Spatial Mixing for Independent Sets in Poisson Random Trees

Varsha Dani∗ Thomas P. Hayes∗ Cristopher Moore†

Abstract

We consider correlation decay in the hard-core model with fugacity λ on a rooted tree T in which the arity of each vertex is independently Poisson distributed with mean d. Specifically, we investigate the question of which parameter settings (d, λ) result in strong spatial mixing, weak spatial mixing, or neither. (In our context, weak spatial mixing is equivalent to Gibbs uniqueness.) For finite fugacity, a zero-one law implies that these spatial mixing properties hold either almost surely or almost never, once we have conditioned on whether T is finite or infinite.

We provide a partial answer to this question, which implies in particular that

1. As d → ∞, weak spatial mixing on the Poisson tree occurs whenever λ < f(d) − o(1) but not when λ is slightly above f(d), where f(d) is the threshold for WSM (and SSM) on the d-regular tree. This suggests that, in most cases, Poisson trees have similar spatial mixing behavior to regular trees.

2. When 1 < d ≤ 1.179, there is weak spatial mixing on the Poisson(d) tree for all values of λ. However, strong spatial mixing does not hold for sufficiently large λ. This is in contrast to regular trees, for which strong spatial mixing and weak spatial mixing always coincide.

For infinite fugacity SSM holds only when the tree is finite, and hence almost surely fails on the Poisson(d) tree when d > 1. We show that WSM almost surely holds on the Poisson(d) tree for d < e^{1/\sqrt{2}}/\sqrt{2} = 1.434..., but that it fails with positive probability if d > e.

∗University of New Mexico, Dept. of Computer Science. email: varshadani@gmail.com, hayes@cs.unm.edu
†Santa Fe Institute. email: moore@santafe.edu
1 Introduction

Spatial mixing, or the decay of correlations between spins in a spin system, is a fundamental question of interest in statistical physics. It is intimately related to temporal mixing for the corresponding Glauber dynamics Markov chain, which means fast convergence to its equilibrium distribution.

There are two flavors of spatial mixing: strong and weak (see Section 2.3 for definitions.) For our purposes, weak spatial mixing is equivalent to Gibbs uniqueness, another fundamental concept from statistical physics.

The hard core model defines a distribution over the independent sets of a graph $G$ in terms of a fugacity $\lambda > 0$. When $G$ is finite and $\lambda = 1$, this is the uniform distribution. More generally, an independent set $S$ has a probability proportional to $\lambda^{|S|}$, so that when $\lambda > 1$, the distribution is biased towards larger independent sets, and when $\lambda < 1$, it is biased towards smaller ones. By convention, when $\lambda = +\infty$, the conditional distribution on finite subgraphs is uniform over all independent sets of maximum size.

In computer science, the problem of sampling from this distribution when $\lambda = 1$ is well-known to be poly-time equivalent to the problem of approximately counting the independent sets of a graph, which is known to be a hard problem in general. We refer the reader to recent work by Sly and Sun [7] for further hardness results.

A seminal paper of D. Weitz [10] found that the infinite regular $d$-ary tree has the same threshold for weak and strong spatial mixing, namely $\lambda = d^d/(d-1)^{d+1} \sim e/d$. More importantly, this is a worst case: every other graph of maximum degree $d + 1$ also exhibits WSM and SSM for all $\lambda$ up to the aforementioned threshold. At the time, this established the strongest positive results for spatial mixing for a wide variety of graphs, including, for instance, the square grid.

Brightwell, Hägström and Winkler [2] showed that there are graphs, even trees, for which the property of WSM is non-monotone as a function of $\lambda$. That is, increasing $\lambda$ can actually decrease the extent to which correlations travel over long distances, and so WSM holds at sufficiently small and sufficiently large $\lambda$, but not in between. They even give a more complicated construction (not a tree) for which the hard-core model exhibits WSM iff $\lambda \in (0, \lambda_1] \cup [\lambda_2, \lambda_3]$ where $\lambda_1 < \lambda_2 < \lambda_3 < \infty$.

Restrepo et al. [6] showed that for some graphs, such as the planar square lattice, SSM occurs at higher $\lambda$ than for the 4-regular tree. Recently, Vera, Vigoda and Yang [8] have shown that the tree of self-avoiding walks on the square lattice contains a subtree which has WSM but not SSM, at a still higher value of $\lambda$, but still below the conjectured critical value for the square lattice. (See [8, Lemmas 4, 7].) This suggests that it may not be such an uncommon phenomenon for WSM to occur without SSM. In Section 2.5 we exhibit an example of an infinite tree which has WSM for all $\lambda > 0$ but does not have SSM for any $\lambda > 4$.

We consider random Poisson trees, in which every vertex has an independent, identically Poisson distributed number of children. This is a natural model because of its connection to sparse Erdős-Rényi random graphs, $G(n, p)$. When $d = \Theta(1)$ and $p = d/n$, for large $n$, the local structure of balls of volume $o(\sqrt{n})$ is well approximated by a Poisson tree.

It is natural, given an infinite graph, to consider the following threshold conjecture: There is a threshold $\lambda_{\text{crit}}$ such that WSM holds if and only if $\lambda < \lambda_{\text{crit}}$. The analogous conjecture with SSM in place of WSM is also interesting. For instance, both conjectures are known to be true with

$$\lambda_{\text{crit}} = \frac{\Delta^\Delta}{(\Delta - 1)^{\Delta + 1}}$$

when $G$ is the infinite regular $\Delta$-ary tree. Note that $\lambda_{\text{crit}}$ is asymptotically $e/\Delta$ as $\Delta \to \infty$. 

1
Brightwell, Häggström and Winkler [2] have constructed other graphs $G$ for which the WSM conjecture is false.

Understanding weak spatial mixing for regular $d$-ary trees is relatively straightforward. Note that, in general, the conditional probability $a_v$ that node $v$ is unoccupied, given that the parent of $v$ is unoccupied, obeys the recurrence

$$a_v = \frac{1}{1 + \lambda \prod_{w} a_w}$$

where $w$ ranges over the children of $v$. Since for the $d$-regular tree all the $a_w$ are equal, the problem boils down to understanding the stability of the fixed point of the iterated function $f_d(a) = (1 + \lambda a_d)^{-1}$.

For random Poisson trees, the situation is more complicated. Since the various subtrees of a node are no longer identical, but merely identically distributed, we now need to, in effect, consider a recurrence relation on distributions rather than on real values.

Intuitively, we may expect a Poisson$(d)$ tree to behave something like a regular $d$-ary tree. We show that this is the case for large $d$, proving that WSM holds for $\lambda = c/d$ if $c < e$ but not if $c > e$. On the other hand, for small $d$, there are several ways in which this is not the case. In particular,

1. There are some settings of the Poisson parameter and fugacity for which there is weak mixing (almost surely) but not strong mixing (with positive probability). In particular, this happens when the expected degree is 1.1 and the fugacity is sufficiently large.

2. For sufficiently small $d$, but still greater than 1, the Poisson tree exhibits WSM for all values of $\lambda$, even $\lambda = \infty$.

3. One might have thought that the phenomenon exploited in [2], where increasing $\lambda$ causes childless nodes to be occupied with high probability, which then cuts off the flow of information from their siblings up through their parent, is pathological. In fact, we will see that, for small enough $d$, this phenomenon is pervasive in Poisson trees.

4. As a consequence, some of our results are non-monotonic, in that for $1.179 < d < 1.434$, we know WSM occurs at $\lambda = \infty$, and for sufficiently small $\lambda$, but we don’t know what happens in between.

Before summarizing our main results, we begin by observing the following zero-one law for spatial mixing on Poisson trees with finite fugacity.

**Theorem 1.1.** For all $d > 1$ and $0 < \lambda < \infty$, conditioned on Poisson$(d)$ being infinite, the probability that the hard-core model on Poisson$(d)$ with fugacity $\lambda$ has WSM (resp. SSM) is either zero or one.

Note that for $d \leq 1$, the Poisson tree Poisson$(d)$ is almost surely finite.

In light of this zero-one law (proved in Section 2.4) for finite $\lambda$, we focus our attention on the question of which parameter settings $(d, \lambda)$ result in SSM, WSM, or neither.

We summarize our results for finite fugacities. See Figure 1 for graphs of some of the functions involved. Overall, our results describe where WSM and SSM occur or do not occur in various regions of the $(d, \lambda)$ plane.
**Theorem 1.2.** The hard-core model with fugacity $\lambda < \infty$ on Poisson($d$) has the following properties, almost surely, conditioned on being infinite.

1. WSM if $d < 1.179...$, for any $0 < \lambda < \infty$.
2. SSM if $\lambda < \begin{cases} \frac{4d^2}{(d^2-1)^2} & \text{when } d < \sqrt{2 + \sqrt{5}} \\ \frac{3+\sqrt{1+4d^2}}{2d^2-4} & \text{otherwise.} \end{cases}$
3. WSM if $\lambda < \frac{e^{-o(1)}}{d}$, as $d \to \infty$.
4. No WSM if $\lambda = \frac{e^{+o(1)}}{d}$, as $d \to \infty$.

Thus, if the WSM threshold conjecture is true for the hard-core model on the Poisson tree, then we have shown that the location of the threshold is, for large $d$, asymptotically the same as for $d$-regular trees. On the other hand, unlike $d$-regular trees, there is a range of parameters for which the Poisson tree exhibits WSM but not SSM. Specifically, for $1 < d < 1.179...$ and for sufficiently large $\lambda$, the Poisson tree almost surely has WSM but not SSM, conditioned on being infinite; see Remark 5.2. We conjecture that the Poisson tree almost surely exhibits SSM up to a threshold that is asymptotically $e/d$, the same as for $d$-regular trees.

We also study spatial mixing properties of the Poisson($d$) tree when the fugacity is infinite. The following theorem summarizes our results for this case.

**Theorem 1.3.** There exists a constant $d^* > 1$ such that for all $1 < d < d^*$, the hard-core model on Poisson($d$) with fugacity $\lambda = +\infty$ exhibits WSM but not SSM, almost surely, conditioned on being infinite. Furthermore, we prove that the largest such $d^*$ is at least $e^{1/\sqrt{2}}/\sqrt{2} = 1.434...$, and at most $e = 2.718...$.

We conjecture that $e$ is the correct value for $d^*$. We prove Theorem 1.3 in Section 3.

