A Second Threshold for the Hard-core Model on a Bethe Lattice

Graham R. Brightwell *
Department of Mathematics
London School of Economics
Houghton St.
London WC2A 2AE England

Peter Winkler†
Bell Labs 2C-365
Lucent Technologies
700 Mountain Ave.
Murray Hill, NJ 07974 U.S.A.

February 14, 2003

Abstract

We determine the approximate value of a critical activity for the hard-core model on the Bethe lattice, which determines whether the unique simple invariant Gibbs measure is extremal. This "recovery threshold" turns out to be different both from the threshold for unique Gibbs measure and (in contrast to the Ising model) from the threshold for recovery of root information purely from statistical information about distant sites.

1 Summary

The hard-core model on a Bethe lattice—a.k.a. random independent sets in a Cayley tree—is arguably the best-understood model of combinatorial phase transition, and the best-understood of *any* model of phase transition other than the Ising model. For all activities, it has just one simple[1] invariant Gibbs measure μ, which can be obtained by taking a branching random walk on a certain two-node graph. Its critical activity λ_1 for having just one Gibbs measure of any description is known exactly (apparently first computed by Kelly [5], in the context of call blocking in communications networks).

*Author's net address: g.r.brightwell@lse.ac.uk

†Author's net address: pw@lucent.com. Supported in part by a grant from the U.S. Office of Naval research.

[1] see below for definition

However, there is a second transition that until now has been something of an enigma. The issue is whether the measure μ is *extremal*, that is, not expressible as a convex combination of distinct Gibbs measures.

It is trivial that μ is extremal below the unique-Gibbs-measure critical activity λ_1, and also known that μ fails to be extremal above the "census threshold", which determines when non-vanishing information about whether the root of the tree is occupied may be obtained from the *number*, i.e., the "census", of occupied sites at some great distance from the root. For both the Ising and hard-core models, and indeed practically all models on the Bethe lattice, the census threshold is easily computed by comparing the square of the branching number of the tree with the second eigenvalue of a natural Markov chain. It follows from a result of Mossel [7] that there *is* a threshold value λ_2: μ is extremal for activities $\lambda < \lambda_2$ and non-extremal for $\lambda > \lambda_2$.

For μ to be non-extremal, it is sufficient (and necessary) that non-vanishing information about whether the root is occupied can be recovered, with positive probability, from the full configuration of distant sites generated by a sample from μ. For the Ising model (see e.g., [3]) recovery is possible only above the census threshold.

For the hard-core model, however, it was previously not known whether μ is extremal anywhere above the unique-Gibbs-measure threshold; on the other hand, no technique was known for recovery below the census threshold.

We show that the "recovery threshold" λ_2 is always strictly above the unique-Gibbs-measure threshold λ_1 and, for a Bethe lattice with branching number $k \geq 29$, λ_2 lies strictly below the census threshold. We leave open the curious possibility that the recovery and census thresholds are the same for some small values of k.

We have learned that our upper bound result has been proved independently by Y. Suhov and U.A. Rozikov [9]. Also, Svante Janson [4] has shown us an alternative method, using the Hellinger integral, of proving almost the same upper bound on λ_2; intriguingly, his method also fails to separate the recovery and census thresholds for small k.

2 Introduction

Let $\mathbb{T}^k$ denote the infinite $(k+1)$-regular tree: $\mathbb{T}^k$ is variously referred to as the *Bethe lattice* or *Cayley tree*. We think of $\mathbb{T}^k$ as having a specified *root* vertex r, and we often describe the tree by saying that each vertex has k "children". Technically, the root itself has $k+1$ children; this is not usually important, and we shall normally ignore it.

The *hard-core* model is a model for "random independent sets" in $\mathbb{T}^k$. There is a positive parameter λ of the model, called the *activity*. In this setting, a *Gibbs measure* (for λ) is a probability measure μ on the space of independent sets I in $\mathbb{T}^k$, satisfying the following condition: for any finite set U of vertices of $\mathbb{T}^k$, and almost every independent set I_0, the probability that $I = I_0$, conditioned on $I \backslash U = I_0 \backslash U$, is proportional to $\lambda^{|I_0 \cap U|}$. In fact, it suffices to check this condition in the case when U consists of a single vertex u: the non-trivial case is when none of the neighbors of u is in I_0, and the condition then states that

$\Pr(u \in I \mid I\backslash\{u\} = I_0\backslash\{u\}) = \lambda/(\lambda+1)$.

