SECOND-ORDER MARKOV RANDOM FIELDS FOR INDEPENDENT SETS ON THE INFINITE CAYLEY TREE

BY DAVID A. GOLDBERG

Georgia Institute of Technology

Recently, there has been significant interest in understanding the properties of Markov random fields (M.r.f.) defined on the independent sets of sparse graphs. When these M.r.f. are restricted to pairwise interactions (i.e., hardcore model), much progress has been made. However, considerably less is known in the presence of higher-order interactions, which arise, for example, in the analysis of independent sets with special properties and the study of resource-constrained communication networks. In this paper, we further our understanding of such models by analyzing M.r.f. with second-order interactions on the independent sets of the infinite Cayley tree. We prove that the associated Gibbsian specification satisfies the celebrated FKG inequality whenever the local potentials defining the Hamiltonian satisfy a log-convexity condition. Under this condition, we give necessary and sufficient conditions for the existence of a unique infinite-volume Gibbs measure in terms of an explicit system of equations, prove the existence of a phase transition and give explicit bounds on the associated critical activity, which we prove to exhibit a certain robustness. For potentials which are small perturbations of those coinciding to the hardcore model at the critical activity, we characterize whether the resulting specification has a unique infinite-volume Gibbs measure in terms of whether these perturbations satisfy an explicit linear inequality. Our analysis reveals an interesting nonmonotonicity with regards to biasing toward excluded nodes with no included neighbors.

1. Introduction. Recently, there has been a significant interest in combining ideas from probability, computer science, physics, statistics and operations research, to shed light on the structure and complexity of combinatorial optimization, counting and sampling problems (cf. [1, 10, 23]). Some of the most well-studied such problems involve the independent sets of a graph. Consider an undirected graph G, which consists of a set of nodes V and edges E, where each edge $e \in E$ is of the form (v_i, v_j) for some $v_i, v_j \in V$. Then the independent sets of G, $\mathcal{I}(G)$, are defined to be the subsets S of V with no internal edges, that is, a set $S \subseteq V$ is an independent set iff for all pairs of nodes $v_i, v_j \in S$, $(v_i, v_j) \notin E$. There are a wealth of results about the complexity and (in)approximability of counting, sampling and optimizing independent sets under various restrictions. We make no attempt to

survey that literature here, instead focusing only on the results most relevant to our own investigations, and refer the interested reader to [36] and the references therein for a recent overview.

1.1. *Infinite-volume Gibbs measures on the Cayley tree and the uniqueness regime.* As our main results will be stated in terms of measures on the Δ-regular infinite Cayley tree T_∞, we begin by briefly reviewing several concepts needed to formally describe such measures, following the exposition given in [13]. We assume the nodes of T_∞ are indexed by the nonnegative integers Z^+, and the tree is rooted at node 0. With a slight abuse of notation, we also let T_∞ denote the corresponding indexed set of nodes. In the spin systems considered in this paper, each node $i \in T_\infty$ is assigned a spin from the set $\{0, 1\}$. Let Ω denote the collection of all $\{0, 1\}$ spin assignments to the nodes of T_∞. For $\omega \in \Omega$ and $S \subseteq T_\infty$, let ω_S denote the restriction of ω to the nodes of S, and Ω_S denote the collection of all $\{0, 1\}$ spin assignments to the nodes of S. For an event A, let $I(A)$ denote the corresponding indicator. For $S \subseteq T_\infty$ and $\omega \in \Omega_S$, let $|\omega| \triangleq \sum_{i \in S} \omega_{\{i\}}$. Also, for a general set S, let $|S|$ denote the cardinality of S.

For every $S \subseteq T_\infty$, we define a potential $\Phi_S : \Omega_S \to \mathbb{R}$, mapping the spins of S to $\mathbb{R}$, where we use $\mathbb{R}$ to denote the positively extended real numbers, that is, including ∞. For all models considered, there will exist a finite radius R such that $\Phi_S = 0$ (i.e., is identically zero) for all S which contain any two nodes i, j at graph-theoretic distance strictly greater than R in T_∞. Let $\boldsymbol{\Phi}$ denote the collection of all potentials, that is, $\{\Phi_S, S \subseteq T_\infty\}$. As a notational convention, let us evaluate all empty summations to zero, and all empty products to unity. For $i, j \in T_\infty$, let $d(i, j)$ denote the graph-theoretic distance between i and j in T_∞. For $S \subseteq T_\infty$ and $i \in T_\infty$, let $d(i, S) \triangleq \inf_{j \in S} d(i, j)$, and $\partial S \triangleq \bigcup_{j \in T_\infty : d(j, S) \le 2R} \{j\} \setminus S$, that is, ∂S denotes the depth-$2R$ boundary surrounding S. For $d \ge 1$, let T_d denote the set of nodes with graph-theoretic distance at most d from 0 in T_∞. For two disjoint subsets $S_1, S_2 \subseteq T_\infty$, and configurations $\omega^1 \in \Omega_{S_1}, \omega^2 \in \Omega_{S_2}$, let $\omega^1 \cdot \omega^2$ denote the composition spin assignment which agrees with ω^1 on S_1 and ω^2 on S_2.

For every $\Lambda \subseteq T_\infty$, we define the Hamiltonian $\mathcal{H}_\Lambda^{\boldsymbol{\Phi}} : \Omega_{\Lambda \cup \partial \Lambda} \to \mathbb{R}$ as $\sum_{S \subseteq T_\infty : S \cap \Lambda \ne \varnothing} \Phi_S(\omega_S)$. A so-called infinite-volume Gibbs measure μ consistent with $\boldsymbol{\Phi}$ is a probability measure μ on Ω (associated with an appropriate probability space and filtration $\mathcal{F}$, see [13] for details), which satisfies certain consistency requirements associated with conditioning on a boundary. In particular, for finite $S \subseteq \Lambda \subseteq T_\infty$, $\omega \in \Omega_S$, and $\eta \in \Omega_{\partial \Lambda}$, let

$$(1) \qquad \mathbb{P}_{\boldsymbol{\Phi}, \Lambda}(S = \omega | \eta) \triangleq \frac{\sum_{v \in \Omega_\Lambda : v_S = \omega} \exp(-\mathcal{H}_\Lambda^{\boldsymbol{\Phi}}(v \cdot \eta))}{\sum_{v \in \Omega_\Lambda} \exp(-\mathcal{H}_\Lambda^{\boldsymbol{\Phi}}(v \cdot \eta))},$$

whenever this ratio is well-defined. For an event A on an appropriate filtration associated with the subset S, we analogously define $\mathbb{P}_{\boldsymbol{\Phi}, \Lambda}(A | \eta) = \sum_{\omega \in A} \mathbb{P}_{\boldsymbol{\Phi}, \Lambda}(S = \omega | \eta)$. For $S \subseteq T_\infty$ and $\omega \in \Omega_S$, we let $\{S = \omega\}$ be the event that the nodes of S

receive the spin-configuration dictated by ω. Then the aforementioned consistency requires that for any finite $S \subseteq \Lambda \subseteq T_\infty$ and $\omega \in \Omega_S$,

$$(2) \qquad \mu(S = \omega) = \sum_{\eta \in \partial \Lambda} \mathbb{P}_{\Phi, \Lambda}(S = \omega | \eta) \times \mu(\partial \Lambda = \eta).$$

Any measure μ satisfying (2), as well as certain other technical conditions (the details of which we omit, instead referring the reader to [13]), is said to be an infinite-volume Gibbs measure consistent with Φ, and we let $\mathcal{G}(\Phi)$ denote the collection of all such measures. As a notational convenience, we denote $\mathbb{P}_{\Phi, T_d}(S = \omega | \eta)$ by $\mathbb{P}_\Phi(S = \omega | \eta)$, where d is to be inferred from context (e.g., η belonging to $\Omega_{\partial T_d}$).

It is well known that under minimal technical conditions $\mathcal{G}(\Phi)$ is a nonempty convex set, where we denote the corresponding set of extreme measures as $\hat{\mathcal{G}}(\Phi)$. If $|\hat{\mathcal{G}}(\Phi)| = 1$, we say that Φ belongs to the *uniqueness regime*, that is, admits a unique infinite-volume Gibbs measure. Furthermore, every such extremal measure can be constructed as a so-called thermodynamic limit of appropriately conditioned finite spin systems, in the following sense. To each $\mu \in \hat{\mathcal{G}}(\Phi)$, we can associate $\omega^\mu \in \Omega$ such that for any finite $S \subseteq T_\infty$ and $\omega \in \Omega_S$,

$$(3) \qquad \mu(S = \omega) = \lim_{d \to \infty} \mathbb{P}_\Phi\big(S = \omega | \omega^\mu_{\partial T_d}\big).$$

In light of (3), nonuniqueness can also be interpreted as nonvanishing dependence on distant boundary conditions.

1.2. *Hardcore model on T_∞.* The hardcore model on T_∞ coincides with the following collection of potentials Φ. For some fixed activity $\lambda > 0$: $\Phi_{\{i\}}(\omega) = -\log(\lambda) I(\omega_{\{i\}} = 1)$ for all $i \in T_\infty$, $\Phi_{\{i,j\}}(\omega) = \infty I(|\omega_{\{i,j\}}| = 2)$ for all pairs of nodes (i, j) which are adjacent in T_∞ and Φ_S is identically zero for all other $S \subseteq T_\infty$. Under local conditioning, this measure puts all probability on spin assignments corresponding to independent sets, assigning an independent set S probability proportional to $\lambda^{|S|}$. When $\lambda = 1$, computing the relevant normalizing constant (i.e., partition function) is equivalent to counting the number of independent sets (a #P-Complete problem in general graphs [45]); as $\lambda \to \infty$, all the probability mass gets put on the largest independent sets, and computing the partition function is analogous to finding the cardinality of the maximum independent set (an NP-Complete problem in general graphs [15]). Such models have a rich history in the physics literature. Models on the infinite lattice were studied early on by several authors (cf. [11, 32, 38]). This work was extended to the three-regular infinite Cayley tree by Runnels in [31], and the general Δ-regular infinite Cayley tree in [39].

1.2.1. *Phase transition and nonuniqueness.* Motivated by the behavior of large particle systems, several of the original investigations of the hardcore model focused on identifying which sets of potentials (here parametrized by λ) belonged to the uniqueness regime. In particular, for each $\Delta \geq 3$, there exists a critical activity $\lambda_\Delta \overset{\Delta}{=} (\Delta - 1)^{\Delta-1}(\Delta - 2)^{-\Delta}$ such that the hardcore model on the infinite Δ-regular Cayley tree admits a unique infinite-volume Gibbs measure iff $\lambda \in (0, \lambda_\Delta]$ (cf. [18]). More recently, it has been shown that this same phase transition also corresponds to the point at which certain Markov chains for sampling from the independent sets of a graph of maximum degree Δ switch from mixing in polynomial time to mixing in exponential time (cf. [24]). Furthermore, it was shown in [45] that for all $\lambda \leq \lambda_\Delta$, the problem of computing $\sum_{S \in \mathcal{I}(G)} \lambda^{|S|}$ admits a Fully Polynomial Time Approximation Scheme (FPTAS) for all graphs of maximum degree Δ. Combined with the results of [36] (and the references therein), which show that no such FPTAS exists for $\lambda > \lambda_\Delta$ unless certain complexity classes collapse, this shows that the aforementioned phase transition has deep connections to computational complexity. This phase transition also has implications for various other applications, for example, the design of communication networks (cf. [18]).

1.3. *Higher-order M.r.f. for independent sets.* Many applications modeled by Gibbs measures defined on the independent sets of graphs involve more complicated dependencies and constraints on the independent sets themselves. This includes several models in physics, for example, models with next-nearest neighbor and/or competing interactions (cf. [42]), kinetically constrained spin models (cf. [20]), and geometrically constrained spin models (cf. [6]), and we refer the interested reader to the recent survey of [29] for many more such examples. Such measures also arise in combinatorial optimization, for example, through the study of subfamilies of independent sets such as those in which every excluded node is adjacent to some minimal number of included nodes (cf. [9, 10]). Closely related models have also arisen in the analysis of resource-constrained communication networks (cf. [17, 22, 28]). In such networking applications, two central questions are:

- What is the distribution of the number of included neighbors of an excluded node? (cf. [2, 5, 19]).
- Does the system exhibit long-range boundary independence? (cf. [18, 28]).

Combining the above, we are led to the following question.

QUESTION 1. *When sampling from the independent sets of the infinite Cayley tree, which distributions can be attained for the number of included neighbors of any given excluded node, while staying in the uniqueness regime?*

A good starting place is the hardcore model, for which the following result is well known. Let $B(n, p)$ denote a standard binomial distribution with parameters n and p.

OBSERVATION 1.   [37] *For the hardcore model on the infinite Cayley tree in the uniqueness regime, every excluded node has a number of included neighbors which follows a binomial distribution. Exactly which binomial distributions can be achieved in this way is dictated by the phase-transition at λ_Δ. In particular, it is possible to induce a $B(\Delta, p)$ distribution on the number of included neighbors of each excluded node for any $p \in (0, (\Delta - 1)^{-1}]$ in the uniqueness regime, and this characterization is tight.*

A natural framework for studying distributions on the independent sets of a graph with more complicated dependencies, reflected in many of the applications discussed above, is that of so-called higher-order M.r.f. (cf. [40]), equivalently spin systems in which the potentials Φ_S defining the Hamiltonian are nonzero for more complicated subsets of T_∞ (i.e., not just individual nodes and edges, which correspond to first-order M.r.f.). We note that such systems can also be analyzed as so-called factor (i.e., graphical) models with long-range interactions, and refer the reader to the excellent survey [44] for an overview.

1.3.1. *Second-order M.r.f. for independent sets.*   In this paper, we will consider so-called second-order M.r.f. for independent sets (cf. [40]), in which potentials are defined on depth-1 neighborhoods, that is, $R = 1$ (which should be assumed throughout). Here, we also assume that the potentials are translation and rotation-invariant. In particular, for $i \in T_\infty$, let $N(i)$ denote the set of neighbors of i in T_∞, as well as i itself, and $N_1(i) \triangleq N(i) \setminus \{i\}$. We will consider sets of potentials $\boldsymbol{\Phi}$ such that for some activity $\lambda > 0$ and strictly positive $(\Delta + 1)$-dimensional vector $\boldsymbol{\theta} = (\theta_0, \ldots, \theta_\Delta)$, and every $i \in T_\infty$,

$$(4) \qquad \boldsymbol{\Phi}_{N(i)}(\omega) = \begin{cases} -\log(\lambda), & \text{if } \omega_{\{i\}} = 1, |\omega_{N(i)}| = 1; \\ -\log(\theta_k), & \text{if } \omega_{\{i\}} = 0, |\omega_{N(i)}| = k; \\ \infty, & \text{otherwise}; \end{cases}$$

while $\boldsymbol{\Phi}_S$ is identically zero for all other $S \subseteq T_\infty$. Thus, in addition to the hardcore constraints and activity parameter λ, we assign a different potential $-\log(\theta_k)$ for each excluded node which is adjacent to exactly k included nodes. To express the dependence on λ and $\boldsymbol{\theta}$, we denote the corresponding set of potentials $\boldsymbol{\Phi}$ by the vector $(\lambda, \boldsymbol{\theta})$. For a given vector $\boldsymbol{\theta}$, let us say that $\boldsymbol{\theta}$ exhibits a phase transition if there exist strictly positive finite $\lambda_1 < \lambda_2$ such that $(\lambda_1, \boldsymbol{\theta})$ belongs to the uniqueness regime, while $(\lambda_2, \boldsymbol{\theta})$ does not belong to the uniqueness regime. For $\boldsymbol{\theta}$ exhibiting a phase transition, let us define the critical activity $\lambda_{\boldsymbol{\theta}} \triangleq \inf\{\lambda > 0 : (\lambda, \boldsymbol{\theta})$ does not belong to the uniqueness regime$\}$. We note that several of the examples mentioned earlier involving independent sets with more complicated dependency structure may be put in the framework of such second-order M.r.f.

