

Random walk with long-range constraints*

Ron Peled[†] Yinon Spinka[‡]

Abstract

We consider a model of a random height function with long-range constraints on a discrete segment. This model was suggested by Benjamini, Yadin and Yehudayoff and is a generalization of simple random walk. The random function is uniformly sampled from all graph homomorphisms from the graph $P_{n,d}$ to the integers \mathbb{Z} , where the graph $P_{n,d}$ is the discrete segment $\{0, 1, \dots, n\}$ with edges between vertices of different parity whose distance is at most $2d + 1$. Such a graph homomorphism can be viewed as a height function whose values change by exactly one along edges of the graph $P_{n,d}$. We also consider a similarly defined model on the discrete torus.

Benjamini, Yadin and Yehudayoff conjectured that this model undergoes a phase transition from a delocalized to a localized phase when d grows beyond a threshold $c \log n$. We establish this conjecture with the precise threshold $\log_2 n$. Our results provide information on the typical range and variance of the height function for every given pair of n and d , including the critical case when $d - \log_2 n$ tends to a constant.

In addition, we identify the local limit of the model, when d is constant and n tends to infinity, as an explicitly defined Markov chain.

Keywords: random graph homomorphism; homomorphism height function; constrained random walk; localization / delocalization transition.

AMS MSC 2010: 82B41; 60C05; 82B20; 82B26; 60D05; 05A16; 60K35.

Submitted to EJP on October 6, 2013, final version accepted on June 13, 2014.

Supersedes arXiv:1310.1583.

1 Introduction

Given two graphs G and H , a *graph homomorphism* from G to H is a function $f: V(G) \rightarrow V(H)$ such that if x and y are neighbors in G , then $f(x)$ and $f(y)$ are neighbors in H . A graph homomorphism from a graph G to \mathbb{Z} is then a map from the vertex set of G to the integers, that maps adjacent vertices to adjacent integers. For a given vertex $v_0 \in G$, we denote by $\text{Hom}(G, v_0)$ the set of all homomorphisms from G to \mathbb{Z} , which map v_0 to 0. Precisely,

$$\text{Hom}(G, v_0) := \{f: V(G) \rightarrow \mathbb{Z} \mid f(v_0) = 0, |f(x) - f(y)| = 1 \text{ when } (x, y) \in E(G)\}.$$

*Research of R.P. and Y.S. supported by an ISF grant and an IRG grant.

[†]Tel Aviv University, Israel. E-mail: peledron@post.tau.ac.il <http://www.math.tau.ac.il/~peledron>

[‡]Tel Aviv University, Israel. E-mail: yinonspi@post.tau.ac.il <http://www.math.tau.ac.il/~yinonspi>

The set $\text{Hom}(G, v_0)$ is non-empty and finite when G is finite, bipartite and connected. Benjamini, Häggström and Mossel [1] initiated the study of random \mathbb{Z} -homomorphisms, that is, uniformly chosen elements of $\text{Hom}(G, v_0)$. Special cases of this model include the simple random walk, when $G = \{0, 1, \dots, n\}$ with nearest-neighbor connections, the random walk bridge, when G is a cycle, and the branching random walk, when G is a tree [3, 2]. The model is sometimes referred to as a G -indexed random walk. The behavior of typical \mathbb{Z} -homomorphisms is poorly understood for general graphs G . Beyond simple and branching random walks, results are available mainly for the hypercube [8, 7], high-dimensional cubic lattices [10] and expander and tree graphs [1, 11]. In particular, the case when $G = \mathbb{Z}_{2n}^2$, a two-dimensional discrete torus, appears completely open. This case is related to the 6-vertex, square-ice and antiferromagnetic 3-state Potts models of statistical physics (see [10]).

Benjamini, Yadin and Yehudayoff [4] suggested the study of this model when $G = T_{n,d}$ is a certain one-dimensional graph with long-range edges, defined below, and where d may depend on n . In this work we study the properties of the model on this graph, as well as its close relative, the graph $P_{n,d}$. Specifically, let $P_{n,d}$, for $n, d \geq 1$, be the graph defined by

$$\begin{aligned} V(P_{n,d}) &:= \{0, 1, \dots, n\}, \\ E(P_{n,d}) &:= \{(i, j) \mid |i - j| \in \{1, 3, \dots, 2d + 1\}\}. \end{aligned} \tag{1.1}$$

Thus, a uniformly chosen random function f from $\text{Hom}(P_{n,d}, 0)$ is a simple random walk conditioned on satisfying $|f(i) - f(j)| = 1$ whenever i, j have different parity and are at distance at most $2d + 1$. Figure 1 shows a typical sample from $\text{Hom}(P_{n,d}, 0)$. Similarly, let $T_{n,d}$, $n \geq 1$ even and $d \geq 1$, be the graph defined by

$$\begin{aligned} V(T_{n,d}) &:= \{0, 1, \dots, n - 1\}, \\ E(T_{n,d}) &:= \{(i, j) \mid \min\{|i - j|, n - |i - j|\} \in \{1, 3, \dots, 2d + 1\}\}. \end{aligned} \tag{1.2}$$

Thus, a uniformly chosen random function f in $\text{Hom}(T_{n,d}, 0)$ is a simple random walk *bridge* conditioned on satisfying $|f(i) - f(j)| = 1$ whenever i, j have different parity and are at distance at most $2d + 1$ on the cycle.

In the rest of the paper we abbreviate \mathbb{Z} -homomorphisms to homomorphisms. We shall loosely refer to homomorphisms on $P_{n,d}$ as being on the line, and to homomorphisms on $T_{n,d}$ as being on the torus.

Our main objects of study are the size of the range of a typical homomorphism on $P_{n,d}$ or $T_{n,d}$ and the variance of the homomorphism at given vertices. For a graph G , the *range* of a function $f: V(G) \rightarrow \mathbb{Z}$ is defined as

$$\text{Rng}(f) := \{f(v) \mid v \in V(G)\}.$$

Benjamini, Yadin and Yehudayoff made the following conjecture.

Conjecture (I4). There exist constants $b, c > 0$ such that if $f_{n,d}$ is uniformly sampled from $\text{Hom}(T_{n,d}, 0)$,

1. If $d(n) - c \log n \rightarrow -\infty$ as $n \rightarrow \infty$ then for any positive integer r , we have

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq r) \xrightarrow[n \rightarrow \infty]{} 0.$$

2. If $d(n) - c \log n \rightarrow \infty$ as $n \rightarrow \infty$ then

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq b) \xrightarrow[n \rightarrow \infty]{} 1.$$

Random walk with long-range constraints

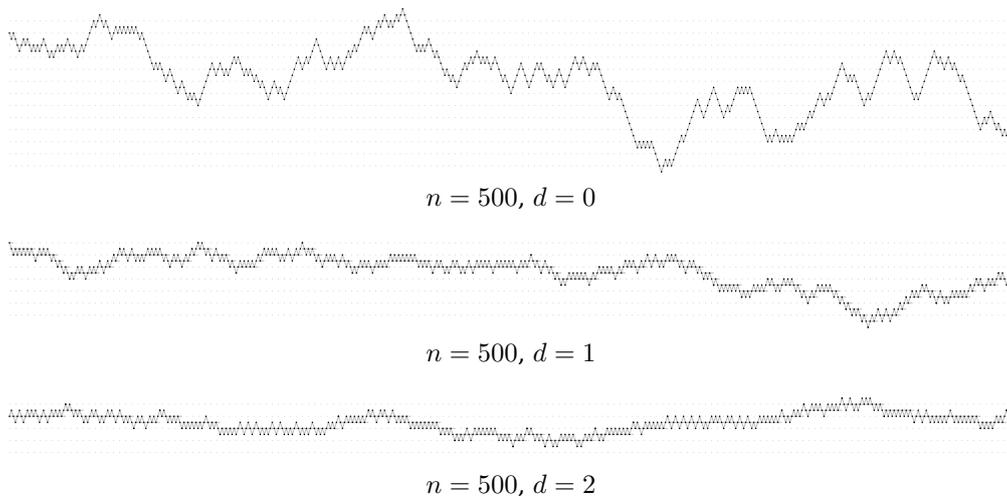


Figure 1: Uniformly sampled homomorphisms in $\text{Hom}(P_{n,d}, 0)$. The case $d = 0$ is just a simple random walk. The simulation uses a Metropolis algorithm (see, e.g., [9, Chapter 3]) and coupling from the past [12].

Our work establishes this conjecture with the precise constants $b = 3$ and $c = 1/\log 2$, both on $T_{n,d}$ and $P_{n,d}$. In addition, we discover that in the subcritical regime, when $d(n) - \log_2 n \rightarrow -\infty$, the size of the range of a typical homomorphism is of order $\sqrt{n2^{-d}}$ and the variance of the homomorphism at vertex k is of order $k2^{-d}$. Moreover, we explore the behavior in the critical regime, when $d(n) - \log_2 n \rightarrow \mu \in \mathbb{R}$, and find that in this case, the size of the range is a tight random variable whose distribution is closely related to the Poisson distribution.

Our results may be intuitively understood as follows. Let $f \in \text{Hom}(P_{n,d})$. It is not difficult to verify that if $f(i+m) - f(i) \geq 3$ then $m \geq 2d + 3$. Figure 2 shows such an event. Moreover, if $m = 2d + 3$ and this event occurs, then necessarily

$$(f(i+1) - f(i), f(i+2) - f(i), \dots, f(i+2d+3) - f(i)) = (1, 2, 1, 2, \dots, 1, 2, 1, 2, 3).$$

However, if this sequence of values is possible for f , then there are at least 2^d other possible candidates of the form

$$(1, 1 + s_1, 1, 1 + s_2, \dots, 1, 1 + s_d, 1, 2, 1), \quad s_i \in \{-1, 1\}.$$

Thus, intuitively, the probability that the homomorphism changes its height by 3 on any given small segment is about 2^{-d} . Therefore, when $n2^{-d} \rightarrow 0$, we will not have any such segment, so that the size of the range of the homomorphism will be bounded by 3. Conversely, when $n2^{-d} \rightarrow \infty$, the expected number of segments with an upward or downward movement of size 3 will be roughly $n2^{-d}$. Since the direction of these movements should be only mildly correlated, we expect the size of the resulting range to be of order $\sqrt{n2^{-d}}$. Our work makes these ideas precise.

2 Main Results

2.1 Homomorphisms on the line

In this section we present results on homomorphisms on the graph $P_{n,d}$, which was defined in (1.1). Throughout this section, $f_{n,d}$ denotes a uniformly chosen homomorphism in $\text{Hom}(P_{n,d}, 0)$.

Random walk with long-range constraints

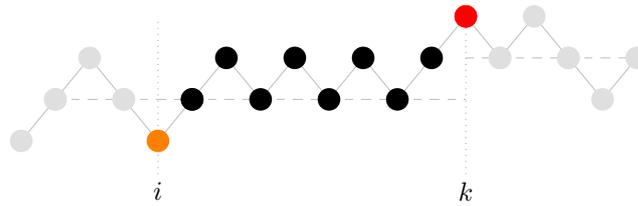


Figure 2: A homomorphism jumps from some value t at vertex i to $t + 3$ at vertex k . The minimal length of such a segment is $k - i = 2d + 3$. In order for this jump to occur, the values at the $d + 1$ vertices, $k - 1, k - 3, \dots, k - 2d - 1$, are forced to be $t + 2$. Here $d = 3$.

We state results regarding the size of the range of a typical homomorphism. As a homomorphism must change its value by exactly one along edges, the range is always of size at least 2. In fact, the range is exactly 2 only for two particular homomorphisms, and at least 3 otherwise. We shall show that the size of the range is 3 plus a term of order $\sqrt{n2^{-d}}$. Hence, we distinguish between three regimes, $n2^{-d} \rightarrow \infty$, $n2^{-d} \rightarrow 0$ and $n2^{-d} \rightarrow \lambda \in (0, \infty)$, termed the subcritical regime, the supercritical regime and the critical regime, respectively.

The supercritical regime. The supercritical regime is when $d(n) - \log_2 n \rightarrow \infty$ (i.e. $n2^{-d(n)} \rightarrow 0$) as $n \rightarrow \infty$. In this case, the large number of constraints prevents a typical homomorphism from growing. In fact, we show that, with high probability, it will take on only 3 values.

Theorem 2.1. For any positive integers n , d and r , we have

$$\mathbb{P}(|\text{Rng}(f_{n,d})| \geq 3 + r) \leq \binom{n}{r} 2^{-dr} \quad \text{and} \quad \mathbb{P}(|\text{Rng}(f_{n,d})| < 3) \leq 2^{1-n/2}.$$

Thus, if $d(n) - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| = 3) \xrightarrow{n \rightarrow \infty} 1.$$

The following corollary gives more precise information about the structure of a typical homomorphism in the supercritical regime. Denote by $V_i := \{2k + i \mid 0 \leq 2k + i \leq n\}$, $i = 0, 1$, the even and odd vertices, respectively, and denote by Ω_0 and Ω_1 the set of homomorphisms which are constant on V_0 and V_1 , respectively. Note that for each $i \in \{0, 1\}$, conditioned on $f \in \Omega_i$, the random vector $(f(x) - f(i) \mid x \in V_{1-i})$ consists of independent uniform signs.

Corollary 2.2. If $d(n) - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then

$$\begin{aligned} \mathbb{P}(\Omega_0 \cup \Omega_1) &= \mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq 3) \xrightarrow{n \rightarrow \infty} 1 \quad \text{and} \\ \mathbb{P}(\Omega_0 \cap \Omega_1) &= \mathbb{P}(|\text{Rng}(f_{n,d(n)})| < 3) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Moreover, if n tends to infinity through odd numbers then $\mathbb{P}(\Omega_0) \rightarrow 1/2$, and if n tends to infinity through even numbers then $\mathbb{P}(\Omega_0) \rightarrow 1/3$.

The corollary implies that a typical homomorphism in the supercritical regime has one of three possible structures. For odd values of n , with probability $1/2 - o(1)$, the homomorphism takes the value 0 on all the even vertices, with probability $1/4 - o(1)$, it takes the value 1 on all the odd vertices, and, with probability $1/4 - o(1)$, it takes the value -1 on all the odd vertices. For even values of n , the probability of each of these

three options is $1/3 - o(1)$. The dependence on the parity of n arises from the difference in the number of even and odd vertices in each case. For odd values of n , $|V_0| = |V_1|$, whereas for even values of n , $|V_0| = |V_1| + 1$.

The subcritical regime. The subcritical regime is when $d(n) - \log_2 n \rightarrow -\infty$ (i.e. $n2^{-d(n)} \rightarrow \infty$) as $n \rightarrow \infty$. Here, the relatively small number of constraints allows a typical homomorphism to grow.

Theorem 2.3. *There exist absolute constants $C, c > 0$ such that for any positive integers n and d , we have*

$$3 + \lfloor c\sqrt{n2^{-d}} \rfloor - 2^{1-n/2} \leq \mathbb{E}[|\text{Rng}(f_{n,d})|] \leq 3 + C\sqrt{n2^{-d}}.$$

Moreover, for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any positive integers n and d , we have

$$\mathbb{P}\left(|\text{Rng}(f_{n,d})| < 3 + \lfloor \delta\sqrt{n2^{-d}} \rfloor\right) \leq \epsilon + 2^{1-n/2}. \tag{2.1}$$

In particular, if $d(n) - \log_2 n \rightarrow -\infty$ as $n \rightarrow \infty$ then for any positive integer r , we have

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq r) \xrightarrow{n \rightarrow \infty} 0.$$

The next theorem quantifies the rate of growth of the variance of the homomorphism.

Theorem 2.4. *There exist absolute constants $C, c > 0$ such that for any positive integers n and d , we have*

$$\max\{ck2^{-d}, 1\} \leq \text{Var}(f_{n,d}(k)) \leq Ck2^{-d} + 4, \quad 1 \leq k \leq n.$$

The critical regime. The critical regime is when $d(n) - \log_2 n \rightarrow -\log_2 \lambda$ (i.e. $n2^{-d(n)} \rightarrow \lambda$) as $n \rightarrow \infty$, for some $\lambda \in (0, \infty)$. In this case, the balance between the number of constraints at each vertex and the amount of time available leads to an interesting limiting behavior. Perhaps surprisingly, it turns out that the parity of n induces an effect which does not disappear in the limit.

Denote by $\mu^{\text{even}}(\lambda)$ the distribution of a Poisson(λ) variable conditioned to be even, and denote by $\mu^{\text{odd}}(\lambda)$ the distribution of a Poisson(λ) variable conditioned to be odd. Define the *parity-biased Poisson distribution* with parameters λ and α to be the following convex combination of $\mu^{\text{even}}(\lambda)$ and $\mu^{\text{odd}}(\lambda)$,

$$\mu(\lambda, \alpha) := \frac{\alpha}{\alpha + \tanh(\lambda)} \cdot \mu^{\text{even}}(\lambda) + \frac{\tanh(\lambda)}{\alpha + \tanh(\lambda)} \cdot \mu^{\text{odd}}(\lambda). \tag{2.2}$$

One may check that

$$\mu(\lambda, \alpha)(r) = Z(\lambda, \alpha)^{-1} \cdot \alpha(r) \cdot \frac{\lambda^r}{r!}, \quad r \geq 0, \tag{2.3}$$

where $\alpha(r) = \alpha$ if r is even and $\alpha(r) = 1$ if r is odd and where $Z(\lambda, \alpha)$ is a normalizing constant. In particular, we see that the Poisson(λ) distribution is obtained as $\mu(\lambda, 1)$.

Let $(S_i \mid i = 0, 1, \dots)$ denote a simple random walk, and let

$$N^\pm(\lambda) \sim \mu\left(\lambda/(2\sqrt{2}), (3/(2\sqrt{2}))^{\pm 1}\right)$$

be independent of $(S_i \mid i \geq 0)$. Then $S_{N+(\lambda)}$ and $S_{N-(\lambda)}$ are simple random walks stopped at independent random times. For a positive integer k , denote $\text{Rng}(S_k) := \{S_i \mid 0 \leq i \leq k\}$.

Theorem 2.5. *If $n2^{-d(n)} \rightarrow \lambda$ as $n \rightarrow \infty$, for some $\lambda \in (0, \infty)$, then*

$$|\text{Rng}(f_{n,d(n)})| \xrightarrow[n \text{ even}]{n \rightarrow \infty} |\text{Rng}(S_{N^+(\lambda)})| + 2 \quad \text{and} \quad |\text{Rng}(f_{n,d(n)})| \xrightarrow[n \text{ odd}]{n \rightarrow \infty} |\text{Rng}(S_{N^-(\lambda)})| + 2.$$

Here, the notation $X_n \Rightarrow X$ as $n \rightarrow \infty$, n even (odd), means that X_n converges in distribution to X when n tends to infinity along even (odd) values. In fact, as the proof shows, one may couple a critical homomorphism to a simple random walk run for N^+ or N^- steps, according to the parity of n .

2.2 Homomorphisms on the torus

In this section we present results for homomorphisms on the graph $T_{n,d}$, which was defined in (1.2). Throughout this section, n is even and $f_{n,d}$ denotes a uniformly chosen homomorphism in $\text{Hom}(T_{n,d}, 0)$.

The supercritical regime. The supercritical regime is when $d(n) - \log_2 n \rightarrow \infty$ (i.e. $n2^{-d(n)} \rightarrow 0$) as $n \rightarrow \infty$. Similarly to the case on the line, the large number of constraints cause a typical homomorphism to take on only 3 values.

Theorem 2.6. *For any positive even integer n and any positive integers d and r , we have*

$$\mathbb{P}(|\text{Rng}(f_{n,d})| \geq 3 + r) \leq \binom{n}{r}^2 2^{-(2d-1)r} \quad \text{and} \quad \mathbb{P}(|\text{Rng}(f_{n,d})| < 3) \leq 2^{1-n/2}$$

Thus, if $d(n) - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| = 3) \xrightarrow[n \rightarrow \infty]{} 1.$$

Similarly to the case of the line, the following corollary gives more precise information about the structure of a typical homomorphism in the supercritical regime. Denote by $V_i := \{2k + i \mid 0 \leq k < n/2\}$, $i = 0, 1$, the even and odd vertices of $T_{n,d}$, respectively, and denote by Ω_0 and Ω_1 the set of homomorphisms which are constant on V_0 and V_1 , respectively. Note that for each $i \in \{0, 1\}$, conditioned on $f \in \Omega_i$, the random vector $(f(x) - f(i) \mid x \in V_{1-i})$ consists of independent uniform signs.

Corollary 2.7. *If $d(n) - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then*

$$\mathbb{P}(\Omega_0 \cup \Omega_1) = \mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq 3) \xrightarrow[n \rightarrow \infty]{} 1 \quad \text{and} \quad \mathbb{P}(\Omega_0) = \mathbb{P}(\Omega_1) \xrightarrow[n \rightarrow \infty]{} 1/2.$$

Thus, a typical homomorphism in the supercritical regime is constant on either the even or odd vertices of $T_{n,d}$, with the two cases being equally likely. The effect induced by the parity of n in Corollary 2.2 does not appear here, as n is always assumed to be even in the case of the torus.

The subcritical regime. The subcritical regime is when $d(n) - \log_2 n \rightarrow -\infty$ (i.e. $n2^{-d(n)} \rightarrow \infty$) as $n \rightarrow \infty$. As before, the relatively small number of constraints allows a typical homomorphism to grow.

Theorem 2.8. *There exist absolute constants $C, c > 0$ such that for any positive even integer n and any positive integer d , we have*

$$3 + \lfloor c\sqrt{n2^{-d}} \rfloor - 2^{1-n/2} \leq \mathbb{E}[|\text{Rng}(f_{n,d})|] \leq C(\sqrt{n2^{-d}} + 1).$$

Moreover, for any $\epsilon > 0$ there exists a $\delta > 0$ such that for any positive even integer n and any positive integer d , we have

$$\mathbb{P}\left(|\text{Rng}(f_{n,d})| < 3 + \lfloor \delta\sqrt{n2^{-d}} \rfloor\right) \leq \epsilon + 2^{1-n/2}.$$

In particular, if $d(n) - \log_2 n \rightarrow -\infty$ as $n \rightarrow \infty$ then for any positive integer r , we have

$$\mathbb{P}(|\text{Rng}(f_{n,d(n)})| \leq r) \xrightarrow{n \rightarrow \infty} 0.$$

The critical regime. The critical regime is when $d(n) - \log_2 n \rightarrow -\log_2 \lambda$ (i.e. $n2^{-d(n)} \rightarrow \lambda$) as $n \rightarrow \infty$, for some $\lambda \in (0, \infty)$. As for the line, this choice of parameters leads to an interesting limiting behavior. In this case, the random homomorphism behaves similarly to a simple random walk *bridge* of length $2N$, where N is an independent random variable whose distribution is a type of biased Poisson distribution. The distribution of N here is biased differently from the case of the line. Specifically, N has the distribution of a Poisson random variable with parameter

$$\lambda' := \frac{\lambda}{4\sqrt{2}}, \tag{2.4}$$

conditioned to be equal to another such independent Poisson random variable.

Denote by $\nu(\lambda')$ the distribution of X conditioned on $X = Y$, where X and Y are independent $\text{Poisson}(\lambda')$ random variables. One may check that

$$\nu(\lambda')(k) = Z(\lambda')^{-1} \cdot \frac{(\lambda')^{2k}}{(k!)^2}, \quad k \geq 0, \tag{2.5}$$

where $Z(\lambda')$ is a normalizing constant.

For a positive even integer k , let $(B_i^k \mid 0 \leq i \leq k)$ denote a simple random walk bridge of length k (that is, a simple random walk conditioned on $B_k^k = 0$), and let $N(\lambda') \sim \nu(\lambda')$ be an independent random variable. Thus, $B^{2N(\lambda')}$ is obtained by first sampling $N(\lambda')$ and then sampling a simple random walk bridge of length $2N(\lambda')$. For a positive even integer k , denote $\text{Rng}(B^k) := \{B_i^k \mid 0 \leq i \leq k\}$.

Theorem 2.9. *If $n2^{-d(n)} \rightarrow \lambda$ as $n \rightarrow \infty$, for some $\lambda \in (0, \infty)$, then*

$$|\text{Rng}(f_{n,d(n)})| \xrightarrow{n \rightarrow \infty} |\text{Rng}(B^{2N(\lambda')})| + 2,$$

where λ' is defined by (2.4).

This theorem is closely related to Theorem 2.5. On the line, the range of a homomorphism in the critical regime is determined by a simple random walk whose length is a parity-biased Poisson random variable. Note that if we condition a simple random walk with a $\text{Poisson}(\mu)$ number of steps to end at its initial value, then the number of steps it takes has distribution $\nu(\mu/2)$. To see this, observe that the number of positive and negative steps of the random walk are independent $\text{Poisson}(\mu/2)$ random variables and we are conditioning that these variables are equal. The same phenomenon continues to hold if we start with a simple random walk taking a parity-biased $\text{Poisson}(\mu, \alpha)$ number of steps. Indeed, the number of steps must be even in order for the walk to end at its initial value, and a parity-biased $\text{Poisson}(\mu, \alpha)$ conditioned to be even is the same as a $\text{Poisson}(\mu)$ conditioned to be even.

2.3 Local limits on the line

In this section we present results for homomorphisms on the graph $P_{n,d}$, which was defined in (1.1), when d is constant and n tends to infinity.

