Rapid mixing beyond worst-case initialisation



Paulina Smolarova

Trinity College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2026

To KSP, for showing me the joy of problem solving.

Abstract

Markov Chain Monte Carlo (MCMC) is a standard algorithmic technique for the design of efficient approximate sampling algorithms. In particular, for spin systems, where the aim is to sample an assignment of finitely many states ("spins") to vertices of an input graph, Glauber dynamics is a standard local Markov chain that in each step resamples the spin of a randomly chosen vertex. Although Glauber dynamics yields simple sampling algorithms, its convergence time (mixing time) is generally not well understood. In regimes where the model exhibits multiple distinct types of typical configurations (phases), Glauber dynamics is known to perform poorly due to the exponential time needed to move between the phases.

While most techniques for bounding mixing times only consider worst-case initialisations, for the purposes of algorithmic design, we may initialise the chain more judiciously. Indeed, to obtain an efficient approximate sampler, it suffices to prove that a Markov chain mixes in polynomial time from a suitable, efficiently computable initial distribution.

As our case study, we consider the random-cluster model with real parameters $q, \beta > 0$. This model is an alternative representation of the ferromagnetic Potts model for integer q, a fundamental spin system originating in statistical physics. Approximate sampling from the Potts model is computationally hard, and remains hard even on bounded-degree graphs in the low-temperature regime (corresponding to large values of the *inverse-temperature* parameter β). Nevertheless, it is conjectured that an efficient algorithm exists for all $\beta > 0$ working with high probability on random regular graphs. In this setting, the behaviour of the model is relatively well understood: there is a critical inverse temperature β_c at which the model undergoes a disorder-order phase transition. The disordered phase is typical at high temperatures ($\beta < \beta_c$), the ordered phase is typical at low temperatures ($\beta > \beta_c$), and both phases are typical at $\beta = \beta_c$. Helmuth, Jenssen and Perkins used the phase structure to obtain an efficient cluster-expansion algorithm for large q and all $\beta > 0$. At the same

time, the presence of the two phases poses a problem for MCMC methods, especially in the window around β_c, where Glauber dynamics experiences exponential slowdown.

We build on recent work by Gheissari and Sinclair, who showed that the slowdown of Glauber dynamics can be avoided by a well-chosen initialisation, and extend their approach to the random-cluster model and, more generally, to monotone models. The idea is to initialise Glauber dynamics in the "extreme state" representing the dominant phase (or a random mixture thereof). We show that, subject to low escape probability from a phase, local mixing, and a weak spatial mixing condition, the Glauber dynamics mixes in polynomial time. This approach also applies to graphs of unbounded degree, including sparse Erdős-Rényi graphs, where existing low-temperature sampling techniques do not apply due to the presence of regions with low expansion.

As the main contribution of this thesis, we apply the method to the random-cluster Glauber dynamics on random graphs. On random Δ-regular graphs, we obtain fast mixing throughout the high-temperature regime ($\beta < \beta_c$) for all integer q, and in the low-temperature regime ($\beta > \beta_c$) for integer $q \geq (5\Delta)^5$. As a corollary, we resolve approximate sampling from the ferromagnetic Potts model on random regular graphs in the high-temperature regime. For sufficiently large real q, our results also extend to all $\beta > 0$, matching the parameter range of the prior cluster-expansion algorithm.

We further apply our approach to sparse Erdős-Rényi graphs $\mathcal{G}(n, d/n)$. To overcome the limitations posed by irregular degrees, we develop a new, simple polymer model for the ordered phase, allowing us to quantify how far a typical sample is from the extreme state representing the phase. This enables us to verify the three conditions and thus obtain fast mixing for all sufficiently large d and then all sufficiently large q at all temperatures. Notably, this yields the first known sampling algorithm on sparse Erdős–Rényi graphs that applies in the low-temperature regime to a problem for which approximate sampling is computationally hard on general graphs.

Acknowledgements

First and foremost, I would like to thank my supervisors, Leslie Goldberg and Andreas Galanis, for their invaluable time, guidance, and expertise. I would also like to extend my thanks to Edith Elkind, Sandra Kiefer, Elias Koutsoupias, and Standa Živný for serving as examiners on my various DPhil examinations, and, in advance, to my DPhil viva examiners for taking the time to review this thesis.

I am also very grateful to everyone who supported me on this journey, in my academic as well as personal life. Special thanks go to Marc for many useful career discussions. I would also like to thank OUSFG for providing a great community during my time in Oxford. On a more personal note, I appreciate everyone who gave me emotional support during this marathon, including but not limited to Zuzka and Darnellia.

Last but not least, I need to mention people without whom I would not have started this journey in the first place. My parents for believing in me and KSP for showing me this wonderful world of computer science. I would not be here without you.

Declaration

I declare that the work presented in this thesis is my own, except where otherwise stated. Parts of this work have been published in peer-reviewed academic journals and conference proceedings. Some parts are available as preprints and are currently under review. All publications have been coauthored with my supervisors, Andreas Galanis and Leslie Ann Goldberg.

Chapters 3 and 4 contain work from the following paper.

[38] 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*, 34(3):359–391, 2025. doi: 10.1017/S0963548324000403

Chapters 3 and 5 contain work from the following paper.

[41] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model. In *43rd International Symposium on Theoretical Aspects of Computer Science (STACS 2026)*, volume 364 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:19, 2026. doi: 10.4230/LIPIcs.STACS.2026.39

The thesis also incorporates material from the corresponding full version, available on arXiv.

[42] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model, 2026. URL `https://arxiv.org/abs/2410.14409`

Chapters 3 and 6 contain work from the following paper.

[40] Andreas Galanis, Leslie Goldberg, and Paulina Smolarova. Low-temperature sampling on sparse random graphs. *Electronic Journal of Probability*, 31:1–40, 01 2026. doi: 10.1214/26-EJP1576

Contents

1 Introduction **1**

1.1 Approximate counting and sampling 1

1.1.1 Hardness of approximate sampling: #BIS and #SAT 2

1.1.2 The Potts model 3

1.1.3 #BIS-hard problems on bounded-degree graphs 4

1.1.4 The Potts model on random regular graphs 6

1.1.4.1 Alternative representation: random-cluster model . . 8

1.2 Our results 11

1.2.1 Bounded-degree graphs 12

1.2.1.1 Mixing of the RC dynamics at all temperatures . . . 12

1.2.1.2 Almost all parameter ranges for the Potts model using planting 14

1.2.2 Beyond bounded-degree graphs 15

1.2.2.1 Fast mixing beyond bounded degree 17

1.3 Proof overview: fast mixing from a ground state 19

1.3.1 Local mixing 24

1.3.2 Weak spatial mixing within the phase 26

1.3.2.1 WSM within the ordered phase on random regular graphs via polymers: proof outline 29

1.3.2.2 Wired boundary using planting: proof outline 31

1.3.2.3 Wired boundary via ordered polymers beyond bounded degree: proof outline 34

1.3.2.4 WSM within the disordered phase 37

1.3.2.5 From WSM to approximation: proof outline of Theorem 1.4 38

2 Preliminaries **40**
2.1 Notation and graph definitions 40
2.2 Probability and algorithmic preliminaries 40
2.2.1 Concentration bounds 42
2.3 Potts model and random-cluster model 43
2.4 Markov chains and mixing times 46

3 Proof of the main strategy **49**
3.1 Mixing within the phase 49
3.1.1 Probability amplification within the phase 50
3.1.2 Mixing within the ordered phase 51
3.1.3 Mixing within the disordered phase and extension to other models 54
3.2 FPRAS for the partition function 55

4 Sampling from the random-cluster model on random regular graphs via Glauber dynamics **63**
4.1 Overview 63
4.1.1 Notation 64
4.1.2 Locally treelike expanders 65
4.1.3 Mixing within the phase 66
4.1.3.1 Local maximum (Key Property 1) 67
4.1.3.2 WSM within the phase (Key Property 2) 67
4.1.3.3 Local mixing (Key Property 3) 68
4.1.4 Partial configurations 69
4.1.5 Boundary conditions 71
4.1.6 Monotonicity 72
4.2 Local maxima and phases of the RC model 73
4.3 WSM within the ordered phase 74
4.3.1 Ordered polymers 76
4.3.2 The coupling process 81
4.3.3 Proof of WSM within the ordered phase 92
4.4 WSM within the disordered phase 93
4.4.1 Disordered polymers 94
4.4.2 Proof of WSM within the disordered phase 95
4.5 Local mixing for the wired RC dynamics 98
4.5.1 Log-Sobolev constant for wired treelike neighbourhoods . . . 98
4.5.2 Mixing on treelike graphs for the wired RC dynamics 100

4.6 Proof of the main theorem 106

5 Sampling and planting of the Potts model on random regular graphs 109

5.1 Overview 109
5.1.1 Notation 110
5.1.1.1 The configuration model 111
5.1.2 Planting 111
5.1.2.1 Phases of the Potts model 111
5.1.2.2 Planting disordered and ordered configurations . . . 113
5.1.2.3 Phases of the random-cluster model via the Potts model 115
5.1.3 Mixing within the phase 117
5.1.4 Weak spatial mixing within a phase 118
5.2 Planting for the Potts and RC models 119
5.2.1 Proof of Lemma 5.2 120
5.2.2 Using planting for the random-cluster model 121
5.3 Proof for WSM within the disordered phase 123
5.4 Proof for WSM within the ordered phase 130
5.4.1 Wired boundary in ordered configurations 131
5.4.2 The giant component in the ordered phase 140
5.4.2.1 Percolation in regular and near-regular graphs 141
5.4.2.2 Branching-process estimates 142
5.4.2.3 Size of the giant component in percolated almost-regular graphs 143
5.4.3 Avoiding the largest component 151
5.4.4 Proof of the WSM within the ordered phase 159
5.5 Proof of the main theorem 160

6 Random-cluster model on sparse random graphs of unbounded degree 163

6.1 Overview 163
6.1.1 Setup and notation 165
6.1.2 Phases of the random-cluster model on $\mathcal{G}(n, d/n)$ 166
6.1.3 WSM within the phase 167
6.1.3.1 Ordered polymers 168
6.2 Graph properties 170
6.2.1 Treelikeness 171
6.2.2 Average degree of connected sets 173

6.2.3 Weak expansion 176
6.2.4 Kernel of $\mathcal{G}(n, d/n)$ 178
6.2.4.1 Expansion of the kernel 181
6.2.4.2 Giant component after edge removal 183
6.3 Phases of the random-cluster model 185
6.4 Weak spatial mixing within the ordered phase 192
6.4.1 Ordered Polymers 193
6.4.2 Proof of WSM within the ordered phase 202
6.5 Weak spatial mixing within the disordered phase 203
6.6 Proof of Theorem 1.3 208
6.6.1 Local mixing on treelike balls 208
6.6.2 Proof of Theorems 6.3 and 6.4 210
6.7 Proof of Theorem 1.4 211

7 Conclusion and open questions **219**
7.1 Bounded-degree graphs 220
7.1.1 All parameter ranges on random regular graphs 220
7.1.2 Fast mixing on all bounded-degree graphs 222
7.2 Sparse graphs of unbounded degree 222
7.2.1 Algorithms for $\mathcal{G}(n, d/n)$ for all parameter ranges 223
7.2.2 Critical threshold for $\mathcal{G}(n, d/n)$ 224
7.3 Extension to other monotone models 225

Bibliography **226**

List of Figures

1.1 Examples of disordered and ordered Potts configurations 5
1.2 Examples of disordered and ordered random-cluster configurations . . 10
1.3 Monotone coupling in the disordered phase 22
1.4 Partially revealed configurations 28
1.5 One step of the bottom-up revealing scheme in a tree 30
1.6 An ordered polymer (regular-graph setting) 31
1.7 An example of the set $S_v(F^{\mathrm{ord}})$ 32
1.8 An issue with the ordered polymer model from Definition 1.6 on irregular-degree graphs 35
1.9 Ordered polymers for $\mathcal{G}(n, d/n)$ 36

4.1 Excess edges in a treelike ball 82
4.2 Illustration of the quantities from Observation 4.33 85

5.1 Construction of the auxiliary graph $\mathbf{H}$ 147

6.1 Construction of new ordered polymers 194

List of Notation

Chapter 1

$\|\cdot - \cdot\|\|_{\mathrm{TV}}$	Total-variation (TV) distance (see (1.1))	2
$\Omega_{G,q}$	Set of vertex q-colourings of G	3
$w(\sigma, \beta)$	Weight of a Potts configuration σ	3
$Z_G(q, \beta)$	Partition function of the Potts and RC models	3
$\mu_G(q, \beta)$	Potts Gibbs distribution	3
β_c	Critical temperature for the Potts and RC models	4
β_u	Uniqueness threshold for the Potts and RC models	5
Ω_G	Edge subsets of G	8
$\pi_G(F; q, \beta)$	RC Gibbs distribution	8
$c(F)$	Number of components in (V, F)	9
$w_G(F; q, \beta)$	Weight of an RC configuration	9
β'_u	Second uniqueness threshold for the RC model	11
$\mathcal{G}(n, d/n)$	Erdős-Rényi graph with edge probability d/n	15
Ω_G^P, π_G^P	Configurations and Gibbs distribution in the phase P.	19
g_P	Ground state of the phase P	19
$B_r(v)$	Radius-r neighbourhood of v	22
$\pi^{g_P}_{B_r(v)}$	RC distribution on $B_r(v)$ where the outside is conditional to be in the ground state g_P	22
$\pi^{+}_{B_r(v)}$	RC distribution on the wired ball	26
$(\boldsymbol{F}^{\mathrm{ord}}, \boldsymbol{F}^{+})$	Coupling of the ordered and the wired RC distributions	26
$(\mathcal{F}_i^{\mathrm{ord}}, \mathcal{F}_i^{+})$	Pair of partially revealed RC configurations at step i of a coupling in (1.4)	27
$Z_G^P(q, \beta)$	Partition function of the phase $P \in \{\mathrm{ord}, \mathrm{dis}\}$	38

Chapter 2

$\mathrm{Bin}(n, p)$	Binomial random variable with parameters n and p	43
p	$1 - \mathrm{e}^{-\beta}$	46
$\hat{p}$	$\frac{p}{(1-p)q+p}$	46
T_{mix}	Mixing time	47

Chapter 4

$\mathcal{G}_{n,\Delta}$	Set of Δ-regular graphs on n vertices	65

$G \sim \mathcal{G}_{n,\Delta}$	A random Δ-regular graph on n vertices	65
$\mathcal{F} : E \to \{0,1\}$	RC configuration	65
$\mathrm{In}(\mathcal{F}), \mathrm{Out}(\mathcal{F})$	In-edges and out-edges of $\mathcal{F}$	65
$G[F], F \subseteq E$	Subgraph (V, F)	65
ϕ_G	Expansion profile of G	65
$\mathcal{G}_{\Delta,\delta}$	Class of Δ-regular expanders	66
$\mathcal{G}_{\Delta,\delta,K}$	Class of K-treelike Δ-regular expanders	66
$\Omega^{\mathrm{ord}}, \Omega^{\mathrm{dis}}$	Sets of disordered/ordered configurations	66
$\mathcal{A} : E \to \{0,1,*\}$	Partial configuration	70
$R(\mathcal{A})$	Revealed edges of $\mathcal{A}$	70
$\Omega*$	Set of all partial configurations	70
$\Omega_{\mathcal{A}}, \Omega_F, \Omega_{\mathcal{A},F}$	Sets of partial configurations with respect to $\mathcal{A}$ and F (see Definition 4.16)	70
$\pi_{\mathcal{A}}, \pi_F, \pi_{\mathcal{A},F}$	Marginal distributions with respect to $\mathcal{A}$ and F (see Definition 4.16)	70
$\xi(\mathcal{A})$	Component structure of the boundary of $\mathcal{A}$	71
$\mathcal{B}_\infty(\mathcal{F})$	Polymer edges of a configuration $\mathcal{F}$	76
$E_u(\gamma)$	Out-edges of a polymer γ	76
$c'(\gamma)$	Number of components in $G[E \setminus E_u(\gamma)]$	76
w_γ^{ord}	Weight of an ordered polymer γ	76
$\partial_G F$	Boundary vertices of revealed edges F	81
$\mathcal{E}$	Excess edges (Definition 4.29)	81
w_γ^{dis}	Weight of a disordered polymer γ	94
$\Omega^{\mathrm{dis}}(\mathcal{F})$	Disordered polymers of $\mathcal{F}$	94

Chapter 5

$\mu_{G,q,\beta}$	Potts Gibbs distribution	110
$\pi_{G,q,p}$	RC Gibbs distribution	110
$G \sim \mathcal{G}_{n,d}$	Random d-regular graph on n vertices	111
$G \sim \mathbb{G}_{n,d}, \boldsymbol{G}$	A d-regular (multi)graph on n vertices chosen uniformly from the configuration model	111
$S_\mu = S_\mu(\theta)$	Potts configurations with colour statistics around μ	112
ν_{d}	Disordered vertex colour statistics	112
ν_{o}	Ordered vertex colour statistics	112
$S_{\mathrm{d}}, S_{\mathrm{o}}$	Colourings with phase colour statistics	113
$Z_{\mathrm{dis}}, Z_{\mathrm{ord}}$	Partition functions for phase colour statistics	113
$(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}), \hat{\boldsymbol{\sigma}}_{\mathrm{d}})$	Planted disordered distribution	114
$(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}})$	Planted ordered distribution	114
$\rho_{\mathrm{d}}, \rho_{\mathrm{o}}$	Disordered/ordered edge statistics	115
$\boldsymbol{F}(G, \sigma)$	RC configuration obtained from σ after the percolation step	115
$\mathfrak{m}_{\mathrm{d}}(\beta), \mathfrak{m}_{\mathrm{o}}(\beta)$	Disordered/ordered RC statistics	116
$\Omega_{G,\mathrm{d}}, \Omega_{G,\mathrm{o}}$	Disordered/ordered phases	116
$\pi_{G,\mathrm{d}}, \pi_{G,\mathrm{o}}$	RC Gibbs distribution in the phase	116

$\pi^{-}_{G,v,\ell}, \pi^{+}_{G,v,\ell}$	RC distribution on a free/wired radius-ℓ ball around v	118
1_e	Subsets of edges including the edge e	118
$\mathcal{C}_1(G)$	Largest component of G	140
$\varphi, \hat{\varphi}, \chi$	Extinction probabilities and giant component density	140
$\tilde{\boldsymbol{G}}_{n,\mathbf{d},m}$	Exact edge model	141

Chapter 6

$C = (V_C, E_C)$	Largest component of $G \sim \mathcal{G}(n, d/n)$	165
n_C	$\lvert V_C \rvert$	165
μ	Conjugate of d	165
$\Omega_C^{\mathrm{ord}}, \Omega_C^{\mathrm{dis}}$	Set of configurations in the ordered and the disordered phase on C	166
$\pi_C^{\mathrm{ord}}, \pi_C^{\mathrm{dis}}$	RC distribution in the ordered and the disordered phase on C	166
(β_0, β_1)	Critical temperature window	166
$E_G(S,T), e_G(S,T)$	Edge set and edge count between S and T in G	170
$e_G(S)$	Inner edges of S in S, i.e. $e_G(S,S)$	170
$G_K = (V_K, E_K)$	Kernel of $G \sim \mathcal{G}(n, d/n)$	178
n_K	$\lvert V_K \rvert$	178
$\mathcal{V}(F, v)$	Inner vertices of polymers of F with respect to v	193
Γ_v^{ord}	Ordered polymers of F with respect to a vertex v	194
$\mathcal{V}_\gamma, w_C^{\mathrm{ord}}(\gamma)$	Inner vertices and weight of a polymer γ	194
$\Gamma^{\mathrm{dis}}(F)$	Set of disordered polymers of F	205
$w_C^{\mathrm{dis}}(\gamma)$	Weight of a disordered polymer γ	205

Chapter 1

Introduction

1.1 Approximate counting and sampling

Counting problems are a natural extension of decision problems. For an NP machine, instead of deciding whether an accepting path exists, we ask for the *number of accepting paths*. This class of problems is known as #P. More generally, we consider counting *weighted combinatorial objects*, such as weighted independent sets (also known as the *hardcore model*), or weighted colourings.[1]

Any counting problem is at least as hard as its decision counterpart, and in fact, for many tractable decision problems, such as DNF or finding a perfect matching, the counting analogue is known to be #P-complete [69]. In fact, #P is believed to be strictly harder than decision, as evidenced by a classic result by Toda [67] which states that the polynomial hierarchy is Turing-reducible to #P.

In light of the hardness of exact counting, we consider a relaxed variant of the problem: *approximate counting*. We say that $\hat{A}$ is a $(1 \pm \varepsilon)$-*approximation* to A if $|\hat{A} - A| \leq \varepsilon |A|$. An approximation problem $\mathcal{P}$ is said to be tractable if it admits a *fully polynomial approximation randomised scheme (FPRAS)*. An FPRAS is a randomised algorithm that takes an input x to $\mathcal{P}$, together with an accuracy parameter $\varepsilon > 0$, and returns, in time $\text{poly}(|x|, 1/\varepsilon)$, with probability $\frac{3}{4}$, an $(1 \pm \varepsilon)$-approximation to $\mathcal{P}(x)$.[2]

Closely related to approximate counting is *approximate sampling*. There, we consider a set of weighted combinatorial objects and the associated distribution where the probability of each object is directly proportional to its weight. The problem of

[1] One can observe that if the objects can be represented by accepting paths, and the weights are rational and efficiently computable, this definition is in fact equivalent to #P up to a multiplicative factor.

[2] In fact, one can achieve success probability $1 - \delta$ for any $\delta > 0$ by running the algorithm $O(\log(1/\delta))$ times independently and taking the median of the answers.

approximate sampling from a distribution π_x, parametrised by its input x, is said to be tractable if it admits a *fully polynomial almost uniform sampler (FPAUS)*, which, informally, is a polynomial-time algorithm whose output distribution can get arbitrarily "close" to the target distribution π_x.

In order to formally define an FPAUS, we first introduce the measure of closeness that we use, *the total-variation (TV) distance*. For two distributions, μ_1, μ_2 with the sample space Ω, their TV-distance is as follows.[3]

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} \coloneqq \tfrac{1}{2} \sum_{\sigma \in \Omega} |\mu_1(\sigma) - \mu_2(\sigma)|. \tag{1.1}$$

Another useful interpretation of the total-variation distance is that two distributions μ_1, μ_2 are within TV distance ε if and only if there is a coupling[4] $(\boldsymbol{\sigma_1}, \boldsymbol{\sigma_2})$ of μ_1 and μ_2 such that $\mathbb{P}(\boldsymbol{\sigma}_1 \neq \boldsymbol{\sigma}_2) \leq \varepsilon$.

Then, an FPAUS for sampling from a distribution μ_x is a randomised algorithm which takes an input x together with an accuracy parameter $\varepsilon \in (0, 1)$, runs in time $\mathrm{poly}(|x|, \frac{1}{\varepsilon})$, and has output distribution within TV distance ε of μ_x.

Approximate counting and sampling are closely linked, and for many canonical problems there are standard techniques for converting an FPAUS to an FPRAS and vice versa; see, e.g., [55].

One common technique for obtaining an FPAUS is *Markov Chain Monte Carlo*, which is based on simulating a Markov chain whose stationary distribution is (close to) the target distribution. The challenge is typically proving that the number of steps until convergence (also known as the *mixing time*) is polynomial (also known as *rapid mixing*). While most existing results concern bounding mixing times from the worst-case initialisations, for algorithm design, it suffices to establish fast mixing from one efficiently computable initial distribution. In this work, we study the performance of MCMC methods for approximate sampling from suitable initialisations and show how we can obtain fast MCMC samplers even in the regimes where worst-case convergence is slow.

1.1.1 Hardness of approximate sampling: #BIS and #SAT

The hardest approximation problems in #P are those equivalent to #SAT: the problem of approximately counting the number of satisfying assignments to a given

[3] For brevity, we also use this notation with random variables instead of their distributions.

[4] A coupling of μ_1 and μ_2 is a joint distribution on a pair of random variables whose respective marginal distributions are μ_1 and μ_2.

boolean formula. Many fundamental problems belong to this class: the counting analogues of NP-hard problems, and even problems for which decision is tractable, such as counting independent sets in general graphs, and *maximal* independent sets in bipartite graphs [46, 27].

In addition to #SAT-hard and fully tractable problems, Dyer, Goldberg, Greenhill and Jerrum [27] identified a third class of problems, represented by #BIS: the problem of approximately counting independent sets in *bipartite graphs*. #BIS is conjectured to be of intermediate complexity, in the sense that it is believed neither to admit an FPRAS nor to be #SAT-hard under approximation-preserving (AP) reductions [27, 33].. Further evidence for this intermediate status is that #BIS is complete under AP-reductions for the logically defined subclass $\#\mathrm{RH}\Pi_1$ of #P [27]; see also [66].

Many fundamental problems are known to be #BIS-equivalent or #BIS-hard. Dyer et al. list several prominent examples, including approximating the number of downsets, path-colourings and certain restrictions on #SAT. We also consider the weighted variant of #BIS, the *bipartite hardcore model* with fugacity $\lambda > 0$. Given a *bipartite graph*, each independent set I is assigned weight $\lambda^{|I|}$, and its probability is proportional to its weight.

1.1.2 The Potts model

The focus of this work is another canonical #BIS-hard problem: the ferromagnetic[5] *Potts model*. In this problem, which originates in statistical physics, we are interested in weighted (non-proper) *vertex q-colourings* of a graph $G = (V, E)$. For an integer number of colours $q \geq 3$ and a real *inverse-temperature* parameter $\beta > 0$, let $\Omega_{G,q} = \{\sigma \mid \sigma : V \to [q]\}$ be the set of all q-colourings, or *(Potts) configurations*, of G. The *weight* of $\sigma \in \Omega_{G,q}$ is $w_G(\sigma; q, \beta) \coloneqq \mathrm{e}^{\beta m(\sigma)}$, where $m(\sigma)$ is the number of *monochromatic edges* under σ. In the counting variant of the Potts model, the aim is to approximate the sum of these weights, $Z_G(q, \beta) \coloneqq \sum_{\sigma \in \Omega_{G,q}} w_G(\sigma; q, \beta)$. Following the standard terminology from statistical physics, $Z_G(q, \beta)$ is called the *partition function*. The associated target distribution on q-colourings is known as the *(Potts) Gibbs distribution*, $\mu_G(\,\cdot\,; q, \beta)$, and is given by $\mu_G(\sigma; q, \beta) \coloneqq \frac{w_G(\sigma;q,\beta)}{Z_G(q,\beta)}$.

The Potts model is an example of a *spin system*: a model where we can view the

[5]Throughout this thesis, whenever we refer to the Potts model, we mean the ferromagnetic Potts model, and we henceforth omit the qualifier "ferromagnetic".

configurations[6] as assignments of finitely many spins (colours for the Potts model) to vertices of an input graph and the weights only depend on the interactions on neighbouring vertices. Independent sets can be also viewed as a spin system, with vertices being assigned to be in or out of the independent set.

We remark that when $q = 2$ the ferromagnetic Potts model is known as the *(ferromagnetic) Ising model*. A classical result by Jerrum and Sinclair [54] gives an FPRAS for approximating its partition function. In contrast, for $q \geq 3$, on general graphs, approximating the partition function and approximately sampling from the Potts Gibbs distribution are #BIS-hard [45].

1.1.3 #BIS-hard problems on bounded-degree graphs

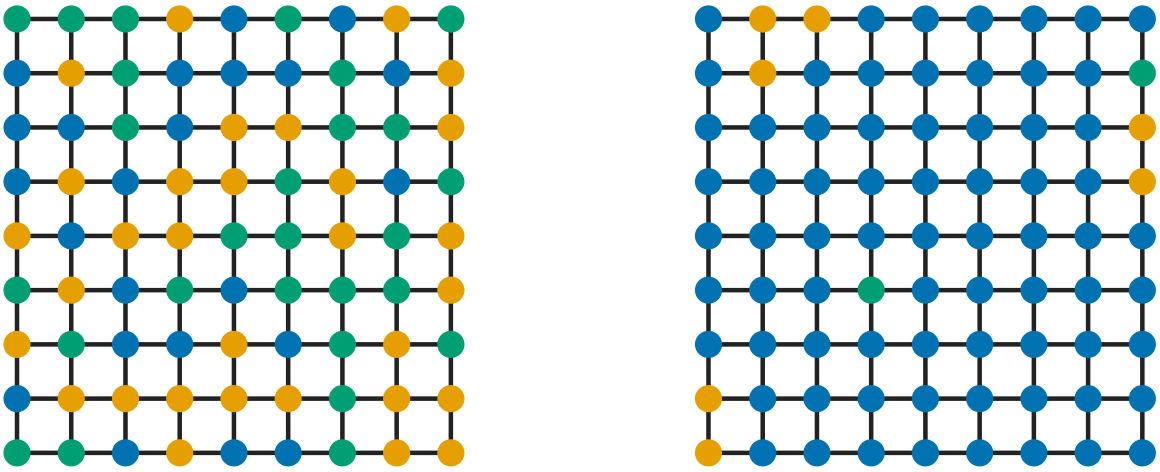
As #BIS-hard problems are not believed to be tractable, a major focus has been on identifying settings where they nevertheless admit an FPRAS (for approximate counting problems) or an FPAUS (for approximate sampling problems). For graph problems, a natural restriction is to bound the maximum degree. However, the hardness of #BIS-hard problems typically persists even on graphs of constant degree in the so-called *low-temperature regime*. In the terminology inherited from statistical physics, low temperature corresponds to stronger interactions between neighbouring vertices (high β) leading to low entropy and long-range interactions[7]. In contrast, at *high-temperatures* (low β), the problems are thought to become easier.

In fact, many models undergo a sharp *phase transition* between these two behaviours, as is well illustrated by the bipartite hardcore model. Recall that in a bipartite hardcore model, an independent set I is sampled with probability proportional to $\lambda^{|I|}$. The low-temperature regime (high λ) corresponds to the system favouring large independent sets whereas for the high-temperature regime (low λ) smaller independent sets are favoured. On bipartite graphs with maximum degree Δ, the transition between these two regimes occurs at the critical value $\lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta-2}}$. Approximate counting remains #BIS-hard in $\lambda > \lambda_c$, whereas for $\lambda \leq \lambda_c$, the model admits fast sampling and approximation algorithms [18, 20].

For the Potts model, the situation becomes more involved. We also see the emergence of a *critical threshold* $\beta_c := \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$, which is linked to the behaviour of the model on *random* Δ*-regular graphs* [34]. With probability $1 - o_n(1)$ over a

[6] We use the term *configuration* for an individual object from the space over which we count or sample.

[7] In the context of this work, a long-range interaction means a local state of a configuration is sensitive to vertices/edges far away in the input graph.



(a) Disordered configuration. (b) Ordered configuration.

Figure 1.1: Examples of disordered and ordered Potts configurations. The lattice is used only for visualisation. In the disordered configuration, the colours occur in roughly equal proportions, whereas in the ordered configuration one colour dominates.

graph chosen uniformly at random among all Δ-regular graphs on n vertices, there are two types of typical colourings, based on $\beta < \beta_c$ or $\beta > \beta_c$. When $\beta < \beta_c$, a typical colouring is *disordered* and colours appear with roughly equal proportions. In contrast, when $\beta > \beta_c$, the model strongly prefers colourings with many monochromatic edges, and as a result, a typical colouring has a majority of the vertices coloured with the same colour, with the remaining colours occurring with roughly the same frequency. These configurations are known to be *ordered*. At $\beta = \beta_c$ the model experiences a *phase coexistence*, and a typical configuration is either ordered or disordered [25]. See Figure 1.1 for illustration.

Galanis, Štefankovič, Vigoda, and Yang [34] use this typical behaviour of ordered configurations to establish a hardness reduction, utilising gadgets based on random regular graphs. In particular, they prove that the Potts model remains #BIS-hard on general bounded-degree graphs when $\beta > \beta_c$, and this hardness also extends to all bounded-degree regular graphs.

In addition to the critical threshold β_c, the Potts model has a second threshold $\beta_u(q, \Delta) < \beta_c$, known as the *uniqueness threshold*[8]. The uniqueness threshold captures the emergence of *long-range interactions*: whether the distribution of the colour of a vertex can be affected by spins arbitrarily far away. Formally, this threshold relates to the behaviour of the Gibbs distribution on $(\Delta - 1)$-ary tree as its height

[8]Only an implicit formula exists, in particular, $\beta_u = \ln\left(1 + \inf_{y>1} \frac{(y-1)(y^{\Delta-1}+q-1)}{(y^{\Delta-1}-y)}\right)$.

grows to infinity when the leaves are fixed to be the same colour. At $\beta < \beta_u$, the influence of the leaves on the root vanishes as the height grows, whereas for $\beta > \beta_u$, the root's colour stays biased towards the colour of the leaves. Since random Δ-regular graphs are locally treelike, this provides intuition for long-range interactions on finite graphs: above β_u, the colour of a vertex can remain correlated with a sufficiently biased configuration far away, whereas in the *uniqueness regime* $\beta < \beta_u$ this influence vanishes with distance [10, 25].

The absence of long-range interactions in the uniqueness regime suggests that approximately sampling and approximating the partition function of the Potts model become tractable for all graphs of maximum degree Δ when $\beta < \beta_u$ [10]. However, unlike the bipartite hardcore model, where algorithms are known for all $\lambda \leq \lambda_c$, this conjecture for the Potts model remains unresolved.

A recent result by Carlson, Davies, Fraiman, Kolla, Potukuchi and Yap [19] gave an approximation algorithm for the partition function in the uniqueness regime $\beta < \beta_u$ when Δ and q are sufficiently large. For all Δ and q, the best known positive sampling result for all graphs of maximum degree Δ applies when $\beta \leq \frac{(1-o_q(1))\log(q)}{\Delta} << \beta_u$. This algorithm is based on a simple Markov chain, known as the *(Potts) Glauber dynamics*, which maintains a q-colouring and updates the colour of one vertex in each step. However, the dynamics already experiences exponential slowdown (also known as *torpid mixing*) for $\beta > \frac{(1+o_q(1))\log q}{\Delta - 1 - 1/(\Delta-1)}$. Moreover, this slowdown occurs with high probability on a random Δ-regular graph [17].

The complexity also remains open for $\beta_u < \beta < \beta_c$. Even on typical random regular graphs, where we expect the typical configurations to be disordered in this regime, the long-range interactions cause the ordered colourings to form local maxima, which pose further bottlenecks for MCMC methods [25].

1.1.4 The Potts model on random regular graphs

Since random regular graphs are an important setting in the study of the complexity of the Potts model on general bounded-degree graphs, a natural question is whether they themselves constitute a tractable setting. That is, are there efficient approximation and sampling algorithms for the Potts model working with high probability[9] over a graph chosen uniformly at random from the set of all n-vertex Δ-regular graphs?

While the structure of typical configurations on random regular graphs is relatively well-understood (see e.g. [25, 34]), it long remained unclear how to obtain a

[9]That is, with probability $1 - o(1)$. We also abbreviate it as w.h.p..

fast algorithm for the low-temperature regime due to the presence of long-range interactions, which typically lead to torpid mixing of simple MCMC methods such as the Glauber dynamics. In particular, recall that a typical Potts configuration for $\beta > \beta_c$ has a majority of vertices assigned the same colour. Hence, the state space divides into q sets of configurations, also known as *phases*, where one particular colour dominates, with the remaining colourings having exponentially small probability. By symmetry, each of the q phases is equally likely [34].

For the Potts Glauber dynamics to move between two colourings with different majority colour, it has to pass through very unlikely states where neither colour has a strong majority, resulting in torpid mixing.

A major breakthrough for the low-temperature Potts model came with the development of the *cluster-expansion algorithm*, introduced by Helmuth, Perkins and Regts [48] for the bipartite hardcore and the Potts model on lattice families. It was later extended to random regular graphs and more generally bounded-degree expanders as well as other spin systems [52, 19]. The cluster-expansion algorithm itself is rather involved, and we do not describe its details here. For our purposes, it suffices to understand its underlying polymer representation and the condition requires for the method to apply. The central idea, built on the statistical-physics framework by Kotecký–Priess [56], is to consider each phase separately. The configurations in each phase are then represented in terms of connected deviations (known as *polymers*) from the "extreme state" of the given phase (known as the *ground state*). The polymers are then assigned weights reflecting how likely they are to occur.

We illustrate this idea on the low-temperature ($\beta > \beta_c$) Potts model. A phase in this case corresponds to the set of configurations with some majority colour $a \in [q]$. The corresponding *ground state* is the monochromatic configuration $\cdot \mapsto a$. Each colouring in the phase can then be represented by the collection of maximal connected regions with colours different from a, which we call *polymers*. Each polymer has an associated *weight* which decays exponentially with the number of bichromatic edges incident to this region, and roughly reflects the probability of that polymer appearing in a colouring. Thus, polymers encode the local deviations (vertices with non-majority colours) of a configuration from its ground state (the all-a monochromatic colouring).

We also mention another algorithmic technique using the polymer representation: the polymer dynamics Markov chain, introduced by Chen, Galanis, Goldberg, Perkins, Stewart and Vigoda [21] for the Potts and bipartite hardcore models, and later used for *any* spin system on regular bipartite graphs with sufficient expansion [35] and even the Potts model on expanders of unbounded degree [36].

In the polymer dynamics a colouring is represented as a collection of polymers (for a fixed phase), and the dynamics adds or removes a polymer in each step, allowing for larger updates than the Glauber dynamics, which only updates the colour of one vertex at a time. Compared to the cluster-expansion algorithms, polymer dynamics yields better, often near-linear running times [21].

A key requirement for both the cluster-expansion algorithm and the polymer dynamics to work is the *Kotecký–Preiss condition* [56]. This is a technical condition which we do not state formally here. Informally, it requires polymer weights to decay exponentially with their size; in our setting, this can be interpreted as saying that the probability of large connected deviations from the ground state decays exponentially with their size.

For the above polymer model for the Potts model, the Kotecký–Preiss condition is known to hold at sufficiently low temperatures on random regular graphs, and more generally on *expander graphs*. Informally, a graph is an expander if every sufficiently small vertex subset has a number of outgoing edges proportional to its size. As a result, in an expander, any maximal region coloured with the non-majority colours induces a number of bichromatic edges linear in the region's size, and thus, the weight of polymers decays exponentially in their size [21].

The cluster-expansion algorithm and the subsequent polymer dynamics represented major progress, yielding efficient and, in the case of polymer dynamics, often near-linear algorithms. However, these approaches remain technically involved. It is therefore still of interest whether there is an effective algorithm based on a simpler Markov chain.

Yet, as we already noted, the Potts Glauber dynamics experiences slow mixing on random regular graphs even in the range where sampling is conjectured to be tractable. Nevertheless, it turns out that one can use a variant of Glauber dynamics exploiting an alternative representation of the Potts model to overcome this slow-down and obtain fast mixing and thus efficient algorithms for a wider range of parameters.

1.1.4.1 Alternative representation: random-cluster model

To this end, we represent the Potts model via *edge subsets* $\Omega_G \coloneqq \{F \subseteq E\}$ of the input graph, which we refer to as the random-cluster (RC) configurations. From a q-colouring, we can obtain an edge subset by including each *monochromatic* edge independently with probability $1 - \mathrm{e}^{-\beta}$. This results in the *random-cluster model (RCM)*. The corresponding distribution is the *random-cluster (RC) Gibbs distribution*, denoted $\pi_G(\,\cdot\,; q, \beta)$. The probability of each subset $F \subseteq E$ is proportional to its

weight $w_G(F;q,\beta) \coloneqq q^{c(F)}(\mathrm{e}^\beta-1)^{|F|}$, where $c(F)$ is the number of components in the subgraph (V,F). From an RC configuration, one can obtain a Potts configuration by colouring each component independently uniformly with one of the q colours. This is known as the *Edwards-Sokal coupling*. Its proof, alongside the formal definition of the random-cluster model, is provided in Preliminaries 2.3. We refer to edges in the random-cluster configuration as the *in-edges* (and say these edges are *in*) and the edges not in the configuration as the *out-edges* (and say these edges are *out*).

We further generalise the model for all non-integer $q > 0$. Note that for $q = 1$ the corresponding distribution is precisely independent bond percolation[10]. Hence, another way to view the random-cluster model is as a percolation weighted by the number of resulting components.

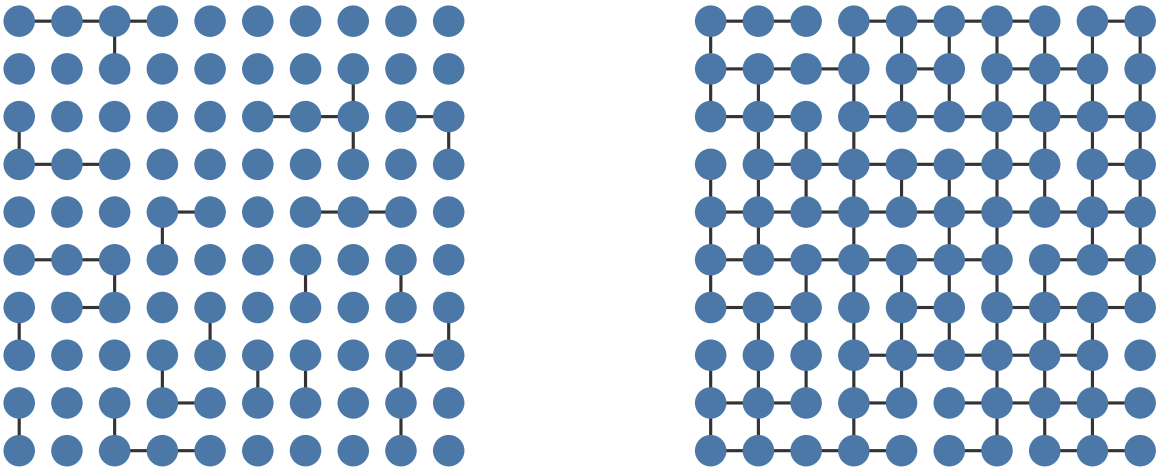
To sample from the random-cluster model (and thus from the Potts model), we consider a variant of the Glauber dynamics, which we also refer to as the *RC dynamics*. This is a Markov chain $(X_t)_{t\geq 0}$, where the states are the edge subsets Ω_G, and in each step, the state of a uniformly chosen edge e is resampled according to the conditional Gibbs distribution (for the full definition see Preliminaries 2.4). It is well-known that the RC dynamics converges to the RC Gibbs distribution (see e.g. [47, Chapter 8]). We remark that the probability of including the chosen edge e in the next step in fact only depends on whether e is a *cut-edge*[11] in the subgraph $(V, X_t \cup \{e\})$. Hence, to make an update it is enough to maintain the component structure of the graph (V, X_t), which can be done efficiently with $O(\log n(\log\log n)^2) = \widetilde{O}(1)$ update time (see e.g. [50]).

In recent years, the RC dynamics has been found to extend efficient MCMC sampling beyond the regimes accessible to the Potts Glauber dynamics. In the uniqueness regime $\beta < \beta_u$, Blanca and Gheissari [10] establish $O(n\log n)$ mixing time of the RC dynamics w.h.p. over random regular graphs for all real $q \geq 1$. The same authors also obtained $O(n^{1+o(1)})$ mixing of the RC dynamics on random regular graphs for the low-temperature regime (sufficiently high β) [12]. Note that in this regime, the Potts Glauber dynamics suffered from torpid mixing due to the exponentially many steps required to change the majority colour, which the RC representation overcomes by maintaining (a subset of) monochromatic edges instead of concrete colours.

However, the RC representation has its own phase structure, which creates a different bottleneck. While for the Potts model, the phases corresponded to majority

[10] Every edge is included in the subset independently, with probability $1-\mathrm{e}^{-\beta}$.

[11] An edge is a cut-edge if removing it from the graph increases the number of components.



(a) Disordered configuration. (b) Ordered configuration.

Figure 1.2: Examples of disordered and ordered random-cluster configurations. In the disordered configuration, the in-edges form small components, whereas in the ordered configuration they form a giant component. The lattice is used only for visualisation.

colours, for the random-cluster model the phases capture the *disordered* and *ordered* configurations.

Recall that a typical Potts colouring for $\beta < \beta_c$ has colours appearing with roughly equal frequencies. For the random-cluster model, this corresponds to configurations containing at most a small fraction of edges, which are all contained in small components (see Figure 1.2a). We call this the *disordered phase*, denoted dis.

On the other hand, for $\beta > \beta_c$, a typical Potts colouring has one colour assigned to a majority of vertices. For the random-cluster model, this corresponds to configurations containing all but at most a small fraction of edges. In turn, these edges then form a giant component[12] and a collection of small components (see Figure 1.1b). This is known as the *ordered phase*, denoted ord[13].

Analogously to the Potts model, at the critical temperature $\beta = \beta_c$, the probability of each phase is bounded away from zero. For all $\beta > 0$, the probability of a configuration outside of a phase is exponentially small [49, 25].

As a consequence, in order for the RC dynamics to cross between the phases, it needs to pass through these exponentially unlikely states, leading to a bottleneck and slow mixing at criticality [49].

[12]Throughout this work, a giant component is a component containing a majority of vertices. We refer to all other components as small components.

[13]We defer the formal definition of phases to Preliminaries 2.3, as the definition varies throughout the chapters.

The critical point is known to be $\beta_c = \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$ for all real $q > 2$ [8]. The stronger result establishing the bottleneck between the disordered and the ordered phase at criticality is known for integer $q \geq 3$ [25] and all real $q \geq \Delta^{\Omega(\Delta)}$ [49].

For integer q, the slowdown is further known to extend beyond criticality. Recall that for $\beta \in (\beta_u, \beta_c)$ the local state can be affected by the state arbitrarily far away. As a result, even though the typical configuration is disordered (all components are small), if the starting state is ordered (contains a giant component), it will take the RC dynamics exponentially many steps to disconnect it. A similar phenomenon happens for $\beta \in (\beta_c, \beta'_u)$[14]. In this regime, if the RC dynamics is initialised with few in-edges (disordered phase), it takes exponentially many steps for a giant component (typical for the ordered phase) to form. As a result, in the *metastability window* $\beta \in (\beta_u, \beta'_u)$ the RC dynamics is slowly mixing because of the time it needs to cross between the phases. We remark that while this behaviour has only been proven for integer q, it is believed to extend to all real $q \geq 3$ [25].

Despite this negative result, Helmuth, Janssen and Perkins found an FPAUS and an FPRAS that work at *all temperatures* β, provided that q is sufficiently large. They represent the configurations in each phase via the connected deviations (polymers) from the "extreme state" of the phase. For the definitions of these polymers, we refer the reader to Section 4.3.1. The algorithm is then based on the cluster-expansion techniques [49]. While efficient, it is highly involved, and therefore, it remains of interest to determine whether a simpler algorithm can be obtained using MCMC methods.

Even though the slow mixing in the metastability window suggests that the RC dynamics cannot yield an efficient algorithm for all temperatures, the slowdown arises from initialising the chain in the *wrong phase* and then having to cross through the bottleneck. However, in the context of algorithm design, we can choose the initial state. Thus, the central question of this work is the following:

Can we show that the (RC) Glauber dynamics mixes rapidly from an (efficiently computable) initial distribution X_0?

1.2 Our results

We answer this question affirmatively, using a remarkably simple initial distribution. The key is to initialise the RC dynamics in the *ground state* of a suitable phase:

[14] $\beta'_u = \beta'_u(\Delta, q)$ is the critical threshold for the existence of a root-to-leaf path of in-edges in a $(\Delta-1)$-ary tree as the height grows to infinity.

when the disordered phase dominates[15], we start the dynamics in $X_0 \coloneqq \emptyset$ (the *all-out* configuration), when the ordered phase dominates, we set $X_0 \coloneqq E$ (the *all-in* configuration), and when the phases coexist we use a suitable random mixture of all-in and all-out.

We first establish fast mixing of the Glauber dynamics with these initialisations on random regular graphs at all temperatures and large q, matching the regime of the cluster-expansion algorithm [49]. We then improve upon this result for the Potts model, completely resolving the high-temperature regime ($\beta < \beta_c$) and significantly improving the lower bound on q for the low-temperature regime ($\beta > \beta_c$). Finally, we also extend our work to sparse Erdős-Rényi graphs $\mathcal{G}(n, d/n)$, obtaining mixing for large d and q at all temperatures. We remark that this yields, to our knowledge, the first sampling and approximation algorithm for $\mathcal{G}(n, d/n)$ applicable in the *low-temperature regime* for a spin system that is #BIS-hard on general graphs.

In the remainder of this chapter, we formally state our results, and then conclude the introduction by providing a proof overview.

1.2.1 Bounded-degree graphs

1.2.1.1 Mixing of the RC dynamics at all temperatures

In our first result, which we present in Chapter 4, we establish fast mixing of the RC dynamics on random regular graphs with the above-described initialisation on the same parameter ranges as the cluster-expansion algorithm by Helmuth, Jenssen and Perkins from [49]: all temperatures β when q is sufficiently large. Recall that the critical temperature β_c where the dominant phase changes from disordered to ordered is $\beta_c = \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$ [8].

Theorem 1.1. *Let $\Delta \geq 5$ be an integer and let $q \geq \Delta^{\Omega(\Delta)}$ be a sufficiently large real. Then, with high probability over a random Δ-regular graph $G = (V, E)$, the following hold.*

- *For all $\beta < \beta_c$, the RC dynamics initialised from $X_0 \coloneqq \emptyset$ (all-out) mixes in $O(n \log n)$ time.*
- *For all $\beta > \beta_c$, the RC dynamics initialised from $X_0 = E$ (all-in) mixes in polynomial time. Moreover, if q is an integer, it mixes in $O(n \log n)$ time.*

[15] We say that a phase *dominates* when the Gibbs distribution assigns probability $1 - \mathrm{e}^{-\Omega(n)}$ to that phase.

- *For $\beta = \beta_c$, the RC dynamics initialised from a suitable random mixture of all-out and all-in mixes in polynomial time. Moreover, this random mixture can be computed in polynomial time.*

We note that the random mixture needed at the critical temperature can be efficiently approximated using the FPRAS from [49], or alternatively, by using a standard self-reducibility method, which we describe in Section 3.2. We also note that constants in the running time does depend on the parameters.

The proof idea is to consider each phase separately. We consider the *Gibbs distribution conditional on being in a phase*. That is, for the disordered phase we consider the distribution $\pi_G^{\text{dis}} = \pi_G(\,\cdot\,; q, \beta \mid \cdot \text{ is disordered})$ and analogously π_G^{ord} for the ordered phase (for the precise definition of the disordered and ordered phases in this setting, see Definition 4.4). Then we show that a typical sample within the phase does not get "too far" from the relevant initialisation (all-out for the disordered phase, and all-in for the ordered phase). We note that this is a rough intuition; for a more detailed proof outline see Section 1.3.

To capture what it means to be "far" from a ground state, we use the notion of *polymers*, as defined by Helmuth, Jenssen and Perkins [49]. Intuitively, these capture how far a disordered configuration is from the all-out ground state, and how far an ordered configuration is from the all-in ground state. For the disordered phase, the polymers are simply the connected components of in-edges, whereas for the ordered phase the definition is more technical and is given later in Definition 1.6.

We use the polymer framework somewhat differently from previous approaches. Whereas the cluster-expansion and polymer dynamics approaches are based on representing the configuration as a collection of polymers, we only use the polymers only as a measure of how far a configuration is from the ground state (the initial state of the RC dynamics). This leads to our variant of the Kotecký–Preiss condition. Recall that, informally, the Kotecký–Preiss condition requires the probability of any fixed polymer to decay exponentially with its size. In our context, our variant of this condition implies the stronger statement that the probability that *any* polymer of a given size occurs also decays exponentially in the size. We use it to show that with sufficiently large probability all polymers are small, which is the key ingredient in our proof of fast mixing. We note that this condition is where we inherit the requirement $q \geq \Delta^{\Omega(\Delta)}$. The proof outline with more details is provided later, in Section 1.3.2.1.

1.2.1.2 Almost all parameter ranges for the Potts model using planting

As our next result, which we present in Chapter 5, we obtain rapid mixing for the Potts model for almost all parameter ranges.

In Theorem 1.1, the high-q requirement arose from controlling the size of polymers, which represented deviations of a typical configuration from the ground state. Here, we instead use a more delicate technique known as *planting*, originally developed for the analysis of random constraint satisfaction problems. This allows us to extend the fast-mixing result to a wider range of q, at least when q is an integer.

The key idea is to use the Potts representation as an intermediate step for analysing a typical RC configuration. Through the Edwards–Sokal coupling (see Preliminaries 2.3), we can view the Potts model as an alternative representation for the random-cluster model (and vice versa) and analyse one through the other. Building on the work of Coja-Oghlan et al. [25], we can then use the *planted model* to analyse Potts colourings. Roughly, rather than first sampling a random regular graph and then analysing a Potts colouring on that graph, planting allows us to consider the graph and colouring jointly, with the colouring constrained to have the statistics typical of the relevant phase. This joint model is substantially easier to analyse.

The information used to construct the planted model is captured by the *colour statistics*. These consist of the proportion of vertices of each colour (the *vertex statistics*) and the proportion of edges whose endpoints receive each pair of colours (the *edge statistics*). Within each phase, these statistics are strongly concentrated around particular values [25] (for the values of these colour statistics, see Section 5.1.2.1). For example, in the disordered phase, colours appear with roughly equal frequencies, making the vertex statistics symmetric. The edge statistics satisfy a corresponding symmetry.

Instead of polymers, for this result we use a different characterisation of what it means for a configuration to be "far" from the relevant ground state (for more details about this adverse event, see Section 1.3.2.2). Rather than directly analysing the probability of this adverse event for an RC configuration on a typical random regular graph, we instead use the planted graph model with these prescribed statistics.

This framework also allows us to consider all $\Delta \geq 3$ (unlike the previous result, which required the degree to be at least 5 due to the limitations of the polymer model from [49]). Formally, we prove the following theorem, where we use $\beta_c = \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$.

Theorem 1.2. *Let $\Delta \geq 3$, $q \geq 3$ be integers and $\beta > 0$ be a real. Then, with high probability over a random Δ-regular graph, the following holds.*

- *If $\beta < \beta_c$, then the RC dynamics initialised from all-out mixes in $O(n \log n)$ time.*
- *If $\beta > \beta_c$ and $q \geq (5\Delta)^5$, the RC dynamics initialised from all-in mixes in $O(n \log n)$ time.*

Since we can obtain a Potts sample from an RC sample, the theorem implies a near-linear time approximate sampling algorithm for the Potts model throughout the high-temperature regime $\beta < \beta_c$. In the low-temperature regime $\beta > \beta_c$, approximate sampling from the Potts model remains open only for $3 \leq q \leq (5\Delta)^5$.

However, unlike Theorem 1.1, this result could not be extended to an approximation algorithm for the partition function. As a result, we were unable to obtain an algorithm for approximating the proportions of the phases at criticality, and thus the appropriate initialisation for $\beta = \beta_c$. The reason for this is technical and discussed further in Section 3.2, but in essence, the issue is that to obtain an approximation algorithm, we need to sample at multiple values of β. The planting analysis gives us fast mixing w.h.p. over the random graph for each fixed β, but does not guarantee it for the same random graph simultaneously at all values of β required for the FPRAS.

1.2.2 Beyond bounded-degree graphs

In light of the successful development of efficient algorithms for #BIS-hard problems on random regular graphs, a natural question is whether these results can be extended to sparse random graphs of *unbounded degree*. The canonical example is the (sparse) Erdős–Rényi random graph $\mathcal{G}(n, d/n)$, for $d = \Theta(1)$.

For $d > 1$, w.h.p. $G \sim \mathcal{G}(n, d/n)$ contains a giant component, while the remaining components have size $O(\log n)$ and are trees up to removing one edge. Although the maximum degree of G grows with n[16], its average degree is concentrated around d. G is also known to satisfy a (weaker) variant of expansion (applying for all sufficiently large sets) [31, 51]. Due to these similarities, it is natural to ask whether positive algorithmic results for random regular graphs can be extended to $G(n, d/n)$.

[16]W.h.p., the maximum degree of $\mathcal{G}(n, d/n)$ is $\Theta(\frac{\log n}{\log \log n})$.

There has been some progress in the high-temperature regime, mostly confined to the uniqueness regime[17]. For the random-cluster model, Blanca and Gheissari [12] establish fast mixing of the RC dynamics for all real $d > 1$, $q \geq 1$ and $\beta < \beta_u$ w.h.p. over $G \sim \mathcal{G}(n, d/n)$. This matches their result for random regular graphs [10].

Other results for the uniqueness regime use *spectral independence*, a technique originally introduced in [4] to bound mixing times of Markov chains. Bezáková, Galanis, Goldberg, and Štefankovič [9] use it to establish fast mixing of Glauber dynamics for multiple models, including Ising model and the bipartite hardcore (weighted independent sets on bipartite graphs) on the (bipartite) Erdős-Rényi graphs. Blanca and Zhang [13] also use spectral independence to establish fast mixing of global Markov chains for the Potts and bipartite hardcore models..

We also remark that for high temperatures, algorithms based on Markov chains (MCMC) are also known for proper q-colourings on unbounded-degree graphs[18] and for the random k-SAT with an unbounded number of variable occurrences [22, 30].

Despite these developments, Erdős-Rényi graphs still pose difficulties for local MCMC methods, especially once long-range interactions are present. Although the average degree is concentrated around d, the graph contains regions where the local average degree can differ substantially from d. This can lead to instances where typical local and global behaviours are at odds. We illustrate this problem for the random-cluster model. The value of β at which the model begins to prefer ordered configurations depends on the degree and q. Hence, even if β falls within the high-temperature regime in terms of the average degree, we can expect regions containing large-degree vertices that are "locally ordered".

In the low-temperature regime this problem is even more prominent. $\mathcal{G}(n, d/n)$ is likely to contain regions of size $O(\log n)$ with average degree substantially smaller than d. In fact, we expect paths of degree-2 vertices with length up to $\Omega(\log n/d)$ [26]; see also the discussion in Section 6.2. These long paths are likely to be split across multiple components of in-edges, leading to local disordered behaviour even when the macroscopic behaviour tends to be ordered. This is a major barrier in adapting low-temperature sampling algorithms from bounded- to unbounded-degree case.

[17] In this context, we use the term uniqueness regime for the parameter ranges characterised by absence of long-range interactions. For the random-cluster model on $\mathcal{G}(n, d/n)$, we use the same formula for β_u as for integer d.

[18] Proper colourings can be viewed as the zero-temperature limit of the antiferromagnetic Potts model. Nevertheless, when q is sufficiently large relative to the average graph degree, the model exhibits behaviour characteristic of a high-temperature regime.

In particular, successful techniques such as cluster expansion and polymer dynamics crucially rely on *small-set expansion* to establish the necessary Kotecký–Preiss condition. Let us illustrate the problem on the Potts model. Recall that, informally, the condition requires the probability of a polymer (a maximal connected subgraph coloured with a non-majority colour) to decay exponentially in its size. However, consider any colouring in which a path consisting of only degree-2 vertices is coloured with a single non-majority colour. This creates only two bichromatic edges, and therefore, the weight of the colouring obtained by recolouring this path with the majority colour only increases by a constant factor. As a result, we can essentially only bound the probability of such a polymer (the path coloured with non-majority colour) by a constant, while its size can grow with n.

At the time of this work, there was no adaptation of the cluster expansion or polymer dynamics that managed to overcome this limitation caused by degree-2 vertices. The only low-temperature result in this direction was the variant of polymer dynamics for unbounded-degree graphs by Galanis, Goldberg and Stewart [36]. However, that required the input graph to have minimum degree at least 3. To the best of our knowledge, prior to this work, no efficient low-temperature sampling algorithm was known on $\mathcal{G}(n, d/n)$ for a spin system (including the Potts model) that is #BIS-hard on general graphs.

1.2.2.1 Fast mixing beyond bounded degree

In our final result, presented in Chapter 6, we extend fast mixing of the RC dynamics from the ground-state initialisations to $\mathcal{G}(n, d/n)$ at all temperatures. We are able to overcome the previous algorithmic limitations caused by regions with low expansion by our particular use of polymers.

Recall that to prove Theorem 1.1, we used the polymer framework from [49] to capture "how far" does a typical configuration within the phase gets from its ground state (the initialisation of the RC dynamics we use). In particular, we used the Koteck'y–Preiss condition from [49], which establishes that the probability that any polymer of a given size appears decays exponentially in the size, to show that with sufficiently high probability there are no large polymers. Crucially, unlike the cluster-expansion and polymer-dynamics approaches, our mixing argument does not require this decay for polymers of every size. It suffices to rule out polymers of size $A \log n/d$, for a sufficiently large constant A. This allows us to avoid the obstruction described

in the previous section: while the long paths of degree-2 vertices cause the Kotecký–Preiss condition to fail for smaller polymers, we do not need to control polymers at those sizes.

To control the size of the large deviations, we establish a *weak expansion* property for $\mathcal{G}(n, d/n)$: every connected set of size $A \log n/d$ has a number of outgoing edges proportional to its size.

This first allows us to establish a capture the behaviour of the phases on $G(n, d/n)$. Recall that on random regular graphs, the threshold where the typical configuration changed from disordered to ordered was the critical temperature $\beta_c = (1+o_q(1))\frac{2\log q}{\Delta}$. For $\mathcal{G}(n, d/n)$, we instead establish the presence of a *critical window* (β_0, β_1), where $\beta_0 \approx (2 - \frac{1}{10})\frac{\log q}{d}$ and $\beta_1 \approx (2 + \frac{1}{10})\frac{\log q}{d}$ (see (6.1) for exact values). Note that the value of β_c with $\Delta = d$ lies in (β_0, β_1).

We show that (provided that d and q are large enough) for all $\beta \leq \beta_0$ the disordered phase dominates (occurs with all but exponentially small probability), for $\beta \geq \beta_1$ the ordered phase dominates, and for $\beta \in (\beta_0, \beta_1)$ the configurations outside of a phase occur with only exponentially small probability.

Then we use the weak expansion, together with the result about phase dominance, to establish fast mixing. For the disordered phase, it suffices to show that all components of in-edges are small with sufficiently high probability (we defer the reader to Section 1.3.2.4). For the ordered phase, the previous polymer framework from [49] no longer accurately captures what it means to be "far from the all-in configuration" due to degree irregularities and regions of low expansion (see discussion in Section 1.3.2.3). This leads to the main technical innovation in Chapter 6, which is a new ordered polymer model, naturally capturing the deviations from the all-in state. Roughly, we show that it suffices to take "almost disconnected small components" as our polymers. For further discussion on our model, see Section 1.3.2.3.

We thus obtain fast mixing at all temperatures. We can convert the resulting approximate sampling algorithm to an approximation algorithm for the partition function, which we describe in Section 3.2, allowing us to compute the proportions of the phases for $\beta \in (\beta_0, \beta_1)$.

Theorem 1.3. *Let d be a large enough real and $q = d^{\Omega(d)}$ be a real. Then, with high probability over $G \sim \mathcal{G}(n, d/n)$, the following hold.*

- *For all $\beta \leq \beta_0$, the RC dynamics initialised from all-out mixes in $O(n \log n)$ time.*
- *For all $\beta \geq \beta_1$, the RC dynamics initialised from all-in mixes in poly(n) time.*

- *For all $\beta \in (\beta_0, \beta_1)$, the RC dynamics initialised from a suitable random mixture of all-in and all-out mixes in polynomial time. Moreover, this random mixture can be computed in polynomial time.*
- *For all $\beta > 0$, there is an FPRAS for approximating the partition function.*

Note that the requirement that d is sufficiently large is needed to establish a sufficient (weak) expansion, and $q = d^{\Omega(d)}$ is inherited from our use of the polymer framework.

As an additional result, we also provide a deterministic algorithm with a significantly better running time for approximating the partition function with an arbitrary *polynomial* precision. Instead of Markov chains, this algorithm is based on obtaining an approximation of the logarithm of the partition function via symbolic integration[19]. We provide a proof outline in Section 1.3.2.5.

Theorem 1.4. *Let $R > 0$ be any real. Then for all sufficiently large reals d and then all sufficiently large q, there is an algorithm which, with high probability over $G \sim \mathcal{G}(n, d/n)$, for any $\beta > 0$, outputs a $(1 + \frac{1}{n^R})$-approximation to $Z_G(\beta, q)$ in time $n^{1+O(1/R)}$.*

1.3 Proof overview: fast mixing from a ground state

The proofs establishing fast mixing throughout this work follow the same high-level outline, which we now provide here.

As a first step, we consider each phase separately. Let $P \in \{\text{dis}, \text{ord}\}$ denote a phase, and let Ω_G^P be the set of configurations in that phase. Informally, the disordered phase contains all configurations with at most some η-fraction of edges, and the ordered phase contains all configurations with at least $(1-\eta)$-fraction of edges. Here, $\eta > 0$ is a small constant, whose value depends on the setting; see Chapters 4-6 for details. Consider the *RC (Gibbs) distribution within the phase P*, denoted $\pi_G^P = \pi_G(\cdot \mid \cdot \in \Omega_G^P)$. We denote the ground state of the phase P (all-out for the disordered and all-in for the ordered phase) by g_P.

[19]Here, symbolic integration refers to an algorithmic method for efficiently evaluating an integral from a suitable symbolic representation of its integrand; we use the method as a black box.

We show, provided that three key properties hold, that the RC dynamics initialised from the ground state g_P gets close[20] to the conditional distribution π_G^P; we call this *fast mixing within the phase*. Now, if the phase P dominates, then this is enough as π_G^P is already close to the unrestricted RC distribution π_G. Otherwise, we choose the starting state with, roughly, the probability of it occurring. We refer the reader to Section 3.2 for more details, including how we obtain this initial mixture. Thus, for the remainder of the introduction, we focus on obtaining *fast mixing within the phase*.

The central idea behind our proof builds on the result by Gheissari and Sinclair [43], who showed how to overcome exponential slowdown for the low-temperature Ising[21] Glauber dynamics on random regular graphs. Analogously to the RC dynamics, the slowdown there is caused by a bottleneck between two phases – corresponding to majority-1 and majority-2 colourings. Gheissari and Sinclair show that the Ising Glauber dynamics initialised from the random mixture of the two ground states (all-1 and all-2) mixes in $O(n \log n)$ time. They later extended this approach to the random-cluster model on lattices [44], where the behaviour of the phases is qualitatively different from that on random regular graphs.

A crucial property of the Ising and the random-cluster models used in our proof is *monotonicity*. For the random-cluster model, this property translates as follows: suppose there are two copies of the RC dynamics with initial states $F_1 \subseteq F_2$. Suppose the same (arbitrary) edge e is chosen by both copies. Then, the probability of including e in the next step in the first RC dynamics (starting from F_1) is at most the probability of including e in the next step of the second RC dynamics (starting from F_2). In other words, by choosing the same edge and the same uniform random variable to make the decision of whether to set the selected edge to be in for both chains, the inclusion $F_1 \subseteq F_2$ can be maintained throughout the run of the RC dynamics [47]. Analogously, for any two Ising configurations with $\sigma_1 \leq \sigma_2$, if two Ising Glauber dynamics are initialised from σ_1, σ_2 respectively, by choosing the same vertex and same uniform random variable to make the decision in each step, the inequality is maintained throughout the run of the dynamics. No such property is known to hold for the Potts model, but we remark that while this has not been the focus of

[20] By close, we mean within $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ TV-distance. For $\varepsilon < \mathrm{e}^{-\Omega(n)}$, we can obtain a sample with ε TV-distance by enumerating through all $2^{|E|} = \mathrm{poly}(1/\varepsilon)$ RC configurations. For more details, see Chapter 3.

[21] The Ising model is the Potts model with $q = 2$. In this model, vertices are assigned colours 1 and 2, inducing a distribution analogous to the Potts model.

our work, monotonicity holds for the low-temperature bipartite hardcore model (see e.g. [58, Chapter 22]), and thus our method also applies for this setting.

Adapting the proof from [43], we reduce fast mixing within the phase reduces to proving establishing properties: a local maximum of the Gibbs distribution within the phase, weak spatial mixing within the phase, and local mixing. The first, technical property is that $(X_t)_{t\geq 0}$ is indeed unlikely to leave the initial phase. Formally:

Key Property 1 (Local maximum)**.** Let $P \in \{\text{ord}, \text{dis}\}$ be a phase. For a graph G and real parameters $q, \beta > 0$, we say that the RC Gibbs distribution $\pi_G(\,\cdot\,; q, \beta)$ has a *local maximum* in the phase P, if the RC dynamics initialised from π_G^P has at most exponentially small probability of escaping Ω_G^P in a single step, i.e.

$$\mathbb{P}(X_1 \notin \Omega_G^P \mid X_0 \sim \pi_G^P) \leq \mathrm{e}^{-\Omega(|E|)} \tag{1.2}$$

For graphs with $|E| = \Theta(n)$, we can replace $\mathrm{e}^{-\Omega(|E|)}$ with $\mathrm{e}^{-\Omega(n)}$. We choose to state this bound more generally; however, all graph families in this work satisfy $|E| = \Theta(n)$ and we use $\mathrm{e}^{-\Omega(n)}$ interchangeably with $\mathrm{e}^{-\Omega(|E|)}$.

Establishing Key Property 1 is typically straightforward and can be obtained as a corollary of the proof that configurations outside of a phase are very unlikely. Intuitively, this property ensures that once the RC dynamics gets close to the RC distribution within the phase, it is unlikely to leave the phase over polynomially many subsequent steps.

The idea behind the remaining two key properties is to use *local behaviour* to establish global convergence. To explain this idea, let us now outline the proof of fast mixing within the phase. We only outline the proof that the RC dynamics $(X_t)_{t\geq 0}$ initialised from g_P gets within $\frac{1}{4}$ total-variation distance from π_G^P. To obtain any $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ precision, we use a technical probability amplification argument and refer the reader to Section 3.1.1. We use that the TV distance between two distributions is at most the probability that two random variables with the respective distributions disagree. We consider a *restricted RC dynamics* $(\hat{X}_t)_{t\geq 0}$, initialised from π_G^P. It is restricted in the sense that whenever it would choose to make an update that makes its state fall outside of the phase P, the Markov chain ignores the update. It can be shown that π_G^P is the stationary distribution of this restricted RC dynamics, and hence, for all $t \geq 0$, $\hat{X}_t \sim \pi_G^P$.

We consider a *monotone coupling* of $(X_t, \hat{X}_t)$, where both copies choose the same edge to update, and use the same random variable to make the update. Due to monotonicity of the random-cluster model, this maintains $X_t \subseteq \hat{X}_t$ for the disordered

phase (as $\emptyset = X_0 \subseteq \hat{X}_0$) and $X_t \supseteq \hat{X}_t$ for the ordered phase (as $E = X_0 \supseteq \hat{X}_0$), provided that no updates have been ignored.

Then, we consider the disagreement on each edge separately. By the union bound, it is enough to show that, for every edge e, the probability that X_t and $\hat{X}_t$ disagree on e after some polynomial number of steps is at most $\frac{1}{4|E|}$. To obtain this bound, we introduce an additional copy of the RC dynamics restricted to a small neighbourhood around e, which is easier to analyse and has the property that whenever X_t and $\hat{X}_t$ disagree on e, so does $\hat{X}_t$ and this third copy of the RC dynamics.

To state this more precisely, we require some additional definitions. For a vertex v and a positive integer r, the *neighbourhood* (or *ball*) of *radius* r *around* v is the induced subgraph on all vertices with distance at most r from v. We denote it by $B_r(v)$. We consider the Gibbs distribution conditional on the outside of the ball being fixed in the ground state and denote it by $\pi^{g_P}_{B_r(v)}$ and the RC dynamics $(X^v_t)_{t\geq 0}$ restricted to the ball, initialised from the ground state g_P (in the sense that if the edge chosen to update falls outside of the ball, the dynamics ignores this update). By construction $\pi^{g_P}_{B_r(v)}$ is the stationary distribution of $(X^v_t)_{t\geq 0}$.

We pick v to be a vertex adjacent to e and use the same coupling as before: X^v_t picks the same edge to update as X_t and $\hat{X}_t$, and the same uniform random variable to make the update. The main idea now is that conditional on $\hat{X}_t$ having ignored no updates, whenever X_t and $\hat{X}_t$ disagree on e, then so do $\hat{X}_t$ and X^v_t (as illustrated in Figure 1.3).

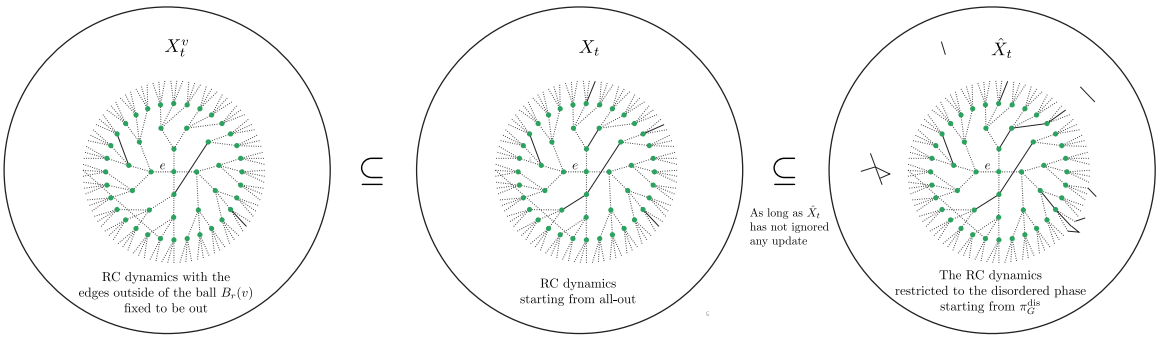


Figure 1.3: Illustration of the monotone coupling in the disordered phase. As long as the RC dynamics restricted to the disordered phase, $\hat{X}_t$, has not ignored an update, the three coupled chains satisfy $X^v_t \subseteq X_t \subseteq \hat{X}_t$. For the ordered phase, the inclusions are reversed. As a result, provided that no updates have been ignored, whenever X_t and $\hat{X}_t$ disagree on e, so do X^v_t and $\hat{X}_t$.

Thus, it suffices to bound the probability that the RC dynamics restricted to the phase, $\hat{X}_t$, and the RC dynamics on the local neighbourhood with the outside

restricted to the ground state, X_t^v, disagree on e. This makes the problem easier, because the RC dynamics X_t^v only changes on a small neighbourhood (which for random graphs is furthermore treelike). Using the triangle inequality and a probability bound, we reduce this disagreement probability to three terms:

a) the probability that an update is ignored within polynomially many steps (captured by Key Property 1),

b) the difference in marginals on e between the RC distribution conditional on being in the phase π_G^P and the RC distribution on the radius-r ball with outside fixed to the ground state g_P, $\pi_{B_r(v)}^{g_P}$ (*weak spatial mixing within the phase*), and

c) the mixing time of the restricted chain X_t^v (*local mixing*).

We next formally state the remaining two key properties:

Key Property 2 (Weak spatial mixing (WSM) within the phase at radius r)**.** Let $P \in \{\text{ord}, \text{dis}\}$ be a phase and $r \geq 1$ an integer. Say that $G = (V, E)$ has *WSM within the phase P at radius r* if, for every $v \in V$ and edge $e \in E$ incident to v, it holds that $|\pi_{B_r(v)}^{g_P}(e \in \cdot) - \pi_G^P(e \in \cdot)| \leq \frac{1}{100|E|}$.

Intuitively, this property says that the influence of edges "far away" is small, provided that the configuration remains sufficiently ordered (respectively, disordered). We formulate WSM within the phase with the specific bound $1/(100|E|)$, as this is the form used by the proof of fast mixing. More generally, one can consider a formulation where the difference in marginals is at most K^r (for some $K < 1$). We use this variant later in the proof of Theorem 1.4; see discussion in Section 1.3.2.3.

In the last key property, we use the notion of the *log-Sobolev constant*. This is a standard quantity relating the decay in entropy to the mixing time, which we mostly import from previous work and use a black box. For the definition and its usage, we refer the reader to Preliminaries 2.4.

Key Property 3 (Local mixing)**.** Let $P \in \{\text{ord}, \text{dis}\}$ be a phase, $r \geq 1$ an integer and $G = (V, E)$ a graph. We say that the RC dynamics has *local mixing in the phase P on radius-r neighbourhoods of G* if, for every $v \in V$, the (RC) Glauber dynamics on the ball $B_r(v)$, where the outside of the ball is conditioned to be in the ground state g_P, mixes in polynomial time. We say that it has *optimal local mixing* if the log-Sobolev constant of the restricted RC dynamics for all v is $\Omega(1/|E(B_r(v))|)$.

Given the three key properties, we obtain fast mixing, as is formalised in the following theorem, whose proof we defer to Chapter 3. We state the properties and theorem for the random-cluster model, but the method can be adapted to any monotone model.

Theorem 1.5. *Let $q, \beta > 0$ be real parameters, $P \in \{\mathrm{ord}, \mathrm{dis}\}$ a phase, and $\mathcal{G}$ a family of graphs such that for all $G \in \mathcal{G}$, $|E(G)| = \Omega(n)$ and there is an integer radius $r \geq 1$ such that all of the following hold.*

- *$\pi_G(\,\cdot\,; q, \beta)$ has a local maximum in the phase P (Key Property 1),*
- *G has WSM within the phase P at radius r (Key Property 2), and*
- *the RC dynamics has local mixing in the phase P on radius-r neighbourhoods of G (Key Property 3).*

Then there exists $T = \mathrm{poly}(n)$, such that for all sufficiently large n, all real $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$, and any n-vertex graph $G \in \mathcal{G}$, the RC dynamics $(X_t)_{t\geq 0}$ initialised from the ground state g_P satisfies

$$\left\| X_{O(T \log \frac{1}{\varepsilon})} - \pi_G^P \right\|_{\mathrm{TV}} \leq \varepsilon.$$

If, moreover, on all graphs $G \in \mathcal{G}$, the RC dynamics has optimal local mixing *in the phase P at radius r, then the above holds with $T = O(|E| \log |E|)$.*

We remark that for sparse graphs (where $|E| = O(n)$), the optimal local mixing gives $O(n \log n)$ mixing time.

1.3.1 Local mixing

We next outline the proofs of Key Property 3. For the most part, the required local-mixing results follow from existing mixing results on trees. In the remaining settings, we obtain polynomial mixing by using a standard canonical path method.

The key property of sparse random graphs used here is that they are w.h.p. *locally treelike*. That is, for every sufficiently small radius $r = O(\log n)$, with high probability over the random graph, every radius-r neighbourhood differs from a tree by at most a constant number of edges.

We use the standard fact that adding or removing a constant number of edges can increase the mixing time of the Glauber dynamics by at most a multiplicative constant (see, e.g., [11, Lemma 6.6]). Thus, it suffices to establish local mixing on

trees resembling the local neighbourhoods of the random graphs: regular trees for random regular graphs and *Poisson trees*[22] for $\mathcal{G}(n, d/n)$.

The constant number of additional edges can be accounted for without changing the asymptotic bounds required for local mixing (see, e.g., [11, Lemma 6.6]). Thus, for the purposes of local mixing, it suffices to consider trees resembling the neighbourhoods of the random graphs: regular trees for random regular graphs and *Poisson trees*[23] for $\mathcal{G}(n, d/n)$. In the latter case, we require the local-mixing result to hold with probability $1 - o(n^{-1})$ over the Poisson tree.

We first consider the disordered phase. Recall, that there the goal is to obtain the mixing time for the RC dynamics on the ball (tree) with the outside edges fixed to be out. We refer to the resulting tree as a *free tree*. Since all edges outside the tree are fixed to be out, the RC distribution on the free tree is the same as if the tree were the entire graph. Moreover, on a tree every edge is a cut-edge and hence, the states of the individual edges are mutually independent [11]. The RC dynamics therefore mixes once all edges have been updated, and the required $\Omega(1/|E|)$ log-Sobolev bound is known; see, e.g., [11]. Thus, local mixing in the disordered phase follows directly from known results.

For the ordered phase, the situation is more involved. Recall that the goal here is to establish fast mixing for the RC dynamics on the ball (tree) when the outside is fixed to be in. On a tree this corresponds to identifying all leaves into a single component. This is known as the *wired boundary condition*, and we refer to the resulting tree as a *wired tree*.

For regular trees and integer q, the required optimal local-mixing result is already known. Martinelli, Sinclair and Weitz [61] established optimal mixing for the Potts Glauber dynamics on trees with monochromatic boundary conditions[24]. Blanca, Chen, Štefankovič and Vigoda [15] subsequently transferred this result to the RC dynamics on wired trees. We use these results as a black box to establish optimal local mixing in the corresponding random-regular-graph setting.

There are two settings in which this argument does not directly apply. First, for non-integer q, there is no corresponding Potts representation, and hence, we cannot

[22] A Poisson tree is generated by a branching process in which each vertex has an independent Poisson(d) number of children. In our application, the required property must hold with probability $1 - o(n^{-1})$ over the Poisson tree.

[23] A Poisson tree is generated by a branching process in which each vertex has an independent Poisson(d) number of children.

[24] A tree with monochromatic boundary condition is a tree where all leaves are conditional to have a single colour. Roughly, this translates to the wired boundary condition for the random-cluster model.

transfer the Potts mixing result to the RC dynamics. Moreover the proof technique from [61] does not directly extend to the random-cluster model, as it heavily relies on the fact that the Potts model is defined via interactions between adjacent vertices, whereas the component structure defining the random-cluster model can depend on edges arbitrarily far away. Second, this technique also fails for $\mathcal{G}(n, d/n)$, where the local neighbourhoods are described by Poisson rather than regular trees, as the proof from [61] relies heavily on symmetry arguments.

In these remaining cases, we establish polynomial local mixing directly for the RC dynamics by adapting the standard canonical-path method (see e.g. [53]). This gives the polynomial local-mixing bound required by Key Property 3, but not the optimal log-Sobolev bound available in the regular-tree, integer-q setting. This is the reason for the discrepancy between the $O(n \log n)$ and polynomial mixing-time bounds in our main results.

1.3.2 Weak spatial mixing within the phase

We next outline the proof of the main technical challenge of this work: WSM within the ordered phase. The proof for the disordered phase follows a similar, but simpler, argument, which we briefly outline in Section 1.3.2.4.

Fix an edge e, a vertex v incident to e and an integer radius r. Recall that the WSM within the ordered phase aims to bound the difference between marginals on e (the probabilities that e is an in-edge) under two distributions: the RC distribution within the ordered phase $\pi_G^{\mathrm{ord}} = \pi_G(\,\cdot \mid \cdot \in \Omega_G^{\mathrm{ord}})$, and the RC distribution on the *wired ball*, i.e. $\pi^+_{B_r(v)}(\,\cdot\,) = \pi_G(\,\cdot \mid E \setminus E(B_r(v)) \subseteq \cdot\,)$. The aim is to prove

$$|\pi_G^{\mathrm{ord}}(e \in \cdot\,) - \pi^+_{B_r(v)}(e \in \cdot\,)| \leq \tfrac{1}{100|E|}. \tag{1.3}$$

Since the state of e is binary, the left-hand side of (1.3) is precisely the TV distance between the distributions of the state of e under π_G^{ord} and $\pi^+_{B_r(v)}$. Recall that the TV distance between two distributions is upper-bounded by the probability that two coupled random variables with the respective distributions disagree.

Therefore, to show (1.3), it is enough to construct a coupling $(\boldsymbol{F}^{\mathrm{ord}}, \boldsymbol{F}^+) \sim (\pi_G^{\mathrm{ord}}, \pi^+_{B_r(v)})$ such that $\mathbb{P}(\boldsymbol{F}^{\mathrm{ord}}$ and $\boldsymbol{F}^+$ disagree on $e) \leq \frac{1}{100|E|}$.

We construct a monotone coupling (see Lemma 2.14) by iteratively revealing the states of the edges in the two configurations, using the same adaptive revealing order. Monotonicity ensures that the probability of including an edge in $\boldsymbol{F}^{\mathrm{ord}}$ is at most the probability of including an edge in $\boldsymbol{F}^+$, so that we obtain $\boldsymbol{F}^{\mathrm{ord}} \subseteq \boldsymbol{F}^+$.

For this purpose, it is convenient to represent a partially revealed RC configuration as a function $\mathcal{F} : E \to \{0, 1, *\}$, where 1 denotes a revealed in-edge, 0 a revealed out-edge, and $*$ an edge whose state has not yet been revealed. We write $R(\mathcal{F}) = \mathcal{F}^{-1}(\{0, 1\})$ for the set of revealed edges. Once all edges have been revealed, $\mathcal{F}^{-1}(\{1\})$ is the resulting RC configuration.

The revealing process produces pairs of partially revealed configurations

$$(\mathcal{F}_0^{\mathrm{ord}}, \mathcal{F}_0^{+}), (\mathcal{F}_1^{\mathrm{ord}}, \mathcal{F}_1^{+}), \dots, (\mathcal{F}_\ell^{\mathrm{ord}}, \mathcal{F}_\ell^{+}), \tag{1.4}$$

where at every step the same edges have been revealed in both configurations. The next edges to reveal are chosen adaptively from the information revealed so far, and their states are sampled according to the appropriate conditional distribution. We use monotone coupling to maintain $\mathcal{F}_i^{\mathrm{ord}} \leq \mathcal{F}_i^{+}$ in each step. At the end, the two resulting sets of in-edges have distributions π_G^{ord} and $\pi_{B_r(v)}^{+}$, respectively.

The aim is to choose the revealing order so that, with probability at least $1 - \frac{1}{100|E|}$, before the edge e itself is revealed, the two partially revealed configurations induce the same conditional distribution on all unrevealed edges. Once this happens, we can couple the remaining revelations identically, and in particular ensure that the two configurations agree on e.

We next explain when these conditional distributions coincide. For a model with local interactions, such as the Potts model, the conditional distribution on the unrevealed region depends only on the revealed colours of vertices adjacent to it. For the random-cluster model, this is not enough: its weight depends on the number of connected components, so the distribution on unrevealed edges can depend on how vertices at the boundary of the unrevealed region are connected through already revealed in-edges.

To be precise, let $\mathcal{F}_1, \mathcal{F}_2$ be two partially revealed configurations with the same set of revealed edges $R = R(\mathcal{F}_1) = R(\mathcal{F}_2)$. We call a vertex a *boundary vertex* if it is incident to both a revealed edge in R and an unrevealed edge in $E \setminus R$. We say that the *boundary component structures of* $\mathcal{F}_1, \mathcal{F}_2$ *agree* if, for every pair of boundary vertices, they are connected by revealed in-edges in $\mathcal{F}_1$ if and only if they are connected by revealed in-edges in $\mathcal{F}_2$. If the boundary component structures agree, then the conditional distributions on the unrevealed edges coincide; see Section 4.1.5 for the formal statement and proof and see Figure 1.4 for illustration.

Rather than trying to make the two boundary component structures agree exactly, we use a stronger condition that is easier to analyse in the ordered phase.

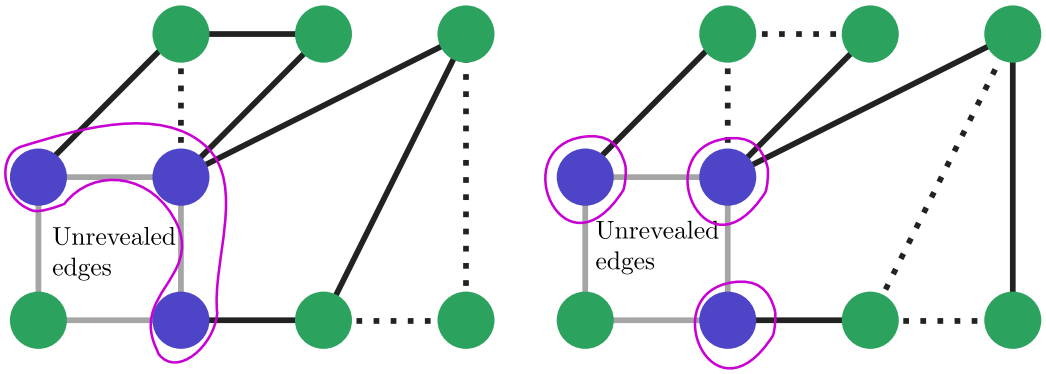


Figure 1.4: Two partially revealed configurations with same revealed edges. Grey edges are unrevealed, solid black edges are revealed in-edges, and dashed black edges are revealed out-edges. The blue vertices are the boundary vertices. In both configurations, the state of the revealed edges incident to the unrevealed edges is the same, but on the left-hand side, all boundary vertices are connected in a single component of in-edges, whereas on the right-hand side they are all in separate components. As a consequence, the conditional distributions on the unrevealed edges do not coincide.

Since an ordered configuration contains a giant component[25] of in-edges, we aim to reveal the configurations until all boundary vertices are connected to this giant component. We call this a *wired boundary*. Formally, a partially revealed configuration $\mathcal{F} : E \to \{0, 1, *\}$ has a *wired boundary in* $B_r(v)$ if all boundary vertices lie in $B_r(v)$, all belong to the *giant* component formed by the revealed in-edges $\mathcal{F}^{-1}(\{1\})$, and no edge incident to v has yet been revealed. If two partially revealed configurations both have a wired boundary, then all their boundary vertices lie in a single component, so their boundary component structures agree. Moreover, by the monotone coupling $\mathcal{F}_i^{\text{ord}} \leq \mathcal{F}_i^{+}$, hence whenever there is a wired boundary in $\mathcal{F}_i^{\text{ord}}$, there is also a wired boundary in $\mathcal{F}_i^{+}$, since adding additional edges cannot disconnect a component.

We remark that there is one additional complication: $\mathcal{F}^{\text{ord}}$ is sampled conditional on the final configuration being ordered. However, with all but very small probability we are able to drop the conditioning after the first set: there we reveal the edges outside of the ball and show with sufficient probability, $\mathcal{F}_0^{\text{ord}}$ already contains enough revealed in-edges guaranteeing that the final configuration will be ordered regardless of the state of the unrevealed edges[26]. This introduces only a small additive error; we refer the reader to Chapter 4 for details.

[25] A giant component is a component containing a majority of vertices. The ordered phase is defined to contain enough edges that a giant component of in-edges is guaranteed to exist for any ordered configuration (see Lemma 4.24 for regular graphs, and Lemma 6.28 for Erdős-Rényi graphs).

[26] A configuration belongs to the ordered phase if it contains at least some $(1-\eta)$-fraction of edges. Therefore, we can drop the conditioning as soon as the partially revealed configuration has at least $(1-\eta)|E|$ in-edges.

Thus, the proof reduces to choosing a revealing order that produces a wired boundary before reaching e, with probability at least $1-\frac{1}{100|E|}$. We first construct an iterative revealing scheme and show that failure to reach a wired boundary implies the presence of a large polymer. We then refine this argument and bound the probability that no suitable revealing scheme exists by the probability that the configuration contains a suitably "badly connected" subgraph.

1.3.2.1 WSM within the ordered phase on random regular graphs via polymers: proof outline

We first consider the proof of WSM within the ordered phase for random regular graphs, used in Theorem 1.1. Recall from the previous section that it suffices to find a revealing scheme which, with sufficiently high probability, reaches a wired boundary before revealing the edge e. We achieve this using a bottom-up revealing scheme and show that it can fail only if the configuration contains a large *ordered polymer*.

We first reveal the edges outside the ball $B_r(v)$. In the terminology of partially revealed configurations, $R(\mathcal{F}_0^{\mathrm{ord}}) = R(\mathcal{F}_0^{+}) = E \setminus E(B_r(v))$. By the definition of the wired-ball distribution, all these revealed edges are in for $\mathcal{F}_0^{+}$, i.e., $\mathcal{F}_0^{+}(f) = 1$ for every $f \in E \setminus E(B_r(v))$. As discussed above, with sufficiently high probability this already reveals enough in-edges to guarantee that the final configuration will be ordered, regardless of the states of the remaining edges. We can therefore remove the conditioning on the ordered phase for the remainder of the revealing process and use the monotone coupling described above.

Let us first suppose that $G[B_r(v)]$ is a tree. If the current partial configuration $\mathcal{F}_i^{\mathrm{ord}}$ does not have a wired boundary, then some boundary vertex is not connected by revealed in-edges to the giant component. Among those boundary vertices, we choose one farthest from v, and reveal the subtree rooted immediately above it. We repeat this procedure until either a wired boundary is obtained or the revealing process reaches v. Figure 1.5 illustrates one step of this process.

We show that if this process fails to obtain a wired boundary before revealing e, then the resulting ordered configuration is "far" from the ground state: roughly, it contains a large region that is poorly connected by in-edges to the giant component. Since the underlying random regular graph is an expander, such a region (or once it is suitably enlarged) has many outgoing out-edges. While a typical ordered configuration might have a linear number of out-edges, they are unlikely to be concentrated around a relatively small connected region.

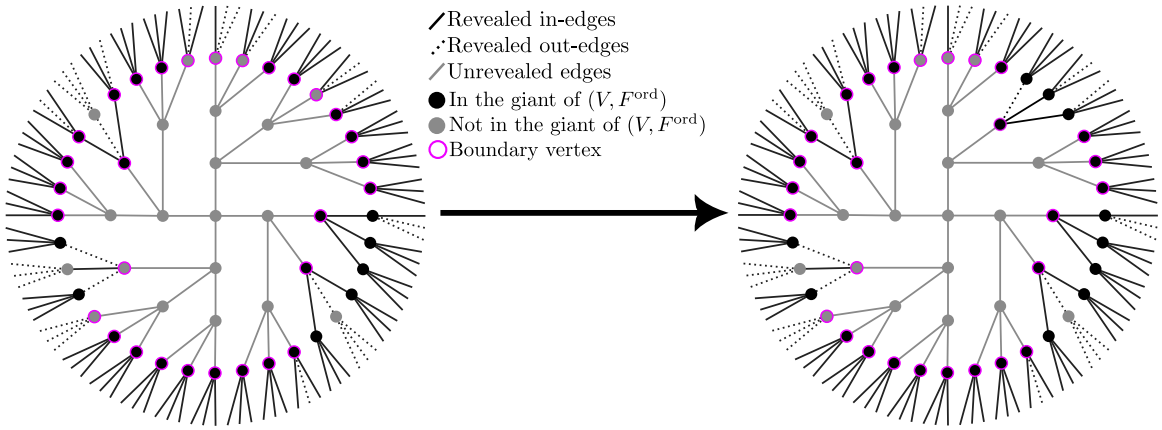


Figure 1.5: One step of the bottom-up revealing scheme in a tree. The unrevealed edges are grey, and the boundary vertices are highlighted in purple. Among the boundary vertices not yet connected to the giant component, we choose one furthest from v and reveal the subtree above it.

