Improved Bounds on the Phase Transition for the Hard-Core Model in 2-Dimensions

Juan C. Vera* Eric Vigoda† Linji Yang‡

May 7, 2014

Abstract

For the hard-core lattice gas model defined on independent sets weighted by an activity \( \lambda \), we study the critical activity \( \lambda_c(Z^2) \) for the uniqueness/non-uniqueness threshold on the 2-dimensional integer lattice \( Z^2 \). The conjectured value of the critical activity is approximately 3.796. Until recently, the best lower bound followed from algorithmic results of Weitz (2006). Weitz presented an FPTAS for approximating the partition function for graphs of constant maximum degree \( \Delta \) when \( \lambda < \lambda_c(T_\Delta) \) where \( T_\Delta \) is the infinite, regular tree of degree \( \Delta \). His result established a certain decay of correlations property called strong spatial mixing (SSM) on \( Z^2 \) by proving that SSM holds on its self-avoiding walk tree \( T_{\text{saw}}(Z^2) \), and as a consequence he obtained that \( \lambda_c(Z^2) \geq \lambda_c(T_4) = 1.675 \). Restrepo et al. (2011) improved Weitz’s approach for the particular case of \( Z^2 \) and obtained that \( \lambda_c(Z^2) > 2.388 \). In this paper, we establish an upper bound for this approach, by showing that SSM does not hold on \( T_{\text{saw}}(Z^2) \) when \( \lambda > 3.4 \). We also present a refinement of the approach of Restrepo et al. which improves the lower bound to \( \lambda_c(Z^2) > 2.48 \).

*Department of Econometrics and Operations Research, Tilburg University, 5000 LE Tilburg, The Netherlands. Email: j.c.veralizcano@uvt.nl.

†School of Computer Science, Georgia Institute of Technology, Atlanta GA 30332. Email: {vigoda,ljyang}@cc.gatech.edu. Research supported in part by NSF grant CCF-1217458.

‡Facebook.
1 Introduction

The hard-core model is a model of a gas composed of particles of non-negligible size and consequently configurations of the model are independent sets \([4, 8]\). For a (finite) graph \(G = (V, E)\) and an activity \(\lambda > 0\) (corresponding to the fugacity of the gas), configurations of the model are the set \(\Omega\) of independent sets of \(G\) where \(\sigma \in \Omega\) has weight \(w(\sigma) = \lambda^{|\sigma|}\). The Gibbs measure is defined as \(\mu(\sigma) = w(\sigma)/Z\) where \(Z = \sum_{\eta \in \Omega} w(\eta)\) is the partition function.

A fundamental question for statistical physics models, such as the hard-core model, is whether there exists a unique or there are multiple infinite-volume Gibbs measures on \(\mathbb{Z}^2\). An equivalent question is whether the influence of the boundary on the origin decays in the limit. More formally, for a box in \(\mathbb{Z}^2\) of side length \(2L + 1\) centered around the origin, let \(p_L^{\text{even}}\) (\(p_L^{\text{odd}}\)) denote the marginal probability that the origin is unoccupied conditional on the even (odd, respectively) vertices on the boundary being occupied. If

\[
\lim_{L \to \infty} \left| p_L^{\text{odd}} - p_L^{\text{even}} \right| = 0
\]

then there is a unique Gibbs measure on \(\mathbb{Z}^2\), and if this limit is \(> 0\) then there are multiple Gibbs measures. It is believed that there is a critical activity \(\lambda_c(\mathbb{Z}^2)\) such that for \(\lambda < \lambda_c(\mathbb{Z}^2)\) uniqueness holds, and for \(\lambda > \lambda_c(\mathbb{Z}^2)\) non-uniqueness holds. For the infinite, regular tree \(T_\Delta\) of degree \(\Delta\) it is easy to show that \(\lambda_c(T_\Delta) = (\Delta - 1)^{-1}/(\Delta - 2)^\Delta\) \([9]\).

There are long-standing heuristic results which suggest that \(\lambda_c(\mathbb{Z}^2) \approx 3.796\) \([8, 2, 10]\). For the upper bound on the critical activity, a classical Peierls’ type argument implies \(\lambda_c(\mathbb{Z}^2) = O(1)\) \([7]\), and Blanca et al. \([5]\) improved this upper bound to show \(\lambda_c(\mathbb{Z}^2) < 5.3646\). Our focus is on the lower bound.

Weitz \([14]\) showed that \(\lambda_c(\mathbb{Z}^2) \geq \lambda_c(T_4) = 27/16 = 1.6875\). His result followed from the algorithmic result. For all graphs with constant maximum degree \(\Delta\), \(\lambda < \lambda_c(T_\Delta)\), Weitz \([14]\) presented an FPTAS for approximating the partition function. A central step in his approach is proving a certain decay of correlations property known as strong spatial mixing (SSM) on the graph \(G\). SSM says that for every \(v \in V\), every \(T \subset V\) and \(S \subset T\), and pair of configurations \(\sigma, \tau\) on \(T\) which only differ on \(S\) (i.e., \(\sigma(T \setminus S) = \tau(T \setminus S)\)) then the difference in the influence of \(\sigma\) and \(\tau\) on the marginal probability of \(v\) decays exponentially in the distance of \(v\) from the difference set \(S\) (see Section 2 for formal definitions of these concepts). In contrast, weak spatial mixing (WSM) only requires that the influence decays exponentially in the distance to the set \(T\). For the hard-core model, since fixing a vertex to be unoccupied or occupied can be realized by removing the vertex or the vertex and its neighbors, it then follows that SSM on a graph \(G\) is equivalent to WSM for all (vertex induced) subgraphs of \(G\).

Weitz constructs a version of the tree \(T_{\text{saw}}(G, v)\) of self-avoiding walks from \(v \in V\), in such a way that SSM on \(T_{\text{saw}}(G, v)\) for all \(v\) implies SSM on \(G\). His variant of the self-avoiding walk tree fixes the leaves of the tree (corresponding to the walk completing a cycle in \(G\)) to be occupied or unoccupied based on a fixed, arbitrary ordering of the neighbors for each vertex. He then shows that SSM holds on the complete tree \(T_\Delta\), and hence SSM holds on all trees of maximum degree \(\Delta\) when \(\lambda < \lambda_c(T_\Delta)\).

Restrepo et al. \([11]\) improve upon Weitz’s approach for \(\mathbb{Z}^2\) by utilizing its structure to build a better “bounding tree” than \(T_\Delta\). They define a set of branching matrices \(M_\ell\) for \(\ell \geq 4\) corresponding to walks in \(\mathbb{Z}^2\) containing no cycles of length \(\leq \ell\) (see Section 3 for a more formal introduction to these notions). The key point is that \(T_{\text{saw}}(Z^2)\) is a subtree of the tree \(T_{M_\ell}\) defined by \(M_\ell\). They then present a decay of correlation proof by using a suitable message passing approach for proving SSM for \(T_{M_\ell}\), and hence for \(T_{\text{saw}}(Z^2)\) as well. They show that SSM holds on \(T_{M_6}\) for \(\lambda < 2.33\), and SSM holds on \(T_{M_8}\) for \(\lambda < 2.388\). Consequently, they establish that \(\lambda_c(\mathbb{Z}^2) > 2.388\).
Our first result establishes a limit to these approaches by showing that SSM does not hold on $T_{\text{saw}}(\mathbb{Z}^2)$. As mentioned earlier, in the construction of $T_{\text{saw}}(\mathbb{Z}^2)$, the assignment for leaves depends on the ordering in $\mathbb{Z}^2$ of neighbors of each vertex. Since $\mathbb{Z}^2$ is vertex-transitive, it is natural to define an ordering that is identical for every vertex (e.g., based on an ordering of the directions $N, S, E,$ and $W$), which we refer to as a homogenous ordering. We prove the following result.

**Theorem 1.** For $T_{\text{saw}}(\mathbb{Z}^2)$, SSM does not hold when $\lambda > 3.4$. Moreover, when $T_{\text{saw}}(\mathbb{Z}^2)$ is based on a homogenous ordering, then SSM does not hold when $\lambda > 3$.

