THE MULTI-STATE HARD CORE MODEL ON A REGULAR TREE*

DAVID GALVIN†, FABIO MARTINELLI‡, KAVITA RAMANAN§, AND PRASAD TETALI¶

Abstract. The classical hard core model from statistical physics, with activity $\lambda > 0$ and capacity $C = 1$, on a graph G, concerns a probability measure on the set $\mathcal{I}(G)$ of independent sets of G, with the measure of each independent set $I \in \mathcal{I}(G)$ being proportional to $\lambda^{|I|}$. Ramanan et al. proposed a generalization of the hard core model as an idealized model of multicasting in communication networks. In this generalization, the *multi-state* hard core model, the capacity C is allowed to be a positive integer, and a configuration in the model is an assignment of states from $\{0, \ldots, C\}$ to $V(G)$ (the set of nodes of G) subject to the constraint that the states of adjacent nodes may not sum to more than C. The activity associated to state i is λ^i, so that the probability of a configuration $\sigma : V(G) \to \{0, \ldots, C\}$ is proportional to $\lambda^{\sum_{v \in V(G)} \sigma(v)}$.

In this work, we consider this generalization when G is an infinite rooted b-ary tree and prove rigorously some of the conjectures made by Ramanan et al. In particular, we show that the $C = 2$ model exhibits a (first-order) phase transition at a larger value of λ than the $C = 1$ model exhibits its (second-order) phase transition. In addition, for large b we identify a short interval of values for λ above which the model exhibits phase co-existence and below which there is phase uniqueness. For odd C, this transition occurs in the region of $\lambda = (e/b)^{1/\lceil C/2 \rceil}$, while for even C, it occurs around $\lambda = (\log b / b(C+2))^{2/(C+2)}$. In the latter case, the transition is first-order.

Key words. Gibbs measures, hard core model, multicasting, phase transition, loss networks

AMS subject classifications. Primary: 82B20, 82B26, 60K35; secondary: 90B15, 05C99.

1. Introduction.

1.1. The Multi-State Hard Core Model. Let $G = (V, E)$ be a finite or countably infinite graph without loops, and let S be a finite set. We refer to the elements of S as *states*. Many stochastic processes on S^V that arise in applications are subject to "hard constraints" that prohibit certain values of S from being adjacent to one another in the graph G. Such processes only attain configurations that lie in a certain feasible subset of S^V. A generic example is the *hard core model*, which has state space $S = \{0, 1\}$ and imposes the constraint that no two adjacent vertices in the graph can both have the state 1. In other words, the set of feasible configurations for the hard core model on a graph G is $\{\sigma \in \{0,1\}^V : \sigma_x + \sigma_y \leq 1 \text{ for every } xy \in E\}$, or, equivalently, the collection of independent sets of the graph G. Processes with such *hard constraints* arise in fields as diverse as combinatorics, statistical mechanics and telecommunications. In particular, the hard core model arises in the study of random independent sets of a graph [5, 7], the study of gas molecules on a lattice [2], and in the analysis of multicasting in telecommunication networks [9, 11, 15].

In this work, we consider a generalization of the hard core model, which we refer

*This work is supported in part by NSF Grants DMS-0401239 and DMS-0701043 (PT), CMMI-0928154 and DMI-0728064 (KR), and DMS-0111298 (DG), by the NSA (DG), and by the European Research Council "Advanced Grant" PTRELSS 228032 (FM).

†Department of Mathematics, University of Notre Dame, Notre Dame IN 46556 (dgalvin1@nd.edu).

‡Dip. Matematica, Universita' di Roma Tre, L.go S. Murialdo 1, 00146 Roma, Italy (martin@mat.uniroma3.it).

§Division of Applied Mathematics, Brown University, Providence RI 02912 (Kavita_Ramanan@brown.edu).

¶School of Mathematics and School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332 (tetali@math.gatech.edu).

to as the *multi-state* hard core model, in which the state space is

$$S_C = \{0, 1, 2, \ldots, C\},$$

for some integer $C \geq 1$, and the set of allowable configurations is given by

$$\Omega_G = \{\sigma \in S_C^V : \sigma_x + \sigma_y \leq C \text{ for every } xy \in E\}.$$

When G is the d-dimensional lattice $\mathbb{Z}^d$, this model was introduced and studied by Mazel and Suhov in [14], motivated by applications in statistical physics. In our work, we focus on the case where G is an infinite rooted b-ary tree (*i.e.,* an infinite graph without cycles in which each vertex has exactly $b+1$ edges incident to it, except for one distinguished vertex called the root which has b edges incident to it), which we denote by $\mathbb{T}^b$.

On the tree, this model was studied by Ramanan et al. in [15] as an idealized example of multicasting on a regular tree network, each of whose edges has the same capacity C. In communications, multicasting arises when, instead of having a simple end-to-end connection, a transmission is made from a single site to a group of individuals [1]. An important performance measure of interest is the probability of packet loss for a given routing protocol [17]. As in [15], here we consider an idealized model in which the routing is simple in the sense that nodes multicast only to their nearest neighbors, and study the impact of the connectivity of the network (*i.e.,* the value of b) and the arrival rate on the blocking (or packet loss) probabilities. The state σ_v of any node or vertex $v \in V$ represents the number of active multicast calls present at that node. Multicast calls are assumed to arrive at each node as a Poisson process with rate λ and require one unit of capacity on each of the $b+1$ edges emanating from that node. If this capacity is available, then the call is accepted and the number of active multicast calls at that node increases by one, while if the required capacity is not available, then the state of the node remains unchanged and the call is said to be blocked or lost. Calls that are accepted require a random amount of service and then depart the system. Service requirements of calls are assumed to be independent and identically distributed (without loss of generality with mean 1), and independent of the arrival process. This model is a special case of a loss network (see [9] for a general survey of loss networks and [12, 15] for connections with this particular model).

For a finite graph G and arrival rate λ, it is well-known that the associated stochastic process has a unique stationary distribution $\mu_{G,\lambda}$ on Ω_G that is given explicitly by

$$\mu_{G,\lambda} \doteq \frac{1}{Z_{G,\lambda}} \prod_{v \in V} \lambda_{\sigma_v} \qquad \text{for } \sigma \in \Omega_G, \tag{1.1}$$

where $Z_{G,\lambda}$ is the corresponding normalizing constant (partition function) $Z_{G,\lambda} \doteq \sum_{\sigma \in \Omega_G} \prod_{v \in V} \lambda_{\sigma_v}$, where the form of λ_i depends on how the multicast calls are served. If the calls are assumed to be served in a first-come first-served manner at each node (see [9]), then we have

$$\lambda_i \doteq \frac{\lambda^i}{i!}, \qquad i = 0, \ldots, C,$$

If they are served using the processor sharing scheduling discipline at each node (see [10]), then we have

$$\lambda_i \doteq \lambda^i, \qquad i = 0, \ldots, C. \tag{1.2}$$

Here we adopt the usual convention that $0! = 1$, so that $\lambda_0 = 1$ in both models, and we will sometimes refer to the arrival rate λ as the *activity*. In this paper (as in [14]) our λ_i's will always be as defined in (1.2). Thus, our exclusive focus will be the study of the multi-state hard core model on a b-ary tree $\mathbb{T}^b$ with activities given by (1.2).

1.2. Gibbs Measures and Phase Transitions. Although there is an explicit expression (1.1) for the stationary distribution on a finite graph, the computational complexity of calculating the normalization constant for large graphs limits the applicability of this formula. Thus, in order to gain insight into the behavior of these measures on large graphs, it is often useful to consider the associated Gibbs measure on an infinite graph. Roughly speaking, a Gibbs measure on an infinite graph G associated with an activity λ is characterized by the property that the distribution of the configuration on any *finite* subset U of V, conditioned on the complement, is equal to the regular conditional probability of the measure $\mu_{G[\bar{U}],\lambda}$ on the restriction $G[\bar{U}]$ of the graph G to the closure $\bar{U} = U \cup \partial U$ of U, given the configuration on the boundary ∂U of U (see Definition 2.1 below for a more precise formulation). It is not hard to show that such a Gibbs measure always exists (in a far more general context, see for example [8]).

However, unlike stationary distributions on finite graphs, the associated Gibbs measures on infinite graphs may not be unique. If there are multiple Gibbs measures associated with a given arrival rate or activity λ, we say that there is *phase coexistence* at that λ. Let T_n denote the finite sub-tree of $\mathbb{T}^b$ with root r and depth n, which contains all vertices in $\mathbb{T}^b$ that are at a distance of at most n from the root r. As is well known (see, for example, Chapter 4 of [8]), for a fixed activity $\lambda > 0$, one way to obtain a Gibbs measure on the tree $\mathbb{T}^b$ rooted at r is as the suitable limit of a sequence of measures, where the nth measure in the sequence is the stationary measure $\mu_{T_n \cup \partial T_n, \lambda}$ on $T_n \cup \partial T_n$ (as defined in (1.1)), conditioned on the boundary ∂T_n being empty (*i.e.,* conditioned on all vertices in the boundary having state 0). We shall refer to this Gibbs measure as the empty boundary condition (b.c.) Gibbs measure (corresponding to the activity λ). In a similar fashion, we define the full b.c. Gibbs measure to be the limit of a sequence of conditioned measures on T_n, but now conditioned on the boundary ∂T_n being full (*i.e.,* conditioned on all vertices in the boundary having state C). Let δ_λ denote the total variation distance of the marginal distributions at the root r under the empty b.c. and full b.c. Gibbs measures corresponding to the activity λ. When λ lies in the region of uniqueness, clearly the empty b.c. Gibbs measure coincides with the full b.c. Gibbs measure, and so $\delta_\lambda = 0$. On the other hand, when λ is in a region of phase coexistence, then $\delta_\lambda > 0$ and it can be shown (due to a certain monotonicity property of our model established in Lemma 2.2 and Proposition 2.3) that the empty b.c. and full b.c. Gibbs measures must necessarily differ. If there exists $\lambda_{cr} = \lambda_{cr}(C)$ for which there is uniqueness for each $\lambda < \lambda_{cr}$ and phase coexistence for every $\lambda > \lambda_{cr}$, then we say that a *phase transition* occurs at λ_{cr}. Moreover, if δ_λ, as a function of λ, is continuous at λ_{cr}, then we say that a *second-order* phase transition occurs, while if δ_λ is discontinuous at λ_{cr}, then we say that a *first-order* phase transition occurs.

