Computational Transition at the Uniqueness Threshold

Allan Sly
Microsoft Research
Redmond, USA
allansly@microsoft.com

***Abstract*—The hardcore model is a model of lattice gas systems which has received much attention in statistical physics, probability theory and theoretical computer science. It is the probability distribution over independent sets I of a graph weighted proportionally to $\lambda^{|I|}$ with fugacity parameter λ. We prove that at the uniqueness threshold of the hardcore model on the d-regular tree, approximating the partition function becomes computationally hard on graphs of maximum degree d.**

Specifically, we show that unless NP=RP there is no polynomial time approximation scheme for the partition function (the sum of such weighted independent sets) on graphs of maximum degree d for fugacity $\lambda_c(d) < \lambda < \lambda_c(d) + \varepsilon(d)$ where

$$\lambda_c = \frac{(d-1)^{d-1}}{(d-2)^d}$$

is the uniqueness threshold on the d-regular tree and $\varepsilon(d) > 0$ is a positive constant. Weitz [36] produced an FPTAS for approximating the partition function when $0 < \lambda < \lambda_c(d)$ so this result demonstrates that the computational threshold exactly coincides with the statistical physics phase transition thus confirming the main conjecture of [30]. We further analyze the special case of $\lambda = 1, d = 6$ and show there is no polynomial time approximation scheme for approximately counting independent sets on graphs of maximum degree $d = 6$, which is optimal, improving the previous bound of $d = 24$.

Our proof is based on specially constructed random bipartite graphs which act as gadgets in a reduction to MAX-CUT. Building on the involved second moment method analysis of [30] and combined with an analysis of the reconstruction problem on the tree our proof establishes a strong version of "replica" method heuristics developed by theoretical physicists. The result establishes the first rigorous correspondence between the hardness of approximate counting and sampling with statistical physics phase transitions.

I. Introduction

The hardcore model is a model from statistical physics representing hardcore interaction of gas particles. It is a probability distribution on independent sets I of a graph weighted as $\frac{1}{Z}\lambda^{|I|}$ where λ is a positive parameter called the fugacity and Z is a normalizing constant called the partition function. Physicists and probabilists have done extensive work towards identifying the phase transitions and other properties of the model.

In computational complexity approximately counting (weighted) independent sets is a central problem. The hardcore model is of key importance as this is exactly the problem of producing an FPRAS (fully polynomial randomized approximation scheme) for Z, the partition function. When λ is small the hardcore model has rapid decay of correlations and the partition function can be approximated either using MCMC or through computational tree methods [36]. For larger fugacities long range dependencies may appear and the problem is known to be hard when λ is sufficiently large.

In this paper we determine a computational threshold where approximating Z becomes hard. Using an ingenious computational tree approach Weitz [36] produced a PTAS for approximating Z when $\lambda < \lambda_c(d)$ where

$$\lambda_c(d) = \frac{(d-1)^{d-1}}{(d-2)^d}$$

is the uniqueness threshold for the hardcore model on the infinite d-regular tree [18]. This corresponds to the point at which long range dependencies become possible in the model (formally defined in Section I-C). Mossel, Weitz and Wormald [30] showed that beyond this phase transition local MCMC algorithms fail and conjectured that it gives the threshold for computations hardness. While such statistical physics phase transitions are believed to coincide with the transition in computational hardness of approximating the partition function for a number of important models no such examples had been proven. Our main result essentially confirms the conjecture of [30] giving the first such rigorous example.

Theorem 1. *For every $d \geq 3$ there exists $\varepsilon(d) > 0$ such that when $\lambda_c(d) < \lambda < \lambda_c(d) + \varepsilon(d)$, unless NP=RP, there does not exist an FPRAS for the partition function of the hardcore model with fugacity λ for graphs of maximum degree at most d.*

While we believe the result holds for all $\lambda > \lambda_c$, for technical reasons (specifically showing that an explicit function of three variables attains its maximum at a prescribed location, see Section I-C1 for details) the result is limited to λ close to criticality. This limitation notwithstanding, it clearly demonstrates the central role played by the uniqueness threshold.

When $\lambda = 1$ the hardcore model is simply the uniform distribution over independent sets and the partition function is simply the number of independent sets and as such this case is of particular interest. When $d \leq 5$ Weitz's result provides a FPRAS as $\lambda_c(d) > 1$. Conversely it is known that with $d \geq 25$ the problem is computationally hard [9]. While the case $d = 6, \lambda = 1$ does not fall within the scope of Theorem 1, using a computer assisted proof, we establish the necessary technical condition and prove the following result.

Theorem 2. *Unless NP=RP for every $d \geq 6$ there does not exist a fully polynomial approximation scheme for counting independent sets on graphs of maximum degree at most d.*

A. Background and Previous Results

Even on graphs of maximum degree 3 the problem of exactly counting independent sets is #P hard [13] and as such one can at most ask when it is possible to approximately count independent sets, that is when an FPRAS exists. As the model is self-reducible, approximate counting is equivalent to approximately sampling from the partition function [31]. This has led to a major line of research in analyzing the performance of MCMC techniques, particularly the Glauber dynamics.

When $\lambda \leq \frac{2}{d-2}$ the Glauber dynamics mixes rapidly [21] which in particular gives an FPRAS for counting independent sets on graphs of maximum degree at most 4 (see [10] for similar bounds). Weitz [36] showed that the hardcore model has a decay of correlation property called strong spatial mixing whenever $\lambda < \lambda_c$ which implies rapid mixing on graphs of sub-exponential growth. Moreover, his paper gives a deterministic polynomial time approximation scheme on all graphs when $\lambda < \lambda_c$ through a computational tree approximation.

Finding the ground state of the hardcore model, the largest independent set, is of course a canonical NP-hard problem and is hard to approximate even on regular graphs of degree 3 [4]. Intuitively the problem of counting becomes harder as λ grows as this places more mass on the larger, harder to find, independent sets and indeed such hardness results have been established. In [21] it was shown that there is no FPRAS (assuming NP$\neq$RP) when $\lambda \leq c/d$ for $c \approx 10000$. In the case of $\lambda = 1$ this was improved to $d \geq 25$ in [9] using random regular bi-partite graphs as basic gadgets in a hardness reduction. They further showed that with high probability the mixing time of the Glauber dynamics on a random bipartite d-regular graph is exponential in the size of the graph. Calculations of [9] led the authors there to speculate that λ_c may be the threshold for hardness but the evidence was not conclusive enough to make such a conjecture.

1) Replica Heuristics: The replica and cavity methods and heuristics have provided powerful tools (often non-rigorous) in the study of a wide range of random optimization problems and predictions for the behavior of spin glasses and dilute mean fields spin systems [24, 25]. Developed by theoretical physcicits, in in some cases these heuristics have been made rigorous, notably the SK model [35], solution space of solutions to random constraint satisfaction problems [1] and the assignment problem [2]. In dilute spin glass models such methods have given rise to powerful new algorithms such as survey propagation (see e.g. [20]).