Our first result gives an approximate count of the number of homomorphisms.

Theorem 2.10. *For any positive integer d there exists a constant $C(d) > 0$ such that*

$$|\text{Hom}(P_{n,d}, 0)| = C(d)\lambda(d)^{n/2}(1 + o(1)) \quad \text{as } n \rightarrow \infty,$$

where $\lambda(d)$ is the unique positive solution to the equation

$$\lambda^{d-1/2}(\lambda - 2) = 1.$$

Remark 2.11. The constant $\lambda(d)$ above satisfies

$$\lambda(d) = 2 + 2^{-d+1/2}(1 - o(1)) \quad \text{as } d \rightarrow \infty.$$

Our next result concerns the local limit of the homomorphism. This local limit lives on $P_{\infty,d}$, the limiting graph of $P_{n,d}$. Precisely, $P_{\infty,d}$, for $d \geq 1$, is the graph defined by

$$\begin{aligned} V(P_{\infty,d}) &:= \{0, 1, 2, \dots\}, \\ E(P_{\infty,d}) &:= \{(i, j) \mid |i - j| \in \{1, 3, \dots, 2d + 1\}\}. \end{aligned} \tag{2.6}$$

For a function g defined on a domain Ω and a set $A \subseteq \Omega$, we write $g|_A$ for the restriction of g to A .

Theorem 2.12 (Local Limit). *For any constant $d \geq 1$, there exists a distribution $\mu_{\infty,d}$ on $\text{Hom}(P_{\infty,d}, 0)$ such that the uniform distribution on $\text{Hom}(P_{n,d}, 0)$ converges to $\mu_{\infty,d}$ as $n \rightarrow \infty$, in the following sense. Let $f_{n,d}$ be a uniformly chosen homomorphism in $\text{Hom}(P_{n,d}, 0)$ and let $f_{\infty,d}$ be sampled from $\mu_{\infty,d}$. Then,*

$$\mathbb{P}(f_{n,d}|_{\{0,1,\dots,r\}} = f) \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(f_{\infty,d}|_{\{0,1,\dots,r\}} = f) \quad \text{for any } r \geq 0 \text{ and } f \in \mathbb{Z}^{\{0,1,\dots,r\}}.$$

Remark 2.13. The random homomorphism $f_{\infty,d}$ is described by an explicit Markov chain on $2d + 2$ states, as shown in Figure 12, through a process which decodes infinite words on the alphabet $\{a, b, A, B\}$ into homomorphisms on $P_{\infty,d}$. See Section 6 for details.

Policy on constants: In the rest of the paper we employ the following policy on constants. We write C, c, C', c' for positive absolute constants, whose values may change from line to line. Specifically, the values of C, C' may increase and the values of c, c' may decrease from line to line.

3 Preliminaries

We gather here a number of general tools which we require for our proof.

Lemma 3.1. *Let E and F be events in a discrete probability space and let $T: E \rightarrow \mathcal{P}(F)$ be a mapping, where $\mathcal{P}(A)$ denotes the collection of all subsets of A . For $f \in F$, define*

$$N(f) := \{e \in E \mid f \in T(e)\}.$$

If for some $p, q > 0$, we have

$$\begin{aligned} \mathbb{P}(T(e)) &\geq \mathbb{P}(e) \cdot p, \quad e \in E, \\ |N(f)| &\leq q, \quad f \in F, \end{aligned} \tag{3.1}$$

then

$$\mathbb{P}(E) \leq \mathbb{P}(F) \cdot \frac{q}{p}.$$

Proof. It is a simple matter to verify that

$$\sum_{e \in E} \mathbb{P}(T(e)) = \sum_{e \in E} \sum_{f \in F} \mathbb{P}(f) \mathbf{1}_{T(e)}(f) = \sum_{f \in F} \sum_{e \in E} \mathbb{P}(f) \mathbf{1}_{T(e)}(f) = \sum_{f \in F} \mathbb{P}(f) \cdot |N(f)|.$$

The result now follows by the assumptions in (3.1). □

Remark 3.2. *The opposite inequalities in (3.1) would yield the analogous result. Namely, if $\mathbb{P}(T(e)) \leq p\mathbb{P}(e)$ for all $e \in E$ and $|N(f)| \geq q$ for all $f \in F$, then $p\mathbb{P}(E) \geq q\mathbb{P}(F)$. Note that when applying this lemma for the uniform distribution, the assumptions in (3.1) become $|T(e)| \geq p$ for all $e \in E$ and $|N(f)| \leq q$ for all $f \in F$, while the conclusion remains the same.*

Lemma 3.3. *Let X be a non-negative, integer-valued random variable. Assume that, for some positive integer n and some $a > 1$, we have $\mathbb{P}(X = k) \geq a \cdot \mathbb{P}(X = k - 1)$, for all $1 \leq k \leq n$. Then*

$$\mathbb{P}(X < n) \leq 1/a.$$

Proof. It is easy to verify (by induction) that

$$\mathbb{P}(X = n) \geq a^k \cdot \mathbb{P}(X = n - k), \quad 1 \leq k \leq n.$$

Thus,

$$\mathbb{P}(X < n) = \sum_{k=0}^{n-1} \mathbb{P}(X = k) = \sum_{k=1}^n \mathbb{P}(X = n - k) \leq \mathbb{P}(X = n) \cdot \sum_{k=1}^n a^{-k} \leq \frac{\mathbb{P}(X = n)}{a - 1}.$$

Therefore,

$$1 \geq \mathbb{P}(X \leq n) = \mathbb{P}(X < n) + \mathbb{P}(X = n) \geq \mathbb{P}(X < n)(1 + (a - 1)) = a \cdot \mathbb{P}(X < n). \quad \square$$

We will use a theorem by Benjamini, Häggström and Mossel [1] to transfer results from the line to the torus. This is an FKG inequality for the measure induced on non-negative homomorphisms by taking pointwise absolute value.

Given a set V , we equip \mathbb{Z}^V with the usual pointwise partial order \preceq , i.e. $f \preceq g$ if $f(v) \leq g(v)$ for all $v \in V$. A function $\phi: \mathbb{Z}^V \rightarrow \mathbb{R}$ is said to be *increasing* if $\phi(f) \leq \phi(g)$ whenever $f \preceq g$.

Theorem 3.4 (FKG inequality for absolute values [1, Proposition 2.3]). *Let G be a finite, bipartite and connected graph, let $v_0 \in V(G)$ and let f be a uniformly chosen homomorphism in $\text{Hom}(G, v_0)$. Then, for any two increasing functions $\phi, \psi: \text{Hom}(G, v_0) \rightarrow \mathbb{R}$, we have*

$$\mathbb{E}[\phi(|f|) \cdot \psi(|f|)] \geq \mathbb{E}[\phi(|f|)] \cdot \mathbb{E}[\psi(|f|)],$$

where $|f|$ is the non-negative homomorphism obtained from f by taking pointwise absolute value.

Consider the event Q that a homomorphism on $P_{n,d}$ is in fact a valid homomorphism on $T_{n,d}$ (by identifying the vertex n with the vertex 0). If we could write $\mathbf{1}_Q(f) = \psi(|f|)$ for some function ψ then we may be able to use the above theorem to transfer results from the line to the torus by conditioning on Q . However, it is not the case that $\mathbf{1}_Q$ is a function of the absolute value of the homomorphism, and so we cannot apply Theorem 3.4 directly. Instead, we make use of Theorem 3.4 in order to prove a similar proposition specialized for our purposes. See Proposition 5.9 in Section 5 for more details.

The following result of Erdős is useful for analyzing homomorphism on the line.

Theorem 3.5 ([6, Theorem 1]). *Let n be a positive integer, let $a_1, \dots, a_n \in \mathbb{R}$ satisfy $|a_i| \geq 1$ for $1 \leq i \leq n$ and let $\epsilon_1, \dots, \epsilon_n \sim U(\{-1, 1\})$ be random independent signs. Denote*

$$S := \epsilon_1 a_1 + \dots + \epsilon_n a_n.$$

Then, for any integer $r > 0$ and any $a \in \mathbb{R}$, we have

$$\mathbb{P}(|S - a| < r) \leq r \cdot \binom{n}{\lfloor n/2 \rfloor} \cdot 2^{-n} \leq \frac{Cr}{\sqrt{n}}.$$

The next proposition, which is a consequence of the previous result, is useful for analyzing homomorphisms on the torus.

Proposition 3.6. *Let n be a positive integer, let $a_1, \dots, a_n \in \mathbb{R}$ satisfy $|a_i| \geq 1$ for $1 \leq i \leq n$ and let π be a uniformly chosen permutation of $\{1, 2, \dots, n\}$. Denote*

$$S_i := a_{\pi(1)} + \dots + a_{\pi(i)}, \quad 1 \leq i \leq n.$$

Then, for any integer $r > 0$, we have

$$\mathbb{P} \left(\max_{1 \leq i \leq n} |S_i| < r \right) \leq \frac{Cr}{\sqrt{n}}.$$

Proof. Let $\epsilon_1, \dots, \epsilon_n \sim U(\{-1, 1\})$ be uniform independent signs. Denote

$$\begin{aligned} a &:= a_1 + \dots + a_n, \\ S' &:= \sum_{i=1}^n \frac{\epsilon_i + 1}{2} a_i = \frac{\epsilon_1 a_1 + \dots + \epsilon_n a_n}{2} + \frac{a}{2}. \end{aligned}$$

Let $T \sim \text{Bin}(n, 1/2)$ be independent of π and observe that

$$S_T \stackrel{d}{=} S',$$

an observation which was pointed out to us by Gady Kozma. Therefore, by Theorem 3.5,

$$\mathbb{P} \left(\max_{1 \leq i \leq n} |S_i| < r \right) \leq \mathbb{P} (|S_T| < r) = \mathbb{P} (|S'| < r) = \mathbb{P} \left(\left| \sum_{i=1}^n \epsilon_i a_i + a \right| < 2r \right) \leq \frac{Cr}{\sqrt{n}}. \quad \square$$

The next lemma presents a simple result on limits of distributions.

Lemma 3.7. *Let $X_\infty, X_1, X_2, \dots$ be non-negative, integer-valued random variables. Assume that $\mathbb{P}(X_\infty = k) > 0$ for all integers $k \geq 0$. If the family $\{X_1, X_2, \dots\}$ is tight, and*

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(X_n = k)}{\mathbb{P}(X_n = k - 1)} = \frac{\mathbb{P}(X_\infty = k)}{\mathbb{P}(X_\infty = k - 1)}, \quad k \geq 1, \quad (3.2)$$

then

$$X_n \xrightarrow[n \rightarrow \infty]{} X_\infty.$$

Proof. Let $\epsilon > 0$ and, using the tightness assumption, choose an integer M such that $\mathbb{P}(X_n > M) \leq \epsilon$ for all $n \in \mathbb{N} \cup \{\infty\}$. Then

$$1 - \epsilon \leq \sum_{k=0}^M \mathbb{P}(X_n = k) \leq 1, \quad n \in \mathbb{N} \cup \{\infty\}.$$

Therefore, by the assumption (3.2),

$$\limsup_{n \rightarrow \infty} \frac{1 - \epsilon}{\mathbb{P}(X_n = 0)} \leq \lim_{n \rightarrow \infty} \sum_{k=0}^M \frac{\mathbb{P}(X_n = k)}{\mathbb{P}(X_n = 0)} = \sum_{k=0}^M \frac{\mathbb{P}(X_\infty = k)}{\mathbb{P}(X_\infty = 0)} \leq \frac{1}{\mathbb{P}(X_\infty = 0)}.$$

Similarly, we have

$$\liminf_{n \rightarrow \infty} \frac{1}{\mathbb{P}(X_n = 0)} \geq \lim_{n \rightarrow \infty} \sum_{k=0}^M \frac{\mathbb{P}(X_n = k)}{\mathbb{P}(X_n = 0)} = \sum_{k=0}^M \frac{\mathbb{P}(X_\infty = k)}{\mathbb{P}(X_\infty = 0)} \geq \frac{1 - \epsilon}{\mathbb{P}(X_\infty = 0)}.$$

Since ϵ is arbitrary, we conclude that $\mathbb{P}(X_n = 0) \rightarrow \mathbb{P}(X_\infty = 0)$ as $n \rightarrow \infty$, which in turn gives that $\mathbb{P}(X_n = k) \rightarrow \mathbb{P}(X_\infty = k)$ as $n \rightarrow \infty$ for all k , by (3.2). \square

4 Homomorphisms on the Line

In this section we will prove the theorems regarding homomorphisms on the line which were stated in Section 2.1. As pointed out in the introduction, it seems unlikely that a homomorphism jumps from some value t to $t \pm 3$ on any given small segment. Figure 2 illustrates a section of a homomorphism for which such a jump occurs. The main idea in our proofs is to identify the vertices at which these jumps occur, as they determine the large scale behavior of the homomorphism. That is, the values of the homomorphism at the jumps contain the global information necessary to determine the range and the variance. To this end, we first define the notion of the (local) average height of a homomorphism at a vertex (this is illustrated by the horizontal dashed line in Figure 3). The average height at a vertex is determined by finding the closest past time at which 3 different values appeared consecutively and taking the midpoint to be the average height. At vertices for which no such time exists (as is the case for the 0 vertex), we set the average height to be 0. One can think of the average height as a process beginning at 0 that “lazily follows” the homomorphism, only to ensure that it is never at a distance greater than 1. With this notion in hand, we define a jump as a change in the average height. Of course any such jump has an associated sign or direction, which is determined by whether the average height increases or decreases.

We later show that the probability of a jump occurring at a given vertex (greater than $2d$) is no more than 2^{-d} . This will show that in the supercritical regime, with high probability, there will not be any jumps (after vertex $2d$). That is, the average height does not change after time $2d$. A moment’s reflection reveals that this means that the homomorphism takes on at most 3 different values (not 4, as it may initially seem).

We do not give a lower bound for the probability of a jump occurring at a given vertex. Instead, we only show that the typical number of jumps is of order $n2^{-d}$, the jumps are approximately equidistributed on the line and that, moreover, the directions of these jumps are weakly correlated. Of course, if the directions of these jumps were truly independent, then the values of the homomorphism at the jumps would constitute a simple random walk. We will show that, at least in terms of the maximum/range of the homomorphism, the behavior is very similar to that of a simple random walk. This will show that the range is of order $\sqrt{n2^{-d}}$ and that the variance at a vertex k is of order $k2^{-d}$.

In the analysis of these so-called jumps, we encounter a minor complication due to the fact that jumps in the same direction can “clump” together. Of course jumps cannot occur consecutively in the sense of two consecutive vertices on the line. So then what is the minimal distance between two jumps? The answer is twofold. The minimal distance between two jumps with different directions is $2d + 3$, while two jumps in the same direction can already occur at distance $2d + 1$. This phenomenon will pop up again and again in our analysis. For example, its manifestation is evident in the Markov chain describing the local limit in Section 6 (see Figure 12).

One meaning of this phenomenon is that if we condition on the event that a jump occurs at two given vertices, say k_1 and k_2 , $k_1 < k_2$, the directions of these jumps are non-negatively correlated. However, conditioning also on the event that a jump does not occur just after the first of these jumps (i.e. at $k_1 + 2d + 1$), their directions become independent. This leads us to consider “chains” of jumps. A chain is just a sequence of minimal-distance same-direction jumps. Now, if we condition on the event that there are chains of given lengths (and not longer) at any number of given vertices, the directions of all these chains will be independent. This will allow us to reduce some of the analysis to a case of independent variables.

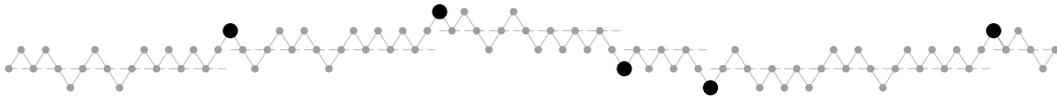


Figure 3: A homomorphism in $\text{Hom}(P_{n,d})$. The big vertices denote the positions of the jumps. The dashed horizontal lines denote the average height. Here $d = 3$.

4.1 Definitions

We consider the graph $P_{n,d}$ whose vertex set is $\{0, 1, \dots, n\}$ and whose edges are (k, m) for $|k - m| = 1, 3, \dots, 2d + 1$. Throughout this section, $\text{Hom}(P_{n,d}) := \text{Hom}(P_{n,d}, 0)$, f is a uniformly sampled homomorphism from $\text{Hom}(P_{n,d})$, the probability space is the uniform distribution on the set $\text{Hom}(P_{n,d})$, and events are subsets of $\text{Hom}(P_{n,d})$.

We define $h(k)$, the (local) average height at vertex k , inductively as follows. Set $h(0) := 0$. For $1 \leq k \leq n$, define

$$h(k) := \begin{cases} h(k-1) & \text{if } |f(k) - h(k-1)| \leq 1 \\ f(k-1) & \text{otherwise} \end{cases},$$

$$\Delta(k) := h(k) - h(k-1).$$

Define the event

$$A_k := \{\Delta(k) \neq 0\}.$$

When A_k occurs, we say that a jump occurred at vertex k (see Figure 3). Note that if k is the position of the first jump then $|f(k)| = 2$, so that, in particular, k is even. Let

$$S := \{1 \leq k \leq n \mid \Delta(k) \neq 0\} \tag{4.1}$$

be the positions of the jumps, and denote by

$$R := |S \setminus \{1, \dots, 2d + 1\}| = \sum_{k=2d+2}^n \mathbf{1}_{A_k} \tag{4.2}$$

the number of jumps after vertex $2d + 1$. Recall that if a jump occurs at vertex k , then the minimal possible value of $k' > k$ at which another jump can occur is $k + 2d + 1$. Let $C_{k,t}$ be the event that there is a chain of t minimal-distance jumps ending at vertex k . That is, for $t \geq 1$ and $(t - 1)(2d + 1) < k \leq n$, we define

$$C_{k,t} := A_k \cap A_{k-2d-1} \cap \dots \cap A_{k-(t-1)(2d+1)}.$$

Let $I = \{s_1, \dots, s_t\} \subset \{1, 2, \dots, n\}$. We say that I is a feasible jump structure if $\{S = I\} \neq \emptyset$. Observe that $\{S = I\} \neq \emptyset$ if and only if $\mathbb{P}(S = I) > 0$ if and only if when we reorder the s_i to satisfy $s_1 < s_2 < \dots < s_t$, we have

$$s_1 \text{ is even and for } 2 \leq j \leq t, s_j - s_{j-1} \text{ is odd and satisfies } s_j - s_{j-1} \geq 2d + 1. \tag{4.3}$$

In addition, we say that a subset $I \subset \{1, \dots, n\}$ is a feasible jump sub-structure if it is a subset of a feasible jump structure, or equivalently, if $\{I \subset S\} \neq \emptyset$. For a feasible jump sub-structure I , the event $\{I \subset S\}$ can be uniquely written as $C_{k_1, t_1} \cap \dots \cap C_{k_m, t_m}$, where

$$t_1 + \dots + t_m = |I|,$$

$$k_j - k_{j-1} > (2d + 1)t_j, \quad 2 \leq j \leq m. \tag{4.4}$$

These conditions ensure that there is no overlap between the different chains, and moreover, that there is some gap between them (since otherwise they would merge into a larger chain). For such I , we define

$$\mathcal{C}(I) := \{(k_j, t_j) \mid 1 \leq j \leq m\},$$

and refer to this as the *chain structure* of I .

4.2 Main lemmas

As the above definitions suggest, the notion of a jump at a given vertex plays an important role in our analysis. It turns out that the behavior of jumps at the first $2d + 1$ vertices differs significantly from that of the other vertices. Hence, it will be a recurring theme throughout Section 4 that these cases are handled separately.

The first two lemmas concern the probability of jumps at given vertices. The first of which shows that jumps at the first $2d + 1$ vertices are not unlikely, while the second shows that elsewhere jumps are unlikely.

Lemma 4.1. *We have*

$$1/4 \leq \mathbb{P}(A_2) \leq 1/2$$

and

$$1/3 \leq \mathbb{P}(A_1 \cup \dots \cup A_{2d+1}) \leq 2/3.$$

Proof. Denote $J := A_1 \cup \dots \cup A_{2d+1}$. We shall show that

$$\mathbb{P}(A_2) \leq \mathbb{P}(J) \leq 2\mathbb{P}(A_2),$$

$$\mathbb{P}(A_2) \leq \mathbb{P}(J^c) \leq 2\mathbb{P}(A_2),$$

from which the result easily follows. Note that, by (4.3), $A_k = \emptyset$ for $k = 1, 3, \dots, 2d + 1$, so that $J = A_2 \cup A_4 \cup \dots \cup A_{2d}$. Clearly $\mathbb{P}(A_2) \leq \mathbb{P}(J)$, as $A_2 \subset J$.

We note the following useful observation. For a homomorphism $f \in \text{Hom}(P_{n,d})$, we have

$$f \in A_2 \iff f(2) = 2f(1) \iff f(2) \neq 0. \tag{4.5}$$

We begin by showing that $\mathbb{P}(J \setminus A_2) \leq \mathbb{P}(A_2)$, from which it follows that $\mathbb{P}(J) \leq 2\mathbb{P}(A_2)$. To this end it suffices to show an injective mapping from $J \setminus A_2$ to A_2 . Consider the mapping $f \mapsto f_1$ from $J \setminus A_2$ to A_2 defined by

$$f_1(k) := \begin{cases} f(k) & \text{if } k \neq 2 \\ 2f(1) & \text{if } k = 2 \end{cases}, \quad 0 \leq k \leq n.$$

One may check that if $f \in J \setminus A_2$ then $f(1) = f(3) = \dots = f(2d + 3)$, so that $f_1 \in A_2$. Recalling (4.5), it is clear that this mapping is invertible, and so we have $\mathbb{P}(J \setminus A_2) \leq \mathbb{P}(A_2)$.

We now show that $\mathbb{P}(A_2) \leq \mathbb{P}(J^c)$. Define a mapping $f \mapsto f_2$ from A_2 to J^c by

$$f_2(k) := \begin{cases} f(k+1) - f(1) & \text{if } 0 \leq k < n \\ f(n-1) - f(1) & \text{if } k = n \end{cases}, \quad 0 \leq k \leq n.$$

Again one may check that this mapping is well-defined (in fact, this mapping can be defined on the entire space). Since it is injective (recall (4.5)), we obtain $\mathbb{P}(A_2) \leq \mathbb{P}(J^c)$.

Finally, we show that $\mathbb{P}(J^c) \leq 2\mathbb{P}(A_2)$. Consider the mapping $T: J^c \rightarrow A_2$ defined by

$$T(f)(k) := \begin{cases} 0 & \text{if } k = 0 \\ f(k-1) + f(1) & \text{if } 1 \leq k \leq n \end{cases}, \quad 0 \leq k \leq n.$$

To see that this mapping is well-defined, recall (4.5), and note that $f \in J^c$ implies that $f(0) = f(2) = \dots = f(2d) = 0$. This is not an injective mapping, however, it satisfies $|T^{-1}(g)| \leq 2$ for $g \in A_2$. Therefore, by Lemma 3.1, we have $\mathbb{P}(J^c) \leq 2\mathbb{P}(A_2)$. \square

The next lemma is concerned with the probability of jumps occurring at given vertices after $2d+1$. It states that this probability is exponentially small in d times the number of jumps. The idea behind the proof is to remove the jumps and replace the freed up areas with segments of constant average height. This allows us to gain entropy by setting the values at every other vertex in each such segment to be the average height ± 1 . See Figure 4.

Lemma 4.2. *For any $t \geq 1$ and for any $2d+1 < s_1 < \dots < s_t \leq n$, we have*

$$\mathbb{P}(A_{s_1} \cap \dots \cap A_{s_t}) \leq 2^{-dt}.$$

Proof. If $I := \{s_1, \dots, s_t\}$ is not a feasible jump sub-structure then there is nothing to prove. Otherwise, we consider the chain structure of I , $\mathcal{C}(I) = \{(k_1, t_1), \dots, (k_m, t_m)\}$, where we have ordered the elements so that the k_j are increasing. Due to our assumption that $s_1 > 2d+1$, we have $k_1 > (2d+1)t_1$. We note that it is enough to prove that for all $1 \leq j \leq m$,

$$\mathbb{P}(C_{k_j, t_j} \mid C_{k_1, t_1} \cap \dots \cap C_{k_{j-1}, t_{j-1}}) \leq 2^{-dt_j}.$$

We prove something stronger. Let $1 \leq k \leq n$ and $t \geq 1$ be such that $k' := k - (2d+1)t - 1 \geq 0$. Then, for any $\xi \in \mathbb{Z}^{\{0, \dots, k'\}}$ such that $\mathbb{P}(f|_{\{0, \dots, k'\}} = \xi) > 0$, we have

$$\mathbb{P}(C_{k, t} \mid f|_{\{0, \dots, k'\}} = \xi) \leq 2^{-dt - \lfloor t/2 \rfloor}. \tag{4.6}$$

In order to show this, we construct a mapping which removes this chain and replaces the freed up segment with a segment of constant average height (see Figure 4). Formally, we proceed as follows. For $m \geq 1$ and $w \in \{-1, 1\}^m$, denote

$$\tilde{w} := (w_1, 0, w_2, 0, \dots, w_m, 0).$$

Define a mapping

$$T_{k, t}: C_{k, t} \times \{-1, 1\}^{dt + \lfloor t/2 \rfloor} \rightarrow \text{Hom}(P_{n, d})$$

by

$$T_{k, t}(f, w)(i) := \begin{cases} f(i) & \text{if } i \leq k' \\ f(k') + \tilde{w}(i - k') & \text{if } k' < i < k - 1 \\ f(i + \delta) - f(k - 1) + f(k') & \text{if } k - 1 \leq i < n \\ f(i - \delta) - f(k - 1) + f(k') & \text{if } i = n \end{cases}, \tag{4.7}$$

where $\delta = 0$ if t is even and $\delta = 1$ if t is odd.

