Posterior sampling for the sparse two-community stochastic block model

30 October 2026

Abstract

We study posterior sampling in the sparse two-community stochastic block model with independent uniform labels and fixed known rational assortative parameters $a > b > 0$. We prove that a polynomial-time sampler exists whenever $(a-b)^2/[2(a+b)]$ exceeds an absolute constant. The expected total variation error tends to zero under the original graph distribution, and the algorithm has deterministic polynomial bounds on its bit complexity and random-bit use on every input graph. The proof combines two constructions. At small contrast, a low-rank Gaussian representation produces a tempered initial law, which is then cooled using quantitative estimates for the disagreement graph of two replicas. At larger contrast, local correction of a recovered labeling produces a core with uniform energy margins; component updates sample its conditional uncertainty, and a centered rejection step restores the exact quadratic interaction contributed by nonedges. A single bounded-excess decomposition serves both initialization and cooling. We also give a locally uniform version and an explicit parameter-precision bound for fixed computable real parameters.

Note. This work was developed using GPT-6 and Pharos, an automated mathematical reasoning system built on Rethlas [JGJ+26], prompted by Zhangsong Li without any mathematical input. We thank Bin Dong, Guoxiong Gao, Jiedong Jiang, Shurui Liu, Zeming Sun, and Bin Wu for developing and refining Pharos, and thank Jihao Liu, Bohan Fang, Jingjun Han, Guchuan Li, Ruochuan Liu, Yujie Luo, Zhenfu Wang, and Yijun Yuan for feedback on the system. Language-model assistance was also used in subsequent proof simplification, exposition, and reference checking.

Contents

1 Introduction **2**
- 1.1 Main result . . . 3
- 1.2 Related work . . . 3
- 1.3 Proof overview . . . 4
- 1.4 Notation and external inputs . . . 5

2 Posterior concentration and initialization **6**

3 Sparse reference measures **8**
- 3.1 Covariance and variance estimates . . . 9
- 3.2 Sampling with two global counts . . . 11

4 Graph decompositions under the replica law **12**
- 4.1 Canonical excision and spectral regularization . . . 12
- 4.2 The conditional law of a disagreement graph . . . 16

5 A tempered initialization **17**

6 Cooling the tempered law **19**

7 A recovered phase and its fluctuations **22**
7.1 A deterministic core construction 23
7.2 Joint fluctuations and component tails 25

8 Component updates and the exact posterior **28**
8.1 A general initialized component sampler 28
8.2 Sampling the ball law 30
8.3 Restoring the quadratic nonedge interaction 31

9 An absolute threshold and real parameters **33**

A Finite-precision implementation **36**

1 Introduction

The stochastic block model is a basic model of a network with latent communities. Much of its algorithmic theory concerns recovery: given the graph, estimate the hidden labels with as few mistakes as possible. Posterior sampling asks for the full conditional distribution of those labels. A sample should reproduce the uncertainty left by the observed graph, including dependence between ambiguous vertices and symmetry between equivalent community names. These requirements remain meaningful when the signal is large enough to recover most labels accurately.

We consider two communities. The labels $X_1, \dots, X_n$ are independent uniform signs. Conditional on X, the unordered edges of a simple graph G are independent, with

$$\mathbb{P}(ij \in E(G) \mid X) = \begin{cases} a/n, & X_i = X_j, \\ b/n, & X_i \neq X_j, \end{cases} \qquad a > b > 0, \quad n > a.$$

The parameters are fixed as $n \to \infty$ and are known to the algorithm. Write

$$d = \frac{a+b}{2}, \qquad \theta = \frac{a-b}{a+b}, \qquad \kappa = d\theta^2 = \frac{(a-b)^2}{2(a+b)}, \qquad \mu_g = \mathcal{L}(X \mid G = g).$$

Thus d is the average-degree parameter and $\theta \in (0,1)$ measures the contrast between the communities. We allow every fixed $a > b > 0$ above a sufficiently large absolute threshold on κ.

Two features of the posterior are central to the argument. First, it is invariant under the global sign change $s \mapsto -s$. Second, the absence of an edge carries information. Define

$$M(s) = \sum_{i=1}^{n} s_i, \qquad F_g(s) = \sum_{ij \in E(g)} s_i s_j,$$

and

$$J_n = \frac{1}{2} \log \frac{a(n-b)}{b(n-a)}, \qquad \gamma_n = \frac{1}{2} \log \frac{n-a}{n-b}, \qquad \tau_n = -\gamma_n.$$

The exact posterior is

$$\mu_g(s) = \frac{1}{Z_g} \exp\left\{ J_n F_g(s) + \frac{\gamma_n}{2} M(s)^2 \right\}, \qquad J_n > 0, \quad \gamma_n < 0. \tag{1.1}$$

Indeed, an observed edge contributes the pair coefficient $\frac{1}{2}\log(a/b)$, a nonedge contributes γ_n, and $\sum_{i<j} s_i s_j = (M(s)^2 - n)/2$. Subtracting the nonedge coefficient on observed edges gives J_n. The uniform prior then gives (1.1). The negative quadratic term couples the total spins of disconnected graph components, so even a decomposition of the observed graph must account for this interaction.

1.1 Main result

For a randomized algorithm run on the observed graph g, let $\widehat{\mu}_g$ denote its conditional output law, including all of its internal randomness. Its random seed is independent of the latent labels conditional on the graph and known parameters. We refer to this as a graph-only algorithm. For finite probability measures, $\mathrm{TV}(\lambda, \pi) = \frac{1}{2}\sum_x |\lambda(x) - \pi(x)|$.

Theorem 1.1. *There is an absolute constant C such that, for every fixed known rational $a > b > 0$ with $\kappa > C$, a graph-only algorithm satisfies*

$$\mathbb{E}_G \mathrm{TV}(\widehat{\mu}_G, \mu_G) = o(1).$$

Its bit complexity and random-bit use have deterministic polynomial bounds on every input graph. The polynomial may depend on the fixed parameters. There is no additional restriction on d.

For every fixed real pair $a > b > 0$ with $\kappa > C$, suppose certified evaluation of $c \in \{a, b\}$ to absolute error 2^{-k} has bit cost $T_c(k)$. There are finite auxiliary rational data, a function $\eta_{a,b}(n) \to 0$, and a polynomial $P_{a,b}$ such that, for dyadic $\epsilon \in (0, 1)$ and all sufficiently large n, the algorithm satisfies

$$\mathbb{E}_G \mathrm{TV}(\widehat{\mu}_{G,\epsilon}, \mu_G) \le \eta_{a,b}(n) + \epsilon.$$

After a fixed finite preprocessing step, it requests one approximation of each parameter with

$$k \le C_{a,b} + \lceil \log_2(n/\epsilon) \rceil$$

and has total bit cost at most

$$T_a(k) + T_b(k) + P_{a,b}(n, \epsilon^{-1}, k).$$

The polynomial also bounds its random-bit use. The preprocessing uses finitely many certified coarse evaluations, and its cost is charged separately.

The statistical guarantee is averaged over the original SBM graph marginal. It therefore also implies convergence in probability of the conditional total variation error. The computational bound holds on every graph, including graphs outside the events used in the proof. For fixed polynomial-time computable real parameters, taking an inverse-polynomial numerical accuracy gives polynomial total cost. For arbitrary computable parameters, the two evaluation costs in the theorem remain part of the conclusion.

The threshold C and the polynomial bounds obtained here are not optimized. The result establishes polynomial-time samplability throughout a parameter range defined by one absolute signal threshold. It does not identify the smallest possible signal threshold or give a practical implementation with competitive constants.

1.2 Related work

Recovery algorithms provide a useful starting point for the phase-based construction. Chin, Rao, and Vu [CRV15] develop spectral initialization and refinement using reserved edges. Mossel, Neeman, and Sly [MNS16] connect belief propagation and robust reconstruction to optimal recovery in

the two-community model. We use a two-half spectral construction in the spirit of [CRV15], with the simultaneous regularization theorem of Le, Levina, and Vershynin [LLV17]. Keeping this short construction makes its guarantee explicit for independent labels, all fixed $b > 0$, and deterministic polynomial bit complexity.

The sparse-reference analysis uses results of Liu, Mohanty, Rajaraman, and Wu [LMRW24]. Their decomposition of a sparse random graph into a spectrally controlled part and a near-forest addresses the high-degree vertices that obstruct a direct operator-norm argument. We use their covariance and tree entropy estimates as external inputs. Our sampling construction also needs quantitative control under the graph law obtained after conditioning on two replicas. We establish this control for sequential edge laws and allow a fixed number of excess edges in each sparse component. This stronger probability bound is sufficient for every stage of the tempering schedule and also supplies the initial sampler.

Component exchanges have a long history in Monte Carlo methods, including the replica-cluster updates of Houdayer [Hou01]. Our phase analysis uses the same elementary exchange symmetry to control linear statistics under the full joint law of exceptional spins and core defects. A component-path argument in the framework of Sinclair [Sin92] then gives a weak Poincaré inequality. Low-temperature polymer algorithms, such as those of Chen et al. [CGG$^+$21], provide a related approach to sampling close to an ordered configuration. Here the exceptional region has its own correlated distribution, and the posterior includes the dense magnetization penalty in (1.1). Both are retained in the construction.

1.3 Proof overview

The proof uses two complementary parameter ranges. A single recovered-phase argument naturally requires $\kappa \geq C_1(1+\log d)$ because connected defect sets must be suppressed relative to their graph-counting entropy. This condition follows from a large absolute lower bound on κ whenever θ is bounded away from zero. Small contrast requires a different construction, since $d = \kappa/\theta^2$ can then be arbitrarily large at a fixed signal-to-noise ratio.

Small contrast. Let $\nu_{\beta,g}(s) \propto \mu_g(s)^\beta$. At $t \asymp \kappa^{-1/2}$, an observable spectral calculation separates a positive semidefinite matrix of rank at most two from a sparse reference interaction and a small residual. The reference law is sampled by dynamic programming with two global counts. Gaussian block resampling adds the small residual, while numerical integration over a two-dimensional Gaussian field restores the low-rank part. The residual is defined exactly after the observable rank-two matrix has been chosen, so the construction samples the intended tempered law.

To cool from t to 1, consider two independent tempered replicas and align their global signs. Likelihood affinities put their disagreement set inside $2e^{-t\kappa/8}n$ vertices, except with exponentially small averaged probability. Conditioning on the replicas changes the graph law. We compute its conditional edge odds directly and obtain uniform upper and lower bounds on every successive internal-edge probability. A quantitative decomposition applies to this law and gives an all-field Poincaré bound on each good disagreement fiber. Averaging over fibers yields an even weak Poincaré inequality with a summable defect. A deterministic temperature schedule, started from the constructed law, then reaches the posterior.

The dependence on the graph-growth constants matters here. Retaining the ratio Δ/D in the external tree entropy inequality yields a weak Poincaré coefficient of order n^2, rather than a large power arising from replacing Δ/D by Δ. This bound is sufficient for $O(n^4 \log n)$ ideal single-site transitions in the cooling stage. The initialization and numerical work have their own polynomial costs.

A recovered phase. Start from a labeling with exponentially small error and perform local energy

ascent. Small-set sparsity prevents the error set from crossing a fixed Hamming barrier. Local stability then confines all remaining errors to the closure of atypical planted vertices. The algorithm constructs an observable exceptional set inside this closure. Its components have logarithmic size, and every small core defect has a uniform energy penalty, regardless of the exceptional spins.

A single witness estimate controls uncertainty components of two independent samples, uniformly in the scalar field, defect fugacity, and component cutoff. Exchanging a whole uncertainty component between replicas preserves their product weight. This gives a susceptibility bound for the joint law and a weak Poincaré estimate for component heat-bath updates. A small initial fugacity fixes the core while leaving only logarithmic exceptional components to sample; increasing the fugacity produces the law within a recovered Hamming ball.

Finally, a sampled estimate of the mean magnetization selects a scalar field. Rejection with probability

$$\exp\{-\tau_n(M(s)-c)^2/2\}$$

restores the quadratic term in (1.1) exactly. The susceptibility bound makes the acceptance probability inverse polynomial. A final independent fair sign combines the two phases. The local-ascent center already has sufficiently small magnetization, so this step requires no exact balancing of its coordinates.

The two constructions have estimates uniform on a sufficiently small fixed parameter neighborhood. This uniformity permits rational approximations that vary with n. Coupling the edge variables at two nearby parameter pairs gives the precision $O(\log(n/\epsilon))$ in Theorem 1.1.

1.4 Notation and external inputs

For spins s, t, let $d_H(s,t)$ be their Hamming distance and put $d_\pm(s,t) = \min\{d_H(s,t), d_H(s,-t)\}$. A function is even if $f(s) = f(-s)$, and $\operatorname{osc}(f) = \max f - \min f$. For a reversible kernel P with stationary law π, use

$$\mathcal{E}_{\pi,P}(f,f) = \langle f, (I-P)f\rangle_\pi = \frac{1}{2}\sum_{x,y}\pi(x)P(x,y)(f(x)-f(y))^2.$$

A weak Poincaré inequality has the form $\operatorname{Var}_\pi f \le A\mathcal{E}_{\pi,P}(f,f) + b\operatorname{osc}(f)^2$. A heat-bath update resamples the chosen coordinates from their exact conditional law. Such a kernel is an average of conditional-expectation projections and is positive semidefinite on $L^2(\pi)$.

Graph balls include their centers. For a connected graph Q, its excess is $\operatorname{exc}(Q) = |E(Q)| - |V(Q)| + 1$. We write Π_U for the coordinate projection onto a set U, $\|\cdot\|$ for the Euclidean operator norm, and $\|\cdot\|_{\mathrm{F}}$ for the Frobenius norm. All logarithms are natural unless a base is displayed. Constants called absolute are independent of a, b, d, θ, n. Other constants and polynomial degrees may depend on the fixed parameters. All computations are capped on every input; the treatment of numerical errors is collected in Appendix A.

We state the spectral input in the exact form used later. Its simultaneity allows the modified vertex set to be chosen from the observed graph.

Lemma 1.2 (Le–Levina–Vershynin, [LLV17], Theorem 2.1)**.** *Let A be the adjacency matrix of a graph on N vertices with independent unordered edges of probabilities p_{ij}, and put $d_{\max} = N\max_{i<j} p_{ij} > 0$. For every $r \ge 1$, with probability at least $1 - N^{-r}$, every symmetric nonnegative matrix A' obtained by reducing edge weights only on edges incident to a set of at most $10N/d_{\max}$ vertices satisfies*

$$\|A' - \mathbb{E}A\| \le C_{\mathrm{LLV}} r^{3/2}(\sqrt{d_{\max}} + \sqrt{d'}), \qquad d' = \max_i \sum_j A'_{ij},$$

where C_{LLV} is absolute.

The remaining external inputs are the tree covariance and entropy inequalities from [LMRW24], the Dobrushin Poincaré inequality from [Wu06], and the Brascamp–Lieb variance inequality [BL76]. They are cited at their points of use. We use standard concentration inequalities in the forms collected in [BLM13].

The posterior geometry and recovery construction are given in §2. The sparse-reference and graph-decomposition estimates occupy §3 and §4; initialization and cooling are treated in §5 and §6. The recovered-phase construction is developed in §7 and §8. Finally, §9 combines the two ranges and proves the real-parameter statement.

2 Posterior concentration and initialization

The posterior identity (1.1) separates the sparse edge interaction from the quadratic penalty on the total spin. We first show that, at large signal-to-noise ratio, the posterior concentrates in two Hamming balls. We then construct a center for these balls. The recovery argument follows the spectral-initialization and reserved-edge-refinement approach of [CRV15]; the spectral estimate is supplied directly by [LLV17, Theorem 2.1].

Lemma 2.1. *Suppose $\kappa \geq 32$ and put $q = e^{-\kappa/16}$. For every $x \in \{-1,1\}^n$,*

$$\mathbb{E}\big[\mu_G\{s : d_{\pm}(s,x) \geq qn\} \mid X = x\big] \leq 4e^{-\kappa q n/8}.$$

Proof. Write $L_s(g) = \mathbb{P}(G = g \mid X = s)$. Since the prior is uniform,

$$\mu_g(s) \leq \min\{1, L_s(g)/L_x(g)\} \leq \sqrt{L_s(g)/L_x(g)}.$$

If $d_H(s,x) = k$, exactly $k(n-k)$ pairs have different label products. Independence of the edges therefore gives

$$\mathbb{E}[\mu_G(s) \mid X = x] \leq h_n^{k(n-k)}, \qquad h_n = \frac{\sqrt{ab}}{n} + \sqrt{(1-a/n)(1-b/n)}.$$

Moreover,

$$1 - h_n \geq \frac{(\sqrt{a}-\sqrt{b})^2}{2n} \geq \frac{\kappa}{2n}.$$

For $qn \leq k \leq n/2$, we have $h_n^{k(n-k)} \leq e^{-\kappa k/4}$ and $\binom{n}{k} \leq (en/k)^k \leq e^{\kappa k/8}$. Summing the two antipodal Hamming layers yields

$$\mathbb{E}[\mu_G\{s : d_{\pm}(s,x) \geq qn\} \mid X = x] \leq 2 \sum_{k \geq \lceil qn \rceil} e^{-\kappa k/8} \leq 4e^{-\kappa q n/8}.$$

□

Proposition 2.2. *There is an absolute C_{rec} such that, for fixed rational $a > b > 0$ with $\kappa > C_{\mathrm{rec}}$, a deterministic algorithm returns a labeling $v = v(G)$ satisfying*

$$\mathbb{P}\{d_{\pm}(v, X) > e^{-\kappa/16} n\} = o(1).$$

Its bit complexity is polynomial on every input graph. In addition, on every compact parameter rectangle $U \subset \{a > b > 0,\ \kappa > C_{\mathrm{rec}}\}$, the failure probability tends to zero uniformly, and its cost is bounded by one polynomial in n and the rational input length.