Random regular bi-partite graphs are widely known to be locally tree-like with only a small number of short cycles. The statistical physics theory makes the following predictions for the hardcore model on typical random bipartite d-regulars. The first is that the model is expected to exhibit spontaneous symmetry breaking for $\lambda > \lambda_c$. When $\lambda < \lambda_c$ correlations decay exponentially and the configuration (independent set) is essentially balanced between the two halves of the bi-partite graph. By contrast when $\lambda > \lambda_c$ the configuration separates its mass unevenly placing $\Omega(n)$ more mass on one side or the other. Configurations with a roughly equal proportion of sites on each side make up only an exponentially small fraction of the distribution. This is intuitively plausible as the largest bi-partite sets will be those containing most of one side of the graph or the other.

The second is that this symmetry breaking splits the configuration space into two "pure states" of roughly equal probability. We will denote the "phase" of the configuration as the side of the graph with more sites. Conditional on the phase the spins of randomly chosen vertices are assumed to be asymptotically independent and the local neighbourhood of the configuration are given by extremal measures. This conditional independence is a crucial element of cavity-method type arguments.

A first moment analysis of [9] suggested that configurations obey the first prediction but their proof proceeded without specifically proving it. In a technical tour de force the prediction was rigorously established for $\lambda_c(d) < \lambda < \lambda_c(d) + \varepsilon(d)$ in [30] using an involved second moment method analysis together with the small graph conditioning method. The restriction to the region $\lambda < \lambda_c(d) + \varepsilon(d)$ is somewhat surprising at first as the problem ought to become easier as λ grows. It is the result of a technical difficulty in estimating the second moment bound. Even establishing this for λ close to the critical value took up fully a third of the proof. As a central part of our proof is a modification of this method the same restriction applies.

Based on establishing the symmetry breaking [30] showed that any local reversible Markov Chain has mixing time exponential in the number of vertices by establishing a bottleneck in the mixing on asymptotically almost all random d-regular bi-partite graphs. This bound is tight as subsequent results [29, Theorem 4] imply rapid mixing on almost all random bi-partite graphs when $\lambda < \lambda_c(d)$. Based on these finding they made the following conjecture.

Conjecture I.1. *(*[30]*) Unless NP=RP for every $d \geq 4$ and $\lambda_c(d) < \lambda$ there does not exist a fully polynomial approximation scheme for the partition function of the hardcore model with fugacity λ for graphs of maximum degree at most d.*

Phase transitions of spin systems have been known to exactly determine the region of rapid mixing in a number of systems including the ferromagnetic Ising model on $\mathbb{Z}^2$ [22] and on the d-regular tree [3]. The first such example on completely general bounded degree graphs was recently established by Mossel and the present author [29] showing rapid mixing of the Glauber dynamics of the ferromagnetic Ising model on graphs of maximum degree d when $(d-1)\tanh\beta < 1$. The threshold $(d-1)\tanh\beta = 1$ is a statistical physics phase transition, the uniqueness threshold for the Ising model on the d-regular tree.

Slow mixing of MCMC algorithms do not by themselves imply hardness of approximating the partition function. Indeed, in the ferromagnetic Ising model the mixing time of local reversible Markov chains may be exponential but nonetheless there is an FPRAS by the famous algorithm of Jerrum and Sinclair [17]. However, unlike the hardcore model or indeed the anti-ferromagnetic Ising model which do exhibit phase transitions, the ground states of the ferromagnetic Ising model are trivially found.

While phase transitions exists on many infinite graphs, it is the uniqueness threshold on the tree that appears to determine the onset of computational hardness in general graphs in a number of models as they represent the extreme case for correlation decay in graphs for many models. Sokal [32] conjectured that uniqueness on the d-regular tree for the hardcore model implies uniqueness on any graph of maximum degree d. This conjecture was established in [36] which further showed that for *any* 2-spin system strong spatial mixing on the d-regular tree implies strong spatial mixing on all graphs of maximum degree d. Indeed for most, although not all, spin systems the regular tree is expected to be the limiting case for extreme correlations amongst all graphs of maximum degree d (see e.g. [33] for more details). The emergence of long range correlations appears to be a necessary prerequisite for hardness of sampling and this motivates the conjectures that the uniqueness threshold on the tree determines the onset of computational hardness.

In this paper we establish a form of the second heuristic prediction on a modified random bipartite graph. We show that on a polynomial sized set of vertices the spins are close to a product measure, conditional on the phase in the L^∞ distance on measures. Being able to treat large numbers of vertices as conditionally independent given the phase plays a key role in our reduction. While some results of this nature have been established previously (see e.g. [8, 27]) this is the first example we are aware of where the number of conditionally independent sites grows polynomially in the size of the graph.

Another recent result which makes use of phase transitions is by Goldberg and Jerrum [14] who show that approximating the partition function for the Potts model is $\#RH\Pi_1$-hard when $q > 2$. Their proof crucially uses the first order phase transition of the Potts model on the complete graph when $q > 2$. The special case $q = 2$ is the Ising model which in P as noted above.

B. Proof Techniques

We now give a sketch of the proof. Some details follow in Section II and a complete proof can be found online at `http://arxiv.org/abs/1005.5584`.

Following the approach of [9] and as suggested in [30] we utilize random bi-partite graphs as basic gadgets in a hardness reduction. In those papers the basic unit of the construction is the random d-regular bipartite graph. To obtain a sharp result we cannot afford to add edges to such graphs (creating degree $d+1$ vertices) so our basic gadgets are bi-partite random graphs, most of whose vertices are degree d but with a small number of degree $d-1$ vertices which are used to connect to other gadgets.

We begin by constructing a graph $\tilde{G}$ which is a random bipartite graph with n vertices of degree d and $m' \approx n^{\theta+\psi}$ vertices of degree $d-1$ where θ, ψ are small positive parameters. We label the sides as "plus" and "minus" and edges are chosen according to random matchings of the vertices on the two sides. We denote the phase of the configuration (the random independent set) to be plus or minus according to the side which has more elements of the set amongst the degree d vertices.

With U denoting the set of vertices of degree $(d-1)$ we consider the random partition functions $Z^{\pm}(\eta)$ giving the sum over $\lambda^{|\sigma|}$ over all configurations with phase $\pm$ and with $\sigma_U = \eta$ where $\eta \in \{0,1\}^U$. We show that in expectation the $\mathbb{E}Z^{\pm}(\eta)$ are essentially proportional to the probabilities of a product measure on U whose marignals are given by the marginals of extremal Gibbs measures for the hardcore model on the $(d-1)$-ary tree. Our proof requires that this holds approximately for the $Z^{\pm}(\eta)$ themselves and adopt the second moment approach of [30] including their use of the small graph conditioning method [37]. While still involved, by estimating ratios of quantities in our model to quantities calculated in [30] we greatly simplify these computations. We are, however, still left with the same technical condition as [30] which we describe in the next subsection.

