Combinatorics, Probability and Computing (2026), **35**, pp. 655–687
doi:10.1017/S0963548326100467

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

On the local convergence of integer-valued Lipschitz functions on regular trees

Nathaniel Butler[1*], Kesav Krishnan[1†], Gourab Ray[1‡] and Yinon Spinka[2§]

[1]University of Victoria, Canada and [2]Tel Aviv University, Israel
Corresponding author: Kesav Krishnan; Email: kesavv8@gmail.com

(Received 24 October 2024 UTC; revised 5 March 2026 UTC; accepted 25 March 2026 UTC;
first published online 6 July 2026 UTC)

Abstract

We study random integer-valued Lipschitz functions on regular trees. It was shown by Peled, Samotij, and Yehudayoff [22] that such functions are localized; however, finer questions about the structure of Gibbs measures remain unanswered. Our main result is that the weak limit of a uniformly chosen 1-Lipschitz function with 0 boundary condition on a d-ary tree of height n exists as $n \to \infty$ if $2 \le d \le 7$, but not if $d \ge 8$, thereby partially answering a question posed by Peled, Samotij and Yehudayoff. For large d, the value at the root alternates between being almost entirely concentrated on 0 for even n and being roughly uniform on $\{-1, 0, 1\}$ for odd n, leading to different limits as n approaches infinity along evens or odds. For $d \ge 8$, the essence of this phenomenon is preserved, which obstructs the convergence. For $d \le 7$, this phenomenon ceases to exist, and the law of the value at the root loses its connection with the parity of n. Along the way, we also obtain an alternative proof of localization. The key idea is a fixed point convergence result for a related operator on ℓ^∞ and a procedure to show that the iterations get into a 'basin of attraction' of the fixed point. We also prove some accompanying analogous 'even-odd phenomenon' type results about M-Lipschitz functions on general non-amenable graphs with high enough expansion (this includes for example the large d case for regular trees). We also prove a convergence result for 1-Lipschitz functions with $\{0, 1\}$ boundary condition. This last result relies on an absolute value FKG for uniform 1-Lipschitz functions when shifted by $1/2$.

Keywords: height functions; phase transitions; localization

2020 MSC Codes: Primary: 60AXX, 60CXX, Secondary: 82BXX

1. Introduction

The main focus of this article is to study the structure of Gibbs measures of uniform M-Lipschitz function on regular trees. Our broader goal is to better understand the structure of Gibbs measures for random surface models with hardcore constraints on graphs of hyperbolic flavour. The theory of Bernoulli percolation on such graphs has seen tremendous progress in the past few decades [4, 20]. In contrast, the results on other popular statistical physics models like the Ising model on such graphs have been few and far between [9, 12–14]. On the other hand, the study of height function models on planar lattices, particularly the structure of Gibbs measures on them, have garnered a lot of attention in recent years [6, 7, 18, 21]. For finite graphs like expanders and wired

*Research supported in part by NSERC 50311-57400.
†Research supported in part by PIMS postdoctoral fellowship.
‡Research supported in part by NSERC 50311-57400.
§Research supported in part by ISF grant 1361/22.

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

regular trees, it is known [1, 8, 22, 23] that the height function models are mostly flat and the height function has a doubly exponential tail. However, finer questions about the existence of local weak limits and its dependence on boundary conditions remain unanswered. Our main focus is to initiate the study of weak local convergence of height functions on regular trees imposing various boundary conditions on the leaves. An even broader goal is to classify all possible extremal Gibbs measures on trees.

Let us fix some notations before getting into our main results. An M-**Lipschitz function** on a graph $G = (V, E)$ is a function $f : V \to \mathbb{Z}$ satisfying that

$$|f(u) - f(v)| \le M \text{ for all } u \sim v,$$

where $u \sim v$ means u is a neighbour of v. If $M = 1$, we drop the M from the terminology and simply refer to the function as a Lipschitz function. A d-ary tree is a connected graph with no cycles such that every vertex has degree $d + 1$ except one vertex, called the root vertex, which has degree d. We denote by $\mathbb{T}_d$ the d-ary tree[1] with the root vertex denoted ρ. Let $\mathbb{T}_d^k$ denote the finite subgraph of $\mathbb{T}_d$ induced by the vertices within distance k from ρ. For a finite set $S \subset \mathbb{Z}$, let $\mathcal{L}_{k,d}(S)$ denote the set of Lipschitz functions on $\mathbb{T}_d^k$ where the values on the leaves are constrained to be in S. Let $\mu_{k,d}^S$ denote the uniform law on $\mathcal{L}_{k,d}(S)$. In the special case $S = \{0\}$, which we refer to as zero boundary conditions, we write $\mu_{n,d}^0$ and $\mathcal{L}_{k,d}^0$ for short.

The main result in this article is the existence of a transition in the convergence behaviour of random 1-Lipschitz functions under zero boundary conditions.

Theorem 1.1. *$\mu_{k,d}^0$ converges as $k \to \infty$ if and only if $2 \le d \le 7$.*

The convergence above is in the local weak sense. Actually, the convergence result for $2 \le d \le 7$ in Theorem 1.1 that we prove is more general: it holds for symmetric weighted boundary conditions on the leaves which satisfy a certain decay condition (we call these the good weight sequences in Section 2.3). The result implies in particular that for any $N \ge 1$, the measures $\mu_{k,d}^{\{-N,\dots,N\}}$ converge to the same limit as that with zero boundary conditions. See Theorem 2.6 for a complete statement of the full result.

In contrast, the next result shows that for other boundary conditions, the above convergence vs. non-convergence phenomenon does not occur.

Theorem 1.2. *$\mu_{k,d}^{\{0,1\}}$ converges as $k \to \infty$ for all $d \ge 2$.*

The above theorem is a consequence of a version of an FKG property for the absolute value of the Lipschitz function *when shifted by* $1/2$.

It is easy to see for any finite S that $\mu_d^S := \lim_{k\to\infty} \mu_{k,d}^S$, when it exists, is a tree-indexed Markov chain. This follows from the observation that the set of $f \in \mathcal{L}_{k,d}(S)$ with a given value at a given vertex, splits into a Cartesian product over the subtrees rooted at that vertex. Furthermore, it is not hard to check that no such Markov chain (obtained as a limit of $\mu_{k,d}^S$) is a non-trivial mixture of other similar Markov chains (simply apply ergodic theorem along a ray). In particular, when $2 \le d \le 7$, the limiting measures $\{\mu_d^{\{i\}}\}_{i\in\mathbb{Z}}$ obtained in Theorem 1.1 and the limiting measures $\{\mu_d^{\{i,i+1\}}\}_{i\in\mathbb{Z}}$ obtained in Theorem 1.2 are Markov chains, none of which is a mixture of the others.

We also prove a general result which works beyond trees and for general M. Let G be an infinite bipartite connected graph with maximum degree d and Cheeger constant h. Recall that the Cheeger constant of a graph is defined as

$$h = h(G) := \inf_{S\subset V, 0<|S|<\infty} \frac{|\partial S|}{|S|},$$

[1] Most of our stated results continue to hold with no changes also for the $(d + 1)$-regular tree, see Section 5.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

where ∂S is the set of vertices in G, which are not in S but have at least one neighbour in S, and $|\cdot|$ denote the cardinality. We establish the existence of (multiple forms of) long-range order for M-Lipschitz functions on G in the regime where $h \gg M$. Observe that since trivially $h \le d$, the condition $h \gg M$ requires that $d \gg M$. Furthermore, note that a d-regular tree has Cheeger constant $d-2$.

Colour the partite classes of G black and white. Given a finite set U of G, and finite sets $S_\circ, S_\bullet \subset \mathbb{Z}$, let $\mathcal{L}_{U,M}(S_\circ, S_\bullet)$ denote the set of all M-Lipschitz functions on $U \cup \partial U$ with the values on ∂U constrained to be in $S_\circ$ or $S_\bullet$ according to their colour. Given integers $a \le b$, let $\nu^{a,b}_{U,M}$ denote the uniform measure on $\mathcal{L}_{U,M}(S_\circ, S_\bullet)$ with $S_\circ = \{a, \dots, b\}$, $S_\bullet = \{b-M, \dots, a+M\}$. The choice of $S_\circ, S_\bullet$ is an artefact of the proof technique: with these boundary conditions, on a graph with large Cheeger constant, a uniform M-Lipschitz function behaves approximately like i.i.d. Uniform$(S_\circ)$ on white vertices and i.i.d. uniform$(S_\bullet)$ on black vertices. Note that the sets $S_\circ$ and $S_\bullet$ are complementary in the sense that if for a given vertex our only knowledge about its value is that it belongs to $S_\circ$, then the possible values at any of its neighbours is precisely described by $S_\bullet$, and vice versa. An exhaustion of G is a sequence of finite sets increasing to G, i.e., $G_1 \subset G_2 \subset \cdots$ and $\cup_{i \ge 1} G_i = G$.

Theorem 1.3. *Assume that $h \ge 6M \log(3d^4(4M+1))$. Let $a \le b$ be integers such that $b - a \le 2M$.*

a. Let (G_k) be an exhaustion. Then $\nu^{a,b}_{G_k,M}$ converges as $k \to \infty$ to a Gibbs measure $\nu^{a,b}_{G,M}$.

b. The measures $\nu^{a,b}_{G,M}$ are distinct for different pairs (a,b).

The limiting measure $\nu^{a,b}_{G,M}$ is necessarily independent of the chosen exhaustion (as one can interleave two exhaustions and then apply the result). Observe that the pushforward of $\nu^{a,b}_{U,M}$ by a parity-preserving automorphism φ of G is $\nu^{a,b}_{\varphi(U),M}$. It follows that the limiting measure $\nu^{a,b}_{G,M}$ is invariant to parity-preserving automorphisms of G. Similarly, the pushforward of $\nu^{a,b}_{G,M}$ by a parity-reversing automorphism is $\nu^{b-M,a+M}_{G,M}$. Thus, $\nu^{a,b}_{G,M}$ is invariant to all automorphisms if and only if $b - a = M$. Consequently, when $b - a \ne M$, the measure $\nu^{a,b}_{G,M}$ behaves differently on even and odd vertices. In fact, the height on any even vertex is close to uniform on $\{a, \dots, b\}$, and the height on any odd vertex is close to uniform on $\{b-M, \dots, a+M\}$ (see Proposition 3.1). Thus, Theorem 1.3 gives a separate proof of the non-convergence result in Theorem 1.1 when d is large enough. Let us also remark that in the case when $b - a = M$, the result stated in the theorem makes sense and holds true even without the assumption that the graph G is bipartite (in this case $\nu^{a,b}_{U,M}$ is the uniform measure on M-Lipschitz functions constrained to taking values in $\{a, \dots, b\}$ on the entire boundary).

Background: Studying Gibbs measures of statistical physics models on trees has a long history. We refer to the book of Rozikov [25] for a comprehensive survey of the results about splitting Gibbs measures on regular trees for models with countable and even uncountable spin values. It is interesting to note that the results in [25] about splitting Gibbs measure for countable spins (see [25, Theorem 8.1]) lead to a unique Gibbs measure. This is clearly not the case for us, as Theorem 1.1 (see also Theorem 2.6) already leads to multiple solutions, and it is plausible that there are many more (see Section 5).

The random homomorphism model is very similar to the uniform Lipschitz model, except that the height difference between adjacent vertices must be in $\{\pm 1\}$ (this makes sense only on bipartite graphs). Both the uniform Lipschitz model and the random homomorphism model on trees have been studied in [1, 22] and on finite expanders in [23]. A recent paper by Lammers and Toninelli [18] exploits log concavity to obtain localization for homomorphisms with very general boundary conditions. See also a recent result by Krueger, Li, and Park [17], which significantly improves the

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

results in [23]. See [16] for a result about the Widom–Rowlinson model on trees, which is a model of similar flavour to ours (it can be thought of as a restricted Lipschitz model where the function cannot take values other than $\{-1, 0, 1\}$). See Peled and Spinka [21, 24] for analogous results in the hypercubic lattice. See also [2, 3, 5] for related results.

As mentioned, the fact that random Lipschitz functions on trees under zero boundary conditions are localized was first proved by Peled, Samotij, and Yehudayoff [22]. In fact, they showed that the height at the root has doubly exponential tails for all $M \geq 1$ and $d \geq 2$. They also point out[2] that when $M \leq cd/\log d$ for some universal constant $c > 0$, more is true – for even n, the height at the root is exponentially concentrated on the single value 0, whereas for odd n, the height at the root is roughly uniformly distributed on $\{-M, \ldots, M\}$. They go on to explain that they expect that such a strong concentration fails when $M \gg d$. Specifically, they raised their suspicion that the height at the root converges in distribution as n tends to infinity (so that there is asymptotically no distinction between even and odd tree heights). While we do not establish that such a transition phenomenon occurs between the cases $M \ll d$ and $M \gg d$ for all M, we do show in Theorem 1.1 that such a transition occurs as d is varied when $M = 1$. Extending this to all M is an interesting problem.

Proof outline: Let us briefly outline the ideas behind the proofs. Theorem 1.1 is the major result in this article and takes up most of the analysis. There are two separate parts, convergence for $2 \leq d \leq 7$ and non-convergence for $d \geq 8$, the latter being much simpler. The basic idea for the convergence part is a contraction argument: we think of the probability distribution at the root as an element of $\ell^1(\mathbb{Z})$, and then we apply iterates of an appropriate operator F. The goal then reduces to proving that the iterates of F (applied to δ_0, i.e., the probability measure assigning unit mass to 0) converge in an appropriate norm. The proof of this would have been somewhat straightforward if the map F turned out to be a contraction (but it isn't). Nevertheless, we show that after a number of iterations, we get into a certain 'basin of attraction', on which the map becomes a contraction. To show the latter, we prove that the operator norm of the derivative of F is strictly less than 1 in this basin. After this, the proof concludes by a simple application of the Banach fixed point theorem.

In practice, directly applying this idea to F soon becomes intractable. The way we work our way through is to look at ratios of the consecutive values, which gets rid of the normalizing factor in a good way. For the non-convergence part of the theorem, we show that the iterates oscillate between two separated 'basins', ruling out the possibility of convergence (though we still expect convergence of $\mu^0_{k,d}$ along even k). The details of these ideas can be found in Section 2.

Theorem 1.2 follows from a certain 'FKG for absolute value' result when the height function is shifted by 1/2. Finally, Theorem 1.3 is a consequence of a variation of a Peierl's type argument following that in [22, 23].

Organization: In Section 2, we describe the setup and the ratio transformation which simplifies the operator F. We also describe heuristically which observables from this operator drives the transition from convergence to non-convergence. In Section 2.1 we define the set which is the 'basin of attraction' and describe our procedure to get into this set. In Section 2.2, we prove the non-convergence part of Theorem 1.1. In Sections 2.3–2.5 we prove the convergence part of Theorem 1.1 modulo some technical estimates, which we push to Section 6. We prove Theorem 1.2 in Section 4 and Theorem 1.3 in Section 3. We list some open questions in Section 5.

[2]This was announced in [22] but was never published. Theorem 1.3 is our attempt to write down a precisely formulated general result and give the details of the proof.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

2. Convergence and non-convergence under 0 boundary conditions

In this section, we prove Theorem 1.1. The proof is separated into the convergence result for $2 \le d \le 7$ (Section 2.3) and the non-convergence result for $d \ge 8$ (Section 2.2). The proofs are based on a recursive approach. We now lay the groundwork for this.

For a probability distribution $z = (z_i)_{i\in\mathbb{Z}}$, define

$$A(z)_i := (z_{i-1} + z_i + z_{i+1})^d, \qquad F(z) := \frac{A(z)}{\sum_{i\in\mathbb{Z}} A(z)_i}.$$

The particular form of $A(z)$ arises from the Lipschitz condition, while $F(z)$ is just a normalization of $A(z)$. We write $F^{(k)}$ for the k-fold composition of F, i.e., $F^{(0)}$ is the identity and $F^{(k+1)} = F \circ F^{(k)}$.