To capture this phenomenon formally, we use the notion of *ordered polymers* introduced by Helmuth, Jenssen and Perkins [49]. The somewhat involved definition below is designed so that every non-giant component of in-edges is contained in an ordered polymer [49]. The obstruction arising from our revealing process is slightly more general, since the poorly connected region may still have some in-edge connections to the giant component; we show in Chapter 4 that failure of the revealing process nevertheless implies the presence of an ordered polymer of size $\Omega(r)$.

Definition 1.6. *Let $G = (V, E)$ be a graph and $F \in \Omega_G^{\mathrm{ord}}$ an ordered configuration.*

Let $\mathcal{B}_0(F) \coloneqq F$, and for all $i \geq 0$, $\mathcal{B}_{i+1}(F) \coloneqq F \cup \{e : e \text{ is incident to } v \in V \text{ with } \geq \frac{5}{9}\Delta \text{ of its edges in } \mathcal{B}_r(F)\}$. Let $\mathcal{B}_\infty(F) \coloneqq \bigcup_{i\geq 0} \mathcal{B}_i(F)$.

The ordered polymers *of F are connected components of edges in $\mathcal{B}_\infty(F)$. We omit the weight of the polymer in this discussion; see Section 4.3.1 for details.*

Helmuth, Jenssen and Perkins [49] establish a Kotecký–Preiss condition for this polymer model, which implies exponential decay in the probability of a large polymer containing a given vertex. Since failure of our revealing process implies the presence of a polymer of size $\Omega(r)$, appropriately choosing $r = \Theta(\log n)$ makes the coupling fail to agree at e with probability at most $\frac{1}{100|E|}$.

Finally, the radius-r neighbourhood of a random regular graph need not itself be a tree. However, we choose $r = \Theta(\log n)$ such that with high probability every radius-r neighbourhood differs from a tree by only a constant number of edges. We can therefore find a forest of depth $\Omega(r)$ within the neighbourhood and apply the same revealing argument there. We refer to Chapter 4 for the precise construction.

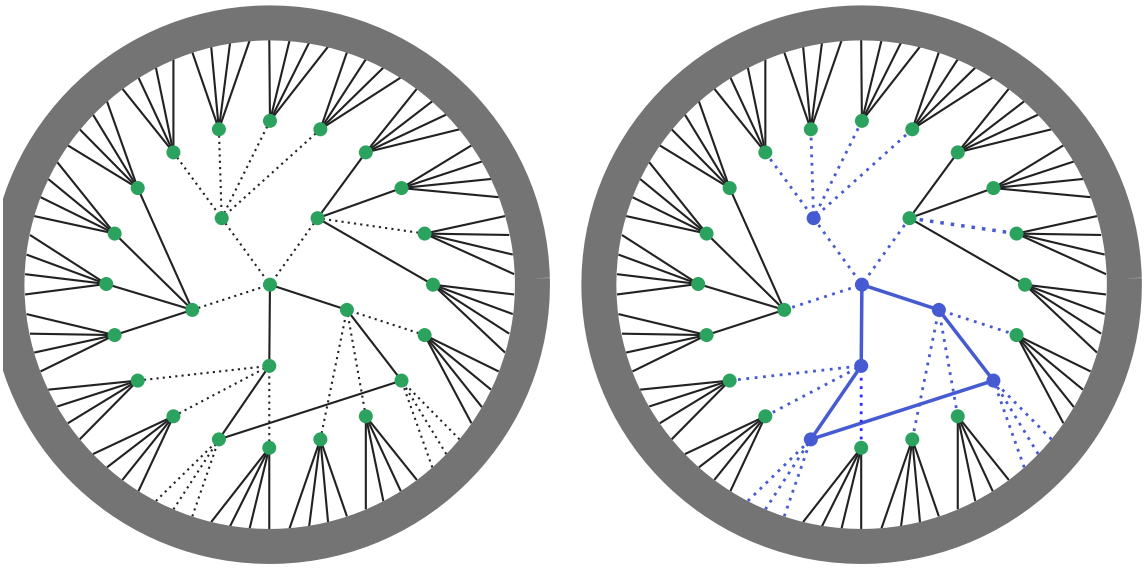


Figure 1.6: Illustration of an ordered polymer. The left-hand side depicts an ordered RC configuration, where solid edges are in, and dashed edges are out. The right-hand side highlights in blue the ordered polymer containing the central region. The two small components (an isolated vertex and a cycle of length 5), alongside all of their incident edges, are contained in a single polymer.

1.3.2.2 Wired boundary using planting: proof outline

Next, we outline the proof of WSM within the ordered phase used for Theorem 1.2. Here, rather than using polymers, we give a simpler characterisation of the obstruction to obtaining a wired boundary. To obtain this characterisation, it is useful to consider *all possible revealing processes* at the same time (instead of one particular strategy), in the sense we next explain.

Fix an ordered configuration $F^{\mathrm{ord}} \in \Omega_G^{\mathrm{ord}}$. Say that a revealing process (1.4) is *consistent with* F^{ord}, if the resulting configuration is exactly F^{ord} (i.e. an edge is revealed to be an in-edge in the process if and only if it belongs to F^{ord}), and moreover we require the first set of revealed edges to be exactly the edges outside of the ball, $E \setminus E(B_r(v))$. The process might be adaptive. We now ask whether there is a revealing scheme consistent with F^{ord} that obtains a wired boundary prior to revealing *any edge incident to* v. Note that we require no edge incident to v to be revealed (rather than solely e), for convenience in the analysis.

As a first step towards obtaining this characterisation, we note that because none of the edges incident to v can be revealed at the time we obtain a wired boundary, it is useful to remove v and its incident edges from the configuration. We write $(V, F^{\mathrm{ord}})\setminus v$ for the resulting graph and denote its giant component by $\kappa_{\overline{v}} = \kappa_{\overline{v}}(F^{\mathrm{ord}})$.

We now define the region around v that cannot be connected to this giant component without using v. Let $S_v(F^{\text{ord}})$ be the vertex set of the maximal connected subgraph of the underlying graph G that contains v and is disjoint from $V(\kappa_{\overline{v}})$ (see Figure 1.7 for illustration). Notice that connectivity in this definition is in G, whereas the giant component is defined in terms of in-edges F^{ord}. Hence, for instance, $S_v(F^{\text{ord}})$ may contain several components of in-edges.

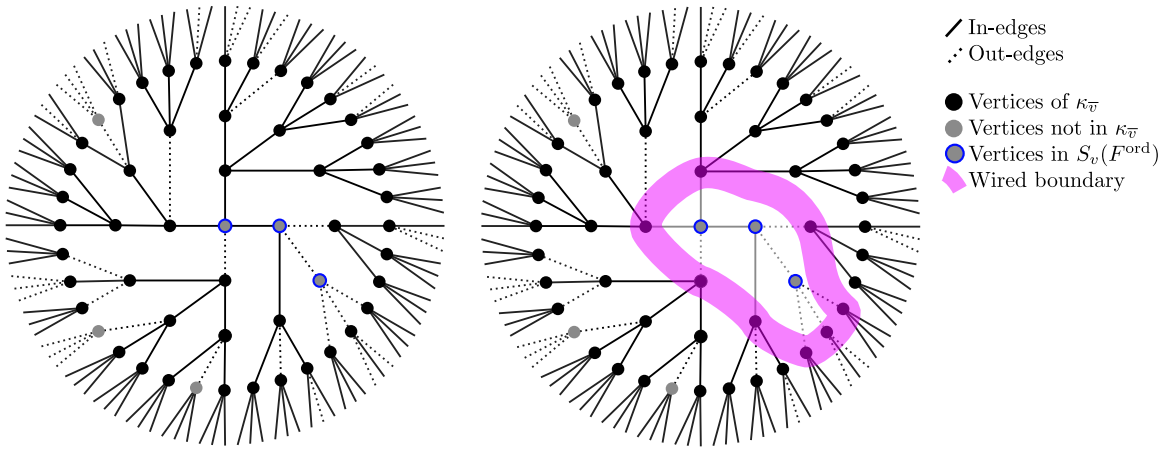


Figure 1.7: An example of the set $S_v(F^{\text{ord}})$. The vertices of $\kappa_{\overline{v}}$ are shown in black, and the vertices of $S_v(F^{\text{ord}})$ are outlined in blue. Note that while $S_v(F^{\text{ord}})$ is connected in G, it might not be connected in (V, F^{ord}). The vertices of $S_v(F^{\text{ord}})$ might be connected to the giant component of (V, F^{ord}) through incident edges of v.

The component $\kappa_{\overline{v}}$ and the set $S_v(F^{\text{ord}})$ give a simple way of understanding the obstruction to a wired boundary. In particular, we show that a revealing process of this form can obtain a wired boundary if and only if

$$S_v(F^{\text{ord}}) \subseteq B_{r-1}(v).$$

In other words, the "adverse event" is precisely that $S_v(F^{\text{ord}})$ reaches the boundary of $B_r(v)$.

We first explain why it is a necessary condition. Suppose for contradiction that $S_v(F^{\text{ord}})$ contains a vertex w at distance r from v, while a revealing process consistent with F^{ord} obtains a wired boundary before revealing any edge incident to v. Consider a path from v to w contained in $S_v(F^{\text{ord}}) \cap B_r(v)$.[27] Since the revealing process starts outside $B_r(v)$ and has not yet revealed any edges incident to v, at the time the wired

[27] While it might be the case that a particular w is not reachable from v via such a path, then there must exist some distance-r vertex for which such a path exists.

boundary is obtained, this path must contain a boundary vertex.[28] Every boundary vertex then belongs to the giant component of revealed in-edges, and hence, to $\kappa_{\overline{v}}$. This contradicts the fact that the path lies in $S_v(F^{\mathrm{ord}})$, which is disjoint from $\kappa_{\overline{v}}$.

To understand why this is a sufficient condition, suppose $S_v(F^{\mathrm{ord}})$ is contained in $B_{r-1}(v)$. By its definition as a maximal subgraph outside of $\kappa_{\overline{v}}$, all neighbours of $S_v(F^{\mathrm{ord}})$ outside $S_v(F^{\mathrm{ord}})$ belong to $\kappa_{\overline{v}}$. Moreover, no in-edge belonging to $\kappa_{\overline{v}}$ is incident to a vertex of $S_v(F^{\mathrm{ord}})$. Therefore, leaving all edges incident to $S_v(F^{\mathrm{ord}})$ unrevealed does not remove any edge of $\kappa_{\overline{v}}$, and hence $\kappa_{\overline{v}}$ remains the giant component of revealed in-edges in the following partial configuration.

$$\mathcal{F}(f) = \begin{cases} * & f \text{ is incident to a vertex in } S_v(F^{\mathrm{ord}}) \\ 1 & f \text{ is not incident to a vertex in } S_v(F^{\mathrm{ord}}) \text{ and } f \in F^{\mathrm{ord}} \\ 0 & f \text{ is not incident to a vertex in } S_v(F^{\mathrm{ord}}) \text{ and } f \notin F^{\mathrm{ord}} \end{cases}$$

The boundary vertices of $\mathcal{F}$ are exactly the neighbours of $S_v(F^{\mathrm{ord}})$ outside of the set. As we noted above, these all belong to the giant component of in-edges. Therefore, $\mathcal{F}$ has a wired boundary.

We can then consider a simple revealing scheme, which first reveals all edges outside of the ball, and then all remaining edges in $R(\mathcal{F})$. After the second step, the process obtains a wired boundary. For illustration, for the configuration depicted in Figure 1.7, this revealing process would obtain exactly the wired boundary marked in purple.

We remark that we omitted a few technical details and corner cases, which we expand on in the full proof. We also note that the formal proof of WSM does not literally search over revealing schemes in this way. Instead, we condition on the possible locations of the wired boundary and bound the resulting disagreement probability. The characterisation above is a convenient way of understanding the event controlled by that calculation; we refer to Chapter 5 for the full formal argument.

It remains to show that, with sufficiently high probability over $\boldsymbol{F}^{\mathrm{ord}} \sim \pi_G^{\mathrm{ord}}$, the set $S_v(\boldsymbol{F}^{\mathrm{ord}})$ is contained in $B_{r-1}(v)$. We prove this by bounding the probability that there exists a simple length-r path from v contained entirely in $S_v(\boldsymbol{F}^{\mathrm{ord}})$. This is where we use the Potts representation and planting. Recall from Section 1.2.1.2 that planting allows us to analyse a typical ordered Potts colouring and then the typical ordered RC distribution on a w.h.p. random regular graph by instead considering the planted model. That is, we consider a uniformly random regular *coloured* graph

[28] We suppress here a minor boundary case concerning edges incident to distance-r vertices. See Chapter 5 for the precise argument.

conditional on prescribed colour statistics. The resulting RC configuration is obtained by taking each monochromatic edge with probability $1 - \mathrm{e}^{-\beta}$ independently.

Rather than directly bounding the probability that such a path exists, we enumerate all potential simple length-r paths from v and bound the expected number that are contained in $S_v(\boldsymbol{F}^{\mathrm{ord}})$. The advantage of the planted model is its symmetry, allowing us to rely on relatively simple combinatorial arguments. In particular, before any further conditioning, every vertex other than v has the same probability of belonging to the giant component $\kappa_{\overline{v}}$. To obtain a lower bound on the size of this component, we adapt a result of Coja-Oghlan et al. [25]; this is where the requirement $q \geq (5d)^5$ arises. We then use the symmetry of the planted model to bound, for each potential path, the probability that it avoids $\kappa_{\overline{v}}$. Summing these probabilities over all potential paths gives a sufficiently small bound on their expected number, and Markov's inequality implies that, with sufficiently high probability, no such path exists. Consequently, $S_v(\boldsymbol{F}^{\mathrm{ord}}) \subseteq B_{r-1}(v)$, giving WSM within the ordered phase.

1.3.2.3 Wired boundary via ordered polymers beyond bounded degree: proof outline

We finish the discussion of WSM within the ordered phase with the setting of Theorem 1.3, where the underlying graph is $G \sim \mathcal{G}(n, d/n)$. Recall from the previous section that the obstruction to obtaining a wired boundary can be characterised by the set $S_v(\boldsymbol{F}^{\mathrm{ord}})$, which is the maximal connected set (in G) of vertices outside of the giant component of in-edges *after removing* v. In particular, we showed that a wired boundary can be obtained by some revealing process if $S_v(\boldsymbol{F}^{\mathrm{ord}})$ is contained in $B_{r-1}(v)$. Hence, it is enough to show that the probability of this event is sufficiently large. For Theorem 1.2, we bound the probability of this event using planting. Here, we instead capture the same obstruction using a (new) polymer model.

In the proof of Theorem 1.1, we bound the probability that a large ordered polymer exists, as defined by Helmuth, Jenssen and Perkins [49] (see Definition 1.6). However, this model is not applicable in this setting.

Recall that its definition iteratively groups together regions containing sufficiently many out-edges. This works well on regular graphs, but can fail in the presence of low-degree vertices, or more generally, regions with expansion at most $\frac{1}{2}$. In particular, a small component of in-edges need not be contained in a single polymer: its surrounding out-edges may instead be split between several polymers, as illustrated in Figure 1.8. Consequently, the absence of large polymers under the previous definition no longer guarantees that $S_v(\boldsymbol{F}^{\mathrm{ord}})$ is small.

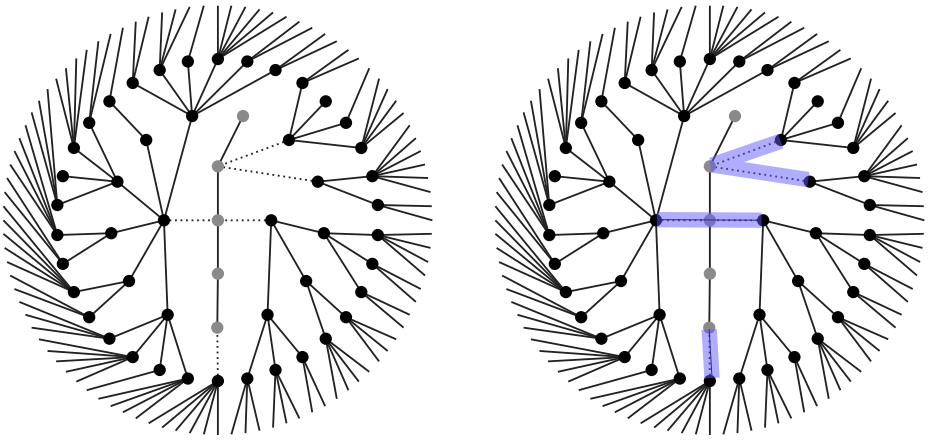


Figure 1.8: Illustration of why the ordered polymer model from Definition 1.6 does not capture small components on graphs with low-degree vertices. On the left, the configuration contains a small component of in-edges with 5 vertices. Under the previous polymer definition, the out-edges surrounding this component are split between three different polymers, highlighted on the right. Thus, the small component is not captured by a single polymer.

Instead of trying to adapt the model from [49], we define a new notion of ordered polymers that directly captures the characterisation from the previous section.

Definition 1.7. *Let $G = (V, E)$ be a graph, $F \subseteq E$ an ordered configuration and $v \in V$ a vertex. Then the* ordered polymers *of F (with respect to v) are the maximal connected subgraphs (in G) of vertices in $V \setminus \kappa_{\overline{v}}$ together with all edges incident to these vertices (*non-singleton polymers*). Any out-edge of F not already contained in a non-singleton polymer forms a* singleton polymer*.*

Roughly, the intuition is that a polymer vertex is a vertex contained either in a non-giant component of in-edges, or in a subgraph that would be disconnected from a giant by removing one vertex. See Figure 1.9 for illustration.

There is also a second motivation for this definition. Recall that the update probability of an edge in the RC dynamics depends on whether or not it is a cut-edge. Every in-edge connecting a small component to the rest of the configuration is therefore relevant, but so is an edge whose removal would disconnect a larger subgraph from the giant component. Our definition is designed to capture both types of regions. Up to minor modifications, the resulting polymers therefore correspond to the structures whose connectivity can affect the local behaviour of the RC dynamics.

The definition directly captures the adverse event from the previous section. Since $S_v(\boldsymbol{F}^{\mathrm{ord}})$ is connected in G and consists entirely of vertices outside $\kappa_{\overline{v}}$, it is contained in a non-singleton polymer. If $S_v(\boldsymbol{F}^{\mathrm{ord}})$ contains a vertex at distance r from v, then

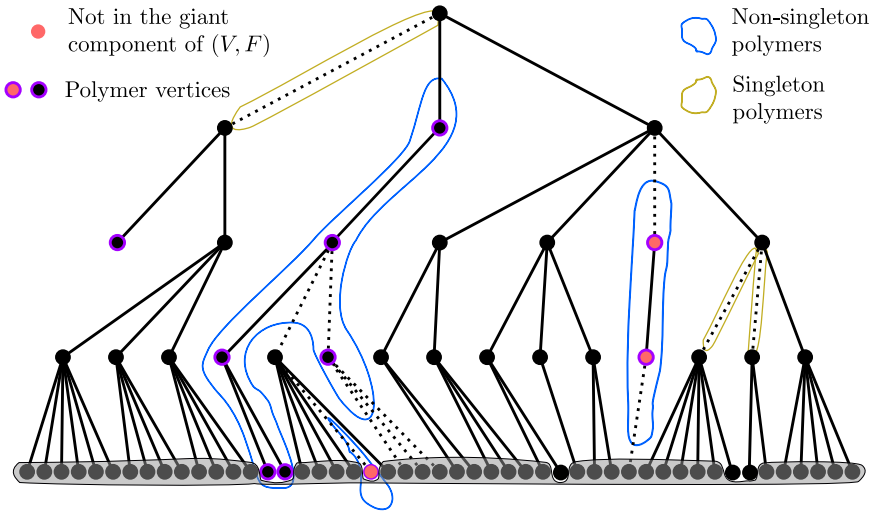


Figure 1.9: Illustration of the ordered polymers used for $\mathcal{G}(n, d/n)$. Vertices outside the giant component of in-edges are highlighted in pink. Each connected component (in G) of such vertices, together with all edges incident to it, forms a non-singleton polymer, outlined in blue. Any remaining out-edge forms a singleton polymer, outlined in yellow. In particular, vertices outside the giant component are always contained in a polymer, including in regions containing low-degree vertices.

it contains at least r vertices, and hence the corresponding polymer has size at least r. It thus remains to show that large polymers are sufficiently unlikely.

As in the random-regular-graph setting, we bound the probability of a large polymer using a relaxed variant of the Kotecký–Preiss condition. That is, we define weights of our ordered polymers that upper bound the probability of a particular polymer appearing in an ordered configuration. Then we establish sufficient expansion of large sets in $\mathcal{G}(n, d/n)$ to show that all polymers of size at least some $\frac{A \log n}{d}$ have weights decaying exponentially in their size. Together with a bound for the number of possible connected sets in a typical instance of $\mathcal{G}(n, d/n)$, we obtain that the probability of any polymer of size $r \geq \frac{A \log n}{d}$ appearing decays as $q^{-\Omega(r)}$. For a well-chosen value of r and all $q \geq d^{\Omega(d)}$, this is at most $\frac{1}{100|E|}$, as required.

Crucially, unlike standard polymer-based algorithms, our argument does not require weights to decay exponentially for polymers with fewer than some $\frac{A \log n}{d}$ vertices. As discussed earlier, these small polymers are precisely where the poor expansion of $\mathcal{G}(n, d/n)$ causes the usual polymer arguments to fail. By requiring the bound only for larger polymers, we are therefore able to overcome this obstruction.

1.3.2.4 WSM within the disordered phase

We conclude the proof overview of WSM by considering the disordered phase. Recall that the aim is to prove, for any $v \in V$ and any edge e incident to v, that

$$|\pi_G^{\mathrm{dis}}(e \in \cdot) - \pi_G(e \in \cdot \mid \cdot \text{ contains no edges of } E \setminus E(B_r(v)))| \leq \tfrac{1}{100|E|}, \qquad (1.5)$$

We call the RC distribution conditional on all edges outside of the ball being out (the second distribution) the *free* RC distribution on the ball.

The argument in Chapters 4 and 6 uses a revealing process, similar to our proof for the ordered phase. Unlike the ordered phase, where we aimed for a wired boundary, a typical configuration in the disordered phase does not have a giant component and instead, all components are typically small. Therefore, we use a different stopping condition for the process: the *free boundary*, where all boundary vertices (in the same terminology as in the previous section) are in separate components of revealed in-edges.

Here, we can use a simple iterative process to reveal $(\boldsymbol{F}^{\mathrm{dis}}, \boldsymbol{F}^{-})$ (where $\boldsymbol{F}^{\mathrm{dis}} \sim \pi_G^{\mathrm{dis}}$ and $\boldsymbol{F}^{-}$ is the free RC distribution on the ball): as long as the process has not obtained free boundary, it picks a boundary vertex that is not in a separate component and reveals all of its incident edges. We show that this process terminates prior to revealing e if the component of v in $(V, \boldsymbol{F}^{\mathrm{dis}})$ is completely contained in $B_{r-1}(v)$. As a result, the two configurations agree on e if all components of in-edges in $(V, \boldsymbol{F}^{\mathrm{dis}})$ are of size at most $r-1$.

This condition is enough for Theorems 1.1 and 1.3. In these settings, we view the connected components of in-edges as polymers and use the relaxed variant of the Kotecký–Preiss condition to show that a component of size at least r occurs with sufficiently small probability. In Chapter 4 we adapt the result from [49]. In Chapter 6, we use the (weak) expansion of $\mathcal{G}(n, d/n)$ to obtain that the probability of a polymer (connected component) with size at least some $\frac{A \log n}{d}$ decays exponentially in its size.

To obtain WSM within the disordered phase for all $q \geq 3$, in the proof of Theorem 1.2 we instead adapt an argument of Blanca and Gheissari [10], which bounds the difference in marginals in (1.5) by a sum of two terms. The first term is the probability that, after revealing the edges outside $B_r(v)$, the number of boundary vertices that are connected to another boundary vertex by the revealed in-edges is more than some constant. We use the planted model to explore the components containing the

boundary vertices to show that with sufficiently high probability, all but a constant number of the boundary vertices belong to distinct components. In turn, this bounds the first term.

The second term captures the probability that, roughly,[29] there are two disjoint paths of in-edges from the boundary of the ball towards v. For this term, we use the argument of Blanca and Gheissari [10], which bounds the probability of this two-path event when the ball is locally treelike and only a constant number of boundary vertices fail to belong to distinct components. Together, the two bounds give (1.5), and hence WSM within the disordered phase.

1.3.2.5 From WSM to approximation: proof outline of Theorem 1.4

We conclude the introduction by providing a brief proof outline of Theorem 1.4. Fix $R > 0$. Our aim is to obtain a $(1 + n^{-R})$-approximation to the partition function in time $n^{1+O(1/R)}$. Recall that the partition function $Z_G(q, \beta)$ is the sum of the weights of all RC configurations and serves as the normalising constant of the RC Gibbs distribution. Since we know that the configurations outside of a phase are very unlikely, it suffices to take into account only the configurations belonging to a phase. We consider each phase separately. For a phase $P \in \{\mathrm{dis}, \mathrm{ord}\}$, we let $Z_G^P(q, \beta)$ denote the *partition function within the phase* P, obtained by summing only over configurations in Ω_G^P. Next we outline how to obtain an approximation to the partition function of a given phase, $Z_G^P(q, \beta)$. The approximation to the overall partition function is then the sum of the two partition functions.

Approximating the partition functions within each phase is also useful for another reason: we need their relative proportions to choose the initial mixture for the RC dynamics when the phases coexist.

The main idea is to express the partition function in terms of edge marginals within the phase, and then use WSM within the phase to approximate these marginals by the marginals of the edges on a free or a wired ball (depending on the phase). We next explain it in more detail.

First of all, we require a stronger quantitative version of WSM within the phase. Instead of requiring the difference between the marginals $\pi^{g_P}_{B_r(v)}(e \in \cdot)$ and $\pi^P_G(e \in \cdot)$ to be at most $\frac{1}{100|E|}$, we require it to decay exponentially with the radius r, namely as $q^{-r/C}$ for some constant C. The arguments outlined above establish this stronger bound for $r = \Omega(\log n/d)$, provided that C is sufficiently large.

[29] The precise definition is rather technical, and thus we refer the reader to Section 5.3 for full details.

Next, rather than approximating Z_G^P directly, we obtain an (additive) approximation to its logarithm. We use the identity

$$\frac{\partial}{\partial\beta}\log Z_G^P(q,\beta) = \frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1}\sum_{e\in E}\pi_G^P(e\in\,\cdot\,).$$

Next, we approximate the edge marginals $\pi_G^P(e\in\,\cdot\,)$ by the marginal distribution on e when the outside of a radius-r ball is conditional to be the ground state of the phase, $\pi_{B_r(v)}^{g_P}(e\in\,\cdot\,)$. By the (stronger variant of) WSM within the phase, this introduces additive error at most $|E|\frac{\mathrm{e}^\beta}{\mathrm{e}^\beta-1}q^{-r/C}$. We choose $r=\frac{A\log n}{d}$ and $q=d^{\Omega(d)}$ large enough, so that this additive error is a sufficiently small inverse polynomial (note that the lower bound on q depends on the required precision).

For the required precision n^{-R}, we choose d and then q sufficiently large so that the radius $r=\frac{A\log n}{d}$ required above gives neighbourhoods of size $n^{O(1/R)}$. Moreover, with high probability, every such neighbourhood can be made into a tree by removing at most one edge. We can therefore compute the local marginals $\pi_{B_r(v)}^{g_P}(e\in\,\cdot\,)$ using dynamic programming in time polynomial in the size of the ball, and hence in time $n^{O(1/R)}$. Roughly, after conditioning on the edges of the single possible cycle, the remaining graph is a tree, where the required partition functions and marginals can be computed inductively from the leaves towards e.

Crucially, the dynamic programming can give the resulting marginals as explicit partial fractions of e^β. Substituting these expressions into the identity above gives an explicit approximation to $\frac{\mathrm{e}^\beta}{\mathrm{e}^\beta-1}\pi_{B_r(v)}^{g_P}(e\in\,\cdot\,)$. We then use symbolic integration as a black box to integrate each of these $|E|$ terms separately. Rather than evaluating the integrand numerically at many different values of β, symbolic integration uses its explicit representation to evaluate the integral in time $\mathrm{poly}(B_r(v)) = n^{O(1/R)}$. By summing these terms for each edge, this gives an additive approximation to $\log Z_G^P$; exponentiating it gives the required multiplicative approximation to Z_G^P.

Lastly, we conclude that this indeed gives us the required running time. W.h.p. $G\sim\mathcal{G}(n,d/n)$ has $O(n)$ edges. For each edge, the dynamic programming algorithm and subsequent symbolic integration take time $n^{O(1/R)}$, resulting in the claimed $n^{1+O(1/R)}$ running time.

Chapter 2

Preliminaries

In this chapter, we provide the background, standard definitions, and results referred to throughout this thesis.

2.1 Notation and graph definitions

When the undirected graph G is clear from context, we use V to denote its vertex set and E to denote its edge set. Throughout this work, let $n = |V|$. Unless stated otherwise, we assume that G is connected.

Note that this assumption is justified, since the Gibbs distribution (both for the Potts and the RC models) on a graph with multiple components is a product distribution of the Gibbs distributions on the individual components. Thus it is enough to consider each component separately, which is the case for $G \sim \mathcal{G}(n, d/n)$ (see Chapter 6 for more details).

Inside the individual chapters, the notation varies slightly. We refer the reader to the List of Notation at the beginning of the thesis for an overview of the main notation throughout the individual chapters.

2.2 Probability and algorithmic preliminaries

In this section, we give some standard notions and probabilistic bounds used throughout the thesis. We first recall the definition of total-variation distance, which we use throughout the thesis to measure the distance between probability distributions.

Definition 2.1. *For two probability distributions μ_1, μ_2 on a finite state space Ω, their* total-variation (TV) distance *is*

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} \coloneqq \frac{1}{2} \sum_{x \in \Omega} |\mu_1(x) - \mu_2(x)| = \max_{A \subseteq \Omega} |\mu_1(A) - \mu_2(A)|.$$

We slightly abuse notation and write $\|\boldsymbol{X}_1 - \boldsymbol{X}_2\|_{\mathrm{TV}}$ for the TV distance between the distributions of random variables $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, and write $\|\boldsymbol{X} - \mu\|_{\mathrm{TV}}$ for the TV distance between the distribution of $\boldsymbol{X}$ and a distribution μ. If $\|\boldsymbol{X} - \mu\|_{\mathrm{TV}} \leq \varepsilon$, we say that $\boldsymbol{X}$ is an *ε-sample* from μ.

We also use an alternative characterisation of total-variation distance in terms of couplings. Recall that a *coupling* of two[1] distributions μ_1, μ_2 is a pair of random variables $(\boldsymbol{X}_1, \boldsymbol{X}_2)$ defined on the same probability space such that $\boldsymbol{X}_1 \sim \mu_1$ and $\boldsymbol{X}_2 \sim \mu_2$; we write $(\boldsymbol{X}_1, \boldsymbol{X}_2) \sim (\mu_1, \mu_2)$. The total-variation distance between μ_1 and μ_2 is exactly the minimum disagreement probability over all possible couplings, i.e.,

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} = \min_{(\boldsymbol{X}_1, \boldsymbol{X}_2) \sim (\mu_1, \mu_2)} \mathbb{P}(\boldsymbol{X}_1 \neq \boldsymbol{X}_2).$$

We note that there is a coupling with disagreement probability equal to the total-variation distance. We call this a *optimal coupling* [58, Proposition 4.7].

In particular, for any coupling $(\boldsymbol{X}_1, \boldsymbol{X}_2) \sim (\mu_1, \mu_2)$,

$$\|\mu_1 - \mu_2\|_{\mathrm{TV}} \leq \mathbb{P}(\boldsymbol{X}_1 \neq \boldsymbol{X}_2).$$

We also use the notion of *stochastic dominance*. A distribution μ_1 stochastically dominates μ_2 if there is a coupling $(\boldsymbol{X}_1, \boldsymbol{X}_2) \sim (\mu_1, \mu_2)$ such that $\boldsymbol{X}_1 \leq \boldsymbol{X}_2$. We may also use stochastic dominance for distributions on sets, in which case we replace the inequality with subset inclusion ($\boldsymbol{X}_1 \subseteq \boldsymbol{X}_2$).

Next, we formally define the notions of approximate sampling and approximation algorithms used throughout the thesis.

Definition 2.2. *Let π_x be a distribution parametrised by an input x. A* fully polynomial almost-uniform sampler (FPAUS) *is a randomised algorithm[2] which takes the input x and an accuracy parameter $\varepsilon \in (0, 1)$ and returns, in time poly$(|x|, 1/\varepsilon)$, an ε-sample from π_x.*

Let $Z(x)$ be a non-negative function. A fully polynomial randomised approximation scheme (FPRAS) *for $Z(x)$ is a randomised algorithm which takes as its input x and an accuracy parameter $\varepsilon > 0$, runs in time poly$(|x|, 1/\varepsilon)$, and returns a random variable $\hat{\boldsymbol{Z}}$ satisfying*

$$\mathbb{P}\left((1-\varepsilon)Z(x) \leq \hat{\boldsymbol{Z}} \leq (1+\varepsilon)Z(x)\right) \geq \frac{3}{4}.$$

[1] Analogously, we can define a coupling of three or more random variables.

[2] Therefore, its output is a random variable.

The constant 3/4 is arbitrary and can be replaced by any constant strictly greater than $\frac{1}{2}$. Moreover, by independently repeating the algorithm $O(\log \frac{1}{\delta})$ times and taking the median of the outputs, the success probability can be amplified to $1 - \delta$.

2.2.1 Concentration bounds

Throughout this work, we repeatedly use the following standard concentration bounds (see e.g. [62, Chapters 3-5]).

Lemma 2.3 (Union bound)**.** *For any events $A_1, \ldots, A_k$ it holds that,*

$$\mathbb{P}\left(\bigcup_{i=1}^{k} A_i\right) \leq \sum_{i=1}^{k} \mathbb{P}(A_i).$$

Lemma 2.4 (Markov's inequality)**.** *Let $\boldsymbol{X}$ be a non-negative random variable. Then, for every $a > 0$,*

$$\mathbb{P}(\boldsymbol{X} \geq a) \leq \frac{\mathbb{E}[\boldsymbol{X}]}{a}.$$

In particular, we use Markov's inequality for the special case where $\boldsymbol{X}$ is a sum of indicator random variables and $a = 1$. Then, the inequality becomes $\mathbb{P}(\boldsymbol{X} > 0) \leq \mathbb{E}[\boldsymbol{X}]$, so we use it to show that with sufficiently high probability none of the adverse events happens (this can also be thought as a variant of the union bound).

Lemma 2.5 (Chebyshev's inequality)**.** *Let $\boldsymbol{X}$ be a random variable with finite variance. Then, for every $a > 0$,*

$$\mathbb{P}(|\boldsymbol{X} - \mathbb{E}[\boldsymbol{X}]| \geq a) \leq \frac{\mathrm{Var}[\boldsymbol{X}]}{a^2}.$$

Lemma 2.6 (Chernoff bound)**.** *Let $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ be independent Bernoulli random variables, let $\boldsymbol{X} = \sum_{i=1}^{n} \boldsymbol{X}_i$, and let $\mu = \mathbb{E}[\boldsymbol{X}]$. Then, for every $\delta > 0$,*

$$\mathbb{P}(\boldsymbol{X} \geq (1+\delta)\mu) \leq \left(\frac{\mathrm{e}^{\delta}}{(1+\delta)^{1+\delta}}\right)^{\mu}.$$

Moreover, for $0 < \delta < 1$,

$$\mathbb{P}(\boldsymbol{X} \leq (1-\delta)\mu) \leq \left(\frac{\mathrm{e}^{-\delta}}{(1-\delta)^{1-\delta}}\right)^{\mu},$$

and

$$\mathbb{P}(\boldsymbol{X} \geq (1+\delta)\mu) \leq \mathrm{e}^{-\mu\delta^2/3}, \qquad \mathbb{P}(\boldsymbol{X} \leq (1-\delta)\mu) \leq \mathrm{e}^{-\mu\delta^2/2}.$$

Since a binomial random variable with parameters n and p has the same distribution as a sum n of i.i.d.[3] Bernoulli(p), we use this bound for $\boldsymbol{X} \sim \mathrm{Bin}(n,p)$ with $\mu = np$.

We also use the following form of a Chernoff bound for the Poisson distribution.

Lemma 2.7 (Poisson Chernoff bound)**.** *Let $\boldsymbol{X} \sim$ Poisson(λ). Then, for every $0 < t < \lambda$,*

$$\mathbb{P}(\boldsymbol{X} \leq \lambda - t) \leq \exp\left(-\frac{t^2}{\lambda + t}\right).$$

Moreover, for every $t > 0$,

$$\mathbb{P}(\boldsymbol{X} \geq \lambda + t) \leq \exp\left(-\frac{t^2}{\lambda + t}\right).$$

2.3 Potts model and random-cluster model

We next include a definition for the Potts and random-cluster models.

Definition 2.8 (Potts model)**.** *For an integer $q \geq 2$, a real $\beta > 0$ and a graph $G = (V, E)$, let $\Omega_{G,q}$ be the set of all q-colourings of V (also known as* Potts configurations*). The* weight *of a Potts configuration $\sigma \in \Omega_{G,q}$ is $w_G(\sigma; q, \beta) \coloneqq \mathrm{e}^{\beta m(\sigma)}$, where $m(\sigma)$ is the number of monochromatic edges under σ. The* partition function *is $Z_G(q, \beta) \coloneqq \sum_{\sigma \in \Omega_{G,q}} w_G(\sigma; q, \beta)$. The (Potts) Gibbs distribution is $\mu_G(\sigma; q, \beta) \coloneqq \frac{w_G(\sigma;q,\beta)}{Z_G(q,\beta)}$.*

For $q = 2$, the model is known as the Ising model*.*

Definition 2.9 (Random-cluster model (RCM))**.** *For real parameters $q, \beta > 0$ and a graph $G = (V, E)$, let $\Omega_G = \{F : F \subseteq E\}$ be the set of random-cluster (RC) configurations. The* weight *of $F \in \Omega_G$ is $w_G(F; q, \beta) \coloneqq q^{c(F)}(\mathrm{e}^{\beta} - 1)^{|F|}$, where $c(F)$ is the number of connected components of the graph (V, F). The* partition function *is $Z_G(q, \beta) \coloneqq \sum_{F \in \Omega_G} w_G(F; q, \beta)$, and the* (RC) Gibbs distribution *is $\pi_G(F) = \pi_G(F; q, \beta) \coloneqq \frac{w_G(F;q,\beta)}{Z_G(q,\beta)}$.*

For a configuration $F \in \Omega_G$ we say that the edges in F are in-edges *(or* in*) and the edges in $E \setminus F$ are* out-edges *(or* out*).*

We also define the necessary notation for the phases. The disordered phase dis consists of all configurations with at most $\eta_{\mathrm{dis}}|E|$ in-edges. The ordered phase ord consists of all configurations with at most $\eta_{\mathrm{ord}}|E|$ out-edges. We do not give the values of $\eta_{\mathrm{dis}}, \eta_{\mathrm{ord}}$ here, as they depend on the parameter ranges and the graph model and therefore vary between the individual chapters

[3]Independent and identically distributed.

Definition 2.10. *For a phase $P \in \{\mathrm{ord}, \mathrm{dis}\}$, Ω_G^P denotes the set of configurations in the phase P, and $\pi_G^P \coloneqq \pi_G(\,\cdot \mid \cdot \in \Omega_G^P)$ denotes the* Gibbs distribution conditional to the phase P *and $Z_G^P \coloneqq \sum_{F \in \Omega_G^P} w_G(F; q, \beta)$ denotes the* partition function restricted to the phase P.

We also define the *ground states* of the phases and the associated wired/free distributions used in the proof of Theorem 1.5.

Definition 2.11. *The* ground state *of the disordered phase is the* all-out *configuration $\emptyset$. The ground state of the ordered phase is the* all-in *configuration E. For a phase $P \in \{\mathrm{dis}, \mathrm{ord}\}$, we denote its ground state by g_P.*

For an integer radius $r \geq 0$ and a vertex $v \in V$, the (RC) Gibbs distribution on the wired ball, *$\pi^+_{B_r(v)} \coloneqq \pi_G(\,\cdot \mid E \setminus E(B_r(v)) \subseteq \cdot\,)$, is the RC distribution conditional on the edges outside of the ball being in. The (RC) Gibbs distribution on the* free ball, *$\pi^-_{B_r(v)}(\,\cdot\,) \coloneqq \pi_G(\,\cdot \mid E \setminus E(B_r(v)) \subseteq E \setminus \cdot\,)$, is the RC distribution conditional on the edges outside of the ball being out.*

When $q \geq 2$ is an integer, the random-cluster model is an alternative representation of the Potts/Ising model and their partition functions coincide. The connection is through the standard Edwards-Sokal coupling, which we outline below (see also [47]).

Lemma 2.12 (Edwards-Sokal coupling)**.** *Let $q \geq 2$ be an integer, $\beta > 0$ a real, and $G = (V, E)$ a graph. Then, the partition functions of the random-cluster model and the Potts model are identical.*

Moreover, one can obtain a Potts sample from an RC sample $\boldsymbol{F} \sim \pi_G(\,\cdot\,; q, \beta)$ by colouring each connected component of (V, F) independently and uniformly with one of the q colours, and one can obtain an RC sample from a Potts sample $\boldsymbol{\sigma} \sim \mu_G(\,\cdot\,; q, \beta)$ by including each monochromatic edge independently with probability $1 - \mathrm{e}^{-\beta}$.

Proof. Consider a joint distribution ν on $\Omega_{G,q} \times \Omega_G$, where the probability of a pair (σ, F) is proportional to the weight

$$w(\sigma, F) = \mathbb{1}\{\text{all edges in } F \text{ are monochromatic under } \sigma\} \cdot (\mathrm{e}^{\beta} - 1)^{|F|}.$$

The idea is to show that its two marginal distributions are precisely the Potts and RC Gibbs distributions, while its conditional distributions give the two sampling procedures in the statement.

First, we show that for any $\sigma \in \Omega_{G,q}$, $w_G(\sigma; q, \beta) = \sum_{F \in \Omega_G} w(\sigma, F)$ and for any $F \in \Omega_G$, $w_G(F; q, \beta) = \sum_{\sigma \in \Omega_{G,q}} w(\sigma, F)$. Assuming these two equalities, we

get the first part of the lemma by noting that the partition function of the Potts model is exactly $\sum_{\sigma\in\Omega_{G,q}} w_G(\sigma;q,\beta)$ and the partition function of the random-cluster model is exactly the sum of all $\sum_{F\in\Omega_G} w_G(F;q,\beta)$. Summing either equality over the remaining configurations shows that the two partition functions are equal.

Next we establish the equalities. First, let $\sigma\in\Omega_{G,q}$ be any colouring. Then for any F with $w(\sigma,F)\neq 0$, it must be the case that F is a subset of monochromatic edges of σ. There are $m(\sigma)$ such edges and therefore precisely $\binom{m(\sigma)}{k}$ of their subsets F with size k. It follows from the binomial theorem that $\sum_{F\subseteq E} w(\sigma,F)=\sum_{k=0}^{m(\sigma)}\binom{m(\sigma)}{k}(\mathrm{e}^{\beta}-1)^k=\mathrm{e}^{\beta m(\sigma)}$. For the other equality, we note that in order for a colouring to have all edges in F monochromatic, vertices in each connected component of (V,F) must have the same colour. There are $c(F)$ components of (V,F) and therefore $q^{c(F)}$ ways to colour them. Thus, we obtain that $\sum_{\sigma\in\Omega_{G,q}} w(\sigma,F)=(\mathrm{e}^{\beta}-1)^{|F|}q^{c(F)}$.

We next consider the marginals of the joint distribution ν. For any $\sigma\in\Omega_{G,q}$, its marginal probability under ν is

$$\frac{\sum_{F\in\Omega_G} w(\sigma,F)}{\sum_{\sigma'\in\Omega_{G,q}}\sum_{F\in\Omega_G} w(\sigma',F)}=\frac{w_G(\sigma;q,\beta)}{Z_G(q,\beta)}=\mu_G(\sigma;q,\beta),$$

where we used the first equality above. Thus, the marginal distribution of σ under ν is precisely the Potts Gibbs distribution. Analogously, using the second equality above, the marginal distribution of F under ν is

$$\frac{w_G(F;q,\beta)}{Z_G(q,\beta)}=\pi_G(F;q,\beta),$$

that is, precisely the RC Gibbs distribution.

We now prove the second part of the lemma by showing that both procedures produce this joint distribution ν. Consider then a random-cluster sample $\boldsymbol{F}\sim\pi_G(\,\cdot\,;q,\beta)$ and the resulting colouring $\boldsymbol{\sigma}(\boldsymbol{F})$ obtained by colouring each of the $c(\boldsymbol{F})$ components independently with one of the q colours. Each of the $q^{c(\boldsymbol{F})}$ colourings of the components is equally likely, and σ can be obtained from $\boldsymbol{F}$ this way if and only if all edges of $\boldsymbol{F}$ are monochromatic under σ. Therefore, for any $F\in\Omega_G$ and $\sigma\in\Omega_{G,q}$,

$$\begin{aligned}&\mathbb{P}((\boldsymbol{F},\boldsymbol{\sigma}(\boldsymbol{F}))=(F,\sigma))\\&\qquad=\frac{(\mathrm{e}^{\beta}-1)^{|F|}q^{c(F)}}{Z_G(q,\beta)}\cdot\frac{\mathbb{1}\{\text{all edges in } F \text{ are monochromatic under } \sigma\}}{q^{c(F)}},\end{aligned}$$

which is exactly the probability of (σ,F) under the joint distribution ν. Therefore, the marginal distribution of σ is exactly the Potts Gibbs distribution.

For the remaining direction, consider a Potts sample $\boldsymbol{\sigma} \sim \mu_G(\,\cdot\,; q, \beta)$ and an edge subset $\boldsymbol{F}(\boldsymbol{\sigma})$ obtained by including each monochromatic edge independently with probability $1-\mathrm{e}^{-\beta}$. For any F that is a subset of monochromatic edges of a q-colouring σ, the probability of obtaining F from σ in this way is $(1-\mathrm{e}^{-\beta})^{|F|}\mathrm{e}^{-\beta(m(\sigma)-|F|)} = (\mathrm{e}^{\beta}-1)^{|F|}\mathrm{e}^{-\beta m(\sigma)}$. Thus, the probability of any pair (σ, F) in the resulting joint distribution $(\boldsymbol{\sigma}, \boldsymbol{F}(\boldsymbol{\sigma}))$ satisfies

$$\begin{aligned}\mathbb{P}((\boldsymbol{\sigma}, \boldsymbol{F}(\boldsymbol{\sigma})) &= (\sigma, F)) \\ &= \frac{\mathrm{e}^{\beta m(\sigma)}}{Z_G(q,\beta)} \cdot \frac{\mathbb{1}\{\text{all edges in } F \text{ are monochromatic under } \sigma\}(\mathrm{e}^{\beta}-1)^{|F|}}{\mathrm{e}^{\beta m(\sigma)}},\end{aligned}$$

which is precisely $\nu(\sigma, F)$. Therefore, the marginal distribution of F is exactly the RC distribution, which concludes the proof. □

2.4 Markov chains and mixing times

Next, we formally define the RC dynamics Markov chain. For this definition, recall that an edge e is a *cut-edge* (or *bridge*) of a graph if removing e increases the number of connected components.

Definition 2.13 (RC Glauber dynamics)**.** *Let $G = (V, E)$ be a graph and $q, \beta > 0$ be real parameters. The* Glauber dynamics *(or the* RC dynamics*) is a discrete random process $(X_t)_{t\geq 0}$ such that for all $t \geq 0$, $X_t \subseteq E$ is a random-cluster configuration, and the 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 in $X_t \cup \{e\}$ then:*
   (a) *With probability $\hat{p} \coloneqq \frac{1-\mathrm{e}^{-\beta}}{(q-1)e^{-\beta}+1}$, set $X_{t+1} \coloneqq X_t \cup \{e\}$*
   (b) *With all remaining probability, set $X_{t+1} = X_t \setminus \{e\}$*
3. *Otherwise,*
   (a) *With probability $p \coloneqq 1-\mathrm{e}^{-\beta}$, set $X_{t+1} \coloneqq X_t \cup \{e\}$*
   (b) *With all remaining probability, set $X_{t+1} = X_t \setminus \{e\}$*

We use the following *monotonicity* property of the random-cluster model for all $q \geq 1$ and $\beta > 0$.

Lemma 2.14. *Let $q \geq 1$ and $\beta > 0$ be reals, and let $G = (V, E)$ be a graph. Let $(X_t)_{t\geq 0}, (X'_t)_{t\geq 0}$ be two copies of the RC dynamics, with $X_0 \subseteq X'_0$, that choose the same edge and use the same uniform random variable to make their decisions in each step. Then, for all $i \geq 0$, it holds that $X_i \subseteq X'_i$.*

Proof. It is enough to show that $X_0 \subseteq X'_0$ implies $X_1 \subseteq X'_1$, as the result then follows by induction.

Let e be the edge chosen for the update. Since both chains use the same uniform random variable to make decisions, it is enough to show that the probability of e being included in X_1 is at most the probability it is included in X'_1.

From the assumption $X_0 \subseteq X'_0$, we obtain that (V, X_0) is a subgraph of (V, X'_0). In particular, whenever e is a cut-edge in (V, X'_0), it is also a cut-edge in (V, X_0).

Since the probability of including the chosen edge in the next step only depends on whether it is a cut-edge, the only case where the transition probabilities are not equal is when e is a cut-edge in (V, X_0) but it is not a cut-edge in (V, X'_0). Then, the probability of including e in X_1 is $\hat{p} = \frac{1-\mathrm{e}^{-\beta}}{(q-1)\mathrm{e}^{-\beta}+1}$ and the probability of including e in X'_1 is $p = 1 - \mathrm{e}^{-\beta}$. It remains to verify that $p \geq \hat{p}$: this is true whenever $(q-1)\mathrm{e}^{-\beta} + 1 \geq 1$, which holds for all $q \geq 1$. □

It is a classic result (see e.g. [47]) that the Glauber dynamics converges to the RC distribution. The *mixing time* is the number of steps needed for convergence, in terms of the size of the graph G. Formally, we define it as follows.

Definition 2.15 (Mixing time)**.** *We say that the* mixing time *of the Glauber dynamics initialised from $\boldsymbol{X_0}$ is:*

$$T_{\mathrm{mix}}(\boldsymbol{X}_0) \coloneqq \min\{T : \|X_T - \pi_G\|_{\mathrm{TV}} \leq \tfrac{1}{4}\}$$

The worst-case mixing time *is $T_{\mathrm{mix}} \coloneqq \max_{X_0 \in \Omega_G} T_{\mathrm{mix}}(X_0)$.*

We say that Glauber dynamics mixes rapidly *(when initialised from $\boldsymbol{X_0}$) if its mixing time (from $\boldsymbol{X_0}$) is polynomial in n.*

Note that while we defined the mixing time only up to a constant TV-distance, in order to obtain an FPAUS, we need to be able to sample with an arbitrary accuracy $\varepsilon > 0$. This is not a barrier: a standard probability amplification argument (see e.g. [58, Section 4.5]) implies that $O(T_{\mathrm{mix}} \log(1/\varepsilon))$ steps suffice to reach a TV distance of at most ε. Since our results establish mixing from a particular ground state rather than worst-case mixing, this standard argument does not apply directly. Instead,

we use monotonicity and Key Property 1 to obtain probability amplification for all $\varepsilon \geq \mathrm{e}^{-\Omega(|E|)}$. For $\varepsilon \leq \mathrm{e}^{-\Omega(|E|)}$, brute-force enumeration over the $2^{|E|}$ RC configurations is sufficient for the approximation guarantees required in this thesis, since then $2^{|E|} = \mathrm{poly}(1/\varepsilon)$. We give the full argument in Chapter 3.

We also remark that $O(n \log n)$ worst-case mixing time is asymptotically optimal for the Glauber dynamics. Until an edge is selected for an update, its state remains unchanged from the initial configuration. Since at each step an edge is chosen uniformly at random, after t steps a fixed edge has not been selected with probability $(1 - 1/|E|)^t$. Thus, it takes order $|E| \log |E|$ steps before all edges have had a chance to update with constant probability. Hence, for the sparse graphs considered in this thesis, where $|E| = \Theta(n)$, the natural lower-bound scale is $\Omega(n \log n)$.

A standard method for establishing worst-case mixing times is the *log-Sobolev* constant, which roughly quantifies the rate of entropy decay under the Markov chain. In particular, a lower bound on the log-Sobolev constant gives an upper bound on the mixing time, through the inequality stated below. For brevity, we write $\pi \coloneqq \pi_G(\,\cdot\,; q, \beta)$ in the rest of this section.

Definition 2.16 (Log-Sobolev constant)**.** *Let $g : \Omega_G \to \mathbb{R}_{\geq 0}$ be any function. Then the entropy of g under π is*

$$\mathrm{Ent}_\pi[g] = \mathbb{E}_{\boldsymbol{X}\sim\pi}[g(\boldsymbol{X})(\ln g(\boldsymbol{X}) - \ln(\mathbb{E}_{\boldsymbol{Y}\sim\pi}[g(\boldsymbol{Y})]))],$$

and the Dirichlet form *is*

$$\mathcal{E}(g, g) \coloneqq \frac{1}{2} \sum_{F_1, F_2 \in \Omega_G} \pi(F_1)\mathbb{P}(X_{t+1} = F_2 \mid X_t = F_1)(g(F_1) - g(F_2))^2.$$

Then the log-Sobolev constant *is*

$$\alpha(\pi) \coloneqq \inf_{\substack{g:\Omega_G \to \mathbb{R}_{\geq 0} \\ \mathrm{Ent}_\pi[g] \neq 0}} \frac{\mathcal{E}(\sqrt{g}, \sqrt{g})}{\mathrm{Ent}_\pi[g]}.$$

We use the following standard inequality as a black box throughout the thesis.

Lemma 2.17 ([10, Fact 6.1])**.** *For any $\gamma < \alpha(\pi)$ and any X_0,*

$$\|X_t - \pi\|_{\mathrm{TV}} \leq \mathrm{e}^{-\gamma t/2}\big(\log \tfrac{1}{\min_{F \in \Omega_G} \pi(F)}\big)^{1/2} \tag{2.1}$$

Chapter 3

Proof of the main strategy

In this chapter, we first provide the full proof of Theorem 1.5, including the proof of probability amplification in this setting. We then show how to obtain an FPRAS for the partition function using rapid mixing within the phase and conclude by providing the details for sampling when neither phase is known to dominate.

3.1 Mixing within the phase

In this section, we prove Theorem 1.5 for the random-cluster model. First, we establish a version of probability amplification, which subsequently allows us to only bound the number of steps until the TV distance is $\frac{1}{4}$. Then, we prove the theorem for the ordered phase, and finally, we comment on how to adapt the proof to the disordered phase and, more generally, for other monotone models.

Throughout the proofs, we use the following *monotone coupling* of two copies of the RC dynamics, $(X_t)_{t\geq 0}$ and $(\hat{X}_t)_{t\geq 0}$. For a phase $P \in \{\text{dis}, \text{ord}\}$, let $(X_t)_{t\geq 0}$ be initialised from the ground state of P, and $(\hat{X}_t)_{t\geq 0}$ be initialised from the RC distribution restricted to the phase P, π_G^P. In each step, both chains choose the same edge to update and use the same uniform random variable for their decisions. In addition, $\hat{X}_t$ is *restricted* to the phase. That is, whenever an update would cause $(\hat{X}_t)_{t\geq 0}$ to leave the phase P, $(\hat{X}_t)_{t\geq 0}$ ignores the update. It is straightforward to verify that π_G^P is the stationary distribution of $(\hat{X}_t)_{t\geq 0}$, and hence that for all t, $\hat{X}_t \sim \pi_G^P$. Thus the TV distance of $(X_t)_{t\geq 0}$ from π_G^P after T steps is at most $\mathbb{P}(X_T \neq \hat{X}_T)$.

This coupling is *monotone*: if $X_t \subseteq \hat{X}_t$ then the inclusion is preserved in the next step, provided the next update is not ignored. The same holds for the opposite inclusion ($X_t \supseteq \hat{X}_t$) (as shown in Lemma 2.14).

3.1.1 Probability amplification within the phase

The next lemma shows that the probability that X_t and $\hat{X}_t$ disagree decays exponentially, except for a small error accounting for the probability of leaving the phase.

Lemma 3.1. *Let $q, \beta > 0$ be real parameters, $P \in \{\mathrm{ord}, \mathrm{dis}\}$ a phase, and $\mathcal{G}$ a family of graphs such that for all $G \in \mathcal{G}$, $|E(G)| = \Omega(n)$ and $\pi_G(\,\cdot\,; q, \beta)$ has a local maximum in phase P (Key Property 1). Let $(X_t)_{t\geq 0}, (\hat{X}_t)_{t\geq 0}$ be the two copies of the RC dynamics under the monotone coupling defined above. Then, for every $T \geq 1$,*

$$\mathbb{P}(X_{2T} \neq \hat{X}_{2T}) \leq \mathbb{P}(X_T \neq \hat{X}_T)^2 + T\mathrm{e}^{-\Omega(n)}.$$

Proof. By conditioning on whether the chains disagree after T steps, we obtain

$$\begin{aligned}\mathbb{P}(X_{2T} \neq \hat{X}_{2T}) \leq\, &\mathbb{P}(X_{2T} \neq \hat{X}_{2T} \mid X_T = \hat{X}_T) \cdot \mathbb{P}(X_T = \hat{X}_T) \\ &+ \mathbb{P}(X_{2T} \neq \hat{X}_{2T} \mid X_T \neq \hat{X}_T) \cdot \mathbb{P}(X_T \neq \hat{X}_T).\end{aligned}$$

For the first quantity, we note that once the chains agree, they will remain coupled until $\hat{X}_t$ ignores an update. Key Property 1 bounds the probability of ignoring an update in a single step by $\mathrm{e}^{-\Omega(|E|)} \leq \mathrm{e}^{-n/C}$, where C is a constant. By the union bound, the probability of at least one update being ignored during the first $2T$ steps, and hence also the first term, is at most $2T\mathrm{e}^{-n/C}$.

For the second quantity, introduce a third coupled copy of the RC dynamics, $(X'_t)_{t\geq T}$, which uses the same monotone coupling, only starts from the T^{th} step and is initialised from the ground state of P (i.e. $X'_T := g_P$ and X'_t is undefined for $t < T$). By monotonicity, as long as no updates were ignored by time T, $\hat{X}_{2T} \subseteq X_{2T} \subseteq X'_{2T}$ (for the ordered phase) or $\hat{X}_{2T} \supseteq X_{2T} \supseteq X'_{2T}$ (for the disordered phase). Thus, conditional on no updates being ignored by time $2T$, the probability of the event $X_{2T} \neq \hat{X}_{2T}$ is at most the probability of $X'_{2T} \neq \hat{X}_{2T}$. As a consequence,

$$\begin{aligned}\mathbb{P}(X_{2T} \neq \hat{X}_{2T} \mid X_T \neq \hat{X}_T) &\leq 2T\mathrm{e}^{-n/C} + \mathbb{P}(X'_{2T} \neq \hat{X}_{2T} \mid \text{ no updates were ignored}) \\ &\leq \frac{4T\mathrm{e}^{-n/C}}{1 - 2T\mathrm{e}^{-n/C}} + \mathbb{P}(X'_{2T} \neq \hat{X}_{2T}),\end{aligned}$$

where, in the second inequality, we remove the conditioning on no updates being ignored. This event has probability at least $1 - 2T\mathrm{e}^{-n/C}$, so removing the conditioning introduces an additional error of at most $\frac{2T\mathrm{e}^{-n/C}}{1-2T\mathrm{e}^{-n/C}}$, giving the stated bound.

Finally, $\hat{X}_T$ has the same distribution as $\hat{X}_{2T}$ (specifically, π_G^P), and X'_{2T} has the same distribution as X_T (since it was only started at the time $t = T$).

We may assume that $T\mathrm{e}^{-n/C} \leq \frac{1}{4}$. Indeed, the lemma is trivial when choosing $4\mathrm{e}^{-n/C}$ as the $\mathrm{e}^{-\Omega(n)}$ term in the statement. Therefore, $2T\mathrm{e}^{-n/C} + \frac{4T\mathrm{e}^{-n/C}}{1-2T\mathrm{e}^{-n/C}} = T\mathrm{e}^{-\Omega(n)}$. We get the proof of the lemma by combining the above inequalities. □

3.1.2 Mixing within the ordered phase

We first prove fast mixing given the key properties for the ordered phase. The proof for the disordered phase is analogous, and we note the differences in the following section. We restate the Theorem 1.5 for convenience.

Theorem 1.5. *Let $q, \beta > 0$ be real parameters, $P \in \{\mathrm{ord}, \mathrm{dis}\}$ a phase, and $\mathcal{G}$ a family of graphs such that for all $G \in \mathcal{G}$, $|E(G)| = \Omega(n)$ and there is an integer radius $r \geq 1$ such that all of the following hold.*

- *$\pi_G(\,\cdot\,; q, \beta)$ has a local maximum in the phase P (Key Property 1),*
- *G has WSM within the phase P at radius r (Key Property 2), and*
- *the RC dynamics has local mixing in the phase P on radius-r neighbourhoods of G (Key Property 3).*

Then there exists $T = \mathrm{poly}(n)$, such that for all sufficiently large n, all real $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$, and any n-vertex graph $G \in \mathcal{G}$, the RC dynamics $(X_t)_{t\geq 0}$ initialised from the ground state g_P satisfies

$$\left\| X_{O(T \log \frac{1}{\varepsilon})} - \pi_G^P \right\|_{\mathrm{TV}} \leq \varepsilon.$$

If, moreover, on all graphs $G \in \mathcal{G}$, the RC dynamics has optimal local mixing *in the phase P at radius r, then the above holds with $T = O(|E| \log |E|)$.*

Proof of Theorem 1.5 for the ordered phase. By Lemma 3.1, it is enough to show that $\mathbb{P}(X_T \neq \hat{X}_T) \leq \frac{1}{4}$ for some $T = \mathrm{poly}(n)$ (or $T = O(n \log n)$ for optimal mixing). We consider each edge separately and use the union bound:

$$\mathbb{P}(X_T \neq \hat{X}_T) \leq \sum_{e \in E} \mathbb{P}(X_T \text{ and } \hat{X}_T \text{ disagree on } e).$$

Therefore, it is sufficient to prove that, for every $e \in E$, $\mathbb{P}(X_t, \hat{X}_t \text{ disagree on } e) \leq \frac{1}{4|E|}$. Throughout the remainder of the proof, fix an edge e and a vertex v incident to e. From Key Properties 2 and 3, there is an integer radius $r \geq 1$ such that G has

WSM within the ordered phase at radius r and the RC dynamics has local mixing in the ordered phase on radius-r neighbourhoods of G.

Let $\pi^+_{B_r(v)} := \pi_G(\cdot \mid \text{all edges outside of } B_r(v) \text{ are in})$ denote the RC distribution on the wired ball around v, as defined in Section 2.3. We next consider a third copy of the RC dynamics using the same monotone coupling as for $(X_t, \hat{X}_t)$, denoted $(X^v_t)_{t\geq 0}$, which is initialised from all-in and *ignores updates outside of the ball*. That is, whenever the monotone coupling chooses an edge outside of the ball $B_r(v)$, then X^v_t ignores this update.

Monotone coupling always ensures that for all t, $X_t \subseteq X^v_t$ (see Lemma 2.14). Also, as we noted at the beginning of the section, as long as $(\hat{X}_t)_{t\geq 0}$ has not ignored any updates by the time t, it holds that $\hat{X}_t \subseteq X_t$. Denote the event "$(\hat{X}_t)_{t\geq 0}$ has not ignored any updates by time t" by $\mathcal{E}_{\leq t}$. Therefore, conditional on $\mathcal{E}_{\leq t}$, it holds that $\hat{X}_t \subseteq X_t \subseteq X^v_t$, and hence whenever X_t and $\hat{X}_t$ disagree on e, so must $\hat{X}_t$ and X^v_t. It follows that

$$\begin{aligned}\mathbb{P}(X_T, \hat{X}_T \text{ disagree on } e) &\leq \mathbb{P}(\overline{\mathcal{E}_{\leq T}}) + \mathbb{P}(X^v_T, \hat{X}_T \text{ disagree on e} \mid \mathcal{E}_{\leq T}) \\ &\leq \mathbb{P}(\overline{\mathcal{E}_{\leq T}}) + |\mathbb{P}(e \in X^v_T \mid \mathcal{E}_{\leq T}) - \mathbb{P}(e \in \hat{X}_T \mid \mathcal{E}_{\leq T})| \\ &\leq 5\mathbb{P}(\overline{\mathcal{E}_{\leq T}}) + |\mathbb{P}(e \in X^v_T) - \mathbb{P}(e \in \hat{X}_T)|,\end{aligned}$$

where the second inequality follows from the fact that conditioned on $\mathcal{E}_{\leq T}$, the chains only disagree on e if $e \in X^v_T$ and $e \notin \hat{X}_T$. The third inequality follows from the standard fact that for any events A, B, it holds that $|\mathbb{P}(A|B) - \mathbb{P}(A)| \leq 2\mathbb{P}(\overline{B})$.

By further applying the triangle inequality and comparing with $\pi^+_{B_r(v)}$, we obtain that

$$\begin{aligned}\mathbb{P}(X_T, \hat{X}_T \text{ disagree on } e) &\leq 5\mathbb{P}(\overline{\mathcal{E}_{\leq T}}) \\ &\quad + |\mathbb{P}(e \in X^v_T) - \pi^+_{B_r(v)}(e \in \cdot)| + |\pi^+_{B_r(v)}(e \in \cdot) - \mathbb{P}(e \in \hat{X}_T)|. \quad (3.1)\end{aligned}$$

Combining Key Property 1 and the union bound gives us $5\mathbb{P}(\overline{\mathcal{E}}_{\leq T}) \leq 5T\mathrm{e}^{-\Omega(n)}$, and the WSM within the ordered phase then gives $|\pi^+_{B_r(v)}(e \in \cdot) - \mathbb{P}(e \in \hat{X}_T)| \leq \frac{1}{100|E|}$.

It remains to show, using local mixing, that $|\mathbb{P}(e \in X^v_T) - \pi^+_{B_r(v)}(e \in \cdot)|$ is sufficiently low for a suitable polynomial T. It is straightforward to verify that $\pi^+_{B_r(v)}$ is the stationary distribution for the chain $(X^v_t)_{t\geq 0}$, and thus in particular,

$$|\mathbb{P}(e \in X^v_T) - \pi^+_{B_r(v)}(e \in \cdot)| \leq \left\| X^v_T - \pi^+_{B_r(v)} \right\|_{\mathrm{TV}}.$$

Ignoring the steps in which an edge outside $B_r(v)$ is selected, $(X_t^v)_{t\geq 0}$ is exactly the RC dynamics on the wired ball. From local mixing we can assume there is $T_v = \text{poly}(n)$ such that after T_v updates *within the ball*, $(X_t^v)_{t\geq 0}$ gets within $|E|^{-3}$ total-variation distance from $\pi^+_{B_r(v)}$.[1]

We choose T large enough such that with high enough probability, at least T_v updates are within the ball. In particular, let

$$T = \max_{v\in V} \left\lceil 30T_v \tfrac{|E|}{|E(B_r(v))|} \log|E| \right\rceil = \text{poly}(n),$$

where we take the maximum over v to establish global mixing time. The number of updates within the ball is a binomial random variable with mean $\frac{|E(B_r(v))|}{|E|}\cdot T$. Hence, by the Chernoff bound, the probability that less than T_v updates are made within the ball is at most $\mathrm{e}^{-30T_v\log|E|/8} \leq |E|^{-3}$ for all n, and thus $|E|$, sufficiently large.

Substituting these bounds back to (3.1), we obtain that for this choice of T,

$$\mathbb{P}(X_T, \hat{X}_T \text{ disagree on } e) \leq 5T\mathrm{e}^{-\Omega(n)} + \tfrac{1}{100|E|} + \tfrac{2}{|E|^3}.$$

This is at most $\frac{1}{5|E|}$ for all n (and thus $|E|$) large enough. This concludes the proof for rapid mixing.

Next, we consider the case that G has optimal local mixing in the ordered phase. Then, there is a constant C such that the log-Sobolev constant for the RC dynamics on the wired ball $B_r(v)$ is strictly greater than $C/|E(B_r(v))|$. Here, we use the relation between the log-Sobolev constant and decay in TV-distance from (2.1) to obtain a $T = O(|E|\log|E|)$.

In particular, let $T := \lceil \max(30, 17/C)|E|\log|E|\rceil$. By the Chernoff bound, with probability $\geq 1-|E|^{-3}$, the number of updates in the ball by the time T is at least $\frac{3|E(B_r(v))|}{C}\log|E|$. Using (2.1), we obtain

$$|\mathbb{P}(e\in X_T^v) - \pi^+_{B_r(v)}(e\in\,\cdot\,)| \leq |E|^{-3} + \mathrm{e}^{-(7/2)\log|E|}\big(\log\tfrac{1}{\min_F \pi^+_{B_r(v)}(F)}\big)^{1/2}.$$

To finish bounding this term, we note that the weight of any configuration is at most $\max((\mathrm{e}^\beta-1),1)^{|E|}\cdot\max(q,1)^n$, thus for any F, $\pi^+_{B_r(v)}(F) \geq (2(\mathrm{e}^\beta-1))^{-|E|}q^{-n}$, and thus $\big(\log\frac{1}{\min_F \pi^+_{B_r(v)}(F)}\big)^{1/2} = O(|E|^{1/2})$. Plugging back to (3.1), we obtain that for this choice of T,

$$\mathbb{P}(X_T, \hat{X}_T \text{ disagree on } e) \leq 5T\mathrm{e}^{-\Omega(n)} + \tfrac{1}{100|E|} + O(\tfrac{1}{|E|^3}),$$

which is at most $\frac{1}{4|E|}$ for all $|E|$ large enough. □

[1]Note that by standard probability amplification (see [58]), if the worst-case mixing time of the chain is some T_{mix}, the dynamics reaches $|E|^{-3}$ TV distance from its stationary distribution in $O(T_{\text{mix}}\log|E|)$ steps.

Remark 3.2. *In Chapters 4–6, we will use a slightly stronger notion of a* wired ball. *Instead of conditioning all edges outside the ball to be in, we tie all boundary vertices (the vertices at distance exactly r from v) into a single component.*

If $G \setminus B_r(v)$ is connected, this is exactly the same notion of a wired ball. Otherwise, we can equivalently introduce artificial edges outside $B_r(v)$ that connect its boundary vertices. In the dynamics on G, these artificial edges can be viewed as fixed to be out, whereas in the wired local dynamics they are fixed to be in. Monotonicity (Lemma 2.14) therefore gives $X_t \subseteq X_t^v$, as required in the proof above.

We defined Key Properties 2 and 3 using the weaker notion in order to keep the formulation consistent between the two phases and readily extendable to other monotone models.

Remark 3.3. *The proof of Theorem 1.5 does not require the same radius r for every vertex. In particular, we can relax Key Properties 2 and 3 by requiring that, for every $v \in V$, there exists a radius r_v at which both WSM within the phase and the required local mixing hold (instead of there being the same radius for all v). Crucially, the same radius r_v must be used for both properties. The proof is unchanged, except that r is replaced by r_v in the corresponding definitions.*

3.1.3 Mixing within the disordered phase and extension to other models

We first explain how to adapt the above proof to the disordered phase. Most of the changes are simply replacing the notation, definitions, and key properties for the ordered phase by their disordered-phase counterparts. In particular, the initial state of $(X_t)_{t\geq 0}$ is now the all-out configuration, and the RC dynamics on the wired ball is replaced by the RC dynamics starting from all-out on the *free ball*, where the outside of the ball is conditioned to be out. The main change in the proof is that the inclusions in the monotone coupling are reversed: conditional on $\hat{X}_t$ having ignored no updates by time t, we have $X_t^v \subseteq X_t \subseteq \hat{X}_t$. With these changes, the remainder of the proof is identical.

More generally, the same proof applies to other monotone models. Consider a model with a phase P, an associated ground state g_P, and a partial order $\preceq$ on its configurations such that every configuration ω satisfies $\omega \preceq g_P$. We further require that the corresponding Glauber dynamics admits a monotone coupling: whenever $\omega_1 \preceq \omega_2$, the coupled updates preserve this ordering. For the random-cluster model,

this partial order is inclusion in the ordered phase and reverse inclusion in the disordered phase.

The proof above then applies with the partial order $\preceq$ in place of inclusion. Similarly, the dynamics on the wired/free ball is replaced by the Glauber dynamics on $B_r(v)$ with the configuration outside the ball fixed to the ground state g_P. Subject to the corresponding versions of the three key properties, the remainder of the argument is unchanged.

We conclude with a technical remark. We formulated the bounds in the definitions of the key properties, as well as in Theorem 1.5, in terms of the number of edges $|E|$, since the RC Glauber dynamics updates edges. For spin systems, where the Glauber dynamics updates the state of vertices instead of edges, it is more natural to formulate these bounds in terms of the number of vertices n. This only requires a small modification to the choice of T in the proof. In particular, we instead use $T = \max_{v \in V} \left\lceil 30 T_v \frac{n}{|V(B_r(v))|} \log n \right\rceil$. For the graph families considered in this thesis, it holds that $|E| = \Theta(n)$ and hence this distinction does not affect the asymptotic mixing time.

3.2 FPRAS for the partition function

We next present the *self-reducibility* technique (see also [9, 21]) for converting an FPAUS to an FPRAS. This is a crucial step for being able to sample at the critical temperature (or critical window), since there we need estimates of the proportions of the phases to obtain an appropriate initial distribution.

In particular, the next lemma shows that, provided we can sample within a phase $P \in \{\text{dis}, \text{ord}\}$ for all β in a suitable interval, we can then also obtain an approximation to the partition function of that phase. Note that the proof is a straightforward adaptation of the self-reducibility technique for the random-cluster model, and the key part of the proof is choosing the parameters appropriately to guarantee the polynomial running time.

Lemma 3.4. *Let $q \geq 1$ and $\beta > 0$ be reals and $\mathcal{G}$ be a class of graphs such that there is an FPAUS for sampling from $\pi_G^{\text{ord}}(\,\cdot\,; q, \beta')$ for all $G \in \mathcal{G}$ and $\beta' \geq \beta$. Then, there is an FPRAS for approximating the partition function restricted to the ordered phase, $Z_G^{\text{ord}}(q, \beta)$, for all $G \in \mathcal{G}$.*

Analogously, suppose that for all $G \in \mathcal{G}$, there is an FPAUS for sampling from the disordered phase $\pi_G^{\text{dis}}(\,\cdot\,; q, \beta')$ for all $G \in \mathcal{G}$ and $\beta' \leq \beta$. Then, there is an FPRAS

for approximating the partition function restricted to the disordered phase, $Z_G^{\mathrm{dis}}(q,\beta)$, for all $G \in \mathcal{G}$.

Proof. First, we prove the lemma for the ordered phase. Let $q \geq 1$, $\beta > 0$ and $\mathcal{G}$ satisfy the assumptions of the lemma. Then, assume that we can obtain, for any $G \in \mathcal{G}$, $\beta' \geq \beta$, and $\varepsilon' \in (0,1)$ an ε'-sample from $\pi_G^{\mathrm{ord}}(\,\cdot\,; q, \beta')$ in time $\mathrm{poly}(n, 1/\varepsilon')$. Now, suppose that we are given a graph $G = (V, E) \in \mathcal{G}$ with m edges and an accuracy parameter $\varepsilon > 0$, and we want to return, in time polynomial in n and $1/\varepsilon$, a $(1 \pm \varepsilon)$-approximation to $Z_G^{\mathrm{ord}}(q,\beta)$. We may assume that G has at least one edge and that $\varepsilon \leq 1$.

The high-level strategy to estimate $Z_G^{\mathrm{ord}}(q,\beta)$ is as follows. We pick a suitable $N = \mathrm{poly}(n, 1/\varepsilon)$ and a sequence of β's, $\beta = \beta_0 < \beta_1 < \cdots < \beta_N$ (to be specified later) whose consecutive values are sufficiently close, and express $Z_G^{\mathrm{ord}}(q,\beta)$ as a telescoping product as follows.

$$Z_G^{\mathrm{ord}}(q,\beta) = Z_G^{\mathrm{ord}}(q,\beta_N) \cdot \prod_{i \geq 0}^{N-1} \frac{Z_G^{\mathrm{ord}}(q,\beta_i)}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})}. \tag{3.2}$$

Next, we approximate the individual terms of the product. We pick β_N large enough so that $Z_G^{\mathrm{ord}}(q,\beta_N)$ is well-approximated by the weight of the all-in configuration $q(\mathrm{e}^{\beta_N} - 1)^m$. It therefore remains to approximate the ratios between the partition functions at consecutive values of β_i. The key observation is that each of these ratios can be expressed as an expectation under the RC distribution at β_{i+1}, and hence can be estimated using samples from this distribution.

$$\rho_i := \frac{Z_G^{\mathrm{ord}}(q,\beta_i)}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})} = \sum_{F \in \Omega_G^{\mathrm{ord}}} \frac{w_G(F; q, \beta_{i+1})(\frac{\mathrm{e}^{\beta_i}-1}{\mathrm{e}^{\beta_{i+1}}-1})^{|F|}}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})} \tag{3.3}$$

$$= \mathbb{E}_{\boldsymbol{F} \sim \pi_G^{\mathrm{ord}}(\,\cdot\,; q, \beta_{i+1})} \left[\left(\frac{\mathrm{e}^{\beta_i} - 1}{\mathrm{e}^{\beta_{i+1}} - 1} \right)^{|\boldsymbol{F}|} \right]. \tag{3.4}$$

We estimate ρ_i by $(\frac{\mathrm{e}^{\beta_i}-1}{\mathrm{e}^{\beta_{i+1}}-1})^{|\hat{\boldsymbol{F}}_i|}$, where $\hat{\boldsymbol{F}}_i$ is a sample obtained from the FPAUS. The idea is to choose small intervals between consecutive β_i's, ensuring that the variance of the quantity in (3.4) is small so we can use polynomially many samples from the FPAUS to estimate the mean.

Subsequently, we take the product over $0 \leq i \leq N-1$ to obtain an estimate of the telescoping product in (3.2) and finally average over $M = \mathrm{poly}(1/\varepsilon)$ independent repetitions. Here, N controls the number and spacing of the intermediate values β_i,

while M is the number of independent repetitions used to obtain the approximation with sufficiently high probability. Formally, the algorithm is as follows.

Algorithm 3.5 (FPRAS for $Z_G^{\mathrm{ord}}(q,\beta)$)**.**

1. *Let* $N := \left\lceil m \log\left(\frac{4\times 2^m \times q^n}{\varepsilon(\mathrm{e}^\beta - 1)}\right)\right\rceil$ *and* $M := \lceil \frac{128\mathrm{e}}{\varepsilon^2} \rceil$. *Then, for* $0 \le i \le N$, *let* β_i *satisfy* $\mathrm{e}^{\beta_i} - 1 := (\mathrm{e}^\beta - 1)\mathrm{e}^{i/m}$.

2. *For* $j = 1, \dots, M$

   (a) *For* $i = 0, \dots, N-1$

      i. *Obtain an* $\frac{1}{8NM}$*-sample from* $\pi_G^{\mathrm{ord}}(\,\cdot\,; q, \beta_{i+1})$, $\hat{\boldsymbol{F}}_{\boldsymbol{i}}^{(\boldsymbol{j})}$, *using the FPAUS.*

      ii. *Let* $\hat{\boldsymbol{\rho}}_{\boldsymbol{i}}^{(\boldsymbol{j})} := \left(\frac{\mathrm{e}^{\beta_i}-1}{\mathrm{e}^{\beta_{i+1}}-1}\right)^{|\hat{\boldsymbol{F}}_{\boldsymbol{i}}^{(\boldsymbol{j})}|}$.

   (b) *Let* $\hat{\boldsymbol{W}}^{(\boldsymbol{j})} := \prod_{i=0}^{N-1} \hat{\boldsymbol{\rho}}_{\boldsymbol{i}}^{(\boldsymbol{j})}$.

3. *Let* $\hat{\boldsymbol{W}} := \frac{1}{M}\sum_{j=1}^{M} \hat{\boldsymbol{W}}^{(\boldsymbol{j})}$.

4. *Return* $\hat{\boldsymbol{Z}} := q(\mathrm{e}^{\beta_N} - 1)^m \hat{\boldsymbol{W}}$.

This gives polynomial running time: the FPAUS produces a sample in time $\mathrm{poly}(n, 8MN) = \mathrm{poly}(n, 1/\varepsilon)$. We need $MN = O((m/\varepsilon)^2 \log(m/\varepsilon))$ samples; thus the overall running time is a polynomial. It hence suffices to show that the algorithm produces a $(1 \pm \varepsilon)$-approximation with probability $\frac{3}{4}$.

For our choice of N, $\mathrm{e}^{\beta_N} - 1 \ge \frac{4\cdot 2^m q^n}{\varepsilon}$. The expected number of out-edges under $\pi_G^{\mathrm{ord}}(\,\cdot\,; q, \beta_N)$ is at most $\frac{q\cdot m}{q+\mathrm{e}^{\beta_N}-1} \le \frac{\varepsilon}{4}$; therefore, by the Markov inequality, the probability that the configuration contains no out-edges is $\pi_G^{\mathrm{ord}}(E; q, \beta_N) \ge 1 - \frac{\varepsilon}{4}$. Hence, $\pi_G^{\mathrm{ord}}(E; q, \beta_N) = \frac{(\mathrm{e}^{\beta_N}-1)^m q}{Z_G^{\mathrm{ord}}(q,\beta_N)} \ge 1 - \frac{\varepsilon}{4}$, giving that $q(\mathrm{e}^{\beta_N} - 1)^m$ is a $(1 \pm \frac{\varepsilon}{4})$-approximation to $Z_G^{\mathrm{ord}}(q, \beta_N)$. Note that if $\mathrm{e}^\beta - 1$ is already large enough, then the trivial approximation $q(\mathrm{e}^{\beta_N} - 1)^m$ suffices, and hence we may assume that $N \ge 1$.

Since $(1-\frac{\varepsilon}{4})^2 \ge 1-\varepsilon$ and $(1+\frac{\varepsilon}{4})^2 \le 1+\varepsilon$, it suffices to show that with probability at least $\frac{3}{4}$, $\hat{\boldsymbol{W}}$ is a $(1 \pm \frac{\varepsilon}{4})$-approximation to $\prod_{i=0}^{N-1} \rho_i$.

We analyse the algorithm assuming access to exact samples and then transfer this bound to the approximate samples. For $0 \le i \le N-1$ and $1 \le j \le M$, let $\boldsymbol{F}_i^{(j)} \sim \pi_G^{\mathrm{ord}}(\,\cdot\,; q, \beta_{i+1})$ be independent random variables, and let $\boldsymbol{W}$ be the estimate obtained by the algorithm in Step 3 if using these samples in Step 2(a)i. By the alternate characterisation of the TV distance, we can couple these samples with the $\frac{1}{8NM}$-samples $\hat{\boldsymbol{F}}_{\boldsymbol{i}}^{(\boldsymbol{j})}$ such that for all i, j, it holds that $\mathbb{P}(\hat{\boldsymbol{F}}_{\boldsymbol{i}}^{(\boldsymbol{j})} \ne \boldsymbol{F}_{\boldsymbol{i}}^{(\boldsymbol{j})}) \le \frac{1}{8NM}$. In

particular, by the union bound, they all coincide with probability at least $\frac{7}{8}$. By conditioning on the event that all samples coincide, we have $\boldsymbol{W} = \hat{\boldsymbol{W}}$. Hence,

$$\mathbb{P}\left(|\hat{\boldsymbol{W}} - \prod_{i=0}^{N-1}\rho_i| > \tfrac{\varepsilon}{4}\prod_{i=0}^{N-1}\rho_i\right) \leq \mathbb{P}(\boldsymbol{W} \neq \hat{\boldsymbol{W}}) + \mathbb{P}\left(|\boldsymbol{W} - \prod_{i=0}^{N-1}\rho_i| > \tfrac{\varepsilon}{4}\prod_{i=0}^{N-1}\rho_i\right).$$

Therefore, it suffices to show that with probability at least $\frac{7}{8}$, $\boldsymbol{W}$ is a $(1 \pm \frac{\varepsilon}{4})$-approximation to $\prod_{i=0}^{N-1}\rho_i$.

For this, we use Chebyshev's inequality. Let $\boldsymbol{\rho_i^{(j)}}$ and $\boldsymbol{W^{(j)}}$ denote the quantities from Steps 2(a)ii and 2b respectively. Then, noting that all $\boldsymbol{\rho_i^{(j)}}$'s are independent, $\boldsymbol{W^{(j)}}$ are i.i.d., and $\mathbb{E}[\boldsymbol{W}] = \prod_{i=0}^{N-1}\rho_i = \frac{Z_G^{\mathrm{ord}}(q,\beta_0)}{Z_G^{\mathrm{ord}}(q,\beta_N)}$, we obtain:

$$\mathbb{P}\left(|\boldsymbol{W} - \prod_{i=0}^{N-1}\rho_i| > \tfrac{\varepsilon}{4}\prod_{i=0}^{N-1}\rho_i\right) \leq \frac{16\mathrm{Var}[\boldsymbol{W}]}{\varepsilon^2\mathbb{E}[\boldsymbol{W}]^2} \leq \frac{16}{\varepsilon^2 M}\left(\frac{Z_G^{\mathrm{ord}}(q,\beta_N)}{Z_G^{\mathrm{ord}}(q,\beta_0)}\right)^2 \cdot \prod_{i=0}^{N-1}\mathbb{E}[(\boldsymbol{\rho_i^{(1)}})^2] \tag{3.5}$$

To bound the variance, we will use the following claim.

Claim 3.6. *For all $0 \leq i \leq N-1$, it holds that $\mathbb{E}[(\boldsymbol{\rho_i^{(1)}})^2] = \frac{Z_G^{\mathrm{ord}}(q,\beta_{i-1})}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})}$, where β_{-1} satisfies $\mathrm{e}^{\beta_{-1}} - 1 = (\mathrm{e}^{\beta} - 1)\mathrm{e}^{-1/m}$.*

To prove the claim, we note that by the choice of the β_i's, we can write $(\boldsymbol{\rho_i^{(1)}})^2$ as $\mathrm{e}^{-2|\boldsymbol{F_i^{(1)}}|/m}$. Then, by writing out the weights, we obtain

$$\mathbb{E}[\mathrm{e}^{-2|\boldsymbol{F_i^{(1)}}|/m}] = \sum_{F \in \Omega_G^{\mathrm{ord}}} \frac{q^{c(F)}(\mathrm{e}^{\beta_{i+1}} - 1)^{|F|}\mathrm{e}^{-2|F|/m}}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})} = \frac{Z_G^{\mathrm{ord}}(q,\beta_{i-1})}{Z_G^{\mathrm{ord}}(q,\beta_{i+1})},$$

where the last equality follows from $(\mathrm{e}^{\beta_{i+1}} - 1)\mathrm{e}^{-2/m} = \mathrm{e}^{\beta_{i-1}} - 1$.

Now, we can plug the bound from the claim back into the r.h.s.[2] of (3.5) to obtain that it is equal to

$$\frac{16}{\varepsilon^2 M} \cdot \left(\frac{Z_G^{\mathrm{ord}}(q,\beta_N)}{Z_G^{\mathrm{ord}}(q,\beta_0)}\right)^2 \cdot \frac{Z_G^{\mathrm{ord}}(q,\beta_{-1}) \cdot Z_G^{\mathrm{ord}}(q,\beta_0)}{Z_G^{\mathrm{ord}}(q,\beta_{N-1}) \cdot Z_G^{\mathrm{ord}}(q,\beta_N)}$$
$$= \frac{16}{\varepsilon^2 M} \cdot \frac{Z_G^{\mathrm{ord}}(q,\beta_{-1})}{Z_G^{\mathrm{ord}}(q,\beta_0)} \cdot \frac{Z_G^{\mathrm{ord}}(q,\beta_N)}{Z_G^{\mathrm{ord}}(q,\beta_{N-1})}. \tag{3.6}$$

Clearly, $\frac{Z_G^{\mathrm{ord}}(q,\beta_{-1})}{Z_G^{\mathrm{ord}}(q,\beta_0)} \leq 1$. Thus, it remains to bound $\frac{Z_G^{\mathrm{ord}}(q,\beta_N)}{Z_G^{\mathrm{ord}}(q,\beta_{N-1})}$. Rewriting the fraction as an expectation (same as in (3.3)), and using that $\frac{\mathrm{e}^{\beta_N}-1}{\mathrm{e}^{\beta_{N-1}}-1} = \mathrm{e}^{1/m}$, we obtain

[2]Right-hand side. Similarly, we use l.h.s. as a shorthand for left-hand side.

$$\frac{Z_G^{\mathrm{ord}}(q,\beta_N)}{Z_G^{\mathrm{ord}}(q,\beta_{N-1})} = \mathbb{E}_{\boldsymbol{F}\sim\pi_G^{\mathrm{ord}}(\,\cdot\,;q,\beta_{N-1})}\left[\mathrm{e}^{|\boldsymbol{F}|/m}\right] \le \mathrm{e}$$

Putting it all together, we get that the failure probability of the algorithm is at most $\frac{1}{8} + \frac{16\mathrm{e}}{\varepsilon^2 M} \le \frac{1}{4}$.

This concludes the proof for the ordered phase. The proof for the disordered phase follows the same strategy using the telescoping product (3.2), together with the provided FPAUS to estimate (3.4), as in Algorithm 3.5. The only key difference is that instead of an increasing sequence of β_i's, we make them decreasing, so that the partition function restricted to the disordered phase at β_N is well approximated by the weight of the all-out configuration, q^n.

In particular, we choose the following values for the algorithm:

$$N = \left\lceil m \log\left(\frac{4(\mathrm{e}^{\beta}-1)2^m}{\min(1,q\varepsilon)}\right)\right\rceil \text{ and for } -1 \le i \le N,\ \mathrm{e}^{\beta_{i+1}}-1 = (\mathrm{e}^{\beta_i}-1)\mathrm{e}^{-1/m}, \quad (3.7)$$

and in the last step, we return $\hat{\boldsymbol{Z}} := q^n \hat{\boldsymbol{W}}$.

Then the remainder of the proof is identical to that of the ordered phase, except for two parts where the proof depended on the precise values of β_i.

First we need to verify that q^n is a $(1\pm\frac{\varepsilon}{4})$-approximation to the partition function restricted to the disordered phase at β_N. From the choice of N and β_i in (3.7), $\mathrm{e}^{\beta_N} - 1 \le \min(1, \frac{q\varepsilon}{2^{m+2}})$. Then, we can upper bound the expected number of in-edges under $\pi_G^{\mathrm{dis}}(\,\cdot\,;q,\beta_N)$ by $(1-\mathrm{e}^{-\beta_N})m \le \frac{\varepsilon}{4}$. By Markov's inequality, $\pi_G^{\mathrm{dis}}(\emptyset;q,\beta_N) = \frac{q^n}{Z_G^{\mathrm{dis}}(q,\beta_N)} \ge 1 - \frac{\varepsilon}{4}$, giving that q^n is a $(1\pm\frac{\varepsilon}{4})$-approximation to $Z_G^{\mathrm{dis}}(q,\beta_N)$. Note that if β is small enough so that the above definition yields $N \le 0$, then the trivial approximation q^n suffices.

The last piece left to prove is that the resulting estimator $\boldsymbol{W}$ indeed returns the appropriate approximation with sufficiently high probability. The values of β_i enter the proof on the right-hand side of (3.5) (changing the ordered to the disordered phase). Using our choice of β_i's from (3.7), we obtain, by a similar calculation as for the ordered phase, that for $0 \le i \le N-1$,

$$\mathbb{E}[(\boldsymbol{\rho_i^{(1)}})^2] = \mathbb{E}[\mathrm{e}^{2|\boldsymbol{F_i^{(1)}}|/m}] = \sum_{F\in\Omega_G^{\mathrm{dis}}} \frac{q^{c(F)}(\mathrm{e}^{\beta_{i+1}}-1)^{|F|}\mathrm{e}^{2|F|/m}}{Z_G^{\mathrm{dis}}(q,\beta_{i+1})} = \frac{Z_G^{\mathrm{dis}}(q,\beta_{i-1})}{Z_G^{\mathrm{dis}}(q,\beta_{i+1})},$$

We thus bound the r.h.s. of (3.5) by (3.6) (changing the ordered for the disordered phase). We finish by noting that $\frac{Z_G^{\mathrm{dis}}(q,\beta_N)}{Z_G^{\mathrm{dis}}(q,\beta_{N-1})} \le 1$ and $\frac{Z_G^{\mathrm{dis}}(q,\beta_{-1})}{Z_G^{\mathrm{dis}}(q,\beta_0)} \le \mathrm{e}$. □

As a corollary of the lemma, we can efficiently approximate the proportion of the phases, and therefore obtain the initial distribution needed to sample at criticality, and also the approximation for the (non-restricted) partition function $Z_G(q,\beta)$.

Corollary 3.7. *Let $q \geq 1$ and $\beta > 0$ be reals, and let $\mathcal{G}$ be a class of graphs such that every $G \in \mathcal{G}$ has $|E| = \Omega(n)$ and admits both*

- *an FPAUS for sampling from $\pi_G^{\mathrm{dis}}(\,\cdot\,;q,\beta')$ for all $\beta' \leq \beta$,*
- *an FPAUS for sampling from $\pi_G^{\mathrm{ord}}(\,\cdot\,;q,\beta')$ for all $\beta' \geq \beta$,*

and moreover, $\pi_G(\Omega_G^{\mathrm{dis}} \cup \Omega_G^{\mathrm{ord}};q,\beta) \geq 1 - \mathrm{e}^{-\Omega(|E|)}$. Then there is an algorithm that, given $G \in \mathcal{G}$ and $\varepsilon \in (0,1)$, runs in time poly$(n,1/\varepsilon)$, returns an ε-sample from $\pi_G(\,\cdot\,;q,\beta)$ and also returns, with probability $\frac{3}{4}$, a $(1\pm\varepsilon)$-approximation to $Z_G(q,\beta)$.