When $C = 1$, the phase transition point $\lambda_{cr}(1)$ on the tree is explicitly computable and is easily seen to be a second-order phase transition (see [9, 18, 19] and also Section 2.2). The behavior is more complicated for higher C. The multi-state hard core model on the d-dimensional lattice $\mathbb{Z}^d$ was studied in [14], where it was shown that when C is odd, there is phase coexistence for all sufficiently large λ, while when C is even, there is a unique Gibbs measure for each sufficiently large λ. If phase coexistence were known to be monotone in the activity (this remains an open problem on $\mathbb{Z}^d$ even when $d = 2$), then the result of Mazel and Suhov would imply that there is no phase transition on $\mathbb{Z}^d$ for even C. On the other hand, numerical experiments for the multi-state hard core model on the regular tree (see Section 3.5 and Figure 5 of [15]) suggest that there is a phase transition on the tree for every C, but that the order of the phase transition depends on the parity of C (being first-order for even C and second-order for odd C). This is particularly interesting as it shows that the parity of the capacity has an effect on the regular tree as well, although the effect is not as pronounced as on the d-dimensional lattice.

The study of phase transitions of models with hard constraints on trees has been the subject of much recent research (see [4, 5, 13]). In [4], the focus is on classifying types of hard constraints (as encoded in a so-called constraint graph) on the basis of whether or not there exists a unique simple invariant Gibbs measure for all activity vectors $(\lambda_i, i \in S)$. For $C > 1$, the model that we present here allows for two 1's to be adjacent, but never allows a 1 to be adjacent to C which, in the language of [4], implies that the associated constraint graph is fertile. From Theorem 8.1 of [4] it follows that there exist some activity vectors for which there exist multiple simple invariant Gibbs measures. However, the emphasis of our work is quite different, as our aim is to identify regions where multiple Gibbs measures (not necessarily simple and invariant) exist for the particular choice of activity vector given in (1.2). Another related work, again motivated by telecommunication networks, is [13], which studies Gibbs measures associated with a three-state generalization of the hard core model. However, the hard constraints considered in [13] are somewhat different from the $C = 2$ case in our model.

1.3. Main Results and Outline. The main contribution of this paper is to make rigorous some of the conjectures made in [15], leading to a better understanding of the multi-state hard core model. Our results may be broadly summarized as follows.

1. For $C = 2$ and every $b \in \mathbb{N}$, $b \geq 2$, we show that the Gibbs measure is unique for larger values of λ than in the usual $C = 1$ hard core model (see Corollary 3.2) and we also show that the phase transition is first-order (see Theorem 3.3). Recall that, in contrast, for $C = 1$, the phase transition is second-order.

2. For large values of b, we identify a rather narrow range of values for λ, above which there is phase co-existence and below which there is uniqueness. Although we do not establish the existence of a unique critical value $\lambda_{cr}(C)$ at which phase transition occurs, we establish a fairly precise estimate of $\lambda_{cr}(C)$ if (as we strongly believe) it exists: when C is odd,

$$\lambda_{cr}(C) \approx \left(\frac{e}{b}\right)^{\frac{1}{\lceil C/2 \rceil}}$$

(see Theorem 4.1), while for C even,

$$\lambda_{cr}(C) \approx \left(\frac{1}{C+2}\frac{\log b}{b}\right)^{\frac{2}{C+2}}$$

(see Theorem 4.5).

3. For all even C and all sufficiently large b (depending on C), the model always exhibits a first-order phase transition (see Section 4.3).

The outline of the paper is as follows. First, in Section 2 we establish a connection between phase coexistence and multiplicity of the fixed points of an associated recursion. This is based on the construction of Gibbs measures as limits of conditional measures on finite trees with boundary conditions, as mentioned above. In Section 3 we provide a detailed analysis of the recursion in the special case $C = 2$. In Section 4 we study the recursion when b is large and identify the phase transition window. Finally, in Section 4.3 we study the asymptotics for large b when C is even and provide evidence of a first-order phase transition. An interesting open problem is to rigorously establish that the phase transition is second-order for all odd C.

2. Gibbs Measures and Recursions.

2.1. Gibbs Measures on Trees. Consider any graph $G = (V, E)$ with vertex set V and edge set $E \subseteq V^{(2)}$ (the set of unordered pairs from V). For any $U \subset V$, the boundary of U is $\partial U \doteq \{x \in V \setminus U : xz \in E \text{ for some } z \in U\}$ and the closure of U is $\overline{U} \doteq U \cup \partial U$. Let $G[U]$ denote the restriction of the graph to the vertex set U. For $\sigma \in S_C^V$, let $\sigma_U = (\sigma_v, v \in U)$ represent the projection of the configuration σ onto the vertex set U. With some abuse of notation, for conciseness, we will write just σ_v for $\sigma_{\{v\}}$ and refer to it as the *state* or, inspired by models in statistical mechanics, the *spin value* at v. For $U \subseteq V$, let $\mathcal{F}(U)$ be the σ-field in S_C^U generated by sets of the form $\{\sigma_v = i\}$ for some $v \in U$ and $i \in S_C$. We now provide a rigorous definition of the Gibbs measure.

DEFINITION 2.1. *A Gibbs measure for the multi-state hard core model associated with the activity λ is a probability measure μ on $(S_C^V, \mathcal{F}(V))$ that satisfies for all $U \subset V$ and μ-a.a. $\tau \in S_C^V$,*

$$\mu(\sigma_U = \tau_U | \sigma_{V \setminus U} = \tau_{V \setminus U}) = \mu_{G[\bar{U}],\lambda}(\sigma_U = \tau_U | \sigma_{\partial U} = \tau_{\partial U}),$$

where $\mu_{G[\bar{U}],\lambda}$ is as defined in (1.1), with λ_i given as in (1.2).

We now specialize to the case when G is a regular, b-ary, rooted tree $\mathbb{T}^b$ with root r. A *child* of a vertex x in $\mathbb{T}^b$ is a neighboring vertex that is further from r than x; the vertices (other than x) that lie along the path from x to r are the *ancestors* of x. We will be concerned with (complete) finite sub-trees T of $\mathbb{T}^b$ rooted at r; such a tree T is determined by a depth n, and consists of all those vertices at distance at most n from r. It has $|T| = (b^{n+1} - 1)/(b - 1)$ vertices, and its *boundary* ∂T consists of the children (in $\mathbb{T}^b$) of its leaves (so that $|\partial T| = b^{n+1}$). The tree consisting of all vertices at distance at most n from the root r will be denoted by T_n.

Given a finite sub-tree T and $\tau \in \Omega_{\mathbb{T}^b}$, we let Ω_T^τ denote the (finite) set of spin configurations $\sigma \in \Omega_{T \cup \partial T}$ that agree with τ on ∂T; thus τ specifies a *boundary condition* on T. For a function $f : \Omega_{T \cup \partial T} \to \mathbb{R}$ we denote by $\mu_{T,\lambda}^\tau(f) = \sum_{\sigma \in \Omega_T^\tau} \mu_{T,\lambda}^\tau(\sigma) f(\sigma)$ the *expectation* of f with respect to the distribution $\mu_{T,\lambda}^\tau(\sigma) \propto \prod_{v \in T} \lambda^{\sigma_v}$. On the configuration space $\Omega_{\mathbb{T}^b}$, we define the partial order $\sigma \prec \eta$ if and only if $\sigma_v \leq \eta_v$ for all v with even distance $d(v, r)$ from the root and $\sigma_v \geq \eta_v$ for all v with odd distance from the root. Given two probability measures on $\Omega_{\mathbb{T}^b}$, we then say that $\mu \prec \nu$ if $\mu(f) \leq \nu(f)$ for any (bounded) function f that is non-decreasing with respect to the above partial order.

Let T be a complete finite tree rooted at r, and let $\mu_{T,\lambda}^0$ and $\mu_{T,\lambda}^C$, respectively, be the empty b.c. and full b.c. measures (corresponding to the two boundary conditions identically equal to 0 and C, respectively, on ∂T). The following monotonicity result is well known (see, for example, Theorem 4.1 of [19]). However, for completeness, we provide an independent proof of this result, which involves a Markov chain argument that constructs a simultaneous coupling of $(\mu^0_{T,\lambda}, \mu^\tau_{T,\lambda}, \mu^C_{T,\lambda})$ such that the required monotonicity conditions are satisfied with probability one.

LEMMA 2.2. *For any* $\tau \in \Omega_{\mathbb{T}^b}$,

$$\begin{array}{ll} \mu^0_{T,\lambda} \prec \mu^\tau_{T,\lambda} \prec \mu^C_{T,\lambda} & \textit{if } d(\partial T, r) \textit{ is even,} \\ \mu^C_{T,\lambda} \prec \mu^\tau_{T,\lambda} \prec \mu^0_{T,\lambda} & \textit{if } d(\partial T, r) \textit{ is odd.} \end{array}$$

Moreover, if $d(\partial T, r)$ *is even (respectively, odd) there is a coupling* $\pi_T = (\sigma^0, \sigma^\tau, \sigma^C)$ *of* $(\mu^0_{T,\lambda}, \mu^\tau_{T,\lambda}, \mu^C_{T,\lambda})$ *such that* $\sigma^0 \prec \sigma^\tau \prec \sigma^C$ *(respectively,* $\sigma^C \prec \sigma^\tau \prec \sigma^0$*) with probability one.*

Proof. We consider only the case when $d(\partial T, r)$ is even, since the other case can be established in an exactly analogous fashion. On $\Omega^0_T \times \Omega^\tau_T \times \Omega^C_T$ we construct an ergodic Markov chain $\{\sigma^0(t), \sigma^\tau(t), \sigma^C(t)\}_{t \in \mathbb{Z}_+}$ such that at any time $t \in \mathbb{Z}_+$ the required ordering relation $\sigma^0(t) \prec \sigma^\tau(t) \prec \sigma^C(t)$ is satisfied, and moreover each replica is itself an ergodic chain that is reversible with respect to the measure $\mu^\cdot_{T,\lambda}$ with the corresponding boundary condition. The stationary distribution π_T of the global chain will then represent the sought coupling of the three measures.

