INTERNATIONAL COMPUTER SCIENCE INSTITUTE

1947 Center St. • Suite 600 • Berkeley, California 94704-1198 • (510) 643-9153 • FAX (510) 643-7684

Fast Convergence of the Glauber Dynamics for Sampling Independent Sets: Part I

Michael Luby * Eric Vigoda †

TR-99-002

January 28, 1999

Abstract

We consider the problem of sampling independent sets of a graph with maximum degree δ. The weight of each independent set is expressed in terms of a fixed positive parameter $\lambda \leq \frac{2}{\delta-2}$, where the weight of an independent set σ is $\lambda^{|\sigma|}$. The Glauber dynamics is a simple Markov chain Monte Carlo method for sampling from this distribution. We show fast convergence of this dynamics. This paper gives the more interesting proof for triangle-free graphs. The proof for arbitrary graphs is given in a companion paper [28]. We also prove complementary hardness of approximation results, which show that it is hard to sample from this distribution when $\lambda > \frac{c}{\delta}$ for a constant $c > 0$.

*International Computer Science Institute Research supported in part by National Science Foundation operating grant NCR-9416101, and United States-Israel Binational Science Foundation grant No. 92-00226

†Computer Science Division, University of California at Berkeley, and International Computer Science Institute. A substational portion of this research was done while the author was visiting DIMACS. Research supported in part by National Science Foundation Fellowship, NSF operating grant NCR-9416101, and DIMACS fellowship.

1 Introduction

The hard-core model in statistical physics is a simple representation of a gas as a graph $G = (V, E)$ [3, 6]. Vertices represent possible sites for the non-negligible sized particles. To prevent particles from overlapping, adjacent sites cannot simultaneously be occupied. We are interested in the hard-core measure which is the probability measure μ_G defined on the set of independent sets $\Omega = \Omega_G$ of $G = (V, E)$ weighted by a parameter $\lambda > 0$. Specifically, for all independent sets $\sigma \in \Omega_G$,

$$\mu_G(\sigma) = \frac{\lambda^{|\sigma|}}{Z_G}$$

where $Z_G = \sum_{\sigma \in \Omega_G} \lambda^{|\sigma|}$ is the normalizing factor, often referred to as the partition function. Equivalent models arise in the Operations Research community when considering properties of stochastic loss systems which model communication networks [15], [17].

Sampling from the hard-core measure and computing the partition function are closely related problems. As described in [13, 26], there are randomized approximation-preserving reductions between the two problems.

We consider a very simple Markov chain MC defined on Ω_G. This chain is referred to as the Glauber dynamics in the statistical physics community [10]. The transition probability matrix P of MC is defined as follows. From an independent set σ:

- Choose a vertex v of G uniformly at random.
- Let
$$\sigma' = \begin{cases} \sigma \cup \{v\} & \text{with probability } \frac{\lambda}{1+\lambda} \\ \sigma \setminus \{v\} & \text{with probability } \frac{1}{1+\lambda} \end{cases}$$
- If σ' is a valid independent set, move to state σ' otherwise remain at state σ.

The stationary distribution of MC is the hard-core measure. To sample from this distribution, we simply start at an arbitrary independent set and follow the random walk defined by P until we reach stationarity. Our main result is that MC quickly converges to its stationary distribution in a manner we formalize later. In particular, for graphs with maximum degree δ, the number of steps of this random walk that are required to achieve stationarity, known as the mixing time, is $O(n \log n\delta)$ for $\lambda < \frac{2}{\delta-2}$ where n is the number of vertices in the input graph G. This implies an efficient method to sample from the hard-core measure, and from the results of [13, 26], an *fpras* (fully-polynomial randomized approximation scheme) to approximate Z. An *fpras* produces an estimate within a $1+\epsilon$ multiplicative factor of the correct answer with probability at least $\frac{3}{4}$, and runs in time polynomial in $\frac{1}{\epsilon}$ and n [14].

In this work, we present the proof for triangle-free graphs and $\lambda < \frac{2}{\delta-2}$. The generalization to arbitrary graphs is much more complicated and laborious. The boundary case of $\lambda = \frac{2}{\delta-2}$ is much easier to handle in the context presented for arbitrary graphs. These extensions are left for a separate note [28].

1.1 Related Work on Sampling

The conference version of this paper [18] showed rapid mixing of a modified Glauber dynamics MC_{edge} for $\lambda \leq \frac{1}{\delta-3}$. In joint work with Michael Mitzenmacher, we discovered a modification of the chain MC_{edge} which leads to rapid mixing for $\lambda \leq \frac{2}{\delta-2}$. We omit the proof since it is weaker than the result we present on the Glauber dynamics. Independently of our work, Dyer and Greenhill [9] also discovered the modification of MC_{edge} leading to the same improved bounds on λ.