Proof. Let $q \geq 1, \beta > 0$ and $\mathcal{G}$ satisfy the conditions in the statement. Let $G \in \mathcal{G}$ and $\varepsilon \in (0,1)$. If $\varepsilon \leq \mathrm{e}^{-\Omega(|E|)}$, then exhaustive enumeration yields a poly$(1/\varepsilon)$-time algorithm. Thus, we may assume that $\varepsilon \leq \frac{1}{10}$[3] and also that,

$$\varepsilon \geq \mathrm{e}^{-\Omega(|E|)} \geq 10(1 - \pi_G(\Omega_G^{\mathrm{dis}} \cup \Omega_G^{\mathrm{ord}};q,\beta)). \tag{3.8}$$

We first approximate the partition functions of the two phases and use their sum to approximate $Z_G(q,\beta)$. We then use their relative sizes to choose the phase from which to sample.

More precisely, we use Lemma 3.4 to obtain, for each phase $P \in \{\mathrm{ord},\mathrm{dis}\}$, in poly$(n,1/\varepsilon)$ time, a quantity $\hat{Z}_G^P$, which is with probability at least $1-\varepsilon/40$ a $(1\pm\frac{\varepsilon}{10})$-approximation to $Z_G^P(q,\beta)$. Note that, using the standard probability amplification for FPRAS, we can obtain this error with only an additional $O(\log\frac{1}{\varepsilon})$ factor in the running time.

First, we claim that $\hat{Z} := \hat{Z}_G^{\mathrm{ord}} + \hat{Z}_G^{\mathrm{dis}}$ is a $(1\pm\varepsilon)$-approximation to $Z_G(q,\beta)$ with probability at least $1 - \varepsilon/20 \geq \frac{3}{4}$. By the union bound, with probability at least $1-\varepsilon/20$, both estimates are $(1\pm\varepsilon/10)$-approximations to the respective partition functions. It follows that

$$\hat{Z}_G^{\mathrm{ord}} + \hat{Z}_G^{\mathrm{dis}} \leq (1+\tfrac{\varepsilon}{10})(Z_G^{\mathrm{ord}}(q,\beta) + Z_G^{\mathrm{dis}}(q,\beta)) \leq (1+\tfrac{\varepsilon}{10})Z_G(q,\beta),$$

and

$$\begin{aligned}\hat{Z}_G^{\mathrm{ord}} + \hat{Z}_G^{\mathrm{dis}} &\geq (1-\tfrac{\varepsilon}{10})(Z_G^{\mathrm{ord}}(q,\beta) + Z_G^{\mathrm{dis}}(q,\beta)) \\ &\geq (1-\tfrac{\varepsilon}{10})(Z_G(q,\beta) - Z_G(q,\beta)\cdot\tfrac{\varepsilon}{10}) \geq (1-\tfrac{\varepsilon}{10})^2 Z_G(q,\beta),\end{aligned}$$

[3]For $\varepsilon > \frac{1}{10}$, we output a better estimate with $\varepsilon' = \frac{1}{10}$. This does not worsen the asymptotic running time since $\frac{1}{\varepsilon}$ is bounded below by a constant.

where the last inequality follows from the lower bound on ε (3.8). Since $(1-\frac{\varepsilon}{10})^2 \geq 1-\frac{1}{5}\varepsilon$, we conclude that $\hat{Z}$ is a $(1\pm\varepsilon)$-approximation to $Z_G(q,\beta)$ with probability at least $1-\varepsilon/20$.

We now use the estimated partition functions to approximate the probabilities with which an exact sample belongs to each phase. The sampling algorithm is as follows:

1. With probability $\frac{\hat{Z}_G^{\mathrm{dis}}}{\hat{Z}_G^{\mathrm{dis}}+\hat{Z}_G^{\mathrm{ord}}}$, return an $\frac{\varepsilon}{5}$-sample from the disordered phase.
2. With the remaining probability, return an $\frac{\varepsilon}{5}$-sample from the ordered phase.

We finalise the proof by showing that this indeed returns an ε-sample from $\pi_G(\,\cdot\,;q,\beta)$.

Consider the optimal coupling with an exact sampler, as defined in Section 2.2. First, we bound the probability that the approximate and exact samplers end up in different phases. We have, for each $P \in \{\mathrm{dis},\mathrm{ord}\}$,

$$\frac{\hat{Z}_G^P}{\hat{Z}_G^{\mathrm{dis}}+\hat{Z}_G^{\mathrm{ord}}} \leq \frac{(1+\varepsilon/10)}{(1-\frac{1}{5}\varepsilon)}\frac{Z_G^P(q,\beta)}{Z_G(q,\beta)} \leq (1+\tfrac{1}{3}\varepsilon)\frac{Z_G^P(q,\beta)}{Z_G(q,\beta)}, \text{ and} \tag{3.9}$$

$$\frac{\hat{Z}_G^P}{\hat{Z}_G^{\mathrm{dis}}+\hat{Z}_G^{\mathrm{ord}}} \geq \frac{(1-\varepsilon/10)}{(1+\frac{1}{10}\varepsilon)}\frac{Z_G^P(q,\beta)}{Z_G(q,\beta)} \geq (1-\tfrac{1}{5}\varepsilon)\frac{Z_G^P(q,\beta)}{Z_G(q,\beta)}, \tag{3.10}$$

where we use $\varepsilon \leq \frac{1}{10}$. This gives that with probability at least $1-2\cdot\frac{1}{3}\varepsilon-\frac{1}{20}\varepsilon = 1-\frac{43}{60}\varepsilon$, we both obtain accurate estimates and choose the same phase as the exact sampler. In that case, the TV distance from the phase is at most $\frac{1}{5}\varepsilon = \frac{12}{60}\varepsilon$. We can conclude that the above procedure produces an $\frac{55}{60}\varepsilon \leq \varepsilon$-sample. □

If each FPAUS in Corollary 3.7 is obtained by running the RC dynamics initialised from the corresponding ground state, then the sampling procedure above is equivalent to initialising the RC dynamics from a random mixture of the all-in and all-out configurations, with probabilities approximately proportional to the partition functions restricted to the ordered and disordered phases, respectively. The mixing time from this initial distribution is bounded by the maximum of the mixing times from the two ground states.

Lemma 3.4, and hence Corollary 3.7, applies to the MCMC samplers obtained from Theorems 1.1 and 1.3. There are, however, several obstacles to applying the same argument in the setting of Theorem 1.2. In particular, the planting framework used there requires β to be fixed before sampling the random regular graph, whereas

Lemma 3.4 requires efficient sampling throughout an interval of temperatures. In particular, this does not guarantee that, for a single random graph G, efficient sampling is possible simultaneously at the multiple temperatures required by Lemma 3.4. Thus, the planting result does not directly imply the FPRAS required for sampling at criticality.

Chapter 4

Sampling from the random-cluster model on random regular graphs via Glauber dynamics

This chapter is based on the following publication:

[38] 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*, 34(3):359–391, 2025. doi: 10.1017/S0963548324000403

An extended abstract of this work also appeared in RANDOM:

[37] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Sampling from the random cluster model on random regular graphs at all temperatures via Glauber dynamics. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023)*, volume 275 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 64:1–64:12, 2023. doi: 10.4230/LIPIcs.APPROX/RANDOM.2023.64

4.1 Overview

In this chapter, we use Theorem 1.5 to prove Theorem 1.1, which establishes fast mixing of the RC dynamics in the same setting and parameter ranges as the previous cluster-expansion algorithm by Helmuth, Jenssen and Perkins [49]. We restate the theorem below.

Theorem 1.1. *Let $\Delta \geq 5$ be an integer and let $q \geq \Delta^{\Omega(\Delta)}$ be a sufficiently large real. Then, with high probability over a random Δ-regular graph $G = (V, E)$, the following*

hold.

- *For all $\beta < \beta_c$, the RC dynamics initialised from $X_0 \coloneqq \emptyset$ (all-out) mixes in $O(n \log n)$ time.*
- *For all $\beta > \beta_c$, the RC dynamics initialised from $X_0 = E$ (all-in) mixes in polynomial time. Moreover, if q is an integer, it mixes in $O(n \log n)$ time.*
- *For $\beta = \beta_c$, the RC dynamics initialised from a suitable random mixture of all-out and all-in mixes in polynomial time. Moreover, this random mixture can be computed in polynomial time.*

Recall that $\beta_c = \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$ is the disordered/ordered phase transition, where the *disordered phase* is the set of configurations that contain at most an η-fraction of edges, and the *ordered phase* is the set of configurations that contain at least a $(1-\eta)$-fraction of edges. Here, η is a small constant whose precise value we give later.

We use that, in the setting of Theorem 1.1, for all $\beta < \beta_c$ an RC configuration is disordered with probability $1 - \mathrm{e}^{-\Omega(n)}$, and for all $\beta > \beta_c$ an RC configuration is ordered with probability $1 - \mathrm{e}^{-\Omega(n)}$. At all temperatures $\beta > 0$, an RC configuration belongs to one of the phases with probability at least $1 - \mathrm{e}^{-\Omega(n)}$ [49].

Therefore, the Theorem reduces to proving two statements: mixing within the disordered phase when starting from all-out for all $\beta \leq \beta_c$, and mixing within the ordered phase from all-in for all $\beta \geq \beta_c$. The two statements directly give us the theorem for $\beta \neq \beta_c$. At criticality, Theorem 1.1 then follows from the two results by Corollary 3.7.

Thus, the remainder of the chapter focuses on proving fast mixing within the two phases. By Theorem 1.5, it suffices to establish the three key properties from Chapter 3. The main technical contribution is the proof of WSM within the ordered phase, for which we introduce the bottom-up revealing scheme.

4.1.1 Notation

We next define the necessary notation before formally stating the main theorems of this chapter. Throughout the chapter, all logarithms are with base e, unless explicitly said otherwise.

For $\Delta \geq 5$ and an integer n such that Δn is an even number[1] let $\mathcal{G}_{n,\Delta}$ denote the set of all Δ-regular graphs with n vertices. We write $G \sim \mathcal{G}_{n,\Delta}$ to denote a graph in $\mathcal{G}_{n,\Delta}$ chosen uniformly at random. Throughout the chapter, when the graph $G = (V, E)$ is clear from context, we write n for the number of vertices and m for the number of edges. For $U \subseteq V$ we write $G[U]$ for the induced graph on the vertex set U, and for $F \subseteq E$, we write $G[F]$ for the graph (V, F).

To be able to work with partially revealed configurations in the revealing scheme, in this chapter we represent RC configurations as functions rather than edge subsets. More precisely, for a graph $G = (V, E)$, let $\Omega = \Omega_G$ be the set of all assignments $\mathcal{F} : E \to \{0, 1\}$; . We typically refer to assignments in Ω as (RC) configurations. For a configuration $\mathcal{F}$, we say that edges mapped to 1 are *in-edges*, and edges mapped to 0 are *out-edges*. We use $\mathrm{In}(\mathcal{F})$ to denote the set of edges e with $\mathcal{F}(e) = 1$, $\mathrm{Out}(\mathcal{F})$ to denote the set of edges e with $\mathcal{F}(e) = 0$, $|\mathcal{F}|$ for the cardinality of $\mathrm{In}(\mathcal{F})$ and $c(\mathcal{F})$ for the number of connected components in the graph $(V, \mathrm{In}(\mathcal{F}))$. Then, for real parameters $q, \beta > 0$, the weight of $\mathcal{F}$ in the RC model is given by $w_G(\mathcal{F}) = q^{c(\mathcal{F})}(e^\beta - 1)^{|\mathcal{F}|}$. We write $\pi_G = \pi_G(\,\cdot\,; q, \beta)$ for the RC distribution, which is proportional to these weights.

4.1.2 Locally treelike expanders

Before formally defining the phases, we introduce two properties of random regular graphs that will be used throughout the chapter. As in [49], instead of considering only random regular graphs, we work slightly more generally with Δ-regular graphs that are both expanding and locally treelike.

The *expansion profile* of an n-vertex graph $G = (V, E)$ for $\varepsilon > 0$ is given by

$$\phi_G(\varepsilon) \coloneqq \min_{S \subseteq V;\ 0 < |S| \leq \varepsilon n} \frac{|E(S, V \backslash S)|}{\Delta |S|}.$$

The *tree-excess* of a connected graph G is given by $|E| - |V| + 1$. For a vertex v in G and integer $r \geq 0$, let $B_r(v)$ denote the set of vertices at distance at most r from v and $S_r(v)$ those at distance exactly r from v. For $K > 0$, a graph G with maximum degree Δ is *locally K-treelike* if for every $v \in V$ and $r \leq \frac{1}{3} \log_{\Delta - 1} |V|$, the graph $G[B_r(v)]$ has tree excess $\leq K$.

The following two lemmas state that a random regular graph is w.h.p. a locally treelike expander.

[1] Since every edge is incident to two vertices, for a n-vertex Δ-regular graph to exist, it must be the case that $n\Delta$ is even.

Lemma 4.1 (see, e.g., [43, Lemma 5.8],[10, Fact 2.3])**.** *For any integer $\Delta \geq 3$, there is $K > 0$ such that w.h.p. $G \sim \mathcal{G}_{n,\Delta}$ is locally K-treelike.*

Lemma 4.2 ([49, Proposition 37])**.** *Fix $\Delta \geq 5$. There is a constant $\delta \in (0, 1/2)$ such that w.h.p. a uniformly random Δ-regular graph G satisfies $\phi_G(1/2) \geq 1/10$ and $\phi_G(\delta) \geq 5/9$.*

Then the classes $\mathcal{G}_{\Delta,\delta}$ and $\mathcal{G}_{\Delta,\delta,K}$ are as follows.

Definition 4.3. *Let $\Delta \geq 5$ be an integer, and $\delta \in (0, 1/2), K > 0$ be reals. $\mathcal{G}_{\Delta,\delta}$ is the class of Δ-regular graphs such that $\phi_G(1/2) \geq 1/10$ and $\phi_G(\delta) \geq 5/9$. $\mathcal{G}_{\Delta,\delta,K}$ is the class of all locally K-treelike graphs $G \in \mathcal{G}_{\Delta,\delta}$.*

By Lemmas 4.1 and 4.2, w.h.p. $G \sim \mathcal{G}_{n,\Delta}$ belongs to $\mathcal{G}_{\Delta,\delta,K}$ for some constant $K > 0$. Therefore, to establish results w.h.p. over random regular graphs, it suffices to show them for all $G \in \mathcal{G}_{\Delta,\delta,K}$ with n sufficiently large.

4.1.3 Mixing within the phase

We can now formally define the phases of the random-cluster model. We use the same definition as in [49].

Definition 4.4. *Fix $\Delta \geq 5$. Fix $\delta \in (0, 1/2)$ satisfying Lemma 4.2. Let $\eta = \min(\delta/5, 1/100)$. For $G \in \mathcal{G}_{n,\Delta}$, the* ordered phase *is the set of configurations $\Omega^{\mathrm{ord}} := \{\mathcal{F} \in \Omega : |\operatorname{In}(\mathcal{F})| \geq (1-\eta)|E|\}$, whereas the* disordered phase *is the set $\Omega^{\mathrm{dis}} := \{\mathcal{F} \in \Omega : |\operatorname{In}(\mathcal{F})| \leq \eta|E|\}$. For $q, \beta > 0$, let $\pi_G^{\mathrm{ord}}, \pi_G^{\mathrm{dis}}$ be the conditional distributions of π_G on $\Omega^{\mathrm{ord}}, \Omega^{\mathrm{dis}}$, respectively.*

We reduce proving Theorem 1.1 to establishing fast mixing within each phase, as stated in the following three theorems.

Theorem 4.5 (Mixing within the ordered phase for real q)**.** *Let $\Delta \geq 5$ be an integer. Then, for all sufficiently large q, there exists $C = C(\Delta)$ such that the following holds for $\beta \geq \beta_c$, w.h.p. over $G \sim \mathcal{G}_{n,\Delta}$. The RC dynamics $(X_t)_{t\geq 0}$ on G starting from all-in satisfies $\operatorname{dist}_{\mathrm{TV}}(X_T, \pi_G^{\mathrm{ord}}) \leq 1/5$ for $T = O(n^C)$.*

Theorem 4.6 (Mixing within the disordered phase for real q)**.** *Let $\Delta \geq 5$ be an integer. Then, for all sufficiently large q and $\beta \leq \beta_c$, w.h.p. over $G \sim \mathcal{G}_{n,\Delta}$, the RC dynamics $(X_t)_{t\geq 0}$ on G starting from all-out satisfies $\operatorname{dist}_{\mathrm{TV}}(X_T, \pi_G^{\mathrm{dis}}) \leq 1/5$ for $T = O(n \log n)$.*

Theorem 4.7 (Optimal mixing within the ordered phase for integer q)**.** *Let $\Delta \geq 5$ be an integer. Then, for all sufficiently large integer q and $\beta \geq \beta_c$, w.h.p. over $G \sim \mathcal{G}_{n,\Delta}$, the RC dynamics $(X_t)_{t\geq 0}$ on G starting from all-in satisfies $\mathrm{dist}_{\mathrm{TV}}(X_T, \pi_G^{\mathrm{ord}}) \leq 1/5$ for $T = O(n \log n)$.*

By Theorem 1.5, to prove these theorems it suffices to establish the three key ingredients, which are the focus for the majority of the chapter. We conclude the proof of Theorems 4.5-4.7 and thus Theorem 1.1 in Section 4.6.

4.1.3.1 Local maximum (Key Property 1)

The first key property, a local maximum in each phase, is a consequence of the following lemma, which is in turn obtained by combining [49, Theorem 1 and Lemma 9]. We give the proof in Section 4.2.

Lemma 4.8. *Let $\Delta \geq 5$ be an integer. Then, for all sufficiently large q, there exists $\beta_c > 0$ satisfying $\beta_c = (1 + o_q(1))\frac{2\log q}{\Delta}$ such that the following holds for any $\beta > 0$ w.h.p. for $G \sim \mathcal{G}_{n,\Delta}$.*

$$\text{if } \beta < \beta_c, \text{ then } \left\|\pi_G - \pi_G^{\mathrm{dis}}\right\|_{\mathrm{TV}} = \mathrm{e}^{-\Omega(n)}; \qquad \text{if } \beta > \beta_c, \text{ then } \left\|\pi_G - \pi_G^{\mathrm{ord}}\right\|_{\mathrm{TV}} = \mathrm{e}^{-\Omega(n)}. \tag{4.1}$$

Moreover, there exists $\zeta = \zeta(\Delta) > 0$ with $\zeta < \eta$ such that

$$\begin{aligned} &\text{for } \beta \leq \beta_c, \quad \pi_G^{\mathrm{dis}}\Big(|\operatorname{In}(\mathcal{F})| \geq \zeta|E|\Big) = \mathrm{e}^{-\Omega(n)}, \quad \text{and} \\ &\text{for } \beta \geq \beta_c, \quad \pi_G^{\mathrm{ord}}\Big(|\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|\Big) = \mathrm{e}^{-\Omega(n)}. \end{aligned} \tag{4.2}$$

4.1.3.2 WSM within the phase (Key Property 2)

The main technical result in this Chapter is establishing WSM within the phase. For a vertex $v \in V$ and an integer $r \geq 1$, recall that $\pi_{B_r^-(v)}$ is the conditional distribution of π where all edges outside $B_r(v)$ are fixed to be out. Analogously, $\pi_{B_r^+(v)}$ denotes the RC distribution where all vertices at distance r from v are tied into a single component. As explained in Remark 3.2, this is a slightly stronger version of the wired ball used in the statement of Key Property 2.

To establish WSM within the phase, we prove the following two theorems. Note that since we defined configurations as functions, we use $e \mapsto \cdot$ for the random indicator variable "e is an in-edge".

Theorem 4.9. *Let $\Delta \geq 5$ and $K \geq 0$ be integers and let $\delta \in (0, 1/2)$ be a real. There exists $M = M(\Delta, K) > 0$ such that for all q sufficiently large and any $\beta \geq \beta_c$ the following holds for all sufficiently large n. Let $G \sim \mathcal{G}_{\Delta,\delta,K}$ be an n-vertex graph. Then G has WSM within the ordered phase at a radius r which satisfies $r \leq \frac{M}{\beta} \log n$.*

That is, for every $v \in V(G)$ and every edge e incident to v, it holds that

$$\left\| \pi_{B_r^+(v)}(e \mapsto \cdot) - \pi_G^{\mathrm{ord}}(e \mapsto \cdot) \right\|_{\mathrm{TV}} \leq \frac{1}{100m}.$$

The upper bound on the radius r in terms of $1/\beta$ ensures that we can remove the dependence on β of the mixing time in Theorem 1.1. We prove WSM within the ordered phase in Section 4.3.

For the disordered phase, we prove WSM within the phase at radius $\frac{1}{3} \log_{\Delta-1} n$ (stated below) in Section 4.4.

Theorem 4.10. *Let $\Delta \geq 5$ and $K \geq 0$ be integers and let $\delta \in (0, 1/2)$ be a real. Then, for all q sufficiently large and any $\beta \leq \beta_c$, and for all sufficiently large n, the following holds. Let $G \sim \mathcal{G}_{\Delta,\delta,K}$ be an n-vertex graph. Then G has WSM within the disordered phase at a radius r which satisfies $r \leq \frac{1}{3} \log_{\Delta-1} n$. That is, for every $v \in V(G)$ and every edge e incident to v, it holds that*

$$\left\| \pi_{B_r^-(v)}(e \mapsto \cdot) - \pi_G^{\mathrm{dis}}(e \mapsto \cdot) \right\|_{\mathrm{TV}} \leq \frac{1}{100m}.$$

4.1.3.3 Local mixing (Key Property 3)

The final ingredient is local mixing for the free and wired RC dynamics on radius-r neighbourhoods. In the *free RC dynamics*, all edges outside $B_r(v)$ are fixed to be out, whereas in the *wired RC dynamics* all vertices at distance r from v are tied into a single component. In both cases, only edges with both endpoints in $B_r(v)$ are updated. As above, the wired dynamics uses the slightly stronger boundary condition justified in Remark 3.2.

Since random regular graphs are w.h.p. locally treelike, there is a constant $K > 0$ such that for all $r \leq \frac{1}{3} \log_\Delta n$, every radius-$r$ ball has tree excess at most K. The following three lemmas give the local-mixing bounds that we need on such neighbourhoods.

First, we give results establishing optimal mixing for the disordered phase and for the ordered phase when q is an integer. Recall that to obtain optimal mixing, we need to show that the log-Sobolev constant (see the definition in the Preliminaries 2.16) for the relevant dynamics on $B_r(v)$ is $\Omega(1/|E(B_r(v))|)$. For the free RC dynamics, this was proven in [10, Lemma 6.5].

Lemma 4.11 ([10, Lemma 6.5])**.** *Let $\Delta \geq 3$ be an integer, and $q, K > 1$, $\beta > 0$ be reals. There exists $C > 0$ such that the following holds for any Δ-regular graph G and integer $r \geq 1$.*

Suppose that $\rho \in V$ is such that $G[B_r(\rho)]$ is K-treelike. Then, with $n_\rho = |B_r(\rho)|$, the log-Sobolev constant for the free RC dynamics on $B_r(\rho)$ is $\geq C/n_\rho$.

We also obtain optimal mixing for the wired RC dynamics *when q is an integer*. The following lemma is a minor modification of the analogous result for the Δ-regular tree from [15] and is proven in Section 4.5.1.

Lemma 4.12. *Let $q \geq 2$ and $\Delta \geq 3$ be integers, and $K, \beta > 0$ be reals. There exists $\tilde{C} > 0$ such that the following holds for every Δ-regular graph $G = (V, E)$ and any integer $r \geq 1$.*

Suppose that $\rho \in V$ is such that $G[B_r(\rho)]$ is K-treelike. Then, with $n_\rho = |B_r(\rho)|$, the log-Sobolev constant for the wired RC dynamics on $B_r(\rho)$ is $\geq \tilde{C}/n_\rho$.

For non-integer q, we only obtain polynomial (as opposed to optimal) local mixing, as the result from [15] relies on the connection with the Potts model. Instead of the log-Sobolev constant, the proof uses the standard canonical path method (see e.g. [53, Chapter 5]). We prove the following lemma in Section 6.6.1.

Lemma 4.13. *Let $\Delta \geq 3$ be an integer, and $q, K > 1$, $\beta > 0$ be reals. There exists $\hat{C} > 0$ such that the following holds for every Δ-regular graph $G = (V, E)$ and any integer $r \geq 1$.*

Suppose that $\rho \in V$ is such that $G[B_r(\rho)]$ is K-treelike. Then, with $n_\rho = |B_r(\rho)|$, the mixing time of the wired RC dynamics on $B_r(\rho)$ is $\leq \hat{C} n_\rho^3 (q^4 \mathrm{e}^\beta)^{\Delta r}$.

Lemma 4.13 is proved in Section 4.5.2. Note that the dependence on β in the exponent is the reason why we required $r \approx \frac{1}{\beta} \log n$ in WSM within the ordered phase.

4.1.4 Partial configurations

To prove WSM within the phase, we use an iterative revealing process. We formalise the state of the process using *partial configurations*, where some edges are not yet *revealed*. In addition to an edge being in or out, we will use the symbol "$*$" to denote that an edge e has not yet been revealed.

Definition 4.14. *A* partial configuration $\mathcal{A}$ *is a function* $\mathcal{A} : E \to \{0, 1, *\}$. *We write* $R(\mathcal{A}) \coloneqq \mathcal{A}^{-1}(\{0, 1\})$ *and refer to it as the set of* revealed edges *in* $\mathcal{A}$. *We denote by* $\mathrm{In}(\mathcal{A})$ *the set of edges* e *with* $\mathcal{A}(e) = 1$ *(the "in" edges) and by* $\mathrm{Out}(\mathcal{A})$ *the set of edges* e *with* $\mathcal{A}(e) = 0$ *(the "out" edges). We use* Ω^* *to denote the set of all partial configurations.*

Note that each configuration is also a partial configuration – these are precisely the partial configurations $\mathcal{A}$ with $R(\mathcal{A}) = E$.

Next we define additional notation for partial configurations, which will be used throughout the revealing process.

Definition 4.15. *Let* $\mathcal{A}_1, \mathcal{A}_2$ *be partial configurations. We say that* $\mathcal{A}_2$ *is a* refinement *of* $\mathcal{A}_1$, *written* $\mathcal{A}_1 \subseteq \mathcal{A}_2$, *whenever* $\mathrm{In}(\mathcal{A}_1) \subseteq \mathrm{In}(\mathcal{A}_2)$ *and* $\mathrm{Out}(\mathcal{A}_1) \subseteq \mathrm{Out}(\mathcal{A}_2)$ *(and hence* $R(\mathcal{A}_1) \subseteq R(\mathcal{A}_2)$ *as well). If* $R(\mathcal{A}_1)$ *and* $R(\mathcal{A}_2)$ *are disjoint, the union of* $\mathcal{A}_1$ *and* $\mathcal{A}_2$, *written* $\mathcal{A}_1 \cup \mathcal{A}_2$, *is the partial configuration with* $\mathrm{In}(\mathcal{A}_1 \cup \mathcal{A}_2) = \mathrm{In}(\mathcal{A}_1) \cup \mathrm{In}(\mathcal{A}_2)$ *and* $\mathrm{Out}(\mathcal{A}_1 \cup \mathcal{A}_2) = \mathrm{Out}(\mathcal{A}_1) \cup \mathrm{Out}(\mathcal{A}_2)$.

The following distributions on configurations and partial configurations will be useful.

Definition 4.16. *Let* $\mathcal{A}$ *be a partial configuration in* Ω^*. *Then* $\Omega_{\mathcal{A}} \coloneqq \{\mathcal{F} \in \Omega \mid \mathcal{A} \subseteq \mathcal{F}\}$ *will denote the set of all configurations that refine* $\mathcal{A}$ *and* $\pi_{\mathcal{A}}$ *is the conditional distribution of* π *in* $\Omega_{\mathcal{A}}$, *i.e., for* $\mathcal{F} \in \Omega_{\mathcal{A}}$, *it holds that* $\pi_{\mathcal{A}}(\mathcal{F}) = \pi(\mathcal{F}) / \sum_{\mathcal{F}' \in \Omega_{\mathcal{A}}} \pi(\mathcal{F}')$. *Similarly, if* $\Omega_{\mathcal{A}} \cap \Omega^{\mathrm{ord}}$ *is non-empty, then* $\pi^{\mathrm{ord}}_{\mathcal{A}}$ *is the conditional distribution of* π^{ord} *in* $\Omega_{\mathcal{A}} \cap \Omega^{\mathrm{ord}}$, *and if* $\Omega_{\mathcal{A}} \cap \Omega^{\mathrm{dis}}$ *is non-empty, then* $\pi^{\mathrm{dis}}_{\mathcal{A}}$ *is the conditional distribution of* π^{dis} *in* $\Omega_{\mathcal{A}} \cap \Omega^{\mathrm{dis}}$.

For an edge subset $F \subseteq E$, *let* $\Omega_F \coloneqq \{\mathcal{A} \in \Omega^* \mid R(\mathcal{A}) = F\}$. *The distribution* π_F *on* Ω_F *is defined as follows. For* $\mathcal{A} \in \Omega_F$, *define* $\pi_F(\mathcal{A}) \coloneqq \pi(\Omega_{\mathcal{A}})$. *Similarly,* π^{ord}_F *and* π^{dis}_F *are the distributions on* Ω_F *so that* $\pi^{\mathrm{ord}}_F(\mathcal{A}) = \pi^{\mathrm{ord}}(\Omega_{\mathcal{A}})$ *and* $\pi^{\mathrm{dis}}_F(\mathcal{A}) = \pi^{\mathrm{dis}}(\Omega_{\mathcal{A}})$ *respectively.*

Finally, let $\mathcal{A} \in \Omega^*$ *be a partial configuration and let* F *be a subset of* E *with* $R(\mathcal{A}) \subseteq F$. *Then* $\Omega_{\mathcal{A},F} \coloneqq \{\mathcal{A}' \in \Omega_F \mid \mathcal{A} \subseteq \mathcal{A}', R(\mathcal{A}') = F\}$. *We define a distribution* $\pi_{\mathcal{A},F}$ *on* $\Omega_{\mathcal{A},F}$ *given by*

$$\pi_{\mathcal{A},F}(\mathcal{A}') \coloneqq \frac{\pi(\Omega_{\mathcal{A}'})}{\sum_{\mathcal{A}'' \in \Omega_{\mathcal{A},F}} \pi(\Omega_{\mathcal{A}''})}.$$

We conclude the section by giving two useful properties of the random-cluster model, in the language of partial configurations.

4.1.5 Boundary conditions

We will be interested in vertices on the *boundary* of a partial configuration, and in the structure of connected components in the graph induced by in-edges. We use the following definition.

Definition 4.17. *Let $\mathcal{A}$ be a partial configuration. The* boundary vertices *of $\mathcal{A}$ are $\partial(R(\mathcal{A})) = \{v \in V \mid v$ is incident to an edge in $R(\mathcal{A})$ and to an edge in $E \setminus E(\mathcal{A})\}$.*

The boundary component set *$\xi(\mathcal{A})$ of $\mathcal{A}$ is the set of equivalence classes corresponding to components of the boundary vertices, i.e., $\xi(\mathcal{A}) = \{\kappa \cap \partial(R(\mathcal{A})) \mid \kappa$ is a component in $G[\mathrm{In}(\mathcal{A})]\}$.*

The next observation says that if the boundary component sets of two partial configurations coincide, then so do marginals on their unrevealed edges.

Observation 4.18. *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two partial configurations such that $R(\mathcal{F}_1) = R(\mathcal{F}_2)$ and $\xi(\mathcal{F}_1) = \xi(\mathcal{F}_2)$. Let $F = E \setminus R(\mathcal{F}_1)$. Then $\pi_{\mathcal{F}_1}$ and $\pi_{\mathcal{F}_2}$ have the same projection on F, meaning that for any partial configuration $\mathcal{A} \in \Omega_F$, it holds that $\pi_{\mathcal{F}_1}(\mathcal{F}_1 \cup \mathcal{A}) = \pi_{\mathcal{F}_2}(\mathcal{F}_2 \cup \mathcal{A})$.*

Proof. Let $\mathcal{A}$ be a partial configuration in Ω_F. Consider the components of $G[\mathrm{In}(\mathcal{F}_1 \cup \mathcal{A})]$ and $G[\mathrm{In}(\mathcal{F}_2 \cup \mathcal{A})]$.

- The components of $G[\mathrm{In}(\mathcal{F}_1 \cup \mathcal{A})]$ containing only vertices of $V_R(\mathcal{F}_1) \setminus \partial\mathcal{F}_1$ do not depend on the partial configuration $\mathcal{A}$. Let C_1 be the number of these components. Similarly, the components of $G[\mathrm{In}(\mathcal{F}_2 \cup \mathcal{A})]$ containing only vertices of $V_R(\mathcal{F}_2) \setminus \partial\mathcal{F}_2$ do not depend on $\mathcal{A}$. Let C_2 be the number of these components.
- The components of $G[\mathrm{In}(\mathcal{F}_1 \cup \mathcal{A})]$ containing only vertices of $V \setminus V_R(\mathcal{F}_1)$ depend on $\mathcal{A}$ but not on F, so the same components are in $G[\mathrm{In}(\mathcal{F}_2 \cup \mathcal{A})]$.
- Since $\xi(\mathcal{F}_1) = \xi(\mathcal{F}_2)$, there is one-to-one correspondence between components of $G[\mathrm{In}(\mathcal{F}_1 \cup \mathcal{A})]$ that contain vertices in $\partial\mathcal{F}_1$ and components of $G[\mathrm{In}(\mathcal{F}_2 \cup \mathcal{A})]$ containing vertices in $\partial\mathcal{F}_2$. Any two components that correspond to each other induce the same component on F.

Thus, for any $\mathcal{A}\colon F \to \{0, 1, *\}$, the number of components in $G[\mathrm{In}(\mathcal{A} \cup \mathcal{F}_1)]$ minus C_1 is equal to the number of components in $G[\mathrm{In}(\mathcal{A} \cup \mathcal{F}_2)]$ minus C_2. Hence $w_G(\mathcal{A} \cup \mathcal{F}_1)$ and $w_G(\mathcal{A} \cup \mathcal{F}_2)$ differ by a constant multiplicative factor, from which the result follows. □

4.1.6 Monotonicity

We conclude the preliminaries by providing a generalisation of Lemma 2.14 allowing us to maintain monotonicity when revealing multiple edges at once.

First, we show the following observation.

Observation 4.19. *Let f be an edge in E. Consider partial configurations $\mathcal{F}_1$ and $\mathcal{F}_2$ with $R(\mathcal{F}_1) = R(\mathcal{F}_2) = E\backslash\{f\}$ and $\mathrm{In}(\mathcal{F}_1) \subseteq \mathrm{In}(\mathcal{F}_2)$. Let $\mathcal{A}_1 \sim \pi_{\mathcal{F}_1}$ and $\mathcal{A}_2 \sim \pi_{\mathcal{F}_2}$. Then $\mathbb{P}(f \in \mathrm{In}(\mathcal{A}_1)) \leq \mathbb{P}(f \in \mathrm{In}(\mathcal{A}_2))$.*

Proof. First, observe that if $G[\mathrm{In}(\mathcal{F}_1)]$ has two components connected by f, then the vertices in these components are connected in $G[\mathrm{In}(\mathcal{F}_2) \cup \{f\}]$.

For $j \in \{0, 1\}$, let $\mathcal{F}_{i,j}$ be $\mathcal{F}_i \cup \{f \mapsto j\}$. Then, by the observation, $c(\mathcal{F}_{2,0}) - c(\mathcal{F}_{2,1}) \leq c(\mathcal{F}_{1,0}) - c(\mathcal{F}_{1,1})$. It follows that

$$\begin{aligned}
\mathbb{P}(f \in \mathrm{In}(\mathcal{A}_1)) &= \frac{\mathbb{P}(\mathcal{F}_{1,1})}{\mathbb{P}(\mathcal{F}_{1,1}) + \mathbb{P}(\mathcal{F}_{1,0})} = \frac{(e^\beta - 1)^{|\mathcal{F}_1^{-1}(1)|+1} q^{c(\mathcal{F}_{1,1})}}{(e^\beta - 1)^{|\mathcal{F}_1^{-1}(1)|+1} q^{c(\mathcal{F}_{1,1})} + (e^\beta - 1)^{|\mathcal{F}_1^{-1}(1)|} q^{c(\mathcal{F}_{1,0})}} \\
&= \frac{(e^\beta - 1) q^{c(\mathcal{F}_{1,1}) - c(\mathcal{F}_{1,0})}}{(e^\beta - 1) q^{c(\mathcal{F}_{1,1}) - c(\mathcal{F}_{1,0})} + 1} \leq \frac{(e^\beta - 1) q^{c(\mathcal{F}_{2,1}) - c(\mathcal{F}_{2,0})}}{(e^\beta - 1) q^{c(\mathcal{F}_{2,1}) - c(\mathcal{F}_{2,0})} + 1} \\
&= \mathbb{P}(f \in \mathrm{In}(\mathcal{A}_2))
\end{aligned}$$

where the inequality follows from $c(\mathcal{F}_{1,1}) - c(\mathcal{F}_{1,0}) \leq c(\mathcal{F}_{2,1}) - c(\mathcal{F}_{2,0})$. □

In the language of [47], Observation 4.19 says that π is 1-monotonic. It implies the following monotonicity result. In particular, Corollary 4.20 gives information about marginals in the case where $R(\mathcal{F}_1) = R(\mathcal{F}_2)$ and $\mathrm{In}(\mathcal{F}_1) \subseteq \mathrm{In}(\mathcal{F}_2)$. The proof follows from Observation 4.19.

Corollary 4.20. *Let f be an edge of E. Consider partial configurations $\mathcal{F}_1$ and $\mathcal{F}_2$ with $R(\mathcal{F}_1) = R(\mathcal{F}_2)$ and $\mathrm{In}(\mathcal{F}_1) \subseteq \mathrm{In}(\mathcal{F}_2)$. Let $\mathcal{A}_1 \sim \pi_{\mathcal{F}_1}$ and $\mathcal{A}_2 \sim \pi_{\mathcal{F}_2}$. Then $\mathbb{P}(f \in \mathrm{In}(\mathcal{A}_1)) \leq \mathbb{P}(f \in \mathrm{In}(\mathcal{A}_2))$.*

Proof. Use [47, Theorem 2.27]. It says that if a RCM measure is 1-monotonic, then the measure is monotonic, i.e. $\pi_{\mathcal{F}_1}(A) \leq \pi_{\mathcal{F}_2}(A)$ for any increasing event – that is any $A \subseteq \Omega$ such that whenever $\mathcal{F}_1', \mathcal{F}_2'$ are configurations with $\mathrm{In}(\mathcal{F}_1') \subseteq \mathrm{In}(\mathcal{F}_2')$ and $\mathcal{F}_1' \in A$, then also $\mathcal{F}_2' \in A$.

Observation 4.19 says that π is 1-monotonic; thus monotonicity follows by the theorem. To conclude, take A to be the increasing event "f is occupied" – we can see that this is increasing as whenever, for configurations $\mathcal{F}_1', \mathcal{F}_2'$, $f \in \mathrm{In}(\mathcal{F}_1')$, and $\mathrm{In}(\mathcal{F}_1') \subseteq \mathrm{In}(\mathcal{F}_2')$, then also $f \in \mathrm{In}(\mathcal{F}_2')$. □

4.2 Local maxima and phases of the RC model

In this section, we prove Lemma 4.8, which we restate for convenience.

Lemma 4.8. *Let $\Delta \geq 5$ be an integer. Then, for all sufficiently large q, there exists $\beta_c > 0$ satisfying $\beta_c = (1 + o_q(1))\frac{2\log q}{\Delta}$ such that the following holds for any $\beta > 0$ w.h.p. for $G \sim \mathcal{G}_{n,\Delta}$.*

$$\text{if } \beta < \beta_c, \text{ then } \left\|\pi_G - \pi_G^{\text{dis}}\right\|_{\text{TV}} = \mathrm{e}^{-\Omega(n)}; \qquad \text{if } \beta > \beta_c, \text{ then } \left\|\pi_G - \pi_G^{\text{ord}}\right\|_{\text{TV}} = \mathrm{e}^{-\Omega(n)}. \tag{4.1}$$

Moreover, there exists $\zeta = \zeta(\Delta) > 0$ with $\zeta < \eta$ such that

$$\begin{aligned} &\text{for } \beta \leq \beta_c, \quad \pi_G^{\text{dis}}\Big(|\operatorname{In}(\mathcal{F})| \geq \zeta|E|\Big) = \mathrm{e}^{-\Omega(n)}, \quad \text{and} \\ &\text{for } \beta \geq \beta_c, \quad \pi_G^{\text{ord}}\Big(|\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|\Big) = \mathrm{e}^{-\Omega(n)}. \end{aligned} \tag{4.2}$$

We will use the following lemma.

Lemma 4.21 ([49, Lemma 8])**.** *Let $\Delta \geq 5$, $\delta \in (0, 1/2)$, and $\zeta \in (0, 1/2)$ be reals. Suppose that $G = (V, E) \in \mathcal{G}_{\Delta,\delta}$ satisfies $|V| \geq 360/(\zeta\delta)$ and $A \subseteq E$ is such that $\zeta|E| \leq |A| \leq (1-\zeta)|E|$. Then $c(A)/|V| + |A|/|E| \leq 1 - \zeta/40$ where $c(A)$ is the number of connected components in $G[A]$.*

We note that [49, Lemma 8] is proved for the special case of $\zeta = \min(1/100, \delta/5)$, however the proof works for any positive $\zeta < 1/2$.

We can now prove Lemma 4.8.

Proof of Lemma 4.8. The claims about the total-variation distance (4.1) are shown in [49, Theorem 1, Items (2), (3), (8)].

Equation (4.2) shows a bit of slack in the definitions of Ω^{dis} and Ω^{ord}; it essentially follows from the same theorem. We prove the first bound; the second follows similarly. Let $q_0 \geq \mathrm{e}^{21\Delta/\eta}$ be sufficiently large that Theorems 1 and 2 of [49] hold, and choose $\zeta = 20\Delta/\log(q_0)$. Note that $\zeta < \eta$, as desired.

We show that ordered configurations with fewer than $(1-\zeta)|E|$ in-edges have exponentially small probability, using the same argument as in [49, Lemma 9]. The idea is that, by expansion, such configurations lose a constant factor in the exponent of their weight.

Let $\Omega' = \{\mathcal{F} \in \Omega^{\text{ord}} : |\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|\}$ and let

$$Z' = \sum_{\mathcal{F}\in\Omega'} w_G(\mathcal{F}) = \sum_{\mathcal{F}\in\Omega'} q^{c(\mathcal{F})}(\mathrm{e}^{\beta} - 1)^{|\operatorname{In}(\mathcal{F})|}.$$

Let $z = \max\{q, (\mathrm{e}^{\beta}-1)^{\Delta/2}\}$. Note that the partition function Z (the sum of weights of all configurations) satisfies $Z \geq z^n$. Hence, the weight of any $\mathcal{F} \in \Omega$ satisfies

$$w_G(\mathcal{F})/Z \leq z^{c(\mathcal{F})+2|\operatorname{In}(\mathcal{F})|/\Delta-n} = z^{n(c(\mathcal{F})/n+|\operatorname{In}(\mathcal{F})|/|E|-1)}.$$

Since $\zeta < \eta < 1/2$, for any $\mathcal{F} \in \Omega'$, $\zeta|E| \leq (1-\eta)|E| \leq |\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|$. By Lemma 4.2, w.h.p., $G \in \mathcal{G}_{\Delta,\delta}$. By Lemma 4.21, for $n \geq 360/(\zeta\delta)$, we have $c(\mathcal{F})/n+|\operatorname{In}(\mathcal{F})|/|E| \leq 1-\zeta/40$ so $Z'/Z \leq 2^{n\Delta/2}z^{n(-\zeta/40)}$, which gives $Z'/Z = \mathrm{e}^{-\Omega(n)}$ since $q > \mathrm{e}^{20\Delta/\zeta}$. To finish, we will show that $Z/Z^{\mathrm{ord}} = O(1)$, so that the probability of a configuration in Ω' satisfies $\pi_G^{\mathrm{ord}}(\Omega') = \frac{Z'}{Z^{\mathrm{ord}}} = \frac{Z'}{Z} \times \frac{Z}{Z^{\mathrm{ord}}} = \mathrm{e}^{-\Omega(n)}O(1) = \mathrm{e}^{-\Omega(n)}$, where Z^{ord} is the partition function in the ordered phase (the sum of weights of all ordered configurations).

If $\beta > \beta_c$, item 3 of [49, Theorem 1] says that there are positive numbers n_0 and ξ such that, for $n \geq n_0$, $(1/n)\log((Z-Z^{\mathrm{ord}})/Z) \leq -\xi$. So for $n \geq n_0$, $Z^{\mathrm{ord}}/Z \geq 1-e^{-\xi n}$, which is at least $1/2$ for $n \geq 1/\xi$.

If $\beta = \beta_c$, then items 1 and 3 of [49, Theorem 2] show that Z^{ord}/Z converges in distribution to a random variable $Q/(Q+1)$ where Q converges to q in probability. Hence, for all sufficiently large q, $Z^{\mathrm{ord}}/Z = \Theta(1)$ (and in particular bounded away from 0 and 1).

Consequently, for all $\beta \geq \beta_c'$, $pi_G^{\mathrm{ord}}(\Omega') = Z'/Z^{\mathrm{ord}} = O(1) \cdot \mathrm{e}^{-\Omega(n)} = \mathrm{e}^{-\Omega(n)}$. The proof for the disordered phase is similar, following the argument from [49, Lemma 9], and using that by [49, Theorem 1, item 2] $Z^{\mathrm{dis}}/Z = \Omega(1)$. □

This immediately gives us Key Property 1: in order for the RC dynamics initialised from the distribution within the ordered phase to leave the phase in one step, it would have to be in a state with exactly $(1-\eta)|E| \leq (1-\zeta)|E|$ edges. By the Lemma 4.8, probability of that is at most $\mathrm{e}^{-\Omega(n)}$. Similarly, in order for the RC dynamics initialised from π^{dis} to leave the disordered phase in one step, it would have to be in a state with exactly $\eta|E| \geq \zeta|E|$ edges, which is also exponentially unlikely by the Lemma.

4.3 WSM within the ordered phase

Let $\Delta \geq 5$ be an integer. Consider any sufficiently large q and any $\beta \geq \beta_c$. For sufficiently large n, choose a "radius" $r \approx \frac{1}{\beta}\log n$ and let $G = (V, E) \sim \mathcal{G}_{n,\Delta}$. Fix a vertex $v \in V$ and an edge e incident to v. We wish to show, with sufficiently high probability, that $\left\|\pi_{B_r^+(v)}(e \mapsto \cdot) - \pi_G^{\mathrm{ord}}(e \mapsto \cdot)\right\|_{\mathrm{TV}} \leq 1/(100|E|)$.

Our goal is essentially to construct a coupling of $\mathcal{F}^+ \sim \pi_{\mathcal{B}_r^+(v)}$ and $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}$, such that $\mathbb{P}(\mathcal{F}^+(e) \neq \mathcal{F}^{\mathrm{ord}}(e))$ is sufficiently small. In order to construct the coupling, we take advantage of the fact that $G[B_r(v)]$ is locally treelike. In fact, we identify a suitable subgraph of $G[B_r(v)]$ without cycles and restrict the coupling to this subgraph.

Consider a breadth-first search from v in $G[B_r(v)]$. Let T_0 be the rooted tree consisting of all forward edges in this breadth-first search. All other edges in $B_r(v)$ are called *excess edges*. W.h.p., since $G \sim \mathcal{G}_{n,\Delta}$, there are at most K excess edges in $B_r(v)$ for some absolute constant $K > 0$. In particular, since G is locally treelike, we can identify integers r_1 and r_2 satisfying $r \geq r_1 > r_2 \geq 0$ such that $E(B_{r_1}(v)) \setminus E(B_{r_2}(v))$ contains no excess edges and $r_1 - r_2 \geq r/(2K) = \Omega(r)$. The fact that $r_1 - r_2 = \Omega(r)$ ensures that $B_{r_1}(v) \setminus B_{r_2}(v)$ is a sufficiently large subgraph of G.

Let us first informally sketch the revealing process that obtains the coupling. The process is formally defined in Section 4.3.2. See Figure 1.5 for an illustration.

In the coupling, we generate a sequence of edge subsets $F_0 \subseteq F_1 \subseteq \cdots \subseteq E$ such that, after iteration i, the edges in F_i are revealed. We also construct two sequences of partial configurations $\mathcal{F}_0^+ \subseteq \mathcal{F}_1^+ \subseteq \cdots \subseteq \mathcal{F}^+$ and $\mathcal{F}_0^{\mathrm{ord}} \subseteq \mathcal{F}_1^{\mathrm{ord}} \subseteq \cdots \subseteq \mathcal{F}^{\mathrm{ord}}$, maintaining the invariant that the revealed edges in $\mathcal{F}_i^{\mathrm{ord}}$ and $\mathcal{F}_i^+$ are exactly the edges in F_i. The coupling will have the crucial property that $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}$ and $\mathcal{F}^+ \sim \pi_{\mathcal{B}_{r_1}^+(v)}$.

- The process starts with iteration $i = 0$. The initial set F_0 of revealed edges is all edges except those in $E(B_{r_1}(v))$. In $\mathcal{F}_0^{\mathrm{ord}}$, these revealed edges are sampled from the projection $\pi_{F_0}^{\mathrm{ord}}$ of π^{ord} onto F_0. It is likely that the configuration $\mathcal{F}_0^{\mathrm{ord}}$ has at least $(1-\eta)|E|$ in-edges. If not, then the coupling terminates (unsuccessfully), generating $\mathcal{F}^{\mathrm{ord}}$ and $\mathcal{F}^+$ from the right distributions. We will show that the probability of this unsuccessful termination is low. On the other hand, if $\mathcal{F}_0^{\mathrm{ord}}$ has at least $(1-\eta)|E|$ in-edges, then the process continues. All configurations refining $\mathcal{F}_0^{\mathrm{ord}}$ are in Ω^{ord}, so the conditional distributions of π and π^{ord} on unrevealed edges are the same, making it possible to maintain monotone coupling $\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}}) \subseteq \mathrm{In}(\mathcal{F}_i^+)$ in subsequent steps. At this point $\mathcal{F}_0^+$ is taken to be the configuration with revealed edges F_0 where all revealed edges are in-edges.

- After iteration $i = 0$, iterations continue with $i = 1, 2, \ldots$ until an edge is revealed whose distance from v is at most r_2 or until the in-edges in $\mathcal{F}_i^{\mathrm{ord}}$ induce a giant component, and this giant component contains all boundary vertices (as defined in Definition 4.17) of $\mathcal{F}_i^{\mathrm{ord}}$. We will show that it is very unlikely that an

edge at distance at most r_2 from v is reached. So it is likely the giant component in $\mathcal{F}_i^{\mathrm{ord}}$ contains all vertices on the boundary of F_i. This is a good situation because the distribution of π conditioned on refining $\mathcal{F}_i^{\mathrm{ord}}$ and the distribution of $\mathcal{F}^+$ conditioned on refining $\mathcal{F}_i^+$ induce the same distribution on edges incident to v, which enables us to show that $\mathbb{P}(\mathcal{F}^+(e) \neq \mathcal{F}^{\mathrm{ord}}(e))$ is sufficiently small.

- The process at iteration $i+1$ is as follows. Let W_i be the set of boundary vertices of $\mathcal{F}_i^{\mathrm{ord}}$ whose components in $G[\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}})]$ are small (contain less than a majority of vertices). If W_i is empty, then all boundary vertices belong to the giant component and the coupling finishes successfully. Otherwise, a vertex $w_i \in W_i$ is chosen to be as far from v as possible. The edges in the subtree of T_0 below the parent of w_i are revealed in F_{i+1}.

We will show that this process, which we formally define in Section 4.3.2, is unlikely to terminate unsuccessfully. To do this, we employ the polymer framework of [49], which we next define.

4.3.1 Ordered polymers

We use the definition of *ordered polymers* from [49, Section 2.4.2]. For a set of edges $A \subseteq E$, let $\mathcal{B}_0(A) = A$, and inductively for $i = 0, 1, 2, \ldots$ define $\mathcal{B}_{i+1}(A)$ to be the set of all edges such that they are either in $\mathcal{B}_i(A)$ or edges that are incident to a vertex that has at least $5\Delta/9$ incident edges in $\mathcal{B}_i(A)$. Let $\mathcal{B}_\infty(A) = \bigcup_{i\in\mathbb{N}} \mathcal{B}_i(A)$. It is shown in [49, Lemma 12] that $|\mathcal{B}_\infty(A)| \leq 10|\mathcal{B}_0(A)|$.

An *ordered polymer* is a connected subgraph $\gamma = (V(\gamma), E(\gamma))$ of G together with an RC configuration on the subgraph γ, $\ell\colon E(\gamma) \to \{0,1\}$ subject to: (i) $|\operatorname{Out}(\ell)| \leq \eta m$, and (ii) $\mathcal{B}_\infty(\operatorname{Out}(\ell)) = E(\gamma)$. In accordance with the notation in [49], we let $E_u(\gamma)$ be the set of out-edges $\operatorname{Out}(\ell)$ of the polymer and $c'(\gamma)$ be the number of components in the graph $G[E\backslash E_u(\gamma)]$ with at most $n/2$ vertices. The *size* of γ is defined as $|E(\gamma)|$, whereas the weight of γ is defined as $w_\gamma^{\mathrm{ord}} := q^{c'(\gamma)}(e^\beta - 1)^{-|E_u(\gamma)|}$.

Given a (partial) configuration $\mathcal{F}$, the set of ordered polymers of $\mathcal{F}$, denoted by $\Gamma(\mathcal{F})$, consists of the connected components of $G[\mathcal{B}_\infty(\operatorname{Out}(\mathcal{F}))]$, each with the labelling on the edges induced by the corresponding assignment in $\mathcal{F}$. Note that, since $|\operatorname{Out}(\mathcal{F})| \leq \eta|E|$, every such polymer satisfies condition (i). Moreover, since the polymers are the connected components of $G[\mathcal{B}_\infty(\operatorname{Out}(\mathcal{F}))]$, condition (ii) is also satisfied.

We will use a couple of polymer properties from [49] to prove the following lemmas. First, we note the following connection between the weight of a configuration $\mathcal{F}$ and the weight of its polymers $\Gamma(\mathcal{F})$.

Lemma 4.22. *Fix $\Delta \geq 5$, $K \geq 0$ and $\delta \in (0, 1/2)$. Let $G \in \mathcal{G}_{\Delta,\delta,K}$ and $\mathcal{F} \in \Omega^{\mathrm{ord}}$. Then $w_G(\mathcal{F}) = q(\mathrm{e}^\beta - 1)^{|E|} \prod_{\gamma \in \Gamma(\mathcal{F})} w_\gamma^{\mathrm{ord}}$.*

Proof. Note that $\mathrm{Out}(\mathcal{F}) = \bigcup_{\gamma \in \Gamma(\mathcal{F})} E_u(\gamma)$, so

$$(\mathrm{e}^\beta - 1)^{|E| - \sum_{\gamma \in \Gamma(\mathcal{F})} |E_u(\gamma)|} = (\mathrm{e}^\beta - 1)^{|\operatorname{In}(\mathcal{F})|}.$$

Hence, by [49, Lemma 21], $q^{\sum_{\gamma \in \Gamma(\mathcal{F})} c'(\gamma)} = q^{c(G[\operatorname{In}(\mathcal{F})]) - 1}$, so the result follows. □

Next we show that with sufficient probability, all polymers are of size $O(\log n)$.

Lemma 4.23. *Let $\Delta \geq 5$ and $K \geq 0$ be integers, $\delta \in (0, 1/2)$ be real. Then, for all sufficiently large q, and all $\beta \geq \log(q^{1.9/\Delta} + 1)$, the following holds for all $G \in \mathcal{G}_{\Delta,\delta,K}$ with $n = |V(G)|$. For $s \coloneqq 2000 \log(n)/\beta$,*

$$\pi_G^{\mathrm{ord}}\big(\mathcal{F} \text{ contains a polymer with size at least } s\big) \leq 2n^{-3/2}.$$

Proof. This proof closely follows the proof of Proposition 15 from [49]. Suppose that q is large enough so that $q^{1.9/\Delta} \geq (2\mathrm{e}\Delta)^{400}$ and fix arbitrary $\beta \geq \log(q^{1.9/\Delta} + 1)$.

Consider an arbitrary ordered polymer γ. For any configuration $\mathcal{F} \in \Omega^{\mathrm{ord}}$ with $\gamma \in \Gamma(\mathcal{F})$, by Lemma 4.22, it holds that $w_G(\mathcal{F}) \leq q(\mathrm{e}^\beta - 1)^{|E|} w_\gamma^{\mathrm{ord}}$. Note that for any $\mathcal{F} \in \Omega^{\mathrm{ord}}$ such that $\gamma \in \Gamma(\mathcal{F})$, there is a configuration $\mathcal{F}'$ with $\Gamma(\mathcal{F}') = \Gamma(\mathcal{F}) \setminus \{\gamma\}$ (obtained by setting $\operatorname{In}(\mathcal{F}') = \operatorname{In}(\mathcal{F}) \cup E_u(\gamma)$), which therefore has weight $w_G(\mathcal{F}') = w_G(\mathcal{F})/w_\gamma^{\mathrm{ord}}$.

Let Ω_γ be the set of all ordered configurations with a polymer γ. We can lower bound Z^{ord} by $\sum_{\mathcal{F} \in \Omega_\gamma} (w_G(\mathcal{F}) + w_G(\mathcal{F}')) = (1 + \frac{1}{w_\gamma^{\mathrm{ord}}}) \sum_{\mathcal{F} \in \Omega_\gamma} w_G(\mathcal{F})$. Therefore,

$$\pi^{\mathrm{ord}}(\gamma \in \Gamma(\cdot)) = \frac{\sum_{\mathcal{F} \in \Omega_\gamma} w_G(\mathcal{F})}{Z^{\mathrm{ord}}} \leq \frac{1}{1 + \frac{1}{w_\gamma^{\mathrm{ord}}}} \leq w_\gamma^{\mathrm{ord}},$$

so we conclude that the probability that any fixed polymer γ is a polymer of $\mathcal{F} \sim \pi_G^{\mathrm{ord}}$ is at most w_γ^{ord}.

Let u be a vertex in $V(G)$ and denote by Γ_u the set of polymers containing u. Let p_u be the probability that u is contained in a polymer of size at least s for $\mathcal{F} \sim \pi_G^{\mathrm{ord}}$. By a union bound over the polymers in Γ_u, we have that $p_u \leq P_u$, where

$$P_u \coloneqq \sum_{\gamma \in \Gamma_u;\ |E(\gamma)| \geq s} w_\gamma^{\mathrm{ord}}.$$

We next show that the weight of a polymer is exponentially decaying in its size. Together with a bound on the number of connected subgraphs with a given size, this allows us to conclude that P_u is sufficiently small.

By [49, Lemma 14], it holds that $c'(\gamma) \le 9|E_u(\gamma)|/(5\Delta)$. The bound on β ensures that

$$w_\gamma^{\mathrm{ord}} = q^{c'(\gamma)}(\mathrm{e}^\beta - 1)^{-|E_u(\gamma)|} \le (\mathrm{e}^\beta - 1)^{(\frac{9}{5\Delta}\cdot\frac{\Delta}{1.9}-1)|E_u(\gamma)|} \le (\mathrm{e}^\beta - 1)^{-|E_u(\gamma)|/20}.$$

Hence we get

$$P_u \le \sum_{\gamma\in\Gamma_u;\ |E(\gamma)|\ge s} (\mathrm{e}^\beta - 1)^{-|E_u(\gamma)|/20}.$$

By [49, Lemma 12], for any ordered polymer γ, its size satisfies $|E(\gamma)| \le 10|E_u(\gamma)|$. Thus,

$$P_u \le \sum_{k\ge s/10} \sum_{\gamma\in\Gamma_u;\ |E_u(\gamma)|=k} (\mathrm{e}^\beta - 1)^{-|E_u(\gamma)|/20}.$$

By [49, Proof of Proposition 15], the number of polymers with k unoccupied edges containing a particular vertex is at most $(2\mathrm{e}\Delta)^{10k}$, hence

$$P_u \le \sum_{k\ge s/10} \big((2\mathrm{e}\Delta)^{10}(\mathrm{e}^\beta - 1)^{-1/20}\big)^k.$$

By the lower bounds on β and q, we have $\mathrm{e}^\beta - 1 \ge q^{1.9/\Delta} \ge (2\mathrm{e}\Delta)^{400}$, so

$$P_u \le \sum_{k\ge s/10} (\mathrm{e}^\beta - 1)^{-k/40} \le \sum_{k\ge s/10} \mathrm{e}^{-\beta k/80}.$$

The last sum is a geometric series with ratio $\mathrm{e}^{-\beta/80} < 1/2$, so it is upper-bounded by twice its largest term. Since $s = \frac{2000}{\beta}\log n$, it follows that $P_u \le 2/n^{5/2}$. By a union bound over the vertices of G, the probability that there is a polymer of $\mathcal{F} \sim \pi_G^{\mathrm{ord}}$ which has size at least s is therefore at most $2/n^{3/2}$, as required. □

Next we state an important property of expanders from [68], which allows us to show that every ordered configuration has a giant component of in-edges.

Lemma 4.24 ([68, Lemma 2.3])**.** *Let $G = (V, E)$ be a regular graph and consider $E' \subseteq E$ with $|E'| \le \theta|E|$ for some $\theta \in (0, \phi_G(1/2))$. Then $(V, E\backslash E')$ has a component of size at least $\big(1 - \frac{\theta}{2\phi_G(1/2)}\big)|V|$.*

We use Lemma 4.24 to establish the existence of a giant component, which we define below.

Definition 4.25. *The size of a connected component of a graph is the number of vertices in the component. A* giant component *in an n-vertex graph is a component whose size is greater than $n/2$.*

Corollary 4.26. *Fix integers $\Delta \geq 5$, $K \geq 0$, and a real number $\delta \in (0, 1/2)$. Let G be a graph in $\mathcal{G}_{\Delta,\delta,K}$ and let $\mathcal{F}$ be a configuration in Ω^{ord} or a partial configuration with $|\operatorname{In}(\mathcal{F})| \geq (1-\eta)|E|$. Then there is a giant component in $G[\operatorname{In}(\mathcal{F})]$ whose size is at least $(1-\delta)|V|$.*

Proof. Apply Lemma 4.24 with $E' = \operatorname{Out}(\mathcal{F})$ and $\theta = \eta = \min(\delta/5, 1/100)$. Note that $|\operatorname{Out}(\mathcal{F})| \leq \eta|E|$ and $\phi_G(1/2) \geq 1/10$. Thus, the lemma says that $G[\operatorname{In}(\mathcal{F})]$ has a component of size at least $\left(1 - \frac{\delta/5}{2 \cdot 1/10}\right) n = (1-\delta)n > n/2$. □

The following technical lemma was based on the proof of [49, Lemma 21] and shows that non-giant components, alongside with "badly connected" parts of the giant component, are contained in polymers. This allows us to relate the failure of the revealing process to the existence of a large polymer.

Lemma 4.27. *Fix $\Delta \geq 5$, $K \geq 0$ and $\delta \in (0, 1/2)$. Suppose $n \geq 2/\delta$. Let $G = (V, E)$ with $|V| = n$ be a graph in $\mathcal{G}_{\Delta,\delta,K}$. Consider $\mathcal{F} \in \Omega^{\mathrm{ord}}$. Let κ be the giant component of $G[\operatorname{In}(\mathcal{F})]$*

1. *If $u \notin \kappa$ and e is incident to u, then e is in a polymer of $\mathcal{F}$.*
2. *Let S be a non-empty set of edges with $|S| < 9\Delta n/200$. Let κ' be a component of $G[\operatorname{In}(\mathcal{F}) \backslash S]$ with $|V(\kappa')| \leq n/2$. Then all but at most $45\Delta|S|$ vertices of κ' are such that all their incident edges belong to a polymer of $\mathcal{F}$.*
3. *For S and κ' as in Item 2, there are at most $50\Delta^2|S|$ connected components in $G[\mathcal{B}_\infty(\operatorname{Out}(\mathcal{F}))]$ containing a vertex of κ', i.e., there are at most $50\Delta^2|S|$ polymers of $\mathcal{F}$ containing vertices of κ'.*

Proof. We start with the proof of Item 1. Consider a non-giant component κ' of $G[\operatorname{In}(\mathcal{F})]$. Let τ be the set of vertices in κ' which have an incident edge that is not in a polymer of $\mathcal{F}$. For a contradiction, assume that τ is non-empty.

By Corollary 4.26, $|\kappa'| \leq \delta|V|$ and thus $|\tau| \leq \delta|V|$. Every edge in the cut $(\tau, V \backslash \tau)$ either connects τ to a vertex of $\kappa' \backslash \tau$ or to a vertex outside the component κ'. By the definition of τ, every vertex of $\kappa' \backslash \tau$ must have all of its incident edges in a polymer. Since κ' is a connected component of in-edges, every edge between a vertex of κ' and a vertex outside κ' must be an out-edge, and therefore belong to a polymer.

Hence, every edge in the cut $(\tau, V\backslash\tau)$ is in some polymer. By the definition of $\mathcal{G}_{\Delta,\delta,K}$, $\phi_G(\delta) \geq 5/9$, and thus $|E(\tau, V\backslash\tau)| \geq 5\Delta|\tau|/9$. Hence, there is a vertex $u \in \tau$ with at least $5\Delta/9$ incident edges in the cut, and thus, from the definition of polymers, all edges adjacent to u are in polymers, contradicting that u is in τ.

Before proving Item 2 of the lemma, we use it to prove Item 3. Let κ' be a non-giant component in $G[\mathrm{In}(\mathcal{F})\backslash S]$. As before, let τ be the set of vertices in κ' which have an incident edge that is not in a polymer of $\mathcal{F}$. By Item 2, it holds that $|\tau| \leq 45\Delta|S|$. Let E_τ be the set of edges that are incident to vertices in τ. Note that $(V(\kappa'), E(\kappa'))$ is connected, and removing an edge from a graph can increase the number of components by at most one. Thus, $(V(\kappa'), E(\kappa') \setminus E_\tau)$ has at most $1+\Delta|\tau|$ components. Any edges in these components are in polymers of $\mathcal{F}$. Thus, there are at most $1 + \Delta|\tau| \leq 50\Delta^2|S|$ components of $G[\mathcal{B}_\infty(\mathrm{Out}(\mathcal{F}))]$ that contain vertices in $V(\kappa')$, proving Item 3.

To finish, we prove Item 2. The argument is very similar to the proof of Item 1, except that we have to account for a few in-edges in the cut $(\kappa', V \setminus \kappa')$. First we bound the size of κ'. Since $\mathcal{F} \in \Omega^{\mathrm{ord}}$, $|\mathrm{In}(\mathcal{F})\backslash S| \geq (1-\eta)|E| - |S|$. Since $\eta \leq 1/100$, $\eta + \frac{|S|}{|E|} < 1/10 \leq \phi_G(1/2)$, hence we can apply Lemma 4.24 with $E' = \mathrm{Out}(\mathcal{F}) \cup S$ and $\theta = |E'|/|E|$. We find that $(V, \mathrm{In}(\mathcal{F})\backslash S)$ has a component of size at least $(1 - \frac{\theta}{2\phi_G(1/2)})n \geq (1-5\theta)n$ so the number of vertices in κ' is at most $5\theta n \leq 5(\eta + |S|/|E|)n \leq \delta n + 10|S|/\Delta$. Thus, $|\tau| \leq \delta n + 10|S|/\Delta$. Next let $\lambda = \lceil 10|S|/\Delta \rceil$ and let U be a subset of τ of size $|\tau| - \lambda \leq \delta n$. Then

$$\begin{aligned}
|E(\tau, V\backslash\tau)| &\geq |E(U, V \setminus \tau)| \\
&\geq |E(U, V \setminus U)| - |E(U, \tau \setminus U)| \\
&\geq \tfrac{5}{9}\Delta|U| - \lambda\Delta = \tfrac{5}{9}\Delta|\tau| - 14\lambda\Delta/9.
\end{aligned}$$

By the definition of τ, all edges in $E(\tau, V(\kappa') \setminus \tau)$ are in polymers of $\mathcal{F}$. Since κ' is a component of $G[\mathrm{In}(\mathcal{F}) \setminus S]$, there are at most $|S|$ edges in $\mathrm{In}(\mathcal{F})$ between τ and $V \setminus V(\kappa')$. Thus the number of edges in $E(\tau, V \setminus \tau)$ that are in polymers is at least $|E(\tau, V\backslash\tau)| - |S|$. This is at least

$$\tfrac{5}{9}\Delta|\tau| - 14\lambda\Delta/9 - |S| \geq \tfrac{5}{9}\Delta|\tau| - 149|S|/9 - 14\Delta/9 \geq \tfrac{5}{9}\Delta|\tau| - 45\Delta|S|/9,$$

where the last inequality follows from $\Delta \geq 5$, $|S| \geq 1$.

Averaged over all vertices $w \in \tau$, the number of edges in $E(\tau, V \setminus \tau)$ that are in polymers and are incident to w is at least $\frac{5}{9}\Delta - 45\Delta|S|/(9|\tau|)$. We may assume τ is non-empty. Otherwise Item 2 directly follows. Thus, there is a vertex $w \in \tau$ such

that the number of edges in $E(\tau, V \setminus \tau)$ that are in polymers and are incident to w is at least $\frac{5}{9}\Delta - 45\Delta|S|/(9\tau)$. Since this number is an integer, the number of edges in $E(\tau, V \setminus \tau)$ that are in polymers and are incident to w is at least

$$W \coloneqq \left\lceil \frac{5}{9}\Delta - \frac{45\Delta|S|}{9|\tau|} \right\rceil.$$

Suppose for contradiction that $|\tau| > 45\Delta|S|$. Then the term that is subtracted in the definition of W is less than $1/9$. On the other hand, if the first term, $5\Delta/9$ is not an integer, then its non-integer part is at least $1/9$. Either way, $W = \lceil 5\Delta/9 \rceil \geq 5\Delta/9$. As in the proof of Item 1, all incident edges of w hence must be in a polymer, contradicting that $w \in \tau$. □

4.3.2 The coupling process

Next, we formally define the coupling process sketched at the beginning of the section. We will require some additional notation.

Definition 4.28. *Let $G = (V, E)$ be a graph. Given vertices $u, v \in V$, let $d_G(v, u)$ be the distance in G from v to u. For $U \subseteq V$ and $F \subseteq E$, denote by $d_G(v, U)$ the minimum distance in G from v to any vertex in U, and by $d_G(v, F)$ the minimum distance in G to any endpoint of any edge in F. Let $\partial_G(F)$ be the set of vertices in V incident to both an edge in F and an edge in $E \setminus F$. For a rooted tree T and a vertex $u \in V(T)$, let T_u denote the subtree of T rooted at u.*

Fix positive integers $\Delta \geq 5$ and $K \geq 0$. Let $M = M(K, \Delta) = 2000 \times 400 \times (2 + K)\Delta \log(\Delta - 1)$. Fix q sufficiently large and $\beta \geq \beta_c$. Fix a graph $G = (V, E) \in \mathcal{G}_{\Delta,\delta,K}$ with $n = |V|$ sufficiently large and fix a vertex $v \in V$. Let $r \coloneqq \min(\frac{\boldsymbol{M}}{\beta}, \frac{1}{3}) \log_{\Delta-1} n$.

Definition 4.29. *Consider a breadth-first search from v in $G[B_r(v)]$. Let T_0 be the rooted tree consisting of all forward edges in this breadth-first search. We refer to edges in $E(T_0)$ as the* tree edges*. We refer to the edges in $\mathcal{E} \coloneqq E(B_r(v)) \setminus E(T_0)$ as* excess edges*.*

Observation 4.30. *For every edge $\{u, u'\}$ in $\mathcal{E}$, it holds that $d_{T_0}(v, u)$ and $d_{T_0}(v, u')$ differ by at most one. For any $u \in B_r(v)$, $d_G(v, u) = d_{T_0}(v, u)$.*

This follows from the properties of breadth-first search: vertices have the same distance from v in T_0 as in G, and the endpoints of any edge in $G[B_r(v)]$ differ in distance from v by at most one.

Since $G \in \mathcal{G}_{\Delta,\delta,K}$, it holds that $|\mathcal{E}| \leq K$ so the ball $B_r(v)$ is close to being a tree. The next observation picks out a piece of $B_r(v)$ without excess edges. See Figure 4.1 for an illustration.

Observation 4.31. *Suppose $r > K+1$. Then there are integers r_1 and r_2 satisfying $r \geq r_1 > r_2 \geq 0$ such that $E(B_{r_1}(v)) \setminus E(B_{r_2}(v))$ contains no edges from $\mathcal{E}$ and $r_1 - r_2 \geq \frac{r}{1+K} - 1$.*

Proof. For every edge e of G, let $d^+(v,e)$ denote the maximum distance in G from v to an endpoint of e. Let $\mathcal{D} = \{d^+_G(e) \mid e \in \mathcal{E}\}$. Note that $|\mathcal{D}| \leq K$. Thus the set $\{1,2,\ldots,\lfloor r \rfloor\} \setminus \mathcal{D}$ can be partitioned into at most $K+1$ intervals of contiguous distances. Thus, there are integers r_1 and r_2' with $r_1 \geq r_2'$ such that the set $\{r_2', r_2'+1, \ldots, r_1\}$ does not intersect $\mathcal{D}$ and $r_1 - r_2' + 1 \geq \frac{\lfloor r \rfloor - K}{K+1} \geq \frac{r-1-K}{K+1} = \frac{r}{K+1} - 1$.

Let $r_2 = r_2' - 1$. We claim that $(E(B_{r_1}(v)) \setminus E(B_{r_2}(v))) \cap \mathcal{E} = \emptyset$. To see this, consider any $e \in \mathcal{E}$. By the choice of r_1 and r_2, either $d^+(v,e) > r_1$ or $d^+(v,e) \leq r_2$. In the former case, an endpoint of e is outside $V(B_{r_1}(v))$, thus $e \notin B_{r_1}(v)$. In the latter case, both endpoints of e are in $B_{r_2}(v)$, thus $e \in B_{r_2}(v)$. That concludes the proof. □

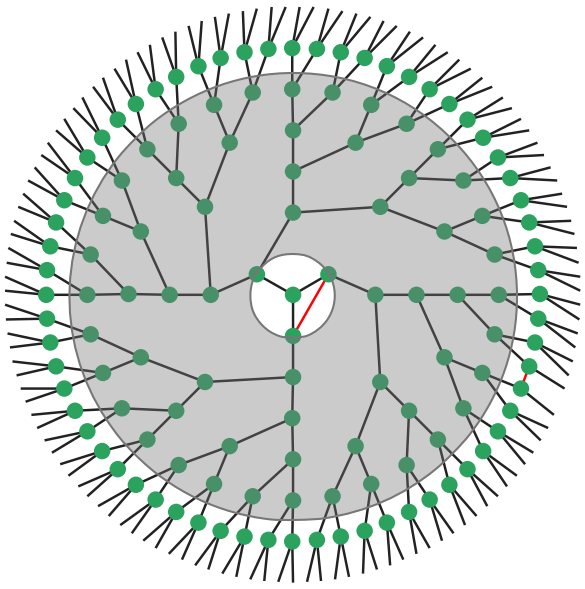


Figure 4.1: A neighbourhood with tree excess 2. The red edges are excess edges. The grey region is $B_{r_1}(v) \setminus B_{r_2}(v)$ and contains no excess edges.

From now on, we assume that n is sufficiently large, so $r > K+1$. We define r_1 and r_2 as in Observation 4.31. Note that $r_1 \geq \frac{r}{K+1} - 1$. Let $T = T_0[B_{r_1}(v)]$ be the tree

T_0 restricted to the first r_1+1 levels. To prove Theorem 4.9, we construct a coupling of $\mathcal{F}^+ \sim \pi_{\mathcal{B}^+_{r_1}(v)}$ and $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}$, such that $\mathbb{P}(\mathcal{F}^+(e) \neq \mathcal{F}^{\mathrm{ord}}(e)) \leq 1/(100|E|)$.

To do this, we generate two sequences of coupled partial configurations, starting by revealing all edges outside the ball $B_{r_1}(v)$ and progressively revealing the edges of T, starting from its leaves. We reveal edges one subtree at a time until either the boundary is contained in a giant component of in-edges, or an edge incident to $B_{r_2}(v)$ is revealed. In the first case, we use Observation 4.18 from Section 4.1.5 to couple edges incident to v perfectly. In the other case, we will be able to show the existence of a large polymer (which is an unlikely event).

Formally, we will construct a sequence of edge subsets $F_0 \subseteq F_1 \subseteq \cdots \subseteq E$, a sequence of vertex subsets $V_0, V_1, \ldots$, and two sequences of partial configurations $\mathcal{F}_0^+ \subseteq \mathcal{F}_1^+ \subseteq \cdots \subseteq \mathcal{F}^+$ and $\mathcal{F}_0^{\mathrm{ord}} \subseteq \mathcal{F}_1^{\mathrm{ord}} \subseteq \cdots \subseteq \mathcal{F}^{\mathrm{ord}}$, maintaining the following invariant for all integers $i \geq 0$:

$$R(\mathcal{F}_i^+) = R(\mathcal{F}_i^{\mathrm{ord}}) = F_i \text{ and } \mathrm{In}(\mathcal{F}_i^{\mathrm{ord}}) \subseteq \mathrm{In}(\mathcal{F}_i^+),$$

and moreover, V_i contains all boundary vertices of $\mathcal{F}_i^{\mathrm{ord}}$.

For each i, all unrevealed edges of some subtree of T are revealed. It will also hold, for each i, that $E \setminus E(B_{r_1}(v)) \subseteq F_i$.

We can now describe the process in detail.

Algorithm 4.32 (Iterative revealing process)**.**

Step 1. Let $i \coloneqq 0$. *Let* $F_0 \coloneqq E \setminus E(B_{r_1}(v))$. *Let* $\mathcal{F}_0^{\mathrm{ord}} \sim \pi_{F_0}^{\mathrm{ord}}$. *Let* $V_0 \coloneqq \{w \in V \mid d_G(v,w) = r_1\}$.

- *(a) If* $|\mathrm{In}(\mathcal{F}_0^{\mathrm{ord}})| < (1-\eta)|E|$*: Generate* $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}_{\mathcal{F}_0^{\mathrm{ord}}}$ *and generate* $\mathcal{F}_0^+ \sim \pi_{\mathcal{B}^+_{r_1}(v)}$. *Let* $\mathcal{F}^+ \coloneqq \mathcal{F}_0^+$. *Terminate (unsuccessfully).*
- *(b) Otherwise: Define* $\mathcal{F}_0^+$ *by* $R(\mathcal{F}_0^+) = \mathrm{In}(\mathcal{F}_0^+) = F_0$.

Step 2. Repeat until $d_G(F_i, v) \leq r_2$*:*

Let W_i *be the set of all vertices in* $\partial_G(F_i)$ *that are in a component of size less than* $n/2$ *in* $G[\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}})]$.

- *(a) If* W_i *is non-empty: Choose* $w_i \in W_i$ *to maximise* $d_G(w_i, v)$. *Let* p_i *be the parent of* w_i *in* T. *Let* $F_{i+1} \coloneqq F_i \cup E(T_{p_i})$. *Generate, with optimal monotone coupling,* $\mathcal{F}_{i+1}^{\mathrm{ord}} \sim \pi_{\mathcal{F}_i^{\mathrm{ord}}, F_{i+1}}$ *and* $\mathcal{F}_{i+1}^+ \sim \pi_{\mathcal{F}_i^+, F_{i+1}}$. *Let* $V_{i+1} \coloneqq (V_i \setminus V(T_{p_i})) \cup \{p_i\}$. *Let* $i \coloneqq i+1$.

(b) *Otherwise: Generate, optimally coupled,* $\mathcal{F}^{\mathrm{ord}} \sim \pi_{\mathcal{F}_i^{\mathrm{ord}}}$ *and* $\mathcal{F}^{+} \sim \pi_{\mathcal{F}_i^{+}}$ *and terminate (successfully).*

Step 3. Generate, optimally coupled, $\mathcal{F}^{\mathrm{ord}} \sim \pi_{\mathcal{F}_i^{\mathrm{ord}}}$ *and* $\mathcal{F}^{+} \sim \pi_{\mathcal{F}_i^{+}}$*. Terminate (unsuccessfully).*

We next derive some useful observations justifying the definition of the process.

Observation 4.33. *Suppose that the process 4.32 does not terminate in Step 1a (with* $i = 0$*). Then it maintains the following invariants for all* $i \geq 0$ *such that* $\mathcal{F}_i^{\mathrm{ord}}, \mathcal{F}_i^{+}, F_i$ *and* V_i *are defined.*

Inv1(i). $R(\mathcal{F}_i^{\mathrm{ord}}) = R(\mathcal{F}_i^{+}) = F_i$, $\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}}) \subseteq \mathrm{In}(\mathcal{F}_i^{+})$, *and* $E \setminus E(B_{r_1}(v)) \subseteq F_i$,

Inv2(i). For all distinct $u_1, u_2 \in V_i$, $V(T_{u_1}) \cap V(T_{u_2}) = \emptyset$,

Inv3(i). $\bigcup_{u \in V_i} V(T_u) = V(F_i) \cap V(T)$ *and* $\bigcup_{u \in V_i} E(T_u) \subseteq F_i$,

Inv4(i). If $d_G(F_i, v) > r_2$, *then* $\partial_G(F_i) \subseteq V_i$,

Inv5(i). If F_{i+1} *is defined, then* $F_i \subsetneq F_{i+1}$, *and*

Inv6(i). If $\mathcal{F}_{i+1}^{\mathrm{ord}}$ *and* $\mathcal{F}_{i+1}^{+}$ *are defined in the process, then* $\mathcal{F}_i^{\mathrm{ord}} \subseteq \mathcal{F}_{i+1}^{\mathrm{ord}}$ *and* $\mathcal{F}_i^{+} \subseteq \mathcal{F}_{i+1}^{+}$

Invariants1, 5 and 6 say that the process is well defined: in each step the set of revealed edges F_i can only increase, and both wired and ordered partial configurations reveal the same set of edges and maintain monotonicity. Inv2-3 say that the revealed edges in the ball form (disjoint) subtrees, with vertices in V_i being the roots of these subtrees. Inv 4 additionally says that, until the process reaches $B_{r_2}(v)$, all boundary vertices belong to V_i. See Figure 4.2 for an illustration.

Proof of Observation 4.33. We do an induction on i. For the base case, let $i = 0$. Inv1(0) holds by the definitions of F_0, $\mathcal{F}_0^{\mathrm{ord}}$, and $\mathcal{F}_0^{+}$. Inv2(0) follows from the fact that any $u \in V_0$ is a leaf of T and thus T_u contains a single vertex, namely, u.

The second part of Inv3(0) is trivial as none of the subtrees has any edges. We show the first part of Inv3(0) by proving that for all $u \in V(T)$, it holds that $u \in V_0$ if and only if $u \in V(F_0)$. First, suppose that $u \in V(F_0) \cap V(T)$. The definitions of F_0 and T imply that $d_G(u, v) = r_1$ so $u \in V_0$, as desired. For the other direction, consider $u \in V_0$. We need to show that G has an edge $\{u, u'\}$ with $d_G(u', v) > r_1$. The vertex u is a leaf of T, hence u has a single incident edge in $E(T)$. Given that G is Δ-regular with $\Delta > 1$, there is some edge $\{u, u'\}$, such that u' is not u's parent in

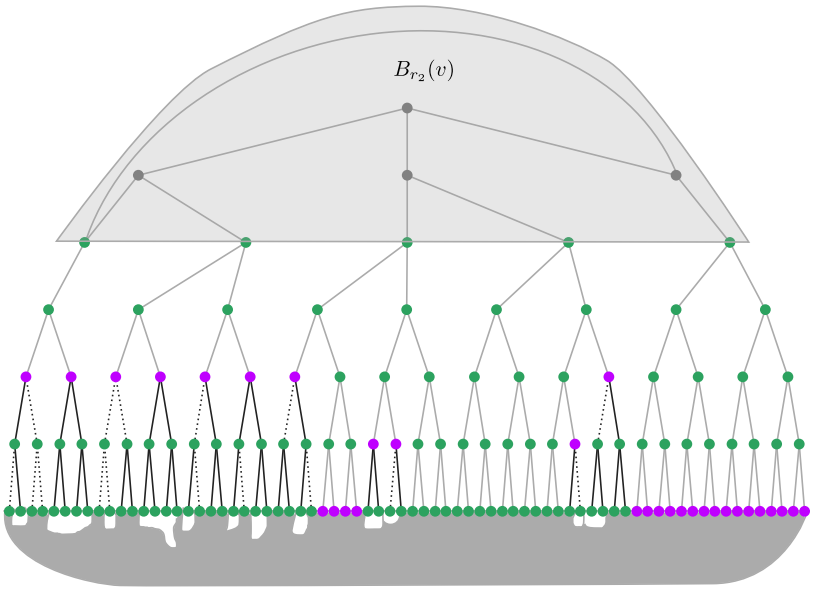


Figure 4.2: Illustration of the quantities appearing in Observation 4.33. The revealed edges F_i are black, whereas the unrevealed edges are grey. The vertices in V_i are marked purple, and $B_{r_2}(v)$ is marked in grey. The revealed edges form subtrees that are rooted in vertices of V_i.

T. By Observation 4.30, $d_G(u', v) \in \{r_1 - 1, r_1, r_1 + 1\}$. By the choice of r_1, r_2, and u', $\{u, u'\} \notin E(B_{r_1}(v)) \setminus E(B_{r_2}(v))$, so $d_G(u', v) = r_1 + 1 > r_1$, and the result follows.

The invariant Inv4(0) follows from the fact that, for any $u \in \partial_G(F_0)$, there are edges $\{u, u_1\} \in F_0$ and $\{u, u_2\} \notin F_0$. Since $F_0 = E \setminus E(B_{r_1}(v))$, it follows that $d_G(u, v) \leq r_1$ and $d_G(u_2, v) \leq r_1$, whereas $d_G(u_1, v) > r_1$. By Observation 4.30, it follows that $d_G(u, v) = r_1$ and thus $u \in V_0$.

For the inductive step, suppose F_{i+1}, $\mathcal{F}^{\mathrm{ord}}_{i+1}$, $\mathcal{F}^{+}_{i+1}$ and V_{i+1} are all defined and suppose that Inv1(i)-Inv4(i) hold. We now show Inv5(i), Inv6(i) and Inv1($i+1$)–Inv4($i+1$).

For Inv5(i), $F_i \subseteq F_{i+1}$ follows by construction. For the strict inequality, it is enough to show that $\{w_i, p_i\} \in F_{i+1} \setminus F_i$. Since F_{i+1} is defined, $d_G(F_i, v) > r_2$, otherwise the process would terminate before the $(i+1)$th iteration of Step 2. Thus by Inv4(i) and by $w_i \in \partial_G(F_i)$, $w_i \in V_i$. So for any $u \in V_i$, Inv2(i) implies that $p_i \notin V(T_u)$. Hence Inv3(i) implies that $p_i \notin V(F_i)$ so $\{w_i, p_i\} \notin F_i$. But $E(T_{p_i}) \subseteq F_{i+1}$, hence $\{w_i, p_i\} \in F_{i+1}$, completing the proof of Inv5(i). Inv6(i) follows by construction.

The first and third parts of Inv1($i+1$) follow by construction. The second part follows from Corollary 4.20 (see Section 4.1.6), using $\mathrm{In}(\mathcal{F}^{\mathrm{ord}}_i) \subseteq \mathrm{In}(\mathcal{F}^{+}_i)$ and the fact that the coupling is optimal. Inv2($i+1$) follows by construction and Inv2(i). Also, the second part of Inv3($i+1$) follows by construction and Inv3(i).

For the first part of Inv3$(i+1)$, by the definition of F_{i+1}, if $u \in V(F_{i+1}) \cap V(T)$, then $u \in V(T_{p_i})$ or $u \in V(F_i) \cap V(T)$. In the first case, note $p_i \in V_{i+1}$. Otherwise, by Inv3(i), there is a vertex $u' \in V_i$ with $u \in V(T_{u'})$. If u' is in $V(T_{p_i})$, then so is u, hence the first case applies. If not, then $u' \in V_{i+1}$. For the converse, for $u \in V_{i+1}$, there are two cases. If $u = p_i$ then $E(T_{p_i}) \subseteq F_{i+1}$ so $V(T_{p_i}) \subseteq V(F_{i+1}) \cap V(T)$. Otherwise, $u \in V_i$, so the result follows by Inv3(i) and Inv5(i).

Finally we prove Inv4$(i+1)$. Suppose that $d_G(F_{i+1}, v) > r_2$ and let $u \in \partial_G F_{i+1}$. Our goal is to show $u \in V_{i+1}$. Let $u_1, u_2 \in V$ be such that $\{u, u_1\} \in F_{i+1}$ and $\{u, u_2\} \notin F_{i+1}$. By Inv1$(i+1)$, $E \setminus E(B_{r_1}(v)) \subseteq F_{i+1}$, and so $u, u_2 \in V(B_{r_1}(v)) = V(T)$. Also, $d_G(u, v) > r_2$ implies $\{u, u_2\} \in E(B_{r_1}(v)) \setminus E(B_{r_2}(v))$. Since $\{u, u_1\} \in F_{i+1}$, $u \in V(F_{i+1})$, so by Inv3$(i+1)$, there is a vertex $u' \in V_{i+1}$ such that $u \in V(T_{u'})$. By the choice of r_1 and r_2, $\{u, u_2\}$ is a tree edge, so if $u \neq u'$, by the second part of Inv3$(i+1)$, $\{u, u_2\}$ would be then in F_{i+1}, contradicting the choice of u_2. Thus $u = u' \in V_{i+1}$. □

We next verify that the process terminates with correct distributions.

Observation 4.34. *The process eventually terminates, with* $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}$ *and* $\mathcal{F}^{+} \sim \pi_{\mathcal{B}^{+}_{r_1}(v)}$.

Proof. By Inv5, the process always terminates. If the process terminates in Step 1a, then, by construction, the resulting configurations are drawn from the correct distributions. Otherwise, Step 2 is executed. Thus, $(1-\eta)|E| \leq |\operatorname{In}(\mathcal{F}_0^{\mathrm{ord}})|$. By Inv6, for every value of i until termination, $(1-\eta)|E| \leq |\operatorname{In}(\mathcal{F}_0^{\mathrm{ord}})| \leq |\operatorname{In}(\mathcal{F}_i^{\mathrm{ord}})|$. This implies that every $\Omega_{\mathcal{F}_i^{\mathrm{ord}}} \subseteq \Omega^{\mathrm{ord}}$ so $\pi^{\mathrm{ord}}_{\mathcal{F}_i^{\mathrm{ord}}} = \pi_{\mathcal{F}_i^{\mathrm{ord}}}$. From this it follows that $\mathcal{F}^{\mathrm{ord}} \sim \pi^{\mathrm{ord}}$. Given how $\mathcal{F}_0^{+}$ is generated in Step 1b, it holds that, for any $\mathcal{F} \in \Omega$, $\pi_{\mathcal{B}^{+}_{r_1}(v)}(\mathcal{F}) > 0$ precisely when $\mathcal{F}_0^{+} \subseteq \mathcal{F}$. Then $\mathcal{F}^{+} \sim \pi_{\mathcal{B}^{+}_{r_1}(v)}$ follows from the way that $\mathcal{F}_i^{+}$ is generated in the process. □

Next we show that if the process terminates successfully, then prior to revealing the central edge e, the boundary vertices are connected in a single, giant component of revealed in-edges.

Observation 4.35. *If the process 4.32 terminates in Step 2b, then there is a non-negative integer* i *such that* $G[\operatorname{In}(\mathcal{F}_i^{\mathrm{ord}})]$ *has a giant component containing all vertices in* $\partial_G(F_i)$. *Also,* $\mathcal{F}^{\mathrm{ord}}$ *and* $\mathcal{F}^{+}$ *agree on the edges that are incident to* v.

Proof. The only way that the process can get to Step 2b, with some index i, is if $d_G(F_i, v) > r_2$ and $W_i = \emptyset$. In this case, all vertices of the boundary $\partial_G(F_i)$ are contained in the giant component of $G[\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}})]$.

Since $r_2 \geq 0$, $d_G(F_i, v) > r_2 \geq 0$, so the edges that are incident to v are not in F_i. Since G is connected, the boundary $\partial_G(F_i)$ is non-empty.

Thus $\xi(\mathcal{F}_i^{\mathrm{ord}})$ consists of a single equivalence class containing all vertices of $\partial_G(F_i)$. By Inv1(i), $\mathrm{In}(\mathcal{F}_i^{\mathrm{ord}}) \subseteq \mathrm{In}(\mathcal{F}_i^{+})$ and $R(\mathcal{F}_i^{\mathrm{ord}}) = R(\mathcal{F}_i^{+})$, so $G[\mathrm{In}(\mathcal{F}_i^{+})]$ has a giant component containing all vertices in $\partial_G(F_i)$ and $\xi(\mathcal{F}_i^{+})$ contains a single equivalence class, containing all vertices in $\partial_G(F_i)$. By Observation 4.18, $\pi_{\mathcal{F}_i^{\mathrm{ord}}}$ and $\pi_{\mathcal{F}_i^{+}}$ have the same projection on $E \backslash F_i$, meaning that for any partial configuration $\mathcal{A} \in \Omega_{E\backslash F_i}$, $\pi_{\mathcal{F}_i^{\mathrm{ord}}}(\mathcal{F}_i^{\mathrm{ord}} \cup \mathcal{A}) = \pi_{\mathcal{F}_i^{+}}(\mathcal{F}_i^{+} \cup \mathcal{A})$. Since the coupling in Step 2b is optimal, $\mathcal{F}^{\mathrm{ord}}$ and $\mathcal{F}^{+}$ agree on the edges incident to v (since they are not in F_i). □

The last observation tracks how the revealing process progresses towards v. If a subtree rooted at w_i is revealed, then there is a path in T from a vertex at distance r_1 from v to w_i, such that every vertex on this path was chosen as some w_j in an earlier iteration of the process

Observation 4.36. *Suppose that Step 2a is executed with index i and let $k = r_1 - d_G(w_i, v)$. There are non-negative integers $j_0, \ldots, j_k$ with $j_k = i$ such that, for every $\ell \in \{0, \ldots, k-1\}$, $j_\ell < j_{\ell+1}$ and $w_{j_{\ell+1}} = p_{j_\ell}$.*

Proof. By construction, $w_i \in \partial_G(F_i)$ and $d_G(F_i, v) > r_2$. By Inv4(i), $\partial_G(F_i) \subseteq V_i \subseteq V(T)$. Hence $d_G(w_i, v) \leq r_1$, so $k \geq 0$. The proof is by induction on k.