Even this approximate conditional independence is not sufficient for our reduction. To this end we construct a new random graph G by appending $(d-1)$-ary trees of height $\psi \log_{d-1} n$ onto U and denote the set of $m \approx n^{\theta}$ roots of the trees as V which are of degree $d-1$. Our proof proceeds to show that, conditional on the phase, σ_V is very close to a product measure. We note that appending the trees reweights the probabilities on configurations σ_U but it does so in a quantifiable way.

By construction the spins σ_V are conditionally independent given σ_U. Moreover, the statistical physics heuristics imply that the configuration of the neighbourhood around σ_V should be given by an extremal semi-translation invariant Gibbs measure on the tree with strong decay of correlation from the root to the leaves of the tree. Based on this intuition, we show that after conditioning on the phase the probability that σ_U has a non-negligible influence on σ_V is doubly exponentially small in the height. Through this we can establish its distribution with bounds in the L^{∞} norm. This is done by bounding the probability that the spins in a distant level influence the root using methods from the "reconstruction problem on the tree" (see e.g. [26, 34]).

The random graphs G constitutes our gadget. Given a graph H on up to $n^{\theta/4}$ vertices we construct H^G by taking a copy of G for each vertex of H. Then for every edge in H we connect $n^{3\theta/4}$ vertices between each side of V in the corresponding copies of G maintaining the maximum degree d. Since the spins in V are almost conditionally independent given the phase we can estimate the effect of adding these edges. An easy calculation shows that the most efficient arrangement is to have connected gadgets have opposite phases. The hardcore model on H^G puts most of its mass on configurations whose phases are solutions to MAX-CUT on H. Hence, by the equivalence of approximate counting and approximate sampling, this gives a randomized reduction to MAX-CUT.

C. Preliminaries

For a finite graph G with edge set $E(G)$ the independent sets are subsets of the vertices containing no adjacent vertices or equivalently elements of the set of configurations

$$I(G) = \{\sigma \in [0,1]^G : \forall (u,v) \in E(G), \sigma_u \sigma_v = 0\}.$$

The *Hardcore Model* is a probability distribution over independent sets of a graph G defined by

$$\mathbb{P}_G(\sigma) = \frac{1}{Z_G(\lambda)} \lambda^{\sum_{v \in G} \sigma_v} \mathbb{1}_{\sigma \in I(G)} \tag{I.1}$$

where $Z_G(\lambda) = \sum_{\sigma \in I(G)} \lambda^{|\sigma|} = \sum_{\sigma \in I(G)} \lambda^{\sum_{v \in G} \sigma_v}$ is a normalizing constant known as the *partition function* and is a weighted counting of the independent sets. When $\lambda = 1$ the hardcore model is the uniform measure on independent sets and $Z_G(1)$ is the number of independent sets of the graph.

The definition of the hardcore model can be extended to infinite graphs by way of the DLR condition which essentially says that for every finite set A the configuration on A is given by the Gibbs distribution given by a random boundary generated by the measure outside of A. Such a measure is called a Gibbs measure and there may be more one or infinitely many such measures (see e.g. [12] for more details). When there is exactly one Gibbs measure we say the model has *uniqueness*. Equivalently, the model has uniqueness if the marginal spin at any vertex is not affected by arbitrary conditioning the spins of sets of distant vertices as the distance goes to infinity. Our main result relates the uniqueness threshold on $\mathbb{T}_d$, the infinite d-regular tree, to the hardness of approximating the partition function on graphs of maximum degree d.

The hardcore model on $\mathbb{T}_d$ undergoes a phase transition at $\lambda_c(d) = \frac{(d-1)^{d-1}}{(d-2)^d}$ with uniqueness when $\lambda \le \lambda_c$ and non-uniqueness when $\lambda > \lambda_c$ [18]. The following picture is described in [30]. For every λ there exists a unique translation invariant Gibbs measure $\mu = \mu_{d,\lambda}$ known as the free measure with occupation density $p^* = \mu(\sigma_\rho)$ for ρ the root of the tree. When $\lambda > \lambda_c$ there also exist two semi-translation invariant (that is invariant under parity preserving automorphisms of $\mathbb{T}_d$) measures μ_+ and μ_- whose occupation densities we denote by $p^+ = \mu_+(\sigma_\rho), p^- = \mu_-(\sigma_\rho)$. These measures are obtained by conditioning on level 2ℓ

(resp. $2\ell+1$) of the tree to be completely occupied and taking the weak limit as $\ell \to \infty$.

It will also be of use to discuss related measures on the infinite $(d-1)$-ary tree $\hat{\mathbb{T}}^d$ rooted at ρ. We define analogously the measures $\hat{\mu}_+$ and $\hat{\mu}_-$ obtained by conditioning on level 2ℓ (resp. $2\ell+1$) of $\hat{\mathbb{T}}_d$ to be completely occupied and taking the weak limit as $\ell \to \infty$. We set q^+ and q^- to be the respective occupation densities $q^+ = \hat{\mu}_+(\sigma_\rho), q^- = \hat{\mu}_-(\sigma_\rho)$ of the root ρ.

The measure $\mu_\pm$ and $\hat{\mu}_\pm$ are naturally related as follows. Let v be a child of ρ and denote $\mathbb{T}_v$ to be the subtree of $\mathbb{T}^d$ rooted at v. There is a natural identification of $\mathbb{T}^d \setminus \mathbb{T}_v$ with the $(d-1)$-ary tree $\hat{\mathbb{T}}^d$ and under this identification the measures satisfy

$$\hat{\mu}_\pm(\sigma \in \cdot) = \mu_\pm(\sigma_{\mathbb{T}^d \setminus \mathbb{T}_v} \in \cdot | \sigma_v = 0). \tag{I.2}$$

In particular since $\sigma_\rho = 1$ implies $\sigma_v = 0$ for an independent set in $\mathbb{T}^d$ it follows that

$$q^\pm = \frac{p^\pm}{1-p^\mp}. \tag{I.3}$$

Furthermore, standard tree recursions for Gibbs measures (see e.g. [30]) establish that

$$q^\pm = \frac{\lambda(1-q^\mp)^{d-1}}{1+\lambda(1-q^\mp)^{d-1}}$$

and consequently by equation (I.3),

$$\frac{q^\pm}{1-q^\pm} = \lambda(1-q^\mp)^{d-1} = \lambda\left(\frac{1-p^\pm-p^\mp}{1-p^\pm}\right)^{d-1}. \tag{I.4}$$