2 Preliminaries

2.1 The Poisson Tree

Let $d > 0$. Consider a recursively generated random tree $T$, where we sample a non-negative integer $X$ from the Poisson distribution with mean $d$, namely,

$$(\forall i \geq 0) \quad \text{Prob}(X = i) = \frac{e^{-d}d^i}{i!}$$

and define $X$ to be the number of children of the root of $T$. Recursively, let each of these children be the root of a subtree sampled independently in the same manner. We call this the Poisson tree of average arity $d$, and denote it by Poisson($d$).

For $d \leq 1$, this tree is almost surely finite. For $d > 1$, the tree is infinite with positive probability, but unlike an infinite $d$-regular tree, it has leaves: indeed, each non-root node has probability $e^{-d}$ to be a leaf. (The root itself is a leaf with probability $e^{-d}(1 + d)$, since it is a also a leaf if it has only one child.)
\( \lambda = \infty \)

(a) For \( \lambda = +\infty \), there is SSM only when \( d \leq 1 \), in which case the tree is almost surely finite. There is WSM for \( d < 1.434... \), but not for \( d > e \).

(b) For finite \( \lambda \), and \( d > 1 \) there is SSM in the shaded region (to the left of the red curve \( d = 1.179... \) and the blue curve that is asymptotic to \( 1/d \)). The threshold for WSM is asymptotic to the purple curve, which is also the threshold for the regular \( d \)-ary tree.

Figure 1: Illustration for Theorems 1.3 and 1.2

Proposition 2.1. Let \( d > 0 \), and let \( T \) be a Poisson tree \( \text{Poisson}(d) \). For \( R \geq 0 \), let \( f(R) \) denote the number of nodes in level \( R \) of \( T \). Then, almost surely,

\[
\lim_{R \to \infty} \frac{f(R)}{R^{2/dR}} = 0.
\]

Proof. By Markov’s inequality, for each \( R \), we have \( \text{Prob} \left( \frac{f(R)}{R^{2/dR}} > R^{-1/2} \right) < R^{-3/2} \). A union bound implies that there are almost surely only finitely many exceptions. \( \square \)
2.2 The Hard-Core Model (Independent Sets)

In Statistical Mechanics, systems involving large numbers of interacting particles are often modeled by a spin system. This is defined in terms of an underlying graph, often an infinite lattice, whose vertices are called sites, each of which can be assigned a spin from some finite set $Q$. A configuration is a function assigning a spin to each site. A Gibbs measure is a probability distribution over configurations that satisfies a consistency criterion on all finite “patches”, or subsets of vertices. Specifically, for each finite subset $\Lambda \subset V$, with boundary $\partial \Lambda = \{v \in \Lambda \mid \exists\{v, w\} \in E_G, w \notin \Lambda\}$, and each boundary condition $\sigma : \partial \Lambda \rightarrow Q$, the conditional distribution of the Gibbs measure restricted to $\Lambda$, conditioned on agreeing with $\sigma$ on $\partial \Lambda$, is prescribed. Although it is known [3] that a Gibbs measure always exists, it is not, in general, guaranteed to be unique. Indeed, many spin systems undergo a phase transition, where some critical threshold for a defining parameter determines whether Gibbs uniqueness holds or not.

In the hard-core model, the spins correspond to a site being “occupied” or “unoccupied”. Adjacent sites are not allowed to both be occupied, and so configurations are independent sets of the graph. Configurations have probabilities that are exponential in the number of occupied sites: an independent set $S$ has probability $\frac{1}{Z} \lambda^{|S|}$, where $\lambda > 0$ is a parameter of the system called the fugacity, and the normalizing constant $Z$ is called the partition function.

We will also be concerned with the case $\lambda = +\infty$, in which, on finite patches, the prescribed distribution is considered to be uniform over all independent sets of the maximum possible size.

2.3 Weak and Strong Spatial Mixing

“Spatial mixing” refers to a phenomenon wherein correlations between spins decay as the distance between the vertices increases.

Let $\Lambda$ be any set of vertices, let $\Psi \supset \Lambda$ be a containing set of vertices and let $\sigma, \tau : \partial \Psi \rightarrow Q$ be two boundary configurations for the larger set. We are interested in the total variation distance between the marginal distributions on configurations over $\Lambda$, conditioned on agreeing with $\sigma$ or $\tau$. Now, consider infinite families of such triples $(\Psi, \sigma, \tau)$, indexed by the positive integers. If

$$\text{dist}(\Lambda, \partial \Psi) \rightarrow \infty \text{ implies } \|\mu_{\Psi}^\sigma - \mu_{\Psi}^\tau\|_\Lambda \rightarrow 0,$$

then we say that weak spatial mixing (WSM) holds. If

$$\text{dist}(\Lambda, \sigma \oplus \tau) \rightarrow \infty \text{ implies } \|\mu_{\Psi}^\sigma - \mu_{\Psi}^\tau\|_\Lambda \rightarrow 0,$$

where $\sigma \oplus \tau$ denotes the set of vertices on which $\sigma$ and $\tau$ disagree, then we say that strong spatial mixing (SSM) holds.

Intuitively, weak spatial mixing requires the effect of changing some spins to decay with distance, assuming all closer vertices are unconstrained, while strong spatial mixing requires the effect to decay even when some of the closer vertices are “frozen” in an adversarial way (which must be the same for both boundary conditions). Obviously, SSM implies WSM.

The above definition of weak spatial mixing is easily seen to be equivalent to Gibbs uniqueness (see [2] Proposition 2.2). We note that several alternative definitions of spatial mixing appear in the literature. In some of these, the rate of decay of correlation is required to be exponential in the distance, rather than merely tending to zero. All of our results apply in this setting as well. In some definitions of spatial mixing, one either restricts attention to the effect on a single vertex, i.e.,
\( \Lambda = \{v\} \), and/or one restricts attention to boundary conditions that disagree on a single boundary vertex. In the case when the convergence rate is required to be exponential, and moreover the graph is such that boundary sizes grow subexponentially, the restriction to a single disagreement doesn’t matter (by a union bound). For trees, however, boundary sizes often grow exponentially, in which case the specific rate of exponential decay of the effect of a vertex would matter. On the other end, there are spin systems where restricting \( \Lambda \) to be a singleton makes WSM hold trivially, even when it does not hold for larger sets \( \Lambda \).

**Remark 2.2.** In the case of independent sets, it is well known that SSM on a graph \( G \) is equivalent to WSM on all subgraphs of \( G \), because any boundary vertices that are frozen to be unoccupied can equivalently be deleted, and any that are occupied can equivalently have all their neighbors deleted.

For independent sets on a tree, there is a simpler characterization of spatial mixing in terms of non-occupation probabilities. Specifically, let \( T \) be a finite tree with a designated root vertex \( r \). For each vertex \( v \), let \( a_v \) denote the conditional probability that \( v \) is unoccupied, conditioned on \( v \)’s parent (if any) being unoccupied. These non-occupation probabilities satisfy the recurrence (1).

When \( T \) is an infinite rooted tree, we will suppose that an adversary has set arbitrary values \( a_z \in [0,1] \) at level \( R + 1 \). In this case, we treat (1) as a recursive definition for \( a_v \), where \( v \) is at distance \( \leq R \) from the root. If for all sequences of boundary conditions, as \( R \to \infty \), \( a_v \) converges to a well-defined limit \( a_v^* \), then we call \( a_v^* \) the non-occupation probability of \( v \).

Since the righthand side of (1) is a decreasing function of each of the \( a_w \), it follows by induction that, for any radius \( R \), the extreme values of any \( a_v \) are induced by the all-zeros and the all-ones boundaries. Thus, when proving the existence of \( a_v^* \), it suffices to consider boundary conditions of this type.

**Proposition 2.3.** For the hard-core model on any infinite tree, the following are equivalent:

1. For all vertices \( v \), there is a well-defined non-occupation probability \( a_v^* \).
2. Weak spatial mixing occurs.
3. There is a unique Gibbs distribution.

Furthermore, when the fugacity, \( \lambda \), is finite, this condition is equivalent to the three above:

4. For the root \( r \), there is a well-defined non-occupation probability \( a_r^* \).

**Proof sketch.** The equivalence of statements 2 and 3 is shown in [9, Proposition 2.2].

To see that statement 3 implies statement 1, let \( v \) be any vertex in the tree. By Gibbs uniqueness, if we consider larger and larger balls centered at \( v \), the effect of the boundary configuration goes to zero, and there is a well-defined marginal distribution on the spins of \( v \) and its parent. Essentially by definition, \( a_v^* \) must equal the probability that \( v \) is unoccupied, conditioned on its parent being unoccupied. Note that the effect of all spins outside the subtree under \( v \) can only influence the spin of \( v \) through the spin of its parent, which we have conditioned on.

---

\([1] Here is a rather contrived example. Start with any 2-spin system for which WSM does not hold. Replace each vertex with a pair of vertices, and decree that if the original vertex had spin 1, the pair have the same spin, but uniformly random 1 or 2. If the original vertex had spin 2, the pair have opposite spins, again uniformly random. We omit the details.
To see that statement 1 implies statement 3, suppose for contradiction that there were two distinct Gibbs measures. Then their marginals must differ on some finite patch $\Lambda \subset V$. Starting with the root vertex $v_0$, let $v_0, v_1, v_2, \ldots$ be a breadth-first traversal of the tree. Then, for some configuration $\sigma$, and some finite $i$, the probability of $\sigma$ restricted to $v_0, \ldots, v_i$, must differ under the two Gibbs measures. Choose $i$ to be minimal with respect to this property. In this case, $\text{Prob}(\sigma(v_i) \mid \sigma(v_0), \ldots, \sigma(v_{i-1}))$ differs under the two distributions. The only way this can happen is if the parent of $v_i$ is unoccupied under $\sigma$, in which case the above conditional probability must equal $a_{v_i}^*$ in both measures, a contradiction.

Statement 4 is a special case of statement 1, corresponding to weak spatial mixing at the root (since the root has no parent). Hence statement 1 implies statement 4. Statement 4 implies statement 1 when $\lambda$ is finite, because the recurrence (1) holds at every vertex $v$, under every boundary condition. It follows that the limit $a_v^*$ cannot exist unless the limits $a_w^*$ exist for every child vertex $w$.