Claim 2.1. *Let $f_k \sim \mu^0_{k,d}$, and let ρ be the root vertex. Then for any $i \in \mathbb{Z}$,*

$$\mathbb{P}(f_k(\rho) = i) = (F^{(k)}(\delta_0))_i.$$

Proof. Note that $F^{(k)}(\delta_0)$ is simply the normalization of $A^{(k)}(\delta_0)$ to a probability measure (since A is homogeneous in z). Let $A^{(0)}$ be the identity. Thus, the claim follows if we show that the number of $f \in \mathcal{L}^0_{k,d}$ having $f(\rho) = i$ equals $(A^{(k)}(\delta_0))_i$. We show this by induction on k. The base of the induction, $k = 0$. is trivial. Fix $k \ge 1$. Since the Lipschitz constraint is a condition on nearest neighbour values, we have that the number of Lipschitz functions f in question is L^d, where L is the number of choices for f in one given subtree. By the Lipschitz condition, L is the number of $f' \in \mathcal{L}^0_{k-1,d}$ having $f'(\rho') \in \{i-1, i, i+1\}$, where ρ' is the root in $\mathbb{T}^{k-1}_d$. By the induction hypothesis, the number of such f' is $A^{(k-1)}(\delta_0)_{i-1} + A^{(k-1)}(\delta_0)_i + A^{(k-1)}(\delta_0)_{i+1}$. This proves the claim. □

Thus, vague convergence of the distribution of $f_k(\rho)$ is determined by the pointwise convergence of $F^{(k)}(\delta_0)$ (tightness is a separate issue which needs to be addressed). In order to understand the latter, our strategy is to perform a change of coordinates which gets rid of the normalizing factor in F and thereby makes things more amenable to analysis. Define a map $\mathsf{R}\colon [0,\infty)^{\mathbb{N}} \to [0,\infty)^{\mathbb{N}}$ by

$$\mathsf{R}(z)_i = \begin{cases} \dfrac{z_i}{z_{i-1}} & \text{if } z_{i-1} \neq 0 \\ 0 & \text{if } z_{i-1} = 0. \end{cases} \tag{2.1}$$

In order for R to be injective, we restrict its domain to the set $\mathcal{E}$ of symmetric probability distributions whose support is an interval or all of $\mathbb{Z}$; i.e., probability distributions z such that $z_i = z_{-i}$ for all $i > 0$, and $z_i = 0$ implies that $z_{i+1} = 0$ for $i > 0$. Let $\mathcal{R}$ be the image of $\mathcal{E}$ under R. Observe that R is indeed injective and hence invertible. Define a new map $\psi : \mathcal{R} \to \mathcal{R}$ by

$$\psi := \mathsf{R} \circ F \circ \mathsf{R}^{-1}.$$

Observe that F preserves $\mathcal{E}$, i.e., $\varphi(z) \in \mathcal{E}$ if $z \in \mathcal{E}$. Hence, ψ preserves $\mathcal{R}$. Thus,

$$\psi^{(k)} = \mathsf{R} \circ F^{(k)} \circ \mathsf{R}^{-1} \qquad \text{and} \qquad F^{(k)} = \mathsf{R}^{-1} \circ \psi^{(k)} \circ \mathsf{R}.$$

Noting that $\mathsf{R}(\delta_0) = \mathbf{0} := (0, 0, \dots)$, and that R^{-1} is continuous (with respect to the pointwise topologies), we see that in order to establish the pointwise convergence of $F^{(n)}(\delta_0)$, it is sufficient to show the pointwise convergence of $\psi^{(n)}(\mathbf{0})$. The converse is also true so that pointwise convergence of these sequences is equivalent, since R is continuous on the subset of $\mathcal{E}$ of fully supported distributions (where the limit must clearly reside), but we will not use this.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Let us record already here the expressions for the coordinates of ψ. Writing $\psi_n(\cdot) = \psi(\cdot)_n$, we have by writing $\mathsf{R}^{-1}(x) = z$,

$$\psi_1(x) = \frac{(z_0 + z_1 + z_2)^d}{(z_{-1} + z_0 + z_1)^d} = \left(\frac{1 + x_1 + x_1x_2}{1 + 2x_1}\right)^d, \tag{2.2}$$

where we used the fact that $z_1 = z_{-1}$ and that $z_0 \neq 0$ since $z \in \mathcal{E}$. Similarly, for $n \geq 2$, we get

$$\psi_n(x) = \frac{(z_{n-1} + z_n + z_{n+1})^d}{(z_{n-2} + z_{n-1} + z_m)^d} = x_{n-1}^d \left(\frac{1 + x_n + x_nx_{n+1}}{1 + x_{n-1} + x_{n-1}x_n}\right)^d. \tag{2.3}$$

We again note that if $z_{n-2} = 0$, then all $z_i = 0$ for $i \geq n - 2$ (since $z \in \mathcal{E}$), and by our definition of R, both expressions above are 0. For the same reason, if $z_{n-1} = 0$ and $z_{n-2} > 0$, then also both expressions above are 0.

Let us now provide a rough explanation of what properties of these functions change from the convergence phase ($d \leq 7$) to the non-convergence phase ($d \geq 8$). For $1 \leq \alpha < 2$ and $0 \leq x \leq 1$, define

$$f(\alpha, x) := \left(\frac{1 + \alpha x}{1 + 2x}\right)^d. \tag{2.4}$$

Letting $(1 + x_2) = \alpha$, we see that ψ_1 is the same as the map $f(\alpha, x)$. It is known [22] that the value of a uniformly sampled Lipschitz function at the root has double exponential decay. For us, this would correspond to the eventual basin of attraction of F. Remember if $\mathbb{P}(f_{k-1}(\rho) = i) = z_i$, then $\mathbb{P}(f_k(\rho) = i) = F(z)_i$. Thus, we are interested in the behaviour of the iterates of F on distributions with doubly exponential decay. Thus, the ratios $x_n = z_n/z_{n-1}$ should be rapidly decaying, and in particular, we can expect x_n to be small already for $n \geq 2$. It is reasonable to look at $f(1, x)$ as a test case. Although $f(1, x)$ is decreasing and convex and has a single fixed point x_*, $|f'(1, x_*)| > 1$ if and only if $d \geq 8$. Furthermore, it turns out that $f \circ f$ starts having multiple fixed points for $d \geq 8$. Also, $|(f \circ f)'| < 1$ for all $0 \leq x \leq 1$ if $2 \leq d \leq 7$, for $d \geq 8$, $(f \circ f)'(x_*) > 1$. These facts can be proved directly, we show a few plots here in Fig. 1. In reality, however, x_2 is not constant but fluctuates. Therefore, to prove convergence, we need to look at the full operator norm.

The rest of this section is organized as follows. We first construct an 'envelope dynamics' in Section 2.1 which gives a decent approximation of the full dynamics of ψ. We then use this to prove non-convergence for $d \geq 8$ in Section 2.2. Finally, we continue on to the longer and more intricate proof of convergence for $2 \leq d \leq 7$ in Section 2.3, for which the envelope dynamics only serves as an initial step to obtain an approximation of the fixed point.

2.1 Envelope dynamics

We now describe subsets of $\mathcal{R}$ which capture much of the essence of the dynamics given by ψ. We manage to describe these sets via three parameters, which give an interval for the range of the first coordinate and a maximum for the rest of the coordinates.

Because of the fast decay, intuition suggests the first coordinate plays the dominant role in the dynamics. In practice, the behaviour of the first coordinate yields appropriate upper bounds on the rest. For numbers $0 \leq a, c \leq b \leq 1$, denote

$$\mathcal{S}_{a,b,c} := \{x \in \mathcal{R} : a \leq x_1 \leq b,\ 0 \leq x_n \leq c \text{ for all } n \geq 2\}.$$

Define $\varphi\colon [0, 1]^3 \to [0, 1]^3$ by

$$\varphi(a, b, c) := \left(\left(\frac{1 + b}{1 + 2b}\right)^d, \left(\frac{1 + a + ac}{1 + 2a}\right)^d, b^d \left(\frac{1 + c + c^2}{1 + b + bc}\right)^d\right). \tag{2.5}$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

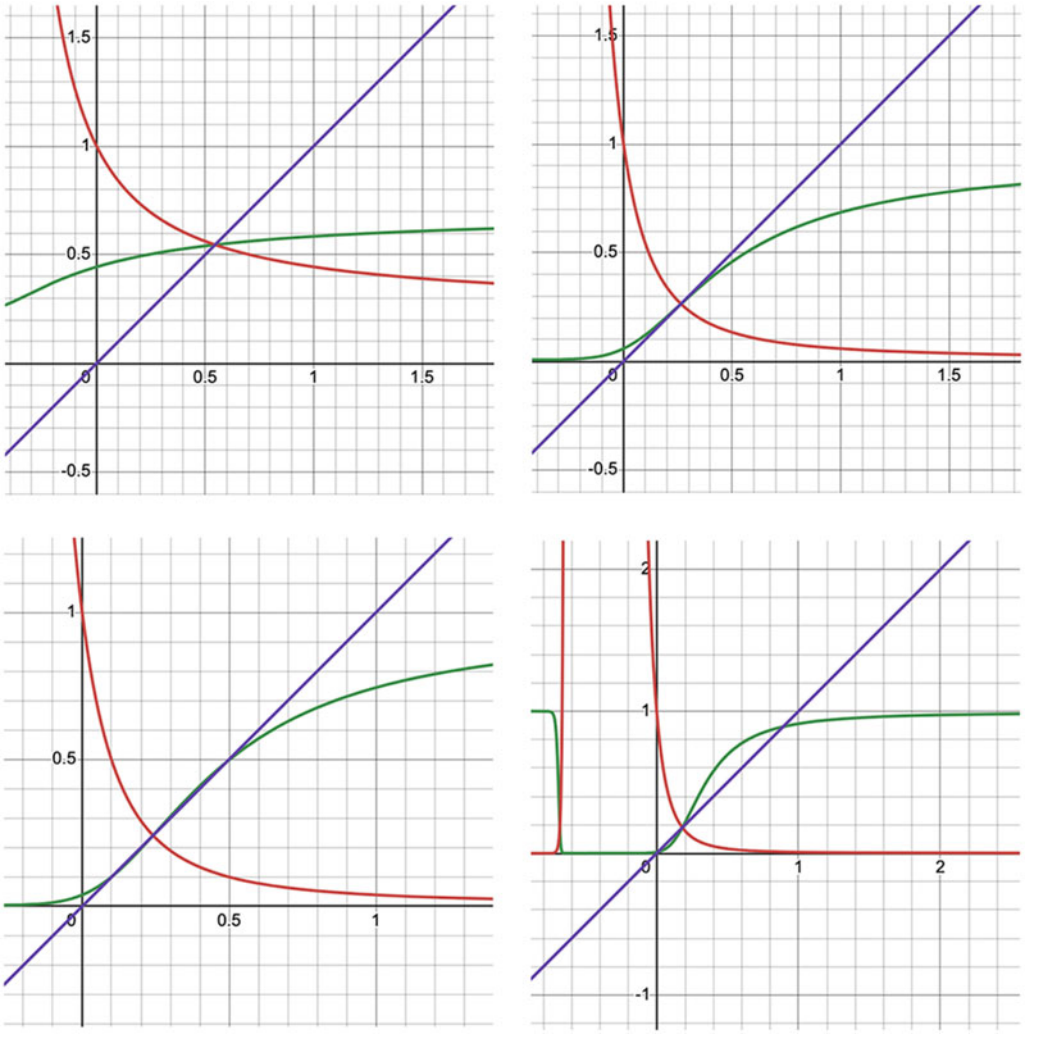


Figure 1. Plots of $f = f(1, x)$ (in red), $f \circ f$ (in green), and $y = x$ (in blue) for different values of d. Top left: $d = 2$; top right: $d = 7$; bottom left: $d = 8$; bottom right: $d = 12$. Multiple fixed points of $f \circ f$ start appearing for $d \geq 8$.

The following lemma describes the evolution of the sets $\mathcal{S}_{a,b,c}$ under applications of ψ. Denote

$$\mathcal{I} := \{(a, b, c) \in [0, 1]^3 : a, c \leq b\}.$$

Lemma 2.2. *Let* $(a, b, c) \in \mathcal{I}$. *Then* $(a', b', c') := \varphi(a, b, c) \in \mathcal{I}$ *and* $\psi(\mathcal{S}_{a,b,c}) \subset \mathcal{S}_{a',b',c'}$.

Before proving Lemma 2.2, it is useful to make some observations. For $0 \leq b, x \leq 1$, define

$$g(b, x) := b^d \left(\frac{1 + x + x^2}{1 + b + bx}\right)^d. \tag{2.6}$$

Note that with these definitions we have

$$\varphi(a, b, c) := (f(1, b), f(1 + c, a), g(b, c)).$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

The following are simple observations. For $x, \alpha, \beta, b > 0$

$$\frac{1+\alpha x}{1+\beta x} \text{ is decreasing (resp. increasing) in } x \text{ if } \alpha < \beta \text{ (resp. } \alpha > \beta). \tag{2.7}$$

$$g(b,x) \text{ is increasing in } b. \tag{2.8}$$

To verify (2.8), simply divide the the numerator and the denominator by b^d and observe that after cancellations, the resulting expression is increasing in b. Before proving the main part of Lemma 2.2, let us show that φ preserves $\mathcal{I}$.

Claim 2.3. *φ preserves $\mathcal{I}$, i.e., $f(\mathcal{I}) \subset \mathcal{I}$.*

Proof. Let $(a,b,c) \in \mathcal{I}$ and set $(a',b',c') := \varphi(a,b,c)$. We need to check that $a' \le b'$ and $c' \le b'$. For the first, we note $f(1,b) \le f(1,a)$ using (2.7) and clearly $f(1,a) \le f(1+c,a)$ For the second, we need to check that $g(b,c) \le f(1+c,a)$. First note that $g(b,c) \le g(1,c)$ by (2.8). Next, to see that $g(1,c) \le f(1+c,a)$, we need to check that $(1+c+c^2)(1+2a) \le (2+c)(1+a+ac)$. Equivalently,

$$1+c+c^2+2a+2ac+2ac^2 \le 2+2a+3ac+c+ac^2,$$

which after cancellations becomes $c^2(1+a) \le 1+ac$, which is trivially true since $c \le 1$. □

Proof of Lemma 2.2. Recalling (2.2), we have

$$\psi_1(x) = \left(\frac{1+x_1+x_1x_2}{1+2x_1}\right)^d \le \left(\frac{1+x_1+cx_1}{1+2x_1}\right)^d \le \left(\frac{1+a+ac}{1+2a}\right)^d.$$

where the second inequality follows from (2.7). Similarly

$$\psi_1(x) = \left(\frac{1+x_1+x_1x_2}{1+2x_1}\right)^d \ge \left(\frac{1+x_1}{1+2x_1}\right)^d \ge \left(\frac{1+b}{1+2b}\right)^d.$$

Fix $n \ge 2$. Recall (2.3). Since $\psi_n(x)$ is clearly increasing in x_{n+1}, we have

$$\psi_n(x) \le x_{n-1}^d \left(\frac{1+x_n+cx_n}{1+x_{n-1}+x_{n-1}x_n}\right)^d.$$

Since the latter expression is increasing in x_{n-1} (this can be seen easily by dividing both the numerator and denominator by x_{n-1}^d) and $b \ge c$ so that $x_{n-1} \le b$, we have

$$\psi_n(x) \le b^d \left(\frac{1+x_n+cx_n}{1+b+bx_n}\right)^d.$$

Since this expression is increasing in x_n (follows from (2.7)), we have

$$\psi_n(x) \le b^d \left(\frac{1+c+c^2}{1+b+bc}\right)^d.$$

thereby completing the proof. □

2.2 Non-convergence for $d \ge 8$

Proposition 2.4. *Fix $d \ge 8$, $a = .4$, $c = .01$, $a' = (.78)^d$. Then*

$$\psi \circ \psi(\mathcal{S}_{a,1,c}) \subset \mathcal{S}_{a,1,c} \qquad \text{and} \qquad \psi(\mathcal{S}_{a,1,c}) \cap \mathcal{S}_{a',1,1} = \emptyset.$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Proof. Denote

$$c' = \left(\frac{1+c+c^2}{2+c}\right)^d,$$

and recall from Lemma 2.2 that

$$\psi(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{0,1,c'}.$$

Note that c' is decreasing in d. Therefore, $c' \le (c')^{8/d} = .0040677 < c$, and hence,

$$\psi(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{0,1,c}.$$

Pick $x \in \mathcal{S}_{a,1,c}$, and let $y = \psi(x)$ and $z = \psi(y)$. Then $\psi(\mathcal{S}_{a,1,c}) \subset \psi(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{0,1,c}$. Therefore, $y_n \le c$ and $z_n \le c$ for all $n \ge 2$. Now applying these bounds, using that $x \mapsto \frac{(1+\beta x)^d}{(1+2x)^d}$ is decreasing in x if $\beta < 2$ and $x_1 \ge a$,

$$y_1 = \frac{(1+(1+x_2)x_1)^d}{(1+2x_1)^d} \le \frac{(1+1.01x_1)^d}{(1+2x_1)^d} \le \frac{(1+1.01a)^d}{(1+2a)^d} = (.78)^d = a'.$$

Now observe that

$$z_1 = \frac{(1+(1+y_2)y_1)^d}{(1+2y_1)^d} \ge \frac{(1+y_1)^d}{(1+2y_1)^d} \ge \frac{(1+(.78)^d)^d}{(1+2(.78)^d)^d} =: \gamma(d). \tag{2.9}$$

For $d = 8, 9$, we compute $\gamma(d)$ directly and see that

$$\gamma(8) = .402\cdots > a \qquad \text{and} \qquad \gamma(9) = .436\cdots > a.$$

For $d \ge 10$, we use the bound

$$\gamma(d) = \left(1 + \frac{(.78)^d}{1+(.78)^d}\right)^{-d} \ge \exp(-d(.78)^d) \ge \exp(-10(.78)^{10}) = 0.434\cdots > a,$$

where we used that $x(.78)^x$ is decreasing for $x > 1/\log(1/.78) \approx 4.025$. □

Theorem 2.5. *Fix $d \ge 8$. The sequence $\psi^{(k)}(\mathbf{0})$ does not converge pointwise.*

Proof. Note, $\psi(\mathbf{0}) = \delta_1$, the sequence which takes value 1 at $n = 1$ and is 0 everywhere else. The theorem is now immediate from the previous proposition as it implies that $\psi^{(2k+1)}(\mathbf{0})_1 \ge .4$ and $\psi^{(2k+2)}(\mathbf{0})_1 \le (0.78)^d \le 0.14$ for all $k \ge 0$. □

2.3 Convergence for $2 \le d \le 7$

We shall prove a generalization of the convergence result stated in Theorem 1.1. We say a sequence of weights $w = (w_i) \in [0,\infty)^{\mathbb{Z}}$ is **good** if it is symmetric (i.e. $w(i) = w(-i)$ for all $i \in \mathbb{Z}$), $w(0) > 0$, and there exists a $N \in \mathbb{N} \cup \{0\}$ and a $c \in [0,1)$ such that

$$w(i) = w(i+1) \quad \text{for all } 0 \le i < N, \qquad \text{and} \qquad w(i+1) \le c \cdot w(i) \quad \text{for all } i \ge N.$$

In other words, the distribution is flat in a symmetric interval around 0 and decays at least at an exponential rate c outside this interval. Let $\mu_{k,d}^w$ be the probability distribution on Lipschitz functions on $\mathbb{T}_{k,d}$ where each Lipschitz function f is weighted by the product of $w(f(v))$ over the leaves v. Note that $\mu_{k,d}^{\{-N,\ldots,N\}}$ is a special case of such a measure.