The theorem follows from considering a tree $T$ that is a subtree of $T_{\text{saw}}(\mathbb{Z}^2)$ and establishing the threshold for WSM on $T$. The tree $T$ that we consider in the homogenous ordering case is quite simple. When $N$ is first in the ordering, the tree is simply the never-go-south tree (see Section 4). For any $T_{\text{saw}}(\mathbb{Z}^2)$ that is based on an inhomogeneous ordering, we are able to find another general subtree for which the WSM does not hold when $\lambda = 3.4$. Such an example gives a strong evidence that in order to prove the SSM for $\mathbb{Z}^2$ when $\lambda$ is close to the conjectured threshold, the self-avoiding walk tree approach might not be appropriate. There are subtrees of the SAW tree of $\mathbb{Z}^2$ that have lower WSM threshold and hence one has to figure out an approach to exclude such trees.

We then present an improvement of the approach of Restrepo et al. [11] for proving SSM for the trees $T_{\text{M}_\ell}$. They consider a particular statistic of the marginal distributions of the vertices, and prove the correlation decay property inductively on the height. The statistics can be viewed as a message passing algorithm, a variant of belief propagation. The messages they consider are a natural generalization of the message which is used to analyze the complete tree up to the tree threshold $\lambda_c(T_{\Delta})$ (which thereby reproves Weitz’s result [14]). They establish a so-called DMS condition as a sufficient condition for these messages to imply SSM holds on the tree under consideration. Some of the limitations of their approach are that to find the settings for the parameters in their messages and the DMS condition, they use a heuristic hill-climbing algorithm which might become trapped in local optima. In addition, verifying their DMS condition is non-trivial.

In this paper, we consider piecewise linear functions for the messages. As a consequence, we can find these functions by solving a linear program. This yields improved results and simpler proofs of the desired contraction property. Consequently, we prove SSM holds for $T_{\text{M}_6}$ when $\lambda \leq 2.45$ (previously, 2.33 by the DMS condition) and SSM holds for $T_{\text{M}_8}$ when $\lambda \leq 2.48$ (previously, 2.388). This establishes the following theorem.

**Theorem 2.** $\lambda_c(\mathbb{Z}^2) > 2.48$.

The rest of the paper is organized in the following way. We formally define WSM and SSM in Section 2 and also present there the self-avoiding walk tree construction used by Weitz [14]. In Section 3, we will introduce branching matrices and present the framework of Restrepo et al. [11] in a manner tailored to our work. In Section 4 we will discuss limitations of Weitz’s approach by showing several counter-examples. Finally, in Section 5 we discuss our linear programming approach for proving SSM, which yields an improvement on the lower bound for the uniqueness threshold of the hard-core model on $\mathbb{Z}^2$.

## 2 Preliminaries

### 2.1 Definitions of WSM and SSM

For a graph $G = (V, E)$ and $S \subset V$, we define the boundary condition $\sigma$ on $S$ to be a fixed configuration on $S$. For a boundary condition $\sigma$, let $p_v(\sigma)$ be the unoccupied probability of vertex $v$ in the Gibbs distribution $\mu$ on $G$ conditional on $\sigma$. We now formally define WSM and SSM.
Definition 1 (Weak Spatial Mixing). For the hard-core model at activity $\lambda$, for finite graph $G = (V,E)$, WSM holds if there exists $0 < \gamma < 1$ such that for every $v \in V$, every $S \subset V$, and every two configurations $\sigma_1, \sigma_2$ on $S$,

$$|p_v(\sigma_1) - p_v(\sigma_2)| \leq \gamma \text{dist}(v,S)$$

where dist$(v,S)$ is the graph distance (i.e., length of the shortest path) between $v$ and (the nearest point in) the subset $S$.

For an infinite graph $G$, we define the WSM threshold for $G$ as

$$WSM(G) = \inf \{ \lambda : \text{WSM does not hold on } G \text{ at activity } \lambda \}. $$

Definition 2 (Strong Spatial Mixing). For the hard-core model at activity $\lambda$, for finite graph $G = (V,E)$, SSM holds if there exists $0 < \gamma < 1$ such that for every $v \in V$, every $S \subset V$, every $S' \subset S$, and every two configurations $\sigma_1, \sigma_2$ on $S$ where $\sigma_1(S \backslash S') = \sigma_2(S \backslash S')$,

$$|p_v(\sigma_1) - p_v(\sigma_2)| \leq \gamma \text{dist}(v,S').$$

Finally, let $SSM(G)$ denote the SSM threshold for $G$, defined analogously to $WSM(G)$ but with respect to SSM.

To contrast the definitions of WSM and SSM, note that in WSM the influence decays exponentially in the distance to the boundary set $S$, whereas in SSM it is exponentially in the distance to the subset $S'$ of the boundary that they differ on. An important observation that we repeat from the Introduction to emphasize it, is that for the hard-core model, for a tree $T$, SSM holds if and only if for all subtrees of $T$ WSM holds.

2.2 Weitz’s SAW Tree

We now detail Weitz’s self-avoiding walk tree construction [14]. Given $G = (V,E)$, we fix an arbitrary ordering $>_w$ on the neighbors of each vertex $w$ in $G$. For each $v \in V$, the tree $T_{\text{saw}}(G,v)$ rooted at $v$ is constructed as follows.

Consider the tree $T$ of self-avoiding walks originating from $v$, including the vertices closing a cycle in the walks as leaves. We assign a boundary condition to the leaves by the following rule. Each leaf closes a cycle in $G$, so say the leaf corresponds to vertex $w$ in $G$ and the path leading to the leaf corresponds to the path $w \rightarrow v_1 \rightarrow \ldots v_\ell \rightarrow w$ in $G$. Then if $v_1 >_w v_\ell$ we fix this leaf to be unoccupied, and if $v_1 <_w v_\ell$ we fix this leaf to be occupied. Since we are in the hard core model, if the leaf is fixed to be unoccupied we simply remove that vertex from the tree. And if the leaf is fixed to be occupied, we remove that leaf and all of its neighbors from the tree, i.e., we remove completely the subtree rooted at the parent of that leaf.

If a boundary condition $\Gamma$ is assigned to a subset $S$ of $G$, then the self-avoiding walk tree can also be constructed consistently to the boundary condition, i.e., for a vertex $w \in S$ of $G$, we assign $\Gamma(w)$ to every occurrence of $w$ in $T_{\text{saw}}(G,v)$. Weitz proves that, for any boundary condition on $G$ and any vertex $v$, the marginal distribution of $v$ on $G$ is the same as the marginal distribution of the root of $T_{\text{saw}}(G,v)$ with the corresponding boundary condition. This further implies the following.

Lemma 1 (Weitz [14]). For a specific $\lambda$, if for all $v$, SSM holds for $T_{\text{saw}}(G,v)$, then SSM holds for $G$. **3**
3 Message Passing Approach for Proving SSM

Let us first recall the recurrence of the marginal distributions on trees for the hard-core model. For now, we fix our infinite tree to be $T$. Let $v$ be a vertex of $T$, and let $T_v$ denote the subtree of $T$ rooted at $v$. Let $N^{-}(v)$ denote the children of $v$ in $T_v$. Let $\alpha_v(\Gamma)$ be the unoccupied probability of vertex $v$ in the subtree $T_v$ rooted at $v$ with boundary condition $\Gamma$. It is straightforward to establish that $\alpha_v(\Gamma)$ satisfy the following recurrence:

$$\alpha_v(\Gamma) = \frac{1}{1 + \lambda \prod_{w \in N^{-}(v)} \alpha_w(\Gamma)}.$$

(1)

There are two special boundary conditions: one is called the odd boundary condition (denoted as $\Gamma_{o,L}$) which occupies all the vertices at level $L$ when $L$ is odd (and unoccupies when $L$ is even); the other is called the even boundary condition (denoted as $\Gamma_{e,L}$) which occupies all the vertices at level $L$ when $L$ is even (and unoccupies when $L$ is odd). These two boundary conditions are the extremal ones, meaning that for any other boundary condition $\Gamma$ for the vertices at distance $L$ from the root $r$ of $T$, $\alpha_r(\Gamma_{o,L}) \leq \alpha_r(\Gamma) \leq \alpha_r(\Gamma_{e,L})$ when $L$ is even (and with the inequalities reversed when $L$ is odd).