2.4 Zero-One Law

In this section, we prove Theorem 1.1. To this end, we say that a boolean predicate, $S$, defined on rooted trees has property $R$ if, for every tree $T$, $S(T)$ holds if and only if $S(T')$ holds for every induced proper subtree $T'$ of $T$. Note that any predicate with property $R$ must hold for every finite tree, by induction.

Examples:

1. “$T$ is finite” has property $R$.

2. When $\lambda < \infty$, the property “The hard-core model for $T$ has WSM,” has property $R$, in light of Proposition 2.3.

3. Similarly, for $\lambda < \infty$, “The hard-core model for $T$ has SSM” also has property $R$.

Lemma 2.4. Let $A$ be a predicate with property $R$. Then, for a random Poisson($d$) tree, conditioned on being infinite, the conditional probability that $A$ holds is either zero or one.

Proof. Let $p$ denote the probability that $A(T)$ holds. Since $A(T)$ holds iff $A(T')$ holds for each of the top-level subtrees $T'$ of $T$, and the number of such subtrees is Poisson distributed with mean $d$, we have

$$p = \sum_{i \geq 0} e^{-d} \frac{d^i}{i!} p^i = e^{d(p-1)}.$$

This equation is easily seen to have the following solutions. $p = 1$ is always a solution. When $d \leq 1$, this is the only solution in $[0, 1]$. When $d > 1$, there is a second solution $p^* < 1$, which equals the probability that Poisson($d$) is finite. Since predicates with property $R$ hold for all finite trees, it follows that $p = p^* + (1 - p^*)q$, where $q$ is the conditional probability of $A(T)$ conditioned on $T$ being infinite. Hence $q$ is 0 or 1, completing the proof. \qed
Theorem 1.1 follows as an immediate corollary in light of the above observation that having WSM (resp. SSM) is a predicate with property $R$.

2.5 Alternating Trees

Consider the infinite rooted tree $T$ with alternating layers of degree $d > 12$ and degree 2 vertices, i.e., the root has $d$ children, each of whom have two children, each of whom have $d$ children and so on. In this section, we examine the question of weak spatial mixing for such trees. Notice that, since $T$ contains a complete binary tree, it does not have SSM for $\lambda > 4$.

Theorem 2.5. $T$ has weak spatial mixing for all $\lambda \leq \frac{d}{4 \ln d}$.

Proof. Consider the function

$$g(x) = \frac{1}{1 + \lambda (1 + \lambda x^2)^{-d}}$$

which determines the values $a_w$ of the nodes $w$ at an even depth $r$ when the values of $a_w$ for $w$ at depth $r + 2$ have been set to $x$. Since $f$ is the composition of two monotone decreasing functions, it is monotone increasing. Observe that

$$g'(x) = -g(x)^2 \cdot \lambda (-d) (1 + \lambda x^2)^{-d-1} \cdot 2\lambda x = 2xg(x)^2\lambda^2 d (1 + \lambda x^2)^{-d-1}.$$  

At the boundary, the adversary can set the $a_w$s to any values in $[0, 1]$. However, recall that by (1), for every level above that, these values will lie in $[\frac{1}{1 + \lambda}, 1]$. Thus $x$ and $g(x)$ are both between $\frac{1}{1 + \lambda}$ and 1, and we have

$$|g'(x)| \leq 2\lambda^2 d \left(1 + \frac{\lambda}{(1 + \lambda)^2}\right)^{-d-1} \leq 2\lambda^2 d e^{-\lambda(d+1)/(1+\lambda)^2}.$$  

Since $d$ is large, when $\lambda < \frac{d}{4 \ln d}$, $|g'(x)|$ is bounded below 1 for all $x$ in $[\frac{1}{1 + \lambda}, 1]$. It follows that $g$ is a contraction mapping with a unique fixed point $a^*$ in $[\frac{1}{1 + \lambda}, 1]$, and moreover, for any $x \in [\frac{1}{1 + \lambda}, 1]$, the sequence $\{a_n\}$ defined recursively by

$$a_0 = x; \quad a_n = g(a_{n-1})$$

converges to $a^*$.

Now suppose that the adversary sets the values of all the nodes at depth $R$ to be either all 0s or all 1s. Then applying (1) results in the same values at all the nodes at depth $R'$ which is the deepest even level above $R$. applying the function $g$ repeatedly from then on, we see that as $R$ goes to infinity, the value at the root, $a_r$ converges to $a^*$. By the monotonicity of (1) with respect to each $a_w$, $a_r$ converges to $a_r$ for all settings of the nodes at depth $R$ by the adversary. It follows that $T$ has weak spatial mixing for all $\lambda < \frac{d}{4 \ln d}$.

On the other hand, $T$ contains the 2-regular tree as a subtree. Thus $T$ does not have strong spatial mixing for any $\lambda > 4$. Thus there is a large range of $\lambda$ for which it has weak, but not strong, spatial mixing.
Now consider the infinite tree $T'$ all of whose vertices at depth $r$ have $d(r)$ children, where

$$d(r) = \begin{cases} 2 & \text{if } r \text{ is odd} \\ 2^r + 1 & \text{if } r \text{ is even} \end{cases}$$

(or any increasing function of $r$ on the even levels should be fine.) As before $T'$ contains the complete infinite binary tree as a subtree, and so has no strong spatial mixing above $\lambda = 4$. However, it is easily seen that $T'$ has weak spatial mixing for all $\lambda$.

### 3 Infinite Fugacity: Maximum Independent Sets

In this section we derive upper and lower bounds on the Weak Spatial Mixing threshold in the infinite fugacity case. We note that at infinite fugacity, the Poisson tree with average degree $d$ does not exhibit Strong Spatial Mixing unless $d < 1$ in which case the tree is almost certainly finite.

When $\lambda = \infty$, equation (1) is potentially indeterminate, so a good first step would be to re-examine the definition of the model. The defining notion is that, for any finite patch with boundary condition, the distribution should be uniform over independent sets of the maximum possible size. However, in order to understand whether this condition leads to a unique Gibbs measure, we still want a recurrence for the probabilities $a_v$, that $v$ is unoccupied, conditioned on its parent being unoccupied.

There are a couple of good ways to deal with the indeterminism in (1). First, we can do arithmetic in the ring $\mathbb{R}[\lambda^{-1}]/(\lambda^{-2})$, where we treat $\lambda^{-1}$ as an infinitesimal, that can be ignored when added to any non-zero real number, and whose square is treated as zero. The expression $1/(1 + \lambda \prod_w a_w)$ evaluates to:

1. 1 whenever two or more of the $a_w$ are infinitesimal,

2. $\lambda^{-1} \prod_w a_w^{-1}$ if none of the $a_w$ are infinitesimal, and

3. $1/(1 + c_{w'} \prod_{w:w\neq w'} a_w)$ if exactly one vertex $w'$ has the infinitesimal value $a_{w'} = c_{w'} \lambda^{-1}$.

The second approach is to treat the above infinitesimals as zeros, but to reconstruct the coefficient $c_{w'}$ in case 3, from the values on the children of $w'$. This gives the formula

$$a_v = \frac{\prod_z a_z}{\prod_z a_z + \prod_{w\neq w'} a_w},$$

where $z$ ranges over the children of the unique child $w'$ with $a_{w'} = 0$.

We will refer to vertex $v$ as “large” when $a_v$ evaluates to a non-zero real number, and as “small” when it evaluates to an infinitesimal (or zero, if you prefer that viewpoint). There is a third possibility, namely that no finite piece of the tree suffices to determine whether $a_v$ is large or small, because of infinite descent; in this case, we say $a_v$ is “unlabeled.” Our rules above now give a particularly easy recursive description of when a node is large, small, or unlabeled:

a. If one or more children of $v$ is small, then $v$ is large.

b. If all children of $v$ are large, then $v$ is small.
Figure 2: An example of Karp-Sipser labeling.
c. Otherwise, no child of $v$ is small, and at least one child is unlabeled. In this case, $v$ is unlabeled.

We call this process Karp-Sipser labeling, since it is a bottom-up version of the Karp-Sipser algorithm \[4\], which generates an independent set $S$ in a graph by choosing a vertex $v$ of degree 1 or 0, placing $v$ in $S$, and removing $v$ and its neighbor, if any, from the graph. See Figure 2.

Starting from the leaves, which are small, one can work upward through the tree, using rules 1 and 2 to assign labels to all the small and large nodes. The nodes that remain unlabeled after this (infinite) process are the ones we called “unlabeled” above. It is easy to see that, by induction, each unlabeled node sits on top of an infinite leafless subtree of unlabeled nodes. The unlabeled nodes in this tree may also have additional children that are labeled large, who in turn have other children, about which we are not concerned.

Now, suppose we cut off our tree at depth $R$, and set a pair of boundary conditions on these nodes, that respects the labeled nodes, and either occupies all or none of the remaining boundary nodes. More precisely, under the first boundary condition, the occupied nodes at depth $R$ are exactly the ones labeled "small,” while under the second boundary condition, the unlabeled nodes are also occupied.

In this case, it is easy to see by induction that, subject to this new boundary condition, all the labeled nodes at depth $< R$ will keep their original labels, and therefore the previously unlabeled nodes at depth $i < R$ will either be all large or all small, depending on the parity of $R - i$ and which of the two boundary conditions was set.

Now let $p_S$, $p_L$ and $p_U$ denote the probabilities that the root is labeled ‘small’, ‘large’ or ‘unlabeled’ respectively. We have

\[ p_S + p_L + p_U = 1 \]

Then, by rules a, b, and c above, $p_S$ is the probability that all the root’s children are large, while $p_L$ is the probability that at least one child of the root is small. Since each child is the root of an independently random subtree, which is distributed just as the entire tree is, the number of children that are large or small or unlabeled is Poisson-distributed with mean $dp_L$ or $dp_S$ or $dp_U$ respectively. This gives

\[ p_L = 1 - e^{-dp_S} \quad \text{and} \quad p_S = e^{-d(p_S+p_U)} = e^{-d(1-p_L)} \quad (2) \]

Together, these imply

\[ p_S = e^{-de^{-dp_S}} \]

Letting $f$ denote the function $f(x) = e^{-dx}$, we see that $p_S$ is a fixed point of $f \circ f$. One fixed point of $f \circ f$ is the (unique) fixed point of $f$. Using Lambert’s $W$ function, where $z = W(z)e^{W(z)}$, this fixed point can be written as $W(d)/d$. In fact, when $d \leq e$ this is the only real fixed point. In that case,

\[ p_S = W(d)/d, \quad \text{and} \quad p_L = 1 - f(p_S) = 1 - p_S \]

so that $p_U = 0$ and the root is labeled with probability 1. When $d > e$, on the other hand, the smallest real fixed point of $f \circ f$ is strictly smaller than $W(d)/d$, and is not a fixed point of $f$. In that case, the smallest fixed point is $p_S$, and hence $p_U > 0$, i.e., with constant probability the root remains unlabeled.

