PHASE TRANSITIONS IN SPIN SYSTEMS: UNIQUENESS, RECONSTRUCTION AND MIXING TIME

A Thesis
Presented to
The Academic Faculty

by

Linji Yang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in
Algorithms, Combinatorics, and Optimization

School of Computer Science
Georgia Institute of Technology
May 2013

PHASE TRANSITIONS IN SPIN SYSTEMS: UNIQUENESS, RECONSTRUCTION AND MIXING TIME

Approved by:

Professor Eric Vigoda, Advisor
School of Computer Science
Georgia Institute of Technology

Professor Dana Randall
School of Computer Science
Georgia Institute of Technology

Professor Daniel Štefankovič
Department of Computer Science
University of Rochester

Professor Prasad Tetali
School of Mathematics
Georgia Institute of Technology

Professor Santosh Vempala
School of Computer Science
Georgia Institute of Technology

Date Approved: 14 March 2013

For my parents

and my wife Pengyi Shi,

who always love and support me.

ACKNOWLEDGEMENTS

Undertaking this Ph.D. has been a life-changing experience for me. This journey would not have been possible to do without the support from many people.

First of all, I want to express my thanks to my thesis advisor, Prof. Eric Vigoda, for his patient encouragement and thoughtful guidance in the last six years. I appreciate all his contributions of time, ideas and funding to make my Ph.D. experience very enjoyable and stimulating. His ability to select and to approach compelling research problems, and his high scientific standards set an example for me.

Besides my advisor, I would like to thank Prof. Dana Randall, Prof. Prasad Tetali and Prof. Santosh Vempala in my department, for their insightful discussions on various research topics, sharp comments on my research and generous sharing of their immense knowledge with me. Taking their courses has always been the most efficient, interesting and comprehensive way of learning. Also, I would like to thank the ACO program director, Prof. Robin Thomas, for his great effort on running and improving this program with high standard of requirements.

During the research, I feel lucky to work with my co-authors Andreas Galanis, Dr. Ricardo Restrepo, Dr. Jinwoo Shin, Prof. Daniel Štefankovič, Prof. Prasad Tetali and Prof. Juan Vera. Many of my research results would not have been possible without their wide knowledge, innovative ideas and intelligent discussions with me. Special thanks to Daniel for inviting me to visit University of Rochester several times and to Juan who invited me to visit Tilburg University, Netherlands.

My most profound thanks go to my parents who have constantly encouraged me and believed in me in the pursuit of my studies.

Finally, I want to express my heartfelt thanks to my wife Pengyi Shi for her love

and encouragement. And, thank you for your support when I have needed it the most.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Spin systems are powerful mathematical models widely used and studied in Statistical Physics and Computer Science. This thesis focuses on two specific spin systems of particular combinatorial interest: colorings and weighted independent sets (hard-core model).

In many spin systems, there exists a phase transition phenomenon: there is a threshold value of a parameter such that when the parameter is on one side of the threshold, the system exhibits the so-called spatial decay of correlation, i.e., the influence from a set of vertices to another set of vertices diminishes as the distance between the two sets grows; when the parameter is on the other side, long range correlations persist. The uniqueness problem and the reconstruction problem are two major threshold problems that are concerned with the decay of correlations in the Gibbs measure from different perspectives.

In Computer Science, the study of spin systems mainly focused on finding an efficient algorithm that samples the configurations from a distribution that is very close to the Gibbs measure. The Glauber dynamics is a typical Markov chain algorithm for conducting such sampling. In many systems, the convergence time of the Glauber dynamics also exhibits a threshold behavior: the speed of convergence experiences a dramatic change around the threshold of the parameter.

The first two parts of this thesis focus on making connections between the phase transition of the convergence time of the dynamics and the phase transition of the reconstruction phenomenon in both colorings and the hard-core model on regular trees. A relatively sharp threshold is established for the change of the convergence time, which coincides with the reconstruction threshold. A general technique of upper

bounding the conductance of the dynamics via analyzing the sensitivity of the reconstruction algorithm is proposed and proven to be very effective for lower bounding the convergence time of the dynamics.

The third part of the thesis provides an innovative analytical method for establishing a strong version of the decay of correlation of the Gibbs distributions for many two spin systems on various classes of graphs. In particular, the method is applied to the hard-core model on the square lattice, a very important graph that is of great interest in both Statistical Physics and Computer Science. As a result, we significantly improve the lower bound on the uniqueness threshold for the square lattice and thereby improve the range of the parameter of interest where the Glauber dynamics has fast convergence (i.e., rapid mixing).

CHAPTER I

INTRODUCTION

1.1 Background and Motivations

Spin systems are powerful mathematical models widely used and studied in Statistical Physics, Applied Probability and Computer Science. The concept of a spin system originated from Statistical Physics as an idealized model to study a physical system at equilibrium such as magnetic materials and lattice gases. Mathematically, a spin system is comprised of a finite graph $G = (V, E)$ with vertices V modeling the particles or individuals and edges E modeling the interactions. A configuration for a vertex or a set of vertices is an assignment of one of the *spins* to each vertex. Different configurations have different likelihoods of appearing in the model of the equilibrium state of the system. The likelihood comes in two ways: one is from the spin of each vertex, each spin may have a different weight; the second way is from the interactions among the vertices, i.e., there are different weights assigned to edges depending on the different spins at the endpoints of the edge. Indeed, this can be made into a formal probability statement and we are able to define the probability space called the Gibbs measure over the set of configurations, which represents the equilibrium state of the finite system.

For spin systems in Statistical Physics, when studying magnetic materials, the underlying graph G typically becomes a lattice (usually a two-dimensional or three-dimensional grid) and each site represents an atom. The states of a site specify the "up" or "down" magnetic moment of the atom and edges are modeling the interaction among the neighboring atoms. For the lattice gas model, sites are the possible places where atoms of non-negligible size can reside and can be either "occupied" or "vacant" with the constraint that no two neighboring sites can be both occupied at the same time so that atoms will not overlap. Researchers have been using this general modeling tool to understand properties of the equilibrium states of various systems including: how the particles/individuals affect each other in the equilibrium states, how the physical system reaches equilibrium, and also whether long range correlations exist.

1.1.1 Phase Transitions

Perhaps, the most important object in the study of spin systems in Statistical Physics is to understand the *phase transition* phenomenon. A *phase transition* is said to occur when a microscopic change in the parameters of the system causes a dramatic macroscopic change in the properties of the system. It is quite common in our daily life, such as when water boils into steam or freezes to ice. A more sophisticated example is the so-called "*spontaneous magnetization*"; if we put a piece of metal into a magnetic field for a while and then remove the external magnetic field, sometimes the metal is still magnetized, i.e., the magnetic moments of most of the atoms stay in the state "up" (or "down") simultaneously. Physicists discovered that depending on the materials, when the temperature is sufficiently high, the magnetization will not persist after removing the external field, while when the temperature is low, magnetization does persist after the removal and spontaneous magnetization happens. There is a critical temperature for spontaneous magnetization to appear; it only happens when the temperature is below a certain threshold value. Generally, why such phase transitions exist is not an easy question to answer. For spontaneous magnetization, this is exactly what Ernst Ising intended to answer in his original study of the now famous Ising model in the 1920s [33].

The (ferromagnetic) Ising model is the simplest and most widely studied example of a spin system. It is defined on a finite graph $G = (V, E)$. In this model, each vertex is assigned a spin from the set $Q = \{-1, +1\}$. A configuration σ is a function from V to Q. Each configuration is weighted by

$$w(\sigma) = \exp\left(\beta \sum_{(u,v)\in E} \sigma(u)\sigma(v)\right),$$

where $\beta > 0$ is a real number called the inverse temperature. The entire space of the configurations with positive weight is denoted as $\Omega = \Omega_G$. In the Ising model, all the assignments have positive weights and hence it is a *soft-constraint* model. The *partition function* $Z = Z_G = \sum_{\sigma\in\Omega} w(\sigma)$ is the total sum of the weights over all the configurations in Ω. Then, one can define a probability space over Ω, i.e., $\mu(\sigma) = w(\sigma)/Z$. This measure is called the *Gibbs measure* and is one of the main subjects we study in this thesis. It is easy to see that, since $\beta > 0$, in this probability space, the configurations with more edges having the same assignments on both endpoints are more likely to appear.

We can assign a *boundary condition* to a spin system by fixing an assignment of spins to a fixed set of vertices (called the boundary). Different boundary conditions will result in different conditional Gibbs measures on the internal vertices which are not on the boundary. For instance, if the graph G is a Δ-regular tree with height h, one can fix the configurations on the leaves, and this derives a conditional probability space for the configurations on the internal (non-leaf) vertices. In the ferromagnetic Ising model, if we assign the vertices L_h on level h to be all $+1$, then in the conditional Gibbs measure on the non-leaf vertices, the vertices are biased to choose $+1$ instead of -1 as their spins since the model favors those configurations that put the same spins on the both endpoints of each edge. On the other hand, if we assign the boundary with all -1 as their spins, then the internal vertices are biased to choose -1. When no boundary condition is assigned to the system, we say the system has free boundary conditions and the distribution is called the free Gibbs measure.

The Ising model on the square lattice $\mathbb{Z}^2$ is the major model for studying spontaneous magnetization. Mathematically, we can characterize and understand the phenomena of spontaneous magnetization in the following way using boundary conditions. Let r be the origin of the square lattice $\mathbb{Z}^2$ and R_n be the finite n by n rectangle (subgraph) of $\mathbb{Z}^2$ centered at the origin r. Now, we fix the configurations on the boundary of the rectangle R_n to all $+1$ (as shown in Figure 1), one of the extremal boundary conditions that biases the conditional distribution of the origin r in favor of $+1$. This is similar to putting an external magnetic field around a piece of metal. Let $p_{+,n}$ denote the conditional probability of the origin r being $+1$ under the all $+1$ boundary condition for the subgraph R_n. Similarly, one can put all -1 around the rectangle and hence define the corresponding conditional probability $p_{-,n}$ of r being $+1$ under the all -1 boundary condition. In this case, the conditional distribution of the origin is biased toward -1, i.e., $p_{-,n} < 1 - p_{-,n}$. It is not hard to see that the effect of any other boundary condition to r is in between the all $+1$ and all -1 boundary conditions: any boundary condition cannot bias the origin r to favor $+1$ more than the all $+1$ boundary condition can do, and neither can it bias the origin to -1 more than the all -1 boundary can do.

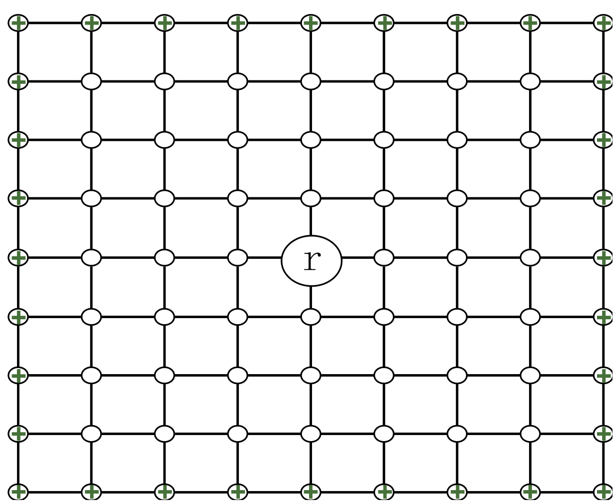

Figure 1: A all + boundary condition for Ising model on the sub-lattice R_9. The conditional probability of the origin r being $+1$ is $p_{+,9}$.

We consider the effect of the boundary condition to r when the boundaries move off to infinity, i.e., let $n \to \infty$, which mimics the removal of the external magnetic field. Hence, let

$$p_+ = \lim_{n\to\infty} p_{+,n} \quad \text{and} \quad p_- = \lim_{n\to\infty} p_{-,n}.$$

Now we ask whether p_+ equals p_-. It turns out that there is a critical inverse temperature $\beta_c(\mathbb{Z}^2)$ such that if $\beta < \beta_c(\mathbb{Z}^2)$ (the temperature is high), then $p_+ = p_-$, and we say the model exhibits *decay of correlation* from the boundary to vertices deep in the interior and when $\beta > \beta_c(\mathbb{Z}^2)$ (the temperature is low), then $p_+ \neq p_-$, and therefore the effect of the boundary condition persists, which corresponds to the persistence of the magnetization after removing the external magnetic field. The value of the critical temperature $\beta_c(\mathbb{Z}^2)$ has been known since 1944 by Onsager [53]. An alternative formulation of the above phase transition phenomenon is known as the *uniqueness problem* asking whether the infinite Gibbs measure on the infinite graph $\mathbb{Z}^2$ is unique or not. The critical value for this phase transition is hence called the *uniqueness threshold.* We will discuss the details of infinite-volume Gibbs measures in Chapter 2.2. Behind those rigorous mathematical formulations for infinite-volume Gibbs measures, it is the existence of decay of correlations from the boundaries to the origin over the finite subgraphs that plays the essential role in the uniqueness problem.

Besides the square lattice, uniqueness problem can be defined for spin systems on other infinite graphs, such as the Δ-regular infinite tree $\mathbb{T}_\Delta$. For the Ising model, the critical value for the uniqueness of infinite Gibbs measure on $\mathbb{T}_\Delta$ is at $\beta_c(\mathbb{T}_\Delta) = \arctan(\frac{1}{\Delta-1})$ (see, e.g., Preston [56]). This critical value is called the *tree uniqueness threshold* for the Ising model. Generally, the uniqueness problem is equivalent to the decay of correlation from the all $+1$ and the all -1 boundary conditions to the origin r as $n \to \infty$. When the spin system is in the uniqueness phase, there is no *long range correlation* among the vertices and hence different boundary conditions have the same limiting influence on internal vertices; on the other hand, when the system is in the non-uniqueness phase, the effect of the worst case boundary conditions persist even when the boundaries are very far away. In this thesis, we will develop a general technique that can be used to improve lower bounds of conjectured uniqueness thresholds for various spin systems.

1.1.2 Glauber Dynamics and Phase Transition of Mixing Time

In Computer Science, the study of spin systems mainly focuses on finding efficient algorithms for sampling configurations from distributions that are very close to the corresponding Gibbs measures. It was shown by Jerrum, Valiant and Vazirani [36] that having an efficient sampling algorithm for the Gibbs measure is equivalent to having a fully polynomial randomized approximation scheme (FPRAS) for the corresponding partition function of the spin system. Typically, computing the partition function of a spin system exactly is known to be #P-complete (see, e.g., [34]). It is also interesting to study the dynamical questions in a finite system: how does the system evolve from an initial pre-assigned configuration to the equilibrium state, i.e., the Gibbs measure, and how long does it take to do so. A key stochastic process to study is called the *Glauber dynamics* which, on the one hand, is considered a basic model for simulating the actual evolution of the physical system and, on the other hand, is the heart of Monte Carlo Markov chain algorithms for performing the sampling from Gibbs measures. The Glauber dynamics is a Markov chain on the state space Ω. The "heat-bath" Glauber dynamics works in the following way: at each step we pick a random vertex v, all vertices other than v maintain the same spins and we update the spin of v according to the Gibbs measure conditioning on the spins of the neighbors of v. The heat-bath Glauber dynamics is an ergodic and reversible Markov chain with the Gibbs measure $\mu = \mu_G$ as its stationary distribution.

There are several ways to measure the convergence speed of Markov chains. The mixing time T_{mix} of the Glauber dynamics is the number of steps the dynamics needs to perform so that starting from any initial configuration, the final distribution is close to the stationary distribution. When the mixing time is polynomial in the number of vertices, the Glauber dynamics naturally gives an efficient sampling algorithm for the Gibbs measure, and in this case, we say the Markov chain is *rapidly mixing*. The relaxation time T_{relax}, defined as the inverse of the spectral gap of the Markov chain, is a different measurement of how fast the dynamics converges to the equilibrium measure. Readers can refer to Chapter 2.3 for formal definitions of the above concepts.

The pursuit of understanding the speed of convergence of Markov chains on spin systems never comes to an end. In the last two decades, numerous results have been proven for both the upper and lower bounds on the mixing time for spin systems under different parameter settings. Generally, it is conjectured that for spin systems on various graphs (such as trees, lattices and graphs with bounded degrees), there is a threshold for the parameter of the system such that the mixing time of the Glauber dynamics experiences drastic change when the parameter varies around the threshold. A major part of this thesis is devoted to develop a methodology that clarifies the threshold behavior of mixing times for certain spin systems on regular trees.

For the Ising model, recent studies (see, e.g., [8, 52, 74, 67]) have been gradually revealing the fundamental reason why such a phase transition for the mixing time could exist, making more explicit connections to the properties of the equilibrium states of the spin systems.

It appears that for general graphs, the mixing time of the Glauber dynamics has a phase transition at the tree uniqueness threshold. For the class of graphs with maximum degree Δ, it was shown by Mossel and Sly [52] that the phase transition for the mixing time of the Glauber dynamics on the Ising model occurs at the tree uniqueness threshold $\beta_c(\mathbb{T}_\Delta) = \operatorname{arctanh}(\frac{1}{\Delta-1})$. They show that when $\beta < \beta_c(\mathbb{T}_\Delta)$, the mixing time of the Glauber dynamics is $O(n \log n)$ where n is the number of vertices in the graph. Previously, it was shown by Gerschenfeld and Montanari [28] that the mixing time of the Glauber dynamics is $\exp(\Omega(n))$ on random Δ-regular graphs with high probability when $\beta > \beta_c(\mathbb{T}_\Delta)$.

However, note that for more restricted classes of graphs, the thresholds for the mixing time of the Glauber dynamics might be different than the tree uniqueness threshold. The phase transition for the mixing time has been known for the Ising model on $\mathbb{Z}^2$ since the early 1990s (see, e.g., [44, 43]). It is shown that when the inverse temperature $\beta < \beta_c(\mathbb{Z}^2)$, in the region that the correlation decays exponentially in the distance from the boundary to the origin, the Glauber dynamics is rapidly mixing. When $\beta > \beta_c(\mathbb{Z}^2)$, the mixing time is lower bounded by $\exp(\Omega(\sqrt{n}))$. In these studies, the connection between the spatial decay of correlation in the Gibbs measure and the mixing time of the Glauber dynamics has been observed and established for the Ising model on the square lattice. The uniqueness threshold of the Ising model on the square lattice is the phase transition threshold of the mixing time for the corresponding Glauber dynamics.

Interestingly, for the Δ-regular tree itself and some of the locally tree-like graphs such as planar graphs, the phase transition for the mixing time does not seem to be at the uniqueness threshold. It is conjectured that the phase transition threshold for the mixing time of the Glauber dynamics in the regular tree case is at the so-called *reconstruction threshold*. This connection is what we investigate next.

1.1.3 The Reconstruction Threshold

For the regular tree $\mathbb{T}_\Delta$, there is another way to generate an infinite Gibbs measure, called *broadcasting* (see, e.g., [51]). As suggested by its name, the process generates spin assignments from the root down to infinity by a homogeneous Markov chain. For example, in the case of Ising model, the broadcasting process works in the following way:

- The root is assigned a spin uniformly at random from $\{-1, +1\}$;
- For each vertex, select the spins for its b children independently at random according to the transition matrix

$$\mathcal{B} = \begin{pmatrix} 1-\epsilon & \epsilon \\ \epsilon & 1-\epsilon \end{pmatrix},$$

  i.e., each child of v will have the same spin as v with probability $1-\epsilon$ and the opposite spin with probability ϵ.

One can easily check that this indeed generates the free Gibbs measure over the whole tree by noticing that if the broadcasting process stops at height h, the distribution of the configurations is exactly the same as the Ising model with $\beta = \frac{1}{2}\ln\frac{1-\epsilon}{\epsilon}$ on the Δ-regular tree of height h with free boundary conditions. An important question related to the broadcasting process is called the *reconstruction problem*. It addresses the question of how much information the root passes to its descendants in the broadcasting process. As we can see during the above broadcasting process, the information of the root's spin is being lost in some sense as the spins pass from the parents to their children. However, when one has all the assignments at the leaves, it may still be possible to recover the root's original spin with a non-trivial probability. The reconstruction problem asks over random boundary conditions generated by the broadcasting process, whether in expectation the assignment of the root can be recovered with a non-trivial probability or not. As compared to the uniqueness problem which concerns the decay of correlation from the worst case boundary conditions (such as the all $+1/-1$ boundary conditions for the Ising model), in the reconstruction problem, it concerns the average decay of correlations for the influence from all the boundary conditions. Readers can refer to Chapter 2.3 for formal details. The reconstruction problem was first studied in Statistical Physics for understanding the free Gibbs measure of the Ising model on the infinite tree. For years, it has also been widely studied in Computational Biology as a method of reconstructing evolutionary trees (see, e.g., [14]), and in Computer Science to model tree communication networks. It turns out that there are again threshold behaviors for reconstruction problems in many spin systems.

A general connection between the reconstruction and the convergence time of the Glauber dynamics was shown by Berger *et al.* [8]. They showed for general spin systems that $O(n)$ relaxation time on the complete tree (with free boundary conditions) implies non-reconstruction on the tree. Recently, for the Ising model, the picture of the mixing time of the Glauber dynamics on the Δ-regular tree is actually completed and the phase transition was shown to happen exactly at the reconstruction threshold. A new work of Ding *et al.* [17] gives very sharp bounds on the mixing time of the Glauber dynamics for the Ising model on the complete tree. They illustrate that the dynamics undergoes a phase transition at the reconstruction threshold by showing that when $\beta = \beta_r = \operatorname{arctanh}(1/\sqrt{\Delta - 1})$, the mixing time is lower bounded by $\Omega(n \log^3 n)$. When $\beta < \beta_r$, the mixing time was already known to be $O(n \log n)$ by Martinelli *et al.* in [46] and when $\beta > \beta_r$, the mixing time is precisely lower bounded by $\Omega(n^{f(\Delta,\beta)})$ for some concrete function $f(\Delta, \beta) > \delta > 1$, as established by Berger *et al.* in [8].

While in the Ising model it seems that the pictures of uniqueness, reconstruction, mixing time and their connections are quite clear for various classes of graphs, it is not the case for other spin systems such as colorings and weighted independent sets. The difficulties arise from the fact that both colorings and weighted independent sets are *hard-constraint* models. They possess different combinatorial properties of their own, which make it impossible to directly apply the existing methods used for Ising models. These two hard-constraint models are the main subjects that we study in this thesis.

1.2 Colorings, Weighted Independent Sets and Our Results

Colorings and weighted independent sets (hard-core model) are the spin systems of particular interest in Combinatorics and Computer Science. In the colorings problem, one is given a finite graph $G = (V, E)$ and a set of k colors. The valid configurations are proper k-colorings of G where each vertex is assigned one of k colors from the set $[k] = \{1, 2, \ldots, k\}$ and for every edge $e \in E$, the colors on the endpoints of e are different. The Gibbs measure is the uniform distribution over all the valid configurations. In the hard-core model, one is given a finite graph $G = (V, E)$ and a real number λ called the activity. The configurations are independent sets (subsets of V where there are no edges between the vertices) of the graph G and each independent set I is weighted by raising the activity $\lambda > 0$ to the power of the number of vertices in I, i.e., $w(I) = \lambda^{|I|}$. The partition function is $Z = \sum_{I \subseteq V} w(I)$ and the Gibbs measure assigns each independent set I with probability $w(I)/Z$. We will formally define these models and concepts in a more general setting later in Chapter 2.1. For k-colorings, Jonasson [37] established that the tree uniqueness threshold is at $k_c(\mathbb{T}_\Delta) = \Delta + 2$. Kelly [38] proved that the tree uniqueness threshold for the hard-core model is at $\lambda_c(\mathbb{T}_\Delta) = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$.

The mixing times of the Glauber dynamics for these two models are heavily studied. For instance, for k-colorings on graphs of maximum degree Δ, Jerrum [35] showed that the Glauber dynamics is rapidly mixing when the number of colors $k \geq 2\Delta$. Later, Vigoda [73] established that the Glauber dynamics is rapidly mixing when $k \geq \frac{11\Delta}{6}$ which is still the best known result for k-colorings on general graphs. When restricting the class of graphs by imposing minimum girth requirements or the planarity, we can establish rapid mixing for smaller k, with respect to Δ [21, 24, 31, 32]. When $k < \Delta + 2$, it is easy to verify that the Glauber dynamics on graphs of maximum degree Δ is not ergodic for some graphs. However, we do not know whether it is true that the Glauber dynamics is always rapidly mixing whenever k is above the tree uniqueness threshold $k_c(\mathbb{T}_\Delta) = \Delta + 2$.

For the hard-core model on graphs of maximum degree Δ, the exact threshold of the phase transition for the Glauber dynamics is not known either. However, recently a remarkable connection was established between the computational complexity of approximating the partition function for graphs of maximum degree Δ and the tree uniqueness threshold $\lambda_c(\mathbb{T}_\Delta)$. On the positive side, Weitz [74] showed a deterministic fully-polynomial time approximation algorithm (FPAS) for approximating the partition function for any graph with maximum degree Δ, when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ and Δ is constant. On the other side, for every $\Delta \geq 3$, it was shown that unless $NP = RP$, there does not exist an FPRAS for the partition function for graphs of maximum degree at most Δ when $\lambda_c(\mathbb{T}_\Delta) < \lambda$, as shown by Sly [67] and later improved by Sly and Sun [68] and Galanis *et al.* [25]. The algorithmic result of Weitz and the hardness of approximation result of Sly suggest that the uniqueness threshold for the infinite Δ-regular tree is the right threshold for the phase transition of the mixing time of the Glauber dynamics on the class of all graphs with maximum degree Δ. The underlying reason is that the Δ-regular tree is considered to be the extremal case for the decay of correlations amongst all graphs of maximum degree Δ.

For these models, to get precise thresholds for the mixing times of the Glauber dynamics in the setting of general graphs is extremely challenging. However, as illustrated earlier in the Ising model, the tree uniqueness threshold is not the threshold for having an efficient sampling algorithm or approximating the partition function in polynomial time for many classes of graphs with more structural information. In this thesis, we develop new analytical methods to study colorings and weighted independent sets on graphs with extra structures (such as square lattices and trees) and hence we are able to prove new bounds for the uniqueness thresholds and mixing times of the Glauber dynamics on these graphs. We now summarize our major contributions below.

1.2.1 Uniqueness Threshold for the Hard-core Model on $\mathbb{Z}^2$

Our first main result focuses on the well-studied particular case of the hard-core model on the square lattice $\mathbb{Z}^2$, and provides a new lower bound for the uniqueness threshold, in particular taking it well above the tree uniqueness threshold $\lambda_c(\mathbb{T}_4)$ which is the best previous result of the lower bound proved by Weitz [74].

Empirical evidence suggests that the critical point $\lambda_c(\mathbb{Z}^2) \approx 3.796$ [26, 6, 57]. But, unlike the Ising model on $\mathbb{Z}^2$ where the critical value of the uniqueness threshold is known, both upper and lower bounds on the possible uniqueness threshold $\lambda_c(\mathbb{Z}^2)$ are significantly far from this conjectured point. The possibility of there being multiple such λ_c is not ruled out, although no one believes that this is the case. From below, van den Berg and Steif [72] used a disagreement percolation argument to prove $\lambda_c(\mathbb{Z}^2) > \frac{p_c}{1-p_c}$ where p_c is the critical probability for site percolation on $\mathbb{Z}^2$. Applying the best known lower bound on $p_c > 0.556$ for $\mathbb{Z}^2$ by van den Berg and Ermakov [71] implies $\lambda_c(\mathbb{Z}^2) > 1.252\ldots$. Prior to that work, an alternative approach aimed at establishing the Dobrushin-Shlosman criterion [19], yielded, via computer-assisted proofs, $\lambda_c(\mathbb{Z}^2) > 1.185$ by Radulescu and Styer [59], and $\lambda_c(\mathbb{Z}^2) > 1.508$ by Radulescu [58].

These results were improved upon by Weitz [74] who showed that $\lambda_c(\mathbb{Z}^2) \geq \lambda_c(\mathbb{T}_4) = 27/16 = 1.6875$. For the upper bound, a classical Peierls' type argument implies $\lambda_c(\mathbb{Z}^2) = O(1)$ [18]. Recently, results of Randall [60] and later improved by Blanca *et al.* [11] show slow mixing of the Glauber dynamics for $\lambda > 5.3646$. The upper bound for the uniqueness threshold $\lambda_c(\mathbb{Z}^2)$ is also established at 5.3646 by them, which is a significant improvement of the upper bound to a relatively small value from what was previously known.

In this thesis, we propose a new analytical way to establish a strong version of the decay of long range correlation (the so-called *strong spatial mixing* condition, see Chapter 2.2 for a formal definition). Our technique refines and builds on the tree of self-avoiding walks approach of Weitz [74], resulting in a technical sufficient criterion for the strong spatial mixing (and hence uniqueness) to hold on the hard-core model. The new criterion achieves better bounds on strong spatial mixing when the graph has extra structures, improving upon what can be achieved by just using the maximum degree. The methodology we adopt is general and applicable to other spin systems for a general class of regular graphs (see, e.g., [65]). Applying our technique to $\mathbb{Z}^2$ we prove that strong spatial mixing holds for all $\lambda < 2.48$, improving upon the work of Weitz that held for $\lambda < 27/16 = 1.6875$. Our results imply a fully-polynomial *deterministic* approximation algorithm for estimating the partition function. In the square lattice, for any vertex v, the number of vertices that are within radius t from v grows polynomially in t, and hence the square lattice is called an *amenable* graph. It was shown by Dyer *et al.* [23] that for amenable graphs, if *strong spatial mixing* holds, then the Glauber dynamics is rapidly mixing. Therefore, for the hard-core model on the square lattice, as a corollary of our decay of correlation results, we improve the range of the activity λ where the Glauber dynamics is rapidly mixing on the square lattice. In summary, we prove the following theorem. This is a joint work with Ricardo Restrepo, Jinwoo Shin, Prasad Tetali and Eric Vigoda [61], published in *Probability Theory and Related Fields, 2012*.

Theorem 1. *The following hold for the hard-core model on $\mathbb{Z}^2$ for all $\lambda \leq \lambda^* = 2.48$:*

1. *Strong spatial mixing holds on $\mathbb{Z}^2$.*
2. *There is a unique infinite-volume Gibbs measure on $\mathbb{Z}^2$.*
3. *For every finite subgraph G of $\mathbb{Z}^2$, Weitz's algorithm [74] gives a fully-polynomial time approximation scheme (FPAS) for approximating the partition function $Z(G)$.*
4. *For every finite subgraph G of $\mathbb{Z}^2$, the Glauber dynamics has $O(n \log n)$ mixing time.*

1.2.2 Phase Transitions of Convergence Times for the Glauber dynamics on Trees

The second main result we present is about the conjectured phase transition for the mixing time of the Glauber dynamics for k-colorings on the Δ-regular tree. Note that for the Δ-regular tree, we always use $b = \Delta - 1$ to denote the branching factor. Sly [66] and Bhatnagar *et al.* [10] show that the reconstruction threshold occurs at $k_r = b(1+o(1))/\ln b$. It was believed that the place where the mixing time transition happens coincides with the reconstruction threshold on regular trees.

Our interest in the reconstruction threshold on trees is its apparent connection to the threshold for the efficiency of certain local algorithms on locally-tree like graphs, such as sparse random graphs $G(n, c/n)$ (where each edge appears with probability c/n for some constant $c > 1$) and planar graphs. For colorings, the reconstruction threshold k_r on the tree is believed to be intimately connected to the threshold for the efficiency of local algorithms for the sampling problem on locally tree-like graphs. The evidence in support of that belief is that the the geometry of the space of solutions on sparse random graphs appears to change dramatically near (and possibly at) the reconstruction threshold; see [2, 28, 49]. Hayes *et al.* [32] recently proved that the Glauber dynamics for any planar graph with maximum degree Δ mixes in polynomial time when $k > 100\Delta/\ln\Delta$, which is a constant factor away from the reconstruction threshold for the regular tree. Their result for the Glauber dynamics of colorings on planar graphs suggests that the reconstruction threshold may have connections to the mixing time of the Glauber dynamics on planar graphs as well.

For k-colorings on the Δ-regular tree, in Goldberg *et al.* [29] a non-trivial lower bound of the mixing time is established when the number of colors $k < \frac{b}{2\ln b}$, which is below the reconstruction threshold. In contrast, Lucier and Molloy [42] show an upper bound of the mixing time using the canonical flow approach (c.f., [41] Chapter 13.5). However, the upper bound is a high degree polynomial of n and hence it is not tight enough to establish the existence of a phase transition. Here, we provide a more precise picture for the phase transition of the mixing time than provided by the results in [29, 42]. Our main result gives (nearly) sharp bounds on the mixing time T_{mix} and relaxation time T_{relax} of the Glauber dynamics of k-colorings for the Δ-regular trees, establishing a phase transition at the critical point $k = b(1 + o_b(1))/\ln b$. This is a joint work with Prasad Tetali, Juan Vera and Eric Vigoda [70], published in *Annals of Applied Probability, 2012*. We prove that:

Theorem 2. *For the Glauber dynamics of k-colorings on the $(b+1)$-regular tree T of n vertices and height $H = \lfloor \log_b n \rfloor$ satisfies the following:*

1. **For all $\epsilon > 0$ and all $k = \frac{(1+\epsilon)b}{\ln b}$:**

$$\begin{aligned} \Omega\left(n \ln n/(b\ \text{poly}(\ln b))\right) \ &\leq\ T_{\text{mix}}\ \leq\ O(n^{1+o_b(1)} \ln n), \\ \Omega(n)\ &\leq\ T_{\text{relax}}\ \leq\ O(n^{1+o_b(1)}). \end{aligned}$$

2. **For all $\epsilon > 0$ and all $k = \frac{b}{(1+\epsilon)\ln b}$:**

$$\begin{aligned} \Omega(n^{1+\epsilon-o_b(1)})\ &\leq\ T_{\text{mix}}\ \leq\ O(n^{1+\epsilon+o_b(1)} \ln n), \\ \Omega(n^{1+\epsilon-o_b(1)})\ &\leq\ T_{\text{relax}}\ \leq\ O(n^{1+\epsilon+o_b(1)}). \end{aligned}$$

where the $o_b(1)$ functions are $O(\ln\ln b/\ln b)$ for the upper bounds, $(1+\epsilon)/b^{\epsilon}$ for the lower bounds when $0 < \epsilon < 1$ and exactly zero for the lower bounds when $\epsilon \geq 1$. The constants in the $\Omega(\cdot)$ and $O(\cdot)$ are universal constants.

Our third main result is for the Glauber dynamics of the hard-core model on the complete tree, where we have the same phase transition phenomena around the reconstruction threshold. For the hard-core model, the existence of the reconstruction threshold follows from Mossel [50, Proposition 20]. By recent works of Bhatnagar *et al.* [9] and Brightwell and Winkler [13], it is known that the critical $\omega_r = (\ln b + (1 + o(1))\ln\ln b)/b$ where ω is the real positive solution of $\lambda = \omega(1+\omega)^b$.

Figuring out the mixing time for the hard-core model on trees is indeed a more complicated question to answer correctly than for k-colorings. Martinelli *et al.* [47] showed that for the hard-core model on the $(b+1)$-regular tree of height h with free boundary condition the relaxation time is $O(n)$ for all λ (and the mixing time is $O(n \log n)$). Hence, for the hard-core model, unlike in the Ising and colorings models, the Glauber dynamics on the tree with free boundary condition does not have connections to the reconstruction threshold. Our interest was whether there is a boundary condition for which there is such a connection.

We prove there is a connection by explicitly constructing a boundary condition for which the relaxation time slows down at the reconstruction threshold. The boundary condition is constructed to mimic the measure generated by the broadcasting process for the hard-core model on the infinite trees. Here is the formal statement of our results. It is a joint work with Ricardo Restrepo, Daniel Stefankovic, Juan Vera and Eric Vigoda [62], published in *SODA 2011.*

Theorem 3. *For the Glauber dynamics of the hard-core model with activity* $\lambda = \omega(1+\omega)^b$ *on the* $(b+1)$*-regular tree of* n *vertices and height* $H = \lfloor \log_b n \rfloor$*, the following hold:*

1. **For all** $\omega \leq \ln b/b$**:**

   For every boundary condition,

$$\Omega(n) \leq \ T_{\text{relax}} \ \leq O(n^{1+o_b(1)}).$$

2. **For all** $\delta > 0$ **and** $\omega = (1+\delta)\ln b/b$**:**

   (a) *For every boundary condition,*

$$T_{\text{relax}} \ \leq O(n^{1+\delta+o_b(1)}).$$

   (b) *There exists a sequence of boundary conditions for all* $H \to \infty$ *such that,*

$$T_{\text{relax}} = \Omega(n^{1+\delta/2-o_b(1)}).$$

Overview of the Organization

The remainder of the thesis is organized as follows. In Chapter 2, we formally define spin systems, Gibbs measure, the Glauber dynamics and related background concepts used in the thesis. In Chapter 3, we introduce a general method that relates reconstruction to a lower bound on the mixing time of the Glauber dynamics on trees. We then apply the method to k-colorings and the hard-core model, to derive the lower bound results in Theorem 2 and Theorem 3. In Chapter 4, we use various coupling techniques to prove the upper bounds given in Theorem 2 and Theorem 3. We also establish a connection between the log-Sobolev constant and the spectral gap of the dynamics for the complete trees that is near optimal in order to sharply bound the mixing time. Finally, we discuss the uniqueness problem for the hard-core model on the square lattice in Chapter 5. Here, we introduce an analytical tool to improve the range where the spatial mixing condition holds for a general class of regular infinite graphs.

CHAPTER II

PRELIMINARIES

In this chapter we formally define the spin systems, Gibbs measures, the uniqueness problem, the reconstruction problem, the Glauber dynamics, mixing time and related concepts.

2.1 Nearest Neighbor Spin Systems

Let $G = (V, E)$ be a (finite) graph and Q be a finite set called spins. Let $k = |Q|$. A configuration σ of the graph is a function from V to Q, i.e., an assignment for each vertex $v \in V$ with a spin $q \in Q$. We can define a probability measure $\mu = \mu_G$ over all the configurations, which is called the Gibbs measure.

Let the external field Λ be a fixed real vector of length k and the interaction matrix B be a fixed k by k matrix with entries in $\mathbb{R} \cup \{\pm\infty\}$. For each $\sigma \in \Omega$, let the weight of the configuration be defined as:

$$w(\sigma) = \exp\left(\sum_{(u,v)\in E} B(\sigma(u), \sigma(v)) + \sum_{v\in V} \Lambda(\sigma(v))\right).$$

The set of all the configurations with positive weights is denoted as $\Omega = \Omega_G$.

Let $Z = \sum_{\sigma\in\Omega} w(\sigma)$ be the partition function which is the total sum of the weights over all the configurations. Then, the Gibbs distribution μ is defined as for $\sigma \in \Omega$, $\mu(\sigma) = w(\sigma)/Z$. Note that, when the graph G is clear in the context, the notations for μ, Ω and Z will always omit G in the subscription for simplicity.

To measure the similarity of two probability distributions, usually we use the notion of total variation distance which is defined as follows.

Definition 4. *For probability distribution μ and ν on the space Ω, the total variation* *distance* $\|\mu - \nu\|_{TV}$ *is defined as*

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{\sigma \in \Omega} |\mu(\sigma) - \nu(\sigma)| = \max_{A \subseteq \Omega} |\mu(A) - \nu(A)| .$$

The ferromagnetic Ising model mentioned in the introduction can be defined using the above notions by setting $Q = \{+1, -1\}$, $\Lambda(+1) = 0, \Lambda(-1) = 0$ and $B(i,j) = ij\beta$ for all $i, j \in Q$. In this thesis, we mainly study the following two hard constraints spin systems: k-colorings and the hard-core model.

Example 2.1.1. *The k-proper coloring is a spin system with $|Q| = k$ spins. Each valid configurations is a proper k coloring of the underlying graph G, i.e., the assignments of two vertices cannot be the same if there is an edge between the vertices. The Gibbs measure is uniform distribution over all the proper colorings. Fitting into the general spin systems setting, the coloring model is the same as setting $B(i,i) = -\infty$, $B(i,j) = 0$ for $i \neq j$ and $\Lambda(i) = 0$.*

Example 2.1.2. *Weighted independent set (hard-core) model, is a spin system with two spins: occupied and unoccupied. Each valid configuration is an independent set of the underlying graph G. An independent set of the graph G is defined as a subset I of vertices V such that there is no edge between the vertices in I. Each independent set I is of weight $\lambda^{|I|}$ where $\lambda > 0$ is called activity. When $\lambda > 1$, the system is in favor of independent sets with larger sizes and when $\lambda < 1$, the system is in favor of smaller independent sets. When $\lambda = 1$, the Gibbs distribution is the uniform distribution over all independent sets. Putting into the context of general spin systems, here the matrix B is defined by the following table:*

	unoccupied	*occupied*
unoccupied	0	0
occupied	0	$-\infty$

And the site activity Λ(unoccupied) $= 0$, Λ(occupied) $= \log(\lambda)$. Generally, we will use $Q = \{0, 1\}$ and let spin 0 represent "unoccupied" and spin 1 represent "occupied".