Dyer and Greenhill [9] and Randall and Tetali [25] have used these bounds on the mixing rate of these modified Glauber dynamics to bound the mixing rates of the simple Glauber dynamics. These comparison approaches give weaker bounds on the mixing rate.

J. Propp and D. Wilson noticed that our proof of rapid mixing and similar types of proofs which use coupling actually show fast convergence of their coupling from the past technique [22, 23]. Their work can be used to generate samples exactly from the hard-core distribution [11].

1.2 Computational Complexity Hardness Results

Exactly computing $|\Omega_G|$ is $\#P$-complete even when restricted to bipartite graphs with maximum degree four [27]. Using standard boosting techniques and hardness of approximation results for independent sets, it is known to be NP-hard to approximate $|\Omega_G|$ for general graphs within a factor of $2^{n^{1-\epsilon}}$ for any $\epsilon > 0$ [26, 12]. Using the same approach, we show it is NP-hard (unless $RP = NP$) to compute Z_G within any polynomial factor when $\lambda > \frac{c}{\delta}$, for a positive constant $c \leq \frac{20(1+\epsilon)}{\epsilon}$ where $1+\epsilon$ is the latest hardness of approximation result for finding the maximum independent set in graphs with maximum degree four. The latest bound on ϵ is $\frac{1}{555}$ from [4].

1.3 Infinite Volume Gibbs Measure

Work on proving rapid mixing of the Glauber dynamics and related Markov chains is closely related to a large body of research in the Statistical Physics community. In this section, we briefly introduce a major topic of interest in the Statistical Physics community and mention its connection to our work.

We refer the reader to [10] for a general introduction to the concepts presented in this section. The following definitions focus on $\mathcal{Z}^d$ but hold for arbitrary lattices. Let Q_L denote the d-dimensional cube of side length L, i.e. Q_L is the induced subgraph of $\mathcal{Z}^d$ on the set of vertices $\{0,\dots,L-1\}^d$. Also consider $\bar{Q_L}$ the complement of Q_L in $\mathcal{Z}^d$, i.e. $\bar{Q_L} = \mathcal{Z}^d \setminus Q_L$.

In statistical physics, the hard-core measure is the associated Gibbs measure for this model. For a fixed independent set $\tau \in \Omega_{\mathcal{Z}^d}$, the Gibbs measure on Q_L is defined as:

$$\mu^{\tau}_{Q_L}(\sigma) = \frac{\lambda^{|\sigma|}}{Z^{\tau}_{Q_L}}$$

where $\{\sigma \cap Q_L\} \cup \{\tau \cap \bar{Q_L}\}$ is a valid independent set on $\mathcal{Z}^d$, and the partition function $Z^{\tau}_{Q_L}$ is the appropriate normalizing constant.

The question of interest is whether there is a unique limiting Gibbs measure as $L \to \infty$, independent of τ [6].

Work on the question of uniqueness has focused on two-dimensions. It is widely believed that there is a critical parameter λ_c such that uniqueness holds for $\lambda \leq \lambda_c$. Simulations suggest λ_c is about 3.79 [2], but rigorous bounds on λ_c are much worse. Most lower bounds rely on showing the Dobrushin-Shlosman [8] condition is satisfied. Satisfying this condition is quite similar to using coupling to prove fast convergence of a Markov chain which allows updates of larger structures of the input graph such as a $k \times k$ subsection of the grid. The best lower bounds using this approach are roughly $\lambda_c > 1.185$ [16, 24, 7]. These proofs are usually computer assisted. A different approach has recently been used by van den Berg and Steif [3]. They relate λ_c to the critical probability p_c for site percolation on the grid. Specifically, they show $\lambda_c > \frac{p_c}{1-p_c}$.

Work on showing that for large enough λ, the limiting Gibbs measure is not unique was done by Dobrushin [6] and rediscovered by Louth [17] in a different context.

The infinite $\delta-$regular tree is the only graph where λ_c is rigorously known exactly. For such graphs, Kelly [15] showed $\lambda_c = \frac{(\delta-1)^{(\delta-1)}}{(\delta-2)^{\delta}}$.

As J. van den Berg explained to us, there is an intimate connection between fast convergence of the Glauber dynamics (or any Markov chain whose transitions depend on a small subgraph of the input graph) and uniqueness of the limiting Gibbs measure. Specifically, our results which show rapid mixing in $O(n \log n)$ time of the Glauber dynamics on the grid for all boundary conditions imply uniqueness of the Gibbs measure. Similarly, using results of Martinelli and Olivieri [19, 20] and Aizenman and Holley [1], most results which show uniqueness (which actually show the stronger condition of so-called weak spatial mixing [19]) imply mixing of the Glauber dynamics in $O(n \log n)$ time for the grid for any boundary condition.