The hardcore model may be viewed as a special case of our model, in which $\theta_k = 1$ for all k. More generally, it follows from a straightforward reduction that the case $\theta_k = \theta_0 \gamma^k$ (for some parameters $\theta_0, \gamma > 0$) also reduces to the hardcore model, albeit with activity $\lambda \theta_0^{-1} \gamma^\Delta$. Recall that a strictly positive sequence $\{x_i, i = 0, \ldots, n\}$ is called *log-convex* if $\frac{x_{i+1}}{x_i} \geq \frac{x_i}{x_{i-1}}$ for all $i \in \{1, \ldots, n-1\}$. If $\boldsymbol{\theta}$ is log-convex, it is natural to define vectors $\underline{\boldsymbol{\theta}}, \overline{\boldsymbol{\theta}}$ such that $\underline{\theta}_k = \theta_0 (\frac{\theta_1}{\theta_0})^k, \overline{\theta}_k = \theta_0 (\frac{\theta_\Delta}{\theta_{\Delta-1}})^k$, where log-convexity ensures that $\underline{\theta}_k \leq \theta_k \leq \overline{\theta}_k$ for all $k \in \{0, \ldots, \Delta\}$, and $\frac{\underline{\theta}_{k+1}}{\underline{\theta}_k} \leq \frac{\theta_{k+1}}{\theta_k} \leq \frac{\overline{\theta}_{k+1}}{\overline{\theta}_k}$ for all $k \in \{0, \ldots, \Delta - 1\}$. Note that $\underline{\theta}$ ($\overline{\theta}$) corresponds to the vector in which all ratios between consecutive entries are lowered (raised) to the lowest (highest) such ratio manifesting in $\boldsymbol{\theta}$. By the aforementioned reduction to the hardcore model, $\lambda_{\underline{\boldsymbol{\theta}}} = \lambda_\Delta \theta_0 (\frac{\theta_0}{\theta_1})^\Delta$; while $\lambda_{\overline{\boldsymbol{\theta}}} = \lambda_\Delta \theta_0 (\frac{\theta_{\Delta-1}}{\theta_\Delta})^\Delta$, where we note that log-convexity ensures $\lambda_{\overline{\boldsymbol{\theta}}} \leq \lambda_{\underline{\boldsymbol{\theta}}}$.

1.4. *FKG inequality.* A powerful tool for analyzing whether a given set of potentials belongs to the uniqueness regime are the so-called correlation inequalities, including the celebrated FKG theorem (cf. [12, 26]). Roughly, the FKG theorem proves that if a probability measure satisfies a certain supermodularity condition known as the FKG inequality, then that measure enjoys certain monotonicity properties. Although the FKG theorem holds in considerable generality, we will only state the inequality and its implications as customized to the specific models considered in this paper, following the exposition given in [16] for a different generalization of the hardcore model. Let us define a partial order $\overset{\sim}{\leq}$ on Ω (and appropriate restrictions) as follows. Let T_∞^e denote the subset of T_∞ consisting of the root 0, and all nodes whose graph-theoretic distance from 0 is even, and $T_\infty^o \overset{\Delta}{=} T_\infty \setminus T_\infty^e$. For $S \subseteq T_\infty$ and $\omega^1, \omega^2 \in \Omega_S$, let us say that $\omega^1 \overset{\sim}{\leq} \omega^2$ if $\omega_{\{i\}}^1 \leq \omega_{\{i\}}^2$ for all $i \in S \cap T_\infty^e$, and $\omega_{\{i\}}^1 \geq \omega_{\{i\}}^2$ for all $i \in S \cap T_\infty^o$. For $S \subseteq T_\infty$ and $\omega^1, \omega^2 \in \Omega_S$, let $\omega^1 \wedge \omega^2 \in \Omega_S$ denote the following spin configuration. $\omega^1 \wedge \omega_{\{i\}}^2 = \min(\omega_{\{i\}}^1, \omega_{\{i\}}^2)$ for $i \in S \cap T_\infty^e$; and $\omega^1 \wedge \omega_{\{i\}}^2 = \max(\omega_{\{i\}}^1, \omega_{\{i\}}^2)$ for $i \in S \cap T_\infty^o$. Similarly, let $\omega^1 \vee \omega^2 \in \Omega_S$ denote the following spin configuration. $\omega^1 \vee \omega_{\{i\}}^2 = \max(\omega_{\{i\}}^1, \omega_{\{i\}}^2)$ for $i \in S \cap T_\infty^e$; and $\omega^1 \vee \omega_{\{i\}}^2 = \min(\omega_{\{i\}}^1, \omega_{\{i\}}^2)$ for $i \in S \cap T_\infty^o$. Note that $\omega^1 \wedge \omega^2 \overset{\sim}{\leq} \omega^1, \omega^2 \overset{\sim}{\leq} \omega^1 \vee \omega^2$. Let $\tilde{\Omega}$ denote the subset of Ω consistent with the hardcore constraints, that is, $\omega \in \tilde{\Omega}$ if $|\omega_{\{i,j\}}| \leq 1$ whenever $d(i, j) = 1$, and define all projective notations (e.g., $\tilde{\Omega}_S$) in analogy with those for Ω. For an event A belonging to an appropriate filtration, let us say that A is *increasing* if $\omega^1 \in A$, $\omega^1 \overset{\sim}{\leq} \omega^2$ implies $\omega^2 \in A$. For example, if S is a finite subset of T_∞^e, then $\{|\omega_S| = |S|\}$, that is, the event that all spins in S are 1, is increasing. In that case, the conditions of the FKG inequality are as follows.

DEFINITION 1 ([12, 16] FKG inequality). The family of potentials $\boldsymbol{\Phi}$ satisfies the FKG inequality on T_∞ under partial order $\overset{\sim}{\leq}$ if for all $d \geq 0$, $\omega^1, \omega^2 \in \Omega_{T_d}$, and

$\eta \in \tilde{\Omega}_{\partial T_d}$,

$$
\begin{aligned}
(5) \quad & \mathbb{P}_{\boldsymbol{\Phi}}\big(T_d = \omega^1 \wedge \omega^2 | \eta\big) \times \mathbb{P}_{\boldsymbol{\Phi}}\big(T_d = \omega^1 \vee \omega^2 | \eta\big) \\
& \geq \mathbb{P}_{\boldsymbol{\Phi}}\big(T_d = \omega^1 | \eta\big) \times \mathbb{P}_{\boldsymbol{\Phi}}\big(T_d = \omega^2 | \eta\big).
\end{aligned}
$$

Then the celebrated FKG theorem is as follows.

THEOREM 1 ([12] FKG theorem). *If $\boldsymbol{\Phi}$ satisfies the FKG inequality on T_∞ under partial order $\tilde{\preceq}$, then for any $d \geq 0$, $\eta \in \tilde{\Omega}_{\partial T_d}$, and increasing events A, B belonging to the appropriate filtration,*

$$
\mathbb{P}_{\boldsymbol{\Phi}}\big(A \cap B | \eta\big) \geq \mathbb{P}_{\boldsymbol{\Phi}}(A|\eta) \times \mathbb{P}_{\boldsymbol{\Phi}}(B|\eta).
$$

It is well known that this monotonicity can be leveraged to reduce the question of uniqueness to the analysis of two special Gibbs measures (cf. [16, 27]). In particular, let $\omega^+ \in \Omega$ denote the spin configuration with $\omega^+_{\{i\}} = 1$ for all $i \in T^e_\infty$, and $\omega^+_{\{i\}} = 0$ for all $i \in T^o_\infty$; and $\omega^- \in \Omega$ denote the spin configuration with $\omega^-_{\{i\}} = 0$ for all $i \in T^e_\infty$, and $\omega^-_{\{i\}} = 1$ for all $i \in T^o_\infty$. Then the following well-known implications of the FKG inequality hold for the family of potentials $(\lambda, \boldsymbol{\theta})$, whenever those potentials indeed satisfy the FKG inequality. Many of these implications hold in considerably greater generality, and we refer the interested reader to [27] for a comprehensive discussion.

THEOREM 2 ([13, 16, 27] Further implications of the FKG inequality). *If the family of potentials $(\lambda, \boldsymbol{\theta})$ satisfies the FKG inequality on T_∞ under partial order $\tilde{\preceq}$, then all of the following implications hold*:

- *There is a unique (up to sets of measure 0) infinite-volume Gibbs measure $\mu^+_{\lambda, \boldsymbol{\theta}}$ such that for every finite $S \subseteq T_\infty$ and $\omega \in \Omega_S$,*

$$
\mu^+_{\lambda, \boldsymbol{\theta}}(S = \omega) = \lim_{d \to \infty} \mathbb{P}_{\lambda, \boldsymbol{\theta}}\big(S = \omega | \omega^+_{\partial T_d}\big),
$$

and a unique (up to sets of measure 0) infinite-volume Gibbs measure $\mu^-_{\lambda, \boldsymbol{\theta}}$ such that for every finite $S \subseteq T_\infty$ and $\omega \in \Omega_S$,

$$
\mu^-_{\lambda, \boldsymbol{\theta}}(S = \omega) = \lim_{d \to \infty} \mathbb{P}_{\lambda, \boldsymbol{\theta}}\big(S = \omega | \omega^-_{\partial T_d}\big);
$$

where all relevant limits appearing in the above definitions exist, and both of these measures are extremal, that is, belong to $\hat{\mathcal{G}}(\lambda, \boldsymbol{\theta})$.
- *For every increasing event A on an appropriate filtration and $\mu \in \mathcal{G}(\lambda, \boldsymbol{\theta})$,*

$$
\mu^-_{\lambda, \boldsymbol{\theta}}(A) \leq \mu(A) \leq \mu^+_{\lambda, \boldsymbol{\theta}}(A).
$$

- *$|\mathcal{G}(\lambda, \boldsymbol{\theta})| = 1$ iff $\mu^+_{\lambda, \boldsymbol{\theta}}(\omega_{\{0\}} = 1) = \mu^-_{\lambda, \boldsymbol{\theta}}(\omega_{\{0\}} = 1)$. Furthermore, if $|\mathcal{G}(\lambda, \boldsymbol{\theta})| = 1$, then the unique such infinite-volume Gibbs measure is translation and rotation-invariant.*

If $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime, we denote the corresponding unique infinite-volume Gibbs measure on T_∞ by $\mu_{\lambda, \boldsymbol{\theta}}^*$. Also, we let $p_+^{\lambda, \boldsymbol{\theta}} \triangleq \mu_{\lambda, \boldsymbol{\theta}}^*(\omega_{\{0\}} = 1)$, $p_k^{\lambda, \boldsymbol{\theta}} \triangleq \mu_{\lambda, \boldsymbol{\theta}}^*(\omega_{\{0\}} = 0, |\omega_{N(0)}| = k)$, and $\mathbf{p}^{\lambda, \boldsymbol{\theta}}$ the corresponding vector. Also, we let $\hat{p}_k^{\lambda, \boldsymbol{\theta}} \triangleq p_k^{\lambda, \boldsymbol{\theta}}(1 - p_+^{\lambda, \boldsymbol{\theta}})^{-1}$ denote the associated conditional distribution for the number of occupied neighbors of an unoccupied node, and $\hat{\mathbf{p}}^{\lambda, \boldsymbol{\theta}}$ the corresponding vector.

1.5. *Our contribution.*   In this paper, we take a step toward answering Question 1, by analyzing M.r.f. with second-order interactions on the independent sets of the infinite Cayley tree. We prove that the associated Gibbsian specification satisfies the FKG inequality whenever the local potentials defining the Hamiltonian satisfy a certain log-convexity condition. Under this condition, we give necessary and sufficient conditions for the existence of a unique infinite-volume Gibbs measure in terms of an explicit system of equations, prove the existence of a phase transition and give explicit lower and upper bounds on the associated critical activity, denoted $\underline{\lambda}_{\boldsymbol{\theta}}$ and $\overline{\lambda}_{\boldsymbol{\theta}}$, respectively, which we prove to exhibit a certain robustness. Interestingly, we find that $\underline{\lambda}_{\boldsymbol{\theta}}$ exhibits a dependence on $(\frac{\theta_{\Delta-1}}{\theta_\Delta})^\Delta$, like $\lambda_{\overline{\boldsymbol{\theta}}}$; while $\overline{\lambda}_{\boldsymbol{\theta}}$ exhibits a dependence on $(\frac{\theta_0}{\theta_1})^\Delta$, like $\lambda_{\underline{\boldsymbol{\theta}}}$. For potentials which are small perturbations of those coinciding to the hardcore model at its critical activity λ_Δ, we perform a perturbative analysis of the system of equations arising from our necessary and sufficient conditions for uniqueness, allowing us to explicitly characterize whether the resulting specification has a unique infinite-volume Gibbs measure in terms of whether these perturbations satisfy an explicit linear inequality. Our analysis reveals an interesting nonmonotonicity with regards to biasing toward excluded nodes with no included neighbors, which implies that the uniqueness regime for our model is incomparable to that suggested by $\lambda_{\overline{\boldsymbol{\theta}}}$ and $\lambda_{\underline{\boldsymbol{\theta}}}$.

1.6. *Outline of paper.*   The rest of the paper proceeds as follows. In Section 2, we make several additional definitions and state our main results. In Section 3, we prove that when $\boldsymbol{\theta}$ is log-convex, $(\lambda, \boldsymbol{\theta})$ satisfies the FKG inequality for all $\lambda > 0$. In Section 4, we rephrase the relevant probabilities and questions of interest in terms of sequences of ratios of partition functions, whose even and odd subsequences we prove to converge, and satisfy a certain system of equations. By proving that the functions arising in this system of equations satisfy certain bounds and monotonicities, we derive our necessary and sufficient conditions for uniqueness. In Section 5, we prove the existence of a phase transition, and provide explicit bounds on the critical activity. In Section 6, we perform a perturbative analysis of the system of equations arising from our necessary and sufficient conditions for uniqueness, allowing us to explicitly characterize whether the resulting specification has a unique infinite-volume Gibbs measure in terms of whether these perturbations satisfy an explicit linear inequality. In Section 7, we summarize our main

results, provide a broader discussion of the potential use of higher-order M.r.f. for analyzing independent sets in graphs and present directions for future research.

2. Main results.

2.1. *Potentials, probabilities and reverse ultra log-concave measures.* Before stating our main results, we formally relate the family of potentials $(\lambda, \boldsymbol{\theta})$ to the resulting occupancy probabilities $p_+^{\lambda,\boldsymbol{\theta}}$, $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ in the uniqueness regime, and review the definition of reverse ultra log-concave measures.

OBSERVATION 2. *If $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime, then the associated occupancy probabilities may be characterized as follows. There exist $c, x \in \mathbb{R}^+$ (depending only on λ and $\boldsymbol{\theta}$) such that $\hat{p}_k^{\lambda,\boldsymbol{\theta}} = c\theta_k \binom{\Delta}{k} x^k$ for $k \in \{0, \ldots, \Delta\}$. For certain natural choices of $\boldsymbol{\theta}$, $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ corresponds exactly to a well-known family of distributions. If $\boldsymbol{\theta} = \mathbf{1}$, then $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ corresponds to a binomial distribution. If $\theta_k = \frac{1}{k!\binom{\Delta}{k}}$ for $k \in \{0, \ldots, \Delta\}$, then $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ corresponds to a truncated Poisson distribution. If $\theta_k = \frac{1}{\binom{\Delta}{k}}$ for $k \in \{0, \ldots, \Delta\}$, then $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ corresponds to a truncated geometric distribution.*

Recall that a strictly positive sequence $\{x_i, i = 0, \ldots, n\}$ is called log-convex if $\frac{x_{i+1}}{x_i} \geq \frac{x_i}{x_{i-1}}$ for all $i \in \{1, \ldots, n-1\}$, reverse ultra log-concave if the sequence $\{\frac{x_i}{\binom{\Delta}{i}}, i = 0, \ldots, n\}$ is log-convex, and convex if for all $i \in \{1, \ldots, n-1\}$, either $x_{i+1} - x_i > x_i - x_{i-1}$ or $x_{i+1} = x_i = x_{i-1}$. We note that our definition of convex does not require nonnegativity, and is slightly nonstandard to preclude degeneracies when the relevant inequalities are not strict. We say that a measure μ with support on $\{0, \ldots, \Delta\}$ is reverse ultra log-concave if the sequence $\{\mu(k), k = 0, \ldots, \Delta\}$ is strictly positive and reverse ultra log-concave. Then the following may be easily verified using Observation 2, and we refer the interested reader to [7] for details and further references regarding reverse ultra log-concave measures.

OBSERVATION 3. *If $\boldsymbol{\theta}$ is log-convex and $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime, then $\hat{\mathbf{p}}^{\lambda,\boldsymbol{\theta}}$ is reverse ultra log-concave. Furthermore, the binomial, truncated Poisson and truncated geometric distributions considered in Observation 2 are all reverse ultra log-concave, with the corresponding choices of $\boldsymbol{\theta}$ log-convex.*

2.2. *Main results.* We now state our main results, and begin by formalizing the connection between log-convexity of the local potentials and the FKG inequality.