For the base case, $k = 0$, it suffices to define $j_0 \coloneqq i$, and the condition in the statement of the observation is vacuous.

For the induction step, fix $k \geq 0$ and assume that the observation holds whenever Step 2a is executed with any index j such that $r_1 - d_G(w_j, v) \leq k$. Suppose that Step 2a is executed with index i such that $r_1 - d_G(w_i, v) = k + 1$. Define $j_{k+1} \coloneqq i$. By Invariant Inv4(i), $w_i \in V_i$. By the construction of V_i, either $d_G(w_i, v) = r_1$, or there is an index $i' < i$ such that $w_i = p_{i'}$. The first case is ruled out by $d_G(w_i, v) = r_1 - (k+1) < r_1$. In the second case, w_i is the parent of $w_{i'}$ in the BFS tree, so $d_G(w_{i'}, v) = d_G(w_i, v) + 1$ and $k = r_1 - d_G(w_{i'}, v)$. We finish by applying the inductive hypothesis with index i'. □

Equipped with these observations, we can now characterise an unsuccessful termination of the process: either the ordered configuration is too close to the boundary of the phase, or it contains a large polymer.

Lemma 4.37. *Fix $\Delta \geq 5$ and $K, M > 0$. Suppose that $\beta \geq 3M$. Suppose that n is sufficiently large so that $r \coloneqq \frac{M}{\beta} \log_{\Delta-1} n > K$. and $|B_r(v)| \leq 9\Delta n/200$. Define r_1 as in Observation 4.31 so that $r_1 \geq \frac{r}{K+1} - 1$. Let $\mathcal{F}^{\mathrm{ord}}$ and $\mathcal{F}^{+}$ be generated by the process. Then at least one of the following conditions holds.*

1. *$\mathcal{F}^{\mathrm{ord}}$ and $\mathcal{F}^{+}$ agree on the edges that are incident to v.*
2. *$|\operatorname{In}(\mathcal{F}^{\mathrm{ord}}) \setminus E(B_{r_1}(v))| < (1-\eta)|E|$.*
3. *$\mathcal{F}^{\mathrm{ord}}$ contains a polymer of size at least $\frac{r}{400\Delta(1+K)} - 1$.*

Proof. By Observation 4.34, the process eventually terminates. If it terminates in Step 1a, then $|\operatorname{In}(\mathcal{F}^{\mathrm{ord}}) \setminus E(B_{r_1}(v))| < (1-\eta)|E|$, thus Condition 2 is satisfied. If the process terminates in Step 2b, then by Observation 4.35, Condition 1 holds.

For the rest of the proof, we assume that the process terminates in Step 3. We will show that Condition 3 holds. Let i' be the index that triggers the condition $d_G(F_{i'}, v) \leq r_2$ in Step 2. Let $i \coloneqq i' - 1$ so that step 2a is run for the last time with index i. By the choice of i', $d_G(F_i, v) > r_2$. Since p_i is an endpoint of an edge in $F_{i'}$, $d_G(p_i, v) \leq r_2$. Since $w_i \in \partial_G(F_i)$, $d_G(v, w_i) > r_2$. Since $d_G(w_i, v) = d_G(p_i, v) + 1$, we conclude that $d_G(w_i, v) = r_2 + 1$. By Observation 4.36 with index i and $k = r_1 - (r_2 + 1)$, there are non-negative integers $\{j_0, \ldots, j_k\}$ with $j_k = i$ such that, for every $\ell \in \{0, \ldots, k-1\}$, $j_\ell < j_{\ell+1}$ and $p_{j_\ell} = w_{j_{\ell+1}}$. Thus $w_{j_{\ell+1}} \in V_{\ell+1}$. Note that $w_{j_0}, \ldots, w_{j_k}$ is a path in T and that w_{j_k} is closest to the root v in this path. By Observation 4.31 and the definition of k, it holds that $k \geq r/(K+1) - 2$. Let $r' = \lfloor k/2 \rfloor$, $J' = \{j_\ell \mid r' \leq \ell \leq k\}$ and $W' = \{w_j \mid j \in J'\}$. Since $V(T) \subseteq V(B_{r_1}(v))$, we have established that for all $w \in W'$ $r_2 + 1 \leq d_G(v, w) \leq r_1$. We now distinguish three cases.

Case 1. For all $j \in J'$, w_j is not in a giant component of $G[\operatorname{In}(\mathcal{F}^{\mathrm{ord}})]$.

Let F' be the set of edges with endpoints in W' and note that F' contains the edges in the path $w_{j_{r'}}, \ldots, w_{j_k}$. By Observation 4.34, $\mathcal{F}^{\mathrm{ord}} \in \Omega^{\mathrm{ord}}$. Applying Item 1 of Lemma 4.27 to each vertex $u \in W'$, we find that every edge in F' is contained in a polymer of $\mathcal{F}^{\mathrm{ord}}$. Since the vertices in W' form a path in G, the edges in F' are all in the same polymer of $\mathcal{F}^{\mathrm{ord}}$. Thus, $\mathcal{F}^{\mathrm{ord}}$ has a polymer of size at least $k/2 \geq r/(2(K+1)) - 1$.

Case 2. There is an index $j \in J'$ such that w_j is in the giant component of $G[\operatorname{In}(\mathcal{F}^{\mathrm{ord}})]$ and the giant component of $G[\operatorname{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{w_j, p_j\}]$.

Since w_j was not in the giant component of revealed in-edges by the time the subtree of its parent was revealed, and we assume it was not connected to the giant

by its parent edge in the tree, it must be connected to the giant *via the outside of the ball*, and as a consequence, it has to be in a subgraph connected to the giant by only a few in-edges. More formally, we make the following claim.

Claim 2a: *Under the conditions of Case 2, there is a non-empty set S of at most $9\Delta|V|/200$ edges of G such that the size of the component κ' of w_j in the graph $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})\backslash S]$ is less than $n/2$ and furthermore $|V(\kappa')| \geq |S|r'$.*

Before proving Claim 2a, we show that it implies Condition 3, completing the proof of Case 2. By Lemma 4.27, item 2, all but at most $45\Delta|S|$ vertices of κ' are such that all their incident edges belong to a polymer. Thus, at least $|S|r' - 45\Delta|S|$ vertices of κ' are such that all their incident edges are in a polymer. By Lemma 4.27, item 3, there are at most $50\Delta^2|S|$ polymers of $\mathcal{F}^{\mathrm{ord}}$ containing vertices of κ'. We conclude that there is a polymer of $\mathcal{F}^{\mathrm{ord}}$ containing at least $\frac{|S|r'-45\Delta|S|}{50\Delta^2|S|}$ vertices of κ' with the property that all of their incident edges are in the polymer.

Since the number of edges incident to a set of size z is at least $\Delta z/2$, there is a polymer of $\mathcal{F}^{\mathrm{ord}}$ containing a vertex of $B_{r_1}(v)$ with size at least

$$\frac{\Delta}{2} \times \frac{|S|r' - 45\Delta|S|}{50\Delta^2|S|} \geq \frac{r' - 45\Delta}{100\Delta} \geq \frac{k/2 - 50\Delta}{100\Delta} \geq \frac{r}{400\Delta(1+K)} - 1.$$

To conclude the proof of Case 2, we prove Claim 2a. Since the process ends in Step 3, Step 1a does not occur, so $|\mathrm{In}(\mathcal{F}_0^{\mathrm{ord}})| \geq (1-\eta)|E|$. Since $\mathcal{F}_0^{\mathrm{ord}} \subseteq \mathcal{F}_j^{\mathrm{ord}}$ from Inv6, $\mathrm{In}(\mathcal{F}_0^{\mathrm{ord}}) \subseteq \mathrm{In}(\mathcal{F}_j^{\mathrm{ord}})$, so $|\mathrm{In}(\mathcal{F}_j^{\mathrm{ord}})| \geq (1-\eta)|E|$. By Corollary 4.26, $G[\mathrm{In}(\mathcal{F}_j^{\mathrm{ord}})]$ has a (unique) giant component, which we call κ_j. Similarly, $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$ has a giant component, which we call κ. Note that $V(\kappa_j) \subseteq V(\kappa)$. Since w_j was chosen in Step 2a, $w_j \in V(\kappa)\backslash V(\kappa_j)$.

Let Γ be the set of all paths from w_j to $V(\kappa_j)$ in the graph $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$. Every path $\gamma \in \Gamma$ contains an edge in $E \setminus F_j$. Let $\{u_1^\gamma, u_2^\gamma\}$ be the first such edge in γ, with $d_G(u_1^\gamma, w_j) < d_G(u_2^\gamma, w_j)$. Let $S \coloneqq \{\{u_1^\gamma, u_2^\gamma\} \mid \gamma \in \Gamma\}$. This is a set that contains the first unrevealed edge on each path from w_j to the vertices of the giant component κ_j.

By the definition of S, any path from w_j to κ_j goes through an edge in S, thus w_j is not connected to κ_j in $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus S]$. Let κ' be the component of w_j in $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus S]$. The component κ' contains no vertices of κ_j. Since $|V(\kappa_j)| > n/2$, we get $|V(\kappa')| < n/2$.

We conclude by proving that $|V(\kappa') \cap B_{r_1}(v)| \geq |S|r'$. The proof follows from a sequence of claims, showing that the edges in S correspond to distinct disjoint subtrees, each containing at least r' vertices of κ'. Recall that for an edge $\{u_1^\gamma, u_2^\gamma\}$ we choose the vertex labelling so that $d_G(u_1^\gamma, w_j) < d_G(u_2^\gamma, w_j)$.

Claim 2b: *For any $\gamma \in \Gamma$, $u_1^\gamma \in \partial_G(F_j)$ and $d_G(v, u_1^\gamma) > r_2$.*

The proof of Claim 2b is as follows. If $u_1^\gamma = w_j$, then $u_1^\gamma \in \partial_G(F_j)$ by the choice of w_j. Otherwise, since $\{u_1^\gamma, u_2^\gamma\}$ is the first edge on γ that is not in F_j, the edges on the path from w_j to u_1^γ are in F_j, so there is at least one edge from F_j incident to u_1^γ, and the edge $\{u_1^\gamma, u_2^\gamma\} \notin F_j$ is also incident to u_1^γ, so $u_1^\gamma \in \partial_G(F_j)$. In either case, $d_G(v, u_1^\gamma) \geq d_G(v, F_j) > r_2$.

Claim 2c: *For any $\gamma \in \Gamma$, $d_G(v, u_1^\gamma) \leq r_1$, $d_G(v, u_2^\gamma) \leq r_1$, $u_1^\gamma \in V_j$, and u_2^γ is the parent of u_1^γ in $V(T)$.*

To prove Claim 2c, consider $\gamma \in \Gamma$. Since $\{u_1^\gamma, u_2^\gamma\} \in E \setminus F_j$, both u_1^γ and u_2^γ are in $V(B_{r_1}(v)) = V(T)$. Since Step 2 is executed with index j, it holds that $d_G(F_j, v) > r_2$. So by Claim 2b and Inv4(j), $u_1^\gamma \in V_j$. By Invariant Inv3(j), $E(T_{u_1^\gamma}) \subseteq F_j$, so the only candidate for u_2^γ is the parent of u_1^γ in T.

Claim 2d: $|S| = |\{u_1^\gamma \mid \gamma \in \Gamma\}|$.

By Claim 2c, the mapping of a vertex $u \in \{u_1^\gamma \mid \gamma \in \Gamma\}$ to its parental edge in T is a bijection between $\{u_1^\gamma \mid \gamma \in \Gamma\}$ and $S = \{\{u_1^\gamma, u_2^\gamma\} \mid \gamma \in \Gamma\}$.

Claim 2e: *For any $\gamma \in \Gamma$, it holds that $|V(\kappa') \cap V(T_{u_1^\gamma})| \geq r'$.*

To prove Claim 2e, consider $\gamma \in \Gamma$. We consider two cases. First, suppose $u_1^\gamma \neq w_j$. By construction $w_j \in V_j$ and from Claim 2c, $u_1^\gamma \in V_j$. By Inv2(j), $V(T_{w_j}) \cap V(T_{u_1^\gamma}) = \emptyset$. At the beginning of the proof, we established $r_2 + 1 \leq d_G(v, w_j) \leq r_1$. By Claims 2b and 2c, $r_2 + 1 \leq d_G(v, u_1^\gamma) \leq r_1$. Since the path γ goes from w_j to u_1^γ without leaving F_j, it goes from w_j to a leaf of T_{w_j} and it finishes by going from a leaf of $T_{u_1^\gamma}$ to u_1^γ. All of these edges are in κ', and there are at least $r_1 - d_G(v, u_1^\gamma)$ of them. Since, by Claim 2b, $u_1^\gamma \in \partial_G(F_j)$ but, by construction, it is not in κ_j, by the definition of W_j, $u_1^\gamma \in W_j$. By the choice of w_j, $d_G(u_1^\gamma, v) \leq d_G(w_j, v) \leq r_1 - r'$. Thus $|V(\kappa') \cap V(T_{u_1^\gamma})| \geq r_1 - (r_1 - r') = r'$.

For the second case, suppose $u_1^\gamma = w_j$. By the assumption of Case 2, w_j is in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{w_j, p_j\}]$, thus S contains an edge other than the edge $\{u_1^{\gamma'}, u_2^{\gamma'}\}$ with $\gamma' \in \Gamma$ that is not equal to the edge $\{w_j, p_j\}$. By Claim 2c, $u_1^{\gamma'} \neq w_j$. Applying the argument from the first case to γ', the path γ' starts by going from w_j to a leaf of T_{w_j} so it contains at least r' edges of T_{w_j}, all of which are in κ_j and thus in κ'.

We now use Claims 2d and 2e to finish the proof that $|V(\kappa')| \geq |S| r'$, which completes the proof of Claim 2a, and hence the proof of Case 2. Consider paths γ and γ' in Γ with $u_1^\gamma \neq u_1^{\gamma'}$. By Inv2($j$), $V(T_{u_1^\gamma}) \cap V(T_{u_1^{\gamma'}}) = \emptyset$, therefore $|V(\kappa')| \geq \sum_{u \in \{u_1^\gamma \mid \gamma \in \Gamma\}} |V(\kappa') \cap V(T_{u_1^\gamma})|$. Claim 2e shows that each term in the sum is at least

r'. Claim 2d shows that $|S|$ is equal to the number of terms in the sum. Therefore we obtain $|V(\kappa')| \geq |S|r'$, as required.

This finishes the proof for the second case. The last remaining case is if every vertex on the path that gets connected to the giant component of in-edges in the resulting ordered configuration does so via the edge to its parent.

Case 3. There is an index $j \in J'$ such that w_j is in the giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$. However, for all $j \in J'$, w_j is not in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{w_j, p_j\}]$.

Let κ be the giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$. Let $j_\ell = \min\{j \in J' \mid w_j \in V(\kappa)\}$.

Claim 3: *For each ℓ' satisfying $\ell \leq \ell' < k$, $w_{j_{\ell'}} \in V(\kappa)$ and $\{w_{j_{\ell'}}, w_{j_{\ell'+1}}\} \in \mathrm{In}(\mathcal{F}^{\mathrm{ord}})$.*

We prove Claim 3 by induction on ℓ'. The base case is $\ell' = \ell$. In this case, $w_{j_\ell} \in V(\kappa)$ is from the definition of j_ℓ. Then since w_{j_ℓ} is in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$ but is not in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})\setminus\{w_{j_\ell}, p_{j_\ell}\}]$, the edge $\{w_{j_\ell}, p_{j_\ell}\} = \{w_{j_\ell}, w_{j_{\ell+1}}\} \in \mathrm{In}(\mathcal{F}^{\mathrm{ord}})$, as required. For the induction step, fix ℓ' satisfying $\ell \leq \ell' < k-1$ and assume Claim 3 for ℓ'. Since Claim 3 implies $w_{j_{\ell'}} \in \kappa$ and $\{w_{j_{\ell'}}, w_{j_{\ell'+1}}\} \in \mathrm{In}(\mathcal{F}^{\mathrm{ord}})$, $w_{j_{\ell'}}$ and $w_{j_{\ell'+1}}$ are in the same component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$, namely, κ. Then since $w_{j_{\ell'+1}}$ is in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})]$ but is not in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}})\setminus\{w_{j_{\ell'+1}}, p_{j_{\ell'+1}}\}]$, the edge $\{w_{j_{\ell'+1}}, p_{j_{\ell'+1}}\} = \{w_{j_{\ell'+1}}, w_{j_{\ell'+2}}\} \in \mathrm{In}(\mathcal{F}^{\mathrm{ord}})$, as required.

Claim 3 implies that for every ℓ' satisfying $\ell \leq \ell' \leq k$, $w_{j_{\ell'}} \in V(\kappa)$. Let $e = \{w_{j_k}, p_{j_k}\}$. Since w_{j_k} is not in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{e\}]$ and the edge e is not on the path $w_{j_\ell}, \ldots, w_{j_k}$, it follows that no vertices in W' are in a giant component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{e\}]$. We now consider two cases.

Case 3a. $k - \ell \geq r'/2$.

Take $S = \{e\}$. Suppose that $|V|$ is sufficiently large that $|S| < 9\Delta|V|/200$. For any $j \in J'$ let κ_j be the component containing w_j in $G[\mathrm{In}(\mathcal{F}^{\mathrm{ord}}) \setminus \{w_j, p_j\}]$. Item 2 of Lemma 4.27 implies that all but at most 45Δ vertices of κ_j are such that all their incident edges belong to a polymer of $\mathcal{F}^{\mathrm{ord}}$. Item 3 implies that there are at most $50\Delta^2$ polymers containing vertices of κ_j. Since $k - \ell \geq r'/2$, we conclude that there are at least $r'/2 - 45\Delta$ vertices of $W' \cap V(\kappa)$ such that all their incident edges are in a polymer, and at most $50\Delta^2$ polymers containing them. Hence, there is a polymer of size at least $\frac{\Delta}{2} \times \frac{r'/2-45\Delta}{50\Delta^2} \geq \frac{r'-90\Delta}{200\Delta} \geq \frac{r}{400\Delta(1+K)} - 1$.

Case 3b. $k - \ell < r'/2$.

We have already seen that $w_{j_{r'}}, \ldots, w_{j_{\ell-1}}$ is a path in G with length at least $\ell - r' \geq r'/2$ and its vertices are not in $V(\kappa)$. By Item 1 of Lemma 4.27, every edge

incident to one of these vertices is contained in a polymer of $\mathcal{F}^{\mathrm{ord}}$. Since the vertices form a path in G, the edges adjacent to them are all in the same polymer of $\mathcal{F}^{\mathrm{ord}}$. Thus, $\mathcal{F}^{\mathrm{ord}}$ has a polymer of size at least $(r'/2) \geq \frac{(k/2)-1}{2} \geq \frac{r}{4(K+1)} - 1$. □

Next we will show that the termination condition from Step 1 of the process is unlikely to happen.

Lemma 4.38. *Let $\Delta \geq 5$ and $K \geq 0$ be integers. Let $\delta \in (0, 1/2)$ be a real number. There are positive real numbers $q_0 \geq 1$ and n_0 such that the following holds for all $q \geq q_0$ and all $\beta \geq \beta_c$. Let $G = (V, E) \in \mathcal{G}_{\Delta,\delta,K}$ with $n = |V| \geq n_0$ and let v be a vertex in V. Let r be a real number satisfying $r \leq \frac{1}{3}\log_{\Delta-1}(n)$. Then $\pi_G^{\mathrm{ord}}(|\operatorname{In}(\mathcal{F})\backslash E(B_r(v))| < (1-\eta)|E|) = e^{-\Omega(n)}$.*

Proof. Let $r_0(n) = \frac{1}{3}\log_{\Delta-1}(n)$ so $r(n) \leq r_0(n)$. Let n_0 be sufficiently large that $\Delta^{r(n_0)+1} \leq (\eta - \zeta)n_0\Delta/2$, where ζ is the constant from Lemma 4.8. By Lemma 4.8,

$$\pi_G^{\mathrm{ord}}\Big(|\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|\Big) = \mathrm{e}^{-\Omega(n)}.$$

However, if $|\operatorname{In}(\mathcal{F})| > (1-\zeta)|E|$, then

$$|\operatorname{In}(\mathcal{F})\backslash E(B_r(v))| > (1-\zeta)|E| - \Delta^{r(n)+1} \geq (1-\zeta)|E| - \Delta^{r(n_0)+1} \geq (1-\eta)|E|. \quad \square$$

The above two lemmas, alongside with Lemma 4.23, show that the process 4.32 terminates unsuccessfully with only very small probability. This allows us to conclude WSM within the ordered phase.

4.3.3 Proof of WSM within the ordered phase

We can now prove Theorem 4.9.

Theorem 4.9. *Let $\Delta \geq 5$ and $K \geq 0$ be integers and let $\delta \in (0, 1/2)$ be a real. There exists $M = M(\Delta, K) > 0$ such that for all q sufficiently large and any $\beta \geq \beta_c$ the following holds for all sufficiently large n. Let $G \sim \mathcal{G}_{\Delta,\delta,K}$ be an n-vertex graph. Then G has WSM within the ordered phase at a radius r which satisfies $r \leq \frac{M}{\beta}\log n$.*

That is, for every $v \in V(G)$ and every edge e incident to v, it holds that

$$\left\|\pi_{B_r^+(v)}(e \mapsto \cdot) - \pi_G^{\mathrm{ord}}(e \mapsto \cdot)\right\|_{\mathrm{TV}} \leq \frac{1}{100m}.$$

Proof. Recall from Definition 4.4 that $\eta = \min\{\delta/5, 1/100\}$. Let $M = 2000 \cdot 400 \cdot (2+K)\Delta\log(\Delta-1)$.

Let $r(n, \beta) \coloneqq \frac{M}{\beta}\log_{\Delta-1}(n)$. Let $\beta_0(q) \coloneqq \log(q^{1.9/\Delta} + 1)$. Let q_0 and n_0 be large enough that

- $\beta_c(q_0) \geq \beta_0(q_0)$.
- Lemmas 4.23, 4.38 and 4.37 apply.
- $\beta_0(q_0) \geq 3M$.
- $2n_0^{-3/2} + f(n_0) \leq 1/(50\Delta n_0)$, where $f(n)$ is the $e^{-\Omega(n)}$ upper bound from Lemma 4.38.

For every β, define $n_1(\beta)$ to be the smallest positive integer such that

- $\frac{r(n_1(\beta),\beta)}{400\Delta(1+K)} - 1 \geq \frac{r(n_1(\beta),\beta)}{400\Delta(2+K)}$.

Now fix $q \geq q_0$ and $\beta \geq \beta_c(q)$.

Consider any $n \geq \max\{n_0, n_1(\beta)\}$. Consider an n-vertex graph $G \in \mathcal{G}_{\Delta,\delta,K}$. Let r_1 be as in Lemma 4.37. Use the process 4.32 to generate $\mathcal{F}^{\text{ord}}$ and $\mathcal{F}^+$. We wish to show that for every edge e incident to v,

$$\left\| \pi_{\mathcal{B}^+_{r_1}(v)}(e \mapsto \cdot) - \pi^{\text{ord}}(e \mapsto \cdot) \right\|_{\text{TV}} \leq 1/(100|E|) = 1/(50\Delta n).$$

By Observation 4.34, $\mathcal{F}^{\text{ord}} \sim \pi^{\text{ord}}$ and $\mathcal{F}^+ \sim \pi_{\mathcal{B}^+_{r_1}(v)}$. By Lemma 4.37, if $\mathcal{F}^{\text{ord}}$ and $\mathcal{F}^+$ do not agree on an edge incident to v, then $\mathcal{F}^{\text{ord}}$ contains a polymer of size at least $\frac{r(n,\beta)}{400\Delta(1+K)} - 1$, or $|\operatorname{In}(\mathcal{F}^{\text{ord}}) \backslash E(B_{r_1}(v))| < (1-\eta)|E|$. Thus, the probability that $\mathcal{F}^{\text{ord}}$ and $\mathcal{F}^+$ do not agree on an edge incident to v is at most the sum of the probabilities of those two events.

Applying Lemma 4.23, noting that $\frac{r(n,\beta)}{400\Delta(1+K)} - 1 \geq \frac{r(n,\beta)}{400\Delta(2+K)}$ and (given the lower bound on M at the start of the proof) that this quantity is at least $2000 \log n/\beta$, the probability that there is a large polymer (the first event) is at most $2n^{-3/2}$.

Note that $r(n,\beta) \leq \frac{1}{3}\log_{\Delta-1}(n)$. Applying Lemma 4.38, the probability that there are not enough in-edges outside of the ball (the second event) is $f(n) = e^{-\Omega(n)}$. The result follows by summing these probabilities. □

4.4 WSM within the disordered phase

The proof of WSM within the disordered phase (Theorem 4.10) follows a similar strategy to that for the ordered phase, but the details are substantially simpler (due to a more straightforward notion of polymers).

4.4.1 Disordered polymers

We will use the *disordered polymers* from [49, Section 2.3]. Given a graph $G = (V, E)$ and a configuration $\mathcal{F} \in \Omega^{\text{dis}}$, a *disordered polymer* of $\mathcal{F}$ is a connected component of $G[\text{In}(\mathcal{F})]$. Since $\mathcal{F} \in \Omega^{\text{dis}}$, each such connected component contains at most $\eta|E|$ edges.

Each disordered polymer $\gamma = (V(\gamma), E(\gamma))$ has an associated weight given by $w_\gamma^{\text{dis}} := q^{1-|V(\gamma)|}(e^\beta - 1)^{|E(\gamma)|}$. We observe the following relation between the weight of a disordered configuration $\mathcal{F} \in \Omega^{\text{dis}}$ and the weights of its polymers.

Observation 4.39. *For $\mathcal{F} \in \Omega^{\text{ord}}$, let $\Gamma^{\text{dis}}(\mathcal{F})$ be the set of disordered polymers of $\mathcal{F}$. Then*

$$w_G(\mathcal{F}) = q^n \prod_{\gamma \in \Gamma^{\text{dis}}(\mathcal{F})} w_\gamma^{\text{dis}}.$$

This observation follows from the fact that each vertex of G is contained in some polymer of $\mathcal{F}$, each in-edge of $\mathcal{F}$ is an edge of some polymer of $\mathcal{F}$, and each polymer of $\mathcal{F}$ corresponds to one component of $G[\text{In}(\mathcal{F})]$.

Based on [49, Proposition 11], we can upper-bound the probability that $\mathcal{F} \sim \pi^{\text{dis}}$ contains a large polymer.

Lemma 4.40. *Let $\Delta \geq 5$ and $K \geq 0$ be integers, $\delta \in (0, 1/2)$ be real. Let C be a constant. Then, for all sufficiently large q, and all $\beta \leq \log(q^{2.1/\Delta} + 1)$, the following holds for all sufficiently large n. For $G \in \mathcal{G}_{\Delta,\delta,K}$ with $n = |V(G)|$, for $s := C \log n$,*

$$\pi_G^{\text{dis}}\big(\mathcal{F} \text{ contains a polymer with at least } s \text{ edges}\big) \leq n^{-3/2}/2.$$

Proof. Let q and n be large enough so that Proposition 11 from [49] applies, and also that $s > 1$ and $q \geq e^{4\Delta(\frac{5}{2C}-1)}$.

For a disordered configuration $\mathcal{F}$, let $\Gamma^{\text{dis}}(\mathcal{F})$ be the set of disordered polymers of $\mathcal{F}$.

First we will bound, for a particular polymer γ with $|E(\gamma)| \geq 1$, the probability that it appears in a disordered configuration, $\mathbb{P}(\gamma \in \Gamma^{\text{dis}}(.))$. Suppose $\mathcal{F} \in \Omega^{\text{dis}}$ has $\gamma \in \Gamma^{\text{dis}}(\mathcal{F})$. By Observation 4.39, $w_G(\mathcal{F}) = q^n \prod_{\gamma \in \Gamma^{\text{dis}}(\mathcal{F})} w_\gamma^{\text{dis}}$. Let $\mathcal{F}'$ be the (disordered) configuration defined by $\text{In}(\mathcal{F}') = \text{In}(\mathcal{F}) \backslash E(\gamma)$. Since $E(\gamma) \neq \emptyset$, $\mathcal{F} \neq \mathcal{F}'$. By construction, $\mathcal{F}' \in \Omega^{\text{dis}}$. Since vertices in γ are not incident to edges in $\text{In}(\mathcal{F}')$, $\Gamma^{\text{dis}}(\mathcal{F}') = (\Gamma^{\text{dis}}(\mathcal{F}) \setminus \{\gamma\}) \cup_{v \in \gamma} \{v\}$. A polymer γ' consisting of a single vertex has weight $w_{\gamma'}^{\text{dis}} = 1$. Thus, $w_G(\mathcal{F}') = w_G(\mathcal{F})/w_\gamma^{\text{dis}}$. Let Ω_γ be the set of all disordered configurations containing γ as a polymer. Then $Z^{\text{dis}} \geq \sum_{\mathcal{F} \in \Omega_\gamma}(w_G(\mathcal{F}) +$

$w_G(\mathcal{F}')) = (1 + 1/w_\gamma^{\mathrm{dis}}) \sum_{\mathcal{F}\in\Omega_\gamma} w_G(\mathcal{F})$. Thus, the probability that $\mathcal{F} \sim \pi^{\mathrm{dis}}$ contains the polymer γ is at most $\frac{1}{1+1/w_\gamma^{\mathrm{dis}}} \leq w_\gamma^{\mathrm{dis}}$.

Next, for $v \in V(G)$ let p_v be the probability that $\mathcal{F} \sim \pi_G^{\mathrm{dis}}$ has a polymer with at least s edges containing the vertex v. By a union bound, $p_v \leq P_v := \sum_{\gamma\ni v:|E(\gamma)|\geq s} w_\gamma^{\mathrm{dis}}$.

We next use the following result from [49] to show that the sum of weights of polymers containing a given vertex decays exponentially. By [49, Equation (13)], the following holds with $\theta = \log(q)/(4\Delta)$.

$$\sum_{\gamma\ni v:|\gamma|>1} e^{(1+\theta)|E(\gamma)|} w_\gamma^{\mathrm{dis}} \leq \tfrac{1}{2}.$$

Thus also

$$e^{(1+\theta)s} P_v = e^{(1+\theta)s} \sum_{\gamma\ni v:|E(\gamma)|\geq s} w_\gamma^{\mathrm{dis}} \leq \sum_{\gamma\ni v:|E(\gamma)|\geq s} e^{(1+\theta)|E(\gamma)|} w_\gamma^{\mathrm{dis}} \leq \tfrac{1}{2}.$$

Hence $P_v \leq (1/2)e^{-(1+\theta)s}$. By a union bound over all vertices, the probability in the statement of the lemma is at most $\sum_v P_v \leq (n/2)\mathrm{e}^{-(1+\theta)s} = (1/2)n^{1-C(1+\theta)}$. The lemma follows by requiring q to be sufficiently large with respect to Δ and C such that $C(1+\theta) - 1 \geq 3/2$.

□

4.4.2 Proof of WSM within the disordered phase

We will prove that for q large enough, all large enough graphs in $\mathcal{G}_{\Delta,\delta,K}$ have WSM within the disordered phase for all $\beta \leq \beta_c$. First, we prove the following lemma showing that a suitable coupling exists.

Lemma 4.41. *Fix $\Delta \geq 5$, $K \geq 0$ integers, and $\delta \in (0, 1/2)$ a real. Then, for all q large enough and $\beta \leq \log(q^{2.1/\Delta} + 1)$, the following holds for all $G = (V, E) \in \mathcal{G}_{\Delta,\delta,K}$ with sufficiently many vertices. Let $v \in V$ and $r := \frac{1}{3}\log_{\Delta-1}|V|$. Then there is an integer r_1 satisfying $r \geq r_1 \geq \frac{r}{K+1} - 1$ and there is a coupling $(\mathcal{F}^-, \mathcal{F}^{\mathrm{dis}})$ such that $\mathcal{F}^- \sim \pi_{\mathcal{B}_{r_1}^-(v)}$ and $\mathcal{F}^{\mathrm{dis}} \sim \pi_G^{\mathrm{dis}}$, and moreover at least one of the following holds:*

1. *$|\mathrm{In}(\mathcal{F}^{\mathrm{dis}}) \backslash E(B_{r_1}(v))| > \eta|E| - |E(B_{r_1}(v))|$.*
2. *For any edge e incident to v, $\mathcal{F}^-(e) = \mathcal{F}^{\mathrm{dis}}(e)$.*
3. *$G[\mathrm{In}(\mathcal{F}^{\mathrm{dis}})]$ has a component with at least $\frac{r}{K+1} - 2$ edges.*

Proof. We use the definition of T_0 and $\mathcal{E}$ from Definition 4.29. By Observation 4.31, we get r_1 and r_2 satisfying $r \geq r_1 > r_2 \geq 0$ such that $r_1 - r_2 \geq \frac{r}{K+1} - 1$ and $E(B_{r_1}(v)) \setminus E(B_{r_2}(v))$ contains no edges from $\mathcal{E}$. Let T be the tree consisting of the first $r_1 + 1$ levels of T_0, i.e. $T_0[B_{r_1}(v)]$.

Construct the coupling as follows: first, let $F_0 \coloneqq E \setminus E(B_{r_1}(v))$ and $\mathcal{F}_0^{\text{dis}} \sim \pi_{F_0}^{\text{dis}}$. If $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| > \eta|E| - |E(B_{r_1}(v))|$, then let $\mathcal{F}^{\text{dis}} \sim \pi_{\mathcal{F}_0^{\text{dis}}}^{\text{dis}}$ and $\mathcal{F}^- \sim \pi_{\mathcal{B}_{r_1}^-(v)}$. Otherwise, let $\mathcal{F}_0^-$ be the partial configuration with $R(\mathcal{F}_0^-) = \operatorname{Out}(\mathcal{F}_0^-) = F_0$, let $F_1 = E \setminus E(B_{r_2+1}(v))$ and generate, using optimal monotone coupling (see Section 4.1.6), $\mathcal{F}_1^{\text{dis}} \sim \pi_{\mathcal{F}_0^{\text{dis}},F_1}$ and $\mathcal{F}_1^- \sim \pi_{\mathcal{F}_0^-,F_1}$. Finally, generate, optimally coupled $\mathcal{F}^- \sim \pi_{\mathcal{F}_1^-}$ and $\mathcal{F}^{\text{dis}} \sim \pi_{\mathcal{F}_1^{\text{dis}}}$.

First note that $\mathcal{F}^{\text{dis}} \sim \pi^{\text{dis}}$ and $\mathcal{F}^- \sim \pi_{\mathcal{B}_{r_1}^-(v)}$.

If $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| > \eta|E| - |E(B_{r_1}(v))|$, this follows by construction. Otherwise, we note that any refinement of $\mathcal{F}_0^{\text{dis}}$ has at most $\eta|E|$ edges, as $|E \setminus R(\mathcal{F}_0^{\text{dis}})| = |E(B_{r_1}(v))|$, thus $\pi_{\mathcal{F}_0^{\text{dis}},F_1} = \pi_{\mathcal{F}_0^{\text{dis}},F_1}^{\text{dis}}$. Similarly, $\pi_{\mathcal{F}_1^{\text{dis}}} = \pi_{\mathcal{F}_1^{\text{dis}}}^{\text{dis}}$. For $\mathcal{F}^-$ the result follows from the fact that a configuration $\mathcal{F}$ is a refinement of $\mathcal{F}_0^-$ if and only if $\pi_{\mathcal{B}_{r_1}^-(v)}(\mathcal{F}) > 0$.

Now we proceed to show that the resulting configurations satisfy one of the conditions. We have three exhaustive cases, and we show that each corresponds to one of the conditions.

Case 1. $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| > \eta|E| - |E(B_{r_1}(v))|$.

Then for any refinement $\mathcal{F}^{\text{dis}}$ of $\mathcal{F}_0^{\text{dis}}$, it holds that $|\operatorname{In}(\mathcal{F}^{\text{dis}}) \setminus E(B_{r_1}(v))| > \eta|E| - |E(B_{r_1}(v))|$, since $E(B_{r_1}(v)) \cap R(\mathcal{F}_0^{\text{dis}}) = \emptyset$, thus the first condition holds.

Case 2. $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| \leq \eta|E| - |E(B_{r_1}(v))|$, and $\xi(\mathcal{F}_1^{\text{dis}})$ is a free boundary. That is, no two vertices in $\partial(F_1)$ are in the same component of $G[\operatorname{In}(\mathcal{F}_1^{\text{dis}})]$.

Since $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| \leq \eta|E| - |E(B_{r_1}(v))|$, we have $R(\mathcal{F}_0^{\text{dis}}) = R(\mathcal{F}_0^-) = F_0$, and $R(\mathcal{F}_1^{\text{dis}}) = R(\mathcal{F}_1^-) = F_1$. Also, $\emptyset = \operatorname{In}(\mathcal{F}_0^-) \subseteq \operatorname{In}(\mathcal{F}_0^{\text{dis}})$, thus by Corollary 4.20, and by the fact that $\mathcal{F}_1^{\text{dis}}$ and $\mathcal{F}_1^-$ were optimally coupled using a monotone coupling, also $\operatorname{In}(\mathcal{F}_1^-) \subseteq \operatorname{In}(\mathcal{F}_1^{\text{dis}})$. So since no two vertices in $\partial_G(R(\mathcal{F}_1^{\text{dis}})) = \partial_G(F_1) = \partial_G(R(\mathcal{F}_1^-))$ are in the same component of $G[\operatorname{In}(\mathcal{F}_1^{\text{dis}})]$, it is also true that no two vertices in $\partial_G(F_1)$ are in the same component of $G[\operatorname{In}(\mathcal{F}_1^-)]$. So $\xi(\mathcal{F}_1^{\text{dis}}) = \xi(\mathcal{F}_1^-)$, thus by Observation 4.18 and by the fact that they were optimally coupled, $\mathcal{F}^-$ and $\mathcal{F}^{\text{dis}}$ agree on the edges of $E \setminus F_1 = E(B_{r_2+1}(v))$. Note that since $r_2 + 1 \geq 1$, all edges incident to v are in $E(B_{r_2+1}(v))$, thus $\mathcal{F}^{\text{dis}}$ and $\mathcal{F}^-$ agree on them. This corresponds to condition two.

Case 3. $|\operatorname{In}(\mathcal{F}_0^{\text{dis}})| \leq \eta|E| - |E(B_{r_1}(v))|$ and there are distinct vertices $u_1, u_2 \in \partial_G(F_1)$ such that u_1 and u_2 are in the same component of $G[\operatorname{In}(\mathcal{F}_1^{\text{dis}})]$.

Consider any path γ between u_1 and u_2 in the $G[\mathrm{In}(\mathcal{F}_1^{\mathrm{dis}})]$. Since no edges in $B_{r_2}(v)$ are revealed in $\mathcal{F}_1^{\mathrm{dis}}$, it must be the case that γ is contained in $E \setminus E(B_{r_2}(v))$. Note that since T_{u_1} and T_{u_2} – the subtrees of T rooted in u_1 and u_2 respectively, are disjoint, $V(\gamma)$ is not contained in $V(T_{u_1})$. Let $e = \{w_1, w_2\}$ be the first edge on γ such that $w_1 \in V(T_{u_1})$ and $w_2 \notin V(T_{u_1})$. Note that for any vertex w with $r_1 > d_G(w, v) > r_2$, all edges incident to w are edges of T, and in particular, if w is a non-leaf vertex in $V(T_{u_1})$, all incident edges to w are in $E(T_{u_1})$.

Then, by construction and the choice of r_1 and r_2, the boundary vertices of $\mathcal{F}_1^{\mathrm{dis}}$ are exactly $\partial_G(F_1) = \{u \in V \mid d_G(u, v) = r_2 + 1\}$. Also, no edges incident to vertices with distance at most r_2 from v were revealed, so it must be the case that $d_G(w_1, v) = r_1$.

Hence, there is a path containing at least $r_1 - (r_2 + 1) \geq \frac{r}{K+1} - 2$ edges in $G[\mathrm{In}(\mathcal{F}_1^{\mathrm{dis}})]$, and thus also in $G[\mathrm{In}(\mathcal{F}^{\mathrm{dis}})]$. So there is a connected component of $G[\mathrm{In}(\mathcal{F}^{\mathrm{dis}})]$ with size at least $\frac{r}{K+1} - 2$, which corresponds to the third condition. □

Next we prove WSM within the disordered phase (Theorem 4.10).

Theorem 4.10. *Let $\Delta \geq 5$ and $K \geq 0$ be integers and let $\delta \in (0, 1/2)$ be a real. Then, for all q sufficiently large and any $\beta \leq \beta_c$, and for all sufficiently large n, the following holds. Let $G \sim \mathcal{G}_{\Delta,\delta,K}$ be an n-vertex graph. Then G has WSM within the disordered phase at a radius r which satisfies $r \leq \frac{1}{3}\log_{\Delta-1} n$. That is, for every $v \in V(G)$ and every edge e incident to v, it holds that*

$$\left\| \pi_{B_r^-(v)}(e \mapsto \cdot) - \pi_G^{\mathrm{dis}}(e \mapsto \cdot) \right\|_{\mathrm{TV}} \leq \frac{1}{100m}.$$

Proof. Let $\beta_1(q) \coloneqq \log(q^{2.1/\Delta} + 1)$. Let ζ be the constant from Lemma 4.8 and let C' be the constant implicit in Equation (4.2) so that this equation guarantees $\pi_G^{\mathrm{dis}}\Big(|\mathrm{In}(\mathcal{F})| \geq \zeta|E|\Big) \leq e^{-C'n}$.

Let q_0 and n_0 be large enough such that all of the following hold.

- $\beta_c(q_0) \leq \beta_1(q_0)$.
- Lemmas 4.8 and 4.41 apply. Lemma 4.40 applies with $C = \frac{1}{3(2+K)\log(\Delta-1)}$.
- $(\eta - \zeta)n_0 \geq \Delta n_0^{1/3}$ and $C'n_0 \geq \log(100\Delta n_0)$ and $n_0 \geq (50\Delta)^2$.
- $\frac{\log_{\Delta-1} n_0}{3(K+1)} - 2 \geq \frac{\log_{\Delta-1} n_0}{3(K+2)}$,

Now fix $q \geq q_0$ and $\beta \leq \beta_c(q)$.

Consider any $n \geq n_0$. Consider an n-vertex graph $G = (V, E) \in \mathcal{G}_{\Delta,\delta,K}$. Let $r \coloneqq \frac{1}{3}\log_{\Delta-1}(n)$. Fix $v \in V$. Let r_1 be as in Lemma 4.41. Use the coupling from

Lemma 4.41 to generate $\mathcal{F}^{\text{dis}}$ and $\mathcal{F}^{-}$. We wish to show that for every edge e incident to v,

$$\left\|\pi_{\mathcal{B}^{-}_{r_1}(v)}(e \mapsto \cdot) - \pi^{\text{dis}}(e \mapsto \cdot)\right\|_{\text{TV}} \leq 1/(100|E|)$$

By Lemma 4.41, if $\mathcal{F}^{\text{dis}}$ and $\mathcal{F}^{-}$ do not agree on all edges incident to v, then at least one of the following two events occurs.

1. $|\operatorname{In}(\mathcal{F}^{\text{dis}})| \geq |\operatorname{In}(\mathcal{F}^{\text{dis}}) \setminus E(B_{r_1}(v))| > \eta|E| - |E(B_{r_1}(v))|$, or
2. $\mathcal{F}^{\text{dis}}$ contains a polymer of size at least $\frac{r}{K+1} - 2 \geq \frac{\log n}{3(K+2)\log(\Delta-1)} = C \log n$.

Thus the probability that $\mathcal{F}^{\text{dis}}$ and $\mathcal{F}^{-}$ do not agree on all edges incident to v (and thus in particular do not agree on e) is at most the sum of the probabilities of these two events.

Since $(\eta - \zeta)n \geq \Delta n^{1/3} \geq |E(B_{r_1}(v))|$, it holds that $\zeta|E| \leq \eta|E| - |E(B_{r_1}(v))|$, so Lemma 4.8 guarantees that the probability of the first event is at most $e^{-C'n} \leq \frac{1}{100\Delta n} = \frac{1}{200|E|}$.

By Lemma 4.40, the probability of the second event is at most $n^{-3/2}/2 \leq \frac{1}{100\Delta n} = \frac{1}{200|E|}$. The result follows by adding up these probabilities. □

4.5 Local mixing for the wired RC dynamics

In this section, we establish the local mixing results for the wired RC dynamics needed to apply Theorem 1.5. Recall that in the context of this chapter, the wired RC dynamics is the RC dynamics on some ball $B_r(v)$, where all vertices distance-r from the vertex v are fixed to be in a single component, and all updates are done inside the ball.

We first establish optimal local mixing for wired RC dynamics on treelike balls when q is an integer in Section 4.5.1 using log-Sobolev constants. We prove a polynomial local mixing result for non-integer q using canonical paths in Section 4.5.2.

4.5.1 Log-Sobolev constant for wired treelike neighbourhoods

Here we discuss how to bound the log-Sobolev constant for the wired RC dynamics for treelike neighbourhoods for integer $q > 1$ using the results from [15] on the tree.

Lemma 4.12. *Let $q \geq 2$ and $\Delta \geq 3$ be integers, and $K, \beta > 0$ be reals. There exists $\tilde{C} > 0$ such that the following holds for every Δ-regular graph $G = (V, E)$ and any integer $r \geq 1$.*

Suppose that $\rho \in V$ is such that $G[B_r(\rho)]$ is K-treelike. Then, with $n_\rho = |B_r(\rho)|$, the log-Sobolev constant for the wired RC dynamics on $B_r(\rho)$ is $\geq \tilde{C}/n_\rho$.

Proof. Let $T = T_{\Delta,r}(\rho)$ denote the Δ-regular tree rooted at ρ with height r. Then, for integer $q \geq 2$ and $\beta > 0$, it is shown in [15, Proof of Lemma 32] that there exists $C > 0$ such that the log-Sobolev constant for the wired RC dynamics on $T_{\Delta,r}(\rho)$ is $\geq C/N$ where $N = |T_{\Delta,r}(\rho)|$.

To translate this into a lower bound on the log-Sobolev constant of $G[B_r(\rho)]$ (which is K-treelike) as in the statement of the lemma, we use the same argument as in [43, Proof of Lemma 4.4], which we now briefly outline.

Since $G[B_r(\rho)]$ is K-treelike, there are at most K excess edges (see Definition 4.29). We introduce some additional wired disjoint trees (i.e. their leaves are fixed to the same single component as the vertices distance-r from ρ) and denote their disjoint union H. The goal is that the disjoint union $G[B_r(\rho)] \uplus H$ can be transformed to $T_{\Delta,r}(\rho)$ by adding and removing at most $3K$ edges. For brevity, call the single component containing all distance-r vertices from ρ, as well as all leaves of H, the *wire*.

Before describing this transformation, we show that it gives the required bound on the log-Sobolev constant.

First of all, since the added trees are only connected to $G[B_r(\rho)]$ via the wire, the RC distribution on their disjoint union is a product distribution, and the corresponding RC dynamics decomposes into the dynamics on the ball and on the collection of wired trees H (since the component structure within $B_r(\rho)$ is independent of the configuration inside H). By the standard fact that entropy tensorises over product spaces (see e.g. [5]), the log-Sobolev constant of the dynamics on $G[B_r(\rho)]$ is at least that of the dynamics on the disjoint union.

Then, it is a standard fact (see e.g. [11, Proof of Lemma 6.6]) that adding or removing a single edge changes the log-Sobolev constant by at most a multiplicative constant. Recall that the log-Sobolev constant on the tree $T_{\Delta,r}(\rho)$ is at least C/N. Therefore, there exists a constant $C' = C(q, \beta, K, \Delta)$ independent of r such that the log-Sobolev constant for the wired RC dynamics on $G[B_r(\rho)] \uplus H$, and hence on $G[B_r(\rho)]$, is at least $(C' \cdot C)/N$. Since there are at most K excess edges in $G[B_r(\rho)]$, if $n_\rho > K$, there is a vertex at distance at most K from ρ whose descendant subtree contains no excess edges, giving that $n_\rho \geq (\Delta - 1)^{r-K}$. In this case, $N = O_{\Delta,K}(n_\rho)$, giving that the log-Sobolev constant for the wired RC dynamics on the ball is of order $\Omega(1/n_\rho)$. If $n_\rho \leq K$, then the log-Sobolev constant is bounded below by a

positive constant, which we can absorb into $\widetilde{C}$ (as there are only a constant number of neighbourhoods of size up to K).

To finish, we describe the decomposition: for each excess edge in $G[B_r(\rho)]$, we add a $(\Delta-1)$-ary wired tree to each of its endpoints. The height of a tree added to a vertex distance-ℓ from ρ is exactly $r-\ell-1$, so when we attach this tree to the vertex by an edge, its leaves are exactly distance-r from ρ. To transform the resulting disjoint union to the regular tree, we need to remove the excess edges (there are at most K of them), and add the edges connecting the wired trees to their respective parents (there are at most $2K$ of them). □

4.5.2 Mixing on treelike graphs for the wired RC dynamics

In this section, we prove Lemma 4.13, which establishes local mixing for the wired RC dynamics when q is an integer.

In the proof, we use a quantity known as the *relaxation time*, denoted τ. We use the standard bound relating the relaxation and mixing times [53],

$$T_{\mathrm{mix}} \leq \tau \left(1 + \tfrac{1}{2}\log\left(\frac{1}{\pi_{\min}}\right)\right), \tag{4.3}$$

where $\pi_{\min} = \min_{\mathcal{F}\in\Omega_G} \pi_G(\mathcal{F})$. We obtain the relaxation time from the standard canonical-paths technique [53, Chapter 5], which we now briefly recall.

In the canonical-paths method, one specifies, for every pair of configurations $\mathcal{A}, \mathcal{A}'$, a *canonical path* $\mathrm{Path}(\mathcal{A}, \mathcal{A}')$, that is, a sequence of configurations $\mathcal{A} = \mathcal{A}_0 \to \mathcal{A}_1 \to \cdots \to \mathcal{A}_k = \mathcal{A}'$. Roughly, the aim is to choose the paths so that a) the RC dynamics is relatively likely to go from $\mathcal{A}_i$ to $\mathcal{A}_{i+1}$ in a single step, b) the transition $\mathcal{A}_i \to \mathcal{A}_{i+1}$ is not used in "too many paths". More precisely, we get the following bound on the relaxation time:

$$\tau \leq m \max_{(\mathcal{F},\mathcal{F}')} \sum_{\substack{\mathcal{A},\mathcal{A}':\\ (\mathcal{F}\to\mathcal{F}')\in\mathrm{Path}(\mathcal{A},\mathcal{A}')}} \frac{\hat{\pi}(\mathcal{A})\hat{\pi}(\mathcal{A}')}{\hat{\pi}(\mathcal{F})\mathbb{P}(\mathcal{F},\mathcal{F}')}, \tag{4.4}$$

where the maximum is over all possible adjacent pairs $\mathcal{F} \to \mathcal{F}'$ used in a canonical path.

Hence, in the proof of Lemma 4.13, which we restate for convenience, the aim is to construct suitable canonical paths $\mathrm{Path}(\mathcal{A}, \mathcal{A}')$ such that the relaxation time bound given in (4.4) is polynomial.

Lemma 4.13. Let $\Delta \geq 3$ be an integer, and $q, K > 1$, $\beta > 0$ be reals. There exists $\hat{C} > 0$ such that the following holds for every Δ-regular graph $G = (V, E)$ and any integer $r \geq 1$.

Suppose that $\rho \in V$ is such that $G[B_r(\rho)]$ is K-treelike. Then, with $n_\rho = |B_r(\rho)|$, the mixing time of the wired RC dynamics on $B_r(\rho)$ is $\leq \hat{C} n_\rho^3 (q^4 \mathrm{e}^\beta)^{\Delta r}$.

Proof. We use a canonical-paths argument, following largely [64, Proof of Lemma 8]. Let $G = (V, E)$ be a Δ-regular graph and $\rho \in V$ be such that $G[B_r(\rho)]$ is K-treelike. Let $n_\rho = n = |B_r(\rho)|$, $m = |E(B_r(\rho)|$. Let $\hat{\pi} = \pi^+_{B_r(\rho)}$ be the stationary distribution of the wired RC dynamics on $B_r(\rho)$.

Consider a BFS tree T for $G[B_r(\rho)]$ starting from ρ, let L be the set of the leaves of T and let $\mathcal{E} = E(G[B_r(\rho)]) \backslash E(T)$ be the set of excess edges (see Definition 4.29). Note that the height of T is at most r, the set $S_r(\rho)$ of vertices at distance r from ρ satisfies $S_r(\rho) \subseteq L$, and there are at most $2K$ leaves of T at depth less than r (by the Δ-regularity of G, each such leaf must be incident to an excess edge of G), so $|L \backslash S_r(\rho)| \leq 2K$. For a vertex u in $B_r(\rho)$, recall that T_u denotes the subtree of T rooted at u, and use $\mathrm{Anc}(u)$ to denote the ancestors of u in T (i.e., the vertices on the path from the root ρ to u, including u).

Consider next a depth-first-search (DFS) traversal of the tree T starting from ρ, and let $v_1, \ldots, v_n$ be the order in which the vertices of T are first visited. We write $v_i < v_j$ whenever $i < j$; moreover, for a subset of vertices S and a vertex u, we write $S < u$ to denote that for each $w \in S$ it holds that $w < u$. Since the tree T has depth r, at any stage of the DFS traversal, there are at most r vertices which are visited but not fully explored (note that a leaf that is visited is automatically explored). Since the degree of any vertex is $\leq \Delta$, it follows that

(a) for $i = 1, \ldots, n$, there are at most Δr tree edges with one endpoint in $\{v_1, \ldots, v_i\}$ and the other in $\{v_{i+1}, \ldots, v_n\}$ (from the DFS traversal).

(b) for an arbitrary vertex u in T and any vertex w which is not an ancestor of u (i.e., $w \notin \mathrm{Anc}(u)$), we have from the DFS traversal either that $V(T_w) < u$ or $V(T_w) > u$.

We order the edges $e_1, \ldots, e_m$ in $G[B_r(\rho)]$ (including excess edges) lexicographically in terms of the order of the vertices. So edges with endpoint v_1 come first (ordered by the other endpoint) and so on.

We now define the canonical paths between configurations. For two configurations $\mathcal{F}, \mathcal{F}' : E \to \{0, 1\}$ that differ on the edges $e_{i_1}, \ldots, e_{i_k}$ (with $i_1 < \cdots < i_k$), define the

path $\mathrm{Path}(\mathcal{F},\mathcal{F}')$ of configurations $\mathrm{Path}(\mathcal{F},\mathcal{F}') = \mathcal{F}_0 \to \mathcal{F}_1 \cdots \to \mathcal{F}_k$ by defining, for $j = 0,\ldots,k$, $\mathcal{F}_j$ to agree with $\mathcal{F}$ on the edges $\{e_{i_{j+1}},\ldots,e_{i_k}\}$ and to agree with $\mathcal{F}'$ on the edges $\{e_{i_1},\ldots,e_{i_j}\}$. Note that $\mathcal{F}_0 = \mathcal{F}, \mathcal{F}_k = \mathcal{F}'$ and both $\mathcal{F},\mathcal{F}'$ and hence $\mathcal{F}_j$ as well agree on $E\backslash\{e_{i_1},\ldots,e_{i_k}\}$.

We bound the relaxation time by (4.4). By our choice of canonical paths, the maximum is taken over all pairs of configurations $\mathcal{F},\mathcal{F}'$ that differ on a single edge, while the sum is over all pairs of configurations $\mathcal{A},\mathcal{A}'$ whose corresponding canonical path $\mathrm{Path}(\mathcal{A},\mathcal{A}')$ includes the transition $\mathcal{F}\to\mathcal{F}'$.

Hence, further on, we focus on bounding the quantity on the r.h.s. of (4.4). Consider an arbitrary pair of configurations $\kappa = (\mathcal{F},\mathcal{F}')$ that differ at a single edge $f = \{v_j, v_{j'}\}$ for some $j < j'$ and suppose further that $f = e_\ell$.

We first bound the sum in (4.4) for a particular κ by "encoding" every pair of configurations $\mathcal{A},\mathcal{A}'$ such that $\mathrm{Path}(\mathcal{A},\mathcal{A}')$ contains κ by a configuration $g_\kappa(\mathcal{A},\mathcal{A}')$. We will show that this encoding is injective, allowing us to bound the sum over such pairs by a sum over configurations.

In particular, for configurations $\mathcal{A},\mathcal{A}'$ such that $(\mathcal{F}\to\mathcal{F}') \in \mathrm{Path}(\mathcal{A},\mathcal{A}')$, define the configuration $g_\kappa(\mathcal{A},\mathcal{A}')$ so that

$$g_\kappa(\mathcal{A},\mathcal{A}') \text{ agrees with } \mathcal{A} \text{ on } \{e_1,\ldots,e_{\ell-1}\} \text{ and with } \mathcal{A}' \text{ on } \{e_\ell,\ldots,e_m\}. \tag{4.5}$$

From the fact that $(\mathcal{F}\to\mathcal{F}') \in \mathrm{Path}(\mathcal{A},\mathcal{A}')$, we also have that

$$\mathcal{A} \text{ agrees with } \mathcal{F} \text{ on } \{e_\ell,\ldots,e_m\}, \text{ and } \mathcal{A}' \text{ agrees with } \mathcal{F}' \text{ on } \{e_1,\ldots,e_{\ell-1}\}. \tag{4.6}$$

It follows that the map $g_\kappa(\cdot,\cdot)$ is injective, i.e., given its value $\mathcal{F}^*$ and the configurations $\mathcal{F},\mathcal{F}'$, there is a unique pair $(\mathcal{A},\mathcal{A}')$ such that $\mathcal{F}^* = g_\kappa(\mathcal{A},\mathcal{A}')$ and $(\mathcal{F}\to\mathcal{F}') \in \mathrm{Path}(\mathcal{A},\mathcal{A}')$. For a configuration on the ball $\mathcal{X} : E(B_r(\rho)) \to \{0,1\}$, let $\hat{c}(\mathcal{X})$ denote the number of components in $(B_r(\rho), \mathrm{In}(\mathbf{X}))$ that do not include any of the vertices of $S_r(\rho)$ (i.e. the number of components excluding the "wired" component containing the vertices distance-r from ρ).

Next, we bound a term in the sum (4.4). We write

$$\frac{\hat{\pi}(\mathcal{A})\hat{\pi}(\mathcal{A}')}{\hat{\pi}(\mathcal{F})\mathbb{P}(\mathcal{F},\mathcal{F}')} = \frac{\hat{\pi}\big(g_\kappa(\mathcal{A},\mathcal{A}')\big)}{\mathbb{P}(\mathcal{F},\mathcal{F}')} \cdot \frac{\hat{\pi}(\mathcal{A})\hat{\pi}(\mathcal{A}')}{\hat{\pi}(\mathcal{F})\hat{\pi}\big(g_\kappa(\mathcal{A},\mathcal{A}')\big)}.$$

First we focus on bounding the second factor in the product on the r.h.s..

Let $s \coloneqq \Delta r + K$ and $t \coloneqq 4\Delta r + 12K$. The main step in the proof is to show that:

$$|\,\mathrm{In}(\mathcal{A})| + |\,\mathrm{In}(\mathcal{A}')| - |\,\mathrm{In}(\mathcal{F})| - |\,\mathrm{In}(g_\kappa(\mathcal{A},\mathcal{A}'))| \le s, \tag{4.7}$$

$$|\hat{c}(\mathcal{A}) + \hat{c}(\mathcal{A}') - \hat{c}(\mathcal{F}) - \hat{c}(g_\kappa(\mathcal{A},\mathcal{A}'))| \le t. \tag{4.8}$$

Assuming these for the moment, we have that

$$\frac{\hat{\pi}(\mathcal{A})\hat{\pi}(\mathcal{A}')}{\hat{\pi}(\mathcal{F})\hat{\pi}\big(g_\kappa(\mathcal{A},\mathcal{A}')\big)} \leq q^t \mathrm{e}^{\beta s}.$$

Recall that $p = 1 - \mathrm{e}^{-\beta}$ and that for $q > 1$ it holds that $\hat{p} \in (p/q, p)$. Therefore, for the RC dynamics, we have that $\mathbb{P}(\mathcal{F}, \mathcal{F}') \geq \frac{\min\{p/q,(1-p)\}}{m} \geq (1-\mathrm{e}^{-\beta})/(qm\mathrm{e}^{\beta})$ using that $\min\{p/q, (1-p)\} \geq (1-\mathrm{e}^{\beta})\mathrm{e}^{-\beta}/q$. Using this and the injectivity of g, we can bound τ by

$$\tau \leq \frac{m^2 q^{t+1}\mathrm{e}^{\beta(s+1)}}{(1-\mathrm{e}^{-\beta})} \sum_{\substack{\mathcal{A},\mathcal{A}':E\to\{0,1\};\\ (\mathcal{F}\to\mathcal{F}')\in\mathrm{Path}(\mathcal{A},\mathcal{A}')}} \hat{\pi}\big(g_\kappa(\mathcal{A},\mathcal{A}')\big) \leq \frac{m^2 q^{t+1}\mathrm{e}^{\beta(s+1)}}{(1-\mathrm{e}^{-\beta})}.$$

The mixing time is at most $\tau(1 + \frac{1}{2}\log(\min_{\mathcal{X}} \hat{\pi}(\mathcal{X})^{-1}))$.

Since $\min_{\mathcal{X}} \hat{\pi}(\mathcal{X}) \geq 2^{-n} q^{-n} e^{-\beta m}$ and $m \leq \Delta n$, the mixing time is at most

$$n_\rho^3 \Delta^2 q^{t+1} \mathrm{e}^{\beta(s+1)\frac{1+\log(q)+\beta\Delta}{1-\mathrm{e}^{-\beta}}}$$

so the statement for the mixing time follows by taking $\hat{C} := 2\Delta^3(q\mathrm{e}^{\beta})^{12K+3}/(1-\mathrm{e}^{-\beta})$.

It remains to show the two inequalities. First, we prove the bound for the edges (4.7) For convenience let $\mathcal{F}^* = g_\kappa(\mathcal{A}, \mathcal{A}')$. To establish the bound, we first prove the following facts for an arbitrary edge $e = \{u, w\}$. Recall that $\{v_j, v_{j'}\}$ is the edge where $\mathcal{F}, \mathcal{F}'$ differ and $j < j'$ (i.e. v_j comes before $v_{j'}$ in the DFS order of the tree T).

1. if $\mathcal{A}(e) = \mathcal{A}'(e)$, then $\mathcal{A}(e) = \mathcal{A}'(e) = \mathcal{F}(e) = \mathcal{F}^*(e)$.
2. if $\mathcal{A}(e) \neq \mathcal{A}'(e)$, then $\big(\mathcal{A}(e), \mathcal{A}'(e)\big) = \big(\mathcal{F}^*(e), \mathcal{F}(e)\big)$ when $u, w \leq v_j$ and $\big(\mathcal{A}(e), \mathcal{A}'(e)\big) = \big(\mathcal{F}(e), \mathcal{F}^*(e)\big)$ when $u, w > v_j$.

Note that Item 2 is non-exhaustive. In particular, it does not cover edges $e = \{u, w\}$ with $\mathcal{A}(e) \neq \mathcal{A}'(e)$ whose endpoints u and w lie on opposite sides of v_j in the DFS ordering. We next show that there are at most s such edges.

Indeed, for Item 1, consider e with $\mathcal{A}(e) = \mathcal{A}'(e)$ and suppose first that $\mathcal{A}(e) = \mathcal{A}'(e) = 1$. Then $\mathcal{F}(e) = \mathcal{F}'(e) = 1$ (since by the construction of $\mathrm{Path}(\mathcal{A}, \mathcal{A}')$, all configurations in it agree on the edges where $\mathcal{A}, \mathcal{A}'$ agree). Moreover, $\mathcal{F}^*(e) = 1$ by construction (since $\mathcal{F}^* = g_\kappa(\mathcal{A}, \mathcal{A}')$ agrees with $\mathcal{A}, \mathcal{A}'$ on the edges where the latter agree). The proof for the second case $\mathcal{A}(e) = \mathcal{A}'(e) = 0$ is analogous. Item 2 follows from (4.5) and (4.6) after observing that $u, w \leq v_j$ implies that $e = \{u, w\} < e_\ell$, while $u, w > v_j$ gives that $e > e_\ell$.

We can now bound (4.7). Let W_j be the tree edges which are incident to a vertex in $\mathrm{Anc}(v_j)$, so that $|W_j| \le \Delta r$. From Item (a), these are the only tree edges that can have one endpoint in $\{v_1, \dots, v_j\}$ and the other among $\{v_{j+1}, \dots, v_n\}$, along with any non-tree edges (i.e., excess edges $\mathcal{E}$). Therefore, from Items 1 and 2 above, it follows that for an edge $e \in E \backslash (W_j \cup \mathcal{E})$ it holds that

$$\mathcal{A}(e) + \mathcal{A}'(e) = \mathcal{F}(e) + \mathcal{F}^*(e), \tag{4.9}$$

which establishes (4.7) since $|W_j \cup \mathcal{E}| \le \Delta r + K$.

To show (4.8), fix an arbitrary root-to-leaf path P passing through v_j, and denote by E_P the edges of the path. Now, consider the slightly "tweaked" configurations $\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*$ obtained from $\mathcal{A}, \mathcal{A}', \mathcal{F}, \mathcal{F}^*$ respectively by setting, for each $\mathcal{X} \in \{\mathcal{A}, \mathcal{A}', \mathcal{F}, \mathcal{F}^*\}$, $\hat{\mathcal{X}}(e) = 1$ for $e \in E_P \cup W_j$, $\hat{\mathcal{X}}(e) = 0$ for $e \in \mathcal{E}$ and $\hat{\mathcal{X}}(e) = \mathcal{X}(e)$ for $e \notin E_P \cup W_j \cup \mathcal{E}$. Note that, for each $\mathcal{X} \in \{\mathcal{A}, \mathcal{A}', \mathcal{F}, \mathcal{F}^*\}$, we have $\big|\,|\ln(\hat{\mathcal{X}})| - |\ln(\mathcal{X})|\big| \le \Delta r + K$, so $\big||\hat{c}(\hat{\mathcal{X}})| - |\hat{c}(\mathcal{X})|\big| \le \Delta r + K$ as well (since flipping an edge can change the component count by at most one). By setting all excess edges to be out, we have also reduced the component-counting problem to the tree T, which makes the components easier to compare.

So, to prove (4.8), it suffices to show

$$|\hat{c}(\hat{\mathcal{A}}) + \hat{c}(\hat{\mathcal{A}}') - \hat{c}(\hat{\mathcal{F}}) - \hat{c}(\hat{\mathcal{F}}^*)| \le 8K. \tag{4.10}$$

To show this, first note that $\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*$ still satisfy Items 1 and 2 above (replacing $\mathcal{A}$ with $\hat{\mathcal{A}}$, and so on) since the only changes are for edges in $E_P \cup W_j \cup \mathcal{E}$ on which the assignments of the tweaked configurations are identical (and hence fall under Item 1). In fact, Items 1 and 2 now apply to all edges for the tweaked configurations (previously, edges in $W_j \cup \mathcal{E}$ were potentially not covered in Item 2, but now are covered since they fall under Item 1). So, the analogue of (4.9) for $\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*$ holds for every edge of $G[B_r(\rho)]$, yielding that

$$|\ln(\hat{\mathcal{A}})| + |\ln(\hat{\mathcal{A}}')| = |\ln(\hat{\mathcal{F}})| + |\ln(\hat{\mathcal{F}}^*)|, \tag{4.11}$$

and, further, for any edge $e = \{u, w\}$ we have that

$$\big(\hat{\mathcal{A}}(e), \hat{\mathcal{A}}'(e)\big) = \begin{cases} \big(\hat{\mathcal{F}}^*(e), \hat{\mathcal{F}}(e)\big) & \text{if } u, w \le v_j \\ \big(\hat{\mathcal{F}}(e), \hat{\mathcal{F}}^*(e)\big), & \text{otherwise.} \end{cases} \tag{4.12}$$

At this stage, it will be convenient to consider the graph T^* obtained from $G[B_r(\rho)]$ by identifying all the leaves of T into a single vertex v^* (recall that the leaves of T consist of the vertices $S_r(\rho)$, which are at depth r, together with at most $2K$ leaves that

are at depth $< r$). For $\mathcal{X} \in \{\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*\}$, let $c^*(\mathcal{X})$ be the number of components in the graph $T^*[\mathrm{In}(\mathcal{X})]$ that do not include v^* and let $\mathcal{C}^*(\mathcal{X}) = (V^*(\mathcal{X}), E^*(\mathcal{X}))$ be the connected component of v^*. Note that any connected component of $(B_r(\rho), \mathrm{In}(\mathcal{X}))$ that does not include a vertex from $S_r(\rho)$ is not connected to v^* in $T^*[\mathrm{In}(\mathcal{X})]$, unless it contains one of the $2K$ leaves at depth $< r$. It follows that $|c^*(\mathcal{X}) - \hat{c}(\mathcal{X})| \leq 2K$, so to prove (4.10) it suffices to show

$$c^*(\hat{\mathcal{A}}) + c^*(\hat{\mathcal{A}}') = c^*(\hat{\mathcal{F}}) + c^*(\hat{\mathcal{F}}^*). \tag{4.13}$$

For $\mathcal{X} \in \{\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*\}$, note that a connected component F of $\mathcal{X}$ that does not include v^* is a subgraph of T (using that $\mathcal{X}(e) = 0$ for $e \in \mathcal{E}$) and hence a tree, so $1 = |V(F)| - |E(F)|$. Summing over all such components, we obtain that

$$c^*(\mathcal{X}) = |V(T^*)| - |V^*(\mathcal{X})| - |\,\mathrm{In}(\mathcal{X})| + |E^*(\mathcal{X})|.$$

Therefore, using (4.11), to finish the proof of (4.13) it suffices to show that

$$|E^*(\hat{\mathcal{A}})| - |V^*(\hat{\mathcal{A}})| + |E^*(\hat{\mathcal{A}}')| - |V^*(\hat{\mathcal{A}}')| = |E^*(\hat{\mathcal{F}})| - |V^*(\hat{\mathcal{F}})| + |E^*(\hat{\mathcal{F}}^*)| - |V^*(\hat{\mathcal{F}}^*)| \tag{4.14}$$

By rewriting the desired equality as a sum of indicators, we reduce the proof to showing that for an arbitrary vertex v and an arbitrary edge e, it holds that

$$\mathbf{1}\{v \in V^*(\hat{\mathcal{A}})\} + \mathbf{1}\{v \in V^*(\hat{\mathcal{A}}')\} = \mathbf{1}\{v \in V^*(\hat{\mathcal{F}})\} + \mathbf{1}\{v \in V^*(\hat{\mathcal{F}}^*)\}. \tag{4.15}$$

$$\mathbf{1}\{e \in E^*(\hat{\mathcal{A}})\} + \mathbf{1}\{e \in E^*(\hat{\mathcal{A}}')\} = \mathbf{1}\{e \in E^*(\hat{\mathcal{F}})\} + \mathbf{1}\{e \in E^*(\hat{\mathcal{F}}^*)\}. \tag{4.16}$$

Clearly (4.15) is true for any $v \in \mathrm{Anc}(v_j)$ since $v \in V^*(\mathcal{X})$ for each $\mathcal{X} \in \{\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*\}$ by construction of the tweaked configurations. So, consider $v \notin \mathrm{Anc}(v_j)$. First, suppose $v < v_j$. Let u be the ancestor of v whose parent w is an ancestor of v_j. The key point is that for each $\mathcal{X} \in \{\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*\}$ we have $w \in V^*(\mathcal{X})$ (since w is an ancestor of v_j) and the edge $e = \{u, w\}$ belongs to W_j , so $\hat{\mathcal{X}}(e) = 1$. Therefore whether $v \in V^*(\mathcal{X})$ is determined by the set of those paths contained in the subtree T_u which start from v and end either at u or a leaf of T_u. Denote by $\mathcal{P}_v$ the set of all such paths. Since $v < v_j$, we have that $u < v_j$ and that u is not ancestor of v_j, so from Item b it holds that $V(T_u) < v_j$. It follows from (4.12) that for every edge $e \in E(T_u)$ we have $(\hat{\mathcal{A}}(e), \hat{\mathcal{A}}'(e)) = (\hat{\mathcal{F}}^*(e), \hat{\mathcal{F}}^*(e))$, so

$$\begin{aligned} v \in V^*(\hat{\mathcal{A}}) &\Leftrightarrow \exists P' \in \mathcal{P}_v :\ E(P') \subseteq \mathrm{In}(\hat{\mathcal{A}}) \\ &\Leftrightarrow \exists P' \in \mathcal{P}_v :\ E(P') \subseteq \mathrm{In}(\hat{\mathcal{F}}^*) \Leftrightarrow v \in V^*(\hat{\mathcal{F}}^*), \end{aligned}$$

and similarly $v \in V^*(\hat{\mathcal{A}}') \Leftrightarrow v \in V^*(\hat{\mathcal{F}})$, proving (4.15) for $v < v_j$. An analogous argument applies when $v > v_j$, then $V(T_u) > v_j$ and, using (4.12) again, $v \in V^*(\hat{\mathcal{A}}) \Leftrightarrow v \in V^*(\hat{\mathcal{F}})$, $v \in V^*(\hat{\mathcal{A}}') \Leftrightarrow v \in V^*(\hat{\mathcal{F}}^*)$, finishing the proof of (4.15). The proof of (4.16) is very similar, after noting that for $\mathcal{X} \in \{\hat{\mathcal{A}}, \hat{\mathcal{A}}', \hat{\mathcal{F}}, \hat{\mathcal{F}}^*\}$ and an edge $e = \{u, w\}$ it holds that $e \in E^*(\mathcal{X})$ iff $\mathcal{X}(e) = 1$ and $u, w \in V^*(\mathcal{X})$.

This completes the proof of (4.13), and hence the proof of (4.10) and (4.8) as well, concluding therefore the lemma. □

4.6 Proof of the main theorem

We now apply Theorem 1.5 using the three key properties established in this chapter: local maxima, WSM within the phase, and local mixing, and thus obtain fast mixing within the phase.

Proof of Theorems 4.5-4.7. Consider $G = (V, E) \sim \mathcal{G}_{n,\Delta}$ with $n = |V|$ and $m = |E|$. Let q be sufficiently large so that Lemma 4.8 and Theorems 4.9 and 4.10 apply. By Lemmas 4.1 and 4.2 there are constants δ and K such that w.h.p. $G \in \mathcal{G}_{\Delta,\delta,K}$. Thus, for the remainder of the proof, we assume $G \in \mathcal{G}_{\Delta,\delta,K}$ is a sufficiently large graph.

First, consider $\beta \geq \beta_c$ and the ordered phase. By Lemma 4.8,

$$\pi_G^{\mathrm{ord}}\big(|\operatorname{In}(\mathcal{F})| \leq (1-\zeta)|E|\big) = \mathrm{e}^{-\Omega(n)} = \mathrm{e}^{-\Omega(m)}.$$

In particular, this is an upper bound for the probability of leaving the phase in one step of the RC dynamics. Therefore, $\pi_G(\,\cdot\,; q, \beta)$ has a local maximum in the ordered phase (Key Property 1).

By Theorem 4.9, for every $v \in V$ there is a radius r_v satisfying $r_v \leq \frac{M}{\beta} \log_d n$ at which WSM within the ordered phase holds (Key Property 2). Note that the radius r_v may depend on v, as justified by Remark 3.3.

For q sufficiently large, and hence $\beta \geq \beta_c(q, \Delta)$ sufficiently large, it holds, for each $v \in V$, that $r_v \leq \frac{1}{3} \log_{\Delta-1} n$. Thus, $B_{r_v}(v)$ is K-treelike for every $v \in V$, and Lemma 4.13 gives the required local mixing at the same radius r_v (Key Property 3). Hence, Theorem 1.5 gives the polynomial mixing-time bound in Theorem 4.5.

For q an integer, we obtain the mixing-time bound in Theorem 4.6 by instead using optimal local mixing at radius-r_v neighbourhoods, established by Lemma 4.12. The other two key properties are unchanged. Applying Theorem 1.5 therefore gives $O(m \log m) = O(n \log n)$ mixing from the ordered ground state, proving Theorem 4.6.

It remains to consider $\beta \leq \beta_c$ and the disordered phase. By Lemma 4.8,

$$\pi_G^{\mathrm{dis}}\big(|\operatorname{In}(\mathcal{F})| \geq \zeta|E|\big) = \mathrm{e}^{-\Omega(n)}.$$

Consequently, π_G^{dis} has a local maximum in the disordered phase (Key Property 1). By Theorem 4.10, for every $v \in V$ there is a radius r_v satisfying $r_v \leq \frac{1}{3}\log_{\Delta-1} n$ at which WSM within the disordered phase holds (Key Property 2). Finally, since G is locally K-treelike, Lemma 4.11 gives optimal local mixing at the same radii r_v (Key Property 3). Thus, Theorem 1.5 gives $O(m \log m) = O(n \log n)$ mixing from the disordered ground state, proving Theorem 4.7. □

We note that, in the above application of Theorem 1.5, the mixing time obtained for Theorem 4.5 is $T = \Theta(n^{2+\log W})$, where $W = \Delta^{2M/\beta_0}\mathrm{e}^{M\Delta(\Delta+1)}$, $\beta_0 > \log(q)/\Delta$, and $M = \Theta(\Delta \log \Delta)$ (see the start of Section 4.3.2). Hence $C(\Delta) = O(\Delta^3 \log \Delta)$, giving the running time $O(n^{C(\Delta)})$ in the statement.

The proof of Theorem 1.1 can now be concluded easily.

Proof of Theorem 1.1. For $\beta \neq \beta_c$, the result follows immediately by Theorems 4.5-4.7, since by Lemma 4.8 we have that $\big\|\pi_G - \pi_G^{\mathrm{ord}}\big\|_{\mathrm{TV}} = \mathrm{e}^{-\Omega(n)}$ when $\beta > \beta_c$ and $\big\|\pi_G - \pi_G^{\mathrm{dis}}\big\|_{\mathrm{TV}} = \mathrm{e}^{-\Omega(n)}$ when $\beta < \beta_c$.

Finally, for $\beta = \beta_c$, we use Corollary 3.7. Instead of the FPRAS from Section 3.2, we can also use the FPRAS for the partition function from [49] to estimate proportions of the phases. □

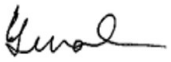
Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	*Sampling from the random cluster model on random regular graphs at all temperatures via Glauber dynamics.*
Publication Status	☒Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	Andreas Galanis, Leslie Ann Goldberg, Paulina Smolarova: Sampling from the random cluster model on random regular graphs at all temperatures via Glauber dynamics. Comb. Probab. Comput. 34(3): 359-391 (2025) An extended abstract also appeared at Andreas Galanis, Leslie Ann Goldberg, Paulina Smolarova: Sampling from the Random Cluster Model on Random Regular Graphs at All Temperatures via Glauber Dynamics. APPROX/RANDOM 2023: 64:1-64:12

Student Confirmation

Student Name:	Paulina Smolarova		
Contribution to the Paper	As is standard in our field, all authors contribute equally and authors are listed in alphabetic order		
Signature		Date	4.6. 2026

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof Leslie Ann Goldberg, Prof Andreas Galanis		
Supervisor comments: Paulina contributed substantially to all aspects		
Signature: Leslie Ann Goldberg Andreas Galanis	Date	4.6. 2026

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 5

Sampling and planting of the Potts model on random regular graphs

The chapter is based on the following publication:

[42] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model, 2026. URL `https://arxiv.org/abs/2410.14409`

An extended abstract of this work also appeared in STACS:

[41] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model. In *43rd International Symposium on Theoretical Aspects of Computer Science (STACS 2026)*, volume 364 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:19, 2026. doi: 10.4230/LIPIcs.STACS.2026.39

5.1 Overview

In this chapter, we use Theorem 1.5 to prove Theorem 1.2, which we restate below. This result establishes fast mixing throughout the high-temperature regime on random regular graphs when q is an integer, and substantially reduces the lower bound on q for the low-temperature regime compared to Theorem 1.1. As a direct consequence, we resolve sampling from the Potts model on random regular graphs for all $q \geq 3$ and $\beta < \beta_c = \log\left(\frac{q-2}{(q-1)^{1-2/\Delta}-1}\right)$.

Theorem 1.2. *Let $\Delta \geq 3$, $q \geq 3$ be integers and $\beta > 0$ be a real. Then, with high probability over a random Δ-regular graph, the following holds.*

- *If $\beta < \beta_c$, then the RC dynamics initialised from all-out mixes in $O(n \log n)$ time.*

- *If $\beta > \beta_c$ and $q \geq (5\Delta)^5$, the RC dynamics initialised from all-in mixes in $O(n \log n)$ time.*

The main technical contribution of this chapter is a delicate analysis of the Potts and RC distributions that goes beyond the threshold β_u, using planting.

Planting is a technique used in the analysis of random constraint satisfaction problems (CSPs) to capture how the space of solutions is correlated with the structure of the random instance (see [1, 24, 6] for some applications). In terms of sampling, however, there is no recipe for converting planting methods to efficient sampling algorithms, and typically some extra sampling step is needed on top (see, for example, [2, 28] for such approaches) even so, the resulting sample usually has some accuracy limitation, i.e., the TV-distance from the target distribution cannot be made arbitrarily small (typically $n^{-\Omega(1)}$).

Here, we instead use planting as a means to obtain WSM within the phase (Key Property 2 in Theorem 1.5), bypassing the $q = \Delta^{\Omega(\Delta)}$ requirement of the polymer method in the previous chapter.

The planting arguments that we use in this chapter follow from careful first and second-moment considerations in [34], which have been carried out for the ferromagnetic Potts model; see also [6] for similar-flavoured "silent" planting results.

5.1.1 Notation

In this chapter, we work with both the Potts and the RC models. Throughout the chapter, $G = (V, E)$ is a graph, $q \geq 3$ is an *integer* and $\beta > 0$ is a real.

We write $\mu = \mu_{G,q,\beta}$ for the *Potts (Gibbs) distribution*. Recall (see Definition 2.8) that this is a distribution on q-colourings of G, where the probability of a q-colouring $\sigma \in [q]^V$ is $\mu(\sigma) = \frac{\mathrm{e}^{\beta m(\sigma)}}{Z_G}$, where $m(\sigma)$ is the number of monochromatic edges under σ and $Z_G = \sum_{\sigma \in [q]^V} \mathrm{e}^{\beta m(\sigma)}$ is the partition function.

For the random-cluster model, rather than working directly with β, we use the parameter $p = 1 - \mathrm{e}^{-\beta}$. We write $\pi = \pi_{G,q,p}$ for the *RC (Gibbs) distribution*. As in Definition 2.9, we write $\Omega_G = \{F \mid F \subseteq E\}$ for the set of RC configurations, and recall that $\pi(F) = q^{c(F)} p^{|F|} (1-p)^{|E \setminus F|} / Z_{RC,G}$ where $c(F)$ is the number of connected components in the graph (V, F), and $Z_{RC,G} = \sum_{F \in \Omega_G} q^{c(F)} p^{|F|} (1-p)^{|E \setminus F|}$ is the partition function for the random-cluster model.

Note that since q is an integer, by Edwards-Sokal coupling (see Section 2.3), we can transform a Potts configuration to an RC configuration by including each monochromatic edge independently with probability p.

5.1.1.1 The configuration model

In this chapter, we work with d-regular graphs with $d \geq 3$. Unlike in the previous chapter, throughout this chapter we write d for the degree of the graph. We write $G \sim \mathcal{G}_{n,d}$ for a uniformly random d-regular graph on n vertices.

We also work in the *configuration model*. Formally, for integers $n, d \geq 1$ with dn even, let $[n] \times [d]$ be a set of "half-edges"; we denote by $G \sim \mathbb{G}_{n,d}$ a (multi)graph on vertex set $[n]$ whose edges are obtained by taking a uniformly random perfect matching of the half-edges; we add an edge between $u, v \in [n]$ in G whenever the half-edges (u, i) and (v, j) are matched together for some $i, j \in [d]$. For brevity, instead of $G \sim \mathbb{G}_{n,d}$ we sometimes use $\boldsymbol{G}$ to denote the random (multi)graph chosen. Any property that holds w.h.p. over $\boldsymbol{G}$ also holds w.h.p. over $G \sim \mathcal{G}_{n,d}$ (see, e.g., [51]), so henceforth we focus on the configuration model $\boldsymbol{G}$. We also use $\mathbb{G}_{n,d}$ for the set of all d-regular multigraphs on the vertex set $[n]$.

5.1.2 Planting

Next, we describe the planting framework for analysing typical configurations within the ordered and the disordered phases. We first define the phases of the Potts model, then introduce the corresponding planted distributions, and finally translate them to the RC model.

5.1.2.1 Phases of the Potts model

Recall, from Section 1.1.3, that the behaviour of a typical Potts configuration depends on the inverse temperature β. For $\beta < \beta_c$, a typical configuration is disordered, and each colour appears with roughly equal proportions. For $\beta > \beta_c$, a typical configuration is ordered and one colour dominates, whereas the remaining colours have roughly equal frequencies.

To make this precise, first define the set of colourings where the colour frequencies are close to a given *vertex colour statistics*. Let $\nu = (\nu_1, \dots, \nu_q)$ be a q-dimensional probability distribution. For an n-vertex graph G, let

$$S_\nu(\theta) = \Big\{\sigma \in [q]^{V(G)} : \sum_{i \in [q]} \Big| |\sigma^{-1}(i)| - n\nu_i \Big| \leq \theta n \Big\},$$

where $\theta = \theta(q, d, \beta) > 0$ is a small constant which we will suppress from notation. Let also

$$Z_\nu(G) = \sum_{\sigma \in S_\nu} \mathrm{e}^{\beta m(\sigma)} \quad \text{and} \quad \Psi = \Psi_{d,q,\beta}(\nu) = \lim_{\theta \downarrow 0} \lim_{n \to \infty} \tfrac{1}{n} \ln \mathbb{E}[Z_\nu(\boldsymbol{G})]$$

In short, the function $\Psi(\nu)$ captures the (logarithm of the) expected contribution to the partition function $Z(\boldsymbol{G})$ of the random graph $\boldsymbol{G}$ given by the configurations in S_ν. The local/global maximisers of Ψ play a key role in determining the phases of the Potts Gibbs distribution.

The colour vector corresponding to the disordered phase is

$$\nu_{\mathrm{d}} = \left(\tfrac{1}{q}, \ldots, \tfrac{1}{q}\right).$$

For the ordered phase, the colour statistics are more complicated, since one colour dominates over the others. The proportion of the majority colour is given by

$$a = \frac{t^d}{t^d + q - 1} \in \left(\frac{1}{q}, 1\right) \text{ where } \mathrm{e}^\beta - 1 = \frac{(t-1)(t^{d-1} + q - 1)}{t^{d-1} - t}. \tag{5.1}$$

Then, up to a permutation of the colours, the ordered colour vector is

$$\nu_{\mathrm{o}} = \left(a, \tfrac{1-a}{q-1}, \ldots, \tfrac{1-a}{q-1}\right) \tag{5.2}$$

For $\beta < \beta_c$, ν_{d} is the global maximiser for Ψ, whereas for $\beta > \beta_c$, ν_{o} (and each of its permutations) is a global maximiser of Ψ. At $\beta = \beta_c$, both ν_{d} and ν_{o} are global maximisers [25].

Interestingly, the vector corresponding to the non-dominant phase may still be a local maximiser of Ψ. Recall the metastability window (β_u, β_u') from Section 1.1.4.1. In this range of β, the RC dynamics experiences slow mixing arising from initialising the chain in the "wrong phase" [25]. This is closely connected to the *local maximisers* of Ψ: for $\beta \in (\beta_u, \beta_c)$, while the disordered vector ν_{d} is the global maximiser, Ψ has a local maximum at the ordered vector ν_{o}.[1] Conversely, for $\beta \in (\beta_c, \beta_u')$, the disordered vector ν_{d} is a local maximiser of Ψ, despite the ordered vector ν_{o} being the global maximiser [25].

[1] In fact, the value of β_u is defined as the infimum value of $\beta > 0$ such that there exists $a \in (1/q, 1)$ which satisfies (5.1); then it is easy to check that such an a exists for all $\beta \geq \beta_u$. For certain values of β (namely when $\beta \in (\beta_u, \beta_u')$), there may be more than one value of $a \in (1/q, 1)$ that satisfies the equation; when this is the case, the value of a that is relevant for the ordered phase is the largest such value.

We can now define the sets of ordered and disordered Potts configurations. Note that, in this chapter, we use a different notation for the phases compared to the rest of the thesis. In particular, we use d as a shorthand for the disordered phase, and o as shorthand for the ordered phase. Let S_{d} and S_{o} denote the sets of configurations S_ν when $\nu = \nu_{\mathrm{d}}$ and $\nu = \nu_{\mathrm{o}}$, respectively. For a graph G, define $Z_{\mathrm{d}}(G)$ and $Z_{\mathrm{o}}(G)$ analogously, and let $\mu_{G,\mathrm{d}} = \mu_G(\cdot \mid S_{\mathrm{d}})$ and $\mu_{G,\mathrm{o}} = \mu_G(\cdot \mid S_{\mathrm{o}})$ denote the conditional Potts Gibbs distributions. The following lemma from [34] gives some sharp concentration estimates for $Z_{\mathrm{d}}(\boldsymbol{G})$ and $Z_{\mathrm{o}}(\boldsymbol{G})$ (derived using moments and the small subgraph conditioning method) and is the key behind our planting approach.

Lemma 5.1 ([34, Theorem 5 & Lemma 16])**.** *Let $d, q \geq 3$ be integers, and $\beta \geq 0$ be real.*[2] *Let $f(n) > 0$ be any function with $f(n) = o(1)$. W.h.p. over $\boldsymbol{G}$, we have:*

$$Z_{\mathrm{d}}(\boldsymbol{G}) \geq f(n)\,\mathbb{E}[Z_{\mathrm{d}}(\boldsymbol{G})] \text{ if } \beta \leq \beta_c \quad \text{and} \quad Z_{\mathrm{o}}(\boldsymbol{G}) \geq f(n)\,\mathbb{E}[Z_{\mathrm{o}}(\boldsymbol{G})] \text{ if } \beta \geq \beta_c.$$

Moreover, $\mu_{\boldsymbol{G}}\big(S_{\mathrm{d}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta < \beta_c$, and $\mu_{\boldsymbol{G}}\big(\tilde{S}_{\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta > \beta_c$, where $\tilde{S}_{\mathrm{o}}$ is the union of S_{o} with its $q-1$ permutations. For $\beta = \beta_c$, it holds that $\mu_{\boldsymbol{G}}\big(S_{\mathrm{d}} \cup \tilde{S}_{\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ and $\mu_{\boldsymbol{G}}(S_{\mathrm{d}}), \mu_{\boldsymbol{G}}(S_{\mathrm{o}}) \geq f(n)$.

The first part of the lemma will be particularly useful for planting, as it allows us to relate the planted distributions to typical configurations from the corresponding phases.

5.1.2.2 Planting disordered and ordered configurations

We now introduce the planting distribution, which will be the key to obtaining Theorem 1.2. Intuitively, instead of a random configuration on a typical random graph, the *planted model* generates a random configuration and a random graph *at the same time*. We can then use the planted model to bound the probability of unlikely events for random configurations on typical random graphs.

To define this formally, we first fix a colouring of the vertices and define a distribution on graphs, where each graph appears with probability proportional to the Potts weight it assigns to this colouring. In particular, recall that $\boldsymbol{G}$ denotes a graph chosen according to the configuration model $\mathbb{G}_{n,d}$; in the expressions below, expectations are taken with respect to the choice of $\boldsymbol{G}$. For $\sigma \in [q]^n$, define a random graph $\hat{\boldsymbol{G}}(\sigma)$ by letting

$$\mathbb{P}\Big[\hat{\boldsymbol{G}}(\sigma) = G\Big] = \frac{\mathbb{P}\left[\boldsymbol{G} = G\right] \mathrm{e}^{\beta m_G(\sigma)}}{\mathbb{E}[\mathrm{e}^{\beta m_{\boldsymbol{G}}(\sigma)}]}. \tag{5.3}$$

[2]Here, we assume that the constant $\theta > 0$, used in the definitions of Z_{d} and Z_{o}, is a sufficiently small constant depending only on d, q, β.

We then consider each phase separately. For $\theta > 0$, recalling the disordered and ordered partition functions $Z_{\mathrm{d}}, Z_{\mathrm{o}}$ from the previous section, we define random configurations $\hat{\boldsymbol{\sigma}}_{\mathrm{d}} = \hat{\boldsymbol{\sigma}}_{\mathrm{d}}(\theta) \in [q]^n$ and $\hat{\boldsymbol{\sigma}}_{\mathrm{o}} = \hat{\boldsymbol{\sigma}}_{\mathrm{o}}(\theta) \in [q]^n$ with distributions

$$\mathbb{P}\big[\hat{\boldsymbol{\sigma}}_{\mathrm{d}} = \sigma\big] = \frac{\mathbf{1}\{\sigma \in S_{\mathrm{d}}\}\,\mathbb{E}[\mathrm{e}^{\beta m_{\boldsymbol{G}}(\sigma)}]}{\mathbb{E}[Z_{\mathrm{d}}(\boldsymbol{G})]}, \quad \mathbb{P}\big[\hat{\boldsymbol{\sigma}}_{\mathrm{o}} = \sigma\big] = \frac{\mathbf{1}\{\sigma \in S_{\mathrm{o}}\}\,\mathbb{E}[\mathrm{e}^{\beta m_{\boldsymbol{G}}(\sigma)}]}{\mathbb{E}[Z_{\mathrm{o}}(\boldsymbol{G})]}. \tag{5.4}$$

Thus, $\hat{\boldsymbol{\sigma}}_{\mathrm{d}}$ and $\hat{\boldsymbol{\sigma}}_{\mathrm{o}}$ choose a colouring from the corresponding phase proportionally to its expected Potts weight. We refer to $\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}), \hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big)$ and $\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}\big)$ as the *disordered* and *ordered planted distributions* respectively.