The results we present in this work are for general graphs and imply bounds on λ_c for arbitrary lattices. Work in the statistical physics community has centered on the grid. For this special class of graphs, their results are stronger than ours and imply fast convergence of the Glauber dynamics for larger λ than we prove.

2 Machinery

2.1 Background

Before getting into the proof, we need to review some background material and machinery. Consider a discrete-time Markov chain (X_t) with transition probability matrix P defined on a finite state space Ω. A classical theorem of stochastic processes states that if P has the following properties:

- *aperiodicity*: $\gcd\{t : P^t_{ii} > 0\} = 1$ for all $i \in \Omega$
- *irreducibility*: there exists a t such that there is a positive probability of going from state i to state j after t steps, i.e. $P^t_{ij} > 0$, for all $i, j \in \Omega$

then the chain (X_t) has a unique limiting distribution, referred to as the stationary distribution π, i.e.

$$\lim_{t \to \infty} P^t_{ij} = \pi_j \text{ for all } i, j \in \Omega$$

In fact, π can easily be determined if the Markov chain is time-reversible, i.e. satisfies all local-balance equations. Specifically, a distribution π is the stationary distribution if it satisfies the following:

$$\pi_i P_{ij} = \pi_j P_{ji} \text{ for all } i, j \in \Omega$$

Our goal is to bound the time until the chain is sufficiently close to the stationary distribution. The traditional bound on the distance from stationarity is the variation distance,

$$\Delta_i(t) = \frac{1}{2} \sum_{j \in \Omega} |P^t_{ij} - \pi_j|$$

We are interested in the mixing time, τ:

$$\tau(\epsilon) = \max_i \min\{t : \Delta_i(t') \leq \epsilon \text{ for all } t' \geq t\}$$

2.2 Coupling

We use coupling to bound the mixing time. Coupling constructs a stochastic process (X_t, Y_t) on $\Omega \times \Omega$ such that separately X_t, Y_t are copies of the original Markov chain and if $X_t = Y_t$, then $X_{t+1} = Y_{t+1}$.

An important tool in our analysis is Bubley and Dyer's path coupling [5]. Our statement of their theorem follows that in [9].

We first need to define the notion of neighbors and paths in Ω. We consider a pair of states $\sigma, \sigma_v \in \Omega$ neighbors if $\sigma_v = \sigma \cup \{v\}, v \notin \sigma$ and denote it by $\sigma \sim \sigma_v$. We call $\tau = (\tau_0, \ldots, \tau_k)$ a simple path if all τ_i are distinct and $\tau_0 \sim \tau_1 \ldots \sim \tau_k$. Then $\rho(\sigma, \eta) = \{\tau : \sigma = \tau_0, \eta = \tau_k, \tau \text{ is a simple path}\}$.

Using path coupling we only need to analyze a coupling for neighboring states. The path coupling theorem is more general than stated here, but this is sufficient for our purposes. In the following theorem σ' (similarly σ'_v) refers to the state of the chain in state σ after one step of the Markov chain.

Theorem 1 *[5] Let Φ be an integer-valued metric defined on $\Omega \times \Omega$ which takes values in $\{0, \ldots, D\}$ and for all $\sigma, \eta \in \Omega$ there exists a $\tau \in \rho(\sigma, \eta)$ such that*

$$\Phi(\sigma, \eta) = \sum_i \Phi(\tau_i, \tau_{i+1})$$

Suppose there exists a $\beta < 1$ and a coupling of the Markov chain MC such that for all $\sigma, \sigma_v \in \Omega$:

$$E[\Phi(\sigma', \sigma'_v)] \leq \beta \Phi(\sigma, \sigma_v)$$

Then the mixing time is

$$\tau(\epsilon) \leq \frac{\log(D\epsilon^{-1})}{1 - \beta}$$

3 Analysis of Glauber Dynamics

In this section, we analyze the Glauber dynamics for triangle-free graphs.

3.1 Potential Function

Consider a pair of states $\sigma \sim \sigma_v$ where vertex v has degree δ_v. Recall that $\sigma_v = \sigma \cup \{v\}$. The obvious idea for a potential function between such states is just δ_v, the Hamming distance weighted by degree. Our potential function is an obvious extension of this. Consider a vertex w which is a neighbor of v. We call w blocked if it has a neighbor which is in both independent sets. Suppose the next move of the Markov chain attempts to add w into the independent set. This might only work in one of the chains, causing an increase in Φ. Notice that this bad situation occurs if w is not blocked. Otherwise, this move is blocked from occurring in both chains. Our potential function is simply the weighted Hamming distance minus a constant $c < 1$ times the number of blocked neighbors of v.