THEOREM 3 (Log-convexity of potentials implies the FKG inequality). *If $\boldsymbol{\theta}$ is log-convex, then for all $\lambda > 0$, $(\lambda, \boldsymbol{\theta})$ satisfies the FKG inequality on T_∞ under partial order $\tilde{\leq}$.*

For a given vector $\boldsymbol{\theta}$, let

$$f_{\boldsymbol{\theta}}(x) \triangleq \frac{\sum_{k=0}^{\Delta-1} \theta_{k+1}\binom{\Delta-1}{k}x^k}{\sum_{k=0}^{\Delta-1} \theta_k\binom{\Delta-1}{k}x^k},$$

and

$$g_{\boldsymbol{\theta}}(x) \triangleq \frac{1}{\sum_{k=0}^{\Delta-1} \theta_k\binom{\Delta-1}{k}x^k}.$$

Using Theorem 3, the implications of the FKG inequality dictated by Theorem 2, and a recursive analysis of certain relevant partition functions, we prove the following necessary and sufficient conditions for uniqueness.

THEOREM 4 (Necessary and sufficient conditions for uniqueness).   *If $\boldsymbol{\theta}$ is log-convex, the system of equations*

(6) $$x = \lambda g_{\boldsymbol{\theta}}(y) f_{\boldsymbol{\theta}}^{\Delta-1}(x);$$

(7) $$y = \lambda g_{\boldsymbol{\theta}}(x) f_{\boldsymbol{\theta}}^{\Delta-1}(y);$$

always has at least one nonnegative solution on $\mathbb{R}^+ \times \mathbb{R}^+$. $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime iff this solution is unique.

Note that Theorem 4 reduces to the well-known characterization for uniqueness in the hardcore model whenever $\theta_k = \theta_0 \gamma^k$ for some $\theta_0, \gamma > 0$, as in this special case $f_{\boldsymbol{\theta}} = \gamma$.

We now prove the existence of a phase transition for every log-convex $\boldsymbol{\theta}$, and provide explicit bounds on the associated critical activity. For log-convex $\boldsymbol{\theta}$, let

$$\psi_{\boldsymbol{\theta}} \triangleq \max_{k=0,\dots,\Delta-2}\left((\Delta-(k+1))\frac{\theta_{k+1}}{\theta_k}\right),$$

$$\underline{\lambda}_{\boldsymbol{\theta}} \triangleq \left(2\psi_{\boldsymbol{\theta}}\theta_0^{-1}\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}\right)^{\Delta-2}\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}+(\Delta-1)\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}-\frac{\theta_1}{\theta_0}\right)\right)\right)^{-1} \geq \underline{\underline{\lambda}}_{\boldsymbol{\theta}}$$

$$\triangleq \frac{\theta_0}{2\Delta^2}\left(\frac{\theta_{\Delta-1}}{\theta_\Delta}\right)^\Delta;$$

and

$$\overline{\lambda}_{\boldsymbol{\theta}} \triangleq \frac{3\theta_0}{\Delta}\exp\left(3\frac{\theta_\Delta}{\theta_{\Delta-1}}\frac{\theta_0}{\theta_1}\right)\left(\frac{\theta_0}{\theta_1}\right)^\Delta.$$

THEOREM 5 (Bounds on the critical activity).   *For any log-convex $\boldsymbol{\theta}$, $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime for all $\lambda < \underline{\lambda}_{\boldsymbol{\theta}}$, and does not does not belong to the uniqueness regime for all $\lambda > \overline{\lambda}_{\boldsymbol{\theta}}$.*

To gain further insight into $\underline{\lambda}_{\boldsymbol{\theta}}$ and $\overline{\lambda}_{\boldsymbol{\theta}}$, we briefly discuss the special case in which $\theta_k = \theta_0 \gamma^k$ for some $\theta_0, \gamma > 0$. In this case, $\boldsymbol{\theta} = \underline{\boldsymbol{\theta}} = \overline{\boldsymbol{\theta}}$, and $\lambda_{\boldsymbol{\theta}} = \lambda_\Delta \theta_0 \gamma^{-\Delta}$, which (with θ_0, γ held fixed) scales like $\theta_0 \frac{e}{\Delta} \gamma^{-\Delta}$ as $\Delta \to \infty$. Furthermore, it may be easily verified that in this case, $\underline{\lambda}_{\boldsymbol{\theta}} = \frac{\theta_0}{2(\Delta-1)} \gamma^{-\Delta}$, $\overline{\lambda}_{\boldsymbol{\theta}} = \frac{3e^3 \theta_0}{\Delta} \gamma^{-\Delta}$. In particular, as $\Delta \to \infty$, both our lower and upper bound scale (up to constant factors independent of θ_0, Δ, γ) like $\frac{\theta_0}{\Delta} \gamma^{-\Delta}$, agreeing with the true asymptotic scaling of the critical activity. When Δ is large but $\boldsymbol{\theta}$ does not have this simple factorized form, we find that $\underline{\lambda}_{\boldsymbol{\theta}}$ exhibits a dependence on $(\frac{\theta_{\Delta-1}}{\theta_\Delta})^\Delta$, like $\lambda_{\overline{\boldsymbol{\theta}}}$; while $\overline{\lambda}_{\boldsymbol{\theta}}$ exhibits a dependence on $(\frac{\theta_0}{\theta_1})^\Delta$, like $\lambda_{\underline{\boldsymbol{\theta}}}$. As log-convexity dictates that $\frac{\theta_{\Delta-1}}{\theta_\Delta} \le \frac{\theta_0}{\theta_1}$, this leads to a potentially exponentially large gap between our lower and upper bounds as one moves away from the special case in which $\theta_k = \theta_0 \gamma^k$. Determining whether the associated phase transition is sharp, and more generally closing the gap between our lower and upper bounds, remain interesting open questions. We also note that similar ideas (albeit connecting log-concavity of local potentials to certain negative association properties of the resulting Gibbs measure) were recently used in [33] to prove the nonexistence of a phase transition for so-called b-matchings on infinite graphs, and exploring further connections between our results and those of [33] remains a direction for future research.

We now comment briefly on several implications of Theorem 5, all of which follow from straightforward algebraic manipulations of $\underline{\lambda}_{\boldsymbol{\theta}}, \overline{\lambda}_{\boldsymbol{\theta}}$ and simple Taylor series expansions. We first show that our lower and upper bounds, and by implication the associated critical activity, exhibit a certain form of robustness.

OBSERVATION 4 (Robustness of bounds).   *If there exists $c \in [0, \Delta]$ such that* $\max(\frac{\theta_\Delta}{\theta_{\Delta-1}}, \frac{\theta_0}{\theta_1}) \le 1 + \frac{c}{\Delta}$, *then* $\underline{\lambda}_{\boldsymbol{\theta}} \ge (2 \exp(c)(1+4c))^{-1} \frac{\theta_0}{\Delta}$, *and* $\overline{\lambda}_{\boldsymbol{\theta}} \le 3 \exp(12 + c) \frac{\theta_0}{\Delta}$. *It follows that, up to constant factors independent of Δ, the critical activity will scale like $\frac{\theta_0}{\Delta}$ (as $\Delta \to \infty$) for any vector $\boldsymbol{\theta}$ which does not deviate too much from the all ones vector.*

We next use our results to bound the critical activity for $\boldsymbol{\theta}$ corresponding to the truncated Poisson distribution.

OBSERVATION 5.   *For $\boldsymbol{\theta}$ such that $\theta_k = \frac{1}{k!\binom{\Delta}{k}}$, one has that $\underline{\lambda}_{\boldsymbol{\theta}} \ge \frac{1}{2}\Delta^{-1}$. In particular, the critical activity is at least $\frac{1}{2}\Delta^{-1}$, as in the hardcore model.*

An explicit description of when equations (6)–(7) have a unique nonnegative solution, and which inclusion/exclusion probabilities can be attained in this way, seems difficult in general. However, we can develop a considerably more in-depth understanding when the relevant potentials are small perturbations of those coinciding to the hardcore model at the critical activity. Let us fix a vector $\mathbf{c} = (\mathbf{c}_0, \dots, \mathbf{c}_\Delta)$. It follows from a simple Taylor series expansion that convex perturbations of the all ones vector yield log-convex $\boldsymbol{\theta}$, as formalized below.

OBSERVATION 6.   *For every vector* $\mathbf{c}$, *there exists* $\varepsilon_{\mathbf{c}} > 0$ *such that for* $h \in (0, \varepsilon_{\mathbf{c}})$, $\mathbf{1} + \mathbf{c}h$ *is log-convex iff* $\mathbf{c}$ *is convex.*

We now define a convenient notion of uniqueness for perturbations around the all ones vector, which we will use in our analysis.

DEFINITION 2 [Direction of (non)uniqueness].   We say that $\mathbf{c}$ is a direction of uniqueness iff there exists $\varepsilon_{\mathbf{c}} > 0$ such that for all $h \in (0, \varepsilon_{\mathbf{c}})$, $(\lambda_{\Delta}, \mathbf{1} + \mathbf{c}h)$ belongs to the uniqueness regime; and a direction of nonuniqueness iff there exists $\varepsilon_{\mathbf{c}} > 0$ such that for all $h \in (0, \varepsilon_{\mathbf{c}})$, $(\lambda_{\Delta}, \mathbf{1} + \mathbf{c}h)$ does not belong to the uniqueness regime.

We now provide an explicit characterization/dichotomy theorem, classifying (almost) all convex vectors as either directions of uniqueness or directions of nonuniqueness. For $j \in \{0, \ldots, \Delta\}$, let $\Lambda_{\Delta, j} \triangleq \binom{\Delta}{j}(\Delta - 2)^{-j}$. Let $\boldsymbol{\pi} = (\pi_0, \ldots, \pi_{\Delta})$ denote the vector such that for $j \in \{0, \ldots, \Delta\}$,

$$\pi_j \triangleq \Lambda_{\Delta, j}\big((\Delta - 2) + (6 - 5\Delta)j + 2(\Delta - 1)j^2\big).$$

Then we prove the following.

THEOREM 6 (Perturbative necessary and sufficient conditions for uniqueness). *A convex vector* $\mathbf{c}$ *is a direction of uniqueness if* $\boldsymbol{\pi} \cdot \mathbf{c} < 0$, *and a direction of nonuniqueness if* $\boldsymbol{\pi} \cdot \mathbf{c} > 0$.

In particular, the hyperplane defined by $\boldsymbol{\pi} \cdot \mathbf{c} = 0$ represents a phase transition in the perturbation parameter space. We note that the question of what happens at the boundary (i.e., $\boldsymbol{\pi} \cdot \mathbf{c} = 0$) seems to require a finer asymptotic analysis, and we leave this as an open question.

We now study some qualitative features of $\boldsymbol{\pi}$, to shed light on the set of convex directions of uniqueness, and reveal an interesting nonmonotonicity of the uniqueness regime.

OBSERVATION 7.   *For all* $\Delta \geq 3$, $\boldsymbol{\pi}_0 > 0$, $\boldsymbol{\pi}_1 < 0$, $\boldsymbol{\pi}_2 < 0$, *and* $\boldsymbol{\pi}_k > 0$ *for all* $k \in \{3, \ldots, \Delta\}$.

That $\boldsymbol{\pi}_1 < 0$, $\boldsymbol{\pi}_2 < 0$, and $\boldsymbol{\pi}_k > 0$ for all $k \in \{3, \ldots, \Delta\}$ makes sense at an intuitive level, since biasing toward excluded nodes which are adjacent to few (many) included nodes should tend to reduce (increase) alternation and long-range correlations. That the cutoff occurs at exactly $k = 2$ can be further justified by noting that the average number of included neighbors of an excluded node in the hardcore model, at the critical activity λ_{Δ}, is $1 + (\Delta - 1)^{-1} \in (1, 2)$.

The counterintuitive feature of Observation 7, which seems to violate the above reasoning, is that $\pi_0 > 0$, that is, biasing toward excluded nodes with no included neighbors leads to nonuniqueness. We note that this effect is perhaps especially surprising in light of Theorem 5, as we now explain. Let $\mathbf{e}_0$ denote the $(\Delta + 1)$-dimensional vector whose first component is a 1, with all remaining components 0. As it is easily verified that $\mathbf{1} + \mathbf{e}_0 h$ is log-convex for all $h \geq 0$, and $\lim_{h \to \infty} \underline{\lambda}_{\mathbf{1}+\mathbf{e}_0 h} = \infty$, we conclude that the associated uniqueness regime exhibits the following nonmonotonicity.

COROLLARY 1 (Nonmonotonicity of uniqueness regime). *For all $\Delta \geq 3$, there exist strictly positive finite constants $a_\Delta < b_\Delta$ such that $(\lambda_\Delta, \mathbf{1} + \mathbf{e}_0 h)$ belongs to the uniqueness regime for $h = 0$ and $h \geq b_\Delta$, and does not belong to the uniqueness regime for $h \in (0, a_\Delta)$.*

Thus, biasing a small amount toward excluded nodes with no included neighbors leads to nonuniqueness, while biasing a large amount toward excluded nodes with no included neighbors leads to uniqueness. This nonmonotonicity also sheds light on the relationship between $\lambda_{\boldsymbol{\theta}}, \lambda_{\overline{\boldsymbol{\theta}}}$ and $\lambda_{\underline{\boldsymbol{\theta}}}$. In particular, it would be natural to conjecture that for general log-convex $\boldsymbol{\theta}$, the critical activity $\lambda_{\boldsymbol{\theta}}$ always belongs to the interval $[\lambda_{\overline{\boldsymbol{\theta}}}, \lambda_{\underline{\boldsymbol{\theta}}}]$, that is, that the critical activity is sandwiched between that for the model in which all ratios between consecutive entries of $\boldsymbol{\theta}$ are raised (lowered) to $\frac{\theta_\Delta}{\theta_{\Delta-1}} (\frac{\theta_1}{\theta_0})$. However, the aforementioned nonmonotonicity demonstrates that such a result cannot hold. Indeed, if such a bound were to hold, it would imply that for all $h > 0$,

$$\lambda_{\mathbf{1}+\mathbf{e}_0 h} \geq \lambda_{\overline{\mathbf{1}+\mathbf{e}_0 h}} = (1+h)\lambda_\Delta > \lambda_\Delta,$$

which Corollary 1 disproves. Furthermore, although it is easily verified that for log-convex $\boldsymbol{\theta}$ one has $\overline{\lambda}_{\boldsymbol{\theta}} \geq \lambda_{\underline{\boldsymbol{\theta}}}$, and $\underline{\lambda}_{\boldsymbol{\theta}} \leq \lambda_{\overline{\boldsymbol{\theta}}}$, in general $\underline{\lambda}_{\boldsymbol{\theta}}$ is incomparable to $\lambda_{\overline{\boldsymbol{\theta}}}$. For example, considering the case that $\theta_k = 1$ for $k \in \{0, \ldots, \Delta - 1\}$ and $\theta_\Delta = \Delta^2$, one can easily compute that

$$\underline{\lambda}_{\boldsymbol{\theta}} = \left(2(\Delta-1)\Delta^{2\Delta-4}(\Delta^2 + (\Delta-1)(\Delta^2-1))\right)^{-1} \geq \frac{1}{2\Delta^{2\Delta}},$$

which can be shown to be strictly greater than $\lambda_{\overline{\boldsymbol{\theta}}} = \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta} \frac{1}{\Delta^{2\Delta}}$ for all $\Delta \geq 8$. However, our previous example involving $\boldsymbol{\theta} = \mathbf{1} + \mathbf{e}_0 h$ demonstrates that the opposite inequality can hold as well. We note that several previous works in the literature on M.r.f. examine various notions of nonmonotonicity (cf. [22]), and better understanding the relevant (non)monotonicities with regards to higher-order M.r.f. remains an interesting open question.

3. Verification of FKG inequality.   In this section, we prove Theorem 3.

PROOF OF THEOREM 3.   It follows from the definition of the Hamiltonian, the symmetry of the potentials $\mathbf{\Phi} = (\lambda, \boldsymbol{\theta})$, and a straightforward algebraic manipulation (cf. [34]) that it suffices to demonstrate that for all $i \in T_\infty$ and $\omega^1, \omega^2 \in \Omega_{N(i)}$,

$$(8) \qquad \mathbf{\Phi}_{N(i)}(\omega^1 \wedge \omega^2) + \mathbf{\Phi}_{N(i)}(\omega^1 \vee \omega^2) \leq \mathbf{\Phi}_{N(i)}(\omega^1) + \mathbf{\Phi}_{N(i)}(\omega^2).$$

Let us fix such a node i. Let $S^{1,1} \triangleq \{j \in N_1(i) : \omega^1_{\{j\}} = 1\}$ (i.e., those neighbors of i with spin 1 in ω^1), $S^{1,0} \triangleq N_1(i) \setminus S^{1,1}$; $S^{2,1} \triangleq \{j \in N_1(i) : \omega^2_{\{j\}} = 1\}$; and $S^{2,0} \triangleq N_1(i) \setminus S^{2,1}$.