To see that WSM holds for the tree $T$, it is enough to show that for the odd and even boundary conditions $\{\Gamma_{o,L}\}_{L \in \mathbb{N}}$ and $\{\Gamma_{e,L}\}_{L \in \mathbb{N}}$, the difference of the marginal probabilities at the root $|\alpha_r(\Gamma_{o,L}) - \alpha_r(\Gamma_{e,L})|$ decay exponentially in $L$.

3.1 Branching matrices

Recall that in order to show that uniqueness holds for $\mathbb{Z}^2$ for a certain $\lambda$, it is enough to show that for the same $\lambda$, SSM holds on a certain tree which is a super-tree of $T_{saw}(\mathbb{Z}^2)$. Due to the regularity of $\mathbb{Z}^2$, in [11], deterministic multi-type Galton-Watson trees are proposed to characterize the candidate super-trees. The trees can be defined by matrices in the following way.

**Definition 3.** Given a $t \times t$ (branching) matrix $M$, $\mathcal{F}_{\leq M}$ is the family of trees which can be generated under the following restrictions:

- Each vertex in tree $T \in \mathcal{F}_{\leq M}$ has its type $i \in \{1, \ldots, t\}$.
- Each vertex of type $i$ has at most $M_{ij}$ children of type $j$.

We use $T_M$ to refer to the tree that is generated by the matrix $M$, specifically, we mean the largest tree in the family $\mathcal{F}_{\leq M}$. The simplest $M$ such that $T_{saw}(\mathbb{Z}^2)$ is in the family $\mathcal{F}_M$ is

$$M = \begin{pmatrix} 0 & 4 \\ 0 & 3 \end{pmatrix}.$$  

In this case, $T_M$ is the complete, regular tree of degree 4. As shown in [11], because of the regularity of $\mathbb{Z}^2$, a more sophisticated set of branching matrices $M'$ we contain $T_{saw}(\mathbb{Z}^2)$ in their family are trees $T_{M'}$ corresponding to all walks of $\mathbb{Z}^2$ truncated when closing a cycle of length less than or equal to a certain constant. Clearly, $T_{M'}$ is a super-tree of $T_{saw}(\mathbb{Z}^2)$, because any path in $T_{M'}$ will only avoid cycles of a certain length whereas paths in $T_{saw}(\mathbb{Z}^2)$ are avoiding all cycles.

When one tries to avoid a cycle of length 4, the matrix becomes

$$M'_4 = \begin{pmatrix} 0 & 4 & 0 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix},$$

4
where each type is simply representing the various stages of completing a cycle of length 4 in a walk. It is easy to verify that $T_{\text{saw}}(\mathbb{Z}^2)$ is in the family $\mathcal{F}_{M_{1}}$.

In $M_{i}'$, we have not yet taken into consideration the effect of the assignments to leaves as detailed in the construction of $T_{\text{saw}}$ in Section 2.2. When we do that, we are able to construct much more sophisticated branching matrices which yield better bounds. Therefore, for $\ell \geq 4$, let $M_{\ell}$ denote the branching matrix generating the tree containing all walks in $\mathbb{Z}^2$ truncated when completing a cycle of length $\leq \ell$, where these leaf vertices are occupied or unoccupied according to the definition in Section 2.2 based on some fixed homogeneous ordering $<_{w}$ of neighbors for every vertex. By taking into account the boundary condition we obtain a smaller tree since when a walk closes a cycle with an occupied assignment to a vertex $u$, this forces the parent of $u$ to be unoccupied, which further trims down the size of the tree. These more sophisticated matrices yield a “tighter” bound on $T_{\text{saw}}(\mathbb{Z}^2)$, however the number of types increase. For example, for $\ell = 4$, whereas $M_{4}'$ has 4 types, $M_{4}$ has 17 types (after some simplifications), see [11] for details of $M_{4}$. For $M_{6}$ there are 132 types, and for $M_{8}$ there are 922 types.

### 3.2 Contraction Principle

For each $t$ by $t$ branching matrix $M$, we would like to derive a condition such that SSM holds for the tree $T_{M}$. Throughout this paper, for each type $i$, we treat the row $M_{i}$ of $M$ as a multi-set and each entry $M_{i}(j)$ of the row denotes the number of elements the set $M_{i}$ has of type $j$. We use $t(w)$ to denote the type of vertex $w \in M_{i}$. The following lemma, which is re-stating Lemma 1 from [11] in a slightly simpler form that is more convenient for our work, provides a sufficient condition for SSM to hold for the tree $T_{M}$. The proof is in Section 7.

**Lemma 2.** Let a branching matrix $M$ be given. Assume there is $0 < \gamma < 1$ such that for each type $i$, there is a positive integrable function $\Psi_{i}$ where

$$
\frac{1 - \alpha_{i}}{\Psi_{i}(\alpha_{i})} \sum_{w \in M_{i}} \Psi_{t(w)}(\alpha_{w}) < \gamma,
$$

for $\alpha_{w}$ in the range $[1/(1 + \lambda), 1]$ for each child $w$ and $\alpha_{i} = (1 + \lambda \prod_{w \in M_{i}} \alpha_{w})^{-1}$ defined in (1) as a function of $\alpha_{w}$’s. Then SSM holds for $T_{M}$, i.e., WSM holds for all trees $T$ in the family $\mathcal{F}_{\leq M}$ with a fixed rate $\gamma < 1$.

### 4 Upper Bound on the SSM Threshold

As described in the introduction, previous approaches for lower bounding $\lambda_{c}(\mathbb{Z}^2)$ are based on proving SSM for $T_{\text{saw}}(\mathbb{Z}^2)$. To provide a bound on the strength of these approaches we upper bound the SSM threshold for $T_{\text{saw}}(\mathbb{Z}^2)$. We will show that for $\lambda \geq 3.4$, SSM does not hold for $T_{\text{saw}}(\mathbb{Z}^2)$ obtained from any edge-ordering used in the vertices. We also show the for any homogeneous ordering SSM does not hold for $T_{\text{saw}}(\mathbb{Z}^2)$ for $\lambda \geq 3$. Note that this does not imply anything about WSM/SSM on $\mathbb{Z}^2$, it simply shows a limitation on the power of the current proof approaches.

To prove that SSM does not hold on $T_{\text{saw}}(\mathbb{Z}^2)$ we define a tree $T$ that is a subtree of $T_{\text{saw}}(\mathbb{Z}^2)$ and prove that WSM does not hold on $T$ for sufficiently large $\lambda$.

#### 4.1 Upper Bound for Homogenous Ordering

We define a branching matrix $D_{H}$ such that $T_{D_{H}}$ corresponds to the never-go-South tree, and prove that WSM does not hold on this tree when $\lambda > 3$. 

5
Since we are assuming a homogeneous ordering, without loss of generality assume that $N$ is smallest in the ordering. We construct $D_H$ by considering those walks on $\mathbb{Z}^2$ that only go $N$, $E$, and $W$. The branching rules can be written in the following finite state machine way:

1. $O \rightarrow N \mid E \mid W$,
2. $N \rightarrow N \mid E \mid W$,
3. $E \rightarrow N \mid E$,
4. $W \rightarrow N \mid W$,

where $O$ corresponds to the origin and is a transient state so can be ignored when analyzing the recurrence. The branching matrix corresponding to the above rule is

$$D_H = \begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}, \quad (3)$$

Lemma 3. Let the order of the edges in each vertex be an homogenous order where $N$ is the smallest in the order. The tree $T_{D_H}$ generated by the branching matrix $D_H$ is a subtree of $\widehat{T_{saw}}(\mathbb{Z}^2)$.

Proof. In Weitz’s construction (see Section 2.2), let $\widehat{T_{saw}}(\mathbb{Z}^2)$ be the tree of self-avoiding walks of $\mathbb{Z}^2$ originating from the origin, including the vertices closing a cycle in the walks as leaves (i.e., we have not yet fixed occupied or unoccupied to vertices of $\widehat{T_{saw}}(\mathbb{Z}^2)$ based on the ordering). The tree $T_{D_H}$ consists of all those self-avoiding walks that never go South, and thus, it is a subtree of $\widehat{T_{saw}}(\mathbb{Z}^2)$.