Specifically, for $c = \frac{\delta\lambda}{\delta\lambda+2}$, our potential function Φ is as follows. We use $\Gamma(v)$ to denote the set of neighboring vertices of v. Denote the set of blocked neighbors of v by

$$B(\sigma, v) = \{w : w \in \Gamma(v), \Gamma(w) \cap \sigma \neq \emptyset\}$$

$$\Phi(\sigma, \sigma_v) = \delta_v - c|B(\sigma, v)|$$

Consider arbitrary states σ, η.

$$\Phi(\sigma, \eta) = \min_{\tau \in \rho(\sigma,\eta)} \sum_i \Phi(\tau_i, \tau_{i+1})$$

This potential function Φ clearly satisfies the following conditions for all $\sigma, \eta \in \Omega$ and thus is a metric:

- $\Phi(\sigma, \eta) \leq \Phi(\sigma, \zeta) + \Phi(\zeta, \eta)$. This is true since $\Phi(\sigma, \eta)$ is defined as a minimum over all paths including those going through ζ.
- $\Phi(\sigma, \eta) = \Phi(\eta, \sigma)$
- $\Phi(\sigma, \eta) \geq 0$ which follows from $c < 1$ and thus $\Phi(\sigma, \sigma_v) \geq \delta_v(1 - c) > 0$ for all $\sigma, \sigma_v \in \Omega$.
- $\Phi(\sigma, \eta) = 0 \leftrightarrow \sigma = \eta$.

3.2 Analysis

Let $\Phi = \Phi(\sigma, \sigma_v)$. We now analyze $E[\Delta\Phi]$. Our coupling is simply the identity, i.e. each chain attempts the same move.

Notice that the only moves which might affect Φ either transition on v, a neighbor of v, or a neighbor of a neighbor of v. Let,

$$E[\Delta^{+x}\Phi] = E[\Delta\Phi|\text{Markov chain attempts to add } x \text{ into the independent set}]$$
$$E[\Delta^{-x}\Phi] = E[\Delta\Phi|\text{Markov chain attempts to remove } x \text{ from the independent set}]$$
$$E[\Delta^{x}\Phi] = \tfrac{\lambda}{1+\lambda}E[\Delta^{+x}\Phi] + \tfrac{1}{1+\lambda}E[\Delta^{-x}\Phi]$$

This gives,

$$E[\Delta\Phi] = \frac{1}{n}\left[E[\Delta^{v}\Phi] + \sum_{w\in\Gamma(v)} E[\Delta^{w}\Phi] + \sum_{x\in\Gamma(\Gamma(v))} E[\Delta^{x}\Phi]\right]$$

Consider a move which

- transitions on v:
  Since all neighbors of v are out of both independent sets, a move which transitions on v works in both chains. Afterwards, both chains are in the same state. Thus,

$$E[\Delta^{+v}\Phi] = E[\Delta^{-v}\Phi] = -\delta_v + c|B(\sigma, v)|$$
$$E[\Delta^{v}\Phi] = -\delta_v + c|B(\sigma, v)|$$

- transitions on w, where w is a neighbor of v:
  Since w is in neither independent set, $E[\Delta^{-w}\Phi] = 0$.

  Consider the move which attempts to add w into the set. Suppose w is not blocked. This move only works in the chain in state σ. To determine the effect of this move for such a w, observe the following:

$$E[\Delta^{+w}\Phi] = \Phi(\sigma_w, \sigma_v) - \Phi(\sigma, \sigma_v)$$
$$\Phi(\sigma_w, \sigma_v) \le \Phi(\sigma_w, \sigma) + \Phi(\sigma, \sigma_v)$$
$$\Phi(\sigma_w, \sigma) = \delta_w - c|B(\sigma, w)|$$

  Combining these give $E[\Delta^{+w}\Phi] \le \delta_w - c|B(\sigma, w)|$.

  Note that,

$$E[\Delta^{w}\Phi] \le \begin{cases} \frac{\lambda}{1+\lambda}(\delta_w - c|B(\sigma, w)|) & \text{if } w \notin B(\sigma, v) \\ 0 & \text{otherwise} \end{cases}$$

- transitions on x, where x is a neighbor of a neighbor of v:
  Suppose x is in both independent sets and consider the move which removes x from both sets. This move may unblock a vertex w. The set of such w are

$$\alpha_x = \{w : w \in B(\sigma, v), \Gamma(w) \cap \sigma = \{x\}\}$$

  We have,

$$E[\Delta^{-x}\Phi] = \begin{cases} |\alpha_x|c & \text{if } x \in \sigma \\ 0 & \text{otherwise} \end{cases}$$

