HIGHER ORDER MARKOV RANDOM FIELDS FOR INDEPENDENT SETS

BY DAVID A. GOLDBERG

Georgia Institute of Technology

It is well-known that if one samples from the independent sets of a large regular graph of large girth using a pairwise Markov random field (i.e. hardcore model) in the uniqueness regime, each excluded node has a binomially distributed number of included neighbors in the limit. In this paper, motivated by applications to the design of communication networks, we pose the question of how to sample from the independent sets of such a graph so that the number of included neighbors of each excluded node has a different distribution of our choosing.

We observe that higher order Markov random fields are well-suited to this task, and investigate the properties of these models. For the family of so-called reverse ultra log-concave distributions, which includes the truncated Poisson and geometric, we give necessary and sufficient conditions for the natural higher order Markov random field which induces the desired distribution to be in the uniqueness regime, in terms of the set of solutions to a certain system of equations. We also show that these Markov random fields undergo a phase transition, and give explicit bounds on the associated critical activity, which we prove to exhibit a certain robustness. For distributions which are small perturbations around the binomial distribution realized by the hardcore model at critical activity, we give a description of the corresponding uniqueness regime in terms of a simple polyhedral cone. Our analysis reveals an interesting non-monotonicity with regards to biasing towards excluded nodes with no included neighbors. We conclude with a broader discussion of the potential use of higher order Markov random fields for analyzing independent sets in graphs.

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 [85],[75],[1],[104]. 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; i.e. a set $S \in 2^V$ is an independent set iff for all pairs of nodes $v_i, v_j \in S$, $(v_i, v_j) \notin E$. Independent sets arise in many applications, ranging from the study of communication networks [103] to computer vision [24], economics [36], and biology [71]; and the problem of finding the maximum independent set of a graph was one of the original NP-Complete problems identified in Garey and Johnson’s classical text on computational intractibility [90]. 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.

AMS 2000 subject classifications: Primary 60K35

Keywords and phrases: independent set, Markov random field, Gibbs measure, phase transition, hardcore model
1.1. **Hardcore model.** The hardcore model, originally studied in the statistical physics community to understand anti-ferromagnetic particle systems, refers to the following family of distributions on the independent sets of a graph $G$. For any independent set $I \in \mathcal{I}(G)$, $\mathbb{P}(I) = \lambda^{|I|}(\sum_{S \in \mathcal{I}(G)} \lambda^{|S|})^{-1},$

where $\lambda$ is a positive activity parameter whose logarithm has an interpretation in terms of the temperature of the system, and $|S|$ denotes the cardinality of $S$. When $\lambda = 1$, computing the relevant normalizing constant / partition function $\sum_{S \in \mathcal{I}(G)} \lambda^{|S|}$ is equivalent to counting the number of independent sets in $G$ (a $\#P$-Complete problem [122]); 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.

Such anti-ferromagnetic models have a rich history in the physics literature. Models in which $G$ is a lattice were studied in [41],[26],[110],[111],[112],[45],[100],[97],[11]. This work was extended to the three-regular infinite tree (i.e. Bethe lattice, infinite Cayley tree) by L.K. Runnels in [101], and the general infinite regular tree in [32]. In these early works, it was observed that the underlying models underwent a *phase transition* as one varied $\lambda$. As this concept will be central to the rest of the paper, we now make this more precise.

1.1.1. **Phase transition.** For a fixed graph $G$ and $\lambda \in \mathbb{R}^+$, let $\mathbb{P}_{\lambda,G}$ denote the measure, on the independent sets of $G$, which assigns independent set $I$ probability $\lambda^{|I|}(\sum_{S \in \mathcal{I}(G)} \lambda^{|S|})^{-1}$. Let $T_d$ denote the rooted (at $r$) depth-$d$ tree in which all non-leaf nodes have degree $\Delta$; and $\mathbb{P}_{\lambda,T_d}(r \in I)$ the probability that the root is included in the corresponding random independent set. Then for each $\Delta \geq 3$, there exists a critical activity $\lambda_\Delta \overset{\Delta}{=} (\Delta - 1)^{\Delta - 1}(\Delta - 2)^{-\Delta}$ such that for all $\lambda \in (0, \lambda_\Delta]$ (i.e. uniqueness regime), $\lim_{d \to \infty} \mathbb{P}_{\lambda,T_d}(r \in I)$ exists; for all $\lambda > \lambda_\Delta$, the limit does not exist. In particular, for $\lambda \leq \lambda_\Delta$, there is an asymptotic independence on the boundary condition at the base of the tree (i.e. parity of $d$); for $\lambda > \lambda_\Delta$, there is a non-vanishing dependence on this boundary condition [72]. It is well-known that the existence of this limit can also be phrased in terms of whether or not an appropriate infinite graph has a unique translation-invariant Gibbs measure [124],[40],[25],[121],[83],[120],[60].

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 $\Delta$ switch from mixing in polynomial time to mixing in exponential time [89]. Furthermore, it was shown in [122] 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 $\Delta$. Combined with the results of [107],[108],[48],[49] which show that no such FPTAS exists for $\lambda > \lambda_\Delta$ unless certain complexity classes collapse (e.g. $NP = RP$), as well as several recent related works [7],[106],[80],[27],[123], this shows that the aforementioned phase transition has deep connections to computational complexity. Questions regarding the existence of such phase transitions also arise in the context of various other applications, e.g. fitting correlation structures in computer vision models [113], since systems exhibiting long-range boundary dependence may be undesirable from a modeling and simulation perspective.

In light of the above difficulties associated with long-range dependencies, and the importance of independent sets in various applications, the following is a natural question.
**Question 1.** What distributions can one construct on the independent sets of large bounded-degree graphs, which do not exhibit long-range correlations?

Certainly, the hardcore model (and its known generalizations) in the uniqueness regime provide one way to generate such distributions, and we refer the reader to the recent survey [99] for an overview. Question 1 has also been approached through the lens of so-called local algorithms and i.i.d. factors of graphs [55], [53], [54], [84], [43], [19], [33], [66], [56], [6], [34] in which one samples from the independent sets of a graph by assigning nodes random weights, and using these weights to select nodes for inclusion in a distributed, localized manner. An alternative (but related) approach is to sample from the independent sets of a graph using greedy algorithms which can be analyzed locally [78], [69], [61], [52], [70]. In these settings, Question 1 is typically approached through the lens of studying how dense an independent set can be sampled before long-range correlations begin to manifest.

Another relevant aspect to Question 1 pertains to the distribution of the number of included neighbors of an excluded node, which arises in the design and optimization of large wireless networks [9], [14], [4]. In particular, we are led to the following instantiation of Question 1.

**Question 2.** What distributions can one construct on the independent sets of large bounded-degree graphs, such that the number of included neighbors of any given excluded node has a given distribution of our choosing? Furthermore, can this be done without inducing long-range correlations?

A good starting place is the hardcore model, for which the following result is well-known [109]. We note that although we have only defined the hardcore model for finite graphs, the model can also be formally defined on infinite graphs, and we refer the reader to [109] for details. Let \( B(n, p) \) denote a standard binomial distribution with parameters \( n \) and \( p \).

**Observation 1.** [109] 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 without inducing long-range correlations is dictated by the phase-transition at \( \lambda_\Delta \). In particular, it is possible to induce a \( B(\Delta, p) \) distribution on the number of included neighbors of each excluded node for any \( p \in (0, (\Delta - 1)^{-1}] \) without inducing long-range correlations.

1.2. Markov random fields. We will now briefly review the family of distributions known as Markov random fields, which generalize the hardcore model, as they will provide the framework with which we tackle Question 2. Recall that a collection of r.v.s \( X_1, \ldots, X_n \) defines a binary Markov random field if the following is true [13], [81]. First, \( X_i \) has support on \( \{0, 1\} \) for all \( i \). Second, to each r.v. \( X_i \), we can associate a subset \( N_i \) of the indices \( \{1, \ldots, n\} \) (called the neighborhood of \( X_i \)), such that for any binary \( n \)-dimensional vector \( x \),

\[
\mathbb{P}\left( \bigcap_{j \in \{1, \ldots, n\} \setminus i} \{X_j = x_j\} \right) > 0 \implies \mathbb{P}(X_i = 1 | \bigcap_{j \in \{1, \ldots, n\} \setminus i} \{X_j = x_j\}) = \mathbb{P}(X_i = 1 | \bigcap_{j \in N_i} \{X_j = x_j\}).
\]

In words, the conditional probability that \( X_i \) takes a given value if we condition on all other r.v.s is the same as the corresponding conditional probability if we only condition on the r.v.s belonging to the neighborhood of \( X_i \). We further suppose that the neighborhood relation is symmetric, i.e.
$i \in N_j$ iff $j \in N_i$. In this case, note that the neighborhood structure of the Markov random field defines an undirected graph, in which there is a node for each variable, and two nodes are adjacent iff they are neighbors.

A closely related concept is that of the Gibbs measure. Recall that the cliques of a graph $G$, $\mathcal{C}(G)$, are defined to be the subsets $S$ of $V$ with all internal edges present; i.e. a set $S \in 2^V$ is a clique iff for all pairs of distinct nodes $v_i, v_j \in S$, $(v_i, v_j) \in E$. For each clique $C \in \mathcal{C}(G)$, we define a non-negative function $\pi_C$ (called the clique potential), which takes as input binary values for all the nodes belonging to $C$, and outputs a non-negative real number. For a graph $G$, let $|G|$ denote the number of nodes. For a binary $|G|$-dimensional vector $x$ and clique $C \in \mathcal{C}(G)$, let $x_C$ denote the projection of the vector $x$ onto the indices of nodes present in the clique $C$. Then given a graph $G$, a binary r.v. $X_i$ associated to each node $v_i \in V$, and a potential function $\pi_C$ for each clique $C \in \mathcal{C}(G)$, we define the associated binary Gibbs measure $\mathbb{P}$ on the r.v.s $X_1, \ldots, X_n$ as follows. For every binary $|G|$-dimensional vector $x$,

$$\mathbb{P}(\bigcap_{i=1}^{|G|}\{X_i = x_i\}) = \left(\sum_{x \in \{0,1\}^{|G|}} \prod_{C \in \mathcal{C}(G)} \pi_C(x_C)\right)^{-1} \prod_{C \in \mathcal{C}(G)} \pi_C(x_C).$$

There is a deep connection between Markov random fields and Gibbs measures. In particular, under certain regularity conditions, the joint distribution of $X_1, \ldots, X_n$ is given by a Markov random field iff the joint distribution of $X_1, \ldots, X_n$ is a Gibbs measure, whose underlying graph is defined by the neighborhood structure of the given Markov random field. The celebrated Hammersley-Clifford Theorem, as well as several follow-up works, make these regularity conditions explicit [65],[62],[13],[59].

1.2.1. Higher order Markov random fields. Note that in the hardcore model, non-trivial clique potentials are given only to cliques of size 1 (individual nodes) and 2 (edges), i.e. the underlying Markov random field is limited to nearest-neighbor (pairwise) interactions. In the language of Markov random fields, this corresponds to letting the neighborhood $N_i$ of the binary variable $X_i$ (set to 1 if node $v_i$ is included in the independent set) be the set of neighbors of node $v_i$ in $G$, so that the cliques of the graph associated with the neighborhood system of the Markov random field are the same as the cliques of the underlying graph $G$ (supposing that $G$ is triangle-free). However, such measures are not capturing the full power of Markov random fields. In particular, the neighborhood $N_i$ of $X_i$ can be any set of indices (not just those corresponding to edges in the underlying graph $G$), so long as all the clique potential functions evaluate to 0 on any configurations which assign the value 1 to two nodes which are adjacent in $G$. For a graph $G$ and node $v$, let $N_d(v)$ denote the set of all nodes at graph-theoretic distance exactly $d$ from $v$, $N_{\leq d}(v) \triangleq \bigcup_{k \leq d} N_k(v)$ the subgraph induced by those nodes at graph-theoretic distance at most $d$, and $N_{\geq d}(v) \triangleq \bigcup_{k \geq d} N_k(v)$ the subgraph induced by those nodes at graph-theoretic distance at least $d$. We use $N(v)$ as shorthand for $N_1(v)$, the set of neighbors of $v$. To be consistent with the literature [113], let us say that a Markov random field is of the $k$-th order with respect to $G$ if the r.v. $X_i$ associated with node $v_i$ of $G$ has neighborhood $N_i$ equal to the set of indices of nodes belonging to $N_{\leq k}(v_i) \setminus v_i$.

For a graph $G$, recall that the girth of $G$ denotes the length of the smallest cycle in $G$. Note that if $G$ is a $\Delta$-regular graph of girth at least $k + 1$, and $k$ is even, the maximal cliques (i.e. cliques not properly contained in other cliques) of the associated $k$-th order Markov random field are exactly
sets of the form $N_{\leq \frac{k}{2}}(v_i)$, i.e. all the depth-$\frac{k}{2}$ neighborhoods. Furthermore, these sets are all rooted trees in which all non-leaf nodes have degree $\Delta$. In this case, the configuration induced on $N_{\leq \frac{k}{2}}(v_i)$ by a binary vector $x$ may be viewed as a rooted (at $v_i$), depth-$\frac{k}{2}$ 0/1 labeled tree. In addition, we say that the associated set of clique potentials is homogenous and isotropic if for all $i, j$ (where $i$ may equal $j$), and any binary $|G|$-dimensional vectors $x, y$, $\pi_{N_{\leq \frac{k}{2}}(v_i)}(x_{N_{\leq \frac{k}{2}}(v_i)}) = \pi_{N_{\leq \frac{k}{2}}(v_j)}(y_{N_{\leq \frac{k}{2}}(v_j)})$ if the rooted 0/1 labeled tree which $x$ induces on $N_{\leq \frac{k}{2}}(v_i)$ is isomorphic to the rooted 0/1 labeled tree which $y$ induces on $N_{\leq \frac{k}{2}}(v_j)$. In this case, for any fixed $k$ and $\Delta$, the associated Markov random field has a finite-size description, i.e. an assignment of non-negative real numbers to isomorphism classes of depth-$\frac{k}{2}$ rooted 0/1 labeled trees in which all non-leaf nodes have degree $\Delta$.

We note that some special cases of higher order homogenous isotropic Markov random fields have already been considered for studying independent sets and generalizations of the hardcore model [99]. This includes work in the statistical physics community which incorporates next-nearest-neighbor and/or competing interactions [44],[77],[118],[58],[57],[2],[3],[30],[8]; work which studies so-called kinetically constrained spin models (e.g. the Kob-Andersen model) [17],[28],[94],[73],[119],[47] and geometrically constrained spin models (e.g. the Biroli-Mezard model) [15],[98],[117],[76], in which a hard density constraint forbids configurations for which a particle has more than some fixed number of occupied neighboring sites; and work on models assigning positive probability only to maximal independent sets, in which every excluded node is incident to at least one included node [35], as well as related combinatorial constraints [86]. Other related work comes from the field of stochastic geometry, in which the distribution of points in space depends on the local geometry of those points [63],[39], as well as stochastic models arising in communication networks [87],[68] and other areas [23],[93]. Related models in which activities are assigned to more general hyperedges are also closely related to this framework [50],[83],[87],[116],[91],[67],[42],[79],[102],[39],[114],[51].