Now we are going to study the conditional Gibbs distribution on a set of vertices, say U, when the configurations on another set of vertices distinct from U, say S, is fixed. To do this, we introduce the concept of boundary condition. A *boundary condition* is a fixed partial assignment to the vertex set V. Formally, let S be a subset of V. Then, a boundary condition $\boldsymbol{\rho}$ is an assignment from S to Q. We can then define the space of valid configurations according to the boundary condition:

$$\Omega_{\boldsymbol{\rho}} = \Omega_{G,\boldsymbol{\rho}} = \{\sigma \in \Omega | \sigma(S) = \boldsymbol{\rho}(S)\}.$$

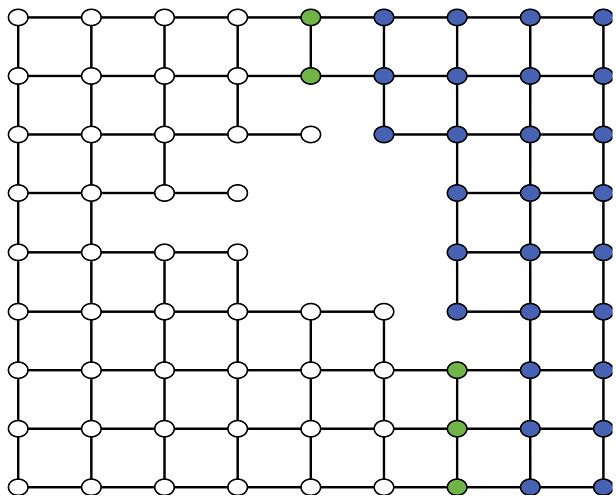

Figure 2: Both blue and green vertices are in the set S and ∂S are the green vertices. The white vertices are hence in the set $V \setminus S$.

We will use the notation ∂S to denote the set of vertices that are on the boundary of S with respect to $V \setminus S$ (See Figure 2 for an example), i.e.,

$$\partial S = \{s \in S | \exists v \in V \setminus S, (v, s) \in E\}.$$

The probability measure $\mu_{\boldsymbol{\rho}} = \mu_{G,\boldsymbol{\rho}}$ is then defined as the conditional probability distribution on $\Omega_{G,\boldsymbol{\rho}}$ for each configuration that agrees with the boundary condition $\boldsymbol{\rho}$ on set S, i.e., for each $\sigma \in \Omega_{\boldsymbol{\rho}}$,

$$\mu_{\boldsymbol{\rho}}(\sigma) = \mu_{G,\boldsymbol{\rho}}(\tau) = \frac{w(\sigma)}{Z_{\boldsymbol{\rho}}},$$

where

$$Z_{\boldsymbol{\rho}} = Z_{G,\boldsymbol{\rho}} = \sum_{\sigma \in \Omega_{G,\boldsymbol{\rho}}} w(\sigma).$$

The conditional Gibbs distribution $\mu_{\boldsymbol{\rho}}$ can be also viewed as the probability distribution on the space $\Omega_{G'}$ with $G' = (V \setminus S, E)$ (e.g., the white vertices and edges in Figure 2). For each configuration $\tau \in \Omega_{G'}$, let

$$\mu_{\boldsymbol{\rho}}(\tau) = \frac{\mu(\tau \circ \boldsymbol{\rho})}{\sum_{\tau} \mu(\tau \circ \boldsymbol{\rho})},$$

where μ is the Gibbs distribution on the graph G. For any two assignments $\eta_1 : S_1 \to Q$ and $\eta_2 : S_2 \to Q$ with $S_1 \cap S_2 = \emptyset$, the concatenated function $\eta_1 \circ \eta_2$ is defined as the assignment from $S_1 \cup S_2$ to Q such that it agrees with η_1 on S_1 and with η_2 on S_2. From this point of view, for any fixed graph G, it is immediate to check that $\mu_{G,\boldsymbol{\rho}}$ actually only depends on the assignments on ∂S.

To extend the definition of the Gibbs measure from finite graphs to infinite graphs, the following well-known DLR (Dobrushin-Landford-Ruelle) compatibility conditions [27] is used.

Definition 5. *A probability measure ν over the configurations on the infinite graph $G = (V, E)$ is a Gibbs measure if, for any subset T of V such that $V \setminus T$ is finite, any boundary condition $\boldsymbol{\rho}$ on ∂T and almost surely every $\tau \in \Omega(G)$ that agrees with $\boldsymbol{\rho}$, we have*

$$\nu(\sigma_{V \setminus T} | \sigma_T = \tau_T) = \mu_{G,\boldsymbol{\rho}}(\sigma_{V \setminus T}),$$

where according to the definition, $\mu_{G,\boldsymbol{\rho}}$ is the Gibbs measure on the finite graph of G restricted on the vertex set $V \setminus T$ with the boundary condition $\boldsymbol{\rho}$.

Note that $\mu_{G,\boldsymbol{\rho}}$ only depends on the assignment on ∂T and when the graph G is locally finite, i.e., the degree of G is bounded, then ∂T is finite and hence $\mu_{G,\boldsymbol{\rho}}$ is enumerable with respect to the assignments to ∂T. In many works, the finite dimension measures $\mu_{G,\boldsymbol{\rho}}$ are called the *specification*. It is clear that all the spin systems we defined previously over the finite graphs give valid specifications and it is well known that for any specification, at least one infinite Gibbs measure always exists [27].

2.2 Uniqueness, Weak and Strong Spatial Mixing

As we saw in the introduction, for the Ising model on infinite tree, we can write a sequence of boundary conditions and it is known that when an appropriate sequence of boundary conditions is chosen, the weak limit of the conditional Gibbs distribution exists and gives a valid infinite Gibbs measure (see, e.g., [27]). When two sequences of boundary conditions are different, the infinite Gibbs measures we obtain from the limits may be different. Hence, there may be coexistence of several infinite Gibbs measures for the same specification μ. One of the central questions in the study of the spin systems is the so-called the *uniqueness problem* asking whether the system admits more than one several infinite Gibbs measures.

In fact, the set of infinite measures forms a simplex and any infinite measure can be written as a convex combination of extremal Gibbs measures. Therefore, to determine whether the infinite Gibbs measure is unique or not, it is sufficient to study those extremal measures. It has been proved that the extremal Gibbs measures can always be defined as the the weak limit of the measures specified by a sequence of boundary conditions on finite subgraphs of the infinite graph. This nice fact gives us a more constructible way to understand and manipulate the infinite Gibbs measures. As we view an extremal measure as the limit of a sequence of finite boundary conditions, we can treat the uniqueness problem as whether different sequences of boundary conditions have substantially different effects to the interior in the limit, i.e., whether the information from the boundaries can affect the probability distributions of the configurations on the internal vertices. For an infinite graph, the Gibbs distribution for the specification μ is unique if and only if the following condition holds. The results mentioned in this paragraph are summarized in [27].

Proposition 6. *A specification μ admits unique Gibbs measure on infinite graph $G = (V, E)$ if and only if for any finite region $S \subset V$, there is a sequence of finite regions $S \subset S_1 \subset S_2 \cdots \subset S_n \subset \ldots$ that $\cup_{n>0} S_n = V$, and for any two configuration σ and η, the following holds:*

$$\lim_{n\to\infty} \left\| \mu^S_{\sigma_n} - \mu^S_{\eta_n} \right\|_{TV} = 0,$$

where for each n, σ_n (same for η_n) corresponds to the boundary conditions σ on $V \backslash S_n$, i.e., the configurations outside S_n are fixed to be σ. And $\mu^S_{\sigma_n}$ is the projection of the measure μ_{σ_n} on $\Omega(S)$.

This inspires the concepts of various *spatial mixing* conditions which are the sufficient conditions for establishing the uniqueness.

Our results about spatial mixing are mainly for two spin systems, especially for the hard-core model, and therefore it is convenient to define them in terms of the hard-core model to simplify the notations.

Let $G = (V, E)$ be a (finite) graph. For $v \in V$, let $\alpha_{\boldsymbol{\rho}}(v) = \alpha_{G,\boldsymbol{\rho}}(v)$ denote the marginal probability of v being set to "unoccupied" in the measure $\mu_{G,\boldsymbol{\rho}}$ with a boundary condition $\boldsymbol{\rho}$ on the vertex set $S \subseteq V$.

The first spatial mixing property is *Weak Spatial Mixing (WSM)*. Here we consider a pair of boundary configurations on a subset S and consider the "influence" on the marginal probability that a vertex v is unoccupied. WSM says that the influence on v decays exponentially in the distance of S from v.

Definition 7 (Weak Spatial Mixing)**.** *For a finite graph $G = (V, E)$, in the spin system with Gibbs distribution μ,* WSM *holds if there is a $0 < \gamma < 1$ such that for every $v \in V$, every $S \subset V$, and every two boundary conditions $\boldsymbol{\rho}, \boldsymbol{\eta}$ on S,*

$$|\alpha_{\boldsymbol{\rho}}(v) - \alpha_{\boldsymbol{\eta}}(v)| \ \leq \ \gamma^{\mathrm{dist}(v,S)},$$

where $\mathrm{dist}(v, S)$ *is the graph distance (i.e., length of the shortest path) between* v *and (the nearest point in) the subset* S.

The second property of interest is *Strong Spatial Mixing (SSM)*. The intuition is that if a pair of boundary configurations on a subset S agree at some vertices in S then those vertices "encourage" v to agree. Therefore, SSM indicates that the influence on v decays exponentially in the shortest distance from v to the subset of vertices where the pair of configurations differs.

Definition 8 (Strong Spatial Mixing)**.** *For a finite graph* $G = (V, E)$*, in the spin system with Gibbs distribution* μ*,* SSM *holds if there is a* $0 < \gamma < 1$ *such that for every* $v \in V$*, every* $S \subset V$*, every* $S' \subset S$*, and every two boundary conditions* $\boldsymbol{\rho}, \boldsymbol{\eta}$ *on* S *where* $\boldsymbol{\rho}(S \setminus S') = \boldsymbol{\eta}(S \setminus S')$*,*

$$|\alpha_{\boldsymbol{\rho}}(v) - \alpha_{\boldsymbol{\eta}}(v)| \;\leq\; \gamma^{\mathrm{dist}(v,S')}.$$

Note that since $\mathrm{dist}(v, T) \leq \mathrm{dist}(v, T \setminus S)$, SSM implies WSM. We can specialize the above notions of WSM and SSM to a particular vertex v, in which case we say that WSM or SSM holds *at* v. If the graph is a rooted tree, we will always assume that the notions of WSM and SSM are considered at the root.

For the hard-core model on a graph $G = (V, E)$, for a subset of vertices S and a fixed configuration ρ on S, the effect of the boundary condition ρ on G is equivalent to modifying the graph G to an induced subgraph G' in the follow way: for each $v \in S$ that is fixed to be unoccupied we remove v from G, and for each $v \in S$ that is fixed to be occupied we remove v and its neighbors $N(v)$ from G. In this way we obtain the following observation which will be useful for proving SSM holds.

Observation 9. *For a graph* $G = (V, E)$ *and* $v \in V$*, in the hard-core model, SSM holds in* G *at vertex* v *iff WSM holds for all the induced subgraphs* G' *of* G *at vertex* v*. To be precise, by induced subgraphs we mean graphs obtained by considering all the vertex induced subgraphs of* G *and taking the component containing* v*.*

2.3 Broadcasting Process and Reconstruction

The broadcasting process is a way to generate an infinite probability measure over the configurations on regular infinite trees starting from the root. It simulates the process of how the information is broadcasted from the root to other vertices in a tree communication network (see, e.g., [51]). Given an infinite tree $\mathbb{T}_\Delta$ of branching factor $b = \Delta - 1$, a set of spins Q with $|Q| = k$ and a k by k stochastic matrix $\mathcal{B}$, the broadcasting process does the following. The assignment to the root is randomly selected from some initial distribution π. We denote the root's assignment as σ_r. Then, each child of the root is assigned with spins independently randomly according to the distribution $\mathcal{B}(\sigma_r, \cdot)$. This procedure is carried on from the root level by level down to the infinite. Usually, we use $\nu = \nu(G, \mathcal{B})$ to denote the probability distribution generated by the broadcasting process. Here we give two examples of the broadcasting processes which we will study later.

Example 2.3.1. *For k-coloring on the complete tree $\mathbb{T}_\Delta$, the broadcasting model is quite straightforward. The initial distribution for the assignment σ_r of the root is $1/k$ for each color. Then, each child of the root is choosing a color from $Q \setminus \{\sigma_r\}$ independently and uniformly randomly. The corresponding stochastic matrix $\mathcal{B}$ is defined as:*

- $\mathcal{B}(i, i) = 0$*;*
- $\mathcal{B}(i, j) = \frac{1}{k-1}$ *for all $i \neq j$.*

The broadcasting process for the hard-core model is a bit complicated.

Example 2.3.2. *Let ω be the real positive solution of the equation*

$$\lambda = \omega(\omega + 1)^b.$$

Initially the root is unoccupied with probability $\frac{\omega+1}{2\omega+1}$ and occupied with probability

$\frac{\omega}{2\omega+1}$. *The 2 by 2 stochastic matrix is then defined as:*

$$\mathcal{B} = \begin{pmatrix} \frac{1}{1+\omega} & \frac{\omega}{1+\omega} \\ 1 & 0 \end{pmatrix}.$$

Let L_n denote the vertices that are at distance n from the root, i.e., the n^{th} level of the tree. We use σ_n to denote the projection (restriction) of the configuration $\sigma \in \Omega$ to the vertex set L_n. We use $\mathbf{P}_n^i$ to denote the conditional probability measure of the configurations on L_n given that the root is assigned with the spin i. The so-called *reconstruction problem* is defined as follows.

Definition 10 (Reconstruction Problem, see, e.g., [51])**.** *The reconstruction problem for $\mathbb{T}_\Delta$ and $\mathcal{B}$ is solvable if there exists $i, j \in Q$ such that*

$$\lim_{n\to\infty} \left\|\mathbf{P}_n^i - \mathbf{P}_n^j\right\|_{TV} > 0,$$

where $\|\mu_1 - \mu_2\|_{TV}$ is the total variation distance between two distributions μ_1 and μ_2.

From the definition, we can see that the reconstruction problem is concerning the difference between two conditional distributions on the configurations of L_n when the broadcasting process starts with different assignments on the root as $n \to \infty$. There are several equivalent definitions of the reconstruction solvability and readers can refer to [51] for more details. We will use the following, which asks whether a typical assignment of the leaves influences the conditional measure at the root.

Definition 11 (Reconstruction Problem)**.** *For the measure ν generated by the broadcasting process $\mathcal{B}$ on $\mathbb{T}_\Delta$, the reconstruction problem is solvable if there exists a spin $q \in Q$ such that*

$$\lim_{n\to\infty} \mathrm{E}_{\sigma\sim\nu}\left[\left|\mu_{\sigma_n}(\eta(r) = q) - \nu(\sigma(r) = q)\right|\right] > 0,$$

where $\mu_{\sigma_n}(\eta(r) = q)$ is defined as the marginal conditional probability of the root r being colored q given the boundary condition σ at the vertices L_n.

Note that the reconstruction problem is not solvable if and only if the corresponding broadcasting measure on the infinite tree is extremal (see, e.g., [27, 51]).

2.4 *Glauber Dynamics and Bounds on Convergence Times*

The (heat bath) Glauber dynamics is a discrete time Markov chain (X_t) for sampling from the Gibbs distribution μ for a specific spin system on a given graph $G = (V, E)$. The transitions $X_t \to X_{t+1}$ of the Glauber dynamics are defined as:

- Choose a vertex v uniformly at random;
- For all $w \neq v$ set $X_{t+1}(w) = X_t(w)$;
- Set $X_{t+1}(v) = i$ with probability $\mu_{\boldsymbol{\rho}}(v = i)$, where $\boldsymbol{\rho}$ is the boundary condition defined on $V \setminus \{v\}$ and $\boldsymbol{\rho}(w) = X_t(w)$ for all $w \in V \setminus \{v\}$.

When a boundary condition $\boldsymbol{\rho}$ is specified for the Glauber dynamics, the state space is restricted to $\Omega_{\boldsymbol{\rho}}$.

Let $P(\cdot,\cdot)$ denote the transition matrix of the Glauber dynamics, and $P^t(\cdot,\cdot)$ denote the t-step transition probability. The total variation distance at time t from initial state σ is defined as

$$\|P^t(\sigma,\cdot) - \pi\|_{TV} := \frac{1}{2}\sum_{\eta} |P^t(\sigma,\eta) - \pi(\eta)|.$$

The mixing time T_{mix} for a Markov chain is then defined as

$$T_{\text{mix}}(\epsilon) = \min_t\{\max_\sigma\{\|P^t(\sigma,\cdot) - \pi\|_{TV}\} \leq \epsilon\}.$$

In many papers, ϵ is often set to be a fixed number, e.g., $\epsilon = \frac{1}{2\mathrm{e}}$. This is due to the well-known fact that $T_{\text{mix}}(\epsilon) \leq T_{\text{mix}}(\frac{1}{2\mathrm{e}}) \log\frac{1}{\epsilon}$. We denote $T_{\text{mix}}(\frac{1}{2\mathrm{e}})$ as T_{mix}.

Let $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\Omega|}$ be the eigenvalues of the transition matrix P. The spectral gap c_{gap} is defined as $1 - \lambda_2$. The relaxation time T_{relax} of the Markov chain is then defined as c_{gap}^{-1}, the inverse of the spectral gap. It is an elementary fact that the mixing time gives a good upper bound on the relaxation time (see, e.g., Theorem 5 in [22]), which we will use in our analysis:

$$T_{\text{relax}} = O(T_{\text{mix}}). \tag{1}$$

Note that our definition of relaxation time following [8, 46] is slightly different from the standard definition, the inverse of the *absolute spectral gap* which is $(1-\max\{|\lambda_2|, |\lambda_{|\Omega|}|\})^{-1}$ (see, e.g., Chapter 13 in [41]). It is a standard fact that by passing to a lazy chain $(\frac{1}{2}P + \frac{1}{2}I)$ which makes no move with half probability, the two definitions are identical. Introducing the laziness to the Glauber dynamics only adds an extra factor of two to the mixing time, therefore it will not affect our asymptotic results.

To lower bound the mixing and relaxation times we analyze the conductance. The conductance of the Markov chain on Ω with transition matrix P is given by $\Phi = \min_{S \subseteq \Omega}\{\Phi_S\}$, where Φ_S is the conductance of a specific set $S \subseteq \Omega$ defined as

$$\Phi_S = \frac{\sum_{\sigma \in S} \sum_{\eta \in \bar{S}} \pi(\sigma) P(\sigma, \eta)}{\pi(S)\pi(\bar{S})}.$$

Roughly speaking, the conductance Φ_S measures the probability of being in the set S and getting out of S in the next step of the transition following the Markov chain transition matrix P. A general way to find a good upper bound on the conductance is to find a set S such that the probability of escaping from S is relatively small. The well-known relationship between the relaxation time and the conductance is established in [39] and [64] and we will use the form

$$T_{\text{relax}} = \Omega(1/\Phi)\,, \tag{2}$$

for proving the lower bounds.

We will also work with the logarithmic Sobolev constant of a (finite) Markov chain. We briefly recall here the variational definition of both the spectral gap and the log-Sobolev constant.

Let f be a function (vector) from Ω to R, π be the stationary distribution over Ω and μ be any probability distribution over Ω. Let $\mathcal{D}(f)$ be the standard *Dirichlet form* of the heat-bath Glauber dynamics defined as:

$$\mathcal{D}(f) = \frac{1}{2}\sum_{\sigma}\sum_{\sigma'}(f(\sigma) - f(\sigma'))^2\pi(\sigma)P(\sigma,\sigma').$$

Let $E_\mu(f)$ be the average of f under the distribution μ, and let $\mathrm{Var}_\mu(f) := E_\mu(f^2) - E_\mu^2(f)$ be the corresponding variance, which can also be written as:

$$\mathrm{Var}_\mu(f) = \frac{1}{2}\sum_{\sigma}\sum_{\sigma'}(f(\sigma) - f(\sigma'))^2\mu(\sigma)\mu(\sigma').$$

Let $\mathrm{Ent}_\mu(f) := E_\mu(f\log f) - E_\mu(f)\log(E_\mu(f))$. When it is clear what the underlying distribution is, we will drop the subscript μ in the notation $\mathrm{Ent}(f)$.

The spectral gap c_{gap} is equivalently defined as (see, e.g., Chapter 13 in [41])

$$c_{gap} = \inf_f \frac{\mathcal{D}(f)}{\mathrm{Var}(f)}.$$

The log-Sobolev constant c_{sob} is defined as (see, e.g., [16]),

$$c_{sob} = \inf_{f\geq 0} \frac{\mathcal{D}(\sqrt{f})}{\mathrm{Ent}(f)},$$

where the infimum in both equations is over non-constant functions f.

From this definition, it is more clear that the relaxation time $T_{\mathrm{relax}} = 1/c_{gap}$ is a good measurement of the convergence speed, since it is not hard to show that, for any function f,

$$\mathrm{Var}_\pi(P^t f) \leq (1 - c_{gap})^{2t}\mathrm{Var}_\pi(f).$$

A similar inequality about the entropy form $\mathrm{Ent}(P^t f)$ holds for the log-Sobolev constant.

As for the upper bounds on the mixing time of the dynamics, one can use the following well-known relationship between the mixing time and the relaxation time (see, e.g., Theorem 12.3 in [41]):

$$T_{\mathrm{mix}} \leq \log\left(\frac{1}{\min_{\sigma\in\Omega}\{\pi(\sigma)\}}\right) T_{\mathrm{relax}}.$$

By applying the above relationship, one usually gets an additional factor of n for the upper bound of the mixing time. To save this factor of n, one may use the following relationship between the mixing time and the inverse of the log-Sobolev constant (see e.g. [16] for more details):

$$T_{\text{mix}} \quad = \quad O\left(c_{sob}^{-1} \ln \ln \frac{1}{\min_{\sigma \in \Omega}\{\pi(\sigma)\}}\right). \tag{3}$$

CHAPTER III

LOWER BOUNDS ON THE RELAXATION TIME AND MIXING TIME

In this chapter we will establish the connection between the reconstruction solvability and the upper bound of the conductance and hence the lower bound of the relaxation time and mixing time of the Glauber dynamics.

This chapter is organized in the following way. First, we will introduce the concept of reconstruction algorithms, which are effective algorithms for recovering the assignments at the root given the configurations on the leaves when the reconstruction problem is solvable. Then, we will show a simple but strong connection between the conductance of the Glauber dynamics and the so-called sensitivity of a reconstruction algorithm. We analyze a reconstruction algorithm (FR) for k-colorings and a reconstruction algorithm (BW) for hard-core model respectively and establish upper bounds for the conductances in both models. As a result, we are able to establish good lower bounds of the relaxation time and the mixing time.

3.1 Conductance and Reconstruction Algorithms

Let us denote the finite Δ-regular tree of height h as T_h when the degree Δ is clear in the context. Let ν_h be the projection of the broadcasting distribution ν (defined in Chapter 2.3) of the entire infinite tree onto the first h levels of the tree, i.e., ν_h is the distribution on the configurations of T_h generated by the broadcasting process with a broadcasting matrix $\mathcal{B}$. For the tree T_h, we use L to denote the leaves L_h for simplicity.

A *reconstruction algorithm* is a function $A : \Omega(L) \to Q \cup Q'$ (ideally efficiently computable) such that $A(\sigma_h)$ and $\sigma(r)$ are positively correlated for random configurations $\sigma \sim \nu_h$. One can imagine that for a configuration σ, only the configurations on the leaves are exposed to the algorithm A and the task of algorithm A is to take the configurations σ_h at the leaves L as the input and tries to guess (compute) the configuration $\sigma(r)$ at the root. When the context is clear, we write $A(\sigma)$ instead of $A(\sigma_h)$. The set Q' is a set of extra symbols that is different from the spins in Q for the reconstruction algorithm to use. Usually, if a good reconstruction algorithm outputs a symbol in Q' for the root, it means that the uncertainty for the configuration at the root is high given the configurations at the leaves.

Under the Gibbs measure ν_h, the *effectiveness* of A is the following measure of the covariance between the algorithm A's output and the marginal at the root of the actual measure:

$$r_{h,A} = \min_{x \in Q} \left[\nu_h(A(\sigma) = \sigma(r) = x) - \nu_h(A(\sigma) = x)\nu_h(\sigma(r) = x) \right].$$

If it is the case that

$$\liminf_{h \to \infty} r_{h,A} = c_0 > 0$$

for some positive constant c_0 (independent of the number of vertices n and height h), then we say that A is an *effective reconstruction algorithm*. In words, an effective algorithm, is able to recover the spin at the root, from the information at the leaves, with a nontrivial success, when $h \to \infty$. Notice that reconstruction (defined in Definition 11) is a necessary condition for the existence of an effective reconstruction algorithm, since

$$\begin{aligned}
&\mathrm{E}_{\sigma \sim \nu_h} \left[\left| \mu_{\sigma_h}(\eta(r) = x) - \nu_h(\eta(r) = x) \right| \right] \\
&\quad \geq \mathrm{E}_{\sigma \sim \nu_h} \left[\left| \mu_{\sigma_h}(\eta(r) = x) - \nu_h(\eta(r) = x) \right| \mathbf{1}(A(\sigma) = x) \right] \\
&\quad \geq \mathrm{E}_{\sigma \sim \nu_h} \left[\mu_{\sigma_h}(\eta(r) = x) \mathbf{1}(A(\sigma) = x) \right] - \nu_h(A(\sigma) = x)\nu_h(\sigma(r) = x) \\
&\quad \geq r_{h,A},
\end{aligned}$$

where $\mathbf{1}(\cdot)$ is the indicator function. We define the *sensitivity* of A, for the configuration $\sigma \in \Omega(T_h)$, as the fraction of vertices v such that switching the spin at v in σ changes the final result of A. More precisely, let $\sigma^{v \to c}$ be the configuration obtained from changing σ at v to spin c. Define the sensitivity as:

$$S_A(\sigma) = \frac{1}{n} \#\{v \in L : \exists c \in Q, A(\sigma^{v \to c}) \neq A(\sigma)\}.$$

The *average sensitivity* $\bar{S}_A$ is hence defined as

$$\bar{S}_A = \min_{x \in Q} \{\bar{S}_{A,x}\}, \text{ and } \bar{S}_{A,x} = \mathrm{E}_{\sigma \sim \nu_h} \left[S_A(\sigma) \mathbf{1}(A(\sigma) = x) \right].$$

It is fine to define the average sensitivity without the indicator function, which only affects a constant factor in the analysis. We are doing so to simplify some of the statements and proofs.

Typically when one proves reconstruction, it is done by presenting an effective reconstruction algorithm. Using the following theorem, by further analyzing the sensitivity of the reconstruction algorithm, one obtains an upper bound on the conductance of the Glauber dynamics.

Theorem 12. *Suppose that A is an effective reconstruction algorithm. Then, the conductance Φ of the Glauber dynamics satisfies $\Phi = O\left(\bar{S}_A\right)$.*

Proof. Throughout the proof let $\pi := \nu_h$. Consider the set $U = \{\sigma : A(\sigma) = 1\}$. Without loss of generality, we can simply assume that spin $x = 1$ is the one that is minimized for $\bar{S}_{A,x}$ as in the definition of $\bar{S}_A$. Recall that P is the transition matrix of the Glauber dynamics. Then,

$$\begin{aligned}
\Phi_U &= \frac{\sum_{\sigma\in U}\pi(\sigma)\sum_{w\in L}\sum_{\tau:\tau(w)\neq\sigma(w),A(\tau)\neq A(\sigma)}P(\sigma,\tau)}{\pi(U)(1-\pi(U))}\\
&\le \frac{\sum_{\sigma\in U}\pi(\sigma)\sum_{w\in L}\sum_{\tau:\tau(w)\neq\sigma(w),A(\tau)\neq A(\sigma)}\left(\frac{1}{n}\pi_\sigma(\tau)\right)}{\pi(U)(1-\pi(U))}\\
&\le \frac{\sum_{\sigma\in U}\pi(\sigma)\frac{1}{n}\sum_{w\in L}\mathbf{1}(\exists c\in Q, A(\sigma^{w\to c})\neq A(\sigma))}{\pi(U)(1-\pi(U))}\\
&\le \frac{\sum_{\sigma\in U}\pi(\sigma)S_A(\sigma)}{\pi(U)(1-\pi(U))}\\
&= \frac{\bar{S}_{A,1}}{\pi(U)(1-\pi(U))}\\
&\le \frac{\bar{S}_A}{r_{h,A}^2} \qquad\qquad \text{by the definition of } r_{h,A} \text{ and } |Q|\ge 2.
\end{aligned}$$

Because the algorithm is effective, we have that $\liminf_{h\to\infty}(r_{h,A}) = c_0 > 0$ and hence for all h big enough, $r_{h,A} > c_0/2$. Therefore, $\Phi_U \le (r_{h,A})^{-2}\bar{S}_A = O(\bar{S}_A)$, and hence,

$$\Phi \le \Phi_U = O\left(\bar{S}_A\right),$$

which completes the proof of the theorem. □

In the following two sections, we will use the above Theorem 12 to establish the upper bound of the conductance and hence prove the lower bounds of the mixing time and relaxation time of the Glauber dynamics in Theorem 2 for k-colorings and Theorem 3 for hard-core models on the Δ-regular trees.

3.2 Colorings

Here we first give a simple reconstruction algorithm for colorings, then we will bound the conductance via analyzing this algorithm. Specifically, we will prove the following theorem for the conductance.

Theorem 13. *For all $\epsilon > 0$, there exists b_0 such that, for all $b > b_0$, for $k = \frac{b}{(1+\epsilon)\ln b}$, the conductance of the Glauber dynamics on the $(b+1)$-regular tree T of n vertices* *and height $H = \lfloor \log_b n \rfloor$ satisfies the following:*

$$\Phi = O\left(n^{-(1+\epsilon-o_b(1))}\right),$$

where the $o_b(1)$ function is $(1+\epsilon)/b^\epsilon$ for the lower bounds when $0 < \epsilon < 1$ and exactly zero for the lower bounds when $\epsilon \geq 1$. The constants in the $O(\cdot)$ are universal constants.

Then, by the relationship between the conductance and the relaxation time in Eq. (2) and the relationship between the relaxation time and the mixing time in Eq. (1), we are able to lower bound both the relaxation time and mixing time by $\Omega\left(n^{1+\epsilon-o_b(1)}\right)$, and hence prove the lower bounds in Theorem 2.

The intuition why the Glauber Dynamics slows down when $k < b/\ln b$ is that for a typical configuration of the complete tree, the configuration of the root is hard to change. The children of the root will have all the colors appearing and hence in order to change the color of the root, the dynamics has to first change the spin of one of the children such that the root has a choice to switch to some other colors. As we will show later, this is impossible since the configuration of the root is actually *frozen* by the configurations on the leaves, i.e., there is a subtree of the complete tree such that if one wants to change the color of the root, then one has to start from the leaves. There is no "free" internal vertex.

The algorithm (denoted as FR, stands for "frozen") is a function that maps each $\sigma \in \Omega(T_H)$ to $Q \cup \{\diamondsuit\}$, where the diamond mark $\diamondsuit$ is a special color called "unknown". It works in a bottom up manner from the configurations on the leaves: for each parent v of the leaves, if its children contains all colors in Q except for $\sigma(v)$, then the algorithm marks v to color $\sigma(v)$ (denote the mark of v as $\mathrm{FR}(\sigma, v)$); otherwise, $\mathrm{FR}(\sigma, v) = \diamondsuit$. Then, the algorithm marks the vertices two level above the leaves in the same manner recursively. The mark of the root is the output of the algorithm. Formally, we describe the algorithm as below. The configuration σ is called *frozen* if

Algorithm 1 FR(σ, v) : Reconstruction Algorithm for Colorings

```
if v is not leaf then
  for w ∈ N⁻(v) do
    FR(σ, w)
  end for
  if |{FR(σ, w) : w ∈ N⁻(v)} \ {◇}| = q − 1 then
    return  FR(σ, v) ← Q \ {FR(σ, w) : w ∈ N⁻(v)}
  else
    return  FR(σ, v) ← ◇
  end if
else
  return  FR(σ, v) ← σ(v)
end if
```

FR$(\sigma, r) = \sigma(r)$ meaning that the root configuration of σ is uniquely determined by the configurations on the leaves.

In coloring $\sigma \in \Omega(T_H)$, we say a vertex v is *frozen* in σ if in the subtree T_v the coloring $\sigma(L(T_v))$ of the leaves of T_v forces the color for v. In other words, v is frozen in σ if: for all $\eta \in \Omega$ where $\eta(L(T_v)) = \sigma(L(T_v))$, we have $\eta(v) = \sigma(v)$ and hence FR$(\sigma, v) = \sigma(v)$. Note, by definition, the leaves are always frozen. If the vertex v is not frozen, then FR$(\sigma, v) = \diamondsuit$. Observe that for a vertex to be frozen, its frozen children must "block" all other color choices. This is formalized in the following observation as in [29].

Observation 14. *A non-leaf vertex v is frozen in coloring σ if and only if, for every color $c \neq \sigma(v)$, there is a child w of v where $\sigma(w) = c$ and w is frozen.*

Using this inductional way of defining a vertex being "frozen" in a coloring, we can further show the following lemma. It is a generalization of Lemma 8 in [29], which only applied to the case $\epsilon \geq 1$.

Lemma 15. *For any $\epsilon \in (0, 1)$, in a random coloring of tree T_H, the probability that a vertex of the tree is not frozen is at most $b^{-\epsilon}$. For the leaves in T_H, by definition, they are always frozen.*

Proof of Lemma 15. The proof is very similar to the proof of Lemma 8 in [29]. We include it here for completeness.

Let U_ℓ be the probability that a vertex at the height ℓ is not frozen. We are going to prove that $U_\ell < b^{-\epsilon}$ by induction.

First of all, by definition, $U_0 = 0$ since they are leaves. Let v be a vertex at height $\ell > 0$. Since the probability that the color of v equals c is independent from the probability that v is frozen, therefore we can just fix the color of v to some $c^* \in \mathcal{C}$, and hence

$$U_\ell = \Pr\left[v \text{ is not frozen in } \sigma \mid \sigma(v) = c^*\right].$$

Let w be a child of v. Again by the same argument using the independency, the probability that w is frozen to color c equals $\frac{1-U_{\ell-1}}{k-1}$. Thus, the probability that all the children of v are either not frozen or not colored by using c is $(1-(1-U_{\ell-1})/(k-1))^b$.

By the union bound and induction, U_ℓ, the probability that there exists a color $c \neq c^*$ such that all the children of v are not frozen to color c, is bounded by:

$$\begin{aligned}
(k-1)\Big(1 - \frac{1-U_{\ell-1}}{k-1}\Big)^b &\leq (k-1)\exp\Big(-\frac{b(1-U_{\ell-1})}{k-1}\Big) \\
&\leq (k-1)\exp\Big(-\frac{b(1-b^{-\epsilon})}{k-1}\Big) \\
&\leq (k-1)\cdot\exp\Big(-(1+\epsilon)(1-b^{-\epsilon})\ln b\Big) \\
&\leq b^{-\epsilon}, \qquad (4)
\end{aligned}$$

where the last inequality holds for large b. Putting everything together, we showed that for every $\ell \leq H$, $U_\ell \leq b^{-\epsilon}$.

Using the same argument, we can prove a slightly stronger result: for any $\epsilon > 0$, $U_\ell \leq \frac{1}{b^\epsilon \ln b}$, and when $\epsilon \geq 1$, $U_\ell \leq \frac{1}{(1+\epsilon)b}$. □

Note that in the reconstruction algorithm FR, we have for any color $c \in Q$, $\mathrm{FR}(\sigma) = c$ implies $\sigma(r) = c$. Then, by a direct calculation, this lemma also implies the fact that

$$\liminf_{h\to\infty} r_{h,\mathrm{FR}} > 0,$$

which means that the algorithm FR is effective.

3.2.1 Upper Bound on the Conductance via Sensitivity Analysis

For the remainder of this section, we will analyze the sensitivity of the reconstruction algorithm FR on the regular tree of height H. Let us denote $T = T_H$ for simplicity.

Let $F_c = F_c(T)$ denote those colorings in $\Omega(T)$ where the root r of T is frozen to color $c \in Q$, i.e.,

$$F_c(T) = \{\sigma \in \Omega(T) : \mathrm{FR}(\sigma, r) = c\}.$$

Note that $\mu(F_c(T)) \leq 1/k$ by the symmetry. For each $c \in Q$, we will analyze the sensitivity of the reconstruction algorithm FR in set $F_c(T)$ to upper bound the conductance.

Fix a color c^*, to this end, we should bound the number of colorings $\sigma \in F_{c^*}$ which can leave the set with one transition, and also the total number of transitions leaving F_{c^*}. To unfreeze the root, one has to recolor a leaf. Thus, we need to bound the number of colorings frozen at the root which can become unfrozen by one recoloring, and in that case, we need to bound the number of leaves which can be recolored to unfreeze the root. For a coloring σ, vertex v and color c, let $\sigma^{v \to c}$ denote the coloring obtained by recoloring v to c.

We capture the colorings on the "frontier" of F_{c^*} as follows. For tree $T = T_H$, coloring $\sigma \in \Omega(T)$, a vertex v and a leaf z of T_v, let $\mathcal{E}^{c^*}_{v,z}$ denote the event that the coloring $\sigma \in \Omega$ is frozen to c^* at the vertex v and there exists a color i where the coloring $\sigma^{z \to i}$ is not frozen at the vertex v, i.e.,

$$\mathcal{E}^{c^*}_{v,z} = \left\{\sigma \in \Omega : \mathrm{FR}(\sigma, v) = c^* \text{ and } \exists i \in Q, \mathrm{FR}(\sigma^{z \to i}, v) = \Diamond \right\}.$$

By definition, this event only depends on the configurations at the leaves of the subtree T_v. In particular, for the root of the tree, let $\mathbf{1}_{\sigma,z,c^*}$ be the indicator of the event $\mathcal{E}^{c^*}_{r,z}$ for a specific configuration σ. According to the definition of the sensitivity

$\bar{S}_{\mathrm{FR},c^*}$, we have

$$\begin{aligned}
\bar{S}_{\mathrm{FR},c^*} &= \mathrm{E}_{\sigma\sim\nu_H}\left[S_{\mathrm{FR}}(\sigma)\mathbf{1}(\mathrm{FR}(\sigma)=c^*)\right] \\
&= \frac{1}{n}\mathrm{E}_{\sigma\sim\nu_H}\left[\sum_{z\in L(T)}\mathbf{1}\big(\exists i\in Q, \mathrm{FR}(\sigma^{v\to i})\neq \mathrm{FR}(\sigma)\big)\cdot\mathbf{1}(\mathrm{FR}(\sigma)=c^*)\right] \\
&= \frac{1}{n}\mathrm{E}_{\sigma\sim\nu_H}\left[\sum_{z\in L(T)}\mathbf{1}_{\sigma,z,c^*}\right] \\
&= \frac{1}{n}\sum_{z\in L(T)}\Pr\left[\mathcal{E}_{r,z}^{c^*}\right], \qquad (5)
\end{aligned}$$

where the probability distribution is ν_H over the configurations on the tree $T=T_H$.

Now if we can prove that

$$\Pr_{\sigma\in\Omega}\left[\mathcal{E}_{r,z}^{c^*}\right]\leq b^{-(1+\epsilon-o(1))H}, \qquad (6)$$

where $o(1)$ is an inverse polynomial of b when $\epsilon<1$ and equals to zero when $\epsilon\geq 1$. This will be clarified later in the proof of Lemma 16. Then by plugging this back into Equation (5) we get

$$\bar{S}_{\mathrm{FR},c^*}\leq \frac{1}{n}\cdot b^H\cdot b^{-(1+\epsilon-o(1))H}\leq 20n^{-1-\epsilon+o(1)}.$$

Therefore, we can conclude that the sensitivity of the reconstruction algorithm FR on the complete tree is upper bounded by $O(n^{-1-\epsilon+o(1)})$, and hence the conductance of the Glauber dynamics is $O(n^{-1-\epsilon+o(1)})$. This proves Theorem 13.

3.2.2 Sensitivity Analysis of a Reconstruction Algorithm: Proof of Equation (6)

Let $\Omega^*=\{\sigma\in\Omega:\sigma(r)=c^*\}$ be the set of colorings where the root is colored c^*. By symmetry and conditioning on Ω^*, it is easy to see that

$$\Pr_{\sigma\in\Omega}\left[\mathcal{E}_{r,z}^{c^*}\right]=\frac{1}{k}\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}_{r,z}^{c^*}\right].$$

Therefore, for the remainder of the proof we condition on the root being colored c^*.