We remark that all this corresponds exactly to the rigorous results on the Karp-Sipser algorithm \[4, 1\]. On $G(n, p = d/n)$, if $d < e$ then the algorithm finds a maximal independent set, except for a core that consists w.h.p. of $O(\log n)$ vertex-disjoint cycles.

11
We are now ready to prove our upper and lower bounds.

3.1 Upper Bound

In this section we will analyze the situation when \( \lambda = \infty \) and \( d > e \). Recall that in this case the root has positive probability \( p_U \) to be unlabeled. Moreover, regardless of the root’s label, the number of children of the root that are, respectively, small, large and unlabeled are independent Poisson random variables with parameter, respectively, \( dp_S, dp_L \) and \( dp_U \).

It follows that with positive probability, the root is unlabeled and has at least two unlabeled children (and no small children.) In this event, based on the parity of \( R \), one boundary condition at depth \( R \) forces both those unlabeled children to be occupied while the other forces them both to be unoccupied. Since the independent set must be of maximum size, if both are occupied then the root is forced to be unoccupied, while if both are unoccupied, the root is forced to be occupied. Since these two alternatives remain possible, independent of \( R \), there is no weak spatial mixing at the root. We have shown the following, which implies the second half of Theorem 1.3

**Theorem 3.1.** For \( \lambda = \infty \) and \( d > e \), with positive probability, the Poisson\((d)\) tree does not have WSM at the root.

We remark that if the Poisson tree is infinite, it almost surely contains some node which is unlabeled and has at least two unlabeled children.

3.2 Lower Bound

In this section we analyze the situation when \( d \leq e \). Recall that in this case \( p_S = W(d)/d \) was the unique fixed point of \( f(x) = e^{-dx} \), so that

\[
p_S = e^{-dp_S}
\]

and by (2), \( p_L = 1 - e^{-dp_S} = 1 - p_S \), i.e., the root is labeled as either ‘small’ or ‘large’ with probability 1. Moreover, this labeling obeys the rules that

- all children of a small node are large, and
- at least one child of a large node is small.

To what do these labels correspond? Intuitively, a vertex being labeled ‘small’ or ‘large’ respectively, corresponds to having non-occupation probabilities (as the root of its subtree) that are small or large respectively. For finite \( \lambda \gg 1 \), roughly speaking, this means \( O(1/\lambda) \) or \( \Theta(1) \) respectively. Note however, that this intuition can sometimes be incorrect, for instance a node with very many children, all “large,” may have a large non-occupation probability, even though it receives a label of “small.” Another example where the above intuition fails is for nodes at the root of a subtree isomorphic to a very long path, specifically one of length \( \omega(\sqrt{\lambda}) \). Although the nodes in this path are labelled with alternating “small” and “large” labels, actually almost all the conditional non-occupation probabilities will be approximately \( 1/\sqrt{\lambda} \).

When \( \lambda \) is infinite, this becomes a distinction of zero vs. non-zero. In other words, conditioned on its parent being unoccupied, (or equivalently, looking at it as the root of its subtree), if vertex \( v \) is labeled ‘large’ then there are maximum independent sets on its subtree which do not contain...
v (i.e., there are configurations in which v is unoccupied), and \(a_v > 0\), whereas if \(v\) is labeled ‘small’ then every maximum independent set contains \(v\) (i.e., \(v\) is occupied in all configurations) and \(a_v = 0\).

Now consider a ‘large’ node \(v\) with two or more ‘small’ children. Looking at the recurrence (1), and the rules for arithmetic in the ring \(\mathbb{R}[\lambda^{-1}]/(\lambda^{-2})\), we see that regardless of the non-occupation probabilities of all the other children of \(v\), \(a_v = 1\).

In other words, if \(v\) has two or more children that are probably occupied, then \(v\) is probably empty, regardless of what other children it has. We say in this situation that \(a_v\) is known. More generally, we say that for ‘large’ \(v\), \(a_v\) is known whenever it is determined by a finite subtree of \(v\)’s descendants. In particular, known \(a_v\)s are rational. For technical reasons, we will not say \(a_v\) is known for all \(v\) that are ‘small’, but rather only those \(v\) all of whose children are known.

Let \(\kappa_L\) and \(\kappa_S\) denote the probability that \(a_v\) is large and known, or small and known, respectively. If \(a_v\) is large, it is known either if it has two or more small children, or if all its children are known and exactly one of them is small. If \(a_v\) is small, then it is known if and only if all its children (which are large) are known. This gives us the equations

\[
\begin{align*}
\kappa_L &= 1 - (1 + dp_S) e^{-dp_S} + d \kappa_S e^{-dp_S} e^{-d(p_L - \kappa_L)} \quad (4) \\
\kappa_S &= e^{-dp_S} e^{-d(p_L - \kappa_L)} \quad (5).
\end{align*}
\]

Simplifying and combining with (3) gives the relations

\[
\begin{align*}
p_L - \kappa_L &= d(p_S^2 - \kappa_S^2) \quad (6) \\
\kappa_S &= p_S e^{-d(p_L - \kappa_L)} \quad (7).
\end{align*}
\]

Rearranging terms and once again using (3) we see that

\[
\kappa_L = 1 - (1 + dp_S)p_S + de^{-2d(1-\kappa_L)}
\]

so that \(\kappa_L\) is a fixed point the function

\[
g(x) := 1 - (1 + W(d)) \frac{W(d)}{d} + de^{-2d(1-\kappa_L)}.
\]

The system of equations (6) and (7) always has \((\kappa_L, \kappa_S) = (p_L, p_S)\) as one solution. Additionally, when \(d\) is sufficiently large, there is a second solution where \(\kappa_L < p_L\) and \(\kappa_S < p_S\), corresponding to the fact that for large enough \(d < e\) there are graphs for which even though the root \(v\) is labeled “large”, the actual value of \(a_v\) is not determined by any finite subtree of the Poisson tree. The threshold where these roots appear is the \(d\) such that

\[
g'(p_L) = 1 = 2d^2 p_S^2,
\]

which with (2) implies

\[
d = \frac{e^{1/\sqrt{2}}}{\sqrt{2}} = 1.434\ldots.
\]
4 Finite $\lambda$ case: Lower bound

In this section we will derive a lower bound on the SSM threshold for the Poisson tree. This proves part 2 of Theorem 1.2.

By Remark 2.2, in order to show SSM, it suffices to show WSM for any subtree of Poisson($d$).

Let $T$ be a subtree of Poisson($d$) and let $r$ be the root of $T$. For $R > 0$, let $T_R$ denote the truncation of $T$ to depth $R$, and let $\partial T_R$ denote the boundary of $T_R$, i.e., the vertices of $T$ at depth $R$. We want to study the influence of the non-occupation probability values at $\partial T_R$ (set adversarially) on the value of $a_r$. For notational convenience we will require the adversary to set the values at $\partial T_R$ from $[\frac{1}{1+\lambda}, 1]$. Since the range of the function $x \mapsto \frac{1}{1+\lambda x}$ on $[0, 1]$ is contained in $[\frac{1}{1+\lambda}, 1]$ when $0 < P \leq 1$, this corresponds to allowing the adversary to set values in $[0, 1]$ on $\partial T_{R+1}$.

Recall, from Proposition 2.3, that to show WSM for $T$, it suffices to show that there is a well-defined non-occupation probability $a_r^*$ at the root $r$ of $T$. This, in turn, would follow if the non-occupation probabilities induced at $r$ by setting the vertices in $\partial T_{R+1}$ to all zeroes or all ones converged to the same value as $R \to \infty$.

Let $w$ be a vertex of $\partial T_R$. Suppose the values at all the other vertices in $\partial T_R$ are fixed, and only the value $a_w$ at $w$ is varied. Let $a_r^{[a_w=1/(1+\lambda)]}$ and $a_r^{[a_w=1]}$ be the values of $a_r$ when $a_w$ is set to $\frac{1}{1+\lambda}$ or 1 respectively. Then by the mean value theorem,

$$\left|a_r^{[a_w=1/(1+\lambda)]} - a_r^{[a_w=1]}\right| \leq \max_{a_w \in [\frac{1}{1+\lambda}, 1]} \left|\frac{\partial a_r}{\partial a_w}\right|.$$ 

Now, if $a_r^0$ and $a_r^1$ are the values of $a_r$ when the vertices at depth $R+1$ have been set respectively to all zeroes or all ones (i.e., the vertices in $\partial T_R$ set to all ones or all $\frac{1}{1+\lambda}$) then by varying the values at the boundary vertices one at a time and applying the triangle inequality, we see that

$$\left|a_r^0 - a_r^1\right| \leq \sum_{w \in \partial T_R} \max_{a_w \in [\frac{1}{1+\lambda}, 1]} \left|\frac{\partial a_r}{\partial a_w}\right|$$

Fix $w \in \partial T_R$ and let $r = w_0, w_1, \ldots, w_{R-1}, w_R = w$ be the path from the root to $w$. Let $a_i = a_{w_i}$ and let $P_i = \prod_x a_x$ where the product is taken over all the children $x$ (if any) of $w_i$ other than $w_{i+1}$. Then for all $i$,

$$a_i = \frac{1}{1 + \lambda a_{i+1} P_i}.$$ 

Note that $a_i \geq \frac{1}{1+\lambda} > 0$ for all $i$. Differentiating $a_i$ with respect to $a_{i+1}$, with some algebraic manipulations, we have

$$\frac{\partial a_i}{\partial a_{i+1}} = \frac{-\lambda P_i}{(1 + \lambda a_{i+1} P_i)^2} = \frac{-a_i (1 - a_i)}{a_{i+1}}.$$ 

Repeatedly applying the chain rule, we see that

$$\frac{\partial a_r}{\partial a_w} = \frac{\partial a_0}{\partial a_R} = \prod_{i=0}^{R-1} \frac{\partial a_i}{\partial a_{i+1}} = \prod_{i=0}^{R-1} \frac{-a_i (1 - a_i)}{a_{i+1}} = (-1)^R \frac{a_0}{a_R} \prod_{i=0}^{R-1} (1 - a_i).$$

Since $\frac{1}{1+\lambda} < a_i \leq 1$,

$$\left|\frac{\partial a_r}{\partial a_w}\right| \leq \frac{a_0}{a_R} \prod_{i=0}^{R-1} (1 - a_i) \leq (1 + \lambda) \prod_{i=0}^{R-1} (1 - a_i).$$

$$14$$
Note that \( a_i \geq \frac{1}{1 + \lambda a_{i+1}} \), so that \((1 - a_i)(1 - a_{i+1}) \leq \frac{\lambda a_{i+1}(1 - a_{i+1})}{1 + \lambda a_{i+1}} \). To bound the partial derivative, we want to maximize this subject to the constraint that \( a_{i+1} \geq \frac{1}{1 + \lambda} \).