We now show that $T_{k, t}$ is well-defined, i.e. that $T_{k, t}(f, w) \in \text{Hom}(P_{n, d})$. For $0 \leq i, j \leq n$, denote

$$\Delta_{i, j} := |T_{k, t}(f, w)(i) - T_{k, t}(f, w)(j)|.$$

For $0 \leq j \leq n$, define the event

$$B_j := \{f(i) = f(j) \text{ when } 0 \leq i \leq n \text{ satisfies } |i - j| \in \{2, 4, \dots, 2d\}\}.$$

We must show that $\Delta_{i, j} = 1$ whenever $|i - j| \in \{1, 3, \dots, 2d+1\}$. Let $0 \leq i, j \leq n$ satisfy $|i - j| \in \{1, 3, \dots, 2d+1\}$ and assume without loss of generality that $i < j$. We shall further assume that $j < n$, the case $j = n$ being similar. If $j \leq k'$, $i \geq k - 1$ or $k' < i < j < k - 1$ then $\Delta_{i, j} = 1$ follows immediately from (4.7) and the fact that

Random walk with long-range constraints

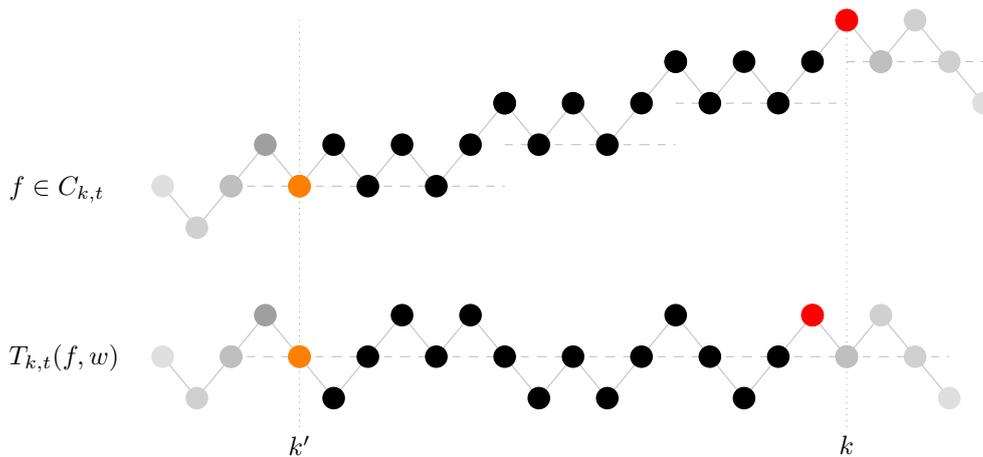


Figure 4: A section of a homomorphism f in the event $C_{k,t}$. Removing the chain between k' and k , and replacing it with fluctuations around the average height, gives the homomorphism $T_{k,t}(f, w)$. The dashed horizontal lines denote the average height. Here $d = 2$, $t = 3$ and $w = (-1, 1, 1, -1, -1, 1, -1)$.

$f \in \text{Hom}(P_{n,d})$. It remains to check the case when $i \leq k'$ and $j > k'$ and the case when $i < k - 1$ and $j \geq k - 1$.

We begin with the first case. Here, we have $\Delta_{i,j} = |f(i) - f(k') - \tilde{w}(j - k')|$. Observe that $C_{k,t} \subset B_{k'}$. Therefore, if i has the same parity as k' then $f(i) = f(k')$ and $|\tilde{w}(j - k')| = 1$ since j has the same parity as k' . Otherwise, i has the opposite parity of k' , and then $|f(i) - f(k')| = 1$ and $\tilde{w}(j - k') = 0$. Thus, $\Delta_{i,j} = 1$.

In the second case, we have $\Delta_{i,j} = |f(j + \delta) - f(k - 1) - \tilde{w}(i - k')|$. Note that $C_{k,t} \subset B_{k-1}$ and that $i - k'$ has the same parity as $j + \delta - k$. One finds in a similar manner as in the first case that $|f(j + \delta) - f(k - 1)| = 1$ and $\tilde{w}(i - k') = 0$ when i has the same parity as k' , and that $f(j + \delta) = f(k - 1)$ and $|\tilde{w}(i - k')| = 1$ when i has the opposite parity of k' . Hence, $\Delta_{i,j} = 1$.

Observe that for any $f \in C_{k,t}$, necessarily,

$$\begin{aligned} & (f(k'), f(k'+1), \dots, f(k), f(k+1)) \\ &= (f(k'), \dots, f(k')) \pm \underbrace{(0, 1, \dots, 0, 1, 0, 1, \dots, 1, 2, \dots, 1, 2, 1, \dots, t-1, t, \dots, t-1, t, t-1, t, t+1, t)}_{2d+1} \end{aligned}$$

Thus, it is easy to see that the mapping is injective. Moreover, the event $\{f|_{\{0, \dots, k'\}} = \xi\}$ is clearly invariant under this mapping, so that

$$\mathbb{P}(C_{k,t} \cap \{f|_{\{0, \dots, k'\}} = \xi\}) \cdot 2^{dt + \lfloor t/2 \rfloor} \leq \mathbb{P}(f|_{\{0, \dots, k'\}} = \xi),$$

proving (4.6). □

Remark 4.3. *The proof shows in fact that the probability of the event $A_{s_1} \cap \dots \cap A_{s_t}$ is bounded by $2^{-dt - (\lfloor t_1/2 \rfloor + \dots + \lfloor t_r/2 \rfloor)}$, where t_1, \dots, t_r are the lengths of the chains corresponding to s_1, \dots, s_t . With a small modification, the proof can be enhanced to give the bound $2^{-dt - \lfloor t/2 \rfloor}$, but we neither prove nor use this.*

Recall the definition of R from (4.2). We would like to obtain inequalities on the probability that R is a given value. We could do this in a similar manner to which the previous lemma was proved. However, for variety, we prefer to employ a more direct combinatorial technique. This approach also has the advantage of introducing

Random walk with long-range constraints

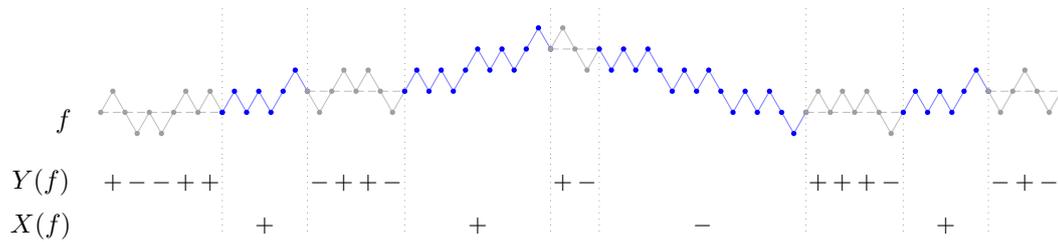


Figure 5: A homomorphism is decomposed into two parts; chains (in blue) and fluctuations (in gray). The chains, which consist of consecutive jumps, contribute to the change in average height, while the fluctuations do not. From the chains, we construct $X(f)$, which contains the direction of each chain. From the fluctuations, we construct $Y(f)$, which contains the direction of each fluctuation. Given the positions of the jumps, $X(f)$ and $Y(f)$ precisely contain the remaining information on the homomorphism. That is, for any choice of X and Y , there exists a unique homomorphism f with $X(f) = X$ and $Y(f) = Y$. Here $d = 2$.

Lemma 4.5, which gives a useful description of the structure of the homomorphisms in $\text{Hom}(P_{n,d})$.

We decompose a homomorphism into two parts (see Figure 5). The first part constitutes the changes in average height (the underlying walk) of the homomorphism, while the second part constitutes the fluctuations around the average height (the segments of constant average height). For a feasible jump sub-structure I , define the *chain points* of I by

$$CP(I) := \bigcup_{(k,t) \in \mathcal{C}(I)} \{k - (2d + 1)t - 1, \dots, k - 1, k\},$$

and the *fluctuation points* of I by

$$FP(I) := \left\{ 1 \leq k \leq n \mid \min \{i > 0 \mid k - i \in CP(I) \cup \{-1\}\} \text{ is even} \right\}.$$

That is, a point is a fluctuation point if its distance from the chain to its left is positive and even. In particular, recalling the definition of S from (4.1), for any homomorphism f and any $k \in FP(S(f))$, f is not at its average height at k . Now, for a homomorphism f , define

$$\begin{aligned} X(f) &\in \{-1, 1\}^{\mathcal{C}(S(f))} \quad \text{and} \\ Y(f) &\in \{-1, 1\}^{FP(S(f))} \end{aligned}$$

by

$$\begin{aligned} X(f)_{(k,t)} &:= f(k) - f(k - 1) \quad \text{and} \\ Y(f)_k &:= f(k) - f(k - 1). \end{aligned} \tag{4.8}$$

Claim 4.4. *For any feasible jump structure I , we have*

$$|FP(I)| = \max \left\{ 0, d + 1 - \frac{\min I}{2} \right\} + \left\lceil \frac{n - |I|}{2} \right\rceil - d|I| - |\mathcal{C}(I)|. \tag{4.9}$$

Proof. Suppose that $\mathcal{C}(I) = \{(k_1, t_1), \dots, (k_m, t_m)\}$. Denote $t := |I| = t_1 + \dots + t_m$ and $s := \min I$. Then

$$|CP(I)| = \sum_{j=1}^m ((2d+1)t_j + 2) = (2d+1)t + 2m,$$

$$|CP(I) \cap \{0, 1, \dots, n\}| = (2d+1)t + 2m - \max\{0, 2d+2-s\}.$$

Therefore, recalling that s is even by (4.3),

$$\begin{aligned} |FP(I)| &= \left\lceil \frac{n - |CP(I) \cap \{0, 1, \dots, n\}|}{2} \right\rceil \\ &= \left\lceil \frac{n - (2d+1)t - 2m + \max\{0, 2d+2-s\}}{2} \right\rceil \\ &= \left\lceil \frac{n-t}{2} \right\rceil - dt - m + \max\{0, d+1-s/2\}. \quad \square \end{aligned}$$

Lemma 4.5. *For any feasible jump structure I , the mapping $f \mapsto (X(f), Y(f))$ is a bijection between $\{S = I\}$ and $\{-1, 1\}^{\mathcal{C}(I)} \times \{-1, 1\}^{FP(I)}$.*

Proof. We shall describe the inverse mapping which maps a pair $(X, Y) \in \{-1, 1\}^{\mathcal{C}(I)} \times \{-1, 1\}^{FP(I)}$ to the homomorphism $f_{X,Y} \in \{S = I\}$. For $0 \leq i \leq n$, let

$$H(X, i) := \sum_{\substack{(k,t) \in \mathcal{C}(I) \\ k < i}} t \cdot X_{(k,t)},$$

be the average height accumulated by chains ending before i . For $(k, t) \in \mathcal{C}(I)$, denote by $k'(k, t) := k - (2d+1)t - 1$ the first vertex of the chain and observe that $H(X, i) = H(X, k'(k, t))$ for all $k'(k, t) \leq i \leq k$. Now, define

$$f_{X,Y}(i) := \begin{cases} H(X, i) + Y_i & \text{if } i \in FP(I) \\ H(X, i) & \text{if } i+1 \in FP(I) \text{ or } i-1 \in FP(I) \\ H(X, i) + X_{(k,t)} C_{i-k'(k,t)} & \text{if } k'(k, t) \leq i \leq k \text{ for some } (k, t) \in \mathcal{C}(I) \end{cases},$$

where $C = (C_0, C_1, \dots)$ is the infinite sequence defined by

$$C := (\underbrace{0, 1, \dots, 0, 1, 0}_{2d+1}, \underbrace{1, 2, \dots, 1, 2, 1}_{2d+1}, \underbrace{2, 3, \dots, 2, 3, 2}_{2d+1}, \dots).$$

See Figure 5. It is not difficult to check that $f_{X,Y} \in \text{Hom}(P_{n,d})$ and that $S(f_{X,Y}) = I$. It remains to check that $X(f_{X,Y}) = X$, $Y(f_{X,Y}) = Y$ and $f_{X(f),Y(f)} = f$. We omit the details. \square

Corollary 4.6. *Conditioned on S , the following properties hold.*

1. X is uniformly distributed over $\{-1, 1\}^{\mathcal{C}(S)}$.
2. The random variables $\{\Delta(s)\}_{s \in S}$ are independent uniform signs conditioned on $\Delta(s) = \Delta(s')$ whenever $s, s' \in S$ satisfy $|s - s'| = 2d + 1$.
3. The difference in average height between two vertices $0 \leq k_0 < k_1 \leq n$ is a sum of independent variables, namely,

$$h(k_1) - h(k_0) = \sum_{(j,t) \in \mathcal{C}(S \cap \{k_0+1, \dots, k_1\})} t \cdot \Delta(j). \quad (4.10)$$

Proof. The first statement is an immediate consequence of Lemma 4.5. The second statement is in turn a consequence of the first statement and of the definition of the chain structure $\mathcal{C}(S)$. For the third statement, since

$$h(k) = \sum_{i=1}^k \Delta(i) = \sum_{\substack{s \in S \\ s \leq k}} \Delta(s), \quad 1 \leq k \leq n,$$

we see that

$$h(k_1) - h(k_0) = \sum_{\substack{s \in S \\ k_0 < s \leq k_1}} \Delta(s) = \sum_{(j,t) \in \mathcal{C}(S \cap \{k_0+1, \dots, k_1\})} t \cdot \Delta(j). \quad \square$$

Corollary 4.7. *Conditioned on $|S|$ and $\min(S \cup \{2d+2\})$, S is uniformly distributed over all feasible jump structures I having $|I| = |S|$ and $\min(I \cup \{2d+2\}) = \min(S \cup \{2d+2\})$.*

Proof. By Lemma 4.5 and Claim 4.4, $\log_2 |\{S = I\}| = |\mathcal{C}(I)| + |FP(I)|$ depends only on $|I|$ and $\min(I \cup \{2d+2\})$. \square

For $r \geq 0$ and $1 \leq i \leq d+1$, denote by $c_i(r)$ the number of feasible jump structures I having $|I \setminus \{1, \dots, 2d+1\}| = r$ and $\min(I \cup \{2d+2\}) = 2i$ (recalling from (4.3) that $\min I$ is even).

Claim 4.8. *For any non-negative integer r , we have*

$$\begin{aligned} c_{d+1}(r) &= \binom{\lfloor (n-r-1)/2 \rfloor - (d-1)r}{r} \quad \text{and} \\ c_i(r) &= \binom{\lfloor (n-r)/2 \rfloor - (d-1)r - i}{r}, \quad 1 \leq i \leq d. \end{aligned} \tag{4.11}$$

Proof. By considering the distance between two consecutive values in I and recalling (4.3), we see that the number of feasible jump structures I having $|I| = r$ and $\min I > 2d+1$ (where we set $\min \emptyset := \infty$) is given by the number of non-negative integer solutions to the equation

$$x_1 + x_2 + \dots + x_r \leq n$$

under the additional constraints that x_1 is even and at least $2d+2$ and, for $2 \leq j \leq r$, x_j is odd and at least $2d+1$. Therefore, after substituting $x_1 = 2y_1 + 2d+2$ and $x_j = 2y_j + 2d+1$ for $2 \leq j \leq r$, we see that $c_{d+1}(r)$ is equal to the number of non-negative integer solutions to the equation

$$2(y_1 + \dots + y_r) \leq n - (2d+1)r - 1,$$

from which the first result easily follows. Similarly, the number of feasible jump structures I having $|I| = r+1$ and $\min I = 2i$ is given by the number of non-negative integer solutions to the equation

$$2i + x_1 + \dots + x_r \leq n$$

under the additional constraint that, for $1 \leq j \leq r$, x_j is odd and at least $2d+1$. Therefore, substituting $x_j = 2y_j + 2d+1$ as before, we see that, for $1 \leq i \leq d$, $c_i(r)$ is equal to the number of non-negative integer solutions to the equation

$$2(y_1 + \dots + y_r) \leq n - (2d+1)r - 2i,$$

from which the second result follows. \square

The next lemma and its corollary give bounds on the distribution of R . Observe that, by Claim 4.8, for any $1 \leq i \leq d+1$, $\mathbb{P}(R = r \mid \min(S \cup \{2d+2\}) = 2i) > 0$ when r satisfies $(2d+1)r + 2d \leq n$.

Lemma 4.9. *For any positive integer r such that $(2d+1)r + 2d \leq n$, we have*

$$\frac{n - Crd}{4r2^d} \leq \frac{\mathbb{P}(R = r \mid \min(S \cup \{2d+2\}) = 2i)}{\mathbb{P}(R = r - 1 \mid \min(S \cup \{2d+2\}) = 2i)} \leq \frac{n}{2r2^d}, \quad 1 \leq i \leq d+1.$$

Proof. By Lemma 4.5, Claim 4.4 and Claim 4.8, we have

$$|\{R = r\} \cap \{\min(S \cup \{2d+2\}) = 2i\}| = c_i(r)b_i(r), \quad 1 \leq i \leq d+1,$$

where $c_i(r)$ is given by (4.11) and

$$\begin{aligned} b_{d+1}(r) &= 2^{\lceil (n-r)/2 \rceil - dr}, \\ b_i(r) &= 2^{\lceil (n-r-1)/2 \rceil - dr - i + 1}, \quad 1 \leq i \leq d. \end{aligned} \tag{4.12}$$

It is easy to see that

$$2^{-d-1} \leq \frac{b_i(r)}{b_i(r-1)} \leq 2^{-d}, \quad 1 \leq i \leq d+1, \tag{4.13}$$

and a computation shows that

$$\frac{n - Crd}{2r} \leq \frac{c_i(r)}{c_i(r-1)} \leq \frac{n}{2r}, \quad 1 \leq i \leq d+1. \tag{4.14}$$

We present this last computation for $i = d+1$. We have

$$\frac{c_{d+1}(r)}{c_{d+1}(r-1)} = \frac{\lfloor (n-r-1)/2 \rfloor - (d-1)r}{r} \cdot \prod_{j=1}^{r-1} \frac{\lfloor (n-r-1)/2 \rfloor - (d-1)r - j}{\lfloor (n-r)/2 \rfloor - (d-1)(r-1) - j + 1}.$$

Since

$$n/2 - Cdr \leq \lfloor (n-r-1)/2 \rfloor - (d-1)r \leq n/2,$$

it suffices to show that the product above is at most 1 and at least $1 - Cdr/(n - Cdr)$. Indeed, every element in the product is clearly at most 1, and hence so is the product. For the other inequality, note that the last element in the product is the smallest, so that the product is at least

$$\left(\frac{\lfloor (n-r-1)/2 \rfloor - dr + 1}{\lfloor (n-r)/2 \rfloor - d(r-1) + 1} \right)^{r-1} \geq \left(1 - \frac{d+1}{n/2 - r/2 - dr + d + 1} \right)^r \geq 1 - \frac{4dr}{n - 4dr}.$$

The statement now follows directly from (4.13) and (4.14). □

Corollary 4.10. *For any positive integer r such that $(2d+1)r + 2d \leq n$, we have*

$$\frac{n - Crd}{4r2^d} \leq \frac{\mathbb{P}(R = r)}{\mathbb{P}(R = r - 1)} \leq \frac{n}{2r2^d}.$$

4.3 Proof of theorems

We are now ready to prove the theorems stated in Section 2.1.

The supercritical regime. We prove Theorem 2.1 and Corollary 2.2. By Lemma 4.2, we have

$$\mathbb{P}(R \geq r) = \mathbb{P} \left(\bigcup_{2d+1 < s_1 < \dots < s_r \leq n} A_{s_1} \cap \dots \cap A_{s_r} \right) \leq \binom{n}{r} 2^{-dr}, \quad r \geq 1.$$

One may easily check that $|\text{Rng}(f)| \leq R + 3$, so that

$$\mathbb{P}(|\text{Rng}(f)| \geq 3 + r) \leq \mathbb{P}(R \geq r) \leq \binom{n}{r} 2^{-dr}, \quad r \geq 1.$$

Moreover, it is easy to describe all homomorphisms which take on at most 3 values. Denote by V_0 and V_1 the even and odd vertices in $\{0, 1, \dots, n\}$, respectively, and denote by Ω_0 and Ω_1 the set of homomorphisms which are constant on V_0 and V_1 , respectively. Then it is clear that $\{|\text{Rng}(f)| \leq 3\} = \Omega_0 \cup \Omega_1$, that $\{|\text{Rng}(f)| < 3\} = \Omega_0 \cap \Omega_1$ and that $|\Omega_0 \cap \Omega_1| = 2$. Also, note that $|V_0| = \lfloor n/2 \rfloor + 1$ and $|V_1| = \lceil n/2 \rceil$, so that $|\Omega_0| = 2^{|V_0|} = 2^{\lfloor n/2 \rfloor + 1}$ and $|\Omega_1| = 2^{|V_1|} = 2^{\lceil n/2 \rceil}$. Therefore,

$$\mathbb{P}(|\text{Rng}(f)| < 3) = \mathbb{P}(\Omega_0 \cap \Omega_1) \leq \frac{|\Omega_0 \cap \Omega_1|}{|\Omega_0 \cup \Omega_1|} = \frac{2}{2^{\lfloor n/2 \rfloor + 1} + 2^{\lceil n/2 \rceil} - 2} \leq 2^{1-n/2},$$

completing the proof of Theorem 2.1. To obtain Corollary 2.2, note that

$$\frac{\mathbb{P}(\Omega_1)}{\mathbb{P}(\Omega_0)} = 2^{\lfloor n/2 \rfloor + 1 - \lceil n/2 \rceil} = \begin{cases} 2 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd} \end{cases}.$$

Hence, if $d - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then, since $\mathbb{P}(\Omega_0 \cup \Omega_1) = \mathbb{P}(|\text{Rng}(f)| \leq 3) = 1 - o(1)$ by Theorem 2.1, we see that $\mathbb{P}(\Omega_0) = 1/2 - o(1)$ if n is odd and $\mathbb{P}(\Omega_0) = 1/3 - o(1)$ if n is even.

The subcritical regime. Before proving the relevant theorems, we need a better understanding of the typical number of jumps.

Lemma 4.11. *For any $\epsilon > 0$, we have*

$$\mathbb{P}(R < \lfloor \epsilon cn 2^{-d} \rfloor \mid \min(S \cup \{2d + 2\}) = 2i) \leq \epsilon, \quad 1 \leq i \leq d + 1.$$

Proof. Let $0 < \epsilon < 1$ and $1 \leq i \leq d + 1$. Lemma 4.9 implies that if c is small enough,

$$\frac{\mathbb{P}(R = r \mid \min(S \cup \{2d + 2\}) = 2i)}{\mathbb{P}(R = r - 1 \mid \min(S \cup \{2d + 2\}) = 2i)} \geq \frac{1}{\epsilon}, \quad 1 \leq r \leq \lfloor \epsilon cn 2^{-d} \rfloor.$$

Lemma 3.3 now yields the result. □

Corollary 4.12. *For any $\epsilon > 0$, if $n 2^{-d} \geq C/\epsilon$ then*

$$\mathbb{P}(R < \epsilon cn 2^{-d} \mid \min(S \cup \{2d + 2\}) = 2i) \leq \epsilon, \quad 1 \leq i \leq d + 1. \tag{4.15}$$

Consequently, if $n 2^{-d} \geq C$ then

$$\mathbb{E}[R \mid \min(S \cup \{2d + 2\}) = 2i] \geq cn 2^{-d}, \quad 1 \leq i \leq d + 1. \tag{4.16}$$

Proof. If $n 2^{-d} \geq C/\epsilon$ then $\epsilon c' n 2^{-d} \leq \lfloor \epsilon cn 2^{-d} \rfloor$, and hence, (4.15) follows from Lemma 4.11. To obtain (4.16), substitute $\epsilon = 1/2$ in (4.15). □

We shall also require a similar inequality for the number of jumps up to a given vertex. For $2d + 1 \leq k \leq n$, define

$$R(k) := |S \cap \{2d + 2, 2d + 3, \dots, k\}| = \sum_{i=2d+2}^k \mathbf{1}_{A_i}.$$

Lemma 4.13. *We have*

$$\mathbb{E}[R(k)] \geq ck 2^{-d} - 1/6, \quad 1 \leq k \leq n.$$

Proof. First note that the statement is trivial when $n2^{-d} < C$. Thus, we may assume that $n2^{-d} \geq C$. Denote $x_1 := \min S - 1$ and denote by x_2, \dots, x_r the distances between consecutive values in S . By Corollary 4.7 and (4.3), conditioned on the event $\{R = r\}$ and on the event $E := \{\min S > 2d\}$ (again, we set $\min \emptyset = \infty$), x_1, \dots, x_r are identically distributed and satisfy $1 + x_1 + \dots + x_r \leq n$. Hence

$$\mathbb{E}[x_i \mid R = r, E] \leq n/r, \quad 1 \leq i \leq r.$$

Since $R(k) < j$ if and only if $1 + x_1 + \dots + x_j > k$, we have by Markov's inequality that

$$\mathbb{P}(R(k) < j \mid R, E) = \mathbb{P}(x_1 + \dots + x_j \geq k \mid R, E) \leq \frac{1}{k} \mathbb{E}[x_1 + \dots + x_j \mid R, E] \leq \frac{jn}{kR}.$$

So

$$\mathbb{E}[R(k) \mid R, E] \geq \mathbb{P}(R(k) \geq \lfloor kR/2n \rfloor \mid R, E) \cdot \lfloor kR/2n \rfloor \geq (1/2) \lfloor kR/2n \rfloor \geq kR/4n - 1/2.$$

Hence, by the assumption that $n2^{-d} \geq C$ and by (4.16),

$$\mathbb{E}[R(k) \mid E] \geq k\mathbb{E}[R \mid E]/4n - 1/2 \geq ck2^{-d} - 1/2.$$

Finally, by Lemma 4.1, we have

$$\mathbb{E}[R(k)] \geq \mathbb{E}[R(k) \mid E] \cdot \mathbb{P}(E) \geq ck2^{-d} - 1/6. \quad \square$$

Lemma 4.14. *We have*

$$\text{Var}(f(k)) \geq 1, \quad 1 \leq k \leq n.$$

Proof. If k is odd then $|f(k)| \geq 1$ and the result follows by the fact that $f(k)$ is symmetric. Henceforth, we assume that k is even.