1.3. Our contribution. In this paper, we provide partial answers to Question 2. We observe that higher order Markov random fields are well-suited to this task, and investigate the properties of these models. For the family of so-called reverse ultra log-concave distributions, which includes the truncated Poisson and geometric, we give necessary and sufficient conditions for the natural higher order Markov random field which induces the desired distribution to be in the uniqueness regime in large regular graphs of large girth, in terms of the set of solutions to a certain system of equations. We also show that these Markov random fields undergo a phase transition, and give explicit bounds on the associated critical activity, which we prove to exhibit a certain robustness. For distributions which are small perturbations around the binomial distribution realized by the hardcore model at critical activity, we give a description of the corresponding uniqueness regime in terms of a simple polyhedral cone. Our analysis reveals an interesting non-monotonicity with regards to biasing towards excluded nodes with no included neighbors. We conclude with a broader discussion of the potential use of higher order Markov random fields for analyzing independent sets in graphs.

1.4. Outline of paper. The rest of the paper proceeds as follows. In Section 2, we formally define all relevant terms and state our main results. In Section 3, we rephrase the probabilities and questions of interest in terms of the relevant partition functions and associated recursions, formally relate higher order Markov random fields to Question 2, and prove our necessary and sufficient conditions for uniqueness under a log-convexity assumption. In Section 4, we prove that the associated higher order Markov random fields undergo a phase transition, and provide explicit bounds on the
critical activity. In Section 5, we give a description (as a polyhedral cone) of the corresponding uniqueness regime for distributions which are small perturbations around the binomial distribution realized by the hardcore model at critical activity. In Section 6, we summarize our main results, provide a broader discussion of the potential use of higher order Markov random fields for analyzing independent sets in graphs, and present directions for future research.

2. Main Results.

2.1. Preliminary definitions and notations.

2.1.1. Relevant Gibbs measures. We now formally define the relevant Gibbs measures, which correspond to second-order homogenous isotropic Markov random fields. As our primary interest will be in regular graphs of large girth, we will always have a fixed reference degree $\Delta$, which will often be implicit. For a fixed activity $\lambda \in \mathbb{R}^+$, and vector $\theta = (\theta_0, \ldots, \theta_\Delta) \in \mathbb{R}^{\Delta+1}$ (strict positivity of both $\lambda$ and $\theta$ is assumed throughout), input graph $G$, node $v \in V$, and independent set $I \in \mathcal{I}(G)$, let

$$\phi_{\lambda, \theta, G}(v, I) \Delta \begin{cases} 
\lambda & \text{if } |N(v)| = \Delta, v \in I, \\
\theta_{|N(v) \cap I|} & \text{if } |N(v)| = \Delta, v \notin I, \\
1 & \text{otherwise;}
\end{cases}$$

and

$$w_{\lambda, \theta, G}(I) \overset{\Delta}{=} \prod_{v \in V} \phi_{\lambda, \theta, G}(v, I);$$

and

$$Z_{\lambda, \theta, G} \overset{\Delta}{=} \sum_{I \in \mathcal{I}(G)} w_{\lambda, \theta, G}(I).$$

Let the Gibbs measure $P_{\lambda, \theta, G}$ denote the probability measure, on $\mathcal{I}(G)$, such that for all $I \in \mathcal{I}(G)$,

$$P_{\lambda, \theta, G}(I) = \frac{w_{\lambda, \theta, G}(I)}{Z_{\lambda, \theta, G}}.$$

2.1.2. (non)Unique infinite-volume Gibbs measures. We now formally define the relevant notions of uniqueness / long-range boundary independence which we will use throughout. We begin by introducing some additional notation, to facilitate describing the measure $P_{\lambda, \theta, G}$ under various conditionings. For an independent set $I \in \mathcal{I}(G)$ and set of nodes $U \in 2^V$, let $I_U$ be the inclusion/exclusion pattern (with respect to $I$) of the nodes belonging to $U$. We note that when referring to boundary conditions, e.g. conditioning on the event $\{I_U = B\}$, and taking associated infima and suprema, we will implicitly restrict ourselves to those boundary conditions with strictly positive probability. Recall that $T_d$ denotes the rooted (at $r$) depth $d$ tree in which all non-leaf nodes have degree $\Delta$. For an independent set $I$ of a rooted depth $d$ tree $T$, let $\partial I \overset{\Delta}{=} I_{N(d-1)}(r)$, i.e. the boundary condition induced on the bottom two layers of the tree. Then we will define long-range boundary dependence in terms of whether

$$\liminf_{d \to \infty} \inf_{B} \mathbb{P}_{\lambda, \theta, T_d} \left( r \notin I, |N(r) \cap I| = k \right) \partial I = B \right)$$

equals

$$\limsup_{d \to \infty} \sup_{B} \mathbb{P}_{\lambda, \theta, T_d} \left( r \notin I, |N(r) \cap I| = k \right) \partial I = B \right)$$.
When these limits indeed coincide for all \( k \in \{0, \ldots, \Delta\} \), we denote the relevant limits as \( p^k_{\lambda, \theta} \), and the associated vector as \( \mathbf{p}^{\lambda, \theta} \). We also let \( p^\Delta_{+} = 1 - \mathbf{p}^{\lambda, \theta} \cdot \mathbf{1} \) denote the corresponding limit of the probability that the root is included.

To be consistent with the literature \([60],[120]\), we formally define uniqueness / long-range boundary independence as follows.

**Definition 1.** (Unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree) We say that the vector \((\lambda, \theta)\) admits a unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree iff for all \( k \in \{0, \ldots, \Delta\} \), (2) equals (3).

We note that (non) existence of a unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree has an analogous interpretation in terms of how the associated measure behaves on any regular graph of sufficiently large girth. In particular, we observe the following, which may be derived from straightforward conditioning arguments.

**Observation 2.** Suppose \((\lambda, \theta)\) admits a unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree. Then for all \( \epsilon > 0 \), there exists \( g_{\epsilon, \lambda, \theta} \) (depending only on \( \epsilon, \lambda, \theta \)) such that for any finite \( \Delta \)-regular graph \( G \) of girth at least \( g_{\epsilon, \lambda, \theta} \), node \( v \in G \), and \( k \in \{0, \ldots, \Delta\} \),

\[
\left| \mathbb{P}_{\lambda, \theta, G} \left( v \notin I, |N(v) \cap I| = k \right) - p^k_{\lambda, \theta} \right| < \epsilon,
\]

and

\[
\sup_{\mathcal{B}} \mathbb{P}_{\lambda, \theta, G} \left( v \notin I, |N(v) \cap I| = k \left| I_{N \geq \left\lfloor \frac{1}{2} (g_{\epsilon, \lambda, \theta} - 3) \right\rfloor (v) = \mathcal{B} \right) \right) - \inf_{\mathcal{B}} \mathbb{P}_{\lambda, \theta, G} \left( v \notin I, |N(v) \cap I| = k \left| I_{N \geq \left\lfloor \frac{1}{2} (g_{\epsilon, \lambda, \theta} - 3) \right\rfloor (v) = \mathcal{B} \right) \right) < \epsilon.
\]

Alternatively, suppose \((\lambda, \theta)\) does not admit a unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree. Then there exists \( \epsilon_{\lambda, \theta} > 0 \), a strictly increasing sequence of positive integers \( \{d_{\lambda, \theta}, i \geq 1\} \) (depending only on \( \lambda, \theta \)), and \( k \in \{0, \ldots, \Delta\} \), such that the following is true. For every \( i \geq 1 \), any finite \( \Delta \)-regular graph \( G \) of girth at least \( 2d_{\lambda, \theta} + 3 \), and all nodes \( v \in G \),

\[
\sup_{\mathcal{B}} \mathbb{P}_{\lambda, \theta, G} \left( v \notin I, |N(v) \cap I| = k \left| I_{N \geq d_{\lambda, \theta} (v) = \mathcal{B} \right) \right) - \inf_{\mathcal{B}} \mathbb{P}_{\lambda, \theta, G} \left( v \notin I, |N(v) \cap I| = k \left| I_{N \geq d_{\lambda, \theta} (v) = \mathcal{B} \right) \right) \geq \epsilon_{\lambda, \theta}.
\]

We note that for closely related models, further connections to regular graphs of large girth have been shown \([108],[89]\), although we do not pursue that here.

### 2.2. Main Results

We now state our main results, which provide partial answers to Question 2. We begin by formally relating the occupancy probabilities attained when one samples from second-order homogenous isotropic Markov random fields in the uniqueness regime to the corresponding vector \( \theta \).
Observation 3. If \((\lambda, \theta)\) admits a unique infinite-volume Gibbs measure on the infinite \(\Delta\)-regular tree, then the probability measure \(\mu\) such that
\[
\mu(k) = (p^\lambda \theta \cdot 1)^{-1} p_k^\lambda \theta, \quad k \in \{0, \ldots, \Delta\}
\]
satisfies
\[
\frac{\mu^2(k)}{\mu(k-1)\mu(k+1)} = \frac{\theta_k^2}{\theta_{k-1} \theta_{k+1}} \frac{(k+1)(\Delta+1-k)}{k(\Delta-k)}, \quad k \in \{1, \ldots, \Delta-1\}.
\]
Equivalently, there exist \(c, x \in \mathbb{R}^+\) (depending on \(\theta\) and \(\lambda\)) such that
\[
\mu(k) = \theta_k (\frac{\Delta}{k}) x^k \text{ for } k \in \{0, \ldots, \Delta\}.
\]

Observation 3 shows that the family of second-order homogenous isotropic Markov random fields provides a natural framework for studying Question 2. We now give several illustrative examples showing that for certain natural choices of \(\theta\), the induced measure \(\mu\) corresponds exactly to a well-known family of distributions.

Example 1. (i) Binomial distribution. If \(\theta = 1\), then \(\mu\) corresponds to a \(B(\Delta, p)\) distribution for some \(p \in (0, 1)\).
(ii) Truncated Poisson distribution. If \(\theta_k = \frac{1}{k! (\Delta^k)}\) for \(k \in \{0, \ldots, \Delta\}\), then \(\mu\) corresponds to a truncated Poisson distribution, i.e.,
\[
\mu(k) = \frac{\theta_k}{x^k}, \quad k \in \{0, \ldots, \Delta\}
\]
for some \(c, x > 0\).
(iii) Truncated geometric distribution. If \(\theta_k = (\frac{\Delta}{k})^{-1}\) for \(k \in \{0, \ldots, \Delta\}\), then \(\mu\) corresponds to a truncated geometric distribution, i.e.,
\[
\mu(k) = cx^k, \quad k \in \{0, \ldots, \Delta\}
\]
for some \(c, x > 0\).

We note that several previous works in the literature on Markov random fields similarly examine how modifying the relevant potentials / activities can change various distributions of interest, for models different from those considered here [83],[102].

We now provide necessary and sufficient conditions for uniqueness under a log-convexity assumption on \(\theta\). 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 \(\{\frac{x_i}{x_i}, i = 0, \ldots, n\}\) is log-convex [31], and convex if \(x_{i+1} - x_i \geq x_i - x_{i-1}\) for all \(i \in \{1, \ldots, n-1\}\). With a slight abuse of notation, we say that a measure \(\mu\) 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. The following may be easily verified using (4) [31].

Observation 4. If \((\lambda, \theta)\) admits a unique infinite-volume Gibbs measure on the infinite \(\Delta\)-regular tree, then the probability measure \(\mu\) such that
\[
\mu(k) = (p^\lambda \theta \cdot 1)^{-1} p_k^\lambda \theta, \quad k \in \{0, \ldots, \Delta\}
\]
is reverse ultra log-concave iff \(\theta\) is log-convex.

We note that several well-known distributions satisfy reverse ultra log-concavity, and arise from a natural choice of log-convex \(\theta\). In particular, one may easily verify the following.

Observation 5. Every distribution considered in Example 1 is reverse ultra log-concave, with the corresponding choice of \(\theta\) log-convex.
For a given vector $\theta$, let
\[
f_\theta(x) \triangleq \sum_{k=0}^{\Delta-1} \frac{\theta_{k+1} (\Delta-1)}{\Delta} x^k,
\]
and
\[
g_\theta(x) \triangleq \left( \sum_{k=0}^{\Delta-1} \theta_k \frac{\Delta-1}{k} x^k \right)^{-1}.
\]
Then our necessary and sufficient conditions for uniqueness are as follows.

**Theorem 1.** The system of equations
\[
(5) \quad x = \lambda g_\theta(y) f_\theta^\Delta(x);
\]
\[
(6) \quad y = \lambda g_\theta(x) f_\theta^\Delta(y);
\]
always has at least one non-negative solution on $\mathbb{R}^+ \times \mathbb{R}^+$. If $\theta$ is log-convex, then $(\lambda, \theta)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree iff this solution is unique.

Note that Theorem 1 reduces to the known characterization for uniqueness in the hardcore model for $\theta = 1$.

For log-convex $\theta$, we further show that the boundary condition in which all nodes on the boundary are included is extremal, i.e. if there is long-range boundary dependence, then this boundary condition will induce such a dependence. Let $B_d$ denote the boundary condition with all nodes in $N_d(r)$ included, and all nodes in $N_{d-1}(r)$ excluded.

**Corollary 1.** If $\theta$ is log-convex, then $(\lambda, \theta)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree iff
\[
(7) \quad \liminf_{d \to \infty} \mathbb{P}_{\lambda, \theta, T_d} \left( r \notin \mathcal{I}, |N(r) \cap \mathcal{I}| = k \right) = \limsup_{d \to \infty} \mathbb{P}_{\lambda, \theta, T_d} \left( r \notin \mathcal{I}, |N(r) \cap \mathcal{I}| = k \right)
\]
for all $k \in \{0, \ldots, \Delta\}$.

We note that several related results with regards to extremal boundary conditions appear throughout the literature. This includes results for so-called repulsive first-order Markov random fields [125], as well as fields which satisfy the positive (negative) lattice conditions, and related notions of positive (negative) association, concavity, and van den Berg-Kesten-Reimer (BKR) / Fortuin-Kasteleyn-Ginibre (FKG) inequalities [46],[92],[115],[22],[96],[102],[60],[114],[83],[64]. It remains an interesting open question to explore the connections between our model and such known results.

We now show that the family of second-order homogenous isotropic Markov random fields undergoes a phase transition with respect to the activity $\lambda$. Let
\[
\psi_\theta \overset{\Delta}{=} \max_{k=0, \ldots, \Delta-2} \left( (\Delta - (k+1)) \frac{\theta_{k+1}}{\theta_k} \right),
\]
\[
\Delta_\theta \overset{\Delta}{=} \left( 2 \psi_\theta \theta_0^{-1} \left( \frac{\theta_\Delta}{\theta_{\Delta-1}} \right)^{\Delta-2} \left( \frac{\theta_\Delta}{\theta_{\Delta-2}} \frac{\theta_{\Delta-1}}{\theta_0} \right) + (\Delta - 1) \left( \frac{\theta_\Delta}{\theta_{\Delta-1}} - \frac{\theta_1}{\theta_0} \right) \right)^{-1},
\]
and

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

Then we prove that the family of second-order homogenous isotropic Markov random fields undergoes a phase transition in the following sense.