Consider the function \( x \mapsto \frac{\lambda x (1 - x)}{1 + \lambda x} \) on the interval \([\frac{1}{1 + \lambda}, 1]\). Differentiating, we see that when \( \lambda \geq \frac{1 + \sqrt{5}}{2} \), it is maximized at \( \frac{\sqrt{1 + \lambda}}{\lambda} \) and that the maximum value is \( 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \). Thus \((1 - a_i)(1 - a_{i+1}) \leq 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \). Applying this to consecutive pairs in \( \prod_{i=0}^{R-1} (1 - a_i) \), we have, for even \( R \)

\[
\left| \frac{\partial a_r}{\partial w} \right| \leq (1 + \lambda) \prod_{i=0}^{R-1} (1 - a_i) \leq (1 + \lambda) \left( 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \right)^{R/2} \tag{10}
\]

On the other hand, if \( \lambda < \frac{1 + \sqrt{5}}{2} \), then the derivative of \( \frac{\lambda x (1 - x)}{1 + \lambda x} \) is never zero in \([\frac{1}{1 + \lambda}, 1]\), and the function is maximized at \( \frac{1}{1 + \lambda} \). Thus \((1 - a_i)(1 - a_{i+1}) \leq \frac{\lambda^2}{(1 + \lambda)(1 + 2\lambda)} \), and once again, applying this to consecutive pairs, for even \( R \),

\[
\left| \frac{\partial a_r}{\partial w} \right| \leq (1 + \lambda) \left( \frac{\lambda^2}{(1 + \lambda)(1 + 2\lambda)} \right)^{R/2} \tag{11}
\]

Let us now re-examine (8). We have

\[
|a_r^0 - a_r^1| \leq \sum_{w \in \partial R^R} \max_{a_w \in [\frac{1}{1 + \lambda}, 1]} \left| \frac{\partial a_r}{\partial w} \right| \leq |\partial T_R| B_{\lambda, R} \tag{12}
\]

where \( B_{\lambda, R} \) is an upper bound on \( \left| \frac{\partial a_r}{\partial w} \right| \).

Since \( T \) is a subtree of a Poisson(\( d \)) tree, it follows from Proposition 2.1 that, almost surely, for all sufficiently large \( R \)

\[
|\partial T_R| \leq R^2 d^R. \tag{13}
\]

If \( \lambda \geq \frac{1 + \sqrt{5}}{2} \) then substituting \( B_{\lambda, R} = (1 + \lambda) \left( 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \right)^{R/2} \) into (12), we have

\[
|a_r^0 - a_r^1| \leq R^2 d^R (1 + \lambda) \left( 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \right)^{R/2} \tag{12}
\]

which goes to 0 as \( R \to \infty \) as long as \( d^2 \left( 1 - \frac{2}{\lambda} (\sqrt{1 + \lambda} - 1) \right) < 1 \), i.e., \( \lambda < \frac{4d^2}{(d^2 - 1)^2} \).

If \( \lambda < \frac{1 + \sqrt{5}}{2} \) then substituting \( B_{\lambda, R} = (1 + \lambda) \lambda^R (1 + \lambda)^{-R/2} (1 + 2\lambda)^{-R/2} \) into (12), we have

\[
|a_r^0 - a_r^1| \leq R^2 d^R (1 + \lambda) \frac{\lambda^R}{(1 + \lambda)^{R/2} (1 + 2\lambda)^{R/2}} = R^2 (1 + \lambda) \left( \frac{d^2 \lambda^2}{(1 + \lambda)(1 + 2\lambda)} \right)^{R/2} \tag{12}
\]

which goes to 0 as \( R \to \infty \) as long as \( d^2 \lambda^2 < (1 + \lambda)(1 + 2\lambda) \), i.e., \( \lambda < \frac{3 + \sqrt{1 + 4d^2}}{2d^2 - 4} \).

The transition point, \( \lambda = \frac{1 + \sqrt{5}}{2} \) corresponds to \( d = \sqrt{2 + \sqrt{5}} \) which is approximately 2.058.

Thus we have shown WSM for independent sets with fugacity \( \lambda \) on any subtree \( T \) of a Poisson(\( d \)) tree, when

\[
\lambda < \begin{cases} \frac{4d^2}{(d^2 - 1)^2} & \text{when } d < \sqrt{2 + \sqrt{5}} \\ \frac{3 + \sqrt{1 + 4d^2}}{2d^2 - 4} & \text{otherwise}. \end{cases}
\]

By Remark 2.2 we have SSM for Poisson(\( d \)) for \( \lambda \) in the same range.
5 Mixing for small $d$

In this section we prove part 1 of Theorem 1.2, which we now restate in an equivalent form.

**Theorem 5.1.** For all $d < 1.179...$, the Poisson$(d)$ tree almost surely has weak spatial mixing for all finite $\lambda > 0$.

**Proof.** Recall our formula for the influence of a leaf $w$ along the path $v = v_0, v_1, \ldots, v_R = w$:

$$\left| \frac{\partial \ln a_v}{\partial \ln a_w} \right| = \prod_{i=1}^{R-1} (1 - a_i).$$

(14)

We claim that the existence of this path tells us nothing about the other branches of the tree that do not survive to depth $R$. In particular, the number of childless children of each $v_i$ for $0 \leq i < R - 1$ is independent, and Poisson-distributed with mean $\mu = de^{-d}$.

The presence of these small leaves gives us a better upper bound on $1 - a_i$. In particular, if $v_i$ has $c_i$ childless children, then

$$1 - a_i \leq 1 - \frac{1}{1 + \lambda \left( \frac{1}{1+\lambda} \right)^{c_i}} = \frac{\lambda}{(1 + \lambda)^{c_i} + \lambda}.$$ 

Thus $w$'s expected influence is at most

$$\mathbb{E} \prod_{i=0}^{R-2} \frac{\lambda}{(1 + \lambda)^{c_i} + \lambda} = \left( \mathbb{E} \frac{\lambda}{c(1 + \lambda)^{c_i} + \lambda} \right)^{R-1} \leq \left( e^{-\mu} \frac{\lambda}{1 + \lambda} + (1 - e^{-\mu}) \frac{\lambda}{1 + 2\lambda} \right)^{R-1}.$$ 

The expected total influence of all the leaves is this times $d^R$, which is exponentially small if

$$e^{-\mu} \frac{\lambda}{1 + \lambda} + (1 - e^{-\mu}) \frac{\lambda}{1 + 2\lambda} < \frac{1}{d}.$$ 

The left-hand side is monotonically increasing with $\lambda$, so this inequality holds as long as

$$\frac{1 + e^{-\mu}}{2} < \frac{1}{d}.$$ 

Substituting $\mu = de^{-d}$, we find that this holds for all $d < 1.179$.

We have made no attempt to optimize the constant in Theorem 5.1.

**Remark 5.2.** Note that for any $d > 1$, there is a $\lambda$ for which Poisson$(d)$ tree lacks strong spatial mixing. The reason (as pointed out to us by Allan Sly) is that it possesses, with positive probability, subgraphs that are “stretched” versions of the infinite binary tree, which branch every $c$ generations for some constant $c$. See Figure 3. Such trees lack weak spatial mixing for sufficiently large $\lambda$, since if

$$f_1(a) = \frac{1}{1 + \lambda a} \quad \text{and} \quad f_2(a) = \frac{1}{1 + \lambda a^2},$$

the function

$$f_1(f_1(\cdots(f_1(f_2(a))))))$$

$c$ times

has a stable period-2 orbit for sufficiently large $\lambda$. 

6 Non-mixing just above the threshold

In this section we will prove that, for sufficiently large but constant \(d\), the Poisson(\(d\)) tree lacks spatial mixing just above the threshold for \(d\)-regular trees. First note that the latter is

\[
\frac{d^d}{(d-1)^{d+1}} = \frac{e}{d} + O(1/d)^2.
\]

Note that for \(z \in [-1, 1]\),

\[
1 - z \leq \frac{1}{1+z} \leq 1 - z + z^2
\]  \(\text{(15)}\)

Let \(v\) be a vertex at level \(L - 1\). By (15) and the definition of \(a_v\), we have

\[
1 - \lambda \prod_w a_w \leq a_v \leq 1 - \lambda \prod_w a_w + \lambda^2 \prod_w a^2_w
\]

where the product is over the children \(w\) of \(v\), which are at level \(L\). Taking expectations, we have

\[
1 - \lambda \mathbb{E} \left[ \prod_w a_w \right] \leq \mathbb{E} a_v \leq 1 - \lambda \mathbb{E} \left[ \prod_w a_w \right] + \lambda^2 \mathbb{E} \left[ \prod_w a^2_w \right]
\]  \(\text{(16)}\)

Let \(a_L\) denote the non-occupation probability of a generic vertex at level \(L\), in a Poisson tree truncated at depth \(R\). (Note that these are independent and identically distributed.) Let \(K \sim \text{Poisson}(d)\) denote the number of children of vertex \(v\), and let \(a_1, a_2, \ldots a_K\) denote the non-occupation probabilities of these children. Then the \(a_i\)'s are independent of each other and \(K\) and each has

\[
\]
expectation $\mathbb{E} a_L$. So

$$
\mathbb{E} \left( \prod_w a_w \right) = \mathbb{E} \left[ \mathbb{E} \left( \prod_{i=1}^{K} a_i \mid K \right) \right] \\
= \mathbb{E} \left( \prod_{i=1}^{K} \mathbb{E} [a_i \mid K] \right) \\
= \mathbb{E} \left( \left[ \mathbb{E} a_L \right]^K \right) \\
= \sum_{k=0}^{\infty} \frac{e^{-d} d^k}{k!} \left[ \mathbb{E} a_L \right]^k \\
= e^{-d(1-\mathbb{E} a_L)}
$$

Similarly,

$$
\mathbb{E} \left( \prod_w a_w^2 \right) = e^{-d(1-\mathbb{E} a_L^2)}
$$

Substituting into (16), we have

$$
1 - \lambda e^{-d(1-\mathbb{E} a_L)} \leq \mathbb{E} a_{L-1} \leq 1 - \lambda e^{-d(1-\mathbb{E} a_L)} + \lambda^2 e^{-d(1-\mathbb{E} a_L^2)}.
$$

If we define

$$
\phi_q(z) = 1 - \lambda e^{-d(1-z)} + q \lambda^2,
$$

we can rewrite (17)

$$
\phi_0(\mathbb{E} a_{L-1}) \leq \mathbb{E} a_L \leq \phi_q(\mathbb{E} a_{L-1}),
$$

where

$$
q = e^{-d(1-\mathbb{E} a_L^2)} \in [0, 1].
$$

The following lemma shows that for $\lambda$ just above $e/d$, even if an adversary controls the second moment $\mathbb{E} a_L^2$ and hence the coefficient $q$ of the quadratic term, this function oscillates between two disjoint intervals. It follows that the expected occupation probability at the root alternates between high and low values based on the parity of the depth of the tree, implying a lack of spatial mixing.