The chain, a standard *Heat Bath* sampler, is defined as follows. Assume that the three current configurations corresponding to $0, \tau$ and C boundary conditions are equal to (α, β, γ) respectively and that they satisfy the ordering relation. Pick uniformly at random $v \in T$ and let (a, b, c) be the maximum spin values in v compatible with the values of (α, β, γ) on the neighbors of v respectively. Due to the ordering assumption either $c \le b \le a$ or the opposite inequalities hold. Then the current three values at v are replaced by new ones, $(\alpha'_v, \beta'_v, \gamma'_v)$, sampled from a coupling of the three distributions on $\{0, 1, \dots, a\}$, $\{0, 1, \dots, b\}$, $\{0, 1, \dots, c\}$ which assign a weight proportional to λ^i to the value i. It is clear that such a coupling can be constructed in such a way that $(\alpha'_v, \beta'_v, \gamma'_v)$ satisfy the opposite ordering of (a, b, c) and thus the global ordering is preserved. ◻

Consider now the sequence $\{T_{2n}\}_{n \in \mathbb{N}}$ with $d(\partial T_{2n}, r) = 2n$. Then, thanks to monotonicity, $\lim_{n\to\infty} \mu^C_{T_{2n},\lambda} = \mu^C_\lambda$ exists (weakly) and it defines the *maximal* Gibbs measure. Similarly $\lim_{n\to\infty} \mu^0_{T_{2n},\lambda} = \mu^0_\lambda$ defines the *minimal* Gibbs measure [8]. Notice that, by construction, $\lim_{n\to\infty} \mu^C_{T_{2n+1},\lambda} = \mu^0_\lambda$ while $\lim_{n\to\infty} \mu^0_{T_{2n+1},\lambda} = \mu^C_\lambda$. Finally, for any other Gibbs measure μ, it holds that $\mu^0_\lambda \prec \mu_\lambda \prec \mu^C_\lambda$.

The main problem is therefore that of deciding when $\mu^C_\lambda = \mu^0_\lambda$. In what follows we establish the following criterion, which is in fact an equivalent criterion, since the other implication is obviously true (see [18, 19]; see also [3, 16] for a similar discussion in the special case of $C = 1$). Let $\mathbb{P}^\tau_{n,\lambda}$ be the marginal of $\mu^\tau_{T_n,\lambda}$ on σ_r given boundary condition τ, and let $\mathbb{P}^C_\lambda$ and $\mathbb{P}^0_\lambda$ be the corresponding marginals for μ^C_λ and μ^0_λ, respectively.

PROPOSITION 2.3. *For every* $\lambda > 0$, *if* $\mathbb{P}^C_\lambda = \mathbb{P}^0_\lambda$ *then* $\mu^C_\lambda = \mu^0_\lambda$.

Proof. Assume $\mathbb{P}^C_\lambda = \mathbb{P}^0_\lambda$. Then, by monotonicity,

$$\lim_{n\to\infty} \|\mathbb{P}^C_{n,\lambda} - \mathbb{P}^0_{n,\lambda}\|_{TV} = 0, \tag{2.1}$$

where $\|\cdot\|_{TV}$ denotes total variation distance. Let A be a local event (*i.e.*, depending only on finitely many spins) and let m be sufficiently large so that A does not depend on the spin configuration outside T_m. Fix $n > m$, and let $\pi_{2n} = (\sigma^0, \sigma^\tau, \sigma^C)$ be the monotone coupling of $(\mu^0_{T_{2n},\lambda}, \mu^\tau_{T_{2n},\lambda}, \mu^C_{T_{2n},\lambda})$ described in Lemma 2.2. Then

$$\begin{aligned}
\|\mu^C_{T_{2n},\lambda}(A) - \mu^0_{T_{2n},\lambda}(A)\| &\le \pi_{2n}(\sigma^C_v \ne \sigma^0_v \text{ for some } v \in T_m) \\
&\le \sum_{\substack{v\in T_m\\ d(v,r)\ even}} \sum_{k=0}^{C} \pi_{2n}(\sigma^C_v \ge k > \sigma^0_v) \\
&\qquad + \sum_{\substack{v\in T_m\\ d(v,r)\ odd}} \sum_{k=0}^{C} \pi_{2n}(\sigma^0_v \ge k > \sigma^C_v) \\
&= \sum_{\substack{v\in T_m\\ d(v,r)\ even}} \sum_{k=0}^{C} \big[\mu^C_{T_{2n},\lambda}(\sigma_v \ge k) - \mu^0_{T_{2n},\lambda}(\sigma_v \ge k)\big] \\
&\qquad + \sum_{\substack{v\in T_m\\ d(v,r)\ odd}} \sum_{k=0}^{C} \big[\mu^0_{T_{2n},\lambda}(\sigma_v \ge k) - \mu^C_{T_{2n},\lambda}(\sigma_v \ge k)\big]\,.
\end{aligned}$$

For simplicity, let us examine an even term $\mu^C_{T_{2n},\lambda}(\sigma_v \ge k) - \mu^0_{T_{2n},\lambda}(\sigma_v \ge k)$ and show that it tends to zero as $n \to \infty$. Let w be the immediate ancestor of v. By conditioning on the spin value at w we can write

$$\begin{aligned}
&\mu^C_{T_{2n},\lambda}(\sigma_v = i) - \mu^0_{T_{2n},\lambda}(\sigma_v = i) \\
&= \sum_{j=0}^{C-i} \mu^C_{T_{2n},\lambda}(\sigma_w = j)\big[\mu^C_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j) - \mu^0_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j)\big] \\
&\quad + \sum_{j=0}^{C-i} \big[\mu^C_{T_{2n},\lambda}(\sigma_w = j) - \mu^0_{T_{2n},\lambda}(\sigma_w = j)\big]\mu^0_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j)\,.
\end{aligned}$$

By iterating upwards until we reach the root, and using (2.1), we see that it is enough to show that

$$\lim_{n\to\infty} \max_{v\in T_m} \max_{i\le C} \max_{j\le C-i} \big|\mu^C_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j) - \mu^0_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j)\big| = 0.$$

Now, let

$$Z^\tau_k(i) := \lambda^i \sum_{\sigma\in\Omega^\tau_{T_k\setminus\{r\}}} \prod_{v\in T_k\setminus\{r\}} \lambda^{\sigma_v} \tag{2.2}$$

denote the partition function (or normalizing constant) on the complete finite tree T_k with boundary conditions τ and $\sigma_r = i$. It is clear that

$$\frac{\mathbb{P}^\tau_{k,\lambda}(i)}{\mathbb{P}^\tau_{k,\lambda}(0)} = \frac{Z^\tau_k(i)}{Z^\tau_k(0)}\,.$$

Therefore,

$$\mu^0_{T_{2n},\lambda}(\sigma_v = i \,|\, \sigma_w = j) = \frac{Z^0_{2n-n_v}(i)}{\sum_{k\le C-j} Z^0_{2n-n_v}(k)} = \frac{\mathbb{P}^0_{2n-n_v}(i)}{\mathbb{P}^0_{2n-n_v}(\sigma_r \le C-j)}\,,$$

where n_v denotes the level of v (the distance from the root). A similar relation holds for the full boundary condition.

The proof is concluded once we observe that $n_v \leq m$ and that

$$\mathbb{P}^0_{2n-n_v,\lambda}(\sigma_r \leq C-j) \geq \mathbb{P}^0_{1,\lambda}(0) > 0.$$

□

2.2. Recursions. Our next step, as in many other spin models on trees, is to set up a recursive scheme to compute the relevant marginals $\mathbb{P}^0_{n,\lambda}$ and $\mathbb{P}^C_{n,\lambda}$. In what follows, for simplicity we count the levels bottom-up and the boundary conditions are at level 0. Moreover, since the recursive scheme is independent of the boundary conditions, and since we will never be considering more than one value of λ at a time, we drop both from our notation.

For $i = 0, \dots, C$, and $n \in \mathbb{N}$, we set

$$Q_n(i) := \frac{\mathbb{P}_n(i)}{\mathbb{P}_n(0)}, \quad R_n(i) := \frac{\sum_{k=0}^{C} Q_n(k)}{\sum_{k=0}^{C-i} Q_n(k)} = \left[1 - \mathbb{P}(\sigma_r > C-i)\right]^{-1}.$$

Thus $R_n(0) = 1$ and $R_n(i) \leq R_n(i+1)$. Moreover, let Z_n be as defined in (2.2), but with τ equal to the empty b.c.Then we obtain the recursive equations

$$\begin{aligned} Z_{n+1}(i) &= \lambda^i \left[\sum_{k=0}^{C-i} Z_n(k)\right]^b, \\ Q_{n+1}(i) &= \lambda^i \left[\frac{\sum_{k=0}^{C-i} Q_n(k)}{\sum_{k=0}^{C} Q_n(k)}\right]^b = \frac{\lambda^i}{R_n^b(i)}, \\ R_{n+1}(i) &= \frac{\sum_{k=0}^{C} \frac{\lambda^k}{R_n^b(k)}}{\sum_{k=0}^{C-i} \frac{\lambda^k}{R_n^b(k)}}. \end{aligned} \tag{2.3}$$

The case when $C = 1$ (the usual hard core model) can therefore be studied by analyzing a one-dimensional recursion governed by the following maps:

$$J(x) := \frac{\lambda}{(1+x)^b}, \qquad J_2(x) := J(J(x)) = \frac{\lambda}{(1+\frac{\lambda}{(1+x)^b})^b}. \tag{2.4}$$

Indeed, J defines the recursion for the quantity $Z_n(1)/Z_n(0)$, while J_2 defines the recursion of this quantity between two levels on the tree. We close this section with a summary of the properties of J and J_2 which, when combined with Proposition 2.3, show that $\lambda_{cr}(1) := b^b/(b-1)^{b+1}$ is the phase transition point for the standard hard core model (see, for example, [9]), and that the phase transition for $C = 1$ is second-order. These properties will turn out to also be useful for our analysis of the higher-dimensional recursions (*i.e.,* when $C \geq 2$). We start with the definition of an S-shaped function.