It is shown in [30, Section 4] and [9, Claim 2.2] that the following hold for $\lambda > \lambda_c$:

1) The solutions to $h(\alpha) = \beta, h(\beta) = \alpha$ with $(\alpha, \beta) \in \mathcal{T} = \{(\alpha,\beta) : \alpha, \beta \geq 0, \alpha+\beta \leq 1\}$ where

$$h(x) = (1-x)\left[1 - \left(\frac{x}{\lambda(1-x)}\right)^{1/d}\right]$$

are exactly (p^+, p^-), (p^-, p^+) and (p^*, p^*). These densities satisfy $p^- < p^* < p^+$ and when $\lambda \downarrow \lambda_c$ we have that $p^*, p^+, p^- \to 1/d$.
2) The points (p^+, p^-) and (p^-, p^+) are the maxima of $\Phi_1(\alpha, \beta)$ in $\mathcal{T}$ where

$$\begin{aligned}
&\Phi_1(\alpha,\beta) = (\alpha+\beta)\log\lambda - \alpha\log\alpha \\
&- \beta\log\beta - d(1-\alpha-\beta)\log(1-\alpha-\beta) \\
&+ (d-1)\left((1-\alpha)\log(1-\alpha) + (1-\beta)\log(1-\beta)\right).
\end{aligned}$$

1) Technical Conditions: We now describe the technical condition necessary for our result. The function in question is

$$\begin{aligned}
&f(\alpha,\beta,\gamma,\delta,\varepsilon) = 2(\alpha+\beta)\log\lambda + H(\alpha) + H_1(\gamma,\alpha) \\
&+ H_1(\alpha-\gamma, 1-\alpha) + H(\beta) + H_1(\delta,\beta) \\
&+ H_1(\beta-\delta, 1-\beta) + d\Big[H_1(\gamma, 1-2\beta+\delta) - H(\gamma)
\end{aligned} \tag{I.5}$$

$$\begin{aligned}
&+ H_1(\varepsilon, 1-2\beta+\delta-\gamma) + H_1(\alpha-\gamma-\varepsilon, \beta-\delta) \\
&- H_1(\alpha-\gamma, 1-\gamma) + H_1(\alpha-\gamma, 1-\beta-\gamma-\varepsilon) \\
&- H_1(\alpha-\gamma, 1-\alpha)\Big]
\end{aligned} \tag{I.6}$$

where $H_1(x,y) = -x(\log x - \log y) + (x-y)(\log(y-x) - \log(y))$ and $H(x) = H(x,1)$ and where f is defined in the range $(\alpha,\beta) \in \mathcal{T}$ and

$$\alpha-\gamma-\varepsilon \geq 0, \beta-\delta \geq 0, 1-2\beta+\delta-\gamma-\varepsilon \geq 0. \tag{I.7}$$

which emerges naturally when calculating the second moment of the partition function.

Condition I.2. *The technical condition is that there exists a constant $\chi > 0$ such that when when $|p^+ - \beta|, |p^- - \alpha| < \chi$ the function $g_{\alpha,\beta}(\gamma,\delta,\varepsilon) = f(\alpha,\beta,\gamma,\delta,\varepsilon)$ attains its unique maximum in the set* (I.7) *at the point $(\gamma^*,\delta^*,\varepsilon^*) = (\alpha^2, \beta^2, \alpha(1-\alpha-\beta))$.*

The following result of [30] establishes Condition I.2 when $\lambda_c < \lambda < \lambda_c(d) + \varepsilon(d)$.

Lemma I.3 ([30, Lemma 6.10, Lemma 5.1])**.** *For each $d \geq 3$ there exists $\chi > 0$ such that when $|\alpha - \frac{1}{d}|, |\beta - \frac{1}{d}| < \chi$ then $g_{\alpha,\beta}(\gamma,\delta,\varepsilon)$ has a unique maximum at $(\gamma^*,\delta^*,\varepsilon^*)$ where $\gamma^* = \alpha^2, \delta^* = \beta^2, \varepsilon^* = \alpha(1-\alpha-\beta)$.*

We give a computer assisted proof which establishes Condition I.2 in the special case of $\lambda = 1$ and $d = 6$. Two other technical conditions we make use of in the proof are that

$$q^+q^-(d-1) < 1 \quad \text{and } q^+ < \frac{3}{5}. \tag{I.8}$$

Both conditions holds in the regions of interest as we have that $q^+, q^- \to \frac{1}{d-1}$ when $\lambda \downarrow \lambda_c$ and $q^+ \approx 0.423, q^- \approx 0.056$ when $\lambda = 1$ and $d = 6$. The first can be shown to hold for all $\lambda > \lambda_c$ with a somewhat involved proof while the latter is unnecessary but somewhat simplifies the proof.

D. Comments and Open Problems

The main open problem, of course, is to remove the $\lambda < \lambda_c(d) + \varepsilon(d)$ condition, ideally with a proof avoiding the second moment analysis. Alternatively, one could try and establish the Condition I.2 for all $\lambda > \lambda_c$ and $d \geq 3$.

Another natural problem is to establish the correspondence between computational hardness and phase transitions in the anti-ferromagnetic Ising model. While calculations of the style of [30] are not available and are likely to be even more challenging, it may be possible to avoid them. Indeed results of [27] already imply conditional local weak convergence of the configuration but not in a strong enough form to complete necessary reduction.

In Section II-A we detail the construction for G and give the reduction of MAX-CUT. In Section III we describe the analysis of the partition functions $Z^{\pm}(\eta)$ using the second moment method. In Section IV we discuss the analysis of the reconstruction problem on the tree and establish the conditional distributions of σ_V.

E. Acknowledgements

A.S. would like to thank Elchanan Mossel for his generous encouragement, guidance, support and advice with this project and also Dror Weitz for helpful discussions. The worked was initiated when the A.S. was a student at UC Berkeley where he was supported by NSF CAREER grant DMS-0548249 and by DOD ONR grant (N0014-07-1-05-06) 1300/08. Thanks to anonymous referees for useful comments.

II. Proof of Theorem 1 and 2

In this section we first describe the construction of our base random graph G which will be the basic gadget in our reduction. We state a theorem describing the properties of the hardcore model on G and then proceed to show how this establishes the reduction for Theorems 1 and 2.

A. Construction of G

We begin by constructing a random bi-partite (multi)graph $\tilde{G} = \tilde{G}(n, \theta, \psi)$ where n is a positive integer and $0 < \theta, \psi < \frac{1}{8}$ are positive constants which will be chosen to depend on λ and d. This graph will be the basis of our construction of G.