To simplify the notation, we denote $B := b^{-(1+\epsilon-o(1))}$. Let $\mathcal{E}_{v,z}$ be the union of $\mathcal{E}^c_{v,z}$ over all $c \in Q$, i.e., $\mathcal{E}_{v,z}$ is the event that includes those configurations σ where v is frozen (to some color) in σ, and there exists a color i where the coloring $\sigma^{z\to i}$ is not frozen at the vertex v. Let $w_0 = r, w_1, \ldots, w_{H-1}, w_H = z$ denote the path in T from the root r down to the leaf z. We will show by induction that,

$$\begin{aligned}
\Pr_{\sigma\in\Omega^*}\left[\mathcal{E}^{c^*}_{r,z}\right] &= \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}_{r,z}\right] \\
&\leq B \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}_{w_1,z}\right] \\
&\leq B^2 \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}_{w_2,z}\right] \\
&\leq B^H \Pr_{\sigma\in\Omega^*}\left[\mathcal{E}_{w_H,z}\right] \\
&= B^H.
\end{aligned}$$

Recall that we always condition on the event that the root is colored to a fixed spin c^*, therefore the first equality above is trivial. For the event $\mathcal{E}_{r,z}$ to occur we need that along the path from the leaf z to the root r, unfreezing each of these vertices will "free" a color for their parent. More precisely, for σ to be in $\mathcal{E}_{r,z}$, w_1 has to be frozen because the color of z only affects the root through w_1, and if w_1 is not frozen then it cannot affect the root becoming unfrozen. In order for the root to become unfrozen by changing the color of the leaf z, it must also occur that w_1 becomes unfrozen at the same time, hence $\sigma \in \mathcal{E}_{w_1,z}$, i.e., $\mathcal{E}_{r,z} \subseteq \mathcal{E}_{w_1,z}$ and more generally, $\mathcal{E}_{w_i,z} \subseteq \mathcal{E}_{w_{i+1},z}$. For each $1 \leq i \leq H$, let $\mathcal{A}_{w_i,z}$ denote the event that includes those configurations σ where no sibling y of w_i satisfies both of the following: $\sigma(y) = \sigma(w_i)$ and σ is frozen at y. By the siblings of w_i, as usual we mean the children (other than w_i) of w_{i-1}. The event $\mathcal{E}_{w_i,z}$ implies the fact that w_{i+1} is the only child that causes w_i simultaneously being frozen and being blocked from using color $\sigma(w_{i+1})$, which means $\mathcal{E}_{w_i,z} \subseteq \mathcal{A}_{w_{i+1},z}$. We will show the following lemma for bounding the probability of $\mathcal{A}_{w_1,z}$.

Lemma 16. *Let $\mathcal{C}^* = \mathcal{C} - c^*$. For a fixed color $c_1 \in \mathcal{C}^*$,*

$$\Pr_{\sigma\in\Omega^*}\left[\mathcal{A}_{w_1,z} \mid \sigma(w_1) = c_1\right] \leq B = b^{-(1+\epsilon-o(1))}.$$

where $o(1)$ is a function that goes to zero when $b \to \infty$.

Proof of Lemma 16. When $\epsilon < 1$, the probability that all the siblings of w_1 are either not frozen or not colored with c_1 is upper bounded by

$$\begin{aligned}\Big(1 - \frac{1-U_{H-1}}{k-1}\Big)^{b-1} &\leq \exp\Big(-\frac{(b-1)(1-b^{-\epsilon})}{k-1}\Big) \\ &\leq b^{-(1+\epsilon)(1-b^{-\epsilon})},\end{aligned}$$

where U_{H-1} is defined in Lemma 15.

Now we can see that $o(1)$ is actually $(1+\epsilon)/b^{\epsilon}$ when $\epsilon < 1$. Note that, when $\epsilon \geq 1$, by the same way it is easy to see that

$$\Big(1 - \frac{1-U_{H-1}}{k-1}\Big)^{b-1} \leq b^{-(1+\epsilon)}.$$

□

Observe that the events $\mathcal{A}_{1,z}$ and $\mathcal{E}_{w_1,z}$ are independent, conditioned on the fixed colors of the root and w_1, because they depend on the configurations of different parts of leaves. Then we have that for each $c_1 \in \mathcal{C}^*$,

$$\begin{aligned}&\mathrm{Pr}_{\sigma\in\Omega^*}\Big[(\sigma(w_1)=c_1)\bigcap\mathcal{E}_{w_1,z}\bigcap\mathcal{A}_{1,z}\Big] \\ &= \mathrm{Pr}_{\sigma\in\Omega^*}\big[\mathcal{E}_{w_1,z} \mid \sigma(w_1)=c_1\big]\cdot\mathrm{Pr}_{\sigma\in\Omega^*}\big[\mathcal{A}_{1,z} \mid \sigma(w_1)=c_1\big]\cdot\frac{1}{k-1} \\ &\leq \frac{B^{H-1}\cdot B}{k-1},\end{aligned} \tag{7}$$

where the last inequality is by the inductive hypothesis applied on the complete tree T_{w_1} of height $H-1$ and Lemma 16.

Finally, by the fact that $\mathcal{E}_{r,z} \subseteq \mathcal{E}_{w_1,z}\bigcap\mathcal{A}_{1,z}$ and (7) above, we have

$$\begin{aligned}\mathrm{Pr}_{\sigma\in\Omega^*}\big[\mathcal{E}_{r,z}\big] &\leq \mathrm{Pr}_{\sigma\in\Omega^*}\Big[\mathcal{E}_{w_1,z}\bigcap\mathcal{A}_{w_1,z}\Big] \\ &= \sum_{c_1\in\mathcal{C}^*}\mathrm{Pr}_{\sigma\in\Omega^*}\Big[(\sigma(w_1)=c_1)\bigcap\mathcal{E}_{w_1,z}\bigcap\mathcal{A}_{w_1,z}\Big] \\ &\leq B^H.\end{aligned}$$

This completes the proof of (6) and hence the proof of the upper bound in Theorem 13.

3.3 Hard-core Model

In this section, we aim at show the following theorem for the conductance of the Glauber dynamics for hard-core models on the Δ-regular tree.

Theorem 17. *For the Glauber dynamics of the hard-core model with activity* $\lambda = \omega(1+\omega)^b$ *on the* $(b+1)$*-regular tree* T_H *with* n *vertices and height* $H = \lfloor \log_b n \rfloor$*, the following holds:*
For all $\delta > 0$ *and* $\omega = (1+\delta)\ln b/b$*, there exists a sequence of boundary conditions for all* $H \to \infty$ *such that,*

$$\Phi = \Omega\left(n^{-(1+\delta/2-o_b(1))}\right).$$

Then again, by the relationship between the conductance and the relaxation time in Eq. (2), we prove the lower bounds in Theorem 3.

Martinelli *et al.* [47] showed that for the hard-core model on T_h with free boundary condition the relaxation time is $O(n)$ for all λ (and the mixing time is $O(n \log n)$). In contrast, recall that from Bhatnagar *et al.* [9] and Brightwell and Winkler [13], it is known that the reconstruction threshold of the broadcasting process for the hard-core model on the infinite b-ary tree is at $\omega_r = (\ln b + (1+o(1)) \ln\ln b)/b$. We first point out that there is no contradiction here about these two facts, by noticing the fact that the free boundary condition gives an infinite measure that is very different from the broadcasting process. Therefore, the rapid mixing result of the Glauber dynamics does not contradict with the conjecture that the reconstruction threshold should overlap with the phase transition threshold for the Glauber dynamics on the complete trees.

Hence, for the hard-core model, the question is even harder: whether there is a boundary condition for which the Glauber dynamics slows down beyond a certain threshold of λ, and whether the threshold is the same as the reconstruction threshold for the broadcasting process. We prove positively that there is a connection by constructing a boundary condition for which the relaxation time slows down at the reconstruction threshold.

There are two major difficulties we were facing: one is to identify a proper subset of the state space with poor conductance, such that the corresponding conductance bound closely matches the relaxation time of the Glauber dynamics. The conductance of such a subset should be sensitive to the boundary conditions, as we already know that the Glauber dynamics is rapidly mixing under properly chosen boundary conditions. The other difficulty, once we realized that the relaxation time of the Glauber dynamics can be nontrivially lower bounded under a nonuniform hard-core model (see Section 3.3.1 for details), is to prove that, when reconstruction happens, such a nonuniform model can be approximated (in the measure sense), by an appropriate sequence of boundary conditions. As a result, then we are able to show Theorem 17 via Theorem 12.

We begin by explaining the high level idea of the proof approach. To that end, we first analyze a variant of the hard-core model in which there are two different activities, the internal vertices have activity λ and the leaves have activity ω. The resulting Gibbs distribution is identical to the measure ν_h which is obtained by following the broadcasting process defined in Section 2.3 from the root to the level h. Thus we refer to the following model as the broadcasting model.

For the tree $T_h = (V, E)$, we look at the following equivalent definition of the distribution ν_h over the set Ω of independent sets of T_h. For $\sigma \in \Omega$, let

$$w'(\sigma) = \lambda^{|\sigma \cap V \setminus L|} \omega^{|\sigma \cap L|},$$

where L are the leaves of T_h and ω is, as before, the positive solution to $\omega(1+\omega)^b = \lambda$. Let $\nu_h(\sigma) = w'(\sigma)/Z'$ where $Z' = \sum_{\sigma \in \Omega} w'(\sigma)$ is the partition function. By simple calculations, the following proposition holds.

Proposition 18. *The measure ν_h defined by the hard-core model with activity λ for* *internal vertices and ω for leaves is identical to the measure defined by the broadcasting process.*

Proof. In fact, we just need to verify that in the hard-core model with activity λ for internal vertices and ω for leaves, the probability p_v of a vertex v being occupied conditioning on its parent is unoccupied is $\omega/(1+\omega)$. This can be proved by induction. The base case is v being a leaf, which is obviously true by the Markovian property of the Gibbs measure. If v is not a leaf, by induction, the probability p_v has to satisfy the following equation

$$p_v = (1 - p_v)\frac{\lambda}{(1+\omega)^b},$$

which solves to $p_v = \omega/(1 + \omega)$. □

The result of Berger *et al.* [8] mentioned before implies that, in the reconstruction region, the relaxation time of the Glauber dynamics on the broadcasting model is $\omega(n)$. We will prove a stronger result, analogous to the desired upper bound of the conductance for Theorem 17.

Theorem 19. *For all $\delta > 0$, the Glauber dynamics for the broadcasting model on the $(b + 1)$-regular tree T_H with n vertices, $H = \lfloor \log_b n \rfloor$ and $w = (1 + \delta) \ln b/b$ satisfies the following:*

$$\Phi = O\left(n^{-(1+\delta/2-o_b(1))}\right) \quad \textit{and} \quad T_{\text{relax}} = \Omega\left(n^{1+\delta/2-o_b(1)}\right),$$

where the $o_b(1)$ function is $O(\ln \ln b/ \ln b)$.

To prove Theorem 19, we analyze the sensitivity of the reconstruction algorithm by Brightwell and Winkler [13, Section 5] which yields the best known upper bounds on the reconstruction threshold. Our goal is to show that the average sensitivity of this algorithm is small. The analysis of the sensitivity of the Brightwell-Winkler (BW) algorithm, which then proves Theorem 19, is presented in Section 3.3.1.

Our main objective remains of constructing a sequence of "bad" boundary conditions under which the Glauber dynamics for the hard-core model slows down in the reconstruction region. An initial approach is to simulate the nonuniform hard-core model on T by by attaching the same tree T' (with boundary conditions) to all of the leaves of a complete tree T, where T' is a (small) complete tree with some boundary condition such that the marginal of the root being occupied is $\omega/(1+\omega)$. In this case, the resulting measure projected onto T is the same as the one in the broadcasting model, and hence we can apply the same approach to upper-bound the conductance of the dynamics on this new augmented tree. However, from a cardinality argument, it is not the case that for every ω there exists a complete tree of finite height with some boundary condition such that the marginal probability of the root being occupied equals $\omega/(1+\omega)$. Alternatively, we give a constructive way to find boundary conditions that approximate the desired marginal probability relatively accurately. This is done in Section 3.4.

Finally, at the end of Section 3.4 we argue that since the error is shrinking very fast from the bottom level under our construction of boundary conditions, we can again analyze the sensitivity of the Brightwell-Winkler algorithm starting from just a few levels above the leaves. This approach yields the upper bound of the conductance of the dynamics stated in Theorem 17.

3.3.1 Lower Bound for Broadcasting: Proof of Theorem 19

Throughout this section we are working with the tree $T_h = (V, E)$, which is the complete tree of height h and branching factor b. We denote L as the leaves of T_h, and for $v \in V$, let $N(v)$ denote the children of v. We will focus on the broadcasting model ν_h where each independent set σ of tree T_h is weighted by $\lambda^{|\sigma \cap V \setminus L|}\omega^{|\sigma \cap L|}$. Recall that $\lambda = \omega(1+\omega)^b$. For simplicity, we identify σ with its characteristic function. We use the following function definition for σ: $\sigma(v) = 1$ if $v \in \sigma$, and $\sigma(v) = 0$ if $v \notin \sigma$.

To prove Theorem 19 we analyze the average sensitivity of the following reconstruction algorithm used by Brightwell and Winkler [13], which we refer to as the BW algorithm. For any configuration $\sigma \in \Omega$ as the input (or it suffices to have the assignment σ_h for the leaves), the algorithm works in the following bottom up manner labeling each vertex starting from the leaves: a vertex v is labeled to occupied if all of its children $N(v)$ are labeled to unoccupied; otherwise, v is labeled to unoccupied if at least one of its children $N(v)$ is occupied. The algorithm will output the labeling of the root as the final result. Formally, it can be described by the following deterministic recursion deciding the labeling of every vertex:

$$\mathrm{BW}(\sigma_h, v) = \begin{cases} \sigma(v) & \text{if v} \in L \\ 1 - \max_{w \in N(v)} \mathrm{BW}(\sigma_h, w) & \text{otherwise} \end{cases}$$

Finally, let $\mathrm{BW}(\sigma) = \mathrm{BW}(\sigma_h, r)$, where r is the root of the tree. Note that, $\mathrm{BW}(\sigma)$ only depends on the initial configuration σ_h of the leaves. The algorithm is proved to be effective in [13] for all $\omega = (1+\delta)\ln b/b$ where $\delta > 0$. Therefore, their algorithm can be used under our framework to lower bound the relaxation time.

In the BW algorithm, by definition, we have that the average sensitivity satisfies:

$$\bar{S}_{\mathrm{BW}} = O\big(n^{-1}\mathrm{E}_{\sigma\sim\nu_h}\left[\#\{z \in L : \mathrm{BW}(\sigma) = 1 \text{ and } \mathrm{BW}(\sigma^z) = 0\}\right]\big). \tag{8}$$

Due to the symmetry of the function $\mathrm{BW}(\sigma_h, v)$ and the measure ν_h, the expectation can be further simplified as follows. Fix a leaf z^*, we have that:

$$\begin{aligned} &\mathrm{E}_\sigma\left[\#\{z \in L : \mathrm{BW}(\sigma) = 1 \text{ and } \mathrm{BW}(\sigma^z) = 0\}\right] \\ &\qquad = b^h \nu_h\left(\mathrm{BW}(\sigma) = 1 \text{ and } \mathrm{BW}(\sigma^{z^*}) = 0\right). \end{aligned} \tag{9}$$

Observe that, for each vertex v and each configuration σ, if $\mathrm{BW}(\sigma_h, v) \neq \mathrm{BW}(\sigma_h^z, v)$, then z^* is a leaf on the subtree rooted at v and moreover, for each child w of v which is not on the path from v to z^*, $\mathrm{BW}(\sigma_h, w) = 0$. This fact leads to the following lemma that we will use to upper bound the right hand side of (9).

Lemma 20. *Let z^* be a leaf of T_h, and let $z^* = u_0, u_1, \ldots, u_h = r$ be the path between z^* and the root of T_h. For each $i > 0$, let*

$$f_i = \nu_{i-1}(\sigma : BW(\sigma) = 0)$$

denote the probability that for the broadcasting model on the complete tree of height $i-1$, the Brightwell-Winkler algorithm outputs 0 for the root. Then,

$$\nu_h(BW(\sigma) = 1 \text{ and } BW(\sigma^{z^*}) = 0) \leq \mathrm{E}_{\sigma\sim\nu_h}\left[\prod_{i>0:\sigma(u_i)=0} (f_i)^{b-1}\right].$$

Proof. Fix a configuration $\sigma \in \Omega$ where $\mathrm{BW}(\sigma) = \mathrm{BW}(\sigma_h, r) = 1$. Let the path $\mathcal{P}$ from z^* to the root r be $u_0 = z^*$,u_1, u_2, $\cdots$, $u_h = r$. Let $\hat{N}(u_i) = N(u_i) \setminus \{u_{i-1}\}$. We want that $\mathrm{BW}(\sigma) = 1$ and $\mathrm{BW}(\sigma^{z^*}) = 0$, i.e., by changing σ only at z^*, the output of the BW algorithm changes from occupied to unoccupied for the labeling of the root. Two necessary conditions for this to occur are the following. First, the output of the BW algorithm along the path $\mathcal{P}$ alternates between occupied and unoccupied, i.e., σ satisfies $\mathrm{BW}(\sigma_h, u_i) = 1 - \mathrm{BW}(\sigma_h, u_{i-1})$ for all $i \geq 1$. Second, for all $i \geq 1$, for all children $w \in \hat{N}(u_i)$, we have $\mathrm{BW}(\sigma_h, w) = 0$. This two conditions ensure that if the configuration at u_i changes then the output of the BW algorithm will change for u_{i-1}. Hence,

$$\nu_h(\mathrm{BW}(\sigma) = 1 \text{ and } \mathrm{BW}(\sigma^{z^*}) = 0) \leq \nu_h(\sigma : \forall i > 0, w \in \hat{N}(u_i), \mathrm{BW}(\sigma_h, w) = 0). \tag{10}$$

To calculate the probability that a random $\sigma \sim \nu_h$ satisfies such conditions, it would be easier if we expose the configurations along the path $\mathcal{P}$. Let $\sigma_{\mathcal{P}}$ be the projection of σ on the path $\mathcal{P}$. Conditioning on a configuration $\sigma_{\mathcal{P}}$ on the path, the events $\mathrm{BW}(\sigma_h, w) = 0$ are independent for all $w \in \bigcup_{i>0} \hat{N}(u_i)$. Note that, given $\sigma(u_i) = 0$, we have for all $w \in \hat{N}(u_i)$, the conditional probability of $\mathrm{BW}(\sigma_h, w) = 0$ equals f_i.

Therefore,

$$\begin{aligned}
&\nu_h(\mathrm{BW}(\sigma) = 1 \text{ and } \mathrm{BW}(\sigma^{z^*}) = 0) \\
&\qquad \le \nu_h(\sigma : \forall i > 0, w \in \hat{N}(u_i), \mathrm{BW}(\sigma_h, w) = 0) \qquad\qquad \text{by (10)} \\
&\qquad = \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \nu_h(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i=1}^{h} \prod_{w \in \hat{N}(u_i)} \Pr_{\sigma \sim \nu_h} [\mathrm{BW}(\sigma_h, w) = 0 \mid \sigma(u_i) = \eta(u_i)] \\
&\qquad \le \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \nu_h(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i>0:\sigma(u_i)=0} \prod_{w \in \hat{N}(u_i)} \Pr_{\sigma \sim \nu_h} [\mathrm{BW}(\sigma_h, w) = 0 \mid \sigma(u_i) = 0] \\
&\qquad = \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \nu_h(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i>0:\sigma(u_i)=0} \prod_{w \in \hat{N}(u_i)} f_i \\
&\qquad = \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \nu_h(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i>0:\eta(u_i)=0} (f_i)^{b-1} \\
&\qquad = \mathrm{E}_{\sigma \sim \nu_h} \left[\prod_{i>0:\sigma(u_i)=0} (f_i)^{b-1} \right].
\end{aligned}$$

□

To use Lemma 20, we derive the following uniform upper bound on the probability f_i, for all i. Note that, since our bounds are asymptotic, we will always assume that the degree b is large enough with respect to δ to make our proofs simpler. In particular, for $\omega = (1+\delta)\ln b/b$ and $\lambda = \omega(1+\omega)^b$, let

$$b_0(\delta) := \min\{b' \ge 10^4 : \exp\left(\frac{2(1.01)(\omega b)^2}{\lambda}\right) \le 1.01 \text{ for all } b > b'\}. \qquad (11)$$

(Note, the extra factor of 2 in the exponential is not needed in the proof of Lemma 20, but is convenient in Section 3.4 for the proof of Proposition 29.) Note that, $b_0(\delta)$ is well-defined since for any fixed δ,

$$\lim_{b\to\infty} \exp\left(\frac{2(1.01)(\omega b)^2}{\lambda}\right) = \lim_{b\to\infty} \exp\left(\frac{2(1.01)\ln b}{b^\delta}\right) < 1.01.$$

Lemma 21. *For all $\delta > 0$, all $b > b_0(\delta)$ and $i \ge 1$, setting $\omega = (1+\delta)\ln b/b$, we have,*

$$f_i \le \frac{(1.01)^{1/b}}{1+\omega}.$$

Proof. We will prove the lemma by induction. We first derive the recurrence of f_i for each i. For the base case $i = 1$, by the definition of the broadcasting model,

$$f_1 = \frac{1}{1+\omega}.$$

When $i = 2$, f_2 is the probability that the complete tree of height two has at least one child that is occupied. This requires to first unoccupy the root with probability $1/(1+\omega)$ and then have at least one child occupied. Therefore,

$$f_2 = \frac{1}{1+\omega}\left(1 - \left(\frac{1}{1+\omega}\right)^b\right).$$

Generally, one can see the recurrence holds for f_{i+1} by looking into two cases of σ sampled from distribution ν_h: Occupying the root r with probability $\omega/(1+\omega)$ in σ and then calculate the conditional probability of having at least one child that is labeled as 1 (occupied) in the BW algorithm. This is the complement of the event that all of the children of r having at least one of their own children reconstruct to occupied in the BW algorithm, given the fact that all children of r are fixed to unoccupied in σ. The probability of this event happening equals $(1-(f_{i-1})^b)^b$. The second case occurs when we do not occupy the root with probability $1/(1+\omega)$ in σ and then the event occurs that at least one child is labeled 1 in the BW algorithm. Therefore, we have

$$f_{i+1} = \frac{\omega}{1+\omega}\left(1 - \left(1-(f_{i-1})^b\right)^b\right) + \frac{1}{1+\omega}\left(1-(f_i)^b\right). \tag{12}$$

By the inductive hypothesis, we have

$$\begin{aligned}
f_{i+1} &\le \frac{\omega}{1+\omega}\left(1 - \left(1-(f_{i-1})^b\right)^b\right) + \frac{1}{1+\omega} && \text{by (12)}\\
&\le \frac{\omega}{1+\omega}\left(1 - \left(1-\frac{1.01\omega}{\lambda}\right)^b\right) + \frac{1}{1+\omega} && \text{by the inductive hypothesis}\\
&\le \frac{1+\frac{1.01\omega^2 b}{\lambda}}{1+\omega} && \text{since } (1-t)^b \ge 1-tb \text{ for } t<1\\
&\le \frac{\exp(1.01\omega^2 b/\lambda)}{1+\omega} && \text{since } (1+t)\le e^t\\
&\le \frac{(1.01)^{1/b}}{1+\omega}. && \text{from the definition of } b_0(\delta) \text{ in (11)}
\end{aligned}$$

□

Now, we combine Lemmas 20 and 21 to prove Theorem 19.

Proof of Theorem 19. Fix a leaf z^* of T_h, and let $\mathcal{P}$ be the path $z^* = u_0, u_1, \ldots, u_h = r$ between z^* and the root of T_h. We upper bound the average sensitivity of the BW algorithm in the following way:

$$\begin{aligned}
\bar{S}_{\mathrm{BW}} &= O\big(\nu_h(\mathrm{BW}(\sigma)=1 \text{ and } \mathrm{BW}(\sigma^{z^*})=0)\big) && \text{by Equations (8) and (9)} \\
&= O\left(\mathrm{E}_{\sigma\sim\nu_h}\left[\left(\frac{1.01\omega(1+\omega)}{\lambda}\right)^{\#\{i:\sigma(u_i)=0\}}\right]\right) && \text{by Lemma 20 and Lemma 21.}
\end{aligned}$$

In this expectation, the number of unoccupied vertices in $\mathcal{P}$ can be trivially lower bounded by $h/2$, since it is impossible to have two consecutive occupied vertices in $\mathcal{P}$. Therefore, the above expectation can be easily bounded by $O^*(n^{-(1+\delta)/2})$ for $h = H$. This is not good enough in our case; to establish the existence of a phase transition we need a bound of the form $O^*(n^{-(1+\delta/2)})$. This improved bound will be a consequence of the following lemma.

Lemma 22. *For all $\delta > 0$, all $b > b_0(\delta)$, setting $\omega = (1+\delta)\ln b/b$, we have,*

$$\mathrm{E}_{\sigma\sim\nu_h}\left[\left(\frac{1.01\omega(1+\omega)}{\lambda}\right)^{\#\{i:\sigma(u_i)=0\}}\right] = O\left(\left[\frac{1.01\omega}{\lambda^{1/2}}\right]^h\right).$$

Lemma 22 is proved in Section 3.5. Then, by the fact that the height of the tree $h = H = \log_b n$, we have, for $\delta > 0$, and $\omega = (1+\delta)\ln b/b$, for all $b > b_0(\delta)$

$$\bar{S}_{\mathrm{BW}} = O\left(\left[\frac{1.01\omega}{\lambda^{1/2}}\right]^h\right) = O\left(n^{-\left[1+\frac{\ln\left(\lambda/(1.01\omega b)^2\right)}{2\ln b}\right]}\right).$$

Now, from the fact in [13, Section 5] that the BW algorithm is effective for all $\delta > 0$, $\omega > (1+\delta)\ln b/b$ and $b > b_0(\delta)$ (A similar statement is proved later in our paper in Proposition 29.), Theorem 12 applies, and the conclusion follows for the conductance, allowing us to conclude that for $\delta > 0$, and $\omega = (1+\delta)\ln b/b$, for all $b \geq b_0(\delta)$,

$$\Phi = O\big(n^{-d}\big) \quad \text{and} \quad T_{\mathrm{relax}} = \Omega\big(n^d\big),$$

where

$$d = \left(1 + \frac{\ln\left(\lambda/(1.01\omega b)^2\right)}{2\ln b}\right).$$

Theorem 19 is a simple corollary by noticing that $d = 1 + \delta/2 - O\left(\frac{\ln\ln b}{\ln b}\right)$. Note that, when $b < b_0(\delta)$, our bound is trivial. $\square$

3.4 "Bad" Boundary Conditions: Proof of Theorem 17

First, we will show that for any ω, there exists a sequence of boundary conditions on the leaves, denoted as $\Gamma_\omega := \{\Gamma_i\}_{i>0}$, one for each complete tree of height $i > 0$, such that if $i \to \infty$, the probability of the root being occupied converges to $\frac{\omega}{1+\omega}$. Later in this section we will exploit such a construction to attain in full the conclusion of Theorem 17.

As a first observation, note that, the Gibbs measure for the hard-core model on T_i with boundary condition Γ is the same as the Gibbs measure for the hard-core model (with the same activity λ) on the tree T obtained from T_i by deleting all of the leaves as well as the parent of each (occupied) leaf $v \in \Gamma$. It will be convenient to work directly with such "trimmed" trees, rather than the complete tree with boundary condition. Having this in mind, our construction will be inductive in the following way. We will define a sequence of (trimmed) trees $\{(L_i, U_i)\}_{i\geq 1}$ such that L_{i+1} is comprised of s_{i+1} copies of U_i and $b - s_{i+1}$ copies of L_i with $\{s_i\}_{i\geq 1}$ properly chosen. Similarly, U_{i+1} is comprised of t_{i+1} copies of U_i and $b - t_{i+1}$ copies of L_i, with $\{t_i\}_{i\geq 1}$ properly chosen. Note that, from the construction we can see that for each $i > 1$, the trees U_i and L_i are always Δ-regular for the internal vertices, therefore, essentially the trimming of the trees only happen at the leaves.

We will show that, for either $T_i^* = L_i$, or $T_i^* = U_i$, it is the case that the 'Q'-value, defined as:

$$Q(T_i^*) = \frac{\mu_{T_i^*}\left(\sigma(r) = 1\right)}{\omega\mu_{T_i^*}\left(\sigma(r) = 0\right)},$$

where $\mu_{T_i^*}(\cdot)$ is the hard-core measure on the trimmed tree T_i^*, satisfies $Q(T_i^*) \to 1$.

Note that if $Q(T_i^*) = 1$, then the probability of the root being occupied is $\omega/(1+\omega)$ as desired. To attain this, we will construct L_i and U_i in such a way that $Q(U_i) \geq 1$ and $Q(L_i) \leq 1$.

The recursion for $Q(L_{i+1})$ can be derived easily as

$$Q(L_{i+1}) = \frac{(1+\omega)^b}{(1+\omega Q(U_i))^{s_{i+1}}(1+\omega Q(L_i))^{b-s_{i+1}}},$$

and a similar equation holds for $Q(U_{i+1})$ by replacing s_{i+1} with t_{i+1}.

To keep the construction simple, we inductively define the appropriate t_i and s_i, so that once L_i and U_i are given, we let t_{i+1} be the minimum choice so that the resulting Q-value is ≥ 1, more precisely, we let:

$$t_{i+1} = \min\{\ell \in [0,b] : \frac{(1+\omega)^b}{(1+\omega Q(U_i))^{\ell}(1+\omega Q(L_i))^{b-\ell}} \geq 1\}. \tag{13}$$

And similarly, we let:

$$s_{i+1} = \max\{\ell \in [0,b] : \frac{(1+\omega)^b}{(1+\omega Q(U_i))^{\ell}(1+\omega Q(L_i))^{b-\ell}} \leq 1\}. \tag{14}$$

The recursion starts with U_1 being the graph of a single node and L_1 being the empty set, so that $Q(U_1) = \lambda/\omega$ and $Q(L_1) = 0$. Observe that, by definition, $s_{i+1} \in \{t_{i+1}, t_{i+1}+1\}$ and that the construction guarantees that the values $Q(L_i)$ are at most 1, and the values $Q(U_i)$ are at least 1. Therefore s_i and t_i are always well-defined for all i. The following simple lemma justifies the correctness of our construction.

Lemma 23.

$$\lim_{i\to\infty} Q(U_i)/Q(L_i) = 1.$$

Proof. It is easy to see that either $t_i = s_i$ (meaning that $Q(L_i) = Q(U_i) = 1$), or $t_i = s_i - 1$, which implies that

$$\frac{Q(U_i)}{Q(L_i)} = \frac{1+\omega Q(U_{i-1})}{1+\omega Q(L_{i-1})} < \frac{Q(U_{i-1})}{Q(L_{i-1})}.$$

Therefore the ratio is shrinking and bounded from below by 1. Suppose the limit is not 1 but some value $q > 1$, which implies that $Q(U_i)/Q(L_i) > q$ for all i. Then we have the following:

$$\begin{aligned}
\frac{Q(U_{i-1})}{Q(L_{i-1})} - \frac{Q(U_i)}{Q(L_i)} &= \frac{Q(U_{i-1})}{Q(L_{i-1})} - \frac{1+\omega Q(U_{i-1})}{1+\omega Q(L_{i-1})} \\
&= \frac{Q(U_{i-1}) - Q(L_{i-1})}{(1+\omega Q(L_{i-1}))Q(L_{i-1})} \\
&\geq \frac{(q-1)Q(L_{i-1})}{Q(L_{i-1})(1+\omega)} && \text{since } Q(U_i)/Q(L_i) > q \\
&= \frac{q-1}{1+\omega} \\
&> 0 && \text{since } q > 1.
\end{aligned}$$

Therefore as long as $q > 1$, we show that the difference between the ratios for each step i is at least some positive constant which is impossible. Hence the assumption is false, and it must be the case that $q = 1$. □

By this lemma, it is easy to check that if we let T_i^* to be equal to either U_i or L_i, then $Q(T_i^*) \to 1$. Indeed, we can show that the additive error decreases exponentially fast. The following lemma indicates that this is the case for $\omega < 1$ (although a similar result holds for any ω).

Lemma 24. *Let ϵ_i^+ be the value of $Q(U_i) - 1$ and let ϵ_i^- be the value of $1 - Q(L_i)$, then*

$$\epsilon_{i+1}^+ + \epsilon_{i+1}^- \leq \omega(\epsilon_i^+ + \epsilon_i^-).$$

Proof. Note, by algebraic manipulations we have:

$$\frac{(1+\omega)^b}{(1+\omega Q(U_i))^j(1+\omega Q(L_i))^{b-j}} = \frac{1}{(1+\frac{\omega}{1+\omega}\epsilon_i^+)^j(1-\frac{\omega}{1+\omega}\epsilon_i^-)^{b-j}}. \tag{15}$$

Now, let k be the largest index j over $[b]$ such that the denominator of the right-hand side of the previous expression is less than 1. Thus, $k+1$ will be the least index such that the denominator is greater than 1. Then by applying Equation (15) for $Q(U_{i+1})$ and $Q(L_{i+1})$,

$$
\begin{aligned}
\epsilon_{i+1}^{+} + \epsilon_{i+1}^{-} &= \frac{1}{(1+\frac{\omega}{1+\omega}\epsilon_i^{+})^{k}(1-\frac{\omega}{1+\omega}\epsilon_i^{-})^{b-k}} - \frac{1}{(1+\frac{\omega}{1+\omega}\epsilon_i^{+})^{k+1}(1-\frac{\omega}{1+\omega}\epsilon_i^{-})^{b-k-1}} \\
&= \frac{\frac{\omega}{1+\omega}(\epsilon_i^{+}+\epsilon_i^{-})}{(1+\frac{\omega}{1+\omega}\epsilon_i^{+})^{k+1}(1-\frac{\omega}{1+\omega}\epsilon_i^{-})^{b-k}} \\
&\leq \frac{\frac{\omega}{1+\omega}(\epsilon_i^{+}+\epsilon_i^{-})}{1-\frac{\omega}{1+\omega}\epsilon_i^{-}} \qquad \text{by the above property of } k+1 \\
&\leq \omega(\epsilon_i^{+}+\epsilon_i^{-}).
\end{aligned}
$$

□

Coming back to the original tree-boundary notation, let Γ_h^1 be the boundary corresponding to the trimming of the tree U_h and let Γ_h^2 be the boundary corresponding to the trimming of the tree L_h. Note that, according to our construction, it is clear that the trees U_h and L_h only differ at the bottom level and hence Γ_h^1 and Γ_h^2 are the boundary conditions on the leaves of the Δ-regular tree T_h.

By our construction, for any vertex v on the tree of height h, the measure from $\mu_{\Gamma_h^1}$ (or $\mu_{\Gamma_h^2}$) projected onto the space of the independent sets of the subtree rooted at v with the boundary condition corresponding to the correct part of Γ and the parent of v being unoccupied is either $\mu_{\Gamma_i^1}$ or $\mu_{\Gamma_i^2}$, where i is the distance of v away from the leaves on T_h. Conditioning on the parent of v being unoccupied, in the broadcast process defined in Chapter 2.3, we would occupy v with probability $\omega/(1+\omega)$. Therefore, in the above construction, the probability v is occupied (or rather unoccupied) is close to the desired probability, and the error will decay exponentially fast with the distance from the leaves. This is formally stated in the following corollary of Lemma 24.

Corollary 25. *Given any $\omega < 1$ and the complete tree of height i, for Γ equal to Γ_i^1 or Γ_i^2 inductively constructed above, we have*

$$
\left| \mu_\Gamma(\sigma(r)=0) - \frac{1}{1+\omega} \right| \leq \omega^{i-1}\lambda/b.
$$

Throughout the rest of this section it is assumed that we are dealing with the boundary conditions $\{\Gamma_h^1\}_{h\in\mathbf{N}}$ and $\{\Gamma_h^2\}_{h\in\mathbf{N}}$ constructed above. We will then show that for every $\omega = (1+\delta)\ln b/b$ under these two boundary conditions, the Glauber dynamics on the hard-core model slows down, whenever $\delta > 0$. As we know from Corollary 25, the error of the marginal goes down very fast, so that roughly we can think of the marginal distribution of the configurations on the tree from the root to the vertices a few levels above the leaves as being close to the broadcasting measure. Note that, Lemma 23 is already sufficient for the rest of our proof even if we did not prove Corollary 25. Corollary 25 provides us a concrete value of the height for each Δ and ω when the approximation is good enough for us to use.

In fact, by following the same proof outline as we did in Section 3.3.1, we are able to prove the same bounds for the hard-core model on regular trees under these boundaries. To do that we need a slight generalization of the reconstruction algorithm and extensions of the corresponding lemmas used in that section to handle the errors in the marginal probabilities.

To generalize the notion of a reconstruction algorithm to the case of a boundary condition we need to add an extra parameter ℓ depending only on ω and b. We will essentially ignore the bottom ℓ levels in the analysis, and we will use that for the top $h-\ell$ levels the marginal probabilities are close to those on the broadcasting tree. We define a *reconstruction algorithm with a parameter* ℓ for the tree T_h with boundary condition Γ as a function $A_\ell : \Omega(L_{h-\ell}) \to \{0,1\}$. The algorithm A_ℓ takes the configurations of the vertices at height $h-\ell$ as the input and tries to compute the configuration at the root. For any $\sigma \in \Omega(T_{h,\Gamma})$, the sensitivity is defined as:

$$S_{\ell,A}(\sigma) = \frac{1}{n}\#\left\{v \in L_{h-\ell} : A_\ell(\sigma^v_{h-\ell}) \neq A_\ell(\sigma_{h-\ell})\right\}.$$

The average sensitivity of the algorithm at height $h-\ell$ with respect to the boundary Γ is defined as:

$$\bar{S}^{\Gamma}_{\ell,A} = \mathrm{E}_\sigma\left[S_{\ell,A}(\sigma)\mathbf{1}(A_\ell(\sigma_{h-\ell}) = 1)\right].$$

And the effectiveness is defined as:

$$r_{\ell,A}^{\Gamma} = \min_{x\in\{0,1\}}[\mu_{h,\Gamma}(A_\ell(\sigma_{h-\ell}) = x \text{ and } \sigma(r) = x) - \mu_{h,\Gamma}(A_\ell(\sigma_{h-\ell}) = x)\mu_{h,\Gamma}(\sigma(r) = x)].$$

We can show the analog of Theorem 12 in this setting.

Theorem 26. *Suppose that A_ℓ is an effective reconstruction algorithm. Then, it is the case that the conductance Φ of the Glauber dynamics for the hard-core model on the tree of height h with boundary condition Γ, satisfies $\Phi = O(\bar{S}_{\ell,A}^{\Gamma})$.*

To bound the average sensitivity for the boundary conditions Γ_h^1 and Γ_h^2 constructed above, we again use the same BW algorithm as we analyzed for the broadcasting tree. As in Equations (8) and (9), it is again enough to bound the probability

$$\mu_{\Gamma_h}(\mathrm{BW}_\ell(\sigma_{h-\ell}) \;=\; 1 \text{ and } \mathrm{BW}_\ell(\sigma_{h-\ell}^{z^*}) \;=\; 0) \tag{16}$$

for a fixed vertex z^* at a distance ℓ from the leaves, although in this case, this probability will not be the same for all z^*. Let the path $\mathcal{P}$ from z^* to the root r be $u_\ell = z^*$, $u_{\ell+1}$, $u_{\ell+2}$, $\cdots$, $u_h = r$.

As in the proof of Lemma 20 in Section 3.3.1, let $\hat{N}(u_i) = N(u_i) \setminus \{u_{i-1}\}$ denote the children of u_i different from u_{i-1}. For $i > \ell$, consider some $w \in \hat{N}(u_i)$. Let $\Gamma(w)$ be the boundary condition Γ_h restricted to the subtree T_w of T_h rooted at the vertex w. These subtrees are of height i. Note that, by our construction of the boundary conditions, $\Gamma(w) = \Gamma_{i-1}^1$ or $\Gamma(w) = \Gamma_{i-1}^2$. Then (16) can be calculated by the following lemma, which is the analog of Lemma 20 for the broadcasting tree.

Lemma 27.

$$\begin{aligned}\mu_{\Gamma_h}(BW_\ell(\sigma_{h-\ell}) \;&=\; 1 \textit{ and } BW_\ell(\sigma_{h-\ell}^{z^*}) \;=\; 0) \\ &\le \mathrm{E}_\sigma\left[\prod_{i>\ell:\sigma(u_i)=0}\;\prod_{w\in\hat{N}(u_i)} \mu_{\Gamma(w)}(\eta : BW_\ell(\eta) = 0)\right],\end{aligned}$$

where the expectation is over the measure μ_{Γ_h}, and for each i and $w \in \hat{N}(u_i)$, configuration η is a random configuration on the subtree rooted at w chosen from the probability measure $\mu_{\Gamma(w)}$.