There are results asserting the existence of Gibbs measures, even in much more general settings; here we are in the fortunate position of being able to construct one. Given any positive activity λ, let w be the unique positive root of $\lambda = w(1+w)^k$. We construct a measure μ on the set of independent sets in $\mathbb{T}^k$ via the following recursive procedure. First we put the root into our independent set I with probability $w/(1+2w)$. If we put any vertex into I, then perforce none of its children is in I. If a vertex is not in I, then each of its children is put in I with probability $w/(1+w)$, all choices made independently. It is straightforward to check that this does give a Gibbs measure for the given activity λ. See [1] for proofs of these facts in a more general hard-constraint setting.

An alternative way to view this measure is to think of a "branching random walk" on the two-node graph H in which node 0, of weight 1, has a loop and node 1, of weight w, does not. At each time-step, each existing particle splits into k particles, and then each new particle takes a step of a random walk on this graph, with transition probabilities given by the node weights. If we follow any particular branch, we are simply observing a node-weighted random walk, and the probability that our particle is at the unlooped node—corresponding to the point of the tree being in the independent set—is the stationary probability $\pi_1 = w/(1+2w)$.

We reserve the symbol μ for this measure, which plays a special role as it is the unique "simple invariant Gibbs measure", for all values of λ. A measure on independent sets in $\mathbb{T}^k$ is *simple* if, for any site $u \in \mathbb{T}^k$ the distributions of

$$I \cap C_1(u), \ldots, I \cap C_{k+1}(u)$$

are mutually independent given $I \cap \{u\}$, where the $C_i(u)$ are the connected components of $\mathbb{T}^k \setminus \{u\}$. A measure is *invariant* if every measurable set of independent sets has the same measure after being shifted by an automorphism of $\mathbb{T}^k$.

Kelly [5] showed that, for $\lambda \le \lambda_1 = k^k/(k-1)^{k+1}$, μ is in fact the *only* Gibbs measure, whereas for λ above this threshold there are other Gibbs measures as well. Specifically, above the threshold there are two "simple semi-invariant" Gibbs measures, one where an independent set contains a greater preponderance of sites at even distance from the root, and the other favoring the sites at odd distance. These two measures are always *extremal*, i.e., they cannot be written as a convex combination of two other Gibbs measures. Above the threshold, there is in fact a large variety of other extremal Gibbs measures; for a start, it is easy to construct one which favors odd sites down one branch from the root, and even sites down the other branches.

Above the threshold λ_1, the original Gibbs measure μ is still present, and Yuri Suhov [10] has raised the question of when it is extremal. Similar questions have been studied for other models—see, for instance, Mossel's survey [8]—and it seems strange that for the hard-core model the question has been neglected up to now.

As we shall show more precisely later, the question of whether μ is extremal can be reformulated along the following lines: suppose we take a sample from μ, look only at the configuration on the vertices at distance d from the root, and then use this information to guess whether the root is occupied (i.e., in the independent set) or not: can we succeed with probability bounded above what we can achieve by guessing at random? Mossel's papers contain many alternative formulations, and different contexts in which this type of problem arises.

One possible strategy is to argue as follows: if the root is occupied, there is a tendency for sites at odd distance from the root to be unoccupied, and sites at even distance to be occupied. If the fraction of vertices at (say) large even distance that are occupied is greater than π_1, we should guess that the root is occupied, and if the fraction is smaller than π_1 then we should guess the root is unoccupied. This argument was used by Kesten and Stigum [6] to prove a result which, for this problem, implies that μ is non-extremal whenever $\eta_2^2 k < 1$, where η_2 is the second largest eigenvalue of the transition matrix of the random walk. We have $\eta_2 = -w/(w+1)$ so, as was observed by Suhov, μ is non-extremal for $w > 1/(\sqrt{k} - 1)$. This is called the *census threshold*; see, for instance, Mossel [7, 8] for more details.

Usually, as above, it is more natural to think of w as being the parameter for this question. In terms of w, the threshold for uniqueness of the Gibbs measure is $w = 1/(k-1)$.

We define the *recovery threshold* to be the infimum of the set of values w such that the simple invariant Gibbs measure μ with parameter w is non-extremal. A result of Mossel, Proposition 12 of [7], shows that μ is non-extremal for any w above the recovery threshold.

