

Low-temperature Sampling on Sparse Random Graphs*

Andreas Galanis

Leslie Ann Goldberg

Paulina Smolarova

12th February 2025

Abstract

We consider sampling in the so-called low-temperature regime, which is typically characterised by non-local behaviour and strong global correlations. Canonical examples include sampling independent sets on bipartite graphs and sampling from the ferromagnetic q -state Potts model. Low-temperature sampling is computationally intractable for general graphs, but recent advances based on the polymer method have made significant progress for graph families that exhibit certain expansion properties that reinforce the correlations, including for example expanders, lattices and dense graphs.

One of the most natural graph classes that has so far escaped this algorithmic framework is the class of sparse Erdős-Rényi random graphs whose expansion only manifests for sufficiently large subsets of vertices; small sets of vertices on the other hand have vanishing expansion which makes them behave independently from the bulk of the graph and therefore weakens the correlations. At a more technical level, the expansion of small sets is crucial for establishing the Kotecky-Priess condition which underpins the applicability of the framework.

Our main contribution is to develop the polymer method in the low-temperature regime for sparse random graphs. As our running example, we use the Potts and random-cluster models on $G(n, d/n)$ for $d = \Theta(1)$, where we show a polynomial-time sampling algorithm for all sufficiently large q and d , at all temperatures. Our approach applies more generally for models that are monotone. Key to our result is a simple polymer definition that blends easily with the connectivity properties of the graph and allows us to show that polymers have size at most $O(\log n)$.

1 Introduction

We consider the problem of sampling from high-dimensional distributions in the so-called low-temperature regime, which is typically characterised by non-local behaviour. A classical example that has been studied in this context is the problem of sampling independent sets on bipartite graphs where it is perhaps intuitive to expect that a typical independent set is correlated with one of the two sides of the bipartition. Another standard example is the ferromagnetic Potts model on (not necessarily proper) q -colourings weighted by the number of monochromatic edges (see below for definitions) where, at low temperatures, a typical colouring is expected to be correlated with one of the q monochromatic colourings.

Starting from [HPR19, JKP20], a series of works have demonstrated that this correlation can be utilised to obtain fast sampling algorithms for certain graph classes that exhibit strong expanding properties; perhaps the most prominent graph example is the class of random regular graphs.

*For the purpose of Open Access, the authors have applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. All data is provided in full in the results section of this paper.

Lattices (such as \mathbb{Z}^d) can also be treated using contour models. By contrast, it is far from clear how to apply this intuition to general graphs. In fact, sampling at low-temperatures is conjectured to be hard on general graphs, captured by so-called $\#$ BIS-hardness in the case of the Potts and independent-set examples.

Perhaps the most natural class of graphs that has so far remained elusive from this algorithmic framework is the class of sparse Erdős-Rényi random graphs like $\mathcal{G}(n, d/n)$ for $d = \Theta(1)$. This class has actually posed challenges even for high-temperature sampling [LMRW24, EF23, BZ23, BGGv24, EHSV18, YZ16, MS13] where the presence of high-degree variables typically complicates the applicability of standard sampling techniques. For low-temperature sampling however the obstacles come primarily from another side, namely from the presence of many small “sparse” parts which have weak, if any, expansion. For example, in $\mathcal{G}(n, d/n)$, this phenomenon not only causes the graph to be disconnected, but even inside the giant component there are *induced* paths of size roughly $\Theta(\log n)$ that therefore have almost no expansion; similarly, one can find various other sparse induced components which are scattered inside the giant.

Our main contribution is to develop the polymer method in the low-temperature regime for sparse random graphs. As our running example, we use the Potts and random-cluster models on $\mathcal{G}(n, d/n)$ for $d = \Theta(1)$ where we show a polynomial-time sampling algorithm for all sufficiently large q and d at all temperatures. Our approach applies more generally for models that are monotone. Key to our result is a simple polymer definition that blends easily with the connectivity properties of the graph and allows us to show that polymers have size at most $O(\log n)$.

To formally state our result, we recall a few standard definitions. The random cluster (RC) model with real parameters $q, \beta > 0$ is a weighted edge model arising from statistical physics, assigning probabilities to the edge subsets of a given graph $G = (V, E)$. For each edge subset $F \subset E$, the associated weight of the *configuration* F is given as

$$w_G(F) = w_{G;q,\beta}(F) = q^{c(F)}(e^\beta - 1)^{|F|},$$

where $c(F)$ is the number of components in the graph (V, F) . We refer to edges in F as the *in-edges* of the configuration (and edges in $E \setminus F$ as the *out-edges*). Denote by Ω_G the set of all configurations, i.e., all edge subsets of G . The Gibbs distribution $\pi_G = \pi_{G;q,\beta}$ is defined by $\pi_G(F) = w_G(F)/Z_G$ for $F \in \Omega_G$, where $Z_G = \sum_{F' \in \Omega_G} w_G(F')$ is the partition function. The random cluster model can be viewed as an equivalent edge-representation of the *Potts model* (when q is an integer), which is supported on vertex assignments $\sigma : V \rightarrow \{1, \dots, q\}$ weighted by $e^{\beta m(\sigma)}$ where $m(\sigma)$ is the number of monochromatic edges under σ .

For the random graph $\mathcal{G}(n, d/n)$ which is our focus here, the qualitative structure is conjectured to align with that on the random regular graph that has been extensively studied. In the so-called high temperature regime (low β) a typical configuration is expected to be correlated with the all-out configuration ($F = \emptyset$) and to consist of many small components – this set of configurations is known as the “disordered phase”; whereas in the so-called low temperature regime (high β) a typical configuration is expected to be correlated with the all-in configuration ($F = E$) and to consist of a giant component (and other small components) – this set of configurations is known as the “ordered phase”. The interesting feature of the random-cluster model is that, on several classes of graphs, there is a critical threshold β_c where the two phases coexist and each have probability bounded below by a constant. This coexistence typically causes severe complications to sampling in a window around β_c , see for example [COGG⁺23, GLP20, GJ97] for various metastability phenomena and [GJ12, GvVY16, CGG⁺16] for computational hardness results that build on this.

Understanding this picture on $\mathcal{G}(n, d/n)$ more precisely is quite a bit more complicated than the random regular graph, even with currently available analysis tools; the key difference is that the

local neighbourhood of a vertex is given by a Poisson tree (rather than a regular tree). For example, the value of the log-partition function $\lim_{n \rightarrow \infty} \frac{1}{n} \log Z_G$ for low-temperatures is likely a rather involved expectation over Poisson trees based on understanding fixed-point equations on the underlying set of trees.¹ Likewise, understanding the performance of Markov chain algorithms on $\mathcal{G}(n, d/n)$ becomes extremely involved; in fact, even getting a fairly precise understanding of what happens on a Poisson tree for low temperatures is open for both the Potts and random cluster models.

Our main result is to obtain an algorithm for $\mathcal{G}(n, d/n)$ by introducing suitable developments to the polymer framework and showing how to use these together with recent MCMC techniques. As a corollary of our techniques, we obtain various tools for the RC distribution on $\mathcal{G}(n, d/n)$ that shed light on the properties of the distribution.

Theorem 1. *Let d be a large enough real. Then, for all sufficiently large reals q , there is a (randomised) algorithm \mathcal{A} such that the following holds whp over $G \sim \mathcal{G}(n, d/n)$, for any inverse temperature $\beta > 0$. On input G , with probability at least $3/4$, the algorithm \mathcal{A} outputs in $\text{poly}(n)$ time a sample $F \in \Omega_G$ whose distribution is within TV-distance $e^{-\Omega(n)}$ from the RC distribution $\pi_{G;q,\beta}$, and an estimate \hat{Z} for the RC partition function $Z_G = Z_{G;q,\beta}$ that satisfies $\hat{Z} = (1 \pm e^{-\Omega(n)})Z_G$.*

The success probability of $3/4$ in Theorem 1 can be powered in the standard way. As we will describe later in more detail (Section 4), our algorithm is based on running the *Glauber dynamics* for the random-cluster model from a suitable initial configuration, though it is a bit more involved than that since, in an interval of temperatures $[\beta_0, \beta_1]$, it needs to approximate the appropriate mixture of the ordered and disordered phases.

The running time of the algorithm in Theorem 1 is $\text{poly}(n)$ where the implicit constant in the exponent scales as $O(\log q)$; this is largely because of a crude polynomial mixing-time bound on Poisson trees with wired boundary (i.e., Poisson trees where the leaves are conditioned to belong to the same component, say by contracting them into a single vertex). It is an open question to obtain sharper bounds with the wired boundary condition, even for regular trees when q is non-integer (cf. [BCvV23, Theorem 5]).

The lower bound on q in Theorem 1 is roughly $\exp(\Omega(d \log d))$, which is essentially the “standard” bound where polymer-based techniques work (at least on bounded-degree graphs). Having q this large facilitates showing closeness of a typical sample to the extreme configurations. In the case of bounded-degree graphs, there has been some progress in lowering the value of q , mainly on random regular graphs [BG21, GGS24a]. For large β these results come from first/second moment methods, though these results do not extend to non-integer q or to our $G(n, d/n)$ setting.

We can improve significantly upon the running time using a different (deterministic) algorithm based on estimating the probability that an edge is an in-edge (based on a certain correlation decay property known as weak spatial mixing (WSM) within the phase, see Section 2.3) and integrating the expected number of in-edges. We expect that this different approach to utilising WSM will be useful in other settings where the underlying graph has tree-like neighbourhoods.

Theorem 2. *Let $R > 0$ be arbitrarily large. For all sufficiently large reals d , for all sufficiently large reals q , there is an algorithm \mathcal{B} such that the following holds whp over $G \sim \mathcal{G}(n, d/n)$, for any temperature $\beta > 0$. On input G , the algorithm \mathcal{B} outputs in $n^{1+\frac{1}{R}}$ time an estimate \hat{Z} for the RC partition function $Z_G = Z_{G;q,\beta}$ that satisfies $\hat{Z} = (1 \pm \frac{1}{n^R})Z_G$.*

¹This is based on the fact that the model is “replica symmetric” [BCM19]; to the best of our knowledge the formula for the log-partition function has not been established yet.

It may help to explain the parameters in Theorem 2. The parameter q controls the accuracy of the algorithm via the decay rate in Lemma 6, and d controls the running time via the length r in Lemma 6 (capturing the size of the polymer). So, for any fixed “target” (captured by $R > 0$), by first taking d large (w.r.t. R), and then taking q large (w.r.t. d), we can make both the running time and the accuracy sufficiently small in terms of the target (as per the theorem statement). See Remark 22 for further comments about the restriction that d be sufficiently large, and how this arises (via Proposition 24) in the proof of Lemma 6.