DEFINITION 2.4. *A twice continuously differentiable function* $f : [0,\infty) \mapsto [0,\infty)$ *is said to be* S-shaped *if it has the following properties:*

1. *it is increasing on* $[0,\infty)$ *with* $f(0) > 0$ *and* $\sup_x f(x) < \infty$*;*
2. *there exists* $\overline{x} \in (0,\infty)$ *such that the derivative* f' *is monotone increasing in the interval* $(0,\overline{x})$ *and monotone decreasing in the interval* $(\overline{x},\infty)$*; in other words,* $\overline{x}$ *satisfies* $f''(\overline{x}) = 0$ *and is the unique inflection point of* f*.*

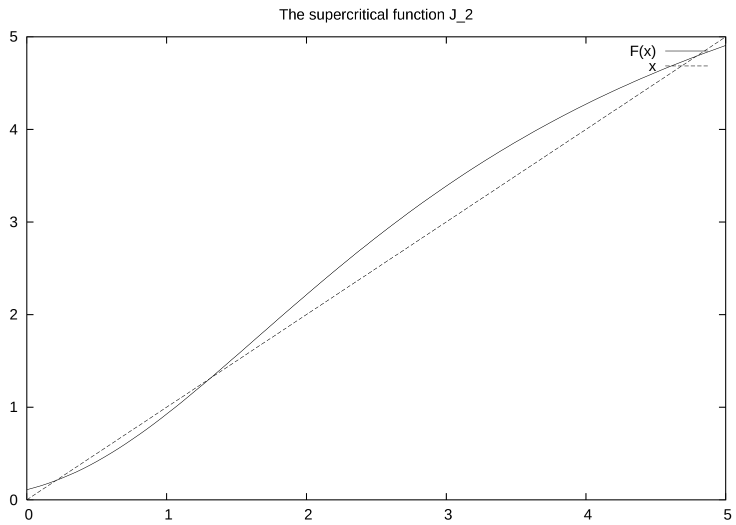

FIG. 2.1. *Graph of the function* $J_2(x)$ *for* $b = 2,\ \lambda = 7$ *(*$\lambda_{cr} = 4$*)*

For future purpose, we observe here that the definition immediately implies that for any $\theta > 0$, and S-shaped function f, θf is also an S-shaped function. It is also easy to verify that any S-shaped function has at most three fixed points in $[0, \infty)$, *i.e.,* points $x \in (0, \infty)$ such that $f(x) = x$. We now summarize the salient properties of J_2 (see e.g. Fig. 2.1), all of which may easily be verified with some calculus.

1. J_2 is an S-shaped function with $J_2(0) = \lambda/(1+\lambda)^b$ and $\sup_x J_2(x) = \lambda$, and a unique point of inflection $x_* \in (0, \infty)$.
2. J has a unique fixed point, x_0, which is also a fixed point of J_2.
3. If $\lambda \le \lambda_{cr}(1)$ then $J_2'(x) \le 1$ for any $x \ge 0$ and x_0 is the unique fixed point of J_2.
4. If $\lambda > \lambda_{cr}(1)$, then J_2 has three fixed points $x_- < x_0 < x_+$, where $J(x_-) = x_+$ and $J(x_+) = x_-$. Moreover $J_2'(x_0) > 1$, $J_2'(x) < 1$ for $x \in [0, x_-] \cup [x_+, +\infty)$ and the three fixed points converge to $x_0(\lambda_{cr}(1))$ as $\lambda \downarrow \lambda_{cr}(1)$.

3. Analysis of the recursions when $C = 2$. When $C = 2$ we have $R_n(1) = \big[1 - \mathbb{P}_n\big(\sigma_r = 2\big)\big]^{-1}$ and (2.3) can be written as:

$$\begin{aligned} R_n(0) &= 1, \\ R_{n+1}(2) &= 1 + \frac{\lambda}{R_n^b(1)} + \frac{\lambda^2}{R_n^b(2)}, \\ R_{n+1}(1) &= \frac{1 + \frac{\lambda}{R_n^b(1)} + \frac{\lambda^2}{R_n^b(2)}}{1 + \frac{\lambda}{R_n^b(1)}} \quad = \frac{R_{n+1}(2)}{1 + \frac{\lambda}{R_n^b(1)}}. \end{aligned} \tag{3.1}$$

On replacing n by $n-1$ in the last equation above, we see that

$$R_n(2) = R_n(1)\left(1 + \frac{\lambda}{R_{n-1}^b(1)}\right).$$

Substituting this back into (3.1), we obtain an exact two-step recursion for $Y_n := R_n(1)$:

$$\begin{aligned} Y_{n+1} &= 1 + \frac{\lambda^2}{\left[1+\frac{\lambda}{Y_n^b}\right]\left[Y_n(1+\frac{\lambda}{Y_{n-1}^b})\right]^b} \\ &= 1 + \frac{\lambda^2}{[Y_n^b+\lambda]\left[1+\frac{\lambda}{Y_{n-1}^b}\right]^b}. \end{aligned} \tag{3.2}$$

It is useful to determine the initial conditions (Y_0, Y_1) for the recursion given the boundary conditions at the 0^{th} level.

$$(Y_0, Y_1) = \begin{cases} (+\infty, 1) & \text{if the b.c. is full } (i.e., \text{identically } C) \\ (1, 1+\frac{\lambda^2}{1+\lambda}) & \text{if the b.c. is empty } (i.e., \text{identically } 0). \end{cases}$$

Numerical calculations of (3.2) using Mathematica strongly suggest that the critical value λ_{cr}, below which the recursion settles to a limit independent of the initial values, takes approximately the following values:

b	λ_{cr}
2	7.2753875
3	3.58029
10	1.107665
100	0.2817409

and that the transition is always first order ($i.e.,$ if $\limsup_n Y_n \neq \liminf_n Y_n$ then their difference is strictly larger than some positive constant δ). Similar observations were made in [15] (see Section 3.4 therein). Here, we provide a rigorous proof of these results.

Let us change variables from Y_n to $X_n := Y_n - 1$ in (3.2). It then follows that

$$X_{n+1} \le \frac{\lambda^2}{\left[\min_{j\ge n} Y_j^b + \lambda\right]\left[1+\frac{\lambda}{(1+X_{n-1})^b}\right]^b} \equiv F_+^{(n)}(X_{n-1}), \tag{3.3}$$

$$X_{n+1} \ge \frac{\lambda^2}{\left[\max_{j\ge n} Y_j^b + \lambda\right]\left[1+\frac{\lambda}{(1+X_{n-1})^b}\right]^b} \equiv F_-^{(n)}(X_{n-1}). \tag{3.4}$$

The maps $F_\pm^{(n)}$ defined above can be rewritten in terms of the map J_2 defined in (2.4) as follows:

$$\begin{aligned} F_-^{(n)}(x) &= \frac{\lambda}{\left(\max_{j\ge n} Y_j^b + \lambda\right)} J_2(x); \\ F_+^{(n)}(x) &= \frac{\lambda}{\left(\min_{j\ge n} Y_j^b + \lambda\right)} J_2(x). \end{aligned}$$

Next, for $\kappa \ge 0$, we define

$$F_\kappa(x) := \frac{\lambda}{\kappa+\lambda} J_2(x), \tag{3.5}$$

so that $F_0 = J_2$. For any $\kappa \geq 0$, F_κ is a strictly positive multiple of J_2 and hence is also an S-shaped function (with the same inflection point x_*). If we denote the fixed points of F_κ by $x_-^{(\kappa)} \leq x_0^{(\kappa)} \leq x_+^{(\kappa)}$ (with the obvious meaning) we see that:

1. if F_κ has a unique fixed point $x_0^{(\kappa)}$ then necessarily $x_0^{(\kappa)} < \min(x_-, x_0)$;
2. since $F'_\kappa(x) = \frac{\lambda}{\kappa+\lambda} J'_2(x)$ necessarily $F'_\kappa(x) \leq 1$ for $x \leq x_-^{(\kappa)}$;
3. the critical value $\lambda_c(\kappa)$ of λ such that F_κ starts to have three fixed points is increasing in κ. In particular,

$$\lambda_c(\kappa) > \lambda_c(0) = \lambda_{cr}(1) = \frac{b^b}{(b-1)^{b+1}};$$

4. if F_κ has three fixed points then necessarily $x_-^{(\kappa)} < x_-$ and $x_0 < x_0^{(\kappa)} < x_+^{(\kappa)}$;
5. the smallest fixed point $x_-^{(\kappa)}$ is continuously differentiable in $\kappa > 0$. Indeed, by the implicit function theorem and the fact that $F'_\kappa(x_-^{(\kappa)}) < 1$, it follows that

$$\begin{aligned} \frac{d}{d\kappa} x_-^{(\kappa)} &= -\frac{\frac{\partial}{\partial\kappa} F_\kappa(x_-^{(\kappa)})}{\frac{\partial}{\partial x} F_\kappa(x_-^{(\kappa)}) - 1} \\ &= -\frac{\frac{1}{\lambda+\kappa} F_\kappa(x_-^{(\kappa)})}{\frac{\lambda}{\lambda+\kappa} J'_2(x_-^{(\kappa)}) - 1} \\ &= -\frac{x_-^{(\kappa)}}{\lambda(1 - J'_2(x_-^{(\kappa)})) + \kappa}. \end{aligned}$$

In what follows, let

$$m := \liminf_n X_n \quad \text{and} \quad M := \limsup_n X_n. \tag{3.6}$$

We are now ready to prove our first result.

Proposition 3.1. *Assume that $\lambda > 0$ is such that F_1 has a unique fixed point. Then $M = m$ and hence the recursion (3.2) has a unique fixed point.*

Proof. Since $Y_n \geq 1$, it follows from (3.3) and (3.5) that $X_{n+1} \leq F_1(X_{n-1})$. Since F_1 is S=shaped and is assumed to have a unique fixed point, this implies that $M \leq x_0^{(1)}$. Moreover, recalling that $m = \liminf_n X_n$, we see that for any $\epsilon > 0$, $X_n \geq m + \epsilon$ for all n large enough. Hence, (3.3) and (3.5) imply that for all large enough n, $X_{n+1} \leq F_\kappa(X_{n-1})$ with $\kappa = (1 + m + \epsilon)^b$. Thus we obtain

$$M \in (0, x_-^{(1+m+\epsilon)^b}). \tag{3.7}$$

Indeed, if F_κ has a unique fixed point, then (3.7) follows immediately. On the other hand, if F_κ has three fixed points then we immediately have $M \in (0, x_-^\kappa) \cup (x_0^\kappa, x_+^\kappa)$. But $M \leq x_0^{(1)}$ and so $M < x_0$ (by property (1) of the F_κ's) and also in this case $x_0 < x_0^{(\kappa)}$ (by property (4) of the F_κ's), giving (3.7).