Now, in the second part of Weitz’s construction, some vertices are deleted from $\widehat{T_{saw}}(\mathbb{Z}^2)$ to obtain $T_{saw}(\mathbb{Z}^2)$. We need to show that no vertex from $T_{D_H}$ is deleted. A vertex is deleted from $\widehat{T_{saw}}(\mathbb{Z}^2)$ because it is an occupied leaf, or is the parent of an occupied leaf. Leaves in $\widehat{T_{saw}}(\mathbb{Z}^2)$ correspond to walks finishing in a cycle, and thus they do not belong to $T_{D_H}$. Now suppose a vertex in $T_{D_H}$ is the parent of a leaf $\zeta$ in $\widehat{T_{saw}}(\mathbb{Z}^2)$. In this case, $\zeta$ corresponds to a path finishing in a cycle in $\mathbb{Z}^2$, say $\cdots \rightarrow w \rightarrow v_1 \rightarrow \ldots v_\ell \rightarrow w$, and the move $v_\ell \rightarrow w$ is a South move. To ensure that we do not fix $\zeta$ to be occupied in $\widehat{T_{saw}}(\mathbb{Z}^2)$, we need to ensure that $v_1 >_w v_\ell$, and for this we just need that $N$ is the smallest in the $>_w$ ordering, which is our assumption.

For the tree $T_{D_H}$ we can establish its WSM threshold as stated in the following result, which immediately implies Theorem 1.

Lemma 4.

$$WSM(T_{D_H}) = 3.$$ 

Proof. The matrix $D_H$ is reducible, as defined in Definition 4, to the following $2 \times 2$ matrix:

$$B = \begin{pmatrix}
1 & 2 \\
1 & 1
\end{pmatrix}. \quad (4)$$

It is easy to check that $B$ and $D_H$ generate the same family of trees. Now the recurrences for the marginal distributions of both types derived from (1) are

$$f(x, y) = \left(\frac{1}{1 + \lambda xy^2}, \frac{1}{1 + \lambda xy}\right). \quad (5)$$

Using some algebra, we are able to determine the fixed points of $f(x, y)$ for $\lambda > 1$

$$(x_0, y_0) = (x_0(\lambda), y_0(\lambda)) = \left(\frac{4\lambda + \sqrt{8\lambda + 1} - 1}{8\lambda}, \frac{\sqrt{8\lambda + 1} - 3}{2(\lambda - 1)}\right).$$
We just need to check the eigenvalues of the Jacobian of the recurrences at the fixed point, see e.g., [12]: If the largest eigenvalue is greater than 1, then the function around the the fixed point is repelling and hence it is impossible for the boundary conditions to converge to this unique fixed point. If the largest eigenvalue is strictly less than 1, the function is contracting to the fixed point in its neighborhood. The Jacobian at the fixed point \((x_0, y_0)\) is the following:

\[
J(\lambda) = \begin{pmatrix} \lambda x_0^2 y_0^2 & 2\lambda y_0 x_0^3 \\ \lambda y_0^2 & \lambda x_0 y_0^2 \end{pmatrix}.
\]  \hspace{1cm} (6)

Denote the trace of \(J(\lambda)\) as \(\text{tr}(J(\lambda)) = \lambda x_0^2 y_0^2 (x_0 + 1)\) and its determinant as \(\det(J(\lambda)) = -\lambda^2 x_0^3 y_0^4\). The largest eigenvalue of \(J(\lambda)\) is then

\[
\rho(\lambda) = \frac{\text{tr}(J(\lambda))}{2} + \left(\frac{\text{tr}(J(\lambda))^2}{4} - \det(J(\lambda))\right)^{1/2},
\]

It is easy to check that \(\rho(\lambda)\) is increasing and that for \(\lambda = 3\), \(x_0(3) = 2/3\), \(y_0(3) = 1/2\) and \(\rho(3) = 1\). \(\square\)

### 4.2 Ordering-Independent Subtree for \(T_{saw}\)

In this section, we will define a branching matrix \(D_G\) such that the generated tree \(T_{D_G}\) is a subtree of \(T_{saw}\) independently on the ordering of edges for each vertex. When in particular the ordering is homogeneous, then \(T_{D_G}\) it is a subtree of the tree \(T_{D_H}\) defined in the previous section. This new tree \(T_{D_G}\) also never goes South (as in \(T_{D_H}\)) but it has further structure to ensure that the leaves in this tree are at least distance two from the leaves of the self-avoiding walk tree \(T_{saw}(\mathbb{Z}^2)\), and therefore for any boundary condition for \(T_{saw}(\mathbb{Z}^2)\) (and hence any ordering of the edges for each vertex) it immediately follows that \(T_{D_G}\) is a subtree of \(T_{saw}(\mathbb{Z}^2)\). To achieve this property, we add to the never-go-South tree the rule that if the walk goes East there must be at least two North steps before it goes West (and similarly, for West to East).

The tree is constructed by the following rules:

0. \(O \rightarrow N \mid E \mid W\),
1. \(N \rightarrow NN \mid NE \mid NW\),
2. \(W \rightarrow WN \mid WW\),
3. \(E \rightarrow EN \mid EE\),
4. \(NN \rightarrow NN \mid NE \mid NW\),
5. \(NW \rightarrow WN \mid WW\),
6. \(NE \rightarrow EN \mid EE\),
7. \(WW \rightarrow WN \mid WW\),
8. \(EE \rightarrow EN \mid EE\),
9. \(WN \rightarrow NW \mid NN\),
10. \(EN \rightarrow NE \mid NN\).

Here the state \(O\) corresponds to the origin, while \(E\), \(W\) and \(N\) correspond to the first edges in the path. Then each of the states corresponds to the last two visited edges. Notice also that states \(O\), \(E\), \(W\) and \(N\) are transient states. We denote the branching matrix for this tree as \(D_G\).

**Lemma 5.** The tree \(T_{D_G}\) generated by the branching matrix \(D_G\) is a subtree of \(T_{saw}(\mathbb{Z}^2)\), independently of the edge ordering used for each vertex.

**Proof.** In any path \(\sigma\) in \(T_{D_G}\) the distance in the grid for any two non-consecutive vertices in \(\sigma\) is at least 2. This will imply that every vertex in \(\sigma\) is distance at least 2 from the leaves of the self-avoiding walk tree originating from the origin of the grid. And thus independently of the boundary condition used in the construction of \(T_{saw}(\mathbb{Z})\), we have that \(T_{D_G}\) is a subgraph of \(T_{saw}(\mathbb{Z}^2)\). \(\square\)

We establish the following bounds on the WSM threshold for the tree \(T_{D_G}\).

**Lemma 6.**

\[3.3 < WSM(T_{D_G}) < 3.4.\]

This lemma is proved in Section 8. Theorem 1 follows from Lemmas 4 and 6.
4.3 Tree with Different Thresholds for SSM and WSM

Brightwell et al. [6] give an example of a tree for which WSM holds but SSM does not hold for the same activity \( \lambda \). Here, we present another example which is more closely related to \( T_{\text{saw}}(\mathbb{Z}^2) \). We show a tree \( T' \), which is a super-tree of \( T_{D_H} \) and subtree of \( T_{\text{saw}}(\mathbb{Z}^2) \), for which WSM holds for some \( \lambda > 3 \).

To construct the tree \( T' \) we allow some South moves in the tree in a certain context. In particular, we only allow that a South move happens when the path contains the following substring: NNEESEN, i.e., a South move is allowed if and only if it is after a sequence of NNEE moves and followed by EN moves. The tree is formally defined by the branching matrix denoted as \( D' \).

The tree family can be formalized in the following finite state machine way:

1. \( E \to E \mid N \)
2. \( W \to N \mid W \)
3. \( N \to NN \mid E \mid W \)
4. \( NN \to NN \mid NNE \mid W \)
5. \( NNE \to N \mid NNEE \)
6. \( NNEE \to N \mid E \mid NNEES \)
7. \( NNEES \to NNEESE \)
8. \( NNEESE \to NNEESEN \)
9. \( NNEESEN \to NN \mid E \)

Let the matrix describing the above rules be denoted as \( D' \).

**Lemma 7.** The tree \( T_{D'} \) that is generated by \( D' \) is a super-tree of \( T_{D_H} \) and is a subtree of \( T_{\text{saw}}(\mathbb{Z}^2) \).

