CS 7535: Markov Chain Monte Carlo Algorithms **Fall 2014**

Lecture 5: Counting independent sets up to the tree threshold

Lecturer: Richard Brooks and Rakshit Trivedi *Date: December 02*

In this lecture, we will discuss the problem of approximately counting/sampling weighted independent sets of a graph G with activity λ, i.e., where the weight of an independent set I is $\lambda^{|I|}$. Specifically, we study a novel analysis presented by Weitz (2006) which provides a deterministic approximation scheme that runs in polynomial time for any graph of maximum degree Δ and $\lambda < \lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$.

It is known that for $\lambda \leq \lambda_c$, the probability that root of a regular tree (with degree Δ) is in the independent set is asymptotically independent of the configurations of leaves far below. The current analysis leverages this independence by focusing on decay of correlations with distance in the above weighted distribution. The algorithmic technique used here draws a comparison between regular trees and general graphs by showing that on any graph of maximum degree Δ, correlations decay at least as fast as they do on the regular tree of degree Δ. The above results also imply fast (polynomial time) mixing Glauber dynamics for the provided λ regime, although this holds true only for graphs that grow sub-exponentially.

In statistical physics, this weighted distribution model for counting/sampling independent sets is referred to as the hard-core model with activity λ. Establishing decay of correlations in distance is of much value to statistical physicists because it means that the Gibbs measure associated with the model is unique, thus yielding a unique macroscopic equilibrium.

5.1 Introduction

5.1.1 Earlier Work, Glauber dynamics and Integer Lattice $\mathbb{Z}^2$

λ (weighting parameter) and Δ (maximum degree) are two parameters in the regime for which various bounds have been proved. Luby and Vigoda (1997) showed that for any $\Delta \geq 4$, it is NP-hard to approximate the sum on weighted distribution of independent sets, for graphs with $\lambda > \frac{c}{\Delta}$.

Dyer and Greenhill (2000) and Vigoda (2001) also provide two important breakthroughs (respectively):

1. There an efficient approximation scheme for uniformly weighted independent sets ($\lambda = 1$) for graphs of $\Delta = 4$.
2. There exists an FPRAS for counting independent sets with $\lambda < \frac{2}{\Delta-2}$.

The above work use Glauber dynamics where in each step a vertex is chosen uniformly at random and its value (occupied or unoccupied) is updated conditioned on the current values of its neighbors. Hayes and Vigoda (2005) shows this Glauber dynamics chain to be rapidly mixing for graphs of girth at least 6 and large degree ($\Delta = \Omega(logn)$).

An important graph of interest for statistical physicist is integer lattice $\mathbb{Z}^2$. The critical activity obtained by simulations for decay of correlations is around 3.79. Van den berg and Ermakov (1996) showed rapid mixing of Glauber dynamics in this regime for $\lambda < 1.255$. This bound on λ was improved by series of results by calculating the Dobrushin-Shlosman condition for larger and larger rectangles and thus obtaining the tighter bound of $\lambda < 1.508$.

5.1.2 Improvements in this work

It is known that any argument that works for general graphs and which establishes decay of correlations with distance as a byproduct is bound to fail for $\lambda > \lambda_c = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^{\Delta}}$. The current work extends the regime of λ shown in Section 5.1.1 to match with the above value of $\lambda \leq \lambda_c$, thus proving the following conjecture proposed by Sokal (2001):

> The limit on λ imposed by the regular tree can be matched, i.e., that the regular tree is the worst-case graph in terms of decay of correlations and that for any graph of maximum degree Δ, correlations decay with distance throughout the regime $\lambda \leq \lambda_c$.

This leads to the development of a deterministic FPTAS for counting independent sets for any graph of maximum degree Δ and $\lambda \leq \lambda_c$. The algorithmic result eliminates the girth and large degree requirements imposed in Hayes and Vigoda (2005).

For uniformly weighted case (i.e. $\lambda = 1$), the current work extends the bound on the maximum degree of graph Δ to 5. This is a tight bound, since there is evidence of the problem being hard for $\Delta > 6$.