Since $\epsilon > 0$ is arbitrary in (3.7), we have in fact $M \leq x_-^{(1+m)^b}$. Similarly, using (3.4) and (3.5), we see that $m \geq x_-^{(1+M)^b}$. We want to conclude that necessarily $m = M$. We write

$$M - m \leq \int_m^M ds \, \left(-\frac{d}{ds} x_-^{((1+s)^b)} \right),$$

and the sought statement will follow if, for example,

$$\sup_{m\le s\le M}\left|\frac{d}{ds}x_-^{((1+s)^b)}\right|<1\,.$$

By properties (1) and (4) of F_κ it follows that $x_-^{(\kappa)}<x_0$, and hence property (4) of J_2 implies $J_2'(x_-^{(\kappa)})\le 1$. When combined with the expression for $dx_-^{(\kappa)}/d\kappa$ given in property (5) of F_κ, this implies that

$$\left|\frac{d}{d\kappa}x_-^{(\kappa)}\right|\le\frac{x_-^{(\kappa)}}{\kappa}\,,$$

and hence that

$$\sup_{m\le s\le M}\left|\frac{d}{ds}x_-^{((1+s)^b)}\right|\le\sup_{m\le s\le M}b\,\frac{x_-^{((1+s)^b)}}{1+s}\le b\,x_-^{(1)}$$

where the last inequality uses the fact that $x_-^{(\kappa)}<x_-^{(1)}$ for any $\kappa>0$.

Thus we have to show that $x_-^{(1)}<1/b$. For this purpose it is enough to show that $F_1(1/b)<1/b$. We compute

$$b\,F_1(1/b)=\frac{\lambda}{(1+\lambda)}\frac{b\lambda}{\left(1+\frac{\lambda}{(1+\frac{1}{b})^b}\right)^b}\,. \tag{3.8}$$

Next, we observe that the map $\lambda\mapsto b\lambda/(1+\frac{\lambda}{(1+\frac{1}{b})^b})^b$ achieves its maximum at $\lambda_{\max}=(1+1/b)^b/(b-1)$, where it is equal to $\frac{b}{b-1}\left[\frac{b^2-1}{b^2}\right]^b$. The latter expression is decreasing in b for $b\ge 2$ and for $b=2$ it is equal to $\frac{18}{16}$. Therefore, if $\lambda/(\lambda+1)<\frac{16}{18}$, *i.e.*, $\lambda<8$, then the r.h.s. of (3.8) is strictly less than one. We now examine the case $\lambda\ge 8$. We write

$$\frac{b\lambda}{(1+\frac{\lambda}{(1+\frac{1}{b})^b})^b}\le\frac{\lambda b}{(1+\frac{\lambda}{e})^b}\le\frac{b\,e^b}{\lambda^{b-1}}\le\frac{b\,e^b}{8^{b-1}}<1,\quad\text{for } b\ge 3\,.$$

Finally the case $b=2$ and $\lambda\ge 8$ is handled directly:

$$\frac{2\lambda}{(1+\frac{\lambda}{(1+\frac{1}{2})^2})^2}=\frac{2\lambda}{(1+\frac{4\lambda}{9})^2}\le\frac{16}{(1+\frac{32}{9})^2}\approx 0.77\,.$$

□

Notice that in the proof of the inequality $x_-^{(1)}<1/b$ we did not use the hypothesis that F_1 has only one fixed point. Moreover, we proved something slightly stronger, namely

$$\text{there exists } \epsilon(b)>0 \text{ such that } 1/b-x_-^{(1)}\ge\epsilon(b) \text{ for } \textit{any } \lambda. \tag{3.9}$$

The following monotonicity property is an immediate consequence of Proposition 3.1. Recall that $\lambda_{cr}(1)=b^b/(b-1)^{b+1}$ is the phase transition point for the usual $(C=1)$ hard core model.

COROLLARY 3.2. *For every $\lambda \leq \lambda_{cr}(1)$, the $C = 2$ multi-state hard core model has a unique Gibbs measure.*

Proof. If J_2 has only one fixed point then the same is true of F_1. By Proposition 3.1 there is then only one fixed point for the recursion (3.2). The result then follows from Lemma 2.2 and Proposition 2.3. □

The next result shows that the phase transition for $C = 2$ is first order. Recall the definitions of M and m given in (3.6) and let $\epsilon(b)$ be as in (3.9).

THEOREM 3.3. *If $m \neq M$ then $M - m > \epsilon(b) > 0$.*

Proof. Suppose $m \neq M$. From Proposition 3.1, it then follows that F_1 (and *a fortiori* J_2) has three fixed points $x_-^{(1)} < x_0^{(1)} < x_+^{(1)}$, with $x_0^{(1)} > x_0$. We now show that $x_0 > 1/b$. Indeed, since $J(x_0) = x_0$ and J is strictly decreasing, it is enough to check that $J(1/b) > 1/b$ or, equivalently, that $\lambda/(1+\frac{1}{b})^b > 1/b$. But $\lambda > \frac{b^b}{(b-1)^{b+1}}$ and clearly

$$\frac{b^b}{(b-1)^{b+1}(1+\frac{1}{b})^b} = \frac{b^{2b}}{\left(b^2-1\right)^b(b-1)} > \frac{1}{b}\,.$$

Since $1/b - x_-^{(1)} \geq \epsilon(b)$ by (3.9), this implies $x_0^{(1)} - x_-^{(1)} > \epsilon(b)$.

Next, since $X_n = \left[1 - \mathbb{P}_n\left(\sigma_r = 2\right)\right]^{-1} - 1$, we infer that X_n is maximized by the empty b.c. and minimized by the full b.c. if n is odd (and vice versa if n is even). Thus, using the recursive inequality $X_{n+1} \leq F_1(X_{n-1})$, we obtain for any odd n, the inequality $X_n \leq U_n$, where $\{U_n, n \text{ odd}\}$ is the sequence that satisfies the recursion $U_{n+2} = F_1(U_n)$, with $U_1 = 0$. In particular, $m \leq x_-^{(1)} \leq \frac{1}{b} - \epsilon(b)$. If now $M \leq m + \epsilon(b) < x_0^{(1)}$ then necessarily $X_n < x_0^{(1)}$ for any n large enough and repeated iterations of $X_{n+1} \leq F_1(X_{n-1})$ imply $M \leq x_-^{(1)}$. At this stage we are back in the framework of the proof of Proposition 3.1 and $m = M$, resulting in a contradiction. □

4. The Large b Asymptotic Regime. In this section we set up and then analyze the recursion for any value of C when b is large. In what follows, $e = \exp(1)$.

For any $j \leq C$ set $j^* = C - j$. Also, for $\lambda < 1$, set $A_\lambda = \sum_{i=0}^{\infty} \lambda^i = (1-\lambda)^{-1}$. Iterating (2.3) we obtain

$$R_{n+2}(j) = 1 + \frac{\displaystyle\sum_{i=j^*+1}^{C} \lambda^i \left(\sum_{k=0}^{i^*} \frac{\lambda^k}{R_n^b(k)}\right)^b}{\displaystyle\sum_{i=0}^{j^*} \lambda^i \left(\sum_{k=0}^{i^*} \frac{\lambda^k}{R_n^b(k)}\right)^b}\,.$$

In turn, this implies that

$$\begin{aligned}
R_{n+2}(j) &\leq 1 + \frac{A_\lambda \lambda^{j^*+1} \left(\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)} \right)^b}{\left(\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)} + \sum_{k=j}^{C} \frac{\lambda^k}{R_n^b(k)} \right)^b} \\
&= 1 + \frac{A_\lambda \lambda^{j^*+1}}{\left(1 + \frac{\sum_{k=j}^{C} \frac{\lambda^k}{R_n^b(k)}}{\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)}} \right)^b} \\
&\leq 1 + \frac{A_\lambda \lambda^{j^*+1}}{\left(1 + A_\lambda^{-1} \frac{\lambda^j}{R_n^b(j)} \right)^b} .
\end{aligned}$$

Therefore, by letting $X_n(j) = R_n(j) - 1$ we have

$$X_{n+2}(j) \leq A_\lambda^2 \lambda^{j^*-j+1} J_2^{(\lambda_j)}(X_n) \equiv F_+^{(j)}(X_n(j)) , \tag{4.1}$$

where $\lambda_j := A_\lambda^{-1} \lambda^j$, and $J^{(\lambda)} = J$ and $J_2^{(\lambda)} = J_2$ are the maps defined in (2.4), but with the λ dependence now denoted explicitly.

In a similar fashion, we obtain a lower bound

$$\begin{aligned}
R_{n+2}(j) &\geq 1 + \frac{\lambda^{j^*+1} \left(\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)} \right)^b}{A_\lambda \left(\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)} + \sum_{k=j}^{C} \frac{\lambda^k}{R_n^b(k)} \right)} \\
&= 1 + \frac{A_\lambda^{-1} \lambda^{j^*+1}}{\left(1 + \frac{\sum_{k=j}^{C} \frac{\lambda^k}{R_n^b(k)}}{\sum_{k=0}^{j-1} \frac{\lambda^k}{R_n^b(k)}} \right)^b} \\
&\geq 1 + \frac{A_\lambda^{-1} \lambda^{j^*+1}}{\left(1 + A_\lambda \frac{\lambda^j}{R_n^b(j)} \right)^b} .
\end{aligned}$$

Therefore, we have

$$X_{n+2}(j) \geq A_\lambda^{-2} \lambda^{j^*-j+1} J_2^{(\lambda'_j)}(X_n) \equiv F_-^{(j)}(X_n(j)) , \tag{4.2}$$

where $\lambda'_j := A_\lambda \lambda^j$.

4.1. The case of C odd. We start by stating the main result of the section. Recall that for $\lambda < 1$, $A_\lambda = (1-\lambda)^{-1}$.