**Theorem 2.** For any fixed log-convex vector $\theta$, $(\lambda, \theta)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for all activities $\lambda < \overline{\lambda}_\theta$, and does not admit a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for all activities $\lambda > \overline{\lambda}_\theta$.

Whether or not this phase transition is sharp remains an interesting open question. We now comment briefly on several implications of Theorem 2, all of which follow from Theorem 2 and straightforward algebraic manipulations of $\Delta \theta, \overline{\lambda}_\theta$. We first show that the critical activity for the models considered exhibits a certain form of robustness.

**Observation 6.** If there exists $c \in [0, \Delta]$ such that $\max \left( \frac{\theta_0}{\theta_{\Delta-1}}, \frac{\theta_0}{\theta_1} \right) \leq 1 + c \Delta$, then $\Delta \theta \geq \left(2 \exp(c)(1 + 5c)\right)^{-1} \frac{\theta_0}{\Delta}$, and $\overline{\lambda}_\theta \leq 3 \exp(12 + c) \frac{\theta_0}{\Delta}$. Applying these bounds to the hardcore model, i.e. $\theta = 1, c = 0, \theta_0 = 1$, we find that $\frac{1}{2} \Delta^{-1} \leq \Delta_1 \leq \overline{\lambda}_1 \leq 3 \exp(12) \Delta^{-1}$. As it is easily verified that $\lim_{\Delta \to \infty} \Delta \lambda_\Delta = e$, we note that in this setting, our bounds are correct up to constant factors. Furthermore, the bounds of Theorem 2 show that (up to constant factors) the critical activity will scale like $\theta_0 \Delta^{-1}$ for any vector $\theta$ which does not deviate too much from the all ones vector.

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

**Observation 7.** For $\theta$ such that $\theta_k = \frac{1}{k! \left( \frac{\Delta}{k} \right)}$, $k \in \{0, \ldots, \Delta\}$, one has that $\Delta \theta \geq \frac{1}{2} \Delta^{-1}$. In particular, the critical activity is at least $\frac{1}{2} \Delta^{-1}$, as in the hardcore model.

A more precise understanding of how the critical activity scales as $\Delta \to \infty$ remains an interesting open question.

An explicit description of when Equations 5 - 6 have a unique non-negative 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 distribution for the number of included neighbors of an excluded node is a small perturbation of the $B(\Delta, (\Delta - 1)^{-1})$ distribution achieved by the hardcore model at the critical activity $\lambda_\Delta$, i.e. $\theta$ is a small perturbation around the all ones vector. Let us fix a vector $c = (c_0, \ldots, c_\Delta)$. It follows from a simple Taylor series expansion that convex perturbations of the all ones vector yield log-convex $\theta$. In particular, one may easily verify the following.

**Observation 8.** There exists $\epsilon_c > 0$ such that for $h \in (0, \epsilon_c)$, $1 + ch$ is log-convex iff $c$ 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 $c$ is a direction of uniqueness iff there exists $\epsilon_c > 0$ such that for all $h \in (0, \epsilon_c)$, $(\lambda_\Delta, 1 + ch)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree; and a direction of non-uniqueness iff there exists $\epsilon_c > 0$ such that for all $h \in (0, \epsilon_c)$, $(\lambda_\Delta, 1 + ch)$ does not admit a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree.

We now provide an explicit characterization / dichotomy theorem, classifying (almost) all convex vectors as either directions of uniqueness or directions of non-uniqueness. For $j \in \{0, \ldots, \Delta\}$, let

$$\Lambda_{\Delta,j} \triangleq \binom{\Delta}{j} (\Delta - 2)^{-j}.$$

Let $\pi = (\pi_0, \ldots, \pi_\Delta)$ denote the vector such that for $j \in \{0, \ldots, \Delta\}$,

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

Then we prove the following.

Theorem 3. A convex vector $c$ is a direction of uniqueness if $\pi \cdot c < 0$, and a direction of non-uniqueness if $\pi \cdot c > 0$.

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

We now study some qualitative features of $\pi$, to shed light on the set of convex directions of uniqueness, and reveal an interesting non-monotonicity of the uniqueness regime.

Observation 9. For all $\Delta \geq 3$, $\pi_0 > 0$, $\pi_1 < 0$, $\pi_2 < 0$, and $\pi_k > 0$ for all $k \in \{3, \ldots, \Delta\}$.

That $\pi_1 < 0$, $\pi_2 < 0$, and $\pi_k > 0$ for all $k \in \{3, \ldots, \Delta\}$ makes sense at an intuitive level, since biasing towards 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 $\lambda_\Delta$, is $1 + (\Delta - 1)^{-1} \in (1, 2)$.

The counterintuitive feature of Observation 9, which seems to violate the above reasoning, is that $\pi_0 > 0$, i.e. biasing towards excluded nodes with no included neighbors leads to non-uniqueness. We note that this effect is perhaps especially surprising in light of Theorem 2, as we now explain. Let $e_0$ denote the $(\Delta + 1)$-dimensional vector whose first component is a 1, with all remaining components 0. As it is easily verified that $1 + e_0 h$ is log-convex for all $h \geq 0$, and $\lim_{h \to \infty} \frac{\Delta}{1 + e_0 h} = \infty$, we conclude that the associated uniqueness regime exhibits the following non-monotonicity.

Corollary 2. For all $\Delta \geq 3$, there exist strictly positive finite constants $a_\Delta < b_\Delta$ such that $(\lambda_\Delta, 1 + e_0 h)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for $h = 0$ and $h \geq b_\Delta$, and does not admit a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for $h \in (0, a_\Delta)$.

In particular, biasing a small amount towards excluded nodes with no included neighbors leads to non-uniqueness, while biasing a large amount towards excluded nodes with no included neighbors leads to uniqueness. We note that several previous works in the literature on Markov random fields examine various notions of non-monotonicity [83], and better understanding the relevant (non) monotonicities with regards to higher order Markov random fields remains an interesting open question.
3. Probabilities, partition functions, and proof of Theorem 1. In this section, we rephrase the probabilities and questions of interest in terms of the relevant partition functions. We also complete the proofs of Observation 3, Theorem 1, and Corollary 1.

3.1. Probabilities as partition functions. For $i \in \{1, \ldots, \Delta\}$, let $T^i_d$ denote the subtree of $T_d$ consisting of $r$, and the subtree rooted at the $i$th child of $r$. Also, for a boundary condition $B$ on $N_{d-1}(r)$ in $T_d$, and $i \in \{1, \ldots, \Delta\}$, let $B^i$ denote the boundary condition which $B$ induces on $N_{d-1}(r)$ in $T^i_d$. Since $\{T^i_d, i = 1, \ldots, \Delta\}$ are all isomorphic, let us denote a generic version of this depth $d$ rooted tree by $T_d$. We denote the root of $T^i_d$ (formerly $r$) by $v_0$, and denote the single child of $v_0$ (formerly a child of $r$) by $v_1$. For a binary vector $x$, let $|x|$ denote $\sum x_i$. For $i, j \in \{0, 1\}$, and a boundary condition $B$ on $N_{d-1}(v_0)$ in $T^i_d$, let $Z_{\lambda, \theta, d}(i, j, B) \triangleq \sum_{r \in I(T^i_d) \setminus \partial I = B} W_{\lambda, \theta, d}(r | N(r))$. Then it follows from the basic properties of independent sets that

$$\mathbb{P}_{\lambda, \theta, T_d}(r \notin I, |N(r) \cap I| = k | \partial I = B)$$

equals

$$\sum_{I \in I(T^i_d) \setminus \partial I = B} W_{\lambda, \theta, d}(I)$$

$$= \sum_{I \in I(T^i_d) \setminus \partial I = B} W_{\lambda, \theta, d}(I) + \sum_{i=0}^{\Delta} \sum_{I \in I(T^i_d) \setminus \partial I = B} W_{\lambda, \theta, d}(I)$$

which is itself equal to

$$\frac{\theta_k \sum_{x \in \{0, 1\}^\Delta} \prod_{i=1}^{\Delta} Z_{\lambda, \theta, d}(0, x^i)}{\lambda \prod_{i=1}^{\Delta} Z_{\lambda, \theta, d}(1, \mathbb{B}^i) + \sum_{i=0}^{\Delta} \theta_i \sum_{x \in \{0, 1\}^\Delta} \prod_{j=1}^{\Delta} Z_{\lambda, \theta, d}(0, x^j)}$$

(8)

where we denote by $p_{\lambda, \theta, d}(B)$. We also let $p_{\lambda, \theta, d}(B)$ denote the associated vector, and $p_{\lambda, \theta, d}(B) \triangleq 1 - p_{\lambda, \theta, d}(B) \cdot 1$ denote the corresponding probability that the root is included. We now derive several recursions for $Z_{\lambda, \theta, d}(i, B)$, to aid in our analysis. First, it will be useful to define multi-dimensional analogues of $f_{\theta, \theta, d}$, to help in deriving the relevant recursions under non-uniform boundary conditions. For $k, n \in \mathbb{Z}^+$ such that $n \geq k$, and a vector $x \in \mathbb{R}^n$, we let $\sigma_k(x)$ denote the $k$th elementary symmetric polynomial on $n$ variables evaluated at $x$. Namely, $\sigma_0(x) = 1$, and

$$\sigma_k(x) = \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \prod_{j=1}^{k} x_{i_j} = \sum_{S \in \mathbb{G}_{k, n}} \prod_{i \in S} x_i.$$
For example, $\sigma_2(x, y, z) = xy + xz + yz$. Also, for $x \in \mathbb{R}^+ (\Delta - 1)$, let

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

and

$$g_\theta(x) \triangleq \left( \sum_{k=0}^{\Delta-1} \theta_k \sigma_k(x) \right)^{-1}.$$

Let $\zeta_{\lambda, \theta, d}(B) \triangleq Z_{\lambda, \theta, d}^{-1}(0, B)$. For $i, j \in \{1, \ldots, \Delta - 1\}$, let $B^{ij}$ denote the boundary condition which $B$ induces on $N_{\geq d-1}(v_0)$ in the subtree of $T'_d$ consisting of $v_0$, $v_1$, the $i$th child of $v_1$, and the $j$th child of the $i$th child of $v_1$. Then we prove the following.

**Lemma 1.** For all $d \geq 5$ and boundary conditions $B$ on $N_{\geq d-1}(v_0)$ in $T'_d$,

$$Z_{\lambda, \theta, d}(1, B) = Z_{\lambda, \theta, d}(0, B) f_\theta(\zeta_{\lambda, \theta, d-1}(B^1), \zeta_{\lambda, \theta, d-1}(B^2), \ldots, \zeta_{\lambda, \theta, d-1}(B^{\Delta-1})),$$

and

$$\zeta_{\lambda, \theta, d}(B) = \lambda g_\theta(\zeta_{\lambda, \theta, d-1}(B^1), \zeta_{\lambda, \theta, d-1}(B^2), \ldots, \zeta_{\lambda, \theta, d-1}(B^{\Delta-1}))$$

$$\times \prod_{i=1}^{\Delta-1} f_\theta(\zeta_{\lambda, \theta, d-2}(B^{i-1}), \zeta_{\lambda, \theta, d-2}(B^{i,2}), \ldots, \zeta_{\lambda, \theta, d-2}(B^{i,\Delta-1})).$$

**Proof.** Note that for $i \in \{0, 1\}$,

$$Z_{\lambda, \theta, d}(i, 0, B) = \sum_{x \in \{0, 1\}^{\Delta-1}} \theta_{|x|+i} \prod_{j=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(0, x_j, B^j),$$

and

$$Z_{\lambda, \theta, d}(0, 1, B) = \lambda \prod_{i=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(1, 0, B^i).$$

Thus

$$Z_{\lambda, \theta, d}(0, B) = \sum_{x \in \{0, 1\}^{\Delta-1}} \theta_{|x|} \prod_{i=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(0, x_i, B^i) / \lambda^{\prod_{i=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(1, 0, B^i)}$$

$$= \lambda^{-\prod_{i=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(0, B^i)} \sum_{x \in \{0, 1\}^{\Delta-1}} \theta_{|x|} \prod_{i=1}^{\Delta-1} \zeta_{\lambda, \theta, d-1}(B^i).$$

Similarly,

$$Z_{\lambda, \theta, d}(1, B) = \lambda^{-\prod_{i=1}^{\Delta-1} Z_{\lambda, \theta, d-1}(0, B^i)} \sum_{x \in \{0, 1\}^{\Delta-1}} \theta_{|x|+1} \prod_{i=1}^{\Delta-1} \zeta_{\lambda, \theta, d-1}(B^i).$$

Combining the above with the definition of $f_\theta$ and $g_\theta$ completes the proof. \qed
3.2. Bounding sequences for \( \sup_B \zeta_{\lambda, \theta, d}(B) \) and \( \inf_B \zeta_{\lambda, \theta, d}(B) \). We now construct bounding sequences for \( \sup_B \zeta_{\lambda, \theta, d}(B) \) and \( \inf_B \zeta_{\lambda, \theta, d}(B) \) when \( \theta \) is log-convex. We will then argue that the upper bounding sequence and the lower bounding sequence are monotone, and thus have limits, which must satisfy the system of equations (5) - (6). If these equations have a unique fixed point, then these upper and lower limits must coincide, proving one direction of Theorem 1. We note that the idea of using bounding sequences to show uniqueness of Gibbs measures has appeared before in the literature, e.g. [50].

Let \( \zeta_{\lambda, \theta, 5} \triangleq 0, \zeta_{\lambda, \theta, 5} \triangleq \lambda \theta_0^{-1}(\theta_{\Delta-1} \theta_{\Delta})^{\Delta-1}, \zeta_{\lambda, \theta, 6} \triangleq \zeta_{\lambda, \theta, 5} \lambda, \zeta_{\lambda, \theta, 6} \triangleq \zeta_{\lambda, \theta, 5}. \) For \( d \geq 7 \), let

\[
\zeta_{\lambda, \theta, d} \triangleq \lambda g_\theta(\zeta_{\lambda, \theta, d-1}) \Delta_{\theta, d-1},
\]

and

\[
\bar{\zeta}_{\lambda, \theta, d} \triangleq \lambda g_\theta(\bar{\zeta}_{\lambda, \theta, d-1}) \Delta_{\theta, d-1}.
\]

We now prove that \( \{\zeta_{\lambda, \theta, d}\} \) and \( \{\bar{\zeta}_{\lambda, \theta, d}\} \) provide lower (upper) bounds for \( \{\zeta_{\lambda, \theta, d}\} \).