Theorem 2.6. *Fix $2 \le d \le 7$. Then $\lim_{k\to\infty} \mu_{k,d}^w$ exists for all good weight sequences w and is independent of w. In particular, $\lim_{k\to\infty} \mu_{k,d}^{\{-N,\ldots,N\}}$ exists for all $N \ge 0$ and is independent of N.*

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Note that $\mu^w_{k,d}$ is not affected by a scaling of w. Thus, we may assume that w is itself a probability distribution. Note in particular that this means it belongs to $\mathcal{E}$ (it is symmetric, and its support is an interval or all of $\mathbb{Z}$). Our broad goal is to prove that the iterates of ψ starting from $\varphi(w)$ converge pointwise. Along the way we also argue that all iterates stay in the set $\mathcal{S}_{0,1,c}$, whose preimage under φ is a tight family of distributions. Consequently, we argue that $f_k(\rho)$ converges in distribution and that $\mu^0_{k,d}$ converges weakly.

The major step is to prove the following.

Proposition 2.7. *Fix $2 \le d \le 7$. Let w be a good sequence of weights normalized to be a probability distribution. There exists $c \in (0,1)$ such that the $\psi^{(k)}(\mathsf{R}(w))$ belongs to $\mathcal{S}_{0,1,c}$ for all k large enough and converges pointwise as $k \to \infty$.*

Proof of Theorem 2.6 assuming Proposition 2.7. It is not hard to check that R^{-1} is continuous on its domain $\mathcal{R}$ (when both the domain and range are endowed with the pointwise topology). Since $F^{(k)}(w) = \mathsf{R}^{-1} \circ \psi^{(k)} \circ \mathsf{R}(w)$, we conclude from Proposition 2.7 that $F^{(k)}(w)$ converges pointwise and that $F^{(k)}(w)$ belongs to $\mathsf{R}^{-1}(\mathcal{S}_{0,1,c})$ for all k large enough. Observe that $\mathsf{R}^{-1}(\mathcal{S}_{0,1,c})$ is a tight family of distributions. Thus, letting $f_k \sim \mu^w_{k,d}$ and recalling Claim 2.1 (or rather its straightforward extension from $\mu^0_{k,d}$ to $\mu^w_{k,d}$), this implies that $f_k(\rho)$ converges in distribution as $n \to \infty$.

It is now not too hard to see that f_n converges locally. We very briefly give the idea. Fix $r \ge 0$ and consider the restriction $f_{k,r}$ of f_k to L_k, where L_r is the subtree of depth r (assume k large). Then, for a given Lipschitz function ξ on L_r, we have that $\mathbb{P}(f_{n,r} = \xi)$ is proportional $\prod_v \mathbb{P}(f_{k-r}(\rho) = \xi(v))$, where the product is over the leaves of L_r. Since this product converges as $k \to \infty$, we see that the distribution of $f_{k,r}$ converges as $n \to \infty$, which shows that f_k converges locally. □

To prove Proposition 2.7 we employ the following strategy. First, we shall show that

$$\text{if } w \text{ is good then there exist } k \ge 0 \text{ and } c \in [0,1) \text{ such that } \psi^{(k)}(\mathsf{R}(w)) \in \mathcal{S}_{0,1,c}. \tag{2.10}$$

This reduces the problem to understanding the iterations of ψ starting from an initial point in $\mathcal{S}_{0,1,c}$. For this we will show that there exist numbers a_*, b_*, c_* with $0 \le a_*, c_* \le b_* \le 1$ such that

$$\text{for every } 0 \le c < 1 \text{ there exists } k \text{ such that } \psi^{(k)}(\mathcal{S}_{0,1,c}) \subseteq \mathcal{S}_{a_*,b_*,c_*}, \tag{2.11}$$

and

$$\psi \text{ preserves the convex set } \mathcal{S}_{a_*,b_*,c_*} \text{ and is a contraction on it.} \tag{2.12}$$

The contraction referred to in (2.12) is with respect to the metric induced by the norm given by

$$\|x\| := \begin{cases} \sup_{i \ge 1} |x_i| & \text{if } d = 2 \\ |x_1| + \sup_{i \ge 2} |x_i| & \text{if } 3 \le d \le 7 \end{cases}. \tag{2.13}$$

Let us remark that the two norms are equivalent (of course, in the case of $d = 2$, this is just the usual ℓ^∞-norm). However, the map ψ is not a contraction with respect to the ℓ^∞ norm when $d \ge 4$ even very near the fixed point (see Table 4), and for this reason, we use a slightly modified norm (we use this norm also for $d = 3$). Let us mention that the ℓ^1-norm is also a natural candidate, but that this would not allow to obtain the full stated result as $\mathcal{S}_{0,1,c}$ is not in ℓ^1.

Let us show how the above easily yields Proposition 2.7.

Proof of Proposition 2.7 assuming (2.10)–(2.12). Let w be a good sequence of weights normalized to be a probability distribution. Denote $x^{(k)} := \psi^{(k)}(R(w))$. By (2.10), there exists k_0 and c such that $x^{(k_0)} \in \mathcal{S}_{0,1,c}$. By (2.11), there exists k_1 such that $x^{(k_0+k_1)} \in \mathcal{S}_{a_*,b_*,c_*}$. By (2.12), there exists $\lambda \in (0,1)$ such that $x^{(k)} \in \mathcal{S}_{a_*,b_*,c_*}$ and $\|x^{(k+1)} - x^{(k)}\| \le \lambda \|x^{(k)} - x^{(k-1)}\|$ for all $k \ge k_0 + k_1$. Since $\mathcal{S}_{a_*,b_*,c_*} \subset \mathcal{S}_{0,1,c_*}$, we have that $x^{(k)} \in \mathcal{S}_{0,1,c_*}$ for all k large enough, and since $\|\cdot\|$ makes $\mathcal{S}_{0,1,c_*}$ a

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Table 1. The fixed point of φ. Simulated in computer by running ten thousand iterations started from $a = 0, b = 1, c = .9$

d	2	3	4	5	6	7	8
a_*	.5192	.4374	.3762	.3294	.2932	.2645	.1027
b_*	.6335	.4649	.3828	.3310	.2935	.2646	.4906
c_*	.1988	.0344	.0060	$9.5 \cdot 10^{-4}$	$1.4 \cdot 10^{-4}$	$1.8 \cdot 10^{-5}$	$1.4 \cdot 10^{-4}$

complete metric space, we have that $x^{(k)}$ converges in this space, which clearly implies pointwise convergence. □

The next two sections are devoted to the proofs of (2.10)–(2.12). In Section 2.4, we prove (2.10) and (2.11), and in Section 2.5, we prove (2.12).

2.4 Getting absorbed into the basin of attraction

***Proof of* (2.10).** For $n \geq 1$, let $\mathcal{A}_n$ be the set of $x \in \mathcal{R}$ such that $x_i = 1$ for $i \leq n$ and $\sup_{i>n} x_i < 1$. Note that $\mathcal{A}_0 \cup \mathcal{A}_1 = \bigcup_{c\in[0,1)} \mathcal{S}_{0,1,c}$. Let w be a good sequence of weights normalized to be a probability distribution. Note that $R(w) \in \bigcup_{n\geq 0} \mathcal{A}_n$. Thus, it suffices to show that

$$\psi(\mathcal{A}_n) \subset \mathcal{A}_{n-1} \qquad \text{for all } n \geq 2.$$

Fix $n \geq 2$ and $x \in \mathcal{A}_n$. Denote $c := \sup_{i>n} x_i < 1$. It is clear from (2.2) and (2.3) that $\psi_i(x) = 1$ for $i \leq n-1$ and that $\psi_i(x) < 1$ for $i \geq n$. Moreover, in the same manner as in the proof of Lemma 2.2, one obtains a uniform upper bound on $\psi_i(x) < 1$ for $i \geq n$, namely,

$$\psi_i(x) \leq \left(\frac{1 + c + c^2}{2 + c}\right)^d.$$

This shows that $\psi(x) \in \mathcal{A}_{n-1}$. □

We now move on to the proof of (2.11). Given Lemma 2.2, the proof of (2.11) boils down to understanding the behaviour of the iterations of φ applied to $\mathcal{S}_{0,1,c}$. It is not too hard to show that the iterations of φ do converge to a fixed point (a_*, b_*, c_*) if we start from $a = 0, b = 1, c > 1/2$ (this part is true for all values of d). Nevertheless to have estimates of the derivative of ψ, we need good estimates on (a_*, b_*, c_*). Iterations in a computer yield the estimates of (a_*, b_*, c_*) laid out in Table 1.

For $d \geq 8$, iterations show that a_* and b_* are far apart, which is a manifestation of the non-convergence (in reality, as we have seen in Section 2.2, the actual value of x_1 alternates between being close to a_* and b_*).

To make the result completely rigorous, we obtain rigorous estimates of a_*, b_*, c_*, outlined in Table 2. While we could have obtained much better estimates (closer to the actual fixed point), it will turn out that these estimates are enough to bound the operator norm of the derivative of ψ (with respect to the norm (2.13) on the base space). For larger values of d (especially $d = 7$), we need slightly better precision.

Proposition 2.8. *Fix $2 \leq d \leq 7$. Let $(\hat{a}_*, \hat{b}_*, \hat{c}_*)$ be as in Table 2. Then, for any $0 \leq c < 1$, $\psi^{(k)}(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{\hat{a}_*,\hat{b}_*,\hat{c}_*}$ for all large enough k.*

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Table 2. Rigorous estimates of a_*, b_*, c_*. Proofs can be found in Section 6

d	2	3	4	5	6	7
$\hat{a}_*$	.5	.4	.3	.3	.27	.26
$\hat{b}_*$	.7	.6	.5	.4	.32	.27
$\hat{c}_*$	.27	.2	.1	.1	.1	.01

Table 3. The fixed point x_* and some partial derivatives of ψ at this point. Simulated in computer by running a million iterations

d	2	3	4	5	6	7
x_1	.5992	.4555	.3803	.3303	.2936	.2646
x_2	.1712	.0327	.0059	$9.5 \cdot 10^{-4}$	$1.4 \cdot 10^{-4}$	$1.8 \cdot 10^{-5}$
x_3	.0222	$3.2 \cdot 10^{-5}$	$1.2 \cdot 10^{-9}$	$1.6 \cdot 10^{-16}$	$6.4 \cdot 10^{-24}$	$5.1 \cdot 10^{-34}$
x_4	$4.7 \cdot 10^{-4}$	$3.2 \cdot 10^{-14}$	$1.8 \cdot 10^{-36}$	$2.5 \cdot 10^{-76}$	$6.8 \cdot 10^{-140}$	$9.1 \cdot 10^{-234}$
$-\frac{\partial \psi_1}{\partial x_1}$	.2655	.4704	.6213	.7467	.8575	.9577
$\frac{\partial \psi_1}{\partial x_2}$	.4220	.4234	.4184	.4100	.3992	.3874
$\frac{\partial \psi_2}{\partial x_1}$	.3357	.1467	$1.8 \cdot 10^{-7}$	.0108	.0022	$3.7 \cdot 10^{-4}$
$\frac{\partial \psi_2}{\partial x_2}$	.1773	.0647	$4.6 \cdot 10^{-9}$	.0036	$6.3 \cdot 10^{-4}$	$9.7 \cdot 10^{-5}$

The proof of Proposition 2.8 is rather technical, so we postpone it to Section 6 on order not to disrupt the flow of the argument. Readers who are content with the computer estimates in Table 1 can skip Section 6 altogether. For estimates of the fixed point and some derivatives evaluated at the fixed point, please see Table 3.

2.5 Contraction via the derivatives

Our goal in this section is to prove (2.12). We do this by bounding the derivatives of ψ. Recall the norm $\|\cdot\|$ defined in (2.13). It is straightforward to check that the Frechét derivative of the map ψ evaluated at x is the linear operator on the vector space spanned by $\mathcal{R}$ given by

$$(D\psi_x(y))_i = (D\psi_i)_x(y) = \sum_{j \geq 1} \frac{\partial \psi_i}{\partial x_j} \cdot y_j.$$

Note that $\partial \psi_i / \partial \psi_j = 0$ unless $|i - j| \leq 1$. Recall that the operator norm of $D\psi_x$ is $\|D\psi_x\| := \sup_{y:\|y\|=1} \|D\psi_x y\|$. Here the norm we use is as in Equation (2.13).

Proposition 2.9. *$\|D\psi_x\| < .99$ for all $x \in \mathcal{S}_{a,b,c}$ with (a, b, c) as in* Table 2.

From here it is simple to deduce (2.12).

***Proof of* (2.12) *assuming Proposition* 2.9.** Let (a, b, c) be as in Table 2. The fact that ψ preserves $\mathcal{S}_{a,b,c}$ is the content of Proposition 2.8. Let $x, y \in \mathcal{S}_{a,b,c}$, and let T denote a bounded linear functional on the vector space spanned by $\mathcal{R}$, such that $\|T\| = 1$ (the operator norm is defined with respect to the norm given in (2.13)). Since ψ is differentiable, via the chain rule we have

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

that

$$D(T \circ \psi) = T \circ D\psi.$$

For $t \in [0,1]$, define

$$x(t) := ty + (1-t)x \qquad \text{and} \qquad r(t) := T \circ \psi(x(t)).$$

By the fundamental theorem of calculus,

$$\begin{aligned} T \circ \psi(y) - T \circ \psi(x) &= \int_0^1 r'(t)dt \\ &= \int_0^1 T \circ D\psi_{x(t)}\left(y - x\right) dt. \end{aligned}$$

Thus, using that $\|T\| = 1$,

$$\begin{aligned} |T \circ \psi(y) - T \circ \psi(x)| &\le \int_0^1 \|T \circ D\psi_{x(t)}\| \cdot \|y - x\| dt \\ &\le \sup_{t \in [0,1]} \|D\psi_{x(t)}\| \cdot \|y - x\|. \end{aligned}$$

Since $\mathcal{S}_{a,b,c}$ is convex, $x(t) \in \mathcal{S}_{a,b,c}$ for all $t \in [0,1]$, so we may replace the upper bound of the norm of the derivative on the path $x(t)$ with the upper bound of the same on $\mathcal{S}_{a,b,c}$. Finally, taking the supremum over all T such that $\|T\| = 1$, we may conclude that

$$\|\psi(x) - \psi(y)\| \le \sup_{y \in \mathcal{S}_{a,b,c}} \|D\psi_y\| \cdot \|x - y\|,$$

which combined with Proposition 2.9 gives that ψ is a contraction on $\mathcal{S}_{a,b,c}$. □

The rest of this section is devoted to the proof of Proposition 2.9. We begin by giving an upper bound for the operator norm of $D\psi_x$ in terms of the partial derivatives of ψ.