The proof of the above lemma is similar to the proof of Lemma 3.

We will prove that the WSM threshold for \( T' = T_{D'} \), the tree generated by \( D' \) is above \( \lambda = 3.01 \), and hence, combined with Lemma 4, we get the following lemma.

**Lemma 8.** For the tree \( T_{D'} \) at \( \lambda = 3.01 \), WSM holds but SSM does not hold.

**Proof.** We only need to show that WSM for \( T_{D'} \) holds when \( \lambda = 3.01 \) and then, since it is a super tree of \( T_{D_H} \) and we prove that \( WSM(T_{D_H}) = 3 \), therefore SSM does not hold for \( T_{D'} \) when \( \lambda = 3.01 \).

To show that WSM holds, it follows the same procedure as the proof of the lower bound in Lemma 6. We can write a recurrence \( r : \mathbb{R}^9 \to \mathbb{R}^9 \) with 9 equations. Let \( \vec{x} \in \mathbb{R}^9 \). For instance, \( r_1(\vec{x}) = \frac{1}{1+\lambda x_1x_3} \), \( r_2(\vec{x}) = \frac{1}{1+\lambda x_2x_3} \) and so on. Starting with \( \vec{x} = (0,0,...,0) \) or \( \vec{x} = (1,1,...,1) \), again by iterating the recurrence, we are able to approximate the fixed point with a very high precision. Then we calculate the Jacobian evaluated at \( r^t(0,0,...,0) \) for a large \( t \). Using Maple we are able to check that the Jacobian is less than 0.996 and we can apply the Bauer-Fike theorem again to bound the spectral radius for establishing the convergence.
5 Linear Program for Lower Bounding SSM Threshold

Here we propose a way to use linear programming to solve the functional inequality (2). Notice that if $\Psi_i$ is positive and bounded for all $i$ then inequality (2) is equivalent to

$$(1 - \alpha_i) \sum_{w \in M_i} \Psi_t(w)(\alpha_w) < \Psi_i(\alpha_i).$$  \hfill (7)

The idea to solve (7) is simple. We will restrict the search for $\Psi_i$ to a family of positive piecewise linear functions with a finite number of discontinuities.

First of all, it is a simple fact that each $\alpha_i$ is in the interval $I = [1/(1 + \lambda), 1]$. We will divide $I$ into a set of $d$ consecutive sub-intervals of the same size. Define

$$X_k = \frac{1}{1 + \lambda} + k\frac{\lambda}{d(1 + \lambda)}, \text{ for } k = 0, \ldots, d - 1.$$

To ease the notation define $Y_k = X_{k+1}$ for $k = 0, \ldots, d - 1$. Note that the intervals $[X_k, Y_k]$ partition $I$. Since the only requirements of $\Psi_i(x)$ are positive and integrable, we restrict the search for $\Psi_i(x)$ to functions of linear form $-a_{i,k}x + b_{i,k}$ in each interval $[X_k, Y_k]$ with $a_{i,k}, b_{i,k} > 0$.

Now, for each type $i$, the functional inequality can be decomposed according to different combinations of the intervals of the variables $\alpha_w$ which are type $i$’s children. For each combination, we are able to write down a set of linear inequalities such that it is a sufficient condition for the functional inequality to hold within that region.

To capture for which sub-intervals should (7) hold, we say that a tuple of indexes $k = (k_0, k_1, k_2, \ldots, k_{\Delta_i})$ is $i$-acceptable if the interval $[X_{k_0}, Y_{k_0}]$ intersects the interval $\left[\frac{1}{1 + \lambda \prod_{j=1}^{\Delta_i} Y_{k_j}}, \frac{1}{1 + \lambda \prod_{j=1}^{\Delta_i} X_{k_j}}\right]$. We have the following theorem.

**Theorem 3.** In order for the functional inequality (2) to hold, it is enough for the following set of linear constraints (a’s and b’s are the variables) to be feasible:

For each $i \in [t]$ and each $i$-acceptable tuple $k$,

$$(1 - X_{k_0}) \sum_{j=1}^{\Delta_i} \left(b_{t(j),k_j} - a_{t(j),k_j} X_{k_j}\right) < (b_{i,k_0} - a_{i,k_0} Y_{k_0}),$$  \hfill (8)

where $\{t(j) : j = 1, \ldots, \Delta_i\} = M_i$ (as multisets).

For each $i \in [t]$ and $k = 0, \ldots, d - 1$,

$$b_{i,k} - a_{i,k} Y_k > 0, \quad 0 \leq a_{i,k} \leq M, \quad 0 \leq b_{i,k} \leq M.$$  \hfill (9)

where $M$ is some (big) constant.

**Proof.** Define $\Psi_i(x) = b_{i,k} - a_{i,k} x$ for all $x \in [X_k, Y_k]$. Linear constraints (9) imply that $\Psi_i$ is non-negative and bounded. Thus it is enough to show (7) holds.

Now fix type $i$ we have $k_w$’s such that $\alpha_w \in [X_{k_w}, Y_{k_w}]$ for each $w \in M_i$. Let $\alpha_i = 1/(1 + \lambda \prod_{w \in M_i} \alpha_w)$, then

$$\frac{1}{1 + \lambda \prod_{w \in M_i} Y_{k_w}} \leq \alpha_i \leq \frac{1}{1 + \lambda \prod_{w \in M_i} X_{k_w}}.$$  

Thus if $k_i$ is such that $X_{k_i} \leq \alpha_i \leq Y_{k_i}$ then the tuple $k = (k_{i,w_1}, \ldots, k_{i,w_{\Delta_i}})$ is $i$-acceptable.
Therefore,
\[
(1 - \alpha_i) \sum_{w \in M_i} \Psi_t(w)(\alpha_w) = (1 - \alpha_i) \sum_{w \in M_i} \left( b_{t(w),k_w} - a_{t(w),k_w}(\alpha_w) \right)
\]
\[
\leq (1 - X_k) \sum_{j=1}^{\Delta_i} \left( b_{t(w),k_w} - a_{t(w),k_w}X_{k_w} \right)
\]
\[
< b_{t,k_i} - a_{i,k_i}Y_{k_i} \quad \text{from (8)}
\]
\[
\leq \Psi_t(\alpha_i).
\]

Consider the branching matrix \( M_\ell \) generating the family of trees avoiding cycles of length \( \leq \ell \). Recall that the tree \( T_{M_\ell} \) which is generated by \( M_\ell \) is a super-tree of \( T_{\text{saw}}(Z^2) \). We show that the system (8)-(9) corresponding to \( M_\ell \) is feasible, proving SSM for \( T_{M_\ell} \) and hence for \( T_{\text{saw}}(Z^2) \).

To solve the feasibility problem, we add a new variable \( v \) in the right hand side of each linear constraint \( ax \leq b \), changing this constraint to \( ax - b \leq v \). We minimize \( v \), which is an upper bound for the maximum violation by \( x \) among all constraints. The original linear system is feasible if and only the linear program has optimal solution \( v < 0 \).

The number of constraints and variables in this LP are huge (almost 10 billion constraints and 1 million variables) when \( d = 200 \) for the matrix \( M_8 \). In order to solve the linear program efficiently, one has significantly to reduce its size. In Section 6, we will discuss about the methods we use to solve this LP. When running the linear programs built for \( M_4 \) we obtain \( \lambda > 2.31 \), and for \( M_6 \) we obtain \( \lambda > 2.45 \), and for \( M_8 \) we obtain \( \lambda > 2.48 \). In this way, we are able to prove that SSM holds for \( \ell \leq 2.48 \). The data for these LP solutions are available in our online appendix [15].

What we obtain from our linear program method are closer to the limit of this approach. Computational experiments suggest the threshold for WSM for \( T_{M_4} \) is at roughly \( \lambda \approx 2.482 \), for \( T_{M_6} \) at \( \lambda \approx 2.653 \), for \( T_{M_8} \) at \( \lambda \approx 2.75 \) and finally for \( T_{M_{10}} \) at \( \lambda \approx 2.82 \). These are thresholds for WSM, and the SSM threshold may in fact be even lower, as occurred for our example in Section 4.3.