**Lemma 2.** Suppose \( \theta \) is log-convex. Then for all \( d \geq 5 \),

\[
\zeta_{\lambda, \theta, d} \leq \inf_B \zeta_{\lambda, \theta, d}(B) \leq \sup_B \zeta_{\lambda, \theta, d}(B) \leq \bar{\zeta}_{\lambda, \theta, d}.
\]

First, it will be useful to prove that when \( \theta \) is log-convex, \( f_\theta \) is a monotone increasing function.

**Lemma 3.** If \( \theta \) is log-convex, then \( \partial_x f_\theta(x) \geq 0 \) for all \( i \in \{1, \ldots, \Delta - 1\} \) and \( x \geq 0 \).

**Proof.** Let us fix some \( i \in \{1, \ldots, \Delta - 1\} \). Then \( \partial_x f_\theta(x) \) will be the same sign as

\[
\sum_{k=0}^{\Delta-1} \theta_k \sigma_k(x) \partial_x \sum_{k=0}^{\Delta-1} \theta_{k+1} \sigma_k(x) - \sum_{k=0}^{\Delta-1} \theta_k \sigma_k(x) \partial_x \sum_{k=0}^{\Delta-1} \theta_{k+1} \sigma_k(x).
\]

Let \( x^i \) denote the \((\Delta - 2)\)-dimensional vector equivalent to \( x \), but with component \( i \) removed. Note that \( \partial_x \sigma_0(x) = 0 \), and for \( k \geq 1 \), \( \partial_x \sigma_k(x) = \sigma_{k-1}(x^i) \). Thus (9) equals

\[
\sum_{j=0}^{\Delta-1} \theta_j \sigma_j(x) \sum_{k=0}^{\Delta-1} \theta_{k+1} \sigma_{k-1}(x^i) - \sum_{j=0}^{\Delta-1} \theta_{j+1} \sigma_j(x) \sum_{k=0}^{\Delta-1} \theta_k \sigma_{k-1}(x^i)
\]

\[
= \left( \sum_{j=0}^{\Delta-1} \theta_j \sum_{|S|=j} \prod_{l \in S} x_l \right) \left( \sum_{j=0}^{\Delta-2} \theta_{j+1} \sum_{|S|=j} \prod_{l \in S} x_l \right)
\]

\[
- \left( \sum_{j=0}^{\Delta-1} \theta_{j+1} \sum_{|S|=j} \prod_{l \in S} x_l \right) \left( \sum_{j=0}^{\Delta-2} \theta_j \sum_{|S|=j} \prod_{l \in S} x_l \right),
\]

14
which itself equals

\[(\sum_{j=0}^{\Delta-2} \theta_j \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) (\sum_{j=0}^{\Delta-2} \theta_{j+2} \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) \]

\[+x_i (\sum_{j=0}^{\Delta-2} \theta_{j+1} \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) (\sum_{j=0}^{\Delta-2} \theta_{j+2} \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) \]

\[-x_i (\sum_{j=0}^{\Delta-2} \theta_{j+1} \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) (\sum_{j=0}^{\Delta-2} \theta_{j+2} \sum_{S \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \prod_{l \in S} x_l) \]

It follows from the definition of \(\sigma_j(x)\) that (10) is a polynomial function of \(x_1, \ldots, x_i-1, x_i+1, \ldots, x_{\Delta-1}\), with degree at most two in any given variable. Equivalently, there exists a unique set of finite constants \(\{u_{S_1, S_2}, S_1, S_2 \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}, S_1 \cap S_2 = \emptyset\}\) such that (10) equals

\[u_{S_1, S_2} \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2.\]

We now compute \(u_{S_1, S_2}\). It follows from the “product of sums” form of (10) that

\[u_{S_1, S_2} = \sum_{A, B \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \theta_{\|A\| \theta_{\|B\|+2}}.\]

For any sets \(A, B, \|A\| + \|B\| - \|A \cap B\| = \|A \cup B\|.\) Thus if \(A \cap B = S_2\), and \(A \cup B = S_1 \cup S_2\), it follows that

\[\|A\| + \|B\| - \|S_2\| = \|S_1 \cup S_2\| = \|S_1\| + \|S_2\| - \|S_1 \cap S_2\| = \|S_1\| + \|S_2\|,\]

since \(S_1 \cap S_2 = \emptyset\). Equivalently,

\[\|B\| = \|S_1\| + 2\|S_2\| - \|A\|,\]

Combining with the fact that for each \(A \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}\) such that \(S_2 \subseteq A\), there is exactly one set \(B \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}\) satisfying \(A \cap B = S_2\) and \(A \cup B = S_1\), i.e. \(B = S_2 \cup (S_1 \setminus A)\), we find that

\[u_{S_1, S_2} = \sum_{A \in 2^{\{1, \ldots, \Delta-1\} \setminus \emptyset}} \theta_{\|A\| \theta_{\|S_1\|+2\|S_2\|-\|A\|+2}}\]

\[= \sum_{j = |S_2|}^{|S_1|+|S_2|} \left(\begin{array}{c} |S_1| \\ j - |S_2| \end{array}\right) \theta_j \theta_{|S_2|+|S_1|-j+2}.\]
Similarly, (11) equals

\[ \sum_{\substack{S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \ni j \geq |S_2|}} \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2, \]

(12) equals

\[ - \sum_{\substack{S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \ni j \geq |S_2|}} \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2, \]

and (13) equals

\[ - \sum_{\substack{S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \ni j \geq |S_2|}} \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2. \]

For \( S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \) and \( j \in [|S_2|, |S_1| + |S_2|] \), let

\[ z^1_{S_1, S_2}(j) \triangleq \theta_j \theta_2 |S_2| + |S_1| - j + 2 - \theta_j \theta_2 |S_2| + |S_1| - j + 1, \]

and

\[ z^2_{S_1, S_2}(j) \triangleq \theta_j + \theta_2 |S_2| - j + 2 - \theta_j + \theta_2 |S_2| + |S_1| - j + 1. \]

Combining the above, it follows that (9) equals

\[ \sum_{\substack{S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \ni j \geq |S_2|}} \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2 \]

(15)

\[ + \sum_{\substack{S_1, S_2 \in 2^{(1, \ldots, \Delta-1)} \setminus \emptyset \ni j \geq |S_2|}} \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2 \]

(16)

Note that \( z^1_{S_1, S_2}(j) = -z^1_{S_1, S_2}(2 |S_2| + |S_1| - j + 1) \) for all \( j \) such that \( j \in [|S_2|, |S_1| + |S_2|] \) and \( 2 |S_2| + |S_1| - j + 1 \in [|S_2|, |S_1| + |S_2|] \), or equivalently \( j \in [|S_2| + 1, |S_1| + |S_2|] \). Combining with the monotonicity (in \( j \)) of the function \( 2 |S_2| + |S_1| - j + 1 \), the fact that the unique solution to the equation \( 2 |S_2| + |S_1| - x + 1 = x \) is \( x = |S_2| + \frac{|S_1| + 1}{2} \) (which may be non-integer), and the fact that whenever \( |S_2| + \frac{|S_1| + 1}{2} \) is an integer one has that \( z^1_{S_1, S_2}(|S_2| + \frac{|S_1| + 1}{2}) = 0 \), we conclude that the coefficient of \( \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2 \) in (15) equals

\[ z^1_{S_1, S_2}(|S_2|) + \sum_{j = |S_2| + 1}^{\frac{|S_1| + 1}{2}} \left( \left( |S_1| \right)^j |S_2| \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2 \right) \]
Similarly, note that \( z_{S_1,S_2}^2(j) = z_{S_1,S_2}^2(2|S_2| + |S_1| - j) \) for all \( j \) such that \( j \in [|S_2|, |S_1| + |S_2|] \) and \( 2|S_2| + |S_1| - j \in [|S_2|, |S_1| + |S_2|] \), or equivalently \( j \in [|S_2|, |S_1| + |S_2|] \). Combining with the monotonicity (in \( j \)) of the function \( 2|S_2| + |S_1| - j \), the fact that the unique solution to the equation \( 2|S_2| + |S_1| - x = x = |S_2| + \frac{|S_1|}{2} \) (which may be non-integer), and the fact that whenever \( |S_2| + \frac{|S_1|}{2} \) is an integer one has that \( z_{S_1,S_2}^2(\frac{|S_1|}{2}) = 0 \), we conclude that the coefficient of \( x_i \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2 \) in (16) equals

\[
(18) \quad \sum_{j = |S_2|}^{2|S_2| + |S_1| + 1} \left( \left( \frac{|S_1|}{j - |S_2|} \right) - \left( \frac{|S_1|}{j - |S_2| - 1} \right) \right) z_{S_1,S_2}^2(j) = 0.
\]

Combining the above, we find that \( \partial_x f_\theta(x) \) is the same sign as

\[
(19) \quad \sum_{S_1,S_2 \in 2^{\mathcal{I}} : S_1 \cap S_2 = \emptyset} \left( \left( \frac{|S_1|}{j - |S_2|} \right) - \left( \frac{|S_1|}{j - |S_2| - 1} \right) \right) z_{S_1,S_2}^2(j) \prod_{j \in S_1} x_j \prod_{j \in S_2} x_j^2.
\]

Note that for all \( j \in [|S_2|, |S_2| + \lfloor \frac{|S_1| + 1}{2} \rfloor] \), one has that \( 2|S_2| + |S_1| - j + 1 \geq j \). Thus the definition of \( z_{S_1,S_2}^1 \) implies that for all \( j \in [|S_2|, |S_2| + \lfloor \frac{|S_1| + 1}{2} \rfloor] \), there exists \( j' \geq j \) such that \( z_{S_1,S_2}^1(j) = \theta_j \theta_{j'+1} - \theta_j \theta_{j'} \). It then follows from the definition of log-convexity that if \( \theta \) is log-convex, then \( z_{S_1,S_2}^1(j) \geq 0 \) for all \( j \in [|S_2|, |S_2| + \lfloor \frac{|S_1| + 1}{2} \rfloor] \). Combining with (19) and the well-known monotonicity properties of the binomial coefficients, i.e. that for all \( k \leq \lfloor \frac{|S_1| + 1}{2} \rfloor \), one has that

\[
(20) \quad \left( \frac{|S_1|}{k} \right) - \left( \frac{|S_1|}{k - 1} \right) \geq 0,
\]

demonstrates that the log-convexity of \( \theta \) implies the desired monotonicity of \( f \). \( \square \)

With Lemma 3 in hand, we now complete the proof of Lemma 2.

**Proof of Lemma 2.** It follows from the monotonicity of \( g_\theta \) that for all \( x \geq 0 \),

\[
(21) \quad g_\theta(x) \leq \theta_0^{-1}.
\]

Also, since

\[
f_\theta(x) = \frac{\sum_{k=0}^{\Delta-1} \theta_{k+1} \sigma_k(x)}{\sum_{k=0}^{\Delta-1} \theta_k \sigma_k(x)} \quad = \quad \frac{\sum_{k=0}^{\Delta-1} \theta_{k+1} \theta_k \sigma_k(x)}{\sum_{k=0}^{\Delta-1} \theta_k \sigma_k(x)}.
\]

17
we conclude that \( f_\theta(x) \) is a convex combination of \( \{\frac{\theta_{i+1}}{\theta_i}, i = 0, \ldots, \Delta - 1\} \), and for all \( x \geq 0 \),

(22) \[
\frac{\theta_1}{\theta_0} \leq f(x) \leq \frac{\theta_\Delta}{\theta_{\Delta-1}}.
\]

With (21)-(22) in hand, we proceed by induction. The base cases \( d = 5, 6 \) follows from non-negativity, Lemma 1, and (21)-(22). Now, suppose the induction is true for all \( k \in \{5, \ldots, d - 1\} \), for some \( d \geq 7 \). Then it follows from Lemma 1, Lemma 3, and the monotonicity of \( g_\theta \) that \( \zeta_{\lambda, \theta, d} \leq \inf_B \zeta_{\lambda, \theta, d}(B) \), and \( \zeta_{\lambda, \theta, d} \geq \sup_B \zeta_{\lambda, \theta, d}(B) \), completing the proof. \( \square \)

We now prove that the above bounding sequences themselves have limits.

**Lemma 4.** Suppose \( \theta \) is log-convex. Then \( \{\zeta_{\lambda, \theta, d}, d \geq 5\} \) is monotone increasing, and \( \{\overline{\zeta}_{\lambda, \theta, d}, d \geq 5\} \) is monotone decreasing.

**Proof.** We proceed by induction, simultaneously on both sequences. The base case, \( d = 5, 6 \), entails demonstrating that \( \zeta_{\lambda, \theta, 5} \leq \zeta_{\lambda, \theta, 6} \leq \zeta_{\lambda, \theta, 7} \) and \( \overline{\zeta}_{\lambda, \theta, 5} \geq \overline{\zeta}_{\lambda, \theta, 6} \geq \overline{\zeta}_{\lambda, \theta, 7} \). That \( \zeta_{\lambda, \theta, 5} \leq \zeta_{\lambda, \theta, 6} \leq \zeta_{\lambda, \theta, 7} \) follows from non-negativity. That \( \zeta_{\lambda, \theta, 5} \leq \overline{\zeta}_{\lambda, \theta, 5} \leq \overline{\zeta}_{\lambda, \theta, 7} \) follows from (21)-(22).

Now, suppose that \( \zeta_{\lambda, \theta, k} \leq \zeta_{\lambda, \theta, k+1}, \overline{\zeta}_{\lambda, \theta, k} \geq \overline{\zeta}_{\lambda, \theta, k+1} \) for all \( k \in \{5, \ldots, d - 1\} \) for some \( d \geq 8 \). Then by the monotonicity of \( f \) and \( g \),

\[
\zeta_{\lambda, \theta, d} = \lambda g_\theta(\zeta_{\lambda, \theta, d-1}) f_\theta^{\Delta-1}(\zeta_{\lambda, \theta, d-2}) \geq \lambda g_\theta(\zeta_{\lambda, \theta, d-2}) f_\theta^{\Delta-1}(\zeta_{\lambda, \theta, d-3}) = \zeta_{\lambda, \theta, d-1}.
\]

Similarly,

\[
\overline{\zeta}_{\lambda, \theta, d} = \lambda g_\theta(\overline{\zeta}_{\lambda, \theta, d-1}) f_\theta^{\Delta-1}(\overline{\zeta}_{\lambda, \theta, d-2}) \leq \lambda g_\theta(\overline{\zeta}_{\lambda, \theta, d-2}) f_\theta^{\Delta-1}(\overline{\zeta}_{\lambda, \theta, d-3}) = \overline{\zeta}_{\lambda, \theta, d-1};
\]

completing the proof. \( \square \)

It follows from Lemma 4 that \( \overline{\zeta}_{\lambda, \theta, \infty} \stackrel{\Delta}{\to} \lim_{d \to \infty} \overline{\zeta}_{\lambda, \theta, d} \) and \( \zeta_{\lambda, \theta, \infty} \stackrel{\Delta}{\to} \lim_{d \to \infty} \zeta_{\lambda, \theta, d} \) both exist. Furthermore, since \( 0 \leq \zeta_{\lambda, \theta, d} \leq \overline{\zeta}_{\lambda, \theta, d} \leq \zeta_{\lambda, \theta, 5} \) for all \( d \geq 5 \), the continuity of \( f_\theta \) and \( g_\theta \) on \( \mathbb{R}^+ (\Delta-1) \) imply the following.