We proceed by a case analysis. First, suppose that for some $l \in \{1, 2\}$, $\mathbf{\Phi}_{N(i)}(\omega^l) = \infty$. In this case, (8) holds trivially. Thus, we subsequently suppose this situation is precluded.

We now treat the case $i \in T^e_\infty$, equivalently $N_1(i) \subseteq T^o_\infty$. Note that from definitions, for $i \in T^e_\infty$, $\omega^1 \wedge \omega^2_{\{i\}} = 1$ iff both $\omega^1_{\{i\}} = 1$ and $\omega^2_{\{i\}} = 1$; $\omega^1 \vee \omega^2_{\{i\}} = 1$ iff either $\omega^1_{\{i\}} = 1$ or $\omega^2_{\{i\}} = 1$; $|\omega^1 \wedge \omega^2_{N_1(i)}| = |S^{1,1} \cup S^{2,1}|$; and $|\omega^1 \vee \omega^2_{N_1(i)}| = |S^{1,1} \cap S^{2,1}|$.

First, suppose $\omega^1_{\{i\}} = \omega^2_{\{i\}} = 1$. In this case, $|S^{1,1}| = |S^{2,1}| = 0$, and thus $|\omega^1 \wedge \omega^2_{N_1(i)}| = |\omega^1 \vee \omega^2_{N_1(i)}| = 0$. We conclude that both the left-hand side (LHS) and right hand side (RHS) of (8) equal $-2\log(\lambda)$, and (8) holds.

Next, suppose $\omega^1_{\{i\}} = 1, \omega^2_{\{i\}} = 0$. In this case, $\omega^1 \wedge \omega^2_{\{i\}} = 0$, $|\omega^1 \wedge \omega^2_{N_1(i)}| = |S^{2,1}|$, $\omega^1 \vee \omega^2_{\{i\}} = 1$, $|\omega^1 \vee \omega^2_{N_1(i)}| = 0$. We conclude that both the LHS and RHS of (8) equal $-\log(\theta_{|S^{2,1}|}) - \log(\lambda)$, and (8) holds. The case $\omega^1_{\{i\}} = 0, \omega^2_{\{i\}} = 1$ follows from a symmetric and identical argument.

Finally, suppose $\omega^1_{\{i\}} = 0, \omega^2_{\{i\}} = 0$. In this case, $\omega^1 \wedge \omega^2_{\{i\}} = 0$, $|\omega^1 \wedge \omega^2_{N_1(i)}| = |S^{1,1} \cup S^{2,1}|$, $\omega^1 \vee \omega^2_{\{i\}} = 0$, $|\omega^1 \vee \omega^2_{N_1(i)}| = |S^{1,1} \cap S^{2,1}|$. We conclude that (8) would hold if it were true that

$$(9) \qquad \log(\theta_{|S^{1,1} \cup S^{2,1}|}) + \log(\theta_{|S^{1,1} \cap S^{2,1}|}) \geq \log(\theta_{|S^{1,1}|}) + \log(\theta_{|S^{2,1}|}).$$

However, (9) follows immediately from the log-convexity of $\boldsymbol{\theta}$, and the well-known connection between convexity and supermodularity (cf. [41]), completing the proof for the case $i \in T^e_\infty$. As the case $i \in T^o_\infty$ follows from a nearly identical argument with the role of $\omega^1 \vee \omega^2$ and $\omega^1 \wedge \omega^2$ reversed, we omit the details. Combining the above completes the proof. $\square$

4. Probabilities, partition functions and proof of Theorem 4.   In this section, we first rephrase the relevant probabilities and questions of interest in terms of sequences of ratios of partition functions, whose even and odd subsequences we prove to converge, and which are amenable to a recursive analysis. We then

prove that the functions arising in the relevant recursions satisfy certain bounds and monotonicities, which we exploit to prove Theorem 4. Without loss of generality, let us assign the neighbors of 0 in T_∞ indices $1, \ldots, \Delta$ in an arbitrary but fixed manner. For $d \geq 2$, let T_d^1 denote the subtree of T_d rooted at node 1, excluding node 1 itself, that is, the collection of nodes $j \in T_d \setminus \{1\}$ such that every undirected path in T_d from j to 0 contains node 1. For $d \geq 2$ and $i, j \in \{0, 1\}$ such that $i + j \leq 1$, let $\eta^{i,j,d} \in \Omega_{\partial T_d^1}$ denote that boundary condition such that $\eta^{i,j,d}_{\{0\}} = i$, $\eta^{i,j,d}_{\{1\}} = j$, $\eta^{i,j,d}_{\{k\}} = 0$ for all $k \in \partial T_d^1$ such that $d(k, 0) = d + 1$, and $\eta^{i,j,d}_{\{k\}} = 1$ for all $k \in \partial T_d^1$ such that $d(k, 0) = d + 2$. Similarly, for $d \geq 2$, let $Z_{\lambda,\boldsymbol{\theta},d}(i, j) \overset{\Delta}{=} \sum_{v \in \Omega_{T_d^1}} \exp(-\mathcal{H}^{\lambda,\boldsymbol{\theta}}_{T_d^1}(v \cdot \eta^{i,j,d}))$. For the special case $d = 1$, as $T_d^1 = \varnothing$, we define $Z_{\lambda,\boldsymbol{\theta},1}(0, 0) = \theta_0 \theta_{\Delta-1}^{\Delta-1}$, $Z_{\lambda,\boldsymbol{\theta},1}(1, 0) = \theta_1 \theta_{\Delta-1}^{\Delta-1}$, $Z_{\lambda,\boldsymbol{\theta},1}(0, 1) = \lambda \theta_\Delta^{\Delta-1}$. For $d \geq 1$ and $i \in \{0, 1\}$, let $Z_{\lambda,\boldsymbol{\theta},d}(i) \overset{\Delta}{=} \frac{Z_{\lambda,\boldsymbol{\theta},d}(i,0)}{Z_{\lambda,\boldsymbol{\theta},d}(0,1)}$, and $\zeta_{\lambda,\boldsymbol{\theta},d} \overset{\Delta}{=} Z^{-1}_{\lambda,\boldsymbol{\theta},d}(0)$, where we also define $\zeta_{\lambda,\boldsymbol{\theta},0} \overset{\Delta}{=} 0$. When there is no ambiguity, we will suppress the notation on $(\lambda, \boldsymbol{\theta})$, simply writing, for example, $Z_d(i, j)$, $Z_d(i)$, ζ_d, f, g. For $d \geq 1$, let $\eta^{-,+,d} \in \Omega_{\partial T_d}$ denote that boundary condition such that $\eta^{-,+,d}_{\{k\}} = 0$ for all $k \in \partial T_d$ such that $d(k, 0) = d + 1$, and $\eta^{-,+,d}_{\{k\}} = 1$ for all $k \in \partial T_d$ such that $d(k, 0) = d + 2$. Note that for even d, $\eta^{-,+,d} = \omega^+_{\partial T_d}$; while for odd d, $\eta^{-,+,d} = \omega^-_{\partial T_d}$. Then for $d \geq 2$ and $k \in \{0, \ldots, \Delta\}$,

$$\mathbb{P}_{\lambda,\boldsymbol{\theta}}\big(\omega_0 = 0, |\omega_{N(0)}| = k | \eta^{-,+,d}\big)$$

(10)
$$= \frac{\binom{\Delta}{k}\theta_k Z^k_d(0, 1) Z^{\Delta-k}_d(0, 0)}{\lambda Z^\Delta_d(1, 0) + \sum_{i=0}^{\Delta} \binom{\Delta}{i}\theta_i Z^i_d(0, 1) Z^{\Delta-i}_d(0, 0)}$$

$$= \frac{\binom{\Delta}{k}\theta_k \zeta^k_d}{\lambda(Z_d(1)\zeta_d)^\Delta + \sum_{i=0}^{\Delta} \binom{\Delta}{i}\theta_i \zeta^i_d,} \overset{\Delta}{=} p^{\lambda,\boldsymbol{\theta},d}_k.$$

We also let $\mathbf{p}^{\lambda,\boldsymbol{\theta},d}$ denote the associated vector, and $p^{\lambda,\boldsymbol{\theta},d}_+ \overset{\Delta}{=} 1 - \mathbf{p}^{\lambda,\boldsymbol{\theta},d} \cdot \mathbf{1}$.

We now derive several recursions for $Z_d(i)$ and ζ_d, to aid in our analysis.

LEMMA 1. *For all $d \geq 1$,*

(11)
$$Z_d(1) = \zeta_d^{-1} f(\zeta_{d-1});$$

and for all $d \geq 2$,

(12)
$$\zeta_d = \lambda g(\zeta_{d-1}) f^{\Delta-1}(\zeta_{d-2}).$$

PROOF. We first treat the cases $d = 1, 2$. That (11) holds for $d = 1$ follows from definitions and the fact that $f(0) = \frac{\theta_1}{\theta_0}$. For $d = 2$, a straightforward calcula-

tion demonstrates that

$$Z_2(0,0) = (\theta_{\Delta-1}^{\Delta-1}\theta_0)^{\Delta-1} \sum_{k=0}^{\Delta-1} \theta_k \binom{\Delta-1}{k} \zeta_1^k,$$

$$Z_2(1,0) = (\theta_{\Delta-1}^{\Delta-1}\theta_0)^{\Delta-1} \sum_{k=0}^{\Delta-1} \theta_{k+1} \binom{\Delta-1}{k} \zeta_1^k,$$

$$Z_2(0,1) = \lambda\theta_1^{\Delta-1}\theta_{\Delta-1}^{(\Delta-1)^2},$$

$$Z_2(1) = \lambda^{-1}\left(\frac{\theta_0}{\theta_1}\right)^{\Delta-1} \sum_{k=0}^{\Delta-1} \theta_{k+1} \binom{\Delta-1}{k} \zeta_1^k,$$

$$Z_2(0) = \lambda^{-1}\left(\frac{\theta_0}{\theta_1}\right)^{\Delta-1} \sum_{k=0}^{\Delta-1} \theta_k \binom{\Delta-1}{k} \zeta_1^k, \qquad \zeta_2 = \lambda f^{\Delta-1}(0)g(\zeta_1);$$

from which (11) and (12) follow.

For $d \geq 3$ and $i \in \{0,1\}$,

$$Z_d(i,0) = \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k+i} Z_{d-1}^k(0,1) Z_{d-1}^{\Delta-1-k}(0,0),$$

and

$$Z_d(0,1) = \lambda Z_{d-1}^{\Delta-1}(1,0).$$

Thus,

$$Z_d(0) = \frac{\sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k}\theta_k Z_{d-1}^k(0,1) Z_{d-1}^{\Delta-1-k}(0,0)}{\lambda Z_{d-1}^{\Delta-1}(1,0)}$$

$$= \lambda^{-1}\left(\frac{Z_{d-1}(0)}{Z_{d-1}(1)}\right)^{\Delta-1} \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_k \zeta_{d-1}^k.$$

Similarly,

$$Z_d(1) = \lambda^{-1}\left(\frac{Z_{d-1}(0)}{Z_{d-1}(1)}\right)^{\Delta-1} \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k+1} \zeta_{d-1}^k.$$

Combining with the definition of f and g completes the proof. □

We next establish some useful properties of f and g.

LEMMA 2. *If θ is log-convex, then for all $x \geq 0$: $\partial_x f(x) \geq 0$, $\partial_x g(x) \leq 0$, $0 \leq g(x) \leq \theta_0^{-1}$, and $\frac{\theta_1}{\theta_0} \leq f(x) \leq \frac{\theta_\Delta}{\theta_{\Delta-1}}$.*

PROOF.   We first prove that $\partial_x f(x) \geq 0$ for all $x \geq 0$. It follows from a straightforward calculation that

$$(13) \quad \partial_x f(x) = (\Delta - 1) \frac{\sum_{i=0}^{\Delta-1} \sum_{j=0}^{\Delta-2} \binom{\Delta-1}{i}\binom{\Delta-2}{j} x^{i+j} (\theta_i \theta_{j+2} - \theta_{i+1}\theta_{j+1})}{(\sum_{k=0}^{\Delta-1} \theta_k \binom{\Delta-1}{k} x^k)^2},$$

and $\sum_{i=0}^{\Delta-1} \sum_{j=0}^{\Delta-2} \binom{\Delta-1}{i}\binom{\Delta-2}{j} x^{i+j} (\theta_i \theta_{j+2} - \theta_{i+1}\theta_{j+1})$ equals

$$(14) \quad \sum_{k=0}^{2\Delta-3} x^k \sum_{i=\max(0,k-(\Delta-2))}^{\min(\Delta-1,k)} \binom{\Delta-1}{i}\binom{\Delta-2}{k-i} (\theta_i \theta_{k-i+2} - \theta_{i+1}\theta_{k-i+1}).$$

We now demonstrate that

$$(15) \quad \sum_{i=\max(0,k-(\Delta-2))}^{\min(\Delta-1,k)} \binom{\Delta-1}{i}\binom{\Delta-2}{k-i} (\theta_i \theta_{k-i+2} - \theta_{i+1}\theta_{k-i+1})$$

is nonnegative for all $k \in [0, 2\Delta - 3]$, completing the proof. We proceed by "pairing up" certain terms appearing in (15), by using the fact that for all $k \in [0, 2\Delta - 3]$ and $i \in [\max(0, k - (\Delta - 2)), \min(\Delta - 1, k)]$,

$$(16) \quad \begin{aligned} &\theta_i \theta_{k-i+2} - \theta_{i+1}\theta_{k-i+1} \\ &\quad = -(\theta_{(k-i+1)}\theta_{k-(k-i+1)+2} - \theta_{(k-i+1)+1}\theta_{k-(k-i+1)+1}). \end{aligned}$$

Let us say that a function f, with domain and range containing the finite set $S \subseteq Z$, is a paired bijection on S if the restriction of f to domain S is a bijection (i.e., every element of S is mapped to some element of S, and every element of S is mapped to by some element of S), and f does not map any element of S to itself. We now demonstrate that for all $k \in [0, 2\Delta - 3]$, the mapping $f(i) = k - i + 1$ is a paired bijection on $S = [\max(0, k - (\Delta - 2)), \min(\Delta - 1, k)] \setminus \{0, \frac{k+1}{2}\}$. We first show $i \in S$ implies $f(i) \in S$. Indeed, $i \geq 0, i \neq 0$ implies $i \geq 1$, and thus $f(i) \leq k$; $i \geq k - (\Delta - 2)$ implies $f(i) \leq \Delta - 1$; $i \leq \Delta - 1$ implies $f(i) \geq k - (\Delta - 2)$; $i \leq k$ implies $f(i) \geq 1$. Furthermore, as $\frac{k+1}{2}$ is the unique (possibly noninteger) solution to $f(x) = x$ (i.e., fixed point), $i \neq \frac{k+1}{2}$ implies $f(i) \neq \frac{k+1}{2}$. Combining the above completes the demonstration that $i \in S$ implies $f(i) \in S$. The proof that f is a paired bijection on S then follows from the fact that f is strictly decreasing and invertible, with unique fixed point $\frac{k+1}{2}$. Furthermore, since f is strictly decreasing and $\lfloor \frac{k}{2} \rfloor < \frac{k+1}{2} \leq \lfloor \frac{k}{2} \rfloor + 1$, it is also true that for $i \in S$, $i \leq \lfloor \frac{k}{2} \rfloor$ iff $f(i) \geq \lfloor \frac{k}{2} \rfloor + 1$. Combining the above with (16) and the fact that $\lfloor \frac{k}{2} \rfloor \leq \min(\Delta - 1, k)$ for all $k \in [0, 2\Delta - 3]$, we conclude that (15) equals

$$(17) \quad \begin{aligned} \sum_{i=\max(1,k-(\Delta-2))}^{\lfloor k/2 \rfloor} &\left(\binom{\Delta-1}{i}\binom{\Delta-2}{k-i} - \binom{\Delta-1}{k-i+1}\binom{\Delta-2}{i-1} \right) \\ &\times (\theta_i \theta_{k-i+2} - \theta_{i+1}\theta_{k-i+1}) \end{aligned}$$

(18)
$$+ I(k \leq \Delta - 2) \binom{\Delta - 2}{k} (\theta_0 \theta_{k+2} - \theta_1 \theta_{k+1})$$

$$+ I\left(\frac{k+1}{2} \in Z^+\right) \binom{\Delta - 1}{\frac{k+1}{2}} \binom{\Delta - 2}{\frac{k-1}{2}}$$

(19)
$$\times (\theta_{(k+1)/2} \theta_{(k+1)/2+1} - \theta_{(k+1)/2+1} \theta_{(k+1)/2}).$$

We now verify that (17)–(19) are nonnegative, and begin with (17). It follows from a straightforward calculation that $\binom{\Delta-1}{i}\binom{\Delta-2}{k-i} - \binom{\Delta-1}{k-i+1}\binom{\Delta-2}{i-1}$ will be the same sign as $\frac{k+1}{i} - 2$, and thus nonnegative for $i \leq \lfloor \frac{k}{2} \rfloor$. Also, the log-convexity of $\boldsymbol{\theta}$ implies that $\theta_i \theta_{k-i+2} - \theta_{i+1} \theta_{k-i+1}$ will be nonnegative if $k - i + 1 \geq i$, which holds for $i \leq \lfloor \frac{k}{2} \rfloor$. Combining the above demonstrates the nonnegativity of (17). The log-convexity of $\boldsymbol{\theta}$ similarly implies the nonnegativity of $\theta_0 \theta_{k+2} - \theta_1 \theta_{k+1}$, and thus also of (18). As (19) is identically zero, combining the above completes the proof. That $\frac{\theta_1}{\theta_0} \leq f(x) \leq \frac{\theta_\Delta}{\theta_{\Delta-1}}$ then follows by letting $x \downarrow 0$ and $x \uparrow \infty$. Noting that the associated monotonicity and bounds for g are straightforward completes the proof of the lemma. $\square$

We now combine Lemmas 1 and 2 to prove that the even and odd subsequences of $\{\zeta_d, d \geq 1\}$ are monotone, and thus converge, where we will later prove that $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime iff these limits coincide. We note that although the monotonicity of certain related sequences follows directly from the FKG theorem and its implied monotonicities, here our analysis fundamentally involves $Z_d(0, 0)$, in which nodes at both even and odd parity have their spins set to 0, which seems to preclude such a direct approach. Instead, we proceed by induction, using the properties of f and g demonstrated in Lemma 2.