  Consider the case when x is in neither independent set. Since the graph is triangle-free, v is not in the neighborhood of x. Thus, the move which attempts to add x into the independent set works in both or neither set. It works in both chains if no neighbor of x is in either independent set, i.e. $\Gamma(x) \cap \sigma = \emptyset$. The only possible effect of such a move is to make a vertex w blocked. The set of such w are

$$\beta_x = \{w : w \in \Gamma(v) \cap \Gamma(x), w \notin B(\sigma, v)\}$$

Thus,

$$E[\Delta^{+x}\Phi] = \begin{cases} -|\beta_x|c & \text{if } \Gamma(x)\cap\sigma = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

Combining these,

$$E[\Delta^{x}\Phi] = \begin{cases} \frac{|\alpha_x|}{1+\lambda}c & \text{if } x\in\sigma \\ -\frac{\lambda|\beta_x|}{1+\lambda}c & \text{if } x\notin\sigma, \Gamma(x)\cap\sigma = \emptyset \\ 0 & \text{otherwise} \end{cases}$$

We now collect terms of $E[\Delta\Phi]$ in a manner that divides the contribution from x over its neighbors w. Note that for any such x, either $E[\Delta^{+x}\Phi]$ or $E[\Delta^{-x}\Phi]$, but not both are non-negative. We can amortize these over those $w\in\alpha_x$ or $w\in\beta_x$. Notice that a blocked (unblocked) vertex w can only be in α_x (β_x, respectively).

For a blocked vertex w in the neighborhood of v, let

$$\Gamma'(w) = \{x : x\in\Gamma(w)\setminus\{v\}, w\in\alpha_x\}$$

$$E[\Delta^{*w}\Phi] = E[\Delta^{w}\Phi] + \sum_{x\in\Gamma'(w)} \frac{1}{|\alpha_x|}E[\Delta^{x}\Phi]$$

Similarly, for an unblocked vertex w in the neighborhood of v, let

$$\Gamma'(w) = \{x : x\in\Gamma(w)\setminus\{v\}, w\in\beta_x\}$$

$$E[\Delta^{*w}\Phi] = E[\Delta^{w}\Phi] + \sum_{x\in\Gamma'(w)} \frac{1}{|\beta_x|}E[\Delta^{x}\Phi]$$

We now have that

$$E[\Delta\Phi] = \frac{1}{n}\left[E[\Delta^{v}\Phi] + \sum_{w\in\Gamma(v)} E[\Delta^{*w}\Phi]\right]$$

We can bound $E[\Delta^{*w}\Phi]$ as follows.

- Suppose w is blocked.
  We know that $E[\Delta^{w}\Phi]=0$. We also know that it can be in α_x for at most one x. Thus,

  $$E[\Delta^{*w}\Phi] \le \frac{c}{1+\lambda}$$

- Suppose w is unblocked.
  For each neighbor x of w (other than v), either $x\in B(\sigma,w)$ and thus contributes to $E[\Delta^{w}\Phi]$ or $x\notin B(\sigma,w)$ which implies $\Gamma(x)\cap\sigma=\emptyset$ and it contributes to $E[\Delta^{x}\Phi]$. From these observations we have,

  $$\begin{aligned} E[\Delta^{*w}\Phi] &= \tfrac{\lambda}{1+\lambda}[\delta_w - c|B(\sigma,w)|] + \textstyle\sum_{x\in\Gamma(w)\setminus B(\sigma,w), x\neq v} -\tfrac{\lambda}{1+\lambda}c \\ &= \tfrac{\lambda}{1+\lambda}[\delta_w - c|B(\sigma,w)| - c(\delta_w - 1 - |B(\sigma,w)|)] \\ &= \tfrac{\lambda}{1+\lambda}[\delta_w - c(\delta_w - 1)] \end{aligned}$$

  where the second equality is from noticing the summation is over a set of size exactly $\delta_w - 1 - |B(\sigma,w)|$. From algebraic manipulations and our definition of c we have that

  $$(2+\lambda)c = \lambda(\delta - c(\delta-1) \ge \lambda(\delta_w - c(\delta_w - 1))$$

  This implies that if w is blocked then $E[\Delta^{*w}\Phi] \le \frac{1}{1+\lambda}(2+\lambda)c$.

Using our bounds on $E[\Delta^{*w}\Phi]$ we have

$$\begin{aligned} n(1+\lambda)E[\Delta\Phi] &= (1+\lambda)\left[E[\Delta^v\Phi] + \sum_{w\in\Gamma(v)} E[\Delta^{*w}\Phi]\right] \\ &\le (1+\lambda)[-\delta_v + c|B(\sigma,v)|] + \sum_{w\in B(\sigma,v)} c + \sum_{w\in\Gamma(v)\setminus B(\sigma,v)} c(2+\lambda)] \\ &= (1+\lambda)[-\delta_v] + \sum_{w\in B(\sigma,v)} c(2+\lambda) + \sum_{w\in\Gamma(v)\setminus B(\sigma,v)} c(2+\lambda) \\ &= -(1+\lambda)\delta_v + \delta_v c(2+\lambda) \\ &= \frac{\delta_v}{\delta\lambda+2}[\lambda(\delta-2)-2] \end{aligned}$$

Therefore,

$$E[\Delta\Phi] \le \frac{1}{n(1+\lambda)}\frac{\delta_v}{\delta\lambda+2}[\lambda(\delta-2)-2]$$

Notice that $E[\Delta\Phi] < 0$ when $\lambda < \frac{2}{\delta-2}$.