**Lemma 6.1.** For fixed $\lambda$, $d$, and $q \in [0, 1]$, let $\phi_q(z)$ be defined as in (18). Let $\lambda = c/d$ where $c > e$ is a constant. Then there are constants $d^*$, $b_1$, and $b_2$ such that, for all $d > d^*$ and all $q \in [0, 1]$,

$$
\forall z > 1 - b_1/d : \phi_q(z) < 1 - b_2/d \\
\forall z < 1 - b_2/d : \phi_q(z) > 1 - b_1/d,
$$

and where $b_1 < b_2$.

**Proof.** Since $\phi_0$ is monotonically decreasing, it has a unique fixed point $z_0 = \phi_0(z_0)$, namely

$$
z_0 = 1 - \frac{b_0}{d} \quad \text{where} \quad b_0 = W(\lambda d) = W(c).
$$
Here $W(x)$ is Lambert’s function, i.e., the unique positive root $y$ of $ye^y = x$. We have $f'_0(x_0) = -W(d\lambda)$. If $c > e$ then $W(c) > 1$, making this fixed point unstable.

To focus on $\phi_q$’s behavior near $z_0$ we change variables, setting $z = z_0 + \delta/d$. Then applying $\phi_q$ to $z$ is equivalent to applying $\psi_q(\delta)$ to $\delta$, where

$$
\psi_q(\delta) = d \cdot (\phi_q(z_0 + \delta/d) - z_0) = -(e^\delta - 1)W(\lambda d) + q\lambda^2 d = -(e^\delta - 1)W(c) + \frac{qc^2}{d}.
$$

Since $g'_0(0) = -W(c)$ and $\psi_0$ is analytic, for any constant $1 < A < W(c)$, there is a constant $\bar{\delta} > 0$ such that

$$
\forall \delta \in [-\bar{\delta}, \bar{\delta}]: g'_0(\delta) < -A.
$$

Therefore, for any $\delta^* < \bar{\delta}$ we have

$$
\forall \delta > \delta^*: \psi_0(\delta) < -A\delta^* \quad \text{and} \quad \forall \delta < -\delta^*: \psi_0(\delta) > A\delta^*.
$$

Choose such an $A$ and such a $\delta^*$ with $\delta^* < b_0$. Finally, since $\psi_0(\delta) \leq \psi_q(\delta) \leq \psi_0(\delta) + c^2/d$, if $d > d^*$ is sufficiently large so that

$$
\frac{c^2}{d} < (A - 1)\delta^*,
$$

the proof is completed by setting $b_1 = b_0 - \delta^*$ and $b_2 = b_0 + \delta^*$.

\[ \square \]

7 Asymptotically Optimal Lower Bound

We saw in Section 6 that asymptotically, for large $d$ the Poisson($d$) tree does not have weak spatial mixing for $\lambda$ just above $e/d$, which is the asymptotic threshold for WSM (and SSM) for the $d$-regular tree. We will now show that below $e/d$ the Poisson($d$) tree almost certainly does have weak spatial mixing. Specifically we will prove the following result, which is equivalent to part 3 of Theorem 1.2.

**Theorem 7.1.** For all $\gamma \in (0, 1)$, for all sufficiently large $d$, the Poisson($d$) tree with activity $\lambda = (1 - \gamma)e/d$ exhibits weak spatial mixing with probability 1.

The proof is fairly involved, and we begin by presenting a summary of the main ideas involved.

**Proof Sketch.** To show WSM we need to show that there is a well defined non-occupation probability $a^*_r$ at the root, i.e., that the sequences $a^0_{r,R}$ and $a^1_{r,R}$ converge to a common limit. As in Section 4 we bound $|a^0_{r,R} - a^1_{r,R}|$ by the sum of the absolute values of the partial derivatives $\partial a_r/\partial a_w$ where $w$ is a vertex at depth $R$. We know that there are almost surely at most $R^2d^R$ such vertices, for all sufficiently large $R$. The improvement in this argument comes from proving a better upper bound on $\prod_v(1 - a_v)$ which controls the size of $|\partial a_r/\partial a_w|$. Here, the product is taken over all vertices $v$ on the path from $r$ to $w$. The main idea is that when $d$ is very large, most of the vertices on the path from $r$ to $w$ are “good” in the sense that they and all their descendants to some depth $h$ have degrees very close to $d$. In other words, each such vertex $v$ is the root of a nearly regular $d$-ary subtree of depth $h$. For large enough $h$, this means that $a_v$ is very close to the fixed point $a^*$ of the function $f_d(x) = (1 + \lambda x^d)^{-1}$, which exists since $\lambda$ is less than the regular $d$-ary threshold. Thus for each good vertex $v$, $(1 - a_v) < c/d$ for some small $c < 1$ and it only remains to show that there are almost surely enough good vertices that, for all sufficiently large $R$, the bound $\prod_v(1 - a_v)$ for each path to depth $R$ beats the $R^2d^R$ such paths.

\[ \square \]
We devote the rest of this section to making the above argument rigorous.

**Remark 7.2.** Unlike the proof in Section 4, this proof does not show strong spatial mixing. Passing to a subtree can destroy the property that most vertices have nearly \( d \)-ary subtrees to some depth (or even that they have degree close to \( d \)). Given the results in Section 4, it is an open question whether SSM holds with high probability for \( \lambda \) between \( 1/d \) and \( e/d \).

The proof of Theorem 7.1 rests heavily of the fact that most of the vertices in the Poisson(\( d \)) tree are roots of subtrees (to some depth) that are almost \( d \)-ary. In order to make precise what we mean by “almost \( d \)-ary”, we will first need some definitions.

**Definition 7.3.** An \( (a, b) \)-tree is an infinite rooted tree in which every vertex at an even depth has \( a \) children and every node at an odd depth has \( b \) children. A truncated \( (a, b) \)-tree is the truncation of an \( (a, b) \)-tree to some finite depth \( R \).

**Definition 7.4.** Let \( 0 < \Delta_1 \leq \Delta_2 \). A rooted tree \( T \) is \( [\Delta_1, \Delta_2] \)-regular if the number of children of every vertex is in \( [\Delta_1, \Delta_2] \).

By an almost \( d \)-ary tree, we will mean a \( [(1-\varepsilon)d, (1+\varepsilon)d] \)-regular tree. In what follows we will show that such a tree behaves like a \( d \)-ary tree, in that if the tree is sufficiently deep, then for almost the same range of \( \lambda \) as for the \( d \)-ary tree, the non-occupation probabilities converge to well defined value at the root.

Our next result gives us a way to find a \( (\Delta_1, \Delta_2) \) tree and a \( (\Delta_2, \Delta_1) \) tree “near” any \( [\Delta_1, \Delta_2] \)-regular tree. See Figure 4 for illustrations.

**Lemma 7.5 (Pruning/Grafting).** Let \( T \) be a \( [\Delta_1, \Delta_2] \)-regular tree with root \( v \) and depth \( R \). Then

1. \( T \) can be transformed into a truncated \( (\Delta_1, \Delta_2) \)-tree \( T' \) of depth \( R \), rooted at \( v \), by pruning (removing children along with their entire subtrees) at even levels and grafting (adding children together with an appropriate subtree) at odd levels.

2. \( T \) can be transformed into a truncated \( (\Delta_2, \Delta_1) \)-tree \( T'' \) of depth \( R \), rooted at \( v \), by grafting at even levels and pruning at odd levels.

Let \( a_v, a'_v \), and \( a''_v \) denote the non-occupation probabilities at the root in \( T, T' \), and \( T'' \) respectively, when all their leaves are set to the same value \( a_0 \in [0, 1] \). Then

\[ a'_v \leq a_v \leq a''_v \]

**Proof.** By induction on depth of \( T \).

Recalling that

\[ f_d(a) = \frac{1}{1 + \lambda a^d}, \]

we wish to prove, for certain values of \( \lambda \), that iterating \( f_{\Delta_1} \circ f_{\Delta_2} \) causes \( a_v \) to converge to a unique fixed point. The following two lemmas establish the existence and uniqueness of this fixed point, and bound its location.
Figure 4: Applying Lemma 7.5. Old subtrees are pruned and new ones grafted on, on alternating levels.