1.1 Discussion and Further Related Work

Most of the known low-temperature algorithms are based on the so-called polymer method which was developed in [HPR19, JKP20]. Intuitively, a polymer represents a local part of a configuration that deviates from a certain extremal/ground state. The success of the method typically relies on the Kotecký-Preiss condition [KP86] that roughly controls the number of polymers and the growth of their weights. This condition guarantees that the log partition function can be approximated by truncating a relevant convergent series expansion (the cluster expansion) and then applying the interpolation techniques of [Bar17, PR17]; see also [CGG⁺21] for MCMC variants. The method has been very successful in the low-temperature regime [HPR19, JKP20, LLLM19, CDF⁺22, HJP23, CDK⁺20, GGS21, JPP23, CDKP24, JPPS24].

The main difficulty in applying the polymer method on $\mathcal{G}(n, d/n)$ is that small polymers, such as induced paths of length $\Theta(\log n)$ or other sparse induced subgraphs present in the graph, can have unusually large weight violating the growth rate required by Kotecký-Preiss type conditions. This difficulty causes undesirable restrictions; for example, [GGS22] worked on a sparse random graph model with degrees ≥ 3 , a condition that ensures the absence of such non-expanding parts and precludes therefore $\mathcal{G}(n, d/n)$. As we will see in the following section, the main contribution of the present paper is to develop the polymer method for sparse random graphs such as $\mathcal{G}(n, d/n)$ by introducing a suitable polymer definition.

The algorithm provided in our proof of Theorem 1 is based on the Glauber dynamics Markov chain for the random-cluster model. For the case of random regular graphs, the mixing properties of Glauber dynamics are well-studied. Blanca and Gheissari [BGG⁺20a] showed that the random cluster Glauber dynamics is mixing in time $O(n \log n)$ for all $q \geq 2$ and $\beta < \beta_u$ where β_u is the uniqueness threshold on the regular tree, and they showed that the same bound applies on $\mathcal{G}(n, d/n)$ as well [BG23b]; see also [BG23a] for mixing-time results on other graph classes that apply for large β . However, for β near the ordered/disordered threshold β_c (satisfying $\beta_c > \beta_u$), the chain undergoes an exponential slowdown due to metastability phenomena and phase coexistence [HJP23, COGG⁺23]. Our results suggest a similar exponential slowdown around criticality.

However, despite the worst-case mixing result, it is still possible to obtain a fast sampling algorithm based on Glauber dynamics using an appropriate initialisation from a ground state, see [GS22, GS23, GGS24b]. Polymer techniques are helpful in this setting since they can be used to show a notion called weak spatial mixing within a phase introduced in [GS22], see Section 2.3 for definitions. In the next section, we discuss more thoroughly how to adapt the polymer method for $\mathcal{G}(n, d/n)$ which is the main bottleneck of our results.

1.2 Overview: old and new polymers

As we have noted, polymers capture small, independent deviations from a certain extremal configuration called the “ground state”. For example, for the random cluster model the natural ground states are the all-in and all-out configurations, and for the ferromagnetic Potts model the ground

states are the monochromatic colourings.

To apply the polymer framework, the *weight* of a polymer needs to be defined in a way that enables us to relate the weight of a configuration to the product of weights of its polymers, and thus to relate the probability of a polymer appearing in a configuration to its weight. For models with only local interactions, such as the Potts model, the characterization of these small deviations, i.e., the definition of polymers, is clear from the model. For instance, for the Potts model on low temperatures (high β), a typical configuration has the majority of vertices having the same colour. Thus the natural choice for polymers of a q -colouring σ are maximal connected sets γ which are not coloured with the majority colour. The weight w_γ measures how much weight is lost by having bichromatic edges within the polymer γ .

However, when considering how a polymer should be defined for a configuration F of the random cluster model, the answer is not as straightforward. For low-temperatures (high β), the natural ground state is the all-in configuration, thus a natural choice would be to consider (connected) sets of all-out edges. However the random cluster model has long-distance interactions, and in particular, the weight (and thus the probability) of a configuration F depends not only on the number of out-edges (or in-edges), but also on the number of components. For a component of (V, F) , it may not be the case that the set of out-edges separating it from the rest of the graph is connected. If multiple polymers are allowed to enclose a single component, it is not clear how to capture that component's contribution to the weight of the configuration using (multiplicative) polymer weights.

One insightful solution to this problem was given for the case of the random regular graph by Helmuth, Jenssen and Perkins [HJP23], where they iteratively added more edges to the polymers, and then used the strong expansion properties of random regular graphs to ensure the desired multiplicative properties. For completeness, we will define their polymers for a configuration F : Let $\mathcal{B}_0 = |E \setminus F|$. For $k \geq 0$, define inductively \mathcal{B}_{k+1} to be the set of edges in \mathcal{B}_k along with all edges incident to vertices that have at least a $\frac{5}{9}$ -fraction of their incident edges in \mathcal{B}_k . Then the set of *polymer edges* is $\mathcal{B}_{\text{inf}}(F) = \bigcup_{k \geq 0} \mathcal{B}_k$ and polymers are connected components of edges in \mathcal{B}_{inf} . Using the strong expansion of random regular graphs, they show that for any ordered configuration (V, F) , there is one giant component containing $> n/2$ vertices, and every other component has all its incident edges contained in a polymer. While the constant $\frac{5}{9}$ in the definition could be lowered to any $\frac{1}{2} + \tau$, to avoid $\mathcal{B}_{\text{inf}}(F) = E$ it has to be greater than a half.

The fact that an ordered configuration has a giant component also applies for $\mathcal{G}(n, d/n)$. However, having the constant $5/9$ (or anything larger than $1/2$) causes problems in graphs where degrees can be small. For instance, a 3-cycle of a 3-regular graph has expansion $\frac{1}{2}$. Suppose that the edges of the 3-cycle are in-edges, but the edges separating it from the rest of graph are out. Then these separating edges are not necessarily contained in a single polymer. This makes it difficult to capture the component's contribution, as explained earlier. These problems persist even if the expansion of sufficiently large subgraphs is above $1/2$ by any margin.

In the case of $\mathcal{G}(n, d/n)$ the same problem is present in a more severe form: while large-enough connected sets can be shown to have expansion $> \frac{1}{2}$ (for d large enough), there are small subgraphs with as many as $\Theta(\log n)$ vertices with extremely low expansion. These small subgraphs can not be captured by these expansion-based polymers and it is therefore not clear how to extend the polymer definition of [HJP23] for the case of $\mathcal{G}(n, d/n)$.

To start working towards a polymer definition, note that expanding properties of $\mathcal{G}(n, d/n)$ that hold for sets of size $\Omega(\log n)$ can be used to show that (V, F) contains a giant component and all the other components are small (provided that $|F|$ is sufficiently large relative to $|E|$). So, a natural-looking solution for the polymer-definition problem is to simply take the polymers of a configuration F to be components formed by the union of the set of out-edges $E \setminus F$ and the set of

edges in E incident to vertices in small components of (V, F) (irrespective of whether they belong to F). While this definition can be endowed with a polymer-weight definition that does satisfy the desired multiplicative properties, these polymers do not capture properly the deviations from the all-in configuration. For instance, suppose there was a connected subgraph S of (V, F) with exactly one in-edge e in the cut (S, \bar{S}) . Then it would make sense that vertices of S belonged to a polymer since the removal of a single edge e carves out S as a component, affecting significantly the marginal probability that an edge in the cut (S, \bar{S}) is an in-edge. Now, if the cut (S, \bar{S}) is large, having S as a component in a configuration (after removal of e) is a large deviation from the ordered state and is extremely unlikely to happen, and the marginal probabilities of edges in the cut should not be so sensitive to the status of a single edge. Hence, such “almost-components” must be enclosed in a polymer.

For our analysis specifically, we care about marginals of edges incident to a particular vertex v , thus we do not want the marginals of edges not in a polymer be sensitive to the state of the edges incident to v . Thus, as we show in Section 3, the natural choice of ordered polymers is to take polymer edges to be the union of the set of out-edges of F and the set of edges of G incident to vertices in subgraphs that would be disconnected by removing this vertex v . In Section 3 we show that with good probability the polymers have size $O(\log n)$ and we use this to prove *weak spatial mixing within the phase*. This is then used for bounding the mixing time from the Glauber dynamics (Theorem 1) and converting to the counting algorithm (Theorem 2).

2 Proof Outline of Theorems 1 and 2

2.1 The largest component of the graph

It is well-known that for a graph G with multiple components, the Gibbs distribution π_G is a product distribution over the individual components, and that the partition function Z_G is the product over the partition function over individual components.

For $d > 1$, $\mathcal{G}(n, d/n)$ whp contains one component of linear size while all remaining components have size $O(\log n)$ and contain at most one cycle, see for example [JLR00, Section 5]. We can therefore brute-force $O(\log n)$ -sized components by enumerating all possible edge configurations in $\text{poly}(n)$ time; in fact, even the mixing time from *worst-case* initial configuration on such a component is going to be at most $\text{poly}(n)$ (see, e.g., [BG21, Lemma 6.7]).

So, we only need to focus on the largest component of $\mathcal{G}(n, d/n)$, which we denote by $C = (V_C, E_C)$. Let $n_C = |V_C|$ be the number of vertices. It is well-known (see e.g. [JLR00, Theorem 5.4]) that n_C is determined by the conjugate $\mu \in (0, 1)$ of d which satisfies $\mu e^{-\mu} = d e^{-d}$. Using this, it is not hard to verify that for large d we have whp $n_C \geq (1 - e^{-d/3})n$ and $|E_C| = \frac{dn}{2} \pm 2ne^{-d/3}$, see Appendix A for details.

For clarity of notation, we use Ω_C to denote the set of configurations, w_C for the weights of configurations, π_C for the Gibbs distribution and Z_C for the partition function for the random cluster model on C . For $F \in \Omega_C$, we use $c(F)$ to refer to the number of components in (V_C, F) .

2.2 Phases of random cluster model on $\mathcal{G}(n, d/n)$

Next, we formally introduce the notion of the *ordered* and *disordered* phases on $C = C(\mathcal{G}(n, d/n))$, as well as some necessary notation.

Definition 3 (Phases). *Let $\eta := 1/1000$. The ordered phase is $\Omega_C^{\text{ord}} := \{F \in \Omega_C \mid |F| \geq (1-\eta)|E_C|\}$. The ordered distribution π_C^{ord} is defined for $F \in \Omega_C^{\text{ord}}$ by $\pi_C^{\text{ord}}(F) := w_C(F)/Z_C^{\text{ord}}$, where*

$Z_C^{\text{ord}} := \sum_{F \in \Omega_C^{\text{ord}}} w_C(F)$. The disordered phase is $\Omega_C^{\text{dis}} := \{F \in \Omega_C \mid |F| \leq \eta|E_C|\}$. The disordered distribution π_C^{dis} is defined for $F \in \Omega_C^{\text{dis}}$ by $\pi_C^{\text{dis}}(F) = w_C(F)/Z_C^{\text{dis}}$, where $Z_C^{\text{dis}} := \sum_{F \in \Omega_C^{\text{dis}}} w_C(F)$.