For integer lattice $\mathbb{Z}^2$, this work improves on all previously known bounds and extends it to $\lambda < 1.6875$.

5.1.3 Setup

Consider a graph $G = (V, E)$ and activity parameter $\lambda > 0$. We want to count (or sample) independent sets of G where the weight of an independent set $I \subset V$ is $\lambda^{|I|}$.

Let $Z \equiv Z_G^\lambda = \sum_I \lambda^{|I|}$ be the partition function, where the summation is over independent sets I of G. We are interested in calculating Z or to sample from distribution in which the probability of outputting independent set I is $\frac{\lambda^{|I|}}{Z}$.

Denote p_v as the probability that a vertex v is chosen in the independent set (referred to as being occupied).

$$p_v \equiv p_\lambda^{G,v} = \frac{\sum_{I \ni v} \lambda^{|I|}}{Z_G^\lambda}.$$

Let $\Lambda \subset V$ be the subset of V with configurations $\sigma_\Lambda, \tau_\Lambda \in \{0,1\}^\Lambda$ ($\sigma_\Lambda(u) = 1$ indicates that u is occupied and 0 indicates that u is unoccupied). Denote $p_v^{\sigma_\Lambda}$ as the probability that v is occupied conditioned on the configuration in Λ being fixed as specified by σ_Λ. Only configurations of Λ which specify independent sets are considered so that the conditional probability is well defined.

Definition 5.1 *Let $\delta : \mathbb{N} \to \mathbb{R}^+$. The distribution over independent sets of $G = (V, E)$ with activity λ is said to exhibit weak spatial mixing with rate $\delta(\cdot)$ if and only if $\forall v \in V,\ \Lambda \in V$, and for any two configurations $\sigma_\Lambda, \tau_\Lambda$ specifying independent sets of Λ,*

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| \leq \delta(dist(v, \Lambda)),$$

where $dist(v, \Lambda)$ is the length of the shortest path between v and Λ.

In statistical physics, G is usually an infinite graph, and weak mixing with a rate that goes to zero is equivalent to the uniqueness of the Gibbs measure (i.e. the existence of a unique macroscopic equilibrium).

Definition 5.2 *Under the same conditions as Definition 5.1, the distribution of over independent sets of G is said to exhibit strong spatial mixing with rate $\delta(\cdot)$, if and only if,*

$$|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| \leq \delta(dist(v, \Delta)),$$

where $\Delta \subseteq \Lambda$ is the subset on which σ_Λ and τ_Λ differ.

Note that, by definition, strong spatial mixing implies weak spatial mixing. This is because Δ is at least the same distance from v as Λ and in the strong mixing definition we are allowed to fix vertices that are close to v as long as we fix them to the same value in both σ and τ. The converse may not be true with a counter example being the ferromagnetic Ising model where fixing vertices close to v may shift p_v to a regime where it is more sensitive to the configuration in Δ.

Definition 5.3 *Let $\hat{\mathbb{T}}^b$ be the infinite regular tree where each vertex has degree $b+1$. If the tree is rooted at any particular vertex then the root has $b+1$ children, while the rest of the vertices have b children each. This graph is usually referred to as the Bethe lattice or Cayley tree. (As used later, $\mathbb{T}^b$ will denote the rooted infinite regular tree in which the root has b children as do the rest of the vertices.)*

5.1.4 Statements

The following two important result statements from the paper will be discussed at length herein:

Theorem 5.4 *(Theorem 2.3): For every positive integer b and any λ, if $\hat{\mathbb{T}}^b$ with activity λ exhibits strong spatial mixing with rate δ then, with the same activity λ, every graph of maximum degree $b+1$ exhibits strong spatial mixing with rate δ.*