### 5.1 Comparison with Previous Approaches

This method has several advantages compared to the method that is proposed in [11] in which a sufficient condition called the DMS condition, is introduced. DMS is a nonlinear matrix inequality obtained by comparing the geometric mean with the arithmetic mean when one analyzes the functional inequality (2) for a specific type of \( \Psi_i \) functions. These functions are the optimal ones when the tree \( T_M \) is a complete regular tree. However, for multi-type branching matrices, they are not necessarily optimal. One has to find the parameters of these functions \( \Psi_i \) in order to satisfy the DMS condition. The parameters for the DMS condition are obtained by a randomized hill-climbing program which may become trapped in a local optima. In contrast, the linear programming method we present here provides the optimal solution for the class of functions being considered.

For the SSM threshold of \( T_{M_\ell} \), our method includes the approximation of a more general class of functions and hence we obtain better lower bounds (see Figure 1). Finally, the mathematical correctness of the linear programming method is very straightforward to check as compared to checking the correctness of the DMS condition. For \( \ell = 4, 6, 8 \), we summarize in the following table, the experimental lower bound for the WSM threshold of \( M_\ell \), the size of the matrix \( M_\ell \), the lower bounds of the SSM threshold for \( M_\ell \) obtained from DMS condition in [11], and the lower bounds of the SSM threshold for \( M_\ell \) obtained from our linear program approach.
6 Reducing the Size of the LP

Initially, when we write down the linear programs (LPs) for the $M_8$ matrix with the size of intervals around $10^{-3}$, the number of constraints and variables is huge, approximately 10 billion constraints and 1 million variables. Solving this LP directly is not possible, as the data will not even fit in memory. Notice that the LPs we create have high constraint-variable ratios. One standard technique to solve such LPs is to write the dual which has a high variable-constraint ratio and apply the column generation method [3]. From the primal point of view, we try to guess the set of tight constraints, by picking a set of primal constraints, solving a smaller LP and checking whether the rest of the constraints are satisfied. When there are violated constraints, several of the most violated constraints are added to the set and we iterate the procedure until the LP is solved.

Using column generation we obtain an LP that can be solved, but running the method takes too long. Next we will present two of our major techniques to reduce the size of the LPs so that we can solve them within a few days.

6.1 Nonhomogeneous interval size

In Theorem 3, we break the intervals into subintervals of the same size. The algorithm was designed to start with a very coarse set of the subintervals with a uniform length and if the LP has no solution, then the algorithm will try to decrease the length and re-solve the new LP. Usually, the algorithm has to make the length as small as $10^{-3}$ for the LP to have a solution. This creates lots of constraints. Notice that, the constraints are tight only in a very small range of the interval $(1 + \lambda, 1)$. Therefore, we can try to break the intervals into subintervals of different sizes.

The goal of breaking intervals is to change the primal constraints so that the objective function $v$ can be achieved at a smaller value. In column generation, shadow prices are used for this purpose. However, here deciding which intervals to break, affects the objective in a nonlinear fashion. Thus, we use a heuristic pricing scheme on the intervals to pick which ones to break. The following briefly describes our heuristic approach.

For each interval, we know how many constraints are involved for that interval and how many of them are violated (i.e., $ax - b > 0$). We sum up the values of $ax - b$ for how much each constraint is violated and then scale this by a factor of the size of the interval to obtain what we define as its price. The algorithm will pick several intervals with the highest prices to break. The reason why we scale by a factor which is a function of the size of the interval is that we do not want to break the intervals that are already very small. In Figure 1 we show a step function $\Phi_i$ for a type $i$ in $D_8$ found by the LPs. One can observe from the figure that most of the intervals have large lengths; in fact, there are some small intervals in the middle as these are the intervals that create tight constraints.

6.2 Reduction of the branching matrices $M_\ell$

Usually, when one applies various methods trying to solve the functional inequality (2), one has to face the fact that the dimension of the matrix $M$ is huge, e.g., $t = 922$ for $\ell = 8$ in [11]. A natural way to generate $M$ is using a DFS program that enumerates all of the types by remembering

| $\ell$ | WSM threshold | Number of Types | $\lambda$ from DMS in [11] | $\lambda$ from LP |
|-------|---------------|-----------------|--------------------------|-----------------|
| 4     | 2.48          | 17              | 2.16                     | 2.31            |
| 6     | 2.65          | 132             | 2.33                     | 2.45            |
| 8     | 2.75          | 922             | 2.38                     | 2.48            |
the history of the self-avoiding walk. However, there are many types in such a matrix that are essentially the “same”. Here we provide a heuristic and rigorous method for finding those types that are the same.

Let $\mathcal{C}$ be a partition of the types in $\mathbf{M}$, i.e., $\mathcal{C} = \{C_1, C_2, ..., C_k\}$ such that $\bigcup_{i=1}^{k} C_i = [t]$. We define the partition to be consistent with $\mathbf{M}$, if for every $i \in [k]$, each pair of types $s, t \in C_i$, the rows $\mathbf{M}_s$ and $\mathbf{M}_t$ are the same with respect to $\mathcal{C}$, that is

$$\sum_{j \in C_{i'}} M_{sj} = \sum_{j \in C_{i'}} M_{tj}, \text{ for all } i' \in [k].$$

**Definition 4.** Given $\mathbf{M}$ and a partition $\mathcal{C}$ of size $k$ which is consistent, we define the $k$-by-$k$ matrix $\mathbf{M}^\mathcal{C}$ by,

$$M^\mathcal{C}_{ij} = \sum_{j \in C_{i'}} M_{sj} \text{ where } s \in C_i.$$

We say $\mathbf{M}$ is reducible to a $k$-by-$k$ matrix $\mathbf{B}$ if there is a consistent partition $\mathcal{C}$ such that $\mathbf{B} = \mathbf{M}^\mathcal{C}$.

**Lemma 9.** For a partition $\mathcal{C}$ of size $k$,

$$F_{\leq \mathbf{M}^\mathcal{C}} = F_{\leq \mathbf{M}} \text{ and } T_{\mathbf{M}^\mathcal{C}} = T_{\mathbf{M}}.$$

**Proof.** The argument is just a standard induction on the height of the tree.

Now the question is how to find a good partition $\mathcal{C}$ easily. For a specific value $\lambda < \text{WSM}(T_{\mathbf{M}})$, let $V_{\lambda}$ be the fixed points of the recurrences of the marginal distributions defined by $\mathbf{M}$. Our conjecture is the following.

**Conjecture 1.** Let the partition $\mathcal{C}(\lambda)$ be the sets of types that have the same value of the fixed points in $V_{\lambda}$, i.e., for each $C_i \in \mathcal{C}(\lambda)$, for all $c \in C_i$, $V_{\lambda}(c)$ are the same. If for all $\lambda$, the partitions $\mathcal{C}(\lambda)$ are identical, then $\mathcal{C}$ is a partition that is consistent of $\mathbf{M}$.

Using the intuition from Conjecture 1 we are able to find good partitions in practice. We simply run a dynamic programming algorithm on the tree $T_{\mathbf{M}}$ to calculate an approximation of the fixed
points in $V_\lambda$. Once the approximation is good enough, we simply make the partition according to this approximation. We then check the consistency of the partition with $M$, and therefore, we know whether the resulting matrix generates the same tree as the original one or not by Lemma 9. Applying this reduction to $M_6$, the number of types goes down from 132 to 34, and for $M_8$ the number of types goes down from 922 to 162. This significant reduction in the size of the matrices greatly reduces the number of constraints and variables in our linear programming formulation. In Section 8, we use this technique to simplify the branching matrix $D_G$ considered in Section 4.2 for proving Theorem 1, and reduce the matrix from 7 types to 3 types.

7 Proof of Contraction Condition Implying SSM

Proof of Lemma 2. We fix a tree $T$ in the family $\mathcal{F}_\leq M$. Assuming Condition (2) holds, we want to show that WSM holds for $T$. Note that Condition (2) is independent of the tree $T$ we choose.