We will use the following two thresholds for β , β_0 and β_1 , defined from the following:

$$e^{\beta_0} - 1 = q^{(2-1/10)/d} \text{ and } e^{\beta_1} - 1 = q^{(2+1/10)/d}. \quad (1)$$

Note that the constant $1/10$ in the definitions above is somewhat arbitrary and could be decreased to any absolute constant $\tau > 0$. The next theorem tells us that a typical configuration on $\mathcal{G}(n, d/n)$ is either ordered or disordered and “far away” from phases’ boundaries. It is proven in Appendix C.

Theorem 4. *Let d be sufficiently large. Then, for all q sufficiently large, the following holds whp over $G \sim \mathcal{G}(n, d/n)$, where F denotes a random configuration from the given distribution.*

(1) For all $\beta > 0$, $\pi_C\left(\frac{9\eta}{10}|E_C| \leq |F| \leq (1 - \frac{9\eta}{10})|E_C|\right) \leq e^{-n}$.

(2) For all $\beta \geq \beta_0$, $\pi_C^{\text{ord}}(|F| \leq (1 - \frac{9\eta}{10})|E_C|) \leq e^{-n}$.

(3) For all $\beta \leq \beta_1$, $\pi_C^{\text{dis}}(\frac{9\eta}{10}|E_C| \leq |F|) \leq e^{-n}$.

(4) For $\beta \geq \beta_1$, $\|\pi_C - \pi_C^{\text{ord}}\|_{\text{TV}} \leq 2e^{-n}$.

(5) For $\beta \leq \beta_0$, $\|\pi_C - \pi_C^{\text{dis}}\|_{\text{TV}} \leq 2e^{-n}$.

2.3 WSM and proof of Theorem 2 via WSM

The main challenge in order to obtain our results is showing WSM within the phase, which we define more formally here for the linear-sized component. Namely, for a vertex $v \in V_C$ and an integer r , let $B_r(v)$ be the ball of radius r around v , i.e. the induced subgraph consisting of all vertices distance $\leq r$ from v . Let $\pi_{B_r^-(v)}$ denote the Gibbs distribution on B with the *free* boundary condition (i.e., conditional on all edges outside of $B_r(v)$ being out) and $\pi_{B_r^+(v)}$ denote the Gibbs distribution on $B_r(v)$ with the *wired* boundary condition (i.e., conditional on all vertices at distance exactly r from v as being wired into a single component). For an edge $e \in E_C$, we use 1_e denote the event that e is an in-edge.

Definition 5. Let $r > 0$ be an integer and $K \in (0, 1)$ be a real.

We say that C has WSM within the disordered phase at distance r with rate K if for every edge $e \in E_C$ and vertex $v \in V_C$ incident to e , it holds that $|\pi_C^{\text{dis}}(1_e) - \pi_{B_r^-(v)}(1_e)| \leq K^r$.

Similarly, C has WSM within the ordered phase at distance r with rate K if for every edge $e \in E_C$ and vertex $v \in V_C$ incident to e , it holds that $|\pi_C^{\text{ord}}(1_e) - \pi_{B_r^+(v)}(1_e)| \leq K^r$.

In Section 3, we show that for all d and q large enough, whp C has WSM within the ordered phase for all $\beta \geq \beta_0$. The proof of WSM within the disordered phase for all $\beta \leq \beta_1$ is deferred to the Appendix B.

Lemma 6. *There is a constant $A > 0$ such that the following holds for all real d sufficiently large. For all q sufficiently large, whp over $G \sim \mathcal{G}(n, d/n)$, for all $\beta \geq \beta_0$, the largest component C of G has WSM within the ordered phase at distance $r = \lceil \frac{A}{d} \log n \rceil$ with rate $K = q^{-1/30}$.*

Using this, we can prove Theorem 2. To outline briefly the main idea, consider arbitrarily large $R > 0$ and $\beta^* \geq \beta_0$; then, our goal is to approximate $Z_{C;q,\beta^*}^{\text{ord}}$ in time $n^{O(1/R)}$ within relative error n^{-R} . We view $Z_{C;q,\beta}^{\text{ord}}$ as a function of β and set $g_C(\beta) := \frac{\partial \log Z_{C;q,\beta}^{\text{ord}}}{\partial \beta}$. Note in particular that

$$\begin{aligned} g_C(\beta) &= \frac{\frac{\partial Z_{C;q,\beta}^{\text{ord}}}{\partial \beta}}{Z_{C;q,\beta}^{\text{ord}}} = \frac{e^\beta}{e^\beta - 1} \frac{\sum_{F \in \Omega_C^{\text{ord}}} |F| (e^\beta - 1)^{|F|} q^{c(F)}}{Z_{C;q,\beta}^{\text{ord}}} \\ &= \frac{e^\beta}{e^\beta - 1} \mathbf{E}_{F \sim \pi_C^{\text{ord}}} [|F|] = \frac{e^\beta}{e^\beta - 1} \sum_{e \in E_C} \pi_{C;q,\beta}^{\text{ord}}(1_e). \end{aligned} \quad (2)$$

Moreover, by setting $\beta_\infty = n^{1/R}$ and integrating, we have that

$$\frac{Z_{C;q,\beta_\infty}^{\text{ord}}}{Z_{C;q,\beta^*}^{\text{ord}}} = \exp \left(\int_{\beta^*}^{\beta_\infty} g_C(\beta) d\beta \right). \quad (3)$$

We will show later that for $\beta = \beta_\infty$ it holds that $Z_{C;q,\beta_\infty}^{\text{ord}} = (1 \pm e^{-n^{\Omega(1/R)}}) q(e^\beta - 1)^{|E_C|}$. So, to obtain the desired approximation to $Z_{C;q,\beta^*}^{\text{ord}}$, it suffices to approximate the integral $I := \int_{\beta^*}^{\beta_\infty} g_C(\beta) d\beta$. The standard way to approximate the integral would be to consider a sequence of β 's between β^* and β_∞ and, for each of them, approximate $g_C(\beta)$ using the WSM guarantee. Namely, by Lemma 6, for an edge $e \in E_C$ and a vertex v incident to it, for the ball $B_e := B_r(v)$ with $r = \lceil \frac{A}{d} \log n \rceil$, it holds that

$$|\pi_{C;q,\beta}^{\text{ord}}(1_e) - \pi_{B_e^+;q,\beta}(1_e)| \leq q^{-r/30} \leq \frac{1}{n^{R+3}}, \quad (4)$$

where the last inequality follows by taking q large enough (with respect to d). Crucially, B_e is a 1-treelike² graph whose size $|V_G(B_e)|$ can be bounded by $d^r \log n = n^{O(1/R)}$; there is a fairly standard recursive way to compute $\pi_{B_e^+;q,\beta}(1_e)$ exactly (cf. Lemma 7 below). However, in order to get the desired $\frac{1}{n^R}$ relative error guarantee for the integral, we would need roughly n^R many $\hat{\beta}$'s, which would lead to a huge running time.

To obtain a faster algorithm, the key observation is that $\pi_{B_e^+;q,\beta}(1_e)$ is an explicit function of β which can be represented using an explicit rational function that can be efficiently computed (using recursions) in time polynomial in $|V_G(B_e)| = n^{O(1/R)}$. Hence we can essentially perform the integration $\int_{\beta^*}^{\beta_\infty} \pi_{B_e^+;q,\beta}(1_e)$ symbolically. A technicality that arises is that we cannot get the exact antiderivative (since in general it will be in terms of algebraic numbers) but rather a numerical approximation to the antiderivative that is within additive $1/2^{n^{O(1/R)}}$ from its true value for all $\beta \in [\beta^*, \beta_\infty]$; all of that however requires only $n^{O(1/R)}$ bits of precision and hence can be carried out in $n^{O(1/R)}$ time.

We next state more formally a few technical lemmas whose proofs are given in Appendix D and then show how to use them and conclude the proof of Theorem 2. We start with the computation of the marginals of edges in 1-treelike graphs.

Lemma 7. *There is an algorithm that, on input an n -vertex 1-treelike graph T rooted at v , an edge e incident to v and an integer $r \geq 1$, computes in time $2^{4r} n^{O(1)}$ rational functions $g^+(q, x), g^-(q, x)$ so that $\pi_{B_r^+(v);q,\beta}(1_e) = g^+(q, e^\beta - 1)$ and $\pi_{B_r^-(v);q,\beta}(1_e) = g^-(q, e^\beta - 1)$ for all $q, \beta > 0$.*

The next lemma captures the integration part of the argument. For an integer t , its size is the number of bits needed to represent it, and for a rational $w = \frac{w_1}{w_2}$ with integers w_1, w_2 , its size is given by adding the sizes of w_1 and w_2 . Moreover, for a polynomial $P(x) = \sum_{i=0}^n c_i x^i$ of degree n and rational coefficients, its size is given by $n + \sum_{i=1}^n \text{size}(c_i)$.

²For a real k , a graph is called k -treelike if it becomes a tree after the removal of at most k edges.

Lemma 8. *There is an algorithm that, on input positive rationals a, b, ε each with size $\leq n$, and integer polynomials $P(x), Q(x)$ with non-negative coefficients and size $\leq n$, computes in time $\text{poly}(n)$ a number \hat{I} so that $\hat{I} = I \pm \varepsilon$ where $I = \int_a^b \frac{P(x)}{Q(x)} dx$.*

Finally, we will need the following estimate for the partition function for very large β .

Lemma 9. *Let d be large enough. For all q sufficiently large, for any arbitrarily small constant $\varepsilon > 0$, whp over $G \sim \mathcal{G}(n, d/n)$ it holds that $Z_{C;q,\beta}^{\text{ord}} = (1 \pm e^{-n^\varepsilon})q(e^\beta - 1)^{|E_C|}$ for any $\beta \geq n^{2\varepsilon}$.*

Proof. TOPROVE 0 □

2.4 Graph properties

Before proceeding to the WSM proofs, we use the following expansion properties of $\mathcal{G}(n, d/n)$ (and its largest component). While the random graph lacks regularity and expansion from small sets, for connected subgraphs of size $\Omega(\log n)$ we can lower bound the average degree and thus obtain a weak expansion property. Since we have two graphs, G and its largest component C , for the clarity of notation, we use V_G for the vertex set of G , and E_G for its edge set. The following propositions are proved in Sections A.2 and A.3.

Proposition 10. *For any $\varepsilon \in (0, 1)$, there exists $A = A(\varepsilon) > 0$ such that for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, any connected vertex set $S_G \subseteq V_G$ with size at least $A \log(n)/d$ has average degree at least $(1 - \varepsilon)d$.*

In the following proposition, and throughout the paper, we use $\deg_G(S)$ to denote the *total degree* of a set $S \subseteq V_G$, that is, $\deg_G(S)$ is the sum of the degrees (in G) of vertices in S . For a subset $T \subseteq V_G$ we write $E_G(S, T)$ to denote the set of edges with one endpoint in S and the other in T . We also use $e_G(S) := |E_G(S, \bar{S})|$.