We now want to use this bound on $E[\Delta\Phi]$ with the path coupling theorem to get a bound on the mixing time. Recall that the path coupling theorem uses a bound on $\beta = \max_{\sigma,\sigma_v}\beta_{\sigma,\sigma_v}$ where

$$E[\Phi(\sigma', \sigma'_v)] = \beta_{\sigma,\sigma_v}\Phi(\sigma,\sigma_v)$$

We want to determine β_{σ,σ_v} in terms of $E[\Delta\Phi(\sigma,\sigma_v)] = E[\Delta\Phi]$ as follows:

$$\begin{aligned} \beta_{\sigma,\sigma_v}\Phi(\sigma,\sigma_v) &= E[\Phi(\sigma',\sigma'_v)] \\ (\beta_{\sigma,\sigma_v}-1)\Phi(\sigma,\sigma_v) &= E[\Phi(\sigma',\sigma'_v)] - \Phi(\sigma,\sigma_v) = E[\Delta\Phi] \\ \beta_{\sigma,\sigma_v} &= 1 + \frac{E[\Delta\Phi]}{\Phi} \end{aligned}$$

Observe that by our definition of Φ we have $\Phi \le \delta_v$. From this observation and our bound on $E[\Delta\Phi]$ we get a bound on β:

$$\beta \le 1 + \frac{1}{n(1+\lambda)}\frac{\lambda(\delta-2)-2}{(\delta\lambda+2)}$$

The path coupling theorem needs a bound on β and for Φ to be integer valued on $\{0,\ldots,D\}$. At the moment Φ can have fractional values since c is not an integer. Simply consider $\Phi' = \frac{\Phi}{c}$ which is integer-valued. Since, $\Phi(\sigma,\sigma_v) \le \delta_v \le \delta$, we have $\Phi(\sigma,\eta) \le n\delta$ for arbitrary $\sigma,\eta \in \Omega$. Thus, $D \le \frac{n\delta}{c}$. Plugging these bounds on β and D into the path coupling theorem we get that when $\lambda < \frac{2}{\delta-2}$,

$$\tau(\epsilon) \le \frac{n(1+\lambda)(\delta\lambda+2)}{2-\lambda(\delta-2)}\log(\frac{n\delta}{c\epsilon})$$

Using the fact that $\delta \ge 3, \lambda \le \frac{2}{\delta-2}$, we get $\lambda \le 2, \delta\lambda \le 6, c \ge \frac{1}{3}$.

We can now simplify the bound on the mixing time. For $\lambda = (1-\alpha)\frac{2}{\delta-2}$, where α is positive,

$$\tau(\epsilon) \le \frac{48n}{\alpha}\log(3n\delta/\epsilon)$$

Theorem 2 *For triangle-free graphs of maximum degree δ, MC mixes in time $O(\frac{n}{\alpha}\log(n\delta/\epsilon))$ when $\lambda = (1-\alpha)\frac{2}{\delta-2}$ for positive $\alpha < 1$.*

We refer the reader to the discusion at the end of section 1 for a very brief discussion relating to the following corollary.

Corollary 3 *For triangle-free lattices of degree δ, the limiting Gibbs measure is unique when $\lambda < (1-\alpha)\frac{2}{\delta-2}$ for fixed positive $\alpha < 1$.*

4 Hardness of Approximation

In this section, we detail how the standard boosting technique along with MAX-SNP-hardness of 4-MIS (finding the maximum independent set in graphs with maximum degree four) implies an NP-hardness result for computing the partition function. Recall that since 4-MIS is MAX-SNP-hard [21], there is an $\epsilon > 0$ such that no algorithm can guarantee an approximation factor better than $1+\epsilon$, unless $P = NP$.

Theorem 4 *Unless $RP = NP$, there is no algorithm to approximately compute Z_G within any polynomial factor when $\lambda > \frac{c}{\delta}$ for some constant $c > 0$.*

Proof. Our proof follows that of [26, Theorem 1.17].

Consider a graph G with maximum degree four. Let r be a positive integer which will be determined later. We boost the graph $G \to G_r$ by replacing each vertex v by an independent set C_v of size r and each edge $e = \{u, v\}$ by the complete bipartite graph $K_{r,r}$ between C_u and C_v. Notice that the maximum degree δ of G_r is $4r$.