Consider the mapping $f \mapsto f_0$ from A_2 to A_2^c defined by

$$f_0(i) := \begin{cases} 0 & \text{if } i = 0 \\ f(i) - f(2) & \text{if } i \geq 1 \end{cases}, \quad 0 \leq i \leq n.$$

One may check that this mapping is indeed well-defined and that it is injective. Since $|f_0(k)| = 2$ when $f(k) = 0$, and since the mapping is injective, we have

$$\begin{aligned} \mathbb{P}(\{f(k) = 0\} \cap A_2) &\leq \mathbb{P}(\{|f(k)| = 2\} \cap A_2^c) \\ &\leq \mathbb{P}(\{f(k) \neq 0\} \cap A_2^c) = 1 - \mathbb{P}(\{f(k) = 0\} \cup A_2). \end{aligned}$$

Therefore,

$$\mathbb{P}(f(k) = 0) + \mathbb{P}(A_2) = \mathbb{P}(\{f(k) = 0\} \cup A_2) + \mathbb{P}(\{f(k) = 0\} \cap A_2) \leq 1.$$

Since, $\mathbb{P}(A_2) \geq 1/4$ by Lemma 4.1, we have

$$\mathbb{P}(f(k) = 0) \leq 3/4.$$

Finally, since $f(k) \neq 0$ if and only if $|f(k)| \geq 2$, we have

$$\text{Var}(f(k)) = \mathbb{E}[f(k)^2] \geq 4 \cdot \mathbb{P}(|f(k)| \geq 2) = 4 \cdot \mathbb{P}(f(k) \neq 0) \geq 1. \quad \square$$

We are now ready to prove Theorem 2.4 and Theorem 2.3. In both proofs, we consider the following modified average height h' . For $1 \leq k \leq n$, define

$$h'(k) := \begin{cases} 0 & \text{if } k \leq 2d + 1 \\ h(k) - h(2d + 1) & \text{otherwise} \end{cases}.$$

Recall that Corollary 4.6 implies that, for any $2d + 2 \leq k \leq n$, conditioned on S , $\{\Delta(j) \mid (j, t) \in \mathcal{C}(S \cap \{2d + 2, \dots, k\})\}$ are independent and

$$h'(k) = \sum_{(j,t) \in \mathcal{C}(S \cap \{2d+2, \dots, k\})} t \cdot \Delta(j), \quad 1 \leq k \leq n. \tag{4.17}$$

Proof of Theorem 2.4. By the above remark, we have

$$\text{Var}(h'(k) \mid S) = \sum_{(j,t) \in \mathcal{C}(S \cap \{2d+2, \dots, k\})} t^2, \quad 1 \leq k \leq n. \tag{4.18}$$

Notice that, conditioned on S , the expectation of $h'(k)$ is zero, so that by the law of total variance,

$$\text{Var}(h'(k)) = \mathbb{E}[\text{Var}(h'(k) \mid S)], \quad 1 \leq k \leq n. \tag{4.19}$$

To obtain an upper bound on $\text{Var}(h'(k))$, we use Lemma 4.2 to obtain

$$\mathbb{P}((j, t) \in \mathcal{C}(S \cap \{2d + 2, \dots, k\})) \leq \mathbb{P}(C_{j,t}) \cdot \mathbf{1}_{\{(2d+1)t < j \leq k\}} \leq 2^{-dt} \cdot \mathbf{1}_{\{j \leq k\}},$$

for any $1 \leq j \leq n$ and $t \geq 1$. Therefore, by (4.18) and (4.19), we have

$$\text{Var}(h'(k)) = \mathbb{E}[\text{Var}(h'(k) \mid S)] \leq \sum_{t=1}^{\infty} \sum_{j=1}^n t^2 2^{-dt} \cdot \mathbf{1}_{\{j \leq k\}} = k \sum_{t=1}^{\infty} t^2 2^{-dt} \leq Ck2^{-d}. \tag{4.20}$$

Finally, since $|f(k) - h(k)| \leq 1$ and $|h(k) - h'(k)| \leq 1$, we have $|f(k)| \leq |h'(k)| + 2$, which gives

$$\begin{aligned} \text{Var}(f(k)) &= \mathbb{E}[f(k)^2] \leq \mathbb{E}[(|h'(k)| + 2)^2] = \mathbb{E}[h'(k)^2 + 4|h'(k)| + 4] \\ &\leq 5 \cdot \mathbb{E}[h'(k)^2] + 4 = 5 \cdot \text{Var}(h'(k)) + 4 \leq Ck2^{-d} + 4. \end{aligned}$$

For the lower bound, we note that

$$\sum_{(j,t) \in \mathcal{C}(S \cap \{2d+2, \dots, k\})} t^2 \geq \sum_{(j,t) \in \mathcal{C}(S \cap \{2d+2, \dots, k\})} t = |S \cap \{2d + 2, \dots, k\}| = R(k).$$

Therefore, by (4.18), (4.19) and Lemma 4.13, we have

$$\text{Var}(h'(k)) = \mathbb{E}[\text{Var}(h'(k) \mid S)] \geq \mathbb{E}[R(k)] \geq ck2^{-d} - 1/6.$$

Since $|f(k) - h(k)| \leq 1$ and $|h(k) - h'(k)| \leq 1$, we have $|f(k)| \geq |h'(k)| - 2$. In particular, $|f(k)| \geq |h'(k)|/3$ when $|h'(k)| \geq 3$. Therefore,

$$\begin{aligned} \text{Var}(f(k)) &= \mathbb{E}[f(k)^2] \geq \mathbb{E}[(h'(k)/3)^2 \cdot \mathbf{1}_{\{|h'(k)| \geq 3\}}] \\ &= \mathbb{E}[h'(k)^2] / 9 - \mathbb{E}[h'(k)^2 \cdot \mathbf{1}_{\{|h'(k)| \leq 2\}}] / 9 \\ &\geq \text{Var}(h'(k)) / 9 - 4/9 \\ &\geq ck2^{-d} - 1/2. \end{aligned}$$

Finally, together with Lemma 4.14, we have

$$\text{Var}(f(k)) \geq \max\{1, ck2^{-d} - 1/2\} \geq \max\{1, c'k2^{-d}\}. \quad \square$$

Proof of upper bound in Theorem 2.3. Denote $\mathcal{C}(S \cap \{2d + 2, \dots, n\}) = \{(k_1, t_1), \dots, (k_m, t_m)\}$, ordering the elements so that the k_j are increasing. Observe that for any $1 \leq j < m$ and any $k_j \leq k \leq k_{j+1}$, we have that $h(k)$ is between $h(k_j)$ and $h(k_{j+1})$. Therefore,

$$\max_{1 \leq k \leq n} |h'(k)| = \max_{1 \leq j \leq m} |h'(k_j)|.$$

In this notation, by (4.17) we have

$$h'(k_j) = \sum_{i=1}^j t_i \cdot \Delta(k_i), \quad 1 \leq j \leq m,$$

where, conditioned on S , $\{\Delta(k_j) \mid 1 \leq j \leq m\}$ are independent. Therefore, we may apply Kolmogorov's maximal inequality to the process $(h'(k_j) \mid 1 \leq j \leq m)$, conditioned on S , to obtain

$$\mathbb{P} \left(\max_{1 \leq k \leq n} |h'(k)| \geq x \mid S \right) = \mathbb{P} \left(\max_{(k,t) \in \mathcal{C}(S \cap \{2d+2, \dots, n\})} |h'(k)| \geq x \mid S \right) \leq \frac{\text{Var}(h'(n) \mid S)}{x^2}.$$

Therefore, by (4.20), we have

$$\mathbb{P} \left(\max_{1 \leq k \leq n} |h'(k)| \geq x \right) = \mathbb{E} \left[\mathbb{P} \left(\max_{1 \leq k \leq n} |h'(k)| \geq x \mid S \right) \right] \leq \frac{\text{Var}(h'(n))}{x^2} \leq \frac{Cn2^{-d}}{x^2}.$$

From this we obtain

$$\mathbb{E} \left[\max_{1 \leq k \leq n} |h'(k)| \right] = \sum_{x=1}^{\infty} \mathbb{P} \left(\max_{1 \leq k \leq n} |h'(k)| \geq x \right) \leq \sum_{x=1}^{\infty} \min \left\{ 1, \frac{Cn2^{-d}}{x^2} \right\} \leq C' \sqrt{n2^{-d}}.$$

To see the last inequality one checks that

$$\begin{aligned} \sum_{x=1}^{\infty} \min \left\{ 1, \frac{Cn2^{-d}}{x^2} \right\} &\leq \sum_{x=1}^{\lceil \sqrt{n2^{-d}} \rceil - 1} 1 + \min \left\{ 1, \frac{Cn2^{-d}}{\lceil \sqrt{n2^{-d}} \rceil^2} \right\} + \sum_{x=\lceil \sqrt{n2^{-d}} \rceil + 1}^{\infty} \frac{Cn2^{-d}}{x^2} \\ &\leq \sqrt{n2^{-d}} + \min \left\{ 1, \frac{Cn2^{-d}}{\lceil \sqrt{n2^{-d}} \rceil^2} \right\} + \int_{\sqrt{n2^{-d}}}^{\infty} \frac{Cn2^{-d}}{x^2} dx \leq C' \sqrt{n2^{-d}}. \end{aligned}$$

Finally, using the fact that

$$|\text{Rng}(f)| \leq 3 + \max_{1 \leq k \leq n} h'(k) - \min_{1 \leq k \leq n} h'(k),$$

we obtain

$$\mathbb{E}[|\text{Rng}(f)|] \leq 3 + 2 \cdot \mathbb{E} \left[\max_{1 \leq k \leq n} |h'(k)| \right] \leq 3 + C\sqrt{n2^{-d}}. \quad \square$$

Proof of lower bound in Theorem 2.3. We begin by showing that the range is large with high probability, when $n2^{-d}$ is large enough. Fix $0 < \epsilon < 1$. Assume that $n2^{-d} \geq C/\epsilon$. By (4.15), there exists a $\delta_1 > 0$, depending only on ϵ , such that

$$\mathbb{P} (R < 2\delta_1 n2^{-d}) \leq \epsilon/4. \tag{4.21}$$

This tells us that typically there are many jumps. We now show that typically there are many distinct chains as well. For $s \geq 1$, let

$$M_s := \sum_{(2d+1)s < k \leq n} \mathbf{1}_{C_{k,s}}$$

be the number of sub-chains of length s . Then, as we shall now show,

$$|\mathcal{C}(S)| \geq \frac{R - M_s}{s - 1}, \quad s \geq 2. \tag{4.22}$$

Indeed, denoting $\mathcal{C}(S \cap \{2d + 2, \dots, n\}) = \{(k_1, t_1), \dots, (k_m, t_m)\}$ and considering the contribution of each chain to M_s , we see that

$$M_s = \sum_{i=1}^m \max\{0, t_i - s + 1\} \geq \sum_{i=1}^m (t_i - s + 1) = R - (s - 1)m.$$

Note that $|\mathcal{C}(S)| \geq m$ now yields (4.22). By Lemma 4.2, we have

$$\mathbb{E}[M_s] = \sum_{(2d+1)s < k \leq n} \mathbb{P}(C_{k,s}) \leq n2^{-ds} \leq n2^{-d}2^{1-s}, \quad s \geq 1.$$

Taking $s = s_0$ large enough, we have by Markov's inequality,

$$\mathbb{P}(M_{s_0} \geq \delta_1 n2^{-d}) \leq \frac{\mathbb{E}[M_{s_0}]}{\delta_1 n2^{-d}} \leq \frac{2^{1-s_0}}{\delta_1} \leq \epsilon/4. \tag{4.23}$$

Therefore, by (4.21), (4.22) and (4.23), we have for $\delta_2 := \delta_1/s_0$ that

$$\mathbb{P}(|\mathcal{C}(S)| \geq \delta_2 n2^{-d}) \geq \mathbb{P}(R \geq 2\delta_1 n2^{-d}, M_{s_0} \leq \delta_1 n2^{-d}) \geq 1 - \epsilon/2. \tag{4.24}$$

Recalling from Corollary 4.6 that, conditioned on S , $h(n)$ is the sum of $|\mathcal{C}(S)|$ independent random variables, we may apply Theorem 3.5 to obtain

$$\mathbb{P}(|h(n)| < r \mid S) \leq \frac{Cr}{\sqrt{|\mathcal{C}(S)|}}, \quad r \in \mathbb{N}.$$

Therefore,

$$\mathbb{P}(|\text{Rng}(f)| < r \mid \mathcal{C}(S)) \leq \mathbb{P}(|h(n)| < r \mid \mathcal{C}(S)) \leq \frac{Cr}{\sqrt{|\mathcal{C}(S)|}}, \quad r \in \mathbb{N}. \tag{4.25}$$

Finally, by (4.24) and (4.25), for any $\delta > 0$, we have

$$\begin{aligned} \mathbb{P}\left(|\text{Rng}(f)| < \lfloor \delta\sqrt{n2^{-d}} \rfloor\right) &\leq \mathbb{P}(|\mathcal{C}(S)| < \delta_2 n2^{-d}) + \mathbb{P}\left(|\text{Rng}(f)| < \lfloor \delta\sqrt{n2^{-d}} \rfloor \mid |\mathcal{C}(S)| \geq \delta_2 n2^{-d}\right) \\ &\leq \epsilon/2 + C\delta/\sqrt{\delta_2}. \end{aligned}$$

Therefore, there exists a $\delta > 0$, depending only on ϵ , such that if $\delta\sqrt{n2^{-d}} \geq 1$ then

$$\mathbb{P}\left(|\text{Rng}(f)| < 3 + \lfloor \delta\sqrt{n2^{-d}} \rfloor\right) \leq \mathbb{P}\left(|\text{Rng}(f)| < \lfloor 4\delta\sqrt{n2^{-d}} \rfloor\right) \leq \epsilon,$$

proving (2.1) when $n2^{-d}$ is large enough. On the other hand, if $\delta\sqrt{n2^{-d}} < 1$ then (2.1) follows immediately from Theorem 2.1, since

$$\mathbb{P}\left(|\text{Rng}(f)| < 3 + \lfloor \delta\sqrt{n2^{-d}} \rfloor\right) = \mathbb{P}(|\text{Rng}(f)| < 3) \leq 2^{1-n/2}.$$

It remains to show the lower bound on the expectation. Note that the statement is trivial when $n \leq 2$, and so we may assume that $n \geq 3$. By taking $\epsilon = 1/4$ in (2.1), noting that $|\text{Rng}(f)| \geq 2$ and by Theorem 2.1, we conclude that

$$\begin{aligned} \mathbb{E}[|\text{Rng}(f)|] &= 3 + \mathbb{E}[(|\text{Rng}(f)| - 3) \cdot \mathbf{1}_{\{|\text{Rng}(f)| \geq 3\}}] + \mathbb{E}[(|\text{Rng}(f)| - 3) \cdot \mathbf{1}_{\{|\text{Rng}(f)| < 3\}}] \\ &= 3 + \mathbb{E}[(|\text{Rng}(f)| - 3) \cdot \mathbf{1}_{\{|\text{Rng}(f)| \geq 3\}}] - \mathbb{P}(|\text{Rng}(f)| < 3) \\ &\geq 3 + (1 - 1/4 - 2^{-1/2})\lfloor \delta\sqrt{n2^{-d}} \rfloor - 2^{1-n/2} \\ &\geq 3 + \lfloor c\sqrt{n2^{-d}} \rfloor - 2^{1-n/2}. \end{aligned} \tag{4.26}$$

□

The critical regime. Here we prove Theorem 2.5. Denote $\lambda := \lim n2^{-d}$ which exists and is a positive number by assumption. The proof of Theorem 2.5 consists of two parts. First, we show that R converges to $N^\pm(\lambda)$ as n tends to infinity through even or odd integers. Next, we show that in this regime the values at the jumps constitute a simple random walk and that this walk determines the range of the homomorphism.

By Lemma 4.2, we have

$$\mathbb{E}[R] = \sum_{k=2d+2}^n \mathbb{P}(A_k) \leq n2^{-d} = \lambda + o(1).$$

Therefore, the expectation of R is uniformly bounded as $n \rightarrow \infty$, and hence, Markov's inequality implies that R is tight as $n \rightarrow \infty$. Using notation as in the proof of Lemma 4.9, we have

$$|\{R = r\} \cap \{\min(S \cup \{2d + 2\}) = 2i\}| = c_i(r)b_i(r), \quad 1 \leq i \leq d + 1.$$

A direct computation shows that for any constant $r \geq 0$, we have

$$c_i(r) \sim \frac{n^r}{2^r r!} \quad \text{as } n \rightarrow \infty,$$

uniformly in $1 \leq i \leq d + 1$, and

$$2^{\lceil (n-r)/2 \rceil} \sim 2^{(n-r)/2} \cdot \gamma(n-r) \quad \text{as } n \rightarrow \infty,$$

where $\gamma(k) := 1$ if k is even and $\gamma(k) := \sqrt{2}$ if k is odd. Denoting by $J := A_1 \cup \dots \cup A_{2d+1}$ the event that a jump occurs prior to vertex $2d + 2$, and recalling (4.12), we obtain

$$\begin{aligned} |\{R = r\} \cap J^c| &\sim \frac{n^r}{r!} \cdot 2^{n/2-(d+3/2)r} \cdot \gamma(n-r), \\ |\{R = r\} \cap A_{2i}| &\sim \frac{n^r}{r!} \cdot 2^{n/2-(d+3/2)r+1/2-i} \cdot \gamma(n-r-1), \end{aligned}$$

uniformly in $1 \leq i \leq d$. Therefore, denoting $\lambda' := \lambda/(2\sqrt{2})$, we have

$$|\{R = r\}| \sim \frac{\lambda'^r}{r!} \cdot 2^{n/2} \cdot \left(\gamma(n-r) + \sqrt{2}\gamma(n-r-1) \right),$$

where we have used the fact that $\sum_{i=1}^d 2^{-i} \rightarrow 1$. Using the tightness of R , we see that

$$\mathbb{P}(R = r) \sim Z(n)^{-1} \cdot \frac{\lambda'^r}{r!} \cdot \left(\gamma(n-r) + \sqrt{2}\gamma(n-r-1) \right),$$

where $Z(n)$ is a normalizing constant. Therefore, recalling the parity-biased Poisson distribution defined in (2.2) and the equation (2.3), we see that

$$R \xrightarrow[n \text{ even}]{n \rightarrow \infty} \mu(\lambda', 3/(2\sqrt{2})) \quad \text{and} \quad R \xrightarrow[n \text{ odd}]{n \rightarrow \infty} \mu(\lambda', 2\sqrt{2}/3), \tag{4.27}$$

completing the first part of the proof.

We remark that it is also possible to obtain the limiting distribution of R conditioned on whether or not a jump occurred at the first $2d + 1$ vertices. We do not make use of this in our paper but we note the final result. A further calculation gives the following formula for the asymptotic probability of J ,

$$\left(\sqrt{2} \cdot \frac{\mathbb{P}(J^c)}{\mathbb{P}(J)} \right)^{(-1)^n} \xrightarrow{n \rightarrow \infty} \frac{\cosh(\lambda') + \sqrt{2} \sinh(\lambda')}{\sinh(\lambda') + \sqrt{2} \cosh(\lambda')},$$

and the following formula for the asymptotic distribution of R given $\mathbf{1}_J$,

$$\mathbb{P}(R = r \mid \mathbf{1}_J) - \mu\left(\lambda', \sqrt{2}^{(-1)^{n+1+1_J}}\right)(r) \xrightarrow[n \rightarrow \infty]{} 0.$$

We now proceed to analyze the range of a typical homomorphism in the critical regime. We begin by showing that the jumps are sparse enough so that it is unlikely to have chains of length greater than one. Let

$$B := \bigcap_{k=2d+2}^n (A_k^c \cup A_{k-2d-1}^c)$$

be the event that there are no two minimal-distance jumps (i.e. jumps at distance $2d + 1$). We wish to show that $\mathbb{P}(B) = 1 - o(1)$. Indeed, by considering the first $2d + 1$ elements in the intersection separately from the rest, Lemma 4.2 implies that $\mathbb{P}(B^c) \leq (2d + 1)2^{-d} + n2^{-2d} = o(1)$ as required. Notice that $\mathbf{1}_B$ is determined by S . Let \mathcal{I} denote the set of all feasible jump structures I such that $\{S = I\} \subset B$. Observe that on the event B , $\mathcal{C}(S) = \{(s, 1) \mid s \in S\}$. Therefore, by Corollary 4.6, for any $I \in \mathcal{I}$, conditioned on $S = I$, $\{\Delta(s) \mid s \in S\}$ are independent uniform signs and, by (4.17),

$$h'(k) = \sum_{s \in S \cap \{2d+2, \dots, k\}} \Delta(s).$$

In other words, for any $I \in \mathcal{I}$, conditioned on $S = I$, $(h'(s) \mid 2d + 1 < s \in S)$ is a simple random walk of length R (without the leading zero). Since

$$\{h'(s) \mid 2d + 1 < s \in S\} \cup \{0\} = \{h'(k) \mid 2d + 1 \leq k \leq n\},$$

for any $I \in \mathcal{I}$,

$$|\{h'(k) \mid 2d + 1 \leq k \leq n\}| \stackrel{d}{=} |\text{Rng}(S_R)| \quad \text{conditioned on } S = I, \quad (4.28)$$

where S_i is an independent simple random walk run for i steps. Define the event

$$E := \{|\{f(k) \mid 0 \leq k \leq 2d\}| > 2\}.$$

It is not difficult to check that

$$|\text{Rng}(f)| = 2 + |\{h'(k) \mid 2d + 1 \leq k \leq n\}| \quad \text{on } E. \quad (4.29)$$

We now show that $\mathbb{P}(E) = 1 - o(1)$. Observe that Lemma 4.5 implies that

$$\mathbb{P}(E^c \mid \min S \geq 4d + 2) = 2 \cdot 2^{-d}.$$

Hence, by Lemma 4.2, and since $J \subset E$, we have

$$\begin{aligned} \mathbb{P}(E^c) &= \mathbb{P}(E^c \cap \{\min S < 4d + 2\}) + \mathbb{P}(E^c \cap \{\min S \geq 4d + 2\}) \\ &\leq \mathbb{P}(2d + 2 \leq \min S < 4d + 2) + \mathbb{P}(E^c \mid \min S \geq 4d + 2) \\ &\leq 2d2^{-d} + 2^{1-d} = o(1). \end{aligned}$$

Finally, Theorem 2.5 follows from (4.27), (4.28), (4.29) and the fact that $\mathbb{P}(E \cap B) = 1 - o(1)$.

5 Homomorphisms on the Torus

In this section we prove the theorems regarding homomorphisms on the torus which were stated in Section 2.2. The ideas and notions previously introduced in Section 4 to handle the case of homomorphisms on the line will still prove to be effective on the torus, although some of them will need to be adapted. For example, the notions of average height, jumps and chains will still be used and they are defined in an analogous manner. One thing which must change, for instance, is how we use these notions and the events that we condition on. Note that, if we condition on the lengths and the positions of the chains, their signs will not be independent, since they must add up correctly. This fact, which is inherently due to the topology of the torus, makes the analysis slightly more complex. Instead, we will show that, conditioned on the lengths and the signs of the chains (but not on their positions), their relative order is uniform. This will allow us to reduce some of the analysis to a case of a uniformly chosen reordering of a sequence of numbers. One aspect which is simpler for homomorphisms on the torus is that there are no boundary effects, i.e., no need to consider the first $2d + 1$ vertices separately.