The following lemma, which we prove in Section 5.2, allows us to relate the planted distributions with the distributions of the (conditional) disordered $\boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}} \sim \mu_{\boldsymbol{G},\mathrm{d}}$ and ordered $\boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{o}} \sim \mu_{\boldsymbol{G},\mathrm{o}}$ configurations of the random graph $\boldsymbol{G}$.

Lemma 5.2. *Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Let $f(n), g(n) > 0$ be functions with $f(n) = o(1)$. Let $\mathcal{E} \subseteq \mathbb{G}_{n,d} \times [q]^n$ be an arbitrary event.*

(1) For $\beta \leq \beta_c$, if $\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}), \hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big) \in \mathcal{E}\big] \leq g(n)$, then, w.h.p. over $\boldsymbol{G}$, it holds that $\mathbb{P}\big[\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}}\big) \in \mathcal{E} \mid \boldsymbol{G}\big] \leq \frac{g(n)}{f(n)}$.

(2) For $\beta \geq \beta_c$, if $\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}\big) \in \mathcal{E}\big] \leq g(n)$, then, w.h.p. over $\boldsymbol{G}$, it holds that $\mathbb{P}\big[\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{o}}\big) \in \mathcal{E} \mid \boldsymbol{G}\big] \leq \frac{g(n)}{f(n)}$.

Thus, to show that an event is unlikely for a typical configuration from one of the phases, it suffices to show that it is sufficiently unlikely under the corresponding planted distribution.

To utilise Lemma 5.2, the key observation is that, once we condition on the *vertex and edge statistics*, the planted models become uniformly distributed. Formally, for a graph $G \in \mathbb{G}_{n,d}$ and a configuration $\sigma \in [q]^n$, the *vertex colour statistics* vector under σ, denoted ν^σ is defined as

$$\text{for a colour } i \in [q], \quad \nu_i^\sigma = |\sigma^{-1}(i)|/n.$$

Similarly, the *edge statistics* matrix under σ, denoted $\rho^{G,\sigma}$ is the $[q] \times [q]$ matrix with the following entries

$$\text{for colours } i, j \in [q], \quad \rho_{ij}^{G,\sigma} = \frac{1}{2|E(G)|} \sum_{u,v \in V(G)} \mathbf{1}\big\{\{u,v\} \in E(G), \sigma(u) = i, \sigma(v) = j\big\} \tag{5.5}$$

Recall that $2|E(G)| = dn$. So, for a colour class $i \in [q]$, out of the $dn\nu_i^\sigma$ half-edges incident to it, there are $dn\rho_{ij}^{G,\sigma}$ ending in a colour class $j \in [q]$ (note, for $i = j$, there

are $\frac{dn}{2}\rho_{ii}^{G,\sigma}$ edges in G that lie inside the colour class i). We have that $\rho^{G,\sigma}$ is symmetric with $\sum_{i,j}\rho_{ij}^{G,\sigma}=1$. Observe further that, conditioned on $\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}=\nu$ and $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}=\rho$ the pair $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}})$ is uniformly distributed over the pairs $(G,\sigma)\in\mathbb{G}_{n,d}\times[q]^n$ with $\nu^{\sigma}=\nu$ and $\rho^{G,\sigma}=\rho$. Similarly, once conditioned on $\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}$ and $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}),\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}$, the pair $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}),\hat{\boldsymbol{\sigma}}_{\mathrm{o}})$ becomes uniformly distributed over the pairs $(G,\sigma)\in\mathbb{G}_{n,d}\times[q]^n$ with these edge statistics.

To put this observation into use, we will need more quantitative information on the vertex and edge colour statistics in the planted models. By definition of $\hat{\boldsymbol{\sigma}}_{\mathrm{d}},\hat{\boldsymbol{\sigma}}_{\mathrm{o}}$, we already know that $\big\|\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}-\nu_{\mathrm{d}}\big\|_1\le\theta$ and $\big\|\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}-\nu_{\mathrm{o}}\big\|_1\le\theta$. Analogously, $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}$ and $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}),\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}$ are concentrated around the vectors ρ_{d} and ρ_{o} where

$$\text{for } i,j\in[q],\quad \rho_{\mathrm{d},ij}=\frac{\mathrm{e}^{\beta\mathbf{1}\{i=j\}}}{\sum_{i',j'\in[q]}\mathrm{e}^{\beta\mathbf{1}\{i'=j'\}}},\quad \rho_{\mathrm{o},ij}=\frac{\mathrm{e}^{\beta\mathbf{1}\{i=j\}}\big(\nu_{\mathrm{o},i}\nu_{\mathrm{o},j}\big)^{(d-1)/d}}{\sum_{i',j'\in[q]}\mathrm{e}^{\beta\mathbf{1}\{i'=j'\}}\big(\nu_{\mathrm{o},i'}\nu_{\mathrm{o},j'}\big)^{(d-1)/d}}. \tag{5.6}$$

We have the following concentration statement, which tells us that it is enough to consider configurations with vertex and edge colour statistics "close" to the average of the phase.

Lemma 5.3 ([25, Lemmas 3.4 & 3.5])**.** *Let $d,q\ge3$ be integers, and $\beta>0$ be real. There exists $c>0$ such that, for any $t=\omega(n^{-1/2})$, it holds that*

(i) For $\beta<\beta_u'$, $\mathbb{P}\Big[\big\|\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}-\nu_{\mathrm{d}}\big\|_1+\big\|\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}-\rho_{\mathrm{d}}\big\|_1\ge t\Big]\le\mathrm{e}^{-ct^2n}$.

(ii) For $\beta>\beta_u$, $\mathbb{P}\Big[\big\|\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}-\nu_{\mathrm{o}}\big\|_1+\big\|\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}),\hat{\boldsymbol{\sigma}}_{\mathrm{o}}}-\rho_{\mathrm{o}}\big\|_1\ge t\Big]\le\mathrm{e}^{-ct^2n}$.

Thus, the planted configurations have vertex and edge colour statistics concentrated around $(\nu_{\mathrm{d}},\rho_{\mathrm{d}})$ and $(\nu_{\mathrm{o}},\rho_{\mathrm{o}})$ in the disordered and ordered phases, respectively. Hence, to study the Potts model within a phase, we can consider vertex and edge colour statistics close to their typical values in that phase, and then analyse a uniformly random pair (G,σ) with these statistics. We next show how to translate these results for the random-cluster model.

5.1.2.3 Phases of the random-cluster model via the Potts model

Recall the Edwards-Sokal coupling (see Section 2.3), which allows us to obtain an RC sample from a Potts sample. For a graph $G\in\mathbb{G}_{n,d}$ and $\sigma\in[q]^n$, let $\boldsymbol{F}(G,\sigma)$ denote the random subset of edges obtained by keeping each monochromatic edge of G under

σ with probability p (we call this the *percolation step*)[3]. By Lemma 2.12, if $\boldsymbol{\sigma} \sim \mu_G$, then $\boldsymbol{F}(G, \boldsymbol{\sigma}) \sim \pi_G$. This in particular holds if G is the random graph $\boldsymbol{G}$.

From Lemma 5.3, we can then figure out, in particular, how many monochromatic edges we should expect from each phase, so we can translate the phases of the Potts model to the random-cluster model. More precisely, set

$$\mathfrak{m}_{\mathrm{d}}(\beta) = p\frac{d}{2}\sum_{i\in[q]} \rho_{\mathrm{d},ii} \quad \text{and} \quad \mathfrak{m}_{\mathrm{o}}(\beta) = p\frac{d}{2}\sum_{i\in[q]} \rho_{\mathrm{o},ii}.$$

Thus, $n\mathfrak{m}_{\mathrm{o}}(\beta)$ and $n\mathfrak{m}_{\mathrm{d}}(\beta)$ give the typical numbers of in-edges in the ordered and disordered phases, respectively.

With $\mathfrak{m}_{\mathrm{o}}(\beta)$ and $\mathfrak{m}_{\mathrm{d}}(\beta)$ in hand, we can provide the precise definition for the ordered and disordered phases of the random-cluster model used in this chapter. Since increasing β favours more monochromatic edges (as can be verified from the expression for ordered and disordered edge statistics in (5.6)), both $\mathfrak{m}_{\mathrm{o}}(\beta)$ and $\mathfrak{m}_{\mathrm{d}}(\beta)$ are increasing in β. Hence, $\mathfrak{m}_{\mathrm{o}}(\beta)$ achieves its minimum on $[\beta_c, \infty)$ at β_c, while $\mathfrak{m}_{\mathrm{d}}(\beta)$ achieves its maximum on $(0, \beta_c]$ at β_c. We therefore use their values at β_c to define fixed thresholds for the two phases.

Definition 5.4 (The ordered/disordered phases)**.** *Let $q, d \geq 3$ be integers. For a graph $G \in \mathbb{G}_{n,d}$ and $\rho \coloneqq \mathfrak{m}_{\mathrm{o}}(\beta_c) - \mathfrak{m}_{\mathrm{d}}(\beta_c) > 0$, define*

$$\Omega_{G,\mathrm{d}} \coloneqq \{F \subseteq E(G) : |F| \leq n\mathfrak{m}_{\mathrm{d}}(\beta_c) + n\rho/4\}, \text{ and}$$
$$\Omega_{G,\mathrm{o}} \coloneqq \{F \subseteq E(G) : |F| \geq n\mathfrak{m}_{\mathrm{o}}(\beta_c) - n\rho/4\}.$$

Let also $\pi_{G,\mathrm{d}} \coloneqq \pi_G(\cdot \mid \Omega_{G,\mathrm{d}})$ and $\pi_{G,\mathrm{o}} \coloneqq \pi_G(\cdot \mid \Omega_{G,\mathrm{o}})$.

Our choice of the phases is justified by the following corollary of Lemma 5.1, which we prove in Section 5.2.

Corollary 5.5. *Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Then, w.h.p. over $G \sim \mathbb{G}_{n,d}$, we have $\pi_G\big(\Omega_{G,\mathrm{d}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta < \beta_c$ and $\pi_G\big(\Omega_{G,\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta > \beta_c$.*

For $\beta = \beta_c$, it holds that $\pi_G\big(\Omega_{G,\mathrm{d}} \cup \Omega_{G,\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ and $\pi_G\big(\Omega_{G,\mathrm{d}}\big), \pi_G\big(\Omega_{G,\mathrm{o}}\big) \geq 1/\ln\ln n$.

[3]Throughout the chapter, a percolated graph is a random subgraph obtained by including each edge independently with some probability p. Thus, a percolation step can be thought of as obtaining a percolated graph from the subgraph of G containing only monochromatic edges.

To analyse the ordered/disordered phases of the random-cluster model using the planted distributions, we will need the following intermediate distributions. Namely, for a graph G, let $\hat{\pi}_{G,\mathrm{d}}$ be the distribution of the edge subset $\boldsymbol{F}(G, \boldsymbol{\sigma}_{G,\mathrm{d}})$, where $\boldsymbol{\sigma}_{G,\mathrm{d}} \sim \mu_{G,\mathrm{d}}$ and $\boldsymbol{F}(G,\sigma)$ is the random RC configuration defined at the beginning of this section. Define $\hat{\pi}_{G,\mathrm{o}}$ analogously, with $\boldsymbol{\sigma}_{G,\mathrm{o}} \sim \mu_{G,\mathrm{o}}$. The proof of the following lemma is given in Section 5.2.2.

Lemma 5.6. *Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Then, w.h.p. over $G \sim \mathbb{G}_{n,d}$, for any event $\mathcal{A} \subseteq 2^{E(G)}$, the following holds.*

(i) if $\beta \leq \beta_c$, then $\pi_{G,\mathrm{d}}(\mathcal{A}) \leq \hat{\pi}_{G,\mathrm{d}}(\mathcal{A}) + \mathrm{e}^{-\Omega(n)}$.

(ii) if $\beta \geq \beta_c$, then $\pi_{G,\mathrm{o}}(\mathcal{A}) \leq \hat{\pi}_{G,\mathrm{o}}(\mathcal{A}) + \mathrm{e}^{-\Omega(n)}$.

Thus, up to an exponentially small error, to upper-bound the probability of an event under the conditional RC distribution it suffices to bound its probability after applying the percolation step to the corresponding conditional Potts distribution. Together with Lemma 5.2, this allows us to obtain properties of the conditional RC distributions by analysing the corresponding planted models. We use this to show mixing within the phase, as stated in the next section.

5.1.3 Mixing within the phase

Having defined the phases of the random-cluster model, we can now give a proof overview for Theorem 1.2. As in the previous chapter, we consider each phase separately and obtain fast mixing by an application of Theorem 1.5. Formally, the proof reduces to establishing the following two theorems, which we prove in Section 5.5.

Theorem 5.7. *Let $d, q \geq 3$ be integers, and $\beta \in (0, \beta_c]$. Then, w.h.p. over $G \sim \mathcal{G}_{n,d}$, for every $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ the RC dynamics $(X_t)_{t\geq 0}$ initialised from all-out satisfies $\big\|X_T - \pi_{G,\mathrm{d}}\big\|_{\mathrm{TV}} \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.*

Theorem 5.8. *Let $d \geq 3, q \geq (5d)^5$ be integers, and $\beta \geq \beta_c$ be a real. Then, w.h.p. over $G \sim \mathcal{G}_{n,d}$, for every $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ the RC dynamics $(X_t)_{t\geq 0}$ initialised from all-in satisfies $\big\|X_T - \pi_{G,\mathrm{o}}\big\|_{\mathrm{TV}} \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.*

To apply Theorem 1.5, recall that we need to establish the three key properties of the RC model: a local maximum in the phase (Key Property 1), WSM within the phase (Key Property 2) and optimal local mixing (Key Property 3).

We obtain the first of the key properties (local maximum) similarly to the previous chapter. Using the phase-concentration results established through planting (Corollary 5.5), we show that a configuration sampled from the relevant phase (the disordered phase for $\beta \leq \beta_c$ and the ordered phase for $\beta \geq \beta_c$) is exponentially unlikely to be close to its boundary. Consequently, the probability of leaving the phase in one step of the RC dynamics is exponentially small, which establishes Key Property 1.

Local mixing follows from the results from the previous chapter (Lemmas 4.12 and 4.11). In this chapter, however, we use radius $\ell = (\frac{1}{2} - \delta) \log_{d-1} n$ for a small positive constant $\delta = \delta(q, \beta, d)$. Although this radius is larger than in the previous chapter, w.h.p. the required local treelikeness continues to hold for all such radii with $\delta > 0$ (see, e.g., [10, Fact 2.3]).

The main outstanding key property is weak spatial mixing within the phase. We conclude the overview by providing a brief proof roadmap.

5.1.4 Weak spatial mixing within a phase

Recall that WSM within the phase (Key Property 2) says that the marginal probability of including an edge under the RC distribution conditioned on the phase is close to its marginal when all edges outside a small neighbourhood are fixed to the corresponding ground state.

To state our theorems formally, we introduce the notation for the local RC distributions used throughout this chapter. Note that this notation differs slightly from the rest of the thesis, and we also define boundary conditions more generally, as will be useful later in the proof of WSM within the disordered phase. Recall that, for an integer ℓ and a vertex v, $B_\ell(v)$ is the graph induced by vertices at distance $\leq \ell$ from v. A boundary condition τ on $B_\ell(v)$ is a subset of the edges in $E \backslash E(B_\ell(v))$; then $\pi^\tau_{G,v,\ell}$ denotes the RC distribution induced on $B_\ell(v)$ conditioned on τ, i.e., for $S \subseteq E(B_\ell(v))$ it holds that $\pi^\tau_{G,v,\ell}(S) = \frac{\pi_G(S \cup \tau)}{\sum_{S' \subseteq E(B_\ell(v))} \pi_G(S' \cup \tau)}$. For WSM within the phase, we use the all-in boundary condition $\tau = E \setminus E(B_\ell(v))$ in the ordered phase and the all-out boundary condition $\tau = \emptyset$ in the disordered phase. In the ordered case, as in the previous chapter, we additionally fix all distance-ℓ vertices into a single component, as justified in Remark 3.2. We use $\pi^+_{G,v,\ell}$ and $\pi^-_{G,v,\ell}$ to denote the resulting wired and free RC distributions, respectively.

For brevity, we also use the following notation: for an edge $e \in E$, let 1_e denote the subsets of E that include e.

We can now state the two main technical results of this chapter, establishing WSM within the disordered and ordered phases, respectively.

Theorem 5.9. *Let $d, q \geq 3$ be integers, and $\beta \leq \beta_c$ be a positive real. Then, there exists $\delta > 0$ such that, w.h.p. over $G \sim \mathbb{G}_{n,d}$, $G = (V, E)$ has WSM within the disordered phase at a distance $\ell = \lfloor (\frac{1}{2} - \delta) \log_{d-1} n \rfloor$. That is, with $m = |E|$, for every $v \in V$ and every edge e incident to v, it holds that $\left| \pi^{-}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{d}}(1_e) \right| \leq \frac{1}{100m}$.*

Theorem 5.10. *Let $d \geq 3, q \geq (5d)^5$ be integers, and $\beta \geq \beta_c$ be a real. Then, there exists $\delta > 0$ such that, w.h.p. over $G \sim \mathbb{G}_{n,d}$, G has WSM within the ordered phase at a distance $\ell = \lfloor (\frac{1}{2} - \delta) \log_{d-1} n \rfloor$. That is, with $m = |E|$, for every $v \in V$ and every edge e incident to v, it holds that $\left| \pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e) \right| \leq \frac{1}{100m}$.*

The proof for the disordered regime $\beta \leq \beta_c$ is given in Section 5.3. Roughly, as in the previous chapter, the key step is to show that the disordered RC configurations tend not to contain large components of in-edges. Following the approach of Blanca and Gheissari [10] for $\beta < \beta_u$, we use a careful revealing process of the planted graph to bound the number of non-trivial boundary components after revealing the configuration outside the ball according to the disordered RC distribution outside of the ball. We then bound the probability that there are "two disjoint paths" from a non-trivial boundary component to the central vertex v (see Lemma 5.16 for the precise definition of this event). Together with the correlation-decay estimates from [10], this gives WSM within the disordered phase.

The proof for the ordered regime $\beta \geq \beta_c$ is given in Section 5.4. The key idea is to consider the giant component of in-edges *after removing the central vertex* v. First, in Section 5.4.2, we show that this component is sufficiently large. Then, in Section 5.4.2, we use the planted model to show that all sufficiently long paths from v intersect this giant component with sufficient probability. We then use this event to characterise the presence of a *wired boundary* for some revealing scheme (as opposed to a particular revealing scheme in the previous chapter) and conclude the proof of Theorem 5.10 in Section 5.4.4.

5.2 Planting for the Potts and RC models

In this section, we prove the results stated in Section 5.1.2 and conclude by showing that in their appropriate temperature regimes, the phases are local maxima (Key Property 1).

5.2.1 Proof of Lemma 5.2

First, we give the proof of Lemma 5.2, restated here.

Lemma 5.2. Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Let $f(n), g(n) > 0$ be functions with $f(n) = o(1)$. Let $\mathcal{E} \subseteq \mathbb{G}_{n,d} \times [q]^n$ be an arbitrary event.

(1) For $\beta \leq \beta_c$, if $\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}), \hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big) \in \mathcal{E}\big] \leq g(n)$, then, w.h.p. over $\boldsymbol{G}$, it holds that $\mathbb{P}\big[(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}}) \in \mathcal{E} \mid \boldsymbol{G}\big] \leq \frac{g(n)}{f(n)}$.

(2) For $\beta \geq \beta_c$, if $\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}\big) \in \mathcal{E}\big] \leq g(n)$, then, w.h.p. over $\boldsymbol{G}$, it holds that $\mathbb{P}\big[(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{o}}) \in \mathcal{E} \mid \boldsymbol{G}\big] \leq \frac{g(n)}{f(n)}$.

Proof of Lemma 5.2. We give the proof for the first item. The proof for the other is completely analogous. Fix arbitrary positive $\beta \leq \beta_c$ and let $\mathcal{Q}$ be the set consisting of $G \in \mathbb{G}_{n,d}$ that satisfy

$$Z_{\mathrm{d}}(G) \geq (f(n))^{1/2}\, \mathbb{E}[Z_{\mathrm{d}}(\boldsymbol{G})],$$

and note that by Lemma 5.1 it holds that $\mathbb{P}[\boldsymbol{G} \in \mathcal{Q}] = 1 - o(1)$. Intuitively, $\mathcal{Q}$ is a set of all graphs where the partition function restricted to the disordered phase is not too small compared to the average graph.

Define then the set of bad events for graph G, $S_{\mathrm{Bad}}(G) \coloneqq \big\{\sigma \in S_{\mathrm{d}} \,\big|\, (G,\sigma) \in \mathcal{E}\big\}$.

We next consider the set of graphs $\mathcal{Q}'$ where these bad events are unlikely. To be precise, let $\mathcal{Q}'$ be the set of $G \in \mathbb{G}_{n,d}$ such that the aggregate weight $Z_{\mathrm{Bad}}(G) = \sum_{\sigma \in S_{\mathrm{Bad}}(G)} \mathrm{e}^{\beta m_G(\sigma)}\ S_{\mathrm{Bad}}(G)$ is less than $\frac{g(n)}{f(n)} Z_{\mathrm{d}}(G)$.

Note that for any graph $G \in \mathcal{Q}'$, we have $\mu_{G,\mathrm{d}}\big((G,\sigma) \in \mathcal{E}\big) = \frac{Z_{\mathrm{Bad}}(G)}{Z_{\mathrm{d}}(G)} \leq \frac{g(n)}{f(n)}$, so to finish the proof, it suffices to show that $\mathbb{P}[\boldsymbol{G} \in \mathcal{Q}'] = 1 - o(1)$.

To do this, consider the aggregate weight of pairs in $\mathcal{E}$. Let $W \coloneqq \sum_{(G,\sigma)\in\mathcal{E}} \mathrm{e}^{\beta m_G(\sigma)} \mathbf{1}\{\sigma \in S_{\mathrm{d}}\}$. By restricting to graphs G in $\mathcal{Q}\backslash\mathcal{Q}'$, we have the lower bound

$$W \geq \sum_{G \in \mathcal{Q}\backslash\mathcal{Q}'} \sum_{\sigma \in S_{\mathrm{Bad}}(G)} \mathrm{e}^{\beta m_G(\sigma)} \geq \frac{g(n)}{f(n)} \sum_{G \in \mathcal{Q}\backslash\mathcal{Q}'} Z_{\mathrm{d}}(G).$$

For graphs $G \in \mathcal{Q}$, we have $Z_{\mathrm{d}}(G) \geq (f(n))^{1/2}\, \mathbb{E}[Z_{\mathrm{d}}(\boldsymbol{G})]$, and therefore we further have

$$W \geq \frac{g(n)}{\sqrt{f(n)}} \big|\mathcal{Q}\backslash\mathcal{Q}'\big|\ \mathbb{E}\big[Z_{\mathrm{d}}(\boldsymbol{G})\big] = \frac{g(n)}{\sqrt{f(n)}} \big|\mathcal{Q}\backslash\mathcal{Q}'\big| \frac{\sum_{G \in \mathbb{G}_{n,d}} \sum_{\sigma \in [q]^n} \mathrm{e}^{\beta m_G(\sigma)} \mathbf{1}\{\sigma \in S_{\mathrm{d}}\}}{\big|\mathbb{G}_{n,d}\big|} \tag{5.7}$$

From the assumption that $\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big)\in\mathcal{E}\big]\le g(n)$ and the definition of $\big(\hat{\boldsymbol{G}}\big(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big)$ we obtain that

$$g(n)\ge\mathbb{P}\big[\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}\big)\in\mathcal{E}\big]=\frac{\sum_{(G,\sigma)\in\mathcal{E}}\mathrm{e}^{\beta m_G(\sigma)}\mathbf{1}\{\sigma\in S_{\mathrm{d}}\}}{\sum_{G\in\mathbb{G}_{n,d}}\sum_{\sigma\in[q]^n}\mathrm{e}^{\beta m_G(\sigma)}\mathbf{1}\{\sigma\in S_{\mathrm{d}}\}}.$$

Combining this with (5.7) and the definition of W, we obtain $\mathbb{P}[\boldsymbol{G}\in\mathcal{Q}\backslash\mathcal{Q}']=\frac{|\mathcal{Q}\backslash\mathcal{Q}'|}{|\mathbb{G}_{n,d}|}\le(f(n))^{1/2}=o(1)$. Since $\mathbb{P}[\boldsymbol{G}\in\mathcal{Q}]=1-o(1)$ from Lemma 5.1, it follows that

$$\mathbb{P}[\mathcal{G}\in\mathcal{Q}']\ge\mathbb{P}[\boldsymbol{G}\in\mathcal{Q}\cap\mathcal{Q}']=\mathbb{P}[\boldsymbol{G}\in\mathcal{Q}]-\mathbb{P}[\boldsymbol{G}\in\mathcal{Q}\backslash\mathcal{Q}']\ge1-o(1).\qquad\square$$

5.2.2 Using planting for the random-cluster model

We next prove the lemmas given in Section 5.1.2.3, and then conclude by establishing that the phases of the RC model are local maxima in their respective regimes.

We first prove the following lemma, which tells us that if we apply a percolation step to a disordered Potts colouring, the resulting RC configuration is likely to be disordered, and the same holds for the ordered phase. Recall that $\tilde{S}_{\mathrm{o}}$ is the union of S_{o} and its $q-1$ permutations.

Lemma 5.11. *Let $d,q\ge3$ be integers, and $\beta>0$ be real. Then, w.h.p. over $G\sim\mathbb{G}_{n,d}$, with $\boldsymbol{\sigma}\sim\mu_G$ and $\boldsymbol{F}=\boldsymbol{F}(G,\boldsymbol{\sigma})$, the following holds.*

For $\beta\le\beta_c$,

$$\mathbb{P}[\boldsymbol{F}\in\Omega_{G,\mathrm{d}}\mid\boldsymbol{\sigma}\in S_{\mathrm{d}}]=1-\mathrm{e}^{-\Omega(n)}\quad\textit{and}\quad\mathbb{P}[\boldsymbol{F}\in\Omega_{G,\mathrm{o}}\mid\boldsymbol{\sigma}\in S_{\mathrm{d}}]=\mathrm{e}^{-\Omega(n)}.$$

For $\beta\ge\beta_c$,

$$\mathbb{P}[\boldsymbol{F}\in\Omega_{G,\mathrm{o}}\mid\boldsymbol{\sigma}\in\tilde{S}_{\mathrm{o}}]=1-\mathrm{e}^{-\Omega(n)}\quad\textit{and}\quad\mathbb{P}[\boldsymbol{F}\in\Omega_{G,\mathrm{d}}\mid\boldsymbol{\sigma}\in\tilde{S}_{\mathrm{o}}]=\mathrm{e}^{-\Omega(n)}.$$

Proof. We prove the first equation; the second is analogous. Fix $\beta\le\beta_c$. By Lemma 5.3, for the disordered planted model $\mathbb{P}\big[\big\|\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}}),\hat{\boldsymbol{\sigma}}_{\mathrm{d}}}-\rho_{\mathrm{d}}\big\|_1=o(1)\big]=1-\mathrm{e}^{-\Omega(n)}$. So Lemma 5.2 yields that $\mathbb{P}\Big[\big\|\rho^{\boldsymbol{G},\boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}}}-\rho_{\mathrm{d}}\big\|_1=o(1)\,\Big|\,\boldsymbol{G}\Big]=1-\mathrm{e}^{-\Omega(n)}$ w.h.p. over $\boldsymbol{G}$. Therefore, for such $\boldsymbol{G}$, conditioned on $\boldsymbol{\sigma}\in S_{\mathrm{d}}$, the number of monochromatic edges under $\boldsymbol{\sigma}$ is at most $n\big(\frac{d}{2}\sum_{i\in[q]}\rho_{\mathrm{d},ii}\big)+n\rho/4=n\frac{1}{p}\mathfrak{m}_{\mathrm{d}}(\beta)+n\rho/4$ w.p.[4] $1-\mathrm{e}^{-\Omega(n)}$. Therefore, after the percolation step, by standard Chernoff bounds (see Lemma 2.6), we have $|\boldsymbol{F}|\le n\mathfrak{m}_{\mathrm{d}}(\beta)+n\rho/4$ w.p. $1-\mathrm{e}^{-\Omega(n)}$. Since $\mathfrak{m}_{\mathrm{d}}(\beta)\le\mathfrak{m}_{\mathrm{d}}(\beta_c)$, by Definition 5.4 of $\Omega_{\boldsymbol{G},\mathrm{d}},\Omega_{\boldsymbol{G},\mathrm{o}}$ we conclude that $\mathbb{P}[\boldsymbol{F}\in\Omega_{\boldsymbol{G},\mathrm{d}}\mid\boldsymbol{\sigma}\in S_{\mathrm{d}}]=1-\mathrm{e}^{-\Omega(n)}$ and $\mathbb{P}[\boldsymbol{F}\in\Omega_{\boldsymbol{G},\mathrm{o}}\mid\boldsymbol{\sigma}\in S_{\mathrm{d}}]=\mathrm{e}^{-\Omega(n)}$. $\square$

[4]With probability.

With this lemma, we can now prove Corollary 5.5, giving us that the phases dominate in their respective regimes.

Corollary 5.5. *Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Then, w.h.p. over $G \sim \mathbb{G}_{n,d}$, we have $\pi_G\big(\Omega_{G,\mathrm{d}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta < \beta_c$ and $\pi_G\big(\Omega_{G,\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ when $\beta > \beta_c$.*

For $\beta = \beta_c$, it holds that $\pi_G\big(\Omega_{G,\mathrm{d}} \cup \Omega_{G,\mathrm{o}}\big) = 1 - \mathrm{e}^{-\Omega(n)}$ and $\pi_G\big(\Omega_{G,\mathrm{d}}\big), \pi_G\big(\Omega_{G,\mathrm{o}}\big) \geq 1/\ln\ln n$.

Proof. Suppose first that $\beta < \beta_c$. The proof for $\beta > \beta_c$ is analogous. By Lemma 5.1,

$$\mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}] = 1 - \mathrm{e}^{-\Omega(n)}$$

for $\boldsymbol{\sigma} \sim \mu_G$, w.h.p. over G. Moreover, by Lemma 5.11, conditioned on $\boldsymbol{\sigma} \in S_{\mathrm{d}}$, the percolation step gives $\boldsymbol{F}(G, \boldsymbol{\sigma}) \in \Omega_{G,\mathrm{d}}$ with probability $1 - \mathrm{e}^{-\Omega(n)}$. Since $\boldsymbol{F}(G, \boldsymbol{\sigma}) \sim \pi_G$ by the Edwards-Sokal coupling, it follows that $\pi_G(\Omega_{G,\mathrm{d}}) = 1 - \mathrm{e}^{-\Omega(n)}$, which concludes the proof for $\beta < \beta_c$.

For $\beta = \beta_c$, Lemma 5.1 gives

$$\mu_G(S_{\mathrm{d}} \cup \tilde{S}_{\mathrm{o}}) = 1 - \mathrm{e}^{-\Omega(n)} \quad \text{and} \quad \mu_G(S_{\mathrm{d}}), \mu_G(S_{\mathrm{o}}) \geq 1/\ln\ln n,$$

where we take $f(n) = 1/\ln\ln n$. Applying Lemma 5.11 to the two phases and using the Edwards-Sokal coupling gives the claimed bounds. □

Next, we prove Lemma 5.6, restated here for convenience. Recall that, for a graph G, $\hat{\pi}_{G,\mathrm{d}}$ is the distribution of $\boldsymbol{F}(G, \boldsymbol{\sigma}_{G,\mathrm{d}})$, where $\boldsymbol{\sigma}_{G,\mathrm{d}} \sim \mu_{G,\mathrm{d}}$ and $\hat{\pi}_{G,\mathrm{o}}$ is defined analogously.

Lemma 5.6. Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Then, w.h.p. over $G \sim \mathbb{G}_{n,d}$, for any event $\mathcal{A} \subseteq 2^{E(G)}$, the following holds.

(i) if $\beta \leq \beta_c$, then $\pi_{G,\mathrm{d}}(\mathcal{A}) \leq \hat{\pi}_{G,\mathrm{d}}(\mathcal{A}) + \mathrm{e}^{-\Omega(n)}$.

(ii) if $\beta \geq \beta_c$, then $\pi_{G,\mathrm{o}}(\mathcal{A}) \leq \hat{\pi}_{G,\mathrm{o}}(\mathcal{A}) + \mathrm{e}^{-\Omega(n)}$.

Proof. We prove the first claim; the proof of the second is analogous. We consider first the case $\beta \leq \beta_c$. We have the lower bound

$$\pi_G(\Omega_{G,\mathrm{d}}) \geq \mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}]\, \mathbb{P}[\boldsymbol{F} \in \Omega_{G,\mathrm{d}} \mid \boldsymbol{\sigma} \in S_{\mathrm{d}}]. \tag{5.8}$$

Similarly, by conditioning on whether $\boldsymbol{\sigma} \in S_{\mathrm{d}}$, $\boldsymbol{\sigma} \in \tilde{S}_{\mathrm{o}}$, or $\boldsymbol{\sigma} \notin S_{\mathrm{d}} \cup \tilde{S}_{\mathrm{o}}$, we have the inequality

$$\pi_G(\mathcal{A} \cap \Omega_{G,\mathrm{d}}) \leq \mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}]\, \mathbb{P}[\boldsymbol{F} \in \mathcal{A} \mid \boldsymbol{\sigma} \in S_{\mathrm{d}}] + \mathbb{P}[\boldsymbol{F} \in \Omega_{G,\mathrm{d}} \mid \boldsymbol{\sigma} \in \tilde{S}_{\mathrm{o}}] + \mathbb{P}[\boldsymbol{\sigma} \notin S_{\mathrm{d}} \cup \tilde{S}_{\mathrm{o}}]. \tag{5.9}$$

For $\beta \leq \beta_c$, by Lemma 5.1 we have $\mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}] \geq \frac{1}{n}$ and $\mathbb{P}[\boldsymbol{\sigma} \notin S_{\mathrm{d}} \cup \tilde{S}_{\mathrm{o}}] = \mathrm{e}^{-\Omega(n)}$; by Lemma 5.11 we also have

$$\mathbb{P}[\boldsymbol{F} \in \Omega_{G,\mathrm{d}} \mid \boldsymbol{\sigma} \in S_{\mathrm{d}}] = 1 - \mathrm{e}^{-\Omega(n)} \quad \text{and} \quad \mathbb{P}[\boldsymbol{F} \in \Omega_{G,\mathrm{d}} \mid \boldsymbol{\sigma} \in S_{\mathrm{o}}] = \mathrm{e}^{-\Omega(n)}. \tag{5.10}$$

Combining the bounds from (5.8), (5.9) and (5.10), we obtain that

$$\pi_{G,\mathrm{d}}(\mathcal{A}) = \frac{\pi_G(\mathcal{A} \cap \Omega_{G,\mathrm{d}})}{\pi_G(\Omega_{G,\mathrm{d}})} \leq \mathbb{P}[\boldsymbol{F} \in \mathcal{A} \mid \boldsymbol{\sigma} \in S_{\mathrm{d}}] + \mathrm{e}^{-\Omega(n)} = \hat{\pi}_{G,\mathrm{d}}(\mathcal{A}) + \mathrm{e}^{-\Omega(n)},$$

where we used $\mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}] \geq 1/n$, so dividing the error terms by $\mathbb{P}[\boldsymbol{\sigma} \in S_{\mathrm{d}}]$ preserves the $\mathrm{e}^{-\Omega(n)}$ bound, and (5.10) to obtain $\big(\mathbb{P}[\boldsymbol{F} \in \Omega_{G,\mathrm{d}} \mid \boldsymbol{\sigma} \in S_{\mathrm{d}}]\big)^{-1} = 1 + \mathrm{e}^{-\Omega(n)}$. □

We can now establish the following corollary, which implies a local maximum (Key Property 1) in the relevant phase.

Corollary 5.12. *Let $d, q \geq 3$ be integers, and $\beta > 0$ be real. Then, for all $\delta > 0$, the following hold w.h.p. over $G \sim \mathbb{G}_{n,d}$.*

If $\beta \leq \beta_c$, for $\boldsymbol{F}_{\mathrm{d}} \sim \pi_{G,\mathrm{d}}$, it holds that $\mathbb{P}(|\boldsymbol{F}_{\mathrm{d}}| \geq n\mathfrak{m}_{\mathrm{d}} + n\delta) = \mathrm{e}^{-\Omega(n)}$; if $\beta \geq \beta_c$, for $\boldsymbol{F}_{\mathrm{o}} \sim \pi_{G,\mathrm{o}}$, it holds that $\mathbb{P}(|\boldsymbol{F}_{\mathrm{o}}| \leq n\mathfrak{m}_{\mathrm{o}} - n\delta) = \mathrm{e}^{-\Omega(n)}$.

Proof. This follows by applying Lemma 5.6 and using the percolation step analysis from Lemma 5.11. In particular, the Chernoff bounds in the proof of Lemma 5.11 give that, for $\hat{\boldsymbol{F}}_{\mathrm{d}} \sim \hat{\pi}_{G,\mathrm{d}}$, it holds that $\mathbb{P}(|\hat{\boldsymbol{F}}_{\mathrm{d}}| \geq n\mathfrak{m}_{\mathrm{d}} + n\delta) = \mathrm{e}^{-\Omega(n)}$. Therefore, from item (i) of Lemma 5.6, we have $\mathbb{P}(|\boldsymbol{F}_{\mathrm{d}}| \geq n\mathfrak{m}_{\mathrm{d}} + n\delta) = \mathrm{e}^{-\Omega(n)}$ as wanted. The proof for $\beta \geq \beta_c$ is analogous. □

This directly gives Key Property 1. Indeed, for one step of the RC dynamics to leave the relevant phase, the configuration must be on its boundary, which by Corollary 5.12 has probability $\mathrm{e}^{-\Omega(n)}$.

5.3 Proof for WSM within the disordered phase

To show WSM in the disordered regime, we follow the strategy outlined in Section 5.1.4. We begin with some relevant definitions. Recall that for $v \in V$ and an integer $\ell \geq 0$, we use $S_\ell(v)$ to denote the set of vertices at distance exactly ℓ from v. For a set of vertices S, we let $E(S)$ be the set of edges with both endpoints in S.

Definition 5.13. *Let $G = (V, E)$ be a graph and $v \in V$ be a vertex. For $K, \ell > 0$, we say that a subset $F \subseteq E$ is* K-shattered *at distance ℓ from v if all but at most K vertices of $S_\ell(v)$ belong to distinct components of the graph $(V, F \setminus E(B_\ell(v)))$.*

Lemma 5.14. *Let $d, q \geq 3$ be integers, and $\beta \leq \beta_c$ be a positive real. For any constant $\varepsilon > 0$, there is $K > 0$ such that the following holds w.h.p. over the random graph $\boldsymbol{G}$. For every vertex $v \in V$, define the event*

$$\mathcal{A}_v = \{F \subseteq E(\boldsymbol{G}) \mid F \text{ is } K\text{-shattered at distance } \ell = \lfloor (\tfrac{1}{2} - 2\varepsilon) \log_{d-1} n \rfloor \text{ from } v\}.$$

Then, for the RC disordered phase of $\boldsymbol{G}$, we have $\pi_{\boldsymbol{G},\mathrm{d}}(\bigcap_v \mathcal{A}_v) \geq 1 - 1/n^3$.

Proof. For a graph $G \in \mathbb{G}_{n,d}$ and $\sigma \in [q]^n$, recall that $\boldsymbol{F}(G, \sigma)$ denotes the random subset of edges obtained by keeping each monochromatic edge of G under σ with probability p. Fix $\varepsilon > 0$ and let $K = \lceil 10/\varepsilon \rceil$. We will consider the disordered planted graph $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d})$ and show that for an arbitrary vertex v it holds that

$$\mathbb{P}\Big[\boldsymbol{F}\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d}\big) \notin \mathcal{A}_v\Big] \leq 2/n^5. \tag{5.11}$$

Assuming this for now, Lemma 5.2 transfers this bound to the conditional Potts distribution. Together with a union bound over the vertices, this gives $\mathbb{P}\big[\boldsymbol{F}\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}}\big) \notin \bigcap_{v \in V} \mathcal{A}_v \mid \boldsymbol{G}\big] \leq \ln n / n^4$ as well. Note that $\boldsymbol{F}\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{d}}\big)$ has distribution $\hat{\pi}_{\boldsymbol{G},\mathrm{d}}$, so Lemma 5.6 yields that $\pi_{\boldsymbol{G},\mathrm{d}}(\bigcap_v \mathcal{A}_v) \geq \hat{\pi}_{\boldsymbol{G},\mathrm{d}}(\bigcap_v \mathcal{A}_v) - \mathrm{e}^{-\Omega(n)} \geq 1 - 1/n^3$.

It remains to prove (5.11). By Lemma 5.3, we have $\big\|\nu^{\hat{\boldsymbol{\sigma}}_\mathrm{d}} - \nu_\mathrm{d}\big\|_1 + \big\|\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d}} - \rho_\mathrm{d}\big\|_1 = O(n^{-1/3})$ w.p. $1 - \mathrm{e}^{-\Omega(n^{1/3})}$. Consider arbitrary ν and ρ with $\big\|\nu - \nu_\mathrm{d}\big\|_1 + \big\|\rho - \rho_\mathrm{d}\big\|_1 = O(n^{-1/3})$. Conditioned on $\nu^{\hat{\boldsymbol{\sigma}}_\mathrm{d}} = \nu$ and $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d}} = \rho$, we have that $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d})$ is uniformly random among all pairs (G, σ) with $\nu^\sigma = \nu$ and $\rho^{G,\sigma} = \rho$.

To proceed, fix σ with $\nu^\sigma = \nu$ and condition on $\hat{\boldsymbol{\sigma}}_\mathrm{d} = \sigma$. Condition further that $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d}} = \rho$ and on the induced graph on $B_\ell(v)$. For convenience, let $\tilde{\boldsymbol{F}}$ denote the (random) set of edges outside $B_\ell(v)$ after the percolation step, i.e., the set $\boldsymbol{F}\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_\mathrm{d}), \hat{\boldsymbol{\sigma}}_\mathrm{d}\big) \setminus E(B_\ell(v))$. We call $(V, \tilde{\boldsymbol{F}})$ the *percolated planted graph* (outside $\mathcal{B}_\ell(v)$). Next, we define an exploration process that will help us to bound the number of boundary components. Recall that we are working in the configuration model, where each vertex has d half-edges and the resulting (multi)graph is obtained from a perfect matching of these half-edges.

Let $T \coloneqq \lfloor n^{1/2-\varepsilon} \rfloor$. For $t = 1, \ldots, T$, we explore the (union of the) components of vertices $u \in S_\ell(v)$ in the graph $(V, \tilde{\boldsymbol{F}})$ by revealing, at each step, an edge incident to an arbitrarily chosen vertex belonging to one of these components. We denote

by A_t the set of *active vertices* in these components that still have unmatched half-edges, so that $|A_t| = 0$ implies that all the components have been fully explored. Let $A_{\le t} = \cup_{t' \le t} A_t$ denote the set of all vertices that have been active up to time t. The process is described precisely in Algorithm 5.15.

At any time t, for $i, j \in [q]$, let $M_{ij,t}$ be the number of unmatched half-edges incident to vertices of colour i that are to be matched to vertices of colour j; this number decreases whenever we reveal an edge connecting colours i and j. The revealing of an edge at time t adjacent to an active vertex w with colour $\sigma(w) = i$ has therefore two stages: we first reveal the other endpoint z in the graph $\hat{\boldsymbol{G}}(\sigma)$, chosen proportionally to the counts $M_{ij,t}$; if $\sigma(w) = \sigma(z)$, we further toss a coin with probability $p = 1 - \mathrm{e}^{-\beta}$ to determine whether the edge belongs to $\tilde{\boldsymbol{F}}$, and add z to the set of active vertices accordingly. The key point here is that we only need to reveal the components involving $S_\ell(v)$ in the graph $(V, \tilde{\boldsymbol{F}})$, rather than the components in $\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{d}})$, which are much bigger. Indeed, as we will see shortly, the former is dominated by a subcritical branching process and therefore at most $O(\log n)$ vertices are revealed for each of the vertices in $S_\ell(v)$.

To make this precise, for a time $t = 1, \dots, T$, let $\mathbf{1}_t$ denote the indicator of the event that the t-th half-edge was matched to a vertex in $A_{\le t}$, the set of vertices that have been active at some point. Let also $\mathbf{1}'_t$ be the indicator that the edge revealed was monochromatic and that it survived the percolation step. Note that $M_{ij,t}$, the number of remaining unmatched half-edges from colour i to j, satisfies $0 \le dn\rho_{ij} - M_{ij,t} \le 2T \le 2n^{1/2-\varepsilon}$, so $M_{ij,t} = dn\rho_{ij} + o(n)$ for all $t \le T$. Similarly, $|A_{\le t}| \le |S_\ell(v)| + T \le 2n^{1/2-\varepsilon}$. So, with $\mathcal{E}_t$ denoting the first t edges revealed, we have the crude bounds

$$\mathbb{P}(\mathbf{1}_t \mid \mathcal{E}_{t-1}) \le \frac{d|A_{\le t}|}{\min_i \sum_j M_{ij,t}} \le \frac{1}{n^{1/2+\varepsilon/2}}, \text{ and}$$
$$\mathbb{P}(\mathbf{1}'_t \mid \mathcal{E}_{t-1}) \le p \max_{i \in [q]} \frac{M_{ii,t}}{\sum_j M_{ij,t}} \le \frac{\mathrm{e}^\beta - 1}{\mathrm{e}^\beta + q - 1} + o(1).$$

It follows that the sum $\sum_{t=1}^T \mathbf{1}_t$ is stochastically dominated above by a sum of T independent coin tosses with probability $\frac{1}{n^{1/2+\varepsilon/2}}$, whereas the sum $\sum_{t=1}^T \mathbf{1}'_t$ is dominated above by a sum of T independent coin tosses with probability $\frac{\mathrm{e}^\beta - 1}{\mathrm{e}^\beta + q - 1} + o(1) < \frac{1-2\kappa}{d-1}$ for some small constant $\kappa = \kappa(q, d, \beta) > 0$ (using that $\mathrm{e}^\beta - 1 < q/(d-2)$ for $\beta \le \beta_c < \beta'_u$). Therefore, recalling that $K = \lceil 10/\varepsilon \rceil$ and $T = \lfloor n^{1/2-\varepsilon} \rfloor$, we have

$$\mathbb{P}\Big(\sum_{t=1}^T \mathbf{1}_t > K \Big) \le \binom{T}{K} \frac{1}{n^{K(1/2+\varepsilon/2)}} \le 1/n^5,$$

Exploration of $(V, \tilde{\boldsymbol{F}})$ starting from $S_\ell(v)$
Let $t = 0$, $T = \lfloor n^{1/2-\varepsilon} \rfloor$, and $A_0 = S_\ell(v)$.
While $|A_t| \neq 0$ or $t \leq T$:

- If $|A_t| = 0$, pick any unmatched half-edge (w, r) and run $\textsc{Match}(w, r)$; increase t by 1 and set $A_t = \emptyset$
- Otherwise, pick $w \in A_t$. For $r = 1, \ldots, d$:

  If the half-edge (w, r) is not matched:
  - increase t by 1;
  - match (w, r) to another half-edge, let $z = \textsc{Match}(w, r)$.
  - if $\sigma(z) = \sigma(w)$, with probability $p = 1 - \mathrm{e}^{-\beta}$ set $A_t \leftarrow A_{t-1} \cup \{z\}$; else set $A_t = A_{t-1}$.
- Remove w and any other vertices from A_t which do not have unmatched half-edges.

$\textsc{Match}(w, r)$ For $i, j \in [q]$, let $M_{ij,t}$ be the number of currently unmatched half-edges from colour i to j.

- Let $i = \sigma(w)$. Choose a colour $j \in [q]$ from the distribution $\{\frac{M_{ij,t}}{\sum_{j'} M_{ij',t}}\}_{j \in [q]}$.
- From the set of unmatched half-edges from colour j to i, choose one u.a.r.[a], say (z, r').
- Match (z, r') with (w, r), and return the vertex z.

[a] Uniformly at random.

Algorithm 5.15: The exploration process of the percolated planted graph used in the proof of Lemma 5.14. This is a combination of first exposing an edge in the planted configuration model, captured by the routine $\textsc{Match}(\cdot, \cdot)$, followed by a percolation step (i.e., including the edge in the RC configuration with probability $p = 1 - \mathrm{e}^{-\beta}$ whenever the edge is monochromatic).

and

$$\mathbb{P}\bigg(\sum_{t=1}^{T}\mathbf{1}'_t \geq \frac{T(1-\kappa)}{d-1}\bigg) = \mathrm{e}^{-\Omega(n^{1/3})},$$

where the second probability bound follows from standard Chernoff bounds (using that $T = \Omega(n^{1/3})$).

Note that the event $|A_T| > 0$ implies that $|A_t| > 0$ for all $t = 0, 1, \dots, T$, which in turn implies that at least one vertex gets removed from the set of active vertices every $d-1$ time steps (or less); indeed, when exploring the neighbours of a vertex $w \in A_t$ at a particular time t, there are at most $d-1$ half-edges incident to w remaining unmatched since one of the d half-edges of w has been previously matched (when w is first added to the set of active vertices). We therefore have the inequality $0 \leq |A_T| \leq |S_\ell(v)| + \sum_{t=1}^{T}\mathbf{1}'_t - \frac{T-(d-1)}{d-1}$. Since $|S_\ell(v)| \leq d(d-1)^\ell \leq dn^{1/2-2\varepsilon} \leq \frac{\kappa T}{d-1}$, it follows that

$$\begin{aligned}\mathbb{P}(|A_T| > 0) &\leq \mathbb{P}\bigg(|S_\ell(v)| - \frac{T-(d-1)}{d-1} + \sum_{t=1}^{T}\mathbf{1}'_t \geq 0\bigg)\\ &\leq \mathbb{P}\bigg(\sum_{t=1}^{T}\mathbf{1}'_t \geq \frac{T(1-\kappa)}{d-1}\bigg) = \mathrm{e}^{-\Omega(n^{1/3})}.\end{aligned}$$

Combining these, we obtain that with probability at least $1-2/n^5$, both $|A_T| = 0$ and $\sum_{t=1}^{T}\mathbf{1}_t < K$ hold. The first implies that by time T all the components of $S_\ell(v)$ in the graph $(V, \tilde{\boldsymbol{F}})$ have been explored, and the second implies that all but at most K of these vertices belong to distinct components. This finishes the proof of (5.11) and hence completes the proof of the lemma. □

Next, we combine Lemma 5.14 with the fact that random regular graphs are locally treelike. Recall that local treelikeness means that every small neighbourhood can be transformed into a tree by removing at most some constant number of edges. We state the exact result we use in the following lemma. We use an argument of Blanca and Gheissari [10] which shows that, when the configuration outside a treelike ball $B_\ell(v)$ is K-shattered, the difference between the marginal probability of an edge incident to v under this boundary condition and under the free boundary condition is bounded by the probability of the existence of (roughly) two root-to-leaf paths of in-edges. The argument in [10] is technically proved for $\beta < \beta_u$ using some uniqueness estimates, but it can be extended to the regime $\beta \leq \beta_c$ by observing that in a treelike ball, a

K-shattered configuration induces a distribution that is comparable to percolation[5] on the tree with parameter $\hat{p}$ (this coincides with the cut-edge update probability in the definition of the random-cluster dynamics). In the lemma below, we use this to bound the probability of this "two-paths" event by $\hat{p}^{2\ell}$. Taking $\ell = (\frac{1}{2} - \delta)\log_{d-1} n$ for a small positive δ then makes the difference in marginals at most $\frac{1}{100m}$, as required for WSM within the disordered phase.[6] The following lemma makes this argument precise and gives the bound on the difference in marginals needed for WSM within the disordered phase. In the following proof, we use *non-trivial boundary component* to denote a boundary component (see 4.1.5) that contains more than one boundary vertex.

Lemma 5.16. *Let $d, q \geq 3$ be integers, and $\beta \leq \beta_c$, $\delta \in (0, 1/2)$ be positive reals.*

There exists a constant $C = C(d, q, \beta, \delta) > 0$ such that w.h.p. over $G \sim \mathbb{G}_{n,d}$ the following holds. For every $v \in V$ and an edge e incident to v, and $\ell = \lfloor(\frac{1}{2} - \delta)\log_{d-1} n\rfloor$,

$$\left|\pi^{-}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{d}}(1_e)\right| \leq C\hat{p}^{2\ell} + \frac{1}{n^3} \text{ where } \hat{p} \coloneqq \tfrac{\mathrm{e}^{\beta}-1}{q+\mathrm{e}^{\beta}-1}.$$

Proof. First of all, it is well-known that w.h.p. $G \sim \mathbb{G}_{n,d}$ is *locally treelike* (see e.g. [10, Fact 2.3]), i.e., that there exists a constant $L = L(\delta, d)$ such that for all v, we can remove at most L edges from $B_\ell(v)$ to make it a tree. Second, by Lemma 5.14, there exists $K = K(d, q, \beta, \delta)$ such that w.h.p. over $G \sim \mathbb{G}_{n,d}$, the probability that for all v, a random disordered configuration is K-shattered at distance ℓ from v is at least $1 - n^{-3}$. Thus, from now on we assume G satisfies these two properties.

Fix a vertex v and an edge e incident to it. With probability at most n^{-3}, a disordered configuration is not K-shattered at distance ℓ from v, which gives the n^{-3} term in the bound. Now assume a configuration is K-shattered and let η be the resulting boundary condition on $B_\ell(v)$. To finish the proof, it is sufficient to show that there is a constant C such that $|\pi^{-}_{G,v,\ell}(1_e) - \pi^{\eta}_{G,v,\ell}(1_e)| \leq C\hat{p}^{2\ell}$.

For the case $\beta < \beta_u$, this bound was proven in [10, Proposition 3.3]. However, parts of their proof use estimates that only hold in uniqueness. We are able to extend their bound for all $\beta \leq \beta_c$ by closely following the start of their argument and then using different estimates which continue to hold beyond the uniqueness threshold.

[5]Here, by percolation on a graph with parameter $\hat{p}$, we mean the random edge subset obtained by including each edge independently with probability $\hat{p}$.

[6]This is notably the threshold up to which the required local treelikeness holds: for $\ell \geq \frac{1}{2}\log_{d-1} n$, radius-$\ell$ neighbourhoods of typical random regular graphs may have a non-constant number of excess edges.

The starting points for the proof are two results from [10]: Lemma 5.3 and equation (5.4). Lemma 5.3 bounds the TV-distance on marginals on e by $\pi^{\eta}_{G,v,\ell}(\Upsilon^{\eta}_{G,v,\ell})$, where $\Upsilon^{\eta}_{G,v,\ell}$ is the event that a configuration $\boldsymbol{F}$ satisfies the following: for all $1 \leq i \leq \ell$, there are two *distinct* vertices $u_1, u_2 \in S_i(v)$ and two corresponding paths $\gamma_{i,1}, \gamma_{i,2}$ such that for both $j \in \{1, 2\}$:

- $\gamma_{i,j}$ is a path of edges in $\boldsymbol{F}$ from u_j to a vertex of $S_\ell(v)$ belonging to a non-trivial boundary component of η, and
- vertices of $\gamma_{i,j}$ are contained in $\bigcup_{i \leq i' \leq \ell} S_{i'}(v)$.

Then in the proof of [10, Proposition 3.3], $\pi^{\eta}_{G,v,\ell}(\Upsilon^{\eta}_{G,v,\ell})$ is further upper-bounded by a probability of an event roughly corresponding to two disjoint paths from a neighbour of v to $S_\ell(v)$. In order to define this event, we need to introduce more notation.

Since G is locally L-treelike, there is a set $E_{\mathcal{E}} \subseteq E(B_\ell(v))$ of at most L edges such that removing $E_{\mathcal{E}}$ makes the graph $G[B_\ell(v)]$ into a tree. Thus there are also $k \leq 2L$ distances $1 \leq d_1 < \cdots < d_k \leq \ell$ such that $S_{d_i}(v)$ contains a vertex from $V(E_{\mathcal{E}})$. Also, let $d_0 = 1, d_{k+1} = \ell$. The following bound is shown in [10, Equation (5.4)]:

$$\pi^{\eta}_{G,v,\ell}(\Upsilon^{\eta}_{G,v,\ell}) \leq K^2 (2L)^{2L} (2^L)^{2L} \sup_{(\Gamma;\Gamma')} \pi^{\eta}_{G,v,\ell}(\Gamma \cup \Gamma' \in \cdot), \tag{5.12}$$

where the supremum is over all pairs of path sequences $(\Gamma; \Gamma') = (\gamma_0, \ldots, \gamma_k; \gamma'_0, \ldots, \gamma'_k)$ such that for each $i = 0, \ldots, k$, γ_i, γ'_i are disjoint paths contained in $\bigcup_{d_i < j < d_{i+1}} S_j(v)$, such that they can be both extended, by edges in $\bigcup_{d_i < j \leq \ell} S_j(v)$ to a path to a non-trivial boundary component of η (note that γ_i, γ'_i are allowed to be empty).

So, it is enough to show that for all such $(\Gamma; \Gamma')$, it holds that $\pi^{\eta}_{G,v,\ell}(\Gamma \cup \Gamma' \in \cdot) \leq C\hat{p}^{2\ell}$ for some constant $C > 0$. Let $(\Gamma; \Gamma')$ be any such pair of path sequences. Consider the distribution on the ball $B_\ell(v)$ with the free boundary and condition further on all the excess edges in $E_{\mathcal{E}}$ being out; for brevity, denote this distribution by $\tilde{\pi}$. Since removing $E_{\mathcal{E}}$ from the ball makes it a tree, $\tilde{\pi}$ is the RC model on a tree, which is the same distribution as including each edge independently with probability $\hat{p}$. Furthermore, $\Gamma \cup \Gamma'$ does not contain any edges in $E_{\mathcal{E}}$. Thus, it is an event with positive probability under $\tilde{\pi}$. Also, $\Gamma \cup \Gamma'$ contains at least $\sum_{i=0}^{k} 2[d_{i+1} - d_i - 2] \geq 2\ell - 8L$ edges, therefore $\tilde{\pi}(\Gamma \cup \Gamma') \leq \hat{p}^{2\ell - 8L}$.

Since $\hat{p}^{-8L}$ is a constant, as a final step, it suffices to show that $\frac{\pi^{\eta}_{G,v,\ell}(\Gamma \cup \Gamma' \in \cdot)}{\tilde{\pi}(\Gamma \cup \Gamma' \in \cdot)}$ is bounded by a constant (not depending on the particular $(\Gamma; \Gamma')$). Consider any configuration $A \subseteq E(B_\ell(v))$. Since $|E_{\mathcal{E}}| \leq L$, then η is K-shattered, and adding an

edge can decrease the number of components by at most one, we get that $|A \setminus E_{\mathcal{E}}| \leq |A| \leq |A \setminus E_{\mathcal{E}}| + L$, and the number of components in the ball for configuration A w.r.t.[7] boundary condition η differs by at most $K+L$ from the number of components for configuration $A \setminus E_{\mathcal{E}}$ w.r.t. the free boundary condition (that is, we compare the number of components when the state of the edges outside the ball is given by η and when all these edges, as well as all excess edges, are out). Thus the weights of A w.r.t. η and of $A \setminus E_{\mathcal{E}}$ w.r.t. the free boundary condition − differ by a factor of at most $q^{L+K}[\max(\mathrm{e}^{\beta}-1,(\mathrm{e}^{\beta}-1)^{-1})]^{L}$.

Given that there are at most $2^{|E_{\mathcal{E}}|} \leq 2^{L}$ configurations $A \subseteq E(B_{\ell}(v))$ agreeing on $A \setminus E_{\mathcal{E}}$, we get that the sums of weights of all configurations satisfying "$\Gamma \cup \Gamma'$ are all in" differ by a constant factor when considering the boundary condition η and when all edges in $E_{\mathcal{E}}$ and all edges outside the ball are fixed to be out, and so do the respective partition functions of these two RC distributions. It hence follows that $\frac{\pi^{\eta}_{G,v,\ell}(\Gamma\cup\Gamma'\in\cdot)}{\tilde{\pi}(\Gamma\cup\Gamma'\in\cdot)}$ is bounded by a constant, which concludes the proof. □

We can now give the proof of WSM within the disordered phase (Theorem 5.9).

Theorem 5.9. *Let $d,q \geq 3$ be integers, and $\beta \leq \beta_c$ be a positive real. Then, there exists $\delta > 0$ such that, w.h.p. over $G \sim \mathbb{G}_{n,d}$, $G=(V,E)$ has WSM within the disordered phase at a distance $\ell = \lfloor(\frac{1}{2}-\delta)\log_{d-1} n\rfloor$. That is, with $m=|E|$, for every $v \in V$ and every edge e incident to v, it holds that $\left|\pi^{-}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{d}}(1_e)\right| \leq \frac{1}{100m}$.*

Proof. Consider $\beta \leq \beta_c$. Since $\beta_c < \beta'_u = \ln(1+\frac{q}{d-2})$, we have that $\hat{p}(\beta) < \hat{p}(\beta'_u) = \frac{1}{d-1}$. Hence there is an $\varepsilon := \varepsilon(q,d,\beta) > 0$ such that $\hat{p}(\beta) = \frac{1}{(d-1)^{1+\varepsilon}}$. Let $\delta := \frac{\varepsilon}{4(1+\varepsilon)}$ and $\ell := \lfloor(\frac{1}{2}-\delta)\log_{d-1} n\rfloor$.

Lemma 5.16 gives that, w.h.p. over $G \sim \mathbb{G}_{n,d}$, for every vertex v and every edge e incident to v, it holds that $\left|\pi^{-}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{d}}(1_e)\right| \leq C\hat{p}^{2\ell} + \frac{1}{n^3}$. By the choice of δ,ε, and ℓ, we obtain that $\hat{p}^{2\ell} \leq n^{-(1+\varepsilon/2)}/(\hat{p}^2)$, so for sufficiently large n, we have $C\hat{p}^{-2}n^{-(1+\varepsilon/2)} + n^{-3} \leq \frac{1}{100m}$. Thus, WSM within the disordered phase holds at radius $\ell = \lfloor(\frac{1}{2}-\delta)\log_{d-1} n\rfloor$, concluding the theorem. □

5.4 Proof for WSM within the ordered phase

In this section, we show WSM within the ordered phase. Recall that the aim is to bound, for every vertex v, every edge e incident to v, and some $\ell = \lfloor(\frac{1}{2}-\delta)\log_{d-1} n\rfloor$,

$$|\pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e)| \leq \tfrac{1}{100m}, \tag{5.13}$$

[7] With respect to.

where $\pi^{+}_{G,v,\ell}$ is the RC distribution where all edges outside of the ball $B_\ell(v)$ are fixed to be in, and all vertices in $S_\ell(v)$ are furthermore fixed to be in a single component.

To show (5.13) for $G \sim \mathbb{G}_{n,d}$ w.h.p., we follow the strategy outlined in Section 5.1.4. We first define the notion of a wired boundary used in this chapter and bound the l.h.s. of (5.13) by the probability of non-existence of the wired boundary. Then we show that a wired boundary occurs if there are no paths of length $\ell \approx \frac{1}{2}\log_{d-1} n$ that avoid the largest component in a typical configuration from $\pi_{G,\mathrm{o}}$ *after removing one vertex.*

5.4.1 Wired boundary in ordered configurations

We next define what it means for an ordered configuration to have a *wired boundary*, and then we bound the difference in marginals in (5.13) by the non-existence of the wired boundary.

Recall that, in the previous chapter, the wired boundary was a property of *partial configurations* in a particular revealing process. Here, instead, we define it as a property of the resulting configuration and, roughly, work with multiple revealing processes at once.

Definition 5.17. *Let $G = (V, E)$ be a graph, $v \in V$ a vertex, $\ell \geq 1$ integer, and let $F \subseteq E$ be a RC configuration. We say that $S \subseteq V(B_\ell(v)) \setminus \{v\}$ is a* wired boundary around v in $B_\ell(v)$*, if*

- *S is a* cut-set separating v *from the outside of the ball. That is, $\mathcal{C}_v = \mathcal{C}_v(S)$, the component of v in $G \setminus S$, is disjoint from $G \setminus B_\ell(v)$; and*
- *in the graph $(V \setminus V(\mathcal{C}_v), \{e \in F : e \cap V(\mathcal{C}_v) = \emptyset\})$, all vertices of S are in the same component.*

For any cut-set $S \subseteq V(B_\ell(v))\setminus\{v\}$ separating v (not necessarily a wired boundary), write $R_S = \{e \in E : e \cap V(\mathcal{C}_v(S)) = \emptyset\}$ for the outside edges*. The edges $E \setminus R_S$ are the* interior edges *of S.*

To use the wired boundary, we will work with the configuration where v is removed. For a configuration $F \subseteq E$, write $F_v = \{f \in F : f \text{ incident to } v\}$.

The next lemma shows that the difference in marginals 5.13 is bounded by the probability of no wired boundary and the probability that the outside of the ball contains sufficiently many edges. Formally, for the second event we write

$$\mathcal{O} \coloneqq \{F \subseteq E : |F \setminus E(B_\ell(v))| \geq n\mathfrak{m}_{\mathrm{o}}(\beta_c) - n\rho/4\}, \tag{5.14}$$

where $\mathfrak{m}_{\mathrm{o}}(\beta_c) - \rho/4$ is the threshold for the ordered phase from Definition 5.4. Note that the probability of $\overline{\mathcal{O}}$ is $\mathrm{e}^{-\Omega(n)}$.

Lemma 5.18. *Let $d, q \geq 3$ be integers, and $\beta \geq \beta_c$ a real. Then w.h.p. over $G \sim \mathbb{G}_{n,d}$ the following holds. For any $v \in V$, any edge e incident to v, and any $\ell \leq \frac{1}{2}\log_{d-1} n$,*

$$\left|\pi^+_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e)\right| \leq \pi_{G,\mathrm{o}}\big(\text{no wired boundary around } v \text{ in } B_\ell(v)\big) + \frac{2\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}{1 - \pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}.$$

Proof. The intuitive idea is that we bound the difference in marginals in (5.13) by the sum of three terms: a) the probability of $\overline{\mathcal{O}}$, b) the probability of no wired boundary, and c) the difference in marginals when a wired boundary exists. Unlike in the previous chapter, we do not know in advance which cut-set will form a wired boundary, and hence which revealing scheme to use. Instead, we condition on the "outermost cut-set" which gets connected in the resulting configuration, and show, using a revealing scheme and monotonicity of the random-cluster model, that the required bound on the difference in marginals holds in each case.

First, by conditioning on $\mathcal{O}$ and using the triangle inequality we obtain that the difference in marginals from (5.13) is at most

$$\begin{aligned}
&|\pi_{G,\mathrm{ord}}(1_e \mid \mathcal{O})(1 - \pi_{G,\mathrm{o}}(\overline{\mathcal{O}})) - \pi_{G,v,\ell}(1_e) + \pi_{G,\mathrm{o}}(\overline{\mathcal{O}})\pi_{G,\mathrm{o}}(1_e \mid \overline{\mathcal{O}})| \\
&\qquad \leq |\pi_{G,\mathrm{ord}}(1_e \mid \mathcal{O}) - \pi_{G,v,\ell}(1_e)| + |\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})(\pi_{G,\mathrm{o}}(1_e \mid \overline{\mathcal{O}}) - \pi_{G,\mathrm{ord}}(1_e \mid \mathcal{O}))| \\
&\qquad\qquad \leq |\pi_{G,\mathrm{ord}}(1_e \mid \mathcal{O}) - \pi_{G,v,\ell}(1_e)| + \pi_{G,\mathrm{o}}(\overline{\mathcal{O}}), \qquad (5.15)
\end{aligned}$$

where in the last inequality, we used that $|\pi_{G,\mathrm{o}}(1_e \mid \overline{\mathcal{O}}) - \pi_{G,\mathrm{ord}}(1_e \mid \mathcal{O})| \leq 1$.

Every configuration in $\mathcal{O}$ is ordered. Hence, we can drop the conditioning on being in the ordered phase, i.e. $\pi_{G,\mathrm{o}}(\cdot \mid \mathcal{O}) = \pi_G(\cdot \mid \mathcal{O})$. The boundary condition + contains all edges outside of the ball. Hence, monotonicity (see 4.1.6) implies that π^+ stochastically dominates $\pi_{G,\mathrm{o}}(\cdot \mid \mathcal{O})$. Consequently $\pi^+(1_e) \geq \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O})$, so it suffices to upper-bound $\pi^+(1_e) - \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O})$.

Next, let $\mathcal{W}$ be the event that there is a wired boundary around v in $B_\ell(v)$. By conditioning on this event, we obtain

$$\begin{aligned}
&\pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O}) \\
&\quad = \pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O},\mathcal{W})\pi_{G,\mathrm{o}}(\mathcal{W} \mid \mathcal{O}) - \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O},\overline{\mathcal{W}})\pi_{G,\mathrm{o}}(\overline{\mathcal{W}} \mid \mathcal{O}) \\
&\qquad\qquad \leq (\pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O},\mathcal{W})) + \pi_{G,\mathrm{o}}(\overline{\mathcal{W}} \mid \mathcal{O}).
\end{aligned}$$

Hence, provided we can establish

$$\pi^{+}_{G,v,\ell}(1_e) \leq \pi_{G,\mathrm{o}}(1_e \mid \mathcal{O},\mathcal{W}) = \pi_G(1_e \mid \mathcal{O},\mathcal{W}), \tag{5.16}$$

the lemma is then obtained using $\frac{\pi_{G,\mathrm{o}}(\overline{\mathcal{W}})}{1-\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})} + \pi_{G,\mathrm{o}}(\overline{\mathcal{O}}) \leq \pi_{G,\mathrm{o}}(\overline{\mathcal{W}}) + \frac{2\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}{1-\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}$.

Therefore, for the remainder of the proof, we focus on proving (5.16). The idea is to find a "witness" for the event $\mathcal{W}$, such that each configuration satisfying $\mathcal{W}$ has exactly one witness. Moreover, conditional on any particular witness, the inequality (5.16) is preserved. This will allow us to partition the state space and obtain (5.16) by summing over all possible witnesses.

To formalise this strategy, for $S \subseteq B_\ell(v)$, a cut-set separating v from the outside of the ball, we define the event E_S as follows:

$$E_S \coloneqq \{F \in \mathcal{O} \mid \text{all vertices of } S \text{ are in the same component of } (V, F \setminus F_v)\}. \tag{5.17}$$

From now on, for brevity, when we write "some E_S holds" or simply E_S, we implicitly assume S is such a cut-set. We prove the following claim.

Claim 1. *$\mathcal{O} \cap \mathcal{W} = \bigcup_S E_S$, where the union is over all cut-sets $S \subseteq B_\ell(v) \setminus \{v\}$ separating v from the outside of the ball.*

If $F \in \mathcal{O} \cap \mathcal{W}$, take S to be a wired boundary. By definition, its vertices are connected without using $\mathcal{C}_v$, and in particular without using v, so $F \in E_S$.

Conversely, suppose $F \in E_S$, and let $\mathcal{C}_S$ be the component containing S in $(V, F \setminus F_v)$. Let $S' = V(\mathcal{C}_S) \cap V(B_\ell(v))$. Clearly, S' is contained in the ball. Moreover, since $\mathcal{C}_S$ is a component in a graph containing no edges incident to v and $v \notin S$, we have $v \notin V(\mathcal{C}_S)$, and thus $v \notin S'$. Since $S \subseteq S'$ and S is a cut-set separating v from the outside of the ball, S' is also such a cut-set. Moreover, S' is connected in $(V \setminus V(\mathcal{C}_v), \{f \in F : f \cap V(\mathcal{C}_v) = \emptyset\})$, where $\mathcal{C}_v$ is the component of v in $G \setminus S'$. This follows since a vertex $w \in \mathcal{C}_S$ is either outside of the ball (and thus not in $\mathcal{C}_v$ since S' is a cut-set separating v from the outside of the ball), or it is in S', thus not in $\mathcal{C}_v$ (since $\mathcal{C}_v$ is a component in $G \setminus S'$). Therefore, all edges of $\mathcal{C}_S$ belong to

$\{f \in F : f \cap V(\mathcal{C}_v) = \emptyset\}$, so S' remains connected in this graph and is thus a wired boundary.

Therefore, from now on we work with $\bigcup E_S$ instead of $\mathcal{W} \cap \mathcal{O}$. The next step is to define a *canonical cut* such that each $F \in \bigcup E_S$ has precisely one canonical cut, and canonical cuts have the following property: if S is a canonical cut, and $F' \subseteq R_S$ is the state of the edges *outside the cut*, then for any configuration $\tilde{F} \subseteq E \setminus R_S$ on the *interior edges* such that $\tilde{F} \cup F' \in E_S$, it must be the case that S will also be the canonical cut-set for this configuration.

We next define the notion formally, and, in a series of claims, we show it satisfies the desired properties, which we then use to show (5.16). Let $F \in \bigcup_S E_S$ be a configuration with a wired boundary, and let $\mathcal{E}(F) = \{S \text{ a cut-set} \mid F \in E_S\}$ be the set of "connected cut-sets" under F. The *canonical cut-set* $S_c(F)$ of F is defined as follows. For every (not necessarily simple[8]) path in G from $S_{\ell+1}(v) \ni w_0, w_1, \ldots, w_k = v$ such that $w_1, \ldots, w_k \in B_\ell(v)$, let w_j be its first vertex belonging to $\bigcup_{S \in \mathcal{E}(F)} S$. Then $S_c(F)$ is the set of all vertices w_j obtained in this way.

Claim 2. *$S_c(F)$ is a cut-set separating v from the outside of the ball and $F \in E_{S_c(F)}$.*

First of all, $S_c(F) \subseteq \bigcup_{S \in \mathcal{E}(F)} S \subseteq V(B_\ell(v)) \setminus \{v\}$. Also, by definition, every path from v to the outside of the ball goes through a vertex in $S_c(F)$. Therefore, $S_c(F)$ is a cut-set separating v from the outside of the ball.

To conclude the claim, we need to show that all vertices of $S_c(F)$ are in a connected component of $(V, F \setminus F_v)$. Let $u, w \in S_c(F)$. If $\exists S \in \mathcal{E}(F)$ such that $u, w \in S$, we are done, as the definition of $F \in E_S$ implies u, w are in the same component of $(V, F \setminus F_v)$. So assume there are *distinct* $S_u, S_w \in \mathcal{E}(F)$ such that $u \in S_u \setminus S_w$ and $w \in S_w \setminus S_u$. Since $u \in S_c(F)$, there exists a path $S_{\ell+1}(v) = u_0, u_1, \ldots, u_k = v$ in G passing through u such that $u = u_i$ for $i = \min\{1 \leq i' < k : u_{i'} \in \bigcup_{S \in \mathcal{E}(F)} S\}$. Notice that a vertex from S_w also belongs to this path, since S_w is a cut-set separating v from the outside of the ball (and hence $G \setminus S_w$ cannot contain any path from v to u_0 in particular), so let i' be an index such that $w' = u_{i'} \in S_w$. By the choice of u and the path, and also since we assumed u, w are in *distinct* $S_u, S_w \in \mathcal{E}(F)$, it must be the case that $i' > i$. We now distinguish two cases.

First, suppose that none of the vertices in $u_{i'}, u_{i'+1}, \ldots, u_k = v$ is in S_u. Now consider a path obtained by concatenating the following three paths: path that goes from a vertex in $S_{\ell+1}(v)$ to w which is (apart from its endpoint) contained in the ball $B_\ell(v)$ and has no vertices of S_u (such a path exists by the choice of w), a path from

[8] A simple path has no repeated vertices. In this section we work with both simple paths and not necessarily simple paths. The latter might contain a vertex multiple times.

w to w' using only the edges in $F \setminus F_v$ (such path exists since $F \in E_{S_w}$), and the path $w' = u_{i'}, w_{i'+1}, \ldots, u_k = v$. A suffix of this path is a (not necessarily simple) path from $S_{\ell+1}(v)$ to v that is (apart from its endpoint) contained in $B_\ell(v)$. Thus, some suffix of this path must contain a vertex of S_u. By construction, the first and third parts of the concatenated path do not contain vertices of S_u. Therefore, a vertex of S_u needs to be contained on the path from w to w' using only edges of $F \setminus F_v$. Thus, this vertex is, in particular, connected to w. Since it is also connected to u in $(V, F \setminus F_v)$ and "being connected to" is a transitive symmetric property, u is connected to w.

Second, suppose there is a vertex of S_u in $u_{i'}, u_{i'+1}, \ldots, u_k = v$. Let j be the greatest index such that $u_j \in S_u$. If there is a vertex of S_w in $u_{j+1}, \ldots, u_k$, we use the same strategy as before (swapping i' for the appropriate index). Otherwise, we construct a path by concatenating the following three paths: the path $S_{\ell+1}(v) \ni u_0, u_1, \ldots, u_i$, the path between u and u_j using only edges of $F \setminus F_v$ (such path exists since $F \in E_{S_u}$) and a path $u_j, u_{j+1}, \ldots, u_k = v$. Some suffix of this path is a path from the boundary of the ball to v, and therefore it has to contain a vertex from S_w. By choice of the first and the third concatenated paths, this vertex of S_w needs to be on the path using the edges of $F \setminus F_v$ between u and u_j. Thus in particular, some vertex of S_w is connected to u via edges of $F \setminus F_v$. Every vertex of S_w is connected to $w \in S_w$ via the edges of $F \setminus F_v$. Hence, by transitivity, u and w are in the same component of $(V, F \setminus F_v)$, which concludes the claim.

Next, let $\mathcal{S}$ be the set of all canonical cuts, i.e.

$$\mathcal{S} = \{S \text{ a cut-set} : \exists F \in E_S \text{ such that } S = S_c(F)\},$$

and then define, for $S \in \mathcal{S}$ and each $F' \subseteq R_S$ (where R_S is the set of edges outside the cut S),

$$Q_{S,F'} = \{F \in E_S : S = S_c(F) \text{ and } \cdot \cap R_S = F'\}.$$

Since every F for which some E_S holds has a canonical cut $S_c(F)$, the non-empty events $Q_{S,F'}$ form a disjoint partition of $\mathcal{O} \cap \mathcal{W}$.

Further, the next two claims show that each event $Q_{S,F'}$ is *increasing* and that the components of vertices outside the cut are determined by F'.

Claim 3. *Fix $S \in \mathcal{S}$ and $F' \subseteq R_S$. Consider any two vertices $w, w' \in V \setminus V(\mathcal{C}_v)$ (where $\mathcal{C}_v$ is the component of v in $G \setminus S$) and any two configurations on the interior edges $\tilde{F}_1, \tilde{F}_2 \subseteq E \setminus R_S$ such that $F' \cup \tilde{F}_1, F' \cup \tilde{F}_2 \in E_S$. Then w, w' are in the same component in $(V, F' \cup \tilde{F}_1)$ if and only if they are in the same component in $(V, F' \cup \tilde{F}_2)$.*

In particular, if the component of w in (V, F') does not contain a vertex in S, then this component is the same in $(V, F' \cup \tilde{F}_1)$ for both $i = 1, 2$.

We show the first direction; the other is identical by symmetry. Let $w = w_0, w_1, \ldots, w_k = w'$ be a path using only edges in $F' \cup \tilde{F}_1$. If the path is contained in F' then we are done. Otherwise, it must use an edge incident to a vertex of $\mathcal{C}_v$, and since S is a cut separating $\mathcal{C}_v$ from (in particular) the vertices w, w', any such path needs to contain a vertex of S. Let $1 \leq j_1 \leq j_2 < k$ be respectively the least and the greatest index such that $w_{j_1}, w_{j_2} \in S$. Then $w = w_0, \ldots, w_{j_1}$ and $w_{j_2}, \ldots, w_k = w'$ are both paths in $F' \subseteq F' \cup \tilde{F}_2$. Moreover, since $F' \cup \tilde{F}_2 \in E_S$, all vertices of S, and thus in particular w_{j_1}, w_{j_2}, are connected by the edges of $F' \cup \tilde{F}_2$. Therefore, there is a path from w_{j_1} to w_{j_2} using the edges of $F' \cup \tilde{F}_2$, and hence also a path from w to w', which concludes the first part of the claim. For the second part, if $w \notin S$ has no path to $\{w'\} \cup S$ in (V, F'), then clearly, for both $i = 1, 2$, it cannot have any path to $\{w'\} \cup S$ in $(V, F' \cup \tilde{F}_i)$ as no edges in $E \setminus R_S \supseteq \tilde{F}_i$ are incident to w. Thus, for any $w \in V \setminus V(\mathcal{C}_v)$ whose component in (V, F') does not contain S (hence also $w \notin S$), this component is preserved in $(V, F' \cup \tilde{F}_i)$, which concludes the proof.

As a direct consequence, we get the following property.

Claim 4. *Let $S \in \mathcal{S}$ and $F' \subseteq R_S$ such that $Q_{S,F'} \neq \emptyset$. Let $\tilde{F} \subseteq E \setminus R_S$ be a subset of interior edges such that $\tilde{F} \cup F' \in E_S$. Then also $\tilde{F} \cup F' \in Q_{S,F'}$, i.e. S is the canonical cut-set of $\tilde{F} \cup F'$.*

Let $S_c(F' \cup \tilde{F}) = S'$ and let $F \in Q_{S,F'}$. We will show $S' \subseteq V \setminus V(\mathcal{C}_v)$. Then, by Claim 3, it holds that S' is connected in (V, F), and thus, by the choice of the canonical cut-set, it must be the case that $S' = S$ (otherwise there would be a path from $S_{\ell+1}(v)$ to v contained in the ball (apart from its endpoint) such that a first vertex of S' comes before a first vertex of S. But that contradicts that S is a canonical cut-set for F).

Consider any $w \in S'$. By definition of a canonical cut-set, there must be a path $S_{\ell+1}(v) = w_0, w_1, \ldots, w_k = v$ contained in the ball (apart from its endpoint) such that $w = w_j$ and j is the least index such that $w_j \in \bigcup_{\tilde{S}: F' \cup \tilde{F} \in E_{\tilde{S}}} \tilde{S}$. Let $w_j \in \tilde{S}$ (and $F' \cup \tilde{F} \in E_{\tilde{S}}$). Then all vertices of $\tilde{S}$ are in the single component of $F' \cup \tilde{F}$. Now, since S is a cut-set, there must be a $j' \geq 1$ such that $w_{j'} \in S$. Let j' be the least such index. Since $F' \cup \tilde{F} \in E_S$, we have $S \in \mathcal{E}(F' \cup \tilde{F})$, and hence, by the minimality of j, it holds that $j' \geq j$. Hence, w is connected to the outside of the ball by a path not containing a vertex of S, thus it cannot be the case that $w \in V(\mathcal{C}_v(S))$. This concludes the proof of the claim.

As a consequence, $Q_{S,F'}$ is *increasing* (on $E \setminus R_S$): for any $\tilde{F}_1 \subseteq \tilde{F}_2 \subseteq E \setminus R_S$, if $F' \cup \tilde{F}_1 \in Q_{S,F'}$, then adding edges (of $E \setminus R_S$) cannot disconnect the component of S, and, by Claim 4, $S_c(F' \cup \tilde{F}_1) = S_c(F' \cup \tilde{F}_2)$. Hence in particular, $Q_{S,F'} = E_S \cap \{F : F \cap R_S = F'\}$.

With this in hand, we write

$$\pi_{G,\mathrm{o}}(1_e \mid \mathcal{O}, \mathcal{W}) - \pi^+_{G,v,\ell}(1_e) = \sum_{\substack{S \in \mathcal{S} \\ F' \subseteq R_S \\ Q_{S,F'} \neq \emptyset}} \pi_{G,\mathrm{o}}(Q_{S,F'} \mid \mathcal{O}, \mathcal{W}) \left(\pi_{G,\mathrm{o}}(1_e \mid Q_{S,F'}) - \pi^+_{G,v,\ell}(1_e)\right).$$

Since probabilities are non-negative, to establish (5.16), it now suffices to show, for each $S \in \mathcal{S}$ and each $F' \subseteq R_S$ with $Q_{S,F'}$ non-empty, that $\pi_{G,\mathrm{o}}(1_e \mid Q_{S,F'}) \geq \pi^+_{G,v,\ell}(1_e)$.

Fix S and F' with non-empty $Q_{S,F'}$. To show the desired inequality, we compare with another auxiliary distribution π^+_S on the interior edges $E \setminus R_S$, where we consider S to be *fixed to a single component* (that is, the number of components is counted on the graph with vertex set $V(R_S)$, with all vertices of S considered to belong to a single component). We finish the proof of the lemma by proving the following:

$$\pi_{G,\mathrm{o}}(1_e \mid Q_{S,F'}) = \pi^+_S(1_e \mid \cdot \cup F' \in Q_{S,F'}) \geq \pi^+_S(1_e) \geq \pi^+_{G,v,\ell}(1_e), \tag{5.18}$$

where the condition in the event $Q_{S,F'}$ under π^+_S does not take into account the external wiring of π^+_S.

The first inequality ($\pi^+_S(1_e \mid \cdot \cup F' \in Q_{S,F'}) \geq \pi^+_S(1_e)$) follows from the fact that the events $Q_{S,F'}$ and 1_e are increasing on the interior edges, and by monotonicity of the random-cluster model (see [47, Theorem 2.27] with combination in Observation 4.19), conditioning on an increasing event cannot decrease the probability of another increasing event.

Next, we prove the equality in (5.18), using Claim 3 and writing out the probabilities of each configuration and showing that they are equal. We will use some notation. For $\tilde{F} \subseteq E \setminus R_S$ such that $F' \cup \tilde{F} \in Q_{S,F'}$, let $c'(F')$ be the number of components of $(V \setminus V(\mathcal{C}_v(S)), F')$ that do not contain a vertex of S, and let $\tilde{c}(\tilde{F} \cup F')$ be the number of components of $(V, F' \cup \tilde{F})$ containing a vertex of $S \cup V(\mathcal{C}_v)$. Recall that $c(\tilde{F} \cup F')$ denotes the number of components of $(V, \tilde{F} \cup F')$.

By Claim 3, any component of (V, F') that does not contain vertices of S is preserved as a component in $(V, \tilde{F} \cup F')$. The single component containing vertices of S and the components contained in $V(\mathcal{C}_v(S))$ are accounted for in $\tilde{c}(\tilde{F} \cup F')$. Thus,

$c(F' \cup \tilde{F}) = \tilde{c}(F' \cup \tilde{F}) + c'(F')$. Since the probabilities are proportional to the weights, we obtain that

$$\pi_G(\tilde{F} \cup F' \mid E_S, \cdot \cap R_S = F') = \frac{q^{\tilde{c}(F' \cup \tilde{F}) + c'(F')}(\mathrm{e}^\beta - 1)^{|F'| + |\tilde{F}|}}{\sum_{\substack{\tilde{F}_1 \subseteq E \setminus R_S \\ \tilde{F}_1 \cup F' \in E_S}} q^{\tilde{c}(F' \cup \tilde{F}_1) + c'(F')}(\mathrm{e}^\beta - 1)^{|F'| + |\tilde{F}_1|}}$$
$$= \frac{q^{\tilde{c}(F' \cup \tilde{F})}(\mathrm{e}^\beta - 1)^{|\tilde{F}|}}{\sum_{\substack{\tilde{F}_1 \subseteq E \setminus R_S \\ \tilde{F}_1 \cup F' \in E_S}} q^{\tilde{c}(F' \cup \tilde{F}_1)}(\mathrm{e}^\beta - 1)^{|\tilde{F}_1|}}.$$

Since for any such $\tilde{F}$, all vertices of S are in the same component of $(V, F' \cup \tilde{F})$ the numerator is the same as the weight of $\tilde{F}$ under π_S^+, i.e. when all vertices of S are treated as belonging to a single component. Since the support of $\pi_S^+(\cdot \mid \cdot \cup F' \in Q_{S,F'})$ is exactly $\{\tilde{F}_1 \subseteq E \setminus R_S : \tilde{F}_1 \cup F' \in E_S\}$, the probability of any $\tilde{F} \subseteq E \setminus R_S$ is the same under $\pi_G(\cdot \mid E_S, F' = \cdot \cap R_S)$, and, in particular, so is the probability of event 1_e.

It remains to show the final inequality in (5.18), for which we use the following iterative revealing scheme for $(\boldsymbol{F}_S, \boldsymbol{F}^+) \sim (\pi_S^+, \pi_{G,v,\ell}^+)$. First, reveal the edges in R_S for $\pi_{G,v,\ell}^+$. The distribution on the remaining interior edges only depends on the boundary (as we showed in the previous chapter 4.1.5). Thus in particular, the distribution on the remaining edges of $\boldsymbol{F}_S$ is the same as if all edges of R_S were fixed to be in, with additional "artificial edges" if necessary, to make S connected. Therefore, regardless of the value of $\boldsymbol{F}^+$ on R_S, by monotonicity (see Section 4.1.6), we can couple the two random configurations so that $\boldsymbol{F}_S \supseteq \boldsymbol{F}^+ \setminus R_S$. Since $e \notin R_S$, this gives that $e \in \boldsymbol{F}^+ \implies e \in \boldsymbol{F}_S$. We conclude $\pi_S^+(1_e) \geq \pi_{G,v,\ell}^+(1_e)$, which finishes the proof of the lemma.

□

We will next show that the wired boundary is implied by an event that no length-ℓ path from v avoids the largest component (breaking ties arbitrarily) of $(V, F \setminus F_v)$. In the next section we show that with all but exponentially small probability, for $F \sim \pi_{G,\mathrm{o}}$, there will be a unique component of $(V, F \setminus F_v)$ with a linear number of vertices, thus the notion of a largest component is unambiguous.

To make the notion precise, let $\mathcal{C}$ be the largest component of $(V, F \setminus F_v)$. We say that $\mathcal{C}$ *avoids a path* P if there is no vertex $u \in \mathcal{C}$ that also belongs to P. For a vertex v and an integer $\ell \geq 1$, we define the following event:

$$\mathcal{A}_{v,\ell} = \left\{ F \subseteq E(G) \;\middle|\; \begin{array}{l} \text{there exists a path of length } \ell \text{ from } v \text{ in } G \text{ which} \\ \text{avoids the largest component of } (V \setminus \{v\}, F \setminus F_v) \end{array} \right\}. \quad (5.19)$$

The complement of $\mathcal{A}_{v,\ell}$ implies the existence of a wired boundary, as proved in the following lemma.

Lemma 5.19. *Let $G = (V, E)$ be a graph, let $v \in V$ be an arbitrary vertex, and let $\ell \geq 1$ be an integer. For any $F \subseteq E$, let F_v be the set of edges in F that are incident to v and let $G_{v,F} \coloneqq (V \setminus \{v\}, F \setminus F_v)$. Suppose that all simple length-ℓ paths from v in G intersect the largest component of $G_{v,F}$, i.e. that $\overline{\mathcal{A}_{v,\ell}}$ holds. Then F has a wired boundary around v in $B_\ell(v)$.*

Proof. Consider $\mathcal{P}$, the set of all (not necessarily simple) paths in $B_\ell(v)$ that start in v and end in a vertex of $S_\ell(v)$. Notice that every path in $\mathcal{P}$ contains a simple path from v to $S_\ell(v)$, whose length is at least ℓ, and hence contains a simple length-ℓ path from v.

Therefore, by assumption, every path in $\mathcal{P}$ intersects $\mathcal{C}_1$, the largest component of $G_{v,F}$. For each path in $\mathcal{P}$, consider the *first* vertex on the path belonging to $\mathcal{C}_1$ (we think of the path as starting in v). Let $\mathcal{S}$ be the set of these vertices. We next show that $\mathcal{S}$ is a wired boundary around v in $B_\ell(v)$.

We can further assume $\mathcal{S}$ is not empty: if $\mathcal{S}$ were empty, then $\mathcal{P} = \emptyset$, which only happens if v is not connected to $G \setminus B_\ell(v)$, and thus an empty set is a trivial wired boundary.

Claim 1. *$\mathcal{S} \subseteq V(B_\ell(v)) \setminus \{v\}$ is a cut-set of G separating v from $V \setminus B_\ell(v)$.*

The paths in $\mathcal{P}$ are contained in the ball $B_\ell(v)$. Also, $v \notin V(G_{v,F})$, hence also $v \notin \mathcal{S}$. Consider any path from v to $V \setminus V(B_\ell(v))$ in G, say $v = w_0, w_1, \ldots, w_r, \ldots, w_{\ell'} \notin B_\ell(v)$. Clearly, some prefix of this path is in $\mathcal{P}$, and thus there is $1 \leq i \leq \ell'$ such that $w_i \in \mathcal{S}$. Therefore, any path from v to the outside of the ball is disconnected in $G \setminus \mathcal{S}$, proving Claim 1.

Next, let $\mathcal{C}_v$ be the component of v in $G \setminus \mathcal{S}$. To conclude that $\mathcal{S}$ is a wired boundary, we need to show that the vertices of $\mathcal{S}$ belong to a single component in $(V \setminus V(\mathcal{C}_v), \{e \in F : e \cap V(\mathcal{C}_v) = \emptyset\})$. In particular, we show that this component is $\mathcal{C}_1$.

Claim 2. *No vertices of $\mathcal{C}_1$ belong to $\mathcal{C}_v$.*

Suppose for contradiction $u \in V(\mathcal{C}_v) \cap V(\mathcal{C}_1)$. By Claim 1, $\mathcal{C}_v$ is a subgraph of $B_\ell(v)$, and thus there exists a path $v = w_0, w_1, \ldots, w_{\ell'} = u$ of length $\ell' \leq \ell$ from v to

u in $\mathcal{C}_v$. Since we assumed $\mathcal{S}$ is not empty and $\mathcal{S} \subseteq V(\mathcal{C}_1)$, we can extend this path to a (not necessarily simple) path $w_0, w_1, \dots, w_{\ell'} = u, w_{\ell'+1}, \dots, w_\ell$ in the ball, which ends in a vertex of $S_\ell(v)$ and contains u. But this is a path in $\mathcal{P}$. By the choice of $\mathcal{S}$ and u, there is $1 \le i \le \ell'$ such that $w_i \in \mathcal{S}$. But then $w_i \notin \mathcal{C}_v$, which contradicts that $w_0, \dots, w_{\ell'}$ is in $\mathcal{C}_v$. Thus, no such u can exist, and Claim 2 follows.