Proposition 11. *There exists a constant $A > 0$, such that for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. Let $S \subseteq V_C$ be a connected vertex set in C with $|S| \geq A \log(n)/d$ and $\deg_G(S) \leq \frac{1}{10} \deg_G(V_C)$. Then $e_G(S) \geq \frac{3}{5} \deg_G(S)$. The same holds if $S \subseteq V_C$ is a connected set in C with $A \log(n)/d \leq |S| \leq n/6$.*

We remark that the choice of $3/5$ in Proposition 11 is arbitrary, and any constant between $1/2$ and 1 could be used. Also, note that the lower bound $A \log(n)/d$ in both propositions is asymptotically optimal in n and d , as it can be shown that G contains an induced path of length $\approx \frac{\log n}{d}$.

Finally, we will use the property that after removing a constant fraction of edges, C still contains a large component. This property is well-known for expanders [Tre16]; to prove it for C we use the fact that its *kernel* (the graph obtained from C by contracting every induced path into a single edge and removing attached trees) is an expander. The constant $1/144$ in the lemma is somewhat arbitrary and follows from the expansion constant for the kernel (see Appendix A).

Lemma 12. *For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. Suppose that $\theta \in (0, \frac{1}{2})$ is a real and $E' \subseteq E_C$ is an edge subset such that $|E'| \leq \frac{|E_C|}{144} (2\theta - \frac{1}{100})$. Then $(V_C, E_C \setminus E')$ has a component with total degree $\geq \deg_G(V_C)(1 - \theta)$, and all other components of $(V_C, E_C \setminus E')$ have total degree at most $\theta \deg_G(V_C)$.*

In light of Lemma 12, we will order components of any subgraph C' of C by their total degree. So we use the phrase “largest component” to refer to the component of C' with largest total degree. A “giant” component is a component whose total degree is more than half of the total degree of C , and “small” otherwise. Thus, by applying Lemma 12 for $\theta = 1/10$, we have that for any $E' \in \Omega_C^{\text{ord}}$ the subgraph (V_C, E') contains a giant component and all the other components are small.

3 Weak spatial mixing within the ordered phase

In this section, we prove WSM within the ordered phase. We bound the difference of the marginals on e by the probability that, roughly, there is a long path avoiding the giant component of (V_C, F) . We then use the new definition of *ordered polymers* that naturally captures this event.

Formally for a vertex v and an integer $r \geq 1$, define $\mathcal{A}_{v,r}$ to be the event, that for a random configuration F , all simple length- r paths from v in G_C intersect the largest component of the graph $(V_C, F) \setminus v$, i.e., the graph (V_C, F) with v removed. Define $\mathcal{A}'_{v,r}$ to be the event that $|F \setminus E_G(B_r(v))| \geq (1 - \eta)|E_C|$. Then, we have that

$$|\pi_{B_r^+(v)}(1_e) - \pi_C^{\text{ord}}(1_e)| \leq 2\pi_C^{\text{ord}}(\neg \mathcal{A}_{v,r}) + 2\pi_C^{\text{ord}}(\neg \mathcal{A}'_{v,r}). \quad (5)$$

This follows by conditioning on the boundary configuration outside of $B_r(v)$. Namely, using the monotonicity of the RC model, this boundary condition is dominated above by the wired boundary condition on $B_r(v)$, where all vertices at the boundary of $B_r(v)$ are conditioned to belong to the same component, and hence $\pi_C^{\text{ord}}(1_e) \leq \pi_{B_r^+(v)}(1_e)$. Likewise, the event $\mathcal{A}'_{v,r}$ gives from Lemma 12 that there is a unique giant component C_F in the graph $(V_C \setminus B_r(v), F \setminus E_C(B_r(v)))$ and $\mathcal{A}_{v,r}$ that there is a set S of vertices inside $B_r(v)$ whose removal disconnects v from the vertices outside of $B_r(v)$ and every $w \in S$ belongs to C_F . Therefore, the event $\mathcal{A}_{v,r} \cap \mathcal{A}'_{v,r}$ implies that there is a set S of vertices inside $B_r(v)$ whose removal disconnects v from the vertices outside of $B_r(v)$ and every $w \in S$ belongs to C_F , i.e., S forms a wired boundary condition closer to v . From monotonicity, it follows that $\pi_{B_r^+(v)}(1_e) \leq \pi_C^{\text{ord}}(1_e \mid \mathcal{A}_{v,r} \cap \mathcal{A}'_{v,r})$. Combining these estimates on $\pi_{B_r^+(v)}(1_e)$ yields (5). (A similar argument can be found in [GGS24a, Lemma 5.7].)

It is relatively easy to bound the probability of $\neg \mathcal{A}'_{v,r}$ under π_C^{ord} . For large enough d , standard estimates for the neighbourhoods of $\mathcal{G}(n, d/n)$ (see also the upcoming Lemma 27) guarantee that whp over $G \sim \mathcal{G}(n, d/n)$, for any $r \leq \frac{A}{d} \log n$, it holds that $|E_G(B_r(v))| \leq \sqrt{n} \log n$ which is at most $\frac{\eta}{10}|E_C|$ for n large enough since $|E_C| = \Omega(n)$ whp. Hence, by Theorem 4(ii), $\pi_C^{\text{ord}}(\neg \mathcal{A}'_{v,r}) \leq e^{-n}$.

It remains to bound the probability of $\mathcal{A}_{v,r}$, for which we use *ordered polymers*.

3.1 Ordered Polymers

We want to bound for each vertex v , and for each $r \sim \log(n)/d$, the quantity $\pi_C^{\text{ord}}(\mathcal{A}_{v,r})$. To do so, we need to control the size of “polymers” with respect to a configuration F and a vertex $v \in V_C$.

We first consider non-giant components in $(V_C, F) \setminus v$; roughly, we will refer to the union of their vertices as the polymer vertices. To form polymers, we will take connected components of these vertices. The details are as follows.

Definition 13 (Ordered polymers). *Fix $v \in V_C$. Consider $F \in \Omega_C^{\text{ord}}$. For any $w \in V_C$, define the graph $H_v(F, w)$ as follows.*

$$H_v(F, w) := \begin{cases} \text{the component of } w \text{ in } (V_C, F), & \text{if } w = v \\ \text{the component of } w \text{ in } (V_C, F) \setminus v, & \text{if } w \neq v. \end{cases}$$

Let $\mathcal{V}(F, v) := \{w \in V_C \mid \deg_G(H_v(F, w)) \leq \frac{1}{2} \deg_G(V_C)\}$. The ordered polymers of F are subgraphs of C . We define two types of ordered polymers, non-singleton polymers and singleton polymers.

- For each maximal set $S \subseteq \mathcal{V}(F, v)$ that is connected in C , there is a non-singleton polymer. The edge set of this polymer consists of all edges in E_C that are incident to vertices in S . The vertex set of this polymer is the set of endpoints of these edges.

- For each edge $e \in E_C \setminus F$ that is not an edge of a non-singleton polymer of F , there is a singleton polymers of F consisting of the single edge e .

Let $\Gamma_v^{\text{ord}}(F)$ denote the set of all ordered polymers of F (with respect to v).

As we show in Lemma 16, a giant exists in (V_C, F) for all $F \in \Omega_C^{\text{ord}}$.

Definition 14. Fix $v \in V_C$. Fix $F \in \Omega_C^{\text{ord}}$ and $\gamma = (U, B) \in \Gamma_v^{\text{ord}}(F)$. The inner vertices of γ are $\mathcal{V}_\gamma := U \cap \mathcal{V}(F, v)$. The out-edges of γ are the edges in $E_{\text{out}}(\gamma) := (E_C \setminus F) \cap B$ and $e_{\text{out}}(\gamma) = |E_{\text{out}}(\gamma)|$ is the number of out-edges of γ . The weight of γ is $w_C^{\text{ord}}(\gamma) := q^{c'(\gamma)}(e^\beta - 1)^{-e_{\text{out}}(\gamma)}$, where $c'(\gamma)$ is the number of components of $(V_C, E_C \setminus E_{\text{out}}(\gamma))$ with total degree at most $(1/2) \deg_G(V_C)$.

Lemma 15 relates the definition of the polymers to the event $\mathcal{A}_{v,r}$. Later, in Lemma 17, we will also relate the weight of a configuration to the product of weights of its polymers.

Lemma 15. Fix $F \in \Omega_C^{\text{ord}}$. Let r be a positive integer. Suppose that there is a simple path v_0, v_1, \dots, v_r in C such that none of v_1, \dots, v_r belongs to the largest component of $(V_C, F) \setminus v_0$. Then there is a polymer $\gamma \in \Gamma_{v_0}^{\text{ord}}(F)$ with $\{v_1, \dots, v_r\} \subseteq \mathcal{V}_\gamma$, hence $|\mathcal{V}_\gamma| \geq r$.

Proof. **TOPROVE 1** □

In order to prove Lemma 17, and to bound the probability that there is a large polymer, we next use Lemma 12 to get an upper bound on a size of an ordered polymer.

Lemma 16. For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, the following holds. Fix any $v \in V_C$. For any $F \in \Omega_C^{\text{ord}}$, $(V_C, F) \setminus v$ contains a component of degree $\geq (1 - \frac{1}{10}) \deg_G(V_C)$. Hence in particular, every polymer of $\gamma \in \Gamma_v^{\text{ord}}(F)$ has $\deg_G(\mathcal{V}_\gamma) \leq \frac{1}{10} \deg_G(V_C)$.

Proof. **TOPROVE 2** □

We use Lemma 16 to relate the weight of a configuration to the product of weights of its polymers.

Lemma 17. For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. For any $F \in \Omega_C^{\text{ord}}$ and $v \in V_C$

$$w_C(F) = q(e^\beta - 1)^{|E_C|} \prod_{\gamma \in \Gamma_v^{\text{ord}}(F)} w_C^{\text{ord}}(\gamma).$$

Proof. **TOPROVE 3** □

It remains to bound the probability that there is a large polymer. To do so, we show that the weights of (sufficiently large) polymers decay, and then relate the weights to the probabilities that polymers occur.

Observe that for any $F \in \Omega_C^{\text{ord}}$, any $v \in V_C$ and any non-singleton polymer $\gamma \in \Gamma_v^{\text{ord}}(F)$, all edges in $E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})$ that are not incident to v are out-edges. To see this, note that, if there is an in-edge from $w \neq v \in \overline{\mathcal{V}_\gamma}$ to $u \in \mathcal{V}_\gamma$, then the vertices u and w would be in the same component of $H_v(F, w)$, but then by the definition of non-singleton polymers, w would be in \mathcal{V}_γ . Since, for any singleton polymer γ , $\mathcal{V}_\gamma = \emptyset$, the same statement applies. Formally we have the following.

Observation 18. Let $F \in \Omega_C^{\text{ord}}$, $v \in V_C$ and $\gamma \in \Gamma_v^{\text{ord}}(F)$. Then $E_{\text{out}}(\gamma) \supseteq E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma}) \setminus E_G(v, \mathcal{V}_\gamma)$, and hence $e_{\text{out}}(\gamma) \geq |E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})| - |E_G(v, \mathcal{V}_\gamma)|$.