LEMMA 3. *If $\boldsymbol{\theta}$ is log-convex, then $\{\zeta_{2d}, d \geq 0\}$ is monotone increasing, and $\{\zeta_{2d+1}, d \geq 0\}$ is monotone decreasing.*

PROOF.   We proceed by induction simultaneously on both sequences. The base case entails demonstrating that $\zeta_3 \leq \zeta_1$, and $\zeta_2 \geq \zeta_0$. It follows from definitions that $\zeta_1 = \lambda \theta_0^{-1} (\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}$. Combining with the bounds for f and g of Lemma 2, and the fact that ζ_3 satisfies (12), demonstrates that $\zeta_3 \leq \zeta_1$. That $\zeta_2 \geq \zeta_0$ follows from nonnegativity, completing the proof of the base case.

Now, suppose that $\{\zeta_{2k}, k = 0, \ldots, d - 1\}$ is monotone increasing, and $\{\zeta_{2k+1}, k = 0, \ldots, d - 1\}$ is monotone decreasing for some $d \geq 2$. Then it follows from Lemma 1, the monotonicity of f and g guaranteed by Lemma 2, and the induction hypothesis that

$$\zeta_{2d} = \lambda g(\zeta_{2(d-1)+1}) f^{\Delta-1}(\zeta_{2(d-1)})$$

$$\geq \lambda g(\zeta_{2(d-2)+1}) f^{\Delta-1}(\zeta_{2(d-2)}) = \zeta_{2(d-1)}.$$

Similarly, the above further implies that

$$\zeta_{2d+1} = \lambda g(\zeta_{2d}) f^{\Delta-1}(\zeta_{2(d-1)+1})$$
$$\leq \lambda g(\zeta_{2(d-1)}) f^{\Delta-1}(\zeta_{2(d-2)+1}) = \zeta_{2(d-1)+1}.$$

Combining the above completes the proof. □

It follows from Lemma 3 that $\overline{\zeta}_{\lambda,\boldsymbol{\theta},\infty} \triangleq \lim_{d\to\infty} \zeta_{\lambda,\boldsymbol{\theta},2d+1}$ and $\underline{\zeta}_{\lambda,\boldsymbol{\theta},\infty} \triangleq$ $\lim_{d\to\infty} \zeta_{\lambda,\boldsymbol{\theta},2d}$ both exist. Furthermore, the continuity of f and g on $\mathbb{R}^+$, combined with (12) and Lemma 3, implies the following.

OBSERVATION 8. $0 < \underline{\zeta}_\infty, \overline{\zeta}_\infty < \infty$, and $(\underline{\zeta}_\infty, \overline{\zeta}_\infty)$ is a solution to the system of equations (6)–(7).

With Observation 8 in hand, we now complete the proof of Theorem 4, as well as Observation 2. Let $L_{\boldsymbol{\theta}}(z) \triangleq \sum_{i=0}^{\Delta} \theta_i \binom{\Delta}{i} z^i$.

PROOF OF THEOREM 4.   We first prove that the system of equations (6)–(7) always has at least one solution (x^*, y^*) on $\mathbb{R}^+ \times \mathbb{R}^+$ for which $x^* = y^*$. Let $\eta(x) \triangleq x - \lambda g(x) f^{\Delta-1}(x)$. Note that $\eta(0) = -\lambda \frac{\theta_1^{\Delta-1}}{\theta_0^\Delta} < 0$. It follows from Lemma 2 that $\eta(\lambda \theta_0^{-1} (\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}) \geq 0$. As η is continuous on $[0, \infty)$, we conclude that there exists $z^* \in \mathbb{R}^+$ such that $\eta(z^*) = 0$, which implies that (z^*, z^*) is a solution to the system of equations.

We now prove that if the system of equations (6)–(7) has a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$, then $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime. Suppose (6)–(7) has a unique solution (x^*, y^*) on $\mathbb{R}^+ \times \mathbb{R}^+$. Then it must be that any nonnegative solution (x, y) to the system of equations satisfies $x = x^* = y^* = y$. By Observation 8, $(\underline{\zeta}_\infty, \overline{\zeta}_\infty)$ is such a solution. Thus, $\underline{\zeta}_\infty = \overline{\zeta}_\infty$, in which case it follows from Theorems 3 and 2, (10), and Lemma 1 that $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime.

We now prove that if the system of equations (6)–(7) does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$, then $(\lambda, \boldsymbol{\theta})$ does not belong to the uniqueness regime. Indeed, suppose that the system of equations does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$. Let S denote the set of all 2-vectors (x, y) such that $0 \leq x \leq y < \infty$, and (x, y) is a solution to the system of equations. Let $\overline{y} \triangleq \sup_{\mathbf{z} \in S} z_2$, that is, the largest number appearing in any solution pair. We first show that $\overline{y}$ is itself part of some solution pair (i.e., it is not just approached). If $|S| < \infty$, this is immediate. If not, consider any sequence of solution vectors $\{\mathbf{z}^i, i \geq 1\}$ such that $\lim_{i\to\infty} z_2^i = \overline{y}$. Since $\{z_1^i, i \geq 1\}$ is uniformly bounded by Lemma 2, the Bolzano–Weierstrass theorem implies that $\{\mathbf{z}^i, i \geq 1\}$ will itself have a convergent subsequence $\{\mathbf{z}^{i_k}, k \geq 1\}$, and let us denote $\lim_{k\to\infty} z_1^{i_k}$ by $\overline{x}$. That $(\overline{x}, \overline{y})$ satisfies the system of equations then

follows from the continuity of f and g. Similarly, let $\underline{x} \triangleq \inf_{\mathbf{z} \in S} z_1$, that is, the smallest number appearing in any solution pair, and $\underline{y}$ the other number appearing in the corresponding solution pair (whose existence is guaranteed by the same argument used above). Note that $\underline{x} < \overline{y}$.

We now prove (by induction) that in this case, $\{\zeta_d, d \geq 0\}$ has a nonvanishing parity-dependence, with even values lying below $\underline{x}$, and odd values lying above $\overline{y}$. We begin with the base cases $d = 0, 1$. The $d = 0$ case follows from nonnegativity. The $d = 1$ case follows from the fact that $\overline{y}$ satisfies (7), combined with Lemma 2 and the definition of ζ_1. Now, proceeding by induction, suppose that for some $d \geq 1$ and all $k \in \{0, \ldots, d-1\}$, $\zeta_{2k} \leq \underline{x}$, and $\zeta_{2k+1} \geq \overline{y}$. Then it follows from Lemma 1, and the monotonicity of f and g implied by Lemma 2, that

$$
\begin{aligned}
\zeta_{2d} &= \lambda g(\zeta_{2d-1}) f^{\Delta-1}(\zeta_{2d-2}) \\
&\leq \lambda g(\overline{y}) f^{\Delta-1}(\underline{x}) \\
&\leq \lambda g(\underline{y}) f^{\Delta-1}(\underline{x}) = \underline{x}, \qquad \text{since } \overline{y} \geq \underline{y},
\end{aligned}
$$

and

$$
\begin{aligned}
\zeta_{2d+1} &= \lambda g(\zeta_{2d}) f^{\Delta-1}(\zeta_{2d-1}) \\
&\geq \lambda g(\underline{x}) f^{\Delta-1}(\overline{y}) \\
&\geq \lambda g(\overline{x}) f^{\Delta-1}(\overline{y}) = \overline{y}, \qquad \text{since } \underline{x} \leq \overline{x},
\end{aligned}
$$

completing the proof.

Finally, we prove that the aforementioned parity dependence of $\{\zeta_d, d \geq 0\}$ implies a nonvanishing parity dependence on the probability that the root is included when conditioning on the appropriate extremal boundary conditions, implying nonuniqueness. It follows from the parity dependence of $\{\zeta_d, d \geq 0\}$, (10), and Lemma 2 that for all $d \geq 1$,

$$
\mathbb{P}_{\lambda,\boldsymbol{\theta}}\big(\omega_0 = 1 | \eta^{-,+,2d}\big)
$$

$$
= \left(1 + \frac{L(\zeta_{2d})}{\lambda f^{\Delta}(\zeta_{2d-1})}\right)^{-1} \geq \left(1 + \frac{L(\underline{x})}{\lambda f^{\Delta}(\overline{y})}\right)^{-1}; \mathbb{P}_{\lambda,\boldsymbol{\theta}}\big(\omega_0 = 1 | \eta^{-,+,2d+1}\big)
$$

$$
= \left(1 + \frac{L(\zeta_{2d+1})}{\lambda f^{\Delta}(\zeta_{2d})}\right)^{-1} \leq \left(1 + \frac{L(\overline{y})}{\lambda f^{\Delta}(\underline{x})}\right)^{-1} < \left(1 + \frac{L(\underline{x})}{\lambda f^{\Delta}(\overline{y})}\right)^{-1}.
$$

Combining with Theorems 3 and 2, along with the fact that $\eta^{-,+,d}$ equals $\omega^+_{\partial T_d}$ for even d, and equals $\omega^-_{\partial T_d}$ for odd d, completes the proof. As it follows that $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime iff $\lim_{d \to \infty} \zeta_d$ exists, combining with (10) also completes the proof of Observation 2. $\square$

5. Existence of phase transition and proof of Theorem 5. In this section, we show the existence of a phase transition for log-convex $\boldsymbol{\theta}$, and provide explicit bounds on the critical activity, completing the proof of Theorem 5. Recall that

$$\psi_{\boldsymbol{\theta}} = \max_{k=0,\ldots,\Delta-2}\left((\Delta-(k+1))\frac{\theta_{k+1}}{\theta_k}\right),$$

$$\overline{\lambda}_{\boldsymbol{\theta}} = \frac{3\theta_0}{\Delta}\exp\left(3\frac{\theta_{\Delta}}{\theta_{\Delta-1}}\frac{\theta_0}{\theta_1}\right)\left(\frac{\theta_0}{\theta_1}\right)^{\Delta},$$

$$\underline{\lambda}_{\boldsymbol{\theta}} = \left(2\psi_{\boldsymbol{\theta}}\theta_0^{-1}\left(\frac{\theta_{\Delta}}{\theta_{\Delta-1}}\right)^{\Delta-2}\left(\frac{\theta_{\Delta}}{\theta_{\Delta-1}}+(\Delta-1)\left(\frac{\theta_{\Delta}}{\theta_{\Delta-1}}-\frac{\theta_1}{\theta_0}\right)\right)\right)^{-1}.$$

PROOF OF THEOREM 5. We first show that for any log-convex $\boldsymbol{\theta}$, $(\lambda,\boldsymbol{\theta})$ belongs to the uniqueness regime for all activities $\lambda < \underline{\lambda}_{\boldsymbol{\theta}}$. We proceed by proving that for $\lambda < \underline{\lambda}_{\boldsymbol{\theta}}$, the update rule for $|\zeta_d - \zeta_{d-1}|$ implied by (12) is a contraction. We first demonstrate that f, g are Lipschitz, and explicitly bound the relevant Lipschitz constants. Note that f, g are differentiable on $\mathcal{R}^+$. We begin by bounding $|\partial_x g(x)|$. For all $x \geq 0$,

$$|\partial_x g(x)| = \frac{\sum_{k=0}^{\Delta-2}(k+1)\theta_{k+1}\binom{\Delta-1}{k+1}x^k}{(\sum_{k=0}^{\Delta-1}\theta_k\binom{\Delta-1}{k}x^k)^2}.$$

Combining with the fact that for all $x \geq 0$,

$$\sum_{k=0}^{\Delta-1}\theta_k\binom{\Delta-1}{k}x^k \geq \max\left(\theta_0, \sum_{k=0}^{\Delta-2}\theta_k\binom{\Delta-1}{k}x^k\right),$$

it follows that

$$|\partial_x g(x)| \leq \theta_0^{-1}\frac{\sum_{k=0}^{\Delta-2}(k+1)\theta_{k+1}\binom{\Delta-1}{k+1}x^k}{\sum_{k=0}^{\Delta-2}\theta_k\binom{\Delta-1}{k}x^k}$$

$$(20) \qquad\qquad = \theta_0^{-1}\frac{\sum_{k=0}^{\Delta-2}\frac{(k+1)\theta_{k+1}\binom{\Delta-1}{k+1}}{\theta_k\binom{\Delta-1}{k}}(\theta_k\binom{\Delta-1}{k}x^k)}{\sum_{k=0}^{\Delta-2}\theta_k\binom{\Delta-1}{k}x^k}$$

$$\leq \theta_0^{-1}\max_{k=0,\ldots,\Delta-2}\frac{(k+1)\theta_{k+1}\binom{\Delta-1}{k+1}}{\theta_k\binom{\Delta-1}{k}} = \theta_0^{-1}\psi_{\boldsymbol{\theta}};$$

where the final inequality follows from convexity.