Lemma 2.10. *For $d = 2$, the operator norm of $D\psi_x$ is*

$$\begin{aligned} \|D\psi_x\| &= \sup_{i \ge 1} \sum_{j \ge 1} \left| \frac{\partial \psi_i}{\partial x_j} \right| \\ &= \max \left\{ \left| \frac{\partial \psi_1}{\partial x_1} \right| + \left| \frac{\partial \psi_1}{\partial x_2} \right|, \sup_{i \ge 2} \left(\left| \frac{\partial \psi_i}{\partial x_{i-1}} \right| + \left| \frac{\partial \psi_i}{\partial x_i} \right| + \left| \frac{\partial \psi_i}{\partial x_{i+1}} \right| \right) \right\}. \end{aligned}$$

For $d \ge 3$, the operator norm of $D\psi_x$ is

$$\begin{aligned} \|D\psi_x\| &\le \max \left\{ \left| \frac{\partial \psi_1}{\partial x_1} \right| + \sup_{i \ge 2} \left| \frac{\partial \psi_i}{\partial x_1} \right|, \sum_{j \ge 2} \left| \frac{\partial \psi_1}{\partial x_j} \right| + \sup_{i \ge 2} \sum_{j \ge 2} \left| \frac{\partial \psi_i}{\partial x_j} \right| \right\} \\ &= \max \left\{ \left| \frac{\partial \psi_1}{\partial x_1} \right| + \left| \frac{\partial \psi_2}{\partial x_1} \right|, \left| \frac{\partial \psi_1}{\partial x_2} \right| + \sup_{i \ge 2} \sum_{j=i-1}^{i+1} \left| \frac{\partial \psi_i}{\partial x_j} \right| \mathbf{1}_{j \ge 2} \right\}. \end{aligned}$$

Proof. We only prove the case of $d \ge 3$ (the case $d = 2$ is similar and more standard as the norm is simply the ℓ^∞-norm). Denote $T := D\psi_x$ and $t_{i,j} := \frac{\partial \psi_i}{\partial x_j}$, so that $(Ty)_i = \sum_j t_{i,j} y_j$. Denote $m_i := \sum_{j \ge 2} |t_{i,j}|$. Let y have norm 1. Write $a := |y_1|$ and $b := 1 - |a|$, and note that $|y_i| \le b$ for all $i \ge 2$. Then

$$|(Ty)_i| = \left| \sum_j t_{i,j} y_j \right| \le |t_{i,1} y_1| + \sum_{j \ge 2} |t_{i,j} y_j| \le a|t_{i,1}| + b m_i.$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Hence,

$$\|Ty\| = |(Ty)_1| + \sup_{i\geq 2} |(Ty)_i| \leq a|t_{1,1}| + bm_1 + \sup_{i\geq 2} (a|t_{i,1}| + bm_i)$$
$$\leq \max\left\{|t_{1,1}| + \sup_{i\geq 2} |t_{i,1}|,\ m_1 + \sup_{i\geq 2} m_i\right\},$$

as desired. □

We now record the expressions for the derivatives.

$$-\frac{\partial \psi_1}{\partial x_1} = \frac{d(1-x_2)(1+x_1+x_1x_2)^{d-1}}{(1+2x_1)^{d+1}}, \tag{2.14}$$

$$\frac{\partial \psi_1}{\partial x_2} = \frac{dx_1(1+x_1+x_1x_2)^{d-1}}{(1+2x_1)^{d}}, \tag{2.15}$$

and for $n \geq 2$,

$$\frac{\partial \psi_n}{\partial x_{n-1}} = \frac{dx_{n-1}^{d-1}(1+x_n+x_nx_{n+1})^{d}}{(1+x_{n-1}+x_{n-1}x_n)^{d+1}}, \tag{2.16}$$

$$\frac{\partial \psi_n}{\partial x_n} = \frac{dx_{n-1}^{d}(1+x_n+x_nx_{n+1})^{d-1}(1+x_{n-1}x_{n+1}+x_{n+1})}{(1+x_{n-1}+x_{n-1}x_n)^{d+1}}, \tag{2.17}$$

$$\frac{\partial \psi_n}{\partial x_{n+1}} = \frac{dx_{n-1}^{d}x_n(1+x_n+x_nx_{n+1})^{d-1}}{(1+x_{n-1}+x_{n-1}x_n)^{d}}. \tag{2.18}$$

The next two lemmas provide bounds on the partial derivatives of ψ. While we assume throughout that (a, b, c) is chosen according to Table 2, we use only minor information about the actual values, and keep bounds general where possible. We shall only plug in the exact values of these parameters at the very end. The simple properties we need are that $c < 1/2$ and that

$$(d-2)a \leq 1 \quad \text{for } 2 \leq d \leq 5, \qquad (d-2)a \geq 1 + dac \quad \text{for } d = 6, 7. \tag{2.19}$$

$$b((1-c)d - 2) \leq 1 \quad \text{for } d = 2, 3, 4. \tag{2.20}$$

$$1 + 2b(1+c) \leq d \quad \text{for } d \geq 3. \tag{2.21}$$

Lemma 2.11. *Let $x \in \mathcal{S}_{a,b,c}$ with (a, b, c) as in Table 2. Then*

$$\left|\frac{\partial \psi_1}{\partial x_1}\right| \leq \frac{d(1+a)^{d-1}}{(1+2a)^{d+1}} \qquad \text{for } 2 \leq d \leq 5. \tag{2.22}$$

$$\left|\frac{\partial \psi_1}{\partial x_1}\right| \leq \frac{d(1-c)(1+a+ac)^{d-1}}{(1+2a)^{d+1}} \qquad \text{for } d = 6, 7. \tag{2.23}$$

$$\left|\frac{\partial \psi_1}{\partial x_2}\right| \leq \frac{db(1+b+bc)^{d-1}}{(1+2b)^{d}} \qquad \text{for } 2 \leq d \leq 4. \tag{2.24}$$

$$\left|\frac{\partial \psi_1}{\partial x_2}\right| \leq \frac{(d-1)^{d-1}}{d^{d-1}(1-c)} \qquad \text{for } 5 \leq d \leq 7. \tag{2.25}$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Proof. Recall the expression for $|\partial\psi_1/\partial x_1|$ from (2.14). Since this expression is decreasing in x_1 (using (2.7)) we have

$$\left|\frac{\partial\psi_1}{\partial x_1}\right| \le \frac{d(1-x_2)(1+a+ax_2)^{d-1}}{(1+2a)^{d+1}} =: h(x_2). \tag{2.26}$$

The function $h(x)$ is maximized at

$$x_* := \frac{(d-2)a-1}{da}.$$

By (2.19), we have that $x_* < 0$ for $2 \le d \le 5$ and $x_* > c$ for $d = 6, 7$. Note that h is increasing on $(-\infty, x_*]$ and decreasing on $[x_*, \infty)$. Overall, h is decreasing on $[0, c]$ for $2 \le d \le 5$ and increasing on $[0, c]$ for $d = 6, 7$. Thus, $|\partial\psi_1/\partial x_1| \le h(0)$ for $2 \le d \le 5$ and $|\partial\psi_1/\partial x_1| \le h(c)$ for $d = 6, 7$. This establishes the claimed bound on $|\partial\psi_1/\partial x_1|$.

Next, recall the expression for $|\partial\psi_1/\partial x_2|$ from (2.15). Since this expression is increasing in x_2,

$$\left|\frac{\partial\psi_1}{\partial x_2}\right| \le \frac{dx_1(1+x_1+cx_1)^{d-1}}{(1+2x_1)^d} =: \tilde{h}(x_1).$$

Differentiating $\tilde{h}$, we get

$$\tilde{h}'(x) = \frac{d(1+x+cx)^{d-2}}{(1+2x)^{d+1}}[x(2-(1-c)d)+1].$$

If $d = 2$, then clearly $\tilde{h}'(x) \ge 0$ for $x \ge 0$, so that $\tilde{h}(x_1) \le \tilde{h}(b)$. For $d \ge 3$, we have a critical point at

$$\tilde{x}_* := \frac{1}{(1-c)d-2}.$$

Note that $2-(1-c)d < 0$ since $c < 1/2$. Thus, $\tilde{x}_*$ is positive and is in fact a maximum of $\tilde{h}$. Furthermore, $\tilde{x}_* > b$ for $d = 3, 4$ by (2.20). Thus, $\tilde{h}(x_1) \le \tilde{h}(b)$ for $d \le 4$, and $\tilde{h}(x_1) \le \tilde{h}(\tilde{x}_*)$ for $d \ge 5$. This establishes the claimed bound on $|\partial\psi_1/\partial x_2|$. □

Lemma 2.12. *Let $x \in \mathcal{S}_{a,b,c}$ with (a, b, c) as in Table 2. For $n \ge 2$,*

$$\left|\frac{\partial\psi_n}{\partial x_{n-1}}\right| \le \frac{4d(d-1)^{d-1}(1+c+c^2)^d}{(d+1)^{d+1}(1+c)^{d-1}} \quad \text{for } d = 2, \tag{2.27}$$

$$\left|\frac{\partial\psi_n}{\partial x_{n-1}}\right| \le \frac{db^{d-1}(1+c+c^2)^d}{(1+b)(1+b+bc)^d} \quad \text{for } 3 \le d \le 7, \tag{2.28}$$

$$\left|\frac{\partial\psi_n}{\partial x_n}\right| \le \left|\frac{\partial\psi_n}{\partial x_{n-1}}\right| \cdot b(1+c+bc). \tag{2.29}$$

$$\left|\frac{\partial\psi_n}{\partial x_{n+1}}\right| \le \frac{db^d c(1+c+c^2)^{d-1}}{(1+b+bc)^d}. \tag{2.30}$$

We remark that it is easy to show that any $x \in \psi(\mathcal{S}_{a,b,c})$ satisfies $x_n \le c^d$ for $n \ge 3$, and that one could use this extra information to improve the bounds given in the lemma. This did not seem, however, to simplify other parts of the proof, and so we did not find any reason to do this. Furthermore, the bounds we stated are uniform in $n \ge 2$ and governed by the case $n = 2$; the bounds can be significantly improved as n increases.

Proof. Set $\beta := 1 + x_n$, and for $x \ge 0$, define

$$h(x) := \frac{x^{d-1}}{(1+\beta x)^{d+1}},$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

so that

$$\left|\frac{\partial \psi_n}{\partial x_{n-1}}\right| = h(x_{n-1}) \cdot d(1 + x_n + x_n x_{n+1})^d.$$

Note that h is maximized at $x_* := \frac{d-1}{2\beta}$, and that it is increasing on $[0, x_*]$ and decreasing on $[x_*, \infty)$. For $d \geq 3$, we have that $x_{n-1} \leq b \leq x_*$ by (2.21), and hence, $h(x_{n-1}) \leq h(b)$ and

$$\left|\frac{\partial \psi_n}{\partial x_{n-1}}\right| \leq \frac{db^{d-1}(1 + x_n + x_n x_{n+1})^d}{(1 + b + bx_n)^{d+1}} \tag{2.31}$$

$$\leq \frac{db^{d-1}(1 + c + c^2)^d}{(1 + b)(1 + b + bc)^d}, \tag{2.32}$$

where we used $1 + b + bx_n \geq 1 + b$, $x_{n+1} \leq c$ and (2.7). For $d = 2$, we do not know whether $x_* \geq b$, so instead we just use that $h(x_{n-1}) \leq h(x_*)$ and plug in the value of x_* to get that

$$\left|\frac{\partial \psi_n}{\partial x_{n-1}}\right| \leq \frac{d\left(\frac{d-1}{2(1+x_n)}\right)^{d-1}(1 + x_n + x_n x_{n+1})^d}{\left(1 + \frac{d-1}{2}\right)^{d+1}} \tag{2.33}$$

$$= \frac{4d(d-1)^{d-1}(1 + x_n + x_n x_{n+1})^d}{(d+1)^{d+1}(1 + x_n)^{d-1}}$$

$$\leq \frac{4d(d-1)^{d-1}(1 + c + c^2)^d}{(d+1)^{d+1}(1 + c)^{d-1}}, \tag{2.34}$$

where we used $x_{n+1} \leq c$ and (2.7).

To upper bound $\partial \psi_n / \partial x_n$, note that

$$\left|\frac{\partial \psi_n}{\partial x_n}\right| = \left|\frac{\partial \psi_n}{\partial x_{n-1}}\right| \cdot \frac{x_{n-1}(x_{n-1}x_{n+1} + x_{n+1} + 1)}{x_n x_{n+1} + x_n + 1},$$

and then, for the second term on the right-hand side, lower bound the denominator by 1 and upper bound the numerator using $x_{n-1} \leq b$ and $x_{n+1} \leq c$. Finally,

$$\left|\frac{\partial \psi_n}{\partial x_{n+1}}\right| = \frac{dx_{n-1}^d x_n (1 + x_n + x_n x_{n+1})^{d-1}}{(1 + x_{n-1} + x_{n-1}x_n)^d} \tag{2.35}$$

$$\leq \frac{db^d x_n (1 + x_n + x_n x_{n+1})^{d-1}}{(1 + b + bx_n)^d} \tag{2.36}$$

$$\leq \frac{db^d c(1 + c + c^2)^{d-1}}{(1 + b + bc)^d}, \tag{2.37}$$

where in the first inequality we used that $|\partial \psi_n / \partial x_{n+1}|$ is increasing in x_{n-1} and that $x_{n-1} \leq b$, and in the second inequality we used $x_n, x_{n+1} \leq c$ and (2.7). □

***Proof of Proposition* 2.9.** Using Lemmas 2.11 and 2.12, we get upper bounds on the partial derivatives $|\partial \psi_i / \partial x_j|$, which when combined with Lemma 2.10 yields an upper bound on the operator norm $\|D\psi_x\|$ (see Table 4). □

***Proof of Theorem* 1.1.** The proof of the non-convergence part is in Section 2.2. The convergence part follows from the more general Theorem 2.6. □

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Table 4. Upper bounds on the partial derivatives and operator norm when $x \in \mathcal{S}_{a,b,c}$ with (a, b, c) as in Table 2. The bounds on the partial derivatives are obtained by plugging in the values of a, b, c into the bounds given by Lemmas 2.11 and 2.12 ($n \geq 2$ in the table) and rounding up. The bounds on the operator norm are then obtained by plugging in the former bounds into the bounds from Lemma 2.10

d	2	3	4	5	6	7
$\frac{\partial \psi_1}{\partial x_1}$	.38	.57	.84	.86	.97	.986
$\frac{\partial \psi_1}{\partial x_2}$	.46	.51	.47	.46	.45	.401
$\frac{\partial \psi_n}{\partial x_{n-1}}$	.43	.26	.09	.03	.01	.001
$\frac{\partial \psi_n}{\partial x_n}$	.43	.21	.06	.02	.01	.001
$\frac{\partial \psi_n}{\partial x_{n+1}}$	.10	.04	.01	.01	.01	.001
$\|D\psi_x\|$	.96	.83	.95	.89	.98	.987

3. Long-range order for graphs with large expansion

In this section, we prove Theorem 1.3. Recall that G is an infinite bipartite connect graph with maximum degree d and Cheeger constant h. We fix a finite subset G' of G. Recall that $\nu^{a,b}_{G',M}$ is the uniform measure on all M-Lipschitz functions on $G' \cup \partial G'$ whose values on $\partial G'$ are constrained to be in $\{a, \ldots, b\}$ for white vertices or in $\{b - M, \ldots, a + M\}$ for black vertices.

We now fix two integers a, b with $0 \leq b - a \leq 2M$. To facilitate the description of these Gibbs measures, we shall consider the following sets:

$$\mathcal{O}(f) := \{v \in G' \text{ is black}, f(v) \notin \{b - M, \ldots, a + M\}\},$$
$$\mathcal{E}(f) := \{v \in G' \text{ is white}, f(v) \notin \{a, \ldots, b\}\}.$$

and define $B(f) := \mathcal{O}(f) \cup \mathcal{E}(f)$. Let $A(f, v)$ denote the distance two connected component of $B(f)$ containing v. To elaborate: connect two vertices if they are at distance at most two, and let $A(f, v)$ denote the connected component of v thus formed. Our goal in the next proposition will be to prove that the size of $A(f, v)$ has exponential tail once h is large enough compared to M.

Proposition 3.1. *Assume that* $h + 1 \geq 2(2M + 1) \log (d^4(4M + 1))$, *and let* $f \sim \nu^{a,b}_{G',M}$. *Then for all* $v_0 \in G'$ *and* $N \geq 1$,

$$\mathbb{P}(|A(f, v_0)| = N) \leq \exp\left(-\frac{(h + 1)N}{2(2M + 1)}\right).$$

The upshot of Proposition 3.1 is that most of the black vertices take values in $\{b - M, \ldots, a + M\}$ and most of the white vertices take values in $\{a, \ldots, b\}$.

Let us illustrate some special cases. If $a = b = 0$ and $M = 1$ (1-Lipschitz case with 0 boundary conditions on even vertices), most of the white vertices take the value 0 and the black vertices take values in $\{-1, 0, 1\}$ (roughly uniformly at random as they are likely to be surrounded by 0s). Also, if $a = 0, b = 1$ (as in Theorem 1.2), then both white and black vertices are likely to take values in $\{0, 1\}$. In general, if $b - a = M$, the white and black vertices play a symmetric role in the measure (and in fact bipartiteness of G is not needed).

***Proof of Proposition* 3.1.** We begin with an elementary fact (see, e.g., [23, Lemma 1.8]). Let $\mathcal{L}$ denote the set of M-Lipschitz functions on G' satisfying the required boundary constraints. Let $\mathcal{E}$ be a subset of $\mathcal{L}$ and suppose that T is a map from $\mathcal{E}$ to subsets of $\mathcal{L}$. Also assume that there exist $s, t > 0$ such that $|T(g)| \geq t$ for all $g \in \mathcal{E}$, and $|\{g \in \mathcal{A} : g' \in T(g)\}| \leq s$ for all $g' \in \mathcal{L}$. Then

$$\mathbb{P}(f \in \mathcal{E}) = \frac{|\mathcal{E}|}{|\mathcal{L}|} \leq \frac{s}{t}.$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Throughout, we write $A(f)$ for $A(f, v_0)$ and observe that $A(f)$ cannot contain any vertex of $\partial G'$ by definition. Fix a distance two connected set A containing v_0 but not intersecting $\partial G'$, and let $\mathcal{E} := \{g \in \mathcal{L} : A(g) = A\}$. We shall construct a map T as described in the previous paragraph so that s/t is small, where s, t are as in the previous paragraph. Consequently,

$$\mathbb{P}(A(f) = A) = \mathbb{P}(f \in \mathcal{E}) \leq \frac{s}{t}.$$

Our goal is to make $\frac{s}{t}$ exponentially small in the size of A with a suitably large constant, so that we can then apply a union bound over all possible realizations of $A(f)$ and conclude.