Theorem 5.5 *(Theorem 2.7): There exists a deterministic algorithm such that for every integer b and any λ, if $\hat{\mathbb{T}}^b$ with activity λ exhibits strong spatial mixing with rate $\delta(l) = O(exp(-\alpha l))$ for some $\alpha > 0$ and shortest path length l, then on input of any graph G of maximum degree $b+1$ the algorithm approximates the partition function Z_G^λ to within a factor of $(1+\epsilon)$ in time t, where $n = |V|$. (The algorithm outputs two numbers $Z1$, $Z2$ such that $Z1 \leq Z_G^\lambda \leq Z2$ and $Z2 \leq Z1(1+\epsilon)$.) Similarly, there is a randomized algorithm that under the same condition (and with the same running time) generates independent sets of G where for every independent set I, the probability that the algorithm outputs I is within a factor $(1 \pm \epsilon)$ from $\frac{\lambda^{|I|}}{Z_G^\lambda}$.*

The following result follows from the above Theorem 5.5:

Corollary 5.6 *(Corollary 2.8): There exists a deterministic algorithm such that for every positive integer b and any $\lambda < \lambda_c(b)$, on input of any graph G of maximum degree $b+1$, the algorithm approximates the partition function Z_G^λ to within a factor of $(1+\epsilon)$ in time polynomial in $\frac{(1+\epsilon)n}{\epsilon}$. Similarly, there is a randomized algorithm that for the same choice of parameters (and same running time) generates independent sets of G with the probability of sampling an independent set I is within a factor $(1 \pm \epsilon)$ from $\frac{\lambda^{|I|}}{Z_G^\lambda}$.*

There are several other result statements either stated or proved as a part of this work in the paper but we will mainly focus on the above two results as they convey the main idea of the paper. Additional supplementary results will be listed in Section 5.4.

5.1.5 Self-Avoiding-Walk Tree Representation T_{saw} of graph G

For a graph $G(V,E)$, the aim is to calculate the probability p_v that a vertex $v \in G$ is occupied. For every vertex $u \in V$ we fix an (arbitrary) enumeration of the edges incident to u. From here onwards, whenever we say that an edge $\{u,w\}$ (or a neighbor w of u) is larger than $\{u,x\}$ (or a neighbor x of u) we interpret this according to the enumeration of the edges incident to u that we fixed here. The description of the tree corresponding to (G,v) includes the tree (as a graph) together with a specification of a subset of its leaves that are fixed to specific values (occupied or unoccupied).

Denote the tree corresponding to (G,v) as $T_{saw}(G,v)$. $T_{saw}(G,v)$ is defined as the tree of all paths originating at v with following conditions:

- On path closing the copy (in the tree) of the vertex closing the cycle (in G) is fixed to occupied if the edge closing the cycle is larger than the edge starting the cycle with the rest of the path ignored.

- If a condition on G fixes the vertex u to a certain value, the corresponding condition on $T_{saw}(G, v)$ fixes all the copies of u to the same value. For each fixed vertex x in $T_{saw}(G, v)$, we also erase the subtree underneath x in order to make sure the resulting condition is well defined.

Figure 5.1 shows a $T_{saw}(G, v)$ representation for a small example graph G. The tree on the left is $T_{saw}(G, v)$, where G is the graph on the right and where the order on the neighbors of each vertex in G is lexicographic. In order to better illustrate the construction we labeled each vertex in the tree with the name of its corresponding vertex in G. Notice that vertices that close cycles are fixed to be either occupied or unoccupied. For example, the bottom-left copy of d is fixed to be occupied because the edge $\{d, f\}$ that closes the cycle is larger than the edge $\{d, e\}$ that starts it.

$T_{saw}(G, v)$ has two types of fixed vertices:

- Structural: These fixed vertices arise from the cycle structure of the graph G and both their composition and values are independent of the condition imposed on G.
- Natural: These fixed vertices correspond to fixed vertices in G and the values they are fixed to are simply copied from their corresponding vertices in G.