We next use that whp G is 1-treelike (see Lemma 28), i.e. that for all d large enough, whp any radius- $\frac{\log n}{d}$ ball in G (and thus in C) can be made into a tree by removing at most a single edge, to show that $E_G(v, \mathcal{V}_\gamma)$ accounts for a small fraction of edges going from \mathcal{V}_γ .

Lemma 19. *For all real d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. For any $v \in V_G$ and for any connected subset $S \subseteq V_G$ that doesn't contain v , $|E_G(v, S)| \leq 2 + \frac{d|S|}{\log n}$.*

Proof. TOPROVE 4 □

Corollary 20. *For all real d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. Let $v \in V_C$, $F \in \Omega^{\text{ord}}$. Then for any $\gamma \in \Gamma_v^{\text{ord}}(F)$, it holds that $e_{\text{out}}(\gamma) \geq |E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})| - 2 - \frac{d|\mathcal{V}_\gamma|}{\log n}$.*

Proof. TOPROVE 5 □

Now we can bound weight of any ordered polymer γ , provided it is large enough, to apply expansion and average degree bounds.

Lemma 21. *There exists a real $A > 0$ such that for all d large enough, and for all real $q > 1$, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds for all $\beta \geq \beta_0$. For any $F \in \Omega_C^{\text{ord}}$, $v \in V_C$ and $\gamma \in \Gamma_v^{\text{ord}}(F)$ with $|\mathcal{V}_\gamma| \geq \frac{A}{d} \log n$, it holds that $w_C^{\text{ord}}(F) \leq q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}$.*

Proof. TOPROVE 6 □

Remark 22. *Note that Lemma 21 crucially requires d to be large enough for sufficient expansion and average degree properties to hold for all connected sets of size at least $A \log n/d$. While expansion properties hold for all $d > 1$ (see e.g. [FR07]), we must treat sets as large as $A \log n/d$ and for these we require large d to obtain the required average degree properties and to show that the weights of the polymers are sufficiently small.*

We next relate the weight of an ordered polymer to the probability of it occurring.

Lemma 23. *For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. For $v \in V_C$ and any $\gamma \in \bigcup_{F \in \Omega_C^{\text{ord}}} \Gamma_v^{\text{ord}}(F)$, it holds that $\pi_C^{\text{ord}}(\gamma \in \Gamma_v^{\text{ord}}(\cdot)) \leq w_C^{\text{ord}}(\gamma)$.*

Proof. TOPROVE 7 □

With this in hand, we can show an analogue of the Kotecký-Preiss condition for our ordered polymers (but only considering those that are large enough). We will use this to bound the probability of a large polymer occurring.

Proposition 24. *There is a real $A > 0$ such that for all d large enough, for all sufficiently large reals q , the following holds whp over $G \sim \mathcal{G}(n, d/n)$ for all real $\beta \geq \beta_0$. For any $v \in V_C$,*

$$\pi_C^{\text{ord}}(\exists \gamma \in \Gamma_v^{\text{ord}}(\cdot) \text{ with } |\mathcal{V}_\gamma| \geq \frac{A}{d} \log n) \leq q^{-\frac{A}{25d} \log n}.$$

Proof. TOPROVE 8 □

Proposition 24, Lemma 15 and equation (5) together with the bound from Theorem 4 now directly give Lemma 6, i.e. WSM within the ordered phase at the distance $r = \frac{A \log n}{d}$ (for some constant A – take twice the constant from the Proposition 24 to take into account that we want to avoid polymers of size $r - 1$).

4 Fast convergence of the RC dynamics and proof of Theorem 1

In order to prove Theorem 1, we use the RC dynamics $(X_t)_{t \geq 0}$ to obtain samples from the ordered and the disordered phases separately. For the sake of completeness, recall that a transition from X_t to X_{t+1} is done as follows:

- (1) Pick an edge $e \in E$ uniformly at random
- (2) If e is a cut edge of the graph $(V, X_t \cup \{e\})$, then with probability $\hat{p} := \frac{e^\beta}{e^\beta + q - 1}$ set $X_{t+1} := X_t \cup \{e\}$. With all remaining probability set $X_{t+1} := X_t \setminus \{e\}$
- (3) Otherwise, with probability $p := 1 - e^{-\beta}$, set $X_{t+1} := X_t \cup \{e\}$. With all remaining probability set $X_{t+1} := X_t \setminus \{e\}$.

Proving mixing of the RC dynamics (X_t) is in general a challenging task, especially since its moves depend on the current component structure. Fortunately, by now there is a well-established strategy [GS22, GGS24b, GGS24a] that we can use to show the convergence bounds of Theorems 25 and 26 based on the polymer sizes. It consists of utilising the following two main ingredients: (i) *weak spatial mixing within a phase* (WSM) that as we saw asserts that the marginal probability of an edge e being an in-edge depends only on a small neighbourhood around e , and (ii) the mixing time of the dynamics on that neighbourhood is polynomial in its size. Ingredient (ii) is fairly straightforward in sparse random graphs since small-distance neighbourhoods are typically tree-like (i.e., they contain at most a constant number of cycles) which makes poly-time bounds relatively simple to obtain.

Theorem 25. *Let d be sufficiently large. Then, for all q sufficiently large, whp over $G \sim \mathcal{G}(n, d/n)$, for all $\beta \leq \beta_1$ and $\varepsilon > e^{-\Omega(n)}$, the random cluster dynamics X_t initialized from all-out satisfies $\|X_T - \pi_C^{\text{dis}}\| \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.*

Theorem 26. *Let d be sufficiently large. Then, for all q sufficiently large, whp over $G \sim \mathcal{G}(n, d/n)$, for all $\beta \geq \beta_0$ and $\varepsilon > e^{-\Omega(n)}$, the random cluster dynamics X_t initialized from all-in satisfies $\|X_T - \pi_C^{\text{ord}}\| \leq \varepsilon$ for $T = \text{poly}(n, \log \frac{1}{\varepsilon})$.*

Proof. **TOPROVE 9** □

Given WSM within the phases and using the monotonicity of the random-cluster model, we complete the proof of Theorems 25 and 26 in Appendix E.

References

- [Bar17] A. Barvinok. *Combinatorics and Complexity of Partition Functions*. Algorithms and Combinatorics. Springer International Publishing, 2017.
- [BCM19] Jean Barbier, Chun Lam Chan, and Nicolas Macris. Concentration of Multi-overlaps for Random Dilute Ferromagnetic Spin Models. *Journal of Statistical Physics*, 180(1-6):534–557, December 2019.
- [BCvV23] Antonio Blanca, Zongchen Chen, Daniel Štefankovič, and Eric Vigoda. The Swendsen–Wang dynamics on trees. *Random Structures & Algorithms*, 62(4):791–831, 2023.

- [BG21] Antonio Blanca and Reza Gheissari. Random-cluster dynamics on random regular graphs in tree uniqueness. *Communications in Mathematical Physics*, 386(2):1243–1287, 2021.
- [BG23a] Antonio Blanca and Reza Gheissari. On the tractability of sampling from the Potts model at low temperatures via Swendsen-Wang dynamics. *CoRR*, abs/2304.03182, 2023.
- [BG23b] Antonio Blanca and Reza Gheissari. Sampling from Potts on random graphs of unbounded degree via random-cluster dynamics. *The Annals of Applied Probability*, 33(6B):4997 – 5049, 2023.
- [BGG⁺20a] Antonio Blanca, Andreas Galanis, Leslie Ann Goldberg, Daniel Štefankovič, Eric Vigoda, and Kuan Yang. Sampling in uniqueness from the Potts and random-cluster models on random regular graphs. *SIAM Journal on Discrete Mathematics*, 34(1):742–793, 2020.
- [BGG⁺20b] Antonio Blanca, Andreas Galanis, Leslie Ann Goldberg, Daniel Štefankovič, Eric Vigoda, and Kuan Yang. Sampling in uniqueness from the Potts and random-cluster models on random regular graphs. *SIAM Journal on Discrete Mathematics*, 34(1):742–793, 2020.
- [BGGv24] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, and Daniel Štefankovič. Fast sampling via spectral independence beyond bounded-degree graphs. *ACM Trans. Algorithms*, 20(1), 2024.
- [BZ23] Antonio Blanca and Xusheng Zhang. Rapid mixing of global markov chains via spectral independence: The unbounded degree case. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2023*, volume 275, pages 53:1–53:19, 2023.
- [CDF⁺22] Charlie Carlson, Ewan Davies, Nicolas Fraiman, Alexandra Kolla, Aditya Potukuchi, and Corrine Yap. Algorithms for the ferromagnetic Potts model on expanders. In *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 344–355, 2022.
- [CDK⁺20] Matthew Coulson, Ewan Davies, Alexandra Kolla, Viresh Patel, and Guus Regts. Statistical Physics Approaches to Unique Games. In *35th Computational Complexity Conference (CCC 2020)*, volume 169, pages 13:1–13:27, 2020.
- [CDKP24] Charlie Carlson, Ewan Davies, Alexandra Kolla, and Aditya Potukuchi. A spectral approach to approximately counting independent sets in dense bipartite graphs. In *51st International Colloquium on Automata, Languages, and Programming (ICALP 2024)*, volume 297, pages 35:1–35:18, 2024.
- [CGG⁺16] Jin-Yi Cai, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Mark Jerrum, Daniel Štefankovič, and Eric Vigoda. #bis-hardness for 2-spin systems on bipartite bounded degree graphs in the tree non-uniqueness region. *Journal of Computer and System Sciences*, 82(5):690–711, 2016.
- [CGG⁺21] Zongchen Chen, Andreas Galanis, Leslie A. Goldberg, Will Perkins, James Stewart, and Eric Vigoda. Fast algorithms at low temperatures via Markov chains. *Random Structures & Algorithms*, 58(2):294–321, 2021.