Proof. Let $\mathrm{BW}_\ell(\sigma_{h-\ell}, w)$ denote the labeling of the algorithm on vertex w from the input configurations σ on the vertices at height $h-\ell$ (i.e., level ℓ). By a similar argument as we did in the proof of Lemma 20,

$$
\begin{aligned}
\mu_{\Gamma_h}(\mathrm{BW}_\ell(\sigma_{h-\ell}) \;=\; 1 \text{ and } \mathrm{BW}_\ell(\sigma^{z^*}_{h-\ell}) \;=\; 0)& \\
\leq \mu_{\Gamma_h}(\sigma : \forall i > \ell, w \in \hat{N}(u_i), \mathrm{BW}_\ell(\sigma_{h-\ell}, w))& \\
\leq \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \mu_{\Gamma_h}(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i>\ell} \prod_{w \in \hat{N}(u_i)} \Pr_\sigma\left[\mathrm{BW}_\ell(\sigma_{h-\ell}, w) = 0 \mid \sigma(u_i) = \eta(u_i)\right]& \\
\leq \sum_{\eta \in \{0,1\}^{|\mathcal{P}|}} \mu_{\Gamma_h}(\sigma : \sigma_{\mathcal{P}} = \eta) \prod_{i>\ell:\eta(u_i)=0} \prod_{w \in \hat{N}(u_i)} \mu_{\Gamma(w)}(\eta : \mathrm{BW}_\ell(\eta) = 0)& \\
= \mathrm{E}_\sigma\left[\prod_{i>\ell:\sigma(u_i)=0} \prod_{w \in \hat{N}(u_i)} \mu_{\Gamma(w)}(\eta : \mathrm{BW}_\ell(\eta) = 0)\right].&
\end{aligned}
$$

□

To bound $\mu_{\Gamma(w)}(\eta : \mathrm{BW}_\ell(\eta) = 0)$ for every $i > \ell$ and $w \in \hat{N}(u_i)$, we proceed along the lines of Lemma 21, but extra care is required to deal with the errors in the marginal probabilities which were bounded in Corollary 25. Here and throughout the remainder of the paper, we define $\ell_0 = \ell(\lambda, b)$ to be

$$\ell_0 = \min\left\{\ell : \left|\frac{\mu_{\Gamma_i^1}(\eta : \eta(r) = 0)}{1/(1+\omega)} - 1\right| \leq \left(\exp\left(\frac{1.01\omega^2 b}{\lambda}\right) - 1\right) \text{ for all } i > \ell\right\}. \quad (17)$$

The existence of such a constant $\ell(\lambda, b)$ is guaranteed by Lemma 23. Moreover, from Corollary 25 we can deduce an explicit value for ℓ_0, provided that $\omega < 1$. For every $i \geq \ell$, let $f_{i,1} = \mu_{\Gamma_i^1}(\eta : \mathrm{BW}_\ell(\eta) = 0)$ and similarly let $f_{i,2} = \mu_{\Gamma_i^2}(\eta : \mathrm{BW}_\ell(\eta) = 0)$. We will use the following lemma to bound $f_{i,1}$ and $f_{i,2}$.

Lemma 28. *For all $\delta > 0$, all $b \geq b_0(\delta)$, there exist $\ell_0 = \ell(\lambda, b)$ such that for all $i > \ell_0$, the following bounds hold:*

$$f_{i,1} \leq \frac{1.01^{1/b}}{1+\omega},$$

and

$$f_{i,2} \leq \frac{1.01^{1/b}}{1+\omega}.$$

Proof. Similar to the proof of Lemma 21, the proof is again by induction. Here we take $\ell = \ell_0$. Let $\bar{t}_i = b - t_i$ and $\bar{s}_i = b - s_i$ for simplicity. Recall that we define t_i and s_i in Equation (13) and (14). Again, we can derive the recurrences in exactly the same way as in Lemma 21.

For the base case $i = \ell$, by the definition the algorithm will label the vertices on level ℓ to be the same as their actual configurations. For instance, for the boundary condition Γ_ℓ^1 on the complete tree of height ℓ, the root is unoccupied with probability $\mu_{\Gamma_\ell^1}(\eta(r) = 0)$ for a random configuration η. Therefore,

$$f_{\ell,1} = \mu_{\Gamma_\ell^1}(\eta(r) = 0), \quad f_{\ell,2} = \mu_{\Gamma_\ell^2}(\eta(r) = 0).$$

For the case $i = \ell + 1$, for a random configuration $\eta \sim \mu_{\Gamma_{\ell+1}^1}$, in order for the root r at level $\ell + 1$ to be labeled as 0 (unoccupied) in the algorithm, at least one child of r should be occupied in η since the algorithm takes input at level ℓ by the definition. This requires us to unoccupy the root with probability $\mu_{\Gamma_{\ell+1}^1}(\eta(r) = 0)$ and then have at least one child occupied in η, which happens with probability $(1 - f_{\ell,1}^{t_{\ell+1}} f_{\ell,2}^{\bar{t}_{\ell+1}})$. Note that, the boundary condition for the subtree rooted at each child is not the same. There are $t_{\ell+1}$ trees with boundary condition Γ_ℓ^1 and $\bar{t}_{\ell+1}$ trees with boundary condition Γ_ℓ^2 by the definition of $\Gamma_{\ell+1}^1$. The same argument holds for $f_{\ell+1,2}$. Therefore,

$$f_{\ell+1,1} = \mu_{\Gamma_{\ell+1}^1}(\eta(r) = 0) \left(1 - f_{\ell,1}^{t_{\ell+1}} f_{\ell,2}^{\bar{t}_{\ell+1}}\right),$$

$$f_{\ell+1,2} = \mu_{\Gamma_{\ell+1}^2}(\eta(r) = 0) \left(1 - f_{\ell,1}^{s_{\ell+1}} f_{\ell,2}^{\bar{s}_{\ell+1}}\right).$$

And for each $i > \ell$, by the same argument as in Lemma 21 and taking the boundary conditions into consideration as we did for $f_{\ell+1,1}$ and $f_{\ell+1,2}$, we have

$$\begin{aligned} f_{i+1,1} &= \mu_{\Gamma_i^1}(\eta(r)=1)\left(1-\left(1-f_{i-1,1}^{t_i}f_{i-1,2}^{\bar{t}_i}\right)^{t_{i+1}}\left(1-f_{i-1,1}^{s_i}f_{i-1,2}^{\bar{s}_i}\right)^{\bar{t}_{i+1}}\right) \\ &\quad + \mu_{\Gamma_i^1}(\eta(r)=0)\left(1-f_{i,1}^{t_{i+1}}f_{i,2}^{\bar{t}_{i+1}}\right), \qquad (18) \\ f_{i+1,2} &= \mu_{\Gamma_i^2}(\eta(r)=1)\left(1-\left(1-f_{i-1,1}^{t_i}f_{i-1,2}^{\bar{t}_i}\right)^{s_{i+1}}\left(1-f_{i-1,1}^{s_i}f_{i-1,2}^{\bar{s}_i}\right)^{\bar{s}_{i+1}}\right) \\ &\quad + \mu_{\Gamma_i^2}(\eta(r)=0)\left(1-f_{i,1}^{s_{i+1}}f_{i,2}^{\bar{s}_{i+1}}\right). \qquad (19) \end{aligned}$$

Our goal now, is to show by induction that $f_{i,1}, f_{i,2} \le \frac{1.01^{1/b}}{1+\omega}$ for all $i \ge \ell = \ell_0$. From the definition of ℓ_0 in (17), the base case is simple:

$$f_{\ell+1,1} \le f_{\ell,1} \le \mu_{\Gamma_\ell^1}(\eta(r)=0) \le \frac{1}{1+\omega}\exp\left(\frac{1.01\omega^2 b}{\lambda}\right),$$

and the last term is less or equal to $\frac{1.01^{1/b}}{1+\omega}$ for $b \ge b_0(\delta)$. Similarly, it is the case that $f_{\ell+1,2} \le f_{\ell,2} \le \frac{1.01^{1/b}}{1+\omega}$. Assuming the inductive hypothesis, by algebraic calculations, we can get from the above recurrence (18) that

$$\begin{aligned} (f_{i+1,1})^b &\le \left[\frac{\omega}{1+\omega}\left(1-\left(1-\frac{1.01\omega}{\lambda}\right)^b\right)+\frac{1}{1+\omega}\right]^b \exp\left(\frac{1.01(\omega b)^2}{\lambda}\right) \\ &\le \frac{\exp(1.01(\omega b)^2/\lambda)}{(1+\omega)^b}\exp\left(\frac{1.01(\omega b)^2}{\lambda}\right) \qquad \text{as in the proof of Lemma 21} \\ &= \frac{\exp(2(1.01)(\omega b)^2/\lambda)}{(1+\omega)^b} \\ &\le \frac{1.01}{(1+\omega)^b} \qquad \text{for } b \ge b_0(\delta)\text{, by the definition of } b_0(\delta) \text{ in (11)} \end{aligned}$$

This proves $f_{i+1,1} \le \frac{1.01^{1/b}}{1+\omega}$ by induction and a similar proof can be done for $f_{i+1,2}$. □

It is also not hard to show that the Brightwell-Winkler algorithm under the same setting as in Lemma 28 is effective.

Proposition 29. *For all $\delta > 0$ and $b > b_0(\delta)$, the BW algorithm is an effective reconstruction algorithm to recover the configuration at the root from the configurations at distance $\ell(\lambda, b)$ from the leaves.*

Proof. We use the same notation as in the proof of lemma 28. Let $\bar{t}_i = b - t_i$ and $\bar{s}_i = b - s_i$, where t_i and s_i are defined in Equation (13) and (14). Then,

$$\begin{aligned}
\Pr_{\sigma\sim\mu_{\Gamma_h^1}}\left[\mathrm{BW}(\sigma) = 0 \mid \sigma(r) = 0\right] &= 1 - \left[\left(f_{h-1,1}\right)^{t_h}\left(f_{h-1,2}\right)^{\bar{t}_h}\right],\\
\Pr_{\sigma\sim\mu_{\Gamma_h^1}}\left[\mathrm{BW}(\sigma) = 1 \mid \sigma(r) = 1\right] &= \left[1 - \left(f_{h-2,1}\right)^{t_{h-1}}\left(f_{h-2,2}\right)^{\bar{t}_{h-1}}\right]^{t_h}\\
&\quad \times \left[1 - \left(f_{h-2,1}\right)^{s_{h-1}}\left(f_{h-2,2}\right)^{\bar{s}_{h-1}}\right]^{\bar{t}_h}.
\end{aligned}$$

These recursions follow easily by noticing that $\mathrm{BW}(\sigma) = 0$ iff it is not true that $\mathrm{BW}(\sigma_i) = 0$ for all $i = 1, \ldots, b$, where σ_i is the restriction of σ to the tree subtended at the i-th children of the root. And also that $\mathrm{BW}(\sigma) = 1$ iff it is not true that $\mathrm{BW}(\sigma_i) = 0$ for all $i = 1, \ldots, b^2$, where σ_i is the restriction of σ to the tree subtended at the i-th grandchildren of the root. Now, from these recurrences and the bounds stated in Lemma 28, we deduce that

$$\begin{aligned}
\mu_{\Gamma_h^1}(\mathrm{BW}(\sigma) = 0, \sigma(r) = 0) &- \mu_{\Gamma_h^1}(\mathrm{BW}(\sigma) = 0)\mu_{\Gamma_h^1}(\sigma(r) = 0)\\
&= \Omega\left(1 - \frac{1.01}{(1+\omega)^b} - \frac{1.01^{1/b}}{1+\omega}\right).
\end{aligned}$$

Now, notice that

$$1 - \frac{1.01^{1/b}}{1+\omega} \geq \frac{1}{b}\left[\frac{(1+\delta)\ln b - 0.01}{(1+\omega)}\right].$$

Also, for $b \geq b_0(\delta)$ as defined in (11), we have that

$$\frac{1.01}{(1+\omega)^b} \leq \frac{1}{b}\left[\frac{0.01}{2(1+\delta)\ln b}\right].$$

Therefore, effectiveness, with rate roughly $\frac{(1+\delta)\ln b}{b}$, holds for all $b \geq b_0(\delta)$. The same result holds for $\mu_{\Gamma_h^2}(\cdot)$. □

Then, we are able to again bound $\bar{S}^{\Gamma}_{\ell,\mathrm{BW}}$ for $\Gamma = \Gamma_h^1$ or Γ_h^2, proving the following theorem, which completes the proof of Theorem 17.

Theorem 30. *Let $\delta > 0$, and let $\omega = (1+\delta)\ln b/b$. For all $b \geq b_0(\delta)$, it is the case that*

$$\Phi = O\big(n^{-d}\big), \quad \textit{where } d = \left(1 + \frac{\ln\left(\lambda/(1.01\omega b)^2\right)}{2\ln b}\right).$$

Proof. We take ℓ as $\ell_0 = \ell(\lambda, b)$, as in Lemma 28. Now, due to Lemma 27, we have that

$$\bar{S}^{\Gamma}_{\ell,\mathrm{BW}} = O\left(\frac{1}{n}\sum_{z^*\in \text{ level } \ell} \mathrm{E}_{\sigma\sim\mu_{\Gamma_h}}\left[\left(\frac{1.01\omega(1+\omega)}{\lambda}\right)^{\#\{i:\sigma(u_i)=0\}}\right]\right).$$

The following lemma bounds the expectation and the proof is presented in the next section.

Lemma 31. *For all $\delta > 0$, all $b > b_0(\delta)$, setting $\omega = (1+\delta)\ln b/b$, we have that for any leaf z^* in level ℓ and the corresponding path $\mathcal{P}$ from the root to z^*,*

$$\mathrm{E}_{\sigma\sim\mu_{\Gamma_h}}\left[\left(\frac{1.01\omega(1+\omega)}{\lambda}\right)^{\#\{i:\sigma(u_i)=0\}}\right] = O\left(\left(1.01\frac{\omega}{\lambda^{1/2}}\right)^h\right).$$

Since ℓ_0 is a constant independent of n and h, just as the argument at the end of Section 3.3.1, we can deduce that for $\delta > 0$, $b > b_0(\delta)$, setting $\omega = (1+\delta)\ln b/b$,

$$\bar{S}_{\mathrm{BW}} = O\left(\left[\frac{1.01\omega}{\lambda^{1/2}}\right]^H\right) = O\left(n^{-\left[1+\frac{\ln\left(\lambda/(1.01\omega b)^2\right)}{2\ln b}\right]}\right).$$

Now, from Proposition 29, the BW algorithm is effective for $\omega > (1+\delta)\ln b/b$, $b > b_0(\delta)$. Therefore, Theorem 26 applies, and the conclusion of Theorem 17 follows (trivially for $b < b_0(\delta)$). □

3.5 Calculating the Expectation

Here we provide all the technical lemmas that are needed in the previous sections for calculating the expectations. Recall that, in the broadcasting model, we fix a leaf z^* and take sample σ from distribution ν_h and we want to calculate the expectation

$$\mathrm{E}_{\sigma\sim\nu_h}\left[\left(\frac{1.01\omega(1+\omega)}{\lambda}\right)^{\#\{i:\sigma(u_i)=0\}}\right],$$

where u_i are vertices on the path $\mathcal{P}$ from z^* to the root. Here we let u_0 be the root and u_h be the leaf z^*, which reverses the order we used for u_i in previous sections. Observe that the random configurations for each u_i are essentially Markovian with respect to i due to the spatial Markov property of hard-core model. Therefore, we will first prove the following results for general Markov chain and then apply it to calculate the expectations for Lemma 22 and Lemma 31.

Let $\zeta_0, \zeta_1, \ldots$ be a Markov process with state space $\{0, 1\}$, such that $\zeta_0 = 0$ and with transition rates $p_{0\to 0} = p$, $p_{0\to 1} = q$, $p_{1\to 0} = 1$, $p_{1\to 1} = 0$. Let $N_h = \#\{1 \le i \le h : \zeta_i = 0\}$.

Lemma 32. *For the Markovian process $\{\zeta\}_{t\le h}$, we have*

1. *for any $a > 0$:*

$$\mathbf{E}\left[a^{N_h}\right] = O\left(\left(\frac{pa}{2}\left[1 + \sqrt{1 + 4q/\left(ap^2\right)}\right]\right)^h\right).$$

2. *Moreover, if $\bar{\zeta}_0, \bar{\zeta}_1, \ldots$ is a 'perturbed' version of the chain, in the sense that the transition rate $p_{0\to 0}$ (and therefore $p_{0\to 1}$) is now inhomogeneous but such that for some $\gamma > 0$, $\left|\frac{p^{(i)}_{0\to 0}}{p} - 1\right| \le \gamma$. Then, if $\bar{N}_h = \#\left\{1 \le i \le h : \bar{\zeta}_i = 0\right\}$, we have that for any $a > 0$:*

$$\mathbf{E}\left[a^{\bar{N}_h}\right] \le (1+\gamma)\,\mathbf{E}\left[\left(a\,(1+\gamma)\right)^{N_h}\right].$$

Proof. Let

$$\tau_1 = \min\{\ell : \zeta_\ell = 1\},$$

and for each $i \ge 1$, let

$$\tau_{i+1} = \min\left\{\ell - \tau_i : \ell \ge \tau_i \text{ and } \zeta_\ell = 1\right\}.$$

Therefore, τ_1 is the index of the first occurrence of state 1 and $\tau_2, \tau_3, \ldots$ are the distance between subsequent occurrences of the state 1 in the sequence. Also, let $\widetilde{\tau} = \min\{h - \ell : \ell \le h \text{ and } \tau_\ell = 1\}$, that is, the distance between h and the last occurrence of the state 1 in the sequence $\zeta_0, \zeta_1, \ldots \zeta_h$. It is easy to see that if $\widetilde{t} \ge 1$ and $0 \le k \le \lfloor h/2 \rfloor$,

$$\Pr\left[N_h = h - k,\, \tau_1 = t_1, \ldots, \tau_k = t_k,\, \widetilde{\tau} = \widetilde{t}\right] = p^{h-2k} q^k,$$

and if $\widetilde{t} = 0$ and $1 \le k \le \lfloor (h+1)/2 \rfloor$,

$$\Pr\left[N_h = h-k,\, \tau_1 = t_1, \ldots, \tau_k = t_k,\, \widetilde{\tau} = \widetilde{t}\right] = p^{h-2k+1} q^k.$$

Thus, adding up over all the possible choices of $t_1, \ldots, t_k, \widetilde{t}$, having in mind the restrictions $t_1 \ge 1$, $t_2 \ge 2$, $\ldots$, $t_k \ge 2$ and $t_1 + \cdots + t_k + \widetilde{t} = h$; we obtain

$$\Pr[N_h = h-k \text{ and } \zeta_h = 0] = \tbinom{h-k}{k} p^{h-2k} q^k \qquad \text{for } 0 \le k \le \lfloor h/2 \rfloor$$
$$\Pr[N_h = h-k \text{ and } \zeta_h = 1] = \tbinom{h-k}{k-1} p^{h-2k+1} q^k \qquad \text{for } 1 \le k \le \lfloor (h+1)/2 \rfloor,$$

therefore

$$\mathbf{E}\left[a^{N_h}\right] = \sum_{k=0}^{\lfloor h/2 \rfloor} \tbinom{h-k}{k} p^{h-2k} q^k a^{h-k} + \sum_{k=1}^{\lfloor (h+1)/2 \rfloor} \tbinom{h-k}{k-1} p^{h-2k+1} q^k a^{h-k} \tag{20}$$

Now, for the first term, we have that

$$\sum_{k=0}^{\lfloor h/2 \rfloor} \tbinom{h-k}{k} p^{h-2k} q^k a^{h-k} = (pa)^h \sum_{k=0}^{\lfloor h/2 \rfloor} \tbinom{h-k}{k} x^k,$$

where $x = \frac{q}{ap^2}$.

Let us define

$$\begin{aligned}
\phi(t) &:= \lim_{h\to\infty} \frac{\ln\left[\binom{h-th}{th} x^{th}\right]}{h} \\
&= (1-t)\,\mathrm{H}\left(\frac{t}{1-t}\right) + t \ln(x),
\end{aligned}$$

where H stands for natural entropy. $\phi(t)$ reaches its maximum at $t^* = \frac{1}{2}(1-\epsilon)$, where $\epsilon = 1/\sqrt{1+4x}$. It is easy to see the following holds.

$$\begin{aligned}
I(h) &= \int_0^{1/2} \mathrm{e}^{h\phi(t)} dt \\
&\approx \sum_{k=0}^{\lfloor h/2 \rfloor} \tbinom{h-k}{k} x^k.
\end{aligned}$$

Using the Laplace's approximation method, it is a standard procedure to approximate the asymptotic sum (integral $I(h)$) when $h \to \infty$ by the saddle point evaluation.

Noticing that the function $\phi''(t^*) = \frac{-4}{\epsilon(1-\epsilon)(1+\epsilon)}$, we have that

$$
\begin{aligned}
I(h) &= O\left(e^{h\phi(t^*)}\sqrt{\frac{2\pi}{h\phi''(t^*)}}\right) &(21)\\
&= O\left(e^{h\phi(t^*)}\right) &(22)\\
&= O\left(\left(\frac{1+\sqrt{1+4x}}{2}\right)^h\right). &(23)
\end{aligned}
$$

For the second term in (20), we have that

$$\sum_{k=1}^{\lfloor (h+1)/2\rfloor} \tbinom{h-k}{k-1} p^{h-2k+1} q^k a^{h-k} = p(pa)^h \sum_{k=1}^{\lfloor (h+1)/2\rfloor} \tbinom{h-k}{k-1} x^k,$$

Using a similar saddle point estimate, we have that

$$\sum_{k=1}^{\lfloor (h+1)/2\rfloor} \tbinom{h-k}{k-1} x^k = O\left(\left(\frac{1+\sqrt{1+4x}}{2}\right)^h\right). \tag{24}$$

Now, combining the asymptotics (23) and (24) into eq. (20) part 1 follows.

For part 2, using the same notation as above, we have that

$$
\begin{aligned}
&\mathbf{Pr}\left[\bar{N}_h = h-k,\, \tau_1 = t_1, \ldots, \tau_k = t_k,\, \widetilde{\tau} = \widetilde{t}\right]\\
&\qquad\leq \begin{cases} (1+\gamma)^{h-k} p^{h-2k} q^k & \text{if } \widetilde{t} \geq 1,\ 0 \leq k \leq \lfloor h/2 \rfloor, \\ (1+\gamma)^{h-k+1} p^{h-2k+1} q^k & \text{if } \widetilde{t} = 0,\ 1 \leq k \leq \lfloor (h+1)/2 \rfloor. \end{cases}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\Pr\left[\bar{N}_h = h-k \text{ and } \zeta_h = 0\right] \leq \tbinom{h-k}{k} (1+\gamma)^{h-k} p^{h-2k} q^k && \text{for } 0 \leq k \leq \lfloor h/2 \rfloor\\
&\Pr\left[\bar{N}_h = h-k \text{ and } \zeta_h = 1\right] \leq \tbinom{h-k}{k-1} (1+\gamma)^{h-k+1} p^{h-2k+1} q^k && \text{for } 1 \leq k \leq \lfloor (h+1)/2 \rfloor
\end{aligned}
$$

This leads to

$$
\begin{aligned}
\mathbf{E}\left[a^{\bar{N}_h}\right] &\leq \sum_{k=0}^{\lfloor h/2 \rfloor} \tbinom{h-k}{k} (1+\gamma)^{h-k+1} p^{h-2k} q^k a^{h-k}\\
&\quad + \sum_{k=1}^{\lfloor (h+1)/2 \rfloor} \tbinom{h-k}{k-1} (1+\gamma)^{h-k+1} p^{h-2k+1} q^k a^{h-k}\\
&= (1+\gamma)\, \mathbf{E}\left[(a(1+\gamma))^{N_h}\right].
\end{aligned}
$$

$\square$

Proof of Lemma 22. Notice that when $\sigma \sim \nu_h$, $\zeta_i := \sigma(u_i)$ is a Markov chain with state space $\{0,1\}$ and transition probabilities $p_{0\to 0} = 1/(1+\omega)$ and $p_{1\to 0} = 1$. To estimate $\mathrm{E}_{\sigma\sim\nu_h}\left[\theta^{\#\{i:\zeta_i=0\}}\right]$ for any $\theta > 0$, we apply the technical result in Part 1 of Lemma 32. In fact, recalling the random variable N_h defined in Lemma 32, we have that

$$\begin{aligned}\mathrm{E}_{\sigma\sim\nu_h}\left[\theta^{\#\{i:\zeta_i=0\}}\right] &= O\left(\mathrm{E}_{\sigma\sim\nu_h}\left[\theta^{\#\{i:\zeta_i=0\}} : \zeta_0 = 0\right]\right)\\ &= O\left(\mathbf{E}\left[\theta^{N_h}\right]\right)\end{aligned}$$

Therefore, plugging in the asymptotic from the lemma for $\theta = \frac{1.01\omega(1+\omega)}{\lambda}$, we get

$$\mathbf{E}\left[\theta^{N_h}\right] = O\left(\left(\frac{1.01\omega}{2\lambda}\left[1+\sqrt{1+\frac{4\lambda}{1.01}}\right]\right)^h\right) \le O\left(\left(\frac{1.01\omega}{\lambda^{1/2}}\right)^h\right).$$

For the last inequality we used the fact that $1+\sqrt{1+4\lambda/1.01} \le 2\lambda^{1/2}$, which holds for $\lambda > (101)^2$, and in particular, when $\omega = (1+\delta)\ln(b)/b$ and $b > b_0(\delta)$ where $b_0(\delta)$ was defined in (11). □

Proof of Lemma 31. The proof goes along the lines of lemma 22. For $\sigma \sim \mu_{h,\Gamma_h}$, $\zeta_i := \sigma(u_i)$ is a inhomogeneous Markov chain with state space $\{0,1\}$ and transition probabilities, for $i \le h - \ell(\lambda, b)$, such that $|\frac{p_{0\to 0}}{1/(1+\omega)} - 1| \le \left(\exp\left(\frac{1.01\omega^2 b}{\lambda}\right) - 1\right)$ (from eq. 17) and $p_{1\to 0} = 1$. Now, to estimate $\mathrm{E}_{\sigma\sim\nu_h}\left[\theta^{\#\{i:\zeta_i=0\}}\right]$ for $\theta > 0$, we apply Part 2 of Lemma 32. This time, recalling the random variables N_h and $\bar{N}_h$ defined in such lemma, we have that

$$\begin{aligned}\mathrm{E}_{\sigma\sim\mu_{h,\Gamma_h}}\left[\theta^{\#\{i:\zeta_i=0\}}\right] &= O\left(\frac{1}{1+\omega}\mathrm{E}_{\sigma\sim\mu_{h,\Gamma_h}}\left[\theta^{\#\{i:\zeta_i=0\}} : \zeta_0 = 0\right]\right)\\ &= O\left(\frac{1}{1+\omega}\mathrm{E}_{\sigma\sim\mu_{h,\Gamma_h}}\left[\theta^{\#\{i\le h-\ell:\zeta_i=0\}} : \zeta_0 = 0\right]\right)\\ &= O\left(\mathbf{E}\left[\theta^{\bar{N}_{h-\ell}}\right]\right)\\ &= O\left(\mathbf{E}\left[\theta^{\bar{N}_h}\right]\right)\\ &= O\left(\mathbf{E}\left[\left(\exp\left(\frac{1.01\omega^2 b}{\lambda}\right)\theta\right)^{N_h}\right]\right)\end{aligned}$$

Now, plugging in the asymptotics for $\theta = \frac{1.01\omega(1+\omega)}{\lambda}$, we get

$$\mathbf{E}\left[\left(\exp\left(\frac{1.01\left(\omega b\right)^2}{\lambda}\right)\theta\right)^{N_h}\right]$$
$$= O\left(\frac{1.01\omega}{2\lambda}\exp\left(\frac{1.01\left(\omega b\right)^2}{\lambda}\right)\left[1+\sqrt{1+\frac{4\lambda}{1.01}\exp\left(-\frac{1.01\left(\omega b\right)^2}{\lambda}\right)}\right]\right)^h$$

Finally we use the inequality

$$1+\sqrt{1+\frac{4\lambda}{1.01}\exp\left(-\frac{1.01\omega^2 b}{\lambda}\right)} \leq 2\lambda^{1/2}\exp\left(-\frac{1.01\omega^2 b}{\lambda}\right),$$

which holds whenever $\omega = (1+\delta)\ln b/b$ and $b > b_0(\delta)$.

$\square$

CHAPTER IV

UPPER BOUNDS VIA COUPLING

Recall that in the Ising model on regular trees, the phase transition for the mixing time of the Glauber dynamics happens exactly at the reconstruction threshold. In last chapter, we show the lower bounds of mixing times of the Glauber dynamics for colorings and hard-core models on trees when the reconstruction problems are solvable. In this chapter, we will prove that the upper bounds for mixing times or relaxation times of the Glauber dynamics are very close to the lower bounds in both reconstruction and non-reconstruction regions. Then we are able to establish similar pictures of phase transitions for the convergence speed of the Glauber dynamics as in the Ising model on trees. The following two theorems summarize the main results of this chapter, which are the restatements of the upper bounds in Theorem 2 and Theorem 3.

Theorem 33. *For the Glauber dynamics on the hard-core model with activity $\lambda = \omega(1+\omega)^b$ on the $(b+1)$-regular tree T of n vertices and height $H = \lfloor \log_b n \rfloor$, the following hold:*

1. **For all $\omega \le \ln b/b$, for every boundary condition,**

$$T_{\text{relax}} \le O(n^{1+o_b(1)}).$$

2. **For all $\delta > 0$ and $\omega = (1+\delta)\ln b/b$, for every boundary condition,**

$$T_{\text{relax}} \le O(n^{1+\delta+o_b(1)}).$$

Theorem 34. *For the Glauber dynamics of k-colorings on the $(b+1)$-regular tree T of n vertices and height $H = \lfloor \log_b n \rfloor$ satisfies the following:*

1. **For all** $\epsilon > 0$ **and all** $k = \frac{(1+\epsilon)b}{\ln b}$ **:**

$$\begin{aligned} T_{\text{mix}} &\leq O(n^{1+o_b(1)} \ln n), \\ T_{\text{relax}} &\leq O(n^{1+o_b(1)}). \end{aligned}$$

2. **For all** $\epsilon > 0$ **and all** $k = \frac{b}{(1+\epsilon)\ln b}$ **:**

$$\begin{aligned} T_{\text{mix}} &\leq O(n^{1+\epsilon+o_b(1)} \ln n), \\ T_{\text{relax}} &\leq O(n^{1+\epsilon+o_b(1)}). \end{aligned}$$

The constants in the $\Omega()$ *and* $O()$ *are universal constants.*

Before we present the proofs of these two theorems, let us briefly review the coupling of two Markov chains. Coupling is a power technique for showing the upper bound of the mixing time is. Given two copies, (X_t) and (Y_t), of the Markov chain at time $t > 0$, recall that a (one-step) coupling of (X_t) and (Y_t), is a joint distribution whose left and right marginals are identical to the (one-step) evolution of (X_t) and (Y_t), respectively. The Coupling Lemma [3] (c.f., Theorem 5.2 in [41]) guarantees that if, there is a coupling and time $t > 0$, so that for every pair (X_0, Y_0) of initial states, $\Pr[X_t \neq Y_t] \leq 1/2\mathrm{e}$ under the coupling, then $T_{\text{mix}} \leq t$. It is our main tool here to prove the bounds.

This chapter is organized as follows. First we will show the upper bound of the relaxation time of the Glauber dynamics for the hard-core model under any boundary condition on trees by using a monotone coupling and the censoring inequality for monotone systems. Then, we are going to show an upper bound for the mixing time and the relaxation time for colorings on trees with the free boundary condition. This combining with the lower bound results in Chapter 3, completes the full picture as described in Theorem 2 and Theorem 3 and establishes the phase transition for the Glauber dynamics around the reconstruction thresholds of both models respectively. Meanwhile, we will improve a connection between the log-Sobolev constant and the spectral gap for the Glauber dynamics of the colorings on trees which makes it possible for us to bound the mixing time from the relaxation time in an accurate way.

4.1 Hard-core Model

Before we show the main idea for our upper bound proofs, we first introduce some notation we use in this section. For a b dimensional vector $\rho = (\rho_1, \ldots, \rho_b)$ where $0 \leq \rho_i \leq 1$ for every $1 \leq i \leq b$, let τ_ρ be the relaxation time of the following Glauber dynamics of the hard-core model on the star graph $G^\star$ with root r and $b = \Delta - 1$ leaves $\{w_1, \ldots, w_b\}$. The dynamics on the star graph $G^\star$ is defined as follows. Given an independent set X_t,

1. Choose a vertex v uniformly at random from $\{r, w_1, \ldots, w_b\}$.

2. If $v = r$, then set

$$X' = \begin{cases} X_t \cup \{v\} & \text{with probability } \lambda/(1+\lambda) \\ X_t \setminus \{v\} & \text{with probability } 1/(1+\lambda). \end{cases}$$

3. If $v = w_i$ is a leaf of $G^\star$, then set

$$X' = \begin{cases} X_t \cup \{w_i\} & \text{with probability } \rho_i \\ X_t \setminus \{w_i\} & \text{with probability } 1-\rho_i. \end{cases}$$

4. If X' is an independent set, then set $X_{t+1} = X'$, otherwise set $X_{t+1} = X_t$.

Let $\tau^\star := \max_\rho\{\tau_\rho\}$ be defined as the worst case relaxation time over all possible choices of ρ. Using the block dynamics approach of Martinelli [43], as used in [8, Section 2.3] (see also [42] for similar results), it is not hard to show that the relaxation time of the above Glauber dynamics is exactly the same as that of the natural block dynamics which updates the configurations of a whole subtree of the root in one step, and hence the following lemma holds.

Lemma 35. *For the complete tree of height H with any boundary condition on the leaves, the relaxation time $T_{\rm relax}$ of the Glauber dynamics of the hard-core model satisfies:*

$$T_{\rm relax} \le (\tau^\star)^H.$$

We omit the proof of the above lemma since it is essentially identical to that in [8, Section 2.3].

Note that, the relaxation time on the complete tree is quite sensitive to the boundary conditions. For example, as mentioned in the Introduction, Martinelli *et al.* [47] show that when the boundary condition is all even (or similarly for all odd), i.e., all of the leaves are occupied when the height is even (respectively, odd) and all of the leaves are unoccupied when the height is odd (even), then the mixing time is $O(n \ln n)$ for all λ. In this paper we are considering all boundary conditions, and in our lower bound, we show there are boundary conditions that slow down the Glauber dynamics. The lower bound on the relaxation time for the Glauber dynamics under those boundary conditions which we show suffer the slow-down roughly matches up with the upper bound we prove here. The following lemma establishes an upper bound for $\tau^\star$.

Lemma 36. *For the Glauber dynamics of the hard-core model on $G^\star$, the worst relaxation time over all the boundary conditions ρ satisfies*

$$\tau^\star \le 100(\lambda+1)(b+1)\ln^2(b+1).$$

Therefore, by Lemmas 35 and 36, for any boundary condition on the leaves, the Glauber dynamics of the hard-core model on the complete tree of height H satisfies,

$$T_{\rm relax} \le (\tau^\star)^H \le \big(100(\lambda+1)(b+1)\ln^2(b+1)\big)^{\log_b n} \le n^d,$$

where

$$d = 1 + \frac{\ln\big(200(\lambda+1)\ln^2(b+1)\big)}{\ln b}.$$

Now, if $\omega \le \frac{\ln b}{b}$, we have that, for some constant $c_0 > 0$,

$$d \le 1 + \frac{c_0 \ln\ln b}{\ln b}.$$

On the other hand, for $\delta > 0$ and $\omega = (1+\delta)\ln b/b$, we instead get, that for some constant $c_1 > 0$,

$$d \le 1 + \delta + \frac{c_1 \ln\ln b}{\ln b}.$$

This proves Theorem 33.

4.1.1 Upper Bound for the Relaxation Time on the Star Graph: Proof of Lemma 36

We will analyze the following coupling $\mathcal{L}$ of two copies (X_t), (Y_t) of the Glauber Dynamics of the hard-core model on $G^\star$. The coupling $\mathcal{L}$ chooses the same random vertex v to update in both chains X_t and Y_t. If $v = r$, the root of $G^\star$, and there is not an occupied leaf in any of the two copies, then r is coupled to be occupied with probability $\lambda/(1+\lambda)$ and unoccupied with probability $1/(1+\lambda)$ in both X_{t+1} and Y_{t+1}. If $v = w_i$, a leaf in $G^\star$, and the root is unoccupied in both X_t and Y_t, then v is also coupled to be occupied or unoccupied in both X_{t+1} and Y_{t+1} with the corresponding probability. If in either X_t or Y_t the neighbors of v contains an occupied vertex, then each copy is updated independently with the corresponding probability.

Given a pair of configurations $\eta, \eta' : G^\star \to \{0,1\}$, we say that $\eta \preceq \eta'$ if $\eta(r) \le \eta'(r)$ and for every $i = 1, \ldots, b$, $\eta(w_i) \ge \eta'(w_i)$. Let $\eta_{\max}$ and $\eta_{\min}$ be the unique maximal and minimal elements in this partial order, respectively. An important property of the coupling $\mathcal{L}$ of the hard-core model in the star is *monotonicity.* Namely, if (X_t, Y_t) are such that $X_t \preceq Y_t$ then after applying one step of the coupled dynamics we have that $X_{t+1} \preceq Y_{t+1}$. More generally for bipartite graphs G, the hard-core model is a *monotone* system (see, e.g., Chapter 22 in [41]) in the sense that, if (X_t) and (Y_t) are two copies of the Glauber dynamics on the hard-core model on G and $x_0 \preceq y_0$, then there exists a one-step coupling $\mathcal{C}$ of (X_t) and (Y_t) such that for all $t \geq 0$,

$$\Pr_{\mathcal{C}}\left[X_t \preceq Y_t \mid X_0 = x_0, Y_0 = y_0\right] = 1.$$

In this case, we say (X_t) is stochastically dominated by (Y_t) and denote it as $X_t \preceq^d Y_t$.

Using monotonicity of the coupling $\mathcal{L}$, we have that:

$$\Pr_{\mathcal{L}}\left[X_t \neq Y_t \mid X_0, Y_0\right] \leq \Pr_{\mathcal{L}}\left[X_t \neq Y_t \mid X_0 = \eta_{\max}, Y_0 = \eta_{\min}\right];$$

that is, the worst case initial configurations, for the coupling probability, are the maximal and minimal configurations.

Therefore, using (1) and the coupling lemma, to prove Lemma 36 it is enough to show

$$\Pr_{\mathcal{L}}\left[X_T \neq Y_T \mid X_0 = \eta_{\max}, Y_0 = \eta_{\min}\right] \leq 1/2e \text{ for } T = 100(1{+}\lambda)(b{+}1)\ln^2(b{+}1). \quad (25)$$

We will use the censoring technique of Peres and Winkler (see Theorem 1.1 in [54]) which intuitively says that in a monotone system, extra moves of the Glauber dynamics will not hurt the total variation distance from current distribution to the stationary distribution. In our case, we will do the coupling analysis in the following way. First of all, instead of bounding the coupling probability for any T-step sequence of vertices in the Glauber dynamics, we want to bound the coupling probability for some "good" sequences which occur with a very high probability. Then, we are able to use the censoring technique to reduce the calculations of coupling probabilities for those "good" sequences to the calculations on just a few fixed sequences of vertices which are the "censored" subsequences of those "good" sequences of vertices. The calculations of the coupling probabilities on those fixed sequences are very easy to conduct.

We now specify the details formally. Throughout this section we assume that the initial states are $X_0 = \eta_{\max}$ and $Y_0 = \eta_{\min}$. Given a sequence $u = (u_1, u_2, \ldots)$ of vertices of $G^\star$, let $X^u = (X^u_t)_{t\geq 0}$ be the Glauber dynamics such that for every $t \geq 1$, the chain is updating the vertices according to the sequence u, i.e., at time t, X_t^u chooses vertex u_t to update. Let $\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \ldots)$ be a sequence of i.i.d uniform random vertices of $G^\star$. Notice that X, the (original) Glauber dynamics satisfies

$$X \stackrel{d}{=} X^{\mathbf{U}}. \tag{26}$$

Given a 0/1 sequence $\gamma = (\gamma_t)_{t\geq 1}$, and a sequence of vertices $u = (u_1, u_2, \ldots)$, we define $X^{u,\gamma} = (X_t^{u,\gamma})_{t\geq 0}$ to be the *censored* version of X^u, which is restricted, in addition, to change the configuration at vertex u_t at time t only if $\gamma_t = 1$ (if $\gamma_t = 0$ then $X_t = X_{t-1}$).