Let $\mathcal{I}$ denote the independent sets of G_r. Each independent set $S \in \mathcal{I}$ is a witness for an independent set $w(S)$ in G where:

$$w(S) = \{v \in V : S \cap C_v \neq \emptyset\}$$

Suppose we have an *fpras* to compute Z_{G_r} and thus can generate a sample S from the hard-core measure [13]. For some constant $c > 0$, we show that for $\lambda > \frac{c}{\delta}$, it is likely that S is a witness to a large independent set in G. Specifically, with probability at least $\frac{1}{2}$, $|w(S)| \geq \frac{m}{1+\epsilon}$, where m is the size of the maximum of the independent set in G. This gives a randomized algorithm to approximate 4-MIS within a factor $1+\epsilon$.

To determine the sufficient parameters, we consider those $S \in \mathcal{I}$ which are witnesses to a small independent set. For $k = \frac{m}{1+\epsilon}$, let

$$S_{G_r} = \sum_{S \in \mathcal{I} : |w(S)| < k} \lambda^{|S|}$$

Notice that if $S_{G_r} \leq \frac{1}{2} Z_{G_r}$ then with probability at least $\frac{1}{2}$ our sample S is a witness to a large independent set. Straightforward bounds on Z_{G_r} and S_{G_r} turn out to be sufficient for our purposes. Since G has an independent set of size m, G_r has at least $\binom{mr}{l}$ independent sets of size l for all $l \leq mr$. Weighting these by λ, we have

$$Z_{G_r} \geq \sum_{l=0}^{mr} \binom{mr}{l} \lambda^l = (1+\lambda)^{mr}$$

We also know,

$$S_{G_r} \leq \sum_{i=0}^{m/(1+\epsilon)} \binom{n}{i} \sum_{j=0}^{ir} \binom{ir}{j} \lambda^j \leq (1+\lambda)^{mr/(1+\epsilon)} 2^n$$

Using these bounds we can determine for which r that $S_{G_r} \leq \frac{1}{2} Z_{G_r}$:

$$(1+\lambda)^{mr/(1+\epsilon)} 2^n \leq \frac{1}{2}(1+\lambda)^{mr}$$

$$2^{n+1} \leq (1+\lambda)^{mr(\epsilon/(1+\epsilon))} \leq e^{mr\lambda(\frac{\epsilon}{1+\epsilon})}$$

We know $m \geq \frac{n}{5}$ because the greedy algorithm finds an independent set of at least such a size. Thus, there is a gap when $r > \frac{c}{\lambda}$, for some $c > 0$. Since $\delta = 4r$, we have $S_{G_r} \leq \frac{1}{2} Z_{G_r}$ when

$\lambda > \frac{c'}{\delta}$, for a constant $c' > 0$. Notice that if we want to optimize the constant, we simply need $\frac{Z_{G_r} - S_{G_r}}{Z_{G_r}} \geq \frac{1}{poly(n)}$. Thus, we have $c' \leq \frac{20(1+\epsilon)}{\epsilon}$ where ϵ is from the hardness of approximation result for 4-MIS.

This shows that unless $RP = NP$, there is no *fpras* to compute Z_{G_r} when $\lambda > \frac{c'}{\delta}$. It is shown in [26] that an algorithm which approximates Z_{G_r} within any polynomial factor can be boosted to obtain an *fpras* for Z_{G_r}. □

5 Comments on Extension to Graphs with Triangles

In this section we give a glimpse of the proof given in [28] for the extension to arbitrary graphs by showing the potential function used in the analysis. The proof still uses coupling, but not path coupling. Instead we define a potential function $\Theta = \Theta(\sigma, \eta)$ between an arbitrary pair of states $\sigma, \eta \in \Omega$. Let D denote the set of disagree vertices, i.e. D is the symmetric difference between σ and η. We also use d_v to denote the number of disagree neighbors of v, i.e. $d_v = |D \cap \Gamma(v)|$.

$$\alpha_v = \left\{ \begin{array}{ll} \delta_v & \text{if } v \in D \\ 0 & \text{otherwise} \end{array} \right\}$$

$$\beta_v = \left\{ \begin{array}{ll} -cd_v & \text{if there exists a neighbor } w \text{ of } v \text{ such that } w \in \sigma, w \in \eta \\ -c(d_v - 1) & \text{if there is no such } w \text{ and } d_v > 1 \\ 0 & \text{otherwise} \end{array} \right\}$$

$$\Theta = \sum_v [\alpha_v + \beta_v]$$

References

[1] M. Aizenman and R. Holley. Rapid convergence to equilibrium of stochastic Ising models in the Dobrushin-Shlosman regime. In *Percolation theory and ergodic theory of infinite particle systems*, volume 8 of *The IMA Volumes in Mathematics and its Applications*, pages 1–11. Springer, New York-Berlin, 1987.