The results we have seen so far show that the recovery threshold lies in the closed interval between $1/(k-1)$ and $1/(\sqrt{k} - 1)$. We are able to narrow this range considerably, and show that the recovery threshold lies between $(\ln k - 2\ln\ln k)/k$ and—for k sufficiently large—$(\ln k + \ln\ln k + 1 + \varepsilon)/k$, where ε is any fixed positive constant. Our methods do not suffice to separate the recovery threshold and the census threshold for $k \leq 28$.

Even for large k, the gap between our bounds is a significant one. If we translate the bounds into bounds on the threshold *activity* $\lambda_2(k)$ for recovery, we see that

$$\frac{1}{\ln k}(1 + o(1)) \leq \lambda_2(k) \leq \ln^2 k(1 + o(1)) \ .$$

We have no idea whether $\lambda_2(k)$ is increasing or decreasing in k, or whether μ is extremal in the appealing special case $\lambda = 1$.

The hard-core model is the most basic "hard-constraint" model, and the Bethe lattice is a particularly convenient graph on which to study it. The simplicity of this model allows for a much more precise analysis than is available in general. It is noteworthy that there are still some interesting open questions about this model.

3 Extremality and Information

From now on, we think of the "random" independent set I sampled from μ as a function $\varphi : \mathbb{T}^k \to \{0, 1\}$, with $\varphi(u) = 1$ corresponding to the vertex u being in I.

The purpose of this section is to make the connection between extremality of μ and the notion that non-vanishing information about whether the root is occupied can be recovered from the restriction of φ to vertices at arbitrarily large distances from the root.

More concretely, we show that μ is extremal if and only if the difference of the expectations of two particular random variables tends to zero. The random variables in question are $\mathbf{A}_0(d)$ and $\mathbf{A}_1(d)$, defined as follows: we sample a hard-core configuration φ on $\mathbb{T}^k$ from μ. We then observe the values of φ at the sites at distance d from the root r, "erase" the rest of the values, and compute the *a posteriori* probability a that $\varphi(r) = 0$ (i.e., the root was unoccupied). Then $\mathbf{A}_0(d)$ is the (random) value of a conditioned on the process actually having begun with $\varphi(r) = 0$; $\mathbf{A}_1(d)$ is defined in the same way, starting with $\varphi(r) = 1$.

The intuition is that, if $\mathbf{A}_0(d) > \mathbf{A}_1(d)$, this represents information suggesting that the root actually was unoccupied in the sample φ. If $\mathbb{E}\mathbf{A}_0(d)$ and $\mathbb{E}\mathbf{A}_1(d)$ remain bounded away from each other as $d \to \infty$, we can interpret this as saying that information about the root is retained in the vertices at arbitrarily distance d from r.

Let X be the event that $\varphi(r) = 0$, so that X has *a priori* probability $\pi_0 = (1+w)/(1+2w)$. Let $\mathbf{Y}_d$ be the (random) configuration found on the sites at distance d from r. Then

$$\mathbb{E}\mathbf{A}_0(d) = \mathbb{E}_{\mathbf{Y}_d|X} \Pr(X|\mathbf{Y}_d = y) = \sum_y \Pr(X|\mathbf{Y}_d = y) \Pr(\mathbf{Y}_d = y|X),$$

where the second expectation, and the sum, are taken over all possible values y of $\mathbf{Y}_d$, working in the probability space arising from conditioning on X.

The following result does not require any special properties of our setting. Similar results have no doubt appeared elsewhere; we include a proof—which is essentially just an application of the Cauchy-Schwarz inequality—for completeness.

Lemma 3.1. *$\mathbb{E}\mathbf{A}_0(d) \geq \pi_0 \geq \mathbb{E}\mathbf{A}_1(d)$, with equality if and only if $\mathbf{Y}_d$ is independent of X.*

Proof. We have

$$\begin{aligned}
0 &\leq \mathbb{E}_{\mathbf{Y}_d}(\Pr(X|\mathbf{Y}_d = y) - \pi_0)^2 \\
&= \sum_y \Pr(\mathbf{Y}_d = y)(\Pr(X|\mathbf{Y}_d = y) - \pi_0)^2 \\
&= \sum_y \Pr(X|\mathbf{Y}_d = y) \Pr(X \wedge \mathbf{Y}_d = y) - 2\pi_0 \sum_y \Pr(X \wedge \mathbf{Y}_d = y) + \pi_0^2 \sum_y \Pr(\mathbf{Y}_d = y) \\
&= \sum_y [\Pr(X|\mathbf{Y}_d = y) \Pr(\mathbf{Y}_d = y|X) \Pr(X)] - 2\pi_0^2 + \pi_0^2 \\
&= \pi_0 \mathbb{E}\mathbf{A}_0(d) - \pi_0^2 = \pi_0(\mathbb{E}\mathbf{A}_0(d) - \pi_0)\ .
\end{aligned}$$