- [COGG⁺23] Amin Coja-Oghlan, Andreas Galanis, Leslie Ann Goldberg, Jean Bernoulli Ravelomanana, Daniel Štefankovič, and Eric Vigoda. Metastability of the Potts ferromagnet on random regular graphs. *Communications in Mathematical Physics*, pages 1–41, 2023.
- [DLP14] Jian Ding, Eyal Lubetzky, and Yuval Peres. Anatomy of the giant component: The strictly supercritical regime. *European Journal of Combinatorics*, 35:155–168, 2014.
- [EF23] Charilaos Efthymiou and Weiming Feng. On the mixing time of Glauber dynamics for the hard-core and related models on $g(n, d/n)$. In *50th International Colloquium on Automata, Languages, and Programming (ICALP 2023)*, volume 261, pages 54:1–54:17, 2023.
- [EHŠV18] Charilaos Efthymiou, Thomas P. Hayes, Daniel Štefankovič, and Eric Vigoda. Sampling random colorings of sparse random graphs. In *Proceedings of the 2018 Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1759–1771, 2018.
- [FK15] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2015.
- [FR07] Nikolaos Fountoulakis and Bruce Reed. The evolution of the mixing rate, 2007.
- [GGS21] Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast algorithms for general spin systems on bipartite expanders. *ACM Trans. Comput. Theory*, 13(4), September 2021.
- [GGS22] Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast mixing via polymers for random graphs with unbounded degree. *Information and Computation*, 285, 2022.
- [GGS24a] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model, 2024.
- [GGS24b] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Sampling from the random cluster model on random regular graphs at all temperatures via glauber dynamics. *Combinatorics, Probability and Computing*, page 1–33, 2024.
- [GJ97] Vivek K Gore and Mark R Jerrum. The Swendsen-Wang process does not always mix rapidly. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 674–681, 1997.
- [GJ12] Leslie Ann Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM*, 59(5):1–31, 2012.
- [GLP20] Reza Gheissari, Eyal Lubetzky, and Yuval Peres. Exponentially slow mixing in the mean-field Swendsen–Wang dynamics. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, 56(1):68 – 86, 2020.
- [GS22] Reza Gheissari and Alistair Sinclair. Low-temperature Ising dynamics with random initializations. In *54th Annual ACM SIGACT Symposium on Theory of Computing*, (STOC ’22), pages 1445–1458, 2022.
- [GS23] Reza Gheissari and Alistair Sinclair. Spatial mixing and the random-cluster dynamics on lattices. In *Proceedings of the 2023 Annual ACM-SIAM Symposium on Discrete Algorithms*, (SODA ’23) , pages 4606–4621, 2023.

- [GvVY16] Andreas Galanis, Daniel Štefankovic, Eric Vigoda, and Linji Yang. Ferromagnetic Potts model: Refined #bis-hardness and related results. *SIAM Journal on Computing*, 45(6):2004–2065, 2016.
- [HJP23] Tyler Helmuth, Matthew Jenssen, and Will Perkins. Finite-size scaling, phase coexistence, and algorithms for the random cluster model on random graphs. *Annales de l’Institut Henri Poincaré (B) Probabilités et statistiques*, 59(2):817–848, 2023.
- [HPR19] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov-Sinai theory. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, (STOC ’19), pages 1009–1020, 2019.
- [JKP20] Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM Journal on Computing*, 49(4):681–710, 2020.
- [JLR00] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. John Wiley & Sons, Inc., 2000 - 2000.
- [JPP23] Matthew Jenssen, Will Perkins, and Aditya Potukuchi. Approximately counting independent sets in bipartite graphs via graph containers. *Random Structures & Algorithms*, 63(1):215–241, 2023.
- [JPPS24] Matthew Jenssen, Will Perkins, Aditya Potukuchi, and Michael Simkin. Sampling, counting, and large deviations for triangle-free graphs near the critical density. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 151–165, 2024.
- [KP86] R. Kotecký and D. Preiss. Cluster expansion for abstract polymer models. *Comm. Math. Phys.*, 103(3):491–498, 1986.
- [LLLM19] Chao Liao, Jiabao Lin, Pinyan Lu, and Zhenyu Mao. Counting independent sets and colorings on random regular bipartite graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2019)*, volume 145, pages 34:1–34:12, 2019.
- [LMRW24] Kuikui Liu, Sidhanth Mohanty, Amit Rajaraman, and David X. Wu. Fast mixing in sparse random Ising models. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 120–128, 2024.
- [Mos71] Joel Moses. Symbolic integration: the stormy decade. *Commun. ACM*, 14(8):548–560, 1971.
- [MS13] Elchanan Mossel and Allan Sly. Exact thresholds for Ising Gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294 – 328, 2013.
- [NR96] C.Andrew Neff and John H. Reif. An efficient algorithm for the complex roots problem. *Journal of Complexity*, 12(2):81–115, 1996.
- [PR17] Viresh Patel and Guus Regts. Deterministic polynomial-time approximation algorithms for functions and graph polynomials. *SIAM Journal on Computing*, 46(6):1893–1919, 2017.

- [Tre16] Luca Trevisan. Lecture notes on graph partitioning, expanders and spectral methods, 2016. URL: <https://lucatrevisan.github.io/books/expanders-2016.pdf>.
- [YZ16] Yitong Yin and Chihao Zhang. Sampling in Potts model on sparse random graphs. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2016)*, volume 60, pages 47:1–47:22, 2016.

A Proofs of properties of $\mathcal{G}(n, d/n)$

In this section we prove the properties from Section 2.4, as well as introduce a few more results used in proving mixing within the disordered phase.

A.1 Preliminaries

To bound the size of radius- r balls, we use the following lemma by Mossel and Sly [MS13]. Note that while their lemma is more general and stated for $r = (\log \log n)^2$, this bound on r was not used for upper bounding the volume.

Lemma 27 ([MS13, Lemma 7]). *Let $d > 1$ be a real. Then whp over $G \sim \mathcal{G}(n, d/n)$, for any integer r and any $v \in V(G)$, $|V(B_r(v))| \leq d^r \log n$.*

Next we show local treelikeness properties of $\mathcal{G}(n, d/n)$. we say that a subgraph is k -treelike if at most k edges can be deleted from it to make it a tree (or a forest).

First we will consider balls of radius $\approx \frac{1}{d} \log n$. Recall that the radius- r ball around a vertex v is an induced subgraph consisting of all vertices with distance $\leq r$ to v .

Lemma 28. *Let $A > 0$ be a real. Then for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, each radius- $\frac{A}{d} \log n$ ball in G is 1-treelike.*

Proof. TOPROVE 10 □

We also show $O(1)$ -treelikness for sets of constant and $O(\log n)$ size.

Lemma 29. *For any $d > 1$, there exists $k > 0$ such that whp over $G \sim \mathcal{G}(n, d/n)$, any connected subset of size $\leq \log n$ is k -treelike.*

Proof. TOPROVE 11 □

Lemma 30. *For any $d > 1$, and an integer $R > 0$, whp over $G \sim \mathcal{G}(n, d/n)$, any connected vertex subset of G with at most R vertices is 1-treelike (i.e we can remove at most one edge to make it a tree).*

Proof. TOPROVE 12 □

The last lemma we will need before going on to proving Propositions 10 and 11, will be to bound the expected number of connected sets by given size. This result is proven in [FR07].

Lemma 31 ([FR07, (3)]). *For all reals $d > 1$ and integers $s > 0$, the expected number of subtrees with s vertices in $\mathcal{G}(n, d/n)$ is at most $n(ed)^s$.*

A.2 Average degree of connected sets

In this subsection, we prove Proposition 10. We split the proof into three parts, depending on the size of the set. For sets of $\Omega(n)$ size, we can prove a stronger result – they do not have to be connected for the average degree bound to hold.

Lemma 32. *Let $\xi < \varepsilon \in (0, 1)$. Then there exists a real $A(\varepsilon, \xi) > 0$ such that for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ it holds that all connected sets of size $\frac{A \log n}{d}$ and at most ξn have average degree at least $(1 - \varepsilon)d$.*

Proof. TOPROVE 13 □

Lemma 33. *Let $\xi, \varepsilon \in (0, 1)$ be arbitrary constants. Then, for all $d \geq \frac{32 \log 2}{(\varepsilon \xi)^2}$, whp over $G \sim \mathcal{G}(n, d/n)$, all vertex sets $S \subseteq V_G$ of size at least ξn and at most $(1 - \xi)n$ have average degree at least $(1 - \varepsilon)d$, and in particular have $|E_G(S, \bar{S})| \geq \frac{d}{n}(1 - \varepsilon)|S|(n - |S|)$.*

Proof. TOPROVE 14 □

Lemma 34. *For $\xi, \varepsilon \in (0, 1)$ with $\xi < \varepsilon/2$, and all $d > \frac{8 \log 2}{(\varepsilon - 2\xi)^2}$, whp over $G \sim \mathcal{G}(n, d/n)$, all vertex sets $S \subseteq V(G)$ of size $\geq (1 - \xi)n$ have $\deg_G(S) \geq (1 - \varepsilon)d|S|$.*

Proof. TOPROVE 15 □

Now the Proposition 10 follows directly by applying Lemmas 32, 33 and 34 we get the proposition with, say $\xi := \varepsilon/3 < \varepsilon/2$.

A.3 Weak Expansion

In this section, we prove Proposition 11. As a first step, we show that the number of interval edges of connected subsets is not too large.

Lemma 35. *For any $\xi \in (0, 1]$, there exists a real $A(\xi) > 0$ such that for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, any connected vertex set $S \subseteq V_G$ with $\frac{A \log n}{d} \leq |S| \leq \xi n$ satisfies $e_G(S) < (2\xi d/3 + 1)|S|$.*

Proof. TOPROVE 16 □

With this lemma in hand, we can prove Proposition 11.

Proposition 11. *There exists a constant $A > 0$, such that for all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. Let $S \subseteq V_C$ be a connected vertex set in C with $|S| \geq A \log(n)/d$ and $\deg_G(S) \leq \frac{1}{10} \deg_G(V_C)$. Then $e_G(S) \geq \frac{3}{5} \deg_G(S)$. The same holds if $S \subseteq V_C$ is a connected set in C with $A \log(n)/d \leq |S| \leq n/6$.*

Proof. TOPROVE 17 □

A.4 Kernel of $\mathcal{G}(n, d/n)$

While $\mathcal{G}(n, d/n)$ (and C) only exhibit weak expansion, its *kernel* – that is, a graph obtained by taking a maximal subgraph of C with minimum degree ≥ 2 and then contracting maximal induced paths into a single edge – is an expander. This will allow to prove Lemma 12, as well as Theorem 4. Note that the difference between C and its maximal subgraph of minimum degree ≥ 2 is a *forest*, and we to it as the *attached trees* in C .

We will denote the kernel as $G_K = (V_K, E_K)$ and write $n_K := |V_K|$. To reason about the kernel, and show its expansion, we are going to use the following contiguous model for the kernel from [DLP14].

Proposition 36 ([DLP14, Theorem 1]). *Let $d > 1$ be a constant and $\mu \in (0, 1)$ be its conjugate (i.e. the unique solution in $(0, 1)$ of $de^{-d} = \mu e^{-\mu}$). Let $n > 1$ be an integer. Consider the following graph model $\mathbb{K}(n, d)$ generated as follows:*

Let $\Delta \sim \mathcal{N}(d - \mu, 1/n)$ and then let D_1, \dots, D_n be i.i.d. $\text{Poisson}(\Delta)$ for conditional on $\sum_i^n D_i \mathbb{1}_{D_i \geq 3}$ even.

The degree sequence of $\mathbb{K}(n, d)$ is $\{D_i\}_{D_i \geq 3}$, and $\mathbb{K}(n, d)$ is a (multi)graph with this degree sequence chosen u.a.r..

Then $\mathbb{K}(n, d)$ is contiguous to the kernel G_K of $\mathcal{G}(n, d/n)$, i.e. whenever a statement holds whp about $\mathbb{K}(n, d)$, it also holds whp about G_K .