- The bipartite graph is constructed in two halves which we will call respectively the plus half and the minus half each with $n + m'$ vertices where $m' = (d-1)^{\lfloor \theta \log_{d-1} n \rfloor + 2\lfloor \frac{\psi}{2} \log_{d-1} \rfloor n}$.
- The vertices of each side are split into two sets $W^{\pm}$ and $U^{\pm}$ of size n and m' respectively. We label the vertices of $U^{\pm}$ by $u_1^{\pm}, \ldots, u_{m'}^{\pm}$.
- We connect $d-1$ edges to each vertex by taking $d-1$ random perfect matchings of $W^+ \cup U^+$ with $W^- \cup U^-$ and adding an edge between each pair of matched vertices.
- We take one more perfect matching of W^+ with W^- and add an edge between each pair of matched vertices.

In this construction the vertices in $W = W^+ \cup W^-$ are of degree d and the vertices in $U = U^+ \cup U^-$ are of degree $d-1$. Note that in this construction there will be multiple edges between vertices with asymptotically constant probability bounded away from 1. However, in the hardcore model multiple edges are irrelevant and we simply treat them as single edges (some degrees will be decreased but this will not affect our proof).

We now complete our construction of $G = G(n, \theta, \psi)$ by adjoining trees onto U^+ and to U^-.

- Construct a collection of $m = (d-1)^{\lfloor \theta \log_{d-1} n \rfloor}$ disconnected $(d-1)$-ary trees of depth $2\lfloor \frac{\psi}{2} \log_{d-1} n \rfloor$ rooted at $v_1^+, \ldots, v_m^+$. The total number of leaves of the trees is m'.
- Adjoin this collection of trees to U^+ by identifying each vertex of U^+ with the leaf of one of the trees. Denote the set of roots as V^+ which are vertices of degree $d-1$.
- Perform the analogous construction on U^- to complete G.

This construction yields a bi-partite graph of maximum degree d with m vertices of degree $d-1$ on each side. We now consider a the Hardcore model $P_G(\sigma)$ on G. Our construction is a modification of the model considered in [30] where they showed that on a.a.a random bi-partite d-regular graphs the probability of "balanced" sets is exponentially small. This is also the case for our construction and we define the phase of the configuration as

$$Y = Y(\sigma) := \begin{cases} +1 & \text{if } \sum_{w \in W^+} \sigma_w \geq \sum_{w \in W^-} \sigma_w, \\ -1 & \text{if } \sum_{w \in W^+} \sigma_w < \sum_{w \in W^-} \sigma_w. \end{cases}$$

We define the product measure Q_V^+ (respectively Q^-) on configurations on $V = V^+ \cup V^-$ so that the spins are iid Bernoulli with probability q^+ (resp. q^-) on V^+ and q^- (resp. q^+) on V^-, i.e.,

$$\begin{aligned} Q_V^{\pm}(\sigma_V) := {} & (q^{\pm})^{\sum_{v \in V^+} \sigma_v} (1-q^{\pm})^{m - \sum_{v \in V^+} \sigma_v} \\ & \cdot (q^{\mp})^{\sum_{v \in V^-} \sigma_v} (1-q^{\mp})^{m - \sum_{v \in V^-} \sigma_v}. \end{aligned}$$

We define Q_U on $U = U^+ \cup U^-$ similarly. With these definitions we establish the following result about hardcore model on G.

Theorem II.1. *For every $d \geq 3$ when $\lambda_c(d) < \lambda$ and when Condition I.2 and equation* (I.8) *hold there exists constants $\theta(\lambda, d), \psi(\lambda, d) > 0$ such that the graph $G(n, \theta, \psi)$ has*

$(2+o(1))n$ *vertices and satisfies the following with high probability:*

- *The phases occur with roughly balanced probability so that*

$$\mathbb{P}_G(Y=+) \geq \frac{1}{n}, \mathbb{P}_G(Y=-) \geq \frac{1}{n}. \tag{II.1}$$

- *The conditional distribution of the configuration on V satisfies*

$$\max_{\sigma_V} \left| \frac{\mathbb{P}_G(\sigma_V | Y=\pm)}{Q_V^{\pm}(\sigma_V)} - 1 \right| \leq n^{-2\theta}. \tag{II.2}$$

The proof of this theorem is deferred to Section IV.

B. Reduction to Max-Cut

We now demonstrate how to use Theorem II.1 to establish a reduction from sampling from the hardcore model to Max-Cut. Let H be a graph on up to $\frac{1}{d-1}n^{\theta/4}$ vertices. With a random bi-partite graph $G = G(n,\theta,\psi)$ constructed as above we define H^G as follows.

- Take the graph comprising $|H|$ disconnected copies of G and identify each copy with with a vertex in H labeling the copies $(G_x)_{x\in H}$. Denote this graph by $\widehat{H}^G$. We let V_x^+ and V_x^- denote the vertices of G_x corresponding to V^+ and V^-.
- For every edge (x,y) in the graph H add $n^{3\theta/4}$ edges between V_x^+ and V_y^+ and similarly add $n^{3\theta/4}$ edges between and V_x^- and V_y^-. This can be done deterministically in such a way that no vertex in $\widehat{H}^G$ has its degree increased by more than 1. Denote the resulting graph by H^G.

The resulting graph has maximum degree d. For each $x \in H$ we let $Y_x = Y_x(\sigma)$ denote the phase of a configuration σ on G_x. Let $\mathcal{Y} = (Y_x)_{x\in H} \in \{0,1\}^H$ denote the vector of phases of the G_x. Denote the partition function given the phase $\mathcal{Y}$ by

$$Z_{H^G}(\mathcal{Y}') = \sum_{\sigma \in I(H^G)} \lambda^{|\sigma|} \mathbb{1}\left(\mathcal{Y}(\sigma) = \mathcal{Y}'\right).$$

Lemma II.2. *Suppose that G satisfies equations* (II.1) *and* (II.2) *of Theorem II.1. Then*

$$\begin{aligned}&\frac{Z_{\widehat{H}^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}} \\ &= \mathbb{P}_G(Y=+)^{\sum_{x\in H} \mathbb{1}_{Y_x'=+}} \cdot \mathbb{P}_G(Y=-)^{\sum_{x\in H} \mathbb{1}_{Y_x'=-}} \\ &\geq n^{-n^{\theta/4}},\end{aligned} \tag{II.3}$$

and

$$\begin{aligned}&\frac{Z_{H^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}(\mathcal{Y}')} \\ &= (C_H + o(1)) \left[\frac{(1-q^+q^-)^2}{(1-(q^+)^2)(1-(q^-)^2)}\right]^{n^{3\theta/4}\mathrm{Cut}(\mathcal{Y}')},\end{aligned} \tag{II.4}$$

where $C_H = \left[(1-(q^+)^2)(1-(q^-)^2)\right]^{n^{3\theta/4}E(H)}$ and where $\mathrm{Cut}(\mathcal{Y}') = \#\{(x,y) \in E(H) : \mathcal{Y}'_x \neq \mathcal{Y}'_y\}$ denotes the number of edges in cut of H induced by $\mathcal{Y}'$.