We will "censor" u in the following way to ease the calculation of the coupling probability. To couple both copies (X_t) with (Y_t) using $\mathcal{L}$ it is enough to get the root to agree in both copies, and then get the leaves to agree. Given a sequence u, we call a "scan" a subsequence $u_{i_0}, u_{i_1}, \ldots, u_{i_b}$ where the root is visited and then all the leaves, that is $u_{i_0} = r$ and $\{u_{i_1}, \ldots, u_{i_b}\} = \{w_1, \ldots, w_b\}$. We define γ^u as a 0/1 sequence maximizing the number of non-overlapping scans in $u_1, \ldots, u_T$ (if there is more than one such sequence just choose an arbitrary one). We say that γ^u is a k-scanning of u if the sequence $(u_t)_{t\leq T: \gamma_t^u = 1}$ consists of at least k scans. Let $\mathcal{S}_k = \{u : \gamma_u \text{ is a } k\text{-scanning}\}$, the set of sequences that contain at least k scans before time T.

Notice that under the coupling $\mathcal{L}$, when the root is unoccupied, to get the leaves to agree it is enough to just update(choose) them. Thus, the coupling probability after one scan is the probability of coupling the root (when updated), which is at least $1/(1+\lambda)$. Therefore,

$$\Pr_{\mathcal{L}}\left[X_T^{u,\gamma_u} \neq Y_T^{u,\gamma_u}\right] \leq \left(1 - \frac{1}{\lambda+1}\right)^k \text{ for all } u \in \mathcal{S}_k. \tag{27}$$

Now to prove (25), let $k = 3(1+\lambda)\ln(b+1)$. We have

$$\begin{aligned} &\mathrm{Pr}_{\mathcal{L}}\left[X_T \neq Y_T\right] \\ &= \sum_u \mathrm{Pr}_{\mathcal{L}}\left[X_T^u \neq Y_T^u\right]\mathrm{Pr}_{\mathbf{U}}\left[u\right] \\ &\leq \mathrm{Pr}_{\mathbf{U}}\left[u \notin \mathcal{S}_k\right] + \sum_{u: u \in \mathcal{S}_k} \mathrm{Pr}_{\mathcal{L}}\left[X_T^u \neq Y_T^u\right]\mathrm{Pr}_{\mathbf{U}}\left[u\right]. \end{aligned} \tag{28}$$

First we bound $\mathrm{Pr}_{\mathbf{U}}\left[u \notin \mathcal{S}_k\right]$. Let τ_u be the first time u contains k consecutive scans (τ_u is a positive random variable which can be equal to ∞). By the coupon collector, $\mathrm{E}_{\mathbf{U}}\left[\tau_u\right] = k(b+1)(1+\ln b)$. Using Markov's inequality we have,

$$\mathrm{Pr}_{\mathbf{U}}\left[u \notin \mathcal{S}_k\right] = \mathrm{Pr}_{\mathbf{U}}\left[\tau_u > 20k(b+1)(1+\ln b)\right] \leq 1/20. \tag{29}$$

Now to bound $\mathrm{Pr}_{\mathcal{L}}\left[X_T^u \neq Y_T^u\right]$ we use the following Censoring Lemma of Peres and Winkler.

Lemma 37 (Theorem 1.1 in [54])**.** *For any u, γ, and t,*

$$X_t^u \preceq^d X_t^{u,\gamma_u} \text{ and } Y_t^{u,\gamma_u} \preceq^d Y_t^u.$$

Also,

$$\left\|\mu_{X_t^u} - \pi\right\|_{TV} \leq \left\|\mu_{X_t^{u,\gamma_u}} - \pi\right\|_{TV} \text{ and, } \left\|\mu_{Y_t^u} - \pi\right\|_{TV} \leq \left\|\mu_{Y_t^{u,\gamma_u}} - \pi\right\|_{TV}. \tag{30}$$

Notice that the above Censoring lemma allows us to bound the variation distance, but only starting at the extremal initial configurations. As the extremal configurations are not necessarily the worst case for variation distance, we can not use the censoring Lemma alone. But, as discussed before, the monotonicity of the local coupling allows us to assume extremal initial configurations.

To bound the coupling probability in terms of the coupling probability of the censored chain, we use as an intermediate proxy the variation distance the following manner.

$$
\begin{aligned}
&\Pr_{\mathcal{L}}\left[X_T^u \neq Y_T^u\right] \\
&\leq \sum_{v\in V} \Pr_{\mathcal{L}}\left[X_T^u(v) \neq Y_T^u(v)\right] \\
&= \sum_{v\in V} \left\|\mu_{X_T^u(v)} - \mu_{Y_T^u(v)}\right\|_{\mathrm{TV}} \\
&\leq (b+1)\left\|\mu_{X_T^u} - \mu_{Y_T^u}\right\|_{\mathrm{TV}} \\
&\leq (b+1)\left(\left\|\mu_{X_T^u} - \pi\right\|_{\mathrm{TV}} + \left\|\mu_{Y_T^u} - \pi\right\|_{\mathrm{TV}}\right) \\
&\leq (b+1)\left(\left\|\mu_{X_T^{u,\gamma_u}} - \pi\right\|_{\mathrm{TV}} + \left\|\mu_{Y_T^{u,\gamma_u}} - \pi\right\|_{\mathrm{TV}}\right) && \text{by the Censoring lemma (30)} \\
&\leq (b+1)(1-1/(\lambda+1))^{3(\lambda+1)\ln(b+1)} && \text{by Proposition 4.7 in [41] and (27)} \\
&\leq 1/(b+1)^2,
\end{aligned}
$$

where the first equality is by the fact that for any monotone coupling of monotone two spin system, when projecting on a specific vertex v, there is actually only one way to couple and hence the probability equals to the total variation distance. The penultimate inequality follows by applying Proposition 4.7 in [41] to the total variation distances $\left\|\mu_{X_T^{u,\gamma_u}} - \pi\right\|_{\mathrm{TV}}$ and $\left\|\mu_{Y_T^{u,\gamma_u}} - \pi\right\|_{\mathrm{TV}}$, then using (27) to bound the probabilities of couplings. Proposition 4.7 in [41] is a well-known inequality that upper bounds the total variation distance between two distributions ν and μ by the probability of $X \neq Y$ for any coupling (X, Y) of μ and ν.

Combining with (28) and (29), we have that:

$$\Pr_{\mathcal{L}}\left[X_T \neq Y_T \mid X_0 = \eta_{\max}, Y_0 = \eta_{\min}\right] \leq \frac{1}{20} + \frac{1}{(b+1)^2},$$

which implies (25) and thus Lemma 36 follows.

4.2 Colorings

Let $\tau^\star$ here be the relaxation time of the Glauber dynamics for k-colorings on the star graph $G^\star$ with $b = \Delta - 1$ vertices of degree 1. Again, by the block dynamics argument, we are able to bound the relaxation time T_{relax} of the dynamics on the whole tree using the bounds for $\tau^\star$.

Theorem 38. *The relaxation time T_{relax} of the Glauber dynamics for k-colorings of the Δ-regular tree with height H satisfies*

$$T_{\mathrm{relax}} \leq (\tau^\star)^H.$$

Therefore, proving the upper bounds in Theorem 34 reduces to the problem of getting tight upper bounds of the relaxation time $\tau^\star$ of the Glauber dynamics on $G^\star$. In [42], the authors used a canonical path argument to bound $\tau^\star = O(b^{3+\epsilon}k)$ for any $\epsilon > 0$. Instead, here we use two different coupling arguments to show the following two theorems for $\tau^\star$.

Theorem 39. *For any $\epsilon > 0$, there exists $b_0 > 0$ such that, for any $b > b_0$, the mixing and relaxation times of the Glauber dynamics on $G^\star$ using $k = \frac{b}{(1+\epsilon)\ln b}$ colors are $O(b^{1+\epsilon}\ln^2 b)$. When $\epsilon = 0$, the mixing and relaxation times are $O(b\ln^4 b)$.*

Theorem 40. *For any $\epsilon > 0$, there exists $b_0 > 0$ such that, for any $b > b_0$, the mixing and relaxation times of the Glauber dynamics on $G^\star$ using $k = \frac{(1+\epsilon)b}{\ln b}$ colors are $O(b\ln b)$.*

It can be shown that the relaxation time is actually $O(b)$ when $k > b/\ln b$, from our analysis. However, unless we can also eliminate the constant factors and thereby show a very sharp bound of at most b, the extra $\ln b$ factor makes little difference to the relaxation time of the dynamics on the whole tree.

The most difficult (and also interesting) case turns out to be when $k < b/\ln b$. We will prove Theorem 39 in Section 4.2.1 and Theorem 40 in Section 4.2.2. Having Theorems 39 and 40 in hand, we can then apply Theorem 38 to get the upper bounds on the relaxation time as stated in Theorem 34. We get

$$T_{\text{relax}} = \begin{cases} O(b\ln b)^H = O\left(n^{1+(\ln\ln b+O(1))/\ln b}\right), & \text{if } k=(1+\epsilon)b/\ln b; \\ O(b\ln^4 b)^H = O\left(n^{1+(4\ln\ln b+O(1))/\ln b}\right), & \text{if } k=b/\ln b; \\ O(b^{1+\epsilon}\ln^2 b)^H = O\left(n^{1+\epsilon+(2\ln\ln b+O(1))/\ln b}\right), & \text{if } k=b/((1+\epsilon)\ln b). \end{cases}$$

To then get the desired upper bounds on the mixing time of the whole tree we need a slightly more advanced tool, the logarithmic Sobolev constant of the Markov chain. By adapting Theorem 5.7 in Martinelli, Sinclair and Weitz [46] to our setting of colorings, we establish and improve (in Section 4.2.4) the following relationship between the inverse of the log-Sobolev constant c_{sob}^{-1} and the relaxation time T_{relax} of the Glauber dynamics on trees.

Theorem 41.

$$c_{sob}^{-1} \leq T_{\text{relax}} \cdot 2b\ln(k).$$

Since the inverse of the log-Sobolev constant gives a relatively tight upper bound on the mixing time (see Inequality (3) in Chapter 2), using Theorem 41 we are able to complete the proofs of the upper bounds in Theorem 34.

4.2.1 Upper Bound on Mixing Time for $k < b/\ln b$

In this section, we upper bound the mixing time of the Glauber dynamics on the star graph $G^\star = (V, E)$ when $k = b/((1+\epsilon)\ln b)$ for any $\epsilon \geq 0$. To be more precise, let $V = \{r, \ell_1, ..., \ell_b\}$, where r refers to the root and $\ell_1, ..., \ell_b$ are the b leaves and $E = \{(r,\ell_1), ..., (r,\ell_b)\}$.

We use the maximal one-step coupling, originally studied for colorings by Jerrum [35] to upper bound the mixing time of the Glauber dynamics on general graphs. For a coloring $X \in \Omega$, let $A_X(v)$ denote the set of available colors of v in the coloring X, i.e., $A_\sigma(v) = \{c \in \mathcal{C} : \forall u \in N(v), \sigma(u) \neq c\}$. The coupling (X_t, Y_t) of the two chains is done by choosing the same random vertex v_t for recoloring at step t and maximizing the probability of the two chains choosing the same update for the color of v_t. Thus, for each color $c \in A_{X_t}(v) \cap A_{Y_t}(v)$, with probability $1/\max\{|A_{X_t}(v)|, |A_{Y_t}(v)|\}$ we set $X_{t+1}(v) = Y_{t+1}(v) = c$. With the remaining probability, the color choices for $X_{t+1}(v)$ and $Y_{t+1}(v)$ are coupled arbitrarily.

We prove the theorem by analyzing the coupling in rounds, where each round consists of $T := 20b\ln b$ steps. Our main result is the following lemma which says that in each round we have a good probability of coalescing (i.e., achieving $X_t = Y_t$).

Lemma 42. *For all $\epsilon \geq 0$, there exists $b_0(\epsilon)$ such that for all $b > b_0(\epsilon)$ if $k = b/((1+\epsilon)\ln b)$ and $T = 20b\ln b$ for all $(x_0, y_0) \in \Omega \times \Omega$, the following holds:*

$$\Pr\left[X_T = Y_T \mid X_0 = x_0, Y_0 = y_0\right] \geq \begin{cases} (20(1+\epsilon)b^{\epsilon}\ln b)^{-1}, & \text{if } \epsilon > 0; \\ (20\ln^3 b)^{-1}, & \text{if } \epsilon = 0. \end{cases}$$

It is then straightforward to prove Theorem 39.

Proof of Theorem 39. For $\epsilon > 0$, let $p_T := (20(1+\epsilon)b^{\epsilon}\ln b)^{-1}$; and for $\epsilon = 0$ let $p_T := (20\ln^3 b)^{-1}$. By repeatedly applying Lemma 42 we have, for all (x_0, y_0),

$$\Pr\left[X_{2iT} \neq Y_{2iT} \mid X_0 = x_0, Y_0 = y_0\right] \leq (1 - p_T)^{2i} \leq 1/2\mathrm{e}$$

for $i = 1/p_T$. Therefore, by applying the Coupling Lemma, the mixing time is $O((1+\epsilon)b^{1+\epsilon}\ln^2 b)$ for $\epsilon > 0$ and $O(b\ln^4 b)$ for $\epsilon = 0$. □

Before formally proving Lemma 42 we give a high-level overview of its proof. We will analyze the maximal one-step coupling on the star graph $G^{\star}$. We say a vertex v "disagrees" at time t if $X_t(v) \neq Y_t(v)$, otherwise we say the vertex v "agrees". We denote the set of disagreeing vertices at time t of our coupled chains by

$$D_t = \{v \in V : X_t(v) \neq Y_t(v)\},$$

and we use $D_t^L = D_t \setminus \{r\}$ to represent the set of disagreeing leaves. When we use the term "with high probability" in this section, it means that the probability goes to 1 as b goes to infinity.

If the coupling selects a leaf ℓ to recolor at time t, then the probability that ℓ disagrees in X_t and Y_t is at most $1/(k-1)$, and with probability at least $(k-2)/(k-1)$, the leaf will use the same color that is chosen uniformly at random from $\mathcal{C} \setminus \{X_t(r), Y_t(r)\}$. We also know that if we simply assign a random color from $\mathcal{C}$ to each leaf, with probability at least $\Omega(1/(b^\epsilon \ln b))$, there is a color in $\mathcal{C}$ that is unused in any leaf. This last point hints at the success probability in the statement of Lemma 42.

We analyze the T-step epoch in three stages. The warm-up round is of length $T_{\mathrm{w}} := 8(b+1)\ln b$ steps. We will show in Lemmas 45 and 46 that with good probability, after the warm up, all of the leaf disagreements will be of the same form in the sense that they will have the same pair of colors.

The next stage is of a random length T_1, which is defined as the first time (after T_{w}) where we are recoloring the root and the root has a common available color in (X_t) and (Y_t). We prove in Lemma 47, that with probability $\Omega(1/b^\epsilon \ln b)$, $T_1 < 4(b+1)\ln b$. We then have probability at least $1/2$ of the root agreeing after it is updated, and then after at most $T_2 := 4(b+1)\ln b$ further steps we are likely to coalesce since we just need to recolor each leaf at least once before the root changes back to a disagreement.

We begin our proof of Lemma 42 with a basic observation about the maximal one-step coupling.

Observation 43. *Let $\mathcal{C}(D_t^L) := \bigcup_{\ell \in D_t^L} \{X_t(\ell), Y_t(\ell)\}$ denote the set of colors that appear in the disagreeing leaves at time t. Then, $A_{X_t}(r) \oplus A_{Y_t}(r) \subseteq \mathcal{C}(D_t^L)$.*

This is simply because those colors that appear on the leaves with agreements are both unavailable in X_t and Y_t for the root. We now analyze the first stage of the T-step epoch.

Proposition 44. *The probability that in $T_0 = 4(b+1)\ln b$ steps, the coupling (X_t, Y_t) (or the Glauber dynamics (X_t)) will recolor the root at most $20 \ln b$ times and recolor every leaf at least once is at least $1 - 2b^{-3}$.*

Proof. Using the union bound the probability that there is a leaf which is not recolored in T_0 steps is at most

$$b\left(1-\frac{1}{b+1}\right)^{4(b+1)\ln b} \leq b^{-3}.$$

Now, let N be the number of times the root is recolored in T_0 steps. The expectation $\mathrm{E}[N]$ is $4\ln b$. Then, by the Chernoff bound (see, e.g., Theorem 4.5 Part 2 in [48])

$$\Pr[N \geq 20\ln b] \leq \Pr[N \geq (1+4)\mathrm{E}[N]] \leq b^{-3}.$$

Therefore the lemma holds by the union bound. □

Then we will prove that in $T_{\mathrm{w}} = 2T_0$ steps, with high probability all of the leaf disagreements are of the same type when $\epsilon > 0$.

Lemma 45. *For any $\epsilon > 0$ and $k > (1+\epsilon)b/\ln b$, for any pair of initial states (x_0, y_0),*

$$\Pr\left[\forall \ell \in D^L_{T_w}, X_{T_w}(\ell) = Y_{T_w}(r) \wedge Y_{T_w}(\ell) = X_{T_w}(r) \mid x_0, y_0\right] \geq 1 - O\left(\frac{1}{b^\epsilon}\right).$$

Proof. The idea is that if we just look at one chain, say (X_t), then after T_0 steps, with high probability the root is frozen. Moreover, the root is likely to continue to be frozen for the remainder of the T_{w} steps since we recolor the root at most $O(\ln b)$ times. In the worst case the root is frozen to a disagreement, say $X_t(r) = 2$ and $Y_t(r) = 1$. Then after recoloring a leaf ℓ at time t' where $t < t' < T_{\mathrm{w}}$, the only possible disagreement is $X_{t'}(\ell) = 1, Y_{t'}(\ell) = 2$. Hence, it suffices to recolor each leaf at least once.

Let $\mathcal{E}$ be the event that in the first T_0 steps, every leaf is recolored at least once and in another $4(b+1)\ln b$ steps, every leaf is recolored again at least once and the root is recolored at most $20\ln b$ times. We are first going to bound that for $t > T_0$,

$$\Pr[|A_{X_t}(r)| > 1 \mid \mathcal{E}] \leq \frac{1}{(1+\epsilon)b^\epsilon \ln b} := p_0, \tag{31}$$

and the same thing happens for Y_t.

Let G_W be the graph with b isolated vertices $\{v_1, ..., v_b\}$, corresponding to the leaves $\{\ell_1, ..., \ell_b\}$. Let (W_t) be a Glauber process on G_W using $k-1$ colors from another color set $\mathcal{C}_W$. We are going to define W_0 and couple (W_t) with (X_t) such that $|A_{X_t}(r)| = |A_{W_t}| + 1$ at any time t, where $A_{W_t} := \{c \in \mathcal{C}_W : \forall v_i, W_t(v_i) \neq c\}$. To do this, for every t we are going to define a bijection $f_t : \mathcal{C} \setminus \{X_t(r)\} \to \mathcal{C}_W$ such that $f_t(X_t(\ell_i)) = W_t(v_i)$ for all i. Notice that if such a bijection exists then $|A_{X_t}(r)| = |A_{W_t}| + 1$.

At time $t = 0$, pick any bijection f_0 from $\mathcal{C}_W$ to $\mathcal{C} \setminus \{X_0(r)\}$. Define W_0 by $W_0(v_i) = f(X_0(\ell_i))$ for all i. We will update f_t only when we choose the root to recolor at time t in the coupling of (W_t) and (X_t). To do the coupling at time $t+1$, we first choose a vertex v in $G^\star$ to recolor:

- if $v = \ell_i$, then we choose a random color c different from $X_t(r)$ to recolor v. Correspondingly, we choose the vertex v_i in G_W to recolor using color $f_t(c)$.

- if $v = r$, then we choose a random color c from $A_{X_t}(r)$ to recolor the root in $G^\star$. Correspondingly, we update the mapping f_t in the following natural way: $f_t(X_{t-1}(r)) = f_{t-1}(c)$, (and $f_t(c)$ is undefined).

Since (W_t) itself is a Glauber process that recolors the vertices of G_W uniformly at random from C_W, conditioning on $\mathcal{E}$, simple calculations yield that for any $t > T_0$,

$$\Pr\left[|A_{W_t}| \geq 1 \mid \mathcal{E}\right] \leq \frac{1}{(1+\epsilon)b^\epsilon \ln b}\,.$$

Then (31) follows by coupling.

Since the same thing happens for (Y_t) and the root is recolored at most $20\ln b$ times, then by the union bound, conditioning on $\mathcal{E}$, the probability that at each time we try to recolor the root after T_0 steps, the root is always frozen in both copies is at least $1 - (40\ln b)(p_0) = 1 - 40/((1+\epsilon)b^\epsilon)$. Finally, by Proposition 44, $\mathcal{E}$ happens with high probability, and hence the lemma holds. $\square$

Note that for the warm-up stage, we need to show, with probability at least $1/\text{poly}(\log b)$, that for $\epsilon \geq 0$, all of the leaf disagreements are of the same type in $O(b \ln b)$ steps. This is easier to prove for the $\epsilon > 0$ case – that this happens with high probability, if we run the dynamics for $T_{\text{w}} = 8(b+1)\ln b$ steps. For the threshold case when $\epsilon = 0$, we will prove a slightly weaker lemma, in the sense that the successful probability will be at least $\Omega(1/\ln^2 b)$.

Lemma 46. *Let $T'_w = T_0 + 2b \ln \ln b$. For $k = b/\ln b$, for any pair of initial states (x_0, y_0),*

$$\Pr\left[\forall \ell \in D^L_{T'_w}, X_{T'_w}(\ell) = Y_{T'_w}(r) \wedge Y_{T'_w}(\ell) = X_{T'_w}(r) \mid x_0, y_0\right] \geq 1/(2\ln^2 b).$$

Proof. We use a different approach to prove this lemma, since it is not true that the root will still always be frozen during T'_{w} steps with high probability.

Let $T_0 = 4(b+1)\ln b$. We first prove that after T_0 steps, with high probability, the number of disagreeing leaves is at most $O(\ln b)$, namely:

$$\Pr\left[|D^L_{T_0}| \geq 4\ln b \mid X_0 = x_0, Y_0 = y_0\right] \leq \frac{2}{b^2}. \tag{32}$$

To prove (32), we construct a simpler process that stochastically upper bounds the number of disagreements. We define the following Markov chain (U_t) on 2-colorings of the graph G_U which consists of b isolated vertices $\{v_1, ..., v_b\}$. We view the set of colors as $\{0, 1\}$. In each step, a random vertex v_i is chosen, then with probability $1/(k-1)$, v_i is recolored to 1, and with probability $1 - 1/(k-1)$, v_i is recolored to 0. Let $D^U_t = \{v \in \{v_1, v_2, ..., v_b\} : U_t(v) = 1\}$. The initial state U_0 is constructed in the following way: for any $i > 0$, $U_0(v_i) = 1$ if and only if $x_0(\ell_i) \neq y_0(\ell_i)$. By associating the b vertices of G_U with the leaves of $G^\star$, we can easily couple the process (U_t) with (X_t, Y_t) such that $|D^U_t| \geq |D^L_t|$.

Let $\mathcal{E}$ denote the event that all of the vertices of G_U are recolored at least once in T_0 steps. Note, $\Pr[\mathcal{E}] \geq 1 - 1/b^2$. Conditioned on $\mathcal{E}$, the expected size of $|D^U_{T_0}|$ is

$b/(k-1) \approx \ln b$. Then we have

$$\begin{aligned} \Pr\left[|D^U_{T_0}| \geq 4\ln b\right] &\leq \Pr\left[|D^U_{T_0}| \geq 4\ln b \mid \mathcal{E}\right] + \Pr\left[\mathcal{E}\right] \\ &\leq \frac{2}{b^2}. \end{aligned}$$

Where, for the last inequality, we have used the Chernoff bounds (see, e.g., Theorem 4.5 Part 2 in [48]). Since $|D^U_t| \geq |D^L_t|$, this proves (32).

Hence, with high probability there are $O(\ln b)$ disagreeing leaves in $G^\star$ at time T_0. Notice that from time T_0, if we recolor all of the disagree leaves before we recolor the root again, then all of the remaining disagreements in the leaves will be of the same type (more precisely, for such a leaf ℓ that becomes a disagreement at time t we will have that $X_t(\ell) = Y_{T_0}(r)$ and $Y_t(\ell) = X_{T_0}(r)$), and this implies the desired conclusion of the lemma. To this end, let $\mathcal{E}_2$ be the event that the root is not chosen from recoloring from time T_0 to T'_{w}. Let $\mathcal{E}_3$ be the event that each leaf in $D^L_{T_0}$ is recolored at least once in the interval of times $[T_0, T'_{\mathrm{w}}]$. By simple calculations, we have that:

$$\Pr\left[\mathcal{E}_2\right] \geq \ln^{-2} b, \qquad \Pr\left[\mathcal{E}_3 \mid \mathcal{E}_2\right] \geq 1 - \frac{O(1)}{\ln b}. \tag{33}$$

Therefore, conditioned on $|D^L_{T_0}| \leq 4\ln b$, from time T_0 to T'_{w} with probability at least $2/(3\ln^2 b)$, both $\mathcal{E}_2$ and $\mathcal{E}_3$ happen, which implies all of the leaf disagreements will be of the same type at time T'_{w}.

In conclusion, combining the above bounds with (32), we proved that with probability at least $1/(2\ln^2 b)$, all of the uncoupled leaves are of the same at time T'_{w}. $\square$

After we succeed in the warm-up stage meaning that all of the leaf disagreements are of the same type, we enter the root-coupling stage, where we try to couple the root. Let T_1 be the first time that there is a common available color in the root and the coupling chain selects the root to recolor, that is

$$T_1 := T_1^{XY} = \min\{t : A_{X_t}(r) \bigcap A_{Y_t}(r) \neq \emptyset \text{ and the root } r \text{ is selected at step } t\}.$$

Lemma 47. *For $\epsilon \geq 0$, for any pair of initial states (x_0, y_0) where all of the leaf disagreements are of the same type (i.e., there is a pair of colors c_1, c_2 such that for all $\ell \in D_0^L$, we have $x_0(\ell) = c_1$ and $y_0(\ell) = c_2$), we have*

$$\Pr\left[T_1^{XY} < 4(b+1)\ln b \mid (X_0, Y_0) = (x_0, y_0)\right] > \frac{1}{4(1+\epsilon)b^\epsilon \ln b}.$$

Proof. First of all, by Proposition 43, $|A_{X_0}(r) \oplus A_{Y_0}(r)| \leq 2$. We are interested in the time t when there is a common color available for the root in (X_t, Y_t).

Let (Z_t) be a Glauber process on the graph G_Z of $b+1$ isolated vertices $\{v_0, v_1, v_2, ..., v_b\}$ in which v_0 corresponds to the root and v_i corresponds to the leaves ℓ_i for any $i > 0$. The color set used in the process (Z_t) is $\mathcal{C}_Z = [k] \setminus \{c_1, c_2\}$. Each step, (Z_t) chooses a random vertex and recolors it with a random color from the set $\mathcal{C}_Z$. Let T_Z be the stopping time on Z satisfying:

$$T_1^Z = \min\{t > 2(b+1)\ln b : |A_{Z_t}| \geq 1 \text{ and } v_0 \text{ is selected at the step } t\},$$

where $A_{Z_t} = \{c \in \mathcal{C}_Z : \forall i \in [1, .., b], Z_t(v_i) \neq c\}$ is the set of unused colors in the vertices $\{v_1, v_2, ...v_b\}$. We want to couple (Z_t) with (X_t, Y_t) in such a way that $T_1^Z \geq T_1^{XY}$ for all the runs, and then if we show that for any initial state z_0, we have

$$\Pr\left[T_1^Z < 4(b+1)\ln b \mid Z_0 = z_0\right] > \frac{1}{4(1+\epsilon)b^\epsilon \ln b}. \tag{34}$$

Then by the coupling, we know that the lemma is also true.

Now we are going to construct the coupling between (Z_t) and (X_t, Y_t) for $t \leq T_1^{XY}$. Let z_0 be the initial state satisfying that for any $i \in [1, .., b]$, if $x_0(\ell_i) = y_0(\ell_i) \in \mathcal{C}_Z$ then $z_0(v_i) = x_0(\ell_i)$, otherwise we give an arbitrary color to the vertex v_i. On each step t, we first randomly select a vertex in $G^\star$ to update in (X_t, Y_t) and accordingly we select the corresponding vertex in G_Z to update in Z_t:

- If the vertex is a leaf ℓ_i:

  (X_t, Y_t) selects a random color c or a disagreement to update. If $c \in \mathcal{C}_Z$ then we give the same color to v_i in Z_t, otherwise we give a random color to v_i.

- If the vertex is the root r:

  Recolor the root on (X_t, Y_t) according to the maximal one-step coupling and pick a random color in $\mathcal{C}_Z$ to recolor v_0 in Z.

Observe that, $A_{Z_t} \subseteq A_{X_t}(r) \bigcap A_{Y_t}(r)$ for any $0 \le t \le T_1^{XY}$, which implies that $T_1^Z \ge T_1^{XY}$ holds with probability 1. Now we will show that (34) holds. Let $\mathcal{E}$ be the event that, in (Z_t), every vertex in the graph G_Z will be recolored at least once within the first $2(b+1)\ln b$ steps. Let t_z be the first time after time $2(b+1)\ln b$ when the dynamics (Z_t) recolors the root. For each color $c \in \mathcal{C}_Z$, define the indicator function $\mathbf{1}_c := \mathbf{1}\{c \neq Z_{t_z}(v_i), \forall 1 \le i \le b\}$. These indicator functions are negatively associated to each other (c.f., Theorem 14 in [20]). It follows by elementary calculations that conditioned on $t_z = t$ for some $t > 2(b+1)\ln b$ and for large enough b, we have

$$\begin{aligned}
&\Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t\right] \\
&\quad \ge \Pr[\mathcal{E}] \cdot \Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t, \mathcal{E}\right] \\
&\quad \ge 0.99\Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t, \mathcal{E}\right] \qquad (\text{since } \Pr[\mathcal{E}] > 1 - 1/b^2) \\
&\quad \ge 0.99\left(1 - \prod_{c \in \mathcal{C}_z} \Pr\left[\mathbf{1}_c = 0 \mid t_z = t, \mathcal{E}\right]\right) \qquad (\text{negative association}) \\
&\quad \ge 0.99\left(1 - \left(1 - \left(1 - \frac{1}{|\mathcal{C}_Z|}\right)^b\right)^{|\mathcal{C}_Z|}\right) \\
&\quad \ge \frac{1}{3(1+\epsilon)b^\epsilon \ln b}.
\end{aligned} \tag{35}$$

Since $\Pr\left[t_z \le 4(b+1)\ln b\right] > 1 - 1/b^2$, by applying (35) we have

$$\begin{aligned}
\Pr\left[T_1^Z < 4(b+1)\ln b \mid Z_0 = z_0\right] &\ge \sum_{t=2(b+1)\ln b}^{4(b+1)\ln b} \Pr\left[A_{Z_t} \neq \emptyset \mid t_z = t\right] \cdot \Pr\left[t_z = t\right] \\
&\ge \frac{\Pr\left[t_z \le 4(b+1)\ln b\right]}{3(1+\epsilon)b^\epsilon \ln b} \\
&\ge \frac{1}{4(1+\epsilon)b^\epsilon \ln b}.
\end{aligned}$$

This completes the proof of Lemma 47. □

We also know that when the root is recolored, if $|A_X(r) \oplus A_Y(r)| \leq 2$ and $|A_X(r) \bigcap A_Y(r)| \geq 1$ holds, then the probability that the root will be recolored to the same color in both X and Y is at least $1/2$. Hence, at time $T_1 = T_1^{XY}$, with probability at least $1/2$ the root will become an agreement. Combining with Lemma 45, we proved that with probability at least $1/O((1+\epsilon)b^\epsilon \ln b)$ when $\epsilon > 0$, starting from arbitrary initial states (x_0, y_0), the root will couple in at most $12(b+1)\ln b$ steps and by that time all the disagreements (if there is any) in the leaves are of the same type. When $\epsilon = 0$, combining with Lemma 46, we get that the probability of the same event happening is at least $1/O(\ln^3 b)$.

The last step is to let all of the disagreements in the leaves go away without changing the root to a disagreement, again with constant probability, after $T_2 = 4(b+1)\ln b$ more steps. Here is the precise statement of the lemma.

Lemma 48. *For $\epsilon \geq 0$, consider a pair of initial states (x_0, y_0) where the root r agrees (i.e., $x_0(r) = y_0(r)$) and all of the leaf disagreements are of the same type (i.e., there is a pair of colors c_1, c_2 such that for all $\ell \in D_0^L$, we have $x_0(\ell) = c_1$ and $y_0(\ell) = c_2$). Then, with probability at least $1/2$ after $T_2 = 4(b+1)\ln b$ steps, we have $X_{T_2} = Y_{T_2}$.*

Proof. First, observe that with high probability after T_2 steps all of the leaves will be recolored at least once. Assuming all of the leaves are recolored at least once, if the root does not become a disagreement within these T_2 steps, then all of the leaves will be agreements. Therefore, we just need to show that the root will not change to a disagreement in T_2 steps with probability at least $3/5$. This is done by a coupling argument.

Let t_2 be the first time when the root becomes a disagreement, that is, $X_{t_2}(r) \neq Y_{t_2}(r)$. Note, since any disagreements on the leaves are colored c_1 in X_0 and c_2 in Y_0, either $X_{t_2}(r) = c_2$ and/or $Y_{t_2}(r) = c_1$. Therefore, we define the stopping times T_2^X

and T_2^Y as follows:

$$T_2^X = \min\{t : X_t(r) = c_2\}, \quad T_2^Y = \min\{t : Y_t(r) = \ c_1\}.$$

We can assume without loss of generality that $X_0(r)$ (and hence $Y_0(r)$) does not equal either c_1 or c_2. Otherwise, by the hypothesis of the lemma, there are no disagreements in the leaves and hence $X_0 = Y_0$. Hence, our goal is to show that

$$\Pr\left[T_2^X \leq T_2 \text{ or } T_2^Y \leq T_2\right] < \frac{2}{5}.$$

And the main step is to show that

$$\Pr\left[T_2^X \leq T_2\right] \ < \ \frac{1}{5}. \tag{36}$$

Let (S_t) be a random subset process on $V(G^\star)$. Each time it picks a vertex v,

- If $v \neq r$, with probability $1/(k-1)$, $S_{t+1} = S_t \cup \{v\}$ and with probability $1 - 1/(k-1)$, $S_{t+1} = S_t \setminus \{v\}$;

- If $v = r$, if $S_t = \emptyset$ then $S_{t+1} = \{r\}$ otherwise $S_{t+1} = S_t$.

Let us define $T^S = \min_t\{t : r \in S_t\}$. We are going to couple (S_t) with (X_t) such that $\{v \in V(G^\star) : X_t(v) = c_2\} \subseteq S_t$. This implies $T^S \leq T_2^X$. And if we can show that $\Pr\left[T^S \leq T_2\right] \leq 1/5$ then we have proved inequality (36).

The coupling (X_t, S_t) is defined as follows. We start with $S_0 = X_0^{-1}(c_2)$, the set of vertices of color c_2 in the initial coloring. Each time both processes picks the same vertex v to update.

- If $v = r$, X_t and S_t act independently at this time.

- If $v \neq r$ and $X_t(r) \neq c_2$, then X_t chooses a random color different from the root to recolor v and if that color is not c_2, $S_{t+1} = S_t \setminus \{v\}$ otherwise $S_{t+1} = S_t \cup \{v\}$.

- If $v \neq r$ and $X_t(r) = c_2$, then X_t chooses a random color different from c_2 to recolor v and if that color is not c_1, $S_{t+1} = S_t \setminus \{v\}$ otherwise $S_{t+1} = S_t \cup \{v\}$.

It is easy to see that this is a valid coupling. More importantly, it satisfies $X_t^{-1}(c_2) \subseteq S_t$.

Now we are going to show that $\Pr\left[T^S \le T_2\right] < 1/5$ holds. It is not hard to show that with probability at least 0.9, the first time when the root is updated is later than $0.1b$ steps. We now condition on this event. The indicators of whether each leaf is in S_t or not during those $0.1b$ steps are negatively associated (c.f., Theorem 14 in [20]). Then by using the Chernoff bound with negative association among the random variables (c.f., Proposition 7 in [20]), it can be shown that with high probability at least $\ge 0.01b$ many different leaves are recolored before the first time we recolor the root. Thus, together with the proof of Proposition 44, we can claim that with probability at least 0.85, before the first t such that $r \in S_t$, at least $0.01b$ many leaves have been recolored and root will be recolored at most $20 \ln b$ times before T_2. Denote this event as $\mathcal{E}$. We have

$$\Pr\left[T^S \le T_2\right] \le \Pr\left[T^S \le T_2 \mid \mathcal{E}\right] + \Pr\left[\bar{\mathcal{E}}\right] \le \Pr\left[T^S \le T_2 \mid \mathcal{E}\right] + 0.15.$$

In fact $\Pr\left[T^S \le T_2 \mid \mathcal{E}\right]$ can be arbitrarily small when b grows, since at each time t we update the root in (S_t), we know that the probability of $S_{t-1} = \emptyset$ is at most $b^{-0.01(1+\epsilon)}$, and we know that the root updates at most $20 \ln b$ times.

In conclusion, we proved inequality (36) and hence the lemma. $\square$

Finally, by combining Lemmas 45, 47 and 48 together, we can conclude that: when $\epsilon > 0$, with probability at least $1/(20(1+\epsilon)b^\epsilon \ln b)$ after $t = T_\mathrm{w} + T_1 + T_2 < T$ steps of the coupling, we have $X_t = Y_t$; when $\epsilon = 0$, from Lemmas 46, 47 and 48, we have that with probability at least $1/(20 \ln^3 b)$ after $t = T'_\mathrm{w} + T_1 + T_2 < T$ steps of the coupling, we have $X_t = Y_t$, which proves Lemma 42.

4.2.2 Upper Bound on Mixing Time for $k > b/\ln b$.

In this section we analyze the upper bound of the mixing time of the Glauber dynamics on the star graph $G^\star$ when $k = (1+\epsilon)b/\ln b$ for $\epsilon > 0$.

We will analyze the maximal one-step coupling using a weighted Hamming distance. The root r will have weight $w(r) = b^{\epsilon/2} > 1$ and the leaves will have weight $w(v) = 1$. For a set of vertices S, let $w(S) = \sum_{v \in S} w(v)$. Recall that the set of disagreeing vertices at time t of our coupled chains is denoted as

$$D_t = \{v \in V : X_t(v) \neq Y_t(v)\}$$

. Let D_t^r denote whether there is a disagreement at the root.

We want to show that the coupling decreases the distance in expectation. Hence, we say a pair of colorings (X_0, Y_0) are η-distance-decreasing if there exists a coupling $(X_0, Y_0) \to (X_1, Y_1)$ such that:

$$\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right] < (1 - \eta) w(D_0).$$

To simplify the analysis of the coupling, we will use the following theorem of Hayes and Vigoda [31] to utilize properties of the stationary distribution. The quantity $\mathrm{diam}(\Omega)$ is the diameter of Ω with respect to the Glauber dynamics. In our case, a trivial bound is $\mathrm{diam}(\Omega) \leq 2b$.

Theorem 49. *[31, Theorem 1.2] Let $\eta > 0$. Suppose $S \subseteq \Omega$ such that every $(X_0, Y_0) \in S \times \Omega$ is η-distance-decreasing, and*

$$\pi(S) \geq 1 - \frac{\eta}{16 \ diam(\Omega)},$$

then the mixing time is

$$T_{\mathrm{mix}} \leq 3\eta^{-1} \lceil \ln(32 \ diam(\Omega)) \rceil.$$

We use S as the set of colorings where the root has many available colors. Along the lines of the Dyer-Frieze [21] local uniformity results, we will prove the following statement about the available colors for the root r in a random coloring.

Lemma 50. *Let X be a random coloring of the star graph on b vertices. For every $\epsilon > 0$, there exists b_0, such that for all $b > b_0$ and $k = (1+\epsilon)b/\ln b$,*

$$\Pr\left[|A_X(r)| > b^{.9\epsilon}\right] > 1 - \exp(-b^{.99\epsilon}/10).$$

Proof of Lemma 50. Fix the color of the root to be c. Let σ be a random coloring conditional on the root receiving color c. We are going to prove that

$$\Pr\left[|A_\sigma(r)| \le b^{0.9\epsilon} \mid \sigma(r) = c\right] < \exp(-b^{0.99\epsilon}/10).$$

For each color $i \in C \setminus \{c\}$, let Z_i be the indicator function that $c \in A_\sigma(r)$. $|A_\sigma(r)| = \sum_{i \in C} Z_i$. By Theorem 14 in [20], the Z_i's are negatively associated with each other once the root is fixed. Note that for b sufficiently large,

$$\begin{aligned} \mathrm{E}\left[|A_\sigma(r)|\right] &\ge k\exp(-b/(k-1)) \\ &\ge b^{0.99\epsilon}. \end{aligned}$$

Now applying the Chernoff bound, which hold for negatively associated random variables (c.f., Proposition 7 in [20]), we have:

$$\Pr\left[|A_\sigma(r)| \le b^{0.9\epsilon} \mid \sigma(r) = c\right] < \exp(-b^{0.99\epsilon}/10).$$

□

Hence, we let the set S be those colorings $X \in \Omega$ where $|A_X(r)| \ge b^{.9\epsilon}$ when we use Theorem 49.