So indeed $\mathbb{E}\mathbf{A}_0(d) \geq \pi_0$, equality requiring that $\Pr(X|\mathbf{Y}_d = y) = \pi_0$ for all possible y. Also

$$\mathbb{E}\mathbf{A}_1(d) = \mathbb{E}_{\mathbf{Y}_d|\overline{X}} \Pr(X|\mathbf{Y}_d = y) = 1 - \mathbb{E}_{\mathbf{Y}_d|\overline{X}} \Pr(\overline{X}|\mathbf{Y}_d = y) \leq 1 - (1 - \pi_0) = \pi_0\ ,$$

as required. □

The random variables $\mathbf{A}_0(d)$, $\mathbf{A}_1(d)$ and $\mathbf{A}(d) := \mathbb{E}_{\mathbf{Y}_d} \Pr(X|\mathbf{Y}_d = y) = \pi_0 \mathbf{A}_0(d) + (1 - \pi_0)\mathbf{A}_1(d)$ are backwards martingales; that is, the expectation of each is its value at $d-1$.

The "Backwards Martingale Convergence Theorem" (see e.g. [2]) tells us that in a sample φ from μ each almost surely has a limit, which we denote by $\mathbf{A}_0$, $\mathbf{A}_1$ and $\mathbf{A}$ respectively.

Since $\mathbf{A}_0(d) \to \mathbf{A}_0$ a.s., and the $\mathbf{A}_0(d)$ are uniformly bounded, we also have $\mathbf{A}_0(d) \to \mathbf{A}_0$ in expectation. By Lemma 3.1, if $\mathbb{E}\mathbf{A}_0(d) - \mathbb{E}\mathbf{A}_1(d) \to 0$, then also $\mathbb{E}\mathbf{A}_0(d) - \pi_0 \to 0$, i.e., $\mathbf{A}_0(d)$ converges to the constant π_0 in expectation, and thus $\mathbf{A}_0 = \pi_0$. Under this hypothesis, we can also conclude that $\mathbf{A}_1$ and $\mathbf{A}$ are equal to π_0. As we now show, this is exactly what we need to conclude that μ is extremal.

Lemma 3.2. *The measure μ is extremal if and only if $\mathbf{A}$ is constant.*

Proof. If $\mathbf{A}$ is not constant then there is a constant c such that the event $W :=$"$\mathbf{A} > c$" has non-trivial probability. This event W is a tail event (does not depend on any finite part of φ) and thus $\mu|W$ and $\mu|\overline{W}$ are (evidently different) Gibbs measures. Since $\mu = \Pr(W) \cdot \mu|W + (1 - \Pr(W)) \cdot \mu|\overline{W}$, μ is not extremal.

On the other hand, suppose $\mathbf{A}$ is constant, say equal to π, and fix some distance d. Let U be the set of sites of $\mathbb{T}^k$ at distance d from the root r, and, for $u \in U$, let $\mathbb{T}^k(u)$ be the subtree of $\mathbb{T}^k$, rooted at u, obtained by deleting the edge from u on its unique path to r. Let $V_{d'}(u)$ be the sites in $\mathbb{T}^k(u)$ at some very large distance d' from u. If φ is chosen from μ then given $\varphi \restriction U$, $\varphi \restriction V_{d'}(u)$ is independent for each $u \in U$. Also, as $d' \to \infty$, for each $u \in U$ the probability that u is unoccupied—given the occupancies of its d'-descendants—tends to π. It follows that the occupancies of sites at distance $d+d'$ from r generate a conditional distribution of φ on the d-neighborhood of r which is, with probability as high as is desired, within arbitrarily small total variation distance of that obtained by occupying each site in U independently with probability $1-\pi$. But then μ must be extremal, because if $\mu = p\sigma + (1-p)\tau$ for some $p \in (0,1)$ where σ and τ are different Gibbs measures, then σ and τ differ on the d-neighborhood of r for some d. ∎

4 The Lower Bound

We now show that $\mathbf{A}_0(d) - \mathbf{A}_1(d)$ does indeed approach 0 in expectation when w is sufficiently small, concluding that μ is extremal in this regime. We assume throughout that k is fixed, and φ is a configuration chosen from μ on the tree $\mathbb{T}^k$. We will be interested in the restriction of φ to the finite tree $\mathbb{T}^k_d$ consisting of all sites within distance d of the root r. For a site u in $\mathbb{T}^k_d$, let $\mathbb{T}^k_d(u)$ denote the tree rooted at u obtained by deleting the edge incident to u on the unique path from u to r.