THEOREM 4.1. *Let $j_c = \lceil \frac{C}{2} \rceil$, and define $\lambda_- := A_\lambda^{-1} \lambda^{j_c}$ and $\lambda_+ := A_\lambda \lambda^{j_c}$. Then the following two properties hold:*

1. *If* $\left(\frac{\gamma}{b}\right)^{\frac{1}{j_c}} \leq \lambda < 1$ *with* $\gamma > e$ *, then, for any* b *large enough depending on* γ*, the smallest fixed point of*

$$x \mapsto A_\lambda^2 J_2^{(\lambda_-)}(x) \tag{4.3}$$

is strictly smaller than the largest fixed point of

$$x \mapsto A_\lambda^{-2} J_2^{(\lambda_+)}(x)\,. \tag{4.4}$$

In particular, there is phase coexistence.

2. *On the other hand, if* $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$ *with* $\gamma' < e$ *then, for every* b *large enough, depending on* γ'*, there is a unique Gibbs measure.*

We start by establishing the first assertion of the theorem. Our proof will make use of the following elementary observation.

LEMMA 4.2. *For* $\gamma > 0$ *the function* $H_\gamma : [0,\infty) \mapsto [0,\infty)$ *defined by*

$$H_\gamma(z) = \gamma\, e^{-\gamma e^{-z}}, \quad z \in [0,\infty),$$

is S*-shaped. In addition, the following two properties hold:*

1. *if* $\gamma \leq e$ *then* H_γ *has one fixed point* $z_0 < 1$*;*
2. *if* $\gamma > e$ *then* H_γ *has three distinct fixed points* $z_- < z_0 < z_+$ *that satisfy*

$$0 \leq z_- \leq \log(\gamma) - \log(\log(\gamma)) < z_0 \leq \log\gamma < z_+\,. \tag{4.5}$$

Proof. The function H_γ is clearly twice continuously differentiable, satisfies $H_\gamma(0) = \gamma > 0$ and $\sup_x H_\gamma(x) = \gamma e^{-\gamma} < \infty$. That it is S-shaped therefore follows from the fact that

$$H_\gamma'(z) = \gamma e^{-z} H_\gamma(z) > 0 \quad \text{and} \quad H_\gamma''(z) = \gamma e^{-z} H_\gamma(z)\big[\gamma e^{-z} - 1\big]\,.$$

Now suppose $\gamma < e$. Then $\sup_z H_\gamma'(z) < 1$ and therefore there exists a unique fixed point z_0. The fact that $z_0 < 1$ follows from the observation that

$$H_\gamma(1) = \gamma e^{-\gamma e^{-1}} < 1\,.$$

On the other hand, if $\gamma = e$ the value $z_0 = \log\gamma$ is the unique fixed point, and satisfies $H_\gamma'(z_0) = 1$. Lastly, for $\gamma > e$, we have the inequalities

$$\begin{aligned} H_\gamma'(\log\gamma) &> 1,\\ H_\gamma(\log\gamma) &> \log\gamma,\\ H_\gamma\big(\log\gamma - \log(\log\gamma)\big) &< \log\gamma - \log(\log\gamma), \end{aligned}$$

where the last inequality holds because $H_\gamma\big(\log\gamma - \log(\log\gamma)\big) = 1$ and $\gamma \mapsto \log\gamma - \log(\log\gamma)$ restricted to the interval $[e,\infty)$ is increasing with $\log(e) - \log(\log(e)) = 1$. Together with the S-shaped property of H, these inequalities immediately imply that H has three fixed points that satisfy (4.5). □

We are now ready to establish the first statement of Theorem 4.1.

Proof. [Proof of Theorem 4.1(1)] Fix $\lambda \in [\left(\frac{\gamma}{b}\right)^{\frac{1}{j_c}}, 1)$ with $\gamma > e$, and for notational conciseness, denote A_λ simply by A. We first show that the asserted inequality between the fixed points of the two maps implies phase coexistence. This is a simple consequence of the fact that, for any boundary condition τ, the sequence $\{X_n^*\}$ defined by

$$X_n^* \equiv X_n(j_c) = \mu_{T_n}^\tau(\sigma_r \geq j_c)/\mu_{T_n}^\tau(\sigma_r \leq j_c), \quad n \in \mathbb{N},$$

obeys the recurrence

$$A^{-2} J_2^{(\lambda_+)}(X_n) \leq X_{n+2}^* \leq A^2 J_2^{(\lambda_-)}(X_n^*)$$

where we have made use of (4.1) and (4.2), together with the duality property $j_c^* + 1 = j_c$. If now $\lfloor \frac{C}{2} \rfloor$ boundary conditions are imposed at the zeroth level then $X_0^* = 0$ and X_n^* will always be smaller than the smallest fixed point of $x \mapsto A^2 J_2^{(\lambda_-)}(x)$. On the other hand, under $\lceil \frac{C}{2} \rceil$ boundary conditions, $X_0^* = 1$ and X_n^* will always be larger than the largest fixed point of $x \mapsto A^{-2} J_2^{(\lambda_+)}(x)$ because the range of this mapping is contained in $[0,1]$ for large b.

We now prove our statement concerning the fixed points of (4.3), (4.4). First, consider the case $\lambda = \left(\frac{\gamma}{b}\right)^{\frac{1}{j_c}}$ and observe that for any $z > 0$,

$$\lim_{b \to \infty} b\, A^{-2} J_2^{(\lambda_+)}(z/b) = \lim_{b \to \infty} b\, A^2 J_2^{(\lambda_-)}(z/b) = H_\gamma(z), \tag{4.6}$$

uniformly on bounded intervals. Next, we define

$$\tilde{x}_- := \frac{\log\gamma - \log(\log\gamma)}{b} \qquad \text{and} \qquad \tilde{x}_+ := \frac{\log\gamma}{b}.$$

From Lemma 4.2, it follows that $H_\gamma(b\tilde{x}_-) < b\tilde{x}_- < b\tilde{x}_+ < H_\gamma(b\tilde{x}_+)$. Together with (4.6), this shows that for any b large enough,

$$A^2 J_2^{(\lambda_-)}(\tilde{x}_-) < \tilde{x}_- < \tilde{x}_+ < A^{-2} J_2^{(\lambda_+)}(\tilde{x}_+),$$

and the first assertion of the lemma follows (for this case) because $A^{-2} J_2^{(\lambda_+)}$ and $A^2 J_2^{(\lambda_-)}$ are S-shaped exactly like H_γ.

We now consider the case $\left(\frac{\gamma}{b}\right)^{\frac{1}{j_c}} \leq \lambda < 1$ and again we compute

$$A^2 J_2^{(\lambda_-)}(\tilde{x}_-) \leq A^2 \frac{\lambda_-}{(1 + \lambda_- e^{-b\tilde{x}_-})^b} = A^2 \frac{\lambda_-}{\left(1 + \lambda_- \frac{\log\gamma}{\gamma}\right)^b}. \tag{4.7}$$

If λ does not tend to zero as $b \to \infty$, then it is obvious that the r.h.s of (4.7) is smaller than $\tilde{x}_-$ for large enough b. If instead $\lim_{b\to\infty} \lambda = 0$ we proceed as follows. The function $f_\gamma(\lambda) = \lambda / \left(1 + \lambda \frac{\log\gamma}{\gamma}\right)^b$ satisfies

$$f_\gamma'(\lambda) = \frac{1}{\left(1 + \lambda \frac{\log\gamma}{\gamma}\right)^{2b}} \left(1 - \frac{b\lambda\log\gamma}{\gamma + \lambda\log\gamma}\right),$$

and hence is decreasing in the interval $(\frac{\gamma}{(b-1)\log\gamma}, \infty)$. Since $\gamma > e$ and our assumption $\lambda \to 0$ implies $A = A_\lambda \approx 1$ for large b, we have the inequality

$$\lambda_- > A^{-1}\gamma/b > \gamma/((b-1)\log\gamma).$$

Thus, we can conclude that the r.h.s of (4.7) is smaller than the same expression with λ_- replaced by $A^{-1}\gamma/b$. After this replacement, the resulting r.h.s of (4.7) is indeed smaller than $\tilde{x}_-$ for all large enough b because of (4.5) and (4.6). In conclusion, we have shown that for any $\left(\frac{\gamma}{b}\right)^{\frac{1}{j_c}} \leq \lambda < 1$ the function $A^2 J_2^{(\lambda_-)}$ has a fixed point smaller than $\tilde{x}_-$.

Next, we examine $A^{-2}J_2^{(\lambda_+)}$. If $\lim_{b\to\infty} b\lambda_+ = \infty$ then it easily follows that for large b, we have $A^{-2}J_2^{(\lambda_+)}(A^{-2}\lambda_+/2) > A^{-2}\lambda_+/2 \gg \tilde{x}_-$. If instead $\lambda_+ \leq C/b$ for some finite constant C, we choose $x_\lambda = \log(b\lambda_+)/b > x_-$ and write

$$A^{-2}J_2^{(\lambda_+)}(x_\lambda) \geq A^{-2}\lambda_+ e^{-b\lambda_+/(1+x_\lambda)^b}\,.$$

By construction, $\lim_{b\to\infty} e^{-b\lambda_+/(1+x_\lambda)^b} = e^{-1}$. Therefore, for sufficiently large b,

$$A^{-2}\lambda_+\, e^{-b\lambda_+/(1+x_\lambda)^b} \geq (1-O(b^{-1}))\lambda_+\, e^{-1} \geq x_\lambda\,,$$

because $\lambda_+ > \gamma/b$ with $\gamma > e$. In conclusion $A^{-2}J_2^{(\lambda_+)}(x)$ has a fixed point strictly bigger than x_- and the existence of a phase transition follows. ∎

We now turn to the proof of the second assertion of Theorem 4.1, namely the absence of a phase transition for $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$, with $\gamma' < e$. For this, we first establish two preliminary results in Lemmas 4.3 and 4.4. For any vertex $y \in T_n$ and $i \in S_C$, we define a probability measure $\mu_y^{(i)}$ on the set of spins at y as follows:

$$\mu_y^{(i)}(\sigma_y = j) \doteq \mathbb{P}(\sigma_y = j \,|\, \sigma_y \leq i^*), \quad j \in S_C, \tag{4.8}$$

with $\mathbb{P}$, as always, depending on λ and a boundary condition on T_n (which for clarity we have suppressed in the notation). Note that if x is a site in T_n that is neighbouring to y, then $\mu_y^{(i)}$ represents the marginal on y of the Gibbs measure (with some boundary condition on the leaves of T_n), conditioned to have i particles at x. Recall that $\|\cdot\|_{\mathrm{TV}}$ denotes the total variation distance.