We need to analyze $\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right]$. Note, when a leaf v is recolored, if the root is a disagreement (i.e., $X_0(r) \ne Y_0(r)$) then with probability $1/(k-1)$ we have

$X_1(v) \neq Y_1(v)$. Hence,

$$
\begin{aligned}
&\mathrm{E}\left[w(D_1^L) \mid X_0, Y_0\right] \\
&= \sum_{v \in V \setminus \{r\}} w(v) \Big[\Pr\left[v \text{ is recolored}\right] \cdot \Pr\left[X_1(v) \neq Y_1(v) \mid v \text{ is recolored}, X_0, Y_0\right] \\
&\quad + (1 - \Pr\left[v \text{ is recolored}\right]) \mathbf{1}[X_0(v) \neq Y_0(v)] \Big] \\
&= \frac{b}{b+1} \frac{\mathbf{1}[r \in D_0]}{k-1} + \left(1 - \frac{1}{b+1}\right) w(D_0^L).
\end{aligned}
$$

There is probability at most $|D_0^L| / \max\{|A_{X_0}(r)|, |A_{Y_0}(r)|\}$ that $X_1(r) \neq Y_1(r)$, when the root r is recolored. Hence, for $X_0 \in S$, we have:

$$
\begin{aligned}
&\mathrm{E}\left[w(D_1^r) \mid X_0, Y_0\right] \\
&\leq w(r) \frac{1}{b+1} \frac{|D_0^L|}{\max\{|A_{X_0}(r)|, |A_{Y_0}(r)|\}} + \left(1 - \frac{1}{b+1}\right) w(D_0^r) \\
&\leq \frac{|D_0^L| b^{-\epsilon/3}}{b+1} + \left(1 - \frac{1}{b+1}\right) w(D_0^r).
\end{aligned}
$$

Therefore, for $(X_0, Y_0) \in S \times \Omega$, we have:

$$
\begin{aligned}
&\mathrm{E}\left[w(D_1) \mid X_0, Y_0\right] \\
&\leq \frac{1}{b+1} \left(\mathbf{1}[r \in D_0] \frac{b}{k-1} + b^{-\epsilon/3} |D_0^L| \right) + \left(1 - \frac{1}{b+1}\right) w(D_0) \\
&\leq w(D_0) + \frac{1}{b+1} \left(-w(D_0) + \mathbf{1}[r \in D_0] w(r) b^{-\epsilon/3} + b^{-\epsilon/3} |D_0^L| \right) \\
&\leq w(D_0) + \frac{1}{b} \left(-1 + b^{-\epsilon/4} \right) w(D_0).
\end{aligned}
$$

Thus, they are η-distance-decreasing for $\eta = (1 - b^{-\epsilon/4})/b$.

Now applying Theorem 49, by Lemma 50 we have the necessary bound on $\pi(S)$, and thus conclude for b sufficiently large we have:

$$
T_{\text{mix}} \leq (6b \ln b)/(1 - b^{-\epsilon/4}) \leq 12 b \ln b.
$$

This completes the proof of Theorem 40.

4.2.3 A Simple Generalization to $k = o(b/\ln b)$.

In all of the previous sections, we assumed $k = Cb/\ln b$ where C is constant. But we are also interested in the case when k is constant, say a hundred colors, and what the mixing time of the Glauber dynamics will be in this case. Let $\alpha = \alpha(k, b) := b/(k \ln b)$. We would also like to see how to generalize the upper bound and lower bound analysis assuming α is any function growing with b, that is when k is $o(b/\ln b)$. Actually, all of our proofs will be the same and we just need to slightly modify the statements.

For the upper bound, we change Lemma 42 and Lemma 47 into the following ones.

Lemma 51. *Let $T = 20b \ln b$. There exists b_0, for all $(x_0, y_0) \in \Omega \times \Omega$, all $\alpha(k, b) \geq 2$, and all $b > b_0$ the following holds:*

$$\Pr\left[X_T = Y_T \mid X_0 = x_0, Y_0 = y_0\right] \geq 1/(20\alpha(k, b) b^{\alpha(k,b)} \ln b).$$

Lemma 52. *For any pair of initial states (x_0, y_0) where all of the leaf disagreements are of the same type, then*

$$\Pr\left[T_1^{XY} < 4b \ln b \mid (X_0, Y_0) = (x_0, y_0)\right] \geq 1/(4\alpha(k, b) b^{\alpha(k,b)-1} \ln b).$$

Then by the same argument as in Section 4.2.1, we are able to show that the relaxation time of the Glauber dynamics on $G^\star$ is upper bounded by $O(\alpha b^\alpha \ln b)$. Thus, the mixing time of the Glauber dynamics on the complete tree is bounded by

$$T_{\text{mix}} = O\left(n^{\alpha + (\ln \alpha + 2 \ln \ln b + 20)/\ln b} \ln n\right),$$

and the relaxation time is bounded by

$$T_{\text{relax}} = O\left(n^{\alpha + (\ln \alpha + 2 \ln \ln b + 20)/\ln b}\right).$$

For the lower bound, we replace Lemma 15 and Lemma 16 into the following lemmas.

Lemma 53. *In a random coloring of the tree T, the probability that a vertex of T is not frozen is at most b^{-1}.*

Lemma 54.

$$\Pr_{\sigma\in\Omega^*}\left[\mathcal{A}^{\sigma}_{w_1,z} \mid \sigma(w_1)=c_1)\right] \le b^{-\alpha(k,b)}.$$

Then, by exactly the same way as in Section 3.2, we can show that the mixing time and the relaxation time of the Glauber dynamics on the complete tree T when $\alpha \ge 2$ is lower bounded by $\Omega(n^{\alpha}) = \Omega(n^{b/(k\ln b)})$.

4.2.4 Bounding the Log-Sobolev Constant: Proof of Theorem 41

In this section we will analyze the log-Sobolev constant c_{sob} of the heat-bath Glauber dynamics on the complete tree by comparing it with the spectral gap c_{gap}. For completeness, we prove Theorem 41, which is an improvement over the proof of Theorem 5.7 in Martinelli, Sinclar and Weitz [46]. In their paper, they proved it for the case of the Ising model on the complete tree with a fixed boundary condition, although they observed that it holds more generally. For convenience, we will use the same notation for the complete tree and its vertices, that is, T_ℓ stands for both the complete tree of height ℓ and its vertices $V(T_\ell)$.

Let $B \subseteq A \subseteq T$ be two subsets of the vertices on tree T. Let $\eta \in \Omega$ be a configuration. Let $E^{\eta}_A(f)$ be the expectation of f under a prefixed distribution μ in the region A with boundary condition η. That is

$$E^{\eta}_A(f) = \sum_{\sigma} \frac{\mu(\sigma)}{Z} f(\sigma)$$

where σ ranges over the configurations that are the same as η outside A (denoted as $\sigma \sim_A \eta$) and Z is the normalizing factor. The quantities Var^{η}_A and Ent^{η}_A are defined similarly. If we drop η then $E_A(f), \mathrm{Var}_A(f), \mathrm{Ent}_A(f)$ become functions from Ω to R. The following are standard facts concerning variance and entropy; the first being the chain rule, and the second follows from the so-called tensoring property over a product distribution – see e.g., Proposition 5.6 of [40]. In the following, we will use the fact that the distribution on configurations over the tree with the root removed, has a product form over the subtrees rooted at the children of the root, to satisfy the hypothesis for the tensoring property.

Proposition 55.

$$Var^{\eta}_{A}(f) = E^{\eta}_{A}(Var_B(f)) + Var^{\eta}_{A}(E_B(f)).$$

$$Ent^{\eta}_{A}(f) = E^{\eta}_{A}(Ent_B(f)) + Ent^{\eta}_{A}(E_B(f)).$$

Proposition 56. *Let $A = \bigcup A_i$ where A_i are disjoint, and suppose that conditioning on the boundary being η, the probability of A_i's being in any configuration for different i's is completely independent. Then*

$$Var^{\eta}_{A}(f) \leq \sum_i E^{\eta}_{A}(Var_{A_i}(f)),$$

and

$$Ent^{\eta}_{A}(f) \leq \sum_i E^{\eta}_{A}(Ent_{A_i}(f)).$$

Lemma 57. *Let $c_{sob}(\ell)$ be the log-Sobolev constant of the heat-bath Glauber dynamics on the complete tree of height $\ell > 0$ with the root being attached to an external vertex with a fixed color, then*

$$c_{sob}(\ell)^{-1} \leq c_{sob}(\ell-1)^{-1} + \alpha \cdot c_{gap}(\ell)^{-1},$$

where $\alpha = \frac{\log(k-2)}{1-2/(k-1)} = c_{sob}(0)^{-1}$.

Proof. Let f be any non-negative function. Let I be the set of vertices in the complete tree T_ℓ without the root, i.e., $I = T_\ell \backslash \{\text{root}\}$. Let us first use Proposition 55 to analyze the $\text{Ent}(f)$:

$$\text{Ent}(f) = E(\text{Ent}_I(f)) + \text{Ent}(E_I(f))$$

We will bound $E(\mathrm{Ent}_I(f))$ and $\mathrm{Ent}(E_I(f))$ separately. For $E(\mathrm{Ent}_I(f))$, by Proposition 56, it can be upper bounded as

$$E(\mathrm{Ent}_I(f)) \leq \sum_v E(\mathrm{Ent}_{T_v}(f)), \tag{37}$$

where v ranges over all the children of the root of T_ℓ and T_v denotes the subtree of T_ℓ rooted at the vertex v. Let $\eta \in \Omega(T_\ell)$, then for a specific $\mathrm{Ent}^\eta_{T_v}(f)$, we then have

$$\mathrm{Ent}^\eta_{T_v}(f) \leq c_{sob}(\ell-1)^{-1}\mathcal{D}_{T_v}(\sqrt{f}), \tag{38}$$

where $\mathcal{D}_{T_v}(\sqrt{f})$ is the corresponding Dirichlet form for the dynamics on the subtree T_v. For the heat-bath Glauber dynamics, since $P(\sigma,\tau) \neq 0$ only if they differ at a single vertex, we can further derive that

$$\begin{aligned}\mathcal{D}_{T_v}(f) &= \frac{1}{2}\sum_{\sigma,\tau}(f(\sigma)-f(\tau))^2\mu(\sigma)P(\sigma,\tau) \\ &= \frac{1}{2}\sum_{x\in T_v} E^\eta_{T_v}(\mathrm{Var}_{\{x\}}(f))\end{aligned} \tag{39}$$

where $\mu(\sigma)$ is the marginal distribution with respect to η.

Then, from (37), (38) and above we have

$$\begin{aligned}E(\mathrm{Ent}_I(f)) &\leq \sum_v E(\mathrm{Ent}_{T_v}(f)) \quad \text{(by (37))} \\ &\leq \sum_v c_{sob}(\ell-1)^{-1}E(\mathcal{D}_{T_v}(\sqrt{f})) \quad \text{(by (38))} \\ &= \sum_v c_{sob}(\ell-1)^{-1}E\big(\sum_{x\in T_v} E^\eta_{T_v}[\mathrm{Var}_{\{x\}}(f)]\big) \quad \text{(by (39))} \\ &= c_{sob}(\ell-1)^{-1}\sum_{x\in I} E(\mathrm{Var}_{\{x\}}(f)) \\ &\leq c_{sob}(\ell-1)^{-1}\mathcal{D}(\sqrt{f}) \quad \text{(by applying (39) again)}\end{aligned}$$

For $\mathrm{Ent}(E_I(f))$, $E_I(f)$ can be viewed as a function from $\{1,2,...,k-1\}$ to R since those $k-1$ values can represent the colors of the root (boundary). Therefore $\mathrm{Ent}(E_I(f))$ is the entropy of the random variable $E_I(f)$ taking $k-1$ values uniformly at random. It is well-known (see e.g. Appendix of [16]) that $\frac{\log(k-2)}{1-2/(k-1)}$ is the inverse of the log-Sobolev constant of the random walk $\mathcal{R}$ on the complete graph K_{k-1}, which jumps to stationarity in one step. Thus, letting $\alpha = \frac{\log(k-2)}{1-2/(k-1)}$, we may upper bound $\mathrm{Ent}(E_I(f))$ as follows.

$$\begin{aligned}
\mathrm{Ent}(E_I(f)) &\leq \alpha \mathcal{D}_{\mathcal{R}}(\sqrt{E_I(f)}) \quad \text{(by the log-Sobolev Inequality)} \\
&= \alpha \mathrm{Var}_{\mathcal{R}}(\sqrt{E_I(f)}) \quad \text{(for the complete graph } P(x,y) = \pi_{\mathcal{R}}(y)\text{)} \\
&= \alpha \mathrm{Var}_T(\sqrt{E_I(f)}) \\
&\leq \alpha\big(E[E_I(f)] - E^2(\sqrt{E_I(f)})\big) \quad \text{(by the definition of the variance)} \\
&\leq \alpha E(\sqrt{f})^2 - E^2(\sqrt{f}) \quad \text{(by the concavity of } \sqrt{x}\text{)} \\
&\leq \alpha c_{gap}(\ell)^{-1} \mathcal{D}(\sqrt{f}) \quad \text{(by the definition of the spectral gap)}
\end{aligned}$$

Putting everything together, we proved

$$\begin{aligned}
\mathrm{Ent}(f) &= E(\mathrm{Ent}_I(f)) + \mathrm{Ent}(E_I(f)) \\
&\leq c_{sob}(\ell-1)^{-1}\mathcal{D}(\sqrt{f}) + \alpha c_{gap}(\ell)^{-1}\mathcal{D}(\sqrt{f}),
\end{aligned}$$

then by the definition of c_{sob}, we get

$$c_{sob}(\ell)^{-1} \leq c_{sob}(\ell-1)^{-1} + \alpha c_{gap}(\ell)^{-1}\,.$$

□

Lemma 58. *Let $c_{sob}(\ell)$ be the spectral gap of the heat-bath Glauber dynamics on the complete tree of height $\ell > 0$ with the root being attached to an external vertex with a fixed color, then for $\ell > 0$, we have $c_{gap}(\ell) \leq c_{gap}(\ell-1)/b$.*

Proof. Let $\mathcal{D}_\ell(f)$ and $\mathrm{Var}_\ell(f)$ be the Dirichlet form and the variance of function $f : \Omega(T_\ell) \to \mathcal{R}$ for the Glauber dynamics on the complete tree of height ℓ with the root attached to an external vertex with a fixed color. Let P_ℓ denote the probability transition of the dynamics, and let π_ℓ denote its unique stationary distribution.

Let g be the eigenfunction such that $c_{gap}(\ell-1) = \mathcal{D}_{\ell-1}(g)/\mathrm{Var}_{\ell-1}(g)$. Now we are going to construct a function $f : \Omega(T_\ell) \to \mathcal{R}$, such that $\mathcal{D}_\ell(f) \leq \mathcal{D}_{\ell-1}(g)$ and $\mathrm{Var}_\ell(f) = \mathrm{Var}_{\ell-1}(g)$. Then, since

$$c_{gap}(\ell) \leq \frac{\mathcal{D}_\ell(f)}{\mathrm{Var}_\ell(f)} \leq \frac{\mathcal{D}_{\ell-1}(g)}{b \cdot \mathrm{Var}_{\ell-1}(g)} = c_{gap}(\ell-1),$$

we prove the lemma.

Let $A \subseteq T_\ell$ be the set of non-leaf vertices of T_ℓ, i.e., $A = T_\ell \setminus L(T_\ell)$, where $L(T_\ell)$ is the set of leaves in the tree T_ℓ. There is a natural correspondence between vertices in A and in $T_{\ell-1}$. The function f is then defined as: for $\sigma \in \Omega(T_\ell)$ and $\sigma' \in \Omega(T_{\ell-1})$, $f(\sigma) = g(\sigma')$ if the configuration σ agrees with σ' on the subset A.

It is straightforward to show that $\mathrm{Var}_\ell(f) = \mathrm{Var}_{\ell-1}(g)$. We will show $\mathcal{D}_\ell(f) \leq \mathcal{D}_{\ell-1}(g)/b$. By definition,

$$\mathcal{D}_\ell(f) = \sum_{\sigma,\eta \in \Omega(T_\ell)} \pi_\ell(\sigma) P_\ell(\sigma,\eta)(f(\sigma) - f(\eta))^2.$$

For a subset of vertices $S \subset T_\ell$, let

$$\Omega(S) = \{\sigma' \in [k]^S : \text{ there exists } \sigma \in \Omega(T_\ell) \text{ where } \sigma(A) = \sigma'\}.$$

Let $\sigma', \eta' \in \Omega(A)$ be colorings of the internal vertices. Let $\phi, \psi \in \Omega(L(T_\ell))$ be colorings of the leaves. Finally, let $\circ$ be the concatenation operator as we previously defined, thus $\sigma' \circ \psi = \sigma \in \Omega(T_\ell)$ where $\sigma(A) = \sigma'$ and $\sigma(L(T_\ell)) = \phi$. Then we can rewrite the Dirichlet form as:

$$\mathcal{D}_\ell(f) = \sum_{\sigma',\eta' \in \Omega(A)} \sum_{\phi,\psi \in \Omega(L(T_\ell))} \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \psi)(f(\sigma' \circ \phi) - f(\eta' \circ \psi))^2.$$

According to the definition of the Glauber dynamics, for configurations $\sigma, \eta \in \Omega(T_\ell)$ which differ at more than one vertex, we have $P_\ell(\sigma,\eta) = 0$. Let $\oplus$ denote the symmetric difference. Now we can rewrite the Dirichlet form as:

$$
\begin{aligned}
&\mathcal{D}_\ell(f)\\
&= \sum_{v \in A} \sum_{\substack{\sigma', \eta' \in \Omega(A): \\ \sigma' \oplus \eta' = \{v\}}} \sum_{\phi \in \Omega(L(T_\ell))} \left[(f(\sigma' \circ \phi) - f(\eta' \circ \phi))^2 \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \phi) \right] \\
&+ \sum_{v \in L(T_\ell)} \sum_{\sigma' \in \Omega(A)} \sum_{\substack{\phi, \psi \in \Omega(L(T_\ell)): \\ \phi \oplus \psi = \{v\}}} \left[(f(\sigma' \circ \phi) - f(\sigma' \circ \psi))^2 \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \sigma' \circ \psi) \right] \\
&= \sum_{v \in A} \sum_{\substack{\sigma', \eta' \in \Omega(A): \\ \sigma' \oplus \eta' = \{v\}}} \left[(g(\sigma') - g(\eta'))^2 \sum_{\phi \in \Omega(L(T_\ell))} \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \phi) \right],
\end{aligned}
$$

since $g(\sigma' \circ \phi) = g(\sigma' \circ \psi) = f(\sigma')$.

Thus we only need to consider when the sole disagreement is at an internal vertex. We can further decompose based on whether the disagreement is an internal vertex of the tree $T_{\ell-1}$, which we denote as I, or a leaf of $T_{\ell-1}$.

For $v \in L(T_{\ell-1})$, the goal is to bound the sum $\sum_\phi \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \phi)$ by $\pi_{\ell-1}(\sigma')/(|T_{\ell-1}|(k-1)b)$, i.e., $\pi_{\ell-1}(\sigma') P_{\ell-1}(\sigma', \eta')/b$. We have the following observation: Fix the vertex v, for each color c such that $\sigma' \oplus \eta' = \{v\}$ and $\eta'(v) = c$, the quantity $Q(c) := \sum_\phi \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \phi)$ are the same, i.e., $Q(c) = Q(c')$ for any two colors $c \neq c'$ because of the symmetry. Therefore, in order to bound $Q(c)$, it is easier to bound $\sum_{c \neq \sigma'(v)} Q(c)$ by $\pi_{\ell-1}(\sigma')/(|T_{\ell-1}|b)$. Then, by taking the average over $k-1$ colors, we are done. It is a straightforward calculation to upper bound the sum of $Q(c)$:

$$
\begin{aligned}
\sum_{c \neq \sigma'(v)} Q(c) &= \pi_{\ell-1}(\sigma') \sum_\phi \frac{\pi_\ell(\sigma' \circ \phi)}{\pi_{\ell-1}(\sigma')} \sum_{c \neq \sigma'(v)} \mathbf{1}\{c \in A_{\sigma' \circ \phi}(v)\} P_\ell(\sigma' \circ \phi, \eta' \circ \phi) \\
&= \pi_{\ell-1}(\sigma') \sum_\phi \frac{\pi_\ell(\sigma' \circ \phi)}{\pi_{\ell-1}(\sigma')} \frac{|A_{\sigma' \circ \phi}(v)| - 1}{|T_\ell||A_{\sigma' \circ \phi}(v)|} \\
&\leq \pi_{\ell-1}(\sigma') \frac{1}{|T_{\ell-1}|b}, \qquad (40)
\end{aligned}
$$

where by definition, $A_{\sigma' \circ \phi}(v)$ is the set of available colors for vertex v in the configuration $\sigma' \circ \phi$.

Recall, I denotes the internal vertices of $T_{\ell-1}$, i.e., $I = V(T_{\ell-1}) \backslash L(T_{\ell-1})$. Similarly, for $v \in I$ we have:

$$\sum_{\sigma',\eta' \in \Omega(A):\sigma' \oplus \eta' = \{v\}} [(g(\sigma') - g(\eta'))^2 \sum_{\phi \in \Omega(L(T_\ell))} \pi_\ell(\sigma' \circ \phi) P_\ell(\sigma' \circ \phi, \eta' \circ \phi)]$$
$$= \sum_{\substack{\sigma',\eta' \in \Omega(T_{\ell-1}): \\ \sigma' \oplus \eta' = \{v\}}} (g(\sigma') - g(\eta'))^2 \pi_{\ell-1}(\sigma') P_{\ell-1}(\sigma', \eta')/b. \quad (41)$$

Combining (40) and (41), and summing over $v \in T_{\ell-1}$ we have shown that $\mathcal{D}_\ell(f) \leq \mathcal{D}_{\ell-1}(g)/b$, which implies the lemma. □

Proof of Theorem 41. Now we apply Lemma 57 inductively, and we get

$$c_{sob}^{-1} = c_{sob}^{-1}(H) \leq \alpha\big(1 + c_{gap}^{-1}(1) + \cdots + c_{gap}^{-1}(\lfloor \log_b n \rfloor)\big).$$

Then by applying Lemma 58 on the spectral gaps, we can conclude that

$$c_{sob}^{-1} \leq b\alpha c_{gap}^{-1}(H) \leq c_{gap}^{-1} \cdot 2b \log k.$$

□

CHAPTER V

A CONDITION FOR SPATIAL MIXING

In this chapter we present a new general approach for proving the strong spatial mixing property of the spin systems on various classes of graphs with regular structures, which for the case of the hard-core model on $\mathbb{Z}^2$, improves the lower bound of the uniqueness threshold to $\lambda_c(\mathbb{Z}^2) > 2.48$. In a subsequent paper, Sinclair *et al.* [65] apply the approach to analyze the anti-ferromagnetic Ising model with arbitrary field, and show that there is a deterministic fully polynomial approximation scheme for the partition function up to the corresponding critical point on the d-regular tree.

There are various algorithmic implications for finite subgraphs of the $\mathbb{Z}^2$ when $\lambda < 2.48$. Our results imply that Weitz's deterministic FPAS is also valid on subgraphs of $\mathbb{Z}^2$ for the same range of λ. Thanks to the existing literature on general spin systems ([44, 45, 15, 23]), our results also imply that the Glauber dynamics has $O(n \log n)$ mixing time for any finite subregion $G = (V, E)$ of $\mathbb{Z}^2$ when $\lambda < 2.48$, where $n = |V|$. It also provides an arguably simpler way to derive the main technical result of Weitz showing that any graph with maximum degree Δ has strong spatial mixing (SSM) when $\lambda < \lambda_c(\mathbb{T}_\Delta)$.

The technique we develop builds upon Weitz's work to get improved results for specific graphs of interest. We focus our attention on what is arguably the simplest, not yet well-understood, case of interest namely the square grid, or the 2-dimensional integer lattice $\mathbb{Z}^2$. Empirical evidence suggests that the critical point $\lambda_c(\mathbb{Z}^2) \approx 3.796$ [26, 6, 57], but rigorous results are significantly far from this conjectured point. The possibility of there being multiple such λ_c is not ruled out, although no one believes that this is the case.

To underline the difficulty in estimating bounds on λ_c, we remark that the *existence* of a (unique) critical activity λ_c remains conjectural and an open problem for $\mathbb{Z}^d$, for $d \geq 2$. In contrast, as mentioned in the introduction for the Ising model, the critical inverse temperature $\beta_c(\mathbb{Z}^2)$ has been known since 1944 [53]; the corresponding critical point for the q-state Potts model (for $q \geq 2$) has only recently been established (by Beffara and Duminil-Copin [7]) to be $\beta_c(q) = \log(1 + \sqrt{q})$, settling a long-standing open problem. The lack of monotonicity in λ in the hard-core model poses a serious challenge in establishing such a sharp result for this model. In fact, Brightwell *et al.* [12] showed that in general such a monotonicity need not hold, by providing an example with a non-regular tree.

5.1 Weitz's Approach

Since our work builds on that of Weitz's, we first describe the self-avoiding walk (SAW) tree representation introduced in [74]. Given $G = (V, E)$, we first fix an arbitrary ordering $>_w$ on the neighbors of each vertex w in G. For each $v \in V$, the tree $T_{\text{saw}}(G, v)$ is constructed as follows. Consider the tree T of self-avoiding walks originating from v, additionally including the vertices closing a cycle as leaves of the tree. The tree has a specially constructed boundary condition, which is described as follow. We fix each leaf of T to be occupied or unoccupied. Suppose a leaf vertex closes a cycle in G, say $w \to v_1 \to \dots v_\ell \to w$, then:

- If $v_1 >_w v_\ell$ we fix this leaf to be unoccupied;
- Otherwise if $v_1 <_w v_\ell$ we fix the leaf to be occupied.

Note that by the natural of the hard-core model, if the leaf is fixed to be unoccupied we simply remove that vertex from the tree. If the leaf is fixed to be occupied, we remove that leaf and all of its neighbors, i.e. we remove the parent of that leaf from the tree. The resulting tree is denoted as $T_{\text{saw}} = T_{\text{saw}}(G, v)$. See Figure 3 for an illustration of T_{saw} for a particular example.

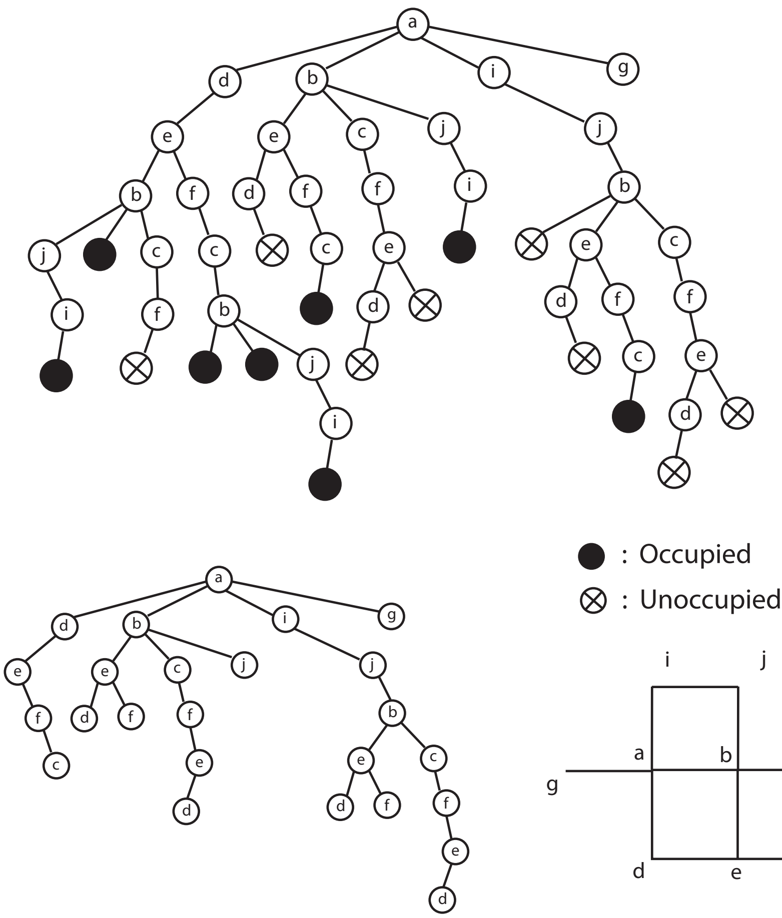

Figure 3: Example of self-avoiding walk tree T_{saw}. The above tree describes $T_{\text{saw}}(G, a)$ with occupied and unoccupied leaves, while the below one is the same tree after removing those assigned leaves. At each vertex, we consider the ordering $N > E > S > W$ of its neighbors where N, E, S, W represent the neighbors in the North, East, South, West directions, respectively.

Weitz [74] proves the following theorem for the hard-core model, which shows that the marginal distribution at the root in $T_{\text{saw}}(G, v)$ is identical to the marginal distribution for v in G. For a graph $G = (V, E)$, a subset $S \subset V$ and configuration ρ on S, for $T_{\text{saw}} = T_{\text{saw}}(G, v)$, let ρ in T_{saw} denote the configuration on S in T_{saw} where for $w \in S$ every occurrence of w in T_{saw} is assigned according to ρ.

Theorem 59 (SAW Tree Representation, Theorem 3.1 in [74])**.** *For any graph* $G = (V, E)$*,* $v \in V$*,* $\lambda > 0$*, and configuration* $\boldsymbol{\rho}$ *on* $S \subset V$*, for* $T = T_{\text{saw}}(G, v)$ *the following* *holds:*

$$\alpha_{G,\boldsymbol{\rho}}(v) = \alpha_{T,\boldsymbol{\rho}}(v).$$

Note, the tree $T_{\text{saw}}(G, v)$ preserves the distance of vertices from v in G, which implies the following corollary.

Corollary 60. *If SSM holds for $T_{\text{saw}}(G, v)$ for all v, then SSM holds for G. And the correlation decay rates are the same in both.*

The reverse implication of Corollary 60 does not hold since there are configurations on S in T_{saw} which are not necessarily realizable in G. Observe that if G has maximum degree Δ, any SAW tree of G is a subtree of the regular tree of degree Δ.

5.1.1 Our Proof Approach

In summary, Weitz [74] first shows (via Theorem 59) that to prove SSM holds on a graph $G = (V, E)$, it suffices to prove SSM holds on the trees $T_{\text{saw}}(G, v)$, for all $v \in V$. Weitz then proves that the regular tree $\mathbb{T}_\Delta$ "dominates" every tree of maximum degree Δ in the sense that, for all trees of maximum degree Δ, SSM holds when $\lambda < \lambda_c(\mathbb{T}_\Delta)$. We refine this second part of Weitz's approach. In particular, for graphs with extra structure, such as $G = \mathbb{Z}^2$, we bound $T_{\text{saw}}(\mathbb{Z}^2)$ by a tree T^* that is much closer to it than the regular tree $\mathbb{T}_\Delta$. We then establish a criterion that achieves better bounds on SSM for trees when the trees have extra structure.

The tree T^* will be constructed in a regular manner so that we can prove properties about it – the construction of T^* is governed by a (progeny) $t \times t$ matrix $\mathbf{M}$, whose rows correspond to t types of vertices, with the entry M_{ij} specifying the number of children of type j that a vertex of type i begets. We will then show a sufficient condition using entries of $\mathbf{M}$ which implies that SSM holds for T^* and for any subgraph of T^*, including $T_{\text{saw}}(\mathbb{Z}^2)$. The construction of T^* is reminiscent of the strategy employed in [4, 55] to upper bound the connectivity constant of several lattice graphs, including $\mathbb{Z}^2$.

The derivation of our sufficient condition has some inspiration from belief propagation algorithms.

As a byproduct of our proof that our new criterion implies SSM for T^*, we get a new (and simpler) proof of the second part of Weitz's approach, namely, that for all trees of maximum degree Δ, SSM holds when $\lambda < \lambda_c(\mathbb{T}_\Delta)$.

5.2 Branching Matrices and Strong Spatial Mixing

As alluded to above, we will utilize more structural properties of self-avoiding walk trees. To this end, we consider families of trees which can be recursively generated by certain rules; we then show that such a general family is also analytically tractable.

5.2.1 Definition of Branching Matrices

We say that the matrix $\mathbf{M}$ is a $t \times t$ branching matrix if every entry M_{ij} is a non-negative integer. We say the maximum degree of $\mathbf{M}$ is

$$\Delta = \Delta(\mathbf{M}) = \max_{1 \le i \le t} \sum_{1 \le j \le t} M_{ij},$$

the maximum row sum. Given a branching matrix $\mathbf{M}$, we define the following family of graphs. In essence, it includes a graph G if the self-avoiding walk trees of G can be generated by $\mathbf{M}$.

Definition 61 (Branching Family)**.** *Given a $t \times t$ branching matrix $\mathbf{M}$, $\mathcal{F}_{\le \mathbf{M}}$ includes trees which can be generated under the following restrictions:*

- *Each vertex in tree $T \in \mathcal{F}_{\le \mathbf{M}}$ has its type $i \in \{1, \dots, t\}$.*
- *Each vertex of type i has at most M_{ij} children of type j.*

In addition, we use the notation $G = (V, E) \in \mathcal{F}_{\le \mathbf{M}}$ if $T_{\text{saw}}(G, v) \in \mathcal{F}_{\le \mathbf{M}}$ for all $v \in V$.

For example, the family $\mathcal{F}_{\leq \mathbf{M}}$ with $\mathbf{M} = [\Delta]$ includes the family of trees with maximum branching Δ. On the other hand, $\mathcal{F}_{\leq \mathbf{M}}$ with $\mathbf{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ describes the family of graphs of maximum degree $\Delta+1$, by assigning the root of tree $T \in \mathcal{F}_{\leq \mathbf{M}}$ to be of type 1 and the other vertices of the tree to be of type 2. Note that if $\mathbf{M}$ has maximum degree Δ, then every $G \in \mathcal{F}_{\leq \mathbf{M}}$ also has maximum degree Δ.

In this framework, Weitz's result establishing SSM for all graphs of maximum degree Δ when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ can be stated as establishing SSM with uniform rate for all $G \in \mathcal{F}_{\leq \mathbf{M}}$ with $\mathbf{M} = \begin{pmatrix} 0 & \Delta \\ 0 & \Delta-1 \end{pmatrix}$; and we are interested in establishing its analogy for general $\mathbf{M}$. To this end, we will use the following notion of SSM for $\mathbf{M}$.

Definition 62. *Given a branching matrix* $\mathbf{M}$*, we say SSM holds for* $\mathbf{M}$ *if SSM holds with uniform rate for all* $G \in \mathcal{F}_{\leq \mathbf{M}}$.

To establish SSM for $\mathbf{M}$, it suffices to prove that SSM holds with uniform rate for all trees in $\mathcal{F}_{\leq \mathbf{M}}$ due to Corollary 60. In addition, note that SSM holds for $\mathbf{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ if and only if it holds for (Δ) since the root of a tree $T \in \mathcal{F}_{\leq \mathbf{M}}$ is the only possible vertex of type 1 in T. In general, it is a simple fact that any transient state of the branching matrix $\mathbf{M}$ has no effect on whether $\mathbf{M}$ has the SSM property or not. Therefore, we can simply reduce any branching matrix to the one with only the recurrent states.

5.2.2 Implications of SSM

We present a new approach for proving SSM for a branching matrix $\mathbf{M}$. There are multiple consequences of SSM for $\mathbf{M}$ as summarized in the following theorem. We first state some definitions needed for stating the theorem.

Following Goldberg *et al.* [30] we use the following variant of amenability for infinite graphs. Here we consider an infinite graph $G = (V, E)$. For $v \in V$ and a non-negative integer d, let $\mathrm{B}_d(v)$ denote the set of vertices within distance $\leq d$ from v, where distance is the length of the shortest path. For a set of vertices S, the (outer) boundary and neighborhood amenability are defined, respectively, as:

$$\partial S := \{w \in V : w \notin S, \text{ and } w \text{ has a neighbor } y \in S\} \quad \text{and} \quad r_d = \sup_{v \in V} \frac{|\partial \mathrm{B}_d(v)|}{|\mathrm{B}_d(v)|}.$$

The infinite graph is said to be *neighborhood-amenable* if $\inf_d r_d = 0$.

Now we can state the following theorem detailing the implications of SSM of interest to us.

Theorem 63. *For a $t \times t$ branching matrix $\mathbf{M}$, if SSM holds for $\mathbf{M}$ then the following hold:*

1. *For every $G \in \mathcal{F}_{\leq \mathbf{M}}$, SSM holds on G.*

2. *For every infinite graph $G \in \mathcal{F}_{\leq \mathbf{M}}$, there is a unique infinite-volume Gibbs measure on G.*

3. *If $\mathbf{M}$ has maximum degree Δ, if $t = O(1)$ and $\Delta = O(1)$, then for every (finite) $G \in \mathcal{F}_{\leq \mathbf{M}}$, Weitz's algorithm [74] gives an FPAS for approximating the partition function $Z(G)$.*

4. *For every infinite $H \in \mathcal{F}_{\leq \mathbf{M}}$ which is neighborhood-amenable, for every finite subgraph $G = (V, E)$ of H, the Glauber dynamics has $O(n^2)$ mixing time. Moreover, if $H = \mathbb{Z}^d$ for constant d, then for every finite subgraph $G = (V, E)$ of H, the Glauber dynamics has $O(n \log n)$ mixing time.*

Proof. Part 1 is by the definition of SSM for $\mathbf{M}$. The uniqueness result follows from the fact that the infinite-volume extremal Gibbs measures on the infinite graph G can be obtained by taking limits of finite measures, see Georgii [27] for an introduction to infinite-volume Gibbs measures, and see Martinelli [43] for Part 2. Part 3 immediately follows from the work of Weitz [74]. Finally, for Part 4, there is a long line of work showing that for the integer lattice $\mathbb{Z}^d$ in fixed dimensions, for the Ising model SSM on $\mathbb{Z}^d$ implies $O(n \log n)$ mixing time of the Glauber dynamics on finite subregions of $\mathbb{Z}^d$, e.g., see Cesi [15] and Martinelli [43] (and the references therein) for recent results on this problem. These results for the Ising model are typically stated for a general class of models, but that class does not include models with hard constraints, such as the hard-core model studied here. Dyer *et al.* [23] showed a simpler proof for the hard-core model that utilizes the monotonicity of the model. We use this result of [23] in Theorem 1 to get $O(n \log n)$ mixing time for subregions of $\mathbb{Z}^2$. Goldberg *et al.* [30, Theorem 8] showed that for k-colorings, if SSM holds for an infinite graph G that is neighborhood-amenable, the Glauber dynamics has $O(n^2)$ mixing time for all finite subgraphs of G. Their proof holds for the hard-core model which implies Part 4. $\square$

5.3 Establishing SSM for Branching Matrices

In this section we present a sufficient condition (and one of its simplified conditions) implying SSM for the family of trees generated by a branching matrix. As a consequence of the approach presented in this section we get a simpler proof of Weitz's result [74] implying SSM for all graphs with maximum degree Δ when $\lambda < \lambda_c(\mathbb{T}_\Delta)$. We then apply the condition presented in this section to $\mathbb{Z}^2$ in Section 5.4.1.

To show the decay of influence of a boundary condition $\boldsymbol{\rho}$, a common strategy is to prove some form of contraction for the 'one-step' iteration given in (42) below. More generally, we will prove such a contraction for an appropriate set of 'statistics' of the unoccupied marginal probability.

Definition 64. *A* statistic *of the univariate parameter $x \in [a, b]$ is a monotone (i.e., strictly increasing or decreasing) function $\varphi : [a, b] \to \mathbb{R}$.*

For a $t \times t$ branching matrix $\mathbf{M}$ we consider a set of t statistics $\varphi_1, \ldots, \varphi_t$, one for each type. For the simpler case when $\mathbf{M} = [\Delta]$ and hence $t = 1$, we have a single statistic φ. Our aim is proving contraction for an appropriate set of statistics of the probability that the root of a tree is unoccupied.

We first focus on the case of a single type, hence, $\mathbf{M} = [\Delta]$, $t = 1$ and there is a single statistic φ. Consider a tree $T = (V, E) \in \mathcal{F}_{\leq \mathbf{M}}$ with root r. For $v \in V$, let $N(v)$ denote the children of v, and let $d(v) := |N(v)|$ the number of children. Let T_v denote the subtree rooted at v. We will analyze the unoccupied probability for a vertex v, but v will always be the root of its subtree. Hence, to simplify the notation, for a boundary condition $\boldsymbol{\rho}$ on $S \subset V$, let $\alpha_{\boldsymbol{\rho}}(v) = \alpha_{T_v,\boldsymbol{\rho}}(v)$, i.e., it always denotes the marginal probability of the root v of the subtree T_v given the boundary condition $\boldsymbol{\rho}$ restricted to T_v.