We begin by fixing an arbitrary configuration y on the leaves of $\mathbb{T}^k_d$. For a site u of $\mathbb{T}^k_d$, let y_u be the restriction of y to the leaves of $\mathbb{T}^k_d(u)$, that is, to the remote descendants of u. Let $q_0(u) = q_0(u; y_u) := \Pr(y_u | \varphi(u) = 0)$, $q_1(u) = q_1(u; y_u) := \Pr(y_u | \varphi(u) = 1)$, and $q(u) = q(u; y_u) := q_0(u; y_u)/q_1(u; y_u)$ (the last expression taking the value $+\infty$ in the extended reals when $q_1(u; y_u) = 0$).

Thus the *a posteriori* probability $a(u) = a(u; y_u)$ that $\varphi(u) = 0$ is just

$$\frac{\pi_0 q_0(u; y_u)}{\pi_0 q_0(u; y_u) + \pi_1 q_1(u; y_u)} = \frac{1}{1 + \frac{\pi_1}{\pi_0 q(u;y_u)}} = \frac{1}{1 + \frac{w}{(1+w)q(u;y_u)}} \ .$$

If u has children $u_1, \ldots, u_k$ then clearly

$$q_1(u) = \prod_{j=1}^{k} q_0(u_j)$$

and

$$q_0(u) = \prod_{j=1}^{k} \left(\frac{1}{1+w} q_0(u_j) + \frac{w}{1+w} q_1(u_j) \right)$$

so that

$$q(u) = \left(\frac{1}{1+w} \right)^k \prod_{j=1}^{k} \left(1 + \frac{w}{q(u_j)} \right) \ .$$

It is convenient to introduce the function

$$b(u) = b(u; y_u) = \frac{1+w/q(u; y_u)}{1+w} \ ,$$

so that the recursion for q translates to

$$b(u) = \frac{1 + w/\prod_j b(u_j)}{1+w} \ ,$$

while

$$a(u) = \frac{1}{\frac{w}{1+w} + b(u)} \ .$$

Proving that $a(u)$ approaches $\pi_0 = (1+w)/(1+2w)$ as the depth increases is thus equivalent to proving that $b(u)$ approaches 1.

We now define random variables $\mathbf{B}_0(u)$, $\mathbf{B}_1(u)$ by $\mathbf{B}_0(u) := b(u; Y_u)$ where Y_u is the (random) restriction of φ to $\mathbb{T}_d^k(u)$ given $\varphi(u) = 0$, and $\mathbf{B}_1(u)$ is the same given $\varphi(u) = 1$. Note that the subscripts are not playing the same role for these random variables that they did for the probabilities q_0 and q_1.

The relationship between $a(u)$ and $b(u)$ translates to relationships between the corresponding random variables:

$$\mathbf{A}_0(u) = \frac{1}{\frac{w}{1+w} + \mathbf{B}_0(u)} \ ; \quad \mathbf{A}_1(u) = \frac{1}{\frac{w}{1+w} + \mathbf{B}_1(u)} \ .$$

The branching random walk construction of φ again enables us to express $\mathbf{B}_0(u)$ and $\mathbf{B}_1(u)$ recursively, as follows:

$$\mathbf{B}_1(u) = \frac{1}{1+w} \left(1 + \frac{w}{\prod_{j=1}^{k} \mathbf{B}_0(u_j)} \right) \ ,$$

where the random variables $\mathbf{B}_0(u_j)$ are independent, and

$$\mathbf{B}_0(u) = \frac{1}{1+w}\left(1 + \frac{w}{\prod_{j=1}^{k}\mathbf{B}_*(u_j)}\right)$$

where $\mathbf{B}_*(u_j) = \mathbf{B}_0(u_j)$ with probability $1/(1+w)$, and $\mathbf{B}_*(u_j) = \mathbf{B}_1(u_j)$ with probability $w/(1+w)$, independently for each j.