Let us describe the map T. For $g \in \mathcal{E}$, let $T(g)$ be the set of $g' \in \mathcal{L}$ satisfying

$$g'(x) \begin{cases} = g(x) & \text{if } x \notin A \cup \partial A \\ \in \{a, \dots, b\} & \text{if } x \in A \cup \partial A, x \text{ white}, \\ \in \{b - M, \dots, a + M\} & \text{if } x \in A \cup \partial A, x \text{ black}. \end{cases} \tag{3.1}$$

The point of this definition is that, as we will see, it allows much more possibilities in $A \cup \partial A$ for $g' \in T(g)$ compared to g. Specifically, we claim T satisfies the requirement with

$$s := (4M + 1)^{|A|}(b - a)^{|\{v \in \partial A : v \text{ white}\}|}(a - b + 2M)^{|\{v \in \partial A : v \text{ black}\}|}, \tag{3.2}$$

and

$$t := (b - a + 1)^{|\{v \in A \cup \partial A : v \text{ white}\}|}(a - b + 2M + 1)^{|\{v \in A \cup \partial A : v \text{ black}\}|}. \tag{3.3}$$

Before proving this claim, let us show how it yields the proposition. We have

$$\frac{s}{t} \leq (4M + 1)^{|A|} \left(\frac{b - a}{b - a + 1}\right)^{|\{v \in A \cup \partial A, v \text{ white}\}|} \left(\frac{a - b + 2M}{a - b + 2M + 1}\right)^{|\{v \in A \cup \partial A, v \text{ black}\}|}$$

Since $x \mapsto \frac{x}{x+1}$ is increasing and $0 \leq b - a \leq 2M$, and using the fact that $|\partial A| \geq h|A|$ (since recall that A cannot contain any vertex in the boundary),

$$\mathbb{P}(A(f) = A) \leq \frac{s}{t} \leq (4M + 1)^{|A|} \left(\frac{2M}{2M + 1}\right)^{|A| + |\partial A|} \leq e^{|A| \log (4M+1) - \frac{(h+1)|A|}{2M+1}} \tag{3.4}$$

Finally, the number of distance two connected sets A with $|A| = N$ and $v_0 \in A$ is at most d^{4N}. Union bounding, we get

$$\mathbb{P}(|A(f)| = N) \leq e^{N(4 \log d + \log (4M+1) - \frac{h+1}{2M+1})} \leq \exp\left(-\frac{(h + 1)N}{2(2M + 1)}\right),$$

as desired.

It remains to prove that T satisfies the required upper and lower bounds with (3.2) and (3.3). We begin with the lower bound. Let $g \in \mathcal{E}$. Since $|x - y| \leq M$ for $x \in \{a, \dots, b\}$ and $y \in \{b - M, \dots, a + M\}$, and since $g(v) \in \{a, \dots, b\}$ for any white $v \in \partial(A \cup \partial A)$ and $g(v) \in \{b - M, \dots, a + M\}$ for any black $v \in \partial(A \cup \partial A)$, we see that every function satisfying (3.1) must necessarily belong to $\mathcal{L}$. Hence,

$$|T(g)| = (b - a + 1)^{|\{v \in A \cup \partial A : v \text{ white}\}|}(a - b + 2M + 1)^{|\{v \in A \cup \partial A : v \text{ black}\}|} = t.$$

We now move on to the upper bound. Fix $g' \in \mathcal{L}$. We need to upper bound the number of $g \in \mathcal{E}$ such that $g' \in T(g)$. Clearly, g is uniquely determined outside $A \cup \partial A$. Note that if a vertex u in A is at distance 2 from the complement of $A \cup \partial A$, then the number of possible values of $g(u)$ is at most $4M + 1$. By choosing the values of g on A one after another, in any order consistent with their distance from the complement of $A \cup \partial A$, we see that there are at most $(4M + 1)^{|A|}$ possible choices for g on A. To show that $|\{g \in \mathcal{E} : g' \in T(g)\}| \leq s$, it remains to show that, given g outside

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

∂A, the number of choices for g on ∂A is at most $\prod_{v\in\partial A}\ell_v$, where

$$l_v := \begin{cases} b-a & \text{if } v \text{ is white} \\ 2M+a-b & \text{if } v \text{ is black.} \end{cases}$$

For this, it suffices to prove a vertex-by-vertex bound, namely, that the number of choices for $g(v)$ is at most ℓ_v for any $v\in\partial A$. To this end, fix $v\in\partial A$, and let $u\in A$ be any neighbour of v (which exists since $v\in\partial A$). Suppose first that v is white. First note that since $v\in\partial A$ and v is white, $g(v)\in\{a,\dots,b\}$. Similarly, since $u\in A$ and u is black, $g(u)\notin\{b-M,\dots,a+M\}$. Consequently,

$$\begin{aligned} a+1\le g(u)-M\le g(v)\le b, &\qquad \text{if } g(u)\ge a+M+1,\\ a\le g(v)\le g(u)+M\le b-1, &\qquad \text{if } g(u)\le b-M-1. \end{aligned}$$

(One consequence of the above inequalities is that v cannot be white if $a=b$.) Therefore, the number of possible values of $g(v)$ is at most ℓ_v, as claimed. Suppose next that v is black. We similarly have that $g(v)\in\{b-M,\dots,a+M\}$ and $g(u)\notin\{a,\dots,b\}$, so that

$$\begin{aligned} b-M+1\le g(u)-M\le g(v)\le a+M &\qquad \text{if } g(u)\ge b+1,\\ b-M\le g(v)\le g(u)+M\le a+M-1 &\qquad \text{if } g(u)\le a-1. \end{aligned}$$

(Again, v cannot be black if $b-a=2M$.) Therefore, the number of possible values of $g(v)$ is again at most ℓ_v, as claimed. □

The following corollary is a simple generalization of (3.4) and is what we will actually need in the proof of Theorem 1.3.

Corollary 3.2. *Let $A_1,A_2,\dots,A_k$ be distance two disjoint, distance two connected sets of total size $N:=\sum_{i=1}^k|A_i|$. Then the probability that $A_1,\dots,A_k$ are distance two connected components of $B(f)$ is at most*

$$\exp\left(N\left(\log(4M+1)-\frac{h+1}{2M+1}\right)\right).$$

Proof of Theorem 1.3. Let (G_k) be an exhaustion. Note that it is enough to prove the following: for all $v\in V$, $R>0$ and for all connected H,H' such that $B(v,R)\subset B(v,R')\subset H\subset H'\subset G$, $\nu^{a,b}_{H,M}$ can be coupled with $\nu^{a,b}_{H',M}$ so that they agree on $B(v,R)$ with probability tending to 1 as $R'\to\infty$.

To construct the coupling, let $f\sim\nu^{a,b}_{H,M}$ and $f'\sim\nu^{a,b}_{H',M}$ be sampled independently, and let $B(f)$ and $B(f')$ be the set of 'atypical' vertices as defined before Proposition 3.1. Let $U=B(v,R)$, and let $A(f,f',U)$ be the union of the distance two connected components of $B(f)\cup B(f')$ intersecting U. Then for any A, using Corollary 3.2 (and $\frac{h+1}{2M+1}\ge\frac{h}{3M}$),

$$\begin{aligned} \mathbb{P}(A(f,f',U)=A) &\le \sum_{A_1\cup A_2=A}\exp\left((|A_1|+|A_2|)\left(\log(4M+1)-\frac{h}{3M}\right)\right)\\ &\le 3^{|A|}\exp\left(|A|\left(\log(4M+1)-\frac{h}{3M}\right)\right). \end{aligned}$$

Here, A_1 and A_2 are $A(f,f',U)\cap B(f)$ and $A(f,f',U)\cap B(f')$, respectively. Since there are at most $(2N)^{|U|}d^{4N}$ possible choices for such an A of a given size N (see Lemma 3.3 below),

$$\begin{aligned} \mathbb{P}(|A(f,f',U)|=N) &\le \exp\left(N\left(\log(3(4M+1))-\frac{h}{3M}+4\log d\right)+|U|\log(2N)\right)\\ &\le \exp\left(-\frac{hN}{6M}+|U|\log(2N)\right). \end{aligned}$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Since U is fixed, this bound decays exponentially in N. In particular, as $R' \to \infty$,

$$\mathbb{P}(A(f,f',U) \subset B(v,R')) \to 1.$$

It remains to show that f and f' can be coupled so that they agree on $B(v,R)$ on the event $\{(A(f,f',U) \subset B(v,R')\}$. We proceed via exploration from the boundary. Let X be the union of the distance two connected components of $B(f) \cup B(f') \cup (G \setminus H)$ containing $G \setminus H$. Note that, by definition of $A(f,f',U)$, we have that X is disjoint from $B(v,R+2)$ whenever $A(f,f',U) \subset B(v,R'-2)$. Denote $\tilde{X} := X \cup \partial X$, $\bar{X} := \tilde{X} \cup \partial \tilde{X}$ and $Y := G \setminus \bar{X}$. Note that X is measurable with respect to $(B(f) \cup B(f')) \cap \bar{X}$. Now condition on X and on $(f|_{\bar{X}}, f'|_{\bar{X}})$ and suppose that X is disjoint from $B(v,R+2)$. We claim that $f|_Y$ and $f'|_Y$ are conditionally distributed according to $\nu^{a,b}_{Y,M}$ (in fact, they are also conditionally independent, but we do not need this), and in particular, $f|_Y$ and $f'|_Y$ can be coupled to agree on Y (which contains $B(v,R)$). We prove this for $f|_Y$ (the proof for f'_Y is identical). Indeed, every even $v \in \partial Y$ has $f(v) \in \{a,\dots,b\}$, since $\partial Y \subset \partial \tilde{X} \subset (B(f) \cup B(f'))^c$. On the other hand, any value in $\{a,\dots,b\}$ is possible for $f(v)$, since any $u \in N(v) \setminus Y \subset \bar{X} \setminus X \subset (B(f) \cup B(f'))^c$, is odd, and satisfies $f(u) \in \{b-M,\dots,a+M\}$. Similarly, any odd vertex in ∂Y must take a values in $\{b-M,\dots,a+M\}$ and is free to take any value there. Since there are no other constraints (beyond the obvious Lipschitz constraint), the conditional law of $f|_Y$ is $\nu^{a,b}_{Y,M}$. This completes the proof of convergence.

Now we show part (b) of the theorem, which is an easy consequence of Proposition 3.1. Let $f \sim \nu^{a,b}_{G,M}$. Fix an odd vertex u and denote $\mathcal{N} := \{f(v) : v \in N(u)\}$. Let $\mathcal{N}^-$ and $\mathcal{N}^+$ be the minimum and maximum elements in $\mathcal{N}$, respectively. Since $f(u)$ is conditionally uniformly distributed in $\{\mathcal{N}^+ - M,\dots,\mathcal{N}^- + M\}$ given $\mathcal{N}$, we see that

$$\begin{aligned}\mathbb{P}(\{\mathcal{N}^-,\dots,\mathcal{N}^+\} \subsetneq \{a,\dots,b\}) &\le (2M+1)\cdot \mathbb{P}(f(u) \notin \{b-M,\dots,a+M\}) \\ &\le (2M+1)\cdot \mathbb{P}(|A(f,u)| \ge 1)) \\ &\le \frac{2M+1}{e^{h/6M}-1}.\end{aligned}$$

On the other hand,

$$\begin{aligned}\mathbb{P}(\{\mathcal{N}^-,\dots,\mathcal{N}^+\} \not\subset \{a,\dots,b\}) &= \mathbb{P}(\mathcal{N}^- < a \text{ or } \mathcal{N}^+ > b) \\ &= \mathbb{P}(f(v) \notin \{a,\dots,b\} \text{ for some } v \in N(u)) \\ &\le \sum_{v \in N(u)} \mathbb{P}(|A(f,v)| \ge 1)) \\ &\le \frac{d}{e^{h/6M}-1}.\end{aligned}$$

Together, we get that

$$\mathbb{P}(\{\mathcal{N}^-,\dots,\mathcal{N}^+\} \ne \{a,\dots,b\}) \le \frac{d+2M+1}{e^{h/6M}-1} \le \frac{d+2M+1}{3d^4(4M+1)-1} < \frac{1}{2},$$

which means that $\{a,\dots,b\}$ is the largest atom of $\{\mathcal{N}^-,\dots,\mathcal{N}^+\}$, so that it is uniquely determined by $\nu^{a,b}_{G,M}$. □

Lemma 3.3. *Let $\mathcal{G}$ be a graph of maximum degree at most Δ. Let $U \subset V(\mathcal{G})$. The number of sets $A \subset V(\mathcal{G})$ of a given size N such that each connected component of A intersects U is at most $(2N)^{|U|}\Delta^{2N}$.*

Proof. Let $U = \{u_1,\dots,u_m\}$ be some enumeration of U. Given a set $B \subset V(\mathcal{G})$ that intersects U, denote $I(B) := \min\{1 \le i \le m : u_i \in B\}$. Given a set A as in the lemma, define $(N_1,\dots,N_m)$ by letting $N_i = |C|$ if there exists a connected component C of A such that $I(C) = i$, and $N_i = 0$ otherwise. Note that $N_1 + \cdots + N_m = N$. The number of possibilities for $(N_1,\dots,N_m)$ is at most

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

$(N+1)^m \le (2N)^m$, and for any such choice, there are at most Δ^{2N_i} possibilities for C_i (since the number of connected sets of size N_i containing u_i is at most Δ^{2N_i}), and hence at most Δ^{2N} possibilities for A. □

4. FKG

In this section, we prove Theorem 1.2. We establish certain monotonicity condition for uniform Lipschitz functions and the absolute value of them. Such results have also been obtained independently by Karrila [15].

Let $G=(V,E)$ be a finite graph, and let ∂ be a subset of its vertices which we call the *boundary* of G. We will work with a slightly more general boundary condition in this section, where the constraints can vary over the vertices of ∂. A *boundary condition* is an assignment $\{\kappa_v\}_{v\in\partial}$ where $\kappa_v \subset \mathbb{Z}$. Let $\mathcal{L}(G,\partial,\kappa)$ be the set of Lipschitz functions f on G having $f(v)\in\kappa_v$ for all $v\in\partial$. We will only consider cases where $|\mathcal{L}(G,\partial,\kappa)|\in(0,\infty)$, and we will call such boundary conditions *admissible*. For an admissible κ, let $\nu_G^{\partial,\kappa}$ denote the law of a uniform element drawn from $\mathcal{L}(G,\partial,\kappa)$.

A function $f:\mathbb{Z}^V\mapsto\mathbb{R}$ is increasing if for any h,h' such that $h_v\le h'_v$ for all $v\in V, f(h)\le f(h')$. We denote by $\nu(f)$ the expectation of f under the law ν.

Proposition 4.1. *(FKG for h) Take a finite graph $G=(V,E)$ with boundary ∂. Suppose κ,κ' are two admissible boundary conditions such that for every $v\in\partial$, $\kappa_v=[a_v,b_v]$ and $\kappa'_v=[a'_v,b'_v]$ and $a_v\le a'_v$, $b_v\le b'_v$ (these integers could be $\pm\infty$). Then*

(CBC) *For every increasing function f, $\nu_G^{\partial,\kappa'}(f)\ge\nu_G^{\partial,\kappa}(f)$.*

(FKG) *For two increasing functions f,g, $\nu_G^{\partial,\kappa}(fg)\ge\nu_G^{\partial,\kappa}(f)\nu_G^{\partial,\kappa}(g)$.*

Proof. Both claims follow from Holley's criterion ([11, Theorem 2.3]). Fix a vertex v and condition on $\{\xi_v : v\in V\setminus\{v\}\}$ in $\mathcal{L}(G,\partial,\kappa)$. Now note that $\max\{|\xi_u-\xi_v| : u\sim v\}$ is in $\{0,1,2\}$. If it is 2, then the conditional value at v is deterministic. If it is 1, then the height at v takes two values with equal probability. Finally, if it is 0, then the height at v takes three values with equal probability. From this observation, it is easy to see that conditioning on $\{\xi'_v : v\in V\setminus\{v\}\}$ with $\xi'\ge\xi$ stochastically increases the height at h. □

A useful tool is monotonicity of the pushforward of the law of $\mathcal{L}(G,\partial,\kappa)$ under the mapping $h\mapsto|h|$. Unfortunately it is not true directly as can be seen in the segment graph with three vertices with two vertices having degree one and one having degree 2 (call it v). If the the height is 0 on the neighbours of v, then the conditional law of the height at v is Bernoulli (2/3). On the other hand, conditioning one of absolute value heights of the neighbours to be 1 makes the conditional law at v Bernoulli (1/2).

It turns out that a different form of absolute value FKG holds, if we do the mapping $h\mapsto|h+0.5|$. Let $\mathcal{L}^{>0}(G,\partial,\kappa)$ denote the set of 1-Lipschitz functions taking values in $\mathbb{N}-0.5:=\{0.5,1.5,\ldots\}$. We also need to be careful with the boundary conditions. We say an admissible boundary condition κ is $|h|$-adapted[3] if we can break up ∂ into $\partial=\partial_{\text{pos}}\sqcup\partial_{\text{sym}}$ such that

- $\kappa_v=-\kappa_v$ for every $v\in\partial_{\text{sym}}$,
- $\kappa_v\subseteq\mathbb{N}+0.5:=\{1.5,2.5,\ldots\}$ for every $v\in\partial_{\text{pos}}$.