Lemma 4.3. *For any $k < i$, we have*

$$\|\mu_y^{(i)} - \mu_y^{(k)}\|_{TV} = \frac{\mu_y^{(0)}\left(\sigma_y \in [i^*+1, k^*]\right)}{\mu_y^{(0)}\left(\sigma_y \leq k^*\right)}\,.$$

Proof. By definition $\mu_y^{(i)}(\sigma_y = j) = \mu_y^{(0)}(\sigma_y = j \,|\, \sigma_y \leq i^*)$. Therefore, also recalling that $k < i$ implies $k^* > i^*$, we have

$$\begin{aligned}
\|\mu_y^{(i)} - \mu_y^{(k)}\|_{\mathrm{TV}} &= \frac{1}{2}\sum_{j=0}^{i^*}\left\|\mu_y^{(i)}(\sigma_y = j) - \mu_y^{(k)}(\sigma_y = j)\right\| + \frac{1}{2}\sum_{j=i^*+1}^{k^*}\mu_y^{(k)}(\sigma_y = j)\\
&= \frac{1}{2}\,\frac{\mu_y^{(0)}(\sigma_y \leq k^*) - \mu_y^{(0)}(\sigma_y \leq i^*)}{\mu_y^{(0)}(\sigma_y \leq k^*)} + \frac{1}{2}\,\frac{\mu_y^{(0)}(i^*+1 \leq \sigma_y \leq k^*)}{\mu_y^{(0)}(\sigma_y \leq k^*)}\\
&= \frac{\mu_y^{(0)}\left(\sigma_y \in [i^*+1, k^*]\right)}{\mu_y^{(0)}\left(\sigma_y \leq k^*\right)}.
\end{aligned}$$

∎

Notice that if x is an ancestor of y then $\mu_y^{(0)}$ is nothing but the Gibbs measure on the tree $\mathbb{T}_y^b$ rooted at y with the boundary conditions induced by those on T_n. If instead y is an ancestor of x then $\mu_y^{(0)}$ becomes a Gibbs measure on the (non regular) tree $T_n \setminus \mathbb{T}_x^b$. However, if x, y are sufficiently below the root of T_n, then $T_n \setminus \mathbb{T}_x^b$ will coincide with a regular tree rooted at y for a large number of levels. That is all that we need to prove uniqueness below $\left(\frac{e}{b}\right)^{\frac{1}{j_c}}$.

In what follows, given any non negative function $b \mapsto f(b)$ of the degree of the tree $\mathbb{T}^b$, we will write $f(b) \approx 0$ if $\lim_{b\to\infty} bf(b) = 0$.

LEMMA 4.4. *Fix $\gamma' < e$ and assume $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$. Then there exists $a < 1$ and $n_0 \in \mathbb{N}$ such that for any $n \geq n_0$ and any boundary condition τ on the leaves of T_n,*

$$\limsup_{b\to\infty} b\mu^\tau(\sigma_r \geq i^* + 1) \leq \begin{cases} 0 & \text{if } i \leq \lfloor \frac{C}{2} \rfloor \\ a & \text{if } i = j_c = \lceil \frac{C}{2} \rceil \end{cases}$$

Proof. It suffices to bound $X_n(i)$ from above for $i \leq \lfloor \frac{C}{2} \rfloor$ or $i = \lceil \frac{C}{2} \rceil$. In the first case, when $i \leq \lfloor \frac{C}{2} \rfloor$, the stated bound follows easily since (4.1) and the assumed bound on λ imply that for some finite constant K,

$$bX_n(i) \leq \lambda^{i^*+1} b \leq K b^{(1-\frac{i^*+1}{j_c})} \approx 0\,.$$

In the second case, when $i = j_c$, set $a_\infty := \limsup_{b\to\infty} b\hat{x}_+(b)$, where $\hat{x}_+(b)$ is the largest fixed point of the S-shaped function $x \mapsto A_\lambda J_2^{(\lambda_-)}(x)$. Due to the assumption $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$, it follows that $a_\infty \leq \gamma'$. Because of (4.1) it is enough to prove that $a_\infty < 1$. Assume the contrary. Then the fixed point equation, together with $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$, readily implies that

$$a_\infty \leq \gamma e^{-\gamma e^{-a_\infty}},$$

which in turn implies that a_∞ must be smaller than the unique fixed point z_0 of the map H. Since $\gamma e^{-\gamma/e} < 1$ if $\gamma < e$ necessarily $z_0 < 1$ and we get a contradiction. Note that in the above proof by contradiction, the hypothesis $a_\infty \geq 1$ enters as follows. If $x > 1 - \delta$, $0 < \delta \ll 1$ then $J_2^{(\lambda)}(x)$ is increasing in λ and so we may safely assume $\lambda = \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$ and not just smaller or equal. ◻

We are now ready to prove uniqueness for $\lambda \leq \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$.

Proof. [Proof of Theorem 4.1(2)] For simplicity we begin with $\lambda = \left(\frac{\gamma'}{b}\right)^{\frac{1}{j_c}}$. In this case, it follows immediately from the basic inequality (4.1) that for any initial condition, any $n \geq 2$ and any b large enough, there exist constants $c_1,\ c_2$ such that

$$X_n^{\lceil C/2 \rceil} \leq c_1 e^{-c_2\, b^\alpha}\,, \tag{4.9}$$

where $\alpha = 1/(j_c + 1)$. In another words, recalling the probability measure $\mu_y^{(i)}$ introduced in (4.8) and using the obvious fact that for any $i \leq C$,

$$\mu_y^{(i)}\left([j_c + 1, C]\right) \leq X_n^{\lceil C/2 \rceil}\,,$$

we get that the probability of having more than j_c particles at y given i particles at x is exponentially small in b.

Now, recall that T_ℓ is the finite-tree of depth ℓ rooted at r, and let τ, τ' be two boundary conditions on the leaves of T_ℓ that differ at only one vertex v_0. Let also $\Gamma = \{v_0, v_1, \dots, v_\ell\}$ be the unique path joining v_0 to the root $r = v_\ell$. We recursively couple the corresponding measures $\mu^\tau \doteq \mu^\tau_{T_\ell,\lambda}$ and $\mu^{\tau'} \doteq \mu^{\tau'}_{T_\ell,\lambda}$ by repeatedly applying the following step. Assume that, for any pair $(\sigma_{v_1}, \sigma_{v_2})$ with $\sigma_{v_1} \neq \sigma'_{v_1}$ we can couple $\mu^\tau(\cdot \mid \sigma_{v_1})$ and $\mu^{\tau'}(\cdot \mid \sigma'_{v_1})$ and call $\nu_{\ell-1}^{\sigma_{v_1},\sigma'_{v_1}}$ the coupled measure. It is understood that $\nu_{\ell-1}^{\sigma_{v_1},\sigma'_{v_1}}$ is concentrated along the diagonal if $\sigma_{v_1} = \sigma'_{v_1}$. Let $\pi_1^{\tau_{v_0},\tau'_{v_0}}$ be the coupling of the marginals on of the two Gibbs measures on v_1 that realizes the variation distance (*i.e.,* $\pi_1^{\tau_{v_0},\tau'_{v_0}}(\sigma_{v_1} \neq \sigma'_{v_1}) = \|\mu^\tau_{v_1} - \mu^{\tau'}_{v_1}\|_{TV}$). Then we set

$$\nu_\ell^{\sigma_{v_1},\sigma'_{v_1}}(\sigma, \sigma') = \pi_1^{\tau_{v_0},\tau'_{v_0}}(\sigma_{v_1}, \sigma'_{v_1}) \nu_{\ell-1}^{\sigma_{v_1},\sigma'_{v_1}}(\sigma_{\mathbb{T}^b_\ell \setminus v_1}, \sigma'_{\mathbb{T}^b_\ell \setminus v_1}) \,.$$

If we iterate the above formula we finally get a coupling $\nu^{\tau,\tau'}$ such that the probability of seeing a discrepancy at the root can be expressed as

$$\sum_{\substack{\sigma_{v_1} \neq \sigma'_{v_1} \\ \eta_{v_2} \neq \eta'_{v_2} \\ \dots}} \pi_1^{\tau_{v_0},\tau'_{v_0}}(\sigma_{v_1}, \sigma'_{v_1})\, \pi_2^{\sigma_{v_1},\sigma'_{v_1}}(\eta_{v_2}, \eta'_{v_2}) \pi_3^{\eta_{v_2},\eta'_{v_2}} \dots \tag{4.10}$$

with self explanatory notation. If we can show that the above expression tends to zero as $\ell \to \infty$ faster than $b^{-\ell}$ uniformly in τ, τ', then uniqueness will follow by a standard path coupling (or triangle inequality) argument (see, for example, [6]).

On the state space $S := [0, \dots, C]^2$ consider a non-homogeneous Markov chain $\{\xi_t\}_{t=0}^\ell$ with transition matrix at time t given by $P_t(\xi, \xi') = \pi_t^\xi(\xi')$ and initial condition $\xi_0 = (\tau_{v_0}, \tau'_{v_0})$. Let also $B = \{(i,j) \in S \ : i \geq j_c + 1\} \cup \{(i,j) \in S^2 \ : j \geq j_c + 1\}$ be the bad set and let $D = \{(i,i) \in S \ : i \in [0, \dots, C]\}$ be the diagonal. Equation (4.10) is then nothing but the probability that the chain does not hit D within time ℓ.