As in [11], we view the boundary condition $\Gamma$ as a continuous parameter. Hence, throughout the remainder of the proof, $\Gamma \in [0, 1]$. Since we are simply aiming to prove WSM on tree $T$, we can view the boundary as all of the vertices a fixed distance $L$ from the root of $T$. Therefore, given a boundary condition $\Gamma$ and for a fixed $L$, we assign the boundary condition by each vertex at depth $L$ being fixed to be unoccupied with probability $\Gamma$ and fixed to occupied with probability $1 - \Gamma$. Note that for $L$ even, $\Gamma = 1$ corresponds to the even boundary, and $\Gamma = 0$ corresponds to the odd boundary.

Let $\alpha_{i,T}(\Gamma, L)$ be the marginal unoccupied probability for a type $i$ vertex $v$ in the tree $T_v$ rooted at $v$ where the boundary condition $\Gamma$ is assigned to the vertices at depth $L$ in $T_v$. Putting this notation into Equation (1), for the tree $T$ we have:

$$\alpha_{i,T}(\Gamma, L) = \frac{1}{1 + \lambda \prod_{w \in M_i} \alpha_{w,T}(\Gamma, L - 1)},$$

(10)

where $\alpha_{w,T}(\Gamma, L)$ equals to 1 if the vertex $w$ is not in the tree $T$, and otherwise is the marginal unoccupied probability of vertex $w$ in tree $T$ with the fractional boundary condition $\Gamma$.

By integrating over $\Gamma$ we can see that if

$$\left| \frac{d \alpha_{i,T}(\Gamma, L)}{d \Gamma} \right| \leq \gamma^L,$$

(11)

then WSM holds for $T$ at the vertex $v$ of type $i$ since the even and odd boundaries correspond to $\Gamma = 1$ and $\Gamma = 0$ (depending on the parity of $L$).

For a vertex $v$ of type $i$, we have the following equation for the derivatives at $\alpha_{i,T}(\Gamma, L)$ with respect to the boundary:

$$\frac{d \alpha_{i,T}(\Gamma, L)}{d \Gamma} = -(1 - \alpha_{i,T}(\Gamma, L))(\alpha_{i,T}(\Gamma, L)) \sum_{w \in M_i} \frac{d \alpha_{w,T}(\Gamma, L - 1)}{d \Gamma} \frac{1}{\alpha_{w,T}(\Gamma, L - 1)}.$$

(12)

From (12) it is sufficient to show for all $i$ and all $\alpha_w \in [1/(1 + \lambda), 1]$,

$$(1 - \alpha_i) \sum_{w \in M_i} \frac{1}{\alpha_w} < \gamma$$

to obtain (11) and hence WSM holds for $T$, where in the inequality $\alpha_i = (1 + \lambda \prod_{w \in M_i} \alpha_w)^{-1}$. Note that, from here we already obtain a condition that implies the WSM holds for all trees $T$ in the family $\mathcal{F}_\leq M$. 

13
However, technically, it is hard to show the contraction of the above inequality due to the nonhomogeneous marginal distributions \( \alpha_w \) from different children vertices as well as the irregular structure of the trees. We instead use a monotonic mapping \( \phi_i \) (the messages from a vertex of type \( i \) to its parent) for each type \( i \), and show that

\[
\left| \frac{d\phi_i(\alpha_i, T(\Gamma, L))}{d\Gamma} \right| \leq \gamma^L, \tag{13}
\]

which also implies that WSM holds for all trees \( T \in F_{\leq M} \).

Setting \( \Psi_i(x) = \left(x \cdot \frac{d\phi_i(x)}{dx}\right)^{-1} \), we have

\[
\frac{1}{\alpha_w} = \Psi_{t(w)}(\alpha_w) \frac{d\phi_{t(w)}(\alpha_w)}{d\alpha_w},
\]

and thus

\[
\frac{d\phi_i(\alpha_i)}{d\Gamma} = -\left(1 - \alpha_i\right) \psi_i(\alpha_i) \sum_{w \in M_i} \frac{\psi_{t(w)}(\alpha_w) \frac{d\phi_{t(w)}(\alpha_w)}{d\Gamma}}{d\alpha_w}.
\]

Notice that to obtain (13), from this last equation we just need Condition (2) to be true.

8 Proof of WSM Threshold for Arbitrary Orderings

In this section we analyze the branching matrix \( D_G \) introduced in Section 4.2, and thereby prove Lemma 6. The matrix can be reduced to a three-state tree, using the method introduced in Section 6.2, to the following rules:

1. \( A \rightarrow A \mid B \mid B \)
2. \( B \rightarrow B \mid C \)
3. \( C \rightarrow A \mid B \)

Therefore, we have a recurrence with three equations:

\[
X(x, y, z) = \frac{1}{1 + \lambda xy^2}, \quad Y(x, y, z) = \frac{1}{1 + \lambda yz}, \quad Z(x, y, z) = \frac{1}{1 + \lambda xy}.
\]

Let \( a, b, c \) be the fixed-point solution of the recurrence. Then, we have the following derivatives evaluated at the fixed point which form the associated Jacobian matrix:

\[
\frac{\partial X}{\partial x} = -\lambda a^2 b^2, \quad \frac{\partial X}{\partial y} = -2\lambda a^3 b, \quad \frac{\partial X}{\partial z} = 0,
\]

\[
\frac{\partial Y}{\partial x} = 0, \quad \frac{\partial Y}{\partial y} = -\lambda b^2 c, \quad \frac{\partial Y}{\partial z} = -\lambda b^3,
\]

\[
\frac{\partial Z}{\partial x} = -\lambda b c^2, \quad \frac{\partial Z}{\partial y} = -\lambda a c^2, \quad \frac{\partial Z}{\partial z} = 0.
\]

We use numerical methods (using Maple) to rigorously prove that when \( \lambda = 3.4 \) the WSM does not hold for the tree \( T_{D_G} \). The mathematical correctness is guaranteed by analyzing the derivatives and using the matrix perturbation theory for bounding the eigenvalues.

Proof of the lower bound in Lemma 6. Let \( \lambda = 3.3 \). We want to show that for any boundary condition on the leaves of the tree, when the height of the tree goes to infinity, there is a unique limit of the conditional probability for the root being unoccupied. We initialize the iteration of the recurrence \( r = (f, g, i) \) with \( (x, y, z) = (0, 0, 0) \) and \( (x, y, z) = (1, 1, 1) \). These are the two extremal cases of the initial values, and hence it suffices to show that they converge to the same point. When \( \lambda = 3.3 \), we can see that after several rounds of iterations by using computer, they seem to converge
to the same point. Let $O_t = r^t(0,0,0)$ and $I_t = r^t(1,1,1)$ where $t$ is the number of the steps in the iteration and when $t = 10^3$, we know that $\|O_t - I_t\|_\infty < 10^{-7} = \epsilon$. To simplify the notation we drop the sub-indices $t$.

Now we know that starting from any boundary condition, at the step $t$ of the iteration of applying the recurrence, the point will be in the cuboid that is defined by $O$ and $I$ as its diagonal vertices, i.e., the eight vertices of the cuboid will be $(O_x, O_y, O_z), (O_x, O_y, I_z), (O_x, I_y, O_z), \ldots, (I_x, I_y, I_z)$. It is easy to see that if any boundary condition converges to a fixed point, then this cuboid must also include that point since $O$ and $I$ are two extremal cases. The only thing left to show is that the Jacobian evaluated at any point in this cuboid has spectral radius strictly less than 1. By using Maple, we can evaluate the spectral radius at $O$ and $I$ and see that they are less than $1 - \epsilon'$ where $\epsilon' > 5 \cdot 10^{-3}$. Let the Jacobian evaluated at $O$ be $J_o$. We know that the Jacobian evaluated at any other point in the cuboid can be written as $J_o + E$ where $E$ is a matrix with $L_\infty$ norm less than $10^{-6}$, which can be rigorously proved by noticing that $r = (f, g, i)$ has small bounded second derivatives within the cuboid.

Finally we turn to the matrix perturbation theory for help. By applying the Bauer-Fike Theorem [1] (see also, Theorem 2.7.5 in [13]) with $L_\infty$ norm, it is easy to prove that the matrix $J_o + E$ has spectral radius less than 1 whenever $\|E\|_\infty < 10^{-3}$.