A straightforward recursive calculation with the partition function leads to the following relation:

$$\alpha_{\boldsymbol{\rho}}(v) = \begin{cases} \frac{1}{1+\lambda} & \text{if } N(v) = \emptyset \\ \frac{1}{1+\lambda \prod_{w \in N(v)} \alpha_{\boldsymbol{\rho}}(w)} & \text{otherwise.} \end{cases} \tag{42}$$

Note, the unoccupied probability always lies in the interval $I := \left[\frac{1}{1+\lambda}, 1\right]$, i.e., for all v, all ρ, $\alpha_{\boldsymbol{\rho}}(v) \in I$.

For $v \in V$, let $m_{\boldsymbol{\rho}}(v) := \varphi(\alpha_{\boldsymbol{\rho}}(v))$ be the 'message' at vertex v. The messages satisfy the following recurrence:

$$m_{\boldsymbol{\rho}}(v) = \varphi\left(\frac{1}{1+\lambda \prod_{w \in N(v)} \alpha_{\boldsymbol{\rho}}(w)}\right) = \varphi\left(\frac{1}{1+\lambda \prod_{w \in N(v)} \varphi^{-1}(m_{\boldsymbol{\rho}}(w))}\right).$$

Our aim is to prove uniform contraction of the messages on all trees $T \in \mathcal{F}_{\leq \mathbf{M}}$. To this end, we will consider a more general set of messages. Namely, we consider messages $m_1, \ldots, m_\Delta$ where for every $1 \leq i \leq \Delta$, $m_i = \varphi(\alpha_i)$ and $\alpha_i \in I := \left[\frac{1}{1+\lambda}, 1\right]$. This set of tuples $\alpha_1, \ldots, \alpha_\Delta \in I$ contains all of the tuples obtainable on a tree.

For $\alpha_1, \ldots, \alpha_\Delta \in I$, let $m_i = \varphi(\alpha_i), 1 \leq i \leq \Delta$, and let

$$F(m_1, \ldots, m_\Delta) := \varphi\left(\frac{1}{1+\lambda \prod_{i=1}^{\Delta} \varphi^{-1}(m_i)}\right).$$

Ideally, we would like to establish the following contraction: there exists a $0 < \gamma < 1$ such that for all $\alpha_1, \ldots, \alpha_\Delta, \alpha'_1, \ldots, \alpha'_\Delta \in I$,

$$|F(m_1, \ldots, m_\Delta) - F(m'_1, \ldots, m'_\Delta)| \leq \gamma \max_{1 \leq i \leq \Delta} |m_i - m'_i|,$$

where $m_i = \varphi(\alpha_i)$ and $m'_i = \varphi(\alpha'_i)$. We will instead show that the following weaker condition suffices. Namely, that the desired contraction holds for all $|\alpha_i - \alpha'_i| \leq \epsilon$ for some $\epsilon > 0$. This is equivalent to the following condition.

Definition 65. *Let $I = \left[\frac{1}{1+\lambda}, 1\right]$. For the branching matrix $\mathbf{M} = [\Delta]$, we say that* Condition $(\star)$ *is satisfied if for all $\alpha_1, \ldots, \alpha_\Delta \in I$, by setting $m_i = \varphi(\alpha_i)$ for $1 \leq i \leq \Delta$, the following holds:*

$$\|\nabla F(m_1, \ldots, m_\Delta)\|_1 = \sum_{i=1}^{\Delta} \left| \frac{\partial F(m_1, \ldots, m_\Delta)}{\partial m_i} \right| < 1. \tag{$\star$}$$

Let us now consider a natural generalization of the above notion for a branching matrix with multiple types. Let $\mathbf{M}$ be a $t \times t$ branching matrix. For $1 \leq \ell \leq t$, let $\Delta_\ell = \sum_{k=1}^{t} M_{\ell k}$ denote the maximum number of children of a vertex of type ℓ. Once again, consider a tree $T = (V, E) \in \mathcal{F}_{\leq \mathbf{M}}$ with root r. For $v \in V$, let $t(v)$ denote its type. As before, $N(v)$ are the children of v, $d(v)$ is the number of children of v, and for a boundary condition $\boldsymbol{\rho}$ on $S \subset V$, $\alpha_{\boldsymbol{\rho}}(v)$ is the unoccupied probability for v in the tree T_v under $\boldsymbol{\rho}$.

The recursive calculation in (42) for α_v in terms of $\alpha_w, w \in N(v)$, still holds. For the case of multiple types, for $v \in V$, let $m_{\boldsymbol{\rho}}(v) := \varphi_{t(v)}(\alpha_{\boldsymbol{\rho}}(v))$ be the message at vertex v. The messages satisfy the following recurrence:

$$m_{\boldsymbol{\rho}}(v) = \varphi_{t(v)} \left(\frac{1}{1 + \lambda \prod_{w \in N(v)} \varphi_{t(w)}^{-1}(m_{\boldsymbol{\rho}}(w))} \right).$$

For each type $1 \leq \ell \leq t$, we consider contraction of messages derived from all $\alpha_1, \ldots, \alpha_{\Delta_\ell} \in I$. We need to identify the type of each these quantities α_i in order to determine the appropriate statistic to apply. The assignment of types needs to be consistent with the branching matrix $\mathbf{M}$. Hence, let $s_\ell : \{1, \dots, \Delta_\ell\} \to \{1, \dots, t\}$ be the following assignment. Let $M_{\ell,\le 0} = 0$ and for $1 \le i \le t$, let $M_{\ell,\le i} = \sum_{k=1}^{i} M_{\ell,k}$. For $1 \le i \le t$, for $M_{\ell,\le i-1} < j \le M_{\ell,\le i}$, let $s_\ell(j) = i$.

For type $1 \le \ell \le t$, for $\alpha_1, \dots, \alpha_{\Delta_\ell} \in I$, set $m_j = \varphi_{s_\ell(j)}(\alpha_j), 1 \le j \le \Delta_\ell$, and let

$$F_\ell(m_1, \dots, m_{\Delta_\ell}) := \varphi_\ell \left(\frac{1}{1 + \lambda \prod_{j=1}^{\Delta_\ell} \varphi_{s_\ell(j)}^{-1}(m_j)} \right).$$

Note,

$$m_{\boldsymbol{\rho}}(v) = F_{t(v)} \left(m_{\boldsymbol{\rho}}(w_1), \dots, m_{\boldsymbol{\rho}}(w_{d(v)}) \right) \quad \text{where} \quad N(v) = \{w_1, \dots, w_{d(v)}\}.^{1} \tag{43}$$

We generalize Condition ($\star$) to branching matrices with multiple types by allowing a weighting of the types by parameters $c_1, \dots, c_t$.

Definition 66. *Let $I = \left[\frac{1}{1+\lambda}, 1\right]$. For a $t \times t$ branching matrix $\mathbf{M}$, we say that* Condition ($\star\star$) *is satisfied if there exist $c_1, \dots, c_t$, such that for all $1 \le \ell \le t$, for all $\alpha_1, \dots, \alpha_{\Delta_\ell} \in I$, by setting $m_i = \varphi_{s_\ell(i)}(\alpha_i)$ for $1 \le i \le \Delta_\ell$, the following holds:*

$$\sum_{i=1}^{\Delta_\ell} c_{s_\ell(i)} \left| \frac{\partial F_\ell\left(m_1, \dots, m_{\Delta_\ell}\right)}{\partial m_i} \right| < c_\ell. \tag{$\star\star$}$$

The following lemma establishes a sufficient condition so that SSM holds for $\mathbf{M}$.

Lemma 67. *For a $t \times t$ branching matrix $\mathbf{M}$, if for every $1 \le \ell \le t$, φ_ℓ is continuously differentiable on the interval $I = \left[\frac{1}{1+\lambda}, 1\right]$ and $\inf_{x \in I} |\varphi_\ell'(x)| > 0$, and if Condition ($\star$) is satisfied for $t = 1$ or Condition ($\star\star$) is satisfied for $t \ge 2$ then SSM holds for $\mathbf{M}$, and hence the conclusions of Theorem 63 follow.*

Proof. For a tree $T = (V, E)$ with root r, let $\alpha_{+_L}(r)$ and $\alpha_{-_L}(r)$ denote the marginal probabilities that the root of T is unoccupied conditional on the vertices at level L (i.e., distance L from the root) being occupied and unoccupied, respectively.

[1] Strictly speaking, F_ℓ requires Δ_ℓ arguments, so for (43) to hold in the case when $d(v) < \Delta_\ell$ we can simply add additional arguments corresponding to $\alpha = 1$, which fixes these additional vertices to be unoccupied (and therefore absent).

The main result for proving Lemma 67 is that there exist $\gamma < 1$ and $L_0 < \infty$ such that for every tree $T \in \mathcal{F}_{\leq \mathbf{M}}$ and every integer $L \geq L_0$,

$$|\alpha_{+_L}(r) - \alpha_{-_L}(r)| \;\leq\; \gamma^L. \tag{44}$$

We first explain why (44) implies Lemma 67 and then we prove (44). Consider a tree $T = (V, E)$ with root r, and a boundary condition $\boldsymbol{\rho}$ on $S \subset V$. Set $L = \mathrm{dist}(r, S)$ as the distance of S to the root of T. The hard-core model on bipartite graphs has a monotonicity of boundary conditions (c.f., [23]) which implies that for odd L, $\alpha_{+_L}(r) \geq \alpha_{\boldsymbol{\rho}}(r) \geq \alpha_{-_L}(r)$, and for even L, $\alpha_{+_L}(r) \leq \alpha_{\boldsymbol{\rho}}(r) \leq \alpha_{-_L}(r)$. Hence, for any pair of boundary conditions $\boldsymbol{\rho}$ and $\boldsymbol{\eta}$ on S,

$$|\alpha_{\boldsymbol{\rho}}(r) - \alpha_{\boldsymbol{\eta}}(r)| \leq |\alpha_{+_L}(r) - \alpha_{-_L}(r)|\,.$$

Therefore, by the definition of WSM in Definition 7, proving (44) implies WSM for T. Since this holds for all $T' \in \mathcal{F}_{\leq \mathbf{M}}$, by Observation 9, it implies SSM for all $T' \in \mathcal{F}_{\leq \mathbf{M}}$, which implies SSM for $\mathbf{M}$.

We now turn our attention to proving (44). Fix a $t \times t$ branching matrix $\mathbf{M}$ and consider a tree $T = (V, E) \in \mathcal{F}_{\leq \mathbf{M}}$ with root r. Given $y \in [0, 1]$, let $\beta_{L,v}(y)$ denote the marginal probability that the root of T_v is unoccupied given all of the vertices at level L (in T_v) are assigned marginal probability y of being unoccupied (conditional on its parent being unoccupied). Intuitively, $\beta_{L,v}(y)$ can be thought as the marginal probability conditioned on a 'fractional' boundary configuration at level L. As in (42), $\beta_{L,r}(y)$ satisfies the following recurrence for $y \in [0, 1]$:

$$\beta_{L,r}(y) = \begin{cases} y & \text{if } L = 0, \\ \frac{1}{1+\lambda} & \text{if } L > 0 \text{ and } N(r) = \emptyset, \\ \frac{1}{1+\lambda \prod_{w \in N(r)} \beta_{L-1,w}(y)} & \text{otherwise.} \end{cases} \tag{45}$$

From (45) and (42), it follows that $\alpha_{+_L}(r) = \beta_{L,r}(1)$ and $\alpha_{-_L}(r) = \beta_{L,r}(0)$. Hence, in order to analyze the messages for $\alpha_{+_L}(r)$ and $\alpha_{-_L}(r)$, we will analyze the messages for $\beta_{L,r}(y)$. Therefore, for $v \in V$, let $m_{L,v}(y) = \varphi_{t(v)}(\beta_{L,v}(y))$. Analogous to (43), we now have that:

$$m_{L,r}(y) = F_{t(r)}\left(m_{L-1,w_1}(y), \ldots, m_{L-1,w_{d(r)}}(y)\right) \qquad \text{where} \quad N(r) = \{w_1, \ldots, w_{d(r)}\}.$$

Observe that for all $y \in [0,1]$, all $L > 0$, all $v \in V$, $\beta_{L,v}(y) \in I = \left[\frac{1}{1+\lambda}, 1\right]$, and hence we can use Condition $(\star\star)$ to analyze $m_{L,r}$.

Using the fact that $\beta_{L,v}(y)$ and $m_{L,v}(y)$ are continuously differentiable for $y \in [0,1]$, we have that for $L > 0$,

$$|\alpha_{+_L}(r) - \alpha_{-_L}(r)| \;=\; |\beta_{L,r}(1) - \beta_{L,r}(0)| \;\leq\; \int_0^1 \left|\frac{\partial \beta_{L,r}(y)}{\partial y}\right| dy \;\leq\; \frac{\int_0^1 \left|\frac{\partial m_{L,r}(y)}{\partial y}\right| dy}{\inf\limits_{x \in I}\left|\varphi'_{t(r)}(x)\right|}.$$

By the hypothesis of Lemma 67, we know that $\left|\varphi'_{t(r)}(x)\right| > 0$. Therefore, to prove the desired conclusion (44), it suffices to prove that there exist constants $K < \infty$ and $\eta < 1$ such that for every tree $T \in \mathcal{F}_{\leq \mathbf{M}}$ with root r, all $L > 0$,

$$\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| \leq c_{t(r)} K \eta^{L-1}. \tag{46}$$

Note that K and η should be independent of T and L, but may depend on $\lambda, \varphi_1, \ldots, \varphi_t$ and $c_1, \ldots, c_t$. The constant K will be the following:

$$K \;:=\; \frac{\lambda \Delta \max\limits_{1 \leq \ell \leq t} \sup\limits_{x \in I} |\varphi'_\ell(x)|}{\min\limits_{1 \leq \ell \leq t} c_\ell},$$

and the constant η will be the constant implicit in Condition $(\star\star)$.

We will show (46) by induction on L. First we verify the base case $L = 1$. In this case,

$$m_{L,r}(y) = \varphi_{t(r)}(\beta_{L,r}(y)) = \varphi_{t(r)}\left(\frac{1}{1 + \lambda y^{d(r)}}\right).$$

Thus,

$$\begin{aligned}
\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| &= \left|\frac{\partial \varphi_{t(r)}\left(\frac{1}{1+\lambda y^{d(r)}}\right)}{\partial y}\right| && \text{since } L=1 \\
&\le \sup_{x\in I}\left|\varphi'_{t(r)}(x)\right| \sup_{y\in[0,1]} \frac{\lambda d(r) y^{d(r)-1}}{\left(1+\lambda y^{d(r)}\right)^2} && \text{by the chain rule} \\
&\le \sup_{x\in I}\left|\varphi'_{t(r)}(x)\right| \lambda d(r) \\
&\le \sup_{x\in I}\left|\varphi'_{t(r)}(x)\right| \lambda \Delta \\
&\le c_{t(r)} K && \text{by the definition of } K.
\end{aligned}$$

This completes the analysis of the base case.

Now we proceed toward establishing the necessary induction step using the inductive hypothesis. We have that

$$\begin{aligned}
\left|\frac{\partial m_{L,r}(y)}{\partial y}\right| &= \left|\frac{\partial F_{t(r)}\left(m_{L-1,w_1}(y), \dots, m_{L-1,w_{d(r)}}(y)\right)}{\partial y}\right| \\
&= \left|\sum_{i=1}^{d(r)} \frac{\partial F_{t(r)}\left(m_1, \dots, m_{d(r)}\right)}{\partial m_i} \cdot \frac{\partial m_{L-1,w_i}(y)}{\partial y}\right| && \text{where } m_i := m_{L-1,w_i}(y) \\
&= \left|\sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \dots, m_{d(r)}\right)}{\partial m_i} \cdot \frac{1}{c_{t(w_i)}} \frac{\partial m_{L-1,w_i}(y)}{\partial y}\right| \\
&= \left|\sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \dots, m_{d(r)}\right)}{\partial m_i}\right| \\
&\qquad \times \max_{1\le i\le d(r)} \frac{1}{c_{t(w_i)}} \left|\frac{\partial m_{L-1,w_i}(y)}{\partial y}\right| && \text{by Hölder's inequality.}
\end{aligned} \tag{47}$$

From ($\star\star$), there exists a universal constant $\eta < 1$ such that

$$\left|\sum_{i=1}^{d(r)} c_{t(w_i)} \frac{\partial F_{t(r)}\left(m_1, \dots, m_{d(r)}\right)}{\partial m_i}\right| < \eta\, c_{t(r)}.$$

Therefore, it follows that

$$\begin{aligned}\left|\frac{\partial m_{L,r}\,(y)}{\partial y}\right| &\le \eta\, c_{t(r)} \cdot \max_{1\le i\le d(r)} \frac{1}{c_{t(w_i)}}\left|\frac{\partial m_{L-1,w_i}(y)}{\partial y}\right| && \text{by (47) and the definition of } \eta \\ &\le c_{t(r)} K \eta^{L-1} && \text{by the inductive hypothesis.}\end{aligned}$$

This completes the proof of (46), and hence that of Lemma 67. □

Condition ($\star\star$) together with Lemma 67 is a natural message-passing way of argument to establish the strong spatial mixing (SSM). The proof is crystal clear and easy to follow. However, sometimes with the understanding of Condition ($\star\star$), it is more convenient to work with a formulation which is more concise, concrete and easy to manipulate. We present it as follows.

For each type i, we treat the row $\mathbf{M}_i$ of $\mathbf{M}$ as a multi-set and each entry $\mathbf{M}_i(j)$ of the row denotes the number of elements the set $\mathbf{M}_i$ has for the type j. We use $t(w)$ to denote the type of vertex $w \in \mathbf{M}_i$. The following lemma, which is re-stating the previous lemma, provides such a condition for SSM to hold for the tree $T_{\mathbf{M}}$.

Lemma 68. *Let a branching matrix* $\mathbf{M}$*, be given. Assume then for each type* i*, there is a positive integrable function* Ψ_i *such that*

$$\frac{1-\alpha_i}{\Psi_i(\alpha_i)} \sum_{w\in\mathbf{M}_i} \Psi_{t(w)}(\alpha_w) < \gamma, \tag{48}$$

for α_w *in the range* $[1/(1+\lambda), 1]$ *for each child* w *and for* α_i *defined in* (42) *as a function of* α_w*'s. Then SSM holds for* $T_{\mathbf{M}}$*, i.e., WSM holds for all trees* T *in the family* $\mathcal{F}_{\le\mathbf{M}}$ *with a fixed rate* $\gamma < 1$.

Proof. We fix a tree $T \in \mathcal{F}_{\le\mathbf{M}}$ and we want to show that WSM holds for T as long as the condition (48) is true. Note that, this condition is independent of the tree T we choose. We assume that our boundary consists of all vertices at a fix distance from the root and each vertex on the boundary is independently set to unoccupied with marginal probability $\Gamma \in [0, 1]$.

Let $\alpha_{L,i}(\Gamma)$ be the marginal unoccupied probability for a type i vertex v, with distance L from the boundary, in the sub-tree T_v rooted at v. Putting this notation into Equation (42), for any fixed tree T we have:

$$\alpha_{L,i}(\Gamma) = \frac{1}{1 + \lambda \prod_{w \in \mathbf{M}_i} \alpha_{L-1,w}(\Gamma)}, \tag{49}$$

where $\alpha_{L-1,w}(\Gamma)$ equals to 1 if the vertex w is not in the tree T, and otherwise is the marginal unoccupied probability of vertex w in tree T_w with the fractional boundary condition Γ.

By integrating over Γ we can see that if

$$\left| \frac{\mathbf{d}\alpha_{L,i}(\Gamma)}{d\Gamma} \right| \leq \gamma^L, \tag{50}$$

then WSM holds for T at the vertex v of type i as discussed in the beginning of Section 5.4.

For a vertex v of type i, we have the following equation for the derivatives at α_i with respect to the boundary:

$$\frac{\mathbf{d}\alpha_{L,i}(\Gamma)}{\mathbf{d}\Gamma} = -(1 - \alpha_{L,i}(\Gamma))(\alpha_{L,i}(\Gamma)) \sum_{w \in \mathbf{M}_i} \frac{d\alpha_{L-1,w}(\Gamma)}{\mathbf{d}\Gamma} \frac{1}{\alpha_{L-1,w}(\Gamma)}. \tag{51}$$

From (51) it is sufficient to show for all i and all $\alpha_w \in [1/(1+\lambda), 1]$,

$$(1 - \alpha_i)(\alpha_i) \sum_{w \in \mathbf{M}_i} \frac{1}{\alpha_w} < \gamma$$

to obtain (50) and hence WSM holds for T, where in the inequality α_i is a function of α_w's as defined in (42). Note that, from here we already obtain a condition that implies the WSM holds for all tree $T \in \mathcal{F}_{\leq \mathbf{M}}$.

However, technically, it is hard to show the contraction of the above inequality due to the nonhomogeneous marginal distributions α_w from different children vertices as well as the irregular structure of the trees. We instead use a monotonic mapping φ_i (the messages from i to its parents) for each type i, and show that

$$\left| \frac{\mathbf{d}\varphi_i(\alpha_{L,i}(\Gamma))}{d\Gamma} \right| \leq \gamma^L, \tag{52}$$

which also implies that WSM holds for all trees $T \in \mathcal{F}_{\leq \mathbf{M}}$.

Setting $\Psi_i(x) = \left(x \cdot \frac{\mathbf{d}\varphi_i(x)}{\mathbf{d}x}\right)^{-1}$, we have

$$\frac{1}{\alpha_w} = \Psi_{t(w)}(\alpha_w)\frac{\mathbf{d}\varphi_{t(w)}(\alpha_w)}{\mathbf{d}\alpha_w},$$

and thus

$$\frac{\mathbf{d}\varphi_i(\alpha_i)}{\mathbf{d}\Gamma} = -\frac{(1-\alpha_i)}{\Psi_i(\alpha_i)}\sum_{w\in\mathbf{M}_i}\Psi_{t(w)}(\alpha_w)\frac{\mathbf{d}\varphi_{t(w)}(\alpha_w)}{\mathbf{d}\Gamma}.$$

Notice that to obtain (52), from this last equation we just need the condition (48) to be true. □

5.3.1 Reproving Weitz's Result of SSM for Trees

In this section, we aim at finding a good choice of statistics. First we find such a statistic for the case $\mathbf{M} = [\Delta]$, i.e., the case of a single type, which enables us to reprove Weitz's result [74] that when $\lambda < \lambda_c(\mathbb{T}_\Delta)$ SSM holds for every tree of maximum degree Δ.

Using Lemma 67 (and the simpler condition ($\star$) for the case of a single type) we obtain a simpler proof of Weitz's result [74] that for every tree T with maximum degree $\Delta + 1$ (hence, for every graph G of maximum degree $\Delta + 1$) and for all $\lambda < \lambda_c(\mathbb{T}_{\Delta+1}) = \Delta^\Delta/(\Delta-1)^{\Delta+1}$, SSM holds on T (and on G).

Theorem 69. *Let us define*

$$\varphi(x) = \frac{1}{s}\log\left(\frac{x}{s-x}\right),$$

where $s = \frac{\Delta+1}{\Delta}$. Then, Condition ($\star$) holds for $\mathbf{M} = [\Delta]$ and $\lambda < \lambda_c(\mathbb{T}_{\Delta+1})$. Consequently, SSM and the conclusions of Theorem 63 hold for $\mathbf{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$ and $\lambda < \lambda_c(\mathbb{T}_{\Delta+1})$.

Proof. Note that the proof we are going to present also works when using Lemma 68. First, a straightforward calculation implies that

$$\left|\frac{\partial F}{\partial m_i}\right| = \frac{1-\alpha}{s-\alpha}(s-\alpha_i),$$

where $\alpha_i = \varphi^{-1}(m_i)$ and $\alpha = \left(1 + \lambda \prod_{i=1}^{\Delta} \alpha_i\right)^{-1}$.

Hence, we have

$$\begin{aligned}
\|\nabla F\|_1 &= \sum_{i=1}^{\Delta} \left| \frac{\partial F}{\partial m_i} \right| \\
&= \sum_{i=1}^{\Delta} \frac{1-\alpha}{s-\alpha}(s-\alpha_i) \\
&\leq \frac{1-\alpha}{s-\alpha}\ \Delta\ \left(s - \left(\prod_{i=1}^{\Delta} \alpha_i \right)^{1/\Delta} \right) \qquad \text{by the arithmetic-geometric mean inequality} \\
&= \frac{1-\alpha}{s-\alpha}\ \Delta\ \left(s - \left(\frac{1-\alpha}{\lambda\alpha} \right)^{1/\Delta} \right). \qquad (53)
\end{aligned}$$

We now use the following technical lemma.

Lemma 70.

$$\max_{x\in[0,1]} \frac{(1-x)\left(1+\frac{1}{\Delta} - (\frac{1-x}{\lambda x})^{\frac{1}{\Delta}}\right)}{1+\frac{1}{\Delta}-x} \ \leq\ \frac{\omega}{1+\omega},$$

where Δ is a positive integer and ω is the unique solution to $\omega(1+\omega)^{\Delta} = \lambda$.

Using the above inequality (53) with Lemma 70, we have that:

$$\|\nabla F\|_1 < 1 \quad \text{if} \quad \frac{\omega}{1+\omega} \cdot \Delta \ <\ 1,$$

where ω is the unique solution of $\omega(1+\omega)^{\Delta} = \lambda$. This leads to the desired condition $\lambda < \lambda_c(\mathbb{T}_{\Delta+1}) = \Delta^{\Delta}/(\Delta-1)^{\Delta+1}$ so that SSM holds for $\mathbf{M} = [\Delta]$. Note that, this is equivalent to SSM for $\mathbf{M} = \begin{pmatrix} 0 & \Delta+1 \\ 0 & \Delta \end{pmatrix}$. This completes the proof of Theorem 69. □

Proof of Lemma 70. Let us define

$$\Phi_{\Delta}(x) = (\frac{1-x}{\lambda x})^{\frac{1}{\Delta}}$$

and

$$f(x) = \frac{(1-x)(1+\frac{1}{\Delta} - \Phi_{\Delta}(x))}{1+\frac{1}{\Delta}-x}.$$

Take the derivative of Φ with respect to x, we have

$$\Phi'_\Delta(x) = -\frac{\Phi_\Delta(x)}{\Delta x(1-x)}.$$

It is easy to see that Φ_Δ is a decreasing function in $[0,1]$ such that $\Phi_\Delta(0) = +\infty$ and $\Phi_\Delta(1) = 0$. Therefore it has a unique fixed point that can be shown to be $\bar{x} = \frac{1}{1+\omega}$. Moreover, it is the case that $\Phi_\Delta(x) > x$ if and only if $x < \bar{x}$. To prove the lemma, we notice that

$$f'(x) = \frac{(1+\frac{1}{\Delta})(\Phi(x)-x)}{\Delta x(1+\frac{1}{\Delta}-x)^2},$$

hence $f'(x) > 0$ for $x < \bar{x}$ and $f'(x) < 0$ for $x > \bar{x}$. This implies that f has a maximum at $\bar{x}$, namely $f(\bar{x}) = \frac{\omega}{1+\omega}$. □

5.3.2 Sufficient Criterion One: The DMS Condition

Theorem 69 suggests choosing $\varphi_j(x) = \frac{1}{s_j}\log\left(\frac{x}{s_j - x}\right)$ with appropriate parameters s_j for a general branching matrix $\mathbf{M}$. Under this choice, we obtain the following condition for SSM.

Definition 71 (DMS Condition). *Given a $t \times t$ branching matrix $\mathbf{M}$ and $\lambda^* > 0$, for $s_1, \ldots, s_t > 1$ and $\boldsymbol{c} = (c_1, \ldots, c_t) > 0$, let $\boldsymbol{D}$ and $\boldsymbol{S}$ be the diagonal matrices defined as*

$$D_{jj} = \sup_{\alpha \in \left[\frac{1}{1+\lambda^*}, 1\right]} \frac{(1-\alpha)\left(1 - \theta_j\left(\frac{1-\alpha}{\lambda^*\alpha}\right)^{1/\Delta_j}\right)}{s_j - \alpha} \qquad \text{and} \qquad S_{jj} = s_j,$$

where

$$\theta_j := \frac{\left(\prod_\ell c_\ell^{M_{j\ell}}\right)^{1/\Delta_j}}{\sum_\ell c_\ell s_\ell M_{j\ell}/\Delta_j} \qquad \text{and} \qquad \Delta_j = \sum_\ell M_{j\ell}.$$

We say the DMS Condition holds for $\mathbf{M}$ and λ^* *if there exist $s_1, \ldots, s_t > 1$ and $\boldsymbol{c} > 0$ such that:*

$$(\boldsymbol{D}\mathbf{M}\boldsymbol{S})\,\boldsymbol{c} < \boldsymbol{c}.$$

Theorem 72. *If the DMS Condition holds for $\mathbf{M}$ and $\lambda^* > 0$, then Condition ($\star\star$) holds with the choice of $\varphi_j(x) = \frac{1}{s_j}\log\left(\frac{x}{s_j - x}\right)$ for all $\lambda \le \lambda^*$. Consequently, SSM and the conclusions of Theorem 63 hold for $\mathbf{M}$ and all $\lambda \le \lambda^*$.*

Proof. First, one can check that

$$\left|\frac{\partial F_j}{\partial m_i}\right| = \frac{1-\alpha}{s_j-\alpha}(s_{j_i}-\alpha_i),$$

where $\alpha_i = \varphi_{j_i}^{-1}(m_i)$ and $\alpha = \frac{1}{1+\lambda\prod_{i=1}^{\Delta_j}\alpha_i}$.

Hence, it follows that

$$\begin{aligned}
\sum_{i=1}^{\Delta_j} c_{j_i}\left|\frac{\partial F_j}{\partial m_j}\right| &= \frac{1-\alpha}{s_j-\alpha}\sum_{i=1}^{\Delta_j} c_{j_i}(s_{j_i}-\alpha_i) \\
&\quad \text{by the arithmetic-geometric mean Inequality} \\
&\le \frac{1-\alpha}{s_j-\alpha}\left(\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} - \Delta_j\left(\prod_{i=1}^{\Delta_j} c_{j_i}\alpha_i\right)^{1/\Delta_j}\right) \\
&= \frac{1-\alpha}{s_j-\alpha}\left(\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} - \Delta_j\left(\prod_{i=1}^{\Delta_j} c_{j_i}\right)^{1/\Delta_j}\left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right) \\
&= \frac{1-\alpha}{s_j-\alpha}\left(1-\theta_j\left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right)\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} \quad \text{by the definition of } \theta_j \\
&\le \frac{1-\alpha}{s_j-\alpha}\left(1-\theta_j\left(\frac{1-\alpha}{\lambda^*\alpha}\right)^{1/\Delta_j}\right)\sum_{i=1}^{\Delta_j} c_{j_i}s_{j_i} \\
&\le D_{jj}\sum_{\ell} M_{j\ell}c_\ell s_\ell \quad \text{by the definition of } D_{jj} \\
&< c_j \quad \text{by the DMS condition.}
\end{aligned}$$

which satisfies the desired condition ($\star\star$) of Lemma 67. This completes the proof of Theorem 72. □

5.3.3 Sufficient Criterion Two: Linear Programming Condition

All of the above analysis use Lemma 67 as the bridge to establish the Strong spatial mixing. Here we propose a more general way to use linear programming to solve the functional inequality (48) in Lemma 68 to establish the SSM. In practice, this method is a lot slower than the DMS condition, however, it does provide a better bound. This is because DMS condition is just a simple guess of the “optimal” functions and since we have a concrete form of the functions, a lots of simplifications can be done through calculations. In contrast, here we almost assume nothing is known and try to use linear programming interpolation to find the optimal functions.

Notice that if Ψ_i is positive and bounded for all i then inequality (48) is equivalent to

$$(1-\alpha_i)\sum_{w\in\mathbf{M}_i}\Psi_{t(w)}(\alpha_w) < \Psi_i(\alpha_i). \tag{54}$$

The idea to solve (54) is simple. We will restrict the search for Ψ_i to a family of positive piecewise linear functions with a finite number of discontinuities.

First of all, it is a simple fact that each α_i is in the interval $I = [1/(1+\lambda), 1]$. We will divide I into a set of d consecutive sub-intervals of the same size. Define

$$X_k = \frac{1}{1+\lambda} + k\frac{\lambda}{d(1+\lambda)}, \text{ for } k = 0, \ldots, d-1.$$

To ease the notation define $Y_k = X_{k+1}$ for $k = 0, \ldots, d-1$. Note that the intervals $[X_k, Y_k]$ partition I. Since the only requirements of $\Psi_i(x)$ are positive and integrable, we restrict the search for $\Psi_i(x)$ to the functions of the linear form $-a_{i,k}x + b_{i,k}$ in each interval $[X_k, Y_k]$ with $a_{i,k}, b_{i,k} > 0$.

Now, for each type i, the functional inequality can be decomposed according to different combinations of the intervals of the variables α_w which are type i's children. For each combination, we are able to write down a set of linear inequalities such that it is a sufficient condition for the functional inequality to hold within that region.

To capture for which sub-intervals should (54) hold, we say that a tuple of indexes $k = (k_0, k_1, k_2, \ldots, k_{\Delta_i})$ is i-acceptable if the interval $[X_{k_0}, Y_{k_0}]$ intersects the interval $\left[\frac{1}{1+\lambda\prod_{j=1}^{\Delta_i} Y_{k_j}}, \frac{1}{1+\lambda\prod_{j=1}^{\Delta_i} X_{k_j}}\right]$. We have the following theorem.

Theorem 73. *In order for the functional inequality* (48) *to hold, it is enough for the following set of linear constraints (a's and b's are the variables) to be feasible:*

For each $i \in [t]$ and each i-acceptable tuple k

$$(1 - X_{k_0}) \sum_{j=1}^{\Delta_i} \left(b_{t(j),k_j} - a_{t(j),k_j} X_{k_j}\right) < \left(b_{i,k_0} - a_{i,k_0} Y_{k_0}\right), \tag{55}$$

where $\{t(j) : j = 1, \ldots, \Delta_i\} = M_i$ (as multisets)

For each $i \in t$ and $k = 0, \ldots, d-1$

$$b_{i,k} - a_{i,k} Y_k > 0, \qquad 0 \le a_{i,k} \le M \qquad 0 \le b_{i,k} \le M. \tag{56}$$

where M is some (big) constant.

Proof. Define $\Psi_i(x) = b_{i,k} - a_{i,k}x$ for all $x \in [X_k, Y_k)$. Linear constraints (56) imply that Ψ_i is non-negative and bounded. Thus it is enough to show (54) holds.

Now fix type i we have k_w's such that $\alpha_w \in [X_{k_w}, Y_{k_w}]$ for each $w \in \mathbf{M}_i$. Let $\alpha_i = 1/(1 + \lambda \prod_{w \in \mathbf{M}_i} \alpha_w)$, then

$$\frac{1}{1 + \lambda \prod_{w \in \mathbf{M}_i} Y_{k_w}} \le \alpha_i \le \frac{1}{1 + \lambda \prod_{w \in \mathbf{M}_i} X_{k_w}}.$$

Thus if k_i is such that $X_{k_i} \le \alpha_i \le Y_{k_i}$ then the tuple $k = (k_i, k_{w_1}, \ldots, k_{w_{\Delta_i}})$ is i-acceptable.

Therefore,

$$\begin{aligned}
(1 - \alpha_i) \sum_{w \in \mathbf{M}_i} \Psi_{t(w)}(\alpha_w) &= (1 - \alpha_i) \sum_{w \in \mathbf{M}_i} \left(b_{t(w),k_w} - a_{t(w),k_w}(\alpha_w)\right) \\
&\le (1 - X_{k_i}) \sum_{j=1}^{\Delta_i} \left(b_{t(w),k_w} - a_{t(w),k_w} X_{k_w}\right) \\
&< b_{i,k_i} - a_{i,k_i} Y_{k_i} && \text{from (55)} \\
&\le \Psi_i(\alpha_i).
\end{aligned}$$

□

5.4 Hard-core Model on the Square Lattice $\mathbb{Z}^2$

In this section, we show how to construct the branch matrices for the two-dimensional integer lattice $\mathbb{Z}^2$ and then apply the previous two conditions to improve the range of λ where the strong spatial mixing holds, which gives the following theorem.

Theorem 74. *There exists a $t \times t$ matrix $\mathbf{M}$ such that $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\leq \mathbf{M}}$ and the strong spatial mixing holds for $\mathcal{F}_{\leq \mathbf{M}}$ when $\lambda < \lambda^* = 2.48$.*

By applying Theorem 63, from the above strong spatial mixing result we prove Theorem 1.

At the end of this section, we will discuss about the obstacles and potential limitations of using the tree recursion methods for showing the uniqueness on graphs.

5.4.1 Lower bound of the Uniqueness Threshold for Hard-core Model on $\mathbb{Z}^2$.

Because of the regularity of $\mathbb{Z}^2$, we could use branching matrices $\mathbf{M}$ such that the tree $T_{\mathbf{M}}$ consist of all walks of $\mathbb{Z}^2$ truncated when closing a cycle of length less than or equal to a certain constant. Clearly, $T_{\mathbf{M}}$ is a super-tree of $T_{\text{saw}}(\mathbb{Z}^2)$, because any path in $T_{\mathbf{M}}$ will only avoid cycles of a certain length whereas paths $T_{\text{saw}}(\mathbb{Z}^2)$ are avoiding all cycles. We first illustrate our approach by showing that Theorem 74 holds with $\lambda^* = 1.8801$ for a simple choice of $\mathbf{M}$, denoted as $\widehat{\mathbf{M}}$. We then explain how to extend the approach to higher λ.

Avoiding Cycles of Length Four

The graph $\mathbb{Z}^2$ is translation-invariant, hence the tree $T_{\text{saw}}(\mathbb{Z}^2, v)$ is identical for every vertex $v \in \mathbb{Z}^2$. Fix a vertex, call it the origin $\boldsymbol{o}$, and let us consider $T_{\text{saw}}(\mathbb{Z}^2) = T_{\text{saw}}(\mathbb{Z}^2, \emptyset)$. Each path from the root of $T_{\text{saw}}(\mathbb{Z}^2)$ corresponds to a self-avoiding walk in $\mathbb{Z}^2$ starting at the origin. Any walk on $\mathbb{Z}^2$ starting at the origin $\boldsymbol{o}$ can be encoded as a string over the alphabet $\{N, E, S, W\}$ corresponding to North, East, South and West. The tree $T_{\text{saw}}(\mathbb{Z}^2)$ contains such strings, truncated the first time the corresponding walk completes a cycle. A relaxed notion of such a tree would be to truncate a walk only when a 4-cycle is completed. Denote such a tree by T_4, and clearly we have that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of T_4. Our first idea is to define a branching matrix $\widehat{\mathbf{M}}$ so that $T_4 \in \mathcal{F}_{\leq \widehat{\mathbf{M}}}$, and hence $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\leq \widehat{\mathbf{M}}}$.

To avoid cycles of length four, it is enough to track the last three steps of the walks. Labeling the paths using $\{N, E, S, W\}$ as mentioned above, their branching rule is easily determined. For example, a path labeled NWW is followed by paths labeled WWS, WWN and WWW which corresponds to adding the directions S, N and W respectively. As another example, a path labeled NWS is followed by paths labeled WSW and WSS corresponding to adding the directions W and S to the path, while adding the direction E would have resulted in a cycle of length 4. The number of types in the corresponding branching matrix is $\leq 4 + 4^2 + 4^3 \leq 5^3$. Indeed, we can reduce the representation of such paths by using isomorphisms between the generating rules among them. This results in 4 types in the following branching matrix $\widehat{\mathbf{M}}$:

$$\widehat{\mathbf{M}} = \begin{pmatrix} 0 & 4 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \tag{57}$$

where the type $i = 0, ..., 3$ of a vertex (walk) in the tree represents the fact that a continuation with a minimum of $4 - i$ additional edges are needed to complete a cycle of length 4.