[2] R.J. Baxter, I.G. Entig, and S.K. Tsang. Hard-square lattice gas. *Journal of Statistical Physics*, 22(4):465–489, 1980.

[3] J. van den Berg and J.E. Steif. Percolation and the hard-core lattice gas model. *Stochastic Processes and their Applications*, 49:179–197, 1994.

[4] Piotr Berman and Marek Karpinski. On some tighter inapproximability results. Technical Report TR98-029, ECCC, 1998.

[5] Russ Bubley and Martin Dyer. Path coupling, Dobrushin uniqueness, and approximate counting. In *38th Annual Symposium on Foundations of Computer Science*, Miami Beach, FL, 19–22 October 1997. IEEE.

[6] R. L. Dobrushin. The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. *Functional Analysis and its Applications*, 2(4):302–312, 1968.

[7] R.L. Dobrushin, J. Kolafa, and S.B. Shlosman. Phase diagram of the two-dimensional Ising antiferromagnet, computer-assisted proof. *Communications in Mathematical Physics*, 102(1):89–103, 1985.

[8] R.L. Dobrushin and S.B. Shlosman. Constructive criterion for the uniqueness of Gibbs field. In *Statistical physics and dynamical systems*, volume 10 of *Progress in Physics*, pages 347–370. Birkhäuser, Boston, 1985.

[9] Martin Dyer and Catherine Greenhill. On Markov chains for independent sets. preprint, October 1997.

[10] Hans-Otto Georgii. *Gibbs measures and phase transitions*, volume 9 of *de Gruyter Studies in Mathematics*. Walter de Gruyter and Co., Berlin-New York, 1988.

[11] O. Häggström and K. Nelander. Exact sampling from anti-montone systems. preprint, 1997.

[12] Johan Hastad. Clique is hard to approximate within $n^{1-\epsilon}$. In *37th Annual Symposium on Foundations of Computer Science*, Burlington, Vermont, 14–16 October 1996. IEEE.

[13] Mark R. Jerrum, Leslie G. Valiant, and Vijay V. Vazirani. Random generation of combinatorial structures from a uniform distribution. *Theoretical Computer Science*, 43(2–3):169–188, 1986.

[14] Richard M. Karp and Michael Luby. Monte-Carlo algorithms for enumeration and reliability problems. In *24th Annual Symposium on Foundations of Computer Science*, pages 56–64, Tucson, Arizona, 7–9 November 1983. IEEE.

[15] Frank Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.

[16] A.B. Kirillov, D.C. Radulescu, and D.F Styer. Vassertein distances in two-state systems. *Journal of Statistical Physics*, 56:931–937, 1989.

[17] Graham Louth. *Stochastic Networks: Complexity, dependence, and routing.* PhD thesis, Univ. Cambridge, 1990.

[18] Michael Luby and Eric Vigoda. Approximately counting up to four (extended abstract). In *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, pages 682–687, El Paso, Texas, 4–6 1997.

[19] Fabio Martinelli and Enzo Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. i. the attractive case. *Communications in Mathematical Physics*, 161(3):447–486, 1994.

[20] Fabio Martinelli and Enzo Olivieri. Approach to equilibrium of Glauber dynamics in the one phase region. ii. the general case. *Communications in Mathematical Physics*, 161(3):487–514, 1994.

[21] Christos H. Papadimitriou and Mihalis Yannakakis. Optimization, approximation, and complexity classes (extended abstract). In *Proceedings of the Twentieth Annual ACM Symposium on Theory of Computing*, pages 229–234, Chicago, Illinois, 2–4 May 1988. ACM.

[22] Jim Propp and David Wilson. Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures and Algorithms*, 9(1):223–252, 1996.

[23] Jim Propp and David Wilson. Coupling from the past: a user's guide. preprint, 1998.

[24] D.C. Radulescu and D.F. Styer. The Dobrushin-Shlosman phase uniqueness criterion and applications to hard squares. *Journal of Statistical Physics*, 49:281–295, 1987.

[25] Dana Randall and Prasad Tetali. Analyzing Glauber dynamics by comparison of Markov chains. In *Third Latin American Symposium on Theoretical Informatics*, Campinas, Brazil, April 1998. UNICAMP.

[26] Alistair Sinclair. *Algorithms for Random Generation and Counting: a Markov chain approach.* Birkhäuser, Boston, 1993.

[27] Salil Vadhan. The complexity of counting. Undergraduate Thesis, Harvard University, 1995.

[28] Eric Vigoda. Fast convergence of the Glauber dynamics for sampling independent sets: Part II. Technical Report TR-99-003, International Computer Science Institute, January 1998.