Using this model, we first use the following series of bounds the size of the kernel. As a first step, we bound the conjugate of d .

Lemma 37. *For all $d \geq 3$, $\mu < e^{-d/2}$.*

Proof. TOPROVE 18 □

With an upper bound on μ , the following two lemmas give us a lower bound on the size, and the number of edges in the kernel (and thus in C).

Lemma 38. *For $d \geq 4$, whp over $K \sim \mathbb{K}(n, d)$, $n_K \geq n(1 - e^{-d/3})$. Thus also whp over $G \sim \mathcal{G}(n, d/n)$, $n_K \geq n(1 - e^{-d/3})$.*

Proof. TOPROVE 19 □

We use the following standard Chernoff bound for real $0 < \delta < \lambda$ and $X \sim \text{Poisson}(\lambda)$.

$$\mathbb{P}(X \leq d - \delta) \leq e^{-\frac{\delta^2}{\delta + d}} \quad (6)$$

Lemma 39. *For any real $d \geq 4$, whp over over $K \sim \mathbb{K}(n, d)$, $\frac{dn}{2} - 2ne^{-d/3} \leq |E_K|$. Thus also whp over $G \sim \mathcal{G}(n, d/n)$, $\frac{dn}{2} - 2ne^{-d/3} \leq |E_K| \leq |E_C|$*

Proof. TOPROVE 20 □

With a lower bound on $|E_K|$, we can now bound $|E_C|/|E_K|$.

Lemma 40. *For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, $|E_C| \leq (1 + e^{-d/3})|E_K|$.*

Proof. TOPROVE 21 □

Using the kernel we also show a stronger variant of the average-degree result for linear-sized sets, by bounding the number of low-degree vertices in the kernel. We use the following notation. For a vertex set $S_G \subseteq V_G$, let the corresponding *kernel-representation* be the set $S_K \subseteq V_K$ containing all vertices which have their counterpart in S_G (i.e. that have degree > 2 in G and are part of the maximal subgraph of minimal degree ≥ 2 of C). Note that $\deg_K(\cdot)$, $E_K(\cdot, \cdot)$ are defined analogously to $\deg_G(\cdot)$, $E_G(\cdot, \cdot)$.

Proposition 41. *For all real d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, the number of vertices in G , (and thus in its largest component C and its kernel G_K) of degree $\leq (1 - d^{-1/3})d$ is at most $2ne^{-\frac{1}{2}d^{1/3}}$. Hence then for any vertex set $S_G \subseteq V_G$*

$$\deg_G(S_G) \geq \deg_K(S_K) \geq (1 - d^{-1/3})d|S_G| - 2dne^{-\frac{1}{2}d^{1/3}}$$

Proof. **TOPROVE 22** □

A.4.1 Expansion of the kernel

Next, we prove strong expansion properties of the kernel. Due to the absence of degree-2 and degree-1 vertices, the expansion property holds for subgraphs without a lower bound on their size.

To handle expansion of $O(\log n)$ sets, we will make use of the following result from [GGS22], which gives expansion properties of small sets whp over random graphs with fixed degree sequences with a) all degrees between 3 and $n^{1/50}$, and b) sum of degree square being $O(n)$. Note that in $K \sim \mathbb{K}(n, d)$, whp all degrees are between 3 and $O(\frac{\log n}{\log \log n})$ (which is at most $n^{1/50}$ for all n large enough). Then sum of squares of degrees of $K \sim \mathbb{K}(n, d)$ is dominated by sum of n i.i.d. $\text{Poisson}(d)^2$ variables. By Chebyshev, with probability $1 - O(1/n)$, that sum is at most $2(d^2 + d)n$, which is (whp) linear in n_K (Lemma 38), thus we can apply their result.

Lemma 42 ([GGS22, Lemma 15]). *Whp over $K \sim \mathbb{K}(n, d)$, for any connected kernel vertex set $S \subseteq V_K$ with $|S| \leq \log^2 n_K$, $|E_K(S, \bar{S})| \geq \deg_K(S)/144$.*

We note that while [GGS22] gives total-degree-expansion for *all* sets of size $\leq n_K/2$, the expansion on sets of larger size is worse than we require. To improve on their result, we make use of the weak expansion of G . Observe that we can directly translate the size of the respective cuts.

Observation 43. *For any vertex set of the kernel, $S_K \subseteq V_K$, let $S_G \subseteq V_G$ be the vertex set containing: all vertices corresponding to those in S_K , all vertices lying on the expanded edges of $G_K[S_K]$, and all vertices belonging to the attached trees adjacent to vertices corresponding to S_K in C . Then $|E_K(S_K, \bar{S}_K)| = |E_G(S_G, \bar{S}_G)|$.*

Proof. **TOPROVE 23** □

With these two results in hand, we can prove expansion of the kernel.

Proposition 44. *For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$, the kernel G_K is a $\frac{1}{144}$ -total-degree-expander, i.e. for every set $S \subseteq V_K$ with $\deg_K(S) \leq \frac{1}{2} \deg_K(G_K)$, $|E_K(S, \bar{S})| \geq \frac{1}{144} \deg_K(S)$.*

Proof. **TOPROVE 24** □

A.4.2 Proof of Lemma 12

Now we are equipped to prove Lemma 12. To do so, we first prove similar statements for the case of degree-expander graphs. The following Lemma is a modification of [Tre16, Lemma 2.3].

Lemma 45 (Analogue of [Tre16, Lemma 2.3]). *Let $H = (V_H, E_H)$ be a α -total degree expander, i.e. for any vertex set $S \subseteq V$ with $\deg_H(S) \leq \frac{1}{2} \deg_H(H)$, $E_H(S, \bar{S}) \geq \alpha \deg_H(S)$.*

Then, for any $\theta \in (0, \alpha)$ and any subset of edges $E' \subseteq E_H$ with $|E'| \leq \theta |E_H|$, it holds that $(V_H, E_H \setminus E')$ contains a component with total degree at least $(1 - \frac{\theta}{2\alpha}) \deg_H(V_H)$.

Proof. TOPROVE 25 □

Since whp the kernel is $\frac{1}{144}$ -total-degree expander, we can apply this lemma for the kernel. We use that the number of edges in C is a small fraction away from the number of kernel edges, and hence get Lemma 12, which we restate for convenience.

Lemma 12. *For all d large enough, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. Suppose that $\theta \in (0, \frac{1}{2})$ is a real and $E' \subseteq E_C$ is an edge subset such that $|E'| \leq \frac{|E_C|}{144} (2\theta - \frac{1}{100})$. Then $(V_C, E_C \setminus E')$ has a component with total degree $\geq \deg_G(V_C)(1 - \theta)$, and all other components of $(V_C, E_C \setminus E')$ have total degree at most $\theta \deg_G(V_C)$.*

Proof. TOPROVE 26 □

B Weak spatial mixing within the disordered phase

In this section, we prove WSM within the disordered phase at a distance $r \sim \frac{\log n}{d}$. We make use of [BG21, Lemma 5.3], which says that in order to bound $|\pi_C^{\text{dis}}(1_e) - \pi_{B_r^-(v)}(1_e)|_{\text{TV}}$, it is enough to bound the event that there is component of in-edges of size $\geq r - 1$. To capture the probability of this event we use *disordered polymers* (analogously to [HJP23]).

Definition 46. *For $F \in \Omega_C^{\text{dis}}$, the disordered polymers of F are connected components of (V_C, F) . Let $\Gamma^{\text{dis}}(F)$ be the set of all disordered polymers of F . For a disordered polymer $\gamma = (U, B) \in \Gamma^{\text{dis}}(F)$, its weight is defined as $w_C^{\text{dis}}(\gamma) := q^{1-|U|} (e^\beta - 1)^{|B|}$.*

Recall that $\beta_1 = \beta_1(q)$ is the temperature satisfying $e^{\beta_1} - 1 = q^{(2+1/10)/d}$. We next show that when $\beta \leq \beta_1$, then weight of all sufficiently large polymers is exponentially decaying in their size.

Lemma 47. *There exists a real $A > 0$, such that for all d sufficiently large, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds for all real parameters $0 < q, 0 < \beta \leq \beta_1$. Let $F \in \Omega_C^{\text{dis}}$, and $\gamma \in \Gamma^{\text{dis}}(F)$. If $|V_G(\gamma)| \geq \frac{A \log n}{d}$, then also*

$$w_C^{\text{dis}}(\gamma) \leq q^{1-|V_G(\gamma)|/10}.$$

Proof. TOPROVE 27 □

Next we show that polymer weight is an upper bound on the the probability of it occurring in the disordered phase.

Lemma 48. *For all connected graphs C , the following holds. For any $F \in \Omega_C^{\text{dis}}$,*

$$w_C(F) = q^{n_C} \prod_{\gamma \in \Gamma^{\text{dis}}(F)} w_C^{\text{dis}}(\gamma),$$

and for any $\gamma \in \bigcup_{F \in \Omega_C^{\text{dis}}} \Gamma^{\text{dis}}(F)$, it holds that $\pi_C^{\text{dis}}(\gamma \in \Gamma^{\text{dis}}(\cdot)) \leq w_C^{\text{dis}}(\gamma)$.

Proof. **TOPROVE 28** □

Now, we can bound that with good probability, there are no large polymers.

Proposition 49. *There exists a real $A > 0$, such that for all sufficiently large d , for all sufficiently large q , whp over $G \sim \mathcal{G}(n, d/n)$ the following holds for all $0 < \beta \leq \beta_1$.*

$$\pi_C^{\text{dis}}(\exists \gamma \in \Gamma^{\text{dis}}(\cdot) \text{ with } |V_G(\gamma)| \geq \frac{A \log n}{d}) \leq q^{-A \log n / (25d)}$$

Proof. **TOPROVE 29** □

Now we can prove WSM within the disordered phase.

Proposition 50. *There exists a real $A > 0$, such that for all d large enough, and then for all q large enough reals, whp over $G \sim \mathcal{G}(n, d/n)$, $C = C(G)$ has WSM within the disordered phase at distance $\lceil \frac{A \log n}{d} \rceil$ with rate $q^{-1/30}$.*

Proof. **TOPROVE 30** □

C Proof of Theorem 4

In this section, we prove Theorem 4 using the following five lemmas. First we show that for ordered phase vanishes for all $\beta \leq \beta_0$.

Lemma 51. *Let d be a large enough real. For all sufficiently large real q , whp over $G \sim \mathcal{G}(n, d/n)$, it holds that $Z_C^{\text{ord}}/Z_C = e^{-n}$ for all $\beta \leq \beta_0$.*

Proof. **TOPROVE 31** □

Analogously, the contribution of the disordered phase becomes negligible for $\beta \geq \beta_1$.

Lemma 52. *Let d be a large enough real. For all sufficiently large real q , whp over $G \sim \mathcal{G}(n, d/n)$ it holds that $Z_C^{\text{dis}}/Z_C = e^{-n}$ for all $\beta \geq \beta_1$.*

Proof. **TOPROVE 32** □

It remains to show that the configurations with $\leq (1 - \frac{9\eta}{10})|E_C|$ and $\geq \frac{9\eta}{10}|E_C|$ edges are very unlikely. First we show this for the Gibbs distribution without conditioning.