5.1 Definitions

We consider the graph $T_{n,d}$, n even, whose vertex set is $V := \{0, \dots, n-1\}$ and whose edges are defined by $i \sim j$ if and only if $\rho(i, j) = 1, 3, \dots, 2d + 1$, where we define the distance ρ between x and y to be

$$\rho(x, y) := \min\{|x - y|, n - |x - y|\}, \quad x, y \in V.$$

We define also the *clockwise distance* ρ^+ from x to y to be

$$\rho^+(x, y) := y - x + n\mathbf{1}_{\{x > y\}} = \min\{k \geq 0 \mid x + k = y \pmod n\}, \quad x, y \in V.$$

Note that $\rho^+(x, y) + \rho^+(y, x) = n$ and that $\rho(x, y) = \min\{\rho^+(x, y), \rho^+(y, x)\}$ for any $x, y \in V$.

Throughout this section, $\text{Hom}(T_{n,d}) := \text{Hom}(T_{n,d}, 0)$, f is a uniformly sampled homomorphism from $\text{Hom}(T_{n,d})$, the probability space is the uniform distribution on the set $\text{Hom}(T_{n,d})$, and events are subsets of $\text{Hom}(T_{n,d})$. We also note that, in this section, addition and subtraction of elements in V are always modulo n .

We would like to define the notion of the (local) average height of a homomorphism at a vertex $x \in V$. To do so, we “look back” just enough in order to define this in a meaningful way. Precisely, for $x \in V$, define the *average height* at x as the unique number $h(x)$ satisfying

$$\text{There exists a } k \geq 1 \text{ for which } \{f(x - i) \mid i = 0, 1, \dots, k\} = \{h(x) - 1, h(x), h(x) + 1\}.$$

This is well defined for any homomorphism f which takes on at least 3 values. There are two specific homomorphisms for which the size of the range is 2, and hence for which this is not well defined. These are f_1^{flat} and f_{-1}^{flat} , where

$$f_i^{\text{flat}}(x) := \begin{cases} 0 & \text{if } x \text{ is even} \\ i & \text{if } x \text{ is odd} \end{cases}, \quad x \in V, i \in \{-1, 1\}.$$

For these homomorphisms we define $h(x) := 0$ for all $x \in V$. For $x \in V$, define

$$\begin{aligned} \Delta(x) &:= h(x) - h(x - 1), \\ A_x &:= \{\Delta(x) \neq 0\}. \end{aligned}$$

Observe that necessarily $\Delta(x) \in \{-1, 0, 1\}$. When A_x occurs, we will say that a *jump* occurred at vertex x . For $x, y \in V$, denote by

$$A_{x,y} := A_x \cap A_y \cap \{\Delta(x) = -\Delta(y)\}$$

the event that there are jumps in opposite directions at x and y . Denote by

$$S^+ := \{x \in V \mid \Delta(x) = 1\} \quad \text{and} \quad S^- := \{x \in V \mid \Delta(x) = -1\}$$

the sets of vertices at which a positive or negative jump occurred, respectively. Let

$$S := S^+ \cup S^- \tag{5.1}$$

be the set of vertices at which we have a jump in either direction. Notice that necessarily $|S^+| = |S^-|$, and define

$$R := |S^+| = |S^-| = |S|/2,$$

the number of jumps in a given direction. Notice that the distance between jumps is at least $2d + 1$, as for homomorphisms on the line. For $x \in V$ and $t \geq 1$, let

$$C_{x,t} := A_x \cap A_{x-2d-1} \cap \cdots \cap A_{x-(t-1)(2d-1)}$$

be the event that there is a chain of t minimal-distance jumps ending at vertex x .

We say that a subset $I \subset V$ is a *feasible jump structure* if $\{S = I\} \neq \emptyset$, i.e. if $\mathbb{P}(S = I) > 0$. We would like to describe this condition solely in terms of the structure of I . To this end, write $I = \{s_1, \dots, s_t\}$, where $0 \leq s_1 < \cdots < s_t < n$ and let $s_0 := s_t$. Similarly to the case of the line, see condition (4.3), the following conditions are necessary for I to be a feasible jump structure.

$$\begin{aligned} \rho^+(s_{j-1}, s_j) &\geq 2d + 1, \quad 1 \leq j \leq t, \\ \rho^+(s_{j-1}, s_j) &\text{ is odd, } \quad 1 \leq j \leq t. \end{aligned} \tag{5.2}$$

In contrast to the case of the line, these conditions alone are not sufficient for I to be a feasible jump structure. This is due to the fact that the torus imposes a topological constraint. Namely, that at the end of the homomorphism the average height must “return” to its initial value. This additional condition, whose precise description (5.4) we postpone to the next section, along with condition (5.2), is necessary and sufficient for I to be a feasible jump structure.

In addition, we say that a subset $I \subset V$ is a *feasible jump sub-structure* if it is a subset of a feasible jump structure, or equivalently, if $\{I \subset S\} \neq \emptyset$. Notice that the definition implies that condition (5.2) is necessary for I to be a feasible jump sub-structure. For any $I \subset V$ satisfying (5.2), by considering the connected components of the subgraph of $T_{n,d}$ induced by I , one may see that the event $\{I \subset S\}$ can be uniquely written as $C_{k_1,t_1} \cap \cdots \cap C_{k_m,t_m}$, where

$$\begin{aligned} 0 &\leq k_1 < \cdots < k_m < n, \\ t_1 + \cdots + t_m &= |I|, \\ \rho^+(k_{j-1}, k_j) &> (2d + 1)t_j, \quad 1 \leq j \leq m, \end{aligned} \tag{5.3}$$

and where we let $k_0 := k_m$ (see Figure 6). These conditions ensure that there is no overlap between the different chains, and moreover, that there is some gap between them (since otherwise they would merge into a larger chain). For a subset $I \subset V$ satisfying (5.2), we define

$$\mathcal{C}(I) := \{(k_j, t_j) \mid 1 \leq j \leq m\},$$

and refer to this as the *chain structure* of I .

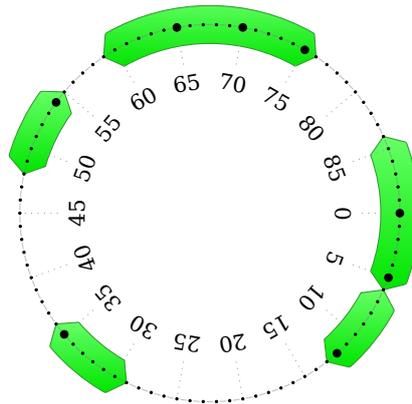


Figure 6: Given a subset $I \subset V$ satisfying (5.2), we construct $\mathcal{C}(I)$, the chain structure of I , by partitioning the elements of I according to the connected components in the subgraph of $T_{n,d}$ induced by I . The elements of I are denoted by large vertices and the chain structure is denoted by blocks surrounding the vertices. In the figure, $n = 90$, $d = 2$ and $I = \{0, 5, 12, 35, 54, 65, 70, 75\}$, and hence, $\mathcal{C}(I) = \{(5, 2), (12, 1), (35, 1), (54, 1), (75, 3)\}$.

5.2 The structure of a homomorphism

In this section, our goal is to give a useful description of the structure of a homomorphism on the torus. Namely, that which is stated in Lemma 5.2 and Lemma 5.3 below. To this end, we would like to decompose a homomorphism into two parts (see Figure 5 and Figure 7). The first part, which we shall denote by X , constitutes the changes in average height (the underlying bridge) of the homomorphism, while the second part, which we shall denote by Y , constitutes the fluctuations around the average height (the segments of constant average height).

We proceed first to define X . Given a subset $I \subset V$ satisfying (5.2), denote the set of feasible sign vectors for I by

$$\mathcal{B}^*(I) := \left\{ \epsilon \in \{-1, 1\}^{\mathcal{C}(I)} \mid \sum_{(k,t) \in \mathcal{C}(I)} \epsilon_{(k,t)} t = 0 \right\}.$$

When $I = \emptyset$, this set contains one element, the function with the empty domain. Note that in order for a subset $I \subset V$ to be a feasible jump structure, it is necessary and sufficient for I to satisfy (5.2) and

$$\mathcal{B}^*(I) \neq \emptyset. \tag{5.4}$$

This last condition is the manifestation of the topological constraint imposed by the torus. It says that the chain structure induced by the position of the jumps is such that it is possible to assign signs to each chain so that the average height “returns” to its initial value when completing an entire loop around the torus.

For a feasible jump structure I and a feasible sign vector $\epsilon \in \mathcal{B}^*(I)$, define the signed chain structure of (I, ϵ) by

$$\mathcal{C}^*(I, \epsilon) := \{(k, \epsilon_{(k,t)} \cdot t) \mid (k, t) \in \mathcal{C}(I)\}. \tag{5.5}$$

Recall the definition of S from (5.1). Define $X \in \{-1, 1\}^{\mathcal{C}(S)}$ by

$$X_{(k,t)} := f(k) - f(k-1), \quad (k, t) \in \mathcal{C}(S),$$

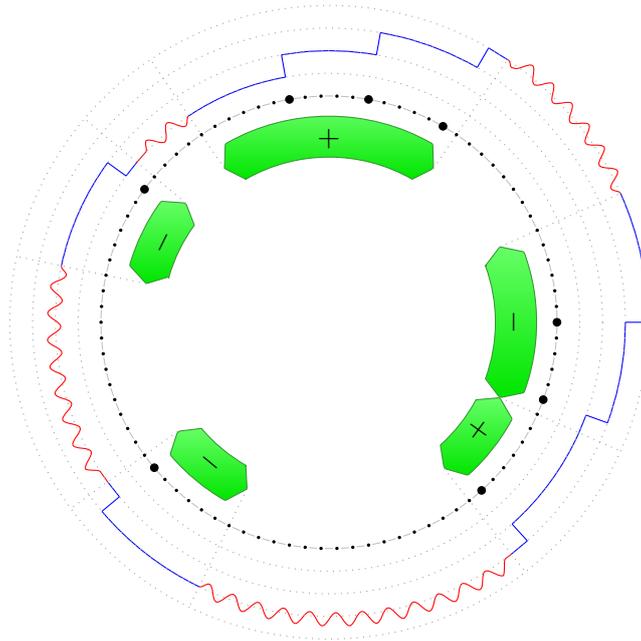


Figure 7: A homomorphism is broken up into sections of fluctuations and chains. The positions, lengths and associated signs of each chain (denoted by blocks with signs inside) make up the signed chain structure $\mathcal{C}^*(S, X)$ of the homomorphism. This information, along with the independent fluctuation values between the chains (denoted by wavy lines), uniquely determines the homomorphism.

and note that $X \in \mathcal{B}^*(S)$. This defines for us the random signed chain structure $\mathcal{C}^*(S, X)$. This random variable contains in a fairly simple manner all the necessary information for determining the range of f . Namely, it gives us the positions, lengths and signs of the chains in f .

We now proceed to define Y . For a non-empty feasible jump structure I , define the fluctuation points of I by

$$FP(I) := \left\{ y \in V \mid \rho^+(y, I) \in 2d + 1 + 2\mathbb{N} \right\},$$

where $\rho^+(y, I) := \min_{s \in I} \rho^+(y, s)$ and $\mathbb{N} := \{1, 2, 3, \dots\}$. That is, a point is a fluctuation point if its clockwise distance to the closest jump in the clockwise direction is odd and at least $2d + 3$. In particular, for any homomorphism f and any $y \in FP(S(f))$, f is not at its average height at y . Now, for a homomorphism f having at least one jump, define $Y \in \{-1, 1\}^{FP(S)}$ by

$$Y_y := f(y) - f(y - 1), \quad y \in FP(S).$$

It will be useful to have the following formula for the number of fluctuation points.

Claim 5.1. For any non-empty feasible jump structure I , we have

$$|FP(I)| = n/2 - (d + 1/2)|I| - |\mathcal{C}(I)|.$$

Proof. We have

$$\left| \{y \in V \mid \rho^+(y, I) < 2d + 3\} \right| = \sum_{(k,t) \in \mathcal{C}(I)} ((2d + 1)t + 2) = (2d + 1)|I| + 2|\mathcal{C}(I)|.$$

Furthermore, the set $\{y \in V \mid \rho^+(y, I) \geq 2d + 3\}$ is a disjoint union of intervals of even length, by (5.2), so that

$$\begin{aligned} |FP(I)| &= |\{y \in V \mid \rho^+(y, I) \geq 2d + 3\}| / 2 \\ &= (n - |\{y \in V \mid \rho^+(y, I) < 2d + 3\}|) / 2 \\ &= n/2 - (d + 1/2)|I| - |C(I)|. \end{aligned} \quad \square$$

The final lemmas show that X, Y and the jump structure S exactly encode the homomorphism.

Lemma 5.2. *For any non-empty feasible jump structure I , the mapping $f \mapsto (X(f), Y(f))$ is a bijection between $\{S = I\}$ and $\mathcal{B}^*(I) \times \{-1, 1\}^{FP(I)}$. Also, the event $\{S = \emptyset\}$ is of size $2^{n/2+1} - 2$.*

This is an immediate consequence of the following lemma.

Lemma 5.3. *For any non-empty feasible jump structure I and any feasible sign vector $\epsilon \in \mathcal{B}^*(I)$, the mapping $f \mapsto Y(f)$ is a bijection between $\{C^*(S, X) = C^*(I, \epsilon)\}$ and $\{-1, 1\}^{FP(I)}$. Also, the event $\{C^*(S, X) = \emptyset\}$ is of size $2^{n/2+1} - 2$.*

Proof. It is not hard to verify that this is indeed a bijection (see Figure 7 for a macroscopic picture and Figure 5 for a microscopic picture). We omit the proof as it is very similar to that of Lemma 4.5.

For the second statement, we note that $\{C^*(S, X) = \emptyset\} = \{S = \emptyset\} = \{h \equiv \text{const}\}$, and hence by considering the events $\{h \equiv 0\}$ and $\{|h| \equiv 1\}$, and recalling that we set $h \equiv 0$ when f takes on only two values, the statement readily follows. \square

5.3 The range

In this section, our goal is to give a more explicit description of the distribution of the range of a homomorphism. Namely, that which is stated in Proposition 5.4 below.

Recall the definition of the signed chain structure from (5.5). Let

$$\bar{W} := \{w \mid (k, w) \in C^*(S, X)\}, \tag{5.6}$$

be the set of lengths and signs of the chains taken with multiplicities, i.e. \bar{W} is a multi-set. For a vector of integers $w = (w_1, \dots, w_m)$, denote by

$$|\text{Rng}(w)| := 1 + \max_{0 \leq j \leq m} \sum_{i=1}^j w_i - \min_{0 \leq j \leq m} \sum_{i=1}^j w_i, \tag{5.7}$$

the size of the smallest interval in \mathbb{Z} which contains all partial sums of w .

Proposition 5.4. *Let $m \geq 1$, let $\bar{w} = \{w_1, \dots, w_m\}$ be a multi-set such that $\mathbb{P}(\bar{W} = \bar{w}) > 0$ and let π be a uniformly chosen permutation of $\{1, 2, \dots, m\}$. Then,*

$$(|\text{Rng}(f)| \text{ conditioned on } \bar{W} = \bar{w}) \stackrel{d}{=} 2 + |\text{Rng}(w_{\pi(1)}, \dots, w_{\pi(m)})|.$$

Proposition 5.4 is a direct consequence of the following two lemmas. The first of these, Lemma 5.5, relates the range of f to a random variable W defined below. The second, Lemma 5.7, describes the distribution of W conditioned on \bar{W} .

Given a set \mathcal{X} and a vector $x \in \mathcal{X}^m$, define the *period* of x by

$$\text{per}(x) := \min\{1 \leq k \leq m \mid \sigma^k(x) = x\},$$

where $\sigma: \mathcal{X}^m \rightarrow \mathcal{X}^m$ is the mapping $(x_1, \dots, x_m) \mapsto (x_2, \dots, x_m, x_1)$ and σ^k is its iteration k times, so that σ^m is the identity map. Define an equivalence relation on \mathcal{X}^m by $x \sim y$ if and only if there exists a k such that $\sigma^k(x) = y$. Denote by $[x] := \{x, \sigma(x), \dots, \sigma^{m-1}(x)\}$ the equivalence class of x . Observe that $|[x]| = \text{per}(x)$. For $x, y \in \mathcal{X}^m$, define

$$x \vee y := ((x_1, y_1), \dots, (x_m, y_m)),$$

and note that $\text{per}(x \vee y) = \text{lcm}(\text{per}(x), \text{per}(y))$.

Write $\mathcal{C}^*(S, X) = \{(k_i, w_i)\}_{i=1}^m$, where $0 \leq k_1 < \dots < k_m < n$. Let $k_0 := k_m$ and define

$$W := [(w_1, \dots, w_m)].$$

That is, W forgets the absolute position of the chains and remembers only their signed length and relative ordering. Note that \bar{W} is determined by W .

We begin by showing that the random variable W governs the range of the homomorphism. For a vector of integers w whose sum is zero, recalling (5.7), we define $|\text{Rng}([w])| := |\text{Rng}(w)|$, and note that this is indeed well-defined by the equivalence class of w .

Lemma 5.5. *We have*

$$|\text{Rng}(f)| = 2 + |\text{Rng}(W)| \quad \text{on the event } \{\bar{W} \neq \emptyset\}.$$

Proof. The partial sums of W correspond to differences in average height between two vertices. Therefore, $|\text{Rng}(W)| = 1 + \max_{x \in V} h(x) - \min_{x \in V} h(x)$. By the definition of the average height, we have $|f(x) - h(x)| \leq 1$ and $\{h(x) - 1, h(x), h(x) + 1\} \subset \text{Rng}(f)$ for any vertex $x \in V$. Therefore, by considering vertices at which the average height is maximal or minimal, we obtain the additional factor of 2 in the above equation. \square

Remark 5.6. *On the event $\{\bar{W} = \emptyset\}$, the size of the range of f is either 2 or 3. However, Lemma 5.2 implies that, conditioned on $\bar{W} = \emptyset$, the probability that the size of the range is 2 is of order $2^{-n/2}$.*

The next lemma is the final ingredient in the proof of Proposition 5.4. The remaining part of this section is devoted to its proof.

Lemma 5.7. *Let $m \geq 1$, let $\bar{w} = \{w_1, \dots, w_m\}$ be a multi-set such that $\mathbb{P}(\bar{W} = \bar{w}) > 0$ and let π be a uniformly chosen permutation of $\{1, 2, \dots, m\}$. Then,*

$$(W \text{ conditioned on } \bar{W} = \bar{w}) \stackrel{d}{=} [w_{\pi(1)}, \dots, w_{\pi(m)}].$$

Write $\mathcal{C}^*(S, X) = \{(k_i, w_i)\}_{i=1}^m$, where $0 \leq k_1 < \dots < k_m < n$. Let $k_0 := k_m$ and define

$$Z := [(\rho^+(k_{i-1}, k_i - (2d + 1)|w_i| - 2)/2, w_i)_{i=1}^m].$$

That is, Z forgets the absolute positions of the chains in $\mathcal{C}^*(S, X)$ and remembers only their distances one to the other (precisely, half the distance from the last vertex of one chain to one vertex before the beginning of the next chain). Note that the first coordinate of each element in Z is necessarily a non-negative integer. Also note that W is determined by Z . The next claim calculates the distribution of Z .

Claim 5.8. *Let $m \geq 1$. Let $w \in (\mathbb{Z} \setminus \{0\})^m$ be such that $w_1 + \dots + w_m = 0$ and let $x \in (\mathbb{N} \cup \{0\})^m$ be such that*

$$(2d + 1)(|w_1| + \dots + |w_m|) + 2m + 2(x_1 + \dots + x_m) = n. \tag{5.8}$$

Then

$$|\{Z = [x \vee w]\}| = \frac{n \cdot \text{per}(x \vee w)}{m} \cdot 2^{x_1 + \dots + x_m}.$$

Proof. Recall conditions (5.2) and (5.4), and note that, together with the assumptions, they imply that the event $\{Z = [x \vee w]\}$ is non-empty. We partition the event $\{Z = [x \vee w]\}$ according to $\mathcal{C}^*(S, X)$. Let $r \geq 1$ be the number of subsets in this partition, so that

$$\{Z = [x \vee w]\} = \bigcup_{i=1}^r \{\mathcal{C}^*(S, X) = \mathcal{C}^*(I_i, \epsilon_i)\},$$

where the (I_i, ϵ_i) are distinct and feasible. By Lemma 5.3, Claim 5.1 and (5.8),

$$|\{\mathcal{C}^*(S, X) = \mathcal{C}^*(I_i, \epsilon_i)\}| = 2^{n/2 - (d+1/2)(|w_1| + \dots + |w_m|) - m} = 2^{x_1 + \dots + x_m},$$

for any $1 \leq i \leq r$, and therefore,

$$|\{Z = [x \vee w]\}| = r \cdot 2^{x_1 + \dots + x_m}.$$

Recalling the definition of Z , we see that $\{Z = [x \vee w]\}$ determines $\mathcal{C}^*(S, X) = \{(k_i, w_i)\}_{i=1}^m$ up to a rotation of the torus. It is not hard to see that if $l = \text{per}(x \vee w)$ then $r = nl/m$. \square

Proof of Lemma 5.7. We show something stronger. Define

$$D := [(x_1, \dots, x_m)] \text{ where } Z = [((x_i, w_i))_{i=1}^m].$$

Let $m \geq 1$, let $\bar{w} = \{w_1, \dots, w_m\}$ be a multi-set and let $x = (x_1, \dots, x_m)$ be such that $\mathbb{P}(\bar{W} = \bar{w}, D = [x]) > 0$. We shall show that

$$(W \text{ conditioned on } \bar{W} = \bar{w} \text{ and } D = [x]) \stackrel{d}{=} [w_{\pi(1)}, \dots, w_{\pi(m)}],$$

where π is a uniformly chosen permutation of $\{1, 2, \dots, m\}$.

Let w be an ordering of \bar{w} . Define

$$\mathcal{Z}(x, w) := \{[x' \vee w'] \mid x' \in [x], w' \in [w]\},$$

and note that $|\mathcal{Z}(x, w)| = \text{gcd}(\text{per}(x), \text{per}(w))$. We have

$$\mathbb{P}(W = [w] \mid \bar{W} = \bar{w}, D = [x]) = \frac{\mathbb{P}(W = [w], D = [x])}{\mathbb{P}(\bar{W} = \bar{w}, D = [x])} = \frac{\sum_{z \in \mathcal{Z}(x, w)} |\{Z = z\}|}{|\{\bar{W} = \bar{w}\} \cap \{D = [x]\}|}.$$

Let $w' \in [w]$ and $x' \in [x]$ be representatives of their equivalence classes. By Claim 5.8, we have

$$|\{Z = [x' \vee w']\}| = \frac{n \cdot \text{per}(x' \vee w')}{m} \cdot 2^{x_1 + \dots + x_m}.$$

Since, $\text{per}(x' \vee w') = \text{lcm}(\text{per}(x'), \text{per}(w'))$, $\text{per}(x') = \text{per}(x)$ and $\text{per}(w') = \text{per}(w)$, we see that $\mathbb{P}(W = [w] \mid \bar{W} = \bar{w}, D = [x])$ is proportional to

$$|\mathcal{Z}(x, w)| \cdot \text{lcm}(\text{per}(x), \text{per}(w)) = \text{per}(x) \cdot \text{per}(w).$$

That is, conditioned on $\bar{W} = \bar{w}$ and $D = [x]$, the probability that W equals $[w]$ is proportional to $\text{per}(w)$. Finally, observe that the same is true for the probability that $[w_{\pi(1)}, \dots, w_{\pi(m)}]$ equals $[w]$. Indeed, one may check that

$$\mathbb{P}([w_{\pi(1)}, \dots, w_{\pi(m)}] = [w]) = \frac{\text{per}(w)}{C(\bar{w})},$$

where $C(\bar{w})$ is a multinomial coefficient depending on \bar{w} . \square

5.4 Proof of theorems

In this section, we are primarily concerned with homomorphisms on the graph $T_{n,d}$. However, we will occasionally also refer to homomorphisms on the graph $P_{n,d}$. We note that in either case, such a homomorphism can be seen as an element of $\mathbb{Z}^{\{0,1,\dots,n\}}$, where $f \in \text{Hom}(T_{n,d})$ is extended to $\{0, 1, \dots, n\}$ by $f(n) := 0$. Therefore, the uniform distributions on $\text{Hom}(P_{n,d})$ and $\text{Hom}(T_{n,d})$ can be seen as distributions on $\mathbb{Z}^{\{0,1,\dots,n\}}$. We shall denote the probability and expectation with respect to each of these distributions by \mathbb{P}_P and \mathbb{E}_P and \mathbb{P}_T and \mathbb{E}_T , respectively. Throughout this section, we will frequently drop the subscript, in which case \mathbb{P} and \mathbb{E} will refer to \mathbb{P}_T and \mathbb{E}_T .

We first state some technical lemmas and propositions whose proofs we defer to the next section. Our first proposition is one which will allow us to transfer some results from the line to the torus. This is an FKG-type inequality for the measure induced on non-negative homomorphisms by taking pointwise absolute value.

Proposition 5.9. *For any increasing function $\phi: \mathbb{Z}^{\{0,1,\dots,n\}} \rightarrow [0, \infty)$, we have*

$$\mathbb{E}_T[\phi(|f|)] \leq 9 \cdot \mathbb{E}_P[\phi(|f|)].$$

The next two lemmas are concerned with the probability of jumps occurring at given vertices. In the case of the line, we were able to obtain in Lemma 4.2 a good upper bound on the probability of having t jumps at any given vertices. In the case of the torus, we are not able to obtain such a general result. The main difficulty is due to the topological constraint imposed by the torus. In particular, if a jump occurs at a given vertex then a jump in the opposite direction must also occur at some other vertex. The next lemma shows that the probability of a chain of consecutive jumps is still unlikely.