Lemma 7.6. Let $\Delta_1, \Delta_2 \geq 2$, and let

$$\lambda(\Delta_1, \Delta_2) = \Delta_2^{\Delta_1} \left( \frac{\Delta_1 + 1}{\Delta_1 \Delta_2 - 1} \right)^{\Delta_1 + 1}. \quad (20)$$

For any $\lambda < \lambda(\Delta_1, \Delta_2)$, there is a unique fixed point $a^*$ such that $(f_{\Delta_1} \circ f_{\Delta_2})(a^*) = a^*$. Moreover, there is a constant $c < 1$ such that

$$\left| (f_{\Delta_1} \circ f_{\Delta_2})^t(a_0) - a^* \right| \leq c^{t-1} \ln(\lambda + 1).$$
Moreover,

\[ c \leq \frac{1}{2} \left( 1 + \frac{\lambda}{\lambda(\Delta_1, \Delta_2)} \right). \]

**Proof.** We will begin by changing variables. First define \( y = \ln a \), in which case \( y \in (-\infty, 0] \) and

\[ g_d(y) = -\ln(1 + \lambda e^{dy}). \]

Note that \( g_{\Delta_1} \circ g_{\Delta_2} \) is monotonically increasing. We will show that, for any \( \lambda < \lambda(\Delta_1, \Delta_2) \), there is a constant \( c < 1 \) such that

\[ \frac{d}{dy} g_{\Delta_1}(g_{\Delta_2}(y)) = g_{\Delta_1}'(g_{\Delta_2}(y))g_{\Delta_2}'(y) \leq c \quad \text{for all} \quad y \leq 0. \]

This implies that the fixed point \( y^* = (g_{\Delta_1} \circ g_{\Delta_2})(y^*) = \ln a^* \) is unique, and that we approach it exponentially quickly as we iterate \( g_{\Delta_1} \circ g_{\Delta_2} \). Rather than finding \( c \) as a function of \( \lambda \), it is analytically simpler to find a \( \lambda \) such that (21) holds for a given \( c \), and then showing that this \( \lambda \) coincides with \( \lambda(\Delta_1, \Delta_2) \) when \( c = 1 \).

It is convenient to do one more change of variables, from \( y \) to \( g^{-1}_{\Delta_2}(y) \) (which is well-defined since \( g_d \) is monotonic). Thus we can focus on

\[ h(y) = g_{\Delta_1}'(y)g_{\Delta_2}'(g^{-1}_{\Delta_2}(y)) = \Delta_1 \Delta_2 e^{\Delta_1 y}(1 - e^y) \frac{\lambda}{1 + \lambda e^{\Delta_1 y}} \]

We will find a \( \lambda \) such that \( h(y) \leq c \) for all \( y \leq 0 \). For any fixed \( y \), \( h(y) \) is a monotonically increasing function of \( \lambda \). Moreover, we can find the \( \lambda \) where \( h(y) = c \), namely

\[ \lambda_c(y) = \frac{ce^{-\Delta_1 y}}{\Delta_1 \Delta_2(1 - e^y) - c}, \]

where we note that if \( \Delta_1 \Delta_2(1 - e^y) < c \) then \( h(y) < c \) for all \( \lambda > 0 \). Taking derivatives, we find that \( \lambda_c(y) \) is minimized at

\[ y_{\min} = \ln \frac{\Delta_1 \Delta_2 - c}{(1 + \Delta_1) \Delta_2}, \]

where

\[ \lambda_c = \lambda_c(y_{\min}) = c(1 + 1) (\frac{\Delta_1}{\Delta_1 \Delta_2 - c})^{\Delta_1 + 1}. \]

Thus if \( \lambda \leq \lambda_c \), we have \( h(y) \leq c \) for all \( y \leq 0 \).

Now note that \( \lambda_c \) is a strictly increasing function of \( c \), and that it ranges from 0 to \( \lambda(\Delta_1, \Delta_2) \) as \( c \) goes from 0 to 1. Thus for any \( 0 \leq \lambda < \lambda(\Delta_1, \Delta_2) \) there is a \( c = c(\lambda) < 1 \) such that \( \lambda = \lambda_c \), and (21) holds. Specifically, an easy calculation shows that \( d^2 \lambda_c/dc^2 \geq 0 \) for \( 0 < c < 1 \), and that

\[ \frac{1}{\lambda(\Delta_1, \Delta_2)} \frac{d\lambda}{dc} \bigg|_{c=1} = \frac{\Delta_1(\Delta_2 + 1)}{\Delta_1 \Delta_2 - 1} \leq 2. \]

(Indeed, this derivative is \( 1 + O(1/\Delta_2) \).) Therefore,

\[ \lambda_c \geq \lambda(\Delta_1, \Delta_2)(1 - 2(1 - c)), \]
and so
\[ c \leq \frac{1}{2} \left( 1 + \frac{\lambda}{\lambda(\Delta_1, \Delta_2)} \right). \]

To complete the proof, each time we iterate \( g_{\Delta_1} \circ g_{\Delta_2} \), any interval shrinks by a factor of \( c \). Since \( g_{\Delta_1} \circ g_{\Delta_2} \) maps \( (-\infty, 0] \) into \( (-\ln(\lambda + 1), 0] \), the width of any interval after \( t \) iterations is at most \( e^{t-1} \ln(\lambda + 1) \). The same bound holds when we change variables back to \( a = e^y \), since \( de^y/dy \leq 1 \) for all \( y \leq 0 \).

Note that when \( \Delta_1 = \Delta_2 = \Delta \), the value of \( \lambda \) defined in Lemma 7.6 becomes the known value for the \( \Delta \)-regular tree,
\[ \lambda(\Delta, \Delta) = \Delta^{\Delta} \left( \frac{\Delta + 1}{\Delta^2 - 1} \right)^{\Delta + 1} = \frac{\Delta^\Delta}{(\Delta - 1)^{\Delta + 1}}. \]

We will also use the following lower bound,
\[ \lambda(\Delta_1, \Delta_2) \geq \Delta_2^{\Delta_1} \left( \frac{\Delta_1 + 1}{\Delta_1 \Delta_2} \right)^{\Delta_1 + 1} = \frac{1}{\Delta_2} \left( 1 + \frac{1}{\Delta_1} \right)^{\Delta_1 + 1} \geq \frac{e}{\Delta_2}. \]  

(23)

**Lemma 7.7.** Let \( \gamma \in (0, 1) \) and let \( \lambda = \frac{(1-\gamma)e}{d} \). Let \( \epsilon = \gamma^2/4 \). There is a constant \( d_0 = d_0(\gamma) \) such that for all \( d > d_0 \), the fixed point \( a^* \) of \( f_{(1-\epsilon)d} \circ f_{(1+\epsilon)d} \) is at least \( 1 - \frac{1}{(1+\epsilon)d} \).

**Proof.** As before, we change variables to \( y = \ln a \), and consider the fixed point of \( g_{(1-\epsilon)d} \circ g_{(1+\epsilon)d} \) where \( g_d(y) = -\ln(1 + \lambda e^y) \). First, we show the conditions of Lemma 7.6 are met. Recall the definition of \( \lambda(\Delta_1, \Delta_2) \) from (20). Since \( \epsilon = \gamma^2/4 < \gamma \) we have
\[ \lambda = \frac{(1-\gamma)e}{d} \leq \frac{(1-\epsilon)e}{d} \leq \frac{e}{(1+\epsilon)d} \leq \lambda((1-\epsilon)d, (1+\epsilon)d), \]

where the last inequality follows from (23).

Now that we know that \( g_{(1-\epsilon)d} \circ g_{(1+\epsilon)d} \) has a unique fixed point, it suffices to show that for
\[ y = \frac{1}{(1+\epsilon)d}, \]
\[ g_{(1-\epsilon)d} \circ g_{(1+\epsilon)d}(y) \geq y. \]  

(24)

In that case, the fixed point \( a^* \) is at least \( e^y \geq 1 + y = 1 - \frac{1}{(1+\epsilon)d} \). Since
\[ -x \leq -\ln(1+x) \leq -x \left( 1 - \frac{x}{2} \right), \]

whenever \( x > 0 \), we have
\[ g_{(1+\epsilon)d}(y) = -\ln(1 + \lambda e^{(1+\epsilon)y}) = -\ln(1 + \lambda/e) \leq -\lambda/e(1 - \lambda/2e) \]

and for any \( z \),
\[ g_{(1-\epsilon)d}(z) = -\ln(1 + \lambda e^{(1-\epsilon)dz}) \geq -\lambda e^{(1-\epsilon)dz}. \]

Since \( g_{(1-\epsilon)d} \) is monotonically decreasing, recalling \( \lambda = \frac{(1-\gamma)e}{d} \), we have
\[ g_{(1-\epsilon)d} \circ g_{(1+\epsilon)d} \left( \frac{-1}{(1+\epsilon)d} \right) \geq g_{(1-\epsilon)d} \left( -\lambda/e(1 - \lambda/2e) \right) \]
\[ \geq -\lambda e^{-d(1-\epsilon)(\lambda/e)(1-\lambda/2e)} \]
\[ = \frac{(\gamma - 1)}{d} e^{1-(1-\epsilon)(1-\gamma)(1-\frac{1-\epsilon}{2\epsilon})}. \]

23
Thus to prove (24) it suffices to show that
\[ \frac{1}{1 + \varepsilon} \geq (1 - \gamma)e^{1 - (1 - \varepsilon)(1 - \gamma)}(1 - \frac{1 - \gamma}{2d}) \]
or equivalently, setting \( \varepsilon = \frac{\gamma^2}{4} \),
\[- \ln(1 + \frac{\gamma^2}{4}) \geq \ln(1 - \gamma) + 1 - (1 - \frac{\gamma^2}{4})(1 - \gamma) \left( 1 - \frac{1 - \gamma}{2d} \right) \] (25)

We choose \( d_0 \) such that \( \frac{1 - \gamma^2}{2d_0} < \gamma^3/3 \). Then, recalling that \( \ln(1 - \gamma) = -\sum_i \gamma^i/i \), for all \( d \geq d_0 \), we have
\[ 1 + \ln(1 - \gamma) + \ln(1 + \frac{\gamma^2}{4}) \leq 1 - \sum_i \frac{\gamma^i}{i} + \frac{\gamma^2}{4} \]
\[ \leq 1 - \gamma - \frac{\gamma^2}{4} - \frac{\gamma^3}{3} \]
\[ \leq 1 - \gamma - \frac{\gamma^2}{4} - \frac{1 - \gamma}{2d} \]
\[ \leq (1 - \frac{\gamma^2}{4})(1 - \gamma) \left( 1 - \frac{1 - \gamma}{2d} \right) \]
which implies (25).

Let \( a_{v,R}^0 \) and \( a_{v,R}^1 \) denote the non-occupation probabilities at the root of the Poisson\((d)\) tree with activity \( \lambda = \frac{(1 - \gamma)e}{d} \) when the vertices at depth \( R \) are all occupied or all unoccupied respectively.