**Observation 10.** \( 0 \leq \zeta_{\lambda, \theta, \infty} \leq \zeta_{\lambda, \theta, \infty} < \infty \), and \( (\zeta_{\lambda, \theta, \infty}, \overline{\zeta}_{\lambda, \theta, \infty}) \) is a solution to the system of equations (5) - (6).

With Observation 10 in hand, we now complete the proof of Theorem 1. We actually prove a slightly stronger statement (in terms of partition functions), from which Observation 3, Theorem 1, and Corollary 1 all immediately follow. Let \( L_\theta(z) \triangleq \sum_{i=0}^{\Delta} \theta_i (\Delta \choose i) z^i \).
Lemma 5. The system of equations (5) - (6) always has at least one solution \((x^*, y^*)\) on \(\mathbb{R}^+ \times \mathbb{R}^+\) for which \(x^* = y^*\). If the system of equations (5) - (6) has a unique solution on \(\mathbb{R}^+ \times \mathbb{R}^+\), then

\[
x^* = \zeta_{\lambda, \theta, \infty} = \lim_{d \to \infty} \inf_{B} \zeta_{\lambda, \theta, d}(B) = \lim_{d \to \infty} \sup_{B} \zeta_{\lambda, \theta, d}(B) = \zeta_{\lambda, \theta, \infty} = \zeta_{\lambda, \theta, \infty}^\Delta.
\]

In this case, \(p_k^{x, \theta}\) exists for all \(k \in \{0, \ldots, \Delta\}\), and equals \(\frac{\theta_k(\Delta_k) \zeta_{\lambda, \theta, \infty}^\Delta}{\lambda f_{\theta}^\Delta(\zeta_{\lambda, \theta, \infty}) + L_{\theta}(\zeta_{\lambda, \theta, \infty})}\). Alternatively, if the system of equations (5) - (6) does not have a unique solution on \(\mathbb{R}^+ \times \mathbb{R}^+\), then the left-hand-side of (7) does not equal the right-hand-side of (7).

Proof. We first prove that the system of equations (5) - (6) always has at least one solution \((x^*, y^*)\) on \(\mathbb{R}^+ \times \mathbb{R}^+\) for which \(x^* = y^*\). Let \(\eta_{\lambda, \theta}(x) = x - \lambda g_{\theta}(x) f_{\theta}^{\Delta-1}(x)\). Note that \(\eta_{\lambda, \theta}(0) = -\lambda \theta_0^{\Delta-1} \theta_0^\Delta < 0\). It follows from (21)-(22) that \(\eta_{\lambda, \theta}(\lambda \theta_0^{\Delta-1}(\theta_0^\Delta)^{\Delta-1}) \geq 0\). As \(\eta_{\lambda, \theta}\) 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 (5) - (6) has a unique solution on \(\mathbb{R}^+ \times \mathbb{R}^+\), then (23) holds. In particular, suppose (5) - (6) has a unique solution \((x^*, y^*)\) on \(\mathbb{R}^+ \times \mathbb{R}^+\). Then it must be that any solution \((x, y)\) to the system of equations satisfies \(x = x^* = y^* = y\). By Observation 10, \((\zeta_{\infty}, \zeta_{\infty})\) is such a solution. Thus \(\zeta_{\infty} = \zeta_{\infty}\), in which case (23) follows from Lemma 2. Furthermore, it then follows from (8) and Lemma 1 that in this case, \(p_k^{x, \theta}\) exists for all \(k \in \{0, \ldots, \Delta\}\), and equals \(\frac{\theta_k(\Delta_k) \zeta_{\lambda, \theta, \infty}^\Delta}{\lambda f_{\theta}^\Delta(\zeta_{\lambda, \theta, \infty}) + L_{\theta}(\zeta_{\lambda, \theta, \infty})}\).

Finally, we prove that if the system of equations (5) - (6) does not have a unique solution on \(\mathbb{R}^+ \times \mathbb{R}^+\), then the left-hand-side of (7) does not equal the right-hand-side of (7). In particular, suppose that (5) - (6) has at least two distinct non-negative solutions. 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 \(\overline{y} = \sup_{z \in S} z_2\), i.e. 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). Indeed, consider any sequence of solution vectors \(\{z^i, i \geq 1\}\) such that \(\lim_{i \to \infty} z^i = \overline{y}\). Since \(\{z^i, i \geq 1\}\) is uniformly bounded, the Bolzano-Weierstrass Theorem implies that \(\{z^i, i \geq 1\}\) will itself have a subsequence \(\{z^{i_k}, k \geq 1\}\) such that \(\{z^{i_k}, k \geq 1\}\) converges, and let us denote this limit by \(\underline{z}\). That \((\underline{z}, \overline{y})\) satisfies the system of equations then follows from the continuity of \(f\) and \(g\). Similarly, let \(\underline{x} = \inf_{z \in S} z_1\), i.e. the smallest number appearing in any solution pair, and \(y\) the other number appearing in the corresponding solution pair (whose existence is guaranteed by the same argument used above). Note that \(\overline{y} = \overline{y}\).

Recall that \(B\) denotes the boundary condition with all nodes at depth \(d\) included. We now prove (by induction) that \(\{\zeta_d(B), d \geq 3\}\) has a non-vanishing parity-dependence, with even values lying below \(\underline{x}\), and odd values lying above \(\overline{y}\). We begin with the base cases \(d = 3, 4\). Note that \(Z_{\lambda, \theta, 3}(0, 0, B) = \theta_0 \theta^{\Delta-1}_{\Delta-1}, Z_{\lambda, \theta, 3}(0, 1, B) = \lambda \theta^{\Delta-1}_{\Delta-1}\), and thus

\[
\zeta_{\lambda, \theta, 3}(B) = \lambda \theta_0^{-1}(\theta_0^{\Delta})^{\Delta-1}.
\]
That \( \zeta_{\lambda, \theta, d} (B) \geq \overline{y} \) then follows from (21)-(22). Similarly,

\[
Z_{\lambda, \theta, d}(0,0,B) = \sum_{x \in \{0,1\}^{\Delta - 1}} \theta_{x} (\lambda \theta_{\Delta - 1}^{-1} |x| (\theta_{0} \theta_{\Delta - 1}^{-1} \Delta - 1 - |x|)
= \theta_{0}^{\Delta - 1} \theta_{\Delta - 1}^{-1} \sum_{x \in \{0,1\}^{\Delta - 1}} \theta_{x} (\lambda \theta_{0}^{-1} (\theta_{\Delta - 1}^{-1} \Delta - 1 - |x|))
= \theta_{0}^{\Delta - 1} \theta_{\Delta - 1}^{-1} \lambda \theta_{0}^{-1} (\lambda \theta_{\Delta - 1}^{-1} \Delta - 1),
\]
and \( Z_{\lambda, \theta, d}(0,1,B) = \lambda \theta_{1}^{\Delta - 1} \theta_{\Delta - 1}^{-1} \). We conclude that

\[
(25) \quad \zeta_{\lambda, \theta, d}(B) = \lambda \left( \frac{\theta_{1}}{\theta_{0}} \right)^{\Delta - 1} \theta_{\theta}^{-1} (\lambda \theta_{0}^{-1} (\theta_{\Delta - 1}^{-1} \Delta - 1))
\leq \lambda f_{\theta}^{\Delta - 1}(\overline{y}) g_{\theta}(\overline{y}) = \overline{y},
\]
with the final inequality following from (21) - (22), the fact that \( \overline{y} = \lambda g_{\theta}(x) f_{\theta}^{\Delta - 1}(\overline{y}) \), and the monotonicity of \( g \). This completes the proof for the base case.

Now, suppose the induction is true for all \( d \in \{3, \ldots, 2k\} \) for some \( k \geq 2 \). Then it follows from Lemma 1, and the monotonicity of \( f_{\theta} \) and \( g_{\theta} \), that

\[
\zeta_{\lambda, \theta, 2k+1}(B) = \lambda g_{\theta} (\zeta_{\lambda, \theta, 2k}(B)) f_{\theta}^{\Delta - 1} (\zeta_{\lambda, \theta, 2k+1}(B))
\geq \lambda g_{\theta} (\overline{y}) f_{\theta}^{\Delta - 1} (\overline{y})
\geq \lambda g_{\theta} (\overline{x}) f_{\theta}^{\Delta - 1} (\overline{y}) = \overline{y}, \quad \text{since} \overline{x} \leq \overline{x};
\]
and

\[
\zeta_{\lambda, \theta, 2k+2}(B) = \lambda g_{\theta} (\zeta_{\lambda, \theta, 2k+1}(B)) f_{\theta}^{\Delta - 1} (\zeta_{\lambda, \theta, 2k+2}(B))
\leq \lambda g_{\theta} (\overline{y}) f_{\theta}^{\Delta - 1} (\overline{x})
\leq \lambda g_{\theta} (\overline{y}) f_{\theta}^{\Delta - 1} (\overline{x}) = \overline{x}, \quad \text{since} \overline{y} \geq \overline{y},
\]
completing the proof.

Since \( \{B^{i}, i = 1, \ldots, \Delta - 1\} \) are identical boundary conditions, (8) and Lemma 1 imply that for \( d \geq 5 \),

\[
p_{+}^{\lambda, \theta, d}(B) = \left( 1 + \frac{L_{\theta}(\zeta_{\lambda, \theta, d}(B))}{\lambda f_{\theta}^{\Delta - 1}(\zeta_{\lambda, \theta, d-1}(B))} \right)^{-1}.
\]
It follows from the monotonicity of \( L_{\theta} \) and \( f_{\theta} \) that for all even \( d \geq 5 \),

\[
p_{+}^{\lambda, \theta, d}(B) \geq \left( 1 + \frac{L_{\theta}(\overline{y})}{\lambda f_{\theta}^{\Delta - 1}(\overline{y})} \right)^{-1},
\]
while for all odd \( d \geq 5 \),

\[
p_{+}^{\lambda, \theta, d}(B) \leq \left( 1 + \frac{L_{\theta}(\overline{y})}{\lambda f_{\theta}^{\Delta - 1}(\overline{y})} \right)^{-1} < \left( 1 + \frac{L_{\theta}(\overline{y})}{\lambda f_{\theta}^{\Delta - 1}(\overline{y})} \right)^{-1} \leq p_{+}^{\lambda, \theta, d+1}(B).
\]
Combining the above completes the proof. \( \square \)
4. Existence of phase transition and proof of Theorem 2. In this section, we show that the family of second-order homogenous isotropic Markov random fields undergoes a phase transition with respect to the activity $\lambda$, by completing the proof of Theorem 2.

Proof of Theorem 2. We first show that for any fixed log-convex vector $\theta$, $(\lambda, \theta)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for all activities $\lambda < \lambda_0$. We proceed by proving that for $\lambda < \lambda_0$, the update rule for $|z_{\lambda,\theta_d} - \zeta_{\lambda,\theta_d}|$ is a contraction. We first prove that $f_{\theta, \theta_d}$ Lipschitz, and explicitly bound the relevant Lipschitz constants. Note that $f_{\theta, \theta_d}$ are differentiable on $\mathcal{R}^+$. Recall that $\psi_{\theta} = \max_{k=0, \ldots, \Delta - 2} \left( (\Delta - (k + 1)) \frac{\theta_{k+1}}{\theta_k} \right)$. Then for all $x \geq 0$,

$$\left| \partial_x g_{\theta}(x) \right| = \frac{\sum_{k=0}^{\Delta-2} (k + 1) \theta_{k+1} (\Delta-1) x^k}{\left( \sum_{k=0}^{\Delta-1} \theta_k (\Delta-1) x^k \right)^2} \leq \theta_0^{-1} \frac{\sum_{k=0}^{\Delta-1} (k + 1) \theta_{k+1} (\Delta-1) x^k}{\sum_{k=0}^{\Delta-1} \theta_k (\Delta-1) x^k} \leq \theta_0^{-1} \max_{k=0, \ldots, \Delta - 2} \frac{(k + 1) \theta_{k+1} (\Delta-1)}{\theta_k (\Delta-1)} = \theta_0^{-1} \psi_{\theta};$$

(26)

$$\left| \partial_x f_{\theta}(x) \right| = \left| \frac{\sum_{k=0}^{\Delta-1} \theta_k (\Delta-1) x^k}{(\sum_{k=0}^{\Delta-1} \theta_k (\Delta-1) x^k)^2} \right| - \frac{(\sum_{k=0}^{\Delta-1} \theta_{k+1} (\Delta-1) x^k)}{(\sum_{k=0}^{\Delta-1} \theta_k (\Delta-1) x^k)^2} \leq \sum_{i=0}^{\Delta-1} \sum_{j=0}^{\Delta-2} (j + 1) (\Delta-1) \theta_{i+1} \theta_{j+1} - \theta_{i+1} \theta_{j+1} x^{i+j} \leq \max_{i \in [0, \Delta - 1]} \theta_{i+1} \theta_{j+1} \frac{(\Delta - (j + 1)) \max \left( \frac{\theta_{j+2}}{\theta_j} - \frac{\theta_{i+1} \theta_{j+1}}{\theta_i \theta_j}, \frac{\theta_{i+1} \theta_{j+1}}{\theta_i \theta_j} - \frac{\theta_{j+2}}{\theta_j} \right)}{\theta_{j+1}} = \max_{j \in [0, \Delta - 2]} \left( (\Delta - (j + 1)) \frac{\theta_{j+1}}{\theta_j} \max \left( \frac{\theta_{j+2}}{\theta_{j+1}}, \frac{\theta_{j+2}}{\theta_j} - \frac{\theta_{j+1}}{\theta_j} \right) \right) \leq \frac{\theta_{\Delta}}{\theta_{\Delta-1}} - \frac{\theta_{1}}{\theta_{0}} \psi_{\theta}.$$
for all $a, b, c, d \in \mathbb{R}$ that for all $d \geq 7,$

$$|\xi_{\lambda, \theta, d} - \xi_{\lambda, \theta, d}| = \lambda |g_{\theta}(\xi_{\lambda, \theta, d-1})J_{\theta}^{\Delta-1}(\xi_{\lambda, \theta, d-2}) - g_{\theta}(\xi_{\lambda, \theta, d-1})J_{\theta}^{\Delta-1}(\xi_{\lambda, \theta, d-2})|$$

$$\leq \lambda \left( 2(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1} \theta_{\Delta-1}^{-1} \psi_{\theta} \xi_{\lambda, \theta, d-1} - \xi_{\lambda, \theta, d-1} \right)$$

$$+ 2\theta_{\Delta-1}^{\Delta-1}(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-2}(\frac{\theta_{\Delta}}{\theta_{\Delta-1}} - \frac{\theta_{1}}{\theta_{0}}) \psi_{\theta} \xi_{\lambda, \theta, d-2} - \xi_{\lambda, \theta, d-2} \right)$$

$$\leq \lambda \Delta_{\theta}^{-1} \max \left( |\xi_{\lambda, \theta, d-1} - \xi_{\lambda, \theta, d-1}|, |\xi_{\lambda, \theta, d-2} - \xi_{\lambda, \theta, d-2}| \right).$$

Thus $\lambda \in (0, \Delta_{\theta})$ implies that there exists $\rho \in (0, 1)$ such that for all $d \geq 7,$

$$|\xi_{\lambda, \theta, d} - \xi_{\lambda, \theta, d}| \leq \rho \max \left( |\xi_{\lambda, \theta, d-1} - \xi_{\lambda, \theta, d-1}|, |\xi_{\lambda, \theta, d-2} - \xi_{\lambda, \theta, d-2}| \right).$$

We conclude that \( \lim_{d \to \infty} |\xi_{\lambda, \theta, d} - \xi_{\lambda, \theta, d}| = 0 \), and the desired claim then follows from Lemma 2.