Lemma 5.10. *For any vertex $x \in V$ and any positive integer t , we have*

$$\mathbb{P}(C_{x,t}) \leq C2^{-dt}.$$

The following lemma shows that having jumps in opposing directions at given vertices is also unlikely.

Lemma 5.11. *For $J, J' \subset V$ let $A_{J,J'} := \bigcap_{x \in J, y \in J'} A_{x,y}$ be the event that there are jumps in one direction at all vertices in J and jumps in the opposite direction at all vertices in J' . Then, for any subsets $J, J' \subset V$ of size m each, we have*

$$\mathbb{P}(A_{J,J'}) \leq 2^{-(2d-1)m}.$$

The last lemma is the analog of Corollary 4.10 on the line. It will allow us to deduce the typical order of magnitude of R .

Lemma 5.12. *For any positive even integer n and any positive integers d and r such that $Cdr \leq n$, we have*

$$\mathbb{P}(R = r) \geq \frac{n^2}{r^2 2^{2d+5}} \cdot \left(1 - \frac{Cdr}{n}\right)^2 \cdot \mathbb{P}(R = r - 1).$$

As in the case of homomorphisms on the line, it is also possible to prove an inequality in the opposite direction, showing that $\mathbb{P}(R = r) \leq C \frac{n^2}{r^2 2^{2d}} \mathbb{P}(R = r - 1)$, but we neither use nor prove this.

The supercritical regime. We prove Theorem 2.6 and Corollary 2.7. By Lemma 5.11 and by the union bound, we have

$$\mathbb{P}(R \geq r) = \mathbb{P}\left(\bigcup_{\substack{J, J' \subset V \\ |J|=|J'|=r}} A_{J,J'}\right) \leq \binom{n}{r}^2 2^{-(2d-1)r}, \quad r \geq 1.$$

One may easily check that $|\text{Rng}(f)| \leq R + 3$, so that

$$\mathbb{P}(|\text{Rng}(f)| \geq 3 + r) \leq \binom{n}{r}^2 2^{-(2d-1)r}, \quad r \geq 1. \tag{5.9}$$

Moreover, it is easy to describe all homomorphisms which take on at most 3 values. Let Ω_0 be the set of homomorphisms which are constant on the even vertices (having the value 0 on the even vertices and 1 or -1 on the odd vertices), and let Ω_1 be the set of homomorphisms which are constant on the odd vertices (having the value ± 1 on the odd vertices, and 0 or ± 2 , respectively, on the even vertices). Then $\{|\text{Rng}(f)| \leq 3\} = \Omega_0 \cup \Omega_1$, $|\Omega_0 \cap \Omega_1| = 2$, and $|\Omega_0| = |\Omega_1| = 2^{n/2}$. Therefore,

$$\mathbb{P}(|\text{Rng}(f)| < 3) = \mathbb{P}(\Omega_0 \cap \Omega_1) \leq \frac{|\Omega_0 \cap \Omega_1|}{|\Omega_0 \cup \Omega_1|} = \frac{2}{2^{n/2+1} - 2} \leq 2^{1-n/2},$$

completing the proof of Theorem 2.6. To obtain Corollary 2.7, recall that $|\Omega_0| = |\Omega_1|$ and note that if $d - \log_2 n \rightarrow \infty$ as $n \rightarrow \infty$ then $\mathbb{P}(\Omega_0 \cup \Omega_1) = \mathbb{P}(|\text{Rng}(f)| \leq 3) = 1 - o(1)$, by Theorem 2.6.

We remark that the bound (5.9) obtained for the probability that the range is large constitutes something of a compromise between two possibilities. With somewhat less work we could have used the FKG-type inequality, Proposition 5.9, to obtain a weaker bound. With somewhat more work we could make a finer analysis of the possible cases in the proof of Lemma 5.11 and obtain a somewhat better bound, with $2d - 1$ replaced by $2d + 1$ or even $2d + 2$. The bound we chose to prove has the benefit that it is already rather good and has a relatively simple proof.

The subcritical regime. We begin by proving the upper bound in Theorem 2.8. Since $\max_{0 \leq k \leq n} |f(k)|$ is an increasing function in $|f|$, we have by Proposition 5.9,

$$\mathbb{E}_T \left[\max_{0 \leq k \leq n} |f(k)| \right] \leq 9 \cdot \mathbb{E}_P \left[\max_{0 \leq k \leq n} |f(k)| \right].$$

By our previous result on the line, Theorem 2.3, we have

$$\mathbb{E}_P \left[\max_{0 \leq k \leq n} |f(k)| \right] \leq \mathbb{E}_P [|\text{Rng}(f)| - 1] \leq 2 + C\sqrt{n2^{-d}},$$

and then, using symmetry,

$$\mathbb{E}_T [|\text{Rng}(f)|] = \mathbb{E}_T \left[1 + \max_{0 \leq k \leq n} f(k) - \min_{0 \leq k \leq n} f(k) \right] \leq 1 + 2 \cdot \mathbb{E}_T \left[\max_{0 \leq k \leq n} |f(k)| \right] \leq 37 + C\sqrt{n2^{-d}}.$$

We now prove the lower bound in Theorem 2.8. The proof is very similar to that of the lower bound in Theorem 2.3, and so we only give an outline. First, we show that for any $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\mathbb{P} \left(|\text{Rng}(f)| < 3 + \lfloor \delta\sqrt{n2^{-d}} \rfloor \right) \leq \epsilon + 2^{1-n/2}.$$

Let $\epsilon > 0$. Note that, by Theorem 2.6, the statement is trivial when $\delta\sqrt{n2^{-d}} < 1$. Hence, we may assume that $n2^{-d} \geq C/\epsilon$. Mimicking the proof of Lemma 4.11 and its corollary, using Lemma 5.12 in place of Lemma 4.9, we find that there exists a $\delta_1 > 0$ such that

$$\mathbb{P} (R < \delta_1 n2^{-d}) \leq \epsilon/4.$$

Continuing as in (4.22) - (4.24), using Lemma 5.10 in place of Lemma 4.2, we obtain

$$\mathbb{P} (|\mathcal{C}(S)| \geq \delta_2 n2^{-d}) \geq 1 - \epsilon/2, \tag{5.10}$$

for some $\delta_2 > 0$. Proposition 5.4 and Proposition 3.6 imply that

$$\mathbb{P}(|\text{Rng}(f)| < r \mid |\mathcal{C}(S)|) \leq \frac{Cr}{\sqrt{|\mathcal{C}(S)|}}, \quad r \in \mathbb{N}. \tag{5.11}$$

Now, putting (5.10) and (5.11) together, we see that there exists a $\delta > 0$ such that

$$\mathbb{P}\left(|\text{Rng}(f)| < 3 + \lfloor \delta \sqrt{n2^{-d}} \rfloor\right) \leq \mathbb{P}\left(|\text{Rng}(f)| < \lfloor 4\delta \sqrt{n2^{-d}} \rfloor\right) \leq \epsilon.$$

Finally, repeating the calculation in (4.26), where we use Theorem 2.6 in place of Theorem 2.1, we obtain

$$\mathbb{E}[|\text{Rng}(f)|] \geq 3 + \lfloor c\sqrt{n2^{-d}} \rfloor - 2^{1-n/2}.$$

The critical regime. Here we prove Theorem 2.9. Denote $\lambda := \lim n2^{-d}$ which exists and is a positive number by assumption. We begin by showing that in the critical regime the jumps are sparse enough so that it is unlikely to have chains of length greater than one. Let

$$B := \bigcap_{x \in V} (A_x^c \cup A_{x+2d+1}^c)$$

be the event that there are no two minimal-distance jumps (i.e. jumps at distance $2d+1$). We wish to show that $\mathbb{P}(B) = 1 - o(1)$. Indeed, by Lemma 5.10, we have

$$\mathbb{P}(B^c) \leq \sum_{x \in V} \mathbb{P}(A_x \cap A_{x+2d+1}) \leq Cn2^{-2d} = o(1).$$

We now find the limiting distribution of R as n tends to infinity. By Lemma 5.2 and Claim 5.1, we have that $\mathbb{P}(R = r \mid B)$ is proportional to

$$\begin{cases} c(n, d, r) \cdot \binom{2r}{r} \cdot 2^{n/2-r(2d+3)} & \text{if } r \geq 1 \\ 2^{n/2+1} - 2 & \text{if } r = 0 \end{cases}, \tag{5.12}$$

where $c(n, d, r)$ is the number of feasible jump structures I having $|I| = |\mathcal{C}(I)| = 2r$.

Claim 5.13. For any $r \geq 1$, we have

$$c(n, d, r) = \frac{n}{2r} \cdot \binom{n/2 - (2d+1)r - 1}{2r - 1}.$$

Proof. Denote by \mathcal{I} the set of all feasible jump structures I having $|I| = |\mathcal{C}(I)| = 2r$. For $v \in V$, let $\mathcal{I}_v := \{I \in \mathcal{I} \mid v \in I\}$. Then,

$$n|\mathcal{I}_0| = \sum_{v \in V} |\mathcal{I}_v| = \sum_{v \in V} \sum_{I \in \mathcal{I}} \mathbf{1}_I(v) = \sum_{I \in \mathcal{I}} \sum_{v \in V} \mathbf{1}_I(v) = \sum_{I \in \mathcal{I}} |I| = 2r|\mathcal{I}| = 2r \cdot c(n, d, r).$$

It remains to compute the size of \mathcal{I}_0 . By considering the distances between consecutive elements in any $I \in \mathcal{I}_0$, and recalling (5.2), (5.3) and (5.4), we see that $|\mathcal{I}_0|$ is given by the number of non-negative integer solutions to the equation

$$x_1 + x_2 + \dots + x_{2r} = n,$$

under the additional constraint that, for $1 \leq j \leq 2r$, x_j is odd and at least $2d + 3$. Therefore, after substituting $x_j = 2y_j + 2d + 3$ for $1 \leq j \leq 2r$, we see that $|\mathcal{I}_0|$ is equal to the number of non-negative integer solutions to the equation

$$y_1 + \dots + y_{2r} = n/2 - (2d + 3)r,$$

from which the result now follows. □

By (5.12) and Claim 5.13, for any fixed $r \geq 1$, we have

$$\frac{\mathbb{P}(R = r \mid B)}{\mathbb{P}(R = r - 1 \mid B)} = \frac{n^2}{r^2 2^{2d+5}} \cdot (1 + o(1)) = \frac{(\lambda')^2}{r^2} \cdot (1 + o(1)),$$

where $\lambda' := \lambda/(4\sqrt{2})$. By Lemma 5.10 and since $\mathbb{P}(B) = 1 - o(1)$, we have

$$\mathbb{E}[R \mid B] = \frac{1}{2} \sum_{x \in V} \mathbb{P}(A_x \mid B) \leq \frac{1}{2\mathbb{P}(B)} \sum_{x \in V} \mathbb{P}(A_x) \leq \frac{Cn2^{-d}}{\mathbb{P}(B)} = O(1).$$

Therefore, conditioned on B , the expectation of R is uniformly bounded as $n \rightarrow \infty$. Hence, Markov's inequality implies that, conditioned on B , R is tight as $n \rightarrow \infty$. Recall the definition of the distribution $\nu(\lambda')$ in (2.5). Let $N(\lambda') \sim \nu(\lambda')$ and note that

$$\frac{\mathbb{P}(N(\lambda') = r)}{\mathbb{P}(N(\lambda') = r - 1)} = \frac{(\lambda')^2}{r^2}, \quad r \geq 1.$$

Thus, Lemma 3.7 implies that, conditioned on B , R converges in distribution to $N(\lambda')$. Finally, since $\mathbb{P}(B) = 1 - o(1)$, we conclude that R converges in distribution to $\nu(\lambda')$.

It remains to understand the range of a homomorphism. Recalling the definition of \bar{W} given in (5.6), we observe that the event B is the same as the event $\{|\bar{W}| = |\mathcal{C}(S)| = 2R\}$, which is the same as the event that \bar{W} consists of R 1's and R (-1)'s. Therefore, by Proposition 5.4, conditioned on B and on R , on the event $\{R > 0\}$, the range of a homomorphism is equal in distribution to two plus the range of a random walk bridge of length $2R$. By Theorem 2.6, conditioned on the event $\{R = 0\}$, the range of a homomorphism is 3 with probability tending to one. This, together with our previous result on the convergence of R in distribution, completes the proof of Theorem 2.9.

5.5 Proof of main lemmas

Proof of Proposition 5.9. Recall that a homomorphism on $P_{n,d}$ or $T_{n,d}$ can be seen as an element of $\mathbb{Z}^{\{0,1,\dots,n\}}$, where $f \in \text{Hom}(T_{n,d})$ is extended to $\{0, 1, \dots, n\}$ by $f(n) := 0$. Therefore, the uniform distributions on $\text{Hom}(P_{n,d})$ and $\text{Hom}(T_{n,d})$ are distributions on $\mathbb{Z}^{\{0,1,\dots,n\}}$, which we denote by \mathbb{P}_P and \mathbb{P}_T respectively. We also denote by $Q := \{f \in \text{Hom}(T_{n,d})\}$ the support of \mathbb{P}_T , so that the measure $\mathbb{P}_P(\cdot \mid Q)$ is just the measure \mathbb{P}_T .

Let $\phi: \mathbb{Z}^{\{0,1,\dots,n\}} \rightarrow [0, \infty)$ be an increasing function. Note that the event $\{f(n) = 0\}$ is a decreasing event in $|f|$. Therefore, we can apply Theorem 3.4 for the functions ϕ and $\psi(f) := \mathbf{1}_{\{f(n)=0\}}$, to obtain

$$\mathbb{E}_P[\phi(|f|) \mid f(n) = 0] \leq \mathbb{E}_P[\phi(|f|)]. \tag{5.13}$$

Notice that sampling a random homomorphism on $P_{n,d}$ conditioned on $\{f(n) = 0\}$ is not equivalent to sampling a random homomorphism on $T_{n,d}$, which is just to say that $Q \neq \{f(n) = 0\}$. However, it is equivalent to sampling a random homomorphism on another graph. Namely, the graph $P'_{n,d}$ obtained from $P_{n,d}$ by identifying the vertex 0 with the vertex n . In order to obtain $T_{n,d}$ from $P'_{n,d}$, we must still add some edges which are missing, for example, the edge between $n - 1$ and 2, and the edge between $n - 2$ and 1. Nonetheless, this observation shows that the measure $\mathbb{P}_P(\cdot \mid f(n) = 0)$ also satisfies the FKG inequality in Theorem 3.4, since it is equivalent to sampling a random homomorphism on $P'_{n,d}$. Define the events

$$J := \{|f(k)| \leq 1, k = 0, 1, \dots, 2d\} \quad \text{and} \\ J' := \{|f(k)| \leq 1, k = n, n - 1, \dots, n - 2d\}.$$

Notice that

$$J \cap J' \subset Q \subset \{f(n) = 0\}.$$

Thus, using (5.13) and the fact that ϕ is non-negative, we obtain

$$\mathbb{E}_P[\phi(|f|) \mid Q] \leq \mathbb{E}_P[\phi(|f|) \mid f(n) = 0] \cdot \frac{\mathbb{P}_P(f(n) = 0)}{\mathbb{P}_P(Q)} \leq \frac{\mathbb{E}_P[\phi(|f|)]}{\mathbb{P}_P(J \cap J' \mid f(n) = 0)}.$$

We now wish to bound $\mathbb{P}(J \cap J' \mid f(n) = 0)$ from below. We first apply Theorem 3.4 to the graph $P'_{n,d}$ to get

$$\mathbb{P}_P(J \cap J' \mid f(n) = 0) \geq \mathbb{P}_P(J \mid f(n) = 0) \cdot \mathbb{P}_P(J' \mid f(n) = 0) = \mathbb{P}_P(J \mid f(n) = 0)^2,$$

where we have used symmetry in the second step. Next, we apply Theorem 3.4 again to the graph $P_{n,d}$ to get

$$\mathbb{P}_P(J \mid f(n) = 0) \geq \mathbb{P}_P(J).$$

Finally, since J is just the event that no jump occurs at the first $2d + 1$ vertices, we have by Lemma 4.1 that $\mathbb{P}_P(J) \geq 1/3$. Therefore, we have shown that

$$\mathbb{E}_T[\phi(|f|)] = \mathbb{E}_P[\phi(|f|) \mid Q] \leq 9 \cdot \mathbb{E}_P[\phi(|f|)]. \quad \square$$

For the proofs of the remaining lemmas, it is convenient to denote by $[x, y]$ the vertices on the arc going from x to y in the clockwise direction. That is, for $x, y \in V$, we define

$$[x, y] := \{z \in V \mid \rho^+(x, z) \leq \rho^+(x, y)\}.$$

Also, for a set $J \subset V$ and an integer t , we let

$$J + t := \{j + t \mid j \in J\}$$

where, as always, addition for vertices on the torus is taken modulo n .

Proof of Lemma 5.10. First, we partition $C_{x,t}$ into two events $C_{x,t}^0$ and $C_{x,t}^1$. Denote $x' := x - (2d + 1)t - 1$ and define

$$\begin{aligned} C_{x,t}^0 &:= C_{x,t} \cap \{|f(x) - f(x' - 1)| = 2 + t\}, \\ C_{x,t}^1 &:= C_{x,t} \cap \{|f(x) - f(x' - 1)| = t\}. \end{aligned}$$

Note that since any $f \in C_{x,t}$ has $|f(x) - f(x')| = 1 + t$, we indeed have $C_{x,t} = C_{x,t}^0 \cup C_{x,t}^1$. Also, observe that

$$C_{x,t}^0 = \{|f(x) - f(x' - 1)| \geq 2 + t\}, \tag{5.14}$$

since for any $f \in \text{Hom}(T_{n,d})$, $|f(x) - f(x' - 1)| \leq 2 + t$ necessarily holds and $|f(x) - f(x' - 1)| = 2 + t$ holds only if there is a chain of length t at x .

We now prove that

$$\mathbb{P}(C_{x,t}^1) \leq \mathbb{P}(C_{x,t}^0). \tag{5.15}$$

By rotating the torus if necessary, we note that it suffices to prove this under the assumption that $x' \neq 1$. This simplifies slightly the following discussion as it avoids issues stemming from the fact that $f(0)$ is normalized to be 0.

Consider the mapping $f \mapsto f_0$ from $C_{x,t}^1$ to $C_{x,t}^0$ defined by

$$f_0(y) := \begin{cases} f(y) & \text{if } y \neq x' - 1 \\ 2f(y + 1) - f(y) & \text{if } y = x' - 1 \end{cases}, \quad y \in V.$$

Random walk with long-range constraints

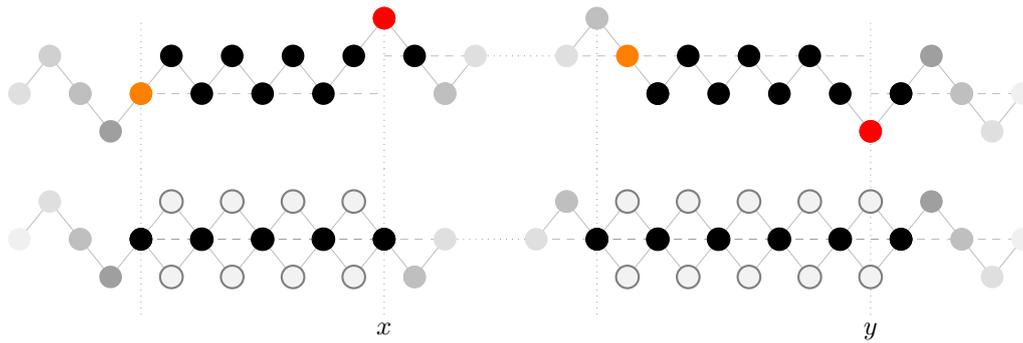


Figure 9: An illustration of the operation of “removing jumps” from a homomorphism. Given a homomorphism having jumps at x and y in opposite directions, we may remove these jumps and gain entropy in the newly formed fluctuation points. Here $d = 3$.

That is, the sign of a chain in $\mathcal{C}(I^{x,y})$ is inherited from its corresponding chain in $\mathcal{C}(I)$. Note that,

$$\sum_{(k,t) \in \mathcal{C}(I^{x,y})} \epsilon^{x,y}(k,t) \cdot t = -\epsilon(C(I,x)) - \epsilon(C(I,y)).$$

Hence, if $\epsilon(C(I,x)) \neq \epsilon(C(I,y))$ then $\epsilon^{x,y} \in \mathcal{B}^*(I^{x,y})$ and, by (5.4), $I^{x,y}$ is a feasible jump structure.

Denote by \mathcal{I} the set of all feasible jump structures. For subsets $J, J' \subset V$, denote by $B(J, J')$ the set of all feasible signed jump structures containing $J \cup J'$ and having different signs on J and J' , i.e.,

$$B(J, J') := \{(I, \epsilon) \mid I \in \mathcal{I}, \epsilon \in \mathcal{B}^*(I), J \cup J' \subset I, \epsilon(C(I,x)) \neq \epsilon(C(I,y)) \forall x \in J, y \in J'\}.$$

Then, we have the equality of events,

$$A_{J,J'} = \{(S, X) \in B(J, J')\}.$$

Therefore, the lemma is equivalent to

$$\mathbb{P}((S, X) \in B(J, J')) \leq 2^{-(2d-1)m}, \quad m \geq 0, J, J' \subset V, |J| = |J'| = m. \quad (5.18)$$

We prove this by induction on m . The induction base, $m = 0$, and the case when $B(J, J') = \emptyset$ are trivial. Suppose that $m \geq 1$ and let $J, J' \subset V$ be such that $B(J, J') \neq \emptyset$. We choose two vertices $x \in J$ and $y \in J'$ such that $[x, y] \cap (J \cup J') = \{x, y\}$ and define a mapping

$$T: B(J, J') \rightarrow B(J \setminus \{x\}, J' \setminus \{y\})$$

by

$$T(I, \epsilon) := (I^{x,y}, \epsilon^{x,y}).$$

Note that the mapping $I \mapsto I^{x,y}$ is injective on $\{I \in \mathcal{I} \mid x, y \in I\}$. Thus, recalling that jumps belonging to the same chain must have the same sign, it is not hard to see that, for any $(I', \epsilon') \in B(J \setminus \{x\}, J' \setminus \{y\})$, we have

$$|T^{-1}(I', \epsilon')| \leq \begin{cases} 1 & \text{if } I' \cap D'_{x,y} \neq \emptyset \\ 2 & \text{if } I' \cap D'_{x,y} = \emptyset \end{cases},$$

where $D'_{x,y} := \{x - 2d - 1, x + 2d, y - 2d - 2, y + 2d + 1\}$. Moreover, by Lemma 5.3 and Claim 5.1, for any $(I, \epsilon) \in B(J, J')$,

$$\mathbb{P}((S, X) = T(I, \epsilon)) = \mathbb{P}((S, X) = (I, \epsilon)) \cdot \begin{cases} 2^{2d+1+|C(I)|-|C(I^{x,y})|} & \text{if } I^{x,y} \neq \emptyset \\ 2^{2d+3}(2 - 2^{1-n/2}) & \text{if } I^{x,y} = \emptyset \end{cases}.$$

Thus, by (5.16) and (5.17), for any $(I, \epsilon) \in B(J, J')$,

$$\mathbb{P}((S, X) = T(I, \epsilon)) \geq \mathbb{P}((S, X) = (I, \epsilon)) \cdot 2^{2d-1} \cdot \begin{cases} 1 & \text{if } I \cap D_{x,y} \neq \emptyset \\ 16 & \text{if } I \cap D_{x,y} = \emptyset \end{cases}.$$

Denote

$$B := \{(I, \epsilon) \in B(J, J') \mid I \cap D_{x,y} = \emptyset\}, \\ B' := \{(I', \epsilon') \in B(J \setminus \{x\}, J' \setminus \{y\}) \mid I' \cap D'_{x,y} = \emptyset\}.$$

Note that T maps B into B' and $B(J, J') \setminus B$ into $B(J \setminus \{x\}, J' \setminus \{y\}) \setminus B'$. Therefore, applying Lemma 3.1 to the restriction of T to B and separately to its restriction to $B(J, J') \setminus B$, we obtain

$$\begin{aligned} \mathbb{P}((S, X) \in B(J, J')) &= \mathbb{P}((S, X) \in B) + \mathbb{P}((S, X) \in B(J, J') \setminus B) \\ &\leq 2^{-(2d+2)} \cdot \mathbb{P}((S, X) \in B') + 2^{-(2d-1)} \cdot \mathbb{P}((S, X) \in B(J \setminus \{x\}, J' \setminus \{y\}) \setminus B') \\ &\leq 2^{-(2d-1)} \cdot \mathbb{P}((S, X) \in B(J \setminus \{x\}, J' \setminus \{y\})). \end{aligned}$$

Thus, (5.18) follows by induction, proving the lemma. □

Proof of Lemma 5.12. The proof utilizes a similar technique as the proof of Lemma 5.11, where this time we aim to add jumps to the jump structure of the homomorphism rather than remove jumps.