Proof. Put $\alpha = (a-b)/2 = d\theta$ and $q = e^{-\kappa/16}$. Split the vertices into the fixed halves $H_1 = \{1, \dots, \lfloor n/2 \rfloor\}$ and $H_2 = [n] \setminus H_1$. For either half H, write $m = |H|$, $\lambda_H = ma/n$, and let A_H be the induced adjacency matrix. Delete edges incident to vertices of degree greater than $64\lambda_H$, obtaining A'_H, and set

$$B_H = A'_H - \frac{d}{n}\mathbf{1}_H\mathbf{1}_H^{\mathsf{T}} + \frac{a}{n}I_H.$$

Choose an integer D clearing the denominators of a and d. Form the integer matrix $T_H = DnB_H$, raise it to the even power $t = 10(1 + \lfloor \log_2 n \rfloor)$, and choose a column of T_H^t with largest squared Euclidean norm. Assign $+1$ to its $\lfloor m/2 \rfloor$ largest coordinates and -1 to the others. All ties are resolved by vertex index. Denote this balanced labeling by W_H.

Conditional on $X = x$, edges in H are independent and have probabilities at most a/n. If at least $k = \lceil m/\lambda_H \rceil$ vertices have degree greater than $64\lambda_H$, some k-set meets more than $32\lambda_H k$ edges. The exponential bound for a sum of Bernoulli variables and a union bound over the sets give probability at most

$$2^m \exp\{-[32 - (e-1)]\lambda_H k\} \le e^{-29m}.$$

Outside this event, at most $10m/\lambda_H$ vertices were modified. Applying [LLV17, Theorem 2.1] with failure exponent 3 gives, uniformly in x,

$$\|A'_H - \mathbb{E}[A_H \mid X = x]\| \le K\sqrt{a}$$

except with probability at most $m^{-3} + e^{-29m}$, where K is absolute. The expectation is

$$\mathbb{E}[A_H \mid X = x] = \frac{d}{n}\mathbf{1}_H\mathbf{1}_H^{\mathsf{T}} + \frac{\alpha}{n}x_H x_H^{\mathsf{T}} - \frac{a}{n}I_H.$$

Thus $B_H = \ell_H u_H u_H^{\mathsf{T}} + E_H$, where $u_H = x_H/\sqrt{m}$, $\ell_H = \alpha m/n$, and $\|E_H\| \le K\sqrt{a}$. Since $a < 2d$ and $\alpha^2 = \kappa d$, increasing C_{rec} makes $\|E_H\|/\ell_H \le 2^{-16}$. The largest eigenvalue is then at least $\ell_H - \|E_H\|$, while all remaining eigenvalues have absolute value at most $\|E_H\|$. The variational characterization also shows that a corresponding unit eigenvector is within squared distance $4\|E_H\|/\ell_H$ of one of $\pm u_H$.

The selected column of B_H^t has norm at least its largest eigenvalue to power t, divided by $\sqrt{m}$; its component orthogonal to the leading eigenvector has norm at most $\|E_H\|^t$. Consequently its normalized vector has squared distance at most 2^{-12} from one of $\pm u_H$, for all sufficiently large n. Its coordinate signs make at most $2^{-12}m$ mistakes. Enforcing balance adds at most this many mistakes plus $|\sum_{i\in H} x_i|/2 + 1$. Hoeffding's inequality [BLM13, Chapter 2] therefore gives an event $\mathcal{G}_H$ with

$$\mathbb{P}(\mathcal{G}_H^c) \le m^{-3} + e^{-29m} + 2e^{-m/(2\cdot 10^6)}$$

on which $d_{\pm}(W_H, X_H) \le m/100$ and $|\sum_{i\in H} X_i| \le m/1000$.

Now use the cross edges to refine each estimate. If $K = [n] \setminus H$, put

$$R_K(i) = \operatorname{sgn}_+\left(\sum_{j\in H} W_H(j)A_{ji}\right), \qquad i \in K,$$

where $\operatorname{sgn}_+(0) = 1$. For the analysis, choose $\tau_H \in \{-1, 1\}$ so that $Y_H = \tau_H W_H$ is closest to X_H, and condition on $\mathcal{F}_H = \sigma(X, A_H)$. On $\mathcal{G}_H$, the signed vote $S_i = X_i \sum_{j\in H} Y_H(j)A_{ji}$ has conditional mean η_i and variance σ_i^2 satisfying, for sufficiently large n,

$$\frac{7\alpha}{16} \le \eta_i \le \frac{9\alpha}{16}, \qquad \sigma_i^2 \le \frac{9d}{16}.$$

Here balance of Y_H controls the d/n term in the mean, and balance of X_H controls the variance. Bernstein's inequality [BLM13, Theorem 2.10 and Corollary 2.11] yields

$$\mathbb{P}(S_i \le 0 \mid \mathcal{F}_H) \le \exp\left\{-\frac{\eta_i^2}{2(\sigma_i^2 + \eta_i/3)}\right\} \le e^{-49\kappa/384} \le q^2.$$

For distinct $i \in K$, these error events use disjoint cross edges and are conditionally independent. The Bernoulli upper-tail bound, with exponential parameter $\log(1/q)$, gives

$$\mathbb{P}\{d_H(R_K, \tau_H X_K) > q|K| \mid \mathcal{F}_H\} \le e^{-q|K|} \quad \text{on } \mathcal{G}_H.$$

Apply this bound in both directions and take a union bound; independence between the two refinement procedures is not needed.

Finally let $c = \operatorname{sgn}_+\langle R_{H_2}, W_{H_2}\rangle$, and output $v_{H_2} = R_{H_2}$ and $v_{H_1} = cR_{H_1}$. On the preceding success event,

$$d_H(R_{H_2}, \tau_{H_1}\tau_{H_2} W_{H_2}) \le (q + 1/100)|H_2| < |H_2|/2,$$

so $c = \tau_{H_1}\tau_{H_2}$ and $d_\pm(v, X) \le qn$.

All algorithmic steps are integer or rational operations. If b_{in} is the parameter encoding length, an entry of T_H^j has $O(j(b_{\text{in}} + \log(n+1)))$ bits. The exponent is $O(\log n)$, so matrix multiplication, column comparisons, sorting, and refinement have polynomial bit cost on every graph. For finitely many small sizes the algorithm may return a fixed labeling. On a compact rectangle U, all finite-size inequalities above hold beyond one threshold, the spectral margin is uniform, and q has a positive lower bound. The displayed failure bounds and arithmetic estimates are therefore uniform on U. □

Lemma 2.3. *Let P be a positive-semidefinite reversible Markov kernel with stationary law π. Suppose*

$$\operatorname{Var}_\pi f \le A\mathcal{E}_{\pi,P}(f,f) + b\operatorname{osc}(f)^2, \qquad A \ge 1, \quad b \ge 0,$$

for every real f. If $\lambda \le M\pi$ pointwise with $M \ge 1$, then

$$\operatorname{TV}(\lambda P^T, \pi) \le \sqrt{M}\sqrt{\tfrac{1}{4}e^{-T/A} + b}.$$

The same conclusion holds if the inequality is only assumed for even functions, provided π, λ are invariant under global sign and $P(-s,-t) = P(s,t)$.

Proof. For $0 \le f \le 1$, put $f_j = P^j f$ and $V_j = \operatorname{Var}_\pi f_j$. The spectrum of P lies in $[0,1]$, so

$$V_j - V_{j+1} = \langle f_j, (I - P^2) f_j\rangle_\pi \ge \mathcal{E}_{\pi,P}(f_j, f_j) \ge (V_j - b)/A.$$

Since $V_0 \le 1/4$ and $\operatorname{osc}(f_j) \le 1$, iteration gives $V_T \le e^{-T/A}/4 + b$. With $h = d\lambda/d\pi$, we have $\mathbb{E}_\pi h^2 \le M$, and Cauchy–Schwarz gives $|\lambda P^T f - \pi f| \le \sqrt{MV_T}$. Taking the supremum over f proves the claim. Under sign symmetry, replace f by $(f(s) + f(-s))/2$; its iterates remain even and both measures have the same expectation of f and its even part. □

3 Sparse reference measures

The small-contrast construction uses a sparse interaction together with a small matrix perturbation. The sparse components need not be trees: a fixed bound on their excess is sufficient both for

sampling and for the variance estimates. This observation lets us use one quantitative graph decomposition throughout the proof.

For a symmetric matrix A and $h \in \mathbb{R}^m$, write

$$\pi_{A,h}(s) \propto \exp\{s^{\mathsf{T}} A s/2 + h^{\mathsf{T}} s\}, \qquad s \in \{-1,1\}^m.$$

Diagonal entries of A only change the normalizing constant. Let H be a graph with components $Q_1, \dots, Q_N$, including singletons, and let U contain every singleton component. Assume that, for some $1 \le D \le \Delta$,

$$|B_j^H(v)| \le \Delta D^{j-1} \quad (j \ge 1), \qquad |B_j^H(v) \cap U| \le D^j \quad (j \ge 0). \tag{3.1}$$

Let J_H, E be symmetric zero-diagonal matrices, with J_H supported on H, and put $R = J_H + E$. Define

$$W_{ii} = 0, \quad W_{ij} = \sum_{v \in Q_i,\ w \in Q_j} |E_{vw}| \ (i \ne j), \qquad q_{\mathrm{blk}} = \|W\|, \quad \delta = \max_i \max_{s \in \{-1,1\}^{Q_i}} \left| \frac{1}{2} s^{\mathsf{T}} E_{Q_i,Q_i} s \right|. \tag{3.2}$$

The two quantities in (3.2) measure, respectively, interactions between distinct components and the perturbation within one component.

3.1 Covariance and variance estimates

We use two results for tree interactions from [LMRW24, Lemmas 4.5, 4.18–4.19 and Appendix A.1]. If H is a forest satisfying (3.1) and $|(J_H)_{vw}| \le 1/(2\sqrt{D})$, then, uniformly in h,

$$\|\Pi_U \operatorname{Cov}_{\pi_{J_H,h}}(s) \Pi_U\| \le 4e, \qquad \operatorname{Var}_{\pi_{J_H,h}} f \le 4e^{1+32\Delta/D} \sum_{i=1}^m \mathbb{E}_{\pi_{J_H,h}} \operatorname{Var}(f \mid s_{-i}). \tag{3.3}$$

For signed tree interactions, a change of variables $s_v \mapsto \zeta_v s_v$ makes every edge interaction nonnegative and only changes the field. Thus the cited results apply with the same bounds. The second inequality is the Poincaré consequence of the cited tree entropy inequality, with a deliberately conservative constant. We use these results directly.

Lemma 3.1. *Suppose every component of H has excess at most e_0, $|(J_H)_{vw}| \le 1/(2\sqrt{D})$, $q_{\mathrm{blk}} \le 1/2$, and $\delta \le 1$. Set $C_{e_0} = 256\, 4^{e_0}$. Then, for every h,*

$$\operatorname{Var}_{\pi_{R,h}} f \le 2 \sum_{j=1}^N \mathbb{E}_{\pi_{R,h}} \operatorname{Var}(f \mid s_{Q_j^c}), \tag{3.4}$$

$$\|\Pi_U \operatorname{Cov}_{\pi_{R,h}}(s) \Pi_U\| \le C_{e_0}, \tag{3.5}$$

$$\operatorname{Var}_{\pi_{R,h}} f \le 8e^{5+2e_0+32\Delta/D} \sum_{i=1}^m \mathbb{E}_{\pi_{R,h}} \operatorname{Var}(f \mid s_{-i}). \tag{3.6}$$

The same assertions hold after multiplying R by any number in $[0,1]$.

Proof. Changing the spins of Q_j changes the conditional log weight on Q_i by a function with oscillation at most $4W_{ij}$. Exponential interpolation between the two conditional laws shows that their total variation distance is at most W_{ij}: for an event A, the derivative is a covariance bounded by $\frac{1}{4}$ times this oscillation. Consequently the block Dobrushin matrix is entrywise bounded by W.

Its spectral radius is at most $\|W\| \le 1/2$. The block Poincaré inequality of [Wu06, Theorem 2.1], applied with the discrete metric on each block, gives (3.4).

Choose a spanning tree in each component. Deleting at most e_0 non-tree edges changes its energy by at most $e_0/(2\sqrt{D})$. Conditional on all other components, the remaining perturbation of its tree law therefore has absolute energy at most

$$\delta' = 1 + \frac{e_0}{2\sqrt{D}} \le 1 + e_0/2.$$

The tree inherits (3.1). Its conditional density ratio relative to the perturbed law lies between $e^{-2\delta'}$ and $e^{2\delta'}$. The usual bounded-perturbation comparison, in the form of [LMRW24, Lemma 3.8], transfers the covariance bound in (3.3) with factor $e^{2\delta'}$ and the Poincaré bound with factor $e^{4\delta'}$. Apply (3.4) to a linear form supported on U and sum over the components. This gives

$$\|\Pi_U \operatorname{Cov}_{\pi_{R,h}}(s)\Pi_U\| \le 8e^{3+e_0} < 256\, 4^{e_0}.$$

Applying the same block inequality to the conditional single-site estimates gives (3.6). Scaling the interaction down preserves all hypotheses. □

The next lemma adds a positive semidefinite interaction through an auxiliary Gaussian field. We state the argument for a general finite support because the conditional-variance identity will also clarify which covariance bound is required.

Lemma 3.2. *Let $\Omega \subseteq \{-1,1\}^m$ be nonempty, let $w(s) > 0$ on Ω, and let L be a real $k \times m$ matrix. Suppose that*

$$\nu_z(s) \propto w(s)e^{z^{\mathsf{T}}Ls}, \qquad L \operatorname{Cov}_{\nu_z}(s)L^{\mathsf{T}} \preceq qI_k \quad (z \in \mathbb{R}^k), \qquad 0 \le q < 1.$$

For $\mu(s) \propto w(s)e^{\|Ls\|^2/2}$, consider the update

$$Z \sim N(Ls, I_k), \qquad s' \sim \nu_Z.$$

Its kernel P is reversible and positive semidefinite, and

$$\operatorname{Var}_\mu f \le \frac{1}{1-q}\mathbb{E}\operatorname{Var}_{\nu_Z} f, \qquad \mathcal{E}_{\mu,P}(f,f) = \mathbb{E}\operatorname{Var}_{\nu_Z} f. \tag{3.7}$$

In particular,

$$\operatorname{TV}(P^T(s,\cdot),\mu) \le \frac{1}{2}\sqrt{\mu(s)^{-1}-1}\, e^{-(1-q)T}.$$

Proof. Use the joint density proportional to $w(s)e^{-\|z\|^2/2+z^{\mathsf{T}}Ls}$. Completing the square gives the stated marginals and conditional laws. In stationarity, the joint mass of two consecutive spins is $\int \nu_z(s)\nu_z(s')\,\eta(dz)$, where η is the field marginal. This proves reversibility, positivity, and the Dirichlet-form identity.