We now introduce a coupling of $\mathbf{B}_0(u)$ and $\mathbf{B}_1(u)$ for which always $\mathbf{B}_0(u) \le \mathbf{B}_1(u)$. Working in from the leaves, we generate at each site two full configurations beyond that site, one with the site occupied and the other unoccupied. When we come to u, we take for its "occupied" configuration the union of the "unoccupied" configurations of its children. For the unoccupied configuration, we take each child independently, choosing its occupied configuration with probability $w/(w+1)$ and its unoccupied configuration otherwise. At each site we do get faithful copies of two samples from μ, each conditioned on a different state of the site. To check that indeed $\mathbf{B}_0(u) \le \mathbf{B}_1(u)$ we return to the recursion above, noting that $\mathbf{B}_*(u)$ is sandwiched between $\mathbf{B}_0(u)$ and $\mathbf{B}_1(u)$ at each site. From now on we use this coupling implicitly, so we think of the $\mathbf{B}_i(u)$ as random variables defined on the same probability space, so in particular $\mathbf{B}_1(u) - \mathbf{B}_0(u)$ is a non-negative random variable for each u.

Note that if u is not a leaf, $\mathbf{B}_0(u) \ge 1/(1+w)$; and thus similarly for $\mathbf{B}_*(u)$.

Our overall aim is to show that $\mathbb{E}(\mathbf{A}_0(u) - \mathbf{A}_1(u))$ tends to 0 as the distance from u to the leaves tends to infinity. Note that

$$\mathbb{E}(\mathbf{A}_0(u) - \mathbf{A}_1(u)) = \mathbb{E}\left(\frac{\mathbf{B}_1(u) - \mathbf{B}_0(u)}{\left(\frac{w}{1+w} + \mathbf{B}_0(u)\right)\left(\frac{w}{1+w} + \mathbf{B}_1(u)\right)}\right) \le \mathbb{E}(\mathbf{B}_1(u) - \mathbf{B}_0(u)) ,$$

where we used that $\mathbf{B}_1(u) \ge \mathbf{B}_0(u) \ge 1/(1+w)$.

Therefore it suffices to show that $\mathbb{E}(\mathbf{B}_1(u) - \mathbf{B}_0(u))$ tends to 0 as the distance from u to the leaves tends to infinity. Our plan is to show that, for suitably small w,

$$\mathbb{E}(\mathbf{B}_1(u) - \mathbf{B}_0(u)) \le \rho\mathbb{E}(\mathbf{B}_1(u_1) - \mathbf{B}_0(u_1)) ,$$

for u_1 a child of u and $\rho < 1$; this will clearly imply the required result.

We have:

$$\mathbb{E}(\mathbf{B}_1(u) - \mathbf{B}_0(u)) = \frac{w}{1+w}\mathbb{E}\left(\frac{1}{\prod_{j=1}^{k}\mathbf{B}_0(u_j)} - \frac{1}{\prod_{j=1}^{k}\mathbf{B}_*(u_j)}\right)$$

$$\le \frac{w}{1+w}\Big[\mathbb{E}\left(\frac{1}{\prod_{j=1}^{k}\mathbf{B}_0(u_j)} - \frac{1}{\mathbf{B}_*(u_1)\prod_{j=2}^{k}\mathbf{B}_0(u_j)}\right) + \cdots +$$

$$\mathbb{E}\left(\frac{1}{\prod_{j=1}^{k-1}\mathbf{B}_*(u_j)\mathbf{B}_0(u_k)} - \frac{1}{\prod_{j=1}^{k}\mathbf{B}_*(u_j)}\right)\Big]$$

$$= \frac{w}{1+w}\left[\mathbb{E}\left(\frac{\mathbf{B}_*(u_1)-\mathbf{B}_0(u_1)}{\mathbf{B}_*(u_1)\prod_{j=1}^k \mathbf{B}_0(u_j)}\right)+\cdots+\mathbb{E}\left(\frac{\mathbf{B}_*(u_k)-\mathbf{B}_0(u_k)}{\prod_{j=1}^k \mathbf{B}_*(u_j)\mathbf{B}_0(u_k)}\right)\right] .$$