We now prove that \((\lambda, \theta)\) does not admit a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree for all activities $\lambda > \lambda_{\theta}$. We first show that for $\lambda = \lambda_{\theta}$, any non-negative solution to the system of equations (5)-(6) of the form $(x, x)$ satisfies

\begin{equation}
(28) \quad x \geq \frac{3 \theta_{0}}{\Delta \theta_{1}}.
\end{equation}

Indeed, it follows from log-convexity that $g_{\theta}(x) \geq (\theta_{0}(1 + \frac{\theta_{\Delta}}{\theta_{\Delta-1}}x)^{\Delta-1})^{-1},$ and thus by (22),

\[ x(1 + \frac{\theta_{\Delta}}{\theta_{\Delta-1}}x)^{\Delta-1} \geq \lambda_{\theta}(\frac{\theta_{1}}{\theta_{0}})^{\Delta-1}. \]

Further applying the fact that $e^z \geq z + 1$ for all $z \in \mathbb{R}$ yields

\[ x \exp \left( \Delta \frac{\theta_{\Delta}}{\theta_{\Delta-1}}x \right) \geq \lambda_{\theta}(\frac{\theta_{1}}{\theta_{0}})^{\Delta-1}. \]

The desired result then follows from a straightforward contradiction argument, the details of which we omit.

We next prove that for any $x \geq \frac{3 \theta_{0}}{\Delta \theta_{1}}$ and $M \geq 1$, $\eta_{\theta}(M, x) \Delta M g_{\theta}(Mx) \leq g_{\theta}(x)$. We proceed by showing that in this case, $\partial_{x} g_{\theta}(M, x) \leq 0$ for all $M \geq 1$. Indeed, since $\partial_{x} g_{\theta}(x) = -g_{\theta}^{2}(x) \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k} x^{k-1},$ it follows form the chain rule that

\begin{align*}
\partial_{M} \eta_{\theta}(M, x) &= -M x g_{\theta}^{2}(Mx) \sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k} (Mx)^{k-1} + g_{\theta}(Mx) \\
&= g_{\theta}(Mx) \left( 1 - \frac{\sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k} x^{k}}{\sum_{k=0}^{\Delta-1} \binom{\Delta-1}{k} \theta_{k} x^{k}} \right).
\end{align*}
It thus suffices to demonstrate that for any \( x \geq \frac{3 \theta_0}{\Delta \theta_1} \) and \( M \geq 1 \),
\[
\sum_{k=0}^{\Delta-1} \frac{(\Delta-1)! \theta_k k (Mx)^k}{\sum_{k=0}^{\Delta-1} (\Delta-1)! \theta_k x^k} \geq 1,
\]
or equivalently that for any \( x \geq \frac{3 \theta_0}{\Delta \theta_1} \),
\[
\sum_{k=2}^{\Delta-1} \left( \frac{\Delta-1}{k} \right) \theta_k (k-1)x^k \geq \theta_0.
\]
As \( \theta_2 \geq \theta_0 \left( \frac{\theta_1}{\theta_0} \right)^2 \), and \( \left( \frac{\Delta-1}{2} \right) \geq \frac{\Delta^2}{9} \) for all \( \Delta \geq 3 \), the desired claim then follows from non-negativity.

We now combine the above two claims to complete the proof, by demonstrating that for all \( \lambda > \overline{\lambda}_{\theta} \), the associated sequence \( \{\zeta_{\lambda, \theta, d}(B), d \geq 3\} \) has a non-vanishing parity-dependence, mirroring our proof of Lemma 5. Recall from Lemma 5 that for \( \lambda = \overline{\lambda}_{\theta} \), the system of equations \((5) - (6)\) always has at least one non-negative solution of the form \((x, x)\). Let us fix any such solution \((x_{\theta}, x_{\theta})\), and note that \( x_{\theta} \geq \frac{3 \theta_0}{\Delta \theta_1} \). We now prove by induction that for all \( \lambda > \overline{\lambda}_{\theta} \), \( \{\zeta_{\lambda, \theta, d}(B), d \geq 3\} \) has a non-vanishing parity-dependence, with even values lying below \( x_{\theta} \), and odd values lying above \( \frac{\lambda \lambda_{\theta}}{\lambda_{\theta}}x_{\theta} \). We begin with the base cases \( d = 3, 4 \). It follows from \((24), (21), \) and \((22)\) that
\[
\zeta_{\lambda, \theta, 3}(B) = \lambda \theta_0^{-1} \left( \frac{\theta_\Delta}{\theta_{\Delta-1}} \right)^{\Delta-1} \geq \frac{\lambda}{\lambda_{\theta}}x_{\theta}.
\]
Similarly, it follows from \((25)\), the fact that \( x_{\theta} \geq \frac{3 \theta_0}{\Delta \theta_1} \), and our proven claims that
\[
\zeta_{\lambda, \theta, 4}(B) = \lambda \left( \frac{\theta_1}{\theta_0} \right)^{\Delta-1} g_{\theta} \left( \lambda \theta_0^{-1} \left( \frac{\theta_\Delta}{\theta_{\Delta-1}} \right)^{\Delta-1} \right)
\leq \frac{\lambda}{\lambda_{\theta}} \lambda_{\theta} \left( \frac{\theta_1}{\theta_0} \right)^{\Delta-1} g_{\theta} \left( \frac{\lambda}{\lambda_{\theta}}x_{\theta} \right)
\leq \lambda_{\theta} g_{\theta}(x_{\theta}) \left( \frac{\theta_1}{\theta_0} \right)^{\Delta-1} \leq x_{\theta}.
\]
Now, suppose the induction is true for all \( d \in \{3, \ldots, 2k\} \) for some \( k \geq 2 \). Then it follows from Lemma 1, and the monotonicity of \( f_{\theta} \) and \( g_{\theta} \), that
\[
\zeta_{\lambda, \theta, 2k+1}(B) = \lambda g_{\theta} \left( \zeta_{\lambda, \theta, 2k}(B) \right) f_{\theta}^{\Delta-1} \left( \zeta_{\lambda, \theta, 2k-1}(B) \right)
\geq \lambda g_{\theta}(x_{\theta}) f_{\theta}^{\Delta-1} \left( \frac{\lambda}{\lambda_{\theta}}x_{\theta} \right)
\geq \lambda g_{\theta}(x_{\theta}) f_{\theta}^{\Delta-1}(x_{\theta}) = \frac{\lambda}{\lambda_{\theta}}x_{\theta};
\]
and
\[ \zeta_{\lambda, \theta, 2k+2}(B) = \lambda g_\theta (\zeta_{\lambda, \theta, 2k+1}(B)) f_\theta^{\Delta-1}(\zeta_{\lambda, \theta, 2k}(B)) \]
\[ \leq \frac{\lambda}{\lambda_\theta} \tilde{g}_\theta (\frac{\lambda}{\lambda_\theta} x_\theta) f_\theta^{\Delta-1}(x_\theta) \]
\[ \leq \tilde{g}_\theta (x_\theta) f_\theta^{\Delta-1}(x_\theta) = x_\theta. \]

completing the proof. That the non-vanishing parity-dependence of \( \{\zeta_{\lambda, \theta, d}(B), d \geq 3\} \) implies \((\lambda, \theta)\)
does not admit a unique infinite-volume Gibbs measure on the infinite \( \Delta \)-regular tree follows identically to the proof of the corresponding result in the proof of Lemma 5, and we omit the details. \( \square \)

5. A perturbative analysis, and proof of Theorem 3. In this section, we develop a perturbative approach to gain insight into the geometry of the uniqueness regime, proving Theorem 3. First, it will be useful to rewrite the system of equations (5) - (6), which will allow us to give necessary and sufficient conditions for uniqueness using known results from the theory of dynamical systems. Note that if \( p_\theta(x) \overset{\Delta}{=} x f_\theta^{-(\Delta-1)}(x) \) is strictly increasing on \([0, \lambda(\frac{\theta_\Delta}{\theta_\Delta-1})^{\Delta-1}\theta_0^{-1}]\),
then it follows from (21) - (22) that \( p_\theta \) has a well-defined and unique inverse \( p_\theta^{-1} \), with domain
a superset of \([0, \lambda \theta_0^{-1}]\) and range a subset of \( \mathbb{R}^+ \), i.e. \( p_\theta^{-1}(p_\theta(x)) = x \). In this case we can define \( q_{\lambda, \theta}(x) \overset{\Delta}{=} p_\theta^{-1}(\lambda g_\theta(x)) \), and we observe that the system of equations (5) - (6) may be rewritten as follows.

**Observation 11.** If \( \theta \) 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 (5) - (6) is equivalent to the system of equations
\[ q_{\lambda, \theta}(q_{\lambda, \theta}(x)) = x, \]
\[ y = q_{\lambda, \theta}(x). \]
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 (21) - (22) that every solution \((x, y)\) to the system of equations (29) - (30) 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 (29) - (30) 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 \([83]\). Recall that for a thrice-differentiable function \( F(x) \) with non-vanishing derivative on some interval \( I \), we define (on \( I \)) the Schwarzsian derivative of \( F \) as the function
\[ S[F] = \frac{\Delta}{\Delta x^3} F - \frac{3}{2} (\frac{\Delta^2}{\Delta x^2} F)^2. \]
For a function $F$ and $n \geq 1$, let $F^{(n)}(x)$ denote the $n$-fold iterate of $F$, i.e. $F^{(n+1)}(x) = F(F^{(n)}(x))$, with $F^{(1)}(x) = F(x)$. Then the following well-known result from dynamical systems is stated in Lemma 4.3 of [83].

**Theorem 4.** Suppose $I = [L, R] \subseteq \mathbb{R}$ 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) $\lim_{n \to \infty} F^{(n)}(x) < 0$ 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 4 to our own setting. Let $r_{\theta}(x) \equiv \frac{\partial_x q_{\theta}(x)}{\partial_x p_{\theta}(x)}$. Then combined with Observation 11, Theorem 4 implies the following.

**Observation 12.** Suppose that $\theta$ is log-convex, $p_{\theta}(x)$ is strictly increasing on $[0, \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1} \theta_0^{-1}]$, and the conditions of Theorem 4 are satisfied with $F = q_{\lambda, \theta} \cdot I = [0, \lambda(\frac{\theta_{\Delta}}{\theta_{\Delta-1}})^{\Delta-1} \theta_0^{-1}]$. Then $(\lambda, \theta)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree iff $r_{\theta}(x_{\lambda, \theta}) \leq \lambda^{-1}$.

**Proof.** We first prove that the system of equations (29) - (30) does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$ iff $|\partial_x q_{\lambda, \theta}(x_{\lambda, \theta})| > 1$. Suppose the system of equations (29) - (30) does not have a unique solution on $\mathbb{R}^+ \times \mathbb{R}^+$. Since Observation 11 implies that the equation $q_{\lambda, \theta}(x) = x$ has a unique solution $x_{\lambda, \theta}$, it follows that there must exist a solution $(x, y)$ to the system of equations (29) - (30) with $x < y$. In this case, $\lim_{n \to \infty} q_{\lambda, \theta}^{(n)}(x)$ does not exist, as the series alternates between $x$ and $y$, and it follows from Theorem 4 that $|\partial_x q_{\lambda, \theta}(x_{\lambda, \theta})| > 1$.

Alternatively, suppose that $|\partial_x q_{\lambda, \theta}(x_{\lambda, \theta})| > 1$. Then it follows from Theorem 4 that $\lim_{n \to \infty} q_{\lambda, \theta}^{(n)}(0)$ does not exist. However, both $Z_{\text{even}} \equiv \lim_{n \to \infty} q_{\lambda, \theta}^{(2n)}(0)$ and $Z_{\text{odd}} \equiv \lim_{n \to \infty} q_{\lambda, \theta}^{(2n+1)}(0)$ both exist. Indeed, this follows from the fact that $q_{\lambda, \theta}^{(1)}$ is decreasing, $q_{\lambda, \theta}^{(2)}$ is increasing, $q_{\lambda, \theta}^{(2)}(0) \geq 0$, and $q_{\lambda, \theta}^{(3)}(0) \leq q_{\lambda, \theta}^{(1)}(0)$, which implies that both relevant sequences are monotone. Noting that the non-existence 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 (29) - (30), completes the desired demonstration.

As it follows from elementary calculus that $\partial_x q_{\lambda, \theta}(x_{\lambda, \theta}) = \frac{\partial_x g_{\theta}(x_{\lambda, \theta})}{\partial_x p_{\theta}(x_{\lambda, \theta})}$, combining the above with Theorem 1 and Observation 11 completes the proof.

We note that $p_{\theta}$ is not necessarily an increasing function for the case of general log-convex $\theta$. Furthermore, even when $p_{\theta}$ is increasing, an analysis of $S[q_{\lambda, \theta}]$ seems difficult, and the associated uniqueness regime of the parameter space seems to be quite complex. However, for the special setting in which $\theta$ belongs to a neighborhood of the all ones vector, in which case the associated
Markov random field 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 [74] for details. Using this framework, we prove the following.

**Lemma 6.** For each fixed convex vector $c$, and $U \in \mathbb{R}^+$, there exists $\delta_{c,U} > 0$ such that the following hold.

1. $g_{1+ch}(x)$ and $p_{1+ch}(x)$ are jointly real analytic functions of $(h, x)$ on $[0, \delta_{c,U}] \times [0, U]$. For each fixed $h \in [0, \delta_{c,U}]$ and all $x \in [0, U]$, $\partial_x g_{1+ch}(x) < 0$, and $\partial_x p_{1+ch}(x) > 0$.
2. For each fixed $h \in [0, \delta_{c,U}]$, $p_{1+ch}(x)$ has a well-defined and unique inverse $p_{1+ch}^−(x)$ with domain a superset of $[0, U]$ and range a subset of $\mathbb{R}^+$. Furthermore, $p_{1+ch}^−(x)$ is a jointly real analytic function of $(h, x)$ on $[0, \delta_{c,U}] \times [0, U]$.
3. $q_{\lambda, 1+ch}(x)$, $\partial_x q_{\lambda, 1+ch}(x)$, and $S[q_{\lambda, 1+ch}](x)$ are all jointly real analytic functions of $(h, x)$ on $[0, \delta_{c,U}] \times [0, U]$. Furthermore, $\partial_x q_{\lambda, 1+ch}(x)$ and $S[q_{\lambda, 1+ch}](x)$ are strictly negative for all $(h, x) \in [0, \delta_{c,U}] \times [0, U]$.