We now bound $|\partial_x f(x)|$. For all $x \geq 0$, it follows from (13) that

$$|\partial_x f(x)| \leq (\Delta-1)\frac{\sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2}\binom{\Delta-1}{i}\binom{\Delta-2}{j}x^{i+j}|\theta_i\theta_{j+2}-\theta_{i+1}\theta_{j+1}|}{(\sum_{k=0}^{\Delta-1}\theta_k\binom{\Delta-1}{k}x^k)^2}.$$

Combining with the fact that nonnegativity implies

$$\left(\sum_{k=0}^{\Delta-1} \theta_k \binom{\Delta-1}{k} x^k\right)^2 = \sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-1} \theta_i \theta_j \binom{\Delta-1}{i}\binom{\Delta-1}{j} x^{i+j}$$

$$\geq \sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2} \theta_i \theta_j \binom{\Delta-1}{i}\binom{\Delta-1}{j} x^{i+j},$$

we conclude that

$$|\partial_x f(x)|$$

$$\leq (\Delta-1)\frac{\sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2}\binom{\Delta-1}{i}\binom{\Delta-2}{j}|\theta_i\theta_{j+2}-\theta_{i+1}\theta_{j+1}|x^{i+j}}{\sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2}\theta_i\theta_j\binom{\Delta-1}{i}\binom{\Delta-1}{j}x^{i+j}}$$

$$= (\Delta-1)\frac{\sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2}\frac{\binom{\Delta-1}{i}\binom{\Delta-2}{j}|\theta_i\theta_{j+2}-\theta_{i+1}\theta_{j+1}|x^{i+j}}{\theta_i\theta_j\binom{\Delta-1}{i}\binom{\Delta-1}{j}x^{i+j}}(\theta_i\theta_j\binom{\Delta-1}{i}\binom{\Delta-1}{j}x^{i+j})}{\sum_{i=0}^{\Delta-1}\sum_{j=0}^{\Delta-2}\theta_i\theta_j\binom{\Delta-1}{i}\binom{\Delta-1}{j}x^{i+j}}$$

$$\leq (\Delta-1)\max_{\substack{i\in[0,\Delta-1]\\j\in[0,\Delta-2]}}\frac{\binom{\Delta-1}{i}\binom{\Delta-2}{j}|\theta_i\theta_{j+2}-\theta_{i+1}\theta_{j+1}|x^{i+j}}{\theta_i\theta_j\binom{\Delta-1}{i}\binom{\Delta-1}{j}x^{i+j}}$$

$$= \max_{\substack{i\in[0,\Delta-1]\\j\in[0,\Delta-2]}}\left((\Delta-(j+1))\left|\frac{\theta_{j+2}}{\theta_j}-\frac{\theta_{i+1}}{\theta_i}\frac{\theta_{j+1}}{\theta_j}\right|\right),$$

where the final inequality follows from convexity. Further noting that the definition of ψ_θ and log-convexity together imply that for all $i \in [0, \Delta-1]$ and $j \in [0, \Delta-2]$,

$$(\Delta-(j+1))\left|\frac{\theta_{j+2}}{\theta_j}-\frac{\theta_{i+1}}{\theta_i}\frac{\theta_{j+1}}{\theta_j}\right|$$

$$= (\Delta-(j+1))\frac{\theta_{j+1}}{\theta_j}\left|\frac{\theta_{j+2}}{\theta_{j+1}}-\frac{\theta_{i+1}}{\theta_i}\right|$$

$$\leq \psi_\theta\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}-\frac{\theta_1}{\theta_0}\right),$$

we may combine the above with the chain rule and Lemma 2 to conclude that for all $x \geq 0$,

$$(21) \qquad |\partial_x f^{\Delta-1}(x)| \leq (\Delta-1)\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}\right)^{\Delta-2}\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}-\frac{\theta_1}{\theta_0}\right)\psi_\theta.$$

It follows from (20), (21), Lemmas 1 and 2, the fact that $|ab - cd| \le |a + c||b - d| + |b + d||a - c|$ for all $a, b, c, d \in \mathbb{R}$, that for all $d \ge 2$,

$$
\begin{aligned}
|\zeta_{d+1} - \zeta_d| &= |\lambda g(\zeta_d) f^{\Delta-1}(\zeta_{d-1}) - \lambda g(\zeta_{d-1}) f^{\Delta-1}(\zeta_{d-2})| \\
&\le \lambda |g(\zeta_d) + g(\zeta_{d-1})||f^{\Delta-1}(\zeta_{d-1}) - f^{\Delta-1}(\zeta_{d-2})| \\
&\quad + \lambda |f^{\Delta-1}(\zeta_{d-1}) + f^{\Delta-1}(\zeta_{d-2})||g(\zeta_d) - g(\zeta_{d-1})| \\
&\le \lambda (2\theta_0^{-1}) \left(|\zeta_{d-1} - \zeta_{d-2}|(\Delta - 1)\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}\right)^{\Delta-2} \left(\frac{\theta_\Delta}{\theta_{\Delta-1}} - \frac{\theta_1}{\theta_0}\right)\psi_\theta \right) \\
&\quad + \lambda \left(2\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}\right)^{\Delta-1} \right) (|\zeta_d - \zeta_{d-1}|\theta_0^{-1}\psi_\theta).
\end{aligned}
$$

Note that Lemma 3, combined with our proof of Theorem 4 (in particular the fact that $\sup_{d\ge 0} \zeta_{2d} \le \underline{x} \le \overline{y} \le \inf_{d\ge 0} \zeta_{2d+1}$), implies that $\{|\zeta_{d+1} - \zeta_d|, d \ge 0\}$ is monotone decreasing. Combining the above, we conclude that

$$
\begin{aligned}
&|\zeta_{d+1} - \zeta_d| \\
&\le 2\lambda\theta_0^{-1}\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}\right)^{\Delta-2}\psi_\theta\left((\Delta-1)\left(\frac{\theta_\Delta}{\theta_{\Delta-1}} - \frac{\theta_1}{\theta_0}\right) + \frac{\theta_\Delta}{\theta_{\Delta-1}}\right)|\zeta_{d-1} - \zeta_{d-2}|.
\end{aligned}
$$

It thus follows from the definition of $\underline{\lambda}_\theta$ that for all $\lambda < \underline{\lambda}_\theta$, there exists $\rho \in (0, 1)$ such that for all $d \ge 2$, $|\zeta_{d+1} - \zeta_d| \le \rho|\zeta_{d-1} - \zeta_{d-2}|$. Combining with our proof of Theorem 4 (in particular the fact that existence of $\lim_{d\to\infty} \zeta_d$ implies uniqueness) completes the proof.

We now prove that (λ, θ) does not belong to the uniqueness regime for all $\lambda > \overline{\lambda}_\theta$. We first show that for $\lambda = \overline{\lambda}_\theta$, any nonnegative solution to the system of equations (6)–(7) of the form (x, x) satisfies

$$
(22) \qquad x \ge \frac{3}{\Delta}\frac{\theta_0}{\theta_1}.
$$

Indeed, it follows from log-convexity that $\theta_k \le \theta_0(\frac{\theta_\Delta}{\theta_{\Delta-1}})^k$, $k = 0, \ldots, \Delta$. Thus, for all $x \ge 0$,

$$
\begin{aligned}
\sum_{k=0}^{\Delta-1} \theta_k \binom{\Delta-1}{k} x^k &\le \theta_0 \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k}\left(\frac{\theta_\Delta}{\theta_{\Delta-1}}x\right)^k \\
&= \theta_0\left(1 + \frac{\theta_\Delta}{\theta_{\Delta-1}}x\right)^{\Delta-1},
\end{aligned}
$$

and

$$
g(x) \ge \left(\theta_0\left(1 + \frac{\theta_\Delta}{\theta_{\Delta-1}}x\right)^{\Delta-1}\right)^{-1} \ge \theta_0^{-1}\exp\left(-\Delta\frac{\theta_\Delta}{\theta_{\Delta-1}}x\right).
$$

Thus, by Lemma 2, any nonnegative solution to the system of equations (6)–(7) of the form (x, x) for $\lambda = \overline{\lambda}_{\boldsymbol{\theta}}$ satisfies

$$(23) \qquad x \exp\left(\Delta \frac{\theta_\Delta}{\theta_{\Delta-1}} x\right) \geq \frac{\overline{\lambda}_{\boldsymbol{\theta}}}{\theta_0}\left(\frac{\theta_1}{\theta_0}\right)^{\Delta-1} = \frac{3}{\Delta} \frac{\theta_0}{\theta_1} \exp\left(3\frac{\theta_\Delta}{\theta_{\Delta-1}} \frac{\theta_0}{\theta_1}\right).$$

To complete the proof of (22), we observe that in light of (23), $x < \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$ would yield a contradiction, since it would imply

$$x \exp\left(\Delta \frac{\theta_\Delta}{\theta_{\Delta-1}} x\right) < \frac{3}{\Delta} \frac{\theta_0}{\theta_1} \exp\left(3\frac{\theta_\Delta}{\theta_{\Delta-1}} \frac{\theta_0}{\theta_1}\right).$$

We next prove that for any $x \geq \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$ and $M \geq 1$,

$$(24) \qquad \eta(M, x) \stackrel{\Delta}{=} M g(Mx) \leq g(x),$$

a property that will allow us to use Lemma 1 to explicitly demonstrate that $\{\zeta_d, d \geq 0\}$ exhibits a parity dependence. We proceed by showing that for any $x \geq \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$ and $M \geq 1$, $\partial_M \eta(M, x) \leq 0$. Since $\partial_x g(x) = -g^2(x) \sum_{k=1}^{\Delta-1} \binom{\Delta-1}{k} \theta_k k x^{k-1}$, it follows from the chain rule that

$$\partial_M \eta(M, x) = -Mx g^2(Mx) \sum_{k=1}^{\Delta-1} \binom{\Delta-1}{k} \theta_k k (Mx)^{k-1} + g(Mx)$$

$$= g(Mx)\left(1 - \frac{\sum_{k=1}^{\Delta-1} \binom{\Delta-1}{k} \theta_k k (Mx)^k}{\sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_k (Mx)^k}\right).$$

By the nonnegativity of g, it thus suffices to demonstrate that for any $x \geq \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$ and $M \geq 1$,

$$(25) \qquad \sum_{k=1}^{\Delta-1} \binom{\Delta-1}{k} \theta_k k (Mx)^k - \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_k (Mx)^k$$

is nonnegative. Note that (25) equals

$$\sum_{k=1}^{\Delta-1} \binom{\Delta-1}{k} \theta_k (k-1)(Mx)^k - \theta_0,$$

which by nonnegativity and the fact that $M \geq 1$ is at least $\binom{\Delta-1}{2} \theta_2 x^2 - \theta_0$. As log-convexity implies $\theta_2 \geq \theta_0(\frac{\theta_1}{\theta_0})^2$, and it is easily verified that $\binom{\Delta-1}{2} \geq \frac{\Delta^2}{9}$ for all $\Delta \geq 3$, we conclude that $x \geq \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$ implies that (25) is at least $\frac{\Delta^2}{9} \theta_0(\frac{\theta_1}{\theta_0})^2(\frac{3}{\Delta} \frac{\theta_0}{\theta_1})^2 - \theta_0 = 0$, completing the proof.

We now use (22) and (24) to prove by induction that for $\lambda > \overline{\lambda}_{\boldsymbol{\theta}}$, $\{\zeta_d, d \geq 0\}$ exhibits a parity dependence, mirroring our proof of Theorem 4. Recall from

our proof of Theorem 4 that for $\lambda = \overline{\lambda}_{\boldsymbol{\theta}}$, the system of equations (6)–(7) always has at least one nonnegative solution of the form (x, x). Let us fix any such solution $(x_{\boldsymbol{\theta}}, x_{\boldsymbol{\theta}})$, and note that $x_{\boldsymbol{\theta}} \geq \frac{3}{\Delta} \frac{\theta_0}{\theta_1}$. We now prove by induction that for all $\lambda > \overline{\lambda}_{\boldsymbol{\theta}}$, $\{\zeta_d, d \geq 0\}$ has a nonvanishing parity-dependence, with even values lying below $x_{\boldsymbol{\theta}}$, and odd values lying above $\frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}}$. The $d = 0$ base case follows from nonnegativity. For the case $d = 1$, recall from definitions that $\zeta_1 = \lambda \theta_0^{-1}(\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}$. However, by virtue of satisfying (6)–(7) with $\overline{\lambda}_{\boldsymbol{\theta}}$ and Lemma 2, we have $x_{\boldsymbol{\theta}} \leq \overline{\lambda}_{\boldsymbol{\theta}} \theta_0^{-1}(\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}$. Combining the above completes the proof for the $d = 1$ case. Now, proceeding by induction, suppose that for some $d \geq 1$ and all $k \in \{0, \ldots, d-1\}$, $\zeta_{2k} \leq x_{\boldsymbol{\theta}}$, and $\zeta_{2k+1} \geq \frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}}$. Then it follows from Lemma 1, the monotonicity of f and g implied by Lemma 2, (22) and (24) that

$$\begin{aligned}
\zeta_{2d} &= \lambda g(\zeta_{2d-1}) f^{\Delta-1}(\zeta_{2d-2}) \\
&\leq \lambda g\left(\frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}}\right) f^{\Delta-1}(x_{\boldsymbol{\theta}}) \\
&\leq \lambda \frac{\overline{\lambda}_{\boldsymbol{\theta}}}{\lambda} g(x_{\boldsymbol{\theta}}) f^{\Delta-1}(x_{\boldsymbol{\theta}}) = \overline{\lambda}_{\boldsymbol{\theta}} g(x_{\boldsymbol{\theta}}) f^{\Delta-1}(x_{\boldsymbol{\theta}}) = x_{\boldsymbol{\theta}},
\end{aligned}$$

with the final inequality following from (24). Similarly,

$$\begin{aligned}
\zeta_{2d+1} &= \lambda g(\zeta_{2d}) f^{\Delta-1}(\zeta_{2d-1}) \\
&\geq \lambda g(x_{\boldsymbol{\theta}}) f^{\Delta-1}\left(\frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}}\right) \\
&= \frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} \overline{\lambda}_{\boldsymbol{\theta}} g(x_{\boldsymbol{\theta}}) f^{\Delta-1}\left(\frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}}\right) \geq \frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} \overline{\lambda}_{\boldsymbol{\theta}} g(x_{\boldsymbol{\theta}}) f^{\Delta-1}(x_{\boldsymbol{\theta}}) = \frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} x_{\boldsymbol{\theta}},
\end{aligned}$$

with the final inequality following from the monotonicity of f and fact that $\frac{\lambda}{\overline{\lambda}_{\boldsymbol{\theta}}} > 1$. This completes the desired induction, demonstrating that $\{\zeta_d, d \geq 0\}$ exhibits the stated parity-dependence. Combining with our proof of Theorem 4 (in particular the fact that nonexistence of $\lim_{d \to \infty} \zeta_d$ implies nonuniqueness) completes the proof. □

6. A perturbative analysis and proof of Theorem 6.

In this section, we perform a perturbative analysis of the system of equations arising from our necessary and sufficient conditions for uniqueness, proving Theorem 6. First, it will be useful to rewrite the system of equations (6)–(7), which will allow us to apply known results from the theory of dynamical systems. Note that if $p_{\boldsymbol{\theta}}(x) \stackrel{\Delta}{=} x f_{\boldsymbol{\theta}}^{-(\Delta-1)}(x)$ is strictly increasing on $[0, \lambda(\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1} \theta_0^{-1}]$, then it follows from Lemma 2 that $p_{\boldsymbol{\theta}}$ has a well-defined and unique inverse $p_{\boldsymbol{\theta}}^{\leftarrow}$, with domain a superset of $[0, \lambda \theta_0^{-1}]$ and range a subset of $\mathbb{R}^+$, that is, $p_{\boldsymbol{\theta}}^{\leftarrow}(p_{\boldsymbol{\theta}}(x)) = x$. In this case, we can define

$q_{\lambda,\theta}(x) \overset{\Delta}{=} p_{\theta}^{\leftarrow}(\lambda g_{\theta}(x))$, and we observe that the system of equations (6)–(7) may be rewritten as follows.

OBSERVATION 9. *If θ is log-convex, and $p_{\theta}(x)$ is strictly increasing on $[0, \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}]$, then on $\mathbb{R}^+ \times \mathbb{R}^+$, the system of equations (6)–(7) is equivalent to the system of equations*:

$$(26) \qquad\qquad q_{\lambda,\theta}(q_{\lambda,\theta}(x)) = x,$$

$$(27) \qquad\qquad\qquad y = q_{\lambda,\theta}(x).$$

Furthermore, $q_{\lambda,\theta}$ is strictly decreasing on $\mathbb{R}^+$, and the equation $q_{\lambda,\theta}(x) = x$ has a unique solution $x_{\lambda,\theta}$ on $\mathbb{R}^+$. Also, it follows from Lemma 2 that every solution (x, y) to the system of equations (26)–(27) on $\mathbb{R}^+ \times \mathbb{R}^+$ satisfies $0 \leq x, y \leq \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}$. In addition, $x \in [0, \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}]$ implies $q_{\lambda,\theta}(x) \in [0, \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}]$.

It is well known from the theory of dynamical systems that under certain additional assumptions on $q_{\lambda,\theta}$, necessary and sufficient conditions for when the system of equations (26)–(27) has a unique solution can be stated in terms of whether the map $q_{\lambda,\theta}$ exhibits a certain local stability at the fixed point $x_{\lambda,\theta}$. We now make this precise, and note that our approach is similar to that taken previously in the literature to analyze related models (cf. [22]). Recall that for a thrice-differentiable function $F(x)$ with nonvanishing derivative on some interval I, we define (on I) the *Schwarzian derivative* of F as the function

$$S[F] \overset{\Delta}{=} \frac{d^3}{dx^3}F \bigg/ \frac{d}{dx}F - \frac{3}{2}\left(\frac{d^2}{dx^2}F \bigg/ \frac{d}{dx}F\right)^2.$$

For a function F and $n \geq 1$, let $F^{\{n\}}(x)$ denote the *n-fold iterate* of F, that is, $F^{\{n+1\}}(x) = F(F^{\{n\}}(x))$, with $F^{\{1\}}(x) = F(x)$. Then the following well-known result from dynamical systems is stated in Lemma 4.3 of [22].