If $V = -\log(d\eta/dz)$ and $g(z) = \mathbb{E}_{\nu_z} f$, differentiation gives

$$\nabla^2 V = I_k - L \operatorname{Cov}_{\nu_z}(s)L^{\mathsf{T}} \succeq (1-q)I_k, \qquad \nabla g = \operatorname{Cov}_{\nu_z}(f, Ls).$$

For every unit vector u, Cauchy–Schwarz bounds $|u^{\mathsf{T}}\nabla g|^2$ by $q\operatorname{Var}_{\nu_z} f$. The Brascamp–Lieb variance inequality [BL76, BK20] therefore yields

$$\operatorname{Var}_\eta g \le \frac{q}{1-q}\mathbb{E}\operatorname{Var}_{\nu_Z} f.$$

Conditional total variance gives (3.7). The stated total variation bound follows from the standard L^2 spectral estimate for a reversible positive-semidefinite kernel; see [LP17, Chapter 12]. □

Proposition 3.3. *Under the assumptions of Lemma 3.1, let K be symmetric, supported on $U \times U$, and satisfy*

$$\|K\| \le \frac{1}{16 C_{e_0}} = \frac{1}{4096\, 4^{e_0}}.$$

Then for every $\alpha \in [0,1]$, every h, and $\lambda = \pi_{\alpha(K+R),h}$,

$$\mathrm{Var}_\lambda f \le 16 e^{5+2e_0+32\Delta/D} \sum_{i=1}^{m} \mathbb{E}_\lambda \mathrm{Var}_\lambda(f \mid s_{-i}).$$

Proof. Put $b = (16C_{e_0})^{-1}$ and $B = \alpha K + 2b\Pi_U$. On U its eigenvalues lie in $[b, 3b]$, so write $B = L^{\mathsf{T}} L$ with L zero outside U. Lemma 3.1 gives $L\,\mathrm{Cov}_{\pi_{\alpha R,h'}}(s) L^{\mathsf{T}} \preceq (3/16) I$ for every h'. The Gaussian representation has spin marginal λ, since the added diagonal only contributes a constant. Apply Lemma 3.2 and then (3.6) inside the conditional expectation. For each coordinate, conditional total variance gives

$$\mathbb{E}_Z \mathbb{E}_{\pi_{\alpha R, h + L^{\mathsf{T}} Z}} \mathrm{Var}(f \mid s_{-i}, Z) \le \mathbb{E}_\lambda \mathrm{Var}(f \mid s_{-i}).$$

Multiplying the reference bound by $16/13$ proves the assertion. If U is empty, $K = 0$ and the reference bound already suffices. □

3.2 Sampling with two global counts

The nonedge interaction will be retained by keeping two integer counts in a tree recursion. Allowing a bounded number of cycles per component does not change polynomial-time computability.

Lemma 3.4. *Let every component of a graph H on $[m]$ have excess at most a fixed integer e_0, and fix $v \in \{-1,1\}^m$. Put*

$$k(s) = |\{i : s_i = 1\}|, \qquad \ell(s) = |\{i : s_i = v_i\}|.$$

For nonnegative rational unary factors, positive rational edge factors, and a nonnegative rational table $g(k,\ell)$, the partition function of

$$W(s) = g(k(s), \ell(s)) \prod_i \lambda_i(s_i) \prod_{ij \in E(H)} \psi_{ij}(s_i, s_j)$$

is exactly computable, including detection of zero partition, in time polynomial in m and the input length for fixed e_0. If it is positive, sampling to total variation error η has a deterministic polynomial bit and random-bit bound in these quantities and $\log(1/\eta)$. The output always belongs to the support.

Proof. First consider a rooted tree. At vertex i, keep the conditional partition table indexed by its spin and by the two counts in its subtree. Incorporating a child multiplies by its edge factor and convolves the two count coordinates. Each table has at most $2(m+1)^2$ entries; direct convolution over all children costs $O(m^5)$ arithmetic operations. Convolve component tables and multiply the final table by g.

For a component of excess at most e_0, choose a spanning tree and one endpoint of each non-tree edge. Pin the spins on this set of at most e_0 vertices. After deleting the pinned vertices, all remaining edges lie in a forest; edges incident to pinned vertices become unary factors. Sum the count tables over at most 2^{e_0} pin assignments, including the pinned vertices' weights and counts. This enumeration is done separately in each component, so its factor is 2^{e_0}, not exponential in the number of components.

Clear the denominators factor by factor. A partial product has bit length at most the total input length, and summing over assignments adds at most m bits. Thus all table entries have polynomial bit length. Conditional partitions with additional pins are obtained by the same procedure. A uniform integer rank in the final integer partition sum then samples by self-reduction, deciding one spin at a time from its two pinned partitions. Rejection from the next larger power of two accepts with probability at least $1/2$. Cap this rejection after $\lceil \log_2(1/\eta) \rceil + 1$ attempts, using the rank-zero supported configuration on failure. The cap changes the law by at most η and gives the stated deterministic bounds. □

The same algorithm applies to computable real factors once their logarithms can be evaluated to the required precision. If there are N_f factors, logarithmic relative error at most $\eta/(10N_f)$ in each positive factor changes the normalized law by at most η. Exact zero factors remain zero. This observation, together with Lemma A.1, will be used without repeating numerical details in each sampling construction.

4 Graph decompositions under the replica law

The disagreement set of two posterior samples depends on the observed graph. We therefore need graph estimates that remain valid after conditioning on this set. We first establish the estimates for graphs whose successive conditional edge probabilities lie in a prescribed interval, and then verify this condition for the actual replica law. All graphs in this section retain their isolated vertices.

4.1 Canonical excision and spectral regularization

For a graph A on $[m]$ and an integer $D > 1$, define

$$\ell_A(v) = \min\{r \geq 0 : |B_j^A(v)| \leq D^j \text{ for every integer } j \geq r\}, \qquad H_D(A) = \bigcup_{v \in [m]} A[B^A_{\ell_A(v)}(v)].$$

The union is an edge union. Write $H = H_D(A)$, let F have the remaining edges, and put

$$T = V'(H) \setminus V'(F), \qquad U = [m] \setminus T, \qquad M_m = \mathbf{1}_m \mathbf{1}_m^\top - I_m,$$

where $V'(J)$ denotes the vertices incident to an edge of J.

Lemma 4.1. *The canonical excision is monotone: $A \subseteq A'$ implies $H_D(A) \subseteq H_D(A')$. Moreover,*

$$\deg_{\max}(F) \leq D, \qquad A_F = \Pi_U A_F \Pi_U, \qquad |B_j^H(v) \cap U| \leq D^j \quad (j \geq 0).$$

The partition is computable in polynomial time. If $S_2(H)$ is the sum of $|Q|^2$ over the nonsingleton components Q of H, then $|V'(H)| \leq S_2(H)$.

Proof. Ball inclusion implies $\ell_A(v) \leq \ell_{A'}(v)$ and hence monotonicity. If $\mathrm{dist}_A(v, w) < \ell_A(v)$, all edges incident to w belong to the selected ball around v. Consequently, for a nonsingleton component, $B_j^H(v)$ misses U when $j < \ell_A(v)$, whereas for $j \geq \ell_A(v)$ it has size at most D^j. A singleton belongs to U. Any vertex incident to F has $\ell_A(v) = 0$, so its degree in A is less than D. This also proves the support assertion. The last inequality follows by summing component sizes. Breadth-first search determines the radii: only $j \leq \lceil \log_D m \rceil$ need be tested. □

Here and below, a bound on successive edge probabilities refers to one fixed ordering of the edges, conditional on those already exposed.

Lemma 4.2. *Suppose a random graph A on $[m]$ has successive edge probabilities in $[\ell, u]$, where $0 < \ell < u < 1$, and let $p \in [\ell, u]$. Set $\Lambda = \max\{1, mu\}$. For every $r \geq 1$, outside an event of probability at most $2m^{-r} + e^{-m}$, the following holds simultaneously for every symmetric matrix A' obtained by reducing edge weights of A only on edges incident to a set of at most $8m/\Lambda$ vertices, and satisfying $\max_i \sum_j A'_{ij} \leq 64\Lambda$:*

$$\|A' - pM_m\| \leq Cr^{3/2}\sqrt{\Lambda} + 2m(u - \ell).$$

On the same event, at most m/Λ vertices of A have degree greater than 64Λ.

Proof. Use independent uniforms in the sequential construction to couple

$$A^- \subseteq A \subseteq A^+, \qquad A^- \sim G(m, \ell), \quad A^+ \sim G(m, u).$$

The difference $A^\circ = A^+ - A^-$ also has independent edges, of probability $u - \ell$. Apply the simultaneous regularization theorem [LLV17, Theorem 2.1], in the form stated in Lemma 1.2, to A^- and A°.

Let $I^+ = \{v : \deg_{A^+}(v) > 64\Lambda\}$. If $|I^+| \geq k := \lfloor m/\Lambda \rfloor + 1$, some set of k vertices touches more than $32\Lambda k$ edges. The number of edges touching a fixed such set is binomial with mean at most Λk. The Chernoff bound and a union bound over the sets give

$$\mathbb{P}(|I^+| > m/\Lambda) \leq e^{-m}.$$

Fix an admissible A', with modified set I, on these three events. Split $A' = B' + R'$ by putting $B'_{ij} = A'_{ij}A^-_{ij}$. The regularization theorem gives

$$\|B' - \ell M_m\| \leq Cr^{3/2}\sqrt{\Lambda}.$$

For the remainder, form a matrix C' by retaining R' on edges incident to $I \cup I^+$ and retaining A° elsewhere. Then $0 \leq R' \leq C'$, the maximum row sum of C' is at most 64Λ, and its modified set has size at most $9m/\Lambda$. Since $m(u - \ell) \leq \Lambda$, the same regularization theorem gives

$$\|C' - (u - \ell)M_m\| \leq Cr^{3/2}\sqrt{\Lambda}.$$

Entrywise domination of nonnegative symmetric matrices implies $\|R'\| \leq \|C'\|$. Combining the two bounds with $0 \leq p - \ell \leq u - \ell$ proves the result, after enlarging C. The external theorem is simultaneous over the modified sets and retained weights, so all conclusions hold simultaneously over A'. □

The next estimate is the quantitative geometric input used both to initialize the sampler and to cool it. It gives a polynomially small failure probability for a single canonical decomposition.

Lemma 4.3. *For every integer $B \geq 1$, there is an integer L_B with the following property. Fix an integer $L \geq L_B$, and suppose the successive edge probabilities of A on $[m]$ are at most L/m. Put*

$$D = 1024L, \quad \alpha = e^{-D/16}, \quad e_0 = 2B + 29, \quad \Delta_m = D + (B + 5)\log m.$$

For all sufficiently large m, with probability at least $1 - m^{-B}$, every component Q of $H = H_D(A)$ satisfies

$$\mathrm{exc}(Q) \leq e_0, \qquad |Q| \leq m^{1/8}, \qquad S_2(H) \leq \alpha m, \qquad |B_j^H(v)| \leq \Delta_m D^{j-1} \quad (j \geq 1).$$

The probability estimate is uniform over all eligible sequential laws. The boundary and degree bounds of Lemma 4.1 hold deterministically.

Proof. First consider $\widehat{A} \sim G(m, L/m)$. We give the estimates for its excision; monotone coupling will then transfer them to A.

A heavy radius-j ball, meaning $|B_j(v)| > D^j$, contains a path of at most j vertices, starting at v, whose ambient degrees sum to more than $j(D-1)$. Indeed, take a breadth-first tree, allow termination at every vertex, and extend terminated paths to depth j by single-child steps. This tree has one leaf for each vertex of the ball. Uniform choices among children assign some leaf probability less than D^{-j}, so the product of the branching numbers along that path exceeds D^j. The arithmetic–geometric mean inequality gives the asserted degree sum after deleting the artificial steps. A positive canonical radius r is at least two, and the ball of radius $r-1$ is heavy. Its selected ball therefore contains a path with ambient-degree sum at least $(D-1)r/2$ and at most $r-1$ vertices.

Consider an induced path in the intersection graph of positive-radius selected balls, and let R be the sum of their radii. Nonconsecutive balls are disjoint. The odd or the even balls have total radius at least $R/2$, so their witness paths are disjoint and have total ambient-degree sum at least $(D-1)R/4$. Joining consecutive centers through ball intersections, and attaching these witnesses, gives a connected set S containing any prescribed vertices in the two end balls, with

$$|S| \le 3R+1, \qquad \sum_{v\in S} \deg_{\widehat{A}}(v) \ge (D-1)R/4. \tag{4.1}$$

For a specified root, the expected number of embedded rooted plane trees with s vertices is at most $(4L)^{s-1}$. Conditional on its tree edges, the sum of ambient degrees of its vertices has exponential moment at most e^{9Ls}: the fixed tree contributes $2(s-1)$, and the other incident edges contribute at most two each. Thus, for $R \ge 2$ and $s \le 3R+1$, the expected number of connected witnesses in (4.1) is at most

$$(4L)^{s-1} e^{9Ls-(D-1)R/4} \le e^{-DR/8}.$$

Summing over s, and also weighting by s when needed, gives

$$\sum_{R\ge R_*} 16R^2 e^{-DR/8} \le e^{-DR_*/16} \qquad (R_* \ge 2). \tag{4.2}$$

It follows that an induced intersection path of total radius at least R_* exists with probability at most $me^{-DR_*/16}$. If none exists, each excision component has diameter less than $2R_*$ and each selected radius is less than R_*. The radius-$2R_*$ ball around any vertex is then light, so each component has size at most D^{2R_*}.

Choose L_B sufficiently large that

$$\frac{32(B+12)\log D}{D} \le \frac{1}{16}, \qquad R_m = \left\lceil \frac{16(B+12)\log m}{D} \right\rceil + 2.$$

Outside probability m^{-B-11}, every excision component has diameter less than $2R_m$ and size at most $D^{2R_m} \le m^{1/8}$. If a component on this event has at least $k = 2B+30$ excess edges, a breadth-first spanning tree and k non-tree edges contain a connected witness with at most $1+4kR_m$ vertices and k excess edges. Enumerating its plane tree and additional edges bounds its probability by

$$\sum_{s\le 1+4kR_m} m(4L)^{s-1} s^{2k} (L/m)^k \le m^{1-k/2} = m^{-B-14}$$

for sufficiently large m. Here the displayed choice of L_B bounds the tree factor by a constant times $m^{k/8}$, and the remaining fixed constants and logarithmic factors by $m^{k/8}$.

Write $\Xi = S_2(H_D(\widehat{A}))$. For a fixed vertex x, every vertex of its nonsingleton excision component is covered by a witness in (4.1) rooted at x. Summing witness sizes using (4.2), starting at $R_* = 2$, gives

$$\mathbb{E}\Xi \le me^{-D/8} = \alpha^2 m.$$

To obtain concentration, encode the graph by its independent upper triangular rows. On the set of graphs whose excision components have size at most $s_* = m^{1/8}$, the function Ξ is $2s_*^2$-Lipschitz in the number of altered rows. To see this, adding an edge changes the excision only inside the final excision component containing that edge. Every newly selected edge lies in a final selected ball containing the added edge: otherwise the balls and induced edges determining that selection would be unchanged. For two graphs differing in rows indexed by I, delete all edges incident to I to obtain a common subgraph. Monotonicity shows that all changes lie in final components meeting I, whose total squared sizes are at most $|I|s_*^2$ in each graph.

Extend Ξ from this set by the infimum of its value plus $2s_*^2$ times row distance, and truncate to $[0, m^2]$. The extension Y remains $2s_*^2$-Lipschitz and agrees with Ξ on the good set. The bounded-differences inequality [BLM13, Theorem 6.2] gives

$$\mathbb{P}(Y - \mathbb{E}Y \ge u) \le \exp\{-u^2/(2ms_*^4)\}.$$

The component-size estimate implies $\mathbb{E}Y \le \alpha^2 m + m^{-B-9} \le \alpha m/2$ for large m. Consequently,

$$\mathbb{P}(\Xi > \alpha m,\ \max_Q |Q| \le s_*) \le e^{-\alpha^2\sqrt{m}/8}.$$

Finally, breadth-first exploration of $\widehat{A}$ is dominated by a Galton–Watson tree with offspring distribution $\mathrm{Poi}(2L)$, for $m \ge 2L$. This follows by dominating each $\mathrm{Bernoulli}(L/m)$ test by a Poisson variable of mean $-\log(1 - L/m)$ and adding dummy tests. If Z_j is the total population through depth j, iterating its generating function gives

$$\mathbb{E}\exp\{Z_j/D^{j-1}\} \le e^{1+4L} \qquad (j \ge 1).$$

For completeness, the logarithmic moment generating function is the $(j+1)$-fold iterate at zero of $x \mapsto h + 2L(e^x - 1)$, with $h = D^{-(j-1)}$. For $j \ge 2$, its first j iterates are at most $2h(4L)^{i-1} \le 1$; the last is at most $1 + 4L$. The case $j = 1$ is immediate. Exponential Markov and a union bound over roots and radii give the stated ball bound outside probability m^{-B-3}. These failure probabilities sum to at most m^{-B} for large m.