Proof: Since the graph $\widehat{H}^G$ consists of a collection of disconnected copies of G, the distribution of a configuration on $\widehat{H}^G$ is given by the product measure of configurations on the $(G_x)_{x\in H}$. In particular the phases are independent and so

$$\begin{aligned}&\frac{Z_{\widehat{H}^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}} \\ &= \mathbb{P}_{\widehat{H}^G}\left(\mathcal{Y}(\sigma) = \mathcal{Y}'\right) \\ &= \mathbb{P}_G(Y=+)^{\sum_{x\in H} \mathbb{1}_{Y_x'=+}} \cdot \mathbb{P}_G(Y=-)^{\sum_{x\in H} \mathbb{1}_{Y_x'=-}} \\ &\geq n^{-n^{\theta/4}},\end{aligned}$$

which establishes equation (II.3). Now the ratio of the partition functions in (II.4) is exactly the probability that the configuration σ sampled under $\mathbb{P}_{\widehat{H}^G}$ is also an independent set for H^G after adding in the extra edges, that is

$$\begin{aligned}&\frac{Z_{H^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}(\mathcal{Y}')} \\ &= \mathbb{P}_{\widehat{H}^G}\left(\sigma \in I(H^G) \mid \mathcal{Y}(\sigma) = \mathcal{Y}'\right) \\ &= \mathbb{P}_{\widehat{H}^G}\left(\forall (v,v') \in \mathcal{E}, \sigma_v\sigma_{v'} \neq 1 \mid \mathcal{Y}(\sigma) = \mathcal{Y}'\right)\end{aligned}$$

where $\mathcal{E} = E(H^G) \setminus E(\widehat{H}^G)$. Now by equation (II.2), conditional on the phase $\mathcal{Y}'$ the spins of $\sigma_{\cup_{x\in H} V_x}$ are asymptotically conditionally independent with probabilities q^+ or q^- depending on the phase. It follows that

$$\begin{aligned}&\mathbb{P}_{\widehat{H}^G}\left(\forall (v,v') \in \mathcal{E}, \sigma_v\sigma_{v'} \neq 1 \mid \mathcal{Y}(\sigma) = \mathcal{Y}'\right) \\ &= (1+o(1)) \prod \mathbb{P}_{\widehat{H}^G}(\sigma_v\sigma_{v'} \neq 1 \mid \mathcal{Y}(\sigma) = \mathcal{Y}'),\end{aligned}$$

where the product is over $(v,v') \in \mathcal{E}$. If $(x,x') \in \mathcal{E}$ then by direction calculations and equation (II.2)

$$\begin{aligned}&\prod_{v\in G_x, v'\in G_{x'}:(v,v')\in\mathcal{E}} \mathbb{P}_{\widehat{H}^G}(\sigma_v\sigma_{v'} \neq 1 \mid \mathcal{Y}(\sigma) = \mathcal{Y}') \\ &= \begin{cases} (1+O(n^{-\theta}))\left((1-(q^+)^2)(1-(q^-)^2)\right)^{n^{3\theta/4}} \\ \quad \text{if } Y_x = Y_{x'}, \\ (1+O(n^{-\theta}))\left((1-q^+q^-)^2\right)^{n^{3\theta/4}} \\ \quad \text{if } Y_x \neq Y_{x'}. \end{cases}\end{aligned}$$

Combining the above estimates we have that

$$\frac{Z_{H^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}(\mathcal{Y}')} = (C_H + o(1)) \left[\frac{(1-q^+q^-)^2}{(1-(q^+)^2)(1-(q^-)^2)}\right]^{n^{3\theta/4}\mathrm{Cut}(\mathcal{Y}')}$$

which completes the proof. ■

Given the previous lemma we now show how to produce the randomized reduction to Max-Cut establishing Theorems 1 and 2.

Theorem 1 and 2: Let H be a graph on at most $\frac{1}{d-1}n^\theta$ vertices. Take an instance of a random graph $G = G(n,\theta,\psi)$ according to the construction in Section II-A. By Theorem II.1 with probability tending to 1 the graph satisfies equations (II.1) and (II.2). Assume that it does and construct the graph H^G which has at most $O(n^{1+\theta})$ vertices and maximum degree d.

Now suppose there exists an FPRAS for the partition function for the hardcore model with fugacity λ on graphs of maximum degree d. We now use the equivalence of approximating the partition function and approximately sampling for the hardcore model described in the introduction. In polynomial time we may approximately sample from the hardcore model on H^G to within δ of the Gibbs distribution in total-variation distance for any $\delta > 0$. Let σ' denote such an approximate sample. We may couple σ' with σ distributed according to the Gibbs measure so that $\mathbb{P}(\sigma' \neq \sigma) \leq \delta$. We now consider the phase of σ. Let $\mathcal{Y}', \mathcal{Y}'' \in \{0,1\}^H$ such that

$$\mathrm{Cut}(\mathcal{Y}') > \mathrm{Cut}(\mathcal{Y}'').$$

Then by Lemma II.2 we have that

$$\frac{\mathbb{P}(\mathcal{Y}(\sigma) = \mathcal{Y}')}{\mathbb{P}(\mathcal{Y}(\sigma) = \mathcal{Y}'')} = \frac{Z_{H^G}(\mathcal{Y}')}{Z_{H^G}(\mathcal{Y}'')} \tag{II.5}$$

$$\geq \frac{(1+o(1))Z_{\widehat{H}^G}(\mathcal{Y}')}{Z_{\widehat{H}^G}(\mathcal{Y}'')}\xi^{n^{3\theta/4}[\mathrm{Cut}(\mathcal{Y}')-\mathrm{Cut}(\mathcal{Y}'')]}$$

$$\geq (1+o(1))n^{-n^{\theta/4}}\xi^{n^{3\theta/4}[\mathrm{Cut}(\mathcal{Y}')-\mathrm{Cut}(\mathcal{Y}'')]} \tag{II.6}$$

where

$$\xi = \frac{(1-q^+q^-)^2}{(1-(q^+)^2)(1-(q^-)^2)}.$$

As we have that $0 < q^- < q^+ < 1$ if follows that $(1-q^+q^-)^2 - (1-(q^+)^2)(1-(q^-)^2) = (q^+ - q^-)^2 > 0$ and hence

$$\xi = \frac{(1-q^+q^-)^2}{(1-(q^+)^2)(1-(q^-)^2)} > 1.$$

Therefore, for large enough n by equation (II.5) it follows that

$$\frac{\mathbb{P}(\mathcal{Y}(\sigma) = \mathcal{Y}')}{\mathbb{P}(\mathcal{Y}(\sigma) = \mathcal{Y}'')} \geq \xi^{\frac{1}{2}n^{3\theta/4}} \geq 4^{n^{\theta/4}}.$$