Lemma 53. *Let d be a sufficiently large real. For all sufficiently large real q , the following holds whp over $G \sim \mathcal{G}(n, d/n)$, for any real $\beta > 0$.*

Let \mathcal{B} be the set of configurations $F \in \Omega_C$ with $\frac{9}{10}\eta \leq \frac{|F|}{|E_C|} \leq 1 - \frac{9}{10}\eta$. Then, $\pi_C(\mathcal{B}) \leq e^{-n}$.

Proof. **TOPROVE 33** □

We now prove analogous lemmas for π_C^{dis} and π_C^{ord} .

Lemma 54. *Let d be a sufficiently large real. For all sufficiently large real q , whp over $G \sim \mathcal{G}(n, d/n)$, the following holds for all real $\beta \leq \beta_1$.*

Let \mathcal{B}_1 be the set of configurations $F \in \Omega_C$ with $|F| \geq \frac{9}{10}\eta|E_C|$. Then, $\pi_C^{\text{dis}}(\mathcal{B}_1) \leq e^{-n}$.

Proof. **TOPROVE 34** □

Lemma 55. *Let d be a sufficiently large real. For all sufficiently large real q , whp over $G \sim \mathcal{G}(n, d/n)$, the following holds for any real $\beta \geq \beta_0$.*

Let \mathcal{B}_0 be the set of configurations $F \in \Omega_C$ with $|F| \leq (1 - \frac{9}{10}\eta)|E_C|$. Then, $\pi_C^{\text{ord}}(\mathcal{B}_0) \leq e^{-n}$.

Proof. **TOPROVE 35** □

Now Theorem 4 follows directly from these five lemmas.

D Remaining proofs for Theorem 2

In this section, we prove Lemmas 7, 8 and 9 from Section 2.3, thus completing the proof of Theorem 2.

Lemma 7. *There is an algorithm that, on input an n -vertex 1-treelike graph T rooted at v , an edge e incident to v and an integer $r \geq 1$, computes in time $2^{4r} n^{O(1)}$ rational functions $g^+(q, x), g^-(q, x)$ so that $\pi_{B_r^+(v); q, \beta}(1_e) = g^+(q, e^\beta - 1)$ and $\pi_{B_r^-(v); q, \beta}(1_e) = g^-(q, e^\beta - 1)$ for all $q, \beta > 0$.*

Proof. TOPROVE 36 □

Lemma 8. *There is an algorithm that, on input positive rationals a, b, ε each with size $\leq n$, and integer polynomials $P(x), Q(x)$ with non-negative coefficients and size $\leq n$, computes in time $\text{poly}(n)$ a number \hat{I} so that $\hat{I} = I \pm \varepsilon$ where $I = \int_a^b \frac{P(x)}{Q(x)} dx$.*

Proof. TOPROVE 37 □

Lemma 9. *Let d be large enough. For all q sufficiently large, for any arbitrarily small constant $\varepsilon > 0$, whp over $G \sim \mathcal{G}(n, d/n)$ it holds that $Z_{C; q, \beta}^{\text{ord}} = (1 \pm e^{-n^\varepsilon}) q(e^\beta - 1)^{|E_C|}$ for any $\beta \geq n^{2\varepsilon}$.*

Proof. TOPROVE 38 □

E Remaining proofs for Theorem 1

In this section, we finish the proof of Theorems 25 and 26.

E.1 Local mixing on treelike balls

In addition to WSM within the phase, we need another proof ingredient: rapid mixing of the Glauber dynamics on treelike balls with wired and free boundary condition.

For the free boundary condition, we can directly use the Lemma 6.7. from [BG23b]:

Lemma 56 ([BG23b, Lemma 6.7]). *Let $B = (U, A)$ be a k -treelike graph. For every $q, \beta > 0$, there exists $\alpha_0(q, \beta, k)$ such that the Glauber dynamics X_t on B satisfies,*

$$\max_{X_0 \in \Omega_B} \|X_t - \pi_{B^+}\|_{\text{TV}} \leq \frac{1}{\sqrt{2}} e^{-\alpha_0 t} \log \left(\frac{1}{\min_{F \in \Omega_B} \pi_{B^-}(F)} \right)^{-1},$$

where π_{B^-} is the Gibbs distribution on B with the free boundary condition (no additional wirings).

We use this Lemma for $B = B_r(v)$. Note that for any configuration F , $\pi_{B_r^-(v)}(F) \geq e^{-O(|V_G(B_r(v))|)}$, and thus this implies $O(N \log N)$ mixing for a free treelike ball with N vertices.

For the case of Glauber dynamics on a wired treelike ball, we use the result of Blanca and Gheisari [BG23a] bounding the spectral gap for Glauber dynamics on a wired tree using its edge-cut-width. We note that while their lemma was stated for trees with bounded degree, that restriction was not needed to get the following bound.

Lemma 57 ([BG23a, Lemma 5.10]). *For any tree T with N vertices and edges $E(T) = \{e_1, \dots, e_{N-1}\}$, define its edge-cut-width to be*

$$\text{CW}(T) = \min_{\sigma} \max_{1 \leq i < N} |V(\{e_{\sigma(j)} : j \leq i\} \cap V(\{e_{\sigma(j)} : j > i\}))|,$$

where the maximum is over all permutations on $\{1, \dots, N-1\}$.

Then for any reals $q \geq 2$ and $\beta > 0$, there is a constant $\theta(q, \beta) > 0$ such that the inverse spectral gap of the Glauber dynamics on T with a wired boundary (all leaves of T belonging to the same component) is $\theta N^2 \exp\{2(\text{CW}(T) + 1) \log q\}$.

We now observe that the edge-cut-width of a tree is at most its height.

Lemma 58. *Let T be a height- r tree. Then $\text{CW}(T) \leq r$.*

Proof. TOPROVE 39 □

To obtain the mixing time for a treelike ball from the mixing on a tree, we use that adding an edge to the graph changes the probabilities by at most a constant factor, and therefore the spectral gap is only changed by a constant factor.

Lemma 59. *For any real $A > 0$, for all d large enough, and then for all q large enough reals, whp over $G \sim \mathcal{G}(n, d/n)$ the following holds. For all integer $r \leq \frac{A \log n}{d}$ and all $v \in V_C$, let $B := B_r(v)$.*

Then the mixing time of the Glauber dynamics on B with wired boundary condition is $O(n^{4\frac{A}{d} \log q})$.

Proof. TOPROVE 40 □

E.2 Proof of Theorems 25 and 26

Here we outline the proof of Theorem 26, the proof of Theorem 25 is analogous. The proofs follow closely those in [GGS24a, GS22, GGS24b], thus we only outline the key ideas.

Before we start, let A be such that for all d and q large enough reals, whp G has WSM within the ordered phase at the radius $r = \lceil \frac{A \log n}{d} \rceil$ with rate $q^{-1/30}$ (Lemma 6).

Let X_t be Glauber dynamics initialized from all-in. We focus on proving that for $T = \text{poly}(n)$ steps, $\|X_T - \pi_C^{\text{ord}}\|_{\text{TV}} \leq \frac{1}{4}$, as then any accuracy $\varepsilon \geq e^{-\Omega(n)}$ can be achieved by probability amplification (for details see [GGS24a, Theorem 3.4]).

To bound the total variation distance, we couple X_t with another copy of the Glauber dynamics \hat{X}_T , which starts from π_C^{ord} , and is restricted to the ordered phase – if it was to make an update outside the phase, it ignores the update instead. We couple the chains by having them choose the same edge to update, and both make the update based on the same $\text{Uniform}(0, 1)$ random variable. One can easily verify that π_C^{ord} is the stationary distribution for the restricted Glauber dynamics, and thus for all $t \geq 0$, $\hat{X}_t \sim \pi_C^{\text{ord}}$.

To bound $\|X_T - \pi_C^{\text{ord}}\|_{\text{TV}} \leq \mathbb{P}(X_T \neq \hat{X}_T)$, we bound the probability of disagreement on each edge separately, and then use union bound. Thus fix an edge e and a vertex v incident to e .

We introduce another coupled copy (using the same coupling as X_t and \hat{X}_t) of *wired* Glauber dynamics X_t^v starting from all-in restricted to the radius- $r = \lceil \frac{A \log n}{d} \rceil$ ball around v , $B_r(v)$, i.e. X_t^v acts as if every vertex distance- r from v is in a single component, and ignores every update that would change an edge outside this ball. Note that the stationary distribution of this chain is exactly $\pi_{B_r^+(v)}$.

Using standard monotonicity arguments (see e.g. [GGS24b, Appendix B]), we can see that conditional on \hat{X}_t having not ignored any update by time t (call this event $\mathcal{E}_{\leq t}$), $\hat{X}_t \subseteq X_t \subseteq X_t^v$. This gives us, that conditional on $\mathcal{E}_{\leq t}$, whenever X_t and \hat{X}_t disagree on e , then also \hat{X}_t and X_t^v disagree on e . With some arithmetic (for details see e.g. the proof of Theorem 1 in [GGS24b]), we obtain that

$$\mathbb{P}(\mathbb{1}\{e \in X_T\} \neq \mathbb{1}\{e \in \hat{X}_T\}) \leq 5\mathbb{P}(\mathcal{E}_{\leq T}) + \|\pi_C^{\text{ord}}(1_e) - \pi_{B_r^+(v)}\|_{\text{TV}} + \|X_t^v - \pi_{B_r^+(v)}\|_{\text{TV}}.$$

The first term can be bound using the Theorem 4(ii) – noting that the probability that an update would be ignored is at most the probability that the current state has $\geq \frac{9\eta}{10}|E_C|$ out-edges, which is whp $\leq e^{-n}$ (for all d and q large enough). Thus the first term is $e^{-\Omega(n)}$. The second term is exactly the WSM within the ordered phase, and provided that q is sufficiently large, it is at most $\frac{1}{n^2}$.

The second term is the TV-distance between the wired Glauber on the ball after T steps and its stationary distribution. By Lemma 59, after $T_v = O(n^{5A \log q/d})$ updates within the ball $B_r(v)$, $\|X_t^v - \pi_{B_r^+(v)}\|_{\text{TV}} \leq \frac{1}{n^3}$ (note that $\log(n^3) = O(n^{A \log q/d})$). Using standard Chernoff bounds, if we take $T \geq 30T_v \frac{|E_C|}{|E_G(B_r(v))|}$, the probability that less than T_v updates were made in $B_r(v)$ within T steps is $\leq e^{-\Omega(\log^3 n)} = O(1/n^3)$.

Summing up, this gives that for all n large enough, $\mathbb{P}(\mathbb{1}\{e \in X_T\} \neq \mathbb{1}\{e \in X_T\}) \leq \frac{1}{4|E_C|}$, and thus the theorem follows from union bound.