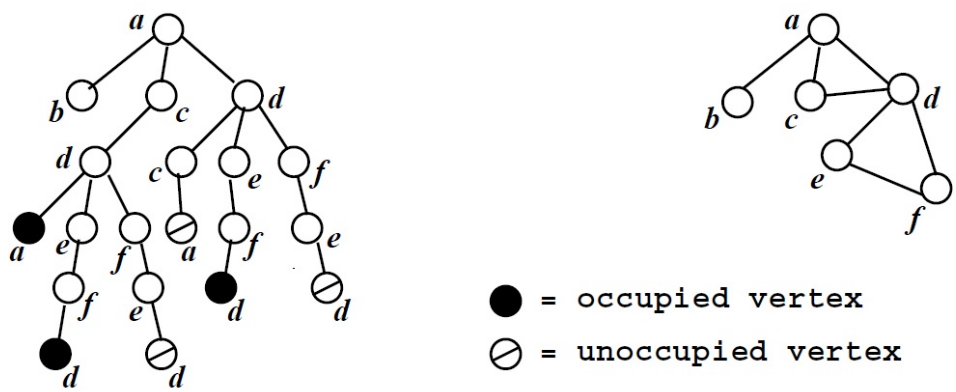

Figure 5.1: The construction of T_{saw}.

5.2 Theorem 2.3: Transferring strong spatial mixing from regular trees to general graphs

Theorem 2.3 follows from Theorem 3.1 which is stated as follows:

Theorem 5.7 *(Theorem 3.1): For any graph $G, \lambda, \Lambda \subset V$, and configuration σ_Λ,*

$$p_v^{\sigma_\Lambda} = \mathbb{P}_v^{\sigma_\Lambda},$$

where $\mathbb{P}_v^{\sigma_\Lambda} \equiv \mathbb{P}_{G,v}^{\sigma_\Lambda}(\lambda)$ stands for the probability that the root of $T_{saw}(G, v)$ is occuipied when imposing the condition σ_Λ.

Note that Theorem 3.1 implies Theorem 2.3 since we would get $|p_v^{\sigma_\Lambda} - p_v^{\tau_\Lambda}| = |\mathbb{P}_v^{\sigma_\Lambda} - \mathbb{P}_v^{\tau_\Lambda}|$, and that for any subset Δ of vertices of G, $\text{dist}(v, \Delta)$ is exactly same as distance between the root of $T_{saw}(G, v)$ and subset of vertices of the tree composed of copies of vertices in Δ.

Proof: For convenience of description, the paper works with ratios of probabilities rather than the probability of occupation itself. Thus, let $R_v \equiv R_{G,v}(\lambda) = p_v/(1-p_v)$ be the ratio between the probability that v is occupied and unoccupied, with $R_v^{\sigma_\Lambda}$ representing the ratio of probabilities conditioned on σ_Λ. For $p_v = 1$, let $R_v = \infty$. Note that for a rooted tree T, once we fix the value at the root, the configurations of the subtrees are all independent of each other. Thus we see that the ratios of the probabilities ($R_T^{\sigma_\Lambda}$) that the root of tree T is occupied and unoccupied is given by,

$$R_T^{\sigma_\Lambda} = \lambda \prod_{i=1}^{d} \frac{1}{1 + R_{T_i}^{\sigma_{\Lambda_i}}},$$

where d is the number of children of the root, T_i is the subtree rooted at the i-th child, $\Lambda_i = \Lambda \cap T_i$, and σ_{Λ_i} is the restriction of σ_Λ to Λ_i. $R_T^{\sigma_\Lambda}$ is a recursive caluculation where the base cases are either when $v \in \Lambda$, in which $R_v = \infty$ if v is fixed to be occupied or $R_v = 0$ if v is fixed to be unoccupied, or when v has no children in which case $R_v = \lambda$.