Since the size of $\{0,1\}^{|H|}$ is only $2^{n^{\theta/4}}$ it follows that with probability at least $1 - 2^{|H|}$ that $\mathrm{Cut}(\mathcal{Y}(\sigma))$ attains the maximum value. Hence with probability at least $1-\delta-o(1)$ the phases $\mathcal{Y}(\sigma')$ of the approximate sample σ' also attains a maximum cut in H. As such we have constructed a randomized polynomial-time reduction from approximating the partition function of the hardcore model to constructing a maximum cut. It follows that unless RP=NP there is no polynomial-time algorithm for approximating the partition function of the hardcore model for $\lambda_c(d) < \lambda < \lambda_c(d)+\varepsilon(d)$ on graphs of maximum degree d or when $\lambda = 1$ on graphs of maximum degree 6 or more. ■

III. The partition function of $\tilde{G}$

In this section we analyse the hardcore model on the random bi-partite graph $\tilde{G}$ and in particular consider the effect of conditioning on the spins in $U = U^+ \cup U^-$. For $\eta \in \{0,1\}^U$ we define $Z_{\tilde{G}}(\eta)$ to be the partition function over configurations whose restriction to U is η, that is

$$Z_{\tilde{G}}(\eta) = \sum_{\sigma \in I(\tilde{G}):\sigma_U = \eta} \lambda^{|\sigma|}.$$

Our analysis borrows heavily on hard computations carried out in [30]. There they considered a random d-regular bipartite graph where each side has n vertices and the edges are chosen according to d independent perfect matchings of the vertices of the sides. They denote $Z^{\alpha,\beta}$ to be the weighted sum over configurations of the graph with αn and βn vertices on the plus and minus sides of the configuration (for α,β such that $\alpha n, \beta n$ are integers). We will denote their quantity by $Z^{\alpha,\beta}_{\mathrm{MWW}}$. In the same spirit define

$$Z^{\alpha,\beta}_{\tilde{G}}(\eta) = \sum_{\sigma:\sigma_U = \eta, \sum_{w\in W^+}\sigma_w = \alpha n, \sum_{w \in W^-}\sigma_w = \beta n} \lambda^{|\sigma|}.$$

Lemma III.1. *For any (α,β) in the interior of $\mathcal{T}$ and all $\eta \in \{0,1\}^U$ we have that:*

$$\mathbb{E}Z^{\alpha,\beta}_{\tilde{G}}(\eta) = (1+O(n^{-1/2}))C^*\left(\lambda\left(\frac{1-\alpha-\beta}{1-\beta}\right)^{d-1}\right)^{\eta^-} \tag{III.1}$$

$$\cdot\left(\lambda\left(\frac{1-\alpha-\beta}{1-\alpha}\right)^{d-1}\right)^{\eta^+}\mathbb{E}Z^{\alpha,\beta}_{\mathrm{MWW}} \tag{III.2}$$

where

$$C^* = \left(\frac{(1-\alpha)(1-\beta)}{1-\alpha-\beta}\right)^{m'}$$

and where $\eta^\pm$ denotes $\sum_{u\in U^\pm}\eta_u$.

A long series of calculations using the second moment method and the small graph conditioning method eventually yields the following bound.

Theorem III.2. *For every $d \geq 3$ and $\lambda > \lambda_c$ such that Condition I.2 holds there exists a positive constant $\varepsilon(d) > 0$ and constants $\theta^*(\lambda, d), \psi^*(\lambda, d) > 0$ such that the partition functions satisfy the following asymptotic almost sure statements,*

$$\sup_{\eta \in \{0,1\}^U} \mathbb{P}\left(Z^{\pm}_{\tilde{G}}(\eta) < \frac{1}{\sqrt{n}} \mathbb{E} Z^{\pm}_{\tilde{G}}(\eta)\right) \to 0. \tag{III.3}$$

IV. Reconstruction on the tree

Our proof now takes a detour through the reconstruction problem on the tree. This problem concerns determining which Gibbs measures on the tree are extremal, or equivalently when the tail σ-algebra is trivial or when point-to-set correlations converge to 0 in the distance of the point to the set [28]. In our setting the measures $\hat{\mu}_{\pm}$ are extremal so we automatically have that non-reconstruction holds. We will use facts about the rate of decay of point-to-set correlations to establish that σ_V is essentially independent of σ_U conditioned on the phase. In most cases the reconstruction problem has been considered in the case of the translation invariant free measure (see [5] for recent progress on the hardcore model) but we will be interested in the case of the semi-translation invariant measures $\hat{\mu}_{\pm}$ on $\hat{\mathbb{T}}_d$ and as such results from the literature do not directly apply here.

The reconstruction problem has for the most part been studied in the case of Markov models on trees with a single transition kernel M. In this theory the key role is played by the λ_* the second eigenvalue of the transition matrix. The famous Kesten-Stigum bound [19, 28] states that there is reconstruction when $\lambda_*^2(d-1) > 1$ while results of [15] show that if non-reconstruction holds and $\lambda_*^2(d-1) < 1$ then point to set correlations decay exponentially quickly. In our setting, however, the Gibbs measure is semi-translation invariant and the Markov model is given by a pair of alternating Markov transition kernels, $M^{\pm}$ defined below.

With minor modifications the proof of [7] (or also [34] or [15]) can be adapted to the semi-translation invariant setting. Here the role of λ_* is played by the second eigenvalue of M^+M^- where

$$M^{\pm} = \begin{pmatrix} 1-q^{\pm} & q^{\pm} \\ 1 & 0 \end{pmatrix}$$

and there is reconstruction when $\lambda_*^2(d-1)^2 > 1$ and exponential decay of correlations when there is non-reconstruction and $\lambda_*^2(d-1)^2 < 1$. The term $(d-1)^2$ is explained by the fact that this this the branching from two levels of the tree. Using the methods of [34] which build on the work of [7] we establish the necessary decay of correlations result. These bounds can also be derived from the work of Martinelli, Sinclair and Weitz [23].

