Slow mixing of Glauber dynamics for the hard-core model on the hypercube

David Galvin *   Prasad Tetali †

Abstract

For $\lambda > 0$, let π_λ be the probability measure on the independent sets of the hypercube $\{0,1\}^d$ in which I is chosen with probability proportional to $\lambda^{|I|}$. We study the Glauber dynamics, or single-site-update Markov chain, whose stationary distribution is π_λ, and show that for values of λ tending to 0 as d grows, the convergence to stationarity is exponentially slow in the volume of the cube. The proof combines a conductance argument with combinatorial enumeration methods.

1 Introduction

We consider the well known hard-core model on a finite graph, with activity λ. Our particular focus is the mixing time of the Glauber dynamics, or single-site update Markov chain, for this model.

In [1] it was shown that for λ growing exponentially with d, the Glauber dynamics for the hard-core model on the discrete torus $[-L,L]^d$ (with the obvious adjacency) mixes exponentially slowly in the surface area of the torus.

In light of a recent result of Galvin and Kahn [3], it is tempting to believe that slow mixing on the torus should hold for much smaller values of λ. The main result of [3] is that the hard-core model on $\mathbf{Z}^d$ exhibits multiple Gibbs phases for $\lambda = \omega(\log^{3/4} d/d^{1/4})$. This suggests that for λ in this range, the typical independent set chosen from the torus according to the hard-core distribution is either predominantly odd (defined in the obvious way: $x = (x_1, \ldots, x_d)$ is odd if $\sum_i x_i$ is odd) or predominantly even. Thus there is an unlikely "bottleneck" set of balanced independent sets separating the predominantly odd sets from the predominantly even ones. This bottleneck should cause the conductance of the Glauber dynamics chain to be small, and so cause its mixing time to be large.

In this paper, by studying a simpler but related graph, we provide some evidence for slow mixing on the discrete torus at small values of λ. We outline a proof that the Glauber dynamics for the hard-core model on the d-dimensional hypercube $\{0,1\}^d$ mixes exponentially slowly in the volume of the cube for $\lambda = \omega(\log d/d^{1/4})$.

2 Statement of the result

For a graph $\Sigma = (V,E)$, write $\mathcal{I}(\Sigma)$ for the collection of independent sets (sets of vertices spanning no edges) in V. For $\lambda > 0$ define the *hard-core measure with activity λ on $\mathcal{I}(\Sigma)$* by

$$\pi_\lambda(\{I\}) = \pi_\lambda(\Sigma)(\{I\}) = \lambda^{|I|}/Z_\lambda(\Sigma) \quad \text{for } I \in \mathcal{I},$$

where $Z_\lambda(\Sigma) = \sum_{I\in\mathcal{I}(\Sigma)} \lambda^{|I|}$. This is the stationary distribution of the Markov chain $\mathcal{M}_\lambda = \mathcal{M}_\lambda(\Sigma)$ with transition probabilities $P_\lambda(I,J)$ $(I,J \in \mathcal{I}(\Sigma))$ given by

$$P_\lambda(I,J) = \begin{cases} 0 & \text{if } |I \triangle J| > 1 \\ \frac{1}{|V|}\frac{\lambda}{1+\lambda} & \text{if } I \subset J \\ \frac{1}{|V|}\frac{1}{1+\lambda} & \text{if } J \subset I \\ 1 - \sum_{J \neq J'} P_\lambda(I,J') & \text{if } I = J. \end{cases}$$

(Here $I \subset J$ means $|I \triangle J| = 1, I \subseteq J$.) This chain is often referred to as the *Glauber dynamics* on $\mathcal{I}(\Sigma)$.

Write $P^t(I,\cdot)$ for the distribution of the chain at time t, given that it started in state I, and set $VAR_I(t) = \sum_{J\in\mathcal{I}(\Sigma)} |P^t(I,J) - \pi_\lambda(J)|$. The *mixing time* $\tau_{\mathcal{M}_\lambda} = \tau_{\mathcal{M}_\lambda}(\Sigma)$ of $\mathcal{M}_\lambda$, which measures the speed at which the chain converges to stationarity, is, as usual,

$$\tau_{\mathcal{M}_\lambda} = \max_{I\in\mathcal{I}(\Sigma)} \min\left\{t : VAR_I(t') \leq \frac{1}{e}\ \forall t' > t\right\}.$$

Our main result concerns $\tau_{\mathcal{M}_\lambda}(Q_d)$, where Q_d is the usual discrete hypercube (the graph on vertex set $\{0,1\}^d$ in which pairs of strings are adjacent iff they differ on exactly one coordinate). Writing M for 2^{d-1}, we have

THEOREM 2.1. *There are constants $c_1, c_2, c_3 > 0$ such that for all d and for $\lambda > c_1 \log d/d^{1/4}$,*

$$\tau_{\mathcal{M}_\lambda}(Q_d) > 2^{c_2 M \frac{\min\{\lambda^2,1\}\log^2(1+\lambda)}{\sqrt{d}(\log(1+\lambda)+c_3\log d)}}.$$

Remark 1*:* Our bound on λ is no doubt not optimal. In the other direction, a result of Luby and Vigoda [5] implies that $\tau_{\mathcal{M}_\lambda}(Q_d)$ is a polynomial in M for $\lambda \leq 2/(d-2)$.