THEOREM 7. *Suppose $I = [L, R] \subseteq \mathcal{R}$ is some closed bounded interval, and F is some function with the following properties*:

  (i) *F has domain I, and range a subset of I.*
 (ii) *The third derivative of F exists and is continuous on I.*
(iii) *The equation $x = F(x)$ has a unique solution x^* on I.*
 (iv) *F is a decreasing function on I.*
  (v) *$S[F](x) < 0$ for all $x \in I$.*

Then $\lim_{n\to\infty} F^{\{n\}}(x)$ exists and equals x^* for all $x \in I$ iff $|\partial_x F(x^*)| \leq 1$ iff $\lim_{n\to\infty} F^{\{n\}}(L) = x^*$.

We now customize Theorem 7 to our own setting. Let $r_{\boldsymbol{\theta}}(x) \triangleq |\frac{\partial_x g_{\boldsymbol{\theta}}(x)}{\partial_x p_{\boldsymbol{\theta}}(x)}|$. Then combined with Observation 9, Theorem 7 implies the following.

OBSERVATION 10. *Suppose that $\boldsymbol{\theta}$ is log-convex, $p_{\boldsymbol{\theta}}(x)$ is strictly increasing on $[0, \lambda(\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}]$, and the conditions of Theorem 7 are satisfied with $F = q_{\lambda,\boldsymbol{\theta}}, I = [0, \lambda(\frac{\theta_\Delta}{\theta_{\Delta-1}})^{\Delta-1}\theta_0^{-1}]$. Then $(\lambda, \boldsymbol{\theta})$ belongs to the uniqueness regime iff $r_{\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}}) \leq \lambda^{-1}$.*

PROOF. We first prove that the system of equations (26)–(27) does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$ iff $|\partial_x q_{\lambda,\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}})| > 1$. Suppose the system of equations (26)–(27) does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$. Since Observation 9 implies that the equation $q_{\lambda,\boldsymbol{\theta}}(x) = x$ has a unique solution $x_{\lambda,\boldsymbol{\theta}}$, it follows that there must exist a solution (x, y) to the system of equations (26)–(27) with $x < y$. In this case, $\lim_{n\to\infty} q_{\lambda,\boldsymbol{\theta}}^{\{n\}}(x)$ does not exist, as the series alternates between x and y, and it follows from Theorem 7 that $|\partial_x q_{\lambda,\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}})| > 1$.

Alternatively, suppose that $|\partial_x q_{\lambda,\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}})| > 1$. Then it follows from Theorem 7 that $\lim_{n\to\infty} q_{\lambda,\boldsymbol{\theta}}^{\{n\}}(0)$ does not exist. However, both $Z_{\text{even}} \triangleq \lim_{n\to\infty} q_{\lambda,\boldsymbol{\theta}}^{\{2n\}}(0)$ and $Z_{\text{odd}} \triangleq \lim_{n\to\infty} q_{\lambda,\boldsymbol{\theta}}^{\{2n+1\}}(0)$ both exist. Indeed, this follows from the fact that $q_{\lambda,\boldsymbol{\theta}}^{\{1\}}$ is decreasing, $q_{\lambda,\boldsymbol{\theta}}^{\{2\}}$ is increasing, $q_{\lambda,\boldsymbol{\theta}}^{\{2\}}(0) \geq 0$, and $q_{\lambda,\boldsymbol{\theta}}^{\{3\}}(0) \leq q_{\lambda,\boldsymbol{\theta}}^{\{1\}}(0)$, which implies that both relevant sequences are appropriately monotone. Noting that the nonexistence of the stated limit implies $Z_{\text{even}} \neq Z_{\text{odd}}$, and the pair $(Z_{\text{even}}, Z_{\text{odd}})$ must be a solution to the system of equations (26)–(27), completes the desired demonstration.

As it follows from elementary calculus that $\partial_x q_{\lambda,\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}}) = \lambda \frac{\partial_x g_{\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}})}{\partial_x p_{\boldsymbol{\theta}}(x_{\lambda,\boldsymbol{\theta}})}$, combining the above with Theorem 4 and Observation 9 completes the proof. □

We note that $p_{\boldsymbol{\theta}}$ is not necessarily an increasing function for the case of general log-convex $\boldsymbol{\theta}$. Furthermore, even when $p_{\boldsymbol{\theta}}$ is increasing, an analysis of $S[q_{\lambda,\boldsymbol{\theta}}]$ seems difficult, and the associated uniqueness regime of the parameter space seems to be quite complex. However, for the special setting in which $\boldsymbol{\theta}$ belongs to a neighborhood of the all ones vector, in which case the associated M.r.f. becomes a perturbation of the hardcore model at criticality, these difficulties can be overcome by expanding the relevant functions using appropriate Taylor series. The theory of real analytic functions provides a convenient framework for proving the validity of these expansions, and we refer the reader to [21] for details. Using this framework, we prove the following.

LEMMA 4. *For each convex vector $\mathbf{c}$, and $U \in \mathbb{R}^+$, there exists $\delta_{\mathbf{c},U} > 0$ such that the following hold:*

(i) $g_{1+\mathbf{c}h}(x)$ and $p_{1+\mathbf{c}h}(x)$ are jointly real analytic functions of (h, x) on $[0, \delta_{\mathbf{c},U}] \times [0, U]$. For each fixed $h \in [0, \delta_{\mathbf{c},U}]$ and all $x \in [0, U]$, $\partial_x g_{1+\mathbf{c}h}(x) < 0$, and $\partial_x p_{1+\mathbf{c}h}(x) > 0$.

(ii) For each fixed $h \in [0, \delta_{\mathbf{c},U}]$, $p_{1+\mathbf{c}h}(x)$ has a well-defined and unique inverse $p^{\leftarrow}_{1+\mathbf{c}h}(x)$ with domain a superset of $[0, U]$ and range a subset of $\mathbb{R}^+$. Furthermore, $p^{\leftarrow}_{1+\mathbf{c}h}(x)$ is a jointly real analytic function of (h, x) on $[0, \delta_{\mathbf{c},U}] \times [0, U]$.

(iii) $q_{\lambda_\Delta, 1+\mathbf{c}h}(x)$, $\partial_x q_{\lambda_\Delta, 1+\mathbf{c}h}(x)$, and $S[q_{\lambda_\Delta, 1+\mathbf{c}h}](x)$ are all jointly real analytic functions of (h, x) on $[0, \delta_{\mathbf{c},U}] \times [0, U]$. Furthermore, $\partial_x q_{\lambda_\Delta, 1+\mathbf{c}h}(x)$ and $S[q_{\lambda_\Delta, 1+\mathbf{c}h}](x)$ are strictly negative for all $(h, x) \in [0, \delta_{\mathbf{c},U}] \times [0, U]$.

PROOF.   We prove (i)–(iii) in order.

(i) The claim with respect to real analyticity follows from the fact that for any fixed U_1, there exists $\delta_{1,U_1} > 0$ such that both $g_{1+\mathbf{c}h}(x)$ and $p_{1+\mathbf{c}h}(x)$ are ratios of nonvanishing polynomials of (h, x) on $[0, \delta_{1,U_1}] \times [0, U_1]$. That there exists $\delta_{2,U_1} > 0$ such that $\partial_x g_{1+\mathbf{c}h}(x) < 0$ and $\partial_x p_{1+\mathbf{c}h}(x) > 0$ for all $(h, x) \in [0, \delta_{2,U_1}] \times [0, U_1]$ then follows from the fact that $\partial_x g_{\mathbf{1}}(x) = -(\Delta - 1)(x + 1)^{-\Delta}$, and $\partial_x p_{\mathbf{1}}(x) = 1$.

(ii) The claim follows from (i) and the inverse function theorem for real analytic functions (cf. [21]).

(iii) The claim with respect to $q_{\lambda_\Delta, 1+\mathbf{c}h}(x)$ and $\partial_x q_{\lambda_\Delta, 1+\mathbf{c}h}(x)$ follows from (ii), and the fact that $\partial_x q_{\lambda_\Delta, \mathbf{1}}(x) = -(\Delta - 1)\lambda_\Delta(x + 1)^{-\Delta}$. As this implies that $\partial_x q_{\lambda_\Delta, \mathbf{1}}(x)$ is strictly negative (and thus nonvanishing), the desired claim with respect to $S[q_{\lambda_\Delta, 1+\mathbf{c}h}](x)$ then follows from the fact that $S[q_{\lambda_\Delta, \mathbf{1}}](x) = \frac{-\Delta(\Delta-2)}{2(x+1)^2}$. □

Combining Observations 10 and 6 with Lemma 4 immediately yields necessary and sufficient conditions for uniqueness when $\boldsymbol{\theta}$ is a convex perturbation of $\mathbf{1}$.

COROLLARY 2.   For each convex vector $\mathbf{c}$, there exists $\delta_{\mathbf{c}} > 0$ such that the following hold for all $h \in [0, \delta_{\mathbf{c}}]$:

(i) $q_{\lambda_\Delta, 1+\mathbf{c}h}(x) - x$ is strictly decreasing on $[0, 2\lambda_\Delta]$, and has a unique zero $x_{\lambda_\Delta, 1+\mathbf{c}h}$ on $[0, 2\lambda_\Delta]$.

(ii) $(\lambda_\Delta, \mathbf{1} + \mathbf{c}h)$ belongs to the uniqueness regime iff $r_{1+\mathbf{c}h}(x_{\lambda_\Delta, 1+\mathbf{c}h}) \leq \lambda_\Delta^{-1}$.

With Corollary 2 in hand, we now complete the proof of Theorem 6. For $l \in \{0, 1\}$ and $\mathbf{c} = (c_0, \ldots, c_\Delta)$, let

$$f_{l,\boldsymbol{\theta}}(x) \triangleq \sum_{i=0}^{\Delta-1} \theta_{i+l} \binom{\Delta - 1}{i} x^i,$$

$$z_{l,\mathbf{c}} \triangleq \sum_{i=0}^{\Delta-1} \binom{\Delta - 1}{i} x_{\lambda_\Delta, \mathbf{1}}^i c_{i+l},$$

$$w_{l,\mathbf{c}} \triangleq \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} i x_{\lambda_\Delta,\mathbf{1}}^{i-1} c_{i+l},$$

and

$$x_{\mathbf{c}} \triangleq \frac{1}{2} \frac{(\Delta-2)^{\Delta-2}}{(\Delta-1)^{\Delta-1}} ((\Delta-1)z_{1,\mathbf{c}} - \Delta z_{0,\mathbf{c}}).$$

Also, let $o(h)$ denote the family of functions $F(h)$ such that $\lim_{h\downarrow 0} h^{-1}F(h) = 0$. With a slight abuse of notation, we will also let $o(h)$ refer to any particular function belonging to this family. Finally, in simplifying certain expressions, we will use the following identities, which follow from a straightforward calculation (the details of which we omit).

LEMMA 5.

$$x_{\lambda_\Delta,\mathbf{1}} = (\Delta-2)^{-1}, \qquad \sum_{i=0}^{\Delta} \Lambda_{\Delta,i} = \left(\frac{\Delta-1}{\Delta-2}\right)^\Delta,$$

$$\sum_{i=0}^{\Delta} i \Lambda_{\Delta,i} = \Delta \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta}, \qquad \sum_{i=0}^{\Delta} i^2 \Lambda_{\Delta,i} = 2\Delta \frac{(\Delta-1)^{\Delta-1}}{(\Delta-2)^\Delta},$$

$$\sum_{i=0}^{\Delta} \pi_i = -\left(\frac{\Delta-1}{\Delta-2}\right)^{\Delta-1}, \qquad \sum_{i=0}^{\Delta} i\pi_i = \Delta\left(\frac{\Delta-1}{\Delta-2}\right)^{\Delta-1}.$$

PROOF OF THEOREM 6. We proceed by analyzing $r_{\mathbf{1}+\mathbf{c}h}(x_{\lambda_\Delta,\mathbf{1}+\mathbf{c}h}) - \lambda_\Delta^{-1}$ as $h \downarrow 0$, and begin by proving that

(28) $$\lim_{h\downarrow 0}(x_{\lambda_\Delta,\mathbf{1}+\mathbf{c}h} - x_{\lambda_\Delta,\mathbf{1}})h^{-1} = x_{\mathbf{c}}.$$

Note that for any fixed $\alpha \in \mathbb{R}$ and $l \in \{0,1\}$,

$$f_{l,\mathbf{1}+\mathbf{c}h}(x_{\lambda_\Delta,\mathbf{1}} + \alpha h)$$

(29)
$$= \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} (1 + c_{i+l}h) \sum_{j=0}^{i} \binom{i}{j} x_{\lambda_\Delta,\mathbf{1}}^{i-j} (\alpha h)^j$$

$$= \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} x_{\lambda_\Delta,\mathbf{1}}^{i}(1 + c_{i+l}h)(1 + ix_{\lambda_\Delta,\mathbf{1}}^{-1}\alpha h) + o(h)$$

$$= (1 + x_{\lambda_\Delta,\mathbf{1}})^{\Delta-1} + ((\Delta-1)(1 + x_{\lambda_\Delta,\mathbf{1}})^{\Delta-2}\alpha + z_{l,\mathbf{c}})h + o(h).$$

We conclude that

$$g_{\mathbf{1}+\mathbf{c}h}(x_{\lambda_\Delta,\mathbf{1}} + \alpha h)$$

(30) $$= (1 + x_{\lambda_\Delta,\mathbf{1}})^{-(\Delta-1)}$$

$$- (1 + x_{\lambda_\Delta,\mathbf{1}})^{-2(\Delta-1)}((\Delta-1)(1 + x_{\lambda_\Delta,\mathbf{1}})^{\Delta-2}\alpha + z_{0,\mathbf{c}})h + o(h),$$

and

$$(31) \qquad f_{1+ch}(x_{\lambda_\Delta,1} + \alpha h) = 1 + (1 + x_{\lambda_\Delta,1})^{-(\Delta-1)}(z_{1,c} - z_{0,c})h + o(h).$$

It follows from (30), (31) and a straightforward calculation (the details of which we omit) that for $\alpha \in \mathbb{R}$,

$$(x_{\lambda_\Delta,1} + \alpha h) - \lambda_\Delta f_{1+ch}^{\Delta-1}(x_{\lambda_\Delta,1} + \alpha h)g_{1+ch}(x_{\lambda_\Delta,1} + \alpha h)$$
$$= 2(\alpha - x_c)h + o(h).$$

Combining with Corollary 2(i), and the fact that $x_{\lambda_\Delta,1} < \lambda_\Delta$, completes the proof.