For a feasible jump structure I , denote

$$U(I) := \begin{cases} \{(x, y) \in V^2 \mid \rho(x, y) \geq 2d + 3, \rho^+(I, x), \rho^+(x, I) + 1, \rho^+(y, I), \rho^+(I, y) - 1 \in D\} & \text{if } I \neq \emptyset \\ \{(x, y) \in V^2 \mid \rho(x, y) \in D, \rho(0, x) \text{ is even}\} & \text{if } I = \emptyset \end{cases},$$

where $\rho^+(x, I) := \min_{s \in I} \rho^+(x, s)$, $\rho^+(I, x) := \min_{s \in I} \rho^+(s, x)$ and

$$D := \{2d + 3, 2d + 5, \dots\}.$$

For a feasible jump structure I and a pair $(x, y) \in U(I)$, define

$$I_{x,y} := (I \cap [y, x]) \cup ((I \cap [x, y]) + 1) \cup \{x, y\}.$$

Note that $I_{x,y}$ satisfies condition (5.2). Moreover, it is easy to see that the chain structure of $I_{x,y}$ satisfies

$$\mathcal{C}(I_{x,y}) = \mathcal{C}(I \cap [y, x]) \cup \mathcal{C}((I \cap [x, y]) + 1) \cup \{(x, 1), (y, 1)\}.$$

Now, for a feasible sign vector $\epsilon \in \mathcal{B}^*(I)$ and a sign $i \in \{-1, 1\}$, define $\epsilon_{x,y,i} \in \mathcal{B}^*(I_{x,y})$ by

$$\epsilon_{x,y,i}(k, t) := \begin{cases} \epsilon(k, t) & \text{if } k \in I \\ \epsilon(k - 1, t) & \text{if } k - 1 \in I \\ i & \text{if } k = x \\ -i & \text{if } k = y \end{cases}, \quad (k, t) \in \mathcal{C}(I_{x,y}).$$

That is, the sign of a chain in $\mathcal{C}(I_{x,y}) \setminus \{(x, 1), (y, 1)\}$ is inherited from its corresponding chain in $\mathcal{C}(I)$ and the sign of the chain at x , which is opposite of that of y , is determined independently. Note that, by (5.4), $I_{x,y}$ is a feasible jump structure. Moreover, since $|I_{x,y}| = |I| + 2$ and $|\mathcal{C}(I_{x,y})| = |\mathcal{C}(I)| + 2$, Lemma 5.3 and Claim 5.1 imply that

$$\mathbb{P}((S, X) = (I_{x,y}, \epsilon_{x,y,i})) = \mathbb{P}((S, X) = (I, \epsilon)) \cdot 2^{-2d-3} \cdot \begin{cases} 1 & \text{if } I \neq \emptyset \\ (2 - 2^{1-n/2})^{-1} & \text{if } I = \emptyset \end{cases}. \quad (5.19)$$

Denote by \mathcal{I} the set of all feasible jump structures. For $r \geq 0$, let B_r denote the set of feasible signed jump structures having $2r$ jumps, i.e.,

$$B_r := \{(I, \epsilon) \mid I \in \mathcal{I}, \epsilon \in \mathcal{B}^*(I), |I| = 2r\}.$$

For $r \geq 1$, define the mapping

$$T_r : B_{r-1} \rightarrow \mathcal{P}(B_r)$$

by

$$T_r(I, \epsilon) := \{(I_{x,y}, \epsilon_{x,y,i}) \mid (x, y) \in U(I), i \in \{-1, 1\}\}.$$

Assume henceforth that $n \geq Cdr$. Then, since the mapping $((x, y), i) \mapsto (I_{x,y}, \epsilon_{x,y,i})$ is injective on $U(I) \times \{-1, 1\}$, we have

$$|T_r(I, \epsilon)| = 2|U(I)| \geq 2(n/2 - Cdr)^2.$$

Therefore, by (5.19),

$$\mathbb{P}((S, X) \in T_r(I, \epsilon)) \geq \mathbb{P}((S, X) = (I, \epsilon)) \cdot \frac{(n - Cdr)^2}{2^{2d+5}} \cdot \begin{cases} 2 & \text{if } r \geq 2 \\ 1 & \text{if } r = 1 \end{cases}.$$

For $(I', \epsilon') \in B_r$, denote

$$N_r(I', \epsilon') := \{(I, \epsilon) \in B_{r-1} \mid (I', \epsilon') \in T_r(I, \epsilon)\}.$$

We have

$$|N_r(I', \epsilon')| \leq r^2 \cdot \begin{cases} 2 & \text{if } r \geq 2 \\ 1 & \text{if } r = 1 \end{cases}.$$

Thus, considering separately the case $r = 1$, Lemma 3.1 implies that for any $r \geq 1$,

$$\mathbb{P}(R = r - 1) = \mathbb{P}((S, X) \in B_{r-1}) \leq \mathbb{P}((S, X) \in B_r) \cdot \frac{r^2 2^{2d+5}}{(n - Cdr)^2} = \mathbb{P}(R = r) \cdot \frac{r^2 2^{2d+5}}{(n - Cdr)^2}. \quad \square$$

6 Local limits on the line

In this section we prove the theorems which were stated in Section 2.3. Throughout this section, the parameter $d \geq 1$ is fixed, and so we drop the d from the notation when convenient. On the other hand, the parameter $n \geq 1$ is allowed to vary, and our main goal is to understand $\text{Hom}(P_{n,d}) := \text{Hom}(P_{n,d}, 0)$ as n grows larger. At first, in Section 6.2, we investigate the asymptotic size of $\text{Hom}(P_{n,d})$ as n tends to infinity. Subsequently, in Sections 6.3 and 6.4, we describe the local limit of such homomorphisms as a probability measure on infinite homomorphisms defined through a Markov chain (see Figure 12).

6.1 Definitions

Given a finite set Π , called an *alphabet*, we denote by Π^* the set of all finite words on Π . That is,

$$\Pi^* := \{(a_1, a_2, \dots, a_t) \mid a_i \in \Pi, t \geq 0\}.$$

For $u, v \in \Pi^*$, we denote the length of u by $|u|$ and the *concatenation* of u and v by $u \circ v$, i.e.,

$$u \circ v := (u_1, \dots, u_{|u|}, v_1, \dots, v_{|v|}).$$

It is clear that concatenation is associative. For $u \in \Pi^*$ with $|u| \geq 1$, let u^- be the word obtained from u by dropping the last element, i.e.,

$$u^- := (u_1, \dots, u_{|u|-1}).$$

Define the derivative operator $D_n: \text{Hom}(P_{n,d}) \rightarrow \{-1, 1\}^n$ by

$$(D_n(f))_k := f(k) - f(k-1), \quad 1 \leq k \leq n. \tag{6.1}$$

Denote by $\mathcal{D} \subset \{-1, 1\}^*$ the set of words on $\{-1, 1\}$ which do not contain $(-1, -1, -1)$ or $(1, 1, 1)$ as a subsequence, and note that $D_n(\text{Hom}(P_{n,d})) \subset \mathcal{D}$. Let Σ be the four letter alphabet

$$\Sigma := \{a, b, A, B\},$$

where

$$\begin{aligned} a &:= (1, -1), \\ b &:= (-1, 1), \\ A &:= (1, 1, -1), \\ B &:= (-1, -1, 1). \end{aligned}$$

These basic sequences will serve as a means to encode homomorphisms into words (see Figure 10). Define $T': \Sigma^* \rightarrow \{-1, 1\}^*$ by

$$T'(x) := x_1 \circ x_2 \circ \dots \circ x_{|x|}.$$

For $x \in \Sigma^*$, define the *weight* of x by

$$w(x) := |T'(x)| = \sum_{k=1}^{|x|} |x_k|. \tag{6.2}$$

Now, we define a mapping $T: \mathcal{D} \rightarrow \Sigma^*$ recursively by the relations

$$\begin{aligned} T(()) &:= (), \\ T((1)) &:= (a), \\ T((-1)) &:= (b), \\ T((1, 1)) &:= (A), \\ T((-1, -1)) &:= (B), \\ T(u \circ v) &:= (u) \circ T(v) \quad \text{for } u \in \Sigma \text{ and } v \in \mathcal{D}. \end{aligned} \tag{6.3}$$

It is not hard to see that T is indeed well-defined (see Figure 11), and that it maps a word $u \in \mathcal{D}$ to the unique word $x \in \Sigma^*$ satisfying $T'(x) = u$ or $T'(x)^- = u$ (in which case $w(x) = |u|$ or $w(x) = |u|+1$, respectively). Also, one should note that $T^{-1}(x) = \emptyset$ if $x \in \Sigma^*$ contains (a, B) or (b, A) as a sub-word or if $x = \emptyset$, and that $T^{-1}(x) = \{T'(x), T'(x)^-\}$ otherwise.

Another observation which will be useful later on is that the recursive relation in the last line of (6.3) may be generalized to hold for certain $u \in \mathcal{D}$.

Claim 6.1. *We have*

$$T(u \circ v) = T(u) \circ T(v) \quad \text{for } u, v \in \mathcal{D} \text{ such that } |u| = w(T(u)). \tag{6.4}$$

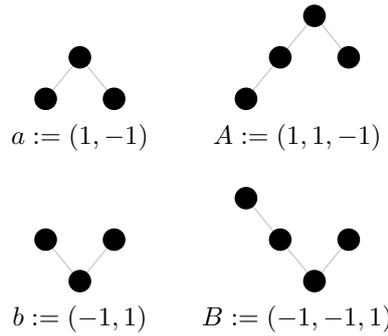


Figure 10: The basic building blocks for encoding a homomorphism into a word on the alphabet $\Sigma := \{a, b, A, B\}$.

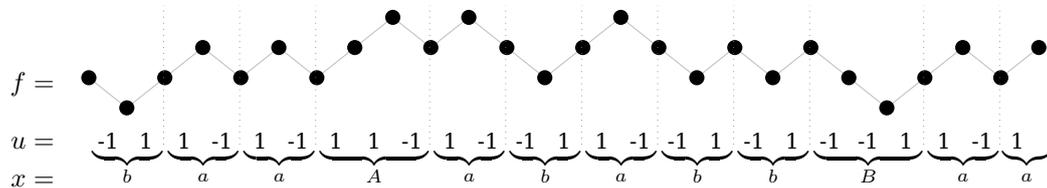


Figure 11: A homomorphism $f \in \text{Hom}(P_{n,d})$ is first viewed as a word $u := D_n(f)$ of length n on the alphabet $\{-1, 1\}$. Then, u is encoded into a word $x := T(u)$ on the alphabet Σ by sequentially reading off the letters from left to right, as defined in the recursive formula in (6.3). If this process exhausts u completely then we end up with a word x of weight exactly n . Otherwise, we remain with a tail of u of length one or two (as is the case in this figure), which is a prefix of at least one element in Σ . In this case, the last letter is chosen in such a way that the weight of the resulting word x is $n + 1$, as defined by the base cases in (6.3).

Proof. We prove the claim by induction on $|u|$. If $|u| = 0$ then there is nothing to prove. Otherwise, $|u| \geq 1$. By the assumption, we have $|u| = w(T(u))$, which implies that u may be decomposed as $u = u' \circ u''$, where $u' \in \Sigma$. Note that this now implies that $|u''| = w(T(u'')) < |u|$, since $T(u) = (u') \circ T(u'')$ by (6.3). Therefore, by (6.3) and induction,

$$\begin{aligned} T(u \circ v) &= T(u' \circ u'' \circ v) = (u') \circ T(u'' \circ v) \\ &= (u') \circ T(u'') \circ T(v) = T(u' \circ u'') \circ T(v) = T(u) \circ T(v). \end{aligned} \quad \square$$

We say a word $x \in \Sigma^*$ is d -legal if it satisfies the conditions

$$\begin{aligned} x_m = A &\Rightarrow x_{m-i} = a, & \forall i \in \{1, \dots, d-1\} \text{ such that } i < m, \\ x_m = B &\Rightarrow x_{m-i} = b, & \forall i \in \{1, \dots, d-1\} \text{ such that } i < m, \\ x_m = A, m > d &\Rightarrow x_{m-d} \in \{a, A\}, \\ x_m = B, m > d &\Rightarrow x_{m-d} \in \{b, B\}. \end{aligned} \quad (6.5)$$

Denote by $\Omega_{n,d}$ the set of d -legal words on Σ of weight n or $n + 1$. That is,

$$\Omega_{n,d} := \{x \in \Sigma^* \mid x \text{ is } d\text{-legal, } w(x) \in \{n, n + 1\}\}.$$

Define

$$L_n := (T \circ D_n)|_{\text{Hom}(P_{n,d})}.$$

Claim 6.2. *The mapping L_n is a bijection between $\text{Hom}(P_{n,d})$ and $\Omega_{n,d}$.*

Proof. It is clear from (6.2) and (6.3) that L_n injectively maps $\text{Hom}(P_{n,d})$ to words on Σ of weight n or $n + 1$. It remains to show that the image of L_n is precisely $\Omega_{n,d}$.

One may easily see that a homomorphism $f \in \text{Hom}(P_{n,1})$ is a homomorphism in $\text{Hom}(P_{n,d})$ if and only if $D_n(f)$ does not contain a sequence of the form

$$u_{\pm}^l := \pm(1, \underbrace{1, -1, \dots, 1, -1}_{2l}, 1, 1),$$

with $1 \leq l \leq d - 1$. Indeed, $D_n(f)$ contains u_{\pm}^l if and only if there exist $0 \leq i, j \leq n$ such that $|i - j| = 2l + 3$ and $|f(i) - f(j)| = 3$. Now, it is also not hard to check that u_{\pm}^l appears in $D_n(f)$ if and only if $L_n(f)$ contains a subword of the form

$$(A, \underbrace{a, \dots, a}_{l-1}, A) \quad \text{or} \quad (B, \underbrace{a, \dots, a}_l, A) \quad \text{or} \quad (b, \underbrace{a, \dots, a}_l, A),$$

depending on the position of u_{\pm}^l in $D_n(f)$. The same is true for u_{\pm}^l with $\{a, A\}$ and $\{b, B\}$ interchanged. Therefore, by (6.5), we see that u_{\pm}^l appears in $D_n(f)$, for some $1 \leq l \leq d - 1$, if and only if $L_n(f)$ is not d -legal.

We have shown that for any $f \in \text{Hom}(P_{n,1})$, $f \in \text{Hom}(P_{n,d})$ if and only if $L_n(f) \in \Omega_{n,d}$. In particular, since $\text{Hom}(P_{n,d}) \subset \text{Hom}(P_{n,1})$, we have $L_n(\text{Hom}(P_{n,d})) \subset \Omega_{n,d}$. For the other direction, let $x \in \Omega_{n,d}$. Either $T'(x)$ or $T'(x)^-$ is of length n . Let $u \in \mathcal{D}$ be this sequence and let $f := D_n^{-1}(u) \in \text{Hom}(P_{n,1})$. Since $L_n(f) = T(u) = x$ is d -legal, we see that $f \in \text{Hom}(P_{n,d})$. Hence, $\Omega_{n,d} \subset L_n(\text{Hom}(P_{n,d}))$, completing the proof. \square

6.2 Counting the homomorphisms

In this section we prove Theorem 2.10. This is done by deriving a recursion formula and investigating its characteristic polynomial.

For $0 \leq k, m \leq d$, define

$$\Omega_{n,d}(k, m) := \{x \in \Omega_{n,d} \mid x_1, \dots, x_k \neq A \text{ and } x_1, \dots, x_m \neq B\}.$$

By symmetry we have $|\Omega_{n,d}(k, m)| = |\Omega_{n,d}(m, k)|$, so we can define

$$c_n(k) := |\Omega_{n,d}(k, d)| = |\Omega_{n,d}(d, k)|, \quad 0 \leq k \leq d - 1.$$

This definition is motivated by the following two lemmas, which show that the $c_n(k)$ satisfy some explicit recursion formulas and that they have a simple relation to $|\text{Hom}(P_{n,d})|$.

Lemma 6.3. *For any $n \geq 3$, we have*

$$|\text{Hom}(P_{n,d})| = 2c_{n-2}(0) + 2c_{n-3}(d - 1).$$

Proof. Note that by (6.5), we have

$$\begin{aligned} \{x \in \Omega_{n,d} \mid x_1 = a\} &= \{(a) \circ x \mid x \in \Omega_{n-2,d}(0, d)\}, \\ \{x \in \Omega_{n,d} \mid x_1 = A\} &= \{(A) \circ x \mid x \in \Omega_{n-3,d}(d - 1, d)\}. \end{aligned} \tag{6.6}$$

Therefore, by partitioning according to the first element and by the symmetry between $\{a, A\}$ and $\{b, B\}$, we obtain

$$\begin{aligned} |\Omega_{n,d}| &= 2|\{x \in \Omega_{n,d} \mid x_1 = a\}| + 2|\{x \in \Omega_{n,d} \mid x_1 = A\}| \\ &= 2c_{n-2}(0) + 2c_{n-3}(d - 1). \end{aligned}$$

The result now follows as $|\Omega_{n,d}| = |\text{Hom}(P_{n,d})|$, by Claim 6.2. \square

Lemma 6.4. For any $n \geq 3$, we have

$$c_n(0) = c_{n-2}(0) + c_{n-2}(d-1) + c_{n-3}(d-1), \tag{6.7}$$

$$c_n(k) = c_{n-2}(k-1) + c_{n-2}(d-1), \quad 1 \leq k \leq d-1. \tag{6.8}$$

Proof. Note that by (6.5), similarly to (6.6), we have

$$\begin{aligned} \{x \in \Omega_{n,d}(0, d) \mid x_1 = a\} &= \{(a) \circ x \mid x \in \Omega_{n-2,d}(0, d)\}, \\ \{x \in \Omega_{n,d}(0, d) \mid x_1 = b\} &= \{(b) \circ x \mid x \in \Omega_{n-2,d}(d, d-1)\}, \\ \{x \in \Omega_{n,d}(0, d) \mid x_1 = A\} &= \{(A) \circ x \mid x \in \Omega_{n-3,d}(d-1, d)\}. \end{aligned}$$

Therefore, by partitioning according to the first element, we obtain

$$\begin{aligned} c_n(0) &= |\Omega_{n,d}(0, d)| \\ &= |\{x \in \Omega_{n,d}(0, d) \mid x_1 = a\}| + |\{x \in \Omega_{n,d}(0, d) \mid x_1 = b\}| + |\{x \in \Omega_{n,d}(0, d) \mid x_1 = A\}| \\ &= |\Omega_{n-2,d}(0, d)| + |\Omega_{n-2,d}(d, d-1)| + |\Omega_{n-3,d}(d-1, d)| \\ &= c_{n-2}(0) + c_{n-2}(d-1) + c_{n-3}(d-1). \end{aligned}$$

In a similar manner, for $1 \leq k \leq d-1$, we have

$$\begin{aligned} c_n(k) &= |\Omega_{n,d}(k, d)| \\ &= |\{x \in \Omega_{n,d}(k, d) \mid x_1 = a\}| + |\{x \in \Omega_{n,d}(k, d) \mid x_1 = b\}| \\ &= |\{(a) \circ x \mid x \in \Omega_{n-2,d}(k-1, d)\}| + |\{(b) \circ x \mid x \in \Omega_{n-2,d}(d, d-1)\}| \\ &= c_{n-2}(k-1) + c_{n-2}(d-1). \end{aligned} \quad \square$$

We express all quantities $c_n(k)$ in terms of $c_n(d-1)$. Substituting $k = d-1$ in (6.8) yields

$$c_n(d-2) = c_{n+2}(d-1) - c_n(d-1).$$

Now substituting $k = d-2$ in (6.8) gives

$$c_n(d-3) = c_{n+2}(d-2) - c_n(d-1) = c_{n+4}(d-1) - c_{n+2}(d-1) - c_n(d-1).$$

Continuing in this manner (by induction), we get for $1 \leq m < d$,

$$c_n(d-m-1) = c_{n+2m}(d-1) - c_{n+2m-2}(d-1) - \dots - c_{n+2}(d-1) - c_n(d-1). \tag{6.9}$$

In particular, for $m = d-1$ this gives,

$$c_n(0) = c_{n+2d-2}(d-1) - c_{n+2d-4}(d-1) - \dots - c_{n+2}(d-1) - c_n(d-1).$$

Substituting this in (6.7) gives

$$c_{n+2d-2}(d-1) = 2c_{n+2d-4}(d-1) + c_{n-3}(d-1), \quad n \geq 3.$$

The characteristic polynomial for this equation is

$$q(\mu) := \mu^{2d-1}(\mu^2 - 2) - 1.$$

Claim 6.5. The polynomial q has $2d+1$ distinct (complex) roots. Exactly one of these, which we denote by μ_0 , is positive. Moreover, $\mu_0 > \sqrt{2}$, while all other roots have modulus less than $\sqrt{2}$.

Proof. Assume that $d \geq 2$ (the case $d = 1$ can be verified directly). It is easy to verify that the derivative of q does not vanish at any zero, so that the roots are simple, and hence there are $2d + 1$ distinct roots. Since $q(\pm\sqrt{2}) = -1$, $q(\sqrt{3}) > 0$ and $q(-2/\sqrt{3}) > 0$, the intermediate value theorem implies that there are roots $\sqrt{2} < \mu_0 < \sqrt{3}$ and $-\sqrt{2} < \mu_1 < -2/\sqrt{3}$. Considering q as a real function, by differentiating, one finds that q has a single minimum and a single maximum, and hence at most 3 real roots. Since $q(-1) = 0$, we see that μ_0 is indeed the unique positive root. For the last part, it suffices to show q has $2d - 1$ roots of modulus at most 1. This is a consequence of Rouché's theorem applied to q and $g(z) := 2z^{2d-1}$ on the disc $D := \{z \in \mathbb{C} \mid |z| \leq r\}$ for any sufficiently small $r > 1$. Indeed, on ∂D , we have $|g(z)| = 2r^{2d-1}$ and $|q(z) + g(z)| \leq r^{2d+1} + 1$, and since $r^{2d+1} + 1 < 2r^{2d-1}$ (using our assumption that $d \geq 2$), Rouché's theorem implies that g and q have the same number of zeros in D . As g clearly has $2d - 1$ zeros in D , this completes the proof. \square

Let μ_0 be the unique positive root of q . We denote $\lambda := \mu_0^2$. That is, λ is the unique positive solution of the equation

$$\lambda^{d-1/2}(\lambda - 2) = 1. \tag{6.10}$$

Claim 6.6. *For any fixed $0 \leq k \leq d - 1$, there exists a constant $r_k > 0$, such that*

$$c_n(k) \sim r_k \lambda^{n/2} \quad \text{as } n \rightarrow \infty.$$

Proof. Denote by $\mu_0 := \sqrt{\lambda}, \mu_1, \dots, \mu_{2d}$ the roots of q . The roots μ_i are distinct, and therefore,

$$c_n(k) = r_k^0 \mu_0^n + \dots + r_k^{2d} \mu_{2d}^n,$$

for some coefficients r_k^i . Indeed, this is the case for $k = d - 1$ as the polynomial q is the characteristic polynomial for the recursion of $c_n(d - 1)$ and it then follows also for $0 \leq k < d - 1$ by (6.9). Now, since any word $x \in \{a, b\}^*$ is d -legal and has $w(x) = 2|x|$, we see that

$$c_n(k) = |\Omega_{n,d}(k, d)| \geq |\{a, b\}^{\lceil n/2 \rceil}| = 2^{\lceil n/2 \rceil} \geq \sqrt{2}^n.$$

Therefore, since $|\mu_i| < \sqrt{2}$ for $1 \leq i \leq 2d$, we must have $r_k^0 > 0$ for all $0 \leq k \leq d - 1$, and then

$$c_n(k) \sim r_k^0 \mu_0^n = r_k^0 \lambda^{n/2}. \tag{6.11} \quad \square$$

We now have all the ingredients to prove Theorem 2.10.

Proof of Theorem 2.10. By Lemma 6.3 and Claim 6.6, we have

$$\begin{aligned} |\text{Hom}(P_n, d)| &= 2c_{n-2}(0) + 2c_{n-3}(d - 1) \\ &\sim 2r_0 \lambda^{(n-2)/2} + 2r_{d-1} \lambda^{(n-3)/2} = C(d) \lambda^{n/2} (1 + o(1)) \quad \text{as } n \rightarrow \infty. \quad \square \end{aligned}$$

Claim 6.6 gives the asymptotic behavior of $c_n(k)$ as $n \rightarrow \infty$ for fixed $d \geq 1$. Specifically, it says that the order of magnitude of $c_n(k)$ is λ^n , where $\lambda = \lambda(d)$ depends on d and is given implicitly by (6.10). The next claim describes the dependence of the constant $\lambda(d)$ on d as $d \rightarrow \infty$.