Claim 2 implies that $\mathcal{C}_1$ is connected in $(V \setminus V(\mathcal{C}_v), \{e \in F : e \cap V(\mathcal{C}_v) = \emptyset\})$, thus in particular all vertices of $\mathcal{S}$ belong to it in this graph as well, so these vertices are all in the same component.

Thus, it follows that $\mathcal{S}$ is a wired boundary around v in $B_\ell(v)$. □

5.4.2 The giant component in the ordered phase

We will bound the probability of $\mathcal{A}_{v,\ell}$ by considering percolation on the planted random graph model $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}})$ for $\mu_{G,\mathrm{o}}$. Since we work with a graph where a is removed, with the aim of bounding the probability of a path belonging to the resulting largest component, it will thus be relevant to consider random graph percolation after removing a path of length ℓ, which results in an "almost d-regular random graph".

For a graph G, let $\mathcal{C}_1(G)$ denote the largest component of G (breaking ties arbitrarily), and more generally, $\mathcal{C}_i(G)$ denote the i-th largest component of G. It is a standard fact (see e.g. [25, Lemma 5.3]) that the size of the largest component in a percolated random d-regular graph is controlled by the extinction probability φ of a Galton-Watson process with offspring distribution $\mathrm{Bin}(d-1, p)$.[9] For $p \in (\frac{1}{d-1}, 1)$, define $\varphi = \varphi(p), \chi = \chi(p)$ and $\hat{\varphi} = \hat{\varphi}(p)$ to be the unique solutions on the interval $(0, 1)$ to the following

$$\varphi = (1 - \varphi + p\varphi)^{d-1}, \ \chi \coloneqq 1 - (p\varphi + 1 - p)^d, \ \hat{\varphi} \coloneqq (1 - \varphi + p\varphi)^{d-2}. \tag{5.20}$$

In particular, the size of the largest component is roughly equal to $\chi n + o(n)$; see e.g. [25, Proposition 5.4] (restated below) and also [3, 57] for more details. We also remark that $\hat{\varphi}$ is the extinction probability of a modified Galton-Watson process in

[9] A Galton-Watson process starts with a single individual (the root). Each individual independently produces a random number of offspring according to the same distribution (in this case the binomial distribution with parameters $d-1$ and p), and the process is repeated independently for each new individual. The process becomes extinct if some generation has no individuals. We say that the process is *subcritical* if the expected number of offspring is less than 1; then the extinction probability is 1. We say that the process is *supercritical* if the expected number of offspring is greater than 1; in this case its extinction probability is strictly less than 1. We call the resulting tree the Galton-Watson (GW) tree.

which the root has offspring distribution $\mathrm{Bin}(d-2,p)$, while all subsequent generations have offspring distribution $\mathrm{Bin}(d-1,p)$; this will be relevant later for our path considerations.

We show a similar result for the size of the giant component for slightly non-regular degree sequences $\boldsymbol{d}$, where we allow a small set S of vertices to have degree less than d; in our WSM setting, we will further be interested in the probability that a particular subset of the vertices S' does not belong to the giant. Since the number of edges in the percolated graph is strongly concentrated around its expectation, it will be convenient to consider a variant of percolation where the number of edges in the resulting graph is fixed to be m, rather than obtained by including each edge independently. For percolation on random graphs with given degree sequence $\boldsymbol{d} = (d_1, \dots, d_n)$, this is captured by the *exact edge model*. For a degree sequence $\boldsymbol{d}$ and an edge count $m \leq \frac{1}{2}\sum d_i$, let $\tilde{\boldsymbol{G}}_{n,\mathbf{d},m}$ be the random graph model obtained by the following procedure: first, each vertex $i \in [n]$ is associated with d_i half-edges, then we choose a uniformly random subset of $2m$ half-edges and then choose a uniformly random perfect matching of the $2m$ selected half-edges. Each matched pair of half-edges then forms an edge between the corresponding vertices, giving the final (multi)graph.

5.4.2.1 Percolation in regular and near-regular graphs

We first consider the case where $\boldsymbol{d} = (d_1, \dots, d_n)$ is the degree sequence of a d-regular graph (i.e., $d_i = d$ for all i). For simplicity, in this uniform setting, we denote the percolation model with $m = pdn/2 + o(n)$ edges by $\tilde{\boldsymbol{G}}_p$. The following bounds were shown in [25].

Lemma 5.20 ([25, Proposition 5.4])**.** *Let $d \geq 3$ be an integer, $p \in (\frac{1}{d-1}, 1)$, and $m = pdn/2 + o(n)$. Then, for any $\delta > 0$, there exists a constant $M > 0$, such that with probability $1 - \mathrm{e}^{-\Omega(n)}$ over $\tilde{\boldsymbol{G}}_p$, it holds that $\mathcal{C}_1(\tilde{\boldsymbol{G}}_p) = (\chi \pm \delta)n$, where χ is as in (5.20). Moreover, $\sum_{i\geq 2} |\mathcal{C}_i(\tilde{\boldsymbol{G}}_p)|^2 \leq Mn$.*

We use Lemma 5.20 to infer bounds for the component sizes for slightly non-regular degree sequences $\boldsymbol{d}$.

Lemma 5.21. *Let $d \geq 3$ be an integer, $p \in (1/(d-1), 1)$, and $\delta, \varepsilon > 0$ reals. Suppose $S \subseteq [n]$ satisfies $|S| = O(n^{1/6})$ and $\boldsymbol{d} = (d_1, \dots, d_n)$ is a degree sequence such that $d_i = d$ for $i \notin S$ and $d_i \in [0, d]$ for $i \in S$. Then, there exists a constant $M(d, p, \delta) > 0$, such that for $m = \frac{1}{2}dpn + o(n)$, it holds that*

$$\mathbb{P}\big(|\mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m})| \geq (\chi - \delta)n \text{ and } |\mathcal{C}_2(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m})| \leq Mn^{2/3}\big) \geq 1 - \mathrm{e}^{-\Omega(n)}.$$

Proof. Observe that, conditional on the number of percolated edges incident to vertices of S being some integer k, the graph $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m} \setminus S$ is distributed as an exact d-regular edge model with $n-|S|$ vertices and $m-k$ edges, as every submatching of size $m-k$ is equally likely by symmetry. Since $|S| = O(n^{1/6})$, we trivially have that $k = O(n^{1/6})$ as well. Thus, for any such k, it holds that $\frac{m-k}{d(n-|S|)/2} = p + o(1) - O(n^{-5/6}) > \frac{1}{d-1}$ for sufficiently large n. Furthermore, since $\chi(p)$ is continuous on $(\frac{1}{d-1}, 1)$, it also holds that $\chi' \coloneqq \chi(\frac{m-k}{d(n-|S|)/2}) = \chi(p) \pm \delta/3$. Therefore, applying Lemma 5.20 to $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m} \setminus S$, we get that with probability $1 - \mathrm{e}^{-\Omega(n)}$, its largest component has size at least $(\chi' - \delta/3)(n - |S|) = (\chi - \delta)n$ for n large enough, and hence so does $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$.

Moreover, by the second part of Lemma 5.20, the second largest component of $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m} \setminus S$ has size $O(n^{1/2})$. Since $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ has at most $O(n^{1/6})$ more edges than $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m} \setminus S$, the non-giant components can get connected to form a component with size at most $O(n^{1/2}n^{1/6}) = O(n^{2/3})$. □

With these bounds on the size of the giant component in hand, we next turn to bounding the probability that a given set of vertices avoids the giant.

5.4.2.2 Branching-process estimates

We will next use the following results about branching processes. Recall that a Galton-Watson process generates a tree, where the numbers of children (the offspring distribution) of the vertices are independent. Unless stated otherwise, the offspring distribution is also identical for each vertex in the tree. The *surviving population at generation r* is the number of vertices at depth r. We say that the process goes *extinct* if the resulting tree is finite. Recall that the process is supercritical if the expected number of offspring is strictly more than 1. In this case, it is a standard fact (see, e.g. [60, Proposition 5.4.]) that the extinction probability of a supercritical branching process is strictly below 1. The next lemma says that, conditional on survival, the size of the surviving population at generation r of a supercritical branching process grows with r.

Lemma 5.22. *Consider a supercritical Galton-Watson process and let Z_r be the size of the surviving population at generation r. Then, for any fixed integer $K > 0$, it holds that $\mathbb{P}(Z_r \geq K \mid Z_r > 0) = 1 - o_r(1)$.*

Proof. Let $\phi = \mathbb{P}(\text{extinction})$ and q_0 be the probability of no offspring in one step; since the process is supercritical, we have $\phi, q_0 < 1$. Suppose for the sake of contradiction that there is $\varepsilon > 0$ such that $\mathbb{P}(Z_r < K \mid Z_r > 0) \geq \varepsilon$ for all $r \geq 0$. Then, we

have $\mathbb{P}(\text{extinction} \mid Z_r > 0) \geq \delta$ where $\delta = \varepsilon q_0^K > 0$. Hence, for every r we have

$$\mathbb{P}(\text{extinction}) \geq \delta(1 - \mathbb{P}(Z_r = 0)) + \mathbb{P}(Z_r = 0).$$

As r grows large, we have that $\mathbb{P}(Z_r = 0) \to \phi$, so by taking limits in the above inequality we obtain that $\phi \geq \delta(1 - \phi) + \phi$, contradicting that $\phi < 1$. □

To account for vertices with irregular degrees, we next consider a slightly modified Galton-Watson process, where the root has a different offspring distribution (the offspring distribution for all subsequent generations is identical).

Corollary 5.23. *Let $0 < d_0 \leq d - 1$ be integers, and $p \in (\frac{1}{d-1})$ be a real. Then consider a Galton-Watson process where the root has offspring distribution $\mathrm{Bin}(d_0, p)$, and all the following generations have offspring distribution $\mathrm{Bin}(d-1, p)$. Let Z_r be the size of the surviving population at generation r. Then, for any fixed integer $K > 0$, it holds that $\mathbb{P}(Z_r \geq K \mid Z_r > 0) = 1 - o_r(1)$.*

Proof. Note that once at the second or later generation, the process is supercritical, so then we can apply Lemma 5.22. Thus we need to consider the case of the root.

Note that if $Z_r > 0$, then there is a non-empty set $\mathcal{C}$ of children of the root whose branching process survives $r - 1$ generations. Since their individual branching processes – denote them $(Z_t^i)_{i \in \mathcal{C}, t \geq 0}$, are independent, $\mathbb{P}(Z_r < K \mid Z_r > 0, \mathcal{C} \text{ survive}) \leq \prod_{i \in \mathcal{C}} \mathbb{P}(Z_{r-1}^i < K \mid Z_{r-1}^i > 0) = o_r(1)$ by the previous lemma. Summing over the possible (non-empty) sets of surviving children, we obtain that the bound holds unconditionally. □

5.4.2.3 Size of the giant component in percolated almost-regular graphs

Having all necessary results, we can now prove the main result of this section. In the following lemma, we consider two types of percolation: the *subcritical case* with $p < \frac{1}{d-1}$, which models the components of in-edges arising (after the percolation step) from subgraphs with non-majority colours, and the *supercritical case* with $p > \frac{1}{d-1}$, which models the giant component in the ordered planted graph. In the first case, we show that with sufficient probability the components are all small, and in the second case we upper-bound the probability that a small vertex set avoids the giant component; we later use this event to establish WSM within the ordered phase.

Lemma 5.24. *Let $d \geq 3$ be an integer and $p \in (0, 1)$. Let $\delta, \varepsilon > 0$ be arbitrarily small constants.*

Suppose that $S \subseteq [n]$ *satisfies* $|S| = O(n^{1/6})$ *and* $\mathbf{d} = (d_1, \dots, d_n)$ *is a degree sequence such that* $d_i = d$ *for* $i \notin S$ *and* $d_i \in [0, d]$ *for* $i \in S$*. Then, there exists a constant* $M(d, p, \delta) > 0$*, such that for* $\tilde{\boldsymbol{G}}_{n,\mathbf{d},m}$ *with* $m = \frac{1}{2}dpn + o(n)$*, the following hold:*

1. *If* $p < \frac{1}{d-1}$*, then* $\mathbb{P}\big[\mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \geq Mn^{1/2}\big] \leq \mathrm{e}^{-\Omega(n)}$*.*
2. *If* $p > \frac{1}{d-1}$*, then* $\mathbb{P}\big[\mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \leq (\chi - \delta)n\big] \leq \mathrm{e}^{-\Omega(n)}$ *and* $\mathbb{P}\big[\mathcal{C}_2(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \geq Mn^{2/3}\big] \leq \mathrm{e}^{-\Omega(n)}$*. Moreover, with* $\varphi = \varphi(p)$ *as in* (5.20)*, for any subset* $S' \subseteq S$*,*

$$\mathbb{P}\big[S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset\big] \leq (1+\varepsilon)^{|S|} \prod_{i \in S'} (p\varphi + 1 - p)^{d_i}.$$

The first item is a straightforward corollary of a result for regular degree sequences. The bound for the second item is more involved; intuitively, the high-level strategy is as follows: we start by revealing the R-neighbourhoods of vertices in S' for a suitably large constant R. We show that with sufficient probability most of these neighbourhoods are acyclic, disjoint from all other revealed neighbourhoods, and contain only degree-d vertices apart from their roots (Claim 1 in the proof), and thus can be modelled by Galton-Watson trees. Since we only want to obtain a lower bound on the probability of belonging to the giant component, it is enough to consider vertices of S' with these "nice" neighbourhoods. The idea is then to obtain this bound by introducing an auxiliary (coupled) d-regular percolated graph $\boldsymbol{H}$ together with a set of "special vertices" of $\boldsymbol{H}$. Every "special vertex" corresponds to a vertex of S' with a "nice" neighbourhood, with the property that whenever it belongs to the giant component of $\boldsymbol{H}$, the corresponding vertex in S' belongs to the giant component of $\tilde{\boldsymbol{G}}_{n,\mathbf{d},m}$ (Claim 2 in the proof). Crucially, since $\boldsymbol{H}$ is a regular percolated graph, the bound on the probability of this event is easier to obtain. The final probability is then obtained by a technical calculation; note that for a vertex $i \in S'$, the factor $(p\varphi + 1 - p)^{d_i}$ roughly captures the probability that the percolation dies out within the radius-R neighbourhood of the vertex i.

Proof. For Item 1 (the subcritical case), the statement is proved in [25, Proposition 5.2] for $S = \emptyset$. Note that while the statement of the proposition is for d-regular degree sequences, as it is noted therein, it also applies to percolation on a graph of maximum degree d (with any degree sequence).

We thus focus on Item 2. The first part of the claim follows from Lemma 5.21. It only remains to bound the probability that none of the vertices in S' belongs to

the largest component. Let $R, K > 0$ be sufficiently large constants and $\varepsilon_0, \varepsilon_1 > 0$ be sufficiently small, to be specified later.

First, consider revealing the depth-R neighbourhoods around vertices of S' in the underlying configuration model, and then the outcome of the percolation on these edges. We say that the revealed R-neighbourhood of $v \in S'$ is *nice* if it is acyclic, disjoint from all the other revealed R-neighbourhoods, and all vertices in it other than v have degree d. The following claim says that there are at least $(1-\varepsilon_0)|S'|$ nice neighbourhoods.

Claim 1: *Let $\varepsilon_0 \in (0,1)$ be a constant. Then there exists a constant $C > 0$ (not depending on n) such that with probability $1 - n^{-C|S'|}$, there are at least $(1-\varepsilon_0)|S'|$ vertices $v \in S'$ with nice neighbourhoods.*

The proof of the claim is standard and similar to the proof of [14, Lemma 15], so we sketch the argument briefly. Consider a BFS revealing process from all the vertices in S' up to depth R. Say that a *bad event* occurs when we reveal a match to a half-edge incident to a vertex that has either already been seen or belongs to S (only vertices in S can have degree $\neq d$). Note that the total size of the revealed neighbourhoods is at most $|S'|d^R$ (at least for R large enough), and thus in each revealing step, regardless of previous outcomes, the probability of a bad event occurring is bounded by $\frac{|S'|d^R+|S|}{n-|S'|d^R-|S|} \leq \frac{2|S|d^R}{n}$ (since $S = O(n^{1/6})$, for large enough n the denominator is at least $n/2$). Hence we can upper-bound the number of bad events by a sum of $|S'|d^R$ i.i.d. Bernoulli random variables with mean $\frac{2|S|d^R}{n}$. Then, by Chernoff bounds (here use the statement as in [14, Lemma 16]),

$$\mathbb{P}(\#\text{bad events} \geq \varepsilon_0|S'|/2) \leq \exp\big\{-\frac{\varepsilon_0|S'|}{2}\log\big(\frac{n}{2e|S|d^R}\big)\big\}.$$

Since $2e|S|d^R = O(n^{1/6})$, the probability of having at least $\varepsilon_0|S'|/2$ bad events is at most $(C'n^{-5\varepsilon_0/12})^{|S'|}$ for some constant $C' > 0$ (independent of n). Note that each bad event makes at most two neighbourhoods not nice, and hence fewer than $\varepsilon_0|S'|/2$ bad events implies that at least $(1-\varepsilon_0)|S'|$ neighbourhoods are nice. Thus, there exists some constant $C = C(\varepsilon_0, d, R, C') > 0$, such that with probability at least $1 - n^{-C|S'|}$, there are at least $(1-\varepsilon_0)|S'|$ nice neighbourhoods, which concludes the claim.

Having proved the claim, let T be the subset of S' with nice neighbourhoods. Further on, we will consider the probability of $S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset$ conditional on $|T| \geq (1-\varepsilon_0)|S'|$.

As a next step in the analysis, we find an auxiliary "hidden"[10] d-regular percolated graph $\boldsymbol{H}$ in the remaining unrevealed graph, with the property that whenever a vertex of $\boldsymbol{H}$ *corresponding to a vertex of* $T \subseteq S'$ belongs to $\mathcal{C}_1(\boldsymbol{H})$, then also $S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \neq \emptyset$. Thus we can upper-bound the probability of $S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset$ by the probability that none of the vertices corresponding to $T \subseteq S'$ in $\boldsymbol{H}$ belongs to $\mathcal{C}_1(\boldsymbol{H})$.

For each $v \in T$, consider the set of corresponding *surviving leaves*, that is, the set of vertices distance-R from v (in the underlying percolation model) that are connected to v after percolation. Let L_v denote the number of surviving leaves corresponding to v.

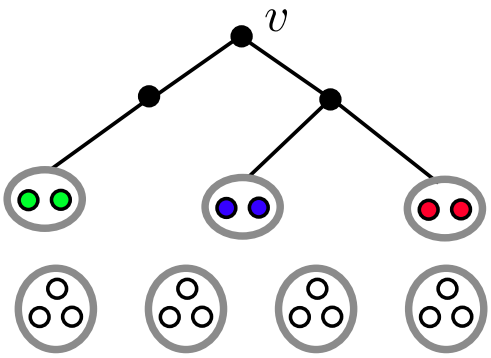
Each of the surviving leaves has $d-1$ half-edges with an unrevealed match (call these *unexposed half-edges*). Notice that all of these L_v leaves are connected in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$, and are in particular connected to v. Hence, whenever one of these leaves is connected to the giant component of $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$, so are all the others. Therefore, it will be convenient to view all the unexposed half-edges belonging to the leaves as half-edges belonging to a single vertex. More specifically, since the aim is to obtain an auxiliary d-regular graph $\boldsymbol{H}$, we "regroup" the unexposed half-edges of surviving leaves corresponding to v into groups of d half-edges, thus obtaining $\lfloor \frac{L_v(d-1)}{d} \rfloor$ new vertices of degree d (we can "discard" the excess half-edges). These new vertices will be the vertices in $\boldsymbol{H}$ corresponding to v. Crucially, whenever any of these new vertices corresponding to v belong to the giant component of $\boldsymbol{H}$, then v must belong to the giant component of $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$.

To make the construction formal, we introduce three types of vertices and half-edges of $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ and $\boldsymbol{H}$. See Figure 5.1 for illustration.

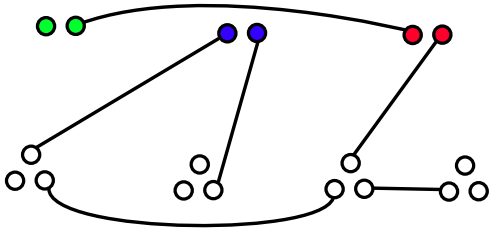
- Vertices (and their corresponding half-edges) in $V \setminus S$ that were not revealed by the R-neighbourhood exploration process. Call these V_0.
- The newly formed vertices corresponding to the "regrouped" half-edges: for each $v \in T$, divide the half-edges corresponding to its surviving leaves into $\lfloor \frac{L_v(d-1)}{d} \rfloor$ new vertices – denote these as V_v.
- All the remaining half-edges and vertices. In particular, this third group consists of unexposed half-edges of S and half-edges of vertices in the R-neighbourhoods that were not yet revealed, which do not fall into the second category of regrouped vertices.

[10] Recall that the exact edge model can be viewed as being generated from a uniformly random matching of the selected half-edges. We construct $\boldsymbol{H}$ using the same random matching, but regroup the half-edges into a different set of vertices; in this sense, $\boldsymbol{H}$ is coupled with $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$, more specifically, with the not-yet-revealed part of the graph.

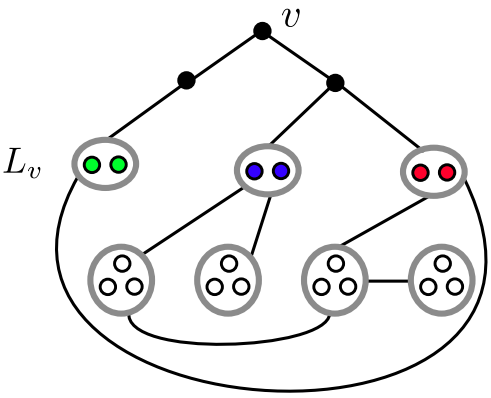
The vertices of $\boldsymbol{H}$ will be $V_0 \cup \bigcup_{v \in T} V_v$, and all of them have degree d. The edges of $\boldsymbol{H}$ are all percolated edges of $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ that are between two half-edges that also belong to a vertex of $\boldsymbol{H}$.



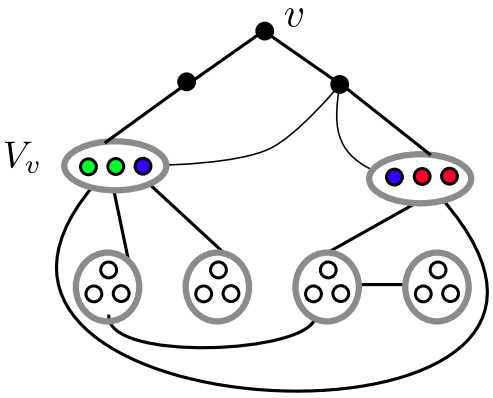
(a) The R-neighbourhood of v.



(b) The matching of the unrevealed half-edges.



(c) The resulting edges in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$.



(d) The same matching after regrouping the half-edges to form $\boldsymbol{H}$.

Figure 5.1: An illustration of a construction of $\tilde{\mathbf{G}}_{n,\mathbf{d},m}$ and the auxiliary graph $\mathbf{H}$ with $S' = \{v\}$. We first reveal the neighbourhood of v (5.1a). Note that some of the half-edges were revealed not to belong to the selected matching (which only contains $2m$ half-edges). The coloured half-edges belong to three surviving leaves in the revealed neighbourhood of v. We use the same underlying matching (5.1b) of the unrevealed half-edges in $\tilde{\mathbf{G}}_{n,\mathbf{d},m}$ and $\mathbf{H}$, but in $\mathbf{H}$ the half-edges of surviving leaves corresponding to the same v are regrouped into vertices of degree d. Thus, the construction changes the assignment of half-edges to vertices while preserving their matching.

Next, we claim that conditional on the number of vertices in $\bigcup_{v \in T} V_v$ and the number of edges in $\tilde{\boldsymbol{G}}_{n,d,m}$ containing a half-edge in the third category, $\boldsymbol{H}$ is distributed as a d-regular exact edge model with $n' = n - O(n^{1/6})$ vertices and $m' = \frac{1}{2}dpn + o(n)$ edges.

For the number of vertices, it follows from $|V_0| \leq n' = |V_0| + |\bigcup_{v \in T} V_v|$, and then

from $|V_0| \geq n - |S| - d^R|S'| = n - O(n^{1/6})$, and $|\bigcup_{v\in T} V_v| \leq d^R|S'| = O(n^{1/6})$.

For the number of edges, note that the number of half-edges in the third category is at most $|S| + |S'|d^{R+1} = O(n^{1/6}) = o(n)$ and the number of already revealed edges is also bounded by $O(n^{1/6})$. Hence we get $m' = \frac{1}{2}dpn + o(n)$.

It remains to show that $\boldsymbol{H}$ is indeed distributed as a d-regular exact edge model. Conditional on the revealed part of the graph and on the resulting number m' of edges in $\boldsymbol{H}$, every matching of m' edges on the given set of half-edges of $\boldsymbol{H}$ is equally likely. This is precisely the distribution of the exact edge model.

The following claim implies that we can upper-bound the probability of the event $S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset$ by the probability of $(\bigcup_{v\in T} V_v) \cap \mathcal{C}_1(\boldsymbol{H}) = \emptyset$, at least as long as the components of $\boldsymbol{H}, \tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ are sufficiently well-behaved (as in the statement of this lemma). The latter probability is easier to bound given the symmetry of $\boldsymbol{H}$.

Claim 2: *Let $v \in T$, and $w \in V_v$ be such that w is in a component of size s in $\boldsymbol{H}$. Then v is in a component of size $\geq s - |S'|d^R$ in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$.*

This claim follows from the correspondence between the vertices of $\boldsymbol{H}$ and $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$. The half-edges incident to w in $\boldsymbol{H}$ correspond, in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$, to surviving leaves associated with v, which are all connected to v. Thus, all vertices in V_0 belonging to the component of w in $\boldsymbol{H}$ also belong to the component of v in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$. Since $|V(\boldsymbol{H})\setminus V_0| \leq |S'|d^R$, the component of v in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ has size at least $s - |S'|d^R$, proving the claim.

Next we bound the probability that none of the vertices in $\bigcup_{v\in T} V_v$ belong to the largest component of $\boldsymbol{H}$. First we introduce some useful conditioning. Let $\mathcal{C}_{good}$ be the event that $\boldsymbol{H}$ and $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ have components satisfying $\mathcal{C}_1(\boldsymbol{H}) \geq \chi n'/2$, $\mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \geq \chi n/3$, and $\mathcal{C}_2(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) \leq Mn^{2/3}$, where M is the constant from Lemma 5.21 (using $\delta = 2\chi/3$). We can assume that n is large enough so $Mn^{2/3} < \chi n'/2 - d^R|S'| \geq \chi n/3$. Note that this holds for all sufficiently large n as $n' = n - O(n^{1/6})$. By Lemma 5.21, $\mathbb{P}(\mathcal{C}_{good}) \geq 1 - \mathrm{e}^{-\Omega(n)}$.

Conditional on $\mathcal{C}_{good}$, the above claim implies that whenever $w \in V_v$ belongs to $\mathcal{C}_1(\boldsymbol{H})$, then $v \in \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m})$. Next we show that for any $\varepsilon_1 > 0$, conditional on T, the following inequality holds:

$$\mathbb{P}(\forall v \in T:\ V_v \cap \mathcal{C}_1(\boldsymbol{H}) = \emptyset \mid T, \mathcal{C}_{good}) \leq \prod_{i\in T}[(\varphi p + 1 - p)^{d_i} + \varepsilon_1] \tag{5.21}$$

In order to prove (5.21), first note that, by symmetry of the d-regular exact edge model, conditional on the size of the largest component of $\boldsymbol{H}$, the vertices belonging to it form a u.a.r. subset of that size.

Write $T = \{i_1, \dots, i_k\}$, and let $1 \le j \le k$. Suppose we condition on the outcome for R-neighbourhoods of $i_1, \dots, i_{j-1}$, that is, the number of surviving leaves, and which of the vertices in $\bigcup_{\ell<j} V_{i_\ell}$ belong to the largest component of $\boldsymbol{H}$. We do not reveal any further information about $\mathcal{C}_1(\boldsymbol{H})$. Hence, by symmetry, conditional on these outcomes and on $\mathcal{C}_{\text{good}}$, the remaining undetermined vertices of the giant are still uniformly distributed among the vertices whose membership has not yet been determined. There are at least $\chi n'/2 - d^R|S'| \ge \chi n/3$ such vertices in the largest component.

Now consider the number of surviving leaves, L_{i_j} in the R-neighbourhood of i_j. By Corollary 5.23, conditional on $L_{i_j} > 0$, the probability that $L_{i_j} < Kd$ is $o_R(1)$[11]. By taking R large enough, this probability is at most $\varepsilon_1/2$. Since the extinction probability of the corresponding branching process is $(p\varphi + 1 - p)^{d_{i_j}}$, it follows that the probability that $L_{i_j} < Kd$ is $(\varphi p + 1 - p)^{d_{i_j}} + \varepsilon_1/2$. Then, conditional on $L_{i_j} \ge Kd$, it holds that $|V_{i_j}| \ge K(d-1)$. Since at least $\chi n/3$ of the at most n vertices whose membership is still undetermined belong to $\mathcal{C}_1(\boldsymbol{H})$, the probability that none of the vertices in V_{i_j} belongs to $\mathcal{C}_1(\boldsymbol{H})$ is at most $(1 - \chi/3)^{K(d-1)}$, which goes to 0 as we increase K. Thus, for K large enough, this probability is at most $\varepsilon_1/2$.

Putting it all together, the probability that $V_{i_j} \cap \mathcal{C}_1(\boldsymbol{H}) = \emptyset$ is at most $(\varphi p + 1 - p)^{d_{i_j}} + \varepsilon_1$. Writing the joint probability as a product of successive conditional probabilities, and using that the above bound does not depend on the particular outcomes of the previous neighbourhoods, gives (5.21).

It remains only to remove the conditioning and account for the errors introduced above. There are three such errors: the probability that fewer than $(1 - \varepsilon_0)|S'|$ neighbourhoods are nice, the probability that $\mathcal{C}_{good}$ fails, and the fact that the product bound in (5.21) is over T rather than S', with an additional ε_1 in each factor. The first two probabilities are negligible, while by taking $\varepsilon_0, \varepsilon_1$ sufficiently small, the latter two losses can be absorbed into the factor $(1 + \varepsilon)^{|S'|}$. We make this precise below.

We now remove the conditioning, accounting separately for the events that $|T| <$

[11] Strictly speaking, in the exact edge model the edge states are not independent. However, since R is constant and only a sublinear number of edges have been revealed, conditional on the R-neighbourhood being nice, the R-neighbourhood in $\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}$ stochastically dominates a supercritical Galton-Watson tree.

$(1-\varepsilon_0)|S'|$ and that $\mathcal{C}_{good}$ fails:

$$\begin{aligned}
&\mathbb{P}(S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset) \\
&\quad \leq \mathbb{P}(S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset \mid |T| \geq (1-\varepsilon_0)|S'|) + \mathbb{P}(|T| < (1-\varepsilon_0)|S'|) \\
&\quad\quad \leq \mathbb{P}(S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset \mid |T| \geq (1-\varepsilon_0)|S'|, \mathcal{C}_{good}) \\
&\quad\quad\quad + \mathbb{P}(\overline{\mathcal{C}_{good}} \mid |T| \geq (1-\varepsilon_0)|S'|) + \mathbb{P}(|T| < (1-\varepsilon_0)|S'|). \quad (5.22)
\end{aligned}$$

We showed $\mathbb{P}(|T| < (1-\varepsilon_0)|S'|) \leq n^{-C|S'|}$, and $\mathbb{P}(\overline{\mathcal{C}_{good}} \mid |T| \geq (1-\varepsilon_0)|S'|) \leq \mathrm{e}^{-\Omega(n)}$. The inequality in (5.21) gives

$$\mathbb{P}(S' \cap \mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m}) = \emptyset \mid |T| \geq (1-\varepsilon_0)|S'|, T, \mathcal{C}_{good}) \leq \prod_{i \in T}[(\varphi p + 1 - p)^{d_i} + \varepsilon_1]. \quad (5.23)$$

The next step is to obtain the following bound on (5.23) which holds for $\varepsilon_0, \varepsilon_1$ sufficiently small and for any $T \subseteq S'$ with $|T| \geq (1-\varepsilon_0)|S'|$, and does not depend on the particular choice of T (which will allow us to drop the conditioning and thus bound (5.22)).

$$\prod_{i \in T}[(\varphi p + 1 - p)^{d_i} + \varepsilon_1] \leq (1 + \varepsilon/2)^{|S'|} \prod_{i \in S'}(\varphi p + 1 - p)^{d_i} \quad (5.24)$$

As a first step towards (5.24), note that for each $i \in S'$, it is the case that $d_i \leq d$. Since it also holds that $(\varphi p + 1 - p) < 1$, it follows that $(\varphi p + 1 - p)^{d_i} + \varepsilon_1 \leq (\varphi p + 1 - p)^{d_i}(1 + \varepsilon_1/(\varphi p + 1 - p)^d)$. Also, we can bound, for all $T \subseteq S'$ with $|T| \geq (1-\varepsilon_0)|S'|$,

$$\frac{\prod_{i \in T}(\varphi p + 1 - p)^{d_i}}{\prod_{i \in S'}(\varphi p + 1 - p)^{d_i}} = \prod_{i \in S' \setminus T}(\varphi p + 1 - p)^{-d_i} \leq (\varphi p + 1 - p)^{-d|S' \setminus T|} \leq (\varphi p + 1 - p)^{-d\varepsilon_0 |S'|}.$$

Combining, we therefore get that

$$\prod_{i \in T}[(\varphi p + 1 - p)^{d_i} + \varepsilon_1] \leq [(1 + \tfrac{\varepsilon_1}{(\varphi p + 1 - p)^d})(\varphi p + 1 - p)^{-d\varepsilon_0}]^{|S'|} \prod_{i \in S'}(\varphi p + 1 - p)^{d_i}.$$

The next step towards (5.24) is to show that

$$(1 + \tfrac{\varepsilon_1}{(\varphi p + 1 - p)^d})(\varphi p + 1 - p)^{-d\varepsilon_0} \leq 1 + \varepsilon/2.$$

Note that by taking $\varepsilon_0, \varepsilon_1 \to 0$, we get $(1 + \frac{\varepsilon_1}{(\varphi p + 1 - p)^d})(\varphi p + 1 - p)^{-d\varepsilon_0} \to 1$, so in particular we can have $\varepsilon_0, \varepsilon_1 > 0$ small enough so that the expression becomes at most $1 + \varepsilon/2$, yielding (5.24).

Since the degrees of vertices in S' are all at most d and $\varphi p+1-p<1$, we get that $\prod_{i\in S'}(\varphi p+1-p)^{d_i}\geq(\varphi p+1-p)^{d|S'|}$. Finally, since $|S'|=O(n^{1/6})$, for all sufficiently large n we have $\mathrm{e}^{-\Omega(n)}\leq n^{-C|S'|}$. Moreover, since $n^{-C}\to 0$, for all sufficiently large n,

$$n^{-C|S'|}+\mathrm{e}^{-\Omega(n)}\leq\Big[\frac{\varepsilon}{2}(\varphi p+1-p)^{d}\Big]^{|S'|}\leq\Big(\frac{\varepsilon}{2}\Big)^{|S'|}\prod_{i\in S'}(\varphi p+1-p)^{d_i}.$$

Therefore, by combining the terms in 5.22, we get that $\mathbb{P}(S'\cap\mathcal{C}_1(\tilde{\boldsymbol{G}}_{n,\boldsymbol{d},m})=\emptyset)$ is at most

$$\Big(\frac{\varepsilon}{2}\Big)^{|S'|}\prod_{i\in S'}(\varphi p+1-p)^{d_i}+(1+\varepsilon/2)^{|S'|}\prod_{i\in S'}(\varphi p+1-p)^{d_i}\leq(1+\varepsilon)^{|S'|}\prod_{i\in S'}(\varphi p+1-p)^{d_i},$$

which completes the proof. □

Next, we use this result in the proof of WSM within the ordered phase.

5.4.3 Avoiding the largest component

Next, we use the previous section to bound the probability of the event $\mathcal{A}_{v,\ell}$, that is, the probability that some length-ℓ simple path from v avoids the largest component of in-edges after v is removed.

To relate the planted model to the results for percolated random graphs, we consider the density of the edges inside the colour classes. The following lemma says that in the ordered regime, one colour class is supercritical, while the others are subcritical. That is, the dominant colour will give rise to a component with linearly many edges, whereas the components of in-edges connecting vertices with non-majority colours will all have size $O(n^{1/2})$. Recall that ν_{o} and ρ_{o} are the ordered vertex and edge statistics for the ordered phase respectively, defined in (5.2) and (5.6).

Lemma 5.25 ([25, Lemma 5.6])**.** *Let $q,d\geq 3$ be integers. For $\beta>\beta_u$ and $p=1-\mathrm{e}^{-\beta}$, the ordered colour statistics ν_{o} and ρ_{o} satisfy $p\frac{\rho_{\mathrm{o},ii}}{\nu_{\mathrm{o},i}}>\frac{1}{d-1}$ for $i=1$, and $p\frac{\rho_{\mathrm{o},ii}}{\nu_{\mathrm{o},i}}<\frac{1}{d-1}$ for $i\in\{2,\dots,q\}$.*

In particular, we can apply Lemma 5.24 to study the largest component of the percolated planted graph $\boldsymbol{F}(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}),\hat{\boldsymbol{\sigma}}_{\mathrm{o}})$ after removing v. To see this, we relate the resulting graph within each colour class to the exact-edge model of Lemma 5.24. Fix the colouring $\hat{\boldsymbol{\sigma}}_{\mathrm{o}}$ and consider the subgraph induced by the vertices of the majority colour. By Lemma 5.3, we can assume (incurring only an exponentially small error) that the number of vertices with the majority colour is $\nu_{\mathrm{o},1}n\pm O(n^{2/3})$, and using a Chernoff bound as in the proof of Corollary 5.12, the number of resulting in-edges

between these vertices is $p \cdot \rho_{\mathrm{o},11} \cdot \frac{dn}{2} + O(n^{2/3})$. Since these in-edges are uniformly distributed, once we condition on the numbers of vertices and edges in this graph, this becomes an exact edge model that satisfies (for all n large enough) the conditions for Item 2 in Lemma 5.24. Conversely, when we consider any of the remaining colours and condition on the number of vertices and in-edges after the percolation step, we get the conditions in Item 1 in Lemma 5.24. As a result, there will be a unique largest component of in-edges of linear size whose endpoints (in $\hat{\boldsymbol{\sigma}}_{\mathrm{o}}$) have the majority colour, while all remaining components contain at most $O(n^{2/3})$. Moreover, this continues to apply after removing the edges of a path of length at most $O(n^{1/6})$: only $O(n^{1/6})$ vertices then have degree different from d. Thus, for the majority colour, the resulting graph is in the setting of Item 2 of Lemma 5.24, which will give us an upper bound on the event that none of the vertices of the path belongs to the largest component.

We use this in the following lemma, which is a key technical result of this chapter.

Lemma 5.26. *Let $d, q \geq 3$ be integers, $\beta \geq \beta_c$ be a real and $p = 1 - \mathrm{e}^{-\beta}$. For any constant $\varepsilon > 0$, and for any integer ℓ satisfying $\ln \ln n \leq \ell \leq n^{1/6}$, the following holds w.h.p. over $G \sim \mathbb{G}_{n,d}$.*

For a vertex $v \in V(G)$, consider the event $\mathcal{A}_{v,\ell}$, as defined in (5.19)*. Then, the ordered RC distribution on G satisfies*

$$\pi_{G,\mathrm{o}}\left(\bigcup_{v \in V(G)} \mathcal{A}_{v,\ell}\right) \leq nd(d-1)^{\ell-1}(A+\varepsilon)^{\ell},$$

where $A = \max_{i \in [q]} \left\{1 - (1 - \sqrt{\hat{\varphi}_1})\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}}\right\}$ with $\rho_{\mathrm{o}}, \nu_{\mathrm{o}}$ as in (5.2) *and $\hat{\varphi}_1 = \hat{\varphi}(p_1)$ as in* (5.20) *for $p_1 = p\frac{\rho_{\mathrm{o},11}}{\nu_{\mathrm{o},1}}$.*

To parse the bound on $\pi_{G,\mathrm{o}}\big(\bigcup_{v \in V(G)} \mathcal{A}_{v,\ell}\big)$, note that the factor $nd(d-1)^{\ell-1}$ is just an upper bound on the total number of paths of length ℓ. The factor A^{ℓ} controls the probability that a particular path P avoids the largest component and follows from an application of Lemma 5.24. The term $\hat{\varphi}_1$ in A is the probability that the Galton-Watson process dies out from a vertex v of the path that belongs to the supercritical colour class 1 (note that v has "remaining" degree $d-2$ other than its neighbours on the path, and that the relevant percolation parameter is given by p_1). The term $\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}}$ accounts for how frequently colour class 1 (the majority colour) appears along any "potential path" P in the planting model, which follows from the symmetry of the planted graph. The exact quantitative dependence of A on $\hat{\varphi}_1$ and $\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}}$ will then follow by a more technical calculation involving eigenvalues, carried out in the proof

below. Later, in Lemma 5.27, we will show that A can be made less than $1/d^5$ by taking q sufficiently large.

Proof. From Lemmas 5.2 and 5.6, it suffices to show for the ordered planted graph $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}})$ that (for all n sufficiently large)

$$\mathbb{P}\Big[\boldsymbol{F}\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}\big) \in \bigcup_{v \in V(G)} \mathcal{A}_{v,\ell}\Big] \le nd\big(d-1\big)^{\ell-1}\big(A + \tfrac{\varepsilon}{2}\big)^{\ell}. \tag{5.25}$$

From this, we obtain by Lemma 5.2 (applied with $g(n)$ being the r.h.s. of (5.25) and $f(n) = 1/\ln(n)$) that

$$\mathbb{P}\big[\boldsymbol{F}\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{o}}\big) \in \bigcup_{v \in V} \mathcal{A}_{v,\ell} \mid \boldsymbol{G}\big] \le n\ln(n)d\big(d-1\big)^{\ell-1}\big(A + \tfrac{\varepsilon}{2}\big)^{\ell}.$$

Since $\boldsymbol{F}\big(\boldsymbol{G}, \boldsymbol{\sigma}_{\boldsymbol{G},\mathrm{o}}\big)$ has distribution $\hat{\pi}_{\boldsymbol{G},\mathrm{o}}$, Lemma 5.6 yields that, w.h.p. over $\boldsymbol{G}$, it holds that $\pi_{\boldsymbol{G},\mathrm{o}}(\bigcup_v \mathcal{A}_{v,\ell}) \le \hat{\pi}_{\boldsymbol{G},\mathrm{o}}(\bigcup_v \mathcal{A}_{v,\ell}) + \mathrm{e}^{-\Omega(n)} \le nd(d-1)^{\ell-1}(A+\varepsilon)^{\ell}$ (where the $\ln n$ factor and the $\mathrm{e}^{-\Omega(n)}$ term are absorbed into the $(A+\varepsilon)^{\ell}$ factor, using that $\ell \ge \ln\ln n$).

By Lemma 5.3, we have $\big\|\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{o}}} - \nu_{\mathrm{o}}\big\|_1 + \big\|\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}} - \rho_{\mathrm{o}}\big\|_1 = O(n^{-1/3})$ w.p. $1 - \mathrm{e}^{-\Omega(n^{1/3})}$. Condition on $\nu^{\hat{\boldsymbol{\sigma}}_{\mathrm{o}}} = \nu$ and $\rho^{\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}} = \rho$ where ν and ρ satisfy $\|\nu - \nu_{\mathrm{o}}\|_1 + \|\rho - \rho_{\mathrm{o}}\|_1 = O(n^{-1/3})$, and note that by the definition of the planted model, the pair $(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}})$ is uniformly random from the pairs having these statistics. We use this to bound the probability of a particular potential path occurring, counting the possible ways to pair the relevant half-edges in the configuration model with coloured vertices, and then to apply results for the exact-edge model to bound the probability that the path avoids the giant component.

Let P be an arbitrary sequence of vertices $(v_0, v_1, \dots, v_\ell)$, together with a q-colouring $\tau = (\tau_0, \dots, \tau_\ell)$. Let also ξ be a set of paired half-edges that turn $v_0, \dots, v_\ell$ into a path in that order. Let $\mathbb{E}$ be the set of random edges in the graph before percolation. Denote by

$$\mathcal{E}_{P,\tau,\xi} = \{\xi \subseteq \mathbb{E},\ \hat{\boldsymbol{\sigma}}_{\mathrm{o}}(v_0) = \tau_0, \dots,\ \hat{\boldsymbol{\sigma}}_{\mathrm{o}}(v_\ell) = \tau_\ell\}$$

the event that $v_0, v_1, \dots, v_\ell$ form a path in $\hat{\boldsymbol{G}}$ (in this order) using the half-edges prescribed by ξ and that they belong to the colour classes that τ prescribes.

We will use the following claim, proved at the end of the proof, which gives an asymptotic expression for $\mathbb{P}(\mathcal{E}_{P,\tau,\xi})$ in terms of ν and ρ.

Claim 1.

$$\mathbb{P}(\mathcal{E}_{P,\tau,\xi}) \sim (dn)^{-\ell}\nu_{\mathrm{o},\tau_0}\prod_{k=0}^{\ell-1}\frac{\rho_{\mathrm{o},\tau_k\tau_{k+1}}}{\nu_{\mathrm{o},\tau_k}}, \tag{5.26}$$

where $f(n) \sim g(n)$ denotes $\frac{f(n)}{g(n)} \to 1$ as $n \to \infty$.

Assuming the claim for now, fix an arbitrary vertex v and let P be an arbitrary sequence $(v_0, v_1, \dots, v_\ell)$ with $v_0 = v$, together with a q-colouring $\tau = (\tau_0, \dots, \tau_\ell)$ and pairing ξ (as above). Condition on the event $\mathcal{E}_{P,\tau,\xi}$. Let $\boldsymbol{F}$ be the edge set obtained by keeping each monochromatic edge with probability p, and $\boldsymbol{F}_v$ be the set of these edges that are incident to v. Then, as we discussed after Lemma 5.25, for the graph $\hat{\boldsymbol{G}}_v(\boldsymbol{F}) \coloneqq (\boldsymbol{V}\backslash\{v\}, \boldsymbol{F}\backslash\boldsymbol{F}_v)$, the subgraph of $\hat{\boldsymbol{G}}_v(\boldsymbol{F})$ consisting of vertices with the majority colour corresponds to the supercritical setting of Item 2 in Lemma 5.24, and the subgraphs consisting of vertices of any other colour correspond to the subcritical setting of Item 1 in Lemma 5.24. We will show this more formally in the proof of the following claim, which bounds the probability that the path specified by $\mathcal{E}_{P,\tau,\xi}$ avoids the largest component.

Claim 2.

$$\mathbb{P}\Big[\begin{array}{c}\textit{the largest component of } \hat{\boldsymbol{G}}_v(\boldsymbol{F}) \\ \textit{avoids the } \tau\textit{-coloured path } P\end{array} \Big| \mathcal{E}_{P,\tau,\xi}\Big] \le (1+\tfrac{\varepsilon}{3})^\ell \hat{\varphi}_1^{\ell_1(\tau)}, \tag{5.27}$$

where $\ell_1(\tau)$ is the number of vertices on the path, apart from v, with colour 1 under τ and $\hat{\varphi}_1 \coloneqq (p_1\varphi_1 + 1 - p_1)^{d-2}$ where $p_1 = p\frac{\rho_{\mathrm{o},11}}{\nu_{\mathrm{o},1}}$ and $\varphi_1 = \varphi(p_1)$ (as in (5.20)).

We assume both claims for now and show how they together conclude the lemma. Fix a colouring τ. There are $\sim n^\ell$ sequences P with $\ell+1$ vertices starting from v, and $d^2(d(d-1))^{\ell-1}$ ways to select the half-edges ξ that turn each such sequence into a path. The factors counting P and ξ combine with the $(dn)^{-\ell}$ factor in Claim 1 as $n^\ell d^2(d(d-1))^{\ell-1}(dn)^{-\ell} = d(d-1)^{\ell-1}$.

Combining (5.26) and (5.27) via a union bound over all possible paths yields that

$$\mathbb{P}\left[\begin{array}{c}\text{the largest component} \\ \text{of } \hat{\boldsymbol{G}}_v(\boldsymbol{F}) \text{ avoids some} \\ \tau\text{-coloured path from } v \text{ in } G\end{array}\right] \le (1+\tfrac{\varepsilon}{3})^\ell d(d-1)^{\ell-1}\nu_{\mathrm{o},\tau_0}\prod_{k=0}^{\ell-1}\frac{\rho_{\mathrm{o},\tau_k\tau_{k+1}}}{\nu_{\mathrm{o},\tau_k}}\hat{\varphi}_1^{\mathbf{1}\{\tau_{k+1}=1\}}. \tag{5.28}$$

At this stage, it remains to take a union bound over all possible colourings τ on $\ell+1$ vertices, i.e., sum the r.h.s. of (5.28) over all τ. The resulting expression can be upper-bounded by a more succinct expression, using that $\frac{\hat{\varphi}_1^{\mathbf{1}\{\tau_\ell=1\}}}{\hat{\varphi}_1^{\mathbf{1}\{\tau_0=1\}}} \le \hat{\varphi}_1^{-1}$:

$$(1+\tfrac{\varepsilon}{3})^\ell d(d-1)^{\ell-1}\hat{\varphi}_1^{-1}\cdot\hat{\nu}_{\mathrm{o}}^T M^\ell 1,$$

where M is the $q\times q$ matrix whose (i,j)-entry is given by $\frac{\rho_{\mathrm{o},ij}}{\hat{\nu}_{\mathrm{o},i}}$ and $\hat{\nu}_{\mathrm{o}}$ is the vector whose i-th entry is equal to $\nu_{\mathrm{o},1}/\hat{\varphi}_1$ if $i=1$ and $\nu_{\mathrm{o},i}$ if $i\in\{2,\dots,q\}$. Note that $M = D^{-1/2}\hat{M}D^{1/2}$ where D is the $q\times q$ diagonal matrix with $\hat{\nu}_{\mathrm{o}}$ on the diagonal and

$\hat{M}$ is the symmetric matrix $\frac{\rho_{\mathrm{o},ij}}{(\hat{\nu}_{\mathrm{o},i}\hat{\nu}_{\mathrm{o},j})^{1/2}}$. In particular, we have that $\hat{M}$ and M have the same eigenvalues, which are all real. Let $\sqrt{\hat{\nu}_{\mathrm{o}}}$ denote the vector whose entries are the square roots of the entries of $\hat{\nu}_{\mathrm{o}}$. Then, we can write

$$\hat{\nu}_{\mathrm{o}}^T M^\ell 1 = \sqrt{\hat{\nu}_{\mathrm{o}}}^T \hat{M}^\ell \sqrt{\hat{\nu}_{\mathrm{o}}} \leq \tfrac{1}{\hat{\varphi}_1} \sqrt{\nu_{\mathrm{o}}}^T \hat{M}^\ell \sqrt{\nu_{\mathrm{o}}}.$$

Now, we will show that entry-wise it holds that $\hat{M}\sqrt{\nu_{\mathrm{o}}} \leq A\sqrt{\nu_{\mathrm{o}}}$, for $A = \max_{i\in[q]}\{1 - \frac{(1-\sqrt{\hat{\varphi}_1})\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}}\}$ (as in the lemma statement).

Note first that $(\hat{M}\sqrt{\nu_{\mathrm{o}}})_i = \sum_j \frac{\rho_{\mathrm{o},ij}}{(\hat{\nu}_{\mathrm{o},i}\hat{\nu}_{\mathrm{o},j})^{1/2}}(\nu_{\mathrm{o},j})^{1/2}$. Since $\hat{\nu}_{\mathrm{o},j} = \nu_{\mathrm{o},j}$ for $j \neq 1$ and $\sum_j \rho_{\mathrm{o},ij} = \nu_{\mathrm{o},i}$ for $i \in [q]$, we have that

$$\begin{aligned}(\hat{M}\sqrt{\nu_{\mathrm{o}}})_1 &= \frac{\hat{\varphi}_1\,\rho_{\mathrm{o},11}}{(\nu_{\mathrm{o},1})^{1/2}} + \sum_{j\neq 1} \frac{\sqrt{\hat{\varphi}_1}\rho_{\mathrm{o},1j}}{(\nu_{\mathrm{o},1})^{1/2}} \\ &= \frac{(\hat{\varphi}_1 - \sqrt{\hat{\varphi}_1})\rho_{\mathrm{o},11}}{(\nu_{\mathrm{o},1})^{1/2}} + \sqrt{\hat{\varphi}_1}(\nu_{\mathrm{o},1})^{1/2} \leq \sqrt{\hat{\varphi}_1} A (\nu_{\mathrm{o},1})^{1/2}.\end{aligned}$$

For $i \neq 1$, we have

$$(\hat{M}\sqrt{\nu_{\mathrm{o}}})_i = \frac{\sqrt{\hat{\varphi}_1}\rho_{\mathrm{o},i1}}{(\nu_{\mathrm{o},i})^{1/2}} + \sum_{j\neq 1} \frac{\rho_{\mathrm{o},ij}}{(\nu_{\mathrm{o},i})^{1/2}} = (\nu_{\mathrm{o},i})^{1/2} - \frac{(1-\sqrt{\hat{\varphi}_1})\rho_{\mathrm{o},i1}}{(\nu_{\mathrm{o},i})^{1/2}} \leq A(\nu_{\mathrm{o},i})^{1/2}.$$

This finishes the proof of $\hat{M}\sqrt{\nu_{\mathrm{o}}} \leq A\sqrt{\nu_{\mathrm{o}}}$. Therefore, we can bound the sum of r.h.s. of (5.28) over all colourings τ by $(1+\frac{\varepsilon}{3})^\ell d(d-1)^{\ell-1}(\hat{\varphi}_1)^{-2}A^\ell(\sqrt{\nu_{\mathrm{o}}}^T\sqrt{\nu_{\mathrm{o}}})$. Since $A \leq 1$ and $\ell \geq \ln\ln n$, the factor $(1+\varepsilon/3)^\ell$ and the constant factors can be absorbed into the ε-slack, giving that the resulting probability is at most $d(d-1)^{\ell-1}(A+\frac{\varepsilon}{2})^\ell$. Taking the union over all possible colourings τ and all vertices v of the events in (5.28) gives the required bound on $\mathbb{P}\Big[\boldsymbol{F}\big(\hat{\boldsymbol{G}}(\hat{\boldsymbol{\sigma}}_{\mathrm{o}}), \hat{\boldsymbol{\sigma}}_{\mathrm{o}}\big) \in \bigcup_{v\in V(G)} \mathcal{A}_{v,\ell}\Big]$.

It remains to prove the two claims. We start with Claim 1, which gives us equation (5.26).

Recall that the graph $\hat{\boldsymbol{G}}$ has $n_i = n\nu_i$ vertices of colour $i \in [q]$ and the number of edges between colour classes i and j with $i \neq j$ is given by $2m_{ij} \coloneqq dn\rho_{ij}$ if $i \neq j$, and $m_{ii} \coloneqq \frac{dn}{2}\rho_{ii}$ if $i = j$ (see (5.5)); we further have that ν and ρ satisfy $\|\nu-\nu_{\mathrm{o}}\|_1 + \|\rho-\rho_{\mathrm{o}}\|_1 = O(n^{-1/3})$. The graph $\hat{\boldsymbol{G}}$ is uniformly distributed conditioned on these statistics. In particular, the number of ways to match the half-edges consistent with the vertex statistics $\{n_i\}$ and edge statistics $\{m_{ij}\}$ is given by

$$\mathcal{N} \coloneqq \binom{n}{n_1,\ldots,n_q} \prod_{i\in[q]} \binom{dn_i}{2m_{i1},\ldots,2m_{iq}} \prod_{i\in[q]} \frac{(2m_{ii})!}{2^{m_{ii}}m_{ii}!} \times \prod_{i<j}(2m_{ij})!.$$

There are $\binom{n}{n_1,\ldots,n_q}$ ways to assign colours to get the right colour statistics. Then, for a colour $i \in [q]$, there are dn_i half-edges which can be split among the q-colour classes

in $\binom{dn_i}{2m_{i1},\dots,2m_{iq}}$ ways so that the j-th class gets $2m_{ij}$ half-edges ($j \in [q]$). For colours $i \neq j$, the number of ways to match the selected half-edges is $(2m_{ij})!$, whereas for $i = j$ it is $(2m_{ii})!! = \frac{(2m_{ii})!}{2^{m_{ii}} m_{ii}!}$.

For the τ-coloured path P, for colours $i, j \in [q]$ with $i \neq j$, let $\ell_i = \ell_i(\tau)$ be the number of vertices of colour i, let $e_{ii} = e_{ii}(\tau)$ be the number of edges whose endpoints have both colour i, and let $2e_{ij} = 2e_{ij}(\tau)$ be the number of edges with one endpoint of colour i and the other colour j (in particular, $e_{ij} = e_{ji}$). Recall also that ξ is the set of paired half-edges of vertices of P that turns P into a path. Note that for $i \in [q]$, every vertex v in the path of colour i uses two half-edges from the colour class i (unless v is one of the endpoints v_0 and v_ℓ, in which case it uses one), so the remaining number of half-edges incident to colour class i is $M_i = dn_i - 2\ell_i + \mathbf{1}\{\tau_0 = i\} + \mathbf{1}\{\tau_\ell = i\}$. Let $m'_{ij} = m_{ij} - e_{ij}$.

Akin to the previous counting, we obtain that the number of ways to colour the remaining vertices and match the remaining edges so that $\mathcal{E}_{P,\tau,\xi}$ occurs is $\mathcal{N}'$ where

$$\mathcal{N}' := \binom{n-(\ell+1)}{n_1-\ell_1,\dots,n_q-\ell_q} \prod_{i\in[q]} \binom{M_i}{2m'_{i1},\dots,2m'_{iq}} \times \prod_{i\in[q]} \frac{(2m'_{ii})!}{2^{m'_{ii}}(m'_{ii})!} \times \prod_{i<j} (2m'_{ij})!.$$

We have

$$\frac{\mathcal{N}'}{\mathcal{N}} = \frac{(n-(\ell+1))!}{n!} \prod_{i\in[q]} \frac{n_i!}{(n_i-\ell_i)!} \prod_{i\in[q]} \frac{(M_i)!}{(dn_i)!} \prod_{i\in[q]} \frac{2^{m_{ii}}(m_{ii})!}{2^{m'_{ii}}(m'_{ii})!} \prod_{i<j} \frac{(2m_{ij})!}{(2m'_{ij})!}. \tag{5.29}$$

First of all, $\frac{(n-(\ell+1))!}{n!} \sim n^{-(\ell+1)}$. Then using that $n_i = n\nu_{\mathrm{o},i} + O(n^{2/3})$ and $\sum_{i\in[q]} \ell_i = \ell + 1$, the second and third factors in (5.29) are respectively,

$$\prod_{i\in[q]} \frac{n_i!}{(n_i-\ell_i)!} \sim n^{\ell+1} \prod_{i\in[q]} (\nu_{\mathrm{o},i})^{\ell_i} \quad \text{and} \quad \prod_{i\in[q]} \frac{(M_i)!}{(dn_i)!} \sim \frac{(dn)^2 \nu_{\mathrm{o},\tau_0}\nu_{\mathrm{o},\tau_\ell}}{\prod_{i\in[q]} (dn\,\nu_{\mathrm{o},i})^{2\ell_i}}$$

.

Note that in the latter expression, the factor $(dn)^2\nu_{\mathrm{o},\tau_0}\nu_{\mathrm{o},\tau_\ell}$ accounts for the presence of the term $\mathbf{1}\{\tau_0 = i\} + \mathbf{1}\{\tau_\ell = i\}$ in the expression for M_i.

Finally, using that $2m_{ij} = dn\rho_{\mathrm{o},ij} + O(n^{2/3})$, we also have that

$$\prod_{i\in[q]} \frac{2^{m_{ii}}(m_{ii})!}{2^{m'_{ii}}(m'_{ii})!} \prod_{i<j} \frac{(2m_{ij})!}{(2m'_{ij})!} \sim \prod_{i,j\in[q]} (dn\rho_{\mathrm{o},ij})^{e_{ij}}.$$

Combining these and using that[12] $\sum_{i,j} e_{ij} = \ell$. we obtain that

$$\mathbb{P}(\mathcal{E}_{P,\tau,\xi}) = \frac{\mathcal{N}'}{\mathcal{N}} \sim (dn)^{-\ell} \nu_{\mathrm{o},\tau_0}\nu_{\mathrm{o},\tau_\ell} \prod_{i,j\in[q]} (\rho_{\mathrm{o},ij})^{e_{ij}} \prod_{i\in[q]} (\nu_{\mathrm{o},i})^{-\ell_i}.$$

[12]To see this, note that every internal vertex of P has degree 2. By considering the edges incident to colour i (and paying attention to endpoints), we have that $2\ell_i - \mathbf{1}\{\tau_0 = i\} - \mathbf{1}\{\tau_\ell = i\} = \sum_{j\in[q]} 2e_{ij}$. Summing over $i \in [q]$ and using that $\sum_i 2\ell_i = 2(\ell+1)$ yields the stated equality.

To obtain the expression in (5.26), we write e_{ij} and ℓ_i as sums of indicators. Then

$$\prod_{i,j\in[q]}(\rho_{\mathrm{o},ij})^{e_{ij}}\prod_{i\in[q]}(\nu_{\mathrm{o},i})^{-\ell_i}=\prod_{k=0}^{\ell-1}\rho_{\mathrm{o},\tau_k\tau_{k+1}}\prod_{k=0}^{\ell}\nu_{\mathrm{o},\tau_k}^{-1}=\frac{1}{\nu_{\mathrm{o},\tau_\ell}}\prod_{k=0}^{\ell-1}\frac{\rho_{\mathrm{o},\tau_k\tau_{k+1}}}{\nu_{\mathrm{o},\tau_k}}.$$

By adding in the factor $(dn)^{-\ell}\nu_{\mathrm{o},\tau_0}\nu_{\mathrm{o},\tau_\ell}$ and noting that $\nu_{\mathrm{o},\tau_\ell}$ cancels out, we obtain the expression (5.26), which concludes the proof of Claim 1.

It remains to prove Claim 2, for which we use Lemma 5.24. We first show that we can indeed consider the exact edge model in this setting and apply the lemma for an appropriate set S': the vertices in our chosen path coloured with the majority colour.

Condition on $\mathcal{E}_{P,\tau,\xi}$ and on the set of edges E_v in $\hat{\boldsymbol{G}}$ incident to v. Let $\mathbf{d}$ be the degree sequence of the graph $\hat{\boldsymbol{G}}\backslash v=(\boldsymbol{V}\backslash v,\boldsymbol{E}\backslash(E_P\cup E_v))$. For a colour $i\in[q]$, let $\boldsymbol{V}_i$ denote the i-th colour class of $\hat{\boldsymbol{G}}\backslash v$ under $\hat{\boldsymbol{\sigma}}_{\mathrm{o}}$, $\mathbf{d}_i$ be the subsequence of $\mathbf{d}$ corresponding to vertices in $\boldsymbol{V}_i$, $\mathbb{E}_i$ be the set of monochromatic edges within $\boldsymbol{V}_i$, and $\boldsymbol{F}_i=\boldsymbol{F}\cap\mathbb{E}_i$. Conditioned on $|\boldsymbol{F}_i|=f_i$, the graph $\boldsymbol{G}_i=\big(\boldsymbol{V}_i,\boldsymbol{F}_i\backslash(E_P\cup E_v)\big)$ has the same distribution as the exact-edge model $\tilde{\boldsymbol{G}}_{n_i,\mathbf{d}_i,f_i}$ where $n_i=|\boldsymbol{V}_i|=n\nu_{\mathrm{o},i}+o(n)$. Note that for all $i\in[q]$, there are at most $\ell+1+d=O(n^{1/6})$ vertices in $\boldsymbol{d}_i$ with degree $<d$ (all other vertices have degree d).

Since $|\mathbb{E}_i|=\frac{dn}{2}\rho_{\mathrm{o},ii}+o(n)$, by a Chernoff bound we have that $|\boldsymbol{F}_i|=p\frac{dn}{2}\rho_{\mathrm{o},ii}+o(n)$ with probability $1-\mathrm{e}^{-\Omega(n)}$. It follows from Lemma 5.25, applied with densities $p_i=p\frac{\rho_{\mathrm{o},ii}}{\nu_{\mathrm{o},i}}$ for $i\in[q]$, that $\boldsymbol{G}_1$ is in the supercritical regime while $\boldsymbol{G}_i$ for $i\in\{2,\dots,q\}$ is in the subcritical regime. Hence, it follows from Lemma 5.24 that w.p. $1-\mathrm{e}^{-\Omega(n)}$, the largest component $\mathcal{C}$ in $\boldsymbol{G}_1$ has size $\Omega(n)$ and the remaining components have all size $O(n^{2/3})$. Moreover, from Item 2 of Lemma 5.24 we obtain that[13]

$$\mathbb{P}[\mathcal{C}\text{ avoids the path }P\mid\mathcal{E}_{P,\tau,\xi},E_v]\le(1+\tfrac{\varepsilon}{3})^{\ell}(p_1\varphi_1+1-\varphi_1)^{(d-2)\ell_1}\le(1+\tfrac{\varepsilon}{3})^{\ell+1}\hat{\varphi}_1^{\ell_1}.\tag{5.30}$$

By contrast, for $i\in\{2,\dots,q\}$, w.p. $1-\mathrm{e}^{-\Omega(n)}$, the largest component in $\boldsymbol{G}_i$ has size $O(n^{1/2})$. Since the union of the $\boldsymbol{G}_i$'s over $i\in[q]$ is the graph $\widetilde{\boldsymbol{G}}_{P,v}=\big(\boldsymbol{V}\backslash\{v\},\boldsymbol{F}\backslash(E_P\cup E_v)\big)$, by a union bound it follows that, w.p. $1-\mathrm{e}^{-\Omega(n)}$, $\mathcal{C}$ is the largest component in $\widetilde{\boldsymbol{G}}_{P,v}$ with $\Omega(n)$ size, while all the others have size $O(n^{2/3})$. Now add the monochromatic edges of the path P, each percolated with probability p, to obtain the graph $\hat{\boldsymbol{G}}_v(\boldsymbol{F})=(\boldsymbol{V}\backslash\{v\},\boldsymbol{F}\backslash\boldsymbol{F}_v)$. Let $\mathcal{C}'$ be the component of $\hat{\boldsymbol{G}}_v(\boldsymbol{F})$ with $\mathcal{C}\subseteq\mathcal{C}'$. Since we added only edges on the path P, we have

$$\mathbb{P}[\mathcal{C}'\text{ avoids the path }P\mid\mathcal{E}_{P,\tau,\xi}]=\mathbb{P}[\mathcal{C}\text{ avoids the path }P\mid\mathcal{E}_{P,\tau,\xi}].\tag{5.31}$$

[13]Note that, because the edges in E_v are excluded from $\hat{\boldsymbol{G}}\setminus v$, all but at most $d+2$ vertices of the path P have degree $d-2$ in $\hat{\boldsymbol{G}}\setminus v$; the contribution of vertices with degree $<d-2$ is at most a constant which is absorbed into the $(1+\frac{\varepsilon}{3})^{\ell}$ factor in (5.30).

Moreover, the length of the path P is $\leq n^{1/6}$, so $|\mathcal{C}'| - |\mathcal{C}| = O(n^{5/6})$ (since the added path edges can join $\mathcal{C}$ to at most $O(n^{1/6})$ components, each of size $O(n^{2/3})$). Similarly, any component in $\hat{\boldsymbol{G}}_v(\boldsymbol{F})$ other than $\mathcal{C}'$ has size $O(n^{5/6})$, and in particular $\mathcal{C}'$ is the largest component in $\hat{\boldsymbol{G}}_v(\boldsymbol{F})$. Combining this with (5.30) and (5.31), we obtain (5.27), which concludes the proof of the lemma. □

We next bound the quantity A appearing in Lemma 5.26 by a suitably small constant so that we can take a union bound over paths; it is for this step that we require $q \geq (5d)^5$.

Lemma 5.27. *Let $d \geq 3$, $q \geq (5d)^5$ and $\beta \geq \beta_c$. Then, for all $i \in [q]$, it holds that $(1 - \sqrt{\hat{\varphi}_1})\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}} > 1 - \frac{1}{(d-1)^5}$.*

Proof. First, we observe that for $\beta = \beta_c$, it holds that $\mathrm{e}^{\beta} = \frac{q-2}{(q-1)^{1-2/d}-1}$ and hence the value of t in (5.1) is $t_c = (q-1)^{2/d}$, which in turn gives $\nu_{\mathrm{o},1} = \frac{q-1}{q}$ and $\nu_{\mathrm{o},i} = \frac{1}{q(q-1)}$ for $i \in \{2, \dots, q\}$. Consider henceforth arbitrary $\beta \geq \beta_c$; then the value of t in (5.1) satisfies $t \geq t_c$.

To show the desired inequality, we use the expressions for $\rho_{\mathrm{o},ij}$ given in (5.6). We have that $\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}}$ is $\frac{\mathrm{e}^{\beta}}{W}\frac{(\nu_{\mathrm{o},1})^{(d-1)/d}}{(\nu_{\mathrm{o},1})^{1/d}}$ if $i = 1$ and $\frac{1}{W}\frac{(\nu_{\mathrm{o},1})^{(d-1)/d}}{(\nu_{\mathrm{o},i})^{1/d}}$ for $i \neq 1$, where $W = \sum_{i,j\in[q]} \mathrm{e}^{\beta\mathbf{1}\{i=j\}}(\nu_{\mathrm{o},i}\nu_{\mathrm{o},j})^{(d-1)/d}$. From (5.2) and (5.1), we have that $\nu_{\mathrm{o},1} = \frac{t^d}{t^d+q-1}$ and $\nu_{\mathrm{o},i} = \frac{1}{t^d+q-1}$. Hence, for $i \neq 1$ it follows that $\mathrm{e}^{\beta}\big(\frac{\nu_{\mathrm{o},i}}{\nu_{\mathrm{o},1}}\big)^{1/d} = \mathrm{e}^{\beta}/t \geq 1$. In particular, it then holds that $\frac{\rho_{\mathrm{o},11}/\nu_{\mathrm{o},1}}{\rho_{\mathrm{o},i1}/\nu_{\mathrm{o},i}} = \mathrm{e}^{\beta}\left(\frac{\nu_{\mathrm{o},i}}{\nu_{\mathrm{o},1}}\right)^{1/d} \geq 1$,. It suffices therefore to show the bound for $i \in \{2, \dots, q\}$; fix an arbitrary such i.

The expression for $\rho_{\mathrm{o},i1}$ given in (5.6) yields $\rho_{\mathrm{o},i1} = \frac{t^{d-1}}{(t^{d-1}+q-1)^2+(\mathrm{e}^{\beta}-1)(t^{2d-2}+(q-1))}$. Using the expression for $\mathrm{e}^{\beta}-1 = \frac{(t-1)(t^{d-1}+q-1)}{t^{d-1}-t}$ from (5.1) and $\nu_{\mathrm{o},i} = \frac{1}{t^d+q-1}$, we obtain

$$\begin{aligned}\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}} &= \frac{t^{d-1}(t^d+q-1)(t^{d-1}-t)}{(t^{d-1}+q-1)\big((t^{d-1}+q-1)(t^{d-1}-t)+(t-1)(t^{2d-2}+q-1)\big)} \\ &= \frac{t^{d-1}(t^{d-1}-t)}{(t^{d-1}+q-1)(t^{d-1}-1)} = 1 - \frac{t^{d+1}+(q-2)t^d-(q-1)t}{(t^{d-1}+q-1)(t^d-t)}.\end{aligned}$$

Hence we can write

$$\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}} = 1 - R \quad \text{where} \quad R \coloneqq \frac{t^{d+1}+(q-2)t^d-(q-1)t}{(t^d-t)(t^{d-1}+q-1)}. \tag{5.32}$$

So, to show the inequality $(1 - \sqrt{\hat{\varphi}_1})\frac{\rho_{\mathrm{o},i1}}{\nu_{\mathrm{o},i}} > 1 - \frac{1}{(d-1)^5}$ it suffices to show that $\frac{1}{(d-1)^5} > \sqrt{\hat{\varphi}_1} + R$.

To upper-bound $\hat{\varphi}_1$, we need first to find the solution of $\varphi_1 = (p_1\varphi_1+1-p_1)^{d-1}$ for $p_1 = p\frac{\rho_{\mathrm{o},11}}{\nu_{\mathrm{o},1}}$. Similarly to above, we have $\rho_{\mathrm{o},11} = \frac{\mathrm{e}^{\beta}t^{2(d-1)}}{(t^{d-1}+q-1)^2+(\mathrm{e}^{\beta}-1)(t^{2d-2}+(q-1))}$, so using the expression for $\mathrm{e}^{\beta}-1$ from (5.1) and cancelling common factors gives $p_1 = \frac{t^{d-1}(t-1)}{t^d-t}$. In particular, $p_1\varphi_1 + 1 - p_1 = t^{-1}$, so $\varphi_1 = t^{-(d-1)}$ satisfies $\varphi_1 = (p_1\varphi_1 + 1 - p_1)^{d-1}$.

Since $t \geq t_c$ for $\beta \geq \beta_c$, we obtain that $\varphi_1 \leq \frac{1}{(q-1)^{2(d-1)/d}}$. It follows that $\hat{\varphi}_1 = (p_1\varphi_1 + 1 - p_1)^{d-2} = \varphi_1^{(d-2)/(d-1)} \leq \frac{1}{(q-1)^{2(d-2)/d}}$, so $\sqrt{\hat{\varphi}_1} \leq \frac{1}{(q-1)^{(d-2)/d}}$. For $q \geq (5d)^5$ and $d \geq 3$, we have $(q-1)^{(d-2)/d} \geq \big((5d)^5 - 1\big)^{(d-2)/d} > 2(d-1)^5$, and hence $\sqrt{\hat{\varphi}_1} < \frac{1}{2(d-1)^5}$.[14]

Next, using the expression for R from (5.32), we have $R \leq R' \coloneqq \frac{t+q-1}{t^{d-1}+q-1}$, which follows from $(t+q-1)(t^d-t) - \big(t^{d+1}+(q-2)t^d-(q-1)t\big) = t^d - t^2 \geq 0$. We have that R' is decreasing in t, so using that $t \geq t_c$, we obtain that $R' \leq \frac{t_c+q-1}{t_c^{d-1}+q-1} = \frac{1}{(q-1)^{(d-2)/d}}$; from the same numerical inequality as above, we therefore conclude that R is less than $\frac{1}{2(d-1)^5}$ for all $d \geq 3$ when $q \geq (5d)^5$. Hence, $\frac{1}{(d-1)^5} > \sqrt{\hat{\varphi}_1} + R$, as needed. □

5.4.4 Proof of the WSM within the ordered phase

Finally, we can conclude WSM within the ordered phase, which we restate for convenience.

Theorem 5.10. *Let $d \geq 3, q \geq (5d)^5$ be integers, and $\beta \geq \beta_c$ be a real. Then, there exists $\delta > 0$ such that, w.h.p. over $G \sim \mathbb{G}_{n,d}$, G has WSM within the ordered phase at a distance $\ell = \lfloor(\frac{1}{2} - \delta)\log_{d-1} n\rfloor$. That is, with $m = |E|$, for every $v \in V$ and every edge e incident to v, it holds that $\big|\pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e)\big| \leq \frac{1}{100m}$.*

Proof. By Lemma 5.27, we have $A = \max_i\{(1 - (1 - \sqrt{\hat{\varphi}_1})\frac{\rho_{o,i1}}{\nu_{o,i}}\} < (d-1)^{-5}$. Let $\varepsilon > 0$ be a small constant (depending on d) such that $A + \varepsilon \leq (d-1)^{-(5+\varepsilon)}$. Let $\delta \coloneqq \frac{\varepsilon}{4(4+\varepsilon)}$ and $\ell = \lfloor(\frac{1}{2} - \delta)\log_{d-1} n\rfloor$. By Lemma 5.18,

$$\big|\pi^{+}_{G,v,\ell}(1_e) - \pi_{G,\mathrm{o}}(1_e)\big| \leq \pi_{G,\mathrm{o}}\big(\text{no wired boundary around } v \text{ in } B_\ell(v)\big) + \frac{2\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}{1-\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})},$$

where $\overline{\mathcal{O}}$ is the event (see (5.14)) that the outside of the ball contains fewer than $n\mathfrak{m}_{\mathrm{o}}(\beta_c) - n\rho/4$ in-edges. For all n large enough it holds that $|E(B_\ell(v))| \leq n^{1/2} \leq n\rho/8$. Hence, $\overline{\mathcal{O}}$ implies $|F| \leq n\mathfrak{m}_{\mathrm{o}}(\beta_c) - n\rho/8$. The probability of this event in the ordered phase is $\mathrm{e}^{-\Omega(n)}$ by Corollary 5.12, giving that $\frac{2\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})}{1-\pi_{G,\mathrm{o}}(\overline{\mathcal{O}})} = \mathrm{e}^{-\Omega(n)}$.

[14]The second inequality is immediate for $d = 3, 4$. For $d \geq 5$, we have $\big((5d)^5 - 1\big)^{(d-2)/d} > \frac{1}{2}(5d)^{5-10/d}$. Moreover, $(5d)^{10/d} \leq 625$, so $\frac{1}{2}(5d)^{5-10/d} = \frac{(5d)^5}{2(5d)^{10/d}} \geq \frac{5}{2}d^5 > 2(d-1)^5$.

Then, by Lemma 5.19, the probability of no wired boundary is at most $\pi_{G,\mathrm{o}}(\mathcal{A}_{v,\ell})$. By Lemma 5.26, this is at most $\mathrm{e}^{-\Omega(n)} + nd(d-1)^{\ell-1}(A+\varepsilon/2)^{\ell}$. So w.h.p. over $G \sim \mathbb{G}_{n,d}$ it holds for every vertex v and edge e incident to v that

$$\left|\pi_{G,\mathrm{o}}(1_e) - \pi^{+}_{G,v,\ell}(1_e)\right| \leq \mathrm{e}^{-\Omega(n)} + \frac{nd}{d-1}(d-1)^{-(4+\varepsilon)\ell} \leq \mathrm{e}^{-\Omega(n)} + \frac{1}{n^{1+\varepsilon/4}}$$

which is at most $\frac{1}{100m}$ for all n large enough. □

5.5 Proof of the main theorem

We now combine the three key properties established in this chapter: local maxima, WSM within the phase, and local mixing, and apply Theorem 1.5 to obtain the $O(n \log n)$ mixing time within each phase, giving Theorems 5.7 and 5.8.

Theorem 5.7. Let $d, q \geq 3$ be integers, and $\beta \in (0, \beta_c]$. Then, w.h.p. over $G \sim \mathcal{G}_{n,d}$, for every $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ the RC dynamics $(X_t)_{t\geq 0}$ initialised from all-out satisfies $\left\|X_T - \pi_{G,\mathrm{d}}\right\|_{\mathrm{TV}} \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.

Proof. First, we have that w.h.p. $G = (V, E) \sim \mathcal{G}_{n,d}$ has the three key properties. Note that we used that any w.h.p. property for $G \sim \mathbb{G}_{n,d}$ is also a w.h.p. property for $G \sim \mathcal{G}_{n,d}$.

1. The probability that the RC dynamics initialised from the disordered phase leaves the phase in one step is $\mathrm{e}^{-\Omega(n)}$ (Key Property 1); this follows from Corollary 5.12).

2. There exists $\delta > 0$ so that G has WSM within the disordered phase at a distance $\ell \leq (\frac{1}{2} - \delta) \log_{d-1} n$ (Key Property 2); this follows by Theorem 5.9.

3. For ℓ as above, there exists a constant $L > 0$ such that for every vertex v, it is the case that $B_\ell(v)$ is L-treelike, see, e.g., [10, Fact 2.3.]. Then, from Lemma 4.11, the RC dynamics has optimal local mixing in the disordered phase at radius ℓ (Key Property 3).

By Theorem 1.5, these three key properties imply $O(n \log n)$ mixing within the disordered phase, giving the theorem. □

Theorem 5.8. Let $d \geq 3, q \geq (5d)^5$ be integers, and $\beta \geq \beta_c$ be a real. Then, w.h.p. over $G \sim \mathcal{G}_{n,d}$, for every $\varepsilon \geq \mathrm{e}^{-\Omega(n)}$ the RC dynamics $(X_t)_{t\geq 0}$ initialised from all-in satisfies $\left\|X_T - \pi_{G,\mathrm{o}}\right\|_{\mathrm{TV}} \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.

Proof. The proof is very similar to the disordered phase. W.h.p. $G \sim \mathcal{G}_{n,d}$ satisfies the three key properties.

1. The probability that the RC dynamics initialised from the ordered phase leaves the phase in one step is $\mathrm{e}^{-\Omega(n)}$ (Key Property 1); this follows from Corollary 5.12).

2. There exists $\delta > 0$ so that G has WSM within the ordered phase at a distance $\ell \leq (\frac{1}{2} - \delta) \log_{d-1} n$ (Key Property 2); this follows by Theorem 5.10.

3. For ℓ as above, there exists a constant $L > 0$ such that for every vertex v, it is the case that $B_\ell(v)$ is L-treelike, see, e.g., [10, Fact 2.3.]. Then, from Lemma 6.56, the RC dynamics has optimal local mixing in the ordered phase at radius ℓ (Key Property 3).

By Theorem 1.5, these three key properties imply $O(n \log n)$ mixing within the ordered phase, giving the theorem. □

Theorem 1.2 then follows directly from Corollary 5.5 (which gives TV-distance $\mathrm{e}^{-\Omega(n)}$ from the dominant phase).

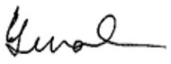
Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	*Planting and MCMC Sampling from the Potts model*
Publication Status	☒Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	An extended abstract appeared at *Andreas Galanis, Leslie Ann Goldberg, Paulina Smolarova: Planting and MCMC Sampling from the Potts Model. STACS 2026: 39:1-39:19* The full version has been invited for submission to *ACM Journal Transactions on Computation Theory (ToCT)*

Student Confirmation

Student Name:	Paulina Smolarova		
Contribution to the Paper	As is standard in our field, all authors contribute equally and authors are listed in alphabetic order		
Signature		Date	4.6. 2026

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof Leslie Ann Goldberg, Prof Andreas Galanis		
Supervisor comments: Paulina contributed substantially to all aspects		
Signature: Leslie Ann Goldberg Andreas Galanis	Date	4.6. 2026

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 6

Random-cluster model on sparse random graphs of unbounded degree

This chapter is based on the following publication:

[40] Andreas Galanis, Leslie Goldberg, and Paulina Smolarova. Low-temperature sampling on sparse random graphs. *Electronic Journal of Probability*, 31:1–40, 01 2026. doi: 10.1214/26-EJP1576

An extended abstract of this work also appeared in ICALP 2025:

[39] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Low-temperature sampling on sparse random graphs. In *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, volume 334 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 83:1–83:20, 2025. doi: 10.4230/LIPIcs.ICALP.2025.83

6.1 Overview

In the previous chapters, we considered the random-cluster model on random regular graphs. In this chapter, we extend our algorithmic results to sparse Erdős-Rényi random graphs $\mathcal{G}(n, d/n)$ for $d = \Theta(1)$, for all temperatures, a setting that has so far proved difficult for sampling algorithms. Recall the two key challenges for lifting sampling algorithms beyond bounded-degree graphs. In the high-temperature setting, the presence of high-degree vertices typically complicates the applicability of standard techniques [9, 13, 29, 30, 59, 64, 70], whereas in the low-temperature regime, the obstacles come primarily from another side, namely from the presence of many small

"sparse" parts which have weak, if any, expansion. This has posed the major challenge in lifting polymer-based techniques, such as cluster expansion, to $\mathcal{G}(n, d/n)$ [36].

Understanding the random-cluster model on $\mathcal{G}(n, d/n)$ is also more complicated than on random regular graphs. While the qualitative picture, with the presence of ordered and disordered phases, is conjectured to be similar, it remains open whether there is a sharp phase transition between the two phases on $\mathcal{G}(n, d/n)$. A key difference is that the local neighbourhood of a vertex is described by a Poisson tree[1], rather than a regular tree. For example, at low temperatures, the limiting log-partition function $\lim_{n\to\infty} \frac{1}{n} \log Z_G$ is believed to involve an expectation over Poisson trees.[2]

In this chapter, we make progress towards establishing the phase transition of the random-cluster model on $\mathcal{G}(n, d/n)$ for d and q large enough. While we do not identify a single critical temperature, we show the existence of a critical window $[\beta_0, \beta_1]$, with β_0 and β_1 as defined in (6.1): below β_0 the disordered phase dominates, above β_1 the ordered phase dominates, and for all temperatures β, a typical configuration falls into one of the phases. We then use this phase structure to establish the key properties required by Theorem 1.5 and obtain the first main result of this chapter.

Theorem 1.3. *Let d be a large enough real and $q = d^{\Omega(d)}$ be a real. Then, with high probability over $G \sim \mathcal{G}(n, d/n)$, the following hold.*

- *For all $\beta \leq \beta_0$, the RC dynamics initialised from all-out mixes in $O(n \log n)$ time.*
- *For all $\beta \geq \beta_1$, the RC dynamics initialised from all-in mixes in poly(n) time.*
- *For all $\beta \in (\beta_0, \beta_1)$, the RC dynamics initialised from a suitable random mixture of all-in and all-out mixes in polynomial time. Moreover, this random mixture can be computed in polynomial time.*
- *For all $\beta > 0$, there is an FPRAS for approximating the partition function.*

For $\beta > \beta_0$, the mixing time is $n^{O(\log q)}$; this is largely because of a crude polynomial mixing-time bound on Poisson trees with wired boundary (in Key Property 3). It is an open question to obtain sharper bounds with the wired boundary condition, even for regular trees when q is non-integer [15, Theorem 5]. The FPRAS for the partition

[1] A tree generated by a Galton-Watson process with offspring distribution Poisson(d).

[2] This is based on the fact that the model is "replica symmetric" [7]; to the best of our knowledge, the formula for the log-partition function has not been established yet.

function and the initial mixture for $\beta \in (\beta_0, \beta_1)$ are obtained using the technique from Section 3.2.

The main technical contribution of this chapter is a new polymer model for the random-cluster model on sparse random graphs in the low-temperature regime, applicable even to graphs with non-uniform and unbounded degrees. Combining this with the approaches from Chapters 4 and 5, we reduce WSM within the phase (Key Property 2) to showing that, with sufficiently high probability, all relevant polymers have size $O(\log n)$. Following [31], we prove a weaker expansion property of $\mathcal{G}(n, d/n)$ for connected sets of size $\Omega(\log n)$ and use it to bound the polymer sizes; this is where the requirement that d and q be sufficiently large arises.

Our second main result uses WSM within the phase to obtain a significantly faster deterministic approximation algorithm for the partition function, based on estimating edge marginals and integrating the expected number of in-edges.

Theorem 1.4. *Let $R > 0$ be any real. Then for all sufficiently large reals d and then all sufficiently large q, there is an algorithm which, with high probability over $G \sim \mathcal{G}(n, d/n)$, for any $\beta > 0$, outputs a $(1 + \frac{1}{n^R})$-approximation to $Z_G(\beta, q)$ in time $n^{1+O(1/R)}$.*

6.1.1 Setup and notation

It is well-known that for a disconnected graph G, the RC distribution is a product distribution over its connected components (that is, the distributions on the individual components are independent), and the partition function of G is the product of the partition functions of the individual components.

For $d > 1$, w.h.p. $\mathcal{G}(n, d/n)$ contains one component of linear size while all remaining components have size $O(\log n)$ and contain at most one cycle (see, for example, [51, Section 5] and also Section 6.2.1). 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 a *worst-case* initial configuration on such a component is at most $\text{poly}(n)$ (see, e.g., [10, 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. [51, Theorem 5.4]) that n_C is determined by the *conjugate* $\mu \in (0, 1)$ of d which satisfies $\mu e^{-\mu} = de^{-d}$. Using the formula for μ, we show later in Section 6.2 that for large d we have w.h.p. $n_C \geq (1 - e^{-d/3})n$ and $|E_C| = \frac{dn}{2} \pm 2ne^{-d/3}$. In other words, C contains almost all vertices and edges of G.

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 RC 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 the graph (V_C, F).

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

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

Definition 6.1 (Phases)**.** *Let* $\eta := 1/1000$. *The* ordered phase *is* $\Omega_C^{\mathrm{ord}} := \{F \in \Omega_C \mid |F| \geq (1-\eta)|E_C|\}$. *The ordered distribution* π_C^{ord} *is defined for* $F \in \Omega_C^{\mathrm{ord}}$ *by* $\pi_C^{\mathrm{ord}}(F) := w_C(F)/Z_C^{\mathrm{ord}}$, *where* $Z_C^{\mathrm{ord}} := \sum_{F \in \Omega_C^{\mathrm{ord}}} w_C(F)$. *The* disordered phase *is* $\Omega_C^{\mathrm{dis}} := \{F \in \Omega_C \mid |F| \leq \eta|E_C|\}$. *The disordered distribution* π_C^{dis} *is defined for* $F \in \Omega_C^{\mathrm{dis}}$ *by* $\pi_C^{\mathrm{dis}}(F) = w_C(F)/Z_C^{\mathrm{dis}}$, *where* $Z_C^{\mathrm{dis}} := \sum_{F \in \Omega_C^{\mathrm{dis}}} w_C(F)$.

We use two thresholds, β_0 and β_1, defined by:

$$\mathrm{e}^{\beta_0} - 1 = q^{(2-1/10)/d} \text{ and } \mathrm{e}^{\beta_1} - 1 = q^{(2+1/10)/d}. \tag{6.1}$$

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 the boundaries of the phases. In particular, as a consequence, the RC dynamics initialised from π_C^{ord} or π_C^{dis}, in the respective parameter regimes, is exponentially unlikely to leave the corresponding phase in a single step. The proof of the theorem is given in Section 6.3.

Theorem 6.2. *Let* d *be sufficiently large. Then, for all* q *sufficiently large, the following holds w.h.p. over* $G \sim \mathcal{G}(n, d/n)$, *where* F *denotes a random configuration from the given distribution.*

(1) For all $\beta > 0$, *it holds that* $\pi_C\Big(\frac{9\eta}{10}|E_C| \leq |F| \leq (1 - \frac{9\eta}{10})|E_C|\Big) \leq \mathrm{e}^{-n}$.

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

(3) For all $\beta \leq \beta_1$, *it holds that* $\pi_C^{\mathrm{dis}}\Big(\frac{9\eta}{10}|E_C| \leq |F|\Big) \leq \mathrm{e}^{-n}$.

(4) For $\beta \geq \beta_1$, *it holds that* $\big\|\pi_C - \pi_C^{\mathrm{ord}}\big\|_{\mathrm{TV}} \leq 2\mathrm{e}^{-n}$.

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

Having defined the phases, we can now state the theorems for mixing within each phase, which will together imply Theorem 1.3.

Theorem 6.3. *Let d be sufficiently large. Then, for all q sufficiently large, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, for all $\beta \leq \beta_1$ and $\varepsilon > \text{e}^{-\Omega(n)}$, the RC dynamics X_t initialised from all-out satisfies $||X_T - \pi_C^{\text{dis}}|| \leq \varepsilon$ for $T = O(n \log n \log \frac{1}{\varepsilon})$.*

Theorem 6.4. *Let d be sufficiently large. Then, for all q sufficiently large, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, for all $\beta \geq \beta_0$ and $\varepsilon > \text{e}^{-\Omega(n)}$, the RC dynamics X_t initialised from all-in satisfies $||X_T - \pi_C^{\text{ord}}|| \leq \varepsilon$ for $T = \text{poly}(n, \log \frac{1}{\varepsilon})$.*

We prove these two theorems using Theorem 1.5, for which we need to establish the three key properties (local maximum, WSM within the phase and local mixing).

Theorem 6.2 implies that the RC dynamics has a local maximum (Key Property 1) in each phase within the appropriate regimes, since to leave a phase in one step, the dynamics needs to be initialised right on its boundary (that is, with precisely $\lfloor \eta |E_C| \rfloor$ in-edges for the disordered phase and $\lfloor (1-\eta)|E_C| \rfloor$ in-edges for the ordered phase), which is exponentially unlikely by Items 2 and 3 of Theorem 6.2. The local mixing (Key Property 3) follows largely from existing results, which we state alongside the proof of Theorems 6.3-6.4, in Section 6.6.

6.1.3 WSM within the phase

The remaining key property, and the main technical challenge, is WSM within the phase (Key Property 2). In this chapter, we use a slightly different bound on the marginals from that used in the rest of the thesis, so we include the definition for completeness. Recall that for an integer $r \geq 1$ and $v \in V(C)$, $\pi_{B_r^-(v)}$ denotes the RC distribution on $B_r(v)$ with the *free* boundary condition (i.e., conditional on all edges outside $B_r(v)$ being out) and $\pi_{B_r^+(v)}$ denotes the RC distribution on $B_r(v)$ with the *wired* boundary condition (i.e., conditional on all vertices at distance exactly r from v being wired into a single component). As in the previous chapter, for an edge $e \in E_C$, we use 1_e to denote the event that e is an in-edge.

Definition 6.5. *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 $\left|\pi_C^{\text{dis}}(1_e) - \pi_{B_r^-(v)}(1_e)\right| \leq K^r$.*

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* $\left|\pi_C^{\mathrm{ord}}(1_e) - \pi_{B_r^+(v)}(1_e)\right| \leq K^r$.

We will show that for all d and then q sufficiently large, w.h.p. C has WSM within the disordered phase for all $\beta \leq \beta_1$, and WSM within the ordered phase for all $\beta \geq \beta_0$, both with rate $K = q^{-1/30}$ and radius $r = \lceil \frac{A}{d} \log n \rceil$ (where A is a constant independent of the parameters). We prove the result for the ordered phase in Section 6.4 and the result for the disordered phase in Section 6.5.