For every $\xi\in\mathcal{L}^{>0}(G,\partial,\kappa)$, connect two vertices u,v by an edge if $\max\{\xi_u,\xi_v\}>0.5$. Let $k(\xi)$ denote the number of connected components which do not intersect B_{pos}. Then clearly for any

[3] Although we shift by 0.5, we keep this notation, which hopefully is not a source of confusion.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

$|h|$-adapted boundary condition κ,

$$\nu_G^{\partial,\kappa}(|h+0.5| = \xi) = \frac{2^{k(\xi)}}{|\mathcal{L}(G,\partial,\kappa)|}, \qquad \xi \in \mathcal{L}^{>0}(G,\partial,\kappa). \tag{4.1}$$

Proposition 4.2 (monotonicity for $|h+0.5|$). *Take a finite graph $G=(V,E)$ with boundary ∂. Let κ, κ' be two $|h|$-adapted boundary conditions with $\partial_{\text{pos}}(\kappa) \subseteq \partial_{\text{pos}}(\kappa')$ and for every $v \in \partial$, $[a_v, b_v] := \kappa_v \cap \{0.5, 1.5, \ldots\}$ and $[a'_v, b'_v] := \kappa'_v \cap \{0.5, 1.5, \ldots\}$ satisfy $a_v \le a'_v$ and $b_v \le b'_v$ (these numbers could be ∞),*

(CBC) *For every increasing function f, $\nu_G^{\partial,\kappa'}[f(|h+0.5|)] \ge \nu_G^{\partial,\kappa}[f(|h+0.5|)]$;*

(FKG) *For any two increasing functions f, g, $\nu_G^{\partial,\kappa}[f(|h+0.5|)g(|h+0.5|)] \ge \nu_G^{\partial,\kappa}[f(|h+0.5|)]\nu_G^{\partial,\kappa}[g(|h+0.5|)]$.*

Proof. The proof is very similar to [7, Proposition 2.2].

Fix a vertex v. Using [10, Theorem 4.8] and references therein, it is sufficient to prove that for any ξ (resp. η) which are restrictions to $V \setminus \{v\}$ of configurations in $\mathcal{L}^{>0}(G,\partial,\kappa)$ (resp. $\mathcal{L}^{>0}(G,\partial,\kappa')$) such that $\xi_v \le \eta_v$ for all v, and every $k \ge 0$,

$$\nu_G^{\partial,\kappa}\big[|h_v+0.5| \ge k \big| |h_{|V\setminus\{v\}}+0.5| = \xi\big] \le \nu_G^{\partial,\kappa'}\big[|h_v+0.5| \ge k \big| |h+0.5|_{|V\setminus\{v\}} = \eta\big]. \tag{4.2}$$

If $\xi_u \ge 2.5$ for some neighbour u of v, or if $\xi_u = 1.5$ for all neighbours u of v, then it is straightforward to see using (4.1) that (4.2) holds, as in this case we are simply reduced to the analysis in Proposition 4.1 (the quantity $k(\xi)$ in (4.1) does not change whatever be the value of $|h_v+0.5|$, consequently $|h_v+0.5|$ is equally likely among all possibilities dictated by the Lipschitz constraint or the constraint imposed by κ, κ' in case v is a boundary vertex).

Now suppose $\xi_u \le 1.5$ for every neighbour of v, and $\xi = 0.5$ for at least one of the neighbours of v (whence $|h_v+0.5| \in \{0.5, 1.5\}$). Notice that all the neighbours u with $\xi_u = 1.5$ must belong to the same component in ξ, call it $N_v(\xi)$. If $v \in \partial$ and $\{0.5, 1.5\} \not\subseteq \kappa_v$ then the value of $|h+0.5|$ at v is deterministically equal to $\kappa_v \cap \{0.5, 1.5\}$. Otherwise, let $k'_v(\xi)$ denote the number of components of $k(\xi)$ containing at least one of the neighbours u of v with $\xi_u = 0.5$ and which do not intersect $\partial_{\text{pos}}(\kappa)$. If $N_v(\xi)$ does not intersect $\partial_{\text{pos}}(\kappa)$, let $k_v(\xi) = k'_v(\xi) + 1$, otherwise $k_v(\xi) = k'_v(\xi)$. Let $\mathcal{B}(\kappa)$ denote the event that at least one of the components of $k(\xi)$ containing at least one of the neighbours of v intersects $\partial_{\text{pos}}(\kappa)$.

$$\nu_G^{\partial,\kappa}\big[|h_v+0.5| = 0.5 \big| |h+0.5|_{|V\setminus\{v\}} = \xi\big] = \begin{cases} 0 \text{ if } v \in \partial_{\text{pos}} \\ 2^{k_v(\xi)}/(2+2^{k_v(\xi)}) \text{ if } \xi \notin \mathcal{B}(\kappa),\ v \notin \partial_{\text{pos}} \\ 2^{k_v(\xi)}/(1+2^{k_v(\xi)}) \text{ if } \xi \in \mathcal{B}(\kappa),\ v \notin \partial_{\text{pos}} \end{cases}$$

The same formula holds for η as well. Now observe that $k_v(\xi) \ge k_v(\eta)$ since $\eta_v \ge \xi_v$ for all v and $\partial_{\text{pos}}(\kappa) \subseteq \partial_{\text{pos}}(\kappa')$. If there is strict inequality $k_v(\xi) > k_v(\eta)$ then the above expression for ξ is at least that for η in all the cases by monotonicity of $x \mapsto x/(1+x)$ and $x \mapsto x/(2+x)$ and the fact that $2^k/(2+2^k) \ge 2^{k'}/(1+2^{k'})$ if $k \ge k'+1$. So assume $k_v(\xi) = k_v(\eta)$. If ξ, η are both in $\mathcal{B}(\kappa)$ or both not in $\mathcal{B}(\kappa)$, we are also done for similar reasons. The only case remaining is if $\eta \in \mathcal{B}(\kappa')$ but $\xi \notin \mathcal{B}(\kappa)$ (recall $\mathcal{B}(\kappa) \subseteq \mathcal{B}(\kappa')$) and $k_v(\xi) = k_v(\eta)$. We claim that this is impossible, as in this case the strict inequality $k_v(\xi) > k_v(\eta)$ must hold. Indeed, in this case there must be a neighbour of v whose cluster intersects $B_{\text{pos}}(\kappa')$ and is not counted in $k_v(\eta)$, but every neighbour cluster is counted in ξ as none of them intersect $B_{\text{pos}}(\kappa)$ on the event $\mathcal{B}(\kappa)$. □

Corollary 4.3. *Let G be an infinite graph with a fixed vertex ρ. Let $(G_k)_{k\ge1}$ be a sequence such that $G_k \subseteq G_{k+1}$ and $\cup G_k = G$. Suppose ∂_k be a subset of vertices of G_k such that $\partial_{k+1} \cap G_k \subseteq \partial_k$, and*

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

the distance between ∂_k and ρ converges to ∞. Let $\kappa_v^k = \{-0.5, 0.5\}$ for all $v \in \partial_k$. Then one of the following two cases hold.

(**Localization**) $\nu_{\partial_k,\kappa^k}^{G_k}$ *converges to a translation-invariant Gibbs measure supported on Lipschitz functions on G.*

(**Delocalization**) $\lim_{m\to\infty} \sup_k \mu_{\partial_k,\kappa^k}^{G_k}(|h(\rho)| > m) > 0$.

Proof. Let $h^k \sim \mu_{\partial_k,\kappa^k}^{G_k}$. Observe that $|h^n + 0.5|$ is stochastically increasing in k, and hence if the delocalization condition does not hold, $\{|h^k + 0.5|\}_{k\geq 1}$ is tight. Consequently, weak limit of $|h^k + 0.5|$ exists as a probability measure on $\mathbb{Z}^V$. Furthermore, observe that for each k, the height function $h^k + 0.5$ is obtained by assigning a uniformly random sign to each connected component of ξ as described in Equation (4.1). Thus, this rule to obtain the height function from the absolute value also persists in the weak limit, and the proof of the corollary is complete. □

Proof of Theorem 1.2. Using Proposition 4.1, $\mu_{k,d}^{\{0,1\}}$ is dominated by $\mu_{k,d}^1$ and dominates $\mu_{k,d}^0$, both of which are tight by Theorem 1.1. Thus $\mu_{k,d}^{\{0,1\}}$ is tight. Thus $\mu_{k,d}^{\{0,1\}}$ is in the localization phase of Corollary 4.3, and the result follows. □

5. Discussion and open questions

In this section, we briefly discuss some properties of the limiting measures μ_d^S for finite subsets $S \subset \mathbb{Z}$. Let us first show how to construct translation-invariant Gibbs measures on d-regular trees rather than d-ary trees, the nice thing about the former is that they are vertex transitive. The convergence result for the d-regular tree actually follows from that of the d-ary tree in a straightforward manner: Let $f_k \sim \mu_{d,k}^S$, and let $\tilde{f}_k \sim \tilde{\mu}_{d,k}^S$ denote the analogous law for the $d+1$-regular tree. Then

$$\mathbb{P}(\tilde{f}_k(\rho) = i) = F_{d+1}(F_d^{(k-1)}(\delta_S))_i$$

where F_d is the map F for d-ary tree as described in the beginning of Section 2. Thus the limit as $k \to \infty$ for the $d+1$-regular tree exists whenever it exists for the d-ary tree. Furthermore if $\pi_d^S := \pi_d^S(\rho)$ denote the law of the height at ρ under μ_d^S, whenever it exists, then the corresponding law for the $d+1$-regular tree is given by $F_{d+1}(\pi_d^S)$. It is also easy to check that the transition matrix for the tree indexed Markov chain for the d-ary tree as well as the $d+1$ regular tree is given by

$$(P_d^S)_{i,j} = \frac{\pi_d^S(j)}{\pi_d^S(i-1) + \pi_d^S(i) + \pi_d^S(i+1)} 1_{|j-i|\leq 1}.$$

and that $F_{d+1}(\pi_d^S)$ is the stationary measure for this transition matrix. Using reversibility of the chain, translation invariance of the limit for the $d+1$-regular tree is immediate.

We now speculate about the most general result possible for general boundary conditions in our setup. We saw that for small d convergence holds for all the cases we considered here. For large d, however, such convergence holds along the even sequences. This leads us to conjecture:

Conjecture 5.1. $\lim_{k\to\infty} \mu_{k,2}^S$ *exists for all finite S. Also,* $\lim_{k\to\infty} \mu_{2k,d}^S$ *exists for all finite S and for all* $d \geq 2$.

We believe that for $S = \{-N, \ldots, N\}$ and $d \geq 8$, a more careful analysis similar to that in Section 6 should be enough to conclude the convergence of the map $\psi \circ \psi$.

Next comes the question of understanding all possible translation-invariant measures we can get as limits of the above type. To summarize the results of this article, we constructed the translation-invariant Gibbs measures $\tilde{\mu}_d^S$, $S = \{-N, \ldots, N\}$ for $2 \leq d \leq 7$ and for $S = \{0, 1\}$, $d \geq 2$.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

We also proved that for $2 \le d \le 7$, $\tilde{\mu}_d^{\{0,1\}} \neq \tilde{\mu}_d^0 = \tilde{\mu}_d^{\{-N,\dots,N\}}$ for all $N \ge 0$. This raised the following research programme.

Question 5.2. *Characterize all translation-invariant measures of the form $\tilde{\mu}_d^S$ where $S \subset \mathbb{Z}$ is finite, whenever they exist. In particular, is there a natural set of properties such that if S_1, S_2 satisfies all of the properties, then it is guaranteed that $\tilde{\mu}_d^{S_1} = \tilde{\mu}_d^{S_2}$?*

Another interesting direction is to consider M-Lipschitz functions for $M > 1$.

Question 5.3. *Extend Theorem 1.2 to $M > 1$: Does $\mu_{k,d,M}^{\{0,\dots,M\}}$ converge for all $d \ge 2$ and $M \ge 2$?*

Question 5.4. *Extend Theorem 1.1 to $M > 1$: Given $M > 1$, show that $\mu_{k,d,M}^0$ converges if and only if $d \le d_c(M)$. Can one at least show convergence for $M \gg d$?*

Another very interesting avenue of research is extremality. The measures $\tilde{\mu}_d^0$ and $\tilde{\mu}_d^{\{0,1\}}$ are both mixing, hence ergodic. However, the question of whether they are extremal remains:

Question 5.5. *For what values of d are the measures $\mu_d^{\{0,1\}}$ and μ_d^0 extremal for $2 \le d \le 7$?*

This is related to the question of extremality of the Ising model, see [19, Section 3].

6. Estimates of fixed point of φ

In this section, we complete the proof of Proposition 2.8. Recall from Lemma 2.2 that

$$\psi(\mathcal{S}_{a,b,c}) \subset \mathcal{S}_{a',b',c'},$$

where

$$(a', b', c') = \varphi(a, b, c) := (f(1, b), f(1 + c, a), g(b, c)),$$

and

$$f(\alpha, x) := \left(\frac{1 + \alpha x}{1 + 2x}\right)^d,$$
$$g(b, x) := b^d \left(\frac{1 + x + x^2}{1 + b + bx}\right)^d.$$

Let us summarize the basic strategy to approximate the fixed point of φ.

- First we prove in Lemma 6.1 that $x \mapsto g(b, x)$ has a unique fixed point in $(0, 1)$.
- Since $\mathcal{S}_{\varphi(0,1,c)} \subset \mathcal{S}_{0,1,g(1,c)}$, we can iterate this bound many times to reach close to $\mathcal{S}_{0,1,c_g}$ where c_g is the unique fixed point of $x \mapsto g(1, x)$ in $(0, 1)$.
- Next we note that

$$\psi \circ \psi(\mathcal{S}_{a,b,c_g})) \subset \mathcal{S}_{\mathrm{i}(c_g,a),\mathrm{j}(c_g,a),c_g}$$

  where i, j are defined as in (6.1). The reason for this approximation is that now it is enough to analyse the one dimensional maps i, j (with c_g fixed).
- We prove in Lemma 6.3 next that i, j have derivatives in $(0, 1)$ which (along with some other basic properties) is enough to prove that each of them has a unique fixed point in $(0, 1)$. We call them a_{c_g}, b_{c_g}. In conclusion, after enough iterates we approximately reach $\mathcal{S}_{a_{c_g},b_{c_g},c_g}$.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

- It turns out that a_{c_g}, b_{c_g}, c_g can be taken to be our estimates $\hat{a}_*, \hat{b}_*, \hat{c}_*$ for $3 \le d \le 6$. Unfortunately for $d = 2, 7$, these estimates are not enough. For these two cases we need another 'round' of approximation as follows.
- In Lemma 6.5, we prove using similar ideas that starting close to $\mathcal{S}_{a_{c_g},b_{c_g},c_g}$, we can reach $\mathcal{S}_{a_{c_g},b_{c_g},\tilde{c}_g}$ where $\tilde{c}_g$ is the unique fixed point of $x \mapsto g(b'_g, x)$ where $b'_g = \max(f(1 + c_g, a_{c_g}), b_g)$.
- Finally, in Lemma 6.6, we prove that starting close to $\mathcal{S}_{a_{c_g},b_{c_g},\tilde{c}_g}$ we reach $\mathcal{S}_{a_{\tilde{c}_g},b_{\tilde{c}_g},\tilde{c}_g}$ where $a_{\tilde{c}_g}, b_{\tilde{c}_g}$ are unique fixed points of $x \mapsto \mathrm{i}(x, \tilde{c}_g), x \mapsto \mathrm{j}(x, \tilde{c}_g)$.
- Now we need to explain how we get hold of the values close enough to the fixed points $c_g, a_{c_g}, b_{c_g}, \tilde{c}_g, a_{\tilde{c}_g}, b_{\tilde{c}_g}$. We prove in Lemma 6.2 that the map $x > g(b, x)$ if and only if $x < c_g$ where c_g is the fixed point of $x \mapsto g(b, x)$. Thus if we can produce two points c_1, c_2 such that $g(b, c_1) > c_1$ and $g(b, c_2) < c_2$ then $c_1 < c_g < c_2$. In practice, we can 'guess' good estimates of c_1, c_2 by looking at the plots and then plug them into $g(b, x)$ to confirm that indeed $g(b, c_1) > c_1$, $g(b, c_2) < c_2$, which makes the estimates completely rigorous. The maps $x \mapsto \mathrm{i}(c, x), x \mapsto \mathrm{j}(c, x)$ also satisfy similar properties which we exploit in a similar manner to get good estimates of a_{c_g}, b_{c_g} and then $a_{\tilde{c}_g}, b_{\tilde{c}_g}$. This part is done in a separate subsection Section 6.1.

We now explain the details of the above strategy.

Lemma 6.1. *For $x > 0, b \in (0, 1]$, $x \mapsto g(b, x)$ is increasing, convex and has a unique fixed point c_0 in $(0, 1)$. Moreover, if $0 \le c < 1$ then for any fixed b, defining $g_b(c) = g(b, c)$, $g_b^{(i)}(c)$ converges to $c_0 = c_0(b)$. Also, $g_b(x) > x$ if and only if $x < c_0(b)$.*

Proof. We have

$$\frac{\partial g}{\partial x} = \frac{db^d(bx^2 + 2bx + 2x + 1)(x^2 + x + 1)^{d-1}}{(bx + b + 1)^{d+1}}$$

and

$$\begin{aligned}\frac{\partial^2 g}{\partial x^2} &= db^d(x^2 + x + 1)^{d-2}(bx + b + 1)^{-d-2} \\ &(b^2((d-1)x^4 + 4(d-1)x^3 + (4d-2)x^2 + 2x + 2) + 2b(2(d-1)x^3 + (5d-4)x^2 \\ &\quad + (2d-1)x + 1) + d(2x+1)^2 - 2x^2 - 2x + 1)\end{aligned}$$

It can be easily checked from these expressions that $\partial g/\partial x$ and $\partial^2 g/\partial x^2$ are both strictly positive for $d \ge 2$.