We then apply this result to the graph G. Let $\mathcal{B}$ denote the set of configurations η which have a large influence on σ_V in the "plus phase",

$$\mathcal{B} = \Big\{\eta \in \{0,1\}^U : \sup_{v \in V} |\mathbb{P}_G(\sigma_v = 1 \mid \sigma_U = \eta) - Q^+_V(\sigma_v = 1)| > \exp(-2\zeta_1 \lfloor \frac{\psi}{2} \log_{d-1} n \rfloor)\Big\},$$

where $\zeta_2(d, \lambda) > 0$ is some appropriately chosen constant. We use Theorem III.2 to relate the distribution of σ_U to the measure on the infinite d-ary tree and combine this with strong decay of correlations results for the reconstruction problem to show that

$$\mathbb{P}(\sigma_U \in \mathcal{B} \mid Y = +) \leq \exp(-\exp(\zeta_2 \ell)). \tag{IV.1}$$

for some $\zeta_1(d, \lambda) > 0$. A similar result holds for the minus phase. This establishes that, given the phase, the effect of σ_U on σ_V is very small with high probability. From (IV.1) we then establish the key equation (II.2).

References

[1] D. Achlioptas and A. Coja-Oghlan, *Algorithmic barriers from phase transitions*, IEEE 49th Annual IEEE Symposium on Foundations of Computer Science, 2008. FOCS'08, 2008, pp. 793–802.

[2] D. J. Aldous, *The $\zeta(2)$ limit in the random assignment problem*, Random Structures and Algorithms **18** (2001), 381–418.

[3] N. Berger, C. Kenyon, E. Mossel, and Y. Peres, *Glauber dynamics on trees and hyperbolic graphs*, Probability Theory and Related Fields **131** (2005), 311–340.

[4] P. Berman and T. Fujito, *On approximation properties of the independent set problem for low degree graphs*, Theory of Computing Systems **32** (1999), 115–132.

[5] N. Bhatnagar, A. Sly, and P. Tetali, *Reconstruction Threshold for the Hardcore Model*, arXiv:1004.3531 (2010).

[6] N. Bhatnagar, J. Vera, E. Vigoda, and D. Weitz, *Reconstruction for colorings on trees*, To appear in SIAM J. on Discrete Math. (2008).

[7] C. Borgs, J. Chayes, E. Mossel, and S. Roch, *The Kesten-Stigum reconstruction bound is tight for roughly symmetric binary channels*, Foundations of Computer Science, 2006. FOCS'06. 47th Annual IEEE Symposium on, 2006, pp. 518–530.

[8] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs*, Ann. Appl. Probab **20** (2010), 565–592.

[9] M. Dyer, A. Frieze, and M. Jerrum, *On Counting Independent Sets in Sparse Graphs*, SIAM Journal on Computing **31** (2002), 1527.

[10] M. Dyer and C. Greenhill, *On Markov Chains for Independent Sets*, Journal of Algorithms **35** (2000), 17–49.

[11] D. Galvin and P. Tetali, *Slow mixing of Glauber dynamics for the hard-core model on regular bipartite graphs*, Random Structures and Algorithms **28** (2006), 427–443.

[12] H. O. Georgii, *Gibbs measures and phase transitions*, Walter de Gruyter, 1988.

[13] C. Greenhill, *The complexity of counting colourings and independent sets in sparse graphs and hypergraphs*, Computational Complexity **9** (2000), 52–72.

[14] L. A. Goldberg and M. Jerrum, *Approximating the partition function of the ferromagnetic Potts model*, Arxiv:1002.0986 (2010).

[15] S. Janson and E. Mossel, *Robust reconstruction on trees is determined by the second eigenvalue*, Annals of Probability **32** (2004), 2630–2649.

[16] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley, 2000.
[17] M. Jerrum and A. Sinclair, *Polynomial-time approximation algorithms for the Ising model*, Automata, Languages and Programming (1990), 462–475.
[18] F. Kelly, *Stochastic models of computer communication systems*, Journal of the Royal Statistical Society. Series B (Methodological) **47** (1985), no. 3, 379–395.
[19] H. Kesten and B. P. Stigum, *Additional limit theorems for indecomposable multidimensional Galton-Watson processes*, The Annals of Mathematical Statistics **37** (1966), 1463–1481.
[20] F. Krząkała, A. Montanari, F. Ricci-Tersenghi, G. Semerjian, and L. Zdeborová, *Gibbs states and the set of solutions of random constraint satisfaction problems*, Proceedings of the National Academy of Sciences **104** (2007), 10318.
[21] M. Luby and E. Vigoda, *Fast convergence of the Glauber dynamics for sampling independent sets*, Random Structures and Algorithms **15** (1999), 229–241.
[22] F. Martinelli and E. Olivieri, *Approach to equilibrium of Glauber dynamics in the one phase region. I. The attractive case*, Comm. Math. Phys. **161** (1994), no. 3, 447–486.
[23] F. Martinelli, A. Sinclair, and D. Weitz, *Fast mixing for independent sets, colorings, and other models on trees*, Random Structures Algorithms **31** (2007), 134–172.
[24] M. Mezard and A. Montanari, *Information, physics, and computation*, Oxford University Press, USA, 2009.
[25] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin glass theory and beyond*, World scientific, Singapore, 1987.
[26] E. Mossel, *Survey: information flow on trees*, Graphs, morphisms, and statistical physics: DIMACS Workshop Graphs, Morphisms and Statistical Physics, March 19-21, 2001, DIMACS Center, 2004, pp. 155–170.
[27] A. Montanari, E. Mossel, and A. Sly, *The weak limit of Ising models on locally tree-like graphs*, ArXiv:0912.0719 (2009).
[28] E. Mossel and Y. Peres, *Information flow on trees*, Annals of Applied Probability **13** (2003), 817–844.
[29] E. Mossel and A. Sly, *Exact Thresholds for Ising-Gibbs Samplers on General Graphs*, arXiv:0903.2906 (2009).
[30] E. Mossel, D. Weitz, and N. Wormald, *On the hardness of sampling independent sets beyond the tree threshold*, Probability Theory and Related Fields **143** (2009), 401–439.
[31] A. Sinclair and M. Jerrum, *Approximate counting, uniform generation and rapidly mixing Markov chains*, Information and Computation **82** (1989), 93–133.
[32] A. Sokal, *A personal list of unsolved problems concerning lattice gases and antiferromagnetic Potts models*, Markov Process. Related Fields **7** (2001), 21–38.
[33] A. Sly, *Uniqueness thresholds on trees versus graphs*, The Annals of Applied Probability **18** (2008), no. 5, 1897–1909.
[34] A. Sly, *Reconstruction for the Potts model*, Proceedings of the 41st annual ACM symposium on Theory of computing, 2009, pp. 581–590.
[35] M. Talagrand, *The Parisi formula*, Annals of Mathematics **163** (2006), 221–264.
[36] D. Weitz, *Counting independent sets up to the tree threshold*, STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing (2006), pp. 140–149.
[37] N. C. Wormald, *Models of random regular graphs*, London Mathematical Society Lecture Note Series (1999), 239–298.