The above version of the WSM bound allows us to obtain Theorem 1.4. Recall that the goal there is to obtain, for a target $R > 0$, a $(1 \pm n^{-R})$-approximation to the partition function in time $n^{1+O(1/R)}$. The rough idea is to approximate the partition function within each phase separately using the edge marginals within that phase. WSM allows us to approximate these marginals by the corresponding marginals on balls around the edges with free or wired boundary conditions, which can be computed in time polynomial in the size of the balls. We pick d large enough so that the radius $r = \lceil \frac{A}{d} \log n \rceil$ is small enough that $|B_r(v)| = n^{O(1/R)}$, with the implicit constant made sufficiently small that any required computation polynomial in the size of $B_r(v)$ takes $O(n^{1/R})$ time. We then pick q large enough so that $K^r = q^{-r/30}$ is small enough to obtain the required approximation error $1 \pm n^{-R}$: this allows us to use $\pi_{B_r^+(v)}(1_e)$ as an additive approximation to $\pi_C^{\mathrm{ord}}(1_e)$, and likewise $\pi_{B_r^-(v)}(1_e)$ as an additive approximation to $\pi_C^{\mathrm{dis}}(1_e)$. We defer the remaining discussion and the proof of Theorem 1.4 to Section 4.5.

6.1.3.1 Ordered polymers

The main technical contribution of this chapter is a new, simple ordered polymer model which directly reflects the characterisation of the wired boundary from the previous chapter and allows us to prove WSM within the ordered phase.

Recall that polymers capture small, independent deviations from the "extreme state" (ground state) of a phase: for the ordered phase of the RC model this is the all-in configuration, while for the disordered phase it is the all-out configuration. The weight of a polymer captures how likely it is to occur; in particular, the weight of a configuration is typically proportional to the product of the weights of its polymers. For the disordered phase, the natural definition is to take polymers to be components of in-edges. However, the definition for the ordered phase of the random-cluster model is more subtle, as the long-range interactions pose a particular challenge: the weight

of a configuration F depends on its number of components, but the set of out-edges separating a component from the rest of the graph need not be connected. If several 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.

For random regular graphs, Helmuth, Jenssen and Perkins [49] overcome this problem by iteratively adding edges to the polymers and using the strong expansion of the underlying graph to obtain the desired multiplicative properties. This is the polymer framework that we used in Chapter 4.

However, this model can fail in the presence of regions with expansion at most $\frac{1}{2}$, i.e., regions for which at most half of the incident edges connect to the rest of the graph. This already poses an issue when vertices of degree 3 and 4 are present, and is especially problematic for $\mathcal{G}(n, d/n)$: while sufficiently large connected sets can be shown to have expansion greater than $\frac{1}{2}$ (for d large enough), there are small subgraphs with as many as $\Theta(\log n)$ vertices with extremely low expansion (see Figure 1.8). These subgraphs need not be captured by the expansion-based polymers, so the polymer definition of [49] does not directly extend to $\mathcal{G}(n, d/n)$.

The weaker expansion properties of $\mathcal{G}(n, d/n)$ are nevertheless sufficient to show that an ordered configuration (V, F) has a giant component and that all remaining components are small, provided that $|F|$ is sufficiently large relative to $|E|$ (see Section 6.2). This suggests an apparently natural solution: take the polymers to be the components formed by the union of the out-edges $E \setminus F$ and all edges incident to vertices in small components of (V, F). While this definition can be endowed with polymer weights satisfying the desired multiplicative properties, these polymers do not properly capture the deviations from the all-in configuration. For instance, suppose there is a connected subgraph S of (V, F) with exactly one in-edge e in the cut $(S, \overline{S})$. Then the vertices of S should belong to a polymer since the removal of a single edge e carves out S as a component, and hence affects the marginal probability that an edge in the cut $(S, \overline{S})$ is an in-edge. Moreover, if the cut $(S, \overline{S})$ is large, the event that S becomes a component after removing e is a large deviation from the all-in configuration and is thus very unlikely. Hence, the marginal probabilities of edges in the cut should not be so sensitive to the status of a single edge, suggesting that such "almost-components" must be enclosed in a polymer.

This matches the intuition developed in Chapter 5. Recall that there we proved that WSM within the ordered phase at radius r holds if every path of length r from each vertex v intersects the largest component of in-edges *after removing* v

(Section 5.4.1). In other words, we care about the components of in-edges after removing a particular vertex v, and in particular need to show that all components other than the largest are sufficiently small. It is thus natural to define the polymer edges as the union of the out-edges of F and the edges of G incident to vertices in the small components created after removing v; then the ordered polymers are connected components of polymer edges. In Section 6.4 we show that with sufficiently high probability these polymers have size $O(\log n)$, which we then use to prove WSM within the ordered phase.

6.2 Graph properties

Before proceeding to the properties of the random-cluster model, we will need to establish some properties of $\mathcal{G}(n, d/n)$ and its largest component C.

First, we show that $G \sim \mathcal{G}(n, d/n)$ is locally treelike in Section 6.2.1. Then, in Section 6.2.2, we show that while $\mathcal{G}(n, d/n)$ has variable degrees, for d large enough the average degree of connected sets of size $\Omega(\log n)$ is concentrated around d. We use this in Section 6.2.3 to establish a variant of expansion that we use in this chapter. Finally, in Section 6.2.4, we consider an auxiliary graph: the *kernel* of the largest component C, obtained by contracting all degree-2 vertices. We use it to show, in Section 6.2.4.2, that any ordered configuration induces a *giant component*, that is, a component of in-edges with more than half of the total degree of C.

To prove properties for the largest component $C = (V_C, E_C)$, it will be useful to work with $G = (V_G, E_G)$ directly. When it is necessary to distinguish between the two graphs, we use the corresponding subscript; thus, the notation below applies equally with G replaced by C. 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 $e_G(S) \coloneqq |E_G(S, S)|$ to denote the number of internal edges in S. A set $S \subseteq V_G$ is *connected* if the induced graph $G[S]$ is connected.

We will repeatedly use the following two lemmas. First, to bound the size of radius-r balls, we use the following lemma by Mossel and Sly [64]. Note that while their result is stated for $r = (\log\log n)^2$, their bound on the volume of the ball does not use this restriction on r.

Lemma 6.6 ([64, Lemma 7])**.** *Let $d > 1$ be a real. Then w.h.p. over $G \sim \mathcal{G}(n, d/n)$, for any integer r and any $v \in V(G)$, it holds that $|V_G(B_r(v))| \leq d^r \log n$.*

Second, we will use the following bound from [31] on the expected number of subtrees in $G \sim \mathcal{G}(n, d/n)$. Note that a *subtree* of G is simply a (not necessarily induced) subgraph that is a tree. Since every connected subgraph of G has a spanning tree, this immediately bounds the number of connected vertex sets.

Lemma 6.7 ([31, (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 $n(\mathrm{e}d)^s$.*

6.2.1 Treelikeness

Next we show local treelikeness properties of $\mathcal{G}(n, d/n)$. In this chapter, we say that a subgraph is *k-treelike* if at most k edges can be deleted from it to make it acyclic (i.e. a tree or a forest). First, we consider balls of radius $\approx \frac{1}{d} \log n$. The proof of the following lemma is based on the proof of [11, Lemma 2.2].

Lemma 6.8. *Let $A > 0$ be a real. Then for all d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, each radius-$\lceil \frac{A}{d} \log n \rceil$ ball in G is 1-treelike.*

Proof. Let $r = \lceil \frac{A}{d} \log n \rceil$ and let $\mathcal{A}$ be the event that some radius-r ball has more than $d^r \log n$ vertices. Pick d large enough so $d^r \log n \leq n^{A \log d/d} d \log n \leq n^{1/5}$.

Then for $v \in V_G$, let $\mathcal{B}_v$ be the event that $|V_G(B_r(v))| \leq n^{1/5}$ and that $B_r(v)$ is *not* 1-treelike. Note that we want to show $\mathbb{P}(\mathcal{A} \cup \bigcup_v \mathcal{B}_v) = o(1)$. By a union bound and Lemma 6.6, it is sufficient to show that, for each v, the probability of $\mathcal{B}_v$ is $o(1/n)$.

To do so, we adapt the breadth-first revealing argument from the proof of [11, Lemma 2.2]. Fix a vertex v and consider the following breadth-first exploration from v, stopped either after all vertices at distance at most r from v have been discovered or once $n^{1/5}$ vertices have been discovered. Initially, v is the only discovered vertex. Starting from v, for each discovered vertex x, in breadth-first order, we reveal all edges from x to vertices that have not yet been discovered. Whenever such an edge is present, its other endpoint is declared discovered. We do not reveal edges whose two endpoints have already been discovered. Thus, the present edges revealed during this procedure form a tree on the set V_B of discovered vertices, while the edges between pairs of vertices in V_B that were not revealed remain independent Bernoulli random variables with parameter d/n.

This way, we revealed a tree rooted at v, and there are at most $\binom{|V_B|}{2} \leq n^{2/5}$ unrevealed potential edges between the vertices in the ball. If $\mathcal{B}_v$ holds, then $V_B = V_G(B_r(v))$ and at least two of these unrevealed edges must be present. Thus, to prove the lemma, it suffices to conclude that the probability that there are two or more edges

between these "unrevealed ball edges" is at most $O(n^{-6/5})$. This follows from a form of Chernoff bound (see [11, 4.3]), which gives that $\mathbb{P}(\mathrm{Bin}(N,p) \geq k) \leq \mathrm{e}^{-k+Np}(\frac{k}{Np})^{-k}$. Hence,

$$\mathbb{P}(\mathcal{B}_v) \leq \mathbb{P}\left(\mathrm{Bin}\left(n^{2/5}, \tfrac{d}{n}\right) \geq 2\right) = O(n^{-6/5}),$$

which concludes the proof.

□

In the following two lemmas we show that connected sets of $O(\log n)$ size are $O(1)$-treelike, and in particular, any connected set of constant size is 1-treelike.

Lemma 6.9. *For any $d > 1$, there exists $k > 0$ such that w.h.p. over $G \sim \mathcal{G}(n, d/n)$, any connected subset of size $\leq \log n$ is k-treelike.*

Proof. Here, $k > 0$ will be a sufficiently large constant that we specify later.

Fix $1 \leq s \leq \log n$. We union-bound over subtrees of G of size s and the events "T is a subtree of G with $e(V(T)) > s - 1 + k$". Fix any tree T of size s. Then note that conditional on T being a subgraph of G, we have that $e(V(T)) \sim s - 1 + \mathrm{Bin}(\binom{s}{2} - (s-1), d/n)$, which is stochastically dominated by $s + \mathrm{Bin}(s^2, d/n)$.

We have $\mathbb{P}(\mathrm{Bin}(s^2, d/n) \geq k) \leq s^{2k}(\frac{d}{n})^k$ by the union bound. Lemma 6.7 gives that the expected number of subtrees of size s is at most $n(\mathrm{e}d)^s$, thus the expected number of subtrees T of G of size s and $> s - 1 + k$ edges between vertices of T is at most

$$\left(\frac{ds^2}{n}\right)^k n(\mathrm{e}d)^s = n\exp\{s\log(\mathrm{e}d) + k\log(\tfrac{ds^2}{n})\} \leq d^k n\exp\{\log(\mathrm{e}d)\log n - k\log(\tfrac{n}{\log^2 n})\}$$

For $k > \log(\mathrm{e}d) + 2$ and all n large enough, this is $\leq \frac{d^k}{\sqrt{n}}$. Then, we union-bound over $\log n$ possible sizes s and use Markov's inequality to obtain that the probability of a "bad subset" is at most its expectation, which is at most $d^k \log n/\sqrt{n} = o(1)$. □

Lemma 6.10. *For any $d > 1$, and an integer $R > 0$, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, any connected vertex subset of G with at most R vertices is 1-treelike.*

Proof. Let X be the number of connected vertex sets of size at most R with more than one extra edge. For any tree of size at most R, let X_T be the event that the tree T is a subgraph of G and the subgraph on its vertices $G[V(T)]$ has more than $|V(T)|$ edges. Then

$$\mathbb{E}[X] \leq \sum_{\substack{T \text{ a tree} \\ |V(T)| \leq R}} \mathbb{P}[X_T] = \sum_{T} \mathbb{P}[X_T \mid T \text{ a subgraph of } G]\mathbb{P}(T \text{ a subgraph of } G).$$

Note that the rest of the $\binom{|T|}{2} - (|T|-1) \leq R^2$ edges with both endpoints in $V(T)$ are independent of the event "T a subgraph of G", thus by the union bound (noting that the probability that $\mathrm{Bin}(R^2, d/n) > 1$ is the probability that among R^2 i.i.d. Bernoulli(d/n) random variables, at least two have value 1)

$$\mathbb{P}(X_T \mid T \text{ a subgraph of } G) \leq \mathbb{P}(\mathrm{Bin}(R^2, d/n) > 1) \leq \binom{R^2}{2}\frac{d^2}{n^2}.$$

By Lemma 6.7, the expected number of trees of size up to R is at most $nR(\mathrm{e}d)^R$. Thus, $\mathbb{E}[X]$ is at most

$$\sum_{\substack{T \text{ a tree} \\ |V(T)| \leq R}} \mathbb{P}(T \text{ a subgraph of } G)\binom{R^2}{2}\left(\frac{d}{n}\right)^2 \leq nR(\mathrm{e}d)^R \cdot \binom{R^2}{2}\left(\frac{d}{n}\right)^2 = \frac{R(\mathrm{e}d)^R d^2 \binom{R^2}{2}}{n}.$$

Since R is a constant, it follows that $\frac{R(\mathrm{e}d)^R d^2 \binom{R^2}{2}}{n} \to 0$. We conclude the lemma from Markov's inequality. □

6.2.2 Average degree of connected sets

Next, we prove that w.h.p. any connected set of $G \sim \mathcal{G}(n, d/n)$ of size $\approx \log n/d$ has its average degree close to d. We split the proof into three parts, depending on the size of the set. First, we establish the result for sets that contain at most a small fraction of vertices.

Lemma 6.11. *Let $0 < \xi < \varepsilon < 1$. Then there exists a real $A = A(\varepsilon, \xi) > 0$ such that for all d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ it holds that all connected sets of size at least $\frac{A \log n}{d}$ and at most ξn have average degree at least $(1-\varepsilon)d$.*

Proof. Fix some real $A > 0$ for now; we choose its precise value later. Since for any connected set of G, there is a subtree of G with the same vertex set, it is enough to show that

$$\sum_{\substack{T \text{ a possible subtree of } G \\ A \log n/d \leq |V_G(T)| \leq \xi n}} \mathbb{E}[\mathbb{1}\{T \text{ is a subtree of } G \text{ and } \deg_G(T) < (1-\varepsilon)d|V_G(T)|\}] = o(1),$$

and then the lemma follows from Markov's inequality.

Fix any possible tree T with $|V_G(T)| = s$. Then the term in the sum corresponding to this tree can be rewritten by conditioning on the event "T is a subtree of G" as

$$\mathbb{E}[\mathbb{1}\{T \text{ a subtree of } G\}] \cdot \mathbb{P}(\deg_G(T) < (1-\varepsilon)ds \mid T \text{ a subtree of } G).$$

Note that conditional on "T a subtree of G",

$$\deg_G(T) = 2e_G(T) + |E_G(V_G(T), \overline{V_G(T)})| \geq |E_G(V_G(T), \overline{V_G(T)})| \sim \mathrm{Bin}(s(n-s), \tfrac{d}{n}).$$

Using $n - s \geq n(1-\xi)$, we get that $|E_G(V_G(T), \overline{V_G(T)})|$ stochastically dominates $\mathrm{Bin}(sn(1-\xi), \frac{d}{n})$. Thus,

$$\mathbb{P}(\deg_G(T) < (1-\varepsilon)ds \mid T \text{ a subtree of } G) \leq \mathbb{P}(\mathrm{Bin}(s(1-\xi)n, d/n) \leq (1-\varepsilon)ds).$$

By the Chernoff bound (Lemma 2.6 with $\delta = \frac{\varepsilon-\xi}{1-\xi}$), this is at most $\exp\{-\frac{ds(\varepsilon-\xi)^2}{2(1-\xi)}\}$.

Then, we sum first over all possible subtree sizes $\frac{A\log n}{d} \leq s \leq \xi n$ and use Lemma 6.7 to bound the expected number of subtrees of size s by $n(\mathrm{e}d)^s$.

Putting it together, we obtain that the expected number of subtrees of G of size $\frac{A\log n}{d} \leq s \leq \xi n$ with total degree less than $(1-\varepsilon)ds$ (and thus average degree less than $(1-\varepsilon)d$) is at most

$$\begin{aligned}
&\sum_{s=\frac{A\log n}{d}}^{\xi n} \mathbb{E}[\#\text{subtrees of } G \text{ with } s \text{ vertices}] \exp\left\{-\frac{ds(\varepsilon-\xi)^2}{2(1-\xi)}\right\} \\
&\leq \sum_{s=\frac{A\log n}{d}}^{\xi n} n(\mathrm{e}d)^s \exp\left\{-\frac{ds(\varepsilon-\xi)^2}{2(1-\xi)}\right\}
\end{aligned}$$

For all sufficiently large d, it holds that $\frac{d(\varepsilon-\xi)^2}{2(1-\xi)} - 1 - \log d \geq \frac{d(\varepsilon-\xi)^2}{3(1-\xi)}$, and hence the sum is

$$O\left(\exp\left\{\log n\left[1 - \frac{A(\varepsilon-\xi)^2}{3(1-\xi)}\right]\right\}\right).$$

For any $A > \frac{3(1-\xi)}{(\varepsilon-\xi)^2}$, this is $o(1)$, hence we conclude the lemma by Markov's inequality. □

We next show the average-degree bound for all sets of size between ξn and $(1-\xi)n$, as well as a lower bound on the size of the resulting cut. The latter result will be useful when establishing (weak) expansion in the next section. Note that for the next two lemmas, we do not require the sets to be connected.

Lemma 6.12. *Let $\xi, \varepsilon \in (0,1)$ be arbitrary constants. Then, for all $d > \frac{32\log 2}{(\varepsilon\xi)^2}$, w.h.p. 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 $|E_G(S, \overline{S})| \geq \frac{d}{n}(1-\varepsilon)|S|(n-|S|)$.*

Proof. For any given set S with size $\xi n \leq |S| \leq (1-\xi)n$, it holds that $\deg_G(S)$ is a sum of two independent random variables: $2\mathrm{Bin}(\binom{|S|}{2}, \frac{d}{n})$ and $\mathrm{Bin}(|S|(n-|S|), \frac{d}{n})$, corresponding to inner edges and the edges going outside of S, respectively. If $2e_G(S) \geq (1-\varepsilon)d|S|^2/n$ and $|E_G(S, \overline{S})| \geq (1-\varepsilon)d|S|(n-|S|)/n$, then $\deg_G(S) \geq (1-\varepsilon)d|S|$. Thus, it is sufficient to upper-bound the probability that one of these events will not happen.

First, we bound $\mathbb{P}(2\mathrm{Bin}(\binom{|S|}{2}, \frac{d}{n}) < \frac{(1-\varepsilon)d|S|^2}{n})$. Let n be large enough so that $(1-\varepsilon)/(\xi n - 1) < \varepsilon/2$ and $n \geq 2/\xi$ hold. Then, by the Chernoff bound (Lemma 2.6, using $\mathbb{E}[\mathrm{Bin}(\binom{|S|}{2}, d/n)] \geq dn\xi^2/4$ and $\delta \geq \varepsilon/2$), this probability is at most $\exp\{-\frac{d\varepsilon^2\xi^2 n}{32}\}$.

Then, for the contribution to the total degree corresponding to the edges going from S to $\overline{S}$, this is $\sim \mathrm{Bin}(|S|(n-|S|), \frac{d}{n})$, and by the Chernoff bound (using $\delta = \varepsilon$ and $\mathbb{E}[\mathrm{Bin}(|S|(n-|S|), d/n)] = d|S|(n-|S|)/n \geq d\xi^2 n$ in Lemma 2.6),

$$\mathbb{P}(\mathrm{Bin}(|S|(n-|S|), \tfrac{d}{n}) < (1-\varepsilon)\frac{d|S|(n-|S|)}{n}) \leq \exp\{-d\varepsilon^2\xi^2 n/2\}.$$

Hence, by the union bound, for any such set S, the probability that either its average degree is less than $(1-\varepsilon)d$ or $|E_G(S, \overline{S})| < (1-\varepsilon)d|S|(n-|S|)/n$ is at most $\exp\{-d\varepsilon^2\xi^2 n/2\} + \exp\{-d\varepsilon^2\xi^2 n/32\}$. There are at most 2^n such sets S, thus by another union bound and by the condition on d, we obtain that the probability that a set with size between ξn and $(1-\xi)n$ has either average degree less than $(1-\varepsilon)d$ or induces a cut of size less than $\frac{(1-\varepsilon)d}{n}|S|(n-|S|)$ is at most $\exp\{-\Omega(n)\} = o(1)$. □

Finally, we prove the average-degree bound for all sufficiently large sets.

Lemma 6.13. *For $\xi, \varepsilon \in (0,1)$ with $\xi < \varepsilon/2$, and all $d > \frac{8\log 2}{(\varepsilon - 2\xi)^2}$, w.h.p. over $G \sim \mathcal{G}(n, d/n)$,* all *vertex sets $S \subseteq V(G)$ of size at least $(1-\xi)n$ have $\deg_G(S) \geq (1-\varepsilon)d|S|$.*

Proof. For any vertex set S, it holds that $\deg_G(S) \geq 2e_G(S) \sim 2\mathrm{Bin}(\binom{|S|}{2}, d/n)$. If $|S| \geq (1-\xi)n$, then $\mathbb{E}[\mathrm{Bin}(\binom{|S|}{2}, d/n)] \geq \frac{d|S|(|S|-1)}{2n} \geq d|S|(1-\xi-1/n)/2$. For $n \geq 1/\xi$, this is at least $d|S|(1-2\xi)/2 > (1-\varepsilon)d|S|/2$. By the Chernoff bound (Lemma 2.6 with $\delta \geq \frac{\varepsilon - 2\xi}{1-2\xi}$) we get

$$\mathbb{P}(e_G(S) \leq (1-\varepsilon)d|S|/2) \leq \exp\left\{-\frac{d|S|(\varepsilon-2\xi)^2}{8(1-2\xi)}\right\} \leq \exp\left\{-dn(\varepsilon-2\xi)^2/8\right\},$$

where the second inequality uses that $\frac{1-\xi}{1-2\xi} \geq 1$.

Thus, union-bounding over at most 2^n possible sets, the probability that there exists a vertex set S with $|S| \geq (1-\xi)n$ and $\deg_G(S) \leq (1-\varepsilon)d|S|$ is at most $\exp\{n[\log 2 - d(\varepsilon - 2\xi)^2/8]\}$, and by the lower bound on d, this is $\exp\{-\Omega(n)\} = o(1)$. □

Combining these three lemmas, we get the following proposition.

Proposition 6.14. *For any $\varepsilon \in (0,1)$, there exists $A = A(\varepsilon) > 0$ such that for all d large enough, w.h.p. 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 $(1-\varepsilon)d$.*

Proof. Let $\xi \coloneqq \frac{\varepsilon}{3} < \frac{\varepsilon}{2}$. Then the proposition follows directly by applying Lemmas 6.11, 6.12 and 6.13. □

6.2.3 Weak expansion

Next, we prove a *weak expansion* property for $G \sim \mathcal{G}(n, d/n)$, that is, we show that for any *connected vertex set* S of size $\approx \frac{\log n}{d}$, the number of edges in the cut $E_G(S, \overline{S})$ is proportional to $|S|$. As a first step, we show that the number of internal edges of connected subsets is not too large.

Lemma 6.15. *For any $\xi \in (0,1]$, there exists a real $A = A(\xi) > 0$ such that for all d large enough, w.h.p. 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. Fix some real $A > 0$ for now; we choose its precise value later. Since every connected set has a spanning tree, it is enough to bound the probability that there exists a subtree T of G, with size as in the lemma, for which $e_G(V_G(T)) \geq (2\xi d/3 + 1)|V_G(T)|$. By Markov's inequality, it suffices to show that the expected number of such subtrees is $o(1)$. For any "potential subtree" T, the probability it is a subtree of G and $e_G(V_G(T)) \geq (2\xi d/3+1)|V_G(T)|$ is at most

$$\mathbb{P}(e_G(V_G(T)) \geq (2\xi d/3+1)|V_G(T)| \mid T \text{ a subtree of } G) \cdot \mathbb{P}(T \text{ a subtree of } G)$$

Fix a size $\frac{A\log n}{d} \leq s \leq \xi n$ and consider the expected number of subtrees with s vertices and at least $(2\xi d/3+1)s$ inner edges. Fix a tree T with s vertices. Then conditional on T being a subgraph of G, it holds that $e_G(V(T)) \sim (s-1) + \mathrm{Bin}(\binom{s}{2} - (s-1), \frac{d}{n})$. Since $\binom{s}{2} - (s-1) \leq s^2/2$, we can bound

$$\mathbb{P}(e_G(V_G(T)) \geq (2\xi d/3+1)s \mid T \text{ a subtree of G}) \leq \mathbb{P}(\mathrm{Bin}(\tfrac{s^2}{2}, \tfrac{d}{n}) \geq \tfrac{2\xi ds}{3}),$$

Noting that $s \leq \xi n$, this is at most $\mathbb{P}(\mathrm{Bin}(\frac{\xi sn}{2}, \frac{d}{n}) \geq \frac{2\xi ds}{3})$. The Chernoff bound (Lemma 2.6 with $\delta \geq \frac{1}{6}$ and $\mu = \xi sd/2$) gives that this is at most $\exp\{-\xi ds/216\}$.

By Lemma 6.7, the expected number of subtrees of size s is at most $n(\mathrm{e}d)^s$. Hence, summing over possible subtree sizes, we get that the expected number of "bad" subtrees is at most

$$\sum_{\frac{A\log n}{d} \leq s \leq \xi n} n[d\mathrm{e}^{1-\xi d/216}]^s.$$

For all d large enough, $d\mathrm{e}^{1-\xi d/216} < \exp\{-\xi d/250\}$, and thus the sum is $O(n^{1-A\xi/250})$, which is $o(1)$ for any $A > \frac{250}{\xi}$. □

With this lemma in hand, we can now prove the *weak expansion* property for all connected sets with size $\Omega(\log n/d)$.

Proposition 6.16. *There exists a constant $A > 0$, such that for all d large enough, w.h.p. 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 $E_G(S, \overline{S}) \geq \frac{3}{5}\deg_G(S)$. 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. By Proposition 6.14 and Lemma 6.15, there exists a real $A > 0$ such that for all d large enough, w.h.p., any connected vertex set S with size at least $\frac{A\log n}{d}$ has $\deg_G(S) \geq \frac{9d|S|}{10}$, and if furthermore $|S| \leq \frac{n}{6}$, then $e_G(S) \leq (1+\frac{d}{9})|S|$.

Since $\deg_G(S) = 2e_G(S) + |E_G(S, \overline{S})|$, we can rearrange and then use the above inequalities to obtain that for any connected vertex set S with $\frac{A\log n}{d} \leq |S| \leq \frac{n}{6}$,

$$\frac{|E_G(S,\overline{S})|}{\deg_G(S)} = 1 - \frac{2e_G(S)}{\deg_G(S)} \geq 1 - \frac{20}{9d} - \frac{20}{81}.$$

For all d large enough, this is at least $\frac{3}{5}$. To finish the proof, note that whenever $|S| \geq \frac{n}{6}$, its total degree is at least $\frac{3dn}{20}$. But w.h.p., $\frac{3}{4}nd > |E_G| \geq |E_C|$, giving $\frac{\deg_G(V_C)}{10} < \frac{3nd}{20} \leq \deg_G(S)$. Thus, whenever S has $\deg_G(S) \leq \deg_G(V_C)/10$, then also $|S| \leq n/6$, thus the expansion applies. □

We remark that the choice of 3/5 in Proposition 6.16 is arbitrary, and any constant between 1/2 and 1 could be used. Also, note that the lower bound $A\log(n)/d$ in Propositions 6.16 and 6.14 is asymptotically optimal in n and d, as it can be shown that G contains a path of only degree-2 vertices of length $\approx \frac{\log n}{d}$.

6.2.4 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 the unique maximal subgraph of C with minimum degree at least 2 and then contracting the degree-2 vertices – satisfies the expansion property for *all small sets*. We then use the expansion of the kernel to show that any sufficiently large edge subset gives rise to a unique *giant component*. This is the key property of configurations in the ordered phase which we will use later to prove Theorem 6.2 as well as WSM within the ordered phase. Note that the vertices and edges of C outside its unique maximal subgraph of minimum degree at least 2 form a forest[3], which we refer to as the *attached trees* in C.

We will denote the kernel as $G_K = (V_K, E_K)$ and write $n_K \coloneqq |V_K|$. To reason about the kernel and show it is expanding, we are going to use the following contiguous model from [26].

Definition 6.17. *Let $d > 1$ be a constant and $\mu \in (0, 1)$ be its* conjugate *(i.e. the unique solution in $(0, 1)$ to $d\mathrm{e}^{-d} = \mu\mathrm{e}^{-\mu}$). Let $n > 1$ be an integer. Consider the graph model $\mathbb{K}(n, d)$ generated as follows:*

Let $\Delta \sim \mathcal{N}(d - \mu, 1/n)$ and then let $D_1, \ldots, D_n$ be i.i.d. Poisson(Δ) random variables, conditional on $\sum_i^n D_i \mathbb{1}_{D_i \geq 3}$ being 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 uniformly at random.

The following proposition tells us that any w.h.p. property for $\mathbb{K}(n, d)$ is a w.h.p. property for the kernel of $\mathcal{G}(n, d/n)$.

Proposition 6.18 ([26, Theorem 1])**.** *Let $d > 1$ be a constant and $\mu \in (0, 1)$ be its* conjugate*. Let $n > 1$ be an integer.*

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

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

Lemma 6.19. *For all $d \geq 3$, it holds that $\mu < \mathrm{e}^{-d/2}$.*

Proof. The function $x \mapsto x\mathrm{e}^{-x}$ is (strictly) decreasing on $x \in (1, \infty)$ and (strictly) increasing on $x \in (0, 1)$, thus there is a unique conjugate for each $d > 1$ in $(0, 1)$ and in particular, note that if for $a \in (0, 1)$ it holds that $a\mathrm{e}^{-a} > d\mathrm{e}^{-d}$, then $\mu(d) < a$. We can easily verify that $\mathrm{e}^{-d/2}\mathrm{e}^{-\mathrm{e}^{-d/2}} > d\mathrm{e}^{-d}$ holds for all $d \geq 3$. □

[3] A forest is an acyclic graph, i.e. a collection of trees.

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 6.20. *For $d \geq 4$, w.h.p. over $K \sim \mathbb{K}(n,d)$, it holds that $n_K \geq n(1-\mathrm{e}^{-d/3})$. Thus also w.h.p. over $G \sim \mathcal{G}(n,d/n)$, it holds that $n_K \geq n(1-\mathrm{e}^{-d/3})$.*

Proof. In the model from Proposition 6.18, we have $\Delta \sim \mathcal{N}(d-\mu, 1/n)$, hence w.h.p. $\Delta \geq d - \mathrm{e}^{-d/2}$ (we used the inequality $\mu < \mathrm{e}^{-d/2}$ from the Lemma 6.19). From now on, assume this is the case.

Then note $n - n_K = \sum_i^n \mathbb{1}\{D_i < 3\}$. Note that while the model is conditioned on $\sum_i^n D_i \mathbb{1}\{D_i \geq 3\}$ being even, this event has a probability bounded below by a constant; thus, it is enough to show that w.h.p. a sum of n i.i.d. Bernoulli($\mathbb{P}(\mathrm{Poisson}(\Delta) < 3)$) r.v.s[4] is at most $n\mathrm{e}^{-d/3}$. Since we assumed $\Delta \geq d - \mathrm{e}^{-d/2}$ and $\mathbb{P}(\mathrm{Poisson}(\lambda) < 3)$ is decreasing in λ, by the definition of Poisson random variable we have

$$\mathbb{P}(\mathrm{Poisson}(\Delta) < 3) \leq \mathrm{e}^{-d+\mathrm{e}^{-d/2}}(1 + d - \mathrm{e}^{-d/2} + (d - \mathrm{e}^{-d/2})^2/2).$$

We can verify that for all $d \geq 4$, there exists a constant (independent of n) $\delta = \delta(d) > 0$ such that the mean of the sum of these n i.i.d. Bernoulli r.v.s is $\leq n\mathrm{e}^{-d/3} - n\delta$. By the Chernoff bound (Lemma 2.6) this implies w.h.p. $n - n_K \leq n\mathrm{e}^{-d/3}$. The result for $G \sim \mathcal{G}(n,d/n)$ follows from the contiguity of $\mathbb{K}(n,d)$ (Proposition 6.18). □

Lemma 6.21. *For any real $d \geq 4$, w.h.p. over $K \sim \mathbb{K}(n,d)$, it holds that $\frac{dn}{2} - 2n\mathrm{e}^{-d/3} \leq |E_K|$. Thus also w.h.p. over $G \sim \mathcal{G}(n,d/n)$, it holds that $\frac{dn}{2} - 2n\mathrm{e}^{-d/3} \leq |E_K| \leq |E_C|$*

Proof. We have $\Delta \sim \mathcal{N}(d-\mu, 1/n)$, thus w.h.p. $d - \mu - \frac{1}{2}\mathrm{e}^{-d/3} < \Delta$. For $d \geq 3$, it holds that $\mu \leq \mathrm{e}^{-d/2} < \mathrm{e}^{-d/3}$, hence $\Delta > d - \frac{3}{2}\mathrm{e}^{-d/3}$.

From the previous lemma, w.h.p. $n - n_K \leq n\mathrm{e}^{-d/3}$. Also, we can write $2|E_K| \geq \sum_i^n D_i - 2(n - n_K)$. It is therefore enough to show that w.h.p. $\sum_i^n D_i \geq n(d - 2\mathrm{e}^{-d/3})$. Since the probability of $\sum_i^n D_i \mathbb{1}\{D_i \geq 3\}$ being even is bounded below by a constant, it is enough to show that w.h.p. $\mathrm{Poisson}(\Delta n) \geq n(d - 2\mathrm{e}^{-d/3})$. Since $\Delta \geq d - \frac{3}{2}\mathrm{e}^{-d/3}$ w.h.p., we have from the Poisson Chernoff bound (Lemma 2.7) that $\mathbb{P}(\mathrm{Poisson}(\Delta n) \leq n(d - 2\mathrm{e}^{-d/3})) = \mathrm{e}^{-\Omega(n)}$. The result for $\mathcal{G}(n,d/n)$ follows from contiguity of $\mathbb{K}(n,d)$. □

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

Lemma 6.22. *For all d large enough, w.h.p. over $G \sim \mathcal{G}(n,d/n)$, it holds that $|E_C| \leq (1 + \mathrm{e}^{-d/3})|E_K|$.*

[4] Random variables.

Proof. We have that w.h.p. $|E_C| \leq \frac{dn}{2} + n\mathrm{e}^{-d/3}$ (since $|E_C| \leq |E_G| = \frac{nd}{2} + o(n)$) and $|E_K| \geq \frac{dn}{2} - 2n\mathrm{e}^{-d/3}$ (by Lemma 6.21). Thus,

$$\frac{|E_C|}{|E_K|} \leq \frac{d/2 + \mathrm{e}^{-d/3}}{d/2 - 2\mathrm{e}^{-d/3}} = 1 + \frac{3\mathrm{e}^{-d/3}}{d/2 - 2\mathrm{e}^{-d/3}}.$$

For all d large enough, $\frac{3}{d/2-2\mathrm{e}^{-d/3}} \leq 1$, hence giving that $|E_C|$ is at most $(1 + \mathrm{e}^{-d/3})|E_K|$. □

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 more than 2 in G and are part of the unique maximal subgraph of minimum degree at least 2 of C). Note that $\deg_K(\cdot)$, $E_K(\cdot,\cdot)$ are defined analogously to $\deg_G(\cdot)$, $E_G(\cdot,\cdot)$.

Proposition 6.23. *For all real d large enough, w.h.p. 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 $2n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$. Hence then for any vertex set $S_G \subseteq V_G$*

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

Proof. We first claim that w.h.p., the number of vertices of degree $\leq (1 - d^{-1/3})d$ in the *kernel* is at most $o(n) + n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$. Note that it is enough, by Proposition 6.18, to show this for $K \sim \mathbb{K}(n, d)$.

There, $\Delta \sim \mathcal{N}(d - \mu, 1/n)$, and hence w.h.p. $\Delta \geq d - \mathrm{e}^{-d/3}$ (recall that by Lemma 6.19, $\mu \leq \mathrm{e}^{-d/2}$ for all $d \geq 3$). Hence, from now on we assume $\Delta \geq d - \mathrm{e}^{-d/3}$.

The number of kernel vertices with degree $\leq (1 - d^{-1/3})d$ is at most $\sum_i^n \mathbb{1}\{D_i \leq (1 - d^{-1/3})d\}$. Since we are proving a w.h.p. property, we can drop the conditioning on $\sum_i^n D_i \mathbb{1}\{D_i \geq 3\}$ being even, so it is enough to bound the sum of n i.i.d. $\mathbb{1}\{\text{Poisson}(d - \mathrm{e}^{-d/3}) < (1 - d^{-1/3})d\}$ r.v.s. Since $d - \mathrm{e}^{-d/3} > (1 - d^{-1/3})d$, we use the Poisson Chernoff bound (Lemma 2.7) to get

$$\mathbb{P}\Big(\text{Poisson}(d - \mathrm{e}^{-d/3}) \leq (1 - d^{-1/3})d\Big) \leq \exp\left(-\tfrac{(d^{2/3} - \mathrm{e}^{-d/3})^2}{d + d^{2/3} - 2\mathrm{e}^{-d/3}}\right) \leq \mathrm{e}^{-\frac{1}{2}d^{1/3}},$$

for all d large enough. Then using a Chernoff bound for the sum of Bernoulli r.v.s, w.h.p. $\sum_i^n \mathbb{1}\{D_i \leq (1 - d^{-1/3})d\} \leq o(n) + n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$, which finishes the claim.

Now, all vertices in G that correspond to kernel vertices with degree at least $(1 - d^{-1/3})d$ also have degree at least $(1 - d^{-1/3})d$. It remains to account for the

remaining vertices. The number of these vertices is $n - n_K \leq n\mathrm{e}^{-d/3}$, where the last inequality follows w.h.p. from Lemma 6.20. Thus the number of vertices with degree $\leq (1-d^{-1/3})d$ in G (and thus in C) is w.h.p. at most $n(\mathrm{e}^{-d/3} + o(1) + \mathrm{e}^{-\frac{1}{2}d^{1/3}})$, which is at most $2n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$ for all d large enough.

For the second part of the proposition, we first note that since every vertex in the kernel-representation of a set S_K has at most the degree of the corresponding vertex in S_G, it is the case that $\deg_G(S_G) \geq \deg_K(S_K)$.

Then, we show that there are enough high-degree vertices in S_K. We have $|S_K| \geq |S_G| - (n - n_K) \geq |S_G| - n\mathrm{e}^{-d/3}$, and all but $n\mathrm{e}^{-\frac{1}{2}d^{1/3}} + o(n)$ of these vertices have degree at least $(1 - d^{-1/3})d$. Thus, we get that for all d large enough,

$$\deg_K(S_K) \geq (1-d^{-1/3})d|S_G| - dn[\mathrm{e}^{-d/3} + o(1) + \mathrm{e}^{-\frac{1}{2}d^{1/3}}] \geq (1-d^{-1/3})d|S_G| - 2dn\mathrm{e}^{-\frac{1}{2}d^{1/3}}.$$

This, together with $\deg_G(S_G) \geq \deg_K(S_K)$ concludes the proof. □

6.2.4.1 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 [36], which gives expansion properties of small sets w.h.p. over random graphs with fixed degree sequences with a) all degrees between 3 and $n^{1/50}$, and b) the sum of degrees squared being $O(n)$. Note that in $K \sim \mathbb{K}(n, d)$, w.h.p. 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, w.h.p. the sum of squares of degrees of $K \sim \mathbb{K}(n, d)$ is dominated by a sum of squares of n independent Poisson($2d$) r.v.s (we use that w.h.p. $\Delta \sim \mathcal{N}(d - \mu, 1/n) \leq 2d$ and drop the parity conditioning, as justified earlier in the section). By Chebyshev's inequality, with probability $1 - O(1/n)$, that sum is at most $8(d^2 + d)n$, which is (w.h.p.) linear in n_K (Lemma 6.20); thus we can apply their result as follows.

Lemma 6.24 ([36, Lemma 15])**.** *W.h.p. over $K \sim \mathbb{K}(n, d)$, for any connected* kernel *vertex set $S \subseteq V_K$ with $|S| \leq \log^2 n_K$, it holds that $|E_K(S, \overline{S})| \geq \deg_K(S)/144$.*

We note that while [36] gives an expansion property for *all* sets of size $\leq n_K/2$, their established fraction of edges going out on sets of larger size is lower 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 6.25. *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 $|E_K(S_K, \overline{S_K})| = |E_G(S_G, \overline{S_G})|$.*

Proof. For any $e \in E_G(S_G, \overline{S_G})$, there is a corresponding $e' = \{v, w\} \in E_K$ such that e is an edge on the path corresponding to e' (as clearly, e cannot be an edge going to an attached tree, since all edges in a tree attached to vertices of S_K in C are in S_G), and an endpoint of e must be a kernel vertex – in particular it is also an endpoint of e', say $v \in S_K$. But then if $e' \notin E_K(S_K, \overline{S_K})$, it must be the case $v, w \in S_K$. Then all vertices on the path corresponding to e' are in S_G, which implies $e \notin E_G(S_G, \overline{S_G})$. Contradiction. Thus $e' \in E_K(S_K, \overline{S_K})$.

Next, we call the path in G corresponding to an edge of G_K its *expanded path*. Then note that there cannot be two edges $e_1 \neq e_2 \in E_G(S_G, \overline{S_G})$ with the same corresponding kernel edge: since S_G is the representation of S_K in C, both e_1, e_2 must have an endpoint in S_K. But these endpoints cannot be identical, as then $e_1 = e_2$ (as they are on a single expanded path). Thus they must lie at opposite ends of the same expanded path. But then, both endpoints of this path are in S_K, so all vertices on this path are in S_G, contradicting that e_1, e_2 are in a cut.

Therefore $|E_K(S_K, \overline{S_K})| \geq |E_G(S_G, \overline{S_G})|$.

For the converse, if $e \in E_K(S_K, \overline{S_K})$, say $e = \{v, w\}$, has $v \in S_K$, then the first edge e' on the expanded path corresponding to e in G going from v is in $E_G(S_G, \overline{S_G})$. This mapping $e \mapsto e'$ is an injection, as it has a unique inverse, since contracting an expanded path gives a unique edge in G_K. □

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

Proposition 6.26. *For all d large enough, w.h.p. 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 $\deg_K(S) \leq \frac{1}{2}\deg_K(G_K)$, it holds that $|E_K(S, \overline{S})| \geq \frac{1}{144}\deg_K(S)$.*

Proof. It is sufficient to show that for any connected set of total degree $\leq \frac{1}{2}\deg_K(G_K)$, the number of edges in the cut is at least a $\frac{1}{144}$-fraction of the total degree. Then, to handle arbitrary (not necessarily connected) vertex sets, combine the results for their connected components.

By Lemma 6.20, w.h.p. $n_K \geq n(1 - \mathrm{e}^{-d/3})$, thus w.h.p. $\log^2 n_K \geq \log n$. Then w.h.p., by contiguity of $\mathbb{K}(n, d)$ and Lemma 6.24, for any $|S_K| \leq \log n$, it holds that $|E_K(S_K, \overline{S_K})| \geq \deg_K(S_K)/144$. Also assume that d is large enough so that

all previously proven w.h.p. results apply, in particular: Propositions 6.14 (with $\varepsilon = \frac{1}{10}$), and 6.16, and 6.15 (with $\xi = 2/3$). Further on, we assume G is such that the properties from these results hold.

Let $S_G \subseteq V_G$ be the set as in Observation 6.25. We divide the proof into four cases, based on the size of S_K and S_G. Note that since S_K is connected, S_G is also connected.

1. If $|S_K| \leq \log n$, then Lemma 6.24 gives $|E_K(S_K, \overline{S}_K)| \geq \deg_K(S_K)/144$.

2. If $\log n \leq |S_K| \leq |S_G| \leq \frac{n}{6}$ (note that the middle inequality always holds), then by Proposition 6.16 and the above observation we get

$$|E_K(S_K, \overline{S_K})| = |E_G(S_G, \overline{S_G})| \geq \frac{3}{5}\deg_G(S_G) \geq \frac{3}{5}\deg_K(S_K),$$

   where the last inequality follows from $\deg_K(S_K) \leq \deg_G(S_G)$.

3. If $\frac{n}{6} < |S_G| \leq \frac{2n}{3}$, then Lemma 6.15 gives $e_G(S_G) \leq (\frac{4}{9}d + 1)|S_G|$, and Proposition 6.14 gives $\deg_G(S_G) \geq \frac{9d}{10}|S_G|$. Thus,

$$\frac{|E_K(S_K, \overline{S_K})|}{\deg_K(S_K)} \geq \frac{|E_G(S_G, \overline{S_G})|}{\deg_G(S_G)} = 1 - \frac{2e_G(S_G)}{\deg_G(S_G)} \geq 1 - \frac{80}{81} - \frac{20}{9d},$$

   where the first inequality follows from $\deg_K(S_K) \leq \deg_G(S_G)$ and the second from the above inequalities. To finish, note that for all d large enough, this is at least $\frac{1}{144}$.

4. Finally, we rule out the case $|S_G| \geq \frac{2n}{3}$.

   By Proposition 6.23, for all d large enough,

$$\deg_K(S_K) \geq (1-d^{-1/3})d|S_G| - 2dn\mathrm{e}^{-\frac{1}{2}d^{1/3}} \geq \frac{2nd}{3}\left[1 - d^{-1/3} - 3\mathrm{e}^{-\frac{1}{2}d^{1/3}}\right] \geq \frac{3dn}{5}.$$

   But also (w.h.p.) $|E_K| \leq |E_G| \leq nd/2 + o(n)$, so for any such set S_K, it holds that $\deg_K(S_K) > \frac{1}{2}\deg_K(G_K)$, hence we can rule this case out. □

6.2.4.2 Giant component after edge removal

Finally, in this section, we will use the expansion of the kernel to prove that after removing a small constant fraction of edges from C, there remains a large component. We adapt the proof of the same property for regular graphs (Lemma 4.24 originally from [68]), first for degree-expander graphs and then for C.

Lemma 6.27 (Analogue of [68, Lemma 2.3])**.** *Let $H = (V_H, E_H)$ be a α-total degree expander, i.e. for any vertex set $S \subseteq V_H$ with $\deg_H(S) \leq \frac{1}{2}\deg_H(H)$, it holds that $E_H(S, \overline{S}) \geq \alpha \deg_H(S)$. 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 $\left(1 - \frac{\theta}{2\alpha}\right)\deg_H(V_H)$.*

Proof. Let $C_1, \ldots, C_k$ be the components of $(V_H, E_H \setminus E')$ and without loss of generality, let C_1 be the component with the highest total degree. We distinguish two cases.

Case 1. $\deg_H(C_1) < \frac{\deg_H(V_H)}{2} = |E_H|$. Then for all $1 \leq i \leq k$, it holds that $\deg_H(C_i) < \frac{\deg_H(V_H)}{2}$, and hence by the expansion property of H, we get $|E'| \geq \frac{1}{2}\sum_{i=1}^{k} \alpha \deg_H(C_i) = \alpha|E_H|$. However, $|E'| \leq \theta|E_H| < \alpha|E_H|$. Contradiction: this case cannot happen.

Case 2. $\deg_H(C_1) \geq \frac{\deg_H(V_H)}{2} = |E_H|$. Then it follows that $\deg_H(C_2 \cup \cdots \cup C_k) \leq \deg_H(V_H)/2$, and thus we can use the expansion property of H to get

$$|E'| \geq \alpha \deg_H(C_2 \cup \cdots \cup C_k) = \alpha(\deg_H(V_H) - \deg_H(C_1)).$$

This gives $\deg_H(C_1) \geq \deg_H(V_H) - |E'|/\alpha \geq (1 - \frac{\theta}{2\alpha})\deg_H(V_H)$. □

Since w.h.p. the kernel is a $\frac{1}{144}$-total-degree expander, we can apply this lemma to 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 the following lemma.

Lemma 6.28. *For all d large enough, w.h.p. 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 $|E'| \leq \frac{|E_C|}{144}\left(2\theta - \frac{1}{100}\right)$. Then $(V_C, E_C \setminus E')$ has a component with total degree at least $\deg_G(V_C)(1-\theta)$, and all other components of $(V_C, E_C \setminus E')$ have total degree at most $\theta \deg_G(V_C)$.*

Proof. Let E'_K be the set of edges in the kernel whose corresponding expanded path (as in Observation 6.25) in G contains *at least one* edge from E'. Notice this does not increase the number of edges, i.e. $|E'_K| \leq |E'|$. Also, for all d large enough, $\frac{1}{100} \geq 2\mathrm{e}^{-d/3}$, which gives $|E'| \leq \frac{|E_C|}{144}(2\theta - 2\mathrm{e}^{-d/3})$. Then, from Lemma 6.22, $|E_C| \leq (1 + \mathrm{e}^{-d/3})|E_K|$, thus

$$|E'_K| \leq \frac{|E_K|}{144}(2\theta - 2\mathrm{e}^{-d/3})(1 + \mathrm{e}^{-d/3}) < \frac{|E_K|}{144}.$$

Hence we can apply Lemma 6.27 to get that $(V_K, E_K \setminus E'_K)$ contains a component of total degree at least

$$\begin{aligned}(1-(\theta-\mathrm{e}^{-d/3})(1+\mathrm{e}^{-d/3}))\deg_K(V_K) &\geq \frac{1-(\theta-\mathrm{e}^{-d/3})(1+\mathrm{e}^{-d/3})}{1+\mathrm{e}^{-d/3}}\deg_G(V_C)\\ &\geq (1-\theta)\deg_G(V_C),\end{aligned}$$

To finish, we note that vertices of this component are also in the same component of $(V_C, E_C \setminus E')$; their degrees do not decrease, and hence there is a component of $(V_C, E_C \setminus E')$ with total degree at least $(1-\theta)\deg_G(V_C)$. Since the total degree of C is $\deg_G(V_C)$, any of the remaining components can have total degree at most $\theta \deg_G(V_C)$. □

In light of Lemma 6.28, 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.

In particular, by applying the lemma for $\theta = 1/10$, we have that for any $F \in \Omega_C^{\mathrm{ord}}$, the subgraph (V_C, F) contains a giant component and all the other components are small.

6.3 Phases of the random-cluster model

In this section, we prove Theorem 6.2 and thus justify our definition of the phases. The proof is split into five lemmas. First, we show that the ordered phase becomes exponentially unlikely for all $\beta \leq \beta_0$.

Lemma 6.29. *Let d be a sufficiently large real. For all sufficiently large real q, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, it holds that $Z_C^{\mathrm{ord}}/Z_C \leq \mathrm{e}^{-n}$ for all $\beta \leq \beta_0$.*

Proof. By Lemmas 6.12, and 6.13, w.h.p. all vertex sets in G with size at least ηn have average degree at least $\frac{3d}{4}$. Also by the second part of Lemma 6.12, w.h.p., for any vertex set $S \subseteq V_G$ with size between ηn and $(1-\eta)n$, it holds that $|E_G(S, \overline{S})| \geq \frac{5d}{6n}|S|(n-|S|)$. Also w.h.p. $n_C \geq (1-\mathrm{e}^{-d/3})n$. Also, w.h.p. $|E_C| \leq \frac{nd}{2} + o(n) \leq \frac{nd}{2} + \eta n$. Assume, henceforth, that G satisfies these properties.

Also by the choice of η and by Lemma 6.28, w.h.p. G is such that for any $F \in \Omega_C^{\mathrm{ord}}$, there is a giant component in (V_C, F) with total degree at least $\frac{9}{10}\deg_G(V_C)$. Thus, the set S of vertices in the non-giant components has total degree $\leq \frac{1}{10}\deg_G(V_C)$.

Note that all edges in $E_G(S,\overline{S})$ are in $E_C \setminus F$. We consider three cases, based on the size of S.

If $|S| \geq \frac{n}{3}$, then $\deg_G(S) \geq \frac{dn}{4}$. But also $\deg_G(S) \leq \frac{\deg_G(V_C)}{10} \leq \frac{dn}{10} + o(n)$, which is a contradiction.

Next consider the case $\eta n \leq |S| \leq \frac{n}{3}$. Then $|E_C \setminus F| \geq |E_G(S,\overline{S})| \geq \frac{5\eta nd}{9}$. But w.h.p. $|E_C| \leq \frac{dn}{2} + o(n) < \frac{5}{9}dn$, so this contradicts $|E_C \setminus F| \leq \eta|E_C|$.

Thus the only remaining possibility is $|S| \leq \eta n$. Then $c(F) \leq |S| + 1 \leq \eta n + 1$ and $|F| \leq |E_C| \leq nd/2 + \eta n$. Thus we obtain that

$$w_C(F) \leq q^{\eta n+1}(\mathrm{e}^\beta - 1)^{nd/2+\eta n} \leq q^{\eta n + \frac{2-1/10}{d}\cdot\frac{nd}{2}+\eta n} = q^{n[1-1/20+2\eta]},$$

where the second inequality follows from $\beta \leq \beta_0$ and since for all $d \geq 3$ and n large enough it holds that $1 + \frac{19}{10d}\eta n \leq \eta n$.

Since $F = \emptyset$ is a configuration with a weight $q^{n_C} \geq q^{n(1-\mathrm{e}^{-d/3})}$, it follows that $Z_C \geq q^{n-n\mathrm{e}^{-d/3}}$. There are at most $2^{|E_C|} \leq 2^{nd/2+\eta n}$ ordered configurations. Therefore,

$$\frac{Z_C^{\mathrm{ord}}}{Z_C} \leq \frac{2^{nd/2+\eta n} q^{n[1-1/20+2\eta]}}{q^{n[1-\mathrm{e}^{-d/3}]}} = \exp\{n[(d/2+\eta)\ln 2 + \log q(-1/20 + 2\eta + \mathrm{e}^{-d/3})]\}.$$

Since $1/20 - 2\eta - \mathrm{e}^{-d/3} > 1/30$ for all d large enough, this is at most $\exp\{n(d/2 + \eta)\ln 2 - n\log q/30\}$, which is at most e^{-n} for all q large enough. □

Analogously, the contribution of the disordered phase becomes exponentially small for all $\beta \geq \beta_1$.

Lemma 6.30. *Let d be a sufficiently large real. For all sufficiently large real q, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ it holds that $Z_C^{\mathrm{dis}}/Z_C \leq \mathrm{e}^{-n}$ for all $\beta \geq \beta_1$.*

Proof. Let $F \in \Omega_C^{\mathrm{dis}}$. Then $|F| \leq \eta|E_C|$, and thus $w_C(F) \leq q^{n_C}(\mathrm{e}^\beta - 1)^{\eta|E_C|}$. Also $|\Omega_C^{\mathrm{dis}}| \leq 2^{|E_C|}$, which is w.h.p. at most $2^{n(d/2+\eta)}$. Since $F = E_C$ is a configuration, $Z_C \geq q(\mathrm{e}^\beta - 1)^{|E_C|}$. Combining the inequalities and using that $\beta \geq \beta_1$, we obtain

$$\frac{Z_C^{\mathrm{dis}}}{Z_C} \leq 2^{(d/2+\eta)n} q^n (\mathrm{e}^\beta - 1)^{-(1-\eta)|E_C|} \leq 2^{(d/2+\eta)n} q^{n - \frac{(1-\eta)(2+1/10)}{d}|E_C|}.$$

Lemma 6.21 gives us that w.h.p. $|E_C| \geq \frac{dn}{2} - 2n\mathrm{e}^{-d/3}$, thus we get that $\frac{Z_C^{\mathrm{dis}}}{Z_C}$ is at most

$$2^{(\frac{d}{2}+\eta)n} q^{n[1-\frac{1}{d}(1-\eta)(2+\frac{1}{10})(\frac{d}{2}-2\mathrm{e}^{-\frac{d}{3}})]} \leq \exp\left\{n\left[(\tfrac{d}{2}+\eta)\ln 2 - \ln q\left(\tfrac{1}{20} - 2\eta - \tfrac{1}{d}5\mathrm{e}^{-\frac{d}{3}}\right)\right]\right\}.$$

By the choice of $\eta = 1/1000$, for all d large enough this is at most $2^{dn}q^{-n/40}$. Hence, for all q large enough, it is at most e^{-n}. □

It remains to show that the configurations with at least $\frac{9\eta}{10}|E_C|$ and at most $(1-\frac{9\eta}{10})|E_C|$ edges are sufficiently unlikely. First we show this for the unconditional RC distribution.

Lemma 6.31. *Let d be a sufficiently large real. For all sufficiently large real q, the following holds w.h.p. 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 $\frac{9}{10}\eta \le \frac{|F|}{|E_C|} \le 1 - \frac{9}{10}\eta$. Then, $\pi_C(\mathcal{B}) \le \mathrm{e}^{-n}$.

Proof. The main part of the proof is to show that for any configuration $F \in \mathcal{B}$ it holds that

$$c(F) - \frac{2}{d}|E_C \setminus F| \le -\rho n \quad \text{with} \quad \rho = \eta/20. \tag{6.2}$$

Let us conclude the proof of the lemma assuming (6.2) for the moment. We first show that (6.2) translates into $w_C(F)/Z_C \le q^{-\rho n/4}$ for all $\beta > 0$.

Indeed, if $\mathrm{e}^\beta - 1 \le q^{2/d}$, we will use the lower bound $|Z_C| \ge q^{n_C} \ge q^{n(1-\mathrm{e}^{-d/3})}$, where the second inequality holds with high probability. Thus, for any $F \in \mathcal{B}$,

$$\frac{w_C(F)}{Z_C} = q^{c(F)-n+n\mathrm{e}^{-d/3}}(\mathrm{e}^\beta - 1)^{|E_C|-|E_C\setminus F|} \le q^{c(F)-\frac{2}{d}|E_C\setminus F|-n+\frac{2}{d}|E_C|+n\mathrm{e}^{-d/3}}.$$

Using that w.h.p. $|E_C| \le |E_G| \le nd/2 + \frac{\eta n}{40}$, we have that $\frac{2}{d}|E_C| - n \le n - n + \frac{\eta n}{20}$. Hence, for all d sufficiently large (recall that $\eta = 1/1000$ is independent of d), we obtain that $w_C(F)/Z_C \le q^{-\rho n/4}$.

Analogously, if $\mathrm{e}^\beta - 1 \ge q^{2/d}$, we use the lower bound $Z_C \ge (\mathrm{e}^\beta - 1)^{|E_C|}$ so that for any $F \in \mathcal{B}$,

$$\frac{w_C(F)}{Z_C} = q^{c(F)}(\mathrm{e}^\beta - 1)^{-|E_C\setminus F|} \le q^{c(F)-\frac{2}{d}|E_C\setminus F|} \le q^{-\rho n}.$$

To finish, note that the size of $\mathcal{B}$ is trivially at most $|\Omega_C| \le 2^{|E_C|} \le 2^{nd/2+\eta n/40}$. Thus we get

$$\frac{1}{Z_C}\sum_{F\in\mathcal{B}} w_C(F) \le 2^{nd/2+\eta n/40} q^{-\rho n/4},$$

which is $\le \mathrm{e}^{-n}$ for all q large enough.

We next return to the proof of (6.2). Fix an arbitrary $F \in \mathcal{B}$. We partition the vertices of C into sets S_1, S_2, S_3, S_4 depending on the size of their components in (V_C, F). Namely, we set S_1 to be the set of vertices in singleton components, S_2 the set of vertices in components of size between 2 and $\log n$, S_3 the set of vertices in components of size more than $\log n$ and at most $\frac{n}{6}$, and S_4 the set of vertices in

components of size more than $\frac{n}{6}$. For $i = 1, \ldots, 4$, let $n_i \coloneqq |S_i|$. We have then the trivial upper bound

$$c(F) \leq n_1 + \frac{n_2}{2} + \frac{n_3}{\log n} + 6. \tag{6.3}$$

Next we give a lower bound on $|E_C \backslash F|$ by analysing how many out-edges are incident to vertices in S_1, S_2, S_3, S_4, treating each case separately. We use the separate bounds for vertex sets in different size ranges proven in Section 6.2; namely, for vertices in S_3 and S_4 we can use the weak expansion, while for vertices in S_2 we use that their components must be treelike.

1. Consider first components of size 1. Every edge adjacent to any $v \in S_1$ must be in $E_C \setminus F$. Thus, aggregating over $v \in S_1$, we have at least $\frac{1}{2}\deg_G(S_1)$ edges in $E_C \setminus F$ adjacent to vertices in S_1; note that the $\frac{1}{2}$-factor accounts for potential double-counting (edges that join two vertices in S_1).

2. Let k be a constant such that w.h.p. all sets of size at most $\log n$ are k-treelike (Lemma 6.9). Also, by Lemma 6.10, w.h.p. any connected set of size at most $2k$ is 1-treelike. We thus further split this case and consider components with less than $2k$ vertices and those with at least $2k$ vertices separately. Consider now any component κ whose size is between 2 and $\log n$.

   If κ has size less than $2k$, then it is 1-treelike, so $e_G(\kappa) \leq |V_G(\kappa)|$. This gives

   $$|E_G(V_G(\kappa), \overline{V_G(\kappa)})| \geq \deg_G(\kappa) - 2|V_G(\kappa)|.$$

   If κ has size between $2k$ and $\log n$, then it is k-treelike, so $e_G(\kappa) \leq k + |V_G(\kappa)| \leq \frac{3}{2}|V_G(\kappa)|$, yielding that

   $$|E_G(V_G(\kappa), \overline{V_G(\kappa)})| \geq \deg_G(\kappa) - 3|V_G(\kappa)|.$$

   Summing over all components κ in S_2, we therefore obtain at least $\frac{1}{2}(\deg_G(S_2) - 3n_2)$ edges in $E_C \backslash F$ that are adjacent to vertices in S_2 (once again, the factor $\frac{1}{2}$ accounts for double-counting).

3. Consider next components with size between $\log n$ and $\frac{n}{6}$. By Proposition 6.16 and by taking d large enough (namely, $d > A$ where A is the absolute constant from the proposition), we have that w.h.p. every connected subset S in G with size $\log n \leq |S| \leq \frac{n}{6}$ satisfies $|E_G(S, \overline{S})| \geq \frac{3}{5}\deg_G(S)$. Therefore, for any component κ of (V_C, F) in this size range, we have at least $\frac{3}{5}\deg_G(\kappa)$ edges going out of it, and therefore belonging to $E_C \backslash F$. Aggregating over κ (and

accounting once again for potential double-counting), we obtain that the total number of out-edges adjacent to vertices in S_3 is at least $\frac{1}{2} \cdot \frac{3}{5} \deg_G(S_3)$.

4. Finally consider components with size more than $\frac{n}{6}$. If $\eta n/4 \leq |S_4| \leq n(1 - \eta/4)$, we get, from the second part of Lemma 6.12, that for all d large enough $|E_G(S_4, \overline{S}_4)| \geq \frac{4d|S_4|(n-|S_4|)}{5n} \geq \frac{d\eta n}{6}$ (since $\eta/4 \leq \frac{1}{30}$). Therefore, the total number of out-edges adjacent to vertices in S_4 is at least $\frac{d\eta n}{6} \cdot \mathbb{1}\{\eta n/4 \leq |S_4| \leq (n - \eta n/4)\}$. Note that if $|S_4|$ is outside the range $\eta n/4$ and $n(1 - \eta/4)$ we use the trivial lower bound 0.

Combining the above, we conclude that

$$\begin{aligned} |E_C \backslash F| &\geq \tfrac{1}{2}\big[\deg_G(S_1 \cup S_2) - 3n_2 + \tfrac{3}{5}\deg_G(S_3) + \tfrac{d\eta n}{6} \cdot \mathbb{1}\{\eta n/4 \leq |S_4| \leq (1-\eta/4)n\}\big] \\ &\geq \tfrac{1}{2}\big[(1 - d^{-1/3})d(n_1 + n_2 + \tfrac{3}{5}n_3 - 2n\mathrm{e}^{-\frac{1}{2}d^{1/3}}) + \tfrac{d\eta n}{6} \cdot \mathbb{1}\{\eta n/4 \leq |S_4| \leq (1-\eta/4)n\}\big], \end{aligned} \tag{6.4}$$

where the second inequality follows from the fact that the number of vertices of G which have degree less than $(1 - d^{-1/3})d$ is less than $2n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$ (Proposition 6.23). From (6.3) and (6.4), we therefore have that

$$\begin{aligned} c(F) - \tfrac{2}{d}|E_C \setminus F| \leq n_1 + (\tfrac{1}{2} + \tfrac{3}{d})n_2 + \tfrac{n_3}{\log n} + 6 - (1 - d^{-1/3})\left[n_1 + n_2 + \tfrac{3}{5}n_3\right] \\ + 2n\mathrm{e}^{-\frac{1}{2}d^{1/3}} - \tfrac{\eta n}{6}\mathbb{1}\{\eta n/4 \leq |S_4| \leq (n - \eta n/4)\}. \end{aligned}$$

For d and n sufficiently large, we have that $\frac{1}{2} - \frac{3}{d} - d^{-1/3} \geq \frac{1}{4}$ and $\frac{3}{5} - \frac{1}{\log n} - \frac{3d^{-1/3}}{5} \geq \frac{1}{4}$, so

$$c(F) - \frac{2}{d}|E_C \setminus F| \leq \frac{n_1}{d^{\frac{1}{3}}} - \frac{n_2 + n_3}{4} - \frac{\eta n}{6} \cdot \mathbb{1}\{\tfrac{\eta n}{4} \leq |S_4| \leq (n - \tfrac{\eta n}{4})\} + 2n\mathrm{e}^{-\frac{1}{2}d^{\frac{1}{3}}} + 6. \tag{6.5}$$

To conclude (6.2), we split the analysis into four cases: first, we will rule out that there are too many singleton components. Second, we consider the case where all but a small fraction of vertices are in linear-sized components. There, we can apply a stronger bound than (6.5). The remaining two cases assume neither of these holds and are distinguished based on the value of $|S_2| + |S_3|$. For both of these cases we use (6.5).

We will use that w.h.p. $|E_C| \geq |E_K| \geq \frac{dn}{2} - 2n\mathrm{e}^{-d/3} > \frac{25}{27}\frac{dn}{2}$ for all d large enough (Lemma 6.21), and also that w.h.p. $|E_C| \leq |E_G| \leq \frac{nd}{2} + \frac{\eta n}{40}$.

Case 1. $\frac{n_1}{n} \geq 1 - \frac{2\eta}{3}$.

Here we show that $|F| < \frac{9}{10}\eta|E_C|$, contradicting our assumption on $|F|$ coming from $F \in \mathcal{B}$ (and hence we can rule out this case). Indeed, no edge incident to a

vertex in S_1 is in F, so $|E_C \setminus F| \geq \frac{1}{2}\deg_G(S_1) \geq \frac{1}{2}d(1-d^{-1/3})(n_1 - 2n\mathrm{e}^{-\frac{1}{2}d^{1/3}})$ where the last inequality follows from the fact that G has at most $2n\mathrm{e}^{-\frac{1}{2}d^{1/3}}$ vertices with degree less than $(1-d^{-1/3})d$ (Proposition 6.23). Hence

$$|F| = |E_C| - |E_C \setminus F| \leq \frac{dn}{2} + \frac{\eta n}{40} - \frac{dn}{2}[(1-\tfrac{2\eta}{3})(1-d^{-1/3}) - 2\mathrm{e}^{-\frac{1}{2}d^{1/3}}].$$

For all d large enough, $\frac{\eta}{40d} + \frac{2\eta}{3} + d^{-1/3} + 2\mathrm{e}^{-\frac{1}{2}d^{1/3}} \leq \frac{5\eta}{6}$, thus also $|F| \leq \frac{dn}{2} \cdot \frac{5\eta}{6}$. But $|E_C| > \frac{25}{27}\frac{dn}{2}$, which gives $|F| < \frac{9}{10}\eta|E_C|$, contradicting $F \in \mathcal{B}$.

Case 2. $n_4 > n(1-\eta/2)$.

Since we have at most 6 components with size more than $n/6$, in this case it holds that $c(F) < n - n(1-\eta/2) + 6 = \eta n/2 + 6$. Since $|F| \leq (1-\frac{9\eta}{10})|E_C|$ we have $|E_C \setminus F| \geq \frac{9}{10}\eta|E_C| > \frac{5\eta}{12}dn$, so for n large enough

$$c(F) - \frac{2}{d}|E_C \setminus F| < 6 + \frac{\eta n}{2} - \frac{10}{12}\eta n = 6 - \frac{\eta n}{3} \leq -\frac{\eta n}{20}.$$

Case 3. $n_2 + n_3 \geq \eta n/3$.

Plugging this into (6.5) (using the trivial bound $n_1 < n$ and using the trivial upper bound 0 for the indicator variable), we obtain that

$$c(F) - \frac{2}{d}|E_C \setminus F| \leq \tfrac{n}{d^{1/3}} - \frac{\eta n}{12} + 2n\mathrm{e}^{-\frac{1}{2}d^{1/3}} + 6$$

Since η does not depend on d, for all d (and n) large enough, $d^{-1/3} + 2\mathrm{e}^{-\frac{1}{2}d^{1/3}} + 6/n \leq \eta/30$, and so for all n large enough, we get (6.2).

Case 4. $n_2 + n_3 < \eta n/3$, $n_1 < n(1-2\eta/3)$, and $n_4 < n(1-\eta/2)$.

Since $n_1 + n_2 + n_3 + n_4 = |V_C|$, we have that $n_4 \geq n_C - n_1 - n_2 - n_3 \geq n - n\mathrm{e}^{-d/3} - \frac{\eta n}{3} - n + \frac{2\eta n}{3}$ which is at least $\eta n/4$ for all d large enough. In this case, (6.5) gives for d and n large that

$$c(F) - \frac{2}{d}|E_C \setminus F| \leq \tfrac{n}{d^{1/3}} - \frac{\eta n}{6} + 2n\mathrm{e}^{-\frac{1}{2}d^{1/3}} + 6 \leq -\eta n/20.$$

This finishes the proof of (6.2) and completes the proof of Lemma 6.31. □

We now prove an analogous result when we replace the unconditional RC distribution by π_C^{dis} and π_C^{ord}.

Lemma 6.32. *Let d be a sufficiently large real. For all sufficiently large real q, w.h.p. 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, $\pi_C^{\mathrm{dis}}(\mathcal{B}_1) \leq \mathrm{e}^{-n}$.*

Proof. We use that w.h.p. (using Lemma 6.21), $\frac{dn}{3} \le |E_C| \le nd/2 + \eta n$. Consider arbitrary $F \in \mathcal{B}_1 \cap \Omega_C^{\text{dis}}$. Let S be the set of vertices that are in components of size at least 2. It suffices to show that for all sufficiently large d and then q, it holds that $\frac{w_C(F)}{Z_C^{\text{dis}}} \le \frac{w_C(F)}{w_C(\emptyset)} \le q^{-\eta n/30}$. Since there are at most $2^{|E_C|} \le 2^{nd/2+\eta n}$ configurations in $\mathcal{B}_1$, for all q large enough, we get that summing over all $F \in \mathcal{B}_1$, the total contribution is at most e^{-n}.

First, we consider a case that there is a component in (V_C, F) with at least $2\eta n$ vertices. Then, $c(F) \le n_C - 2\eta n + 1$. Since $F \in \Omega_C^{\text{dis}}$, it is the case that $|F| \le \eta |E_C| \le \eta(nd/2 + \eta n)$. Since $\beta \le \beta_1$, we also get $(\mathrm{e}^\beta - 1)^{|F|} \le q^{21\eta n(1/2+\eta/d)/10}$. Hence,

$$\frac{w_C(F)}{w_C(\emptyset)} \le q^{1+n_C-n_C-2\eta n}(\mathrm{e}^\beta - 1)^{\eta|E_C|} \le q^{1-\eta n(2-21(1/2+\eta/d)/10)} \le q^{-\eta n/30},$$

where the last inequality follows for all d and n sufficiently large.

It remains to consider the case where all components of S in (V_C, F) are of size at most $2\eta n$. Then, since all components of vertices of S have at least two vertices, it holds that $c(F) \le n_C - |S| + |S|/2 = n_C - |S|/2$. Thus, the goal will be to lower-bound the size of S. We do it by showing that the number of internal edges in the components of S is relatively small, forcing $|S|$ to be large compared to the number of in-edges (since the only in-edges of F are between vertices of S). By Lemma 6.15, for all sufficiently large d, there is a constant A such that any connected set S' of size at least $\frac{A}{d}\log n$ and at most $2\eta n$ satisfies $e_G(S') \le (1 + 4\eta d/3)|S'|$. For smaller components, we use treelikeness. Let k be such that, w.h.p., any connected set of size at most $\frac{A}{d}\log n$ is k-treelike (Lemma 6.9). Then, by Lemma 6.10, all connected sets of size at most k are 1-treelike. As a result, for all connected sets S' of size at most $\frac{A}{d}\log n$, $e_G(S') \le 2|S'|$. As a result, if S' are the vertices of a component of (V_C, F), then $e_G(S') \le 2|S'|(1 + 2\eta d/3)$.

Note that the edges of F must be internal edges in some component of size at least 2. Summing over all of these components, we get $|F| \le 2|S|(1 + 2\eta d/3)$.

Since also $|F| \ge \frac{9\eta}{10}|E_C|$, we must have $|S| \ge \frac{9\eta|E_C|}{20d(2\eta/3+1/d)}$. Combining with the bound for $c(F)$, we obtain

$$\frac{w_C(F)}{w_C(\emptyset)} = q^{n_C - \frac{9\eta|E_C|}{40d(2\eta/3+1/d)} - n_C}(\mathrm{e}^\beta - 1)^{\eta|E_C|} \le q^{-\frac{\eta|E_C|}{10d}(\frac{9}{4(2\eta/3+1/d)} - 21)},$$

where the last inequality follows from $\beta \le \beta_1$. For all d large enough $\frac{9}{4(2\eta/3+1/d)} - 21 \ge 1$. Since $|E_C| \ge \frac{dn}{3}$, we get that $\frac{w_C(F)}{w_C(\emptyset)} \le q^{-n\eta/30}$, which finishes the proof. □

Lemma 6.33. *Let d be a sufficiently large real. For all sufficiently large real q, w.h.p. 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, $\pi_C^{\mathrm{ord}}(\mathcal{B}_0) \leq \mathrm{e}^{-n}$.

Proof. Consider arbitrary $F \in \mathcal{B}_0$. First of all, we showed in the proof of Lemma 6.29, for all d large enough, w.h.p. G is such that for any ordered F, there are at most ηn vertices in small components. Hence in particular $c(F) \leq 1 + \eta n$. Let S be the set of vertices in non-giant components of (V_C, F). We now bound $w_C(F)/Z_C^{\mathrm{ord}}$ by considering two cases.

If $c(F) - 1 \leq \eta n/2$, then using $Z_C^{\mathrm{ord}} \geq (\mathrm{e}^{\beta} - 1)^{|E_C|}$ and $\beta \geq \beta_0$ we get

$$\frac{w_C(F)}{Z_C^{\mathrm{ord}}} \leq \frac{w_C(F)}{w_C(E_C)} \leq q^{\eta n/2}(\mathrm{e}^{\beta} - 1)^{-|E_C \setminus F|} \leq q^{\eta n/2 - \frac{19}{10d}\frac{9\eta}{10}|E_C|}.$$

w.h.p. $|E_C| \geq \frac{dn}{2} - 2n\mathrm{e}^{-d/3} \geq \frac{9}{19}dn$ for all d large enough, thus giving $\frac{w_C(F)}{Z_C^{\mathrm{ord}}} \leq q^{\eta n(1/2 - 81/100)} \leq q^{-\eta n/10}$.

For the remaining case, assume $\eta n/2 < c(F) - 1 \leq \eta n$ (this is exhaustive since we noted that $c(F) - 1 \leq \eta n$). Hence also $|S| \geq c(F) - 1 > \eta n/2$. By the second part of the Lemma 6.12, for all sufficiently large d, w.h.p. G is such that any S (note that we do not require S to be connected), $|E_G(S, \overline{S})| \geq \frac{5d}{6n}(n - |S|)|S| \geq \frac{5d(1-\eta)}{6}(c(F) - 1)$. Noting that all edges in the cut $E_G(S, \overline{S})$ are out-edges of F, we get that

$$\frac{w_C(F)}{Z_C^{\mathrm{ord}}} \leq \frac{w_C(F)}{w_C(E_C)} \leq q^{c(F)}(\mathrm{e}^{\beta} - 1)^{-\frac{5d(1-\eta)}{6}c(F)} \leq q^{(c(F)-1)[1 - \frac{5\cdot 19}{6\cdot 10}(1-\eta)]},$$

where the last inequality follows from $\beta \geq \beta_0$. Next, using that $c(F) - 1 > \eta n/2$ and $\eta = 1/1000$, we obtain that this is at most $q^{-\eta n/10}$.

To finish, we take the $q^{-\eta n/10}$ bound from both cases and bound the aggregate sum over at most $2^{|E_C|} \leq 2^{dn}$ such configurations. For all q sufficiently large, this gives $\pi_C^{\mathrm{ord}}(\mathcal{B}_0) \leq \mathrm{e}^{-n}$. □

Now Theorem 6.2 follows directly from these five lemmas: Lemma 6.31 proves Item 1. Items 2 and 3 follow from Lemmas 6.33 and 6.30 respectively, and, finally Items 4 and 5 follow from Lemmas 6.29-6.31.

6.4 Weak spatial mixing within the ordered phase

In this section, we prove WSM within the ordered phase. As we showed in Chapter 5, the difference in marginals on e in the WSM within the ordered phase (Definition 6.5)

can be bounded by a sum of two terms: first, the probability that there are too few in-edges outside of the ball, and second, the probability that there is a long path avoiding the giant component of in-edges after removing v.

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 C intersect the largest component of the graph $(V_C, F)\backslash v$, 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, combining Lemmas 5.18 and 5.19, we have that

$$\left|\pi_{B_r^+(v)}(1_e) - \pi_C^{\mathrm{ord}}(1_e)\right| \leq 2\pi_C^{\mathrm{ord}}(\neg\mathcal{A}_{v,r}) + \frac{2\pi_C^{\mathrm{ord}}(\neg\mathcal{A}'_{v,r})}{1-\pi_C^{\mathrm{ord}}(\neg\mathcal{A}'_{v,r})}. \tag{6.6}$$

We note that the event $\mathcal{O}$ (our event $\mathcal{A}'_{v,r}$) in Lemma 5.18 was defined for a different threshold. However, the same proof applies here, since that we only need that conditional on $\mathcal{O}$, any resulting configuration is ordered.

It is relatively easy to bound the probability of $\neg\mathcal{A}'_{v,r}$ under π_C^{ord}. For large enough d, Lemma 6.6 gives that w.h.p. 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 $|E_C| = \Omega(n)$ w.h.p.. Hence, by Theorem 6.2(2), $\pi_C^{\mathrm{ord}}(\neg\mathcal{A}'_{v,r}) \leq \mathrm{e}^{-n}$.

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

6.4.1 Ordered Polymers

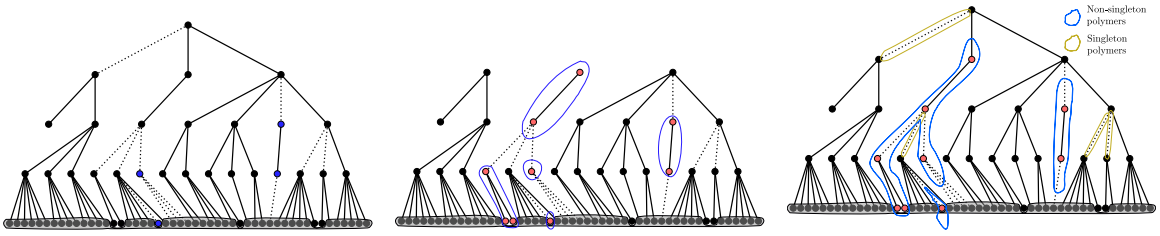
We want, for each vertex v, and for each $r \sim \log(n)/d$, to bound the quantity $\pi_C^{\mathrm{ord}}(\neg\mathcal{A}_{v,r})$. To do so, we need to control the size of "polymers" with respect to a configuration F and a vertex $v \in V_C$.

We first consider non-giant components in $(V_C, F)\backslash v = (V_C \setminus \{v\}, \{e \in F : e \text{ is not incident to } v\})$; 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; see Figure 6.1 for illustration.

Definition 6.34 (Ordered polymers)**.** *Fix $v \in V_C$. Consider $F \in \Omega_C^{\mathrm{ord}}$. For any $w \in V_C$, define the graph $H_v(F, w)$ as follows.*

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

Let $\mathcal{V}(F, v) \coloneqq \{w \in V_C \mid \deg_G(H_v(F, w)) \leq \frac{1}{2}\deg_G(V_C)\}$. We define two types of ordered polymers*, non-singleton polymers and singleton polymers. In both cases, an ordered polymer is a non-empty subgraph of C.*



(a) $H_v(F, v)$, the component containing v in (V_C, F), consists of nodes marked in black.

(b) $H_v(F, w)$, the components in $(V_C, F) \setminus \{v\}$ for $w \neq v$. Small components are denoted in blue and consist of orange vertices, while the unique giant component is formed of black nodes.

(c) The resulting ordered polymers. The non-singleton polymers are marked in blue, and the singleton polymers are marked in yellow. The vertices in $\mathcal{V}(F, v)$ are marked in orange.

Figure 6.1: Illustration of the construction of ordered polymers with respect to v. In all three panels, solid lines represent in-edges of F, while dashed lines represent out-edges.

- *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 polymer *of F consisting of the single edge e.*

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

As we showed in Lemma 6.28, there is a unique giant component (containing a majority of the total degree of C) in (V_C, F) for all $F \in \Omega_C^{\mathrm{ord}}$, while all remaining components are small and their vertices are in $\mathcal{V}(F, v)$. By the above definition, every small component in (V_C, F) is a subgraph of some polymer. This justifies the following definition of polymer weights.

Definition 6.35. *Fix $v \in V_C$. Fix $F \in \Omega_C^{\mathrm{ord}}$ and $\gamma = (U, B) \in \Gamma_v^{\mathrm{ord}}(F)$. The* inner vertices *of γ are $\mathcal{V}_\gamma := U \cap \mathcal{V}(F, v)$. The* out-edges *of γ are the edges in $E_{\mathrm{out}}(\gamma) := (E_C \setminus F) \cap B$ and $e_{\mathrm{out}}(\gamma) = |E_{\mathrm{out}}(\gamma)|$ is the number of out-edges of γ. The weight of γ is $w_C^{\mathrm{ord}}(\gamma) := q^{c'(\gamma)}(\mathrm{e}^\beta - 1)^{-e_{\mathrm{out}}(\gamma)}$, where $c'(\gamma)$ is the number of components of $(V_C, E_C \setminus E_{\mathrm{out}}(\gamma)))$ with total degree at most $(1/2)\deg_G(V_C)$.*

Next, Lemma 6.36 relates the definition of the polymers to the event $\mathcal{A}_{v,r}$. In particular, it says that to bound $\pi_C^{\mathrm{ord}}(\neg\mathcal{A}_{v,r})$, it suffices to bound the probability that there is a polymer of size r or more.

Hence, this will be the focus of the remainder of this section. We show how polymer weights relate to the weight of an ordered configuration and that we can bound the probability of a particular polymer occurring in an ordered configuration by its weight. Finally, we conclude the section by proving that the sum of weights of all large enough ordered polymers is sufficiently small, and thus so is the probability that some large polymer occurs.

Lemma 6.36. *Fix $F \in \Omega_C^{\mathrm{ord}}$. 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 $\gamma \in \Gamma_{v_0}^{\mathrm{ord}}(F)$ with $\{v_1, \dots, v_r\} \subseteq \mathcal{V}_\gamma$, hence $|\mathcal{V}_\gamma| \geq r$.*

Proof. Consider $i \in \{1, 2, \dots, r\}$. Since $v_i \neq v_0$, it holds that $H_{v_0}(F, v_i)$ is the component of v_i in $(V_C, F) \setminus v_0$. By assumption, v_i is not in the largest component of $(V_C, F) \setminus v_0$ and thus $H_{v_0}(F, v_i)$ is not the largest component in $(V_C, F) \setminus v_0$, so $\deg_G(H_{v_0}(F, v_i)) \leq \frac{1}{2} \deg_G(V_C)$ by Lemma 6.28. Now note that the $v_1, \dots, v_r$ form a path in C, so the vertices form a connected set in C. By Definition 6.34, there is $\gamma \in \Gamma_{v_0}^{\mathrm{ord}}(F)$ with $\{v_1, \dots, v_r\} \subseteq \mathcal{V}_\gamma$. □

Hence the goal is to show that with sufficiently high probability, there is no polymer γ with $|\mathcal{V}_\gamma| \geq r$.

First we use Lemma 6.28 to show that the ordered polymers are indeed well-defined and it will not happen that after removing v, all vertices are in $\mathcal{V}(F, v)$. In particular, we show that all ordered polymers have at most a small fraction of the total degree of C.

Lemma 6.37. *For all d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, the following holds. Fix any $v \in V_C$. For any $F \in \Omega_C^{\mathrm{ord}}$, it is the case that $(V_C, F) \setminus v$ contains a component of total degree at least $(1 - \frac{1}{10}) \deg_G(V_C)$. Hence in particular, every polymer $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$ has $\deg_G(\mathcal{V}_\gamma) \leq \frac{1}{10} \deg_G(V_C)$.*

Proof. W.h.p., the maximum degree of G is $O(\frac{\log n}{\log \log n}) \leq (\frac{1/5 - 1/100}{144} - \eta)|E_C|$ (see e.g. [32, Theorem 3.4][5]). Since $|F| \geq (1-\eta)|E_C|$ and $\eta = 1/1000$, it follows that the graph $(V_C, F) \setminus v$ is obtained from C by removing at most $\frac{|E_C|}{144}(\frac{2}{10} - \frac{1}{100})$, so by Lemma 6.28, it has a component κ of total degree at least $(1 - \frac{1}{10}) \deg_G(V_C)$. By Definition 6.34,

[5] This also follows directly from the Chernoff bound: for $t = \frac{C \log n}{\log \log n}$, it holds that $\mathbb{P}(\mathrm{Bin}(n - 1, d/n) \geq t) \leq (ed/t)^t = n^{-C + o(1)}$. Thus, taking $C > 1$ sufficiently large and union-bounding over all n vertices gives the claim.

none of the vertices in κ is in $\mathcal{V}(F,v)$. Thus, $\deg_G(\mathcal{V}(F,v)) \le \frac{1}{10}\deg_G(V_C)$, and so in particular any component of $\mathcal{V}(F,v)$ has total degree at most $\frac{1}{10}\deg_G(V_C)$.

Finally, we note that for any singleton polymer $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$, from Definition 6.35, $\mathcal{V}_\gamma = \emptyset$. □

We use Lemma 6.37 to relate the weight of a configuration to the product of the weights of its polymers.

Lemma 6.38. *For all d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, the following holds. For any $F \in \Omega_C^{\mathrm{ord}}$ and $v \in V_C$,*

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

Proof. Write

$$q(e^\beta - 1)^{|E_C|} \prod_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} w_C^{\mathrm{ord}}(\gamma) = q^{1+\sum_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} c'(\gamma)} (e^\beta - 1)^{|E_C| - \sum_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} e_{\mathrm{out}}(\gamma)}.$$

Every out-edge of F is in exactly one ordered polymer of F, thus $|E_C \setminus F| = \sum_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} e_{\mathrm{out}}(\gamma)$. It remains to show that $c(F) = 1 + \sum_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} c'(\gamma)$.

By Lemma 6.37, w.h.p. G is such that for all $F \in \Omega_C^{\mathrm{ord}}$, it is the case that (V_C, F) contains a giant component with total degree more than $\frac{1}{2}\deg_G(V_C)$. The 1-term in the above equality corresponds to the giant component. Now consider any non-giant component κ of (V_C, F). Its vertices must be in $\mathcal{V}(F,v)$, and thus in $\mathcal{V}_\gamma$ for some $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$, hence also $E_C(V(\kappa), \overline{V(\kappa)}) \subseteq E_{\mathrm{out}}(\gamma)$. Since κ is connected in (V_C, F), it must be connected in $(V_C, E_C \setminus E_{\mathrm{out}}(\gamma))$, as $F \subseteq E_C \setminus E_{\mathrm{out}}(\gamma)$, thus it is accounted for in $c'(\gamma)$. Therefore $c(F) \le 1 + \sum_{\gamma \in \Gamma_v^{\mathrm{ord}}(F)} c'(\gamma)$.

Now consider κ to be a non-giant component of $(V_C, E_C \setminus E_{\mathrm{out}}(\gamma))$ for some $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$. Since (V_C, E_C) is connected, $E_{\mathrm{out}}(\gamma)$ contains an out-edge e incident to some vertex w of κ. Also $F \subseteq E_C \setminus E_{\mathrm{out}}(\gamma)$, hence the component of w in (V_C, F), call it κ', must be a subgraph of κ. Clearly, κ' must also be a non-giant component, thus all of its vertices are in $\mathcal{V}(F,v)$. Since non-singleton polymers are maximal connected subgraphs consisting of vertices in $\mathcal{V}(F,v)$, there is a unique polymer γ' containing vertices in $V(\kappa') \subseteq V(\kappa)$. Since e is an out-edge incident to w, then $e \in E_{\mathrm{out}}(\gamma')$. Thus e is not in a singleton polymer in $\Gamma_v^{\mathrm{ord}}(F)$, so γ must be a non-singleton polymer with an endpoint of e contained in $\mathcal{V}_\gamma$. But then $\mathcal{V}_\gamma \cup \{w\}$ is a connected subset in (V_C, E_C), and thus in particular so is $\mathcal{V}_\gamma \cup \mathcal{V}_{\gamma'}$. Since polymers of F have pairwise disjoint sets of inner vertices and edges and γ is a polymer of F with $\mathcal{V}_\gamma \cap \mathcal{V}_{\gamma'} \neq \emptyset$, it must be the case that $\gamma = \gamma'$ and so also $\kappa = \kappa'$. Thus each non-giant component

κ of $(V_C, E_C \setminus E_{\text{out}}(\gamma))$ for any $\gamma \in \Gamma_v^{\text{ord}}(F)$ is also a non-giant component of (V_C, F) and for any $\gamma' \in \Gamma_v^{\text{ord}}(F)$ with $\gamma' \neq \gamma$, it holds that κ is not a non-giant component of $(V_C, E_C \setminus E_{\text{out}}(\gamma'))$. This finishes the second part of the equality and thus the proof. □

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. Finally, we bound the number of possible polymers and sum their weights.

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 $\mathcal{V}_\gamma$. Since, for any singleton polymer γ it holds that $\mathcal{V}_\gamma = \emptyset$, the same statement applies. Formally, we have the following.

Observation 6.39. *Let $F \in \Omega_C^{\text{ord}}$, let $v \in V_C$ and let $\gamma \in \Gamma_v^{\text{ord}}(F)$. 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 by Lemma 6.8, for all sufficiently large d, w.h.p. every radius-$\frac{A}{d} \log n$ ball is 1-treelike.

We use this to show that $E_G(v, \mathcal{V}_\gamma)$ accounts for a small fraction of edges going from $\mathcal{V}_\gamma$.

Lemma 6.40. *For all real d large enough, w.h.p. 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. Let $r := \frac{\log n}{d}$. By Lemma 6.8, for all d sufficiently large, w.h.p. any radius-r ball in G can be made into a tree by removing at most a single edge. Let $v \in V_G$ be any vertex. If $|E_G(v, S)| \leq 2$, we are done. So assume otherwise (which obviously implies that there is a neighbour of v in S).

Consider the radius-r ball around v in the induced graph on $S \cup \{v\}$, and denote it B. Since B is a subgraph of $B_r(v)$, there is an edge set $\mathcal{E} \subseteq E_G(B)$ with $|\mathcal{E}| \leq 1$, such that removing $\mathcal{E}$ from B makes it a tree.

Let U be the set of neighbours of v in S. Since S is connected and does not contain v, there is a spanning tree T of $G[S]$. Note that since $B \setminus \mathcal{E}$ is a tree where all vertices of U are adjacent to v, any path between two vertices of U with length $\leq 2r$ in $G[S]$ has to contain the edge in $\mathcal{E}$ (as the radius-$(r-1)$ neighbourhoods of vertices

of U are pairwise disjoint in $G[S]$ after removing $\mathcal{E}$). Therefore, if a component of $T \setminus \mathcal{E}$ contains $\ell \geq 2$ vertices from U, then its size is at least ℓr. Removing the edge (if any) of $\mathcal{E}$ from T disconnects it into at most 2 components. Since we assumed $|U| = |E_G(v, S)| > 2$, at least $|E_G(v, S)| - 1$ of the vertices in U are in components of $T \setminus \mathcal{E}$ together. Thus, summing up we obtain $|S| \geq (|U| - 1)r$. Rearranging we get that $|E_G(v, S)| \leq 1 + \max(1, \frac{|S|}{r}) \leq 2 + \frac{d|S|}{\log n}$. □

As a result, we lower-bound the number of out-edges going out of a polymer. For sufficiently large polymers, the resulting bound is proportional to the number of inner vertices.

Corollary 6.41. *For all real d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ the following holds. Let $v \in V_C$ and let $F \in \Omega^{\mathrm{ord}}$. Then for any $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$, it holds that $e_{\mathrm{out}}(\gamma) \geq |E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})| - 2 - \frac{d|\mathcal{V}_\gamma|}{\log n}$.*

Proof. By Observation 6.39, any in-edges in $E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})$ are from $\mathcal{V}_\gamma$ to v. If $v \in \mathcal{V}_\gamma$, all edges in $E_G(\mathcal{V}_\gamma, \overline{\mathcal{V}_\gamma})$ are out, so the inequality follows. If $v \notin \mathcal{V}_\gamma$, recall that, for a singleton polymer γ, $\mathcal{V}_\gamma = \emptyset$ and for a non-singleton polymer γ, $\mathcal{V}_\gamma$ is connected in C. Apply Lemma 6.40 to show that $|E_G(v, \mathcal{V}_\gamma)| \leq 2 + d|\mathcal{V}_\gamma|/\log(n)$. □

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

Lemma 6.42. *There exists a real $A > 0$ such that for all d large enough, and for all real $q > 1$, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ the following holds for all $\beta \geq \beta_0$. For any configuration $F \in \Omega_C^{\mathrm{ord}}$, any vertex $v \in V_C$ and any polymer $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$ with $|\mathcal{V}_\gamma| \geq \frac{A}{d} \log n$, it holds that*

$$w_C^{\mathrm{ord}}(\gamma) \leq q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}.$$

Proof. By Propositions 6.14 and 6.16, there is a real $A > 0$ such that for all d large enough, w.h.p. over G, for any connected vertex subset $S \subseteq V_C$ with $|S| \geq \frac{A \log n}{d}$, its average degree is at least $\frac{35}{36}d$. Moreover, if $\deg_G(S) \leq \frac{1}{10} \deg_G(V_C)$, then $|E_G(S, \overline{S})| \geq \frac{3}{5} \deg_G(S)$. Thus, for the rest of this proof, assume that G satisfies these properties.