For $b < 1$, since $g(b, 0) = b^d > 0$ and $g(b, 1) < g(1, 1) = 1$ and g is strictly increasing and convex, it must be the case that g has a unique fixed point in $(0, 1)$. If $b = 1$, although $g(1) = 1$, and we further note that

$$\frac{\partial g}{\partial x}|_{x=1} = \frac{2d}{3} > 1$$

if $d \ge 2$. Thus g has a unique fixed point in $(0, 1)$ in this case also. This takes care of the 'also' part in the lemma as well.

The fact that $g^{(i)}(c)$ converges to the fixed point c_0 is immediate from the fact that $c_0 < g(x) < x$ if $x > c_0$ and $c_0 > g(x) > x$ if $x < c_0$, and hence the iterates monotonically converge to c_0. □

We now introduce a convenient notion of convergence of sets of the form $\mathcal{S}_{a,b,c}$. We say $\mathcal{S}_{a,b,c} \xrightarrow{\psi} \mathcal{S}_{a_0,b_0,c_0}$ if for all $x \in \mathcal{S}_{a,b,c}$ and all $\varepsilon > 0$, there exists a K such that for all $k \ge K$, $\psi^{(k)}(x) \in$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

$\mathcal{S}_{a_0-\varepsilon,b_0+\varepsilon,c_0+\varepsilon}$. Similarly, we define $\mathcal{S}_{a,b,c} \xrightarrow{\psi^2} \mathcal{S}_{a_0,b_0,c_0}$ for all $x \in \mathcal{S}_{a,b,c}$ and all $\varepsilon > 0$, there exists a K such that for all $k \geq K$, $\psi^{(2k)}(x) \in \mathcal{S}_{a_0-\varepsilon,b_0+\varepsilon,c_0+\varepsilon}$.

Lemma 6.2. *Fix $c < 1$. We have*

$$\mathcal{S}_{0,1,c} \xrightarrow{\psi} \mathcal{S}_{0,1,c_g},$$

where c_g is the unique fixed point of $g(1,\cdot)$ in $(0,1)$ (as guaranteed by Lemma 6.1).

Proof. This is immediate from Lemmas 2.2 and 6.1. □

Now we plan to bootstrap this estimate into an iteration of the first two coordinates. To that end, define

$$\mathsf{i}(c,x) := f(1,f(1+c,x)), \qquad \mathsf{j}(c,x) := f(1+c,f(1,x)). \tag{6.1}$$

To motivate these definitions, note that $\mathsf{i}(c,a)$ is precisely the first coordinate of $\varphi^{(2)}(a,b,c)$. It turns out that $\mathsf{j}(c,b)$ serves as a bound on the second coordinate of $\varphi^{(2)}(a,b,c)$ when $c = c_g$.

Lemma 6.3. *For all $2 \leq d \leq 7$, there exists a $0 < c_g < c_d$ such that for all $c < c_d$, both $x \mapsto \mathsf{i}(c,x)$ and $x \mapsto \mathsf{j}(c,x)$ have unique fixed points a_c, b_c in $(0,1)$. Furthermore, both $|\partial \mathsf{i}/\partial x| < 1$ and $|\partial \mathsf{j}/\partial x| < 1$ in $[0,1]$. Consequently for any $x \in [0,1]$,*

$$\mathsf{i}^{(k)}(x) \to a_c \text{ and } \mathsf{j}^{(k)}(x) \to b_c \text{ as } k \to \infty.$$

Furthermore, $\mathsf{i}(c,x) > x$ if and only if $x \in [0,a_c)$ and $\mathsf{j}(c,x) > x$ if and only if $x \in [0,b_c)$.

We postpone the proof of Lemma 6.3 to later in order to maintain the flow of the argument.

Lemma 6.4. *Let c_g be as in Lemma 6.2 and a_{c_g}, b_{c_g} as in Lemma 6.3 for $c = c_g$. Then for all $c < 1$*

$$\mathcal{S}_{0,1,c} \xrightarrow{\psi^2} \mathcal{S}_{a_{c_g},b_{c_g},c_g}.$$

as $k \to \infty$.

Proof. Applying Lemma 6.2, we see that for all $\delta \geq 0$ so that $c_g + \delta < 1$ for all large enough k

$$\psi^{(k)}(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{0,1,c_g+\delta}$$

Recall

$$\varphi(a,b,c_g+\delta) = (f(1,b), f(1+c_g+\delta,a), g(b,c_g+\delta))$$

Since $g(b,x)$ is increasing in b, and $g(1,x) \leq x$ if and only if $x \geq c_g$, $g(b,c_g+\delta) \leq g(1,c_g+\delta) \leq c_g+\delta$. Thus we have

$$\mathcal{S}_{\varphi(a,b,c_g+\delta)} \subseteq \mathcal{S}_{f(1,b),f(1+c_g+\delta,a),c_g+\delta}$$

Now note

$$\begin{aligned}&\varphi(f(1,b),f(1+c_g+\delta,a),c_g+\delta)\\ &\quad = (f(1,f(1+c_g+\delta,a)),f(1+c_g+\delta,f(1,b)),g(f(1+c_g,a),c_g+\delta)).\end{aligned}$$

Note again that $g(f(1+c_g,a),c_g+\delta) \leq g(1,c_g+\delta) \leq c_g+\delta$. Thus we get

$$\mathcal{S}_{\varphi(f(1,b),f(1+c_g+\delta,a),c_g+\delta)} \subset \mathcal{S}_{\mathsf{i}(c_g+\delta,a),\mathsf{j}(c_g+\delta,b),c_g+\delta}.$$

Overall

$$\begin{aligned}\psi \circ \psi(\mathcal{S}_{a,b,c_g+\delta}) &\subset \psi(\mathcal{S}_{\varphi(a,b,c_g+\delta)}) \subset \psi(\mathcal{S}_{f(1,b),f(1+c_g+\delta,a),c_g})\\ &\subset \mathcal{S}_{\varphi(f(1,b),f(1+c_g+\delta,a),c_g+\delta)} \subset \mathcal{S}_{\mathsf{i}(c_g+\delta,a),\mathsf{j}(c_g+\delta,b),c_g+\delta}.\end{aligned}$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Thus iterating, and using Lemma 6.3, eventually the sequence enters $\mathcal{S}_{a_{c_g(\delta)}-\varepsilon,b_{c_g(\delta)}+\varepsilon,c_g+\delta}$ for arbitrarily small ε, where $a_{c_g(\delta)}, b_{c_g(\delta)}$ are the unique fixed points of $x \mapsto \mathrm{i}(c_g+\delta, x)$ and $\mathrm{j}(c_g+\delta, x)$ in $(0,1)$ respectively. By continuity, $a_{c_g(\delta)} \to a_{c_g}$, $b_{c_g(\delta)} \to b_{c_g}$ as $\delta \to 0$. This allows us to conclude the lemma. □

Although the approximation (a_{c_g}, b_{c_g}, c_g) for the fixed point of $\varphi(a,b,c)$ is good enough for $3 \le d \le 6$, we need another round of iteration to get closer to the fixed point of $\varphi(a,b,c)$ for $d \in \{2,7\}$. To that end, define

$$b'_g = \max(f(1+c_g, a_{c_g}), b_{c_g}).$$

Let $\tilde{c}_g$ be the fixed point of the map $x \mapsto g(b'_g, x)$ (which is guaranteed to exist by Lemma 6.2).

Lemma 6.5. *We have for all $c < 1$*

$$\mathcal{S}_{0,1,c} \xrightarrow{\psi^2} \mathcal{S}_{a_{c_g}, b_{c_g}, \tilde{c}_g}$$

as $k \to \infty$. Furthermore, $\tilde{c}_g < c_g < c_d$.

Proof. Since $b'_g < 1$, $\tilde{c}_g = g(b'_g, \tilde{c}_g) < g(1, \tilde{c}_g)$ which implies $\tilde{c}_g < c_g$.

Next, by Lemma 6.4, for all δ,

$$\psi^{(2k)}(\mathcal{S}_{0,1,c}) \subset \mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, c_g+\delta}.$$

for all large enough k. Pick δ small enough so that $c_g + \delta < c_d$ and $b''_g < 1$ where

$$b''_g = \max(f(1+c_g+\delta, a_{c_g}-\delta), b_{c_g}+\delta). \tag{6.2}$$

Note that

$$g(b''_g, c_g) < g(1, c_g) = c_g$$

Hence by continuity we can pick a further small δ if required to conclude that $g(b''_g, c_g+\delta) \le c_g$.

We first show that after two more steps of iteration, we enter $\mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, c_g}$. Indeed after one iteration

$$\psi(\mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, c_g}) \subset \mathcal{S}_{f(1,b_{c_g}+\delta), f(1+c_g+\delta, a_{c_g}-\delta), g(b_{c_g}-\delta, c_g)}$$

by the choice of k. Now note since $b \mapsto g(b,x)$ is increasing in b,

$$g(b_{c_g}-\delta, c_g+\delta) \le g(b''_g, c_g+\delta) \le c_g$$

by the choice of δ. By the same logic and the choice of k, after another iteration, we are in $\mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, c_g}$.

We will show that for all $c \le c_g$,

$$\psi^{(2)}\big(\mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, c}\big) \subseteq \mathcal{S}_{a_{c_g}-\delta, b_{c_g}+\delta, g(b''_g, g(b''_g, c))} \text{ and } g\big(b''_g, g(b''_g, c)\big) \le c_g. \tag{6.3}$$

Indeed by Lemma 6.2, this is enough since the fixed point of $x \mapsto g(b''_g, x)$ tends to that of $x \mapsto g(b'_g, x)$ as $\delta \to 0$ by continuity.

Since $b \mapsto g(b,x)$ is increasing (see (2.6)) and $c \mapsto f(1+c,x)$ is increasing,

$$g(b_{c_g}+\delta, c) \le g(b''_g, c), \qquad f(1+c, a_{c_g}-\delta) \le f(1+c_g, a_{c_g}-\delta).$$

Thus we have

$$\mathcal{S}_{\varphi(a_{c_g}-\delta, b_{c_g}+\delta, c)} \subseteq \mathcal{S}_{f(1,b_{c_g}+\delta), f(1+c_g, a_{c_g}-\delta), g(b''_g, c)}.$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Now

$$\varphi(f(1,b_{c_g}+\delta),f(1+c_g,a_{c_g}-\delta),g(b''_g,c))$$
$$=(f(1,f(1+c_g,a_{c_g}-\delta)),f(1+g(b''_g,c),f(1,b_{c_g}+\delta)),g(f(1+c_g,a_{c_g}-\delta),g(b''_g,c)))$$

Now note,

$$f(1,f(1+c_g,a_{c_g}-\delta))=\mathrm{i}(c_g,a_{c_g}-\delta)\geq a_{c_g}-\delta$$

since $\mathrm{i}(c_g,x)>x$ if and only if $x<a_{c_g}$. Also since $b\mapsto g(b,c)$ and $c\mapsto g(b,c)$ are both increasing, and $g(1,x)<x$ if and only if $x>c_g$,

$$g(b''_g,c)\leq g(b''_g,c_g)\leq c_g,\qquad g(f(1+c_g,a_{c_g}-\delta),g(b''_g,c))\leq g(b''_g,g(b''_g,c)).$$

and hence since $c\mapsto f(1+c,x)$ is increasing as well,

$$f(1+g(b''_g,c),f(1,b_{c_g}+\delta))\leq f(1+c_g,f(1,b_{c_g}+\delta))\leq b_{c_g}+\delta.$$

where for the last equality we used the face that $j(c_g,x)<x$ if and only if $x>b_{c_g}$. Combining, we get

$$\mathcal{S}_{\varphi(f(1,b_{c_g}+\delta),f(1+c_g,a_{c_g}-\delta),g(b''_g,c))}\subseteq\mathcal{S}_{a_{c_g}-\delta,b_{c_g}+\delta,g(b''_g,g(b''_g,c))}.$$

We now need to show that $g(b''_g,g(b''_g,c))\leq c_g$, which is clear since

$$g(b''_g,g(b''_g,c))\leq g(b''_g,c_g)\leq c_g$$

by our choice of δ. The proof is complete. □

Next we bootstrap this to improve upon a_{c_g},b_{c_g}. Let $a_{\tilde{c}_g},b_{\tilde{c}_g}$ be the fixed points of $x\mapsto\mathrm{i}(\tilde{c}_g,x)$ and $x\mapsto\mathrm{j}(\tilde{c}_g,x)$ respectively. Then

Lemma 6.6. *We have for all $c<1$*

$$\mathcal{S}_{0,1,c}\xrightarrow{\psi^2}\mathcal{S}_{a_{\tilde{c}_g},b_{\tilde{c}_g},\tilde{c}_g}.$$

as $k\to\infty$.

Proof. First fix $\delta'>0$ so that $\tilde{c}_g+\delta'<c_g$. Then choose δ small enough such that letting b''_g be as in Equation (6.2), $\tilde{c}_g+\delta'$ is bigger than the fixed point of $x\mapsto g(b''_g,x)$.

By Lemma 6.5, we see that for all k large enough

$$\psi^{(2k)}(\mathcal{S}_{0,1,c})\subset\mathcal{S}_{a_{c_g}-\delta,b_{c_g+\delta},\tilde{c}_g+\delta'}$$

By continuity of φ, it is enough to show that for all $a\geq a_{c_g}-\delta,b\leq b_{c_g}+\delta$,

$$\psi^{(2)}(\mathcal{S}_{a,b,\tilde{c}_g+\delta'})\subseteq\mathcal{S}_{\mathrm{i}(\tilde{c}_g+\delta,a),\mathrm{j}(\tilde{c}_g+\delta,b),\tilde{c}_g+\delta'},\qquad\mathrm{i}(\tilde{c}_g,a)\geq a_{c_g}-\delta,\quad\mathrm{j}(\tilde{c}_g,b)\leq b_{c_g}+\delta.$$

Note that

$$\varphi(a,b,\tilde{c}_g+\delta')=(f(1,b),f(1+\tilde{c}_g+\delta',a),g(b,\tilde{c}_g+\delta'))$$

and

$$g(b,\tilde{c}_g+\delta')\leq g(b_{c_g}+\delta,\tilde{c}_g+\delta')\leq g(b''_g,\tilde{c}_g+\delta')\leq\tilde{c}_g+\delta'.$$

since $x>g(b''_g,x)$ if and only if x is at least the fixed point of $x\mapsto g(b''_g,x)$ and δ' is chosen to satisfy this. Thus

$$\psi(\mathcal{S}_{a,b,\tilde{c}_g})\subseteq\psi(\mathcal{S}_{f(1,b),f(1+\tilde{c}_g+\delta',a),\tilde{c}_g+\delta'}).$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Again observe,

$$\varphi(f(1,b),f(1+\tilde{c}_g+\delta',a),\tilde{c}_g+\delta')=(\mathsf{i}(\tilde{c}_g+\delta',a),\mathsf{j}(\tilde{c}_g+\delta',b),g(f(1+\tilde{c}_g+\delta',a),\tilde{c}_g+\delta'))$$

Now observe that since $a \geq a_{c_g}-\delta$ and $x \mapsto f(1+c,x)$ is decreasing, $c \mapsto f(1+c,x)$ is increasing, and $\tilde{c}_g+\delta' < c_g$

$$f(1+\tilde{c}_g+\delta',a) \leq f(1+c_g+\delta, a_{c_g}-\delta) \leq b''_g$$

Thus

$$g(f(1+\tilde{c}_g+\delta',a),\tilde{c}_g+\delta') \leq g(b''_g,\tilde{c}_g+\delta') \leq \tilde{c}_g+\delta'.$$

Thus

$$\psi \circ \psi(\mathcal{S}_{a,b,\tilde{c}_g}) \subset \mathcal{S}_{\mathsf{i}(\tilde{c}_g+\delta',a),\mathsf{j}(\tilde{c}_g+\delta',b),\tilde{c}_g+\delta'}$$

Thus we are only left to prove the inequalities involving i, j. Also, for the same reason,

$$f(1+\tilde{c}_g+\delta',a) \leq f(1+c_g, a_{c_g}-\delta)$$

and thus

$$\mathsf{i}(\tilde{c}_g+\delta',a) \geq f(1,f(1+c_g,a_{c_g}-\delta)) = \mathsf{i}(c_g,a_{c_g}-\delta) \geq a_{c_g}-\delta.$$

Similarly,

$$\mathsf{j}(\tilde{c}_g+\delta',b) \leq \mathsf{j}(c_g,b_{c_g}+\delta) \leq b_{c_g}+\delta.$$

and we are done. □

6.1 Finding approximations analytically

We now explain how to analytically find approximations for $c_g, a_{c_g}, b_{c_g}, c_g$ and $\tilde{c}_g, a_{\tilde{c}_g}, b_{\tilde{c}_g}$. Recall the strategy as outlined at the beginning of this section: all the functions $x \mapsto g(b,x)$, $x \mapsto \mathsf{i}(c,x), x \mapsto \mathsf{j}(c,x)$ have unique fixed point in $[0,1)$, and furthermore, the function evaluated at x is at least x if and only if x is smaller than the fixed point. This can be utilized to approximate the parameters in the following sequence of steps.

- Find $c_{g,1}, c_{g,2}$ so that $g(1,c_{g,1}) > c_{g,1}$ and $g(1,c_{g,2}) < c_{g,2}$.
- Next, find $a_{c_{g,1}}, a_{c_{g,2}}$ such that $\mathsf{i}(c_{g,1}, a_{c_{g,1}}) < a_{c_{g,1}}$ and $\mathsf{i}(c_{g,2}, a_{c_{g,2}}) > a_{c_{g,2}}$. Similarly, find $b_{c_{g,1}}$, $b_{c_{g,2}}$ such that $\mathsf{j}(c_{g,1}, b_{c_{g,1}}) > b_{c_{g,1}}$ and $\mathsf{j}(c_{g,2}, b_{c_{g,2}}) < b_{c_{g,2}}$.