To mimic this calculation for general graphs, R_v is to be written in terms of R_{u_i}, where u_i are neighbors of v. However, there is a problem in that u_i might depend on each others values, even with v_i fixed. To overcome this scenario, for the vertex v of R_v, let G' be the same as G except that v is replaced by d vertices $v_1, \ldots, v_d$, where d is the degree of v. Each v_i has a single edge connecting to u_i, the i-th neighbor of v. Furthermore, each v_i is assigned a new activity $\lambda^{1/d}$. Notice that when all of the v_i's are occupied or unoccupied, G' has the same weights as G when v is occupied or unoccupied, respectively. Using this fact, we can write $R_{G,v}^{\sigma_\Lambda}$ as the ratio of the probability of all v_i's in G' being occupied and all being unoccupied,

$$R_{G,v}^{\sigma_\Lambda} = \prod_{i=1}^{d} R_{G',v}^{\sigma_\Lambda \tau_i},$$

where $\sigma_\Lambda \tau_i$ is the concatenation of the two configurations σ_Λ and τ_i, where τ_i is the configuration of the added vertices (for $i \neq j$) where v_j is occupied for $j < i$ and v_j is unoccupied for $j > i$. Now, with v_i only connected to u_i in G', we get

$$R_{G',v_i}^{\sigma_\Lambda \tau_i} = \frac{\lambda^{1/d}}{1 + R_{(G' \setminus v_i),u_i}^{\sigma_\Lambda \tau_i}}$$

and hence,

$$R_{G,v}^{\sigma_\Lambda} = \lambda \frac{1}{1 + R_{(G' \setminus v_i),u_i}^{\sigma_\Lambda \tau_i}}$$

which is a recursive calculation similar to that of $R_T^{\sigma_\Lambda}$. Since the number of unfixed vertices reduces after each recursive call, the calculation will terminate. Furthermore, the calculations carried out by both $R_T^{\sigma_\Lambda}$ and $R_{(G,v)}^{\sigma_\Lambda}$ are the exactly the same since they both are in fact the tree of paths in G starting at u_i, except that whenever v is visited, the corresponding vertex in the tree is fixed to either occupied or unoccupied depending on whether the path reached v from a neighbor smaller or greater than i, thus proving Theorem 3.1. ■

5.3 Algorithmic Implications

This section provides description and analysis of the algorithm that uses results from Theorem 2.3 for approximating the partition function Z_G^λ as claimed in Theorem 2.7.

5.3.1 Algorithm

We note that in order to calculate Z it is enough to calculate the probability of the empty set since this probability is exactly $\frac{1}{Z}$. In order to generate a random independent set we can choose v to be occupied with probability p_v and unoccupied otherwise, and continue to generate the rest of the configuration conditioned on the chosen value at v.

In the above section, we described a recursive procedure for calculating p_v which may have exponential running time but when the tree exhibits strong spatial mixing with exponential decay, a light modification can be used:

If we can output two numbers $p1$, $p2$ such that $p1 \leq p_v \leq p2$ and $p2 \leq (1 + \frac{\epsilon}{(1+\epsilon)n})p1$ then, by the same reduction as above we can generate a random independent set such that the probability of outputting I is within a factor $(1 \pm \epsilon)$ from $\frac{\lambda^{|I|}}{Z_G^\lambda}$. A recursive procedure is used to calculate lower and upper bounds on p_v. The recursive calls return a lower and an upper bound on $R_{G',v_i}^{\sigma_\Lambda \tau_i}$ for each i. The lower bounds are then used to compute an upper bound on$R_{G,v}^{\sigma_\Lambda}$ and vice versa. The procedure has three stopping rules:

1. If v is fixed by σ_Λ then both the lower and upper bounds are set to the same value as described in above section.
2. If v has no neighbors then both the lower and upper bounds are set to the same value as described in above section.
3. If the stack of the recursion is l levels deep, where l is a parameter of the algorithm, set the lower and upper bounds on R to 0 and ∞, respectively.

5.3.2 Analysis

The algorithm outputs two numbers $p1, \ p2$ such that $p_1 \leq p_v \leq p_2$. Consider the algorithm is run using level parameter l.

The upper bound p_2 is exactly the probability that the root of $T_{saw}(G,v)$ is occupied conditioned on all the vertices that are not already fixed at level l below the root being occupied (respectively unoccupied) if l is even (respectively odd). The lower bound p_1 is exactly the probability of the root being occupied under the negated condition.