Indeed, what we want to show is for every $\lambda < 3.3$, we can prove that the spectral radius of the Jacobian is less than 1. However, there is no monotonicity of the WSM property with respect to $\lambda$ here, therefore merely showing that when $\lambda = 3.3$ WSM holds is not enough. The simplest thing to do is again by using Maple to check. We are able to do the above proof for a finite set of values for $\lambda < 3.3$. If our set is dense enough with pairwise distance less than $10^{-4}$, we are able to use the matrix perturbation theory again. Notice that for each $\lambda$ and $\lambda'$, we can write the recurrence $f = \frac{4}{1 + \lambda' xy}$ as $f = \frac{4}{1 + \lambda \frac{1}{x} \frac{1}{y}}$ and hence transfer the changes on $\lambda$ to the changes on the fixed point with the same (or even smaller) error term on the $L_\infty$ norm. Details of the proof are included in an online appendix in Maple available at [15].

\[\Box\]

\textit{Proof of the upper bound in Lemma 6.} Let us denote $F(x, y, z), G(x, y, z)$ and $I(x, y, z)$ as the two-level recurrence of $f$, $g$ and $i$, i.e., $F(x, y, z) = f(f, g, i), G(x, y, z) = g(f, g, i)$ and $I(x, y, z) = i(f, g, i)$. We call the recurrence with $F$, $G$ and $I$ as $R$. We will show that when $\lambda = 3.4$, there are two distinct attractive fixed points $P_1 = (a_1, b_1, c_1)$ and $P_2 = (a_2, b_2, c_2)$ of $R$. The proof is again done by a numerical method.

It is impossible to exactly solve the recurrence and write the fixed points in terms of $\lambda$ here. Therefore, our first step is again to find a good approximation for fixed points. As we saw in the computational experiments, iterating the recurrence from $(x, y, z) = (0, 0, 0)$ and $(x, y, z) = (1, 1, 1)$ will result in two distinct “converging points” $Q_1 = (0.6930312016, 0.5908216326, 0.5238283832)$ and $Q_2 = (0.5486919556, 0.4872669643, 0.4180328313)$, although we cannot say that $Q_1$ and $Q_2$ are the attractive fixed points. We intend to prove that $Q_1$ and $Q_2$ are close to two attractive fixed points respectively and those two fixed points are $P_1$ and $P_2$.

The way to show $Q_1$ (or $Q_2$) is very close to an attractive fixed point is by using the Brouwer fixed point theorem and checking it using numerical methods. We can put a small cuboid $C_1$ centered at $Q_1$. Let the eight vertices of $C_1$ be $Q_1 + (\pm 1.38, \pm 1.01, \pm 1.12)\epsilon$ where $\epsilon = 10^{-5}$. What we hope is that the recurrence $R$ map the cuboid $C_1$ entirely to its interior and then we are able to use the fixed point theorem to assert that there is a fixed point inside the cuboid. To check $R(C_1) \subseteq C_1$ can be done numerically: we can check for a discrete set $C_d \subseteq C_1$ whether $R(C_d)$ is in $C_1'$, where $C_1' \subseteq C_1$ is another cuboid around $Q_1$ with the minimum distance from the boundary of $C_1'$ to the boundary of $C_1$ being upper bounded from above by $12\epsilon'$ where $\epsilon' = 10^{-6}$. The set $C_d$ we choose is dense enough with pairwise $L_\infty$ distances less than $3\epsilon'$ and it is easy to prove symbolically
that the $L_1$ norm of the first order derivatives
\[
\left| \frac{\partial F}{\partial x} \right| + \left| \frac{\partial F}{\partial y} \right| + \left| \frac{\partial F}{\partial z} \right| , \quad \left| \frac{\partial G}{\partial x} \right| + \left| \frac{\partial G}{\partial y} \right| + \left| \frac{\partial G}{\partial z} \right| \quad \text{and} \quad \left| \frac{\partial I}{\partial x} \right| + \left| \frac{\partial I}{\partial y} \right| + \left| \frac{\partial I}{\partial z} \right| ,
\]
are bounded from above by 4 in $C_1$. Then, we are able to use the Fundamental theorem of calculus to assert that the entire $R(C_1) \subseteq C_1$ since the maximum error for the interpolations of all points $F$, $G$ and $I$ is upper bounded by $12\epsilon'$. The same proof method also works for $Q_2$. Actually the proof of the lemma finishes here since we already show that there are two separated regions where the recurrence could end up with depending on the initial inputs from the leaves, which means that WSM does not hold for $\lambda = 3.4$ on this tree.

Applying the Brouwer fixed point theorem for $C_1$ and $C_2$, we know that there is a fixed point $P_1$ of $R$ which is close to $Q_1$ and a fixed point $P_2$ of $R$ which is close to $Q_2$. We want to further prove that it is an attractive fixed point of $R$. What we can directly calculate are the eigenvalues of the Jacobian of $R$ at $Q_1$ (or $Q_2$) since it is just a cubic polynomial. We can see that all the eigenvalues has lengths bounded below $1 - 10^{-2}$. Since $P_1$ is close to $Q_1$, again by the continuity argument, we are able to say that the eigenvalues for the Jacobian of $R$ at $P_1$ is below 1 and hence $P_1$ is an attractive fixed point of $R$. \hfill \square

9 Conclusions

Current techniques for proving lower bounds on $\lambda_c(Z^2)$ analyze SSM on $T_{\text{saw}}(Z^2)$. This paper shows that this approach will not be sufficient to reach the conjectured threshold of 3.79... One problem in this approach is that boundary conditions obtainable on $T_{\text{saw}}(Z^2)$ are not necessary realizable on $Z^2$. Some of the boundary conditions are more “extremal” than the one that is on $Z^2$ which yields a lower weak spatial mixing threshold. Finding a way to exclude certain boundary conditions for $T_{\text{saw}}(Z^2)$ would be an extremely interesting direction.

References

[1] F. L. Bauer and C. T. Fike. Norms and exclusion theorems. Numerische Mathematik, 2:137-141, 1960.

[2] R. J. Baxter, I. G. Enting and S. K. Tsang. Hard-square lattice gas. Journal of Statistical Physics, 22(4):465-489, 1980.

[3] M. S. Bazaraa, J. J. Jarvis and H. D. Sherali. Linear Programming and Network Flows: 3rd Ed., John Wiley & Sons, 2011.

[4] J. van den Berg and J. E. Steif. Percolation and the hard-core lattice gas model. Stochastic Processes and their Applications, 49(2):179–197, 1994.

[5] A. Blanca, D. Galvin, D. Randall, and P. Tetali. Phase Coexistence and Slow Mixing for the Hard-Core Model on $Z^2$, 2012. Preprint available from the arXiv at: http://arxiv.org/abs/1211.6182

[6] G. R. Brightwell, O. Häggström, and P. Winkler. Nonmonotonic behavior in hard-core and Widom-Rowlinson models. Journal of Statistical Physics, 94(3):415–435, 1999.

[7] R. L. Dobrushin. The problem of uniqueness of a Gibbsian random field and the problem of phase transitions. Functional Analysis and its Applications, 2(4):302-312. 1968.
[8] D. S. Gaunt and M. E. Fisher. Hard-Sphere Lattice Gases. I. Plane-Square Lattice. *Journal of Chemical Physics*, 43(8):2840-2863, 1965.

[9] F. P. Kelly. Loss Networks. *Annals of Applied Probability*, 1(3):319-378, 1991.

[10] Z. Récz. Phase boundary of Ising antiferromagnets near $H = H_c$ and $T = 0$: Results from hard-core lattice gas calculations. *Physical Review B*, 21(9):4012-4016, 1980.

[11] R. Restrepo, J. Shin, P. Tetali, E. Vigoda, and L. Yang. Improved Mixing Condition on the Grid for Counting and Sampling Independent Sets. To appear in *Probability Theory and Related Fields*, 2012. A preliminary version appears in the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS), 140-149, 2011.

[12] C. Robinson. *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*, Second Edition. Boca Raton, FL, CRC Press, 1999.

[13] D. S. Watkins. *The matrix eigenvalue problem: GR and Krylov Subspace methods*. SIAM, 2007.

[14] D. Weitz. Counting independent sets up to the tree threshold. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing* (STOC), 140–149, 2006.

[15] Data is available from: [http://www.cc.gatech.edu/~vigoda/hardcore.html](http://www.cc.gatech.edu/~vigoda/hardcore.html)