**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 $g_{1+ch}(x)$ and $p_{1+ch}(x)$ are ratios of non-vanishing polynomials of $(h, x)$ on $[0, \delta_{1,U_1}] \times [0, U_1]$. That there exists $\delta_{2,U_1} > 0$ such that $\partial_x g_{1+ch}(x) < 0$ and $\partial_x p_{1+ch}(x) > 0$ for all $(h, x) \in [0, \delta_{2,U_1}] \times [0, U_1]$ then follows from the fact that $\partial_x g_1(x) = -(\Delta - 1)(x + 1)^{-\Delta}$, and $\partial_x p_1(x) = 1$.

(ii). The claim follows from (i) and the inverse function theorem for real analytic functions [74].

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

Combining Observation 12 with Lemma 6 immediately yields necessary and sufficient conditions for uniqueness when $\theta$ is log-convex and in a neighborhood of 1.

**Corollary 3.** For each fixed convex vector $c$, there exists $\delta_c > 0$ such that the following hold for all $h \in [0, \delta_c]$.

1. $q_{\lambda, 1+ch}(x) - x$ is strictly decreasing on $[0, 2\lambda_\Delta]$, and has a unique zero $x_{\lambda, 1+ch}$ on $[0, 2\lambda_\Delta]$.
2. $(\lambda_\Delta, 1 + ch)$ admits a unique infinite-volume Gibbs measure on the infinite $\Delta$-regular tree iff $\tau_{1+ch}(x_{\lambda, 1+ch}) \leq \lambda_\Delta^{-1}$.

With Corollary 3 in hand, we now complete the proof of Theorem 3. For $l \in \{0,1\}$, and $\theta = (\theta_0, \ldots, \theta_\Delta), c = (c_0, \ldots, c_\Delta) \in \mathbb{R}^+ (\Delta+1)$, let

\[
\begin{align*}
    f_{l, \theta}(x) &\approx \sum_{i=0}^{\Delta-1} \theta_{i+l} \binom{\Delta-1}{i} i^i, \\
    z_{l,c} &\approx \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} x_{\lambda, 1}^{i+l} c_{i+l}, \\
    w_{l,c} &\approx \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} i x_{\lambda, 1}^{i-1} c_{i+l}.
\end{align*}
\]
and 
\[ x_c = \frac{1}{2} \frac{(\Delta - 2)\Delta^2}{(\Delta - 1)\Delta - 1} \left( (\Delta - 1)z_{1,c} - \Delta z_{0,c} \right). \]

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 7.**

\[ x_{\lambda,1} = (\Delta - 2)^{-1}, \sum_{i=0}^{\Delta} A_{\Delta,i} = \frac{(\Delta - 1)}{\Delta - 2}, \sum_{i=0}^{\Delta} i^2 A_{\Delta,i} = 2\Delta \frac{(\Delta - 1)\Delta^2 - 1}{(\Delta - 2)\Delta}, \]

\[ \sum_{i=0}^{\Delta} \pi_i = -\frac{(\Delta - 1)}{\Delta - 2}, \sum_{i=0}^{\Delta} i\pi_i = \Delta \frac{(\Delta - 1)}{\Delta - 2}. \]

**Proof of Theorem 3.** We proceed by analyzing \( r_{1+ch}(x_{\lambda,1} + ch) - \lambda \Delta^1 \) as \( h \downarrow 0 \), and begin by proving that

\[ \lim_{h \downarrow 0} (r_{1+ch}(x_{\lambda,1} + ch) - x_{\lambda,1})h^{-1} = x_c. \]

Note that for any fixed \( \alpha \in \mathbb{R}, \)

\[ f_{1+ch}(x_{\lambda,1} + ah) = \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} (1 + c_i h) \sum_{j=0}^i \binom{i}{j} x_{\lambda,1}^{i-j} (\alpha h)^j \]

\[ = \sum_{i=0}^{\Delta-1} \binom{\Delta-1}{i} x_{\lambda,1}^i (1 + c_i h) (1 + i x_{\lambda,1}^{-1} \alpha h) + o(h) \]

\[ = (1 + x_{\lambda,1})^{\Delta-1} + ((\Delta - 1)(1 + x_{\lambda,1})^{\Delta-2} \alpha + z_{1,c})h + o(h). \]

We conclude that

\[ g_{1+ch}(x_{\lambda,1} + ah) = (1 + x_{\lambda,1})^{-(\Delta-1)} - (1 + x_{\lambda,1})^{-2(\Delta-1)}((\Delta - 1)(1 + x_{\lambda,1})^{\Delta-2} \alpha + z_{0,c})h + o(h), \]

and

\[ f_{1+ch}(x_{\lambda,1} + ah) = 1 + (1 + x_{\lambda,1})^{-(\Delta-1)}(z_{1,c} - z_{0,c})h + o(h). \]

It follows from \( (33), (34) \), and a straightforward calculation (the details of which we omit) that for \( \alpha \in \mathbb{R}, \)

\[ (x_{\lambda,1} + ah) - \lambda f_{1+ch}(x_{\lambda,1} + ah) g_{1+ch}(x_{\lambda,1} + ah) = 2(\alpha - x_c)h + o(h). \]

Combining with Corollary 3.1(i), and the fact that \( x_{\lambda,1} < \lambda \Delta \), completes the proof.

Next, we use \( (31) \) to prove that

\[ \partial_x p_{1+ch}(x_{\lambda,1} + ch) + \lambda \partial_x g_{1+ch}(x_{\lambda,1} + ch) = -\frac{1}{2} \frac{(\Delta - 2)}{(\Delta - 1)} \Delta \pi \cdot ch + o(h). \]
Indeed, it follows from (31) that

\[
\partial_x f_{1,1+ch}(x_{\lambda,1}+ch) = \sum_{i=1}^{\Delta-1} \left( \frac{\Delta-1}{i} \right) i(1 + c_i + ch) \sum_{j=0}^{i-1} \left( \frac{i-1}{j} \right) x_{\lambda,1}^{i-1-j} (x_{c,h})^j + o(h) \\
= \sum_{i=1}^{\Delta-1} \left( \frac{\Delta-1}{i} \right) i x_{\lambda,1}^i (1 + c_i + ch) (x_{\lambda,1}^{-1} + (i-1)x_{\lambda,1}^{-2} x_{c,h} + o(h),
\]

which itself equals

\[
(36) \quad (\Delta - 1)(1 + x_{\lambda,1})^{\Delta-2} + \left( x_c(\Delta - 1)(\Delta - 2)(1 + x_{\lambda,1})^{\Delta-3} + w_{1,c} \right) h + o(h).
\]

It follows from (31) - (36), and a straightforward calculation (the details of which we omit), that

\[
\partial_x g_{1+ch}(x_{\lambda,1}+ch) = -g_{1+ch}^2(x_{\lambda,1}+ch) \partial_x f_{0,1+ch}(x_{\lambda,1}+ch),
\]

which itself equals

\[
(37) \quad -(\Delta-1)(1+x_{\lambda,1})^{-\Delta} + (1+x_{\lambda,1})^{-(2\Delta-1)} \left( -(1+x_{\lambda,1}) w_{0,c} + \Delta(\Delta-1)(1+x_{\lambda,1})^{\Delta-2} x_c + 2(\Delta-1) z_{0,c} \right) h + o(h);
\]

\[
\partial_x f_{1+ch}(x_{\lambda,1}+ch) \text{ equals}
\]

\[
\hat{g}_{1+ch}^2(x_{\lambda,1}+ch) \left( f_{0,1+ch}(x_{\lambda,1}+ch) \partial_x f_{1,1+ch}(x_{\lambda,1}+ch) - f_{1,1+ch}(x_{\lambda,1}+ch) \partial_x f_{0,1+ch}(x_{\lambda,1}+ch) \right),
\]

which itself equals

\[
(1 + x_{\lambda,1})^{-\Delta} \left( (1 + x_{\lambda,1})(w_{1,c} - w_{0,c}) + (\Delta - 1)(z_{0,c} - z_{1,c}) \right) h + o(h);
\]

and

\[
\partial_x p_{1+ch}(x_{\lambda,1}+ch) = f_{1+ch}^{-\Delta-1}(x_{\lambda,1}+ch) - (\Delta - 1)x_{\lambda,1}+ch f_{1+ch}^{-\Delta}(x_{\lambda,1}+ch) \partial_x f_{1+ch}(x_{\lambda,1}+ch),
\]

which itself equals

\[
(38) \quad 1 - (\Delta - 1)x_{\lambda,1}(1 + x_{\lambda,1})^{-(\Delta-1)}(w_{1,c} - w_{0,c}) h + o(h).
\]

Combining (37) - (38) with Lemma 7 and simplifying, we conclude that the left-hand side of (35) equals

\[
(39) \quad \frac{(\Delta - 2)^{\Delta-2}}{2(\Delta - 1)\Delta} \left( -(\Delta-2)^3 z_{0,c} + \Delta(\Delta-1)(\Delta-2) z_{1,c} + 2(\Delta-1)(\Delta-2) w_{0,c} - 2(\Delta-1)^2 w_{1,c} \right) h + o(h).
\]

It follows from the definition of \( \pi \) and a further straightforward algebraic manipulation that (39) equals \(-\frac{1}{2}(\Delta - 2)\pi \cdot ch + o(h)\), completing the desired demonstration.

Combining (31), (35), and Corollary 3 with the fact that \( r_{1+ch}(x_{\lambda,1}+ch) \leq \lambda_{\Delta}^{-1} \) iff the left-hand side of (35) is non-negative, completes the proof of the theorem. \( \square \)
6. Conclusion. In this paper, we posed the question of how to sample from the independent sets of large bounded-degree graphs so that the number of included neighbors of each excluded node has a given distribution of our choosing. We found that higher order Markov random fields were well-suited to this task, and investigated the properties of these models. For the family of so-called reverse ultra log-concave distributions, which includes the truncated Poisson and geometric, we gave necessary and sufficient conditions for the natural higher order Markov random field which induces the desired distribution to be in the uniqueness regime in large regular graphs of large girth, in terms of the set of solutions to a certain system of equations. We observed that the associated set of models corresponds to the family of second-order homogenous isotropic Markov random fields with log-convex clique potentials. We also showed that these Markov random fields undergo a phase transition, identified the extremal boundary conditions, and gave explicit bounds on the associated critical activity, which we proved to exhibit a certain robustness. For distributions which are small perturbations around the binomial distribution realized by the hardcore model at critical activity, we gave a description of the corresponding uniqueness regime in terms of a simple polyhedral cone. Furthermore, our analysis revealed an interesting non-monotonicity with regards to biasing towards excluded nodes with no included neighbors.

This work leaves many interesting directions for future research. The full power of higher order Markov random fields 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 homogenous isotropic Markov random fields with log-convex clique potentials, or the setting of log-concave clique potentials (which includes the restriction to maximal independent sets), where the relevant recursions also simplify. It is also an open question to understand which sets of occupancy probabilities can be acheived by higher order Markov random fields (in the uniqueness regime). Can one use higher order Markov random fields (in the uniqueness regime) to sample from denser independent sets than can be attained using the hardcore model at the critical activity? On a related note, can one given necessary and sufficient conditions for uniqueness for higher (beyond second) order Markov random fields (analogous to Theorem 1), and precisely how do such conditions relate to the positive (negative) lattice conditions and related notions of positive (negative) association and concavity (convexity)? It would also be interesting to study higher order Markov random fields for related combinatorial problems, e.g. graph coloring, as well as for graphs which are not regular and of large girth.

The algorithmic implications of phase transitions for higher order Markov random fields also remain open questions. In particular, one would expect a “complexity transition” at the uniqueness threshold with respect to approximately computing the relevant partition function and sampling from the associated distributions, as has been recently established for first order Markov random fields [122],[107],[108],[48],[7],[106],[80],[27],[123]. Another possible direction would be to investigate the connections between our work and the recent work on computing partition functions using belief propagation and related message-passing algorithms [29],[16],[105],[102].

Finally, it is open to investigate the connection between higher order homogenous isotropic Markov random fields and recent research on sparse graph limits [82]. This includes work studying notions of convergence in large random graphs [10],[51],[20] as well as other notions of convergent graph sequences [12],[21],[18],[66],[37],[38],[88]. It would also be interesting to study the relationship between higher order homogenous isotropic Markov random fields and the related notions of local algorithm and i.i.d. factor [55],[53],[54],[52],[84],[43],[19],[33],[66]. We note that such investigations also relate to various earlier studies of automorphism invariant distributions on infinite graphs.
More generally, we conclude with the following question.

**Question 3.** To what extent are higher order homogenous isotropic Markov random fields in the (appropriately defined) uniqueness regime capable of (approximately) encoding those distributions on independent sets which exhibit long-range boundary independence?

**Acknowledgements.** The author would like to thank David Gamarnik, Kavita Ramanan, Justin Salez, and Prasad Tetali for several stimulating discussions.

**REFERENCES**

[1] D. Achlioptas and A. Coja-Oghlan, *Algorithmic barriers from phase transitions*, Foundations of Computer Science, 2008., IEEE, 2008, pp. 793–802.

[2] H. Akin, N. Ganikhodjaev, and S. Temir, *Modulated phase of a Potts model with competing binary interactions on a Cayley tree*, Journal of Statistical Physics 137 (2009), no. 4, 701–715.

[3] H. Akin, N. Ganikhodjaev, and S. Temir, *Potts model with two competing binary interactions*, Turk. J. Math 31 (2007), 229–238.

[4] I. Akyildiz, S. Weilian, Y. Sankarasubramaniam, and E. Cayirci, *Wireless sensor networks: a survey*, Computer networks 38 (2002), no. 4, 393–422.

[5] D. Aldous and R. Lyons, *Processes on unimodular random networks*, Electron. J. Probab 12 (2007), no. 54, 1454–1508.

[6] A. Backhausz, B. Szegedy, and B. Virag, *Ramanujan graphings and correlation decay in local algorithms*, arXiv preprint arXiv:1305.6784 (2013).

[7] A. Bandopadhyay and D. Gamarnik, *Counting without sampling, new algorithms for enumeration problems using statistical physics*, Random Structures and Algorithms 33 (2008), no. 4.

[8] N.C. Bartelt and T.L. Einstein, *Triangular lattice gas with first-and second-neighbor exclusions: Continuous transition in the four-state Potts universality class*, Physical Review B 30 (1984), no. 9, 5339.

[9] D. Basin, S. Capkun, J. Hubaux, P. Lafourcade, P. Papadimitratos, M. Poturalski, and P. Schaller, *Secure neighborhood discovery: A fundamental element for mobile ad hoc networking*, Communications Magazine, IEEE 46 (2008), no. 2, 132–139.

[10] M. Bayati, D. Gamarnik, and P. Tetali, *Combinatorial approach to the interpolation method and scaling limits in sparse random graphs*, Proceedings of the 42nd ACM symposium on Theory of computing, ACM, 2010, pp. 105–114.