The running time of the algorithm with parameter l is in the order of the size of $T_{saw}(G,v)$ restricted to its first l levels and the upper bound can be given as $O(b^l) = O((\frac{(1+\epsilon)n}{\epsilon})^{\frac{(lnb)}{\alpha}})$ as claimed in Theorem 2.7.

5.4 Other Results

There are several other results discussed in the paper but due to brevity of these notes, we will just state it here. The most important result in addition to the above discussed Theorem 2.3 is the proof that weak spatial mixing implies strong spatial mixing. The paper proves a stronger result related to the monotonicity in the activity λ. Following statements cover these results:

Theorem 5.8 *(Theorem 2.4):For every positive integer b and any λ, if $\hat{\mathbb{T}}^b$ with activity λ exhibits weak spatial mixing with rate δ, then it also exhibits strong spatial mixing with rate $\frac{(1+\lambda)(\lambda+(1+\lambda)^{b+1})}{\lambda}\delta$.*

Theorem 5.9 *(Theorem 4.1):Fix an arbitrary $\lambda \geq 0$. Let $\vec{\lambda}$ be an assignment of activities to the vertices of $\hat{\mathbb{T}}^b$ such that $0 \leq \vec{\lambda}(v) \leq \lambda$ for every $v \in \hat{\mathbb{T}}^b$. Then, for every l,*

$$\frac{R_l^E(\vec{\lambda})}{R_l^O(\vec{\lambda})} \leq \frac{R_l^E(\lambda)}{R_l^O(\lambda)}$$

Lemma 5.10 *(Lemma 4.2):For every integer $l \geq 1$ and any assignment of activities $0 \leq \vec{\lambda} \leq \lambda$ to the vertices of $\mathbb{T}^b$,*

$$\frac{R_l^E(\vec{\lambda})}{R_l^O(\vec{\lambda})} \leq \frac{R_l^E(\lambda)}{R_l^O(\lambda)};$$

$$\frac{1 + R_l^E(\vec{\lambda})}{1 + R_l^O(\vec{\lambda})} \leq \frac{1 + R_l^E(\lambda)}{1 + R_l^O(\lambda)}.$$

Theorem 2.4 follows from Theorem 4.1 from noticing that $\frac{R^E}{R^O} - 1 = \frac{p^E - p^O}{p^O(1-p^E)}$ and that for $l \geq 2, p_l^E(\lambda) \leq \frac{\lambda}{1+\lambda}$ and $p_l^O(\lambda) \geq \frac{\lambda}{\lambda + (1+\lambda)^{b+1}}$.

The regime of λ for which $\hat{\mathbb{T}}^b$ exhibits weak spatial mixing with that rate that goes to zero is well known from Kelly (1985) and Spitzer (1975):

Proposition 5.11 *(Proposition 2.5): $\forall b \in \mathbb{Z}^+, \hat{\mathbb{T}}^b$ with activity λ exhibits weak spatial mixing with a rate δ that goes to zero exponentially fast if and only if $\lambda < \lambda_c(b) = \frac{b^b}{(b-1)^{(b+1)}}$.*

Corollary 5.12 *(Corollary 2.6): For every positive integer b and any $\lambda \leq \lambda_c(b)$, there exists a decaying rate δ such that for every graph G of maximum degree $b+1$, G with activity λ exhibits strong spatial mixing with rate δ. (In particular, for any graph of maximum degree $b+1$, the Gibbs measure is unique for $\lambda \leq \lambda_c(b)$.) Furthermore, the rate δ can be taken to decay exponentially fast (with constants that depend on b and λ) if $\lambda < \lambda_c(b)$.*

Connections between strong spatial mixing and the rate of convergence of the Glauber dynamics were the subject of a number of papers in Statistical Physics as well as in Computer Science. The following is a partial summary that suffices current scope:

Theorem 5.13 *(Theorem 2.9): If G is a graph that grows subexponentially and if G with activity parameter λ exhibits strong spatial mixing with exponential decay then the mixing time of the Glauber dynamics for sampling independent sets of G with activity λ is $O(n^2)$.*

Corollary 5.14 *(Corollary 2.10): If G is a graph that grows subexponentially and has maximum degree $b+1$ then the Glauber dynamics on G mixes in time $O(n^2)$ for any $\lambda < \lambda_c(b)$.*

5.5 Extensions

In this section we describe different extensions that have been proposed in the regime of counting/sampling independent sets and integer lattices. Specifically, two recent papers provide an improved lower and an upper bound on λ.