Next, we use (28) to prove that

$$(32) \qquad \begin{aligned} \partial_x p_{1+ch}(x_{\lambda_\Delta,1+ch}) &+ \lambda_\Delta \partial_x g_{1+ch}(x_{\lambda_\Delta,1+ch}) \\ &= -\frac{1}{2}\left(\frac{\Delta-2}{\Delta-1}\right)^\Delta \pi \cdot ch + o(h). \end{aligned}$$

Indeed, it follows from (28) that

$$\partial_x f_{l,1+ch}(x_{\lambda_\Delta,1+ch})$$
$$= \sum_{i=1}^{\Delta-1} \binom{\Delta-1}{i} i(1 + c_{i+l}h) \sum_{j=0}^{i-1} \binom{i-1}{j} x_{\lambda_\Delta,1}^{i-1-j}(x_c h)^j + o(h)$$
$$= \sum_{i=1}^{\Delta-1} \binom{\Delta-1}{i} i x_{\lambda_\Delta,1}^i (1 + c_{i+l}h)(x_{\lambda_\Delta,1}^{-1} + (i-1)x_{\lambda_\Delta,1}^{-2} x_c h) + o(h),$$

which itself equals

$$(33) \qquad \begin{aligned} (\Delta-1)(1 &+ x_{\lambda_\Delta,1})^{\Delta-2} \\ &+ \big(x_c(\Delta-1)(\Delta-2)(1 + x_{\lambda_\Delta,1})^{\Delta-3} + w_{l,c}\big)h + o(h). \end{aligned}$$

It follows from (28)–(33), and a straightforward calculation (the details of which we omit), that

$$\partial_x g_{1+ch}(x_{\lambda_\Delta,1+ch}) = -g_{1+ch}^2(x_{\lambda_\Delta,1+ch})\partial_x f_{0,1+ch}(x_{\lambda_\Delta,1+ch}),$$

which itself equals

$$(34) \qquad \begin{aligned} -(\Delta-1)(1 &+ x_{\lambda_\Delta,1})^{-\Delta} + (1 + x_{\lambda_\Delta,1})^{-(2\Delta-1)} \\ &\times \big(-(1 + x_{\lambda_\Delta,1})w_{0,c} + \Delta(\Delta-1)(1 + x_{\lambda_\Delta,1})^{\Delta-2}x_c + 2(\Delta-1)z_{0,c}\big)h \\ &+ o(h); \end{aligned}$$

$\partial_x f_{1+ch}(x_{\lambda_\Delta,1+ch})$ equals

$$\begin{aligned} g_{1+ch}^2(x_{\lambda_\Delta,1+ch})\big(f_{0,1+ch}(x_{\lambda_\Delta,1+ch})\partial_x f_{1,1+ch}(x_{\lambda_\Delta,1+ch}) \\ - f_{1,1+ch}(x_{\lambda_\Delta,1+ch})\partial_x f_{0,1+ch}(x_{\lambda_\Delta,1+ch})\big), \end{aligned}$$

which itself equals

$$(1 + x_{\lambda_\Delta, \mathbf{1}})^{-\Delta}\big((1 + x_{\lambda_\Delta, \mathbf{1}})(w_{1,\mathbf{c}} - w_{0,\mathbf{c}}) + (\Delta - 1)(z_{0,\mathbf{c}} - z_{1,\mathbf{c}})\big)h + o(h);$$

and

$$\partial_x p_{\mathbf{1+c}h}(x_{\lambda_\Delta, \mathbf{1+c}h})$$
$$= f_{\mathbf{1+c}h}^{-(\Delta-1)}(x_{\lambda_\Delta, \mathbf{1+c}h})$$
$$\quad - (\Delta - 1)x_{\lambda_\Delta, \mathbf{1+c}h} f_{\mathbf{1+c}h}^{-\Delta}(x_{\lambda_\Delta, \mathbf{1+c}h})\partial_x f_{\mathbf{1+c}h}(x_{\lambda_\Delta, \mathbf{1+c}h}),$$

which itself equals

$$(35) \qquad 1 - (\Delta - 1)x_{\lambda_\Delta, \mathbf{1}}(1 + x_{\lambda_\Delta, \mathbf{1}})^{-(\Delta-1)}(w_{1,\mathbf{c}} - w_{0,\mathbf{c}})h + o(h).$$

Combining (34)–(35) with Lemma 5 and simplifying, we conclude that the LHS of (32) equals

$$(36) \qquad \begin{aligned} &\frac{(\Delta - 2)^{\Delta-2}}{2(\Delta - 1)^\Delta}\big(-(\Delta - 2)^3 z_{0,\mathbf{c}} + \Delta(\Delta - 1)(\Delta - 2)z_{1,\mathbf{c}} \\ &\quad + 2(\Delta - 1)(\Delta - 2)w_{0,\mathbf{c}} - 2(\Delta - 1)^2 w_{1,\mathbf{c}}\big)h + o(h). \end{aligned}$$

It follows from the definition of $\boldsymbol{\pi}$ and a further straightforward algebraic manipulation that (36) equals $-\frac{1}{2}(\frac{\Delta-2}{\Delta-1})^\Delta \boldsymbol{\pi} \cdot \mathbf{c}h + o(h)$, completing the desired demonstration.

Combining (28), (32) and Corollary 2 with the fact that $r_{\mathbf{1+c}h}(x_{\lambda_\Delta, \mathbf{1+c}h}) \le \lambda_\Delta^{-1}$ iff the LHS of (32) is nonnegative completes the proof. $\square$

7. Conclusion. In this paper, we investigated second-order M.r.f. for independent sets on the infinite Cayley tree, a generalization of the hardcore model which arises in statistical physics, combinatorial optimization and operations research, with an eye toward understanding which distributions can be attained for the number of included neighbors of an excluded node, while staying in the uniqueness regime. We proved that the associated Gibbsian specification satisfies the FKG inequality whenever the local potentials defining the Hamiltonian satisfy a certain log-convexity condition, which leads to so-called reverse ultra log-concave distributions for the number of included neighbors of an excluded node. Under this condition, we gave necessary and sufficient conditions for the existence of a unique infinite-volume Gibbs measure in terms of an explicit system of equations, proved the existence of a phase transition, and gave explicit lower and upper bounds on the associated critical activity, which we proved to exhibit a certain robustness. For potentials which are small perturbations of those coinciding to the hardcore model at its critical activity, we performed a perturbative analysis of the system of equations arising from our necessary and sufficient conditions for uniqueness, allowing us to explicitly characterize whether the resulting specification has a unique

infinite-volume Gibbs measure in terms of whether these perturbations satisfy an explicit linear inequality. Our analysis revealed an interesting nonmonotonicity with regards to biasing toward excluded nodes with no included neighbors, which we used (in conjunction with our lower and upper bounds) to compare the uniqueness regime for our model to a related model in which the associated potentials have a simple factorized form.

This work leaves many interesting directions for future research. The full power of higher-order M.r.f. for sampling from independent sets in sparse graphs, and the associated uniqueness regime, remains poorly understood. Several questions build immediately on the models considered in this paper, such as developing a deeper understanding of the uniqueness regime for second-order M.r.f. with log-convex potentials, and more generally higher-order M.r.f. which also satisfy the FKG inequality. Analyzing settings in which the FKG inequality no longer holds (at least for the partial order considered in this paper), for example, second-order M.r.f. with log-concave potentials (which includes the restriction to maximal independent sets, and for which f_{θ} is monotone decreasing instead of increasing), remains an open challenge. It is also an open question to understand which sets of occupancy probabilities can be achieved by higher-order M.r.f. (in the uniqueness regime). Can one use higher-order M.r.f. (in the uniqueness regime) to sample from denser independent sets than can be attained using the hardcore model at the critical activity? It would also be interesting to study higher-order M.r.f. for related combinatorial problems, for example, graph coloring, as well as for more general sparse graphs. Indeed, although our results can be easily extended to the setting of regular graphs of large girth using standard techniques, proving results for general bounded-degree graphs (as was done by Weitz for the hardcore model in [45]) seems to require fundamentally new ideas. Similarly, the algorithmic implications of phase transitions for higher-order M.r.f. also remain open questions. In particular, one would expect a "complexity transition" at the uniqueness threshold with respect to approximately computing the relevant partition functions, as has been recently established for first-order M.r.f. (cf. [35]).

Finally, it is open to investigate the connection between higher-order M.r.f. and research on bernoulli shifts, i.i.d. factors of graphs, and local algorithms (cf. [4, 43]), which have played a prominent role recently in developing algorithms for finding dense independent sets in sparse graphs (cf. [8, 14]). For example, it has been proven that under certain additional technical assumptions, certain M.r.f. can(not) be well-approximated (in an appropriate sense) by i.i.d. factors of graphs (cf. [3, 43]). However, a complete understanding of this and related questions seems beyond the reach of current techniques. The converse, that is, questions regarding whether an i.i.d. factor of graphs can be well-approximated by a finite-order M.r.f. in the uniqueness regime, seem to have received less attention in the literature, beyond the special case in which the underlying graph is a line and the M.r.f. reduces to a Markov chain (cf. [25, 30]). Such a connection could open the door to, for example, searching the space of M.r.f. (which are easily parametrized

on sparse graphs) to find specifications which sample from dense independent sets while remaining in the uniqueness regime. We conclude with the following related question.

QUESTION 2. *To what extent are higher-order M.r.f. in the uniqueness regime capable of* (*approximately*) *encoding those distributions on independent sets which exhibit long-range independence*, *such as i.i.d. factors of graphs?*

Acknowledgements. The author would like to thank Rob van den Berg, David Gamarnik, Kavita Ramanan, Justin Salez and Prasad Tetali for several stimulating discussions. The author especially thanks Rob van den Berg for a very helpful conversation regarding the applicability of the FKG theorem.

REFERENCES

[1] ACHLIOPTAS, D. and COJA-OGHLAN, A. (2008). Algorithmic barriers from phase transitions. In *Foundations of Computer Science* 793–802. IEEE, New York.

[2] AKYILDIZ, I., WEILIAN, S., SANKARASUBRAMANIAM, Y. and CAYIRCI, E. (2002). Wireless sensor networks: A survey. *Computer Networks* **38** 393–422.

[3] BACKHAUSZ, A. and SZEGEDY, B. (2014). On large girth regular graphs and random processes on trees. Preprint. Available at arXiv:1406.4420.

[4] BACKHAUSZ, Á., SZEGEDY, B. and VIRÁG, B. (2015). Ramanujan graphings and correlation decay in local algorithms. *Random Structures Algorithms* **47** 424–435. MR3385741

[5] BASIN, D., CAPKUN, S., HUBAUX, J., LAFOURCADE, P., PAPADIMITRATOS, P., POTURALSKI, M. and SCHALLER, P. (2008). Secure neighborhood discovery: A fundamental element for mobile ad hoc networking. *Communications Magazine*, *IEEE* **46** 132–139.

[6] BIROLI, G. and MÉZARD, M. (2001). Lattice glass models. *Phys. Rev. Lett.* **88** 025501.

[7] CHEN, W. Y. C. and GU, C. C. Y. (2009). The reverse ultra log-concavity of the Boros–Moll polynomials. *Proc. Amer. Math. Soc.* **137** 3991–3998. MR2538559

[8] CSÓKA, E., GERENCSÉR, B., HARANGI, V. and VIRÁG, B. (2015). Invariant Gaussian processes and independent sets on regular graphs of large girth. *Random Structures Algorithms* **47** 284–303. MR3382674

[9] DALL'ASTA, L., PIN, P. and RAMEZANPOUR, A. (2009). Statistical mechanics of maximal independent sets. *Phys. Rev. E* **80** 061136.

[10] DING, J., SLY, A. and SUN, N. (2013). Maximum independent sets on random regular graphs. Preprint. Available at arXiv:1310.4787.

[11] DOMB, C. (1960). On the theory of cooperative phenomena in crystals. *Advances in Physics* **9** 245–361.

[12] FORTUIN, C. M., KASTELEYN, P. W. and GINIBRE, J. (1971). Correlation inequalities on some partially ordered sets. *Comm. Math. Phys.* **22** 89–103. MR0309498

[13] FRIEDLI, S. and VELENIK, Y. (2014). Equilibrium statistical mechanics of classical lattice systems: A concrete introduction.

[14] GAMARNIK, D. and SUDAN, M. (2014). Limits of local algorithms over sparse random graphs [extended abstract]. In *ITCS'14—Proceedings of the* 2014 *Conference on Innovations in Theoretical Computer Science* 369–375. ACM, New York. MR3359490

[15] GAREY, M. R. and JOHNSON, D. S. (1979). *Computers and Intractability*. W. H. Freeman and Co., San Francisco, CA. MR0519066

[16] HÄGGSTRÖM, O. (1997). Ergodicity of the hard-core model on $\mathbf{Z}^2$ with parity-dependent activities. *Ark. Mat.* **35** 171–184. MR1443040

[17] HELLINGS, T., BORST, S. C. and VAN LEEUWAARDEN, J. S. H. (2011). Tandem queueing networks with neighbor blocking and back-offs. *Queueing Syst.* **68** 321–331. MR2834203

[18] KELLY, F. P. (1991). Loss networks. *Ann. Appl. Probab.* **1** 319–378. MR1111523

[19] KHABBAZIAN, M. and BHARGAVA, V. K. (2008). Localized broadcasting with guaranteed delivery and bounded transmission redundancy. *IEEE Trans. Comput.* **57** 1072–1086. MR2447106

[20] KOB, W. and ANDERSEN, H. C. (1993). Kinetic lattice-gas model of cage effects in high-density liquids and a test of mode-coupling theory of the ideal-glass transition. *Phys. Rev. E* **48** 4364–4377.

[21] KRANTZ, S. and PARKS, H. (2002). *A Primer of Real Analysis and Functions*. Springer, Berlin.

[22] LUEN, B., RAMANAN, K. and ZIEDINS, I. (2006). Nonmonotonicity of phase transitions in a loss network with controls. *Ann. Appl. Probab.* **16** 1528–1562. MR2260072

[23] MÉZARD, M., PARISI, G. and ZECCHINA, R. (2002). Analytic and algorithmic solution of random satisfiability problems. *Science* **297** 812–815.

[24] MOSSEL, E., WEITZ, D. and WORMALD, N. (2009). On the hardness of sampling independent sets beyond the tree threshold. *Probab. Theory Related Fields* **143** 401–439. MR2475668

[25] ORNSTEIN, D. S. (1973). An application of ergodic theory to probability theory. *Ann. Probab.* **1** 43–65. MR0348831

[26] PEMANTLE, R. (2000). Towards a theory of negative dependence. *J. Math. Phys.* **41** 1371–1390. MR1757964

[27] PRESTON, C. J. (1976). *Random Fields. Lectures Notes in Math* **534**. Springer, Berlin.

[28] RAMANAN, K., SENGUPTA, A., ZIEDINS, I. and MITRA, P. (2002). Markov random field models of multicasting in tree networks. *Adv. in Appl. Probab.* **34** 58–84. MR1895331

[29] ROZIKOV, U. A. (2013). Gibbs measures on Cayley trees: Results and open problems. *Rev. Math. Phys.* **25** 1330001, 112. MR3024014

[30] RUDOLPH, D. J. and SCHWARZ, G. (1977). The limits in $\overline{d}$ of multi-step Markov chains. *Israel J. Math.* **28** 103–109. MR0460596

[31] RUNNELS, L. K. (1967). Phase transition of a Bethe lattice gas of hard molecules. *J. Math. Phys.* **8** 2081–2087.

[32] RUNNELS, L. K. and COMBS, L. L. (1966). Exact finite method of lattice statistics. I. Square and triangular lattice gases of hard molecules. *J. Chem. Phys.* **45** 2482–2492.

[33] SALEZ, J. (2013). The cavity method for counting spanning subgraphs subject to local constraints. *Random Structures Algorithms* **43** 377–397.

[34] SHEFFIELD, S. (2005). Random surfaces. *Astérisque* **304** vi+175. MR2251117

[35] SLY, A. and SUN, N. (2012). The computational hardness of counting in two-spin models on d-regular graphs. In 2012 *IEEE 53rd Annual Symposium on Foundations of Computer Science—FOCS* 2012 361–369. IEEE Computer Soc., Los Alamitos, CA. MR3186623

[36] SLY, A. and SUN, N. (2014). Counting in two-spin models on d-regular graphs. *Ann. Probab.* **42** 2383–2416. MR3265170

[37] SPITZER, F. (1975). Markov random fields on an infinite tree. *Ann. Probab.* **3** 387–398. MR0378152

[38] TEMPERLEY, H. N. V. (1959). Application of the Mayer method to the melting problem. *Proc. Phys. Soc.* **74** 183–195. MR0110256

[39] THOMPSON, C. J. (1982). Local properties of an Ising model on a Cayley tree. *J. Stat. Phys.* **27** 441–456. MR0659804

[40] TJELMELAND, H. and BESAG, J. (1998). Markov random fields with higher-order interactions. *Scand. J. Stat.* **25** 415–433. MR1650011

[41] TOPKIS, D. M. (1998). *Supermodularity and Complementarity*. Princeton Univ. Press, Princeton, NJ. MR1614637

[42] VANNIMENUS, J. (1981). Modulated phase of an Ising system with competing interactions on a Cayley tree. *Zeitschrift Für Physik B Condensed Matter* **43** 141–148.

[43] VAN DEN BERG, J. and STEIF, J. E. (1999). On the existence and nonexistence of finitary codings for a class of random fields. *Ann. Probab.* **27** 1501–1522. MR1733157

[44] WAINWRIGHT, M. J. and JORDAN, M. I. (2008). Graphical models, exponential families, and variational inference. *Foundations and Trends® in Machine Learning* **1** 1–305.

[45] WEITZ, D. (2006). Counting independent sets up to the tree threshold. In *STOC'06: Proceedings of the* 38*th Annual ACM Symposium on Theory of Computing* 140–149. ACM, New York. MR2277139

GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332
USA
E-MAIL: dgoldberg9@isye.gatech.edu