We first show that for any ordered polymer γ with $|\mathcal{V}_\gamma| \geq \frac{A}{d} \log n$,

$$e_{\mathrm{out}}(\gamma) \geq \tfrac{4}{7} \deg_G(\mathcal{V}_\gamma). \tag{6.7}$$

By Lemma 6.37, $\deg_G(\mathcal{V}_\gamma) \leq \frac{1}{10} \deg_G(V_C)$. By Corollary 6.41, $e_{\mathrm{out}}(\gamma) \geq \frac{3}{5} \deg_G(\mathcal{V}_\gamma) - 2 - \frac{d|\mathcal{V}_\gamma|}{\log n}$. Finally, $\mathcal{V}_\gamma$ is a connected subset of V_C with $|\mathcal{V}_\gamma| \geq \frac{A}{d} \log n$, hence $\frac{2 + d|\mathcal{V}_\gamma|/\log n}{\deg_G(\mathcal{V}_\gamma)} \leq$

$\frac{2+d|\mathcal{V}_\gamma|/\log n}{35d|\mathcal{V}_\gamma|/36}$. Since $|\mathcal{V}_\gamma| = \Omega(\log n)$, this fraction is at most $1/35$ for sufficiently large n. Thus, $\frac{e_{\text{out}}(\gamma)}{\deg_G(\mathcal{V}_\gamma)} \geq \frac{3}{5} - \frac{1}{35} = \frac{4}{7}$, concluding the proof of (6.7).

Recall from Definition 6.35 that $c'(\gamma)$ is the number of components of $(V_C, E_C \setminus E_{\text{out}}(\gamma))$ with total degree at most $(1/2)\deg_G(V_C)$, that is, the number of small components of $(V_C, E_C \setminus E_{\text{out}}(\gamma)))$.

We first argue that $c'(\gamma) \leq |\mathcal{V}_\gamma|$. Let κ be an arbitrary small component of $(V_C, E_C \setminus E_{\text{out}}(\gamma))$. We will show that $V_G(\kappa) \subseteq \mathcal{V}_\gamma$. Since $F \subseteq E_C \setminus E_{\text{out}}(\gamma)$, all vertices in κ are contained in small components of (V_C, F). Thus, from the polymer definition, all vertices of κ are contained in $\mathcal{V}(F, v)$. Since C is connected, the component κ of $(V_C, E_C \setminus E_{\text{out}}(\gamma))$ has a vertex incident to some edge $e \in E_{\text{out}}(\gamma)$. Since γ is a non-singleton polymer, the edges of $E_{\text{out}}(\gamma)$ are all incident to vertices in $\mathcal{V}_\gamma$, and thus in particular there is $w \in \mathcal{V}_\gamma$ incident to e. Without loss of generality, we may assume that $w \in V_G(\kappa)$.[6] We have also shown already that all vertices of κ are contained in $\mathcal{V}(F, v)$. By construction, $\mathcal{V}_\gamma$ is a maximal subset of $\mathcal{V}(F, v)$ that is connected in C. Since κ contains a vertex w of $\mathcal{V}_\gamma$, all of its vertices are contained in $\mathcal{V}_\gamma$, so $V_G(\kappa) \subseteq \mathcal{V}_\gamma$. Hence, $c'(\gamma) \leq |\mathcal{V}_\gamma|$.

Next, using (6.7), the average degree bound from Proposition 6.14, and the fact that $c'(\gamma)$ is at most $|\mathcal{V}_\gamma|$, we get that

$$\frac{c'(\gamma)}{e_{\text{out}}(\gamma)} \leq \frac{|\mathcal{V}_\gamma|}{\frac{4}{7}\deg_G(\mathcal{V}_\gamma)} \leq \frac{1}{\frac{4}{7}\frac{35}{36}d} = \frac{9}{5d} \tag{6.8}$$

To finish the proof, plug (6.8) and then (6.7) into the definition of a polymer weight (Definition 6.35) and use the fact that $\mathrm{e}^\beta - 1 \geq q^{(2-1/10)/d}$ for $\beta \geq \beta_0$ (see (6.1)) to obtain:

$$w_C^{\text{ord}}(\gamma) = q^{c'(\gamma)}(\mathrm{e}^\beta - 1)^{-e_{\text{out}}(\gamma)} \leq [q^{\frac{9}{5d} - \frac{2-1/10}{d}}]^{e_{\text{out}}(\gamma)} = q^{-e_{\text{out}}(\gamma)/(10d)} \leq q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}.$$

□

As a result, the weight of sufficiently large polymers decays exponentially in their size.

Note that Lemma 6.42 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. [31]), we must treat

[6] If not, then let u be the other endpoint of e. Since it is adjacent to e, then $u \in V_G(\kappa)$, so it is in $\mathcal{V}(F, v)$. But since $u \in \mathcal{V}(F, v)$ and u is an endpoint of e, it is a vertex of γ, so by the definition of $\mathcal{V}_\gamma$, it follows that $u \in \mathcal{V}_\gamma$. In this case, we can just use u as the vertex w. So we can assume $w \in V_G(\kappa)$.

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, which allows us to bound the probability of a polymer with size r or more by a sum of weights of all polymers of size r or more.

Lemma 6.43. *For all d large enough, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ the following holds. For $v \in V_C$ and any $\gamma \in \bigcup_{F \in \Omega_C^{\mathrm{ord}}} \Gamma_v^{\mathrm{ord}}(F)$, it holds that $\pi_C^{\mathrm{ord}}(\gamma \in \Gamma_v^{\mathrm{ord}}(\cdot)) \leq w_C^{\mathrm{ord}}(\gamma)$.*

Proof. Fix $v \in V_C$ and let $\gamma \in \bigcup_{F \in \Omega_C^{\mathrm{ord}}} \Gamma_v^{\mathrm{ord}}(F)$. If $e_{\mathrm{out}}(\gamma) = 0$, then since C is connected, $w_C^{\mathrm{ord}}(\gamma) = 1$, and the inequality is trivial. Thus, for the rest of the proof, assume that $E_{\mathrm{out}}(\gamma) \neq \emptyset$. Let $\Omega_C^{\mathrm{ord}}(\gamma) = \{F \in \Omega_C^{\mathrm{ord}}; \gamma \in \Gamma_v^{\mathrm{ord}}(F)\}$ and note that

$$\pi_C^{\mathrm{ord}}(\gamma \in \Gamma_v^{\mathrm{ord}}(\cdot)) = \frac{\sum_{F \in \Omega_C^{\mathrm{ord}}(\gamma)} w_C(F)}{\sum_{F \in \Omega_C^{\mathrm{ord}}} w_C(F)}.$$

For every ordered F with $\gamma \in \Gamma_v^{\mathrm{ord}}(F)$, define $F' \coloneqq F \cup E_{\mathrm{out}}(\gamma) \neq F$. We claim that $w_C(F') = w_C(F)/w_C^{\mathrm{ord}}(\gamma)$. Clearly, $|F'| = |F| + |E_{\mathrm{out}}(\gamma)|$ since all edges in $E_{\mathrm{out}}(\gamma)$ are out-edges.

If γ is a singleton polymer, then neither of its endpoints is in a small component of (V_C, F), thus setting its single edge to be an in-edge does not change the number of small components, so $c(F) = c(F')$, and it follows that $w_C(F') = w_C(F)(\mathrm{e}^{\beta} - 1) = w_C(F)/w_C^{\mathrm{ord}}(\gamma)$.

Hence, assume that γ is a non-singleton polymer with $E_{\mathrm{out}}(\gamma) \neq \emptyset$ and consider the components of (V_C, F) and (V_C, F'). From Lemma 6.37, there is exactly one giant component in each of (V_C, F) and (V_C, F'). We claim that the components of (V_C, F') are the giant component, together with the small components of (V_C, F) that do *not* contain vertices of $\mathcal{V}_\gamma$.

To prove the claim, first suppose κ is a small component of (V_C, F) not containing vertices of $\mathcal{V}_\gamma$. We use $V_G(\kappa)$ for the set of vertices of κ. Then no edge from $E_G(V_G(\kappa), \overline{V_G(\kappa)})$ is in $E_{\mathrm{out}}(\gamma)$, thus κ is a component of (V_C, F').

For the other direction, let κ be a small component of (V_C, F'). The vertices of κ are in small components of (V_C, F), so by the definition of $\mathcal{V}(F, v)$, $V_G(\kappa) \subseteq \mathcal{V}(F, v)$. We wish to show that κ is a small component of (V_C, F) that does not contain vertices of $\mathcal{V}_\gamma$.

Suppose for contradiction that $\mathcal{V}_\gamma \cap V_G(\kappa) \neq \emptyset$. Then, by the definition of non-singleton polymers, $V_G(\kappa) \subseteq \mathcal{V}_\gamma$, resulting in $E_G(V_G(\kappa), \overline{V_G(\kappa)}) \subseteq E_{\text{out}}(\gamma)$, contradicting the fact that κ is a small component of (V_C, F'). For the rest of the proof, we can therefore assume that $\mathcal{V}_\gamma \cap V_G(\kappa) = \emptyset$.

To finish the proof of the claim, we wish to show that κ is a small component of (V_C, F). Since (V_C, F) is a subgraph of (V_C, F'), there is a non-empty subgraph κ' of κ which is a component of (V_C, F). Suppose for contradiction that $V_G(\kappa') \subsetneq V_G(\kappa)$. Since κ is connected in F', there exists an edge $e = \{w, u\}$ where $w \in V_G(\kappa) \setminus V_G(\kappa')$ and $u \in V_G(\kappa')$. But then e must be an in-edge for F' but not for F, hence $e \in E_{\text{out}}(\gamma)$. By the definition of non-singleton polymers, either w or u is in $\mathcal{V}_\gamma$. Both w and u are in $V_G(\kappa)$ so we have contradicted $\mathcal{V}_\gamma \cap V_G(\kappa) = \emptyset$, which we have already shown. Thus $\kappa = \kappa'$, which concludes the proof of the claim.

By the claim, $c(F) - c(F')$ is the number of small components of (V_C, F) that contain vertices of $\mathcal{V}_\gamma$. By the definition of polymers this is exactly $c'(\gamma)$, so by the definition of polymer weights $w_C(F') = q^{c(F')}(\mathrm{e}^\beta - 1)^{|F'|} = \frac{w_C(F)(\mathrm{e}^\beta - 1)^{e_{\text{out}}(\gamma)}}{q^{c'(\gamma)}} = \frac{w_C(F)}{w_C^{\text{ord}}(\gamma)}$.

For every $F \in \Omega_C^{\text{ord}}(\gamma)$, both F and F' are in Ω_C^{ord} (since $F \subseteq F'$). Furthermore, $F' \notin \Omega_C^{\text{ord}}(\gamma)$ and for any distinct $F_1, F_2 \in \Omega_C^{\text{ord}}(\gamma)$, F_1' and F_2' are distinct. Thus

$$\pi_C^{\text{ord}}(\gamma \in \Gamma_v^{\text{ord}}(\cdot)) = \frac{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} w_C(F)}{\sum_{F \in \Omega_C^{\text{ord}}} w_C(F)} \leq \frac{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} w_C(F)}{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} (w_C(F) + w_C(F'))}.$$

Hence we can write

$$\begin{aligned}\pi_C^{\text{ord}}(\gamma \in \Gamma_v^{\text{ord}}(\cdot)) &\leq \frac{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} w_C(F)}{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} (w_C(F) + w_C(F'))} \\ &\leq \frac{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} w_C(F)}{\sum_{F \in \Omega_C^{\text{ord}}(\gamma)} w_C(F)(1 + 1/w_C^{\text{ord}}(\gamma))},\end{aligned}$$

which is $\frac{1}{1+1/w_C^{\text{ord}}(\gamma)} \leq w_C^{\text{ord}}(\gamma)$. □

With this in hand, we can show an analogue of the Kotecký–Preiss condition [56] for our ordered polymers (but only considering those that are large enough).[7] We use this to bound the probability that a large polymer occurs.

Proposition 6.44. *There is a real $A > 0$ such that for all d large enough, for all sufficiently large reals q, the following holds w.h.p. over $G \sim \mathcal{G}(n, d/n)$ for all real $\beta \geq \beta_0$. For any $v \in V_C$,*

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

[7] The standard Kotecký–Preiss condition requires, roughly, that the sum of the weights of polymers containing a given vertex decays exponentially with the size of the polymers. Here, we establish a weaker bound, holding only for sizes at least $A \log n/d$.

Proof. By Proposition 6.14 and Lemma 6.42, there is a constant A such that w.h.p. any connected vertex set with size at least $\frac{A\log n}{d}$ has average degree at least $\frac{9}{10}d$, and any ordered polymer γ with $|\mathcal{V}_\gamma| \geq \frac{A\log n}{d}$ has $w_C^{\mathrm{ord}}(\gamma) \leq q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}$. Also, w.h.p. the bound from Lemma 6.43 holds.

Let $\Gamma_v^{\mathrm{ord}} = \bigcup_{F\in\Omega_C^{\mathrm{ord}}} \Gamma_v^{\mathrm{ord}}(F)$. By a union bound and Lemma 6.43 it is enough to show

$$\sum_{\gamma\in\Gamma_v^{\mathrm{ord}};\, |\mathcal{V}_\gamma|\geq A\log n/d} w_C^{\mathrm{ord}}(\gamma) \leq q^{-A\log n/(25d)}.$$

Start by using the bound on the polymer weights from Lemma 6.42 to show that

$$\sum_{\gamma\in\Gamma_v^{\mathrm{ord}};\, |\mathcal{V}_\gamma|\geq A\log n/d} w_C^{\mathrm{ord}}(\gamma) \leq \sum_{\gamma\in\Gamma_v^{\mathrm{ord}};\, |\mathcal{V}_\gamma|\geq A\log n/d} q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}.$$

Then note that any γ in this sum has $|\mathcal{V}_\gamma| \geq A\log n/d$. Hence, by Proposition 6.14, it has average degree at least $\frac{9}{10}d$ hence degree at least $\frac{9A}{10}\log n$. Thus, the sum is at most

$$\sum_{D\geq \frac{9A}{10}\log n} \sum_{\gamma\in\Gamma_v^{\mathrm{ord}};\, \deg_G(\mathcal{V}_\gamma)=D} q^{-\deg_G(\mathcal{V}_\gamma)/(20d)}$$

Now we bound the number of polymers of the given total degree D. There are $\leq n(2e)^{2D-1}$ connected subgraphs of total degree D in any graph (see e.g. [36, Lemma 6]). For any connected subset $S \subseteq V_C$ of total degree D, there are at most D edges incident to its vertices. Thus, there are at most 2^D polymers γ with $\mathcal{V}_\gamma = S$ (each of its edges can be either in or out). Thus the sum is at most

$$\sum_{D\geq 9A\log n/10} n(2\mathrm{e})^{2D-1}2^D q^{-D/(20d)}.$$

For all q sufficiently large, $8\mathrm{e}^2 q^{-1/(20d)} \leq q^{-1/(22d)}$, so the sum is $O(nq^{-9A\log n/(220d)})$. For q sufficiently large, this is $\leq q^{-A\log n/(25d)}$. □

6.4.2 Proof of WSM within the ordered phase

We can now conclude WSM within the ordered phase.

Lemma 6.45. *There is a constant $A > 0$ such that the following holds for all real d sufficiently large. For all q sufficiently large, w.h.p. 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 $K = q^{-1/30}$.*

Proof. By Equation (6.6),

$$\left|\pi_{B_r^+(v)}(1_e) - \pi_C^{\mathrm{ord}}(1_e)\right| \le 2\pi_C^{\mathrm{ord}}(\neg\mathcal{A}_{v,r}) + \frac{2\pi_C^{\mathrm{ord}}(\neg\mathcal{A}'_{v,r})}{1-\pi_C^{\mathrm{ord}}(\neg\mathcal{A}'_{v,r})}.$$

As we noted at the beginning of the section, by Theorem 6.2, the second term is $\mathrm{e}^{-\Omega(n)}$, which is at most $q^{-r/25}$ for all n large enough. By Proposition 6.44 together with Lemma 6.36, $2\pi_C^{\mathrm{ord}}(\neg\mathcal{A}_{v,r}) \le 2q^{-A\log n/(25d)} \le 2q\cdot q^{-r/25}$. It follows that for all n sufficiently large, $(2q+1)q^{-r/25} \le q^{-r/30}$, which concludes the proof.

□

6.5 Weak spatial mixing within the disordered phase

In this section, we prove WSM within the disordered phase at distance $r \approx \frac{\log n}{d}$. We make use of the results from Chapter 4 and the fact that $B_r(v)$ is 1-treelike. Recall that Lemma 4.41 says that to bound $\left\|\pi_C^{\mathrm{dis}}(1_e) - \pi_{B_r^-(v)}(1_e)\right\|_{\mathrm{TV}}$, it is enough to bound the event that there is a component of in-edges of size at least $\Omega(r)$ and the event that the outside of the ball contains too many edges. Here, we slightly strengthen the result and show that in fact it suffices to bound the probability that either there are too many in-edges outside of the ball (which is unlikely by Theorem 6.2) or that there is a component with at least r vertices (which we bound using *disordered polymers* from [49]).

Lemma 6.46. *Let $r \ge 1$ be an integer, $H = (V_H, E_H)$ be a graph and $v \in V_H$ be a vertex. Then there is a coupling $(\boldsymbol{F}^-, \boldsymbol{F}^{\mathrm{dis}}) \sim (\pi_{B_r^-(v)}, \pi_H^{\mathrm{dis}})$, such that at least one of the following happens:*

1. *$|\boldsymbol{F}^{\mathrm{dis}} \setminus E_H(B_r(v))| > \eta|E_H| - |E_H(B_r(v))|$.*
2. *$\boldsymbol{F}^-$ and $\boldsymbol{F}^{\mathrm{dis}}$ agree on all edges incident to v.*
3. *$(V_H, \boldsymbol{F}^{\mathrm{dis}})$ has a component with at least r vertices.*

Proof. Define the coupling as follows. First, reveal the outside of the ball $B_r(v)$. Let $\boldsymbol{F}_0^{\mathrm{dis}} \subseteq E_H \setminus E_H(B_r(v))$ be the set of in-edges revealed this way for $\boldsymbol{F}^{\mathrm{dis}}$.

Then, if $|\boldsymbol{F}_0^{\mathrm{dis}}| \le \eta|E_H| - |E_H(B_r(v))|$, then regardless of the state of edges in the ball, the resulting configuration is disordered. Otherwise, the first event happens. Thus, henceforth, assume that $|\boldsymbol{F}_0^{\mathrm{dis}}| \le \eta|E_H| - |E_H(B_r(v))|$. As a result, we can drop the conditioning on being disordered and hence use monotonicity (Lemma 2.14)

to maintain $\boldsymbol{F}^- \subseteq \boldsymbol{F}^{\text{dis}}$ (as all edges revealed for $\boldsymbol{F}^-$ in this step are revealed to be out).

Second, we reveal the components of *non-trivial boundary vertices*[8]. Let S_0 be the set of vertices at distance r from v whose component in $(V_H, \boldsymbol{F}_0^{\text{dis}})$ contains a different vertex at distance r from v. We then perform the following procedure for $i = 0, 1, \ldots$.

Algorithm 6.47.

1. *If S_i is empty, reveal all remaining edges and terminate.*
2. *Otherwise, pick an arbitrary $u_i \in S_i$. Reveal all of its incident unrevealed edges, maintaining $\boldsymbol{F}^- \subseteq \boldsymbol{F}^{\text{dis}}$. Let S_{u_i} be the set of neighbours of u_i that are joined to u_i by a revealed in-edge and that have some unrevealed edge incident to them. Let $S_{i+1} = (S_i \cup S_{u_i}) \setminus \{u_i\}$.*

   Increment i by one and repeat.

We observe two invariants. First, if a vertex is in S_i, then it is connected via a path of in-edges to a vertex in S_0. This follows by induction, since a vertex is only added to S_{i+1} if there is an in-edge in $\boldsymbol{F}^{\text{dis}}$ joining it to a vertex in S_i and, by the induction hypothesis, each vertex of S_i is connected via in-edges to a vertex in S_0.

The second invariant is that if, after i steps, there is a boundary vertex (i.e. a vertex that has an incident revealed edge and an incident unrevealed edge), then either this vertex is not connected to any other boundary vertex via in-edges (i.e. it is in a trivial boundary component) or it is in S_i. Indeed, suppose that after i iterations of Algorithm 6.47, u is a boundary vertex that is not in a trivial boundary component. Note that by assumption u has to have a revealed in-edge incident to it. First assume that there is an in-edge incident to u that was revealed during some previous iteration of Algorithm 6.47. Let $j < i$ be such that u_j is incident to u, and this edge was revealed to be an in-edge. But then, $u \in S_{j+1}$. Since u is still a boundary vertex after i iterations, it has not been selected in any of the intervening iterations, and hence $u \in S_i$. Otherwise, there is an incident in-edge of u in $\boldsymbol{F}_0^{\text{dis}}$. Since all edges outside of the ball were revealed and u also has an incident unrevealed edge, u has to be a vertex at distance r from v. Since we assumed that u has no incident in-edge revealed inside the ball and u is in a non-trivial boundary component, it must be connected to another vertex at distance r from v via in-edges in $\boldsymbol{F}_0^{\text{dis}}$ (as

[8]Recall, from Section 4.1.5, that boundary components are the equivalence classes on the boundary vertices induced by the relation "are in the same component of revealed in-edges". A non-trivial boundary component contains two or more boundary vertices.

all in-edges outside of the ball in the resulting configuration $\boldsymbol{F}^{\mathrm{dis}}$ are in $\boldsymbol{F}_0^{\mathrm{dis}}$). Thus, $u \in S_0$. As in the first case, since u is still a boundary vertex after i iterations, it has not yet been selected, and therefore $u \in S_i$.

As a consequence, whenever the condition in Step 1 of the revealing process is satisfied, all boundary components are trivial (monotonicity ensures that this is then the case in both $\boldsymbol{F}^{-}$ and $\boldsymbol{F}^{\mathrm{dis}}$). Thus, by Observation 4.18, the remaining edges can be coupled so that $\boldsymbol{F}^{-}$ and $\boldsymbol{F}^{\mathrm{dis}}$ agree on all edges incident to v, provided none of these edges was revealed prior to the termination in Step 1. Now, if an incident edge of v was revealed, then for some $j \geq 0$, the vertex u_j is a neighbour of v. But then, by the first invariant, u_j is connected via a path of in-edges of $\boldsymbol{F}^{\mathrm{dis}}$ to a vertex at distance r from v. In particular, there must be a path of in-edges in $(V_H, \boldsymbol{F}^{\mathrm{dis}})$ of length at least $r-1$, and hence a component with at least r vertices. Thus Item 3 holds. □

With this in hand, we use the same strategy as in Chapter 4, and capture the probability of this event using the *disordered polymers* from [49]. For completeness, we provide a definition here using the notation from this chapter and prove the analogues of results in Section 4.4 for Erdős-Rényi graphs. In particular, after the definition, we first show that the weights of sufficiently large polymers decay, then prove that the probability of a given polymer appearing is at most its weight, and finally conclude that with sufficiently high probability there are no large polymers (components of in-edges).

Definition 6.48. *For $F \in \Omega_C^{\mathrm{dis}}$, the* disordered polymers *of F are* connected components *of (V_C, F). Let $\Gamma^{\mathrm{dis}}(F)$ be the set of all disordered polymers of F. For a disordered polymer $\gamma = (U, B) \in \Gamma^{\mathrm{dis}}(F)$, its weight is defined as $w_C^{\mathrm{dis}}(\gamma) \coloneqq q^{1-|U|}(\mathrm{e}^{\beta}-1)^{|B|}$.*

Recall that $\beta_1 = \beta_1(q,d)$ satisfies $\mathrm{e}^{\beta_1} - 1 = q^{(2+1/10)/d}$. We next show that when $\beta \leq \beta_1$, the weights of all sufficiently large polymers decay exponentially in their size.

Lemma 6.49. *There exists a real $A > 0$, such that for all d sufficiently large, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ the following holds for all real parameters $q > 1$ and $0 < \beta \leq \beta_1$. Let $F \in \Omega_C^{\mathrm{dis}}$, and $\gamma \in \Gamma^{\mathrm{dis}}(F)$. If $|V_G(\gamma)| \geq \frac{A \log n}{d}$, then also*

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

Proof. We divide the proof into two cases. First consider the case $|V_G(\gamma)| \geq \frac{n}{6}$. Then also note that since $\gamma \in \Gamma^{\mathrm{dis}}(F)$, it holds that $|e_G(\gamma)| \leq |F| \leq \eta |E_C| \leq 3\eta dn/4$ (the last inequality follows w.h.p. over G from $|E_C| \leq |E_G| = nd/2 + o(n)$).

Thus, using that $\mathrm{e}^\beta - 1 \le q^{(2+1/10)/d}$,

$$w_C^{\mathrm{dis}}(\gamma) \le q^{1-n/6+\frac{2+1/10}{d}(3\eta dn/4)} \le q^{1-n[1/6-2\eta]}.$$

Since $\eta < 1/30$, it holds that $1/6 - 2\eta \ge 1/10$. Thus $w_C^{\mathrm{dis}}(\gamma) \le q^{1-n/10} \le q^{1-|V_G(\gamma)|/10}$, where the last inequality follows from $|V_G(\gamma)| \le n$.

Now let A be the constant from Lemma 6.15 (with $\xi = 1/6$) and thus consider the case $\frac{A\log n}{d} \le |V_G(\gamma)| < n/6$. By the lemma, w.h.p. it holds that any such γ has $e_G(V_G(\gamma)) \le (d/9+1)|V_G(\gamma)|$. Since we have $e_G(\gamma) \le e_G(V_G(\gamma)) \le (d/9+1)|V_G(\gamma)|$, we combine the inequalities to obtain:

$$w_C^{\mathrm{dis}}(\gamma) \le q^{1-|V_G(\gamma)|(1-\frac{21}{10d}(d/9+1))} \le q^{1-|V_G(\gamma)|(23/30-21/10d)}.$$

For all d large enough this is at most $q^{1-|V_G(\gamma)|/10}$. □

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

Lemma 6.50. *For all connected graphs C, the following holds. For any $F \in \Omega_C^{\mathrm{dis}}$,*

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

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

Proof. For the first part, expand out the r.h.s. using polymer weight definition,

$$q^{n_C} \prod_{\gamma \in \Gamma^{\mathrm{dis}}(F)} w_C^{\mathrm{dis}}(\gamma) = q^{n_C + |\Gamma^{\mathrm{dis}}(F)| - \sum_{\gamma\in\Gamma^{\mathrm{dis}}(F)} |V_G(\gamma)|} (\mathrm{e}^\beta - 1)^{\sum_{\gamma\in\Gamma^{\mathrm{dis}}(F)} e_G(\gamma)}$$

The disordered polymers of F are exactly the connected components on (V_C, F), thus $|\Gamma^{\mathrm{dis}}(F)| = c(F)$ and $\sum_{\gamma\in\Gamma^{\mathrm{dis}}(F)} |V_G(\gamma)| = n_C$. Every in-edge of F is in a polymer and polymers of F do not overlap, thus also $\sum_{\gamma\in\Gamma^{\mathrm{dis}}(F)} e_G(\gamma) = |F|$. Hence the first part of the lemma follows.

Next, notice that any disordered polymer containing no edges has weight 1. So for any $\gamma = (\{v\}, \emptyset)$, the inequality in the lemma is trivial. So suppose γ is a polymer with at least one edge. Then for any $F \in \Omega_C^{\mathrm{dis}}$ with $\gamma \in \Gamma^{\mathrm{dis}}(F)$, let $F' \coloneqq F \setminus E_G(\gamma)$. Note that $F' \ne F$ and (for fixed γ) the map $F \mapsto F'$ is injective. Also, $\gamma \notin \Gamma^{\mathrm{dis}}(F')$. The components of (V_C, F') are exactly the components of (V_C, F), except for the vertices of γ (which are in a single component γ in (V_C, F), but in isolated vertices in (V_C, F')). Thus $w_C(F') = w_C(F) q^{|V_G(\gamma)|-1} (\mathrm{e}^\beta - 1)^{-e_G(\gamma)} = w_C(F)/w_C^{\mathrm{dis}}(\gamma)$.

Let $\Omega_C^{\mathrm{dis}}(\gamma) = \left\{F \in \Omega_C^{\mathrm{dis}} \mid \gamma \in \Gamma^{\mathrm{dis}}(F)\right\}$. Then we can write

$$\pi_C^{\mathrm{dis}}(\gamma \in \Gamma^{\mathrm{dis}}(\cdot)) = \frac{\sum_{F\in\Omega_C^{\mathrm{dis}}(\gamma)} w_C(F)}{\sum_{F\in\Omega_C^{\mathrm{dis}}} w_C(F)} \leq \frac{\sum_{F\in\Omega_C^{\mathrm{dis}}(\gamma)} w_C(F)}{\sum_{F\in\Omega_C^{\mathrm{dis}}(\gamma)}(w_C(F)+w_C(F'))}$$
$$\leq \frac{1}{1+1/w_C^{\mathrm{dis}}(\gamma)} \leq w_C^{\mathrm{dis}}(\gamma).$$

□

Now, we can state that with sufficiently high probability, there are no large polymers.

Proposition 6.51. *There exists a real $A > 0$, such that for all sufficiently large d, for all sufficiently large q, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ the following holds for all $0 < \beta \leq \beta_1$.*

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

Proof. By Lemmas 6.15 and 6.49, there is $A > 0$ such that for all d large enough, w.h.p. $G \sim \mathcal{G}(n, d/n)$ is such that for all $\beta \in (0, \beta_1]$, any polymer γ with at least $\frac{A\log n}{d}$ vertices has weight at most $q^{1-|V_G(\gamma)|/10}$, and also any *connected* vertex set $S \subseteq V_C$ of size at least $\frac{A\log n}{d}$ has $e_G(S) \leq (2d/3+1)|S|$. Also, by the union bound and Markov inequality applied to Lemma 6.7, w.h.p. G is such that for each $1 \leq s \leq n$ there are at most $n^{5/2}(\mathrm{e}d)^s$ connected subsets of size s. Further on, assume G is such that the above holds. By the previous lemma and by the union bound, it is sufficient to prove that

$$\sum_{\substack{\gamma\in\bigcup_{F\in\Omega_C^{\mathrm{dis}}}\Gamma^{\mathrm{dis}}(F)\\ |V_G(\gamma)|\geq A\log n/d}} w_C^{\mathrm{dis}}(\gamma) \leq q^{-A\log n/(25d)}.$$

For any connected subset S of size at least $\frac{A\log n}{d}$, it holds that $e_G(S) \leq (2d/3+1)|S|$, hence there are at most $2^{(2d/3+1)|S|}$ possible connected subgraphs of G with the vertex set S. Hence the number of polymers with $s \geq \frac{A}{d}\log n$ vertices is at most $n^{5/2}(\mathrm{e}d)^s 2^{(2d/3+1)s}$. By Lemma 6.49, for any polymer with $s \geq \frac{A}{d}\log n$ vertices, its weight is at most $q^{1-s/10}$. Thus, summing over all possible sizes gives us that the sum is at most

$$\sum_{s\geq(A/d)\log n} n^{5/2}(\mathrm{e}d)^s 2^{(2d/3+1)s} q^{1-s/10} = qn^{5/2} \sum_{s\geq(A/d)\log n} [2^{2d/3+1}\mathrm{e}dq^{-1/10}]^s.$$

Finally, provided that q is sufficiently large, $2^{2d/3+1}\mathrm{e}dq^{-1/10} \leq q^{-1/20}$, and for q large enough, the sum can be also bounded by $O(n^{5/2}q^{-\frac{A\log n}{d}/20}) \leq q^{-A\log n/(25d)}$. □

Now we can prove WSM within the disordered phase.

Proposition 6.52. *There exists a real $A > 0$, such that for all d large enough, and then for all q large enough reals, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, it holds that $C = C(G)$ has WSM within the disordered phase at distance $\lceil \frac{A \log n}{d} \rceil$ with rate $q^{-1/30}$.*

Proof. By Proposition 6.51 and Theorem 6.2(3), there exists a real $A > 0$ such that for all d and q large enough, w.h.p. G is such that $\pi_C^{\mathrm{dis}}(\exists \gamma \in \Gamma^{\mathrm{dis}}(\cdot) \text{ with } |V_G(\gamma)| \geq \frac{A \log n}{d}) \leq q^{-A \log n/(25d)}$ and also $\pi_C^{\mathrm{dis}}(|\cdot| \geq \frac{9\eta}{10}|E_C|) \leq \mathrm{e}^{-n}$. Let $r := \lceil \frac{A \log n}{d} \rceil$. By Lemma 6.6 together with Lemma 4.1, w.h.p. G is such that for all v, $|E_G(B_r(v))| \leq |V_G(B_r(v))| \leq d^r \log n$. Since also w.h.p. $|E_C| = \Omega(n)$ (Lemma 6.21), $|E_G(B_r(v))| \leq \eta|E_C|/10$ for all sufficiently large d and n. Fix $e \in E_C$ and a vertex v incident to e.

Since the total-variation distance between two distributions is upper-bounded by the disagreement probability under any coupling, Lemma 6.46 gives

$$\left\| \pi_C^{\mathrm{dis}}(1_e) - \pi_{B_r^-(v)} \right\|_{\mathrm{TV}} \leq \pi_C^{\mathrm{dis}}(|\cdot \setminus E_G(B_r(v))| > \eta|E_C| - |E_G(B_r(v))|) + \pi_C^{\mathrm{dis}}((V_C, \cdot) \text{ has a component with more than } r \text{ vertices})$$

Since $|E_G(B_r(v))| \leq \eta|E_C|/10$, the first event implies $|\boldsymbol{F}^{\mathrm{dis}}| > 9\eta|E_C|/10$. Thus, by Theorem 6.2(3), its probability is at most e^{-n}.

Then, since polymers are exactly the components of in-edges, the second term is exactly $\pi_C^{\mathrm{dis}}(\exists \gamma \in \Gamma^{\mathrm{dis}}(\cdot) \text{ with } |V_G(\gamma)| \geq r \geq \frac{A \log n}{d})$. By Proposition 6.51, this is at most $q^{-A \log n/(25d)}$.

Putting it together, we obtain $\left\| \pi_C^{\mathrm{dis}}(1_e) - \pi_{B_r^-(v)} \right\|_{\mathrm{TV}} \leq \mathrm{e}^{-n} + q^{-A \log n/(25d)} \leq q^{-r/30}$, where the last inequality holds for all n large enough. This concludes the proof of WSM within the disordered phase. □

6.6 Proof of Theorem 1.3

In this section, we finish the proofs of Theorems 6.3 and 6.4 and thus obtain Theorem 1.3.

6.6.1 Local mixing on treelike balls

Before we can proceed to the proof of mixing within the phase, we need to establish the remaining key property from Theorem 1.5: rapid mixing of the RC dynamics on treelike balls with wired and free boundary conditions.

For the free boundary condition, we can directly use Lemma 6.6 from [11][9] which gives us optimal local mixing for treelike neighbourhoods.

Lemma 6.53 ([11, Lemma 6.6])**.** *Let $B = (U, A)$ be a k-treelike graph. For every $q, \beta > 0$, there exists $\alpha_0(q, \beta, k)$ such that the log-Sobolev constant of the free RC dynamics X_t on B is at least $\alpha_0/|A|$.*

For the case of RC dynamics on a wired treelike ball, we use the result of Blanca and Gheissari [12] bounding the spectral gap for RC dynamics on a wired tree using its *edge-cut-width*. For any tree T with N vertices and edges $E(T) = \{e_1, \ldots, e_{N-1}\}$, its *edge-cut-width* is

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

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

We note that while their lemma was stated for trees with bounded degree, that restriction was not needed to get the following bound on the relaxation time (referred to as inverse spectral gap in [12]). Recall from Chapter 4, Equation (4.3), that the relaxation time is a standard quantity that can be used to bound the mixing time of Markov chains.

Lemma 6.54 ([12, Lemma 5.10])**.** *Let T be a tree with N vertices. Then for any reals $q \ge 2$ and $\beta > 0$, there is a constant $\theta = \theta(q, \beta) > 0$ such that the relaxation time of the RC dynamics on T with a* wired boundary *(all leaves of T belonging to the same component) is at most $\theta N^2 \exp\{2(\mathrm{CW}(T) + 1) \log q\}$.*

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

Lemma 6.55. *Let T be a height-r tree. Then $\mathrm{CW}(T) \le r$.*

Proof. Use the permutation obtained by a depth-first search (DFS) traversal of T starting from its root. Let σ be the order in which the edges are explored. We will show that $\max_{1 \le i < N} |V(\{e_{\sigma(j)} : j \le i\}) \cap V(\{e_{\sigma(j)} : j > i\})| \le r$.

Let $1 \le i < N - 1$. Then, after exploring i edges in the DFS traversal, we have three types of vertices: those that have all of their edges processed, those that have none of their edges processed, and those that have both a processed and an unprocessed edge incident to them (call these *active vertices*). Clearly, the first two vertex types are not in $V(\{e_{\sigma(j)} : j \le i\}) \cap V(\{e_{\sigma(j)} : j > i\})$.

[9] Lemma 6.6 in [11] is stated for the continuous-time RC dynamics; for the version of the RC dynamics used in this thesis, the log-Sobolev constant is scaled by a factor of $|A|$.

It hence remains to bound the number of active vertices (after i processed edges in the DFS traversal). For this, note that in a DFS traversal of a tree, active vertices can only lie on a path from a root to the deepest active vertex.

Since the length of a longest path from a root is r, we have $\mathrm{CW}(T) \le r$. □

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 relevant probabilities by at most a constant factor. The comparison argument of [12, Lemma 6.10] then gives that the relaxation time changes by only a constant factor.

Lemma 6.56. *For any real $A > 0$, for all d large enough, and then for all sufficiently large real q, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, the following holds. For all integer $r \le \frac{A \log n}{d}$ and all $v \in V_C$, let $B \coloneqq B_r(v)$.*

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

Proof. By Lemma 6.8, for all d large enough, w.h.p. any such B is 1-treelike. By Lemma 6.6, w.h.p. any such B has size $\le d^r \log n \le n^{A \log d/d} \log n$.

Let T be a breadth-first-search spanning tree of B rooted at v. Since B is 1-treelike, T is obtained from B by removing at most one edge, and has height at most r. By the previous two lemmas, the relaxation time for RC on T with a wired boundary condition satisfies

$$\tau_T = O(|V_G(B)|^2 \mathrm{e}^{2r \log q}) = O(n^{2A(\log d + \log q)/d} \log^2 n).$$

Then, using the comparison argument from [12, Lemma 6.10], the relaxation time τ_B of the RC dynamics on B with a wired boundary condition is at most $C(q, \beta)\tau_T$ (for some constant $C(q, \beta)$ independent of n).

We then have an additional factor of $O(n^{A \log d/d} \log n)$, since the mixing time is at most (see Equation 4.3) $\tau_B(1 + \frac{1}{2}\log(\frac{1}{\min_F \pi_{B^+}(F)}))$, and that for any configuration F on B, we have that $\log \frac{1}{\pi_{B^+}(F)} = \Omega(|V(B)|^{-1})$. Then, we just note that for all q large enough,

$$O(n^{A(3\log d + 2\log q)/d} \log^3 n) \le O(n^{4A \log q/d})$$ □

.

6.6.2 Proof of Theorems 6.3 and 6.4

Next we prove mixing within the phase.

Proof of Theorems 6.3 and 6.4. We obtain fast mixing using Theorem 1.5. It suffices to show that Ingredients 1–3 hold for $G \sim \mathcal{G}(n, d/n)$ with high probability.

Ingredient 1 for both phases follows from Theorem 6.2(2)-(3). Weak spatial mixing within a phase at radius $r \coloneqq \frac{A \log n}{d}$, for A a constant (Ingredient 2), is proven in Sections 6.4 and 6.5.

Finally, w.h.p. over $G \sim \mathcal{G}(n, d/n)$, it holds that G has local mixing in the ordered phase on radius-r neighbourhoods of G (Lemma 6.56) and optimal mixing in the disordered phase (Lemma 6.53).

Applying Theorem 1.5, we obtain polynomial mixing within the ordered phase and $O(|E| \log |E|) = O(n \log n)$ mixing within the disordered phase (since with high probability $|E| = \Theta(n)$). □

Then, we obtain Theorem 1.3 as a straightforward corollary.

Proof of Theorem 1.3. For $\beta \leq \beta_0$ and $\beta \geq \beta_1$, it follows from Theorems 6.2, 6.3 and 6.4 that we can get an $e^{-\Omega(n)}$-sample by initialising the RC dynamics from all-out or all-in, respectively. To obtain an approximation for the partition function in this range of β, use Lemma 3.4.

Finally, the fast mixing and an approximation algorithm for the partition function at $\beta \in (\beta_0, \beta_1)$ follow from Corollary 3.7. □

6.7 Proof of Theorem 1.4

We conclude the chapter by giving a proof of Theorem 1.4. We approximate the partition function using the partition functions for the two phases. We first outline how to efficiently approximate the ordered partition function; the argument for the disordered phase is analogous. We then state and prove the necessary technical lemmas before finishing the proof.

Let $R > 0$ be an arbitrarily large constant and let $\beta^* \geq \beta_0$ be a real; then, our goal is to approximate $Z^{\mathrm{ord}}_{C;q,\beta^*}$ in time $n^{O(1/R)}$ within relative error n^{-R}. We view $Z^{\mathrm{ord}}_C = Z^{\mathrm{ord}}_{C;q,\beta}$ as a function of β and set $g_C(\beta) \coloneqq \frac{\partial \log Z^{\mathrm{ord}}_{C;q,\beta}}{\partial \beta}$. Note in particular that

$$\begin{aligned} g_C(\beta) &= \frac{\frac{\partial Z^{\mathrm{ord}}_{C;q,\beta}}{\partial \beta}}{Z^{\mathrm{ord}}_{C;q,\beta}} = \frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1} \frac{\sum_{F \in \Omega^{\mathrm{ord}}_C} |F| (\mathrm{e}^\beta - 1)^{|F|} q^{c(F)}}{Z^{\mathrm{ord}}_{C;q,\beta}} \\ &= \frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1} \mathbb{E}_{F \sim \pi^{\mathrm{ord}}_C}\big[|F|\big] = \frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1} \sum_{e \in E_C} \pi^{\mathrm{ord}}_{C;q,\beta}(1_e). \end{aligned} \tag{6.9}$$

We relate the partition function at β^* to its value at the much larger inverse temperature $\beta_\infty \coloneqq n^{1/R}$, where it will be easy to approximate. Integrating g_C gives

$$\frac{Z^{\mathrm{ord}}_{C;q,\beta_\infty}}{Z^{\mathrm{ord}}_{C;q,\beta^*}} = \exp\left(\int_{\beta^*}^{\beta_\infty} g_C(\beta)\,\mathrm{d}\beta\right). \tag{6.10}$$

We will show later that for $\beta = \beta_\infty$ it holds that $Z^{\mathrm{ord}}_{C;q,\beta_\infty} = (1\pm \mathrm{e}^{-n^{\Omega(1/R)}})q(\mathrm{e}^{\beta_\infty}-1)^{|E_C|}$. So, to obtain the desired approximation to $Z^{\mathrm{ord}}_{C;q,\beta^*}$, it suffices to approximate the integral $I \coloneqq \int_{\beta^*}^{\beta_\infty} g_C(\beta)\,\mathrm{d}\beta$.

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 bound from WSM within the ordered phase. Namely, by Lemma 6.45, for an edge $e \in E_C$ and a vertex v incident to it, for the ball $B_e \coloneqq B_r(v)$ with $r = \lceil \frac{A}{d}\log n\rceil$, it holds that

$$\left|\pi^{\mathrm{ord}}_{C;q,\beta}(1_e) - \pi_{B_e^+;q,\beta}(1_e)\right| \le q^{-r/30} \le \tfrac{1}{n^{R+3}}, \tag{6.11}$$

where the last inequality follows by taking q large enough (with respect to d). Crucially, B_e is a 1-treelike 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 6.57 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, rather than evaluating the integrand separately at many values of β, we can essentially perform the integration $\int_{\beta^*}^{\beta_\infty} \pi_{B_e^+;q,\beta}(1_e)$ symbolically, i.e., directly manipulate its rational-function representation to compute an antiderivative (a function whose derivative is the integrand). A technicality that arises is that we cannot in general represent this antiderivative exactly, since it may involve algebraic numbers. Instead, we compute a numerical approximation to the antiderivative that is within additive $1/2^{n^{O(1/R)}}$ of its true value for all $\beta \in [\beta^*, \beta_\infty]$; as we formalise below, this requires only $n^{O(1/R)}$ bits of precision and hence can be carried out in $n^{O(1/R)}$ time.

We next state and prove some technical lemmas necessary to conclude the proof of Theorem 1.4. We start with the computation of the marginals of edges in 1-treelike graphs. We will apply this lemma with $T = B_r(v)$ for $r = \lceil \frac{A\log n}{d}\rceil$.

Lemma 6.57. *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 $r \ge 1$, computes in time*

$2^{4r}n^{O(1)}$ *rational functions* $g^+(q,x), g^-(q,x)$ *so that* $\pi_{B_r^+(v);q,\beta}(1_e) = g^+(q, \mathrm{e}^\beta - 1)$ *and* $\pi_{B_r^-(v);q,\beta}(1_e) = g^-(q, \mathrm{e}^\beta - 1)$ *for all* $q, \beta > 0$.

Proof. We argue first about $\pi_{B_r^+(v);q,\beta}(1_e)$. We consider first the case of trees, and then explain how to extend the result to 1-treelike graphs. So, assume that T is a tree rooted at v, and let L be the vertices at the boundary of $B_r(v)$. For a vertex u in T, let T_u be the subtree of T rooted at u (up to the boundary of $B_r(v)$), and let V_u denote its vertices. Moreover, let $Z^+(u)$ be the partition function of the subtree T_u rooted at u, where the vertices in $L \cap V_u$ are wired; let also $Z_1^+(u)$ be the contribution to $Z^+(u)$ coming from configurations in T_u where there is a path of in-edges from u to $L \cap V_u$ and set $Z_0^+(u) = Z^+(u) - Z_1^+(u)$.

Note that $\pi_{B_r^+(v);q,\beta}(1_e)$ equals $(\mathrm{e}^\beta - 1)\hat{Z}^+(v)/Z^+(v)$ where $\hat{Z}^+(v)$ is the analogue of $Z^+(v)$ for the tree T' obtained by contracting the edge e (onto v), where $\mathrm{e}^\beta - 1$ accounts for the contribution of the contracted edge. So, it suffices to explain how to compute $Z^+(u)$ for an arbitrary $u \in T$. Let $u_1, \ldots, u_k$ be the children of u. Then, with $t := \frac{\mathrm{e}^\beta - 1}{q} + 1$ the quantities Z_0^+, Z_1^+ defined above satisfy the following recurrences, see e.g. [11, Fact 3.2] or [14, Lemma 33]:

$$Z_1^+(u) = q \prod_{i=1}^{k} \frac{(q(t-1)+1)Z_1^+(u_i) + tZ_0^+(u_i)}{q} - q \prod_{i=1}^{k} \frac{Z_1^+(u_i) + tZ_0^+(u_i)}{q},$$

and

$$Z_0^+(u) = q^2 \prod_{i=1}^{k} \frac{Z_1^+(u_i) + tZ_0^+(u_i)}{q}.$$

Using these, we can therefore compute $Z^+(u) = Z_0^+(u) + Z_1^+(u)$ (as a function of q and $x = \mathrm{e}^\beta - 1$) recursively in time $n^{O(1)}$. Since these recurrences involve only rational operations in q and $x = \mathrm{e}^\beta - 1$, they produce the desired rational function.

In the case of 1-treelike graphs that are not trees, we view $B_r^+(v)$ as a cycle C with hanging trees from its vertices (whose leaves are wired); note that the cycle has at most $2r$ edges. To find the partition function of $B_r(v)$ with wired boundary, we first use the tree-algorithm from above for each of the hanging trees to compute $Z_0^+(u), Z_1^+(u)$ for each vertex u of the cycle. Then, we consider all possible configurations on the cycle (of which there are at most 2^{2r}), and for each one of them we consider all the vertex-to-boundary possibilities for the vertices u on the cycle, i.e., whether u is connected by a path to the boundary of its hanging tree (of which there are at most 2^{2r}). For each of these edge configurations and vertex-to-boundary settings, we take the product of corresponding variables from $\{Z_0^+(u), Z_1^+(u)\}_{u \in C}$ (divided by $q^{|C|}$ to

avoid double-counting the wired component) multiplied by the contribution of the in-edges on the cycle and the new components formed; aggregating all of these gives us the partition function of $B_r(v)$ in time $2^{4r}n^{O(1)}$.

For $\pi_{B_r^-(v);q,\beta}(1_e)$ the argument is much simpler since the partition function for an n-vertex tree under free boundary is just $q(\mathrm{e}^\beta - 1 + q)^{n-1}$. If $B_r(v)$ has a cycle, then the above approach can be easily adapted. □

The next lemma captures the integration part of the argument. For the statement of the next lemma, we use the following standard encoding-size conventions. 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 $n + \sum_{i=1}^n \mathrm{size}(c_i)$.

Lemma 6.58. *There is an algorithm that, on input positive rationals a, b, ε each with size at most n, and integer polynomials $P(x), Q(x)$ with non-negative coefficients and size at most n, computes in time poly(n) a number $\hat{I}$ so that $\hat{I} = I \pm \varepsilon$ where $I = \int_a^b \frac{P(x)}{Q(x)}\mathrm{d}x$.*

Proof. Since ε has size at most n, it satisfies $\varepsilon \geq 2^{-O(n)}$, and hence it suffices to produce an additive $1/2^{\Theta(n)}$ approximation to I.

Henceforth, we will assume without loss of generality that the degree of P is less than the degree of Q (otherwise, we can perform polynomial division, integrate the resulting polynomial part directly, and consider the remaining proper rational function). We may also focus on the case that all roots of $Q(x)$ have multiplicity 1, since there are standard symbolic integration techniques to reduce the case with higher-multiplicity roots to the case where all roots have multiplicity-1 in poly(n) time (e.g. Hermite's reduction, see [63] for details). Hence we can write the partial fraction decomposition of $g(x) = \frac{P(x)}{Q(x)}$ in the form

$$g(x) = \sum_{i=1}^{n_0} \frac{a_i}{b_i x + c_i} + \sum_{i=1}^{n_1} \frac{a_i' x + b_i'}{c_i' x^2 + d_i' x + e_i'}$$

for some reals $a_i, b_i, c_i, a_i', b_i', c_i', d_i', e_i'$ which can be determined using the roots of Q and integers $n_0, n_1 \leq n$. The advantage of this decomposition is that each of its terms has an explicit antiderivative. The roots of Q determine the terms appearing in this decomposition, after which the corresponding coefficients can be computed; see [63] for details.

Since $Q(x)$ is an integer polynomial which can be represented with n bits, its roots can be additively approximated with precision $\varepsilon' = 1/2^{\Theta(n)}$ in time $\mathrm{poly}(n, \log \frac{1}{\varepsilon'}) = \mathrm{poly}(n)$, see for example the algorithm in [65].

Therefore, we can approximate $a_i, b_i, c_i, a'_i, b'_i, c'_i, d'_i, e'_i$ with precision $1/2^{\Theta(n)}$. Note that the exponential dependence of the precision on n is needed to account for negative numbers that might arise in the computations and the magnitude of the coefficients of $P(x), Q(x)$. The exponential precision will also ensure that the approximation will be good enough later on for all $x \in [a, b]$ (note that a, b use n bits). Putting everything together, we obtain an approximation to the partial fraction decomposition of $g(x)$, i.e., a function $\hat{g}(x)$ of the form $\sum_{i=1}^{k} \frac{\hat{a}_i}{\hat{b}_i x + \hat{c}_i} + \sum_{j=1}^{k} \frac{\hat{a}'_j x + \hat{b}'_j}{\hat{c}'_j x^2 + \hat{d}'_j x + \hat{e}'_j}$ which satisfies $|g(x) - \hat{g}(x)| \leq 1/2^{\Theta(n)}$ for all $x \in [a, b]$. Since a and b have size at most n, the length of the interval $[a, b]$ is at most $2^{O(n)}$; thus, by choosing the constant in the above $\Theta(n)$ sufficiently large, replacing g by $\hat{g}$ changes the integral by at most $2^{-\Theta(n)}$. Now the integral $\int_a^b \hat{g}(x)\mathrm{d}x$ can be symbolically integrated and then numerically evaluated with additive precision $2^{-\Theta(n)}$. Together with the above approximation error, this gives $\hat{I}$ with $|\hat{I} - I| \leq 2^{-\Theta(n)}$, completing the proof. □

Finally, we will need to show that for very large β, we can approximate the ordered partition function by the weight of the all-in configuration, and also for very low β, we can approximate the partition function by the weight of the all-out configuration.

Lemma 6.59. *Let d be large enough. For all q sufficiently large, for any constant $\varepsilon > 0$, w.h.p. over $G \sim \mathcal{G}(n, d/n)$ it holds that $Z^{\mathrm{ord}}_{C;q,\beta} = \big(1 \pm \mathrm{e}^{-n^{\varepsilon}}\big) q (\mathrm{e}^{\beta} - 1)^{|E_C|}$ for any $\beta \geq n^{2\varepsilon}$, and that $Z^{\mathrm{dis}}_{C;q,\beta} = \big(1 \pm n^{-1/\varepsilon}\big) q^{n_C}$ for any β with $\mathrm{e}^{\beta} - 1 \leq n^{-(1/\varepsilon+2)}$.*

Proof. Consider arbitrarily small $\varepsilon > 0$. We claim that for $F \sim \pi^{\mathrm{ord}}_C$, it holds that $\mathbb{E}[|F|] \geq |E_C| - \mathrm{e}^{-n^{\varepsilon}}$. Assuming this for the moment, by Markov's inequality we have that $\mathbb{P}\big(|E_C \backslash F| \geq 1\big) \leq \mathrm{e}^{-n^{\varepsilon}}$. Therefore in the sum $Z^{\mathrm{ord}}_C = \sum_{F \in \Omega^{\mathrm{ord}}_C} (\mathrm{e}^{\beta} - 1)^{|F|} q^{c(F)}$ terms other than $F = E_C$ have at most $\mathrm{e}^{-n^{\varepsilon}}$ relative contribution, yielding the desired bound on $Z^{\mathrm{ord}}_{C;q,\beta}$.

To prove the bound on the expectation of F, we use that, by Theorem 6.2, it holds that $\big\|\pi^{\mathrm{ord}}_C - \pi_C\big\|_{\mathrm{TV}} = \mathrm{e}^{-n}$. Thus, it suffices to show that $\sum_{e \in E_C} \pi_C(1_e) \geq |E_C| - \mathrm{e}^{-n^{\varepsilon}}$. This is true for $\beta \geq n^{2\varepsilon}$ since the probability that e is an in-edge is at least $\frac{\mathrm{e}^{\beta}}{\mathrm{e}^{\beta} + q - 1} \geq 1 - \frac{\mathrm{e}^{-n^{\varepsilon}}}{|E_C|}$ for all n sufficiently large.

For the disordered case, similarly, by Theorem 6.2 we have $\big\|\pi^{\mathrm{dis}}_C - \pi_C\big\|_{\mathrm{TV}} = \mathrm{e}^{-n}$. Under π_C, the probability that a particular edge is an in-edge is at most $1 - \mathrm{e}^{-\beta} \leq$

$\mathrm{e}^{\beta}-1 \le n^{-(2+1/\varepsilon)}$. Hence, for all n sufficiently large, $\mathbb{E}_{F\sim\pi_C^{\mathrm{dis}}}[|F|] \le 2|E_C|(n^{-(2+1/\varepsilon)} + \mathrm{e}^{-n}) \le n^{-1/\varepsilon}$. Hence, by Markov's inequality, $\pi^{\mathrm{dis}}_{C;q,\beta_{-\infty}}(|\cdot| \ge 1) \le n^{-\frac{1}{\varepsilon}}$. Therefore, in $Z_C^{\mathrm{dis}} = \sum_{F\in\Omega_C^{\mathrm{dis}}}(\mathrm{e}^{\beta}-1)^{|F|}q^{c(F)}$, the terms other than $F=\emptyset$ have at most $n^{-1/\varepsilon}$ relative contribution, yielding the desired bound on $Z^{\mathrm{dis}}_{C;q,\beta}$. □

We can now conclude Theorem 1.4, which we restate for convenience.

Theorem 1.4. *Let $R > 0$ be any real. Then for all sufficiently large reals d and then all sufficiently large q, there is an algorithm which, with high probability over $G \sim \mathcal{G}(n, d/n)$, for any $\beta > 0$, outputs a $(1+\frac{1}{n^R})$-approximation to $Z_G(\beta, q)$ in time $n^{1+O(1/R)}$.*

Proof. We follow the argument outlined earlier in order to obtain $\hat{Z}^{\mathrm{ord}}$, an $(1 \pm \frac{1}{n^{R'}})$-approximation for $Z^{\mathrm{ord}}_{C;q,\beta}$ when $\beta \ge \beta_0$, where $R' \coloneqq R+1$. An analogous argument gives an approximation $\hat{Z}^{\mathrm{dis}}$ for $Z^{\mathrm{dis}}_{C;q,\beta}$ for $\beta \le \beta_1$ using WSM within the disordered phase (see below for the differences in the proofs). Then $\hat{Z}^{\mathrm{dis}}$ is the desired approximation when $\beta < \beta_0$, the sum $\hat{Z}^{\mathrm{dis}} + \hat{Z}^{\mathrm{ord}}$ is used when $\beta \in [\beta_0, \beta_1]$, and $\hat{Z}^{\mathrm{ord}}$ is used when $\beta > \beta_1$. By Theorem 6.2, disregarding configurations outside of the relevant phase(s) incurs only an additional $\mathrm{e}^{-\Omega(n)}$ error; for all n sufficiently large, the resulting quantity is therefore a $(1\pm 1/n^R)$-approximation to the overall partition function.

So fix a target $\beta^* \ge \beta_0$ and focus on approximating $Z^{\mathrm{ord}}_{C;q,\beta^*}$. Set $\beta_\infty = n^{1/R'}$. If $\beta^* \ge \beta_\infty$, then by Lemma 6.59 (with $\varepsilon = \frac{1}{2R'}$), $q(\mathrm{e}^{\beta^*}-1)^{|E_C|}$ gives the desired approximation $\hat{Z}^{\mathrm{ord}}$, so we may assume that $\beta^* < \beta_\infty$. From (6.9) and (6.10) we have that

$$\frac{Z^{\mathrm{ord}}_{C;q,\beta_\infty}}{Z^{\mathrm{ord}}_{C;q,\beta^*}} = \exp\left(\int_{\beta^*}^{\beta_\infty} g_C(\beta)\,\mathrm{d}\beta\right) \quad \text{where} \quad g_C(\beta) = \frac{\mathrm{e}^{\beta}}{\mathrm{e}^{\beta}-1}\sum_{e\in E_C}\pi^{\mathrm{ord}}_{C;q,\beta}(1_e).$$

By Lemma 6.59, we can approximate $Z^{\mathrm{ord}}_{C;q,\beta_\infty}$ with relative error $\mathrm{e}^{-n^{\Omega(1/R')}}$ so to obtain the desired approximation to $Z^{\mathrm{ord}}_{C;q,\beta^*}$, it suffices to approximate the integral $I \coloneqq \int_{\beta^*}^{\beta_\infty} g_C(\beta)\,\mathrm{d}\beta$ within additive error $\frac{1}{n^{R'+1}}$.

For an edge $e \in E_C$ and a vertex v incident to it, let $B_e = B_r(v)$ with $r = \lceil \frac{A}{d}\log n\rceil$ as in (6.11) where we saw that WSM gives, for q large enough w.r.t. d,

$$\left|\pi^{\mathrm{ord}}_{C;q,\beta}(1_e) - \pi_{B_e^+;q,\beta}(1_e)\right| \le q^{-r/30} \le \tfrac{\mathrm{e}^{\beta^*}-1}{\mathrm{e}^{\beta^*}}\tfrac{1}{n^{R'+3}}.$$

The r-neighbourhoods for $G \sim \mathcal{G}(n, d/n)$ are 1-treelike w.h.p. (see Lemma 6.8), so by Lemma 6.57 we can compute in time $n^{O(1/R')}$ a rational function[10] $g_e(x)$ so that

[10]Here we assume for convenience that q is rational; if q is irrational, the argument can be modified to use instead a rational approximation $\hat{q}$ to q satisfying $|\hat{q}-q| \le 1/2^{n^{\Theta(1/R')}}$.

$\pi_{B_e^+;q,\beta}(1_e) = g_e(\mathrm{e}^\beta - 1)$. Therefore for $x = \mathrm{e}^\beta - 1$ we have that

$$\left| g_C(\beta) - \frac{x+1}{x} \sum_{e \in E_C} g_e(x) \right| \leq \frac{|E_C|}{n^{R'+3}},$$

giving that

$$\left| I - \sum_{e \in E_C} \int_{x^*}^{x_\infty} \frac{g_e(x)}{x} \mathrm{d}x \right| \leq \frac{|E_C|(\beta_\infty - \beta^*)}{n^{R'+3}} \leq \frac{1}{n^{R'+1}}, \tag{6.12}$$

where the last inequality follows from the fact that w.h.p. $|E_C|(\beta_\infty - \beta^*) \leq n^{1+1/R'} \leq n^2$ (we can assume that $R' > 1$).

Then, using the algorithm in Lemma 6.58, for each $e \in E_C$ we estimate the integral $\int_{x^*}^{x_\infty} \frac{g_e(x)}{x} \mathrm{d}x$ within additive error $1/n^{R'+3}$ in time $n^{O(1/R')}$, since each of $g_e(x)/x$, as well as x_∞, has size $n^{O(1/R')}$. Summing over all edges, this introduces an additional additive error of at most $\frac{|E_C|}{n^{R'+3}}$. Combining this with (6.12), we obtain an estimate $\hat{I}$ satisfying $|\hat{I} - I| \leq \frac{|E_C|(\beta_\infty - \beta^* + 1)}{n^{R'+3}} \leq \frac{2}{n^{R'+1}}$. Consequently, $\mathrm{e}^{-\hat{I}} = (1 \pm O(n^{-(R'+1)}))\mathrm{e}^{-I}$.

Since $q(\mathrm{e}^{\beta_\infty} - 1)^{|E_C|}$ is a $(1 \pm \mathrm{e}^{-n^{\Omega(1/R')}})$-approximation to $Z^{\mathrm{ord}}_{C;q,\beta_\infty}$, it follows that $q(\mathrm{e}^{\beta_\infty} - 1)^{|E_C|}\mathrm{e}^{-\hat{I}}$ is a $(1 \pm \frac{1}{n^{R'}})$-approximation of $Z^{\mathrm{ord}}_{C;q,\beta^*}$ in time $n^{1+O(1/R')}$. This proves the result for the ordered phase.

We now briefly remark on the differences for the disordered phase. Let $\beta_{-\infty}$ satisfy $\mathrm{e}^{\beta_{-\infty}} - 1 = n^{-(R'+4)}$. If $\beta^* \leq \beta_{-\infty}$, then Lemma 6.59 (with $\varepsilon = 1/(R'+2)$) shows that q^{n_C} gives the desired approximation. We may therefore assume that $\beta^* > \beta_{-\infty}$ and approximate the integral of $\frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1} \sum_{e \in E_C} \pi^{\mathrm{dis}}_{C;q,\beta}$ from $\beta_{-\infty}$ to β^*.

The argument is otherwise analogous, using WSM within the disordered phase and the free-boundary rational functions from Lemma 6.57. The only difference is that for the WSM within the disordered phase (Proposition 6.52), we pick q large enough so that the resulting bound gives $\left|\pi^{\mathrm{dis}}_{C;q,\beta}(1_e) - \pi_{B_e^-;q,\beta}(1_e)\right| \leq q^{-r/30} \leq \mathrm{e}^{-\beta^*} \frac{1}{n^{2R'+7}}$.

Note that here we require a slightly stronger bound than in the ordered phase. This is so we can obtain that, for all $\beta \in [\beta_{-\infty}, \beta^*]$, the approximation by marginals in the ball with free boundary condition, $\frac{\mathrm{e}^\beta}{\mathrm{e}^\beta - 1} \sum_{e \in E_C} \pi_{B_e^-;q,\beta}(1_e)$ has at most $\frac{\mathrm{e}^{\beta^*}}{\mathrm{e}^{\beta_{-\infty}} - 1} |E_C| q^{-r/30} \leq |E_C| n^{-R'-3}$ additive error. The rest of the proof is then identical to the ordered phase: applying Lemma 6.58, we approximate each of the resulting integrals within additive error $n^{-R'-3}$ in time $n^{O(1/R')}$. Together with the WSM error above, this gives total additive error at most $|E_C| n^{-R'-3} + (\beta^* - \beta_{-\infty})|E_C| n^{-R'-3} \leq \frac{1}{n^{R'+1}}$ (where in the last inequality we use that $\beta^* \leq \beta_1 = O(1)$ and that w.h.p. $|E_C| = O(n)$). By exponentiating and using the estimate q^{n_C} for $Z^{\mathrm{dis}}_{C;q,\beta_{-\infty}}$ from Lemma 6.59, we obtain the desired $(1 \pm 1/n^{R'})$-approximation in total time $n^{1+O(1/R)}$. □

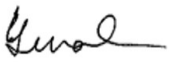
Statement of Authorship for joint/multi-authored papers for PGR thesis

To appear at the end of each thesis chapter submitted as an article/paper

The statement shall describe the candidate's and co-authors' independent research contributions in the thesis publications. For each publication there should exist a complete statement that is to be filled out and signed by the candidate and supervisor **(only required where there isn't already a statement of contribution within the paper itself).**

Title of Paper	*Low-Temperature Sampling on Sparse Random Graphs*
Publication Status	☒Published □ Accepted for Publication □Submitted for Publication □Unpublished and unsubmitted work written in a manuscript style
Publication Details	An extended abstract appeared at *Andreas Galanis, Leslie Ann Goldberg, Paulina Smolarova: Low-Temperature Sampling on Sparse Random Graphs. ICALP 2025: 83:1-83:20* The full version is under review for Electronic Journal of Probability

Student Confirmation

Student Name:	Paulina Smolarova		
Contribution to the Paper	As is standard in our field, all authors contribute equally and authors are listed in alphabetic order		
Signature		Date	4.6. 2026

Supervisor Confirmation

By signing the Statement of Authorship, you are certifying that the candidate made a substantial contribution to the publication, and that the description described above is accurate.

Supervisor name and title: Prof Leslie Ann Goldberg, Prof Andreas Galanis		
Supervisor comments: Paulina contributed substantially to all aspects		
Signature: Leslie Ann Goldberg Andreas Galanis	Date	4.6. 2026

This completed form should be included in the thesis, at the end of the relevant chapter.

Chapter 7

Conclusion and open questions

In this thesis, we explored the feasibility of local Markov Chain Monte Carlo methods in regimes where the worst-case initialisations lead to slow convergence.

Specifically, our focus was on the random-cluster Glauber dynamics. This simple Markov chain is known to converge to the random-cluster model Gibbs distribution; however, its mixing time is slow around the critical temperature due to the bottleneck separating the two phases. We showed that for polynomial-time mixing, it suffices to initialise the Glauber dynamics in the ground state of the dominating phase (or a random mixture of the ground states), and this initialisation can even lead to optimal mixing times in some regimes.

Our approach builds on the framework first introduced by Gheissari and Sinclair [43] for the Ising model. In particular, it reduces establishing fast mixing from ground-state initialisations to verifying three properties: low escape probability from a phase, rapid mixing within local neighbourhoods, and a weak spatial mixing condition. The key technical challenge was establishing weak spatial mixing within the phase, that is, showing that the influence of distant edges decays sufficiently fast when configurations are restricted to be within the phase. Establishing this property required new arguments specific to the different models and parameter regimes considered in this thesis. We formalised these conditions in Theorem 1.5, proven in Chapter 3. The theorem generalises the method from [43] into a unified framework for establishing rapid mixing beyond worst-case analysis for monotone systems.

The primary contribution of this thesis was to apply this framework to the random-cluster Glauber dynamics on random graphs. We obtained fast mixing of the Glauber dynamics for all temperatures and a wide range of q on random regular graphs. Our method furthermore extends to graph families of unbounded degree satisfying weaker expansion properties than those required by previous polymer-based approaches. As

a result, we obtained the first efficient algorithm for sampling from the random-cluster model at low temperatures for the sparse Erdős-Rényi graphs.

7.1 Bounded-degree graphs

In Chapters 4 and 5, we proved rapid mixing of the RC dynamics and hence obtained efficient approximate sampling algorithms applicable to random Δ-regular graphs with high probability.

Our first approach, developed in Chapter 4, used the polymer framework of Helmuth, Jenssen and Perkins [49]. We reduced the weak spatial mixing condition to proving that there are no large polymers. Thus, we obtained a fast sampling algorithm at *all temperatures* for $q = \Delta^{\Omega(\Delta)}$, matching the regime of the cluster expansion algorithm from [49]. Our algorithm has near-linear running time in the high-temperature regime ($\beta < \beta_c$) for all q, and in the low-temperature regime ($\beta > \beta_c$) for integer q.

To relax the restriction $q = \Delta^{\Omega(\Delta)}$, in Chapter 5, we replaced the polymer framework [49] with a more delicate analysis using planting of the Potts model, building on the work of Coja-Oghlan, Galanis, Goldberg, Ravelomanana, Štefankovič and Vigoda [25]. This allowed us to substantially relax the requirements on integer q. In particular, we established optimal $O(n \log n)$ mixing of the RC dynamics for all integer $q, \Delta \geq 3$ in the high-temperature regime ($\beta < \beta_c$), and for integers $\Delta \geq 3$ and $q \geq (5\Delta)^5$ in the low-temperature regime ($\beta > \beta_c$). This result completely resolves sampling for the Potts model for $\beta < \beta_c$ and leaves only the range $3 \leq q < (5\Delta)^5$ open in the low-temperature regime.

7.1.1 All parameter ranges on random regular graphs

A natural follow-up question is whether fast mixing within the phase extends to all real $q \geq 3$. The current lower bounds on q were inherited from the limitations of our methods, and we do not expect them to be tight. We needed $q \geq (5\Delta)^5$ for integer q in Chapter 5 to establish a strong bound on the size of the giant component, and $q = \Delta^{\Omega(\Delta)}$ in Chapter 4 to obtain a variant of the Kotecký–Preiss condition [56]. We note that the bounds on q obtained in our analysis become less restrictive as β moves away from criticality.

We pose the following conjecture.

Conjecture 7.1. *Let $\Delta \geq 3$ be an integer, and let $q \geq 3$ and $\beta > 0$ be reals. Then, with high probability over a uniformly random Δ-regular graph $G = (V, E)$, the following holds:*

- *For $\beta < \beta_c$, the mixing time of the RC dynamics initialised from $X_0 = \emptyset$ is polynomial.*
- *For $\beta > \beta_c$, the mixing time of the RC dynamics initialised from $X_0 = E$ is polynomial.*
- *For $\beta = \beta_c$, the mixing time of the RC dynamics initialised from a random mixture of $\emptyset$ and E, with probabilities proportional to the relative proportions of the phases, is polynomial.*

Answering the conjecture positively immediately yields an FPAUS for the Potts model at all temperatures $\beta \neq \beta_c$ on random regular graphs. To obtain a fast sampling algorithm for $\beta = \beta_c$, we additionally need to efficiently approximate the relative weights of the two phases in order to choose the initial mixture. For this, we require the conjecture to hold w.h.p. over random regular graphs simultaneously for all temperatures $\beta > 0$, so that we can apply the standard FPAUS-to-FPRAS conversion, whose proof we gave in Section 3.2.

We can also ask for a strengthening of the conjecture, namely, *optimal mixing* within the phase.

Conjecture 7.2. *Conjecture 7.1 holds with $O(n \log n)$ mixing time.*

The main barrier to obtaining optimal mixing, even for regimes where polynomial mixing is otherwise established, is a lack of optimal mixing-time bounds for the Glauber dynamics on a regular *wired* tree for non-integer q. For integer q, we can use a comparison with the *block dynamics* for the Potts model, that is, a Markov chain which in each step resamples the spins of some subgraph (as opposed to the Potts Glauber dynamics, which resamples a spin of a single vertex). There are standard techniques [61] translating mixing bounds for the block dynamics into mixing bounds for the Potts Glauber dynamics. In turn, fast mixing of the Potts Glauber dynamics can be translated into fast mixing of the RC dynamics [15]. However, the block-dynamics comparison [61] does not extend directly to the random-cluster model, since it crucially relies on weights being defined via interactions between adjacent vertices. Therefore, this approach requires a comparison with the Potts model and hence only applies when q is an integer.

We note that there has been recent progress in this regard: Blanca, Gheissari, Park and Zhang [16] establish near-optimal mixing on a wired tree for $\beta > \beta'_u$ and all real $q \geq \Omega(\log \Delta)$, and use this to obtain $O(n^{1+o(1)})$ worst-case mixing time for the random-cluster dynamics in these parameter ranges. However, in the critical window $\beta \in (\beta_c, \beta'_u)$, our results remain the best known.

7.1.2 Fast mixing on all bounded-degree graphs

While the problem of approximate sampling from the Potts and random-cluster models on all bounded-degree graphs is #BIS-hard for $\beta > \beta_c$, the situation for $\beta \leq \beta_c$ remains open. In contrast, for the bipartite hardcore model, fast mixing of the Glauber dynamics is known up to and including the computational-hardness threshold.

Although an exponential slowdown of the RC dynamics is known in the non-uniqueness regime for general bounded-degree graphs, it has not been investigated whether the slow mixing persists for well-chosen initialisations. Even in the uniqueness regime, where efficient non-MCMC algorithms are known for the Potts model on all graphs of maximum degree Δ when Δ is sufficiently large [19], it remains open whether the RC dynamics mixes rapidly on all bounded-degree graphs. Together with our results showing that a suitable initialisation can circumvent worst-case slow mixing on random graphs, this motivates the following conjecture.

Conjecture 7.3. *Let $\Delta \geq 3$ be an integer, and let $q \geq 3$ and $0 < \beta < \beta_c$ be reals. Then there exists $C = C(\Delta, q, \beta)$ such that for* all *graphs $G = (V, E)$ of maximum degree Δ, the mixing time of the Glauber dynamics initialised from $\emptyset$ is $O(|V|^C)$.*

We remark that resolving the conjecture for integer q would already be a major breakthrough.

7.2 Sparse graphs of unbounded degree

Further, in Chapter 6, we extended our results to sparse Erdős-Rényi graphs $\mathcal{G}(n, d/n)$. In particular, we established fast mixing from a suitable ground state for all temperatures when d and q are sufficiently large, which we then used to obtain an FPRAS for the partition function. Notably, our algorithms constitute the first low-temperature sampling and approximation algorithms for the random-cluster model, and more generally #BIS-hard spin systems, on $\mathcal{G}(n, d/n)$. Prior to our work, the only low-temperature algorithms on sparse graphs of unbounded degree for #BIS-hard problems required the graph to be an expander and thus were restricted to

graphs with minimum degree at least three. At high temperatures, the previously known algorithms for $\mathcal{G}(n, d/n)$ were restricted to the uniqueness regime.

A major technical contribution of this work was the development of a new simple polymer model for the random-cluster model, in which polymers are defined directly in terms of the vertices of "small" components of in-edges and the components that would arise after removing a single vertex. This simplifies the previous iterative definition from [49]. Our model furthermore applies to a broader class of graphs with weaker expansion and unbounded degree.

We also obtained a new efficient deterministic algorithm for approximating the partition function of the random-cluster model within an arbitrary inverse-polynomial error n^{-R} (provided that d and q are large enough with respect to R) in time $n^{1+O(1/R)}$. This approach was not based on MCMC methods, but instead combined weak spatial mixing within the phase on locally treelike graphs with symbolic integration. It would be interesting to investigate whether the method can be applied to other models for which analogous weak spatial mixing results are available.

7.2.1 Algorithms for $\mathcal{G}(n, d/n)$ for all parameter ranges

Our results relied on first taking d to be a large constant to obtain sufficient expansion and then $q = d^{\Omega(d)}$ to establish a variant of the Kotecký–Preiss condition [56]. These constraints, however, are unlikely to be optimal. We thus pose the following conjecture.

Conjecture 7.4. *Let $d, q \geq 3$ and $\beta > 0$ be reals. Then there is a small window (β_0, β_1) such that with high probability over $(V, E) = G \sim \mathcal{G}(n, d/n)$, the following holds:*

- *For $\beta \leq \beta_0$, the mixing time of the RC dynamics initialised from $X_0 = \emptyset$ is polynomial.*
- *For $\beta \geq \beta_1$, the mixing time of the RC dynamics initialised from $X_0 = E$ is polynomial.*
- *For $\beta \in (\beta_0, \beta_1)$, the mixing time of the RC dynamics initialised from a random mixture of $\emptyset$ and E, with probabilities proportional to the relative proportions of the phases, is polynomial. Moreover, these proportions can be efficiently approximated.*

Relaxing the requirements on q would be a major breakthrough even for q integer, as it would substantially extend the range of parameters for which efficient algorithms are known for the Potts model on $\mathcal{G}(n, d/n)$. In particular, it is of interest whether the planting technique from Chapter 5 applies to $\mathcal{G}(n, d/n)$ for arbitrary $q, d \geq 3$. This would involve adapting the planting framework from [25] to this setting. This appears more feasible in the disordered regime, but more difficult in the ordered regime, where the variability of the Poisson-tree neighbourhoods is expected to lead to larger fluctuations.

One can also consider a strengthening of Conjecture 7.4 as follows.

Conjecture 7.5. *Conjecture 7.4 holds with $O(n \log n)$ mixing time.*

While the conjecture asks for optimal $O(n \log n)$ mixing time, already $O(n^{1+o(1)})$ mixing would be a significant improvement over the polynomial mixing time established in this thesis in the low-temperature regime. The main barrier to improving the low-temperature mixing time was the lack of tight mixing bounds for the RC dynamics on wired treelike neighbourhoods, where we were only able to obtain a high-polynomial mixing time.

Recent partial progress towards this conjecture was made in joint work with Chen, Galanis, Goldberg, Mifsud, and Zhang [23]. By adapting the block-dynamics approach from [61] to irregular trees (and in particular to Poisson trees, which model the neighbourhoods of $\mathcal{G}(n, d/n)$), we obtained near-linear mixing time for the RC dynamics on a wired tree and consequently improved the mixing time for $\beta \geq \beta_1$ in Theorem 1.3 to $O(n^{1+o(1)})$, for all sufficiently large d and sufficiently large integer q.

7.2.2 Critical threshold for $\mathcal{G}(n, d/n)$

Our results establish a small critical interval (β_0, β_1) rather than a sharp critical threshold for the random-cluster model on $\mathcal{G}(n, d/n)$. While we used the same β_0, β_1 throughout all parameter regimes considered in Chapter 6, by taking d and q sufficiently large w.r.t. an arbitrarily small constant $\varepsilon > 0$, the critical interval can be taken between $\log(1 + q^{(2-\varepsilon)/d})$ and $\log(1 + q^{(2+\varepsilon)/d})$. Thus, the critical interval can be taken to be an arbitrarily small interval around $(1 + o_q(1))\frac{2}{d} \log q$, which is the asymptotic value of the critical threshold β_c for random regular graphs [49, 8]. We therefore conjecture that a sharp critical threshold also exists for $\mathcal{G}(n, d/n)$ and coincides with that of random regular graphs.

Conjecture 7.6. *For all real $d, q \geq 3$, the random-cluster model on $\mathcal{G}(n, d/n)$ experiences a sharp phase transition between the disordered and ordered phases at*

$$\beta_c = \log \frac{q-2}{(q-1)^{1-2/d} - 1}.$$

Establishing the conjecture even for the Potts model (integer q) would already be a major step towards understanding the random-cluster model on Erdős-Rényi graphs.

7.3 Extension to other monotone models

Our results established fast mixing from ground states for the random-cluster model for a wide range of parameters on sparse random graphs. Notably, the relaxed Kotecký–Preiss condition used in Chapters 4 and 6 also allowed us to obtain algorithms for Erdős-Rényi graphs, a setting that posed a barrier to other standard low-temperature sampling techniques. This suggests that the framework may also be useful in settings beyond the random-cluster and Ising models [43].

In particular, since Theorem 1.5 applies more generally to monotone models, it would be of interest to investigate whether our method can be applied to the bipartite hardcore model on sparse random graphs of unbounded degree. In this model, there are two coexisting phases at low temperatures, corresponding to independent sets predominantly containing vertices on one or the other side of the bipartition. The proof of Theorem 1.5 is easily adapted (see Section 3.1.3), but applying the theorem would require establishing the three key properties for the bipartite hardcore model. A separate algorithmic challenge is to efficiently estimate the relative weights of the phases in order to choose the appropriate initialisation. Existing approximation algorithms make this step substantially more involved, potentially losing some of the simplicity and running-time advantage of the Glauber-dynamics approach. Alternatively, at the cost of allowing only inverse-polynomial error (as opposed to arbitrarily small error), our deterministic approximation approach based on WSM within the phase should also apply in this setting.

Bibliography

[1] Dimitris Achlioptas and Amin Coja-Oghlan. Algorithmic barriers from phase transitions. In *2008 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 793–802, 2008. doi: 10.1109/FOCS.2008.11.

[2] Ahmed El Alaoui, Andrea Montanari, and Mark Sellke. Sampling from the Sherrington-Kirkpatrick Gibbs measure via algorithmic stochastic localization. In *63rd IEEE Annual Symposium on Foundations of Computer Science, FOCS 2022*, pages 323–334, 2022.

[3] Noga Alon, Itai Benjamini, and Alan Stacey. Percolation on finite graphs and isoperimetric inequalities. *The Annals of Probability*, 32(3):1727 – 1745, 2004.

[4] Nima Anari, Kuikui Liu, and Shayan Oveis Gharan. Spectral independence in high-dimensional expanders and applications to the hardcore model. *SIAM Journal on Computing*, 53(6):FOCS20–1–FOCS20–37, 2024. doi: 10.1137/20M1367696.

[5] Cécile Ané, Sébastien Blachère, Djalil Chafai, Pierre Fougères, Ivan Gentil, Florent Malrieu, Cyril Roberto, and Grégory Scheffer. Sur les inégalités de sobolev logarithmiques. *HAL*, 2000(0), 2000. URL `http://dml.mathdoc.fr/item/hal-00180778`.

[6] Victor Bapst, Amin Coja-Oghlan, and Charilaos Efthymiou. Planting colourings silently. *Combinatorics, Probability and Computing*, 26(3):338–366, 2017.

[7] 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.

[8] Ferenc Bencs, Márton Borbényi, and Péter Csikvári. Random cluster model on regular graphs. *Communications in Mathematical Physics*, pages 1–46, 2022.

[9] Ivona Bezáková, Andreas Galanis, Leslie Ann Goldberg, and Daniel Štefankovič. Fast sampling via spectral independence beyond bounded-degree graphs. *ACM Trans. Algorithms*, 20(1), January 2024. doi: 10.1145/3631354.

[10] Antonio Blanca and Reza Gheissari. Random-cluster dynamics on random regular graphs in tree uniqueness. *Communications in Mathematical Physics*, 386 (2):1243–1287, 2021.

[11] 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.

[12] Antonio Blanca and Reza Gheissari. On the tractability of sampling from the Potts model at low temperatures via random-cluster dynamics. *Probability Theory and Related Fields*, 191:1121–1168, 06 2024. doi: 10.1007/s00440-024-01289-x.

[13] Antonio Blanca and Xusheng Zhang. Rapid mixing of global Markov chains via spectral independence: the unbounded degree case. *Algorithmica*, 88(1), 2026. doi: 10.1007/S00453-025-01353-5.

[14] 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.

[15] Antonio Blanca, Zongchen Chen, Daniel Štefankovič, and Eric Vigoda. The Swendsen–Wang dynamics on trees. *Random Structures & Algorithms*, 62(4): 791–831, 2023.

[16] Antonio Blanca, Reza Gheissari, Heehyun Park, and Xusheng Zhang. Uniqueness and mixing in the low-temperature random-cluster model on trees and random graphs, 2026. URL `https://arxiv.org/abs/2604.20693`.

[17] Magnus Bordewich, Catherine Greenhill, and Viresh Patel. Mixing of the Glauber dynamics for the ferromagnetic Potts model. *Random Structures & Algorithms*, 48(1):21–52, 2016. doi: doi.org/10.1002/rsa.20569.

[18] 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.

[19] Charlie Carlson, Ewan Davies, Nicolas Fraiman, Alexandra Kolla, Aditya Potukuchi, and Corrine Yap. Algorithms for the ferromagnetic Potts model on expanders. *Combinatorics, Probability and Computing*, 33:487–517, 7 2024. doi: 10.1017/S0963548324000087.

[20] Xiaoyu Chen, Zongchen Chen, Yitong Yin, and Xinyuan Zhang. Rapid mixing at the uniqueness threshold. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, STOC '25, page 879–890. Association for Computing Machinery, 2025. doi: 10.1145/3717823.3718260.

[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.

[22] Zongchen Chen, Andreas Galanis, Leslie Ann Goldberg, Heng Guo, Andrés Herrera-Poyatos, Nitya Mani, and Ankur Moitra. Fast sampling of satisfying assignments from random k-SAT with applications to connectivity. *SIAM Journal on Discrete Mathematics*, 38(4):2750–2811, 2024. doi: 10.1137/23M1595722.

[23] Zongchen Chen, Andreas Galanis, Leslie Ann Goldberg, Xandru Mifsud, Paulina Smolarova, and Xusheng Zhang. Fast mixing for low-temperature Potts models via Poisson trees, 2026. URL `https://arxiv.org/abs/2607.29495`.

[24] Amin Coja-Oghlan, Florent Krzakala, Will Perkins, and Lenka Zdeborová. Information-theoretic thresholds from the cavity method. *Advances in Mathematics*, 333:694–795, 2018.

[25] 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.

[26] Jian Ding, Eyal Lubetzky, and Yuval Peres. Anatomy of the giant component: the strictly supercritical regime. *European Journal of Combinatorics*, 35:155–168, 2014.

[27] Martin E. Dyer, Leslie Ann Goldberg, Catherine S. Greenhill, and Mark Jerrum. The relative complexity of approximate counting problems. *Algorithmica*, 38(3): 471–500, 2004. doi: 10.1007/S00453-003-1073-Y.

[28] Charilaos Efthymiou. On sampling symmetric Gibbs distributions on sparse random graphs and hypergraphs. In *49th International Colloquium on Automata, Languages, and Programming (ICALP 2022)*, volume 229, pages 57:1–57:16, 2022.

[29] 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.

[30] 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.

[31] Nikolaos Fountoulakis and Bruce A. Reed. The evolution of the mixing rate of a simple random walk on the giant component of a random graph. *Random Structures & Algorithms*, 33(1):68–86, 2008. doi: 10.1002/rsa.20210.

[32] Alan Frieze and Michał Karoński. *Introduction to Random Graphs*. Cambridge University Press, 2015.

[33] Andreas Galanis, Leslie Ann Goldberg, and Mark Jerrum. Approximately Counting H-Colorings is #BIS-Hard. *SIAM Journal on Computing*, 45(3):680–711, 5 2016. doi: 10.1137/15M1020551.

[34] 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.

[35] Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast algorithms for general spin systems on bipartite expanders. *ACM Transactions on Computation Theory*, 13(4), 2021. doi: 10.1145/3470865.

[36] Andreas Galanis, Leslie Ann Goldberg, and James Stewart. Fast mixing via polymers for random graphs with unbounded degree. *Information and Computation*, 285, 2022. doi: 10.1016/J.IC.2022.104894.

[37] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Sampling from the random cluster model on random regular graphs at all temperatures via Glauber dynamics. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (APPROX/RANDOM 2023)*, volume 275 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 64:1–64:12, 2023. doi: 10.4230/LIPIcs.APPROX/RANDOM.2023.64.

[38] 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*, 34(3):359–391, 2025. doi: 10.1017/S0963548324000403.

[39] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Low-temperature sampling on sparse random graphs. In *52nd International Colloquium on Automata, Languages, and Programming (ICALP 2025)*, volume 334 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 83:1–83:20, 2025. doi: 10.4230/LIPIcs.ICALP.2025.83.

[40] Andreas Galanis, Leslie Goldberg, and Paulina Smolarova. Low-temperature sampling on sparse random graphs. *Electronic Journal of Probability*, 31:1–40, 01 2026. doi: 10.1214/26-EJP1576.

[41] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model. In *43rd International Symposium on Theoretical Aspects of Computer Science (STACS 2026)*, volume 364 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 39:1–39:19, 2026. doi: 10.4230/LIPIcs.STACS.2026.39.

[42] Andreas Galanis, Leslie Ann Goldberg, and Paulina Smolarova. Planting and MCMC sampling from the Potts model, 2026. URL `https://arxiv.org/abs/2410.14409`.

[43] Reza Gheissari and Alistair Sinclair. Low-temperature Ising dynamics with random initializations. *The Annals of Applied Probability*, 33(5):3916 – 3957, 2023. doi: 10.1214/22-AAP1911.

[44] Reza Gheissari and Alistair Sinclair. Spatial mixing and the random-cluster dynamics on lattices. *Random Structures & Algorithms*, 64(2):490–534, 2024. doi: https://doi.org/10.1002/rsa.21191.

[45] Leslie Ann Goldberg and Mark Jerrum. Approximating the partition function of the ferromagnetic Potts model. *Journal of the ACM*, 59(5):1–31, 2012.

[46] Leslie Ann Goldberg, Rob Gysel, and John Lapinskas. Approximately counting locally-optimal structures. *Journal of Computer and System Sciences*, 82(6): 1144–1160, 2016. ISSN 0022-0000. doi: 10.1016/j.jcss.2016.04.001.

[47] Geoffrey Grimmett. *The Random-Cluster Model*, volume 333. Springer, 2006.

[48] Tyler Helmuth, Will Perkins, and Guus Regts. Algorithmic Pirogov–Sinai theory. *Probability Theory and Related Fields*, 176:851–895, 4 2020. doi: 10.1007/s00440-019-00928-y.

[49] 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 Poincare (B) Probabilites et statistiques*, 59(2):817–848, 2023.

[50] Shang-En Huang, Dawei Huang, Tsvi Kopelowitz, Seth Pettie, and Mikkel Thorup. Fully dynamic connectivity in $O(\log n(\log \log n)^2)$ amortized expected time. *TheoretiCS*, Volume 2, 5 2023. doi: 10.46298/THEORETICS.23.6.

[51] Svante Janson, Tomasz Łuczak, and Andrzej Ruciński. *Random graphs*. John Wiley & Sons, Inc., 2000 - 2000.

[52] Matthew Jenssen, Peter Keevash, and Will Perkins. Algorithms for #BIS-hard problems on expander graphs. *SIAM Journal on Computing*, 49(4):681–710, 2020.

[53] Mark Jerrum. *Counting, Sampling and Integrating: Algorithms and Complexity*. Springer Science & Business Media, 2003.

[54] Mark Jerrum and Alistair Sinclair. Polynomial-time approximation algorithms for the Ising model. *SIAM Journal on Computing*, 22(5), 1993. doi: 10.1137/0222066.

[55] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43(C):169–188, 1 1986. doi: 10.1016/0304-3975(86)90174-X.

[56] Roman Kotecký and David Preiss. Cluster expansion for abstract polymer models. *Communications in Mathematical Physics*, 103(3), 1986. doi: 10.1007/BF01211762.

[57] Michael Krivelevich, Eyal Lubetzky, and Benny Sudakov. Asymptotics in percolation on high-girth expanders. *Random Structures and Algorithms*, 56(4): 927–947, 2020.

[58] David A Levin and Yuval Peres. Markov Chains and Mixing Times. *American Mathematical Society*, 2009.

[59] 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.

[60] Russell Lyons and Yuval Peres. *Branching Processes, Second Moments, and Percolation*, pages 131–173. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2017.

[61] Fabio Martinelli, Alistair Sinclair, and Dror Weitz. Glauber dynamics on trees: boundary conditions and mixing time. *Communications in Mathematical Physics*, 250(2), 2004. doi: 10.1007/s00220-004-1147-y.

[62] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[63] Joel Moses. Symbolic integration: the stormy decade. *Commun. ACM*, 14(8): 548–560, 1971. doi: 10.1145/362637.362651.

[64] Elchanan Mossel and Allan Sly. Exact thresholds for Ising–Gibbs samplers on general graphs. *The Annals of Probability*, 41(1):294 – 328, 2013. doi: 10.1214/11-AOP737.

[65] C. Andrew Neff and John H. Reif. An efficient algorithm for the complex roots problem. *Journal of Complexity*, 12(2):81–115, 1996.

[66] Sanjeev Saluja, K. V. Subrahmanyam, and Madhukar N. Thakur. Descriptive complexity of #P functions. *Journal of Computer and System Sciences*, 50(3): 493–505, 6 1995. doi: 10.1006/JCSS.1995.1039.

[67] Seinosuke Toda. PP is as hard as the polynomial-time hierarchy. *SIAM Journal on Computing*, 20(5):865–877, 1991. doi: 10.1137/0220053.

[68] Luca Trevisan. Lecture notes on graph partitioning, expanders and spectral methods, 2016. URL: `https://lucatrevisan.github.io/books/expanders-2016.pdf`.

[69] Leslie G. Valiant. The complexity of enumeration and reliability problems. *SIAM Journal on Computing*, 8(3):410–421, 1979. doi: 10.1137/0208032.

[70] 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.