For a general sequential law, the same uniforms couple $A \subseteq \widehat{A}$. Lemma 4.1 gives $H_D(A) \subseteq H_D(\widehat{A})$. Component size, excess, and ball bounds pass to subgraphs; splitting components decreases the sum of their squared sizes. This proves the uniform assertion. □

We shall also use a version of Lemma 4.2 at the larger degree threshold $D = 1024L$. Fix $L_* \ge \max\{2, L_{32}\}$. If $m > 2L_*$, put $L = \lceil \max\{L_*, mu\} \rceil$. Outside probability $2(32m)^{-32} + e^{-32m}$, every edge partition with $\deg_{\max}(F) \le 1024L$ and $|V'(H)| \le m/L$ satisfies

$$\|A_F - pM_m\| \le C_*\sqrt{\max\{1, mu\}} + 64m(u - \ell). \tag{4.3}$$

Indeed, extend the graph to $32m$ vertices using independent $\mathrm{Bernoulli}(p)$ edges outside its original block, and apply Lemma 4.2 with upper endpoint $\max\{u, L_*/m\}$. Replace the original block by A_F and delete the other edges touching vertices above its degree threshold. The modified set has size at most twice the permitted unit $m/\max\{L_*, mu\}$, and all row sums are at most $2048\max\{L_*, mu\} \ge 1024L$. Compression gives (4.3). A violating original graph has no extension in the simultaneous regularization event, so the stated failure bound holds under its marginal law.

4.2 The conditional law of a disagreement graph

For $\beta \in [0,1]$, let

$$\nu_{\beta,g}(s) = \frac{L_s(g)^\beta}{\sum_x L_x(g)^\beta}.$$

The graph always has its original SBM marginal. Conditional on $G = g$, let σ, τ be independent with law $\nu_{\beta,g}$.

Lemma 4.4. *Put $r_- = b/(n-b)$, $r_+ = a/(n-a)$, and*

$$\ell = \frac{r_-^3/r_+^2}{1 + r_-^3/r_+^2}, \qquad u = \frac{r_+^3/r_-^2}{1 + r_+^3/r_-^2}.$$

Conditional on the disagreement set V, its common exterior spins, and all edges outside $\binom{V}{2}$, the graph $G[V]$ has successive edge probabilities in $[\ell, u]$. For fixed g, V and common exterior, the orientation $s = \sigma_V = -\tau_V$ has the full-cube law

$$\lambda_{\beta,V}(s) \propto \exp\left\{ 2\beta J_n \sum_{ij \in E(g[V])} s_i s_j + \beta\gamma_n \left(\sum_{i \in V} s_i \right)^2 \right\}. \tag{4.4}$$

Both conclusions remain true after replacing τ by $-\tau$ whenever $d_H(\sigma,\tau) > n/2$. If $\theta \le 1/30$, $n \ge 2a$, and $p = d/n$, then

$$\ell < p < u < 2p, \quad u - \ell \le 60p\theta, \quad J_n \le 3\theta, \quad |\gamma_n| \le 3p\theta, \quad |\gamma_n + pJ_n| \le p\theta^3. \tag{4.5}$$

Proof. Write $Z_\beta(g) = \sum_s L_s(g)^\beta$. The joint mass of (g, σ, τ) is

$$2^{-n} Z_1(g) \frac{L_\sigma(g)^\beta L_\tau(g)^\beta}{Z_\beta(g)^2}.$$

For graphs g^0, g^1 differing by addition of one edge, $L_s(g^1)/L_s(g^0)$ belongs to $\{r_-, r_+\}$, while $Z_\beta(g^1)/Z_\beta(g^0)$ lies in $[r_-^\beta, r_+^\beta]$. The full conditional edge odds therefore lie in

$$[r_-^{1+2\beta}/r_+^{2\beta},\ r_+^{1+2\beta}/r_-^{2\beta}] \subseteq [r_-^3/r_+^2,\ r_+^3/r_-^2].$$

Conditional expectation preserves the corresponding probability interval under the coarser exposure in the statement.

On a fixed fiber, the crossing edge terms cancel between the two replicas. If m_y is the exterior magnetization and $m_s = \sum_{i \in V} s_i$, the two dense terms sum to $\beta\gamma_n(m_y^2 + m_s^2)$, proving (4.4). Sign alignment multiplies each pair mass by two if $|V| < n/2$, by one if $|V| = n/2$, and by zero otherwise. This factor is constant on each remaining fiber and preserves both conditional laws.

For (4.5), set $z = p/(1-p) < 1$. Then

$$r_- = z\frac{1-\theta}{1+z\theta}, \quad r_+ = z\frac{1+\theta}{1-z\theta}, \quad \gamma_n = -\operatorname{arctanh}(z\theta), \quad J_n = \operatorname{arctanh}\theta + \operatorname{arctanh}(z\theta).$$

Thus $ze^{-3\theta} \le r_- < z < r_+ \le ze^{3\theta}$; applying $x \mapsto x/(1+x)$ gives the first two estimates. The remaining bounds follow from the integral formula for arctanh and $0 \le \operatorname{arctanh} x - x \le 4x^3/9$ for $0 \le x \le 1/2$. In the last bound the linear terms cancel because $(1-p)z = p$. □

5 A tempered initialization

We first sample at inverse temperature of order $\kappa^{-1/2}$. The observed interaction is decomposed into a bounded-excess reference measure, a small matrix, and a positive matrix of rank at most two. The first two terms are handled by Gaussian block resampling, and the final two-dimensional mixture is integrated directly.

Proposition 5.1. *There are absolute $c_{\mathrm{init}} > 0$ and C_{init} such that the following holds. Fix rational parameters with $\theta \le 1/30$, $\kappa \ge C_{\mathrm{init}}$, and a rational*

$$\frac{c_{\mathrm{init}}}{2\sqrt{\kappa}} \le t \le \frac{c_{\mathrm{init}}}{\sqrt{\kappa}}.$$

For every requested accuracy $\varepsilon \in (0,1)$ there is a graph-only sampler for $\nu_{t,G}$ with expected total variation error at most $\varepsilon + O(n^{-32})$. Its bit and random-bit costs have deterministic polynomial bounds in n and ε^{-1} for fixed parameters. On a sufficiently small compact parameter neighborhood, the same auxiliary choices, error bound, and polynomial apply uniformly, with polynomial dependence also on the rational input length.

Proof. Take $B = 32$ and $e_0 = 93$ in Lemma 4.3. Choose an integer $L \ge \max\{L_{32}, a\}$, of order a once C_{init} is large, and set $D_0 = 1024L$. Use the canonical partition (H, F, T, U) of the observed graph at this threshold. Write $p = d/n$, and define

$$\lambda_0 = t(pJ_n + \gamma_n), \qquad \lambda_1 = tpJ_n\theta,$$

$$S = t(J_n A_F + \gamma_n \Pi_U M_n \Pi_U) + (\lambda_0 + \lambda_1)\Pi_U,$$
$$E = t\gamma_n(M_n - \Pi_U M_n \Pi_U), \qquad R = tJ_n A_H + E.$$

For $n \ge 2a$, both λ_0 and λ_1 are nonnegative. For λ_0, this follows from monotonicity of $\operatorname{arctanh}(x)/x$ and $(1-p)\,p\theta/(1-p) = p\theta$.

Choose a sufficiently small absolute $b_0 > 0$, depending only on e_0, such that $3b_0 C_{e_0} \le 3/16$, where $C_{e_0} = 256\,4^{e_0}$ is the covariance constant in Lemma 3.1. From the observed matrix S, compute a positive semidefinite matrix C of rank at most two, supported on $U \times U$, whose error is within $b_0/100$ of the best such approximation in operator norm. One may use a rational approximate orthogonal diagonalization, retain the two largest positive diagonal entries, and refine until the off-diagonal error is below $b_0/1000$. If a rank-two positive semidefinite witness C_* obeys $\|S - C_*\| \le \delta$, the discarded eigenvalues lie in $[-\delta - o(b_0), \delta + o(b_0)]$, which proves the approximation assertion without requiring eigenvalue separation. The certified implementation is covered by Lemma A.1.

Fix this matrix once and put $K = S - C$. The decomposition

$$t(J_n A_g + \gamma_n M_n) = R + K + C - (\lambda_0 + \lambda_1)\Pi_U \tag{5.1}$$

is exact. The last term is diagonal and does not change the spin law. The algorithm accepts only if the component excess of H is at most e_0, the boundary-ball bounds hold, and

$$tJ_n\sqrt{D_0} \le \tfrac{1}{2}, \qquad q_{\mathrm{blk}} \le \tfrac{1}{2}, \qquad \delta \le 1, \qquad \|K\| \le b_0, \tag{5.2}$$

with the notation of Lemma 3.1. These conditions have finite certificates. In particular, writing $m_i = |Q_i|$ and $u_i = |Q_i \cap U|$, the intercomponent entries and the within-component energy are bounded by

$$W_{ij} = t|\gamma_n|(m_i m_j - u_i u_j), \qquad \delta = \frac{t|\gamma_n|}{2}\max_i\{m_i^2 - u_i^2 - (m_i - u_i)\}.$$

Use $\|W\|_F$ in place of q_{blk} for the test, and fixed certified precision with a strict acceptance margin. Rejection produces an arbitrary fixed spin.

We first describe sampling on a certified graph. Lemma 3.1 gives

$$\|\Pi_U \operatorname{Cov}_{\pi_{R,h}} \Pi_U\| \le C_{e_0} \qquad (h \in \mathbb{R}^n).$$

The law $\pi_{R,h}$ can be sampled and its partition function computed to arbitrary prescribed precision by Lemma 3.4. Indeed, take the reference signs to be $+1$ on U and -1 on T. If k counts plus spins and ℓ counts agreement with these signs, then

$$M(s) = 2k - n, \qquad \sum_{i \in U} s_i = k + \ell - n,$$

so the dense remainder depends only on these counts:

$$\frac{1}{2} s^\top E s = \frac{t\gamma_n}{2}\big\{(2k-n)^2 - (k+\ell-n)^2 - |T|\big\}.$$

The component excess e_0 is fixed, so pinning endpoints of its non-tree edges costs only a constant factor in the count dynamic program.

For $\alpha \in [0,1]$, add αK by the Gaussian sampler of Lemma 3.2, using the positive matrix $B_\alpha = \alpha K + 2b_0 \Pi_U$. The diagonal shift is immaterial, and

$$\|B_\alpha^{1/2} \operatorname{Cov}_{\pi_{R,h}} B_\alpha^{1/2}\| \le 3b_0 C_{e_0} \le 3/16.$$

The resulting Gaussian block kernel has gap at least $13/16$, uniformly in h and α. Each transition samples a Gaussian field and then the count reference law. Its mixing time from a fixed spin is polynomial because the logarithm of the reciprocal minimum spin mass is polynomially bounded for all fields used below.

These samplers also estimate the partition function $Z_1(h)$ of $\pi_{R+K,h}$. Along a mesh in α of spacing at most $1/\max\{1, \lceil b_0 n \rceil\}$, consecutive partition ratios are expectations of $\exp\{s^\top K s/(2N)\}$, which lies in $[e^{-1/2}, e^{1/2}]$. Polynomially many independent samples estimate each ratio to prescribed relative accuracy, with any prescribed inverse-polynomial failure probability. Multiplication, starting from the count-DP value $Z_0(h)$, gives a relative partition estimate with polynomial cost. All numerical errors are allocated using Lemma A.1.

Factor $C = L^\top L$ with two rows, allowing a zero row. Since the entries of S are uniformly bounded for fixed parameters, $\|C\|$ and each row sum $\sum_i |L_{ji}| \le \sqrt{n\|C\|}$ are polynomially bounded on every input. The Hubbard–Stratonovich representation [LMRW24, Theorem 3.12] makes $\nu_{t,g}$ the spin marginal of

$$e^{-\|y\|^2/2} \exp\{s^\top (R+K) s/2 + y^\top L s\}, \qquad y \in \mathbb{R}^2.$$

Conditional on s, y is $N(Ls, I_2)$. A polynomial-radius box therefore contains all but $\varepsilon/10$ of the mass. On this box, the log joint density is uniformly Lipschitz with a polynomial constant. A mesh of inverse-polynomial spacing in n and ε^{-1} changes its normalized spin marginal by at most $\varepsilon/10$. The number of cells is polynomial because the dimension is two. Estimate their weights $|Q| e^{-\|y_Q\|^2/2} Z_1(L^\top y_Q)$ to small relative error, sample the resulting categorical law, and sample the conditional spin at its chosen midpoint. A union bound for the partition estimates and the allocated numerical errors gives total error at most ε on every certified input. This construction retains the exact residual K from (5.1); numerical approximations do not redefine the target.

It remains to verify the certificate. Conditional on any planted labeling x, the graph is an independent-edge graph with probabilities at most a/n. Lemma 4.3 gives, outside probability n^{-32},

$$\max_Q |Q| \le n^{1/8}, \qquad S_2(H) \le e^{-D_0/16} n, \qquad \mathrm{exc}(Q) \le e_0.$$

The modified set $V'(H)$ has size at most $10n/a$, after increasing the absolute parameter floor. Applying [LLV17, Theorem 2.1] with exponent 32 and maximum retained degree $D_0 = O(a)$ gives, outside another probability n^{-32},

$$\|A_F - \mathbb{E}[A_g \mid X = x]\| \le C\sqrt{a}.$$

The conditional mean is $pM_n + p\theta(xx^\top - I_n)$. Hence

$$S = \lambda_0 \mathbf{1}_U \mathbf{1}_U^\top + \lambda_1 (\Pi_U x)(\Pi_U x)^\top + tJ_n \Pi_U (A_F - \mathbb{E}[A_g \mid X = x])\Pi_U.$$

The first two terms are the rank-two positive witness. The last has norm $O(t\theta\sqrt{d}) = O(c_{\mathrm{init}})$, so the constructed residual satisfies $\|K\| \le b_0/4$ when c_{init} is sufficiently small. Likewise,

$$tJ_n\sqrt{D_0} = O(c_{\mathrm{init}}), \qquad \|W\|_F \le t|\gamma_n| n \sqrt{e^{-D_0/8} + 2e^{-D_0/16}}, \qquad \delta \le \tfrac{1}{2} t|\gamma_n| n^{1/4}.$$

The second quantity is $O(c_{\mathrm{init}}\sqrt{d}\, e^{-cd})$ and the third tends to zero. They are therefore at most $1/4$ and $1/2$, respectively, after increasing C_{init} and then n. Every condition in (5.2) has a fixed strict margin on this event, so fixed-precision certification succeeds. This proves the expected error bound.

For local uniformity, choose t in the interior of its permitted interval and fix L above the largest a in a sufficiently small compact neighborhood. All certificate margins persist, all component and regularization estimates have the same failure probabilities, and all fields and numerical budgets admit common polynomial bounds there. The argument therefore gives the stated uniformity. The finitely many excluded sizes use a bounded fallback. □

6 Cooling the tempered law

Let $R_{\beta,g}$ be one random-scan single-site heat-bath update for $\nu_{\beta,g}$ followed by multiplication by an independent fair sign. On even functions, its Dirichlet form is the ordinary single-site Dirichlet form.

Proposition 6.1. *There are absolute $\theta_{\mathrm{w}} > 0$ and $C_{\mathrm{w}} \ge C_{\mathrm{init}}$ such that, if $\theta \le \theta_{\mathrm{w}}$ and $\kappa \ge C_{\mathrm{w}}$, the following holds. Fix t as in Proposition 5.1, and put $\rho_t = e^{-t\kappa/8}$. There is a polynomial A_n, of degree at most two, and functions $b_{n,\beta}(g) \in [0, 1/2]$ such that for every even f,*

$$\mathrm{Var}_{\nu_{\beta,g}} f \le A_n \mathcal{E}_{\nu_{\beta,g}, R_{\beta,g}}(f, f) + b_{n,\beta}(g)\, \mathrm{osc}(f)^2, \qquad t \le \beta \le 1,$$

for every graph g, and

$$\sup_{\beta \in [t,1]} \mathbb{E} b_{n,\beta}(G) \le 3n^{-16} + 4e^{-t\kappa\rho_t n/4}.$$

Proof. Choose the absolute constants so that $t \le 1/6$ and $t\kappa \ge 8$. For every planted labeling x and every s,

$$\nu_{\beta,G}(s) \le \min\{1, (L_s(G)/L_x(G))^\beta\} \le (L_s(G)/L_x(G))^t.$$

The Bernoulli affinity on a changed pair is at most $e^{-\kappa/(2n)}$. Interpolating its likelihood moment between exponents zero and $1/2$ gives

$$\mathbb{E}_{G|X=x}(L_s(G)/L_x(G))^t \le e^{-t\kappa k(n-k)/n}, \qquad k = d_H(s,x).$$

The same Hamming-layer summation as in Lemma 2.1, using $\log(en/k) \le t\kappa/4$ for $k > \rho_t n$, yields

$$\mathbb{E}_{X,G}\nu_{\beta,G}\{d_\pm(s,X) > \rho_t n\} \le 4e^{-t\kappa\rho_t n/4}. \tag{6.1}$$

Two replicas therefore satisfy $d_\pm(\sigma,\tau) \le 2\rho_t n$ outside probability $8e^{-t\kappa\rho_t n/4}$.