Now since each numerator is a non-negative random variable, and all the denominators are bounded below by $1/(1+w)^{k+1}$, the above is bounded by

$$\frac{w}{1+w}(1+w)^{k+1}\left[\mathbb{E}(\mathbf{B}_*(u_1)-\mathbf{B}_0(u_1))+\cdots+\mathbb{E}(\mathbf{B}_*(u_k)-\mathbf{B}_0(u_k))\right]$$

$$= w(1+w)^k k\frac{w}{1+w}\mathbb{E}(\mathbf{B}_1(u_1)-\mathbf{B}_0(u_1)) .$$

If we can arrange for $w^2(1+w)^{k-1}k=\rho<1$, then we have

$$\mathbb{E}(\mathbf{B}_1(u)-\mathbf{B}_0(u)) \le \rho\mathbb{E}(\mathbf{B}_1(u_1)-\mathbf{B}_0(u_1)),$$

so $\mathbb{E}(\mathbf{B}_1(u)-\mathbf{B}_0(u))$ approaches 0 (exponentially fast) as the depth of the tree increases, as desired. If in particular we take $w=(\ln k - 2\ln\ln k)/k$, we have $(1+w)^{k-1}<e^{wk}=k/(\ln k)^2$, and thus

$$w^2(1+w)^{k-1}k < \frac{(\ln k)^2}{k^2}\frac{k}{(\ln k)^2}k = 1 .$$

For large k, the above argument suffices to show that the recovery threshold is above the threshold $w=1/(k-1)$ for uniqueness of the Gibbs measure—indeed, this works whenever the value $w=1/(k-1)$ achieves $w^2(1+w)^{k-1}k<1$, which is when $k\ge 5$.

However, for $k=2,3,4$, we need a sharper lower bound on $\mathbf{B}_0(u)$ to beat the unique-Gibbs-measure threshold. Let $b_{\min} := \liminf \min(b(u))$, the $\liminf$ taken as the depth of the tree increases, and the min over all possible hard-core configurations; the recursion for b shows that $b_{\min}$ is the least positive solution of the pair of equations

$$b_{\min} = \frac{1+w/b_{\max}^k}{1+w} : \quad b_{\max} = \frac{1+w/b_{\min}^k}{1+w} .$$

As $\mathbf{B}_0 \ge b_{\min}-\varepsilon$ for any fixed $\varepsilon>0$ and sufficiently large depth, from the above development we see that it suffices to show that

$$k\left(\frac{w}{1+w}\right)^2\left(\frac{1}{b_{\min}}\right)^{k+1} < 1 . \tag{1}$$

For $k=2$, we obtain that

$$b_{\min} = \frac{1}{2}(w+w^2-\sqrt{w(w^3+2w^2+w-4)}) .$$

and (1) is satisfied for $w\le 1.0278$. Since the unique-Gibbs-measure threshold here is at $w=1$, we see that the recovery threshold is strictly higher. (For comparison, the census threshold is $w=1/(\sqrt{2}-1)\approx 2.414$.)

For $k=3$ and $w=0.541$, numerical investigations show that $b_{\min}\ge 0.78$, which satisfies (1), and for $k=4$ and $w=0.381$, we get $b_{\min}\ge 0.7889$, again satisfying (1). As these values are above the unique-Gibbs-measure threshold $w=1/(k-1)$ in both cases, we have separated the recovery threshold λ_2 from λ_1 for all values of k.

5 Upper Bound

To get our upper bound for λ_2, we employ a particular, quite simple, algorithm which recovers a "parsimonious" solution: in our case, a configuration which has the fewest possible 0-to-0 transitions. The algorithm works in from the leaves, labeling as "unoccupied" any parent of a child previously labeled "occupied", and "occupied" any site all of whose children have been labeled "unoccupied".

To keep track of the effectiveness of this algorithm we associate with each level of the tree the probabilities of the two types of labeling errors. To that end let $p_0(d)$ be the probability that a site at distance d from the leaves is labeled unoccupied, conditioned on its actually being occupied; and the reverse for $p_1(d)$. This results in the following recursion:

$$p_0(d{+}1) = 1 - (1 - p_1(d))^k \ ;$$

$$p_1(d{+}1) = \left(\frac{w}{1{+}w} p_0(d) + \frac{1}{1{+}w}(1 - p_1(d))\right)^k \ ,$$

where $p_0(0) = p_1(0) = 0$. Note that, if there is any fixed point (p_0, p_1) of these equations with $p_0 \neq 1 - p_1$, then (for instance by deliberately altering our initial data at the leaves), use of the algorithm yields a non-vanishing correlation between the label at the root and the actual value, so that the simple invariant Gibbs measure μ is not-extremal.