[11] A. Bellemans and R. K. Nigam, *Phase transitions in the hard-square lattice gas*, Phys. Rev. Lett. 16 (1966), 1038–1039.

[12] I. Benjamini and O. Schramm, *Recurrence of distributional limits of finite planar graphs*, Selected Works of Oded Schramm, Springer, 2011, pp. 533–545.

[13] J. Besag, *Spatial interaction and the statistical analysis of lattice systems*, Journal of the royal statistical society, series B 36 (1974), no. 2, 192 – 236.

[14] V. Bhargava and M. Khabbazian, *Localized broadcasting with guaranteed delivery and bounded transmission redundancy*, Computers, IEEE Transactions on 57 (2008), no. 8, 1072–1086.

[15] G. Biroli and M. M´ezard, *Lattice glass models*, Phys. Rev. Lett. 88 (2001), 025501.

[16] M. Black, D. Huttenlocher, X. Lan, and S. Roth, *Efficient belief propagation with learned higher-order Markov random fields*, Computer Vision–ECCV 2006 (2006), 269–282.

[17] T. Bodineau and C. Toninelli, *Activity phase transition for constrained dynamics*, Communications in Mathematical Physics 311 (2012), 357–396.

[18] B. Bollobás and O. Riordan, *Sparse graphs: metrics and random models*, Random Structures & Algorithms 39 (2010), no. 1, 1–38.

[19] C. Borgs, M. Lelarge, and J. Salez, *Matchings on infinite graphs*, Probability Theory and Related Fields (2012), 1–26.

[20] C. Borgs, J. Chayes, and D. Gamarnik, *Convergent sequences of sparse graphs: A large deviations approach*, arXiv preprint arXiv:1302.4615 (2013).

[21] C. Borgs, J. Chayes, J. Kahn, and L. Lovász, *Left and right convergence of graphs with bounded degree*, Random Structures & Algorithms 42 (2013), no. 1, 1–28.
[22] C. Borgs, J. Chayes, and D. Randall, *The van den Berg-Kesten-Reimer inequality: a review*, Perplexing Problems in Probability: Festschrift in Honor of Harry Kesten (1998), 159–173.

[23] C. Brandt, N. Immorlica, G. Kamath, and R. Kleinberg, *An analysis of one-dimensional Schelling segregation*, Proceedings of the 44th symposium on Theory of Computing (New York, NY, USA), STOC ’12, ACM, 2012, pp. 789–804.

[24] W. Brendel and S. Todorovic, *Segmentation as maximum-weight independent set*, Proceedings of 24th Annual Conference on Neural Information Processing Systems (NIPS), 2010.

[25] G.R. Brightwell and P. Winkler, *Graph homomorphisms and phase transitions*, Journal of Combinatorial Theory, Series B 77 (1999), no. 2, 221 – 262.

[26] D.M. Burley, *A lattice model of a classical hard sphere gas*, Proceedings of the Physical Society 75 (1960), no. 2, 262.

[27] J. Cai, X. Chen, H. Guo, and P. Lu, *Inapproximability after uniqueness phase transition in two-spin systems*, Combinatorial Optimization and Applications, Springer, 2012, pp. 336–347.

[28] N. Cancrini, F. Martinelli, C. Roberto, and C. Toninelli, *Kinetically constrained spin models*, Probability Theory and Related Fields 140 (2008), 459–504, 10.1007/s00440-007-0072-3.

[29] V. Chandrasekaran, M. Chertkov, D. Gamarnik, D. Shah, and J. Shin, *Counting independent sets using the Bethe approximation*, SIAM Journal on Discrete Mathematics 25 (2011), no. 2, 1012–1034.

[30] S.C. Chang and R. Shrock, *Ground-state entropy of the Potts antiferromagnet with next-nearest-neighbor spin-spin couplings on strips of the square lattice*, Physical Review E 62 (2000), no. 4, 4650.

[31] W. Chen and C. Gu, *The reverse ultra log-concavity of the Boros-Moll polynomials*, Proceedings of the American Mathematical Society 137 (2009), no. 12, 3991–3998.

[32] T. Colin, *Local properties of an Ising model on a Cayley tree*, Journal of Statistical Physics 27 (1982), 441–456.

[33] E. Csoká, *Local algorithms with public randomisation on sparse graphs*, arXiv preprint arXiv:1202.1565 (2012).

[34] E. Csoka, B. Gerencser, V. Harangi, and B. Virag, *Invariant gaussian processes and independent sets on regular graphs of large girth*, arXiv preprint arXiv:1305.3977 (2013).

[35] L. Dall’Asta, P. Pin, and A. Ramezanpour, *Statistical mechanics of maximal independent sets*, Phys. Rev. E 80 (2009).

[36] L. Dall’Asta, P. Pin, and A. Ramezanpour, *Optimal equilibria of the best shot game*, Journal of Public Economic Theory 13 (2011), no. 6, 885–901.

[37] A. Dembo and A. Montanari, *Ising models on locally tree-like graphs*, The Annals of Applied Probability 20 (2010), no. 2, 565–592.

[38] A. Dembo, A. Montanari, A. Sly, and N. Sun, *The replica symmetric solution for Potts models on d-regular graphs*, arXiv preprint arXiv:1207.5500 (2012).

[39] D. Dereudre, R. Drouilhet, and H. Georgii, *Existence of Gibbsian point processes with geometry-dependent interactions*, Probability Theory and Related Fields 153 (2012), 643–670.

[40] R. Dobrushin and S. Shlosman, *Constructive criterion for the uniqueness of a Gibbs field*, Birkhauser, 1985.

[41] C. Domb, *On the theory of cooperative phenomena in crystals*, Advances in Physics 9 (1960), no. 35, 245–361.

[42] M. Dyer, L. Goldberg, and M. Jerrum, *A complexity dichotomy for hypergraph partition functions*, Computational Complexity 19 (2010), 605–633.

[43] G. Elek and G. Lippner, *Borel oracles. an analytical approach to constant-time algorithms*, Proc. Amer. Math. Soc, vol. 138, 2010, pp. 2939–2947.

[44] J. Essam, J. Lin, and P. Taylor, *Potts model on the Bethe lattice with mixed interactions*, Physical Review E 52 (1995), no. 1, 44.

[45] M. Fisher and D. Gaunt, *Hard-sphere lattice gases. i. plane-square lattice*, The Journal of Chemical Physics 43 (1965), no. 8, 2840–2863.

[46] C. Fortuin, P. Kasteleyn, and J. Ginibre, *Correlation inequalities on some partially ordered sets*, Communications in Mathematical Physics 22 (1971), no. 2, 89–103.

[47] G. Fredrickson and H. Andersen, *Kinetic Ising model of the glass transition*, Phys. Rev. Lett. 53 (1984), 1244–1247.

[48] A. Galanis, Q. Ge, D. Stefankovic, E. Vigoda, and L. Yang, *Improved inapproximability results for counting independent sets in the hard-core model*, Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques (Leslie Goldberg, Klaus Jansen, R. Ravi, and Jose Rolim, eds.), Lecture Notes in Computer Science, vol. 6845, Springer Berlin / Heidelberg, 2011, pp. 567–578.
[76] F. Krzakala, M. Tarzia, and L. Zdeborová, *Lattice model for colloidal gels and glasses*, Phys. Rev. Lett. 101 (2008), 165702.

[77] D. Landau, *Critical and multicritical behavior in a triangular-lattice-gas Ising model: Repulsive nearest-neighbor and attractive next-nearest-neighbor coupling*, Physical Review B 27 (1983), no. 9, 5604.

[78] J. Lauer and N. Wormald, *Large independent sets in regular graphs of large girth*, J. Comb. Theory Ser. B 97 (2007), no. 6, 999–1009.

[79] M. Lelarge, *A new approach to the orientation of random hypergraphs*, Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA ’12, SIAM, 2012, pp. 251–264.

[80] L. Li, P. Lu, and Y. Yin, *Approximate counting via correlation decay in spin systems*, Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA ’12, SIAM, 2012, pp. 922–940.

[81] S.Z. Li, *Markov random field modeling in image analysis*, Springer, 2009.

[82] L. Lovász, *Large networks and graph limits*, vol. 60, AMS Bookstore, 2012.

[83] B. Luen, K. Ramanan, and I. Ziedins, *Nonmonotonicity of phase transitions in a loss network with controls*, Annals of Applied Probability 16 (2006), no. 3, 1528 – 1562.

[84] R. Lyons and F. Nazarov, *Perfect matchings as iid factors on non-amenable groups*, European Journal of Combinatorics 32 (2011), no. 7, 1115 – 1125.

[85] M. Mézard, G. Parisi, and R. Zecchina, *Analytic and algorithmic solution of random satisfiability problems*, Science 297 (2002), no. 5582, 812–815.

[86] M. Mezard and M. Tarzia, *Statistical mechanics of the hitting set problem*, Phys. Rev. E 76 (2007), 041124.

[87] P. Mitra, K. Ramanan, A. Senhupta, and I. Ziedins, *Markov random fields of multicasting in tree networks*, Annals of Applied Probability 34 (2002), no. 1, 58 – 84.

[88] A. Montanari, E. Mossel, and A. Sly, *The weak limit of Ising models on locally tree-like graphs*, Probability Theory and Related Fields (2012), 1–21.

[89] E. Mossel, D. Weitz, and N. Wormald, *On the hardness of sampling independent sets beyond the tree threshold*, Probability Theory and Related Fields 143 (2009), 401–439, 10.1007/s00440-007-0131-9.

[90] D.S. Johnson M.R. Garey, *Computers and intractability*, 1979.

[91] C. Nair and P. Tetali, *The correlation decay (cd) tree and strong spatial mixing in multi-spin systems*, arXiv preprint math/0701494 (2007).

[92] C.M. Newman, *Asymptotic independence and limit theorems for positively and negatively dependent random variables*, Lecture Notes-Monograph Series (1984), 127–140.

[93] P. Olofsson, *Branching processes with local dependencies*, The Annals of Applied Probability 6 (1996), no. 1, 238–268.

[94] G. Parisi and F. Zamponi, *Mean-field theory of hard sphere glasses and jamming*, Rev. Mod. Phys. 82 (2010), 789–845.

[95] R. Pemantle, *Automorphism invariant measures on trees*, The Annals of Probability (1992), 1549–1566.

[96] R. Pemantle, *Towards a theory of negative dependence*, Journal of Mathematical Physics 41 (2000), no. 3, 1371–1390.

[97] F.H. Ree and D.A. Chesnut, *Phase transition of a hard-core lattice gas. the square lattice with nearest-neighbor exclusion*, The Journal of Chemical Physics 45 (1966), no. 11, 3983–4003.

[98] O. Rivoire, G. Biroli, O.C. Martin, and M. Mezard, *Glass models on Bethe lattices*, The European Physical Journal B - Condensed Matter and Complex Systems 37 (2003), 55–78.

[99] U. Rozikov, *Gibbs measures on cayley trees: Results and open problems*, Reviews in Mathematical Physics 25 (2013), no. 01.

[100] L. K. Runnels and L. L. Combs, *Exact finite method of lattice statistics. i. square and triangular lattice gases of hard molecules*, The Journal of Chemical Physics 45 (1966), no. 7, 2482–2492.

[101] L.K. Runnels, *Phase transition of a Bethe lattice gas of hard molecules*, Journal of Mathematical Physics 8 (1967), no. 10, 2081–2087.

[102] J. Salez, *The cavity method for counting spanning subgraphs subject to local constraints*, Random Structures and Algorithms 43 (2013), no. 3, 377–397.

[103] D. Shah, J. Shin, and P. Tetali, *Medium access using queues*, Proceedings of FOCS 2011, 2011.

[104] D. Sherrington, *Physics and complexity*, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 368 (2010), no. 1914, 1175–1189.
[105] J. Shin, Complexity of bethe approximation, International Conference on Artificial Intelligence and Statistics, 2012, pp. 1037–1045.
[106] A. Sinclair, P. Srivastava, and M. Thurley, Approximation algorithms for two-state anti-ferromagnetic spin systems on bounded degree graphs, Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, SODA ’12, SIAM, 2012, pp. 941–953.
[107] A. Sly, Computational transition at the uniqueness threshold, Foundations of Computer Science (FOCS), 2010 51st Annual IEEE Symposium on, oct. 2010, pp. 287 –296.
[108] A. Sly and N. Sun, The computational hardness of counting in two-spin models on d-regular graphs, Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on, IEEE, 2012, pp. 361–369.
[109] F. Spitzer, Markov random fields on an infinite tree, The Annals of Probability 3, no. 3, pp. 387–398.
[110] H. Temperley, Application of the Mayer method to the melting problem, Proceedings of the Physical Society 74 (1959), no. 2, 183.
[111] _, Improvements in the lattice model of a liquid, Proceedings of the Physical Society 77 (1961), no. 3, 630.
[112] _, Statistics of a two-dimensional lattice gas. i, Proceedings of the Physical Society 80 (1962), no. 4, 813.
[113] H. Tjelmeland and J. Besag, Markov random fields with higher-order interactions, Scandinavian Journal of Statistics 25 (2001), no. 3, 415–433.
[114] J. van den Berg and A. Gandolfi, BK-type inequalities and generalized random-cluster representations, Probability Theory and Related Fields (2012), 1–25.
[115] J. van den Berg and H. Kesten, Inequalities with applications to percolation and reliability, Journal of applied probability (1985), 556–569.
[116] J. van den Berg and J.E. Steif, Percolation and the hard-core lattice gas model, Stochastic Processes and their Applications 49 (1994), no. 2, 179 – 197.
[117] V. van Kerrebrouck, Dynamical behavior of the Biroli-Mezard model, Journal of Physics: Conference Series 40 (2006), no. 1, 59.
[118] J. Vannimenus, Modulated phase of an Ising system with competing interactions on a Cayley tree, Zeitschrift für Physik B Condensed Matter 43 (1981), no. 2, 141–148.
[119] M. Weigt and A. K. Hartmann, Glassy behavior induced by geometrical frustration in a hard-core lattice-gas model, EPL (Europhysics Letters) 62 (2003), no. 4, 533.
[120] D. Weitz, Mixing in time and space for discrete spin systems, Ph.D. thesis, UNIVERSITY OF CALIFORNIA, 2004.
[121] _, Combinatorial criteria for uniqueness of Gibbs measures, Random Structures and Algorithms 27 (2005), no. 4, 445–475.
[122] D. Weitz, Counting independent sets up to the tree threshold, Proceedings of the 38th ACM Symposium on the Theory of Computing (STOC), 2006.
[123] Y. Yin and C. Zhang, Approximate counting via correlation decay on planar graphs, arXiv preprint arXiv:1207.3564 (2012).
[124] S. Zachary, Countable state space Markov random fields and Markov chains on trees, The Annals of Probability 11, no. 4, pp. 894–903.
[125] _, Bounded, attractive and repulsive Markov specifications on trees and on the one-dimensional lattice, Stochastic processes and their applications 20 (1985), no. 2, 247–256.