We next bound the orientation variance on disagreement sets of size $m \le 2\rho_t n$. Use the interval $[\ell, u]$ of Lemma 4.4. If $m \ge \sqrt{n}$, set

$$L = \lceil \max\{L_*, mu\} \rceil, \quad D = 1024L, \quad e_* = 93, \quad \Delta = D + 37\log m,$$

where the absolute L_* is at least $\max\{2, L_{32}\}$. Only finitely many L occur for fixed parameters. Lemma 4.3 and (4.3) give their conclusions outside probability at most $3n^{-16}$. Decompose the orientation interaction exactly as

$$2(J_n A + \gamma_n M_m) = K + J_H + E,$$

$$K = 2(J_n A_F + \gamma_n \Pi_U M_m \Pi_U), \qquad J_H = 2J_n A_H, \qquad E = 2\gamma_n(M_m - \Pi_U M_m \Pi_U).$$

Using (4.5) and (4.3), with $mp\theta^2 = \kappa m/n$, gives

$$\|K\| \le C\big(\theta + \sqrt{\kappa m/n} + \kappa m/n\big). \tag{6.2}$$

The component perturbation bounds are

$$q_{\mathrm{blk}} \le 6\theta L\sqrt{e^{-128L} + 2e^{-64L}}, \qquad \delta \le 3d\theta n^{-3/4},$$

and

$$2J_n\sqrt{D} \le 192\sqrt{L_* + 1}\,\theta + 192\sqrt{2\kappa m/n}.$$

Choose θ_{w} sufficiently small, and then C_{w} sufficiently large, so that

$$2\kappa e^{-c_{\mathrm{init}}\sqrt{\kappa}/16}$$

is smaller than an absolute constant sufficient for all three bounds. Since $m/n \le 2\rho_t$ and $t \ge c_{\mathrm{init}}/(2\sqrt{\kappa})$, every hypothesis of Proposition 3.3 holds, uniformly in $\beta \in [0,1]$ and any external field. Its single-site variance coefficient is at most

$$16e^{5+2e_*+32\Delta/D} \le Cn,$$

because $32 \cdot 37/D \le 1$. Keeping the factor D^{-1} in the tree entropy bound is useful here: the variance coefficient grows at most linearly with n.

For $m < \sqrt{n}$, couple the graph below $G(m, 2d/n)$. A component of size at least 40 contains a tree on 40 vertices, so the tree union bound gives

$$\mathbb{P}(\max_Q |Q| \ge 40) \le m(8dm/n)^{39} \le (8d)^{39} n^{-19}.$$

Otherwise use $H = A$, let U be its singleton vertices, and take $D = 1$, $\Delta = 40$, and excess bound 40^2. The bulk norm and block influence are $O_{a,b}(n^{-1/2})$, and the within-component perturbation

is $O_{a,b}(n^{-1})$. Proposition 3.3 again gives a fixed single-site variance coefficient for large n. The empty fiber has zero variance. Thus there is a deterministic notion of a good graph on m vertices such that, uniformly over the eligible sequential laws,

$$\mathbb{P}(A \text{ is not good}) \le 3n^{-16}, \qquad \mathrm{Var}_\lambda \varphi \le C_n \sum_{i=1}^{m} \mathbb{E}_\lambda \mathrm{Var}_\lambda(\varphi \mid s_{-i}), \tag{6.3}$$

where $C_n = Cn$ for a sufficiently large absolute integer C, and λ is the orientation law on a good fiber.

Fix g, align the second replica by its global sign, and write V for their disagreement set and y for the common exterior. Lemma 4.4 allows (6.3) to be applied to the conditional graph law. Let $\epsilon_{n,\beta}(g)$ be the aligned-pair probability that $|V| > 2\rho_t n$ or that $g[V]$ is not good. Equations (6.1) and (6.3) give

$$\mathbb{E}\epsilon_{n,\beta}(G) \le 3n^{-16} + 8e^{-t\kappa\rho_t n/4}.$$

For an even f, define $\Psi(s) = f(y,s) - f(y,-s)$ on the fiber. It is odd and hence centered under (4.4). Moreover,

$$\mathrm{Var}_{\nu_{\beta,g}} f = \tfrac{1}{2}\mathbb{E}_{\mathrm{al}}[f(\sigma) - f(\tau)]^2.$$

Applying (6.3) on good fibers and the oscillation bound on the others gives

$$\mathrm{Var}_{\nu_{\beta,g}} f \le \frac{C_n}{2} \sum_i \mathbb{E}_{\mathrm{al}}\big[\mathbf{1}_{\{i\in V\}} \mathrm{Var}(\Psi \mid s_{V\setminus\{i\}})\big] + \frac{\epsilon_{n,\beta}(g)}{2}\,\mathrm{osc}(f)^2.$$

To compare the first term with ordinary heat bath, condition on the two configurations outside i. Let p_i, q_i be their conditional plus probabilities and a_i, c_i the corresponding increments of f. Before alignment, the opposite-sign contribution to the fiber variance equals

$$\frac{p_i(1-p_i)q_i(1-q_i)}{p_i(1-q_i) + (1-p_i)q_i}(a_i + c_i)^2.$$

The denominator is at least each of $p_i(1-p_i)$ and $q_i(1-q_i)$. Alignment multiplies masses by at most two, so integration bounds this contribution by $8\mathbb{E}_{\nu_{\beta,g}}\mathrm{Var}(f \mid s_{-i})$. Consequently the asserted inequality holds with $A_n = 4nC_n$ and $b_{n,\beta} = \epsilon_{n,\beta}/2$. The expectation bound follows. □

Theorem 6.2. *There are absolute $\theta_{\mathrm{w}} > 0$ and C_{w} such that, for every fixed rational $a > b > 0$ with $\theta \le \theta_{\mathrm{w}}$ and $\kappa \ge C_{\mathrm{w}}$, there is a graph-only posterior sampler with deterministic polynomial bit and random-bit bounds on every input graph and*

$$\mathbb{E}_G \mathrm{TV}(\widehat{\mu}_G, \mu_G) = o(1).$$

No additional condition on d is imposed. About every parameter pair with $\theta < \theta_{\mathrm{w}}$ and $\kappa > C_{\mathrm{w}}$, there is a compact rational rectangle on which the same auxiliary choices give a common error $\psi(n) + \varepsilon$, with $\psi(n) \to 0$, and a common polynomial bound in n, the rational input length, and ε^{-1}.

Proof. The single-site kernel is an average of conditional-expectation projections. It is positive semidefinite and commutes with the global-sign projection. Thus $R_{\beta,g}$ is also positive semidefinite. The even weak Poincaré estimate of Proposition 6.1, together with Lemma 2.3, gives, for any symmetric law $\lambda \le M\nu_{\beta,g}$,

$$\mathrm{TV}(\lambda R_{\beta,g}^T, \nu_{\beta,g}) \le \sqrt{M}\left(\tfrac{1}{2}e^{-T/(2A_n)} + \sqrt{b_{n,\beta}(g)}\right).$$

Initialize using Proposition 5.1 to accuracy n^{-4} and symmetrize by a fair sign. Choose

$$N = 4n^2, \qquad \beta_j = t + (1-t)j/N \quad (0 \le j \le N), \qquad T = 20A_n\lceil \log_2(n+1)\rceil.$$

The spin-dependent log likelihood $H_g(s) = J_n F_g(s) + \gamma_n M(s)^2/2$ has oscillation at most $2n^2$ on every graph for $n \ge 2a$ and $\theta \le 1/30$. Hence

$$\nu_{\beta_{j-1},g} \le e\,\nu_{\beta_j,g}.$$

At stage j, run T transitions of $R_{\beta_j,g}$. Compare this stage with the same kernel started from the exact preceding target; total variation contraction carries the previous error forward. Thus warmth is required only of consecutive exact targets. Summing the displayed warm-start bound and using Jensen's inequality gives expected endpoint error at most

$$n^{-4} + O(n^{-32}) + \frac{\sqrt{e}}{2} N n^{-10} + \sqrt{e}\, N\sqrt{3n^{-16} + 4e^{-t\kappa\rho_t n/4}} = o(1).$$

The endpoint target is $\nu_{1,g} = \mu_g$. Since $A_n = O(n^2)$, this cooling stage uses $O(n^4 \log n)$ heat-bath transitions; the initialization and the arithmetic costs are additional.

At site i, the exact conditional log odds are

$$2\beta_j \left(J_n \sum_{k:\, ik \in E(g)} s_k + \gamma_n \sum_{k \ne i} s_k \right).$$

Lemma A.1 evaluates these probabilities and implements the site and sign choices with deterministic bit budgets. Taking each row error at most $n^{-4}/(NT)$ adds at most n^{-4} to the endpoint error. All caps hold on arbitrary graphs, including those on which the initializer rejects its certificate.

For local uniformity, fix an interior t and the excision degree used in Proposition 5.1 on a sufficiently small compact rectangle. There are finitely many disagreement-degree choices L on this rectangle. All strict small-interaction inequalities persist, $t\kappa\rho_t$ has a positive common lower bound, and A_n has a common polynomial majorant. The preceding proof therefore has a common vanishing statistical error and a common polynomial implementation bound. Replacing the numerical tolerance by a smaller value if needed gives the stated $\psi(n) + \varepsilon$ guarantee. □

7 A recovered phase and its fluctuations

We next give a sampler in the regime

$$\kappa \ge C_{\mathrm{ph}}(1 + \log d), \tag{7.1}$$

where C_{ph} is an absolute constant, chosen sufficiently large below. The construction starts from the recovery algorithm in Proposition 2.2. Its purpose is to locate a phase; the subsequent sampler must still generate the uncertainty within that phase. Throughout this section, put

$$r = e^{-\kappa/64}, \qquad \epsilon = r^2, \qquad \rho = 2r^3, \qquad h_{\max} = \tau_n(2rn+1).$$

All assertions concern sufficiently large n, with the threshold allowed to depend on the fixed parameters. We may take $C_{\mathrm{ph}} \ge 2^{30}$ and increase it a finite number of times. This choice is deliberately conservative.

7.1 A deterministic core construction

For a vertex set D and an integer $t \geq 1$, let $\mathrm{cl}_t(D)$ be its closure under repeatedly adding vertices with at least t neighbors already present. This closure does not depend on the order of additions. For a spin w, its signed neighbor sum at v is $w_v \sum_{u \sim v} w_u$.

Run the recovery algorithm and call its output $\widehat{z}$. Repeatedly flip the least vertex with negative signed neighbor sum, stopping when none remains; denote the resulting spin by z. Define

$$B_0 = \{v : \deg_g(v) \notin [d/4, 4d]\} \cup \left\{ v : z_v \sum_{u \sim v} z_u < \theta d/4 \right\},$$
$$B = \mathrm{cl}_{\lceil \theta d/32 \rceil}(B_0), \qquad K = [n] \setminus B. \tag{7.2}$$

Every flip increases the integer F_g by at least two, so there are at most $|E(g)|$ flips. Thus the construction has deterministic polynomial bit cost on every graph. The small magnetization obtained below is sufficient for the later scalar-field correction.

Proposition 7.1. *Assume (7.1). More generally, the initializer may be any graph-only rule satisfying*

$$\mathbb{P}\{d_{\pm}(\widehat{z}, X) > \rho n\} = o(1).$$

With probability $1 - o(1)$, the output of (7.2) satisfies

(i) $d_{\pm}(z, X) \leq rn/4$, $|M(z)| \leq rn/8$, *and* $|B| \leq rn/8$;

(ii) every component of $g[B]$ has size at most $\lceil 512 \log n/\kappa \rceil$, and the maximum degree of $g[K]$ is at most $4d$;

(iii) for every $\eta \in \{-1, +1\}^B$, every nonempty connected $S \subset K$ of size at most rn, and every $|h| \leq h_{\max}$, writing $y = (z_K, \eta)$ and letting y^S negate y on S, we have

$$\exp\{J_n[F_g(y^S) - F_g(y)] + h[M(y^S) - M(y)]\} \leq e^{-\kappa|S|/8}. \tag{7.3}$$

Proof. The argument uses small-set sparsity and the fact that atypical planted neighbor sums are rare. Set

$$\alpha = \frac{\theta d}{2048}, \qquad t = \left\lceil \frac{\theta d}{64} \right\rceil, \qquad p_0 = 3e^{-\kappa/16}.$$

The parameter assumption implies, with ample slack,

$$\begin{gathered} \alpha \geq 8, \quad r \leq \theta/128, \quad 64\epsilon \leq r, \quad p_0 \leq \epsilon/4, \quad 8192r^2 < \theta, \\ \log(ed/\alpha) \leq \kappa/256, \qquad \log(64e/r) \leq \kappa/32, \qquad 4ea\sqrt{p_0} \leq e^{-\kappa/128} < 1/2. \end{gathered} \tag{7.4}$$

Indeed, $d \geq \kappa$, $a \leq 2d$, and $\log(1/\theta) \leq \frac{1}{2} \log d$; taking logarithms reduces all these comparisons to (7.1).

We first condition on a planted vector with $|M(X)| \leq n^{2/3}$. The probability of the complementary event tends to zero by Hoeffding's inequality. Conditional on this vector, the edges are independent and have probabilities at most $a/n \leq 2d/n$. We use the standard Chernoff and Bernstein inequalities throughout; see [BLM13, Chapter 2].

Write $e_g(U)$ for the number of edges inside U. For $1 \leq j \leq rn$, Chernoff's inequality and a union bound give

$$\mathbb{P}\{\exists U : |U| = j,\ e_G(U) > \alpha j \mid X\} \leq \left[\frac{en}{j} \left(\frac{edj}{\alpha n} \right)^{\alpha} \right]^j \leq (j/n)^{3j}.$$

The last inequality follows from (7.4) by putting $v = \log(n/j) \geq \kappa/64$. Summing first over $j \leq \sqrt{n}$ and then over $j > \sqrt{n}$ proves that, with probability $1 - o(1)$,

$$e_G(U) \leq \alpha|U| \qquad (|U| \leq rn). \tag{7.5}$$

Let $j_* = \lceil rn/64 \rceil$ and $T_X(U) = \sum_{ij \in \partial_G U} X_i X_j$. For a fixed j_*-set, writing $A_U = \sum_{i \in U} X_i$, we have

$$\mathbb{E}[T_X(U) \mid X] = \frac{d}{n}\{\theta j_*(n - j_*) + A_U(M(X) - A_U)\} \geq \frac{31}{32}\theta d j_*.$$

Its conditional variance is at most $33dj_*/32$ for all sufficiently large n. Bernstein's inequality bounds the probability that $T_X(U) < \theta dj_*/4$ by $e^{-\kappa j_*/8}$. A union bound using (7.4) therefore gives

$$T_X(U) \geq \theta d|U|/4 \qquad (|U| = j_*) \tag{7.6}$$

with probability $1 - o(1)$. The same Chernoff argument shows that every set of size $j_{\mathrm{in}} = \lceil \rho n \rceil$ is incident to at most $8dj_{\mathrm{in}}$ edges, with probability $1 - o(1)$: the conditional mean for a fixed set is at most $2dj_{\mathrm{in}}$, whereas the union-bound entropy is at most $j_{\mathrm{in}} \log(en/j_{\mathrm{in}})$ and $\log(en/j_{\mathrm{in}}) \leq 1 + 3\kappa/64$.

Define the planted exceptional set

$$A_X = \{v : \deg_G(v) \notin [d/4, 4d]\} \cup \left\{ v : X_v \sum_{u \sim v} X_u < \theta d/2 \right\}. \tag{7.7}$$

For a deterministic set U with $|U| \leq rn$ and $v \in U$, let $C_v(U)$ be the union of the three events

$$\deg_{U^c}(v) < d/4, \qquad \deg_{U^c}(v) > 4d - \theta d/32, \qquad X_v \sum_{u \notin U \colon u \sim v} X_u < 17\theta d/32. \tag{7.8}$$

The conditional external-degree mean lies in $[31d/32, 33d/32]$, the signed-sum mean is at least $31\theta d/32$, and both variances are at most $33d/32$. These bounds follow directly from $|M(X)| \leq n^{2/3}$, $|U| \leq rn$, and $r \leq \theta/128$. Bernstein's inequality, applied with deviations $23d/32$, $94d/32$, and $7\theta d/16$, yields

$$\mathbb{P}(C_v(U) \mid X) \leq p_0. \tag{7.9}$$

For fixed U, these events are mutually independent and independent of the edges inside U, conditional on X. Moreover,

$$v \in A_X, \quad |N_G(v) \cap U| \leq \theta d/32 \quad \Longrightarrow \quad C_v(U). \tag{7.10}$$

The singleton bounds give $\mathbb{E}(|A_X| \mid X) \leq p_0 n$. Two membership indicators share only their possible connecting edge, so their covariance has absolute value at most a/n. Hence

$$\mathrm{Var}(|A_X| \mid X) \leq (1 + a)n, \qquad \mathbb{P}\{|A_X| > \epsilon n \mid X\} = o(1).$$

On (7.5), every nonempty D with $2|D| \leq rn$ satisfies

$$|\operatorname{cl}_t(D)| < 2|D|. \tag{7.11}$$

Otherwise, at the first size $2|D|$, the added vertices supply $t|D|$ distinct internal edges, charged to the later endpoint. This contradicts $t > 2\alpha$. The empty set has empty closure.