Lemma 6.7. *We have*

$$c_{g,1} \leq c_g \leq c_{g,2}$$
$$a_{c_{g,2}} \leq a_{c_g} \leq a_{c_{g,1}}$$
$$b_{c_{g,1}} \leq b_{c_g} \leq b_{c_{g,2}}.$$

Proof. The first item is immediate from Lemma 6.1. For the inequalities involving a, let $\tilde{a}_{c_{g,1}}$ denote the unique fixed point of $x \mapsto \mathsf{i}(c_{g,1},x)$ in $(0,1)$. By choice and properties of i, $a_{c_{g,1}} > \tilde{a}_{c_{g,1}}$. But

$$\tilde{a}_{c_{g,1}} = \mathsf{i}(c_{g,1},\tilde{a}_{c_{g,1}}) = f(1,f(1+c_{g,1},\tilde{a}_{c_{g,1}})) \geq \mathsf{i}(c_g,\tilde{a}_{c_{g,1}})$$

and hence $a_{c_{g,1}} \geq \tilde{a}_{c_{g,1}} > a_{c_g}$. The other inequalities follow exactly the same line of logic, which we skip. □

Now define

$$b'_{g,1} = \max(f(1+c_{g,1},a_{c_{g,1}}), b_{c_{g,1}}) \qquad b'_{g,2} = \max(f(1+c_{g,2},a_{c_{g,2}}), b_{c_{g,2}}),$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Table 5. Approximations of c_g, a_{c_g}, b_{c_g}

$d \to$	2	3	4	5	6	7
$a_{c_g,2}$	.48	.4	.3495	.308	.2734	.2332
$b_{c_g,2}$	.754	.54	.43	.362	.318	.3006
$c_{g,2}$	.466	.17	.074	.0342	.0165	.0081

Table 6. Approximations of $\tilde{c}_g, a_{\tilde{c}_{g,2}}, b_{\tilde{c}_{g_2}}$ $d = 2, 7$

$d \to$	2	7
$b'_{g,2}$	.755	.3007
$a_{\tilde{c}_{g,2}}$	.51	.26435
$b_{\tilde{c}_{g,2}}$	.664	.26475
$\tilde{c}_{g,2}$	.267	3.6×10^{-5}

and then we judiciously choose $\tilde{c}_{g,1}$ such that $g(b'_{g,1}, \tilde{c}_{g,1}) > \tilde{c}_{g,1}$ (by judicious, we mean as usual as close as the fixed point as possible) and similarly choose $\tilde{c}_{g,2}$ to be such that $g(b'_{g,1}, \tilde{c}_{g,2}) < \tilde{c}_{g,2}$. Then define $a_{\tilde{c}_{g,i}}, b_{\tilde{c}_{g,i}}$ for $i \in \{1, 2\}$ be such that

$$\mathrm{i}(\tilde{c}_{g,1}, a_{\tilde{c}_{g,1}}) < a_{\tilde{c}_{g,1}} \qquad \mathrm{i}(\tilde{c}_{g,2}, a_{\tilde{c}_{g,2}}) > a_{\tilde{c}_{g,2}}, \qquad \mathrm{j}(\tilde{c}_{g,1}, b_{\tilde{c}_{g,1}}) > b_{\tilde{c}_{g,1}}, \qquad \mathrm{j}(\tilde{c}_{g,2}, b_{\tilde{c}_{g,2}}) < b_{\tilde{c}_{g,2}}.$$

The motivation for these definitions is the same as before: choose an approximation of the fixed point so that the inequalities are in the good direction. Then

Lemma 6.8.

$$\tilde{c}_{g,1} \le \tilde{c}_g \le \tilde{c}_{g,2}$$
$$a_{\tilde{c}_{g,2}} \le a_{\tilde{c}_g} \le a_{\tilde{c}_{g,1}}$$
$$b_{\tilde{c}_{g,1}} \le b_{\tilde{c}_g} \le b_{\tilde{c}_{g,2}}$$

Proof. Note $b'_{g,1} \le b'_g$ by choice and Lemma 6.7. Let $c'_{g,1}$ be the fixed point of $x \mapsto g(b'_{g,1}, x)$. Thus

$$\tilde{c}_{g,1} < c'_{g,1} = g(b'_{g,1}, c'_{g,1}) \le g(b'_g, c'_{g,1})$$

which implies $c'_{g,1} \le \tilde{c}_g$ which is the fixed point of $x \mapsto g(b'_g, x)$. Thus $\tilde{c}_{g,1} \le c'_{g,1} \le \tilde{c}_g$ as desired. The other direction follows exactly similarly by noting $b'_{g,2} \ge b'_g$.

The inequalities involving a and b can be proven in a similar way to those in Lemma 6.7 using the bounds for c, we leave the details to the reader. □

In reality, we would need a good upper bound for $c_g, \tilde{c}_g, b_{c_g}, \tilde{b}_{c_g}$ and a good lower bound for $a_{c_g}, a_{\tilde{c}_g}$. We record these bounds in the following tables.

Since we only do even iterations in Tables 5 and 6, we do one more iteration of φ to the values in Table 5 for $3 \le d \le 6$ and to that in Table 6 for $d = 2, 7$.

Taking the minimum over rows involving estimates of a_* and maximum over the rows involving the estimates of b_*, c_* from Tables 5 and 7 for $3 \le d \le 6$ and Tables 6 and 7 for $d = 2, 7$ we obtain the final table, Table 8.

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Table 7. Approximations for odd round of iterations

$d \to$	2	3	4	5	6	7
$a^o_{c_{g,2}}$	.52	.41	.3493	.307	.273	.26435
$b^o_{c_{g,2}}$	.665	.55	.43	.362	.318	.26476
$c^o_{g,2}$	.233	.063	.02	.0016	.0003	1.7616×10^{-5}

Table 8. Final approximations

$d \to$	2	3	4	5	6	7
$\hat{a}_*$	.51	.4	.3493	.307	.273	.26435
$\hat{b}_*$	.665	.55	.43	.362	.318	.26476
$\hat{c}_*$	.267	.17	.074	.0342	.0165	3.6×10^{-5}

***Proof of Proposition* 2.8.** The reader can check the values in Table 8 for $\hat{a}_*$ is bigger than that in Table 2 and that of $\hat{b}_*, \hat{c}_*$ are smaller than that in Table 2. □

6.2 Proof of Lemma 6.3

In this section, we write $f_a(x) = f(a, x)$

The derivative is

$$f'_a(x) = -\frac{(2-a)d}{(2x+1)^2}(f(x))^{\frac{d-1}{d}} = -(2-a)d\frac{(1+ax)^{d-1}}{(1+2x)^{d+1}} < 0 \text{ for } 1 \le a < 2 \tag{6.4}$$

$$f''_a(x) = \frac{df(x)(2-a)(a+(2-a)d+4ax+2)}{(2x+1)^2(ax+1)^2} > 0 \text{ for } 1 \le a < 2 \tag{6.5}$$

This for this range of a, f is decreasing.

***Proof of Lemma* 6.3.** Note that with the new notation used in this section

$$\mathsf{i}(x) = f_1 \circ f_{1+c}(x); \qquad \mathsf{j}(x) = f_{1+c} \circ f_1(x)$$

One can easily plot the derivatives of these functions for a range of choices of c and check the results for $0 < x < 1$, and perhaps that is the most practical way to be convinced of the lemma. We could not find a straightforward algebraic way to prove this lemma, so we employ the following rather undesirable case by case analysis.

First observe that by (6.4), f_1 and f_{1+c} are both decreasing and hence both i and j are increasing in $[0, 1]$. We will show that for all $2 \le d \le 7$, $|\mathsf{i}'(x)| < 1$ for all $x \in (0, 1)$. This ensures that both i, j are contractions in $(0, 1)$, thus the iterations of i and j must converge to a fixed point. Furthermore by Rolle's theorem, this fixed point must be unique.

Our strategy is as follows. Since both f_1, $|f'_1|$ are decreasing in $[0, 1]$, and since

$$|f'_{1+c}(x)| \le d\frac{(1+(1+c)x)^{d-1}}{(1+2x)^{d+1}} \le d\frac{(1+(1+c_d)x)^{d-1}}{(1+2x)^{d+1}},$$

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

for any interval $[a_1, a_2] \subset [0,1]$, we conclude

$$0 \le \mathrm{i}'(x) \le |f_1'(f_{1+c}(a_2))||f_{1+c}'(a_1)| \le |f_1'(f_1(a_2))| d \frac{(1+(1+c_d)a_1)^{d-1}}{(1+2a_1)^{d+1}} \text{ for } a_1 \le x \le a_2$$

$$0 \le \mathrm{j}'(x) \le |f_{1+c}'(f_1(a_2))||f_1'(a_1)| \le |f_1'(a_1)| d \frac{(1+(1+c_d)f_1(a_2))^{d-1}}{(1+2f_1(a_2))^{d+1}} \text{ for } a_1 \le x \le a_2$$

Denote

$$\xi(x) = \xi_{c_d}(x) = d\frac{(1+(1+c_d)x)^{d-1}}{(1+2x))^{d+1}}, \qquad \eta(x) = d\frac{(1+x)^{d-1}}{(1+2x)^{d+1}}.$$

With this new notation, we have

$$0 \le \mathrm{i}'(x) \le \eta(f_1(a_2))\xi(a_1) \text{ for } a_1 \le x \le a_2$$
$$0 \le \mathrm{j}'(x) \le \xi(f_1(a_2))\eta(a_1) \text{ for } a_1 \le x \le a_2$$

Clearly $\xi(x) \le \eta(x)$ for all $x \in [0,1]$, thus we can further approximate

$$\max\{\mathrm{i}'(x), \mathrm{j}'(x)\} \le \xi(f_1(a_2))\xi(a_1) \text{ for } a_1 \le x \le a_2 \tag{6.6}$$

It is now enough to pick a $c_d > c_g$ and then find a partition $[\alpha_1, \beta_1) \cup [\alpha_2, \beta_2) \cup \ldots \cup [\alpha_k, \beta_k]$ of $[0,1]$ so that $\xi(f_1(\beta_i))\xi(\alpha_i) < 1$ for all $1 \le i \le k$. This can be done by hand. We provide the partition below for $2 \le d \le 7$ and leave it to the diligent reader to check (6.6) holds for each partition. It is interesting to note that the partitions get finer as we move closer to the critical value of d (which is equal to 7 in this case).

- For $d = 2$, we choose $c_d = .47$. The crudest choice partition $a_1 = 0, a_2 = 1$ suffices.
- For $d = 3$, we take $c_d = .18$, and we need the partition
$$[0, .15] \cup (.15, .65] \cup (.65, 1].$$
- For $d = 4$, we take $c_d = .08$, and we partition
$$[0, .08) \cup [.08, .2) \cup [.2, .41) \cup [.41, 1].$$
- For $d = 5$, we take $c_d = .04$, and we partition
$$[0, .05) \cup [.05, .1) \cup [.1, .16) \cup [.16, .23] \cup (.23, .33] \cup (.33, .5] \cup (.5, 1].$$
- For $d = 6$, we take $c_d = .02$, and we partition
$$[0, .04) \cup [.04, .08) \cup [.08, .11) \cup [.11, .13] \cup (.13, .16] \cup (.16, .19] \cup (.19, .23] \cup (.23, .27] \cup (.27, .5] \cup [.5, .9) \cup [.9, 1].$$
- For $d = 7$, we take $c_d = .009$, and we partition
$$(0, .03] \cup [.03, .07) \cup [.07, .1) \cup [.1, .12) \cup [.12, 14) \cup [.14, .15)$$
$$\cup \left(\cup_{i=0}^{6} [.15 + .01i, .16 + .01i] \right)$$
$$\cup [.22, .225) \cup [.225, .23) \cup [.23, .238) \cup [.238, .245) \cup [.245, .253)$$
$$\cup [.253, .262) \cup [.262, .272) \cup [.272, .28) \cup [.28, .29) \cup [.29, .3) \cup [.3, .31)$$
$$\cup [.31, .325) \cup [.325, .34) \cup [.34, .365) \cup [.365, .4)$$
$$\cup [.4, .45) \cup [.45, .55) \cup [.55, .85) \cup [.85, 1),$$

which completes the proof. □

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press

Acknowledgements

We thank Omer Angel, Jinyoung Park, and Ron Peled for several illuminating discussions. We thank the anonymous referees for their suggestions, which improved the paper. KK would like to thank the Salmon Coast Field Station for hosting him while a portion of this work was completed.

Data availability statement

There is no additional data required to replicate the findings of this article.

Competing interests

The authors have no competing interests to declare.

References

[1] Benjamini, I., Häggström, O. and Mossel, E. (2000) On random graph homomorphisms into **Z**. *J. Combin. Theory Ser. B* **78**(1) 86–114.
[2] Benjamini, I. and Peres, Y. (1994) Tree-indexed random walks on groups and first passage percolation. *Probab. Theory Related Fields* **98**(1) 91–112.
[3] Benjamini, I. and Schechtman, G. (2000) Upper bounds on the height difference of the Gaussian random field and the range of random graph homomorphisms into $\mathbb{Z}$. *Random Struct. Algor.* **17**(1) 20–25.
[4] Benjamini, I. and Schramm, O. (2011) Percolation beyond Z^d, many questions and a few answers [mr1423907]. In *Selected Works of Oded Schramm*, Vol. **1, 2**, Sel. Works Probab. Stat., New York: Springer, pp. 679–690.
[5] Benjamini, I., Yadin, A. and Yehudayoff, A. (2007) Random graph-homomorphisms and logarithmic degree. *Electron. J. Probab.* **12**(32) 926–950.
[6] Chandgotia, N., Peled, R., Sheffield, S. and Tassy, M. (2018) Delocalization of uniform graph homomorphisms from $\mathbb{Z}^2$ to $\mathbb{Z}$. *Commun. Math. Phys.* **387**, 621–647 (2021).
[7] Duminil-Copin, H., Harel, M., Laslier, B., Raoufi, A. and Ray, G. (2022) Logarithmic variance for the height function of square-ice. *Commun. Math. Phys.* **396**(2) 1–36.
[8] Galvin, D. (2003) On homomorphisms from the Hamming cube to **Z**. *Israel J. Math.* **138** 189–213.
[9] Georgii, H.-O. (2011) *Gibbs Measures and Phase Transitions*, Vol. **9**. Walter de Gruyter.
[10] Georgii, H.-O., Häggström, O. and Maes, C. (2001) The random geometry of equilibrium phases. *Ph. Transit. Crit. Phenom.* **18** 1–142.
[11] Grimmett, G. R. (2006) *The random-cluster model*, Vol. **333**. Springer Science & Business Media.
[12] Häggström, O., Schonmann, R. H. and Steif, J. E. (2000) The Ising model on diluted graphs and strong amenability. *Ann. Probab.* **28**(3) 1111–1137.
[13] Hutchcroft, T. (2023) Continuity of the Ising phase transition on nonamenable groups. *Comm. Math. Phys.* **404**(1) 227–286.
[14] Jonasson, J. and Steif, J. E. (1999) Amenability and phase transition in the Ising model. *J. Theoret. Probab.* **12**(2) 549–559.
[15] Karrila, A. M. (2023) Logarithmic delocalization of random Lipschitz functions on honeycomb and other lattices. arXiv preprint arXiv:2309.14895.
[16] Kissel, S., Külske, C. and Rozikov, U. A. (2019) Hard-core and soft-core Widom-Rowlinson models on Cayley trees. *J. Stat. Mech. Theory Exp.* **2019**(4) 043204.
[17] Krueger, R. A., Li, L. and Park, J. (2026) Lipschitz functions on weak expanders. *J. Lond. Math. Soc.* **113**(3) Paper No. e70500, 32 pp.
[18] Lammers, P. and Toninelli, F. (2024) Height function localisation on trees. *Combin. Probab. Comput.* **33**(1) 50–64.
[19] Lyons, R. (2017) Factors of IID on trees. *Combin. Probab. Comput.* **26**(2) 285–300.
[20] Lyons, R. and Peres, Y. (2017) *Probability on trees and networks*, Vol. 42. Cambridge University Press.
[21] Peled, R. (2017) High-dimensional Lipschitz functions are typically flat. *Ann. Probab.* **45**(3) 1351–1447.
[22] Peled, R., Samotij, W. and Yehudayoff, A. (2013) Grounded Lipschitz functions on trees are typically flat. *Electron. Commun. Probab* **18**(55) 9.
[23] Peled, R., Samotij, W. and Yehudayoff, A. (2013) Lipschitz functions on expanders are typically flat. *Combin. Probab. Comput.* **22**(4) 566–591.
[24] Peled, R. and Spinka, Y. (2020) Long-range order in discrete spin systems. arXiv preprint arXiv:2010.03177.
[25] Rozikov, U. A. (2013) *Gibbs Measures on Cayley Trees*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

Cite this article: Butler N, Krishnan K, Ray G, and Spinka Y (2026). On the local convergence of integer-valued Lipschitz functions on regular trees. *Combinatorics, Probability and Computing* **35**, 655–687. https://doi.org/10.1017/S0963548326100467

https://doi.org/10.1017/S0963548326100467 Published online by Cambridge University Press