We are now ready to prove

**Theorem 7.8.** For all \( \gamma \in (0, 1) \), for all sufficiently large \( d \), for all \( \delta \in (0, 1) \), there exists \( R_0 \) such that
\[ \Pr \left( \left( \forall R \geq R_0 \right) |a_{v,R}^0 - a_{v,R}^1| \leq e^{-\gamma^2 R/56} \right) \geq 1 - \delta. \]

Fix \( \gamma \in (0, 1) \). Let \( \lambda = \frac{(1 - \gamma)e}{d} \), and, as before, let \( \varepsilon = \frac{\gamma^2}{4} \). Denote \( h = 1 + \left[ \frac{2 \log(1 - \gamma/2)}{\log(1 - \gamma)} \right] \).

We’ll call a vertex \( u \) in the Poisson\((d)\) tree good if its subtree to depth \( 2h \) is \( [(1 - \varepsilon)d, (1 + \varepsilon)d] \)-regular. Note that for a Poisson random variable \( X \) with mean \( d \), and \( 0 < \varepsilon \leq 1 \), the following Chernoff bound holds:
\[ \Pr(|X - d| > \varepsilon d) \leq 2e^{-\varepsilon^2 d/3} \]
(This follows, e.g., from [5, Theorem 5.4 and inequalities (4.2), (4.5)].) Applying this to the vertex degrees in the subtree of depth \( 2h \) rooted at \( u \), and taking a union bound, we find
\[ \Pr(\text{u is good}) \geq 1 - 2((1 + \varepsilon)d)^{2h+1}e^{-\varepsilon^2 d/3} \geq 1 - e^{-\varepsilon^2 d/4} \]
for all sufficiently large \( d \).

**Lemma 7.9.** If \( u \) is a good vertex then, subject to any boundary condition at least \( 2h \) levels below \( u \), we have \( 1 - a_u \leq \frac{1}{e^{\varepsilon^2 d}} \)
Proof. Since the tree of depth $2h$ rooted at $u$ has even depth, $a_u$ is minimized when all its descendents at depth $2h$ below it are set to 0. Let $a_u^0$ be this minimum value, and let $a_u'$ be the non-occupation probability at $u$ of the $((1-\varepsilon)d, (1+\varepsilon)d)$ alternating tree of height $2h$ rooted at $u$, when all its leaves are set to 0.

By pruning and grafting (Lemma 7.5), we know that $a_u^0 \geq a_u'$. By Lemma 7.7, the fixed point $a_*$ of $f_{(1-\varepsilon)d} \circ f_{(1+\varepsilon)d}$ is at least $1 - \frac{1}{(1+\varepsilon)d}$.

Let $c$ be the constant from Lemma 7.6 for $f_{(1-\varepsilon)d} \circ f_{(1+\varepsilon)d}$. Since $c > 0$, by (22)\
\[
\lambda = c((1 + \varepsilon)d)^{(1-\varepsilon)d} \left( \frac{(1-\varepsilon)d + 1}{(1-\varepsilon)d(1+\varepsilon)d - c} \right)^{(1-\varepsilon)d+1} 
\geq \frac{c}{(1 + \varepsilon)d} \left( \frac{(1-\varepsilon)d + 1}{(1-\varepsilon)d} \right)^{(1-\varepsilon)d+1} 
\geq \frac{c e}{(1 + \varepsilon)d}
\]
whence it follows that\
\[
c \leq \frac{(1 + \varepsilon)d \lambda}{e} = (1 - \gamma)(1 + \varepsilon) \leq 1 - \gamma/2,
\]
since by definition, $\varepsilon = \gamma^2/4$. By our choice of $h$, it follows that $c^{h-1}e < \gamma^2/8 = \varepsilon/2$.

Since $a_u' = f_{(1-\varepsilon)d} \circ f_{(1+\varepsilon)d}(0)$, by Lemma 7.6 it follows that\
\[
\left| a_u' - a_* \right| \leq c^{h-1} \ln(1 + \lambda) 
\leq c^{h-1} \lambda 
= c^{h-1} e (1 - \gamma) 
\leq \frac{\varepsilon/2}{(1 + \varepsilon)d}.
\]
Rearranging terms, we see that\
\[
a_u \geq a_u^0 \geq a_u' \geq a_* \geq 1 - \frac{\varepsilon}{2(1 + \varepsilon)d} = 1 - \frac{1 + \frac{1}{2} \varepsilon}{(1 + \varepsilon)d}.
\]
Finally,\
\[
1 - a_u \leq \frac{1 + \frac{1}{2} \varepsilon}{(1 + \varepsilon)d} = \frac{1}{d} \left( 1 - \frac{\varepsilon}{2(1 + \varepsilon)} \right) \leq \frac{e^{-\varepsilon/3}}{d}
\]
whence the lemma follows. 

Consider any path $P$ from the root to a leaf at depth $R$ in the truncated Poisson($d$) tree. Fix $j \in \{0, 1, \ldots, 2h-1\}$. Let $P_j = \{u \in P | \text{depth}(u) \equiv j \pmod{2h}\}$. For $u \in P_j$, the events that $u$ is bad are independent.

Let $X_{P,j}$ denote the number of bad $u$ in $P_j$. Then $E X_{P,j} \leq (R/(2h))e^{-\varepsilon^2d/4}$, and by Chernoff’s bound, for any $\alpha > 1$,\
\[
\Pr \left( X_{P,j} \geq \alpha \frac{R}{2h} e^{-\varepsilon^2d/4} \right) \leq \left( \frac{e}{\alpha} \right)^{\alpha R/(2h) e^{-\varepsilon^2d/4}}
\]
Choosing \( \alpha = \varepsilon e^{2d/4} \log(d) \), which is exponential in \( d \), we see that the right hand side becomes

\[
\left( \frac{e}{\alpha} \right) \varepsilon R / 8h \log(d)
\]

In particular, for sufficiently large \( d \), this is less than

\[
e^{-\varepsilon^3dR / 40h \log(d)}
\]

This is so tiny that, even if we take a union bound over all \( R \), all \( j \leq 2h \), and the “first” \( \frac{4R^2}{3}dR \) paths of length \( R \) from the root, the resulting probability bound still can be made smaller than \( \delta / 2 \).

Applying Markov’s inequality to the expected number of nodes at depth \( R \), we get that, with probability \( \geq 1 - \delta / 2 \), there are at most \( 4R^2 \delta dR \) of these, for \( R \geq R_0(\delta) \). Thus, our union bound actually covered all the vertices at depth \( R \).

Let \( X_P = \sum_j X_{P,j} \) denote the total number of bad nodes on the path \( P \). Assuming the above “good” event, we have for all \( R \geq R_0 \), and all paths \( P \) of length \( R \), that \( X_P \leq \alpha R \exp(-\varepsilon^2d/4) \).

By Lemma 7.9 we have

\[
\left| \frac{\partial a_v}{\partial a_w} \right| \leq (1 + \lambda) \prod_{u \in P} (1 - a_u)
\]

By Lemma 7.9, it follows that

\[
|a^0_v - a^1_v| \leq |\partial T_R| d^{-R} \exp(-\varepsilon R / 13) \leq \exp(-\varepsilon R / 14),
\]

as desired, again assuming our good event, and noting that this implied \( |\partial T_R| \leq d^R \text{poly}(R) \). This completes the proof of Theorem 7.8.

Theorem 7.8 says that for any \( \gamma \in (0, 1) \), for sufficiently large \( d \) the Poisson\((d)\) tree with activity \( (1-\gamma)e^d \) exhibits weak spatial mixing at the root, with probability 1. In other words, with probability 1, there is a well-defined value \( a_v \), where \( v \) is the root. Moreover, since each node \( w \) is the root of its own Poisson\((d)\) subtree, whose structure determines \( a_w \), and there are only countably many nodes, it follows that, with probability 1, every node \( w \) has a well-defined value \( a_w \).

Since \( a_w \) is the probability that \( w \) is unoccupied, conditioned on its parent \( p(w) \) being unoccupied, it follows that the occupation probabilities satisfy the recurrence

\[
\Pr(w \in X) = (1 - a_w)(1 - \Pr(p(w) \in X)),
\]

and hence, by induction on depth\((w)\), these probabilities are well-defined, i.e. the Poisson\((d)\) tree exhibits weak spatial mixing at all vertices, with probability 1. This completes the proof of Theorem 7.1.
8 Conclusion

Acknowledgments. This work was supported in part by the National Science Foundation under Grant No. PHYS-1066293 and the hospitality of the Aspen Center for Physics. This work was also partially supported by NSF grants CCF-1150281 and CCF-1219117.

References

[1] J. Aronson, A. M. Frieze, and B. Pittel. “Maximum matchings in sparse random graphs: Karp-Sipser revisited.” Random Struct. Algorithms 12(2): 111–177 (1998).

[2] G. R. Brightwell, O. Häggström and P. Winkler. Nonmonotonic behavior in hard-core and Widom-Rowlinson models. J. Stat. Phys. 94(3–4) 415–435 (1999).

[3] H.-O. Georgii. Gibbs Measures and Phase Transitions. de Gruyter, Berlin (1988).

[4] R. M. Karp and M. Sipser. “Maximum Matchings in Sparse Random Graphs.” In: Proc. FOCS 1981: 364–375.

[5] M. Mitzenmacher and E. Upfal. Probability and Computing: Randomized Algorithms and Probabilistic Analysis. Cambridge University Press, New York, NY, USA (2005).

[6] R. Restrepo, J. Shin, P. Tetali, E. Vigoda and L. Yang. Improved Mixing Condition on the Grid for Counting and Sampling Independent Sets. Probability Theory and Related Fields (2012) 1–25. Preliminary version in: Proc. FOCS 2011.

[7] A. Sly and N. Sun. The Computational Hardness of Counting in Two-spin Models on d-regular Graphs. Extended abstract in: Proc. FOCS 2011. Longer version at arXiv:1203.2602 [math.PR]

[8] J. Vera, E. Vigoda and L. Yang. Improved Bounds on the Phase Transition for the Hard-Core Model in 2-Dimensions. Manuscript, in submission to RANDOM 2013.

[9] D. Weitz. “Combinatorial Criteria for Uniqueness of Gibbs Measures” Random Structures and Algorithms, 27(4) 445–475 (2005).

[10] D. Weitz. “Counting independent sets up to the tree threshold.” In: Proc. STOC 2006.