We now intersect these events with recovery success and orient X so that $d_H(\widehat{z}, X) \leq \rho n$. The graph events are invariant under this global sign choice. The incident-edge bound gives

$$F_G(X) - F_G(\widehat{z}) \leq 16dj_{\mathrm{in}} \leq 32d\rho n.$$

All local-ascent iterates have at least this energy. An iterate with exactly j_* errors would, by (7.6), have energy deficit at least $\theta d r n/128 > 32 d\rho n$. Since the number of errors changes by one at each flip, it remains below j_*.

Put $W = \mathrm{cl}_t(A_X)$ and $E_0 = \{v : z_v \neq X_v\}$. By (7.11), $|W| \leq 2\epsilon n$. If $E_0 \setminus W$ were nonempty, local stability would imply, for each of its vertices,

$$0 \leq z_v \sum_{u \sim v} z_u \leq -\theta d/2 + 2|N_G(v) \cap E_0|.$$

Since a vertex outside W has fewer than $\theta d/64$ neighbors in W, every vertex of $E_0 \setminus W$ would have more than $15\theta d/64$ neighbors in that set. This contradicts (7.5). Thus

$$E_0 \subset W, \qquad |W| \leq 2\epsilon n. \tag{7.12}$$

In particular, $d_H(z, X) \leq 2\epsilon n$ and $|M(z)| \leq n^{2/3} + 4\epsilon n \leq rn/8$ for all sufficiently large n. The last inequality uses the fixed strict comparison $4\epsilon < r/8$, rather than an asymptotic relation between the fixed parameters.

Outside W, the spins z and X agree and the z-margin is greater than $\theta d/2 - 2\theta d/64 > \theta d/4$. Thus $B_0 \subset W$. Since the closure threshold defining B is at least t, also $B \subset W$. Consequently, for every $v \in K$,

$$\deg_G(v) \leq 4d, \qquad z_v \sum_{u \sim v} z_u \geq \theta d/4, \qquad |N_G(v) \cap B| < \theta d/32. \tag{7.13}$$

We also need logarithmic exceptional components. If U is a component of $G[W]$ of size $j \leq rn$, at most $j/32$ of its vertices were added to the closure: charge their t earlier edges and use (7.5). At most a further $j/32$ vertices have more than $\theta d/32$ neighbors in U. By (7.10), at least $15j/16$ vertices satisfy $C_v(U)$. Choose $\lceil j/2 \rceil$ of them and a spanning tree on U. Conditional on X, the probability of a fixed witness is at most $(a/n)^{j-1} p_0^{j/2}$. The labeled-tree count and a union bound yield

$$\mathbb{P}\{\text{such a witness of size } j \mid X\} \leq \binom{n}{j} j^{j-2} (a/n)^{j-1} 4^j p_0^{j/2} \leq n e^{-\kappa j/128}.$$

Summing for $j \geq \lceil 512 \log n/\kappa \rceil$ gives $O(n^{-3})$. This calculation is made under the original conditional edge law, before imposing any recovery or graph event. It therefore combines with those events by a union bound. Since $B \subset W$, it proves the component assertion.

Finally, for $S \subset K$ with $|S| \leq rn$ and $y = (z_K, \eta)$, (7.5) and (7.13) give

$$\sum_{ij \in \partial_G S} y_i y_j \geq \frac{\theta d}{4}|S| - 2e_G(S) - 2e_G(S, B) \geq \frac{\theta d}{8}|S|.$$

For $n \geq 2a$, $J_n \geq \theta$ and $\tau_n \leq 2d\theta/n$. Thus $h_{\max} \leq 5d\theta r \leq J_n \theta d/16$ for large n, and the logarithm in (7.3) is at most $-J_n \theta d|S|/8 \leq -\kappa|S|/8$. Integrating the conditional estimates over X completes the proof. □

7.2 Joint fluctuations and component tails

We analyze the joint distribution of the exceptional spins and the core defects. For $s \in \{-1, +1\}^n$, set

$$D_z(s) = \{v \in K : s_v \neq z_v\}.$$

Let Ω_m be the set of spins for which every component of $g[D_z(s)]$ has at most m vertices. For $0 < u \le 1$, define

$$Q_{h,u}^{(m)}(s) \propto \mathbf{1}_{\Omega_m}(s) e^{J_n F_g(s) + hM(s)} u^{|D_z(s)|}. \tag{7.14}$$

For a positive radius parameter R, let

$$P_h^R(s) \propto \mathbf{1}_{\{d_H(s,z) \le Rn\}} e^{J_n F_g(s) + hM(s)}. \tag{7.15}$$

These laws omit the quadratic nonedge interaction; it will be restored exactly in the final sampling step.

Proposition 7.2. *Fix a rational $R \in [3r/4, 7r/8]$ and an integer $S \ge 1$. Put*

$$c_S = \left\lceil \frac{\max\{4096, 128(2S+16)\}}{\kappa} \right\rceil, \quad L_S = c_S \lceil \log_2 n \rceil, \quad K_0 = 12 + \frac{8192}{\kappa}.$$

On graph/construction events of probability $1 - o(1)$, simultaneously for $|h| \le h_{\max}$, $0 < u \le 1$, and every integer $\lfloor Rn \rfloor \le m \le \lfloor rn \rfloor$, we have

$$Q_{h,u}^{(m)}\{d_H(s,z) > Rn\} \le e^{-\kappa r n/64}, \tag{7.16}$$

$$(Q_{h,u}^{(m)} \otimes Q_{h,u}^{(m)})\{c_{\max}(g[B \cup D_z(\sigma) \cup D_z(\tau)]) \ge L_S\} \le n^{-2S-10} + 2e^{-\kappa r n/256}. \tag{7.17}$$

Here $c_{\max}$ denotes the largest component size, with value zero for the empty graph. On the same events,

$$\operatorname{Var}_{P_h^R}(v \cdot s) \le K_0 \log n \, \|v\|_2^2 \qquad (v \in \mathbb{R}^n, |h| \le h_{\max}). \tag{7.18}$$

Proof. Fix an output satisfying Proposition 7.1. Conditional on $s_B = \eta$, a spin in Ω_m is determined by the connected components of its core defect set. Call these components polymers. Distinct polymers are disjoint and have no core edge between them. If $y_\eta = (z_K, \eta)$, their activities are

$$w_{\eta,h,u}(A) = u^{|A|} \exp\{J_n[F_g(y_\eta^A) - F_g(y_\eta)] + h[M(y_\eta^A) - M(y_\eta)]\} \le e^{-\kappa|A|/8}.$$

The unnormalized joint weight of η and a compatible collection Γ is

$$e^{J_n F_g(y_\eta) + hM(y_\eta)} \prod_{A \in \Gamma} w_{\eta,h,u}(A).$$

In particular, the marginal of η includes its conditional polymer partition function. All estimates below hold uniformly in η, and are averaged over this actual marginal.

Deleting specified occupied polymers injects their collections into all compatible collections. Hence their joint occupation probability is at most the product of their activities. With $\Delta = \max\{2, \lceil 4d \rceil\}$, the number of connected j-vertex core sets containing a given vertex is at most $\Delta^{2(j-1)}$: a deterministic depth-first traversal of a spanning tree encodes the set by a walk of length $2(j-1)$. The parameter assumption gives

$$\Delta^2 e^{-\kappa/16} \le \tfrac{1}{2}, \quad \Delta^2 e^{-3\kappa/32} \le \tfrac{1}{2}, \quad 2e^{-\kappa/16} \le \kappa/48, \quad 2e^{-3\kappa/32} \le \kappa r/512.$$

For a fixed $I \subset K$, root each occupied polymer meeting I at its least vertex in I. Their total size is at least $|I|$. Splitting each activity bound into two equal exponential factors, then dropping compatibility, gives

$$Q_{h,u}^{(m)}\{I \subset D_z(s) \mid \eta\} \le e^{-\kappa|I|/16} \left(1 + \sum_{j \ge 1} \Delta^{2(j-1)} e^{-\kappa j/16}\right)^{|I|} \le e^{-\kappa|I|/24}.$$

After averaging over η and assigning each vertex to one of two independent replicas, we obtain

$$(Q_{h,u}^{(m)} \otimes Q_{h,u}^{(m)})\{I \subset D_z(\sigma) \cup D_z(\tau)\} \le e^{-\kappa|I|/32} =: \zeta^{|I|}. \tag{7.19}$$

The event is empty if I meets B, so the bound also holds for arbitrary $I \subset [n]$.

Expanding the product of polymer occupation indicators and using the same deletion bound yields

$$\mathbb{E}_{Q_{h,u}^{(m)}} e^{\kappa|D_z|/32} \le \exp\left\{\sum_A e^{-\kappa|A|/8}(e^{\kappa|A|/32} - 1)\right\} \le e^{2ne^{-3\kappa/32}} \le e^{\kappa rn/512}.$$

Consequently,

$$Q_{h,u}^{(m)}\{|D_z| > rn/4\} \le e^{-\kappa rn/256}. \tag{7.20}$$

Since $|B| \le rn/8$ and $R \ge 3r/4$, leaving the ball requires more than $5rn/8$ core defects. The same exponential-moment estimate gives (7.16). Every spin in the ball belongs to Ω_m, so

$$P_h^R = Q_{h,1}^{(m)}(\,\cdot \mid d_H(s,z) \le Rn), \qquad \mathrm{TV}(P_h^R, Q_{h,1}^{(m)}) \le e^{-\kappa rn/64}. \tag{7.21}$$

We next prove a single graph-dependent witness bound, uniform in the field, fugacity, and cutoff. Retain W, A_X, and the external tests from the proof of Proposition 7.1. On its good event, $B \subset W$, $|W| \le rn/8$, and (7.5) holds. For two independent replicas put

$$A^+ = W \cup D_z(\sigma) \cup D_z(\tau).$$

By (7.20), the probability that $|A^+| > rn$ is at most $2e^{-\kappa rn/256}$. If U is a component of $g[A^+]$ of size $j \le rn$, the closure-edge and internal-degree counts used above show that at least $15j/16$ vertices either satisfy $C_v(U)$ or belong to the union of the replica defects. Choose $\lceil j/2 \rceil$ such vertices and a spanning tree on U.

For each $j \ge 2$, define Ξ_j on the original graph experiment by summing, over all these choices and partitions $I_0 \cup I_1$ of the selected vertices, the quantities

$$\mathbf{1}_{\{T \subset G\}} \prod_{v \in I_0} \mathbf{1}_{C_v(U)}\, \zeta^{|I_1|}.$$

For a fixed good environment, (7.19) bounds the remaining replica event. Therefore, for every integer $2 \le \ell \le \lfloor rn \rfloor$,

$$\mathbb{P}\{g[A^+] \text{ has a component of size in } [\ell, rn]\} \le \sum_{j=\ell}^{\lfloor rn \rfloor} \Xi_j. \tag{7.22}$$

The right side contains no h, u, or m.

Conditional only on a planted vector with $|M(X)| \le n^{2/3}$, the internal tree edges and external tests are independent. Since $p_0 \le \zeta$, each witness has expectation at most $(a/n)^{j-1}\zeta^{\lceil j/2 \rceil}$. There are at most 4^j choices of the partitioned vertices and j^{j-2} trees on a specified set. Thus

$$\mathbb{E}[\Xi_j \mid X] \le n(4ea\sqrt{\zeta})^j \le ne^{-\kappa j/128}, \qquad \mathbb{E}\left[\sum_{j=\ell}^{\lfloor rn \rfloor} \Xi_j \,\middle|\, X\right] \le 2ne^{-\kappa\ell/128}. \tag{7.23}$$

No recovery success or good graph event has been conditioned on in this expectation. Taking $\ell = L_S$ and using Markov's inequality shows that the right side of (7.22) is at most n^{-2S-10} except on an

event of probability $O(n^{-5})$. This proves (7.17), since $B \subset W$. Taking instead $\ell_0 = \lceil 4096 \log n/\kappa \rceil$ shows, on another event of probability $1 - o(1)$, that the pair-component tail at ℓ_0 is at most

$$\delta_0 = n^{-10} + 2e^{-\kappa r n/256}.$$

Both events are uniform in h, u, m.

For completeness, the covariance bound uses component exchanges under this joint law, a version of the replica-cluster operation of [Hou01]. Abbreviate $Q = Q_{h,1}^{(m)}$. For independent $\sigma, \tau \sim Q$, let $A = B \cup D_z(\sigma) \cup D_z(\tau)$ and set

$$\xi_U = \sum_{i \in U} v_i(\sigma_i - \tau_i), \qquad U \in \mathrm{Comp}(g[A]).$$

Exchanging the two restrictions on one entire component U preserves $Q(\sigma)Q(\tau)$. Vertex factors are preserved coordinatewise, internal edge factors are exchanged, and every boundary edge meets a reference spin in both replicas. The cutoff is also preserved: every core defect component stays inside its uncertainty component and cannot merge through the boundary. The exchange negates ξ_U and leaves the other component contributions unchanged. All cross terms consequently vanish, and

$$\mathrm{Var}_Q(v \cdot s) = \frac{1}{2}\mathbb{E}\sum_U \xi_U^2 \le 2\mathbb{E}\sum_U |U| \sum_{i \in U} v_i^2 \le 2(\ell_0 + n\delta_0)\|v\|_2^2.$$

Since $|v \cdot s| \le \sqrt{n}\|v\|_2$, transferring to the ball by (7.21) adds at most $4ne^{-\kappa r n/64}\|v\|_2^2$. This proves (7.18). Although the proof used X, the conclusions concern only the graph and construction output; their failure probability in that marginal experiment is at most the extended failure probability just bounded. □

8 Component updates and the exact posterior

8.1 A general initialized component sampler

We first isolate the sampling argument from the graph estimates. Let g, z, B, K, Ω_m, and D_z be as above, but now allow arbitrary positive vertex and edge factors:

$$w(s) = \prod_v \phi_v(s_v) \prod_{vw \in E(g)} \psi_{vw}(s_v, s_w), \qquad \pi_u(s) \propto \mathbf{1}_{\Omega_m}(s) w(s) u^{|D_z(s)|}.$$

For $1 \le L \le m$, let $\mathcal{C}_L$ consist of all nonempty connected vertex sets of size at most L, and let $N_L = |\mathcal{C}_L|$. The kernel P_u chooses a set in $\mathcal{C}_L$ uniformly and performs its exact heat-bath update.

Lemma 8.1. *Put*

$$\delta_u = (\pi_u \otimes \pi_u)\{c_{\max}(g[B \cup D_z(\sigma) \cup D_z(\tau)]) > L\}.$$

Then P_u is reversible and positive semidefinite, and

$$\mathrm{Var}_{\pi_u} f \le n N_L \mathcal{E}_{\pi_u, P_u}(f, f) + \frac{\delta_u}{2}\,\mathrm{osc}(f)^2. \tag{8.1}$$

Proof. Call a pair (σ, τ) good when all its uncertainty components have size at most L. Order these components by their least vertex, and replace the restrictions of σ by those of τ in this order, skipping unchanged restrictions. The path has at most n moves. At an intermediate state α, let ω be the complementary assignment: it uses the original σ restrictions on processed components

and the original τ restrictions on the others. The component-exchange argument in Proposition 7.2 gives

$$A(\alpha,\omega)=A(\sigma,\tau),\qquad \pi_u(\sigma)\pi_u(\tau)=\pi_u(\alpha)\pi_u(\omega),$$

and both intermediate configurations remain in Ω_m.