- Vera, Vigoda and Yang (2014) study the critical activity $\lambda_c(\mathbb{Z}^2)$ for the uniqueness/non-uniqueness threshold on the 2-dimensional integer lattice $\mathbb{Z}^2$. For the hard-core lattice gas model defined on independent sets weighted by an activity λ. In this paper, they use the ideas from Weitz (2006) and Restrepo et. al. (2011) to improve the lower bound on λ to $\lambda < \lambda_c = 2.48$.

- Most of the work in the regime of phase coexistence has focused on uniqueness of Gibbs measure and tried to improve lower bounds on λ. Blanca et. al. (2014) presents a first explicit work in the opposite direction where they concentrate on the existence of multiple Gibbs states. In this work they extend the characterization of fault lines to show that local Markov chains will mix slowly when $\lambda > 5.36464$ on lattice regions with periodic (toroidal) boundary conditions and when $\lambda > 7.1031$ with non-periodic (free) boundary conditions.

A future work shown in Wietz's paper is the model of proper colorings, where the goal is to show that $b+2$ colors (the threshold for weak spatial mixing on the regular tree) are enough for spatial mixing on any graph of maximum degree $b+1$. Yin (2014) shows that for $q \geq \alpha d + \beta$ with $\alpha > 2$ and sufficiently large $\beta = O(1)$, with high probability proper q-colorings of random graph $G(n, d/n)$ exhibit strong spatial mixing with respect to an arbitrarily fixed vertex.

5.6 References

1. D. WEITZ. Counting independent sets up to the tree threshold. *Proc. 38th ACM Symp. on the Theory of Computing (STOC)*, 140149, 2006.

2. M. LUBY and E. VIGODA. Approximately counting up to four. *Proc. 29th ACM Symp. on Theory of Computing*, 1997, pp. 682687.

3. M. DYER and C. GREENHILL. On Markov chains for independent sets. *J. Algorithms 35*, (2000), pp. 1749.

4. E. VIGODA. A note on the Glauber dynamics for sampling independent sets. *Electronic Journal of Combinatorics*, 8(1) (2001).

5. T. HAYES and E. VIGODA. Coupling with the stationary distribution and improved sampling for colorings and independent sets. *Proc. 16th ACM-SIAM Symp. on Discrete Algorithms*, 2005, pp. 971979.

6. J. VAN DEN BERG and A. ERMAKOV. A new lower bound for the critical probability of site percolation on the square lattice. *Random Structures and Algorithms*, 8 (1996), pp. 199212.

7. A. D. SOKAL. A personal list of unsolved problems concerning lattice gases and antiferromagnetic Potts models. *Markov Process*, Related Fields 7 (2001), pp. 2138.

8. J. VERA, E. VIGODA, L. YANG. Improved Bounds on the Phase Transition for the Hard-Core Model in 2-Dimensions. *arXiv:1306.0431.*

9. R. RESTREPO, J. SHIN, P. TETALI, E. VIGODA, and L. YANG. Improved Mixing Condition on the Grid for Counting and Sampling Independent Sets. *Probability Theory and Related Fields*, 156(1-2):75-99, 2011.

10. A. BLANCA, D. GALVIN, D. RANDALL, and P. TETALI. Phase Coexistence and Slow Mixing for the Hard-Core Model on Z2. *In Proceedings of the 17th International Workshop*, RANDOM, 379-394, 2013.

11. Y. YIN. Spatial Mixing of Coloring Random Graphs. *arXiv:1402.4556.*