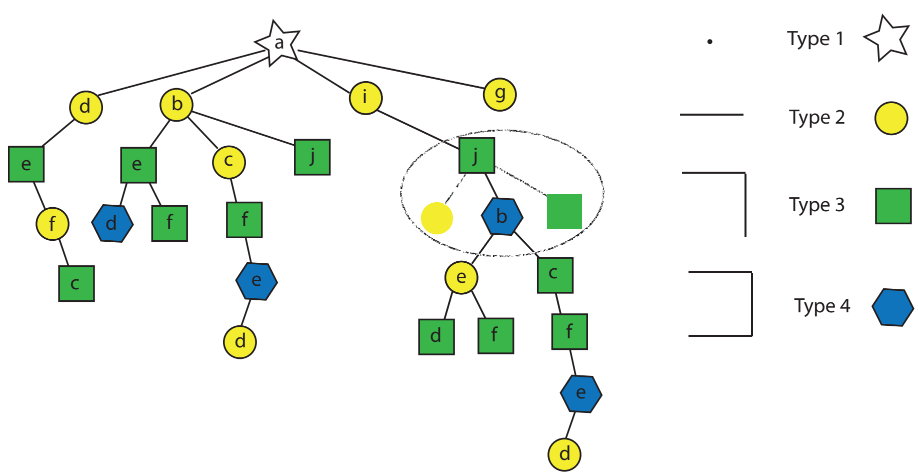

Figure 4: Assignment of the four types from matrix $\widehat{\mathbf{M}}$ defined in (58) to the self-avoiding walk tree T_{saw} from Figure 3. In the circled area, we also draw redundant leaves at vertex j which may appear in the branching rule, but not in T_{saw}.

See Figure 4 for an illustration of this branching matrix $\widehat{\mathbf{M}}$. One can verify that this branching matrix captures, inter alia, the self-avoiding walk trees from $\mathbb{Z}^2$:

Observation 75. *For any finite subgraph $G = (V, E)$ of $\mathbb{Z}^2$ and $v \in V$, $T_{\text{saw}}(G, v) \in \mathcal{F}_{\leq \widehat{\mathbf{M}}}$.*

For this branching matrix, one can check that the (DMS) condition of Theorem 72 holds with $\lambda^* = 1.8801$, $\boldsymbol{S} = \mathbf{Diag}(1.040, 1.388, 1.353, 1.255)$ and $\boldsymbol{c} = (0.266037, 0.100891, 0.100115, 0.0973861)$. As a consequence, we can conclude that Theorem 74 and hence Theorem 1 holds for $\widehat{\mathbf{M}}$ and $\lambda^* = 1.8801$. One can also apply the linear programming condition to the matrix, this will result in $\lambda^* = 1.92$.

Generalization of the Matrix $\widehat{\mathbf{M}}$

The primary reason why the branching matrix $\widehat{\mathbf{M}}$ improves beyond the tree-threshold of $\lambda < \lambda_c(\mathbb{T}_4) = 27/16 = 1.6875$ is that the average branching factor of any $T \in \mathcal{F}_{\leq \widehat{\mathbf{M}}}$ is significantly smaller than that of the regular tree of degree 4.

In $\widehat{\mathbf{M}}$, we have not yet taken into consideration the effect of the assignments of closing a cycle in the walks. When we do that, we are able to construct much more complex branching matrices providing better bounds. The trees generated by these matrices become tighter than $\widehat{\mathbf{M}}$ as a super tree of $T_{\text{saw}}(\mathbb{Z}^2)$ since when a walk closes a cycle with an occupied assignment to a vertex u, this forces the parent of u to be unoccupied, which further trims down the size of the tree.

Starting with T_4, prune the leaves as is done in the construction of $T_{\text{saw}}(\mathbb{Z}^2)$, i.e., including the effect of the boundary conditions. Denote the new tree as T_4'. Clearly we still have that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of T_4'. Let us illustrate the difference between T_4 and the pruned tree T_4'. We first fix an underlying order for the neighbors of each vertex. To this end, say $N > E > S > W$ and this prescribes an ordering of the neighbors of each vertex. Consider a leaf vertex v' in the tree T_4 corresponding to the vertex v in $\mathbb{Z}^2$ and to the path ρ in $\mathbb{Z}^2$. Since v' is a leaf vertex in T_4, ρ must end with a cycle at v, say $WNES$. Since v was exited in the West direction at the beginning of the 4-cycle, and since $W < N$, the leaf vertex v' would be labeled occupied in Weitz's construction, thus resulting in the removal of v' and its parent in the construction of T'_4. Note, every vertex w' in T_4 of type WNE has a child v' of type NES, and consequently w' (and its subtree) will be removed from the tree in the pruning process to construct T'_4. Thus, after removing vertices of type WNE (and similarly, WSE, SEN and ENW) from T_4, it is still the case that $T_{\text{saw}}(\mathbb{Z}^2)$ is a subtree of the resulting tree (T'_4). This highlights why T'_4 has a significantly smaller average branching factor than T_4.

We can define a branching matrix $\mathbf{M}_2$, with 17 types (as illustrated in Figure 5 and explained in Appendix A), such that $T'_4 \in \mathcal{F}_{\leq \mathbf{M}_2}$, and hence $T_{\text{saw}}(\mathbb{Z}^2) \in \mathcal{F}_{\leq \mathbf{M}_2}$. We can prove the DMS Condition is satisfied for $\mathbf{M}_2$ at $\lambda^* = 2.1625$ and the LP condition is satisfied at $\lambda^* = 2.30$, which significantly improves upon our initial bound resulting from considering T_4.

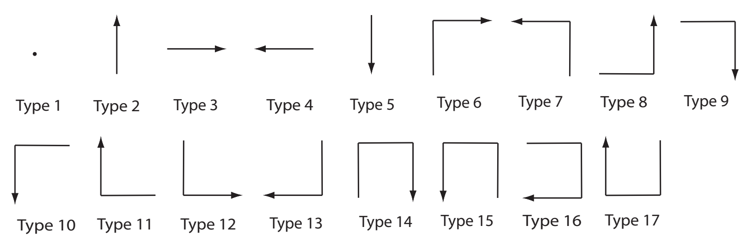

Figure 5: Shapes that the seventeen types (or labels) represent for $\mathbf{M}_2$ where $T'_4 \in \mathcal{F}_{\leq \mathbf{M}_2}$.

A natural direction for improved results is to consider branching matrices corresponding to avoidance of larger cycles, while also accounting for the removal of vertices prescribed by the construction of Weitz. For even $\ell \geq 4$, we use $\mathbf{M}_{\ell/2}$ to denote the branching matrix generating the tree that contains all walks in $\mathbb{Z}^2$ truncated when completing a cycle of length $\leq \ell$, where these leaf vertices are occupied or unoccupied according to Weitz's construction based on some fixed homogeneous ordering $<_w$ of neighbors for every vertex. By taking into account the boundary condition we obtain a smaller tree, however the number of types increases. For $\mathbf{M}_3$ there are 132 types, and for $\mathbf{M}_4$ there are 922 types. Hence the computations become increasingly difficult for larger ℓ. We can make the following general observation from our construction.

Observation 76. *For any finite subgraph $G = (V, E)$ of $\mathbb{Z}^2$ and $v \in V$, $T_{\text{saw}}(G, v) \in \mathcal{F}_{\leq \mathbf{M}_i}$ for any $i \geq 2$.*

Discussions about the DMS condition and the LP condition

The following table summarizes the threshold λ^* we obtain for each $\mathbf{M}_i$ using both the DMS condition and linear programming method:

Table 1: Bounds on λ^* for each branching matrix $\mathbf{M}_i$

Max length of Cycles	Effect of Boundary	# of Types	DMS λ^*	LP λ^*
4	No	4	1.88	1.92
4	Yes	17	2.16	2.30
6	Yes	132	2.33	2.46
8	Yes	922	2.38	2.48

This proves that Theorem 74 holds for any $\lambda < 2.48$. The complete data for results in the above table is given in the Online Appendix at http://www.cc.gatech.edu/~vigoda/hardcore.html. Note that one can further improve the bound on λ by using more types for higher i and hence Theorem 74 will hold with the corresponding activity λ^*. However, the number of types grows exponentially in the length of cycles we avoid, which stops us from obtaining better results. In Section 5.4.3, we will provide a numerical method to reduce the number of types in the matrix significantly, where we also propose an interesting conjecture for reducing the number of types.

For the DMS condition, fast algorithms such as binary search can be use to find the parameters due to the concavity of the candidate functions. Checking the DMS condition for a given choice of parameters would have been a straightforward task, were it not for the irrationality of the coefficients D_{jj}. However, one can establish rigorous upper bounds for D_{jj}, based on concavity of the function (of α) used in the definition of D_{jj}, in a suitable range of the parameters. These details will be discussed further in Section 5.4.2.

As we can see, solving and checking the parameters for the linear programming condition requires substantially larger amount of computing resource, but it does provides a better bound and there is no irrational number issue involved. This does not mean that the DMS condition has no value at all. Indeed, from the comparison between them, we are able to see how much we gain by using a near "optimal" method and understand that it seems impossible to have a simple close form for the optimal statistics φ_i unlike in the one-type case.

One may also wonder how much we are able to prove about λ^* if we are able to do the computation for avoiding big cycles. We will then discuss about limitations of our method by showing several examples in Section 5.4.4 and Section 5.4.5, which potentially implies that the results obtained by the current methods cannot be improved too much to get close to the conjectured threshold $\lambda_c(\mathbb{Z}^2)$.

5.4.2 Verification of the DMS Condition

For any such matrix $\mathbf{M}_i$, the verification of the DMS Condition relies on (i) 'guessing' appropriate values for the parameters $\boldsymbol{S}$ and $\boldsymbol{c}$ and (ii) formally verifying that DMS Condition holds for the chosen $\boldsymbol{S}$ and $\boldsymbol{c}$. In choosing desirable $\boldsymbol{S}$ and $\boldsymbol{c}$, we employed a heuristic random walk algorithm aided with binary search.

To verify that the DMS Condition holds for a given *rational* matrix $\boldsymbol{S}$ and vector $\boldsymbol{c}$ is straightforward, provided we can obtain a rational upper bound for each type j for the function:

$$f_j(\alpha) = \frac{(1-\alpha)\left(1-\theta_j\left(\frac{1-\alpha}{\lambda\alpha}\right)^{1/\Delta_j}\right)}{s_j-\alpha}.$$

Indeed, due to the concavity of this function for $0 < \theta_j \le 1$, $s_j > 51/50$ and $\lambda > 27/16$,

[2] it is always possible to find a *provable* upper bound for f_j in such a regime. This can be done, for example, by describing a suitable 'envelope' for f_j consisting of a piecewise linear function of the form:

$$g_j(\alpha) = \begin{cases} B_\ell & \text{if } \alpha < \alpha_\ell \\ \min\{b_\ell(\alpha - \alpha_\ell) + B_\ell, b_u(\alpha - \alpha_u) + B_u\} & \text{if } \alpha_\ell < \alpha < \alpha_u \\ B_u & \text{if } \alpha > \alpha_u \end{cases}$$

where α_ℓ, α_u are points such that $b_\ell > f_j'(\alpha_\ell) > 0$, $b_u < f_j'(\alpha_u) < 0$, $B_\ell > f_j(\alpha_\ell)$ and $B_u > f_j(\alpha_u)$. It is clear for any such function that $g_j(\alpha) > f_j(\alpha)$, thus we obtain a provable upper bound for f_j using g_j.

For every $\mathbf{M}_i$ in the above table, we provide $\boldsymbol{S}$ and $\boldsymbol{c}$, along with appropriate envelopes that lead to upper bounds $\hat{D}_{jj}$ for the corresponding D_{jj}. Then we verify that the DMS Condition holds for the given values of λ by replacing D_{jj} with $\hat{D}_{jj}$. For $i = 2$, see Appendix A for the matrix $\mathbf{M}_2$ and the corresponding $\boldsymbol{S}$, $\boldsymbol{c}$, α_ℓ and α_u. For $i = 3, 4$ these values ($\mathbf{M}$, $\boldsymbol{S}$, $\boldsymbol{c}$, α_ℓ and α_u) are given in the Online Appendix at http://www.cc.gatech.edu/∼vigoda/hardcore.html.

5.4.3 Reduction of the Branching Matrices $\mathbf{M}_\ell$

Usually, when one applies various methods trying to solve the functional inequality (48), one has to face the fact that the dimension of the matrix $\mathbf{M}_i$ is huge, e.g., $t = 922$ for $i = 4$. A natural way to generate $\mathbf{M}$ is using a DFS program that enumerates all the types by remembering the history of the self avoiding walk. However, there are many types in such a matrix that are essentially the "same". Here we provide a heuristic method for finding those types that are the same.

Let $\mathcal{C}$ be a partition of the types in $\mathbf{M}$, i.e., $\mathcal{C} = \{C_1, C_2, ..., C_k\}$ such that $\biguplus_{i=1}^k C_i = [t]$. We define the partition to be consistent with $\mathbf{M}$, if for every $i \in [k]$,

[2] This is a nontrivial algebraic fact. It can be proved by transforming the second derivatives condition to a set of integer polynomial constraints and using the "**resolve**" function in MATHEMATICA for the satisfiability of the constraints, which is rigorous by the Tarski-Seidenberg Theorem [69] for the real polynomial systems [1] and the so-called cylindrical algebraic decomposition [5].

each pair of types $s, t \in C_i$, the rows $\mathbf{M}_s$ and $\mathbf{M}_t$ are the same with respect to $\mathcal{C}$, that is

$$\sum_{j \in C_{i'}} \mathbf{M}_{sj} = \sum_{j \in C_{i'}} \mathbf{M}_{tj}, \text{ for all } i' \in [k].$$

Definition 77. *Given* $\mathbf{M}$ *and a partition* $\mathcal{C}$ *of size* k *which is consistent, we define the* k*-by-*k *matrix* $\mathbf{M}^{\mathcal{C}}$ *by,*

$$\mathbf{M}^{\mathcal{C}}_{ii'} = \sum_{j \in C_{i'}} \mathbf{M}_{sj} \textit{ where } s \in C_i.$$

We say $\mathbf{M}$ *is* reducible *to a* k*-by-*k *matrix* $\mathbf{B}$ *if there is a consistent partition* $\mathcal{C}$ *such that* $\mathbf{B} = \mathbf{M}^{\mathcal{C}}$.

Lemma 78. *For a partition* $\mathcal{C}$ *of size* k*,*

$$\mathcal{F}_{\leq \mathbf{M}^{\mathcal{C}}} = \mathcal{F}_{\leq \mathbf{M}} \quad \textit{and} \quad T_{\mathbf{M}^{\mathcal{C}}} = T_{\mathbf{M}}.$$

Proof. The argument is just a standard induction on the tree. □

Now the question is how to find a good partition $\mathcal{C}$ easily. For a specific value $\lambda < \mathrm{WSM}(T_{\mathbf{M}})$, Let V_λ be the fixed points of the recurrences of the marginal distributions defined by $\mathbf{M}$. Our conjecture is

Conjecture 79. *Let the partition* $\mathcal{C}(\lambda)$ *be the sets of types that have the same value of the fixed points in* V_λ*, i.e., for each* $C_i \in \mathcal{C}(\lambda)$*, for all* $c \in C_i$*,* $V_\lambda(c)$ *are the same. If for all* λ*, the partitions* $\mathcal{C}(\lambda)$ *are identical, then* $\mathcal{C}$ *is a partition that is consistent of* $\mathbf{M}$.

Using the intuition from Conjecture 79 we are able to find good partitions in practice. We simply run a dynamic programming algorithm on the tree $T_{\mathbf{M}}$ to calculate an approximation of the fixed points in V_λ. Once the approximation is good enough, we simply make the partition according to this approximation. We then check the consistency of the partition with $\mathbf{M}$, and therefore, we know whether the resulting matrix generates the same tree as the original one or not by Lemma 78. Applying this reduction to $\mathbf{M}_3$, the number of types goes down from 132 to 34, and for $\mathbf{M}_4$ the number of types goes down from 922 to 162. This significant reduction in the size of the matrices greatly reduces the number of constraints and variables in our linear programming formulation.

5.4.4 Upper Bound on the SSM Threshold for $\mathbb{Z}^2$

As we know, previous approaches for lower bounding $\lambda_c(\mathbb{Z}^2)$ work by establishing WSM on $\mathbb{Z}^2$ by proving SSM on $T_{\text{saw}}(\mathbb{Z}^2)$. Hence, to provide an upper bound on these approaches we want to upper bound the SSM threshold for $T_{\text{saw}}(\mathbb{Z}^2)$. We will show that for $\lambda > 3$ that SSM does not hold on $T_{\text{saw}}(\mathbb{Z}^2)$ obtained from any homogeneous ordering. Note that this does not imply anything about WSM/SSM on $\mathbb{Z}^2$, it simply gives a limit for the current proof approaches.

To prove that SSM does not hold on $T_{\text{saw}}(\mathbb{Z}^2)$ we define a tree T^* that is a subtree of $T_{\text{saw}}(\mathbb{Z}^2)$ and prove that WSM does not hold on T^* when $\lambda > 3$.

The tree T^* is easy to construct. We consider those walks on $\mathbb{Z}^2$ that only go North, East and West. The branching rules can be written in the following finite state machine way:

1. $N \to N \mid E \mid W$
2. $E \to N \mid E$
3. $W \to N \mid W$

We The branching matrix corresponding to the above rule is

$$\boldsymbol{D} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \tag{58}$$

where row/column 1,2 and 3 corresponding to North, East and West respectively.

Proposition 80. *The tree $T_{\boldsymbol{D}}$ generated by the branching matrix $\boldsymbol{D}$ is a sub-tree of $T_{\text{saw}}(\mathbb{Z}^2)$.*

Proof. We will define an ordering $>_w$ for each $w \in \mathbb{Z}^2$ such that $T_{\boldsymbol{D}}$ is a subtree of T_{saw} in a natural way. In Weitz construction, let $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$ be the tree of self-avoiding walks of $\mathbb{Z}^2$ originating from the origin, including the vertices closing a cycle in the walks as leaves. The tree $T_{\boldsymbol{D}}$ consists of all those self-avoiding walks that never go South, and thus, it is a sub-tree of $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$.

Now, in the second part of Weitz construction, some vertices are delete from $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$ to obtain $T_{\text{saw}}(\mathbb{Z}^2)$. We need to show that not vertex from $T_{\boldsymbol{D}}$ is deleted. A vertex is deleted from $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$ because is an occupied leaf, or is the parent of an occupied leaf. Leaves in $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$ correspond to walks finishing in a cycle, and thus they do not belong to $T_{\boldsymbol{D}}$. Now suppose a vertex in $T_{\boldsymbol{D}}$ is the parent of a leave ζ in $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$. In this case, ζ corresponds to a path finishing in a cycle in $\mathbb{Z}^2$, say $\cdots \to w \to v_1 \to \ldots v_\ell \to w$, and the move $v_\ell \to w$ is a South move. To ensure that we do not occupy ζ in $\widehat{T_{\text{saw}}}(\mathbb{Z}^2)$, we need to ensure that $v_1 >_w v_\ell$, and for this we just need to define for each vertex w in $\mathbb{Z}^2$, the North children as the smallest one in the $>_w$ - ordering. □

Note that, this proof works for the trees $T_{\text{saw}}(\mathbb{Z}^2)$ defined under any homogeneous ordering. In any such $T_{\text{saw}}(\mathbb{Z}^2)$, it is easy to find a tree that is isomorphic to $T_{\boldsymbol{D}}$. For this tree $T_{\boldsymbol{D}}$ we can establish its WSM threshold as stated in the following result, which immediately shows that SSM does not hold for $T_{\text{saw}}(\mathbb{Z}^2)$ when $\lambda > 3$.

Lemma 81. *The weak spatial mixing threshold for the tree $T_{\boldsymbol{D}}$ defined by $\boldsymbol{D}$ is at $\lambda = 3$.*

Proof. The matrix $\boldsymbol{D}$ is reducible, as defined in Definition 77, to the following 2×2

matrix:

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}. \tag{59}$$

It is easy to check that $\mathbf{B}$ and $\boldsymbol{D}$ generate the same family of trees. Now the recurrences for the marginal distributions of both types derived from (42) are

$$f(x,y) = \left(\frac{1}{1+\lambda x y^2}, \frac{1}{1+\lambda x y} \right). \tag{60}$$

Using some algebra, we are able to determine the fixed points of $f(x,y)$ for $\lambda > 1$

$$(x_0, y_0) = (x_0(\lambda), y_0(\lambda)) = \left(\frac{4\lambda + \sqrt{8\lambda+1} - 1}{8\lambda}, \frac{\sqrt{8\lambda+1} - 3}{2(\lambda - 1)} \right).$$

We just need to check the eigenvalues of the Jacobian of the recurrences at the fixed point, see e.g., [63]: If the largest eigenvalue is greater than 1, then the function around the the fixed point is repelling and hence it is impossible for the boundary conditions to converge to this unique fixed point. If the largest eigenvalue is strictly less than 1, the function is contracting to the fixed point in its neighborhood. The Jacobian at the fixed point (x_0, y_0) is the following:

$$J(\lambda) = \begin{pmatrix} \lambda x_0^2 y_0^2 & 2\lambda y_0 x_0^3 \\ \lambda y_0^3 & \lambda x_0 y_0^2 \end{pmatrix}. \tag{61}$$

Denote the trace of $J(\lambda)$ as $\mathrm{tr}(J(\lambda)) = \lambda x_0 y_0^2 (x_0 + 1)$ and its determinant as $\det(J(\lambda)) = -\lambda^2 x_0^3 y_0^4$. The largest eigenvalue of $J(\lambda)$ is then

$$\rho(\lambda) = \frac{\mathrm{tr}(J(\lambda))}{2} + \left(\frac{\mathrm{tr}(J(\lambda))^2}{4} - \det(J(\lambda)) \right)^{1/2},$$

It is easy to check for $\lambda = 3$, $x_0(3) = 2/3$, $y_0(3) = 1/2$ and $\rho(3) = 1$. $\square$

5.4.5 Tree with Different Thresholds for SSM and WSM

Brightwell *et al.* [12] give an example of hard-core model on trees that the WSM does not imply the SSM for the same activity λ. Here, we present another example using sub-trees of $T_{\text{saw}}(\mathbb{Z}^2)$. We can show a tree T' which is a supertree $T_{\boldsymbol{D}}$ for which WSM holds for some $\lambda > 3$.

To construct the tree T' we allow some South moves in the tree in a certain context. In particular, we only allow that a South move happens when the path contains the following substring: NNEESEN, i.e., a South move is allowed if and only if it is after a sequence of NNEE moves and followed by another EN moves. The tree family can be formalized in the following finite state machine way:

1. $E \to E \mid N$
2. $W \to N \mid W$
3. $N \to NN \mid E \mid W$
4. $NN \to NN \mid NNE \mid W$
5. $NNE \to N \mid NNEE$
6. $NNEE \to N \mid E \mid NNEES$
7. $NNEES \to NNEESE$
8. $NNEESE \to NNEESEN$
9. $NNEESEN \to NN \mid E$

Let the matrix describing the above rules be denoted as $\boldsymbol{D}_1$.

Lemma 82. *The tree $T_{\boldsymbol{D}_1}$ that is generated by $\boldsymbol{D}_1$ is a super-tree of $T_{\boldsymbol{D}}$ and is a sub-tree of $T_{\text{saw}}(\mathbb{Z}^2)$.*

The proof of the above proposition is similar to Theorem 81 except for that one has to make the ordering more carefully so that everything works out. One can numerically verify that the WSM threshold for the tree generated by $\boldsymbol{D}_1$ is above $\lambda = 3.01$.

Lemma 83.

$$WSM(T_{\boldsymbol{D}_1}) > 3.01.$$

The rigorous proof is tedious but doable by straightforward calculations using Mathematica.

APPENDIX A

BRANCHING MATRIX OF 17 TYPES

As we explained in Section 5.4.1, for constructing $\mathbf{M}_2$ (i.e., the branching matrix that represents the structure of trees of walks in $\mathbb{Z}^2$, avoiding cycles of length 4 accounting removals of vertices), it is enough to track the last three steps of self-avoiding walks. It is straightforward to verify that under the assignments of labels (or types) to walks as explained below, it is the case that the tree T_4' (cf. Section 5.4.1 for its definition) is contained in the family $\mathcal{F}_{\leq \mathbf{M}_2}$ where $\mathbf{M}_2$ is given as per (62).

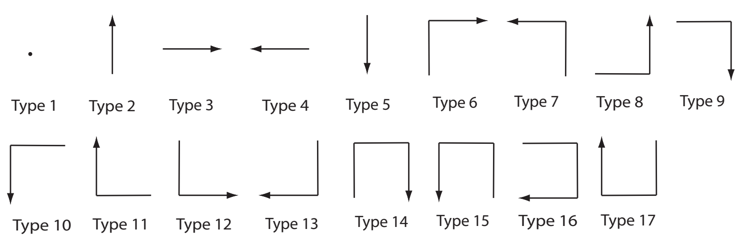

Figure 6: Shapes that the seventeen types (or labels) represent for $\mathbf{M}_2$.

- If the terminal edges of the path are NES (NWS, ESW, SWN), then assign to it the label 14 (15, 16, 17, respectively).
- Else if the terminal edges are NE (NW, EN, ES, WS, WN, SE, SW), then assign to it the label 6 (7, 8, 9, 10, 11, 12, 13 respectively).
- Else if the terminal edges are N (E, W, S), then assign to it the label 2 (3, 4, 5 respectively).
- Otherwise (necessarily a path of length 0), assign the label 1.

$$
\mathbf{M}_2 = \begin{pmatrix}
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix} \tag{62}
$$

Hence, for any finite subgraph $G = (V, E)$ of $\mathbb{Z}^2$, $T_{\text{saw}}(G, v)$ is contained in the family $\mathcal{F}_{\leq \mathbf{M}_2}$. Setting $\boldsymbol{S}, \boldsymbol{c}, (\alpha_\ell, \alpha_u)$ as detailed below, and proceeding as described in Section 5.4.1, we conclude that the DMS condition holds for $\lambda^* = 2.1625$. Therefore, for all $\lambda \leq 2.1625$, SSM holds with uniform rate $\gamma \in (0, 1)$ for subgraphs of $\mathbb{Z}^2$.

$$(\alpha_\ell, \alpha_u) = \begin{pmatrix} 0.829647,\ 0.829704 \\ 0.579816,\ 0.579891 \\ 0.606359,\ 0.606453 \\ 0.638897,\ 0.638974 \\ 0.597921,\ 0.598001 \\ 0.627943,\ 0.627993 \\ 0.629560,\ 0.629638 \\ 0.635604,\ 0.635656 \\ 0.622289,\ 0.622374 \\ 0.638467,\ 0.638555 \\ 0.634357,\ 0.634417 \\ 0.637607,\ 0.637659 \\ 0.623116,\ 0.623202 \\ 0.658641,\ 0.658723 \\ 0.643530,\ 0.643595 \\ 0.666218,\ 0.666285 \\ 0.636866,\ 0.636945 \end{pmatrix} \qquad \boldsymbol{c} = \begin{pmatrix} 0.287690 \\ 0.101247 \\ 0.101683 \\ 0.102709 \\ 0.098046 \\ 0.100724 \\ 0.096228 \\ 0.105722 \\ 0.096053 \\ 0.101926 \\ 0.104926 \\ 0.100006 \\ 0.092699 \\ 0.100489 \\ 0.104415 \\ 0.097162 \\ 0.106246 \end{pmatrix}$$

$$\boldsymbol{S} = \mathbf{Diag}(1.040, 1.494, 1.432, 1.363, 1.45, 1.392, 1.412, 1.282, 1.407,$$
$$1.271, 1.28, 1.288, 1.442, 1.25, 1.256, 1.254, 1.272).$$

REFERENCES

[1] "Real polynomial systems." http://reference.wolfram.com/mathematica/tutorial/RealPolynomialSystems.html.

[2] Achlioptas, D. and Coja-Oghlan, A., "Algorithmic barriers from phase transitions," in *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, pp. 793–802, IEEE Computer Society, 2008.

[3] Aldous, D., "Random walks on finite groups and rapidly mixing Markov chains," in *Seminaire de Probabilites XVII, Lecture Notes in Math. 986*, pp. 243–297, SV, 1983.

[4] Alm, S. E., "Upper bounds for the connective constant of self-avoiding walks," *Combinatorics, Probability and Computing*, vol. 4, pp. 115–136, 1993.

[5] Arnon, D. S., Collins, G. E., and McCallum, S., "Cylindrical algebraic decomposition I: The basic algorithm," *SIAM Journal on Computing*, pp. 865–877, 1984.

[6] Baxter, R. J., Enting, I. G., and Tsang, S. K., "Hard-square lattice gas.," *Journal of Statistical Physics*, vol. 22, pp. 465–489, 1980.

[7] Beffara, V. and Duminil-Copin, H., "The self-dual point of the two-dimensional random-cluster model is critical for $q \geq 1$," *Probability Theory and Related Fields*, vol. 153, pp. 511–542, 2012.

[8] Berger, N., Kenyon, C., Mossel, E., and Peres, Y., "Glauber dynamics on trees and hyperbolic graphs," *Probability Theory and Related Fields*, vol. 131, pp. 311–340, 2005.

[9] Bhatnagar, N., Sly, A., and Tetali, P., "Reconstruction threshold for the hardcore model," in *Proceedings of the 13th International Conference on Approximation, and 14 the International Conference on Randomization, and Combinatorial Optimization: Algorithms and Techniques*, pp. 434–447, Springer-Verlag, 2010.

[10] Bhatnagar, N., Vera, J., Vigoda, E., and Weitz, D., "Reconstruction for colorings on trees," *SIAM Journal on Discrete Mathematics*, vol. 25, no. 2, pp. 809–826, 2011.

[11] Blanca, A., Galvin, D., Randall, D., and Tetali, P., "Phase coexistence and slow mixing for the hard-core model on $\mathbb{Z}^2$," *http://arxiv.org/abs/1211.6182.*

[12] BRIGHTWELL, G. R., HÄGGSTRÖM, O., and WINKLER, P., "Nonmonotonic behavior in hard-core and Widom-Rowlinson models," *Journal of Statistical Physics*, vol. 94, no. 3, pp. 415–435, 1999.

[13] BRIGHTWELL, G. R. and WINKLER, P., "A second threshold for the hard-core model on a Bethe lattice," *Random Structures and Algorithms*, vol. 24, no. 3, pp. 303–314, 2004.

[14] CAVENDER, J. A., "Taxonomy with confidence," *Mathematical Biosciences*, vol. 40, no. 3-4, pp. 271–280, 1978.

[15] CESI, F., "Quasi–factorization of the entropy and logarithmic sobolev inequalities for gibbs random fields," *Probability Theory and Related Fields*, vol. 120, no. 4, pp. 569–584, 2001.

[16] DIACONIS, P. and SALOFF-COSTE, L., "Logarithmic Sobolev inequalities for Markov chains," *Annals of Applied Probability*, vol. 6, no. 3, pp. 695–750, 1996.

[17] DING, J., LUBETZKY, E., and Y.PERES, "Mixing time of critical Ising model on trees is polynomial in the height," *Communications in Mathematical Physics*, vol. 295, no. 1, pp. 161–207, 2010.

[18] DOBRUSHIN, R. L., "The problem of uniqueness of a gibbsian random field and the problem of phase transitions," *Functional Analysis and its Applications*, vol. 2, no. 4, pp. 302–312, 1968.

[19] DOBRUSHIN, R. L. and SHLOSMAN, S. B., "Constructive unicity criterion," in *Statistical Mechanics and Dynamical Systems*, pp. 347–370, Birkhäuser, New York, 1985.

[20] DUBHASHI, D. and RANJAN., D., "Balls and bins: A study in negative dependence," *Random Strctures and Algorithms*, vol. 13, pp. 99–124, 1998.

[21] DYER, M. E. and FRIEZE, A., "Randomly coloring graphs with lower bounds on girth and maximum degree," *Random Strctures and Algorithms*, vol. 23, pp. 167–179, 2003.

[22] DYER, M. E., GOLDBERG, L., JERRUM, M., and MARTIN, R., "Markov chain comparison," *Probability Surveys*, pp. 89–111, 2006.

[23] DYER, M. E., SINCLAIR, A., VIGODA, E., and WEITZ, D., "Mixing in time and space for lattice spin systems: A combinatorial view," *Random Structures and Algorithms*, vol. 24, no. 4, pp. 461–479, 2004.

[24] DYER, M. E., FRIEZE, A., HAYES, T. P., and VIGODA, E., "Randomly coloring constant degree graphs," in *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, (Washington, DC, USA), pp. 582–589, IEEE Computer Society, 2004.

[25] Galanis, A., Stefankovic, D., and Vigoda, E., "Inapproximability of the partition function for the antiferromagnetic Ising and hard-core models," *http://arxiv.org/abs/1203.2226*.

[26] Gaunt, D. S. and Fisher, M. E., "Hard-sphere lattice gases. I. Plane-square lattice.," *Journal of Chemical Physics*, vol. 43, no. 8, pp. 2840–2863, 1965.

[27] Georgii, H. O., *Gibbs Measures and Phase Transitions.* De Gruyter Studies in Mathematics, 9, De Gruyter, 1988.

[28] Gerschenfeld, A. and Montanari, A., "Reconstruction for models on random graphs," in *Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science*, pp. 194–204, IEEE Computer Society, 2007.

[29] Goldberg, L., Jerrum, M., and Karpinski, M., "The mixing time of Glauber dynamics for coloring regular trees," *Random Structures and Algorithms*, vol. 36, no. 4, pp. 464–476, 2010.

[30] Goldberg, L., Martin, R., and Paterson, M., "Strong spatial mixing for lattice graphs with fewer colours," *SIAM Journal on Computing*, vol. 35, no. 2, pp. 486–517, 2005.

[31] Hayes, T. P. and Vigoda, E., "Coupling with the stationary distribution and improved sampling for colorings and independent sets," *Annals of Applied Probability*, vol. 16, pp. 1297–1318, 2006.

[32] Hayes, T. P., Vera, J. C., and Vigoda, E., "Randomly coloring planar graphs with fewer colors than the maximum degree," in *Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pp. 450–458, ACM, 2007.

[33] Ising, E., "Beitrag zur theorie des ferromagnetismus," *Zeitschrift für Physik*, vol. 31, pp. 253–258, 1925.

[34] Jerrum, M., *Counting, Sampling and Integrating: Algorithms and Complexity.* Lectures in mathematics, Birkhäuser, 2003.

[35] Jerrum, M., "A very simple algorithm for estimating the number of k-colorings of a low-degree graph," *Random Structures and Algorithms*, vol. 7, no. 2, pp. 157–166, 1995.

[36] Jerrum, M., Valiant, L. G., and Vazirani, V. V., "Random generation of combinatorial structures from a uniform distribution," *Theoretical Computer Science*, vol. 43, pp. 169–188, 1986.

[37] Jonasson, J., "Uniqueness of uniform random colorings of regular trees," *Statistics and Probability Letters*, vol. 57, pp. 243–248, 2002.

[38] Kelly, F., "Loss networks," *Annals of Applied Probability*, vol. 1, pp. 319–378, 1991.

[39] Lawler, G. and Sokal, A., "Bounds on the l^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality," *Transactions of the American Mathematical Society*, vol. 309, pp. 557–580, 1988.

[40] Ledoux, M., "The concentration of measure phenomenon," in *Mathematical Surveys and Monographs (Vol. 89)*, American Mathematical Society, 2001.

[41] Levin, D. A., Peres, Y., and Wilmer, E. L., *Markov chains and mixing times*. American Mathematical Society, 2009.

[42] Lucier, B. and Molloy, M., "The Glauber dynamics for colorings of bounded degree trees," *SIAM Journal on Discrete Mathematics*, vol. 25, pp. 827–853, 2011.

[43] Martinelli, F., "Lectures on Glauber dynamics for discrete spin models," *Lectures on Probability Theory and Statistics (Saint-Flour, 1997), Lecture Notes in Mathematics*, vol. 1717, pp. 93–191, 1998.

[44] Martinelli, F. and Olivieri, E., "Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case," *Communications in Mathematical Physics*, vol. 161, pp. 447–486, 1994.

[45] Martinelli, F. and Olivieri, E., "Approach to equilibrium of Glauber dynamics in the one phase region. II. The general case," *Communications in Mathematical Physics*, vol. 161, pp. 487–514, 1994.

[46] Martinelli, F., Sinclair, A., and D.Weitz, "The Ising model on trees: Boundary conditions and mixing time," *Communications in Mathematical Physics*, vol. 250, no. 2, pp. 301–334, 2004.

[47] Martinelli, F., Sinclair, A., and Weitz, D., "Fast mixing for independent sets, colorings and other models on trees," in *Proceedings of the 15th annual ACM-SIAM symposium on Discrete algorithms*, pp. 456–465, SIAM, 2004.

[48] Mitzenmacher, M. and Upfal, E., *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[49] Montanari, A., Restrepo, R., and Tetali, P., "Reconstruction and clustering in random constraint satisfaction problems," *SIAM Journal on Discrete Mathematics*, vol. 25, pp. 771–808, 2011.

[50] Mossel, E., "Reconstruction on trees: Beating the second eigenvalue," *Annals of Applied Probabability*, vol. 11, pp. 285–300, 2001.

[51] Mossel, E., "Survey: Information flow on trees," Graphs, Morphisms and Statistical Physics. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pp. 155–170, 204.

[52] Mossel, E. and Sly, A., “Exact thresholds for Ising-Gibbs samplers on general graphs,” *Annals of Probability*, vol. 41, no. 1, pp. 294–328, 2013.

[53] Onsager, L., “Crystal statistics. I. A two-dimensional model with an order-disorder transition,” *Physics Review Letters*, vol. 65, pp. 117–149, 1944.

[54] Peres, Y. and Winkler, P., “Can extra updates delay mixing?,” *http://arxiv.org/abs/1112.0603*.

[55] Pönitz, A. and Tittman, P., “Improved upper bounds for self-avoiding walks in $\mathbb{Z}^d$.,” *Electronic Journal of Combinatorics*, vol. 7, p. R21, 2000.

[56] Preston, C. J., *Gibbs states on countable sets.* Cambridge Tracts in Mathematics, No. 68., Cambridge University Press, London, 1974.

[57] Rácz, Z., “Phase boundary of Ising antiferromagnets near $h = h_c$ and $t = 0$: Results from hard-core lattice gas calculations,” *Physical Review B*, vol. 21, pp. 4012–4016, 1980.

[58] Radulescu, D. C., *A computer-assisted proof of uniqueness of phase for the hard-square lattice gas model in two dimensions.* PhD thesis, Rutgers University, New Brunswick, NJ, USA, 1997.

[59] Radulescu, D. C. and Styer., D. F., “The Dobrushin-Shlosman phase uniqueness criterion and applications to hard squares,” *Journal of Statistical Physics*, vol. 49, pp. 281–295, 1987.

[60] Randall, D., “Slow mixing of Glauber dynamics via topological obstructions,” in *Proceedings of the 17th Annual ACM-SIAM symposium on Discrete Algorithm*, (New York, NY, USA), pp. 870–879, ACM, 2006.

[61] Restrepo, R., Shin, J., Tetali, P., Vigoda, E., and Yang, L., “Improved mixing condition on the grid for counting and sampling independent sets,” *Probability Theory and Related Fields*, pp. 1–25, March, 2012.

[62] Restrepo, R., Stefankovic, D., Vera, J. C., Vigoda, E., and Yang, L., “Phase transition for Glauber dynamics for independent sets on regular trees,” in *Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 945–956, SIAM, 2011.

[63] Robinson, C., *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos.* CRC Press, Inc, 1999.

[64] Sinclair, A. and Jerrum, M., “Approximate counting, uniform generation and rapidly mixing Markov chains,” *Information and Computation*, vol. 82, pp. 93–133, 1989.

[65] Sinclair, A., Srivastava, P., and Thurley, M., "Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs," in *Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 941–953, SIAM, 2012.

[66] Sly, A., "Reconstruction of random colourings," *Communications in Mathematical Physics*, vol. 288, pp. 943–961, 2009.

[67] Sly, A., "Computational transition at the uniqueness threshold," in *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pp. 287–296, IEEE Computer Society, 2010.

[68] Sly, A. and Sun, N., "The computational hardness of counting in two-spin models on d-regular graphs," in *Proceedings of the 2012 IEEE 53rd Annual Symposium on Foundations of Computer Science*, pp. 361–369, IEEE Computer Society, 2012.

[69] Tarski, A., *A Decision Method for Elementary Algebra and Geometry.* University of California Press, 1951.

[70] Tetali, P., Vera, J. C., Vigoda, E., and Yang, L., "Phase transition for the mixing time of the Glauber dynamics for coloring regular trees," *Annals of Applied Probability*, vol. 22, no. 6, pp. 2210–2239, 2012.

[71] van den Berg, J. and Ermakov, A., "A new lower bound for the critical probability of site percolation on the square lattice," *Random Structures and Algorithms*, vol. 8, pp. 199–212, 1996.

[72] van den Berg, J. and Steif, J. E., "Percolation and the hard-core lattice gas model," *Stochastic Processes and their Applications*, vol. 49, pp. 179–197, 1994.

[73] Vigoda, E., "Improved bounds for sampling colorings," *Journal of Mathematical Physics*, vol. 41, no. 3, pp. 1555–1569, 2000.

[74] Weitz, D., "Counting independent sets up to the tree threshold," in *Proceedings of the 38th Annual ACM Symposium on Theory of Computing*, pp. 140–149, ACM, 2006.