Fix a directed move (α,β) and its component C. The map from endpoint pairs using this move to their companions ω immediately before it is injective: α,ω determine the ordered components, and C determines which have already been processed. Also $\omega_C=\beta_C$. Every neighbor of C outside C is a reference core vertex in both exteriors. Under either exterior, every assignment on C is admissible because $|C|\le m$ and no outside defect is adjacent to it. Hence the conditional laws on C agree. Writing $q_C=\pi_u(s_C=\beta_C\mid s_{C^c}=\alpha_{C^c})$, we obtain

$$\sum_{(\sigma,\tau):\,\text{path uses }(\alpha,\beta,C)}\pi_u(\sigma)\pi_u(\tau)\le\pi_u(\alpha)\sum_{\omega\ \text{eligible}}\pi_u(\omega)\le\pi_u(\alpha)q_C.$$

In the middle sum only eligible companions are included; conditioning on each eligible exterior gives the final inequality. The stationary transition mass for this block choice is $\pi_u(\alpha)q_C/N_L$. The standard multicommodity-flow comparison [Sin92], with paths of length at most n, bounds the good-pair part of the variance by $nN_L\mathcal{E}_{\pi_u,P_u}(f,f)$. The omitted pairs contribute at most $\delta_u\,\mathrm{osc}(f)^2/2$, giving (8.1). Each heat-bath operator is an orthogonal conditional-expectation projection, so their average is reversible and positive semidefinite. □

Proposition 8.2. *Suppose every component of $g[B]$ has size at most L, the maximum core degree is at most $\Delta\ge 2$, and*

$$\frac{w(y_\eta^A)}{w(y_\eta)}\le e^{-a_0|A|}$$

for every exceptional assignment η and every connected core set A of size at most m, where $\Delta^2e^{-a_0}\le 1/2$. Suppose also that $\sup_{0<u\le 1}\delta_u\le\delta$. Let ν fix the core at z_K and give the exceptional spins mass proportional to $w(z_K,\eta)$. Given $0<\varepsilon_0<1$, put

$$\varrho=1+1/n,\qquad J=\min\{j:\varrho^j\ge 2n/\varepsilon_0\},\qquad u_j=\varrho^{-j}.$$

Starting from ν, run T updates successively at $u_{J-1},\dots,u_0=1$. Its output law λ satisfies

$$\mathrm{TV}(\lambda,\pi_1)\le\varepsilon_0+\frac{\sqrt{e}}{2}Je^{-T/(2nN_L)}+J\sqrt{e\delta/2},\qquad J\le 1+2n\log(2n/\varepsilon_0).\tag{8.2}$$

The initial law factors over the components of $g[B]$. If the factors are positive rationals and an enumeration cap is at least N_L, the procedure has a finite-bit implementation with arbitrarily prescribed additional TV error ξ, at cost polynomial in

$$n,\ b,\ J,\ JT,\ N_L,\ 2^L,\ \log(1/\xi),$$

where b is their total bit length. All computation and random-bit bounds are deterministic.

Proof. The conditional polymer deletion bound and the connected-set count give

$$\pi_u\{D_z\ne\varnothing\mid s_B=\eta\}\le n\sum_{j\ge 1}\Delta^{2(j-1)}u^je^{-a_0j}\le 2nu.$$

This holds uniformly in η. Conditional on $D_z=\varnothing$, the law is exactly ν, so $\mathrm{TV}(\nu,\pi_{u_J})\le\varepsilon_0$. With the core fixed, the edge factors factor over the exceptional components; their laws can be sampled by enumerating at most 2^L assignments per component.

Consecutive laws have the density-ratio bound

$$\frac{\pi_{u_{j+1}}(s)}{\pi_{u_j}(s)} = \frac{\varrho^{-|D_z(s)|}}{\mathbb{E}_{\pi_{u_j}}\varrho^{-|D_z|}} \le \varrho^n < e.$$

Apply Lemma 2.3 to (8.1), with initial density bound e, coefficient nN_L, and defect $\delta/2$. The resulting one-stage error is at most $\frac{\sqrt{e}}{2}e^{-T/(2nN_L)} + \sqrt{e\delta/2}$. Contraction of TV under Markov kernels and a telescoping sum prove (8.2). The bound on J follows from $\log(1+1/n) \ge 1/(2n)$.

For the implementation, all connected sets can be enumerated without repetition by reverse search. The parent of a connected set of size greater than one deletes the largest vertex other than its minimum whose deletion leaves it connected. Such a vertex exists by taking a nonroot leaf of a spanning tree. Starting at singletons, test each adjacent candidate addition and recurse precisely when its parent is the current set. This costs $O(n^2L^3(N_L+1))$ graph operations, and stopping at any prescribed output cap bounds the cost on every graph. A heat-bath row requires at most 2^L assignments; membership in Ω_m is checked by a graph search.

The fugacity $u_j = (n/(n+1))^j$ has bit length $O(J\log(n+1))$. Products and sums defining the conditional rows therefore have bit length polynomial in the displayed quantities. Lemma A.1 supplies the required finite-bit categorical draws and bounds the accumulated TV error by ξ. The same construction applies to the initial exceptional-component draws and the uniform block choice. A cap smaller than N_L is used only as a detectable fallback, with no sampling claim on that input. □

8.2 Sampling the ball law

Proposition 8.3. *Under (7.1), fix a rational $R \in [3r/4, 7r/8]$ and an integer $S \ge 1$. On events of probability $1-o(1)$, given the constructed g, z, B and any rational c with $|c| \le 2rn+1$, a graph-only initialized algorithm samples $P^R_{-\tau_n c}$ within TV error n^{-S}. Its bit and random-bit costs are deterministically polynomial in n and the bit length of c, on every input graph. The error bound is uniform in all such c on the same good event.*

Proof. Take $m = \lfloor Rn \rfloor$ and use $L = L_S$ from Proposition 7.2. For the observed-edge and field weight $w(s) = e^{J_n F_g(s) - \tau_n c M(s)}$, Proposition 7.1 verifies the activity hypothesis of Proposition 8.2 with $a_0 = \kappa/8$. The uniform pair defect is at most $2n^{-2S-10}$ for all sufficiently large n.

To bound the block enumeration, conditional on the planted labels a fixed tree of size j occurs with probability at most $(a/n)^{j-1}$. Thus

$$\mathbb{E}N_{L_S}(G) \le n + \sum_{j=2}^{L_S} \binom{n}{j} j^{j-2}(a/n)^{j-1} \le nL_S(ea)^{L_S}.$$

Choose a fixed integer $D_S \ge 4c_S(1+a)+10$. Markov's inequality gives $\mathbb{P}\{N_{L_S}(G) > n^{D_S}\} = o(1)$. Enumerate only up to this cap. Also check $m \ge L_S$, the core-degree bound, and the bound L_S on exceptional-component sizes; use a fixed output on any failure. These checks succeed on the good events and make all remaining work deterministically bounded.

In Proposition 8.2, take

$$\varepsilon_0 = n^{-S-3}, \qquad J = \min\{j : (1+1/n)^j \ge 2n^{S+4}\}, \qquad T = \lceil 4nN_{L_S}(g)(S+10)\lceil \log_2 n \rceil \rceil .$$

Then $J = O_S(n\log n)$, and (8.2) is at most n^{-S-2} for large n. The initial exceptional components and all update blocks have at most $L_S = O(\log n)$ vertices, so their enumeration costs are polynomial.

For finite precision, replace the edge factor by

$$e^{2J_n} = \frac{a(n-b)}{b(n-a)}$$

on agreeing spins and by 1 on disagreeing spins. This is rational for rational parameters. Replace the field factor $e^{-2\tau_n c}$ on a plus spin by a positive rational approximation with logarithmic error at most

$$\eta \le \frac{n^{-S-2}}{16n^2(JT+1)}.$$

Every full log weight then changes by at most $n\eta$; normalization changes each conditional row, and the full initial law, by at most $4n^2\eta$ in TV. Lemma A.1, with a further error budget $n^{-S-2}/4$ for categorical draws, bounds the total numerical error by n^{-S-2}. Since $|\tau_n c|$ is bounded by a parameter-dependent constant, the required scalar precision and its bit cost are polynomial in $\log n$ and the bit length of c.

Finally, map outputs outside the ball to z. For the exact law $Q^{(m)}_{-\tau_n c,1}$, this changes the conditional ball law by at most the tail in (7.16); the map contracts all preceding TV errors. Their sum is at most n^{-S} for sufficiently large n. □

8.3 Restoring the quadratic nonedge interaction

Theorem 8.4. *There is an absolute C_{ph} such that, for every fixed known rational $a > b > 0$ satisfying (7.1), a graph-only algorithm has deterministic polynomial bit and random-bit costs on every graph and conditional output law $\widehat{\mu}_G$ satisfying*

$$\mathbb{E}_G \mathrm{TV}(\widehat{\mu}_G, \mu_G) = o(1).$$

At every parameter pair strictly inside this regime, the construction is locally uniform: on a sufficiently small compact parameter neighborhood one may use a common rational R, common polynomial caps, and one statistical error sequence tending to zero. For rational inputs in that neighborhood the cost is polynomial also in their bit length.

Proof. Fix a rational R strictly between $3r/4$ and $7r/8$, set $m = \lfloor Rn \rfloor$, and run (7.2). In addition to the preceding observable checks, require

$$|M(z)| \le m/4.$$

This holds on Proposition 7.1's event for sufficiently large n, because $|M(z)| \le rn/8$ and $R \ge 3r/4$. Otherwise return a fixed output. Put

$$A_n = |M(z)| + 2m, \qquad p = \lceil (a-b)K_0 \rceil + 10, \qquad S = p + 12,$$

where K_0 is from Proposition 7.2. Every spin in the ball has $|M(s)| \le A_n$, and the observable imbalance test gives $A_n \le 9m/4 \le 63rn/32 < 2rn$. Thus all fields queried below lie in the range already covered by the preceding propositions.

For $c \in [-A_n, A_n]$, define

$$f(c) = c - \mathbb{E}_{P^R_{-\tau_n c}} M.$$

The endpoint signs are opposite. Differentiating the finite partition sum gives

$$f'(c) = 1 + \tau_n \mathrm{Var}_{P^R_{-\tau_n c}} M, \qquad 1 \le f'(c) \le 1 + (a-b)n =: D_n, \tag{8.3}$$

where we used $\tau_n \le (a-b)/n$ for $n \ge 2a$ and $\mathrm{Var} M \le n^2$. Hence there is a unique zero.

Use rational bisection to find an approximate zero. At each midpoint c, draw $N = n^8$ independent samples using Proposition 8.3 and let $\overline{M}$ be their mean magnetization. Stop if $|c - \overline{M}| \le 1/2$; otherwise retain the half indicated by its sign. Make at most

$$H = \min\{j : 2^j \ge 16 A_n D_n\} = O_{a,b}(\log n)$$

queries, taking the final midpoint if there was no early stop. For exact samples, conditional on every preceding query history, Chebyshev's inequality gives a mean-estimation error greater than $1/8$ with probability at most $64n^{-6}$. On the complement of these errors, all bisection signs are correct; an early stop gives $|f(c)| \le 5/8$, and the final interval gives $|f(c)| \le 1/8$ by (8.3). Coupling the implemented draws to the exact ones, using the uniform accuracy of Proposition 8.3, proves

$$|c - \mathbb{E}_{P^R_{-\tau_n c}} M| \le 1 \tag{8.4}$$

except with probability $64Hn^{-6} + HNn^{-S}$.

Given this rational c, propose $s \sim P^R_{-\tau_n c}$ and accept with probability

$$a_c(s) = \exp\{-\tau_n (M(s) - c)^2/2\}. \tag{8.5}$$

The product of proposal and acceptance weights is

$$e^{J_n F_g(s) - \tau_n c M(s)} a_c(s) = e^{-\tau_n c^2/2} e^{J_n F_g(s) - \tau_n M(s)^2/2}.$$

Thus the accepted law is exactly

$$\pi_g^z = \mu_g(\,\cdot \mid d_H(s,z) \le Rn) \tag{8.6}$$

by (1.1). This identity holds for every c; centering is needed only to control the acceptance probability. Jensen's inequality, (7.18) with $v = \mathbf{1}$, and (8.4) give

$$\mathbb{E}_{P^R_{-\tau_n c}} a_c(s) \ge \exp\left\{-\frac{\tau_n}{2}\mathbb{E}(M-c)^2\right\} \ge \exp\left\{-\frac{(a-b)K_0}{2}\log n - \frac{a-b}{2n}\right\} \ge n^{-p}$$

for large n. Cap rejection at $T_{\mathrm{acc}} = n^{p+2}$ independent trials and return z if none succeeds. The exact failure probability is at most e^{-n^2}; conditional on success, the first accepted sample has law (8.6).

Use the ball sampler at accuracy n^{-S} for each proposal and implement (8.5) with additive scalar-evaluation and Bernoulli errors at most n^{-S} each. The exponent is nonnegative and $O_{a,b}(n)$, so Lemma A.1 gives these accuracies in polynomial bit time. Coupling along the deterministic trial cap adds at most $3T_{\mathrm{acc}} n^{-S}$. Uniformly on the good graph/construction event, the output before choosing a global sign is consequently within

$$64Hn^{-6} + Hn^8 n^{-S} + e^{-n^2} + 3T_{\mathrm{acc}} n^{-S} = o(1) \tag{8.7}$$

of π_g^z. The target does not depend on the bisection transcript, so this remains true after averaging that transcript. Every loop is capped even if a statistical estimate fails.

Finally, apply an independent fair global sign. The balls of radius Rn around z and $-z$ are disjoint and have equal posterior mass. Symmetrizing (8.6) gives the posterior conditioned on their union. Write $q = e^{-\kappa/16} = r^4$. On the recovery event,

$$d_\pm(s, X) < qn \quad \Longrightarrow \quad d_\pm(s,z) < qn + rn/4 < Rn.$$

Lemma 2.1 therefore bounds the expected posterior mass outside the union by

$$\mathbb{P}\{d_{\pm}(z, X) > rn/4\} + 4e^{-\kappa q n/8} = o(1).$$

The global-sign kernel contracts TV. Combining this bound with (8.7), the environmental failure probability, and convexity of TV when averaging the construction seed proves the theorem.

We record why the uniformity assertion follows from the proof. Choose a compact neighborhood on which R remains strictly between the two allowed endpoints and all inequalities used above retain their strict slack. The parameters $d, \theta, \kappa, a - b$ then have finite upper bounds and positive lower bounds as appropriate. Proposition 2.2 supplies a common recovery error sequence. All further environmental estimates are explicit Chernoff, Bernstein, Chebyshev, and witness bounds with constants uniform on that neighborhood. In particular, the planted-bad-set estimate is $O(1/n)$ with a uniform constant, and (7.23) is uniform. Use upper envelopes when choosing the integers c_S, D_S, p, S and the numerical loop caps. The scalar precision requirements and rational arithmetic bounds are then polynomial in the input bit length with common exponents. No uniformity is inferred merely from the pointwise existence of a recovery algorithm. □

9 An absolute threshold and real parameters

We now combine the two sampling constructions. The small-contrast construction works without an upper bound on d, whereas the recovered-phase construction requires κ to dominate $1 + \log d$. These conditions cover all sufficiently large κ: when the contrast is bounded below, $d = \kappa/\theta^2$ and the logarithmic condition follows from a large absolute lower bound on κ.

To pass from rational to real parameters, the statistical error must be uniform on a neighborhood. Pointwise convergence for each fixed rational pair would not suffice, since the rational approximation used by the algorithm changes with n. The quantitative estimates established above give the required uniformity.

Lemma 9.1. *There is an absolute C with the following property. Fix (a_*, b_*) with $a_* > b_* > 0$ and $\kappa(a_*, b_*) > C$. There are a compact rectangle U with rational endpoints containing (a_*, b_*) in its interior, fixed auxiliary rational data, a function $\psi_U(n) \to 0$, and an algorithm $\mathcal{A}_{n,a,b,\zeta}$ such that, for every rational $(a, b) \in U$, every dyadic $0 < \zeta < 1$, and all sufficiently large n,*

$$\mathbb{E}_{a,b}\mathrm{TV}\big(\mathcal{A}_{n,a,b,\zeta}(G, \cdot), \mu_G^{a,b}\big) \le \psi_U(n) + \zeta.$$

One polynomial in n, the rational encoding length of (a, b), and ζ^{-1} bounds the bit and random-bit costs on every graph. The rectangle and auxiliary data can be selected from finitely many certified evaluations of a_*, b_*.

Proof. Choose a positive rational h_0 such that $3h_0 < \theta_{\mathrm{w}}$, where θ_{w} is the small-contrast threshold in Theorem 6.2. Let C_{ph} be the constant in Theorem 8.4. Choose C above all preceding absolute signal thresholds and sufficiently large that, for every $x > C$,

$$x > C_{\mathrm{ph}}\big(1 + \log x + 2\log(1/h_0)\big).$$

Such a choice is possible because $x/\log x \to \infty$. If $\theta > h_0$, then $d < \kappa/h_0^2$, so $\kappa > C$ implies the hypothesis of Theorem 8.4. If $\theta < 3h_0$, Theorem 6.2 applies. These two open parameter ranges overlap.

A certified interval for θ of length less than $h_0/4$ selects a branch. If its upper endpoint is at most $2h_0$, select the small-contrast construction; otherwise its lower endpoint exceeds $7h_0/4$, and

select the recovered-phase construction. Choose a sufficiently small compact rational rectangle U around the true parameters, entirely inside the selected range and $\{a > b > 0,\ \kappa > C\}$. Write $a_+ = \sup_U a$, $b_- = \inf_U b$, and $\kappa_- = \inf_U \kappa$. These quantities have finite rational upper or lower bounds with strict slack.