For b large enough (depending only on $\gamma' < e$) the two key properties of the chain, which immediately follow from Lemmas 4.3 and 4.4 and the inequality (4.9), are the following:

$$\sup_t \sup_{\xi \in B^c} P_t(\xi, D^c) \leq \frac{a}{b}, \quad a < 1 \tag{4.11}$$

$$\sup_t \sup_\xi P_t(\xi, B) \leq c_1 e^{-c_2\, b^\alpha}, \quad \alpha > 0 \,. \tag{4.12}$$

Notice that it is not difficult to show that

$$\sup_t \sup_{\xi \in B} P_t(\xi, D^c) \approx \lambda \gg 1/b \,.$$

In other words, the probability of not entering the diagonal D in one step is suitably small (*i.e.,* smaller than a/b, $a < 1$) only if we start from the good set B^c. Using (4.11) and (4.12), we can immediately conclude that

$$\begin{aligned}
\mathbb{P}(\xi_t \notin D \text{ for all } 0 \leq t \leq \ell) &\leq \sum_{k=0}^{\ell} \binom{\ell}{k} \left(c_1 e^{-c_2\, b^\alpha}\right)^k \left(\frac{a}{b}\right)^{\ell-2k-1} \\
&\leq \frac{b}{a} \left(\frac{b}{a} c_1 e^{-c_2\, b^\alpha} + \frac{a}{b}\right)^\ell \,.
\end{aligned} \tag{4.13}$$

The "-1" in the exponent of a/b above takes into account the fact that we may start at x_0 in the bad set B, while the extra "$-k$" in the exponent accounts for the fact that for any transition from B to B^c we do not necessarily have a good coupling bound. It is clear that the right hand side of (4.13) tends to zero faster than $b^{-\ell}$ as $\ell \to \infty$ because $a < 1$. ∎

4.2. The case of C even. Throughout this discussion, we assume C even and we set $j_c = \frac{C}{2} + 1$. Notice that $j_c = (\frac{C}{2})^* + 1$.

Theorem 4.5. *Assume $\lambda = \left(\gamma \frac{\log b}{b}\right)^{\frac{1}{j_c}}$ with $\gamma > 1/(C+2)$. Then, for any large enough b there is phase coexistence. If instead $\gamma < \frac{1}{C+2}$, for any large enough b there is a unique Gibbs measure.*

Proof. Fix $\gamma > \frac{1}{C+2}$ and assume $\lambda = \left(\gamma \frac{\log b}{b}\right)^{\frac{1}{j_c}}$. We will show that the largest fixed point of $F_-^{(\frac{C}{2})}$ is strictly larger than the smallest fixed point of $F_+^{(\frac{C}{2})}$. By the usual argument that is enough to prove phase coexistence.

Pick α halfway between $1/(C+2)$ and γ and compute the value $F_-^{(\frac{C}{2})}(\frac{\alpha \log b}{b})$ for large b. From the definition we get

$$F_-^{(\frac{C}{2})}(\frac{\alpha \log b}{b}) \approx \frac{\gamma \log b}{b}\, e^{-b^{\frac{1}{C+2}-\alpha}} \approx \frac{\gamma \log b}{b} \gg \frac{\alpha \log b}{b}\,.$$

Therefore there exists a fixed point of $F_-^{(\frac{C}{2})}$ greater than $\frac{\alpha \log b}{b}$. On the other hand

$$F_+^{(\frac{C}{2})}\left(2\gamma \frac{\log b}{b} e^{-b^{1/(C+2)}}\right) \approx \gamma \frac{\log b}{b} e^{-b\lambda^{\frac{C}{2}}} \ll 2\gamma \frac{\log b}{b} e^{-b^{1/(C+2)}}\,, \tag{4.14}$$

so that $F_+^{(\frac{C}{2})}$ has a fixed point smaller than $2\gamma \frac{\log b}{b} e^{-b^{1/(C+2)}}$, and now the first statement of the theorem follows.

Assume now $\gamma < \frac{1}{C+2}$. In that case, using (4.1), we infer that, for any boundary condition and any large enough b,

$$\mu_{T_n}^{\tau}\left(\sigma_r \geq \tfrac{C}{2} + 1\right) \leq X_n^{(C/2)} \leq e^{-b^a}\,, \quad a = \frac{1}{C+2} - \gamma$$

The proof of uniqueness follows now exactly the same lines of the odd case with the difference that now the bad set is $B = \{C/2+1, \ldots, C\}$ and (4.11), (4.12) are changed into

$$\sup_t \sup_{\xi \in B^c} P_t(\xi, D^c) \leq c_1 e^{-c_2\, b^{\alpha}}\,, \quad \alpha > 0 \tag{4.15}$$

$$\sup_t \sup_{\xi} P_t(\xi, B) \leq c_1 e^{-c_2\, b^{\alpha}}\,, \quad \alpha > 0\,. \tag{4.16}$$

∎

4.3. First-order phase transitions for C even and large b. We now turn to showing that for all even C and large enough b (depending on C), the phase transition established in Theorem 4.5 is first-order. At the end of Section 4.2 we showed that as λ varies, for example, in the interval

$$\left[\left(\frac{\log b}{b}\right)^{\frac{2}{C+2}}, \left(\frac{3\log b}{b}\right)^{\frac{2}{C+2}}\right],$$

the values of

$$m(\lambda) := \limsup_{n\to\infty} \left[\mu^C_{T_n}(\sigma_r > C/2) - \mu^0_{T_n}(\sigma_r > C/2)\right]$$

vary between 0 and $\Omega(2\frac{\log b}{b})$. (Recall that the superscripts C and 0 indicate full b.c. and empty b.c., respectively.) Notice that, by monotonicity, the $\limsup_n$ above is attained over the sequence of even n's and that $\mu^C_{T_{2n}}(\sigma_r > C/2)$ is decreasing in n.

Here, we argue that in the above interval $m(\lambda)$ cannot be continuous. The starting point is the observation that, because of (4.14), for all

$$\lambda \in \left[\left(\frac{\log b}{b}\right)^{\frac{2}{C+2}}, \left(\frac{3\log b}{b}\right)^{\frac{2}{C+2}}\right]$$

the smallest fixed point of $F_+^{(\frac{C}{2})}$ is exponentially small in b^α for some $\alpha > 0$. Thus, in particular, there exist constants c_1, c_2 such that

$$\mu^0_{T_n}(\sigma_r > C/2) \le c_1 e^{-c_2\, b^\alpha}, \quad \forall n \ge 1\,.$$

Fix now $\delta < 1$ and assume that for some n_0,

$$\mu^C_{T^b_{2n_0}}(\sigma_r > C/2) \le \frac{\delta}{b}\,.$$

By monotonicity that implies

$$\sup_{n\ge 2n_0} \sup_\tau \mu^C_{T_n}(\sigma_r > C/2) \le \frac{\delta}{b}\,.$$

Thus we can proceed with the previously described coupling argument with (4.15) and (4.16) replaced by

$$\sup_{t\ge 2n_0} \sup_{\xi\in B^c} P_t(\xi, D^c) \le \frac{\delta}{b}, \tag{4.17}$$

$$\sup_{t\ge 2n_0} \sup_{\xi} P_t(\xi, B) \le c_1 e^{-c_2\, b^\alpha}\,. \tag{4.18}$$

and we may conclude that $m(\lambda) = 0$.

In other words we have shown that $m(\lambda) > 0$ implies that for all n,

$$\mu^C_{T_{2n}}(\sigma_r > C/2) > \frac{\delta}{b},$$

so that

$$m(\lambda) \ge \frac{\delta}{b} - c_1 e^{-c_2\, b^\alpha}\,.$$

It follows now that the phase transition is first-order.

Acknowledgments. The authors thank Microsoft Research, particularly the Theory Group, for its hospitality, and for facilitating this collaboration. The authors are also thankful to Christian Borgs, Roman Kotecký and Ilze Ziedins for useful discussions in the early stages of this research.

REFERENCES

[1] T. Ballardie, P. Francis and J. Crowcroft. Core-based trees (CBT): An architecture for scalable inter-domain multicast routing, *SIGCOMM 93*, 1993.
[2] R. J. Baxter. *Exactly Solved Models in Statistical Mechanics*, Academic Press, 1982.
[3] J. van den Berg and J. E. Steif. Percolation and the hard-core lattice gas model, *Stoch. Proc. Appl.* 49:179-197, 1995.
[4] G. R. Brightwell and P. Winkler. Graph homomorphisms and phase transitions, *J. Comb. Theory (Series B)* 77:415-435, 1999.
[5] G. R. Brightwell and P. Winkler. Hard constraints and the Bethe lattice: adventures at the interface of combinatorics and statistical physics, *Proc. ICM 2002*, Higher Education Press, Beijing, IIIi:605 – 624, 2002.
[6] R. Bubley and M. Dyer, Path coupling: A technique for proving rapid mixing in Markov chains, *Proc. IEEE FOCS* (1997), 223–231.
[7] D. Galvin and J. Kahn. On phase transition in the hard-core model on $\mathbb{Z}^d$, *Comb. Prob. Comp.* **13** (2004), 137–164.
[8] H.-O. Georgii. *Gibbs Measures and Phase Transitions*, de Gruyter, Berlin, 1988.
[9] F. Kelly. Loss networks, *Ann. Appl. Probab.* **1** no. 3 (1991), 319–378.
[10] L. Kleinrock. *Queueing Systems, Volume 2: Computer Applications*, Wiley, New York, 1976.
[11] G. Louth. *Stochastic networks: complexity, dependence and routing*, Cambridge University (thesis), 1990.
[12] B. Luen, K. Ramanan and I. Ziedins. Nonmonotonicity of phase transitions in a loss network with controls, *Annals of Applied Probability*, **16**, 3:1528–1562, 2006.
[13] J. Martin, U. Rozikov and Y. Suhov. A three state hard-core model on a Cayley tree, *Journal of Nonlinear Mathematical Physics*, **12**, 3:432–448, 2005.
[14] A. Mazel and Y. Suhov. Random surfaces with two-sided constraints: an application of the theory of dominant ground states, *J. Stat. Phys.* 64:111-134, 1991.
[15] K. Ramanan, A. Sengupta, I. Ziedins and P. Mitra. Markov random field models of multicasting in tree networks, *Advances in Applied Probability*, **34**, 1:1-27, 2002.
[16] F. Spitzer. Markov random fields on an infinite tree, *Ann. Prob.* 3:387-398, 1975.
[17] Y. Yang, and J. Wang, On blocking probability of multicast networks, *IEEE Transactions on Communications*, **46**, 7:957–968, 1998.
[18] S. Zachary, Countable state space Markov random fields and Markov chains on trees, *Ann. Prob.* 11:894-903, 1983.
[19] S. Zachary, Bounded, attractive and repulsive Markov specifications on trees and on the one-dimensional lattice, *Stoch. Proc. Appl.* 20:247-256, 1985.