Claim 6.7. *The unique positive solution λ of (6.10) satisfies*

$$\lambda = \lambda(d) = 2 + 2^{-d+1/2} (1 - o(1)) \quad \text{as } d \rightarrow \infty.$$

Proof. Writing $\lambda = 2 + \delta$, we have by (6.10) that

$$\delta = (2 + \delta)^{-d+1/2}.$$

Thus, since $\delta > 0$, we have

$$\delta \leq 2^{-d+1/2},$$

and therefore,

$$2^{d-1/2}\delta = (1 + \delta/2)^{-d+1/2} \rightarrow 1 \quad \text{as } d \rightarrow \infty. \quad \square$$

6.3 Infinite homomorphisms

Denote by $P_{\infty,d}$ the graph on the vertex set $\{0, 1, 2, \dots\}$ with the edge set $\{(i, j) \mid |i - j| = 1, 3, \dots, 2d + 1\}$. Note that $\text{Hom}(P_{\infty,d}) := \text{Hom}(P_{\infty,d}, 0)$ is an infinite set of homomorphisms. For a homomorphism $f \in \text{Hom}(P_{n,d})$ (where possibly $n = \infty$) and an integer $r \geq 0$, we denote by $B_r(f)$ the restriction of f to the first $r + 1$ vertices, so that

$$B_r(f) := f|_{\{0,1,\dots,\min\{r,n\}\}} \in \text{Hom}(P_{\min\{r,n\},d}).$$

An infinite word x on Σ is d -legal if it satisfies (6.5), as for finite words. Denote by $\Omega_{\infty,d}$ the set of infinite d -legal words on Σ . That is,

$$\Omega_{\infty,d} := \{x \in \Sigma^{\mathbb{N}} \mid x \text{ is } d\text{-legal}\}.$$

The mapping D_n defined in (6.1) extends to the case $n = \infty$ in an obvious way. The mapping T defined in (6.3) can also be extended to map the infinite words $D_{\infty}(\text{Hom}(P_{\infty,d}))$ to $\Omega_{\infty,d}$ by the same recursion formula. Then, following the proof of Claim 6.2, we see that

$$L_{\infty} := (T \circ D_{\infty})|_{\text{Hom}(P_{\infty,d})}$$

is a bijection between $\text{Hom}(P_{\infty,d})$ and $\Omega_{\infty,d}$.

6.4 The local limit as a Markov chain

The main goal of this section is to prove Theorem 2.12. To this end, we will describe a Markov chain (see Figure 12) on the state space

$$\tilde{\Sigma} := \{a_1, \dots, a_d, b_1, \dots, b_d, A, B\},$$

which will allow us to generate words in $\Omega_{\infty,d}$, and hence also homomorphisms in $\text{Hom}(P_{\infty,d})$ through the bijection L_{∞} . Loosely speaking, the idea of this Markov chain is that the state a_k (b_k) represents the fact that a streak of k consecutive a 's (b 's) has been accumulated. Likewise, the state A (B) represents the fact that a jump has occurred in the positive (negative) direction.

Consider the above Markov chain (see Figure 12) on the state space $\tilde{\Sigma}$ with the transition probabilities p and the initial state distribution π as described below.

$$\begin{aligned} p(A, b_1) &= p(a_d, a_d) := \lambda^{-1}, \\ p(A, a_1) &:= \begin{cases} 0 & \text{if } d \geq 2 \\ \lambda^{-1} & \text{if } d = 1 \end{cases}, \\ p(a_k, b_1) &:= \frac{\lambda - 1}{\lambda^k(\lambda - 2) + \lambda}, \quad 1 \leq k \leq d, \end{aligned} \tag{6.11}$$

where λ is the unique positive solution to (6.10). The analogous relations hold with the roles of $\{a, A\}$ and $\{b, B\}$ interchanged. Figure 12 shows the legal transitions (i.e., transitions having positive probability). The probability of unspecified legal transitions

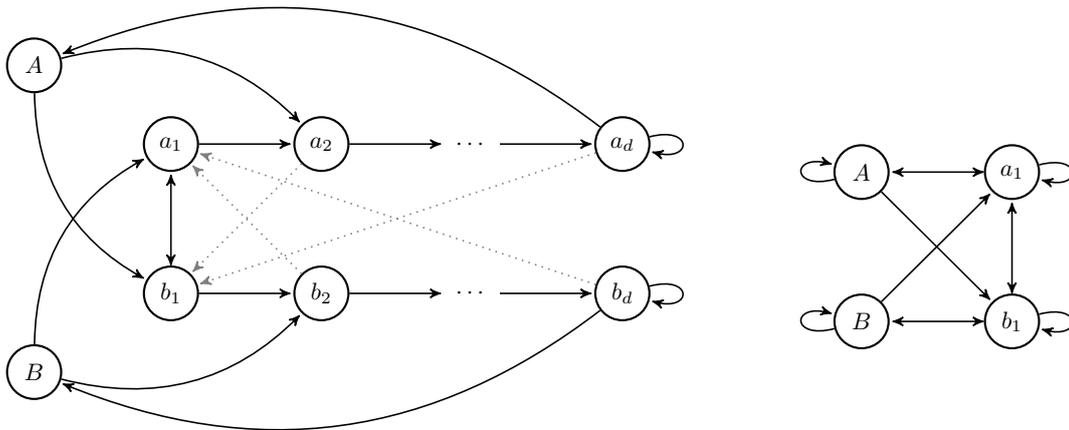


Figure 12: The Markov chains describing the local limit when $d \geq 2$ (on the left) and when $d = 1$ (on the right). The allowed transitions are those determined by (6.5).

are determined by the condition $\sum_s p(s', s) = 1$. The initial state distribution π is given by

$$\pi(a_d) = \pi(b_d) := \frac{\lambda^{-1/2}}{2} \quad \text{and} \quad \pi(A) = \pi(B) := \frac{1 - \lambda^{-1/2}}{2}.$$

It is interesting to note that, since $\lambda > 2$ and using (6.10), we have

$$\frac{1}{2} > \lambda^{-1} = p(a_1, b_1) > p(a_2, b_1) > \dots > p(a_d, b_1) = \frac{\lambda - 1}{\lambda + \sqrt{\lambda}} > \frac{1}{2 + \sqrt{2}},$$

which expresses the fact that there is a small but growing tendency to continue in the same direction.

Running this chain for an infinite amount of time and considering its trajectory as an infinite word on $\tilde{\Sigma}$, we may obtain an infinite word W_∞ on Σ by dropping the subscripts of the letters in $\tilde{\Sigma}$. More precisely, let $\tilde{W}(1), \tilde{W}(2), \dots$ be a Markov chain on $\tilde{\Sigma}$ with transition probabilities as in (6.11) and such that $\tilde{W}(1) \sim \pi$. Define $\phi: \tilde{\Sigma} \rightarrow \Sigma$ by $\phi(a_i) := a$, $\phi(b_i) := b$, $\phi(A) := A$ and $\phi(B) := B$. Then W_∞ is defined by $W_\infty(k) := \phi(\tilde{W}(k))$ for $k \geq 1$. Recalling (6.5), it is clear that this process generates a d -legal word, i.e. that $W_\infty \in \Omega_{\infty,d}$. Denote by

$$f_\infty := L_\infty^{-1}(W_\infty)$$

the infinite homomorphism corresponding to this word. Recall the definition of $B_r(f)$ from Section 6.3. Let f_n be a uniformly chosen homomorphism in $\text{Hom}(P_{n,d})$. Theorem 2.12 will follow when we show that

$$\mathbb{P}(B_r(f_n) = f) \xrightarrow{n \rightarrow \infty} \mathbb{P}(B_r(f_\infty) = f) \quad \text{for any } r \geq 1 \text{ and } f \in \text{Hom}(P_{r,d}). \quad (6.12)$$

For $n \geq 1$, define

$$W_n := L_n(f_n).$$

The following lemma links the uniform distribution on homomorphisms to the above Markov chain. For any $1 \leq n \leq \infty$ and any word $x \in \Omega_{n,d}$, there exists a unique trajectory $(s_1, s_2, \dots, s_{|x|})$ such that $s_i \in \tilde{\Sigma}$, $s_1 \in \{a_d, b_d, A, B\}$ and $p(s_i, s_{i+1}) > 0$, which generates the word x by the process of dropping the subscripts of the symbols in $\tilde{\Sigma}$. For a finite word x , define $\text{State}(x) := s_{|x|}$ to be the final state of this trajectory. Let $P_k: \Sigma^* \rightarrow \Sigma^k$ denote the truncation to length k .

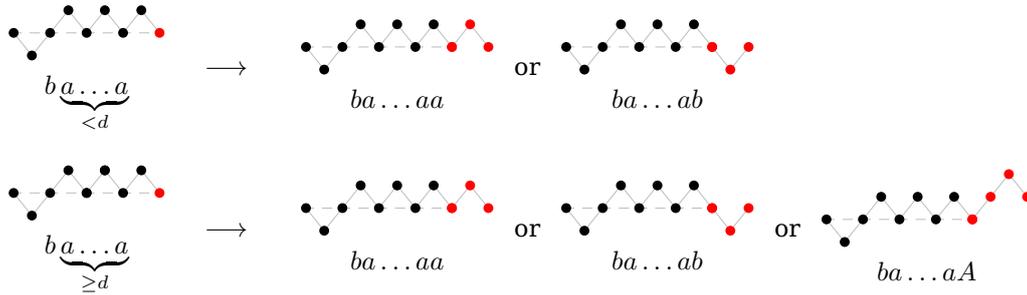


Figure 13: The possible transitions from state a_k for $1 \leq k < d$ (on the top) and for $k = d$ (on the bottom), as determined by (6.5). The transitions from state A are analogous, and the transitions from states b_k and B are symmetric.

Lemma 6.8. For any $u \in \Sigma$ and $x \in \Sigma^*$ such that x and $x \circ (u)$ are d -legal, we have

$$\begin{aligned} \mathbb{P}(W_n(1) = u) &\xrightarrow[n \rightarrow \infty]{} \pi(\text{State}(u)) \\ \mathbb{P}(W_n(|x| + 1) = u \mid P_{|x|}(W_n) = x) &\xrightarrow[n \rightarrow \infty]{} p(\text{State}(x), \text{State}(x \circ (u))). \end{aligned} \tag{6.13}$$

Proof. For a d -legal word $x \in \Sigma^*$, define $M(x) := i - 1$ if $\text{State}(x) \in \{a_i, b_i\}$ for $1 \leq i \leq d$ and $M(x) := 0$ if $\text{State}(x) \in \{A, B\}$. Then, using Claim 6.2, we have for any $n \geq w(x)$ that

$$|\{P_{|x|}(W_n) = x\}| = c_{n-w(x)}(d - M(x) - 1).$$

By Claim 6.6, we have $c_n(d - 1) \sim a\lambda^{n/2}$, for some constant $a > 0$, and then (6.9) gives

$$c_n(d - m - 1) \sim a\lambda^{n/2+m} - a\lambda^{n/2+m-1} - \dots - a\lambda^{n/2+1} - a\lambda^{n/2}, \quad 1 \leq m < d.$$

Thus,

$$c_n(d - m - 1) \sim a\lambda^{n/2} \frac{\lambda^m(\lambda - 2) + 1}{\lambda - 1}, \quad 0 \leq m < d. \tag{6.14}$$

Therefore, if $\text{State}(x) = a_i$ with $1 \leq i \leq d$, then $\text{State}(x \circ (b)) = b_1$ and

$$\begin{aligned} \mathbb{P}(W_n(|x| + 1) = b \mid P_{|x|}(W_n) = x) &= \frac{|\{P_{|x|+1}(W_n) = x \circ (b)\}|}{|\{P_{|x|}(W_n) = x\}|} \\ &= \frac{c_{n-w(x)-2}(d - 1)}{c_{n-w(x)}(d - i)} \sim \frac{\lambda - 1}{\lambda^i(\lambda - 2) + \lambda} = p(a_i, b_1). \end{aligned}$$

The remaining cases are handled similarly by taking the relevant ratios. This proves the second part of (6.13). For the first part, we will also need to compute the size of $\text{Hom}(P_{n,d})$. By Lemma 6.3, (6.14) and (6.10), we have

$$|\text{Hom}(P_{n,d})| = 2c_{n-2}(0) + 2c_{n-3}(d - 1) \sim 2a\lambda^{n/2-1} \frac{\lambda^{1/2} + 1}{\lambda - 1}. \tag{6.15}$$

Therefore, since $\text{State}((a)) = a_d$,

$$\mathbb{P}(W_n(1) = a) = \frac{|\{P_1(W_n) = (a)\}|}{|\text{Hom}(P_{n,d})|} = \frac{c_{n-2}(0)}{|\text{Hom}(P_{n,d})|} \sim \lambda^{-1/2}/2 = \pi(a_d).$$

The remaining cases are again handled similarly. This proves the first part of (6.13). \square

We continue by observing, using Claim 6.1, that for any $r \geq 1$ and $f \in \text{Hom}(P_{r,d})$ there exists a $k \geq 1$ and a set $X(f) \subset \Sigma^k$ such that we have the equality of events,

$$\{B_r(f_n) = f\} = \{P_k(W_n) \in X(f)\}, \quad r \leq n \leq \infty. \tag{6.16}$$

For instance, one may take $X((0, -1, 0, 1, 0)) = \{(b, a)\}$ and $X((0, -1, 0, 1, 0, 1)) = \{(b, a, a), (b, a, A)\}$. In addition, Lemma 6.8 implies that

$$\mathbb{P}(P_k(W_n) = x) \xrightarrow{n \rightarrow \infty} \mathbb{P}(P_k(W_\infty) = x) \text{ for any } k \geq 0 \text{ and } x \in \Sigma^k. \tag{6.17}$$

This follows directly from (6.13) when x is d -legal, and it follows trivially when x is not d -legal since the probabilities involved are zero.

Finally, putting together (6.16) and (6.17), we conclude that for any $r \geq 1$ and $f \in \text{Hom}(P_{r,d})$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}(B_r(f_n) = f) = \lim_{n \rightarrow \infty} \sum_{x \in X(f)} \mathbb{P}(P_{|x|}(W_n) = x) = \sum_{x \in X(f)} \mathbb{P}(P_{|x|}(W_\infty) = x) = \mathbb{P}(B_r(f_\infty) = f),$$

proving (6.12), as required.

We remark that it is now simple to derive an exact formula for the probability that $B_r(f_\infty) = f$ for certain homomorphisms $f \in \text{Hom}(P_{r,d})$. Specifically, let $f \in \text{Hom}(P_{r,d})$ satisfy $w(L_r(f)) = r$. For such f , one may take $X(f) = \{L_r(f)\}$. Hence, denoting $x := L_r(f)$ and $m := M(x)$ (defined in the proof of Lemma 6.8), we have using (6.14) and (6.15) that

$$\begin{aligned} \mathbb{P}(B_r(f_\infty) = f) &= \mathbb{P}(P_{|x|}(W_\infty) = x) = \lim_{n \rightarrow \infty} \mathbb{P}(P_{|x|}(W_n) = x) \\ &= \lim_{n \rightarrow \infty} \frac{c_{n-r}(d-m-1)}{|\text{Hom}(P_{n,d})|} = \frac{1}{2} \lambda^{1-r/2} \frac{\lambda^m(\lambda-2)+1}{\lambda^{1/2}+1}. \end{aligned}$$

7 Discussion and Open Problems

7.1 A continuous model

One may consider a continuous variant of the graph homomorphisms considered here. Given a finite connected graph $G = (V, E)$ and a vertex $v_0 \in V$, let

$$\text{Lip}(G, v_0) := \{f: V \rightarrow \mathbb{R} \mid f(v_0) = 0, |f(u) - f(v)| \leq 1 \text{ when } (u, v) \in E\}.$$

Thus, elements of $\text{Lip}(G, v_0)$ may be regarded as real-valued Lipschitz functions on the graph, normalized to equal 0 at v_0 . There is a natural uniform measure on $\text{Lip}(G, v_0)$ obtained by regarding a function $f \in \text{Lip}(G, v_0)$ as a vector in $\mathbb{R}^{V \setminus \{v_0\}}$ and using normalized Lebesgue measure there. Hence, one may speak of a uniformly sampled function from $\text{Lip}(G, v_0)$. In statistical physics terminology, this models a random surface whose energy is defined via the Hammock potential (see, e.g., [5]).

Naively, one may expect the behavior of a uniformly chosen function f from $\text{Lip}(P_{n,d}, 0)$ to be rather similar, perhaps up to constants, to that of a uniformly chosen function from $\text{Hom}(P_{n,d}, 0)$. In particular, one may expect that $\text{Var}(f(n)) \approx n2^{-d}$ when $n2^{-d} \geq 1$, say. However, a different intuition comes from the following consideration. A standard heuristic in statistical physics is that (continuous) models of random surfaces should behave similarly to the Gaussian free field. The Gaussian free field is again a real-valued function $g: V \rightarrow \mathbb{R}$, satisfying $g(v_0) = 0$, and sampled from a distribution whose density is proportional to

$$\exp \left(-\beta \sum_{(u,v) \in E} (g(u) - g(v))^2 \right),$$

Random walk with long-range constraints

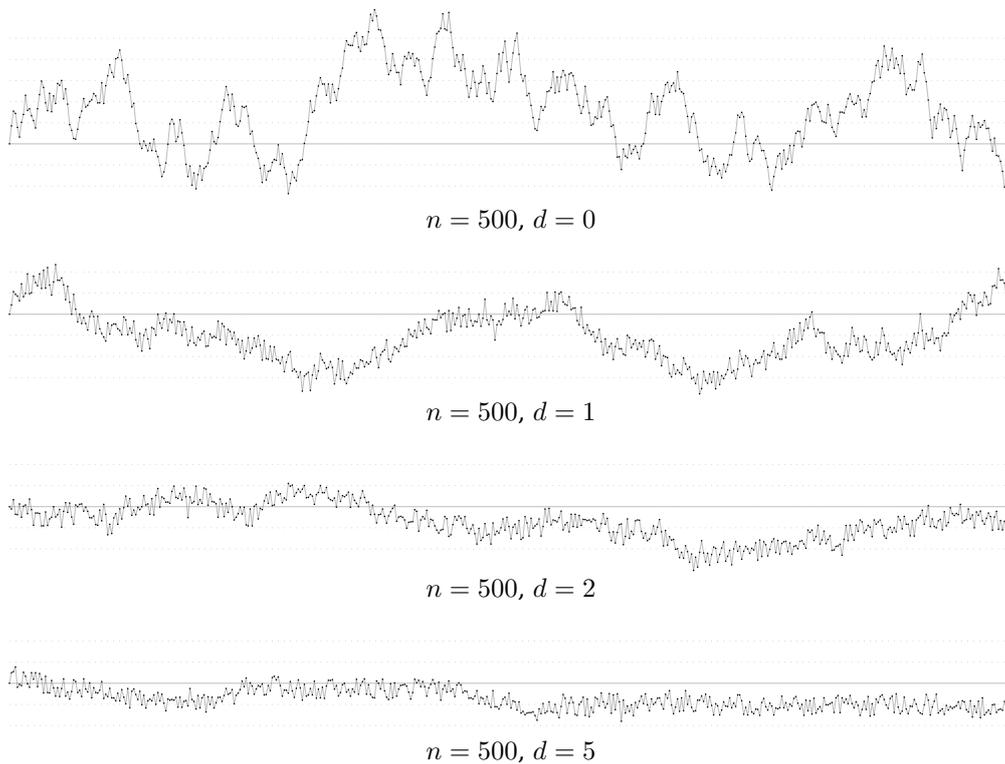


Figure 14: Uniformly sampled functions in $\text{Lip}(P_{n,d}, 0)$. The case $d = 0$ is just a random walk with independent uniform increments in $[-1, 1]$. The simulation uses a Metropolis algorithm (see, e.g., [9, Chapter 3]) and coupling from the past [12].

with $\beta \in (0, \infty)$ a parameter. Analysis of the variance of the Gaussian free field on a graph is made simple by the observation that its distribution is a multivariate Gaussian. When $G = P_{n,d}$ and $v_0 = 0$ one obtains that $\text{Var}(g(n)) \approx nd^{-3}/\beta$. Thus it is not clear whether one should expect a function f sampled uniformly from $\text{Lip}(G, v_0)$ to satisfy $\text{Var}(f(n)) \approx n2^{-d}$ or $\text{Var}(f(n)) \approx nd^{-\alpha}$. We conjecture the latter to be the truth. Thus, we expect a significant difference in behavior between the homomorphism model considered in this paper and its continuous counterpart. Consideration of the complete graph suggests that, when comparing the Gaussian free field to the continuous Lipschitz model on a regular graph, one should take β to be one over the degree. As $P_{n,d}$ is nearly a $(2d + 2)$ -regular graph, this leads to the following conjecture.

Conjecture 7.1. *There exist absolute constants $C, c > 0$ such that the following holds for any positive integers n and d . If f is uniformly sampled from $\text{Lip}(P_{n,d}, 0)$ then*

$$c(nd^{-2} + 1) \leq \text{Var}(f(n)) \leq C(nd^{-2} + 1).$$

In particular, the threshold function $d(n)$ separating the regime of localization from the regime of delocalization is polynomial in n , rather than logarithmic in n as is the case for the homomorphism model. Figure 14 shows a uniformly sampled function in $\text{Lip}(P_{n,d}, 0)$. We remark that when considering this model it is natural to consider the non-bipartite graph $\tilde{P}_{n,d}$, which is the discrete segment $\{0, 1, \dots, n\}$ with edges between vertices at distance at most $d + 1$, regardless of their parity.

7.2 The scaling limit

In this paper we explored the properties of a random homomorphism for given n and d , and also the local limit of the homomorphism when d is fixed and n tends to infinity. Another limit of interest is the scaling limit. As in many models of random walk, one may expect that in the subcritical regime, when the range of a homomorphism in $\text{Hom}(P_{n,d}, 0)$ tends to infinity as n tends to infinity, the homomorphism has a Brownian motion scaling limit. This is the content of the next conjecture.

Conjecture 7.2. *There exists a function $\sigma: \mathbb{N} \rightarrow (0, \infty)$ such that the following holds. Let $f_{n,d}$ be a uniformly chosen homomorphism in $\text{Hom}(P_{n,d}, 0)$. Define $B_{n,d}: [0, 1] \rightarrow \mathbb{R}$ to be the continuous function defined by*

$$B_{n,d} \left(\frac{i}{n} \right) := \frac{f_{n,d}(i)}{\sigma(d)\sqrt{n}}$$

and interpolated linearly between these points. If $d(n) - \log_2 n \rightarrow -\infty$ as $n \rightarrow \infty$, then $B_{n,d(n)}$ converges in distribution as $n \rightarrow \infty$ to a standard Brownian motion on $[0, 1]$.

An educated guess for the function σ may be obtained as follows. Recall the local limit $f_{\infty,d}$ from Section 6. The fact that $f_{\infty,d}$ may be described via a Markov chain simplifies the analysis of its scaling limit. Define

$$\sigma'(d)^2 := \frac{(\lambda(d) - 2)(\lambda(d) - 1)}{4 + (2d + 1)(\lambda(d) - 2)},$$

where $\lambda(d)$ is defined in Theorem 2.10. Observe that, by Claim 6.7,

$$\sigma'(d)^2 = 2^{-d-3/2}(1 - o(1)) \quad \text{as } d \rightarrow \infty.$$

Then, defining the continuous function

$$B'_{n,d} \left(\frac{i}{n} \right) := \frac{f_{\infty,d}(i)}{\sigma'(d)\sqrt{n}},$$

interpolated linearly between these points, it may be shown that for any fixed d the process $B'_{n,d}$ converges in distribution as $n \rightarrow \infty$ to a standard Brownian motion on $[0, 1]$. Thus, it seems plausible that the $\sigma(d)$ of the above conjecture equals $\sigma'(d)$.

References

- [1] Itai Benjamini, Olle Häggström, and Elchanan Mossel, *On random graph homomorphisms into \mathbf{Z}* , J. Combin. Theory Ser. B **78** (2000), no. 1, 86–114. MR-1737627
- [2] Itai Benjamini and Yuval Peres, *Markov chains indexed by trees*, Ann. Probab. **22** (1994), no. 1, 219–243. MR-1258875
- [3] ———, *Tree-indexed random walks on groups and first passage percolation*, Probab. Theory Related Fields **98** (1994), no. 1, 91–112. MR-1254826
- [4] Itai Benjamini, Ariel Yadin, and Amir Yehudayoff, *Random graph-homomorphisms and logarithmic degree*, Electron. J. Probab. **12** (2007), no. 32, 926–950. MR-2324796
- [5] Herm Jan Brascamp, Elliot H. Lieb, and Joel L. Lebowitz, *The statistical mechanics of anharmonic lattices*, Proceedings of the 40th Session of the International Statistical Institute (Warsaw, 1975), Vol. 1. Invited papers, vol. 46, 1975, pp. 393–404 (1976). MR-0676341
- [6] Paul Erdős, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. **51** (1945), 898–902.
- [7] David Galvin, *On homomorphisms from the Hamming cube to \mathbf{Z}* , Israel J. Math. **138** (2003), 189–213. MR-2031957

- [8] Jeff Kahn, *Range of cube-indexed random walk*, Israel J. Math. **124** (2001), 189–201. MR-1856513
- [9] David Asher Levin, Yuval Peres, and Elizabeth Lee Wilmer, *Markov chains and mixing times*, AMS Bookstore, 2009. MR-2466937
- [10] Ron Peled, *High-dimensional Lipschitz functions are typically flat*, arXiv:1005.4636, to appear in Annals of Probability (2010).
- [11] Ron Peled, Wojciech Samotij, and Amir Yehudayoff, *Lipschitz functions on expanders are typically flat*, Combinatorics, Probability and Computing **22** (2013), 566–591. MR-3073490
- [12] James Gary Propp and David Bruce Wilson, *Exact sampling with coupled Markov chains and applications to statistical mechanics*, Proceedings of the Seventh International Conference on Random Structures and Algorithms (Atlanta, GA, 1995), vol. 9, 1996, pp. 223–252. MR-1611693