tensor T supported on the observed clauses with values b_a, so that $\mathbb{E}[T] \propto x^{\otimes k}$, and partition the k modes into two groups of size $k/2$ to obtain a matrix whose leading singular vectors estimate tensor powers of the signal [16; 9; 10]. Spectral separation occurs at $m \gtrsim \widetilde{\Theta}(n^{k/2})$, matching the polynomial-time threshold for constant-degree methods; for odd k one uses an unbalanced unfolding, and related square constructions give the corresponding spectral relaxations. This baseline is simple and efficient but provides *no* smooth sample–time trade-off, and it treats the data as noisy tensor entries without exploiting the algebraic cancellation structure of XOR constraints, exactly the structure that HCE uses.

Trace methods, even covers and low-degree polynomials. The matrix $A^{(l)}$ is closely related to the l-th step of the trace method for refuting random CSPs [1; 15], whose spectral antecedents go back to Goerdt & Krivelevich [6] and Friedman et al. [5], and to the "square" spectral construction

of Hopkins et al. [10] for tensor PCA. Proposition 7 locates HCE within this lineage as a low-degree spectral estimator in the sense of Hopkins [8] and Kunisky et al. [14]. The witness counts of Appendix J are an explicit version of the $\mathbb{F}_2$-cycle counts of Allen et al. [1] and Raghavendra et al. [15]; even-cover counting also underlies the hypergraph Moore bound [12] and Feige-type refutation [4; 7]. Closest to our work, Schmidhuber & Hastings [17] show that the Kikuchi hierarchy is *sharp* for random k-XOR, without polylogarithmic losses. Accordingly, our novelty claim is not a new statistical regime, a proved equivalence with one Kikuchi slice, or uniform runtime superiority; it is the exact final-boundary compression, its cycle interpretation, the $n \times n$ spectral endpoint at every level, the explicit witness-count constants, and the empirical memory frontier.