For the small-contrast construction, fix one rational starting temperature in the interior of its admissible interval throughout U, and fix the degree cutoffs using a_+. The graph-decomposition estimates are quantitative, and the simultaneous regularized spectral estimate [LLV17, Theorem 2.1] is uniform for edge probabilities at most a_+/n. The reference-measure certificates have fixed strict margins. Consequently Theorem 6.2, including its local-uniformity assertion, supplies one error function tending to zero and one polynomial operation bound throughout U. In particular, no qualitative convergence rate depending on a changing rational input is used.

For the recovered-phase construction, Proposition 2.2 gives one uniform recovery error on U. Its radius $e^{-\kappa/16}n$ is smaller than the radius $2e^{-3\kappa/64}n$ allowed in the core construction. Fix one rational R strictly between $3e^{-\kappa/64}/4$ and $7e^{-\kappa/64}/8$ throughout a smaller rectangle U. Every remaining exponent and cutoff can be selected using upper and lower bounds on the fixed compact set. The exceptional-set estimates, component-witness bounds, and bounded enumeration therefore have a common vanishing failure probability. The local-uniformity assertion in Theorem 8.4 gives the desired statistical bound. The magnetization estimate for the terminal local-ascent configuration gives the required charge-restoration interval uniformly on U.

For either branch, first fix the statistical accuracies used by the Markov chains, the interpolation schedule, and any outer rejection procedure. These are inverse powers of n with fixed exponents on U. They are included in $\psi_U(n)$. Allocate the separate numerical budget ζ over the deterministic operation cap, as in Lemma A.1. Since $b_- > 0$ and a, b stay in a bounded set, the logarithms defining J_n, γ_n have uniformly controlled positive arguments for $n \geq 2a_+$. All required finite-precision calculations consequently have a common polynomial cost in the rational input length and the prescribed accuracy. Failed certificates or exceeded caps return a fixed configuration. Their probabilities are part of $\psi_U(n)$, while the caps hold on every graph and every history.

Finally, all auxiliary choices use strict inequalities between continuous functions of the parameters. Enumerating rational rectangles and rational auxiliary data, and certifying these inequalities by interval arithmetic, finds suitable data after finitely many parameter evaluations. This does not require a modulus of convergence for ψ_U. Increasing ψ_U to one at the finitely many excluded sizes completes the construction. □

Lemma 9.2. *Let P and Q be two joint laws of (X, G) on finite spaces, with posteriors μ_g and μ'_g, respectively. Then*

$$\mathbb{E}_{G\sim P_G}\mathrm{TV}(\mu_G, \mu'_G) \leq 2\mathrm{TV}(P, Q).$$

For any randomized graph-based rule $\mathcal{A}$,

$$\mathbb{E}_{P_G}\mathrm{TV}(\mathcal{A}(G,\cdot), \mu_G) \leq \mathbb{E}_{Q_G}\mathrm{TV}(\mathcal{A}(G,\cdot), \mu'_G) + 3\mathrm{TV}(P, Q).$$

For the stochastic block model at two admissible parameter pairs with $|a - a'|, |b - b'| \leq \delta$, one has $\mathrm{TV}(P, Q) \leq n\delta/2$.

Proof. Write p_g, q_g for the graph marginals and introduce $R(x, g) = p_g\mu'_g(x)$. Disintegration and the triangle inequality give

$$\sum_g p_g\mathrm{TV}(\mu_g, \mu'_g) = \mathrm{TV}(P, R) \leq \mathrm{TV}(P, Q) + \mathrm{TV}(Q_G, P_G) \leq 2\mathrm{TV}(P, Q).$$

The function $g \mapsto \mathrm{TV}(\mathcal{A}(g,\cdot),\mu'_g)$ takes values in $[0,1]$, so changing its marginal law costs at most $\mathrm{TV}(P_G,Q_G)$. This proves the second inequality. For the last claim, couple the planted labels identically and use the same uniform variable for each edge decision. Each of the $\binom{n}{2}$ pairs differs with probability at most δ/n, and the union bound proves the assertion. □

Theorem 9.3. *Let $a > b > 0$ be fixed real numbers with $\kappa > C$, where C is the absolute constant in Lemma 9.1. Suppose certified evaluation of $c \in \{a,b\}$ to absolute error 2^{-k} has bit cost $T_c(k)$. There are finite auxiliary rational data, a function $\eta_{a,b}(n) \to 0$, and a polynomial $P_{a,b}$ such that, for every dyadic $0 < \epsilon < 1$, the sampler satisfies, for all sufficiently large n,*

$$\mathbb{E}_{a,b}\mathrm{TV}(\widehat{\mu}_{G,\epsilon},\mu_G) \le \eta_{a,b}(n) + \epsilon.$$

After fixed preprocessing, it requests one approximation of each parameter at precision

$$k \le C_{a,b} + \lceil \log_2(n/\epsilon) \rceil$$

and has bit cost at most

$$T_a(k) + T_b(k) + P_{a,b}(n,\epsilon^{-1},k).$$

The polynomial also bounds its random-bit use on every input graph. The finite preprocessing cost is included separately.

Proof. Take the rectangle U and algorithm from Lemma 9.1. Certified coarse evaluations give a smaller rational rectangle containing (a,b) whose closure lies in the interior of U. They therefore provide a positive lower bound on the distance to the boundary of U. Obtain dyadic approximations a', b' with error at most $\delta = 2^{-k}$, choosing k so that $(a',b') \in U$ and

$$\delta \le \frac{\epsilon}{16n}.$$

A fixed additive constant in k ensures both conditions. Certified rational outputs may be rounded to dyadics, increasing this constant by a fixed amount. Their encoding length is $O_{a,b}(k)$.

Run $\mathcal{A}_{n,a',b',\epsilon/4}$ on the observed graph from the original model. By Lemma 9.2,

$$\mathbb{E}_{a,b}\mathrm{TV}(\mathcal{A}_{n,a',b',\epsilon/4}(G,\cdot),\mu_G^{a,b}) \le \psi_U(n) + \frac{\epsilon}{4} + \frac{3n\delta}{2} \le \psi_U(n) + \epsilon.$$

The same ψ_U applies to the rational pair varying with both n and ϵ. Set $\eta_{a,b} = \psi_U$. Substituting its rational encoding length and the requested accuracy into the common polynomial bound in Lemma 9.1 gives the stated cost. Every subsequent computation uses a', b'; no further evaluation of the original real parameters is required. □

Proof of Theorem 1.1. For fixed rational a, b, apply Lemma 9.1 with $\zeta_n = 2^{-\lceil 4\log_2 n \rceil}$. Since $\zeta_n^{-1} \le 2n^4$, the bit and random-bit costs are polynomial, and the expected total-variation error tends to zero. The two overlapping contrast ranges in the proof of the lemma impose no additional condition on d.

For fixed real parameters use Theorem 9.3. In particular, if the parameters are polynomial-time computable, choosing the same inverse-polynomial numerical accuracy gives polynomial total cost. For arbitrary computable parameters, their evaluation costs remain in the bound; computability by itself gives no polynomial-time estimate for obtaining the required digits. □

A Finite-precision implementation

This appendix collects the numerical facts used by both samplers. The statistical arguments specify polynomial iteration caps and strict certificate margins. Numerical accuracy can then be allocated once over the full operation count, including adaptive choices of fields. Throughout, the polynomial may depend on a fixed compact parameter neighborhood, but not on the input graph.

Lemma A.1. *The following operations have deterministic polynomial bit complexity in the dimension, the rational input length, the indicated magnitude bounds, and the requested number of accuracy bits.*

(i) *Certified evaluation of* $\log x$ *for positive rational* x*, nonnegative square roots, and* e^y *for polynomially bounded* $|y|$*. Positive factors may be computed with a specified logarithmic relative error.*

(ii) *Sampling a rational law on* N *outcomes to total variation error* η*. This uses* $O(\log(N/\eta))$ *fair bits and never selects a zero-probability outcome. A standard Gaussian vector in polynomial dimension admits a finite-bit approximation coupled to it with any prescribed coordinate error outside an event of probability at most* η*.*

(iii) *Given a symmetric matrix* S *whose entries admit polynomial-time certified evaluations, a bound on* $\|S\|$*, a fixed rank* r*, and* $\eta > 0$*, constructing a rational positive semidefinite matrix* C *of rank at most* r *such that*

$$\|S - C\| \le \inf_{A \succeq 0,\ \operatorname{rank}(A) \le r} \|S - A\| + \eta.$$

A factor L *with* $L^{\mathsf{T}}L = C$ *can subsequently be evaluated to any prescribed precision. No eigenvalue separation is required. Positive semidefiniteness of rational symmetric matrices can be tested exactly in polynomial time.*

If an ideal finite-state algorithm makes at most N_* *conditional sampling calls, and each implemented call has conditional row error at most* η/N_* *for every preceding history, the output error is at most* η*. In particular, polynomially many calls with polynomial magnitude bounds admit deterministic polynomial bit and random-bit caps at any inverse-polynomial numerical accuracy.*

Proof. We record the precision choices to make the computational model explicit. The symmetric-matrix algorithms used below are standard; see [GVL13, Chapters 2, 8–9]. For logarithms, scale x into $[1, 2]$ and use

$$\log y = 2 \sum_{j \ge 0} \frac{((y-1)/(y+1))^{2j+1}}{2j+1}.$$

The ratio is at most $1/3$; the same series at $y = 2$ computes the scaling contribution. For exponentials, reduce the argument to absolute value at most one, use the factorial series, and square back with guard bits. Rational bisection computes nonnegative square roots, including near zero. These procedures give certified errors and polynomial encoding lengths on the stated magnitude ranges.

For a categorical distribution, draw a uniform dyadic midpoint and select the first cumulative probability strictly greater than it. With b bits, disagreement with a continuous uniform draw occurs only in intervals meeting a cumulative-probability boundary, and has probability at most $N2^{-b}$. Choose $b \ge \lceil \log_2(N/\eta) \rceil + 1$. For k Gaussian coordinates, truncate at $R = O(\sqrt{\log(k/\eta)})$, paying probability at most $\eta/2$, and invert the truncated distribution function on a dyadic grid.

Its inverse has Lipschitz constant at most $2Re^{R^2/2}$, so polynomially many bits give the specified coordinate error. Integrating the exponential series over $[-R, R]$ supplies certified distribution values. This is a coupling in coordinate distance; no total variation approximation of a continuous law by a discrete law is asserted.

For the matrix assertion, approximate S by a rational symmetric matrix B, and use the maximal-pivot Jacobi procedure of [GVL13, §8.5] to obtain a rational orthogonal O for which the off-diagonal Frobenius norm of $O^{\mathsf{T}}BO$ is at most a chosen δ. To retain exact orthogonality, approximate each rotation's half-angle tangent by a dyadic u and use the block

$$\frac{1}{1+u^2}\begin{pmatrix} 1-u^2 & 2u \\ -2u & 1-u^2 \end{pmatrix}.$$

If $F \geq 1$ bounds $\|B\|_{\mathrm{F}}$, approximate each exact rotation to operator error at most $\delta/(8Fm^2)$. The Jacobi contraction estimate then bounds the next off-diagonal norm by $(1 - 1/(m(m-1)))$ times the current norm plus $\delta/(4m^2)$. Hence $O(m^2 \log(F/\delta))$ rotations suffice. The rational denominators acquire only a polynomial sum of bit lengths over these iterations; exact orthogonality bounds the magnitudes of intermediate entries. The rotation formula uses rational operations and square roots and requires no lower bound on a nonzero pivot.

Dimensions $m \leq 1$ are treated directly. Let $\widetilde{S} = O\,\mathrm{diag}(O^{\mathsf{T}}BO)O^{\mathsf{T}}$. If $\|B - S\| \leq \delta$, then $\|\widetilde{S} - S\| \leq 2\delta$. Retaining the largest r positive diagonal entries gives a rational positive semidefinite C of rank at most r. The spectral variational principle gives

$$\|\widetilde{S} - C\| \leq \inf_{A \succeq 0,\ \mathrm{rank}(A) \leq r} \|S - A\| + 2\delta.$$

Thus $\delta \leq \min\{\eta/4, 1/4\}$ suffices. The selected columns of O, multiplied by square roots of the retained diagonal entries, give the factor L. A rational positive-semidefiniteness test uses symmetric elimination: a negative pivot rejects, a positive pivot permits a Schur complement, and a zero pivot requires the remaining off-diagonal row to vanish. Standard exact elimination has polynomial bit complexity; see [Sch86, Chapter 3].

Finally, couple the ideal and implemented decisions whenever their histories agree. A union bound over at most N_* calls bounds the probability of any discrepancy by η. This also applies to adaptive choices, since the row bound is conditional on the complete previous history. □

Here are the two error estimates used to apply the lemma. If unnormalized weights have log errors bounded by δ uniformly on their common support, normalization puts their density ratio in $[e^{-2\delta}, e^{2\delta}]$, and hence bounds their total variation distance by $e^{2\delta} - 1$. The same bound applies to every conditional row. If a coupled Gaussian field approximation has ℓ_1 error at most δ outside an event of probability τ, and the reference sampler has error ϵ_{ref}, the total row error is at most

$$\tau + \epsilon_{\mathrm{ref}} + e^{2\delta} - 1.$$

For the Gaussian block sampler, matrices whose spectrum lies in a fixed interval $[b, 3b]$ can be square-rooted by a convergent matrix binomial series after centering at $2bI$. On the Gaussian truncation event, all sampled fields have polynomial magnitude. Choosing coefficient precision so that the resulting ℓ_1 field error meets the displayed budget completes the implementation.

All stopping rules in the paper are fixed caps. An inconclusive certificate, an exceeded enumeration cap, or a failed capped rejection returns the specified supported fallback. On events where the statistical proof guarantees success, these branches either do not occur or have the stated small probability. This is why numerical accuracy can be chosen independently of the environmental failure probability.

References

[BK20] Franck Barthe and Bo'az Klartag. Spectral gaps, symmetries and log-concave perturbations. *Bulletin of the Hellenic Mathematical Society*, 64:1–31, 2020. arXiv:1907.01823.

[BL76] Herm Jan Brascamp and Elliott H. Lieb. On extensions of the Brunn–Minkowski and Prékopa–Leindler theorems, including inequalities for log concave functions, and with an application to the diffusion equation. *Journal of Functional Analysis*, 22(4):366–389, 1976.

[BLM13] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press, 2013.

[CGG+21] Zongchen Chen, Andreas Galanis, Leslie Ann Goldberg, Will Perkins, James Stewart, and Eric Vigoda. Fast algorithms at low temperatures via Markov chains. *Random Structures & Algorithms*, 58(2):294–321, 2021.

[CRV15] Peter Chin, Anup Rao, and Van Vu. Stochastic block model and community detection in sparse graphs: A spectral algorithm with optimal rate of recovery. In *Proceedings of the 28th Conference on Learning Theory*, volume 40 of *Proceedings of Machine Learning Research*, pages 391–423, 2015.

[GVL13] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, fourth edition, 2013.

[Hou01] J. Houdayer. A cluster Monte Carlo algorithm for 2-dimensional spin glasses. *European Physical Journal B*, 22:479–484, 2001.

[JGJ+26] Haocheng Ju, Guoxiong Gao, Jiedong Jiang, Bin Wu, Zeming Sun, Shurui Liu, Leheng Chen, Yutong Wang, Yuefeng Wang, Zichen Wang, Wanyi He, Peihao Wu, Liang Xiao, Ruochuan Liu, Bryan Dai, and Bin Dong. Automated conjecture resolution with formal verification. arXiv:2604.03789, 2026.

[LLV17] Can M. Le, Elizaveta Levina, and Roman Vershynin. Concentration and regularization of random graphs. *Random Structures & Algorithms*, 51(3):538–561, 2017. Theorem numbering refers to arXiv:1506.00669v2.

[LMRW24] Kuikui Liu, Sidhanth Mohanty, Amit Rajaraman, and David X. Wu. Fast mixing in sparse random Ising models. arXiv:2405.06616v2, 2024.

[LP17] David A. Levin and Yuval Peres. *Markov Chains and Mixing Times*. American Mathematical Society, second edition, 2017. With contributions by Elizabeth L. Wilmer.

[MNS16] Elchanan Mossel, Joe Neeman, and Allan Sly. Belief propagation, robust reconstruction and optimal recovery of block models. *Annals of Applied Probability*, 26(4):2211–2256, 2016.

[Sch86] Alexander Schrijver. *Theory of Linear and Integer Programming*. John Wiley & Sons, 1986.

[Sin92] Alistair Sinclair. Improved bounds for mixing rates of Markov chains and multicommodity flow. *Combinatorics, Probability and Computing*, 1(4):351–370, 1992.

[Wu06] Liming Wu. Poincaré and transportation inequalities for Gibbs measures under the Dobrushin uniqueness condition. *Annals of Probability*, 34(5):1960–1989, 2006.