Remark 2*:* We can prove an analog of Theorem 2.1 for any family of regular graphs with bounded co-degree (every pair of vertices having a bounded number of common neighbours) and reasonable expansion.

*Microsoft Research, 1 Microsoft Way, Redmond WA 98052.

†Georgia Tech, Atlanta GA 30332-0160; research supported in part by NSF grant DMS-0100289.

3 Outline of the proof

We assume throughout that d is large enough to support our assertions. For simplicity we assume $\lambda \leq 1$; we may deal with $\lambda > 1$ similarly. The proof combines a conductance argument (introduced to the study of Markov chains in [4], and used in [1]) with a combinatorial enumeration argument based on ideas of Sapozhenko [6]. This argument was used in [7] to show that $|\mathcal{I}(Q_d)| \sim 2\sqrt{e}2^M$. (See also [2, 3] for recent applications.)

Write $\mathcal{E}$ and $\mathcal{O}$ for the bipartition classes of Q_d and set $\mathcal{S} = \{I \in \mathcal{I}(Q_d) : |I \cap \mathcal{E}| = |I \cap \mathcal{O}|\}$. The removal of $\mathcal{S}$ splits $\mathcal{I}(Q_d)$ into two sets of equal measure with the property that the Glauber dynamics cannot pass between the two without visiting $\mathcal{S}$. By a well known conductance argument (see [1, 4]), it follows that

$$\tau_{\mathcal{M}_\lambda}(Q_d) \geq (1 - \pi_\lambda(\mathcal{S})^2)/4\pi_\lambda(\mathcal{S}).$$

Thus to prove Theorem 2.1 it suffices to show that

$$\pi_\lambda(\mathcal{S}) < 2^{-c_2 M \frac{\lambda^2 \log^2(1+\lambda)}{\sqrt{d}(\log(1+\lambda)+c_3 \log d)}}/8. \tag{3.1}$$

Set $\mathcal{S}^{triv} = \{I \in \mathcal{S} : |I \cap \mathcal{E}| \leq \alpha M\}$ and $\mathcal{S}^{nt} = \mathcal{S} \setminus \mathcal{S}^{triv}$, where $\alpha = \lambda^2/100$. Using the trivial lower bound $Z_\lambda(Q_d) \geq 2(1+\lambda)^M$ we have

$$\begin{aligned}\pi_\lambda(\mathcal{S}^{triv}) &\leq \sum_{i=0}^{\alpha M} \lambda^{2i} \binom{M}{i}^2 /2(1+\lambda)^M \\ &< 2^{-M/2}. \end{aligned} \tag{3.2}$$

We now turn to $\mathcal{S}^{nt}$. For $A \subseteq Q_d$, define the *closure* of A to be $[A] = \{v \in Q_d : \partial\{v\} \subseteq \partial A\}$, where ∂A denotes the set of vertices adjacent to A. Say that $I \in \mathcal{I}(Q_d)$ is *small on* $\mathcal{E}$ if $|[I \cap \mathcal{E}]| \leq M/2$, and set

$$\mathcal{S}^{nt}_{\mathcal{E}} = \{I \in \mathcal{S}^{nt} : I \text{ is small on } \mathcal{E}\}.$$

Define *small on* $\mathcal{O}$ and $\mathcal{S}^{nt}_{\mathcal{O}}$ similarly. That it is easier to work with $[A]$ than with A is a crucial idea introduced in [6]. A simple argument, based on the fact that Q_d has a perfect matching, shows that any $I \in \mathcal{I}(Q_d)$ must be small on at least one of $\mathcal{E}$, $\mathcal{O}$, and so

$$\pi_\lambda(\mathcal{S}^{nt}) \leq 2\max\{\pi_\lambda(\mathcal{S}^{nt}_{\mathcal{E}}), \pi_\lambda(\mathcal{S}^{nt}_{\mathcal{O}})\} = 2\pi_\lambda(\mathcal{S}^{nt}_{\mathcal{E}}).$$