For a general analysis, it is easier to work with the simpler inequalities:

$$p_0(d{+}1) \leq k p_1(d) \ ;$$

$$p_1(d{+}1) \leq \left(\frac{1 + w p_0(d)}{1{+}w}\right)^k \ .$$

In particular, suppose that w and k satisfy:

$$\left(\frac{1 + \frac{1}{k}}{1{+}w}\right)^k \leq \frac{1}{wk^2} \ . \tag{2}$$

Then always $p_1 \leq \frac{1}{wk^2}$ and $p_0 \leq \frac{1}{wk}$. So, if inequality (2) holds and also $wk > 2$ (say), then we can recover information about the root from the leaf configuration, so we are below the recovery threshold.

Inequality (2) holds if $ewk^2 \leq (1{+}w)^k$, which in turn holds if $w \geq (\ln k + \ln\ln k + 1 + \varepsilon)/k$, for any fixed $\varepsilon > 0$ and sufficiently large k. Therefore the 'critical' value of w lies below $(\ln k + \ln\ln k + 1 + \varepsilon)/k$, for any $\varepsilon > 0$. We remark again that this is below the census threshold, showing that, at least for large k, using this simplistic algorithm is more powerful than counting the occupied leaves.

The upper bound is similar in form to our lower bound of $(\ln k - 2\ln\ln k)/k$ on the critical value of w. However, in terms of the critical *activity* λ_2, we get only an upper bound around $(\ln k)^2$ for λ_2, as opposed to our lower bound near $(\ln k)^{-1}$. We have no idea whether the true value of λ_2—assuming there is a unique threshold—is increasing or decreasing in k.

For small k, this approach seems to be surprisingly ineffective. In particular, for $k = 2$, the equations defining $p_0(d{+}1)$ and $p_1(d{+}1)$ only have a fixed point (p_0, p_1) with $p_0 + p_1 < 1$ when $w \geq 4$. For $w < 4$, numerical experiments suggest that it is indeed not the case that $p_0(d)$ and $p_1(d)$ stay small as d grows. The conclusion would be that census reconstruction is more powerful than parsimonious reconstruction for small k. For $k \geq 36$, the value $w = 1/(\sqrt{k} - 1)$ of the census recovery threshold satisfies (2), so we do have separation between recovery and census recovery. For $k = 29, \ldots, 35$, numerical investigations reveal multiple fixed points of the original recursive equations for $w = 1/(\sqrt{k} - 1)$, so in these cases too the recovery threshold is below the census threshold.

Acknowledgments

The authors were encouraged to work on this problem by attendance at a talk by Yuri Suhov, of the University of Cambridge, citing the "census threshold" upper bound for the recovery threshold λ_2 and asking whether λ_2 was equal to this value, to λ_1, or (as we have shown) some value in between. We also benefited greatly from conversations with Svante Janson, who showed us an alternative method of bounding λ_2 from above using the Hellinger integral, and with Elchanan Mossel. All of these events, including the research reported in this paper, took place at the Isaac Newton Institute in Cambridge, England, during the Programme in Computation, Probability and Combinatorics, July–December 2002.

References

[1] G.R. Brightwell and P. Winkler, Graph homomorphisms and phase transitions, *J. Comb. Theory (Series B)* **77** (1999), 221–262.

[2] W. Durrett, *Probability: Theory and Examples*, Wadsworth, Belmont CA (1992).

[3] W. Evans, C. Kenyon, Y. Peres and L.J. Schulman, A critical phenomenon in a broadcast process, *Ann. Appl. Prob.* **10** #2 (May 2000), 410–433.

[4] S. Janson, personal communications, (2002/2003).

[5] F.P. Kelly, Stochastic models of computer communication systems, *J. R. Statist. Soc. B* **47** #3 (1985), 379–395.

[6] H. Kesten and J. Stigum, Additional limit theorems for indecomposable multidimensional Galton-Watson processes, *Ann. Math. Stat.* **37** (1966), 1463–1481.

[7] E. Mossel, Reconstruction on trees: Beating the second eigenvalue, *Ann. Appl. Probab.* **11** #1 (2001), 285–300.

[8] E. Mossel, Survey: Information flow on trees, preprint (2001).

[9] Y. Suhov and U.A. Rozikov, A hard-core model on a Cayley tree: an example of a loss network, preprint (2003).

[10] Y. Suhov, talk at the Isaac Newton Institute, Cambridge UK, August 2002.