For each $M/2 \geq a \geq \alpha M$ and $g \geq a$ set

$$\mathcal{A}(a,g) = \{A \subseteq \mathcal{E} : |[A]| = a, |N(A)| = g\}.$$

Note that by an isoperimetric inequality in the cube (see e.g. [7]), there is a constant c_4 such that $\mathcal{A}(a,g) = \emptyset$ unless $g > (1 + c_4/\sqrt{d})a$. We have

$$\begin{aligned}\pi_\lambda(\mathcal{S}^{nt}_{\mathcal{E}}) &\leq \sum_{a,g} \pi_\lambda(\mathcal{A}(a,g))(1+\lambda)^{M-g}/(1+\lambda)^M \\ &\leq M^2 \max_{a,g} \pi_\lambda(\mathcal{A}(a,g))(1+\lambda)^{-g}.\end{aligned}$$

We now come to the main lemma, whose (omitted) proof draws on ideas of Sapozhenko [6, 7]. It may be considered a weighted version of a lemma from [6].

LEMMA 3.1. *There are constants $c_1, c_3, c_5 > 0$ such that for any a, g and $\lambda > c_1 \log d/d^{1/4}$ we have*

$$\pi_\lambda(\mathcal{A}(a,g)) \leq 2^{c_5 M \frac{\log d}{d^2} - \frac{(g-a)\log^2(1+\lambda)}{2\log(1+\lambda)+c_3 \log d}}(1+\lambda)^g.$$

The idea of the proof is as follows. For every $1 \leq \psi = o(d)$, we construct, using a combination of probabilistic and algorithmic arguments, a set $\mathcal{B}(a,g)$ of size at most $2^{O(M\frac{\log d}{d^2} + (g-a)\frac{\log d}{\psi})}$ with two properties. The first is that each $A \in \mathcal{A}(a,g)$ is "approximated" in an appropriate sense by some $B \in \mathcal{B}(a,g)$. The second is that for each $B \in \mathcal{B}(a,g)$, the measure of the set of A's in $\mathcal{A}(a,g)$ that B can approximate is at most $(1+\lambda)^{g-\gamma(g-a)}$, where $\gamma = \frac{d\log(1+\lambda) - c_3\psi\log d}{d(\log(1+\lambda)+c_3\log d)}$. Optimizing over the choice of ψ, we obtain Lemma 3.1.

Since $g - a \geq 0$ always, we maximize the exponent of the bound in Lemma 3.1 by taking $g - a$ as small as possible. For the range of values under consideration we have $g - a > c_4 a/\sqrt{d} > c_4\lambda^2 M/100\sqrt{d}$, and so

$$\begin{aligned}\pi_\lambda(\mathcal{S}^{nt}_{\mathcal{E}}) &\leq M^2 2^{c_5 M\frac{\log d}{d^2} - \frac{c_4 M \lambda^2 \log^2(1+\lambda)}{200\log(1+\lambda)+c_3\log d}} \\ &\leq 2^{-c_2 M \frac{\lambda^2\log^2(1+\lambda)}{\sqrt{d}(\log(1+\lambda)+c_3\log d)}}/32,\end{aligned}$$

This gives $\pi_\lambda(\mathcal{S}^{nt}) \leq 2^{-c_2 M \frac{\lambda^2\log^2(1+\lambda)}{\sqrt{d}(\log(1+\lambda)+c_3\log d)}}/16$ which, combined with (3.2) gives (3.1) and Theorem 2.1.

References

[1] C. Borgs, J. Chayes, A. Frieze, J.H. Kim, P. Tetali, E. Vigoda, V. Vu, Torpid Mixing of some Monte Carlo Markov Chain algorithms in Statistical Physics, *FOCS '99*, 218–229.

[2] D. Galvin, On homomorphisms from the Hamming cube to **Z**, to appear.

[3] D. Galvin and J. Kahn, On phase transition in the hard-core model on $\mathbf{Z}^d$, to appear.

[4] M. Jerrum and A. Sinclair, The Monte Carlo Markov Chain method: an approach to approximate counting and integration, in *Approximation Alorithms for NP-hard problems*, PWS, 1996.

[5] M. Luby and E. Vigoda, Fast convergence of the Glauber dynamics for sampling independent sets, *Random Structures and Algorithms* **15** (1999), 229–241.

[6] A. Sapozhenko, On the number of connected subsets with given cardinality of the boundary in bipartite graphs, *Metody Diskret. Analiz.* **45** (1987), 42–70. (Russian)

[7] A. Sapozhenko, The number of antichains in ranked partially ordered sets, *Diskret. Mat.* **1** (1989), 74–93. (Russian; translation in Discrete Math. Appl. 1 no. 1 (1991), 35–58.)